THE ROLE OF POLYPHOSPHATE ACCUMULATING ORGANISMS IN ENVIRONMENTAL PHOSPHORUS CYCLING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sheila M. Saia

August 2017

THE ROLE OF POLYPHOSPHATE ACCUMULATING ORGANISMS IN ENVIRONMENTAL PHOSPHORUS CYCLING

Sheila M. Saia, PhD

Cornell University 2017

ABSTRACT

Phosphorus (P) is a limiting nutrient in freshwater ecosystems and excess P from anthropogenic sources impairs water quality. Strategies to manage P pollution depend on abiotic and well as biotic mechanisms. For example, no till farming practices physically reduce transport of sediment-bound P into nearby waterbodies and specialized wastewater treatment plants (WWTPs) utilize biological (i.e., microbial) mechanisms to remove P from influent waters. With respect to the latter, the main actors of these specialized WWTPs are a group of organisms known as polyphosphate accumulating organisms (PAOs). PAOs are well studied in the context of WWTPs and a limited number of studies have identified them in the natural environment. However, very little is known about their ecological role as well as their influence on P cycling and transport in natural systems. Therefore, the overall goal of this work was to expand our understanding of PAOs in soils and streams.

We started this exploration with a review of PAOs in engineered and natural systems—we discussed knowledge gaps and ways studies from these distinct contexts may build on one another. This review also included a discussion of the potential role of PAOs in agricultural systems. We proposed studies to explore the impacts of PAOs in terms of major agricultural

challenges and discussed how these studies may inform our understanding of PAOs in engineered and natural systems. Next, we conducted a laboratory experiment to explore the role of PAO-mediated P cycling in stream biofilms under alternating aerobic and anaerobic conditions. We demonstrated cyclical patterns of P uptake under aerobic conditions and release during anaerobic conditions, which is consistent with the known behavior of PAOs in WWTPs. We also verified larger percentages of cells with stored intracellular polyphosphate granules under aerobic conditions compared to anaerobic conditions, which we expected given our understanding of the PAO phenotype in WWTPs. However, we observed concurrent patterns in cation uptake and release, which may indicate abiotic precipitation/dissolution of P with these cations and/or biotic uptake/release of these cations to balance the negative charge of intracellular polyP. Next, we explored whether the soil wetness index (SWI), a static index used to predict landscape scale soil moisture patterns, can predict the occurrence of mobile forms of P as well as PAO associated functional genes (i.e., *ppk1*, *ppk2*, and *ppx*). We showed that SWI predicted mobile P (i.e., dilute $CaCl_2$ extractable P) and there was a depletion of mobile P from wetter parts of the landscape. This is consistent with our expectations of PAO behavior; PAOs release mobile forms of P under saturated conditions that are transported off-site. More specifically related to PAOs, we found that SWI was not a good predictor of the relative abundance of polyP functional genes. We observed a general decrease in the relative abundance polyP functional genes versus mobile P concentrations in NY, was consistent with our hypothesis, but this trend was only statistically significant in the case of *ppk2*. In PA, the relationship between the relative abundance of polyP functional genes was not significant and general trends were inconsistent with our hypothesis. Therefore, these results suggest the limited role of PAO-mediated P cycling along the SWI gradients identified and the potential masking of

PAOs by other P controls (e.g. landscape position and management). Future research may consider how the role of biotic and abiotic processes masks the role of PAOs in soil P cycling. These includes the impact of iron reducing bacteria or chemical iron reduction/dissolution with P along a SWI gradient. Despite this limited support for PAOs, we identified contigs harboring both *ppk1* and *ppx* genes that were within the same phyla as known PAOs as well as many unstudied, putative PAOs. Last, we studied whether PAOs played a discernable role in P cycling associated with the decomposition of leaf litter in the stream and on the forest floor. We observed an increase in leaf P concentrations in the stream and a decrease in leaf P concentrations on the forest floor over time. Unexpectedly, we did not observe a concurrent increase in the relative abundance of PAO-associated functional genes over time in the stream. Rather these genes remained constant. *ppk1* and *ppx* relative abundances also remained constant in the forest floor but the relative abundance of *ppk2* genes increased over time. While these trends did not provide support for PAOs control on P cycling in leaf litter decomposition, we identified contigs harboring both *ppk1* and *ppx* genes that were within the same phyla as known PAOs as well as many unstudied, putative PAOs just as we observed in the soil study. In the case of both soil and leaf litter experiments, future studies may consider using microscopy and molecular biology tools to verify whether putative PAOs exhibit the phenotype of PAO established in engineered systems.

Overall, we found support for PAO-mediated P cycling in stream biofilms but only limited support for their impact on P cycling in soils along a SWI gradient as well as on decomposing leaf P patterns over time. We provided additional thoughts at the end of the review paper and in chapter five on how these studies can be modified to test for the potential role of PAOs in

agricultural P management as well as the influence of hydrology and nutrient demand/availability on PAO-mediated P cycling.

BIOGRAPHICAL SKETCH

Sheila Marie Saia was born in Buffalo, New York on August 21, 1985. She graduated in 2007 with a Bachelor's of Science degree in Bioengineering from the University of Binghamton. In the autumn of 2007, she headed to Chiba, Japan where she taught English at a private school for the next year. Upon returning home to the US, she accepted an Americorps sponsored volunteering coordinator position at Buffalo Niagara Riverkeeper. Her positive experience at Buffalo Niagara Riverkeeper motivated her to pursue an advanced degree in environmental engineering. She received her Masters of Science degree in August 2012 from the Department of Biological and Environmental Engineering at Cornell University and stayed on in this same department for her PhD. Sheila is dedicated to improving water quality through research and community involvement and has made a point to do so throughout her tenure as a graduate student. Specifically, she served as a member of the Hydrology Section Student Subcommittee (H3S) of the American Geophysical Union (AGU) where she helped organize and promote leadership training for student members. She mentored high school and undergraduate students on independent research projects and volunteered with Cornell Cooperative Extension in Auburn, NY to promote community education and involvement around local water quality issues caused by anthropogenically sourced phosphorus. She developed science workshops to educate middle school girls and first generation college students about how humans influence water quality and quantity. Most recently, she contributes articles to the FeminaSci blog (www.feminasci.com). When she is not studying hydrology, biogeochemistry, and microbes, she enjoys spending time with friends, cooking, rock climbing, bicycling, and crafting (in that order).

DEDICATION

For my grandmother, Mary Jane Hensen.

ACKNOWLEDGEMENTS

Thanks to my PhD committee members—Todd Walter, Jay Regan, Patrick Sullivan, and Ruth Richardson—for their support and mentorship. I would also like to thank the other researchers who provided comments on my thesis including: Tony Buda, Hunter Carrick, Katy Hofmeister, Claudia Rojas, Miranda Stockton, and Ed Hall. Many thanks to my partner, Chuck Pepe-Ranney, for his help with sequence data analysis as well as his love, pep-talks, and for joining me on weekend climbing breaks. Thank you to my parents for encouraging me to pursue an advanced degree in the sciences and appreciate all things hydrologic. Thanks to my sister for her honesty, support, and for inspiring me to be creative. Thank you to my grandparents—Mary Jane and Harold Hensen as well as James and Norma Saia—who nurtured a tightknit family and taught me how to work hard. Thanks to my Great Aunt Rita Bartlett for her never-ending kindness and to the rest of my family in Buffalo, NY for their love. Thanks to the many friends and peers in the US and abroad who have supported me both before and during my time at Cornell. Thank you to my SMEO friends. I am especially grateful to have worked with and befriended Chelsea Morris, who was always willing to listen to my ideas and provide constructive feedback. Thanks to my peers in the Soil & Water Lab. I could not have completed this work without your feedback, friendship, and help in the lab and field. Thank you to Shree Giri for teaching me all I know about proper laboratory technique and always being there to lend a hand with the phosphorus analyzer. Thanks to all my climbing friends who always knew when I needed to take a break from research. Thank you to my fellow FeminaSci members for encouraging me explore what it means to be a female scientist in 2017. I would like to thank my best friend Shelley Wilson for always being game for an adventure and reminding me to be courageous.

TABLE OF CONTENTS

ABSTRACT 3
BIOGRAPHICAL SKETCH 7
DEDICATION 8
ACKNOWLEDGEMENTS 9
TABLE OF CONTENTS 10
LIST OF FIGURES AND TABLES 12
CHAPTER 1 21

LINKING WASTEWATER ENGINEERING AND THE NATURAL SCIENCES TO IMPROVE OUR UNDERSTANDING OF POLYPHOSPHATE ACCUMULATING ORGANISM (PAO)-MEDIATED PHOSHPHORUS CYCLING: A REVIEW 21
ABSTRACT 21
INTRODUCTION 22
BIOTIC CONTROLS ON PHOSPHORUS CYCLING IN ENGINEERED SYSTEMS 24
BIOTIC CONTROLS ON PHOSPHORUS CYCLING IN NATURAL SYSTEMS 47
CONSIDERING BIOTIC CONTROLS IN AGRICULTURAL SYSTEMS 59
SUMMARY 66
ACKNOWLEDGMENTS 67
REFERENCES 68
CHAPTER 2 89

EVIDENCE FOR POLYPHOSPHATE ACCUMULATING ORGANISMS (PAO)-MEDIATED PHOSHPHORUS CYCLING IN STREAM BIOFILMS UNDER ALTERNATING AEROBIC/ANAEROBIC CONDITIONS 89
ABSTRACT 89
INTRODUCTION 90
METHODS 94
RESULTS 102
DISCUSSION 110
CONCLUSIONS 117
ACKNOWLEDGEMENTS 118
REFERENCES 120

CHAPTER 3 125

THE IMPACT OF HYDROLOGY AND SOIL PHOSPHORUS CONCENTRATIONS ON THE RELATIVE ABUNDANCE OF POLYPHOSPHATE CYCLING GENES 125

ABSTRACT 125

INTRODUCTION 126

METHODS 132

RESULTS & DISCUSSION 143

CONCLUSIONS 193

ACKNOWLEDGMENTS 193

REFERENCES 195

CHAPTER 4 204

TEMPORAL TRENDS OF DECOMPOSING LEAF LITTER PHOSPHORUS CONCENTRATIONS AND THE RELATIVE ABUNDANCE OF POLYPHOSPHATE CYCLING GENES 204

ABSTRACT 204

INTRODUCTION 205

METHODS 210

RESULTS & DISCUSSION 217

CONCLUSIONS 246

ACKNOWLEDGMENTS 246

REFERENCES 247

CHAPTER 5 254

CONCLUSIONS 254

FUTURE DIRECTIONS 257

REFERENCES 262

LIST OF FIGURES AND TABLES

Figure 1. 1 Schematic of research flows between engineered, natural, and agricultural systems disciplines with respect to biotic controls on P cycling. Arrows between each system represent limited (grey) and potential (white, dashed outline) opportunities for information sharing. 23

Figure 1. 2 (A) Anaerobic and (B) aerobic metabolism of model EBPR PAO, *Candidatus* Accumulibacter phosphatis (adapted from Seviour et al. 2003, Seviour and Nielsen 2010, and Skennerton et al. 2014). M^+ represents metal cations and P_i represents phosphate..... 30

Figure 2. 1 A. Location of the study site in the Cascadilla Creek Watershed in Tompkins County, New York. B. Photograph of Cascadilla Creek. C.—Stream cobbles in sealed containers with native stream water used in the experiment. 95

Figure 2. 2 Experimental design schematic for treatment 1 (T1) and treatment 2 (T2). Environmental measurements and water samples collection times are shown as vertical lines. See the methods section for further description. 97

Figure 2. 3 Observed (points) and modeled (lines) PO_4^{3-}-P concentrations in the surrounding water normalized by mass of wet biofilm in the tub as a function of time for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). Black bar indicates periods when T1 was anaerobic. Unmarked periods for T1 were aerobic. Samples at 0 h were taken in the stream and the remaining samples were taken in the laboratory. 103

Figure 2. 4 Observed (points) and modeled (lines) concentrations of Fe^{2+} (A), total S (B), total Mn (C), Ca (D), K (E), and Mg (F) in the surrounding water normalized by mass of wet biofilm in the tub as a function of time for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). Black bar indicates periods when T1 was anaerobic. Unmarked periods for T1 were aerobic. Samples at 0 h were taken in the stream and the remaining samples were taken in the laboratory..... 107

Figure 2. 5 Percentage of biofilm cells (n = 20 fields/treatment) with stored polyphosphate (polyP) at the end of the experiment in treatment 1 (anaerobic) and treatment 2 (aerobic). B.—Microscope image from a T2 biofilm sample at the end of the experiment and visualized with a 4',6-diamindino-2-phenylindole (DAPI)–DNA filter where light blue indicates the presence of DNA. C.—Microscope image taken a T2 biofilm sample at the end of the experiment and visualized with a DAPI–polyP filter where yellow indicates the presence of polyP. Arrow indicates an example cell with stored polyP granules. 109

Figure 2. 6 Observed (points) and modeled (lines) pH of the surrounding water as a function of time for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). 112

Figure 3. 1 Monthly precipitation at New York (NY) and Pennsylvania (PA) field sites for (A) 2015—the year of this study—and (B) 30-year normals. 134

Figure 3. 2 Description of shotgun metagenomic sequencing data workflow used in this study. Final output tables used in subsequent analyses are bolded/shaded green. 141

Figure 3. 3 (A) Soil moisture, measured as percent volumetric moisture content (% VMC), aggregated for all sample months versus location. (B) Soil moisture versus sample month for New York (NY) and Pennsylvania (PA). 144

Figure 3. 4 Soil moisture, measured as percent volumetric moisture content (% VMC), versus soil wetness index for New York (NY) and Pennsylvania (PA) locations taken during each sample month. Best fit linear mixed effects model (LMEM) results (Model 5.4 in Table 3.5) shown as a black line. 147

Figure 3. 5 Phosphorus concentration (mg/kg) versus extraction type for New York (NY) and Pennsylvania (PA). In-set (top left-hand corner) shows a zoomed-in view of $CaCl_2$-P extraction results. 152

Figure 3. 6 $CaCl_2$-P concentration (mg/kg) versus sample month for (A) New York (NY) and (B) Pennsylvania (PA). Note difference in y-axis scale. 153

Figure 3. 7 Phosphorus concentration of $CaCl_2$-P (A and B), Ox-P (C and D), Total-P (E and F) versus soil wetness index for New York (NY; A, C, E) and Pennsylvania (PA; B, D, F). Best fit linear mixed effects model (LMEM) results (Model 9.2, Table 3.9) shown as a black line. 157

Figure 3. 8 Number of unique contigs with various gene combinations in samples from (A) New York (NY) and (B) Pennsylvania (PA). Note: the total number of samples sequenced in NY (n=30) was double the total number of samples sequenced in PA (n=15). 162

Figure 3. 9 Length distributions of contigs harboring various gene combinations from samples taken in New York (NY; A and C) and Pennsylvania (PA; B and D). Note the difference in scale between A and B. The shaded regions in plots A and B refer to the zoomed-in (0-5,000 bp) plots shown in C and D. 164

Figure 3. 1 0 Number of unique contigs classified into major bacterial and archaeal phyla/classes for (A) New York (NY) and (B) Pennsylvania (PA). 168

Figure 3. 1 1 Number of unique contigs classified into major bacterial and archaeal phyla/classes for New York (NY; A-C) and Pennsylvania (PA; D-E) harboring only one of either *ppk1* (A and D), *ppk2* (B and E), and *ppx* (C and F) genes. 169

Figure 3.12 Number of unique contigs harboring both *ppk1* and *ppx* genes classified into major bacterial and archaeal phyla/classes for (A) New York (NY) and (B) Pennsylvania (PA). 172

Figure 3.13 Relative gene abundance (FPKM) of *ppk1* (A and B), *ppk2* (C and D), and *ppx* (E and F) versus sample event ordered from lowest to highest SWI and from May to October. New York (NY) results shown in (A), (C), and (E). Pennsylvania (PA) results shown in (B), (D), and (F)..... 176

Figure 3.14 Observed relative abundance (FPKM) colored by sample site versus soil wetness index for in New York (NY; A, C, and E) and Pennsylvania (PA; B, D, and F) for *ppk1* (A and B), *ppk2* (C and D), and *ppx* (E and F) genes. 177

Figure 3.15 Relative gene abundance (FPKM) of (A) *ppk1*, (B) *ppk2,* and (C) *ppx* aggregated by location; whether samples were taken in New York (NY) or Pennsylvania (PA)..... 179

Figure 3.16 Observations and best fit linear mixed effects model (LMEMs, Table 3.16) of relative abundance (FPKM) versus $CaCl_2$-P concentration for (A) *ppk1*, (B) *ppk2,* and (C) *ppx* genes. Note that the only significant ($p<0.05$) trend is for *ppk2* gene results taken from NY. .. 182

Figure 3.17 Observed relative abundance (FPKM) colored by sample site versus $CaCl_2$-P concentration for in New York (NY; A, C, and E) and Pennsylvania (PA; B, D, and F) for *ppk1* (A and B), *ppk2* (C and D), and *ppx* (E and F) genes. 184

Figure 3.18 Heatmap of relative abundance z-scores for *ppk1* genes found in NY samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O). 187

Figure 3.19 Heatmap of relative abundance z-scores for *ppk1* genes found in PA samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O). 188

Figure 3.20 Heatmap of relative abundance z-score for *ppk2* genes found in NY samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O). 189

Figure 3.21 Heatmap of relative abundance z-scores for *ppk2* genes found in PA samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O). 190

Figure 3.22 Heatmap of relative abundance z-scores for *ppx* genes found in NY samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O). 191

Figure 3.23 Heatmap of relative abundance z-scores for *ppx* genes found in PA samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O). 192

Figure 4.1 Map of the Mianus River Watershed..... 211

Figure 4.2 Description of shotgun metagenomic sequencing data workflow used in this study. Final output tables used in subsequent analyses are bolded/shaded green. 215

Figure 4.3 Observations and linear model (LM) results of ash free dry mass (AFDM) versus days from start of experiment for leaf litter bags placed on the forest floor (F) and in the stream (S) treatments. 219

Figure 4.4 Observations and generalized additive model (GAM) results of leaf P concentration (ppm/g AFDM) versus days from start of experiment for leaf litter bags placed on the forest floor and in the stream treatments. 223

Figure 4.5 (A) Rinse water soluble reactive phosphorus (SRP) concentration and (B) leaf P concentration versus soaking time for laboratory wash-off experiment. Samples at ~ 1 min (~0 hours) were rinsed with water and immediately drained. 225

Figure 4.6 Number of unique contigs with various gene combinations in samples from (A) forest floor and (B) stream. Note: the total number of samples sequenced is the same for both treatments (n=16). 228

Figure 4.7 Length distributions of contigs harboring various gene combinations from samples collected from the stream (A and C) and on the forest floor (B and D). Note the difference in scale between A and B. The shaded regions in plots A and B refer to the zoomed-in (0-25,000 bp) plots shown in C and D..... 230

Figure 4.8 Number of unique contigs classified into major bacterial and archaeal phyla/classes for leaf litter samples collected from the (A) forest floor and (B) stream. 232

Figure 4.9 Number of unique contigs classified into major bacterial and archaeal phyla/classes for leaf litter samples collected on the (A) forest floor (A-C) and in the stream (D-E) harboring only one of either *ppk1* (A and D), *ppk2* (B and E), and *ppx* (C and F) genes. 234

Figure 4.10 Number of unique contigs harboring both *ppk1* and *ppx* genes classified into major bacterial and archaeal phyla/classes for leaf litter samples collected from the (A) forest floor and (B) stream. Number of unique contigs per phylum/class are shown to the right of each bar. 235

Figure 4.11 Observations and generalized additive model (GAM) results (Models 8.2, 8.4, and 8.6 in Tables 4.8 and 4.9) of relative abundance (FPKM) of (A) *ppk1*, (B) *ppk2*, and (C) *ppx* genes over time for leaf litter bags collected from the forest floor and stream. 240

Figure 4.12 Heatmap of the relative abundance z-score for *ppk1* genes clustered by similar relative abundance profiles between sample ID and contig ID. Key on top shows treatment type and collection time for each sample. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. 243

Figure 4.13 Heatmap of the relative abundance z-score for *ppk2* genes clustered by similar relative abundance profiles between sample ID and contig ID. Key on top shows treatment type and collection time for each sample. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. 244

