1 Microbial genetic and enzymatic responses to an anthropogenic phosphorus gradient
2 within a subtropical peatland
3

4 Elise Morrison¹, S. Newman², H.S. Bae¹, Z. He³, J. Zhou³, K. R. Reddy¹, and A. Ogram^{1*}

5 ¹Soil and Water Science Department, University of Florida, Gainesville, FL 32611 USA

6 ² Everglades Systems Assessment Section, South Florida Water Management District,
7 West Palm Beach, FL USA 33406 USA

8 ³Institute for Environmental Genomics, University of Oklahoma, Norman, OK 73019 USA
9

10 *Corresponding author:

11 Andrew Ogram

12 2181 McCarty Hall

13 Soil and Water Science Department

14 University of Florida

15 Gainesville, FL 32611-0290

16 USA

17

18 Email: aogram@ufl.edu

19 Phone: 352-294-3134
20

21 Keywords: Everglades, peat, nutrient limitation, phosphatase genes, GeoChip, enzyme
22 assays

ABSTRACT

Many of the world's peatlands are subject to agricultural runoff, which may impact fundamental biogeochemical processes by altering the concentrations of limiting nutrients. In this study, the distribution of specific microbial genes associated with phosphorus (P) and nitrogen (N) metabolism were compared with potential enzyme activities related to P and N acquisition at five stations along a nutrient gradient in soils of Water Conservation Area 2A (WCA-2A) of the Florida Everglades, USA. Quantitative PCR was used to compare the relative concentrations of genes encoding alkaline phosphatases (*phoX* and *phoD*) with those encoding dinitrogenase reductase (*nifH*). Combined phosphatase and phosphodiesterase (E_P) activities were compared with leucine aminopeptidase activities (E_N), yielding a measure of potential microbial investment in P acquisition relative to N acquisition. The significant inverse relationship observed between bicarbonate extractable organic P concentrations and ratios of gene copy numbers of *phoX:nifH* ($p=0.049$, $R^2=0.77$), and *phoD:nifH* ($p=0.043$, $R^2=0.79$), combined with the significant inverse relationship between total P and E_P/E_N ($p=0.021$, $R^2=0.87$), suggest that there is a greater community selection towards P acquisition relative to N acquisition as bicarbonate extractable organic P and total P decrease, suggesting a shift from P limitation to N limitation along the transect. The total number of unique genes detected by the functional microarray GeoChip 3.0 was greatest in an intermediate site, suggesting that the alleviation of nutrient limitation yielded increased functional diversity. The general agreement between genetic and enzymatic data suggests that assessing microbial nutrient demands with molecular techniques is

feasible, although future work is needed to apply genetic information as indicators of nutrient enrichment.

1. INTRODUCTION

Wetlands provide numerous important ecosystem services, such as regulation of water quality and quantity, provision of critical habitats for diverse species (including endangered species), and recreation opportunities. Many wetlands, including the Florida Everglades, are subject to nutrient enrichment that impacts many of these services. The majority of the Florida Everglades is underlain by limestone, such that ecosystem productivity was historically limited by phosphorus (P) availability. Much of the original water flow into the Everglades was diverted to the Gulf of Mexico and the Atlantic Ocean to control flooding, and approximately 280,000 ha south of Lake Okeechobee were drained for development of the Everglades Agricultural Area (EAA) (McCray et al., 2012).

Fertilizer runoff from the EAA into what is now the Water Conservation Areas (WCAs; Fig. S1) contained significant levels of nutrients, notably phosphate and sulfate (Castro et al., 2002). The input of phosphate into the phosphate-limited marsh resulted in broad ecosystem level changes, including a change in the dominant vegetation from ridges dominated by low-P adapted sawgrass (*Cladium jamaicense* Crantz) to densely distributed cattail (*Typha* spp.) (Hagerthey et al., 2008; Belanger et al., 1989). The increased primary productivity resulted in increased peat accumulation and associated microbially mediated processes, such

as microbial respiration, and methane production (Holmes et al., 2014), in addition to the loss of habitat for numerous species (McCormick et al., 2009).

WCA-2A is characterized by a gradient in water and soil concentrations of P, with highest concentrations of 53 μg per kg soil in peat nearest to the EAA declining to 5 μg per kg in the interior of the WCA (Turner and Newman, 2005). As expected, the activities and structures of the resident microbial communities are characteristic of the nutrient status along the gradient, and exhibit characteristics of a shift from P –limitation within the center of the marsh to N-limitation in the northern, P-impacted soils (Corstanje et al., 2007). Of note is the existence of a transition zone between the P-limited and N-limited zones.

The nutrient limitation gradient in WCA-2A has provided an outstanding natural laboratory in which to study the impacts of nutrient enrichment on the structures and activities of microbial communities and microbially mediated biogeochemical cycles. Much of the work previously conducted in WCA-2A has involved identification of indicators of nutrient enrichment, including potential activities of microbial enzymes involved in phosphorus acquisition. The relationships between the activities of enzymes involved in phosphorus (e.g., phosphatase) and nitrogen (e.g., amino peptidase) acquisition have been shown to be generally good indicators of the shift from P- to N-limitation in the soils of WCA-2A (Corstanje et al., 2007; Penton and Newman., 2008). Indicators of nutrient enrichment can be used in developing management strategies as well as providing basic information on how microbial communities react to shifting nutrient limitations. Although extensive work has been conducted on microbial nutrient

dynamics, particularly with regards to global cycling (DeLong and Karl, 2006; Arrigo, 2005), extracellular enzymes (Gusewell and Freeman 2005; Kang and Freeman 1999) and carbon dynamics (Limpens et al., 2008; Freeman 2001a; Freeman 2001b), little work has been conducted to evaluate genetic indicators of nutrient availability within peatland soils.

One of the fundamental ways that microbial communities react to shifting nutrient limitations, such as that observed in WCA-2A, is through changes in community structure. Changes in community structure are reflected in the distribution of specific taxonomic and functional groups of microorganisms, which can be characterized at the genetic level (Zhou et al., 2015; Handelsman 2004; Torsvik and Øvreås 2002). Shifts in the relative concentrations of specific functional genes characteristic of methanogens and sulfate reducing prokaryotes along the gradient have been well documented (Bae et al., in press; Bae et al., 2014; Castro et al., 2002; Castro et al., 2005), as have the distribution of specific phylotypes of primary and secondary fermenters via 16S rRNA gene sequence analysis (Uz et al. 2007, Chauhan et al., 2004).

A significant gap in our understanding of the fundamental mechanisms through which WCA-2A microbial communities respond and adapt to P-enrichment is how the community changes with respect to P acquisition at the genetic level. Increasing phosphatase activities with decreasing P availability in WCA-2A is well known, but nothing is known of potential differences in the genes that encode those enzymes, or of the microorganisms that harbor those genes. Previous work has been conducted on the distribution of genes encoding relatively high and low efficiency

phosphatases in marine systems (Sebastian and Ammerman, 2009); however, these studies have only been extended to a few terrestrial studies (Zaheer et al., 2009; Tan et al., 2013; Fraser et al., 2015). In addition, little is known of the potential transformation of reduced organophosphorus compounds such as phosphonates and phytic acid to serve as alternative sources of P in P-limited soils.

To address these gaps in knowledge, we assessed the relative abundances of genes related to P and N cycling along the P gradient in WCA-2A. Abundances of *phoX* and *phoD* (genes encoding high affinity phosphatases), *nifH* (gene encoding dinitrogenase reductase), and the abundances of an array of functional genes related to P-metabolism, including those encoding phytase (*phy*) and polyphosphate kinase (*ppk*) were compared with enzyme assays, including phosphatase (indicator of P-limitation) and leucine aminopeptidase (indicator of N- metabolism) and soil P-fractions at five stations along the nutrient gradient of WCA-2A. This suite of techniques was used to address our overarching hypothesis: that microbial communities will exhibit a response to changes in P-limitation at the genetic level. We found relationships between available P forms, functional genes, and related enzyme activities, suggesting that P enrichment alters the functional diversity of the Everglades' resident microbial communities with respect to P-cycling. Additionally, the phylogenetic diversity of phosphatase (*phoX*) genes, and the functional diversity shown with GeoChip analysis, indicates that nutrient shifts along the WCA-2A gradient impacts the diversity of P-cycling genes.

2. MATERIALS AND METHODS

2.1 Sites, Sampling, and Isolation of Environmental DNA for 2013

WCA-2A is located in the southern part of Florida, USA (Figure S1), and receives half of its water from rainfall, and half from canals and other structures (Turner and Newman, 2005). Agricultural runoff into WCA-2A has led to elevated P and calcium levels, resulting in an anthropogenic P gradient that runs north to south (Turner and Newman, 2005). Five sites were selected along this P gradient and ranged from nutrient impacted (F1 and F3), intermediate (F4), and unimpacted (U3 and E5) (Figure S1). Three replicate soil cores from 0 to 2 cm depth were taken in August of 2013 using a 10cm thin-walled, stainless steel sharpened edge corer from each of these five sites. Cores were visually inspected to determine the extent of the unconsolidated flocculent layer. The flocculent (if present) was removed and stored for another study. After removal of the flocculent, the top 0 to 2 cm was measured, and sectioned. Each 0 to 2 cm section was homogenized and subsampled. Subsamples (0 to 2cm) were immediately frozen on dry ice in the field, and transferred to -80°C upon returning to the lab (less than four hours). Frozen subsamples were homogenized and ground using liquid nitrogen. Environmental DNA was isolated from the homogenized, ground soils using the PowerSoil DNA isolation kit (MoBio Laboratories, Inc., Carlsbad, CA, USA). The remaining soil samples were homogenized, and all three replicates were composited and stored at 4°C for further analysis of enzyme activities and P fractionation. Samples were stored at 4°C for 2 to 3 days prior to conducting assays for phosphatase, β -

glucosidase, and leucine aminopeptidase, and 6 days prior to phenol oxidase and peroxidase assays.

2.2 Sites, Sampling, and Isolation of Environmental DNA for 2009

Soils from three of the sites described above were collected for GeoChip analysis in 2009. Sites F1, F4, and U3 were sampled for flocculent and soil from 0 to 2 cm depth. Three replicate soil cores from 0 to 2 cm depth were taken using a 10 cm thin-walled, stainless steel sharpened edge corer from each of these three sites. Cores were visually inspected to determine the extent of the unconsolidated flocculent layer. The flocculent (if present) was removed and the top 0 to 2 cm were measured and sectioned. Each 0 to 2 cm section was homogenized and frozen in the field on dry ice. Upon returning to the lab, they were transferred to a -80°C freezer. Triplicate cores were composited, and DNA was isolated from the composited sample and used for further GeoChip analysis (described in Section 2.7).

2.3 Phosphorus Fractionation

Soils from the composited cores were analyzed for P fractionation at the University of Florida's Wetland Biogeochemistry Lab. Phosphorus extractions were conducted using the method described by Ivanoff et al. (1998), which quantifies different operationally defined pools of P via chemical fractionation. This fractionation scheme identifies the following organic P (P_o) pools: bicarbonate extractable P_o , microbial biomass P, hydrochloric (HCl) acid extractable P_o , fulvic acid-associated P_o , humic acid-associated P_o , residual P_o , and total P_o . The

fractionation scheme identifies the following inorganic P (P_i) pools: bicarbonate extractable P_i , HCl-extractable P_i , and total P_i . Bicarbonate extractable P_o , bicarbonate extractable P_i , and microbial biomass P are generally thought to be the most readily available pools of P; HCl extractable P_o , HCl extractable P_i , and fulvic acid-associated P_o are considered moderately labile; while humic acid-associated P_o and residual P_o are held to be the least available pools. Total P is calculated as the sum of all P_o and P_i pools.

2.4 Enzyme Assays

Soils from the composited cores were analyzed for enzyme activities at the South Florida Water Management District, West Palm Beach, FL, USA. Assays were conducted to determine the activities of phosphatase, β -glucosidase, leucine aminopeptidase, phenol oxidase, and peroxidase using the protocols described in Penton and Newman (2007). The assay for phosphodiesterase was conducted with the protocol for phosphatase described in Penton and Newman (2007), but with bis-4-Methylumbelliferyl phosphate rather than 4-methyleumbellieferyl phosphate as the substrate.

2.5 qPCR

Quantitative PCR (qPCR) was employed to determine the abundance of *phoX*, *phoD*, *nifH*, and 16S rRNA genes within environmental DNA samples. All qPCR systems were run using Sybr®Green-based detection chemistry (Life Technologies™, Carlsbad, CA, USA). The PCR primers developed by Sebastian and