Figure 4.14 Heatmap of the relative abundance z-score for *ppx* genes clustered by similar relative abundance profiles between sample ID and contig ID. Key on top shows treatment type and collection time for each sample. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. 245

Table 1. 1 A summary of PAOs found in the literature. Abbreviations: laboratory-scale sequencing batch reactor (SBR), full-scale wastewater treatment plant (WWTP), uncharacterized wastewater (Mixed). 26

Table 1. 2 Proteins associated with PAO-related phosphorus cycling functional genes. 34

Table 1. 3 Fluorescence in-situ hybridization (FISH) probes utilized in the literature for identification of PAO-related organisms (adapted from Seviour et al. 2010). Abbreviations: polyphosphate accumulating organisms (PAO), not determined (ND). 39

Table 1. 4 PAO PCR primer sequences identified in the literature. 41

Table 2. 1 Land cover (LC) percentages for the ungauged study watershed, Cascadilla Creek, and nearby gauged watersheds, Six Mile Creek and Fall Creek. 95

Table 2. 2 Generalized additive model results for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). SE = standard error, TS = total S, TMn = total Mn. 104

Table 2. 3 Linear model results for PO_4^{3-}-P vs concentrations of Ca, K, Mg, total Mn (TMn), Fe^{2+}, and total S (TS) measured in the surrounding water for treatment 1 (T1; alternating anaerobic/aerobic conditions). SE = standard error, AIC = Akaike Information Criterion. 106

Table 2. 4 Mean (standard error, n = 2) concentration of polyphosphate (polyP) and total P (TP) per unit mean biofilm surface area at the start and end of the experiment for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). 109

Table 3. 1 Description of New York (NY) and Pennsylvania (PA) study sites. Abbreviations: soil wetness index (SWI). 135

Table 3. 2 Protein family (Pfam) database models used in this study. 142

Table 3. 3 Linear mixed effects model comparisons for soil moisture, measured as percent volumetric moisture content (% VMC). ‘Location’ indicates whether samples were collected in New York (NY) or Pennsylvania (PA) and ‘Month’ indicates the sample month. Best fit models are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC). 144

Table 3. 4 Linear mixed effects model outputs for best fit model (Model 3.3 in Table 3.3) for percent volumetric moisture content (% VMC). ‘Location’ indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df). 145

Table 3. 5 Linear mixed effects model comparisons soil moisture, measured as percent volumetric moisture content (% VMC), versus soil wetness index (SWI). ‘Location’ indicates whether samples were from New York (NY) or Pennsylvania (PA) and ‘Month’ indicates the sample month. Best fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC). 146

Table 3. 6 Linear mixed effects model outputs for best fit model (Model 5.4 in Table 3.5) for percent volumetric moisture content (% VMC) versus soil wetness index (SWI). ‘Location’ indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df). 146

Table 3. 7 Linear mixed effects model comparisons for soil extractions ($CaCl_2$-P, Ox-P, and Total-P). ‘Location’ indicates whether samples were collected in New York (NY) or Pennsylvania (PA) and ‘Month’ indicates the sample month. Best fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC). 150

Table 3. 8 Best fit linear mixed effects model (LMEM) outputs (Models 7.1, 7.3, 7.4, and 7.7 in Table 3.7) for soil extractions ($CaCl_2$-P, Ox-P, and Total-P). ‘Location’ indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df). 151

Table 3. 9 Linear mixed effects model comparisons for soil extractions ($CaCl_2$-P, Ox-P, and Total-P) versus soil wetness index (SWI). ‘Location’ indicates whether samples were from New York (NY) or Pennsylvania (PA) and ‘Month’ indicates the sample month. Best fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC). 155

Table 3. 1 0 Best fit linear mixed effects model (LMEM) outputs (Models 9.2, 9.4, 9.10, and 9.17 in Table 3.9) for soil extractions ($CaCl_2$-P, Ox-P, and Total-P) versus soil wetness index (SWI). ‘Location’ indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df). 156

Table 3. 1 1 Shotgun metagenome assembly results for samples taken in New York (NY) and Pennsylvania (PA). 161

Table 3. 1 2 Number of ppk1, ppk2, and ppx genes annotated in this study for soil samples taken from locations in New York (NY) and Pennsylvania (PA). 162

Table 3.13 Contigs with unique combinations of ppk and ppx genes as determined by shotgun metagenome analysis and subsequent comparison of protein sequence to existing sequences in the pBLAST database. 165

Table 3.14 Relative gene abundance (RA; in units of FPKM) linear mixed effects model (LMEM) comparisons versus variables of interest for ppk1, ppk2, and ppx. 'Location' indicates whether samples were collected in New York (NY) or Pennsylvania (PA). Best fit models for each soil extraction are bolded. Abbreviations: soil wetness index (SWI), degrees of freedom (df), Aikaike information criteria (AIC). 174

Table 3.15 Linear mixed effects model summaries for single predictor models of relative gene abundance (RA; in units of FPKM) presented in Table 3.13 for ppk1, ppk2, and ppx. Abbreviations: soil wetness index (SWI), standard error (SE), and degrees of freedom (df). ... 175

Table 3.16 Best fit linear mixed effects model (LMEM) summaries (Models 11.5, 11.16, and 11.27 in Table 3.14) for predicting relative abundance (RA; in units of FPKM) of ppk1, ppk2, and ppx genes. 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA). Abbreviations: standard error (SE) and degrees of freedom (df). 181

Table 4.1 Protein family (Pfam) database models used in this study. 216

Table 4.2 Linear model (LM) and generalized additive model (GAM) comparisons for ash free dry mass (AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC)..... 218

Table 4.3 Best fit linear model (LM) outputs (Model 2.3 in Table 4.2) for ash free dry mass (AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: standard error (SE). 218

Table 4.4 Best fit general additive model (GAM) outputs (Model 2.5 in Table 4.2) for ash free dry mass (AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: standard error (SE). The percent deviance explained by this GAM was 85.5%. 218

Table 4.5 Linear model (LM) and generalized additive model (GAM) comparisons for leaf phosphorus (leaf P; ppm P/g of AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC). 222

Table 4.6 Best fit general additive model (GAM) output (Model 5.5 in Table 4.5) for leaf phosphorus (leaf P; ppm P/g AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: standard error (SE). The percent deviance explained by this GAM was 70.2%. 222

Table 4. 7 Shotgun metagenome assembly results for leaf litter samples included in this study (both forest floor and stream treatments were assembled together). 226

Table 4. 8 Number of ppk1, ppk2, and ppx genes annotated in this study for leaf litter samples taken from on the forest floor and in the stream. 227

Table 4. 9 Contigs with unique combinations of ppk and ppx genes from leaf litter samples collected from the forest floor and in the stream as determined by shotgun metagenome analysis and subsequent comparison of protein sequence to existing sequences in the pBLAST database. Abbreviations: not given (NG). 236

Table 4. 1 0 Relative abundance (RA; in units of FPKM) generalized additive model (GAM) comparisons versus time and treatment (i.e., whether leaf litter was place in the stream or on the forest floor) for ppk1, ppk2, and ppx genes. fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df) and Aikaike information criteria (AIC). 239

Table 4. 1 1 Best fit generalized additive model (GAM) summaries (Models 10.2, 10.4, and 10.8 in Table 4.10) for predicting relative abundance (RA; in units of FPKM) of ppk1, ppk2, and ppx genes. Abbreviations: standard error (SE). 239

CHAPTER 1

LINKING WASTEWATER ENGINEERING AND THE NATURAL SCIENCES TO IMPROVE OUR UNDERSTANDING OF POLYPHOSPHATE ACCUMULATING ORGANISM (PAO)-MEDIATED PHOSHPHORUS CYCLING: A REVIEW

ABSTRACT

Excessive anthropogenic inputs of phosphorus (P) to the landscape continue to have negative consequences on water resources. The influence of abiotic controls on P cycling has been widely studied and applied to water quality management strategies, but less attention has been paid to understand and apply biotic (here, microbial) controls on P cycling. Concurrent research in engineered systems such as wastewater treatments plants and the natural environment (i.e., streams, lakes, soils, estuaries, and oceans) have made considerable progress when it comes to understanding the role of microbial controls on P cycling. However, there has been very little synthesis of this research across these diverse disciplines and application of it to manage water quality. In this review, we present the current knowledge of microbial-mediated P cycling in engineered systems—focusing on polyphosphate accumulating organisms (PAOs). Next, we discuss research in natural systems and outline potential roles of PAOs in the environment. For both systems, we highlight knowledge gaps and suggest tools for addressing them. Finally, we discuss how a broader view of microbial controls on P cycling may benefit agricultural systems. Overall, the goal of this review is to draw attention to microbial mechanisms influencing P cycling for the purpose of promoting discussion between disciplines, identifying PAOs in new environments, informing sustainable P management, and ultimately, improving water resources.

KEYWORDS

phosphorus, water quality, best management practices, microbial processes, polyphosphate, polyphosphate accumulating organisms, enhanced biological phosphorus removal, agriculture

INTRODUCTION

Compared to biotic (here, microbial) controls on phosphorus (P) cycling, abiotic controls are more widely studied and applied to water quality management strategies. However, researchers of engineered systems such as wastewater treatments plants have made considerable progress when it comes to understanding and managing microbial controls on P cycling. Meanwhile, numerous studies in the natural environment have also contributed to a deeper understanding of biotic controls on P but this understanding has not yet been applied across the landscape to manage water quality. With the exception of three reviews on the role of microbial-mediated controls on P cycling in aquatic systems (Davelaar 1993; Hupfer et al. 2007; McMahon and Read 2013), there have been no studies synthesizing this research across engineered and broader natural systems. As anthropogenic P sources on landscapes continue to impact water quality of lakes, rivers, and streams (Carpenter et al. 1998; Carpenter 2005; Dodds et al. 2009; Dubrovsky et al. 2010; Hudnell 2010; Kleinman et al. 2011a; Kleinman et al. 2011b; Obersteiner et al. 2013; Jarvie et al. 2015; Garcia et al. 2016; Jarvie et al. 2017), collaborative research across diverse disciplines may be an important step in developing more holistic strategies for improving water quality.

Key elements of this review are highlighted in Figure 1.1. First, we discuss advancements made in the field of wastewater engineering (i.e., engineered systems); we focus on the role of polyphosphate accumulating organisms (PAOs) in wastewater treatment plants that specialize in microbially-based P removal. We then review concurrent advances in natural systems such as streams, lakes, estuaries, oceans, and soils with respect to evidence for similar biotic controls on P. We synthesize knowledge gaps in both these systems and suggest tools that may push each forward while promoting feedback between one another. Finally, we suggest ways to extend our knowledge of biotic controls on P cycling to manage P in agricultural systems. Agricultural systems are a unique combination of engineered (i.e., managed) and natural systems; therefore, further studies of microbially-mediated P cycling in this environment may be informative to both engineered and natural systems (Figure 1.1).

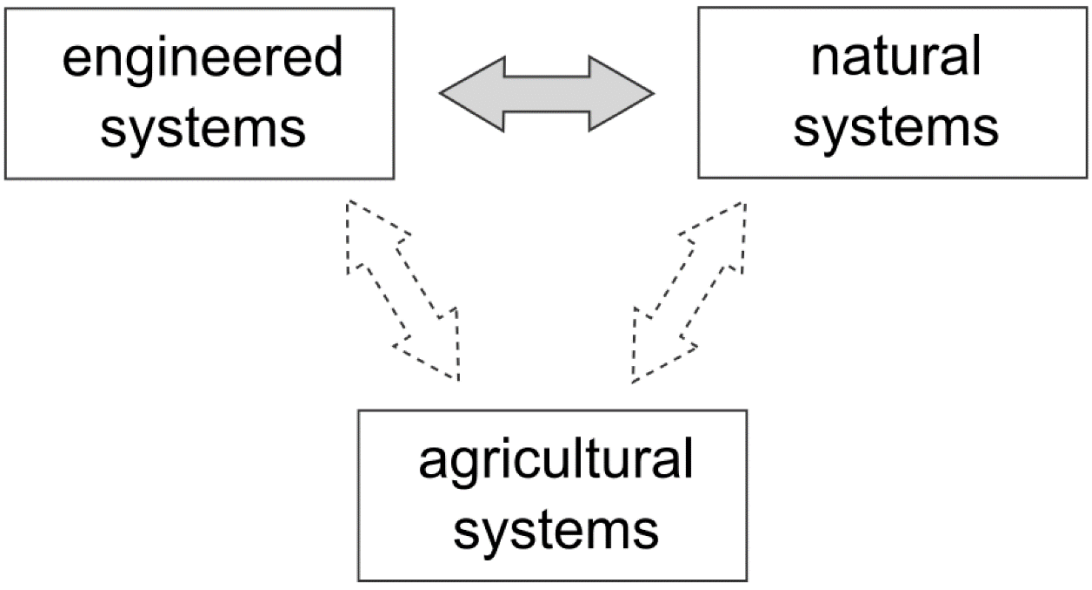

Figure 1. 1 Schematic of research flows between engineered, natural, and agricultural systems disciplines with respect to biotic controls on P cycling. Arrows between each system represent limited (grey) and potential (white, dashed outline) opportunities for information sharing.

BIOTIC CONTROLS ON PHOSPHORUS CYCLING IN ENGINEERED SYSTEMS

Urban areas are the second largest source of P to streams after agriculture (Carpenter et al. 1998; Easton et al. 2007; Alexander et al. 2008; Dubrovsky et al. 2010), and engineered systems such as wastewater treatment plants (WWTPs) play an important role in removing urban point P sources. Prior to the 1970s, WWTPs focused on removing urban point sources of carbon (C) and nitrogen (N). However, in the 1970s, researchers discovered a WWTP configuration that greatly improved P removal rates in addition to removing N and C (see review in Seviour et al. 2003). This configuration, referred to as enhanced biological P removal (EBPR), favors the growth of bacteria that store P intracellularly and thereby transfer P from solution to the waste solids (Barnard 1975; Barnard 1976; Seviour et al. 2003). For example, the P content (dry weight) of EBPR sludge is typically about 5-7%, as compared to 1.2% in conventional sludge (Yuan et al. 2012). This is accomplished in EBPR without the addition of liming salts such as Ca, Al, or Fe used by conventional WWTPs to chemically precipitate out P. Therefore, despite strict operational requirements, EBPR is a more economical means of removing P from effluent waters compared to conventional approaches (Barnard 1975; Oehmen et al. 2007).

Key Factors for Effective Enhanced Biological Phosphorus Removal

EBPR processes are unique because of their configuration as well as the organisms purposefully enriched by this configuration. A typically EBPR WWTP includes a few key operating conditions: (1) an upstream anaerobic zone with an electron donor (e.g., a form of C such as acetate) that is kept strictly anaerobic, (2) an aerobic zone, (3) recycling of a fraction of the suspended biomass back to the anaerobic zone after aerobic conditions, and (4) wasting of the remaining suspended biomass (Seviour et al. 2003). While optimal operating conditions were

originally developed empirically rather than based on an understanding of microbially processes, it is now commonly accepted that alternating between anaerobic and aerobic conditions selects for microbes that take up inorganic P (i.e., phosphate) in excess of normal cellular levels (Seviour et al. 2003)—coined 'luxury uptake' (Yall et al. 1972). Phosphate can make up 15-20% (dry weight) of microbial biomass during the aerobic stage (Crocetti et al. 2000; Ohtake et al. 2001). The key microorganisms enriched by this configuration, referred to as polyphosphate accumulating organisms (PAOs), have been intensively studied because of their major role in removing P from influent waters of EBPR WWTPs around the world (Hesselmann et al. 1999; Crocetti et al. 2000; García Martín et al. 2006; He et al. 2007; Kunin et al. 2008; Albertsen et al. 2012; Skennerton et al. 2014; Mao et al. 2014; Mao et al. 2015). PAOs made up approximately 33% of the bacterial community in a full-scale Danish EBPR WWTP (Albertsen et al. 2012) and other studies in the USA and Denmark have estimated PAOs make up from < 1 to 36% of the bacterial community in full-scale EBPR WWTPs (Zilles et al. 2002a; Kong et al. 2005; He et al. 2007). This stands in contrast to laboratory-scale SBRs that can achieve enrichments of up to 75-90% CAP (McMahon et al. 2002; Lu et al. 2006). The most commonly studied PAO is known by the provisional scientific name *Candidatus* Accumulibacter phosphatis (CAP; Hesselmann et al. 1999) and is a relative of *Rhodocyclus (Betaproteobacteria).* However, since the late 1990's, other non-CAP PAOs have been found in engineered and natural systems using microscopy and molecular biology methods (Table 1.1). To date, CAP has not been isolated in pure culture (Gebremariam et al. 2011; McMahon et al. 2013). *Tetrasphaera* PAOs with the ability to accumulate polyP have been isolated in pure culture, but these organisms did not store PHA (Maszenan et al. 2000).

Table 1. 1 A summary of PAOs found in the literature. Abbreviations: laboratory-scale sequencing batch reactor (SBR), full-scale wastewater treatment plant (WWTP), uncharacterized wastewater (Mixed).

Class/Phylum	Genus	Carbon Source	Reactor Type /Environment	Reference
Alphaproteobacteria	--	Mixed	WWTP	Zilles et al. 2002a
Alphaproteobacteria	--	Mixed	SBR	Kawaharasaki et al. 1999
Alphaproteobacteria	*Defluviicoccus**	Acetate	SBR	Nobu et al. 2014
Betaproteobacteria	--	Mixed	WWTP	Zilles et al. 2002a
Betaproteobacteria	--	Acetate	SBR	Stante et al. 1997
Betaproteobacteria	Accumulibacter	Mixed	WWTP	Beer et al. 2006; Albertsen et al. 2012; Nguyen et al. 2012; Mao et al. 2015
Betaproteobacteria	Accumulibacter	Mixed	SBR	Liu et al. 2001; Günther et al. 2009; Mao et al. 2014; Skennerton et al. 2014
Betaproteobacteria	Accumulibacter	Acetate	SBR	Hesselmann et al. 1999; Crocetti et al. 2000; McMahon et al. 2002; Garcia-Martin et al. 2006; Kim et al. 2010
Betaproteobacteria	Accumulibacter	Propionate	SBR	Garcia-Martin et al. 2006
Betaproteobacteria	Accumulibacter	Mixed	Estuary	Kunin et al. 2008; Peterson et al. 2008
Betaproteobacteria	Accumulibacter	Mixed	Freshwater	Kunin et al. 2008; Peterson et al. 2008
Betaproteobacteria	Accumulibacter	Mixed	Soil	Kunin et al. 2008
Betaproteobacteria	Accumulibacter	Mixed	Sediment	Kunin et al. 2008
Betaproteobacteria	*Dechloromonas*	Mixed	WWTP	Zilles et al. 2002b; Kong et al. 2007
Betaproteobacteria	*Propinoibacter*	Acetate	SBR	Crocetti et al. 2000
Betaproteobacteria	*Rhodocyclus*	Mixed	WWTP	Zilles et al. 2002a; 2002b; Kong et al. 2004; 2005
Betaproteobacteria	*Rhodocyclus*	Acetate	SBR	Hesselmann et al. 1999; Crocetti et al. 2000; McMahon et al. 2002
Gammaproteobacteria	--	Mixed	WWTP	Zilles et al. 2002a
Gammaproteobacteria	--	Mixed	SBR	Liu et al. 2001
Gammaproteobacteria	*Acinetobacter*	Mixed	WWTP	Wagner et al. 1994; Streichan et al. 1990
Gammaproteobacteria	*Halomonas*	Mixed	WWTP	Nguyen et al. 2012
Gammaproteobacteria	*Pseudomonas*	Mixed	WWTP	Nguyen et al. 2012
Gammaproteobacteria	*Pseudomonas*	Mixed	SBR	Günther et al. 2009
Gammaproteobacteria	*Thiomargarita*	Mixed	Sediment pore water	Schulz and Schulz 2005
Gammaproteobacteria	*Beggiatoa*	Acetate	Pure culture	Brock and Schulz-Vogt 2011
Gammaproteobacteria	*Thiothrix*	Acetate	SBR	Rubio-Ricón et al. 2017

*This organism was described as a competitor to PAOs, a glycogen accumulating organism (GAO), but had genes coding for intracellular polyphosphate storage unlike most *Defluviicoccus*-related GAOs

Table 1.1 (continued)

Class/Phylum	Genus	Carbon Source	Reactor Type /Environment	Reference
Actinobacteria	--	Mixed	WWTP	Auling et al. 1991; Beer et al. 2006
Actinobacteria	--	Mixed	SBR	Bark et al. 1993
Actinobacteria	*Arthrobacter*	Mixed	Plant rhizosphere	Li et al. 2013
Actinobacteria	*Gordonia*	Mixed	WWTP	Beer et al. 2006
Actinobacteria	*Microlunatus*	Mixed	Pure culture	Kawakoshi et al. 2012
Actinobacteria	*Microlunatus*	Mixed	SBR	Nakamura et al. 1991; 1995a; 1995b; Kawaharasaki et al. 1998
Actinobacteria	*Microlunatus*	Mixed	WWTP	Beer et al. 2006
Actinobacteria	*Tetrasphaera*	Mixed	WWTP	Maszenan et al. 2000; Nguyen et al. 2011; Kong et al. 2005; Albertsen et al. 2012; Kristiansen et al. 2013
Actinobacteria	*Tetrasphaera*	Mixed	SBR	Günther et al. 2009
Bacilli	*Bacillus*	Mixed	Plant rhizosphere	DebRoy et al. 2013
Gemmatimonadetes	*Gemmatimonas*	Mixed	SBR	Zhang et al. 2003
Cyanobacteria	*Leptolyngbya*	--	Marine, SBR	Zhang et al. 2015, Oyserman et al. 2017
Melainabacteria		Mixed	WWTP	Soo et al. 2014

Polyphosphate Accumulating Organism Metabolism

Since the EBPR process was first introduced, we have a better—but not complete—understanding of PAO metabolism. In this context, model PAOs like CAP display the phenotype of utilizing polyphosphate (polyP) and poly-β-hydroxyalkanoate (PHA) polymers under alternating anaerobic/aerobic conditions (Seviour et al. 2003; Seviour and McIlroy 2008). PolyP is an important molecule to many diverse organisms including bacteria, archaea, fungi, plants, and animals (Kornberg 1995; Zhang et al. 2002; Hupfer et al. 2007; Rao et al. 2009; Seviour and Nielsen 2010; Achbergerová and Nahálka 2011) as it serves as an energy source and a P reservoir, is needed for growth and biofilm formation, is a strong ion chelator, can buffer against alkali conditions, regulates gene expression in organisms under stress, and regulates virulence factors (Kornberg 1995; Brown and Kornberg 2008; Rao et al. 2009; Seviour and Nielsen 2010; Achbergerová and Nahálka 2011; Kulakovskaya et al. 2012). However, while many organisms are capable of accumulating polyP, it is commonly understood that their ability to synthesize large amounts of polyP under aerobic conditions and use this as an energy source under anaerobic conditions enables PAOs to outcompete other EBPR heterotrophs with more limited fermentative processes (Gebremariam et al. 2011).

While metabolic mechanisms separating PAOs from non-PAOs are still being debated and discovered (e.g., Gebremariam et al. 2011; Rubio-Rincón et al. 2017), researchers explain the metabolism of CAP (the model PAO) in effective EBPR WWTPs as follows. During anaerobic conditions (Figure 1.2A), PAOs take up short chain volatile fatty acids (VFAs) and store them as PHAs (Seviour et al. 2003; Seviour and Nielsen 2010; He and McMahon 2011b). For example, CAP stores acetate as the PHA known as poly-β-hydroxybutyrate (PHB). This ability to take up

C and store it as PHB under anaerobic conditions gives CAP an advantage over other organisms that cannot store C anaerobically (Yuan et al. 2012). Also during anaerobic periods, intracellular polyP and glycogen concentrations decrease because these polymers are used as a source of energy and electrons to convert VFA to PHB (Seviour et al. 2003; Seviour and McIlroy 2008; Seviour and Nielsen 2010). This decrease in intracellular polyP is accompanied by an increase in extracellular phosphate as phosphoanhydride bonds between phosphate molecules of the polyP chain are broken and unbound phosphate is transported out of the cell into the bulk water of the WWTP (Seviour et al. 2003; García Martín et al. 2006; Seviour and Nielsen 2010; Skennerton et al. 2014; Oyserman et al. 2016). During aerobic periods (Figure 1.2B), CAP uses the energy released from the respiration of PHB to replace polyP and glycogen stores (Seviour et al. 2003; Seviour and McIlroy 2008; Seviour and Nielsen 2010). The increase in intracellular polyP is accompanied by a decrease in extracellular phosphate as CAP transport phosphate into the cell—removing it from the bulk water before exiting the WWTP—to build polyP chains (Seviour et al. 2003; Seviour and Nielsen 2010).

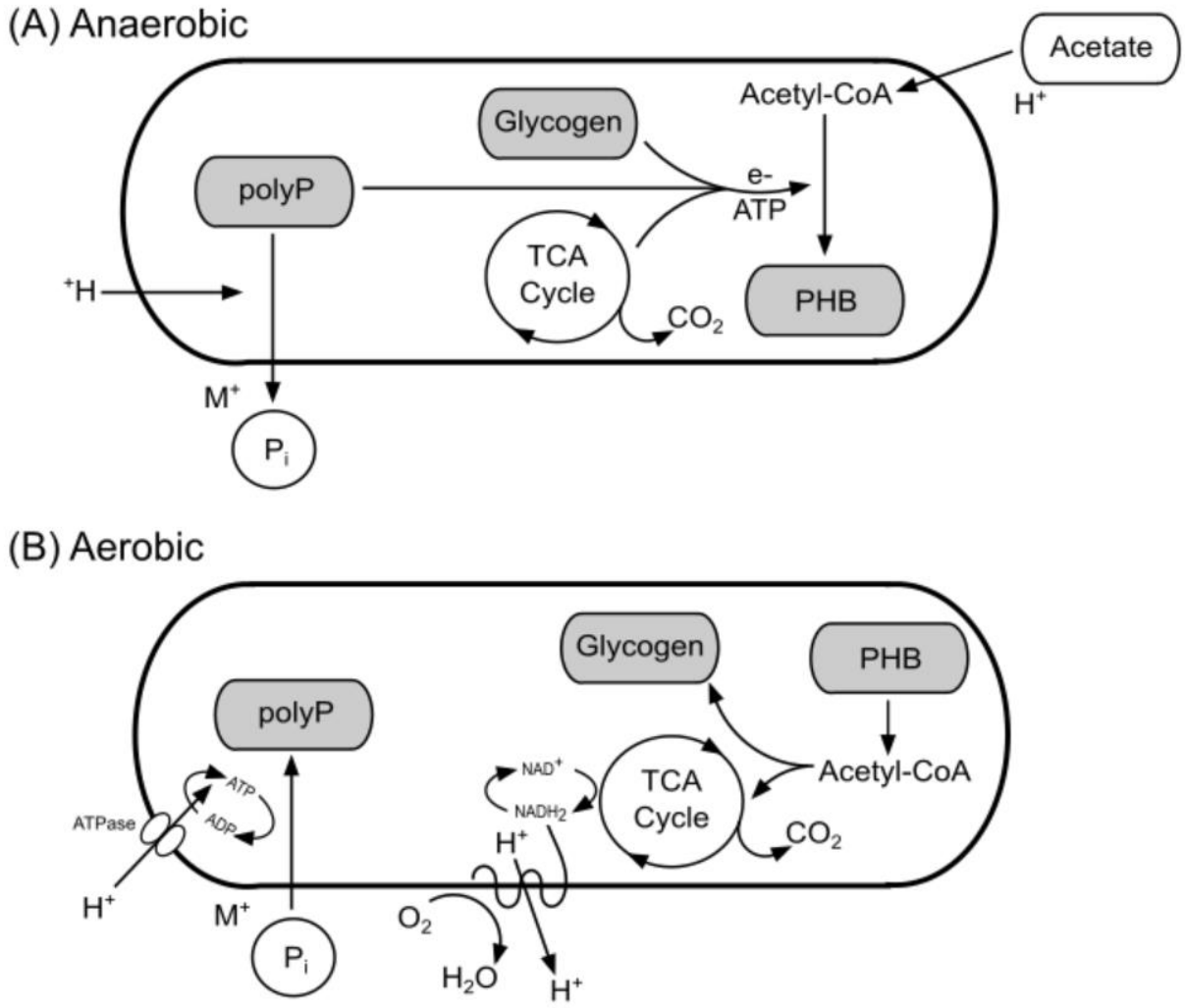

Figure 1. 2 (A) Anaerobic and (B) aerobic metabolism of model EBPR PAO, *Candidatus* Accumulibacter phosphatis (adapted from Seviour et al. 2003, Seviour and Nielsen 2010, and Skennerton et al. 2014). M^+ represents metal cations and P_i represents phosphate.

Debates concerning the specifics of CAP metabolism focus on processes occurring under anaerobic conditions (He and McMahon 2011b; Skennerton et al. 2014): (1) the source of ATP for PHA synthesis, (2) the source of reducing power for PHA synthesis, and (3) the pathway for degrading glycogen (called glycolysis). Many metabolic models have been proposed to address the first two points (Seviour et al. 2003; Oehmen et al. 2007; Seviour and McIlroy 2008; Seviour

and Nielsen 2010) but the two most commonly discussed are the Comeau-Wentzel model (Comeau et al. 1986; Wentzel et al. 1986) and the Mino model (Mino et al. 1987). The Comeau-Wentzel model recognizes only polyP as providing energy to convert ADP to ATP for making PHAs under anaerobic conditions, whereas, both polyP and glycogen can be used as a source of energy in the Mino model (Seviour and Nielsen 2010). These two models also differ in how the anaerobic reducing power, in the form of NADH, is generated to convert acetate or another C source to PHA. In the Comeau-Wentzel model, NADH can only be generated by the tricarboxylic acid cycle (TCA; Seviour and Nielsen 2010). However, in the Mino model, glycolysis as well as the TCA cycle can be used to generate NADH (Seviour and Nielsen 2010). It has also been suggested that reducing power may come from either a glyoxylate shunt or split TCA cycle (Oehmen et al. 2007; He and McMahon 2011b; Skennerton et al. 2014). The use of the anaerobic TCA cycle may be more likely in full-scale WWTPs where VFAs are more limited compared to laboratory-scale SBRs (Lanham et al. 2013; Lanham et al. 2014). In addition to substrate availability, the type of organism and its length of aerobic cycle may influence the pathway for generating reducing power (Flowers et al. 2013; Lanham et al. 2014). In terms of glycolysis pathways, PAOs utilize either the Entner-Douderoff (ED) or Embden-Meyerholf-Parnas (EMP) pathway (Oehmen et al. 2007; Seviour and McIlroy 2008; He and McMahon 2011a; Skennerton et al. 2014). CAP strains were shown to only harbor EMP pathway genes (García Martín et al. 2006) but this may not always be the case of non-CAP PAOs (Skennerton et al. 2014).

Genetic and phenotypic diversity may explain the ongoing debates concerning the specifics of PAO metabolism beyond that of CAP; rather than a single metabolic model, many markedly

different metabolic models may exist (Mino et al. 1998; Crocetti et al. 2000; Seviour et al. 2003; Seviour and McIlroy 2008; Kawakoshi et al. 2012; Kristiansen et al. 2013; Mao et al. 2014; Skennerton et al. 2014; Rubio-Rincón et al. 2017). For example, *Tetrasphaera*-related PAOs, which can accumulate intracellular glycogen and polyP as well as denitrify and ferment, are still considered PAOs even though they do not accumulate PHAs under anaerobic conditions (Kristiansen et al. 2013). The PAO *Microlunatus phosphovorus* released phosphate under anaerobic conditions, stored polyP under aerobic conditions (Nakamura et al. 1995a) and was able to take up a wide range of C substrates under anaerobic conditions (Nakamura et al. 1995b). In a SBR initially dominated by CAP, researchers demonstrated the emergence of *Thiothrix caldifontis* after exposing the SBR to sulfide (Rubio-Rincón et al. 2017). *Thiothrix caldifontis* took up P during aerobic conditions and released it during anaerobic conditions, removed P more effectively than CAP, and used both PHA and intracellular polysulfide as energy sources for growth (Rubio-Rincón et al. 2017). As we learn more about the diversity of PAOs, we must strive to better define what distinguishes a PAOs from a non-PAOs as well as how EBPR WWTP microbial communities (PAOs and non-PAOs) contribute to effective P removal in engineered systems.

Polyphosphate Accumulating Organism-Related Functional Genes

In addition to details concerning PAO metabolism, studies have uncovered key functional genes linked to biotic uptake and release of P (Table 1.3). The *ppk* gene codes for a polyphosphate kinase (PPK) that catalyzes the reversible reaction of ATP to ADP to form intracellular polyP (Ahn and Kornberg 1990; Akiyama et al. 1992). There are two main *ppk* genes: *ppk1* and *ppk2* (Zhang et al. 2002; Rao et al. 2009; Kawakoshi et al. 2012). While not recognized as a PAO, the

nucleotide sequence for *ppk1* was first isolated from *E. coli* (Akiyama et al. 1992) and since then other studies have identified PPKs in a wide range of bacterial, archaeal, and eukaryotic organisms (Trelstad et al. 1999; Zhang et al. 2002; Rao et al. 2009; Kawakoshi et al. 2012). *ppk* has been shown to be an ideal marker for bacterial strain diversity because it is highly conserved (Tzeng and Kornberg 1998). With respect to PAOs, CAP genomes have a single copy of *ppk* and *ppk* has been shown to evolve faster than CAP 16S rRNA genes (Kunin et al. 2008; He and McMahon 2011a). PPK1 is likely a membrane-bound protein with four domains concentrated in regions where the inner and outer cell membranes come together (Ahn and Kornberg 1990). PPK2 is another major PPK enzyme, which differs from PPK1 in its ability to catalyze the making of polyP from both GTP and ATP as well as enzyme cofactors; PPK1 has an affinity for Mg^{2+} while PPK2 has an affinity for Mn^{2+} (Zhang et al. 2002; Rao et al. 2009). The nucleotide sequence of *ppk2* was first isolated from *Pseudomonas aeruginosa* (Zhang et al. 2002). Some organisms have both *ppk1* and *ppk2* while others have only one or the other (Zhang et al. 2002; Rao et al. 2009). Overall, CAP *ppk* genes have been found in full-scale Canadian, United States, British, Danish, Chinese, Singaporean, Japanese and Australian EBPR WWTPs indicating their ubiquity worldwide (Kunin et al. 2008; Albertsen et al. 2012; Mao et al. 2015). However, we found no studies quantifying the abundance and diversity of non-CAP PAO *ppk* genes in EBPR WWTPs. Therefore, given that it is a conserved gene and is found in many diverse organisms, understanding the abundance and diversity of the *ppk* gene in both CAP and non-CAP PAOs may prove important to understanding and monitoring P cycling in engineered systems like EBPR WWTPs as well as in natural systems.

Table 1. 2 Proteins associated with PAO-related phosphorus cycling functional genes.

Abbreviation	Protein	Function and Key Traits	References
PPK1	polyphosphate (polyP) kinase	Catalyzes the de-phosphorylation of ATP to make polyP. Has an affinity for Mg^{2+}.	Ahn and Kornberg 1990, Akiyama et al. 1992, Trelstad et al. 1999, Zhang et al. 2002, Zhu et al. 2005, Rao et al. 2009, Kawakoshi et al. 2012
PPK2	polyphosphate kinase	Catalyzes the de-phosphorylation of ATP and GTP to make polyP. Has an affinity for Mn^{2+}.	Ishige et al. 1998, 2002, Zhang et al. 2002, Nocek et al. 2008, Rao et al. 2009, Kawakoshi et al. 2012
PPX1	exopolyphosphatase	Liberates the terminal phosphate molecule in a polyP chain. Requires Mg^{2+} and KCl.	Reizer et al. 1992, Akiyama et al. 1993, Wurst and Kornberg 1994, Kornberg 1995, Zago et al. 1999, Ohtake et al. 2001, Rangarajan et al. 2006, Lichko et al. 2006, Rao et al. 2009, Kawakoshi et al. 2012
PPX2/GPPA	exopolyphosphatase/ pentaphosphate phosphohydrolase	Liberates the terminal phosphate in a polyP chain and important in cellular stress response by hydrolyzing pppGpp to ppGpp.	Keasling et al. 1993, Reizer et al. 1993, Zago et al. 1999, Rao et al. 2009, Kawakoshi et al. 2012, Alcántara et al. 2014
PPN	endopolyphosphatase	Cleaves phosphate from polyP chains (not selective to terminal phosphate). Has only been found in archaea, fungi, and mammals.	Kumble and Kornberg 1996, Shi and Kornberg 2005, Lichko et al. 2006, Rao et al. 2009
PAP	AMP phosphotransferase	Catalyzes reaction of AMP to ADP using energy from breakdown of polyP.	Rao et al. 2009, Mao et al. 2014, Skennerton et al. 2014
PIT	low-affinity phosphate transporter	Binds phosphate and brings it into the cell.	Seviour and McIlroy 2008; Mao et al. 2014
PST	high-affinity phosphate transporter	Binds phosphate and brings it into the cell.	Seviour and McIlroy 2008; Mao et al. 2014

The *ppx* gene codes for an exopolyphosphatase (PPX) responsible for breaking off the terminal phosphate molecules of a polyP chain under conditions of excess phosphate (Akiyama et al. 1993; Kornberg 1995; Zago et al. 1999; Ohtake et al. 2001). There are two main *ppx* genes: *ppx1* and *ppx2/gppA* (Rao et al. 2009). Some organisms have both *ppx* genes (e.g., *Lactobacillus*; Alcántara et al. 2014, *E. coli*; Akiyama et al. 1993; Keasling et al. 1993; Reizer et al. 1993), but this trend may vary by organism and has not been well characterized for PAOs in the literature. PPX1 is a dimer protein with four domains that requires both Mg^{2+} and K^{+} (Akiyama et al. 1993; Rangarajan et al. 2006; Rao et al. 2009). PPX1 preferentially acts on longer chains of polyP (i.e., 500 phosphate molecules or longer), does not act on ATP, and cannot be inhibited by ADP or ATP (Akiyama et al. 1993). The *ppx2/gppA* gene codes for exopolyphosphatase enzyme known as pentaphosphate phosphohydrolase (PPX2/GPPA) that inhibits polyP accumulation at the enzymatic level through stress response nucleotides ppGpp and pppGpp (Reizer et al. 1993; Zago et al. 1999; Rao et al. 2009). While not recognized as a PAO, PPX2/GPPA was first purified in a mutant strain of *E. coli,* is regulated by the ppx/*gppA* gene, and can cause the release of phosphate by breaking polyP chains or by hydrolyzing pppGpp to ppGpp (Keasling et al. 1993). PPX2/GPPA is thought to be less active than PPX1, prefers longer polyP chains (i.e., 1000 residues or longer), and is inhibited by the presence of short- and medium-length polyP chains (Keasling et al. 1993). We found only a few studies on *ppx1* and *ppx2/gppA* genes, none of which assess their abundance and diversity in the EBPR systems. Some studies have found that P starvation induces transcription of *ppk* and *ppx* genes but whether this can be applied to PAOs is unknown (Seviour et al. 2003). Due to their role in polyP breakdown under anaerobic conditions—an important defining metabolic characteristic of putative PAOs—further study of

ppx genes is important for improving our understanding of microbial controls on P cycling in engineered systems.

Other P cycling genes of interest include *pap*, *pit*, *pst*, and *ppn*. Associated protein functionality and key traits are summarized in Table 1.3. As we mentioned previously for *ppk* and *ppx*, more work needs to be done to characterize the abundance, diversity, and role of these genes in P cycling with respect to engineered systems. While we only focus on P-related genes here, little is known about functional genes regulating other important PAO polymers: PHA and glycogen (Seviour et al. 2003).

Research Methods and Opportunities in Engineered Systems

Researchers have used a combination of microscopy and molecular biology tools to study the (1) metabolism, (2) functional genes, and (3) communities (diversity, structure, and function) of PAOs in EBPR WWTPs. Here, we outline some of the more popular tools used as well as how they may be extended to further our understanding of these three focus areas.