205 Ammerman (2009) were used to target the prokaryotic *phoX* alkaline phosphatase
 206 gene for the F1, F3, F4, E5, and U3 sites. The forward primer was: 5'-
 207 GARGAGAACWTCCACGGYTA-3' and the reverse primer was: 5'-
 208 GATCTCGATGATRTGRCCRAAG-3'. Cycling conditions for *phoX* were as follows: one
 209 cycle of 95°C for 3 minutes, forty cycles of 95°C for 30 seconds, 52°C for 30 seconds,
 210 72°C for 30 seconds, 80°C for 30 seconds, followed by one cycle of 95°C for 15
 211 seconds, 60°C for one minute, and 95°C for 15 seconds. Degenerate PCR primers
 212 developed by Sakurai et al. (2008) were used to quantify *phoD* for all sites. The
 213 forward primer was 5'-CAGTGGGACGACCACGAGGT-3' and the reverse primer was
 214 5'-GAGGCCGATCGGCATGTCG-3'. Cycling conditions for *phoD* were as follows: one
 215 cycle of 94°C for 3 minutes, forty cycles of 94°C for 30 seconds, 57°C for 30 seconds,
 216 72°C for 30 seconds, and 80°C for 15 seconds, followed by one cycle of 95°C for 15
 217 seconds, 60°C for one minute, and 95°C for 15 seconds. Primers from Mehta et al.
 218 (2003) were used to quantify *nifH*. The forward primer was: 5'-
 219 GGHAARGGHGGHATHGGNAARTC-3' and the reverse primer was: 5'-
 220 GGCATNGCRAANCCVCCRCANAC-3'. Cycling conditions for *nifH* were as follows: one
 221 cycle of 95°C for 3 minutes, six cycles of 95°C for 30 seconds, 60°C for 30 seconds
 222 (with -1°C per cycle), and 72°C for 1 minute, then 32 cycles of 95°C for 30 seconds,
 223 55°C for 30 seconds, 72°C for one minute, and 83°C for 15 seconds, followed by one
 224 cycle of 95°C for 15 seconds, 65°C for one minute, and 95°C for 15 seconds. Primers
 225 from Harms et al. (2003) were used to quantify 16S rRNA genes. The forward
 226 primer was: 5'-ATGGCTGTCGTCAGCT-3' and reverse primer was: 5'-
 227 ACGGGCGGTGTGTAC-3'. Cycling conditions for amplification of 16S rRNA genes

were as follows: one cycle of 95°C for 3 minutes, 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, 70°C for 30 seconds, followed by one cycle of 95°C for 15 seconds, 60°C for one minute, and 95°C for 15 seconds. The copy number of each reaction was determined using the standard curve and Ct value of the reaction, and copy numbers per gram of dry soil were calculated and normalized using the NORMA-Gene algorithm (Heckmann, 2011). Melting temperatures and qPCR efficiencies were determined to confirm qPCR fidelity.

2.6 phoX sequencing and sequence analysis

Primer sets described above, from Sebastian and Ammerman (2009), were used to construct *phoX* clone libraries for the F1, F4 and U3 sites. The forward primer was: 5'-GARGAGAACWTCCACGGYTA-3' and the reverse primer was: 5'-GATCTCGATGATRTGRCCRAAG-3'. PCR was run with the cycling conditions described above for *phoX*. PCR products were isolated using QIA-Quick Gel Extraction Kit (QIAGEN Inc., Germantown, MD, USA), cloned into pCR™ 4-TOPO® vector and transformed into electro competent *E. coli* using a TOPO® TA Cloning Kit® (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. Clones were sequenced at the University of Florida's Interdisciplinary Center for Biotechnology Research (ICBR) Sanger Sequencing Core. DNA sequences were translated to protein in MEGA 5.2.2 (Tamura et al, 2011). Translated amino acid sequences were searched in GenBank for similarities with known sequences. Sequences that did not share high sequence similarity to known phosphatases were removed prior to alignment. Reference sequences currently unassigned to either

Type I or II *phoX* were found in GenBank and included in the Maximum Likelihood tree, along with Types I and II *phoX* reference sequences reported by Zaheer et al. (2009). A Maximum Likelihood tree was run with 1,000 permutations using MEGA 5.2.2. Remaining sequence analyses were conducted in Mothur (Schloss *et al.* 2009). Operational Taxonomic Units (OTUs) were defined using the average neighbor algorithm in Mothur with default parameter settings using the cluster.classic command. Rarefaction curves were generated in Mothur using the rarefaction.single command with a 99% cutoff. The coverage, abundance-based coverage estimators (ACE), Chao1 richness estimator, and Shannon diversity index were calculated in Mothur using the summary.single command.

All nucleic acid sequences were deposited in GenBank under the accession numbers KP659264 to KP659308.

2.7 Microbial community DNA isolation and purification for GeoChip 3.0 analysis

DNAs isolated from F1, F4, and U3 soils collected in October 2009 (described above, Section 2.2) were shipped to the Environmental Genomics Institute of the University of Oklahoma, USA. The community DNA was extracted using a freeze-grinding method as described previously, and purified using a Promega Wizard® DNA Clean-Up System (Madison, WI, USA) according to the manufacturer's directions. DNA quality was evaluated by the absorbance ratios of A260/A280 and A260/A230 ratios using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). If the ratio of A260/280 was more than 1.7 and the ratio of A260/230 was greater than 1.8, DNA was processed for further analyses.

The PicoGreen method was used for quantifying DNA concentration using a FLUOstar Optima (BMG Labtech, Jena, Germany). Purified DNA was stored at -80°C until it was used.

2.8 Microbial community DNA direct labeling, microarray hybridization and scanning

DNA (3 µg) was labeled with fluorescent dye Cy5 (GE Healthcare, Piscataway, NJ, USA) by the random priming method. Labeled DNA was purified with a QIAquick purification kit (Qiagen) and then dried in a SpeedVac (45°C, 45 min; ThermoSavant). Dried labeled DNA was suspended in hybridization buffer (50 µl; 40% formamide, 5×SSC, 0.1% SDS, 0.1 µg·µl⁻¹ Salmon sperm DNA) and denatured at 98 °C for 3 min, and then kept at 65°C until hybridization. Hybridizations were performed at 42°C for 10 hours using a MAUI 12-Bay Hybridization System (BioMicro Systems Inc, Salt Lake City, US). After hybridization, arrays were scanned with a ScanArray 500 microarray scanner (PerkinElmer, Boston, MA, USA) at 633nm using a laser power of 90% and a photomultiplier tube (PMT) gain of 75%. Scanned images were processed using ImaGene, version 6.1 (BioDiscovery, El Segundo, CA, USA).

2.9 GeoChip data pre-processing

Raw data obtained using ImaGene (BioDiscovery, El Segundo, CA, USA) were uploaded to the University of Oklahoma's Center for Environmental Genomics' microarray data manager (<http://ieg.ou.edu/microarray/>) and pre-processed using the data analysis pipeline with the following major steps: (i) Spots flagged as poor-

quality by ImaGene 6.1 and with a signal to noise ratio [SNR, $SNR = (\text{Signal Intensity-Background}) / \text{Standard deviation of background}$] less than 2.0 were removed. (ii) The normalized intensity of each spot was calculated by dividing the signal intensity of each spot by the mean intensity of the effective spots of the array. (iii) If any of replicates had (Signal Intensity-Mean Signal Intensity) more than two of the standard deviation, this replicate was removed as an outlier. Preprocessed GeoChip data were used for further statistical analysis.

3. RESULTS AND DISCUSSION

3.1 Phosphorus Fractionation

As expected, total inorganic phosphorus (TP_i) and total organic phosphorus (TP_o) concentrations were greatest at the F1 site and lowest at the U3 and E5 sites (Table 1). The values presented in Table 1 indicate clear P enrichment in the F1 and F3 sites; however, P limitation at the U3 site is not as distinct. The F4 site exhibited the greatest amount of microbial biomass phosphorus (MBP) compared to all other sites, an observation that is consistent with an earlier study (Corstanje et al., 2007). The results of the P fractionation agreed with what is currently known about the WCA-2A nutrient gradient, and were consistent with P enrichment in the F1 site and P-limitation in the U3 site (Corstanje et al., 2007).