Microscopy can be used to formally confirm PAO phenotype of EBPR PAOs (e.g., Hung et al. 2001) as well as track membership and function of PAO communities in time and space (e.g., Mao et al. 2014; Lawson et al. 2015). Transmission electron microscopy (TEM) with energy dispersive x-ray analysis has been used to verify the presence of polyP granules in EBPR systems (Streichan et al. 1990). Besides electron microscopy, non-fluorescent stains such as

Neisser's blue and toluidine blue (Crocetti et al. 2000; Schulz and Schulz 2005) as well as fluorescent stains including 4',6-diamidino-2-phenylindole (DAPI), calcium indicator Fura-2, Nile Blue A, and tetracycline (Hesselmann et al. 1999; Crocetti et al. 2000; Hung et al. 2002; Hupfer et al. 2008; Aschar-Sobbi et al. 2008; Günther et al. 2009; Diaz and Ingall 2010) have been used to verify the presence of intracellular polyP granules. We focus on DAPI, which is commonly used to identify DNA for cell and bacterial counts, but has also been used to visualize and quantify polyP granules in PAOs (e.g., Hung et al. 2002). DAPI-polyP visualization is possible because the emission spectrum for DAPI-polyP is shifted from DAPI-DNA; the maximum emission of DAPI-DNA is at approximately 490 nm, while the maximum emission of DAPI-polyP is at 550 nm (Hung et al. 2002; Aschar-Sobbi et al. 2008; Diaz and Ingall 2010). By adjusting the excitation wavelength from 360 nm to 415 nm, users can detect DAPI-polyP concentrations down to the ng/ml range (Aschar-Sobbi et al. 2008). DAPI also stains lipids but the fluorescence is weak and fades quickly (Streichan et al. 1990). Potential negative controls for DAPI-polyP analysis include *E. coli* BL21 and potential positive controls for DAPI polyP analysis include *Pseudomonas putida* KT2440 and *Acinetobacter calcoaceticus* (Kulakova et al. 2011). The direct count ratio of DAPI-polyP stained cells to DAPI-DNA stained cells can serve as a simple method of quantifying total PAOs, including those that have not been phylogenetically identified (e.g., Saia et al. 2017). Several researchers have used a fluorescent plate reader to quantify relative DAPI-polyP fluorescence between samples (Kulakova et al. 2011; Martin and Van Mooy 2013; Rier et al. 2016) but care should be taken to rule out interference of calcium or other salts in samples because this approach lacks single-cell resolution (Aschar-Sobbi et al. 2008; Diaz and Ingall 2010; Kulakova et al. 2011; Martin and Van Mooy 2013).

Fluorescent *in-situ* hybridization (FISH) can be used to identify PAOs via the hybridization of fluorescent oligonucleotide probes to 16S rRNA or 23S rRNA sequences of interest (Seviour et al. 2003; Seviour and Nielsen 2010). FISH probes have been designed for many known EBPR PAOs (Table 1.4). However, additional research is needed to develop FISH probes for less common PAOs. Researchers have also used quantitative FISH (qFISH) measurements to compare the number of directly counted probed PAOs to probed total bacteria (Kong et al. 2007; Albertsen et al. 2012; Lanham et al. 2014) or to understand which PAOs are more dominant in EBPR communities (e.g., Beer et al. 2006). Lastly, FISH can be used in conjunction with other assays to verify the PAO phenotype. Researchers have used FISH and DAPI to identify PAOs and verify whether they store polyP under aerobic conditions in EBPR WWTPs (e.g., Crocetti et al. 2000). Other studies used microautoradiography-FISH (MAR-FISH) to simultaneously verify intracellular P in PAOs from an EBPR laboratory-scale reactor (Kim et al. 2010). Catalyzed reporter deposition-FISH (CARD-FISH) or helper probes (Fuchs et al. 2000; Amann et al. 2001) can improve probe visibility (Hupfer et al. 2008; Seviour and Nielsen 2010). The pros and cons of using FISH have been reviewed by others (Seviour et al. 2003; Seviour and Nielsen 2010).

Researchers have used DAPI in combination with flow cytometry to sort out and characterize CAP PAOs as well as new EBPR organisms that store polyP (Zilles et al. 2002a; Zilles et al. 2002b; Günther et al. 2009; Kim et al. 2010). However, opportunities exist to apply DAPI and flow cytometry more broadly as well as to combine them with next generation sequencing technologies (discussed below) to design PCR primers and FISH probes for the purpose of identifying non-CAP PAOs.

Table 1. 3 Fluorescence *in-situ* hybridization (FISH) probes utilized in the literature for identification of PAO-related organisms (adapted from Seviour et al. 2010). Abbreviations: polyphosphate accumulating organisms (PAO), not determined (ND).

Probe Name	Sequence (5'-3')	Formamide (%)	Target	Reference
ALF1b	GCTGCCTCCCGTAGGAGT	20	Alphaproteobacteria	Manz et al. 1992
BET42[a]	GCCTTCCCACTTCGTTT	35	Betaproteobacteria	Manz et al. 1992
GAM42[a]	GCCTTCCCACATCGTTT	35	Gammaproteobacteria	Manz et al. 1992
RHC175	TGCTCACAGAATATGCGG	30	*Rhodocyclus*/Accumulibacter	Hesselmann et al. 1999
RHC439	CNATTTCTTCCCCGCCGA	30	Most *Rhodocyclaceae*	Hesselmann et al. 1999
Rc988	AGGATTCCTGACATGTCAAGGG	ND	*Rhodocyclus* group	Crocetti et al. 2000
PAO462[c]	CCGTCATCTACWCAGGGTATTAAC	35	Most Accumulibacter	Crocetti et al. 2000
PAO651[c]	CCCTCTGCCAAACTCCAG	35	Most Accumulibacter	Crocetti et al. 2000
PAO846[c]	GTTAGCTACGGCACTAAAAGG	35	Most Accumulibacter	Crocetti et al. 2000
Acc-I-444	CCCAAGCAATTTCTTCCCC	35	PAO clade IA and other Type I clades	Flowers et al. 2009
Acc-II-444	CCCGTGCAATTTCTTCCCC	35	PAO clade IIA, IIC, and IID	Flowers et al. 2009
Actino-1011	TTGCGGGGCACCCATCTCT	30	*Tetrasphaera*-relatives	Liu et al. 2001
Actino-221[a]	CGCAGGTCCATCCCAGAC	30	*Tetrasphaera*-relatives	Kong et al. 2005
Actino-658[a]	TCCGGTCTCCCCTACCAT	40	*Tetrasphaera*-relatives	Kong et al. 2005
Tet1-266	CCCGTCGTCGCCTGTAGC	25	*Tetrasphaera*-relatives	Nguyen et al. 2011
Tet2-892	TAGTTAGCCTTGCGGCCG	5	*Tetrasphaera*-relatives	Nguyen et al. 2011
Tet2-174	GCTCCGTCTCGTATCCGG	20	*Tetrasphaera*-relatives	Nguyen et al. 2011
Tet3-654	GGTCTCCCCTACCATACT	35	*Tetrasphaera*-relatives	Nguyen et al. 2011
Tet3-19	CAGCGTTCGTCCTACACA	0	*Tetrasphaera*-relatives	Nguyen et al. 2011
BET135	ACGTTATCCCCCACTCAATGG	45	*Dechloromonas*-relatives	Kong et al. 2007
MIC179	GAGCAAGCTCTTCTGAAACCG	10	*Microlunatus phosphovorus*	Kawaharasaki et al. 1998
G123T	CCTTCCGATCTCTATGCA	40	*Thiothrix*-relatives	Kanagawa et al. 2000, Rubio-Rincón et al. 2017
EUB338	GCTGCCTCCCGTAGGAGT	60	Most Bacteria[b]	Amann et al. 1990
EUB338-II	GCAGCCACCCGTAGGTGT	60	Most Bacteria[b]	Daims et al. 1999
EUB338-III	GCTGCCACCCGTAGGTGT	60	Most Bacteria[b]	Daims et al. 1999

[a]Competitor probes required. [b]Use EUB338, EUB338-II, and EUB338-III together to obtain an estimate of total bacteria. [c]Use PAO462, PAO651, PAO846 together to obtain an estimate of total PAOs.

Molecular biology tools such as polymerase chain reaction (PCR) primers (Table 1.5) can be used to identify known PAOs, measure PAO functional gene abundances (via quantitative PCR; qPCR), and determine strain diversity (Flowers et al. 2013; Mao et al. 2014; Mao et al. 2015; Mao et al. 2016; Zhang et al. 2016). One study used 16S rRNA gene and *ppk1* qPCR primers in conjunction with other molecular biology techniques to estimate the abundance of various CAP clades in WWTPs as well as how they were genetically related to one another (Mao et al. 2015). Another study used qPCR in combination with FISH to monitor CAP groups over time in a laboratory reactor (Oyserman et al. 2016). qPCR is culture-independent, has a low detection limit compared to other methods, and users can test for and potentially control biases due to poor cell lysis (Seviour and Nielsen 2010). The pros and cons of using qPCR are reviewed by others (Seviour and Nielsen 2010). There is only one PCR *ppk* primers available for non-CAP PAOs and very few studies using these non-CAP primers in engineered systems (Table 1.1). Additionally, we could find no studies assessing the abundance and diversity of *ppx* genes in EBPR WWTPs. This may be due to, as their name suggests, a focus on P accumulation rather than P breakdown. However, as we seek out new non-CAP PAOs, a shift in focus to both P accumulation and breakdown may be necessary. Overall, more work is needed to design and apply a wider range of qPCR primers for non-CAP PAOs to study functional gene abundance, lineage, and diversity.

Table 1. 4 PAO PCR primer sequences identified in the literature.

Primer Pair	Sequence (5'-3')	Gene Target	Reference
NLDE-0199F	CGTATGAATTTTCTTGGTATTTATTGTACTAATCTngaygarttyt	Most *ppk1*	McMahon et al. 2002, 2007
TGNY-1435R	GTCGAGCAGTTTTTGCATGAwarttnccngt		
ACCppk1-254F	TCACCACCGACGGCAAGAC	CAP *ppk1*	McMahon et al. 2002, 2007,
ACCppk1-1376R	ACGATCATCAGCATCTTGGC		Kunin et al. 2008
ppk274f	ACCGACGGCAAGACSG	CAP *ppk1*	Kunin et al. 2008
ppk1156r	CGGTAGACGGTCATCTTGAT		
ppk734f	CTCGGCTGCTACCAGTTCCG	CAP *ppk1*	Kunin et al. 2008
ppk1601r	GATSCCGGCGACGACGTT		
Acc-ppk1-763f	GACGAAGAAGCGGTCAAG	CAP Clade 1A *ppk1*	He et al. 2007, He and McMahon 2011a
Acc-ppk1-1170r	AACGGTCATCTTGATGGC		
Acc-ppk1-974f	TGATGCGCGACAATCTCAAATTCAA	CAP Clade 1A *ppk1*	Zhang et al. 2016
Acc-ppk1-1113r	AATGATCGGATTGAAGCTCTGGTAG		
Acc-ppk1-372f	TGAAGGCATTCGCTTCCT	CAP Clade 1B *ppk1*	Zhang et al. 2016
Acc-ppk1-653r	AAGCAGTATTCGCTGTC		
Acc-ppk1-362f	AGCTGGCGAGTGAAGGCATTCG	CAP Clade 1C *ppk1*	Zhang et al. 2016
Acc-ppk1-758r	AACAGGTTGCTGTTGCGCGTGA		
Acc-ppk1-634f	TGCGACAGCGAATACAG	CAP Clade 1D *ppk1*	Zhang et al. 2016
Acc-ppk1-848r	ACTTCGAGGCGGACG		

Table 1.4 (continued)

Primer Pair	Sequence (5'-3')	Gene Target	Reference

Acc-ppk1-893f	AGTTCAATCTCACCGACAGC	CAP Clade 2A *ppk1*	He et al. 2007, He and McMahon 2011a
Acc-ppk1-997r	GGAACTTCAGGTCGTTGC		
Acc-ppk1-870f	GATGACCCAGTTCCTGCTCG	CAP Clade 2B *ppk1*	He et al. 2007
Acc-ppk1-1002r	CGGCACGAACTTCAGATCG		
Acc-ppk1-254f	TCACCACCGACGGCAAGAC	CAP Clade 2C *ppk1*	He et al. 2007
Acc-ppk1-460r	CCGGCATGACTTCGCGGAAG		
Acc-ppk1-375f	GGGTATCCGTTTCCTCAAGCG	CAP Clade 2D *ppk1*	He et al. 2007
Acc-ppk1-522r	GAGGCTCTTGTTGAGTACACGC		
Acc-ppk1-757f	TTCGTGGACGAGGAAGA	CAP Clade 2E *ppk1*	Zhang et al. 2016
Acc-ppk1-1129r	ATTGTTCGAGCAACTCGATG		
Acc-ppk1-410f	CCGAGCAACGCGAATGG	CAP Clade 2G *ppk1*	Zhang et al. 2016
Acc-ppk1-514r	TGTTGAGTACGCGCGGGA		
Acc-ppk1-701f	ACTCCTTCGTATTCCTCTCT	CAP Clade 2H *ppk1*	Zhang et al. 2016
Acc-ppk1-928r	TCATCGCTTCGGAGCA		
Acc-ppk1-688f	AGTGATTATGCTTTCGTCTTTC	CAP Clade 2I *ppk1*	Zhang et al. 2016
Acc-ppk1-946r	TGAACTGTCCGAGCAGGA		
CAP438f	GGTTAATACCCTGWGTAGAT	CAP 16S	Kunin et al. 2008
CAP846r	GTTAGCTACGGCACTAAAAGG		

Table 1.4 (continued)

Primer Pair	Sequence (5'-3')	Gene Target	Reference
518f	CCAGCAGCCGCGGTAAT	CAP 16S	He et al. 2007, He and McMahon 2011a
PAO-846r	GTTAGCTACGGCACTAAAAGG		

16S-Acc-1Af	TTGCTTGGGTTAATACCCTGA	CAP Clade 1A 16S	He et al. 2010
16S-Acc-1Ar	CTGCCAAACTCCAGTCTTGC		
16S-Acc-2Af	TTGCACGGGTTAATACCCTGT	CAP Clade 2A 16S	He et al. 2010
16S-Acc-2Ar	CTCTGCCAAACTCCAGCCTG		
Pse136f	TAGTAGTGGGGGATAACGTC	*Halomona*-related 16S	Nguyen et al. 2012
1492r	GCYTACCTTGTTACGACTT		

An alternative molecular biology tool that can be used to expand our knowledge of functional genes in both CAP and non-CAP PAOs is shotgun metagenomic sequencing. This next generation sequencing approach is contrasted to amplicon sequencing, which is commonly used to carry out targeted sequencing of genes such as 16S rRNA, 18S rRNA, and internal transcribed spacer (ITS) regions between rRNA genes for bacterial, eukaryotic, and fungal community analysis, respectively (Zimmerman et al. 2014). In the case of studying PAOs in EBPR WWTPs, our limited understanding of functional genes (discussed above) necessitates approaches that do not rely on complete apriori knowledge of the DNA sequence for a gene of interest. Otherwise put, we can use shotgun metagenomic approaches to identify, quantify relative abundances, and compare previously unknown DNA sequences for a functional gene of interest. Shotgun metagenomic techniques are generally classified into either read- or assembly-based approaches. Assembly-based approaches have benefits over read-based approaches, which compare short (<300 bp) reads generated in sequencing directly to reference databases for the purpose of annotation (i.e., assigning them to a known functional gene and/or organism). The processes of assembling short reads into longer ones (1) can remove errors incorporated during sequencing, (2) results in more annotations, (3) decreases the rate of false positives, and may (4) overcome some of the issues of primer bias or chimeras associated with amplicon sequencing (Wommack et al. 2008; Howe et al. 2014; Freitas et al. 2015; Gao et al. 2016). Assembly-based approaches also have their challenges, which include (1) high computational requirements and costs as well as (2) an emphasis on dominant community members (Howe et al. 2014). However, recently published tools help implement shotgun metagenomic assembly (Howe and Chain 2015) and run targeted assembly with specific functional genes in mind (Wang et al. 2015). It is likely that the number of tools to aid in shotgun metagenomic assembly will increase over time as more

researchers adopt these technologies (Howe et al. 2014). Several EBPR studies have used shotgun metagenomics to assemble genomes of PAOs, verify known metabolic pathways, and highlight potential metabolic capabilities (Kristiansen et al. 2013; Skennerton et al. 2014; Soo et al. 2014; Mao et al. 2016). Shotgun metagenomics has been used to explore the genetic traits of PAO metagenomes (García Martín et al. 2006; Albertsen et al. 2012) and compare genomes of different organisms within EBPR systems (Flowers et al. 2013). Next steps may include using assembled shotgun metagenomes to develop new FISH probes or qPCR primers for non-CAP PAOs. More specifically, researchers may focus on characterizing and quantifying the relative abundance of non-CAP *ppk* and *ppx* genes since both are required for polyP synthesis and degradation, respectively. New computational tools like Kaiju (Menzel et al. 2016) allow users to assign taxonomies to assembled shotgun metagenomic sequencing reads but have not yet been applied to EBPR communities. Overall, there are many opportunities to apply this sequencing approach along with other tools discussed previously to study the abundance and diversity of various P-associated functional genes in EBPR WWTPs.

Amplicon sequencing-based studies have primarily focused on exploring the relatedness of bacterial communities in EBPR settings (Zhang et al. 2003) but more research is needed to study non-CAP PAOs (Gebremariam et al. 2011). Also, work on PAOs beyond the bacterial domain is needed to ensure characterization of eukaryotic and archaeal PAOs (as well as PAO symbionts and competitors) in engineered systems. In this vein, one study demonstrated the relationship between PAO bacteria and non-PAO eukaryotes over time in a laboratory photosynthetic reactor (Oyserman et al. 2017). Another documented the lineages of bacterial glycogen accumulating organism (GAO)–competitors to PAOs—in a laboratory EBPR reactor (Nobu et al. 2014). In

summary, many studies have focused on CAP community diversity, structure, and function in EBPR systems. However, more research is needed to utilize the tools we discussed above (or a combination of them) to study non-CAP bacterial, eukaryotic, and archaeal communities in EBPR WWTPs—including both PAOs and non-PAOs.

Support for Polyphosphate Accumulating Organisms in Natural Systems

In addition to advancing our understanding of key biotic mechanisms controlling P cycling in engineered systems, studies of EBPR bacterial communities have uncovered potential traits that may prove important for PAO survival in the natural settings. Using a combination of laboratory studies and molecular biology approaches, researchers have found that PAOs can use many different C sources—not just acetate. For example, specific CAP groups have the capacity to use ethanol, rather than acetate (Skennerton et al. 2014). Non-CAP PAOs can use C substrates such as glucose (Nakamura et al. 1995b; Nguyen et al. 2011; Kristiansen et al. 2013), sugar alcohols (Nakamura et al. 1995b), amino acids (Nakamura et al. 1995b; Nguyen et al. 2011), and propionate (Lemos et al. 2003). Molecular biology approaches have revealed that certain CAP clades have N and C fixation genes (García Martín et al. 2006; Skennerton et al. 2014), denitrification genes (Kristiansen et al. 2013; Skennerton et al. 2014), genes for low (*pit*) and high (*pst*) affinity phosphate transporters (García Martín et al. 2006; Mao et al. 2014; Skennerton et al. 2014), and genes coding for flagella (García Martín et al. 2006). Beyond gene presence, one EBPR study demonstrated that a specific CAP group was capable of coupled nitrate reduction and phosphate uptake (Flowers et al. 2009). *Thiothrix caldifontis* use PHA and polysulfide as an energy source during anaerobic conditions (Rubio-Rincón et al. 2017). There is also evidence supporting symbiotic relationships between CAP and phototrophic eukaryotes in

an SBR exposed to alternating light/dark periods; *Cyanobacteria* supplies O_2 to drive polyP storage by CAP (Oyserman et al. 2017). When combined, these findings provide support for the idea that EBPR PAO may have originated from or may be well suited for mobility and survival in oligotrophic, natural systems (García Martín et al. 2006; Skennerton et al. 2014). These findings further highlight the need for (1) additional research on PAO metabolism, functional genes, and communities as well as (2) discussions and collaborations between researchers in engineered and natural systems as depicted in Figure 1.1.

BIOTIC CONTROLS ON PHOSPHORUS CYCLING IN NATURAL SYSTEMS

Concurrently to studies in EBPR WWTPs, researchers in the natural sciences documented the impacts of excess anthropogenic P on freshwater systems (e.g., Schindler 1977) and came to the understanding that P is typically a limiting nutrient in these environments (Schindler 1977; Carpenter et al. 1998; Elser et al. 2007; Schindler 2012; Dodds and Smith 2016). Like researchers who studied EBPR systems, natural scientists were motivated to uncover mechanisms helped reduce excess P as well as the negative impacts of subsequent eutrophication caused by P additions. From an environmental perspective, the negative impacts of eutrophication include decreased water transparency, potential growth of toxin producing cyanobacteria, hypoxic (i.e. low O_2) or anoxic (i.e. no O_2) conditions, and fish die-offs (Carpenter et al. 1998; Seviour et al. 2003; Carpenter 2005; Dodds et al. 2009; Hudnell 2010). Others have also documented the negative economic impacts of eutrophication. These include increased spending on drinking water treatment, recreational space closures, loss of waterfront real-estate, increased spending to manage threatened and endangered species, and decreased fish

and wildlife production (Seviour et al. 2003; Carpenter 2005; Dodds et al. 2009; MacDonald et al. 2016).

Early-on, researchers highlighted the influence of biotic (here microbial) controls on P cycling in natural systems (e.g., Fleischer 1978) and drew on research in engineering and medical fields to suggest a potential microbial mechanism for observed P patterns (e.g., Barnard 1976; Kornberg 1995). When specific biological mechanisms were addressed, they centered on the ability of microorganisms to make and break stored intracellular polyP. In the specific case of EBPR PAOs, researchers proposed that alternating anaerobic/aerobic conditions in the upper layers of soils and sediments as well as near the hypolimnion/redoxcline of lakes lead to polyP degradation/synthesis, respectively, by EBPR-like PAOs (Gächter et al. 1988; Davelaar 1993; Reddy et al. 1999; Schulz and Schulz 2005; Pett-Ridge and Firestone 2005; Hupfer et al. 2007; Hupfer et al. 2008; Hupfer and Lewandowski 2008; Peterson et al. 2008; Diaz et al. 2012; McMahon and Read 2013; McParland et al. 2015). Otherwise put, they suggested that naturally occurring PAOs, like their EBPR counterparts, outcompete other organisms because of their ability to break down intracellular polyP (i.e., release phosphate) during anaerobic conditions and store intracellular polyP (i.e., uptake phosphate) during aerobic conditions. This hypothesis assumes that PAOs like the ones found in EBPR WWTPs are (1) present in the natural environment and (2) influence P cycling under fluctuating environmental conditions. Therefore, we synthesize evidence from freshwater, marine, and soil systems that supports the presence and impact of PAOs in natural settings. We also highlight knowledge gaps, tools to overcome these gaps, and how new information uncovered by these tools may be beneficial to the co-advancement of research and water quality management in engineered and natural systems.

Freshwater Systems

Freshwater systems discussed here include the water column and sediments of streams and lakes. In streams, wetlands, and shallow regions of lakes, alternating anaerobic and aerobic periods are primarily driven by diel cycles of primary production; O_2 levels in the water column and upper sediment layers increase during the day due to photosynthesis, while during the night, the cessation of photosynthesis combined with continued respiration decreases O_2 levels (Dodds 2003; Cohen et al. 2013). In the rare case of freshwater tidal wetlands, alternating anaerobic/aerobic periods are also driven by the tide; the tide supplies O_2 rich water and recedes with water that has a lower concentration of O_2 due to respiration within the wetland (Findlay and Fischer 2013). In deeper parts of lakes, alternating anaerobic/aerobic periods are primarily driven by changes in the depths of the oxic epilimnion and anoxic hypolimnion due to wind and waves (McMahon and Read 2013).

Many studies have noted the inverse relationship between O_2 and phosphate concentrations in the water column and near the water column-sediment boundary (Carlton and Wetzel 1988; Gächter et al. 1988; Cohen et al. 2013; Read et al. 2014; Sherson et al. 2015; Saia et al. 2017) and have connected these patterns of diel P uptake and release with biotic processes. For example, a study of diel P cycling in the Ichetucknee River, Florida, USA determined that biological processes made up 66% of P uptake and release with the remaining being attributed to Ca-P precipitation and dilution (Cohen et al. 2013). Earlier laboratory studies subjecting the upper layers of sediment to alternating anaerobic/aerobic conditions demonstrated microbially-mediated P release into the water column during anaerobic conditions and microbially-mediated P uptake into the water column during aerobic conditions (Fleischer 1978; Carlton and Wetzel

1988; Gächter et al. 1988). However, specific microbial mechanisms were not identified in these studies.

Later studies focused on PAO-mediated control of P cycling in freshwater systems by highlighting the ability for organisms to store polyP (e.g., Rier et al. 2016) and/or the identification of known EBPR PAOs (Kunin et al. 2008; Peterson et al. 2008; Graham et al. 2017). Microscopy-based approaches have been used to identify intracellular polyP granules in sediment bacteria using (Uhlmann and Bauer 1988; Hupfer and Gächter 1995; Hupfer et al. 2004) and stream biofilms (Locke 2015; Rier et al. 2016; Saia et al. 2017). Other studies have used ^{31}P nuclear magnetic resonance (^{31}P-NMR) spectroscopy (Hupfer and Gächter 1995; Read et al. 2014) or other extractions (Eixler et al. 2005; Martin and Van Mooy 2013; Price and Carrick 2013; Price and Carrick 2014) to measure intracellular polyP in freshwater systems. One ^{31}P-NMR study demonstrated that polyP concentrations in a eutrophic lake taken over time were not significantly linked to O_2 availability but were highly variable in the anoxic hypolimnion compared to the oxic epilimnion (Read et al. 2014). CAP genes were identified in sediment and water samples in Wisconsin, USA and California, USA (Kunin et al. 2008; Peterson et al. 2008; Martins et al. 2011) and a study of Portuguese lake sediments found CAP made up < 0.3% of the bacterial community (Martins et al. 2011). A very small number of studies in freshwater systems document the PAO phenotype of phosphate storage as polyP under aerobic conditions and phosphate release from the breakdown of polyP during anaerobic conditions (Martins et al. 2011; Amirbahman et al. 2013; Saia et al. 2017). Of these, one identified known EBPR PAOs (Martins et al. 2011). These studies support suggestions that polyP accumulation and breakdown is ubiquitous (Kornberg 1995; Hupfer et al. 2007; Rao et al. 2009; Seviour and Nielsen 2010) and

highlight (1) the presence of known EBPR PAOs as well as unknown PAOs, (2) support of the EBPR PAO phenotype in natural systems, and (3) the potential impact of PAOs on P cycling in freshwater environments.

Marine Systems

Marine systems discussed here include estuaries, coastal waters, and the open ocean. In near-shore areas such as estuaries and the coast, alterations between anaerobic and aerobic conditions are greatly influenced by a combination of anthropogenic nutrient inputs—including their acceleration of primary production (Diaz and Rosenberg 2008)—and the mixing of stratified layers of the water column (Helm et al. 2011). In the open ocean, much like the deeper regions of lakes, alterations in anaerobic/aerobic conditions are driven by the mixing of stratified chemoclines caused by waves and wind (Helm et al. 2011).

Compared to freshwater systems, the characterization of microbially-mediated P cycling in marine environments as related to PAOs is just beginning (Bjorkman 2014; Karl 2014). A number of studies have identified intracellular polyP granules in microorganisms found in marine waters and sediments (Schulz and Schulz 2005; Brock and Schulz-Vogt 2011; Diaz et al. 2012; Martin et al. 2014; Zhang et al. 2015; Diaz et al. 2016). Researchers observed polyP accumulation by filamentous cyanobacteria (*Leptolyngbya sp.)* symbionts within marine sponges (Zhang et al. 2015). This same study verified the presence of *ppk* genes in these filamentous using degenerate PCR primers that were originally developed for EBPR PAOs. Another study found that phytoplankton accumulated more polyP in P-depleted regions of the Sargasso Sea

compared to regions that were more P-rich (Martin et al. 2014). Only a few researchers have linked O_2 availability in the water column and sediments with some of the advancements learned from PAO-mediated P cycling in engineered systems. Namely, one study found that phosphate concentrations were ~3x greater in the redoxcline of a coastal basin compared to the surface (McParland et al. 2015). Another found that a decline in polyP in water samples from a coastal inlet was correlated with the emergence of hypoxia (Diaz et al. 2012). One study observed that the giant sulfur bacteria—identified as *Thiomargarita namibiensis*—accumulated polyP under oxic sediment conditions and released phosphate under anoxic sediment conditions similar to known EBPR PAOs (Schulz and Schulz 2005). While not focusing specifically on PAOs, researchers collected water column samples along an urban estuary transect (i.e., from inland to the open ocean) and verified the presence of known EBPR PAO bacterial classes as well as observed higher abundances of P metabolism-associated genes in the open ocean using, respectively, amplicon and shotgun metagenomic sequencing approaches (Jeffries et al. 2016). By applying shotgun metagenomic sequencing techniques on samples taken globally from the open ocean, researchers found that the abundances of *ppk*, *ppx*, and *pst* genes were inversely proportional to P availability in marine environments (Temperton et al. 2011).

While not as numerous, studies of P cycling in marine systems expand our view of the many diverse microorganisms participating in P uptake and release as well as how changing environmental conditions (here, O_2 and P availability) influence these biotic controls. Specifically, several studies noted the inverse relationship between P availability and polyP accumulation as well as P availability and P functional gene abundance, which supports the

hypothesis that polyP storage provides organisms with the ability to conserve energy and nutrients for future use.

Soils

It is fairly well understood that soil water saturation is proportional to O_2 limitation because the soil pores restrict gas movement such that O_2 demand is larger than O_2 diffusion (Smith and Tiedje 1979; Silver et al. 1999; Chacon et al. 2008). Therefore, hydrological processes that alter soil wetting/drying events lead to subsequent alterations in anaerobic and aerobic conditions (Silver et al. 1999; Pett-Ridge and Firestone 2005; Chacon et al. 2008; Burgin and Groffman 2012; Peralta et al. 2014). Besides O_2 availability, scientists have suggested that hydrological processes may also influence PAO-facilitated P cycling and PAO survival in soils (Davelaar 1993; Pett-Ridge and Firestone 2005). Specifically, as soils become saturated by rainfall, O_2 diffusion is reduced, soils become anaerobic, and PAOs release phosphate that was previously stored as intracellular polyP. The converse may occur in drying soils; soils dry out, become aerobic, PAOs take up phosphate, and store it as intracellular polyP.

In terms of studies supporting the presence and role of PAO-mediated P cycling in soils, there are fewer examples as compared to all other natural systems. Three studies found evidence of CAP in soil (Kunin et al. 2008; Valdivia 2009; Archibald 2010) and two studies identified new PAOs in plant rhizosphere soil (DebRoy et al. 2013; Li et al. 2013). Bacteria occupying similar phylogenetic classes as known PAOs have been identified in soil (Pett-Ridge and Firestone 2005; DeAngelis et al. 2010). With respect to the role of O_2 availability and PAO-mediated P

cycling, soil mesocosms undergoing four hour alternating anaerobic/aerobic conditions had more *Betaproteobacteria* phylogenic assignments compared to the control (Pett-Ridge and Firestone 2005). The model EBPR PAO, CAP, is a member of this class and while this association does not require these soil *Betaproteobacteria* assignments to be CAP, this finding highlights the potential to discover new PAOs in soil.

Research Methods and Opportunities in Natural Systems

Like engineered systems, more work is needed to expand our knowledge of PAO metabolism, functional genes, as well as community diversity, structure, and function in natural systems. Specific research areas include the study of (1) non-CAP PAOs (including bacteria, eukaryotes, and archaea), (2) uncommonly studied functional genes (e.g., *ppx*), and (3) microbial responses to changing environmental conditions (i.e., O_2 and P availability). These microbial responses include anything from changes microbially-mediated P uptake and release, to changes in functional gene abundance, to changes in PAO community diversity, structure, and function. Many, if not all, of the tools discussed previously with respect to engineered systems can be applied to fill in these knowledge gaps. There are two additional tools—not discussed previously—that may be useful in verifying the PAO phenotype and test hypotheses related to the impact of O_2 availability on their role in P cycling in natural systems.

First, nanometer-scale secondary ion mass spectrometry (NanoSIMS) connects high resolution microscopy methods with elemental analysis and can be used to study environmental samples (Herrmann et al. 2007; Mueller et al. 2013; Kruse et al. 2015). For a review of NanoSIMS and

comparison to other microscopy techniques, including TEM and MAR-FISH, see (Gao et al. 2016). NanoSIMS has been used to verify the presence of intracellular P stored as polyP (Sulu-Gambari et al. 2016) and has been used to precisely investigate the 2D and 3D spatial distribution of P in soil (Kruse et al. 2015). It cannot be used to discern specific P molecules (Kruse et al. 2015) but can be used to study the spatial correlation between P and metals such as Fe, Al, Ca, and Mg (Hoppe et al. 2013; Kruse et al. 2015). Therefore, in addition to verifying the presence of intracellular polyP granules, TEM and NanoSIMS technique may help expand our understanding of the role these metals play in stabilizing the negative charge of intracellular polyP. It may also help us understand the role of these metals in protein function; PPK1, which synthesizes intracellular polyP chains, requires Mg (Table 1.3).

Second, to assess knowledge gaps concerning the variation of P cycling in time and space, researchers may look toward recent advances in continuously logging water quality sensors (e.g., Cohen et al. 2013; Abell et al. 2013). In addition to being used to separate out abiotic and biotic controls on P, water quality sensors can be used to identify sources and pathways of nutrients in watersheds, quanitfy coupling between different nutrient cycles (e.g. between C and P) and measure water quality parameters across multiple scales including those that are more approapirate for microbial processes (Pellerin et al. 2016; Rode et al. 2016). However, it is important to calibrate sensors properly and be aware of potential barriers to their use (Fares et al. 2016; Rode et al. 2016).

With respect PAO responses to changing environmental conditions and how this may influence water quality, opportunities exist to demonstrate whether and how PAOs play a role in the P cycling of natural systems. In addressing the ecological role of PAOs, we can learn from previous studies exploring the role of microbes in N and C cycling. Specifically, a central question to the field of environmental microbial ecology is whether understanding the diversity, structure, and function of microbial communities will enable us to predict ecosystem scale processes (Wallenstein and Hall 2012; Graham et al. 2014; Bier et al. 2015; Graham et al. 2016; Shade 2017). Several researchers have proposed general frameworks for linking microbial- and ecosystem- scale processes (e.g., Schimel and Gulledge 1998; Wallenstein and Hall 2012; Prosser 2013; Nemergut et al. 2014; Bier et al. 2015; Martiny et al. 2015). Microbes were originally left out of traditional theories of community ecology (Poisot et al. 2013). However, ways to integrate them have been proposed (Prosser et al. 2007) and may benefit our understanding of PAO community structure and survival in natural systems. In terms of experimental design, steps must be taken to ensure that microbial and environmental measurement timescales are compatible (Bier et al. 2015; Battin et al. 2016), that measurement bias for/against active microbial community members is understood (Schimel and Gulledge 1998; Jones and Lennon 2010; Lennon and Jones 2011; Carini et al. 2016), and environmental variables (e.g., pH and temperature) that may influence microbial community diversity, structure, and function are accounted for in the experimental design (Schimel and Gulledge 1998; Fierer and Jackson 2006; Lauber et al. 2009; Rousk et al. 2011; Bier et al. 2015; Battin et al. 2016; Dinh et al. 2016; Oliverio et al. 2016). Given the goal of the experiment, it may also be necessary to consider microbial P cycling in the context of C and N cycling (Burgin et al. 2011; Bernhardt et al. 2017; Iho et al. 2017) as well as the impact of minerals (e.g., Fe- and Al-oxides), cations

(e.g., Mg^{2+} and Ca^{2+}), and organic matter (OM) on P availability under alternating anaerobic and aerobic conditions (Hem 1963; Fleischer 1978; Bloom 1981; Carlton and Wetzel 1988; Dillon and Molot 1997; Hongve 1997; Lyons et al. 1998; Guppy et al. 2005; Hupfer and Lewandowski 2008; White et al. 2008; Withers and Jarvie 2008; Gerke 2010a; Bird et al. 2011; Tye et al. 2016).

When designing an experiment, researchers may also benefit from asking: Is the goal of the study to describe a microbial community/trait or will the study test a hypothesis? Discovery-based approaches, which help describe microbial communities and their function, can improve current methods of P gene quantification and extend our general understanding of the diversity and abundance of organisms and their P cycling genes in soils. They can also be used to develop PCR primers or FISH probes for new PAOs since current primers and probes mainly target only CAP PAOs (Table 1.4 and Table 1.5). Hypothesis-based approaches enable us to test for the influence of microbes on P availability given changes in environmental conditions like redox/O_2 availability by defining hypotheses (Prosser 2013; Poisot et al. 2013), using statistical techniques to test for significant linkages between microbial- and ecosystem-scale processes (Schimel and Gulledge 1998; Rocca et al. 2015; Bier et al. 2015; Bernhardt et al. 2017), and avoiding type II errors (Bier et al. 2015). A meta-analysis of over 400 studies found very few that used statistical approaches to test for a link between gene abundance determined via qPCR and C and N cycling rates (Rocca et al. 2014). Determining whether PAO-facilitated P cycling impacts water quality necessitates approaches that quantitatively test for this impact. Researchers are also encouraged to think about the specific statistical analysis methods they use. Correlation analysis is common but multivariate statistics and structural equation modeling may assist interpretations (Bier et al.

2015). The GUide to STatistical Analysis in Microbial Ecology (GUSTA ME; Buttigieg and Ramette 2014) is a useful resource for microbial ecologists with questions about statistical analysis.

If researchers decide to use next generation sequencing techniques, they may look to previous studies for guidance on methods (Riesenfeld et al. 2004; Kozich et al. 2013; Ekblom and Wolf 2014; Zimmerman et al. 2014; Pallen 2016), bioinformatics (Cock et al. 2009; Loman and Watson 2013; Howe et al. 2014; Shade and Teal 2015), and reproducible research (da Veiga Leprovost et al. 2014; Howe and Chain 2015; Shade and Teal 2015; Perez-Riverol et al. 2016). For soil studies, the RefSoil database can be used to annotate assembled shotgun soil metagenomes (Choi et al. 2016), which may prove important for studying PAO phyla and functional genes. As mentioned previously, there is only one study quantifying a wide range (i.e., not just CAP-related genes) of ppk and ppx genes in the environment (Temperton et al. 2011). Additionally, both amplicon and shotgun metagenomic sequencing approaches may improve our understanding of eukaryotic and archaeal PAOs. Finally, given the overwhelming amounts of data generated using next generation sequencing techniques, determining the ecological role of PAOs may necessitate testable hypotheses focusing on microbial traits such as the frameworks discussed by others (Wallenstein and Hall 2012; Fierer et al. 2014; Martiny et al. 2015).

Mechanistic modeling may also prove helpful in extending the impact of PAOs to larger scales. We can learn a great deal from the approaches of studies such as those incorporating microbial processes into global C and N models (Todd-Brown et al. 2012; Wieder et al. 2013; Graham et

al. 2014; Manzoni et al. 2014; Reed et al. 2014; Powell et al. 2015; Graham et al. 2016). Commonly used watershed P transport models such as the Soil and Water Assessment Tool (SWAT; Gassman et al. 2007) do not explicitly incorporate microbial processes but doing so may prove helpful in evaluating the connection between microbial processes, ecosystem P processes, and water quality.

Overall, there has been some feedback between natural and engineered systems with respect to the hypothesis that EBPR-like PAOs are present in the natural environment and are influencing P cycling; however, there are many opportunities to study PAO-mediated P cycling in natural systems. For example, soil microbial communities are considered some of the most complex (Dunbar et al. 2002; Gans et al. 2005; Tringe et al. 2005; Fierer and Jackson 2006; Schloss and Handelsman 2006; Howe et al. 2014; Hug et al. 2016; Nesme et al. 2016) and, when included, soil organisms greatly expanded the tree of life (Hug et al. 2016). Microbial soil habitats are also some of the most nutrient limited (Young and Crawford 2004) so the ability to store P and C at times of need, might be beneficial for PAOs. Therefore, research targeting PAO communities in the environment will broaden our overall understanding of PAO metabolism, functional genes, communities, and potentially enable us to apply this knowledge to (1) improve the functionality of EBPR WWTP and (2) reduce the impact of environmental P pollution.