The P extraction method used only provides information about operationally defined P pools (Turner et al., 2005); however, ^{31}P NMR was conducted for three of these sites in a previous study, and showed that phosphodiester, such as nucleic

acids and phospholipids, were the dominant P form within this system (Turner and Newman, 2005). The gene *phoD* encodes an alkaline phosphatase that has been shown to be involved in phosphate release from phosphodiester compounds in cyanobacteria (Tiwari, Microbiology 2015). It is possible that *phoD* gene abundance (Table 3) may be related to the relative abundance of phosphodiesters in this system; however, the environmental controls on abundance and expression of *phoD* are not fully understood in soil systems at this time.

3.2 Enzyme Assays

Potential enzyme activities have been used in many systems, including the Everglades, to assess nutrient limitations. In this study, enzymes associated with P and N acquisition, and with C decomposition, were studied. In general, enzyme activities were highest in the F4 site and lowest in the U3 site, with the exception of phosphatase and phosphodiesterase activities, which were greatest in the F4, E5 and U3 sites (Table 2). Phosphatase and phosphodiesterase activities were significantly correlated ($p=0.0101$, $R^2=0.92$), as would be expected. Phosphatase activities were inversely correlated with bicarbonate-extractable P_i ($p=0.0043$, $R^2=0.95$), which also agrees with studies confirming that phosphatase activity increases as available P decreases (Chen et al., 2002). Phosphodiesterase activities were inversely related to bicarbonate-extractable P_i , although this was not a statistically significant trend.

Leucine aminopeptidase activity, indicative of organic N metabolism, was greatest in the F4 site and lowest in the U3 site. Enzyme activities for phosphatase

and phosphodiesterase (E_P) were compared against leucine aminopeptidase activity (E_N), yielding a measure of potential microbial investment in P acquisition relative to N acquisition (Table 3). A significant negative correlation was also observed between $E_P:E_N$ and TP ($p=0.0214$, $R^2=0.87$), indicating that as TP concentrations decrease, microbial investment in P acquisition increases relative to N acquisition. The ratios of $E_P:E_N$ were greatest in the U3 and E5 sites, suggesting that the low TP concentrations at these sites are driving greater microbial investment in P acquisition relative to N. It is possible that there is a threshold value of P that influences microbial investment in P acquisition relative to N acquisition. It is likely that the U3 and E5 have reached that threshold, as evidenced by the higher values in $E_P:E_N$ at these sites relative to all other sites. When the enzyme activities for phosphatase and phosphodiesterase (E_P) were compared against phenol oxidase and β -glucosidase activities (E_C), no significant relationships were found with TP concentrations; however, a significant inverse relationship was seen between $E_P:E_C$ and bicarbonate extractable inorganic P ($p=0.66$, $R^2=0.0249$). This observation is consistent with the expectation that, as the available P pool increases, there would be less phosphatase activity relative to C enzymes.

β -glucosidase catalyzes the release of glucose from cellobiose (Turner et al., 2002), and its activities were greatest in the F1 site and lowest in the U3 site. Phenol oxidase activities, critical in decomposition of lignin in aerobic soils, were greatest in the U3 site and lowest in the F1 site. Peroxidase activity, also involved in lignin decomposition in aerobic soils, was greatest in the F1 site and lowest in the F4 and U3 sites.

A significant inverse relationship was also observed between phenol oxidase activities and humic acid- associated P concentrations ($p=0.034$, $R^2=0.82$). The oxidation of organic matter can release P_i (Reddy and DeLaune, 2008), such that the oxidation of organic matter via phenol oxidase may result in a decline in the concentration of humic acid P. In addition to providing carbon to microbial communities, depolymerization of complex organic phosphorus may be a mechanism whereby P_i is made available to P-limited microbial biomass.

3.3 *phoX*, *phoD*, and *nifH*, and 16S rRNA gene quantification

Concentrations of genes associated with phosphatase and N fixation may provide information as to the potential gene activity, and may be related to microbial demand for P and N at the different sites. We investigated the concentrations of genes encoding two high efficiency alkaline phosphatases (*phoX* and *phoD*) and a gene involved in N fixation (*nifH*; encoding dinitrogenase reductase) in our study sites. The gene *nifH* is complementary to our leucine aminopeptidase assay, but does not reflect similar activities. The gene *nifH* was chosen for our studies rather than the gene encoding leucine aminopeptidase because a PCR system for the leucine aminopeptidase gene has not been fully developed. However, both *nifH* abundance and leucine aminopeptidase activities are complementary approaches to assess the potential microbial investment in N acquisition. Concentrations of genes do not necessarily correlate with activities, however, such that numbers of specific genes should be interpreted in the context of enzyme activities and other properties. Therefore, our results only represent the

abundance of a gene (measured as gene copy numbers per g of soil), rather than the expression of these genes. It should also be noted that our enumeration of *phoX* copy numbers relied on PCR primers that amplify a significant proportion of non-*phoX* sequences (see section 3.3, above), such that these numbers overestimate the true copy numbers of *phoX* by varying degrees in the different samples, such that the true numbers at F4 may be overestimated at most by a factor of two, and less in the other samples.

Gene copy numbers per gram of dry soil for 16S rRNA genes and *phoX* were highly significantly correlated ($p=0.0001$, $R^2=0.99$), as were 16S rRNA gene and *nifH* copy numbers gram soil⁻¹ ($p=0.007$, $R^2=0.93$), and *phoX* and *nifH* gene copy numbers gram soil⁻¹ ($p=0.005$, $R^2=0.95$) (Table 3). Additionally, there was a positive correlation between 16S rRNA gene, *phoX*, and *nifH* copy numbers gram soil⁻¹ and readily available, bicarbonate-extractable P_i ($p=0.0098$, $R^2=0.92$; $p=0.0061$, $R^2=0.94$; and $p=0.019$, $R^2=0.88$, respectively). It is likely that the high availability of P at the enriched site allowed for increases in microbial numbers, as evidenced by the high 16S rRNA gene copy numbers at the enriched sites, thereby leading to general enrichment of many of the genes studied here. Although 16S rRNA gene copy per cell can vary depending on the type of bacterium, 16S rRNA gene copies have been used as a proxy for bacterial cell numbers in soils (Vetrovsky et al, 2013).