CONSIDERING BIOTIC CONTROLS IN AGRICULTURAL SYSTEMS

Agricultural systems integrate traits of engineered and natural systems; conventional agriculture takes place within the context of natural systems and is ‘managed’ to achieve a specific goal

(e.g., maximum crop yield per acre) much like WWTPs are managed to ensure treatment of water to a predetermined standard. Therefore, we propose that agricultural systems represent an ideal setting to apply new knowledge of biotic controls on P cycling to improve and protect water quality (Figure 1.1). In addition, most P making its way to water bodies originates from terrestrial anthropogenic sources (Carpenter et al. 1998; Smil 2000; Bennett et al. 2001; Dodds et al. 2009; Dubrovsky et al. 2010; Maavara et al. 2015; Metson and Bennett 2015). Globally, P has accumulated in upland soils and river sediments at 3x the rate of pre-industrial times for both developing and developed nations (Bennett et al. 2001) and anthropogenic sources can account for over 50% of the average total P flux in rivers (Maavara et al. 2015). Of all anthropogenic sources, agriculture contributes a substantial—and often the largest—fraction of P loading to water bodies due to applications of inorganic fertilizer and manure to farmland (Bennett et al. 2001; Alexander et al. 2008; Dubrovsky et al. 2010). For example, croplands, pasturelands, and rangelands cumulatively delivered approximately 80% of P to the Gulf of Mexico compared to all other land uses (Alexander et al. 2008).

Current Agricultural Strategies for Reducing Phosphorus Pollution

Agricultural water quality management strategies typically focus on reducing P transport by behavioral and abiotic means (i.e., chemically- and physically-based controls). The behavioral approach encourages farmers to use the "4R's": right source (i.e. balance nutrient inputs with crop needs), right timing, right placement, and right (application) rate (Good and Beatty 2011). It is recommended that P forms (e.g. inorganic fertilizer and manure) are applied at a time and place in the landscape where the probability of runoff generation is low (i.e. rainfall probability is low) and where the application rate of P balances what the crop needs with what the soil

provides. Chemically-based management strategies utilize soil amendments rich in Ca, Fe, and Al to adsorb P (Moore and Miller 1994; Stout et al. 1998; Callahan et al. 2002; Bryant et al. 2012; Buda et al. 2012). Bioreactors with biochar, which has a high surface that immobilizes phosphate, may also provide some chemically-based P retention for surface and subsurface flows (Bock et al. 2015; Sharrer et al. 2016). Because most of the soil P pool is bound to OM, clay, and minerals, physically-based P management strategies focus primarily on reducing soil and sediment transport (Sharpley and Menzel 1987; Carpenter et al. 1998; Bennett et al. 2001; Gregory et al. 2007; Brady and Weil 2008; Prestigiacomo et al. 2016). Specific physically-based strategies for reducing erosion and/or promoting the settling of sediments include: the installation of constructed wetlands (Bergström et al. 2015), vegetated buffer strips (Tomer et al. 2014; Rittenburg et al. 2015), vegetated drainage ditches (Moore et al. 2010), incorporating manure applications (Kleinman et al. 2009; Kovar et al. 2011; Bergström et al. 2015), aerating soils to improve water infiltration (Johnson et al. 2011), and planting cover crops to halt sediment transport during the non-growing season (Kleinman et al. 2005; Kovar et al. 2011; Bergström et al. 2015). Chemical and physical controls have also been combined to reduce P losses. For example, adding gypsum to a vegetated buffer improved P retention (Watts and Torbert 2009).

However, chemically- and physically- based management strategies do not always reduce P transport. At low application rates soil carbon (C) amendments reduce P losses in runoff because they improve soil aggregate stability; however, over application can cause soil to become a P source (McDowell and Sharpley 2003). Cover crops can lead to P accumulation at the surface if crops are left on the field (Jarvie et al. 2015; Jarvie et al. 2017). This leaves the surface soils at risk for becoming a P source during large storm events. No-tillage can enhance P transport due to

preferential flow through macropores unless soils are well drained (Kleinman et al. 2009; Verbree et al. 2010; Smith et al. 2015) and tile drains, used to improve infiltration rates, can increase soluble P losses if the trade-offs of coupled management strategies are not considered (Kleinman et al. 2015; Jarvie et al. 2015; Jarvie et al. 2017). Despite their application for long-term P removal, woodchips may leach P during the start-up phase of agricultural bioreactors (Sharrer et al. 2016). Vegetated buffers and riparian soils can become sources of P (Lyons et al. 1998; Young and Briggs 2008; Walter et al. 2009; Vidon et al. 2010; Young and Ross 2016) as can wetlands that are overwhelmed with high flows (Mitsch et al. 1995) or have accumulated a lot of sediment-bound P (Hill and Robinson 2012). Other studies have shown that arresting soil erosion is not enough to reduce P pollution because soluble P forms continue to be an issue (Jarvie et al. 2015; Smith et al. 2015; Garcia et al. 2016; Williams et al. 2016; Iho et al. 2017; Jarvie et al. 2017).

Biological management strategies of P in agricultural settings typically rely on the crop to reduce P transport from the field to downstream water bodies. As mentioned previously, the 4R approach balances P supply to the crop (i.e., amount of P fertilizer applied) with the P demand of the crop (i.e., amount of P the crop requires for growth) so very little P is left over to be transported to downstream water bodies (Good and Beatty 2011). Another strategy utilizes fungal and bacterial amendments, which excrete compounds that mobilize soil-bound P, to increase P availability to crops (Rodríguez and Fraga 1999; Jeffries et al. 2003; Javot et al. 2007; Hayat et al. 2010; Baas et al. 2016; Tapia-Torres et al. 2016). In addition to physically halting the transport of sediment as discussed previously, wetland vegetation serves as a P sink—although assimilation likely only lasts until vegetation decomposes (Richardson 1985; Reddy et

al. 1999). Wetland soils research also addresses the short-term role of microorganisms in P retention (Richardson 1985; Kellogg and Bridgham 2003; Noe et al. 2003; Scinto and Reddy 2003). In a similar vein, researchers have noted this limited bio-accumulation of P in streams; biofilms accumulating P are sloughed off under high flows (Dodds 2003). As discussed in detail above, researchers have studied the impact of biological controls on P retention natural systems but there are still many open questions concerning specific organisms, their metabolic pathways for retaining P, and whether/how these biological processes may aid water quality management.

The Role of Hydrology and Hydrologic Theory

As previously discussed, hydrological processes supply water, and thus, alter O_2 availability in soils and other natural settings. Hydrology may also supply limiting nutrients that microorganisms require (Hill et al. 2000; Groffman et al. 2009; Vidon et al. 2010; Bernhardt et al. 2017) and shape the diversity, structure, and function of microbial communities (Pett-Ridge and Firestone 2005; DeAngelis et al. 2010; Peralta et al. 2014). This is the guiding principle behind biogeochemical "hot spots", which were first described by McClain et al. (2003) as regions of the landscape with proportionally higher biogeochemical reaction rates compared to the surrounding landscape mainly, but not exclusively, due to their hydrologic connectivity in space. Additionally, a wealth of research has been done to characterize the location and timing of areas in the landscape that undergo seasonal wetting/drying cycles (Beven and Kirkby 1979; Frankenberger et al. 1999; Easton et al. 2008b; Easton et al. 2008a; Buchanan et al. 2014)—termed variable source areas (VSAs) by hydrologists (Hewlett and Hibbert 1967; Beven and Kirkby 1979; Dunne and Black 1970; Walter et al. 2000). VSAs commonly form in humid, well vegetated regions with steep slopes and shallow soils and may increase or decrease in size over

time and in space. VSAs are the main source of runoff in humid, hilly, regions such as the northeastern United States (Dunne and Black 1970; Walter et al. 2000).

There is a great deal of evidence suggesting the overlap of P sources with frequently saturating areas of the landscape (Walter et al. 2000; Hooda et al. 2000; Macrae et al. 2005; Gburek et al. 2002; Vidon et al. 2010; Franklin et al. 2013; Buchanan et al. 2013; Wilson et al. 2016) but specific mechanisms controlling P mobility in these areas of the landscape are poorly understood (Walter et al. 2009; Vidon et al. 2010; Bernhardt et al. 2017). This includes short-term mechanisms related to the wetting/drying of soils (including frequency thereof) as well as long-term, seasonal mechanisms. In terms of microbial controls on nutrient cycling, the majority of "hot spot" and VSA research has focused on how microbes influence C and N cycling (e.g., Hill et al. 2000; Pett-Ridge and Firestone 2005; DeAngelis et al. 2010; Ouyang and Li 2013; Peralta et al. 2014). More research on P cycling is needed as well as the interactions between P, C, N, and other micronutrients (Vidon et al. 2010). While not focusing on specific metabolic mechanisms, several studies have looked at the role of microbes on P cycling in soils undergoing alternating periods of saturation/redox (Grierson et al. 1998; Silver et al. 1999; Turner and Haygarth 2001; Olander and Vitousek 2000; Chacon et al. 2008; Blackwell et al. 2010; Yevdokimov et al. 2016) but more work must be done to determine the extent of PAO representation in soil microbial communities and how they influence P cycling as it applies to water quality management.

Given the important role of hydrology in influencing microbial processes and the transport of P, the strength of current hydrologic theory to consider abiotic and biotic controls on nutrient cycling (i.e., hot spot and VSA theory), and the limitations of current hydrologic theories with respect to P cycling, we suggest that early hypotheses and experimental designs addressing the role of PAOs in agricultural systems may wish to consider the various tools discussed above (e.g., microscopy, qPCR, metagenomic sequencing) in the context of already established hydrological theories. For example, researchers may wish to consider how the impact of PAO-mediated P cycling varies in time and space (1) across O_2 availability/soil moisture gradients, (2) across co-nutrient (e.g., C) gradients, (3) alongside other microbial process (e.g., denitrification), and (4) alongside abiotic controls on P.

Example Research Focus Areas and Questions for Agricultural Systems

1. *Legacy Soil P* - Legacy soil P is an ongoing issue in historically cultivated landscapes and necessitates more aggressive P management strategies because the continuous re-release of P to nearby water bodies can last for > 10 years even after fertilizer application has been completely stopped (Sharpley and Rekolainen 1997; Carpenter 2005; Gregory et al. 2007; Kleinman et al. 2011a; Sharpley et al. 2011; Sharpley et al. 2013; Jarvie et al. 2013; Haygarth et al. 2014; Powers et al. 2016; Cassidy et al. 2017). *Questions* - What role do PAOs play in liberating P from soils rich in legacy P? How do PAOs interact with phosphate solubilizing microbes to determine P availability? How do the roles of these microorganisms in P cycling compare to abiotic processes? Can PAOs be cultivated in existing management structures (e.g., bioreactors) to 'catch' legacy P before it is transported to water bodies? What is the role of PAOs in natural P barriers (e.g., stream biofilms) to legacy P loads? How do storm events impact this storage? How

can we incorporate this storage into management strategies and downstream algal bloom predictions?

2. *Rhizosphere Ecosystems and Irrigation* - Two studies documented the presence of PAOs in rhizosphere soil (DebRoy et al. 2013; Li et al. 2013). *Questions* - What percentage of rhizosphere soil microbial communities are made up of PAOs and what role (if any) do they play in crop growth? How does this vary with different crops or combinations of crops? How does their role compare to abiotic controls on P? Do crop irrigation strategies influence the abundance and function of PAO communities? What is the impact of PAOs on water quality under different irrigation regimes? How does this impact compare to abiotic controls on P?

3. *P Recycling* – Global mineral P stores are geographically limited (Jarvie et al. 2015) and many locations are beginning to consider P recycling programs for agricultural (Withers et al. 2014; Jarvie et al. 2015) as well as urban (Metson and Bennett 2015) systems. *Questions* - Can PAOs play a role in P recycling within an agricultural field? How do we optimize this role while also considering abiotic controls on P cycling?

SUMMARY

Anthropogenic P sources continue to impact the water quality of lakes, rivers, and streams (Carpenter et al. 1998; Carpenter 2005; Dodds et al. 2009; Dubrovsky et al. 2010; Hudnell 2010; Kleinman et al. 2011a; Sharpley et al. 2013; Obersteiner et al. 2013; Jarvie et al. 2013; Jarvie et al. 2015; Garcia et al. 2016; Jarvie et al. 2017) and it is plausible that effective management of

non-point source anthropogenic P has been difficult to achieve due to our limited understanding of the role microbial controls on P cycling in soils. In this review, we synthesize current evidence of PAO-mediated P cycling in engineered and natural systems (i.e., streams, lakes, soils, estuaries, and oceans), highlight knowledge gaps, and suggest tools as well as approaches to overcome these gaps. There is a great deal of opportunity to explore the presence and ecological role of PAOs in natural systems. Further, the integration of these findings with existing knowledge of engineered and agricultural systems may provide a more holistic approach for reducing P pollution and improving water quality in the environment.

ACKNOWLEDGMENTS

The author would like to thank Edward Hall, Anthony Buda, Hunter Carrick, Brian Rahm, Miranda Stockton, and Claudia Rojas for their critical feedback. This publication was developed under STAR Fellowship Assistance Agreement no. FP917670-01-0 awarded by the US Environmental Protection Agency (EPA). It has not been formally reviewed by the EPA. The views expressed in this publication are solely those of SMS and EPA does not endorse any products or commercial services mentioned in this publication.

REFERENCES

Abell, J.M., D.P. Hamilton and J.C. Rutherford. 2013. Quantifying temporal and spatial variations in sediment, nitrogen and phosphorus transport in stream inflows to a large eutrophic lake. Environmental Science Processes and Impacts. 15:1137–1152.

Achbergerová, L. and J. Nahálka. 2011. Polyphosphate - an ancient energy source and active metabolic regulator. Microbial Cell Factories. 10(1):63.

Archibald, J. 2010. Dissolved phosphorus dynamics in shallow groundwater. Master's Thesis, Cornell University.

Ahn, K. and A. Kornberg. 1990. Polyphosphate kinase from *Escherichia coli*. Biochemistry. 265(20):11734–11739.

Akiyama, M., E. Crooke and A. Kornberg. 1993. An exopolyphosphatase of *Escherichia coli*. The Journal of Biological Chemistry. 268(1):633–639.

Akiyama, M., E. Crookes and A. Kornberg. 1992. The polyphosphate kinase gene of *Escherichia coli*. The Journal of Biological Chemistry. 267(31):22556–22561.

Albertsen, M., L.B.S. Hansen, A.M. Saunders, P.H. Nielsen and K.L. Nielsen. 2012. A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. The ISME journal. 6(6):1094–1106.

Alcántara, C., A. Blasco, M. Zúñiga and V. Monedero. 2014. Accumulation of polyphosphate in *Lactobacillus* spp. and its involvement in stress resistance. Applied and Environmental Microbiology. 80(5):1650–1659.

Alexander, R.B., R.A. Smith, G.E. Schwarz, E.W. Boyer, J. V. Nolan and J.W. Brakebill. 2008. Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin. Environmental Science and Technology. 42(3):822–830.

Amann, R., B.M. Fuchs and S. Behrens. 2001. The identification of microorganisms by fluorescence in situ hybridization. Current Opinion in Biotechnology. 12(3):231–236.

Amirbahman, A., B.A. Lake and S.A. Norton. 2013. Seasonal phosphorus dynamics in the surficial sediment of two shallow temperate lakes: a solid-phase and pore-water study. Hydrobiologia. 701(1):65–77.

Aschar-Sobbi, R., A.Y. Abramov, C. Diao, M.E. Kargacin, G.J. Kargacin, R.J. French and E. Pavlov. 2008. High sensitivity, quantitative measurements of polyphosphate using a new DAPI-based approach. Journal of Fluorescence. 18(5):859–866.

Auling, G., F. Pilz, H.J. Busse, S. Karrasch, M. Streichan and G. Schon. 1991. Analysis of the polyphosphate-accumulating microflora in phosphorus- eliminating, anaerobic-aerobic activated sludge systems by using diaminopropane as a biomarker for rapid estimation of *Acinetobacter* spp. Applied and Environmental Microbiology. 57(12):3585–3592.

Baas, P., C. Bell, L.M. Mancini, M.N. Lee, R.T. Conant and M.D. Wallenstein. 2016. Phosphorus mobilizing consortium Mammoth P TM enhances plant growth. PeerJ. 4:e2121.

Bark, K., P. Kampfer, A. Sponner and W. Dott. 1993. Polyphosphate-dependent enzymes in some coryneform bacteria isolated from sewage-sludge. FEMS Microbiology Letters. 107(2–3):133–138.

Barnard, J.L. 1976. A review of biological phosphorus removal in the activated sludge process. Water SA. 2(3):136–144.

Barnard, J.L. 1975. Biological nutrient removal without the addition of chemicals. Water Research. 9:485–490.

Battin, T.J., K. Besemer, M.M. Bengtsson, A.M. Romani and A.I. Packmann. 2016. The ecology and biogeochemistry of stream biofilms. Nature Reviews Microbiology. 14(4):251–263.
Beer, M., H.M. Stratton, P.C. Griffiths and R.J. Seviour. 2006. Which are the polyphosphate accumulating organisms in full-scale activated sludge enhanced biological phosphate removal systems in Australia? Journal of Applied Microbiology. 100(2):233–243.
Bennett, E.M., S.R. Carpenter and N.F. Caraco. 2001. Human impacts on erodible phosphorus and eutrophication: a global perspective. BioScience. 51(3):227–234.
Bergström, L., H. Kirchmann, F. Djodjic, K. Kyllmar, B. Ulén, J. Liu, H. Andersson, H. Aronsson, G. Börjesson, P. Kynkäänniemi, A. Svanbäck and A. Villa. 2015. Turnover and losses of phosphorus in Swedish agricultural soils: long-term changes, leaching trends, and mitigation measures. Journal of Environmental Quality. 44(2):512–523.
Bernhardt, E.S., J.R. Blaszczak, C.D. Ficken, M.L. Fork, K.E. Kaiser and E.C. Seybold. 2017. Control points in ecosystems: moving beyond the hot spot hot moment concept. Ecosystems. 20(2):1–18.
Beven, K.J. and M.J. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin. 24(1):43–69.
Bier, R.L., E.S. Bernhardt, C.M. Boot, E.B. Graham, E.K. Hall, J.T. Lennon, D.R. Nemergut, B.B. Osborne, C. Ruiz-Gonzalez, J.P. Schimel, M.P. Waldrop and M.D. Wallenstein. 2015. Linking microbial community structure and microbial processes: an empirical and conceptual overview. FEMS Microbiology Ecology. 91(10):1–11.
Bird, L.J., V. Bonnefoy and D.K. Newman. 2011. Bioenergetic challenges of microbial iron metabolisms. Trends in Microbiology. 19(7):330–340.
Bjorkman, K.M. 2014. Polyphosphate goes from pedestrian to prominent in the marine P-cycle. Proceedings of the National Academy of Sciences. 111(22):7890–7891.
Blackwell, M.S.A., P.C. Brookes, N. de la Fuente-Martinez, H. Gordon, P.J. Murray, K.E. Snars, J.K. Williams, R. Bol and P.M. Haygarth. 2010. Phosphorus solubilization and potential transfer to surface waters from the soil microbial biomass following drying-rewetting and freezing-thawing 1st ed., Elsevier Inc.
Bloom, P.R. 1981. Phosphorus adsorption by an aluminum-peat complex. Soil Science Society of America Journal. 45(2):267–272.
Bock, E., N. Smith, M. Rogers, B. Coleman, M. Reiter, B. Benham and Z.M. Easton. 2015. Enhanced nitrate and phosphate removal in a denitrifying bioreactor with biochar. Journal of Environment Quality. 44(2):605–613.
Brady, N.C. and R.R. Weil. 2008. The Nature and Properties of Soils 14th ed.,
Brock, J. and H.N. Schulz-Vogt. 2011. Sulfide induces phosphate release from polyphosphate in cultures of a marine *Beggiatoa* strain. International Society for Microbial Ecology Journal. 5(3):497–506.
Brown, M.R.W. and A. Kornberg. 2008. The long and short of it - polyphosphate, PPK and bacterial survival. Trends in Biochemical Sciences. 33(6):284–290.
Bryant, R.B., A.R. Buda, P.J. a. Kleinman, C.D. Church, L.S. Saporito, G.J. Folmar, S. Bose and A.L. Allen. 2012. Using flue gas desulfurization gypsum to remove dissolved phosphorus from agricultural drainage waters. Journal of Environment Quality. 41(3):664.
Buchanan, B.P., J.A. Archibald, Z.M. Easton, S.B. Shaw, R.L. Schneider and M. Todd Walter. 2013. A phosphorus index that combines critical source areas and transport pathways using a travel time approach. Journal of Hydrology. 486:123–135.