Unlike *phoX*, *phoD* was not significantly correlated with 16S rRNA gene copy numbers or the concentrations of the measured P fractions. However, *phoD* was greatly enriched in the very low-P E5 site relative to the other sites (Table 3), suggesting that *phoD* is controlled by different factors than *phoX*, which was

specifically enriched at the higher available-P F1. As previously mentioned, genes for P acquisition have been more thoroughly described in marine and agricultural systems, hence there is a paucity of primers developed for P cycling genes within freshwater wetland soils. Therefore, this study used primer sets from systems other than freshwater wetlands. The primer set used for *phoD* was developed by Sakurai et al (2008) for Andisol soils, (see section 3.4, below), which could explain the differences between *phoX* and *phoD* gene copy number gram soil⁻¹ seen at each site, because the primer sets were developed for different systems. Alternatively, it may be that *phoX* and *phoD* are selected by different environmental factors, such as phosphodiester concentrations, and that their gene abundances may not be correlated in peatland soils. As mentioned previously, the alkaline phosphatase encoded by *phoD* has been shown to have phosphodiesterase activity in cyanobacteria, while the alkaline phosphatase encoded by *phoX* has been shown to have phosphomonoesterase activity in cyanobacteria. It is possible that differences in substrate availability, or different environmental controls, may be effecting differences in their gene abundances at these sites. However, further work is necessary to determine what driver is influencing their distribution and abundance in these soils.

The general correlations between 16S rRNA genes and some of the functional genes studied here were not equally proportional between all genes, such that correlations between specific genes yielded insight into factors controlling their specific enrichment (Table 3). Significant inverse relationships were observed between bicarbonate extractable P_o concentrations and *phoX:nifH* (p=0.049,

$R^2=0.77$), *phoD:nifH* ($p=0.043$, $R^2=0.79$), and *(phoX+phoD):nifH* ($p=0.043$, $R^2=0.79$) (Table 3; Fig. S2). The high correlations between *phoX* and *phoD*, and *nifH* are driven by the very high ratios at the low P site E5. It is possible that there is also a P threshold for gene ratios (similar to that seen with E_P/E_N above) and that the E5 site has reached this threshold, resulting in an enrichment of *phoX* and *phoD* relative to *nifH*. Additionally, these results indicate that as bicarbonate extractable P_o increases, the abundance of microbial alkaline phosphatase genes declines relative to genes involved in N fixation, suggesting greater community selection for N acquisition relative to P acquisition. This is consistent with the conclusion that N becomes limiting as P-limitation is alleviated along this transect (Corstanje et al., 2007).

3.4 PhoX sequence analysis

phoX encodes a high efficiency phosphatase that has been studied in P-limited marine systems (Sebastian and Ammerman, 2009), but little work has been reported on its distribution or importance in freshwater wetlands such as the Everglades. The *phoX* gene product, PhoX, is generally thought to be important in accessing phosphate in low-P environments, and more important in those systems than the better studied *phoA* phosphatase, which is more important in relatively high P environments (Sebastian and Ammerman, 2009). Knowledge of the diversity and phylogeny of PhoX along the P-gradient in WCA-2A will provide insight into the importance of this alternative phosphatase in P-limited wetlands.

A potential limitation of our study is the lack of an extensive database of freshwater and terrestrial sequences that could be used to design *phoX* PCR primers. The primer set used for *phoX* was taken from Sebastian and Ammerman (2009), which were based largely on sequences from marine systems. Consequently, the primers were not highly specific to the *phoX* in the metagenomes of WCA -2A, resulting in non-specific amplification products present in clone libraries. This required that sequences that did not share similarity with *phoX* be removed from clone libraries from all sites prior to analysis. Of the *phoX* sequences retrieved, 31% of sequences were removed for F1, 78% of sequences were removed for F4, and 50% of sequences were removed for U3. Regardless of the proportions of sequences removed, rarefaction analysis indicated that sufficient numbers of clones were sequenced to account for most of the *phoX* diversity, except in F1 (Table 3 and discussed below).

Deduced PhoX amino acid sequences from WCA-2A shared between 85 and 99% sequence similarities with PhoX from cultured and uncultured bacteria (Fig 1). Results from the Maximum Likelihood tree showed that the WCA-2A PhoX sequences clustered into both PhoX Types I and II, as defined by Zaheer et al. (2009). The majority of sequences from WCA-2A clustered with Type II PhoX. Coverages at the 98% cutoff were 54.5% for F1, 71.4% for F4, and 87.5% for U3 (Table 3). Rarefaction curves for F4 and U3 approached an asymptote, but F1 was not close to an asymptote (Figure S3). This suggests that the sampling intensity was insufficiently deep to capture the true diversity of *phoX* at the F1 site. Estimates of Chao richness and Shannon Diversity were greatest at the F1 site and lowest in the

F4 site (Table 3). It is likely that the high diversity at this site is related to the high number of bacterial cells at the site, as indicated by the higher 16S rRNA values at this P-enriched site. Recent fertilization studies have shown that P enrichment can increase the diversity of P mineralizing groups in pasture soils (Tan et al., 2013), such that the greater *phoX* diversity in F1 may have simply been due to the enrichment of bacterial numbers at the P enriched sites.

The ecological significance of the greater abundance of Type II PhoX is currently unknown, and warrants further investigation. Genotypes of *phoX* are widely distributed among the Bacteria (Fig. 1); however, the phylogeny of *phoX* is not well defined at this time, particularly in peatland soils. The majority of sequences obtained in this study cluster with the alphaproteobacteria, although their affiliation to the level of genus is not possible at this time. Of note, a recent metagenomic analysis of these samples indicates that the proteobacteria represent the dominant Bacterial class, with the alphaproteobacteria contributing a significant amount to the total representation of Bacterial orders (Morrison and Ogram, in preparation). The alphaproteobacteria include members that are facultative anaerobes, consistent with what would be expected of a wetland with a fluctuating water table. Further sequencing and characterization of the bacterial communities along this gradient is being undertaken to further elucidate the significance of phosphatase gene phylogenies (of both *phoX* and *phoD*) at these sites.

3.5 Functional Microarray Analysis

The functional microarray GeoChip 3.0 includes 27,812 probes directed toward 292 functional genes, including genes associated with biogeochemical cycling, pollutant metabolism, and energy generation (He et al., 2010). Our analysis in this study focused on those groups of genes involved in P-metabolism (Fig. 2; Table S1): phn; phy; ppk; and ppx. These data represent relative abundances based on signal intensities that were normalized to the total intensities determined for the arrays for each sample. The group of genes included within phn includes those associated with phosphonate metabolism; phosphonates belong to a class of organophosphorus compounds characterized by a C-P bond. The group of genes described by phy are involved in the production of phytase, an enzyme that metabolizes the plant P-storage polymer, phytic acid (inositol hexaphosphate). Genes included in the group ppk include those that are associated with polyphosphate kinase, the enzyme that catalyzes the polymerization of polyphosphate, a common phosphate storage polymer. Those genes in the group ppx are associated with exopolyphosphatase, an enzyme that depolymerizes polyphosphate.

It should be noted that the samples used for GeoChip 3.0 analysis were collected three years prior to those collected for the other analyses reported here (2009 vs. 2013). However, water chemistry data were compared for all sites from 2009 to 2013, and no large changes occurred at these sites within the time period studied (Newman, unpublished data). There can be seasonal differences with respect to the distribution of microbial functional genes related to methanogenesis

and sulfate reduction (Castro et al., 2005; Bae et al., in press) within the sites; however, the sites change consistently with the seasonal impacts.