Buchanan, B.P., M. Fleming, R.L. Schneider, B.K. Richards, J. Archibald, Z. Qiu and M.T. Walter. 2014. Evaluating topographic wetness indices across central New York agricultural landscapes. Hydrology and Earth System Sciences. 18(8):3279–3299.
Buda, A.R., G.F. Koopmans, R.B. Bryant and W.J. Chardon. 2012. Emerging technologies for removing nonpoint phosphorus from surface water and groundwater: introduction. Journal of Environment Quality. 41(3):621–627.
Burgin, A.J. and P.M. Groffman. 2012. Soil O2 controls denitrification rates and N2O yield in a riparian wetland. Journal of Geophysical Research. 117(1):1–10.
Burgin, A.J., W.H. Yang, S.K. Hamilton and W.L. Silver. 2011. Beyond carbon and nitrogen: how the microbial energy economy couples elemental cycles in diverse ecosystems. Frontiers in Ecology and the Environment. 9(1):44–52.
Buttigieg, P.L. and A. Ramette. 2014. A guide to statistical analysis in microbial ecology: a community-focused, living review of multivariate data analyses. FEMS Microbiology Ecology. 90(3):543–550.
Callahan, M.P., P.J.A. Kleinman, A.N. Sharpley and W.L. Stout. 2002. Assessing the efficacy of alternative phosphorus sorbing soil amendments. Soil Science. 167(8):539–547.
Carini, P., P.J. Marsden, J.W. Leff, E.E. Morgan, M.S. Strickland and N. Fierer. 2016. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nature Microbiology. 2:16242.
Carlton, R.G. and R.G. Wetzel. 1988. Phosphorus flux from lake sediments: effect of epipelic algal oxygen production. Limnology and Oceanography. 33(4):562–570.
Carpenter, S.R. 2005. Eutrophication of aquatic ecosystems: bi-stability and soil phosphorus. Proceedings of the National Academy of Sciences. 102(29):10002–10005.
Carpenter, S.R., N.F. Caraco, D.L. Correll, R.W. Howarth, A.N. Sharpley and V.N. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications. 8(3):559–568.
Cassidy, R., D.G. Doody and C.J. Watson. 2017. Impact of legacy soil phosphorus on losses in drainage and overland flow from grazed grassland soils. Science of The Total Environment. 575(1):474–484.
Chacon, N., N. Dezzeo, M. Rangel and S. Flores. 2008. Seasonal Changes in Soil Phosphorus Dynamics and Root Mass along a Flooded Tropical Forest Gradient in the Lower Orinoco River, Venezuela. Biogeochemistry. 87(2):157–168.
Choi, J., F. Yang, R. Stepanauskas, E. Cardenas, A. Garoutte, R. Williams, J. Flater, J.M. Tiedje, K.S. Hofmockel, B. Gelder and A. Howe. 2016. Strategies to improve reference databases for soil microbiomes. The ISME Journal. :1–6.
Cock, P.J.A., T. Antao, J.T. Chang, B.A. Chapman, C.J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski and M.J.L. De Hoon. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 25(11):1422–1423.
Cohen, M.J., M.J. Kurz, J.B. Heffernan, J.B. Martin, R.L. Douglass, C.R. Foster and R.G. Thomas. 2013. Diel phosphorus variation and the stoichiometry of ecosystem metabolism in a large spring-fed river. Ecological Monographs. 83(2):155–176.
Comeau, Y., K. Hall, R. Hancock and W. Oldham. 1986. Biochemical model for enhanced biological phosphorus removal. Water Research. 20(12):1511–1521.
Crocetti, G.R., P. Hugenholtz, P.L. Bond, A. Schuler, J. Keller, D. Jenkins, L. Linda, R.G. Keller and L.L. Blackall. 2000. Identification of polyphosphate-accumulating organisms and

design of 16S rRNA-directed probes for their detection and quantitation. Applied and Environmental Microbiology. 66(3):1175–1182.
Daims, H., A. Brühl, R. Amann, K.H. Schleifer and M. Wagner. 1999. The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. Systematic and Applied Microbiology. 22(3):434–44.
da Veiga Leprovost, F., V.C. Barbosa, E.L. Francisco, Y. Perez-Riverol and P.C. Carvalho. 2014. On best practices in the development of bioinformatics software. Frontiers in Genetics. 5:199.
Davelaar, D. 1993. Ecological significance of bacterial polyphosphate metabolism in sediments. Hydrobiologia. 253(1):179–192.
DeAngelis, K.M., W.L. Silver, A.W. Thompson and M.K. Firestone. 2010. Microbial communities acclimate to recurring changes in soil redox potential status. Environmental Microbiology. 12(12):3137–3149.
DebRoy, S., P. Mukherjee, S. Roy, A.R. Thakur and S. Raychaudhuri. 2013. Draft genome sequence of a phosphate-accumulating Bacillus sp .,. Genome Announcements. 1(1):1–2.
Diaz, J., E. Ingall, C. Benitez-Nelson, D. Paterson, M.D. de Jonge, I. Ian McNulty and J.A. Brandes. 2008. Marine polyphosphate: a key player in geologic phosphorus sequestration. Science. 320:652–655.
Diaz, J.M., K.M. Bjorkman, S.T. Haley, E.D. Ingall, D.M. Karl, A.F. Longo and S.T. Dyhrman. 2016. Polyphosphate dynamics at Station ALOHA, North Pacific subtropical gyre. Limnology and Oceanography. 61(1):227–239.
Diaz, J.M. and E.D. Ingall. 2010. Fluorometric quantification of natural inorganic polyphosphate. Environmental Science and Technology. 44(12):4665–4671.
Diaz, J.M., E.D. Ingall, S.D. Snow, C.R. Benitez-Nelson, M. Taillefert and J.A. Brandes. 2012. Potential role of inorganic polyphosphate in the cycling of phosphorus within the hypoxic water column of Effingham Inlet, British Columbia. Global Biogeochemical Cycles. 26(2):1–14.
Diaz, R.J. and R. Rosenberg. 2008. Spreading dead zones and consequences for marine ecosystems. Science. 321(5891):926–929.
Dillon, P.J. and L. a. Molot. 1997. Effect of landscape form on export of dissolved organic carbon, iron, and phosphorus from forested stream catchments. Water Resources Research. 33(11):2591–2600.
Dinh, M.-V., T. Schramm, M. Spohn and E. Matzner. 2016. Drying-rewetting cycles release phosphorus from forest soils. Journal of Plant Nutrition and Soil Science. 179(5):670–678.
Dodds, W.K. 2003. The role of periphyton in phosphorus retention in shallow freshwater aquatic systems. Journal of Phycology. 39(5):840–849.
Dodds, W.K., W.W. Bouska, J.L. Eitzmann, T.J. Pilger, K.L. Pitts, A.J. Riley, J.T. Schloesser and D.J. Thornbrugh. 2009. Eutrophication of U.S. freshwaters: analysis of potential economic damages. Environmental Science and Technology. 43(1):12–19.
Dodds, W.K. and V.H. Smith. 2016. Nitrogen, phosphorus, and eutrophication in streams. Inland Waters. 6(2):155–164.
Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.A.M. Gronberg, P.A. Hamilton, K.J. Hitt, D.K. Mueller, M.D. Munn, B.T. Nolan, L.J. Puckett, M.G. Rupert, T.M. Short, N.E. Spahr,

L.A. Sprague and W.G. Wilber. 2010. The quality of our Nation's waters - nutrients in the Nation 's streams and groundwater , 1992 – 2004,
Dunbar, J., S.M. Barns, L.O. Ticknor and C.R. Kuske. 2002. Empirical and Theoretical Bacterial Diversity in Four Arizona Soils Empirical and Theoretical Bacterial Diversity in Four Arizona Soils. Applied and Environmental Microbiology. 68(6):3035–3045.
Dunne, T. and R.D. Black. 1970. Partial area contributions to storm runoff in a small New England watershed. Water Resources Research. 6(5):1296–1311.
Easton, Z.M., D.R. Fuka, M.T. Walter, D.M. Cowan, E.M. Schneiderman and T.S. Steenhuis. 2008a. Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas. Journal of Hydrology. 348(3–4):279–291.
Easton, Z.M., P. Gérard-Marchant, M.T. Walter, a. M. Petrovic and T.S. Steenhuis. 2007. Hydrologic assessment of an urban variable source watershed in the northeast United States. Water Resources Research. 43(3):1–18.
Easton, Z.M., M.T. Walter and T.S. Steenhuis. 2008b. Combined monitoring and modeling indicate the most effective agricultural best management practices. Journal of Environmental Quality. 37(5):1798–1809.
Eixler, S., U. Selig and U. Karsten. 2005. Extraction and detection methods for polyphosphate storage in autotrophic planktonic organisms. Hydrobiologia. 533(1):135–143.
Ekblom, R. and J.B.W. Wolf. 2014. A field guide to whole-genome sequencing, assembly and annotation. Evolutionary Applications. 7(9):1026–1042.
Elser, J.J., M.E.S. Bracken, E.E. Cleland, D.S. Gruner, W.S. Harpole, H. Hillebrand, J.T. Ngai, E.W. Seabloom, J.B. Shurin and J.E. Smith. 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecology Letters. 10(12):1135–1142.
Fares, A., R. Awal and H. Bayabil. 2016. Soil water content sensor response to organic matter content under laboratory conditions. Sensors. 16(8):1239.
Fierer, N., A. Barberán and D.C. Laughlin. 2014. Seeing the forest for the genes: Using metagenomics to infer the aggregated traits of microbial communities. Frontiers in Microbiology. 5:614.
Fierer, N. and R.B. Jackson. 2006. The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences. 103(3):626–631.
Findlay, S. and D. Fischer. 2013. Ecosystem attributes related to tidal wetland effects on water quality. Ecology. 94(1):117–125.
Fleischer, S. 1978. Evidence for the anaerobic release of phosphorus from lake sediments as a biological process. Naturwissenschaften. 434:109–110.
Flowers, J.J., S. He, S. Malfatti, T.G. del Rio, S.G. Tringe, P. Hugenholtz and K.D. McMahon. 2013. Comparative genomics of two *Candidatus* Accumulibacter clades performing biological phosphorus removal. The ISME Journal. 7(12):2301–2314.
Flowers, J.J., S. He, S. Yilmaz, D.R. Noguera and K.D. McMahon. 2009. Denitrification capabilities of two biological phosphorus removal sludges dominated by different *Candidatus* Accumulibacter clades. Environmental Microbiology Reports. 1(6):583–588.
Frankenberger, J.R., E.S. Brooks, M.F.T. Walter, M.F.T. Walter and T.S. Steenhuis. 1999. A GIS-based variable source area hydrology model. Hydrological Processes. 13(6):805–822.
Franklin, D.H., J.L. Steiner, S.E. Duke, D.N. Moriasi and P.J. Starks. 2013. Spatial considerations in wet and dry periods for phosphorus in streams of the fort cobb

watershed, United States. Journal of the American Water Resources Association. 49(4):908–922.
Freitas, T.A.K., P.-E. Li, M.B. Scholz and P.S.G. Chain. 2015. Accurate read-based metagenome characterization using a hierarchical suite of unique signatures. Nucleic Acids Research. 43(10):1–14.
Fuchs, B.M., F.O. Glöckner, J. Wulf and F.O. Glo. 2000. Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes. Applied and Environmental Microbiology. 66(8):3603–3607.
Gächter, R., J.S. Meyer and A. Mares. 1988. Contribution of bacteria to release and fixation of phosphorus in lake sediments. Limnology and Oceanography. 33(6 part 2):1542–1558.
Gans, J., M. Wolinsky and J. Dunbar. 2005. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science. 309(5739):1387–1390.
Gao, D., X. Huang and Y. Tao. 2016. A critical review of NanoSIMS in analysis of microbial metabolic activities at single-cell level. Critical Reviews in Biotechnology. 36(5):884–890.
Garcia, A.M., R.B. Alexander, J.G. Arnold, L. Norfleet, M.J. White, D.M. Robertson and G. Schwarz. 2016. Regional effects of agricultural conservation practices on nutrient transport in the Upper Mississippi River Basin. Environmental Science and Technology. 50(13):6991–7000.
García Martín, H., N. Ivanova, V. Kunin, F. Warnecke, K.W. Barry, A.C. McHardy, C. Yeates, S. He, A. a Salamov, E. Szeto, E. Dalin, N.H. Putnam, H.J. Shapiro, J.L. Pangilinan, I. Rigoutsos, N.C. Kyrpides, L.L. Blackall, K.D. McMahon and P. Hugenholtz. 2006. Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. Nature Biotechnology. 24(10):1263–1269.
Gassman, P.P.W., M.M.R. Reyes, C.C.H. Green and J.J.G. Arnold. 2007. The Soil and Water Assessment Tool: historical development, applications, and future research directions. Transactions of the ASAE. 50(4):1211–1250.
Gburek, W.J., C.C. Drungil, M.S. Srivivasan, B.A. Needelman and D.E. Woodward. 2002. Variable-source-area controls on phosphorus transport: Bridging the gap between research and design. Journal of Soil and Water Conservation. 57(6):534–543.
Gebremariam, S.Y., M.W Beutel, D. Christian, T.F. Hess. 2011. Research advances and challenges in the microbiology of enhanced biological phosphorus removal—a critical review. Water Environmental Research. 83(3):195– 219.
Gerke, J. 2010. Humic (Organic Matter)-Al(Fe)-Phosphate Complexes. Soil Science. 175(9):417–425.
Good, A.G. and P.H. Beatty. 2011. Fertilizing nature: A tragedy of excess in the commons. PLOS Biology. 9(8):1–9.
Graham, E.B., W.R. Wieder, J.W. Leff, S.R. Weintraub, A.R. Townsend, C.C. Cleveland, L. Philippot and D.R. Nemergut. 2014. Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes. Soil Biology and Biochemistry. 68:279–282.
Graham, E.B. et al. 2016. Microbes as engines of ecosystem function: when does community structure enhance predictions of ecosystem processes? Frontiers in Microbiology. 7:214.
Graham, E.B., A.R. Crump, C.T. Resch, S. Fansler, E. Arntzen, W. Kennedy, J.K. Fredrickson and J.C. Stegen. 2017. Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. Environmental Microbiology. 19(4):1552–1567.

Gregory, S., A.W. Allen, M. Baker, K. Boyer, T. Dillaha and J. Elliot. 2007. Realistic expectations of timing between conservation and restoration actions and ecological responses. In Managing agricultural landscapes for environmental quality: strengthening the science base. Ankeny, IA: Soil and Water Conservation Society, :111–142.

Grierson, P.F., N.B. Comerford and E.J. Jokela. 1998. Phosphorus mineralization kinetics and response of microbial phosphorus to drying and rewetting in a Florida Spodosol. Soil Biology and Biochemistry. 30(10–11):1323–1331.

Groffman, P.M., K. Butterbach-Bahl, R.W. Fulweiler, A.J. Gold, J.L. Morse, E.K. Stander, C. Tague, C. Tonitto and P. Vidon. 2009. Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. Biogeochemistry. 93(1–2):49–77.

Günther, S., M. Trutnau, S. Kleinsteuber, G. Hause, T. Bley, I. Röske, H. Harms and S. Müller. 2009. Dynamics of polyphosphate-accumulating bacteria in wastewater treatment plant microbial communities detected via DAPI (4',6'-diamidino-2- phenylindole) and tetracycline labeling. Applied and Environmental Microbiology. 75(7):2111–2121.

Guppy, C.N., N.W. Menzies, P.W. Moody and F.P.C. Blamey. 2005. Competitive sorption reactions between phosphorus and organic matter in soil: a review. Australian Journal of Soil Research. 43(2):189.

Hayat, R., S. Ali, U. Amara, R. Khalid and I. Ahmed. 2010. Soil beneficial bacteria and their role in plant growth promotion: a review. Annals of Microbiology. 60(4):579–598.

Haygarth, P.M., H.P. Jarvie, S.M. Powers, A.N. Sharpley, J.J. Elser, J. Shen, H.M. Peterson, N. Chan, N.J.K. Howden, T. Burt, F. Worrall, F. Zhang and X. Liu. 2014. Sustainable phosphorus management and the need for a long-term perspective: the legacy hypothesis. Environmental Science and Technology. 48(15):8417–8419.

He, S., D.L. Gall and K.D. McMahon. 2007. *Candidatus* Accumulibacter population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes. Applied and Environmental Microbiology. 73(18):5865–5874.

He, S. and K.D. McMahon. 2011a. *Candidatus* Accumulibacter gene expression in response to dynamic EBPR conditions. The ISME journal. 5(2):329–340.

He, S. and K.D. McMahon. 2011b. Microbiology of *Candidatus* Accumulibacter in activated sludge. Microbial Biotechnology. 4(5):603–619.

Helm, K.P., N.L. Bindoff and J.A. Church. 2011. Observed decreases in oxygen content of the global ocean. Geophysical Research Letters. 38(23):1–6.

Hem, J. 1963. Chemical equilibria and rates of manganese oxidation - Chemistry of Manganese in Natural Water. US Geological survey water-supply paper 1667-A. :71.

Herrmann, A.M., K. Ritz, N. Nunan, P.L. Clode, J. Pett-Ridge, M.R. Kilburn, D. V. Murphy, A.G. O'Donnell and E.A. Stockdale. 2007. Nano-scale secondary ion mass spectrometry - a new analytical tool in biogeochemistry and soil ecology: a review article. Soil Biology and Biochemistry. 39(8):1835–1850.

Hesselmann, R.P., C. Werlen, D. Hahn, J.R. van der Meer and a J. Zehnder. 1999. Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. Systematic and Applied Microbiology. 22(3):454–465.

Hewlett, J.D. and A.R. Hibbert. 1967. Factors affecting the response of small watersheds to precipitation in humid regions. In W. E. Sopper & H. W. Lull, eds. Forest Hydrology. Oxford, UK: Pergamon Press, :272–290.

Hill, A.R., K.J. Devito, S. Campagnolo and K. Sanmugadas. 2000. Subsurface denitrification in a forest riparian zone: Interactions between hydrology and supplies of nitrate and organic carbon. Biogeochemistry. 51(2):193–223.
Hill, C.R. and J.S. Robinson. 2012. Phosphorus flux from wetland ditch sediments. Science of the Total Environment. 437:315–322.
Hongve, D. 1997. Cycling of iron, manganese, and phosphate in a meromictic lake. Limnology and Oceanography. 42(4):635–647.
Hooda, P.S., A.R. Rendell, A.C. Edwards, P.J.A. Withers, M.N. Aitken and V.W. Truesdale. 2000. Relating soil phosphorus indices to potential phosphorus release to water. Journal of Environment Quality. 29(1995):1166–1171.
Hoppe, P., S. Cohen and A. Meibom. 2013. NanoSIMS: technical aspects and applications in cosmochemistry and biological geochemistry. Geostandards and Geoanalytical Research. 37(2):111–154.
Howe, A. and P.S.G. Chain. 2015. Challenges and opportunities in understanding microbial communities with metagenome assembly (accompanied by IPython Notebook tutorial). Frontiers in Microbiology. 6:678.
Howe, A.C., J.K. Jansson, S.A. Malfatti, S.G. Tinge, J.M. Tiedje and C.T. Brown. 2014. Tackling soil diversity with the assembly of large, complex metagenomes. Proceedings of the National Academy of Sciences. 111(13):4904–4909.
Hudnell, H.K. 2010. The state of U.S. freshwater harmful algal blooms assessments, policy and legislation. Toxicon. 55(5):1024–1034.
Hug, L.A., B.J. Baker, K. Anantharaman, C.T. Brown, A.J. Probst, C.J. Castelle, C.N. Butterfield, A.W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D.A. Relman, K.M. Finstad, R. Amundson, B.C. Thomas and J.F. Banfield. 2016. A new view of the tree of life. Nature Microbiology. 1(5):1–6.
Hung, C., J. Peccia, J.L. Zilles and D.R. Noguera. 2002. Physical Enrichment Accumulating Activated Polyphosphate in Organisms Sludge of. Water Environment Research. 74(4):354–361.
Hupfer, M. and R. Gächter. 1995. Polyphosphate in lake sediments: 31P NMR spectroscopy as a tool for its identification. Limnology and Oceanography. 40(3):610–617.
Hupfer, M., S. Gloess and H. Grossart. 2007. Polyphosphate-accumulating microorganisms in aquatic sediments. Aquatic Microbial Ecology. 47(3):299–311.
Hupfer, M., S. Glöss, P. Schmieder and H.P. Grossart. 2008. Methods for detection and quantification of polyphosphate and polyphosphate accumulating microorganisms in aquatic sediments. International Review of Hydrobiology. 93(1):1–30.
Hupfer, M. and J. Lewandowski. 2008. Oxygen controls the phosphorus release from lake sediments - A long-lasting paradigm in limnology. International Review of Hydrobiology. 93(4–5):415–432.
Hupfer, M., B. Rube and P. Schmieder. 2004. Origin and diagenesis of polyphosphate in lake sediments: a 31P-NMR study. Limnology and Oceanography. 49(1):1–10.
Iho, A., L. Ahlvik, P. Ekholm, J. Lehtoranta and P. Kortelainen. 2017. Optimal Phosphorus Abatement Redefined: Insights From Coupled Element Cycles. Ecological Economics. 137:13–19.
Ishige, K., A. Kameda, T. Noguchi and T. Shiba. 1998. The polyphosphate kinase gene of *Pseudomonas aeruginosa*. DNA Research. 5(3):157–162.