In general, F4 harbored much greater numbers of unique functional genes (780 unique genes) detected by GeoChip 3.0 than either the F1 or U3 sites, with 148 and 98 unique genes, respectively. The F4 site also exhibited the greatest abundance of the P cycling genes detected by the GeoChip (Fig. 2), with the exception of phytase, which was only detected in the U3 site. It is possible that the nutrient status of this site, which is considered to be a transition point between N and P limitation (Corstanje et al., 2007), allows for increased functional diversity. This functional diversity includes a greater suite of metabolic pathways, including those for the use of alternate P forms, such as phosphonates. Additionally, the high microbial biomass P (Corstanje et al., 2007) at the F4 site could be maintained by greater polyphosphate storage and turnover, as evidenced by the higher *ppk* and *ppx* gene abundances. Interestingly, previous work at these sites only documented pyrophosphate (a polyphosphate of $n=2$) at the F4 site (Turner and Newman, 2005) and did not report evidence of long-chain polyphosphates at any of the study sites. This suggests that polyphosphate turnover may be occurring rapidly at these sites, thereby limiting the ability to detect these compounds with ^{31}P NMR.

The greater abundance of functional genes in F4 is consistent with the higher enzyme activities observed for this site (Table 2); however, these results are in contrast to the estimated copy numbers of the genes encoding high efficiency alkaline phosphatases, *phoX* and *phoD* (Table 3), which were highest in F1. It is possible that these differences may be due to changes between the sites from 2009

and 2013, Water chemistry data for all sites from 2009 to 2013 were stable; however, there may have been changes in other factors that could have effected changes in microbial gene abundance.

The greater relative numbers of total unique functional genes and P cycling genes detected by GeoChip 3.0 at the F4 site is consistent with previous work showing the greatest microbial biomass and metabolic efficiency at the F4 site (Corstanje et al., 2007). Despite the diversity of functional genes in F4, the alternate P pathway encoded by *phy*, was only found in the U3 site. However, it was found in low abundance, and its ecological significance at the site still remains unknown and warrants further investigation.

The Geochip data at each site suggested that there are more types of P cycling genes within the unenriched U3 site; however, the microarray only contained probes for four categories of P cycling genes (*phn*, *phy*, *ppk*, and *ppx*) and as such, was limited in its capacity to detect the potentially wide variety of P cycling genes that might be present at our sites (He et al., 2010). Therefore, further analysis of P cycling genes, such as the phosphatase genes presented in this paper, is necessary to broaden our detection and understanding of microbial P cycling genes along P gradients.

4. CONCLUSIONS

The findings presented here provide evidence of microbial community response at the genetic level to changes in P-limitation. The relationships seen between available P forms and genes that encode specific enzymes responsible for

P-acquisition, in conjunction with related enzyme assays, indicate that P enrichment effects changes in the functional diversity of the Everglades' resident microbial community with respect to P-cycling. The phylogenetic diversity of phosphatase (*phoX*) genes, and the functional diversity shown with GeoChip analysis, indicates that nutrient shifts along the WCA-2A gradient may impact the diversity of P-cycling genes.

The multi-faceted approach applied here indicated a complex response of microbial communities in WCA-2A to shifting nutrient limitations. The compositions and activities of the microbial communities were dependent on their specific location along the nutrient gradient, and yielded insight into some of the fundamental mechanisms through which microbial communities adapt to shifting nutrient limitations, and how microbial communities adapt to anthropogenic nutrient enrichment.

The P-gradient in WCA-2A revealed several characteristics of the microbial communities' responses to shifting P availabilities and limitations. Relatively high ratios of phosphatase to nitrogen acquisition at both the genetic and enzyme activity levels suggest greater selection for P acquisition relative to N acquisition in low-P sites. These findings further support the concept that anthropogenic P inputs are effecting a shift towards N limitation within the high-P sites. In the low-P sites (U3 and E5), microbial communities appeared to be subject to conditions that fell below a threshold P concentration. The high enzyme ($E_P:E_N$) and gene ratios (*phoX:nifH*, *phoD:nifH*) observed at E5 relative to the other sites suggest that soil microbial selection for P acquisition relative to N acquisition may be subject to threshold P

concentrations, above which enzymatic and genetic responses may not be detected. However, experimental manipulation of threshold P concentrations, combined with stoichiometric analyses, would be of value to confirm this trend.

At the transition (F4) site, an increase in bacterial functional diversity was observed, including increases in the numbers of genes associated with mineralization of alternative P forms, such as phosphonates (phn), and increased numbers of genes related to polyphosphate storage (ppk) and cycling. The relative numbers of total genes detected and genes associated with phosphate cycling by the functional microarray GeoChip 3.0, and phosphatase and phosphodiesterase enzyme activities, were highest in F4. In addition to the enzyme activities related to organic phosphate attack, leucine aminopeptidase activities were highest at this site. These data support previous work that found the greatest microbial activities were at the F4 site, likely due to the site being an intermediate between P and N limitations (Corstanje et al., 2007). The structures and functions of microbial communities in transition from one limiting nutrient to another warrants further study.

In the P-enriched sites (F1 and F3), high P concentrations were associated with an increase in bacterial numbers (represented by 16S rRNA gene copy numbers), resulting in relatively proportional increases in *phoX* copy numbers and *phoX* diversity.

Shifts in concentrations of genes encoding phosphatases relative to genes involved in N fixation were correlated with bicarbonate extractable P_o concentrations within these sites. These findings were supported by analysis of the

activities of phosphatase, phosphodiesterase, and leucine aminopeptidase along this gradient. The agreement of genetic and enzymatic data suggests that assessing microbial nutrient demands with molecular techniques is feasible, although future work is needed to develop the technique, particularly with additional experimental manipulation, to elucidate the complexity of these linked biogeochemical cycles and provide a more mechanistic understanding of microbial P cycling within this unique peatland.

5. ACKNOWLEDGEMENTS

This work was partially funded by a grant from the National Science Foundation (DEB 0841596). Thanks to Kristin Seitz, South Florida Water Management District for conducting enzyme assays.

REFERENCES

- Arrigo, K.R., 2005. Marine microorganisms and global nutrient cycles (vol 437, pg 349, 2005). *Nature* 438(7064), 122-122.
- Bae, H.-S., Dierberg, F.E., Ogram, A., 2014. Syntrophs dominate sequences associated with the mercury methylation-related gene *hgcA* in the Water Conservation Areas of the Florida Everglades. *Applied and Environmental Microbiology* 80(20), 6517-6526.
- Bae, H.-S., Holmes, E., Chanton, J., Reddy, K.R., and Ogram, A. Distribution, activities, and interactions of methanogens and sulfate reducing prokaryotes in the Florida Everglades. *Appl. Environ. Microbiol.*, in press.
- Belanger, T.V., Scheidt, D.J., Platko, J.R., II, 1989. Effects of nutrient enrichment on the Florida Everglades. *Lake Reservoir Manage.* 5(1), 101-112.
- Castro, H., Reddy, K.R., Ogram, A., 2002. Composition and function of sulfate-reducing prokaryotes in eutrophic and pristine areas of the Florida Everglades. *Applied and Environmental Microbiology* 68(12), 6129-6137.
- Castro, H., Newman, S., Reddy, K.R., Ogram, A., 2005. Distribution and stability of sulfate-reducing prokaryotic and hydrogenotrophic methanogenic assemblages in nutrient-impacted regions of the Florida everglades. *Applied and Environmental Microbiology* 71(5), 2695-2704.

651

652 Chauhan, A., Ogram, A., Reddy, K.R., 2004. Syntrophic-methanogenic associations
653 along a nutrient gradient in the Florida Everglades. Applied and Environmental
654 Microbiology 70(6), 3475-3484.

655

656 Chen, C.R., Condron, L.M., Davis, M.R., Sherlock, R.R., 2002. Phosphorus dynamics in
657 the rhizosphere of perennial ryegrass (*Lolium perenne* L.) and radiata pine (*Pinus*
658 *radiata* D. Don.). Soil Biology & Biochemistry 34(4), 487-499.

659

660 Corstanje, R., Reddy, K.R., Prenger, J.P., Newman, S., Ogram, A.V., 2007. Soil microbial
661 eco-physiological response to nutrient enrichment in a sub-tropical wetland.
662 Ecological Indicators. 7(2), 277-289.

663

664 DeLong, E.F., Karl, D.M., 2006. Genomic perspectives in microbial oceanography (vol
665 437, pg 336, 2005). Nature 439(7079), 1014-1014.

666

667 Fraser, T., Lynch, D.H., Entz, M.H., Dunfield, K.E., 2015. Linking alkaline phosphatase
668 activity with bacterial *phoD* gene abundance in soil from a long-term management
669 trial. Geoderma 257/258, 115-122.

670

671 Freeman, C., Evans, C.D., Monteith, D.T., Reynolds, B., Fenner, N., 2001a. Export of
672 organic carbon from peat soils. Nature 412(6849), 785-785.

673

674 Freeman, C., Ostle, N., Kang, H., 2001b. An enzymic 'latch' on a global carbon store -
675 A shortage of oxygen locks up carbon in peatlands by restraining a single enzyme.
676 Nature 409(6817), 149-149.
677

678 Gusewell, S., Freeman, C., 2005. Nutrient limitation and enzyme activities during
679 litter decomposition of nine wetland species in relation to litter N : P ratios.
680 Functional Ecology 19(4), 582-593.
681

682 Handelsman, J., 2004. Metagenomics: Application of genomics to uncultured
683 microorganisms. Microbiology and Molecular Biology Reviews 68(4), 669-+.
684

685 Hagerthey, S.E., Newman, S., Rutchey, K., Smith, E.P., Godin, J., 2008. Multiple regime
686 shifts in a subtropical peatland: community-specific thresholds to eutrophication.
687 Ecol. Mono. 78(4), 547-565.
688

689 Harms, G., Layton, A.C., Dionisi, H.M., Gregory, I.R., Garrett, V.M., Hawkins, S.A.,
690 Robinson, K.G., Sayler, G.S., 2003. Real-time PCR quantification of nitrifying bacteria
691 in a municipal wastewater treatment plant. Environ. Sci. Technol. 37(2), 343-351.
692

693 He, Z., Deng, Y., Van Nostrand, J.D., Tu, Q., Xu, M., Hemme, C.L., Li, X., Wu, L., Gentry,
694 T.J., Yin, Y., Liebich, J., Hazen, T.C., Zhou, J., 2010. GeoChip 3.0 as a high-throughput
695 tool for analyzing microbial community composition, structure and functional
696 activity. ISME J. 4(9), 1167-1179.

697

698 Heckmann, L.-H., Sorensen, P.B., Krogh, P.H., Sorensen, J.G., 2011. NORMA-Gene: A
699 simple and robust method for qPCR normalization based on target gene data. BMC
700 Bioinform. 12, 250.

701

702 Holmes, M.E., Chanton, J.P., Bae, H.S., Ogram, A., 2014. Effect of nutrient enrichment
703 on delta(CH₄)-C-13 and the methane production pathway in the Florida Everglades.
704 J. Geophys. Res. Biogeosci. 119(7), 1267-1280.

705

706 Ivanoff, D.B., Reddy, K.R., Robinson, S., 1998. Chemical fractionation of organic
707 phosphorus in selected histosols. Soil Sci. 163(1), 36-45.

708

709 Kang, H.J., Freeman, C., 1999. Phosphatase and arylsulphatase activities in wetland
710 soils: annual variation and controlling factors. Soil Biology & Biochemistry 31(3),
711 449-454.

712

713 Limpens, J., Berendse, F., Blodau, C., Canadell, J.G., Freeman, C., Holden, J., Roulet, N.,
714 Rydin, H., Schaepman-Strub, G., 2008. Peatlands and the carbon cycle: from local
715 processes to global implications - a synthesis. Biogeosciences 5(5), 1475-1491.

716

717 McCormick, P.V., Newman, S., Vilchek, L.W., 2009. Landscape responses to wetland
718 eutrophication: loss of slough habitat in the Florida Everglades, USA. Hydrobiologia
719 621, 105-114.

720 McCray, J.M., Wright, A.L., Luo, Y., Ji, S., 2012. Soil Phosphorus Forms Related to
 721 Extractable Phosphorus in the Everglades Agricultural Area. *Soil Science* 177(1), 31-
 722 38.
 723
 724 Mehta, M.P., Butterfield, D.A., Baross, J.A., 2003. Phylogenetic diversity of
 725 nitrogenase (*nifH*) genes in deep-sea and hydrothermal vent environments of the
 726 Juan de Fuca ridge. *Appl. Environ. Microbiol.* 69(2), 960-970.
 727
 728 Penton, C.R., Newman, S., 2007. Enzyme activity responses to nutrient loading in
 729 subtropical wetlands. *Biogeochemistry* 84(1), 83-98.
 730
 731 Penton, C.R., Newman, S., 2008. Enzyme-based resource allocated decomposition
 732 and landscape heterogeneity in the Florida Everglades. *J. Environ. Qual.* 37(3), 972-
 733 976.
 734
 735 Reddy, K., and DeLaune, R.D., 2008. *Biochemistry of Wetlands: Science and*
 736 *Applications*. First ed. CRC Press, London.
 737
 738 Sakurai, M., Wasaki, J., Tomizawa, Y., Shinano, T., Osaki, M., 2008. Analysis of
 739 bacterial communities on alkaline phosphatase genes in soil supplied with organic
 740 matter. *Soil Sci. Plant Nut.* 54(1), 62-71.
 741

742 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B.,
 743 Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B.,
 744 Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source,
 745 platform-independent, community-supported software for describing and
 746 comparing microbial communities. *Appl. Environ. Microbiol.* 75(23), 7537-7541.
 747
 748 Sebastian, M., Ammerman, J.W., 2009. The alkaline phosphatase PhoX is more widely
 749 distributed in marine bacteria than the classical PhoA. *ISME J.* 3(5), 563-572.
 750
 751 Tan, H., Barret, M., Mooij, M.J., Rice, O., Morrissey, J.P., Dobson, A., Griffiths, B., O'Gara,
 752 F., 2013. Long-term phosphorus fertilisation increased the diversity of the total
 753 bacterial community and the *phoD* phosphorus mineraliser group in pasture soils.
 754 *Biology and Fertility of Soils* 49(6), 661-672.
 755
 756 Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5:
 757 Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary
 758 Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* 28(10), 2731-2739.
 759
 760 Tiwari, B., Singh, S., Kaushik, M.S., Mishra, A.K., 2015. Regulation of organophosphate
 761 metabolism in cyanobacteria. A review. *Microbiology* 84(3), 291-302.
 762
 763 Torsvik, V., Ovreas, L., 2002. Microbial diversity and function in soil: from genes to
 764 ecosystems. *Current Opinion in Microbiology* 5(3), 240-245.

765

766 Turner, B.L., Hopkins, D.W., Haygarth, P.M., Ostle, N., 2002. Beta-glucosidase activity
 767 in pasture soils. *Appl. Soil Ecol.* 20(2), 157-162.

768

769 Turner, B.L., Cade-Menun, B.J., Condon, L.M., Newman, S., 2005. Extraction of soil
 770 organic phosphorus. *Talanta* 66(2), 294-306.

771

772 Turner, B.L., Newman, S., 2005. Phosphorus cycling in wetland soils: The importance
 773 of phosphate diesters. *J. Environ. Qual.* 34(5), 1921-1929.

774

775 Uz, I., Chauhan, A., Ogram, A.V., 2007. Cellulolytic, fermentative, and methanogenic
 776 guilds in benthic periphyton mats from the Florida Everglades. *Fems Microbiology*
 777 *Ecology* 61(2), 337-347.

778

779 Vetrovsky, T., Baldrian, P., 2013. The variability of the 16S rRNA gene in bacterial
 780 genomes and its consequences for bacterial community analyses. *PLOS One* 8(2), 10.

781

782 Zaheer, R., Morton, R., Proudfoot, M., Yakunin, A., Finan, T.M., 2009. Genetic and
 783 biochemical properties of an alkaline phosphatase PhoX family protein found in
 784 many bacteria. *Environ. Microbiol.* 11(6), 1572-1587.

785

786 Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S.G., Alvarez-Cohen, L., 2015. High-
787 Throughput Metagenomic Technologies for Complex Microbial Community Analysis:
788 Open and Closed Formats. *Mbio* 6(1).

FIGURE LEGENDS

Figure 1: Figure 2: Phosphorus cycling gene abundance detected via Geochip 3.0 functional microarray. Genes for *phn* (phosphonate utilization), *phy* (phytase), *ppk* (polyphosphate kinase), *ppx* (exopolyphosphatase) were detected (values from composited triplicate field replicates)

Figure 2: *phoX* maximum likelihood tree. Numbers at nodes represent bootstrap values for 1000 re-samplings. Sequences obtained from this study are denoted by site (F1, F4, U3).

1
2
3
4

5
6

Table 1: Phosphorus fractions found at each site (values from composited triplicate field replicates)

Availability	Phosphorus fraction	µg g ⁻¹				
		F1	F3	F4	U3	E5
Readily labile	NaHCO ₃ P _i	29	30	3	3	0
	NaHCO ₃ P _o	42	55	32	41	7
	Microbial Biomass P	504	399	394	36	157
Moderately labile	HCl P _i	297	209	221	33	75
	Fulvic acid P	123	239	163	20	52
Nonlabile	Humic Acid P	409	268	145	87	84
	Residual P	135	73	51	22	37
	Total P	1539	1272	1010	242	411
	Total P _i	325	239	224	36	76
	Total P _o	1214	1033	785	205	336

Table 2: Enzyme activities at each site (values from composited triplicate field replicates)

Enzyme	Enzyme Activity ($\mu\text{mol g afdm}^{-1} \text{ h}^{-1}$)				
	F1	F3	F4	U3	E5
Phosphatase	0.32	0.07	3.80	2.66	3.42
Phosphodiesterase	0.29	0.07	5.10	2.23	3.34
β -glucosidase	0.28	0.24	0.27	0.01	0.08
Leucine aminopeptidase	0.27	0.47	2.82	0.13	0.28
Phenol oxidase	240	353	360	541	475
Peroxidase	104	22	0	1	0

Table 3: Mean gene abundance from three cores (standard deviation), gene and enzyme ratios, and diversity indices along the transect

Site	Gene abundance 10 ¹⁰ copy number g soil ⁻¹				<i>phoX:16S</i> <i>rRNA</i>	<i>phoD:16S</i> <i>rRNA</i>	<i>phoX:nifH</i>	<i>phoD:nifH</i>	E _P :E _N	<i>phoX</i> Chao Richness	<i>phoX</i> Shannon Diversity	<i>phoX</i> ACE	<i>phoX</i> Coverage (%)
	16S <i>rRNA</i>	<i>phoX</i>	<i>phoD</i>	<i>nifH</i>									
F1	268 (38)	0.39 (0.06)	0.29 (0.05)	1.2 (0.54)	0.0015	0.0011	0.32	0.24	2.26	192	2.8	110. 2	54.5
F3	410 (50)	0.53 (0.17)	0.1 (0.05)	1.5 (0.50)	0.0013	0.0002	0.36	0.07	0.30	NA	NA	NA	NA
F4	83 (37)	0.15 (0.04)	0.07 (0.07)	0.65 (0.21)	0.0018	0.0008	0.23	0.10	3.16	9	1.6	5.8	71.4
U3	44 (11)	0.11 (0.02)	0.02 (0.004)	0.21 (0.03)	0.0026	0.0005	0.55	0.10	37.62	25.5	2.3	7.2	87.5
E5	1.2 (1.6)	0.04 (0.02)	0.63 (0.83)	0.01 (0.002)	0.0314	0.5449	4.18	72.45	24.14	NA	NA	NA	NA

1
2 Figure 1: *phoX* maximum likelihood tree. Numbers at nodes represent bootstrap
3 values for 1000 re-samplings. Sequences obtained from this study are denoted by
4 site (F1, F4, U3). Numbers in parentheses indicate numbers of unique sequences
5 obtained within the clades.
6
7

8

9

10

11 Figure 2: Phosphorus cycling gene abundance detected via Geochip 3.0 functional
 12 microarray. Genes for *phn* (phosphonate utilization), *phy* (phytase), *ppk*
 13 (polyphosphate kinase), *ppx* (exopolyphosphatase) were detected (values from
 14 composited triplicate field replicates)