Ishige, K., H. Zhang and A. Kornberg. 2002. Polyphosphate kinase (PPK2), a potent, polyphosphate-driven generator of GTP. Proceedings of the National Academy of Sciences. 99(26):16684–16688.
Jarvie, H.P., L.T. Johnson, A.N. Sharpley, D.R. Smith, D.B. Baker, T.W. Bruulsema and R. Confesor. 2017. Increased soluble phosphorus loads to Lake Erie: unintended consequences of conservation practices? Journal of Environment Quality. 46(1):123–132.
Jarvie, H.P., A.N. Sharpley, D. Flaten, P.J. a Kleinman, A. Jenkins and T. Simmons. 2015. The pivotal role of phosphorus in a resilient water–energy–food security nexus. Journal of Environmental Quality. 44(4):1049–1062.
Jarvie, H.P., A.N. Sharpley, P.J. a Withers, J.T. Scott, B.E. Haggard and C. Neal. 2013. Phosphorus mitigation to control river eutrophication: murky waters, inconvenient truths, and "postnormal" science. Journal of Environmental Quality. 42(2):295–304.
Javot, H., N. Pumplin and M.J. Harrison. 2007. Phosphate in the arbuscular mycorrhizal symbiosis: transport properties and regulatory roles. Plant, Cell, and Environment. 30(3):310–322.
Jeffries, P., S. Gianinazzi, S. Perotto, K. Turnau and J.M. Barea. 2003. The contribution of arbuscular mycorrhizal fungi in sustainable maintenance of plant health and soil fertility. Biology and Fertility of Soils. 37(1):1–16.
Johnson, K.N., P.J.A. Kleinman, D.B. Beegle, H.A. Elliott and L.S. Saporito. 2011. Effect of dairy manure slurry application in a no-till system on phosphorus runoff. Nutrient Cycling in Agroecosystems. 90(2):201–212.
Jones, S.E. and J.T. Lennon. 2010. Dormancy contributes to the maintenance of microbial diversity. Proceedings of the National Academy of Science. 107(13):5881–5886.
Kanagawa, T., Y. Kamagata, S. Aruga, T. Kohno, M. Horn and M. Wagner. 2000. Phylogenetic Analysis of and Oligonucleotide Probe Development for Eikelboom Type 021N Filamentous Bacteria Isolated from Bulking Activated Sludge Phylogenetic Analysis of and Oligonucleotide Probe Development for Eikelboom Type 021N Filamentous Bacteria. Applied and Environmental Microbiology. 66(11):5043–5052.
Karl, D.M. 2014. Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. Annual Reviews in Marine Science. 6:279–337.
Kawaharasaki, M., T. Kanagawa, H. Tanaka and K. Nakamura. 1998. Development and application of 16S rRNA targeted oligonulceotide probe for the detection of the phosphorus-accumulating bacterium *Microlunatus phosphovorus* in an enhanced biological phosphorus removal process. Water Science and Technology. 37(4–5):481–484.
Kawaharasaki, M., H. Tanaka, T. Kanagawa and K. Nakamura. 1999. In situ identification of polyphosphate-accumulating bacteria in activated sludge by dual staining with rRNA-targeted oligonucleotide probes and 4',6-diaimidino-2-phenylindol (DAPI) at a polyphosphate-probing concentration. Water Research. 33(1):257–265.
Kawakoshi, A., H. Nakazawa, J. Fukada, M. Sasagawa, Y. Katano, S. Nakamura, A. Hosoyama, H. Sasaki, N. Ichikawa, S. Hanada, Y. Kamagata, K. Nakamura, S. Yamazaki and N. Fujita. 2012. Deciphering the genome of polyphosphate accumulating *Actinobacterium Microlunatus phosphovorus*. DNA Research. 19(5):383–394.
Keasling, J.D., L. Bertsch and A. Kornberg. 1993. Guanosine pentaphosphate phosphohydrolase of *Escherichia coli* is a long-chain exopolyphosphatase. Proceedings of the National Academy of Sciences. 90(15):7029–7033.

Kellogg, L.E. and S.D. Bridgham. 2003. Phosphorus retention and movement across an. Biogeochemistry. 63(3):299–315.
Kim, J.M., H.J. Lee, S.Y. Kim, J.J. Song, W. Park and C.O. Jeon. 2010. Analysis of the fine-scale population structure of *Candidatus* Accumulibacter phosphatis in enhanced biological phosphorus removal sludge, using fluorescence in situ hybridization and flow cytometric sorting. Applied and Environmental Microbiology. 76(12):3825–3835.
Kleinman, P.J.A., A.N. Sharpley, A.R. Buda, R.W. Mcdowell and A.L. Allen. 2011a. Soil controls of phosphorus in runoff: management barriers and opportunities. Canadian Journal of Soil Science. 91(3):329–338.
Kleinman, P.J.A., A.N. Sharpley, R.W. McDowell, D.N. Flaten, A.R. Buda, L. Tao, L. Bergstrom and Q. Zhu. 2011b. Managing agricultural phosphorus for water quality protection: principles for progress. Plant and Soil. 349(1–2):169–182.
Kleinman, P.J.A., A.N. Sharpley, L.S. Saporito, A.R. Buda and R.B. Bryant. 2009. Application of manure to no-till soils: phosphorus losses by sub-surface and surface pathways. Nutrient Cycling in Agroecosystems. 84(3):215–227.
Kleinman, P.J.A., D.R. Smith, C.H. Bolster and Z.M. Easton. 2015. Phosphorus fate, management, and modeling in artificially drained systems. Journal of Environmental Quality. 44(2):460–466.
Kleinman, P.J. a, P. Salon, a N. Sharpley and L.S. Saporito. 2005. Effect of cover crops established at time of corn planting on phosphorus runoff from soils before and after dairy manure application. Journal of Soil and Water Conservation. 60(6):311–322.
Kong, Y., J.L. Nielsen and P.H. Nielsen. 2005. Identity and ecophysiology of uncultured Actinobacterial polyphosphate-accumulating organisms in full-scale enhanced biological phosphorus removal plants. Applied and Environmental Microbiology. 71(7):4076–4085.
Kong, Y., J.L. Nielsen and P.H. Nielsen. 2004. Microautoradiographic study of Rhodocyclus -related polyphosphate-accumulating bacteria in full-scale enhanced biological phosphorus removal plants. Applied and Environmental Microbiology. 70(9):5383–5390.
Kong, Y., Y. Xia, J.L. Nielsen and P.H. Nielsen. 2007. Structure and function of the microbial community in a full-scale enhanced biological phosphorus removal plant. Microbiology. 153(12):4061–4073.
Kornberg, A. 1995. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. Journal of Bacteriology. 177(3):491–496.
Kovar, J.L., T.B. Moorman, J.W. Singer, C.A. Cambardella and M.D. Tomer. 2011. Swine manure injection with low-disturbance applicator and cover crops reduce phosphorus loss. Journal of Environment Quality. 40(2):329–336.
Kozich, J.J., S.L. Westcott, N.T. Baxter, S.K. Highlander and P.D. Schloss. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MIseq illumina sequencing platform. Applied and Environmental Microbiology. 79(17):5112–5120.
Kristiansen, R. et al. 2013. A metabolic model for members of the genus *Tetrasphaera* involved in enhanced biological phosphorus removal. The ISME Journal. 7(3):543–554.
Kruse, J., M. Abraham, W. Amelung, C. Baum, R. Bol, O. Kühn, H. Lewandowski, J. Niederberger, Y. Oelmann, C. Rüger, J. Santner, M. Siebers, N. Siebers, M. Spohn, J. Vestergren, A. Vogts and P. Leinweber. 2015. Innovative methods in soil phosphorus research: a review. Journal of Plant Nutrition and Soil Science. 178(1):43–88.

Kulakova, A.N., D. Hobbs, M. Smithen, E. Pavlov, J.A. Gilbert, J.P. Quinn and J.W. McGrath. 2011. Direct quantification of inorganic polyphosphate in microbial cells using 4???-6-diamidino-2-phenylindole (DAPI). Environmental Science and Technology. 45(18):7799–7803.

Kulakovskaya, T. V., V.M. Vagabov and I.S. Kulaev. 2012. Inorganic polyphosphate in industry, agriculture and medicine: Modern state and outlook. Process Biochemistry. 47(1):1–10.

Kumble, K.D. and A. Kornberg. 1996. Endopolyphosphatases for long chain inorganic polyphosphate in yeast and mammals. Journal of Biological Chemistry. 271(43):27146–27151.

Kunin, V., S. He, F. Warnecke, S.B. Peterson, H. Garcia Martin, M. Haynes, N. Ivanova, L.L. Blackall, M. Breitbart, F. Rohwer, K.D. McMahon and P. Hugenholtz. 2008. A bacterial metapopulation adapts locally to phage predation despite global dispersal. Genome Research. 18(2):293–297.

Lanham, A.B., A. Oehmen, A.M. Saunders, G. Carvalho, P.H. Nielsen and M.A.M. Reis. 2013. Metabolic versatility in full-scale wastewater treatment plants performing enhanced biological phosphorus removal. Water Research. 47(19):7032–7041.

Lanham, A.B., A. Oehmen, A.M. Saunders, G. Carvalho, P.H. Nielsen and M.A.M. Reis. 2014. Metabolic modelling of full-scale enhanced biological phosphorus removal sludge. Water Research. 66:283–295.

Lauber, C.L., M. Hamady, R. Knight and N. Fierer. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Applied and Environmental Microbiology. 75(15):5111–5120.

Lawson, C.E., B.J. Strachan, N.W. Hanson, A.S. Hahn, E.R. Hall, B. Rabinowitz, D.S. Mavinic, W.D. Ramey and S.J. Hallam. 2015. Rare taxa have potential to make metabolic contributions in enhanced biological phosphorus removal ecosystems. Environmental Microbiology. 17(12):4979–4993.

Lemos, P.C., L.S. Serafim, M.M. Santos, M. a M. Reis and H. Santos. 2003. Metabolic pathway for propionate utilization by phosphorus-accumulating organisms in activated sludge: 13 C labeling and in vivo nuclear magnetic resonance. Applied and Environmental Microbiology. 69(1):241–251.

Lennon, J.T. and S.E. Jones. 2011. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nature Reviews Microbiology. 9(2):119–130.

Li, X., H. Yuan, J. Yang and B. Li. 2013. Genome sequence of the polyphosphate-accumulating organism Arthrobacter sp. strain PAO19 isolated from maize rhizosphere soil. Genome Announcements. 1(4):11039–11040.

Lichko, L.P., T. V. Kulakovskaya and I.S. Kulaev. 2006. Inorganic polyphosphate and exopolyphosphatase in the nuclei of Saccharomyces cerevisiae: dependence on the growth phase and inactivation of the PPX1 and PPN1 genes. Yeast. 23:735–740.

Liu, W.T., A.T. Nielsen, J.H. Wu, C.S. Tsai, Y. Matsuo and S. Molin. 2001. In situ identification of polyphosphate- and polyhydroxyalkanoate-accumulating traits for microbial populations in a biological phosphorus removal process. Environmental Microbiology. 3(2):110–122.

Locke, N. 2015. Classification of polyphosphate-accumulating bacteria in benthic biofilms. Pennsylvania State University, State College, Pennsylvania.

Loman, N. and M. Watson. 2013. So you want to be a computational biologist? Nature Biotechnology. 31(11):996–998.

Lu, H., Oehmen, A., Virdis, B., Keller, J., and Yuan, Z. 2006. Obtaining highly enriched cultures of *Candidatus* Accumulibacter phosphates through alternating carbon sources. Water Research. 40: 3838–3848.

Lyons, J.B., J.H. Gorres and J.A. Amador. 1998. Spatial and temporal variability of phosphorus retention in a riparian forest soil. Journal of Environment Quality. 27(4):895–903.

Maavara, T., C.T. Parsons, C. Ridenour, S. Stojanovic, H.H. Dürr, H.R. Powley and P. Van Cappellen. 2015. Global phosphorus retention by river damming. Proceedings of the National Academy of Sciences. 112(51):15603–15608.

MacDonald, G.K., H.P. Jarvie, P.J.A. Withers, D.G. Doody, B.L. Keeler, P.M. Haygarth, L.T. Johnson, R.W. McDowell, M.K. Miyittah, S.M. Powers, A.N. Sharpley, J. Shen, D.R. Smith, M.N. Weintraub and T. Zhang. 2016. Guiding phosphorus stewardship for multiple ecosystem services. Ecosystem Health and Sustainability. 2(12):1–12.

Macrae, M.L., T.E. Redding, I.F. Creed, W.R. Bell and K.J. Devito. 2005. Soil, surface water and ground water phosphorus relationships in a partially harvested Boreal Plain aspen catchment. Forest Ecology and Management. 206(1–3):315–329.

Manz, W., R. Amann, W. Ludwig, M. Wagner and K.-H. Schleifer. 1992. Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. Systematic and Applied Microbiology. 15(4):593–600.

Manzoni, S., S.M. Schaeffer, G. Katul, A. Porporato and J.P. Schimel. 2014. A theoretical analysis of microbial eco-physiological and diffusion limitations to carbon cycling in drying soils. Soil Biology and Biochemistry. 73:69–83.

Mao, Y., K. Yu, Y. Xia, Y. Chao and T. Zhang. 2014. Genome reconstruction and gene expression of *Candidatus* Accumulibacter phosphatis Clade IB performing biological phosphorus removal. Environmental Science and Technology. 48(17):10363–10371.

Mao, Y., D.W. Graham, H. Tamaki and T. Zhang. 2015. Dominant and novel clades of *Candidatus* Accumulibacter phosphatis in 18 globally distributed full-scale wastewater treatment plants. Scientific Reports. 5:11857.

Mao, Y., Z. Wang, L. Li, X. Jiang, X. Zhang, H. Ren and T. Zhang. 2016. Exploring the shift in structure and function of microbial communities performing biological phosphorus removal. PLOS ONE. 11(8):1–16.

Martin, P., S.T. Dyhrman, M.W. Lomas, N.J. Poulton and B.A.S. Van Mooy. 2014. Accumulation and enhanced cycling of polyphosphate by Sargasso Sea plankton in response to low phosphorus. Proceedings of the National Academy of Sciences. 111(22):8089–8094.

Martin, P. and B.A.S. Van Mooy. 2013. Fluorometric quantification of polyphosphate in environmental plankton samples: extraction protocols, matrix effects, and nucleic acid interference. Applied and Environmental Microbiology. 79(1):273–281.

Martins, G., A. Terada, D.C. Ribeiro, A.M. Corral, A.G. Brito, B.F. Smets and R. Nogueira. 2011. Structure and activity of lacustrine sediment bacteria involved in nutrient and iron cycles. FEMS Microbiology Ecology. 77(3):666–679.

Martiny, J.B.H., S.E. Jones, J.T. Lennon and A.C. Martiny. 2015. Microbiomes in light of traits: a phylogenetic perspective. Science. 350(6261):649–656.

Maszenan, A.M., R.J. Seviour, B.K.C. Patel, P. Schumann, J. Burghardt, Y. Tokiwa and H.M. Stratton. 2000. Three isolates of novel polyphosphate-accumulating Gram-positive cocci, obtained from activated sludge, belong to a new genus, *Tetrasphaera* gen. nov., and description of two new species, *Tetrasphaera japonica* sp. nov. and *Tetrasphaera*

australiensis sp. no. International Journal of Systematic and Evolutionary Microbiology. 50(2):593–603.
McClain, M.E., E.W. Boyer, C.L. Dent, S.E. Gergel, N.B. Grimm, P.M. Groffman, S.C. Hart, J.W. Harvey, C. a. Johnston, E. Mayorga, W.H. McDowell and G. Pinay. 2003. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. Ecosystems. 6(4):301–312.
McDowell, R.W. and a N. Sharpley. 2003. The effects of soil carbon on phosphorus and sediment loss from soil trays by overland flow. Journal of Environmental Quality. 32(1):207–214.
McMahon, K.D., M.A. Dojka, N.R. Pace, D. Jenkins and J.D. Keasling. 2002. Polyphosphate kinase from activated sludge performing enhanced biological phosphorus removal. Applied and Environmental Microbiology. 68(10):4971–4978.
McMahon, K.D. and E.K. Read. 2013. Microbial contributions to phosphorus cycling in eutrophic lakes and wastewater. Annual review of microbiology. 67:199–219.
McParland, E., C.R. Benitez-Nelson, G.T. Taylor, R. Thunell, A. Rollings and L. Lorenzoni. 2015. Cycling of suspended particulate phosphorus in the redoxcline of the Cariaco Basin. Marine Chemistry. 176:64–74.
Menzel, P. and A. Krogh. 2015. Kaiju : Fast and sensitive taxonomic classification for metagenomics. Nature Communications. 7:11257.
Metson, G.S. and E.M. Bennett. 2015. Phosphorus cycling in Montreal's food and urban agriculture systems. PLOS ONE. 10(3):1–18.
Mino, T., V. Arun, Y. Tsuzuki and T. Matsuo. 1987. Effect of phosphorus accumulation on acetate metabolism in the biological phosphorus removal process. In R. Ramandori, ed. Biological Phosphorus Removal from Wastewaters: Advances in Water Pollution Control. Oxford, UK: Pergamon Press, :27–38.
Mino, T., M.C.M. Van Loosdrecht and J.J. Heijnen. 1998. Microbiology and biochemistry of the enhanced biological phosphate removal process. Water Research. 32(11):3193–3207.
Mitsch, W.J., J.K. Cronk, X. Wu, R.W. Nairn and D.L. Hey. 1995. Phosphorus retention in constructed freshwater riparian marshes. Ecological Applications. 5(3):830–845.
Moore, M.T., R. Kröger, M.A. Locke, R.F. Cullum, R.W. Steinriede, S. Testa, R.E. Lizotte, C.T. Bryant and C.M. Cooper. 2010. Nutrient mitigation capacity in Mississippi Delta, USA drainage ditches. Environmental Pollution. 158(1):175–184.
Moore, P. a. and D.M. Miller. 1994. Decreasing phosphorus solubility in poultry litter with aluminum, calcium, and iron amendments. Journal of Environment Quality. 23(2):325.
Mueller, C.W., P.K. Weber, M.R. Kilburn, C. Hoeschen, M. Kleber and J. Pett-Ridge. 2013. Advances in the analysis of biogeochemical interfaces: NanoSIMS to investigate soil microenvironments 1st ed., Elsevier Inc.
Nakamura, K., K. Masuda and E. Mikami. 1991. Isolation of a new type of polyphosphate accumulating bacterium and its phosphate removal characteristics. Journal of Fermentation and Bioengineering. 71(4):258–263.
Nakamura, K., S. Ishikawa and M. Kawaharasaki. 1995a. Phosphate uptake and release activity in immobilized polyphosphate-accumulating bacterium *Microlunatus phosphovorus* strain NM-1. Journal of Fermentation and Bioengineering. 80(4):377–382.
Nakamura, K., A. Hiraishi, Y. Yoshimi, M. Kawaharasaki, K. Masuda and Y. Kamagata. 1995b. *Microlunatus phosphovorus* gen. nov., sp. nov., a new Gram-positive polyphosphate-

accumulating bacterium isolated from activated sludge. International Journal of Systematic Bacteriology. 45(1):17–22.
Nemergut, D., A. Shade and C. Violle. 2014. When, where and how does microbial community composition matter? Frontiers in Microbiology. 5:2012–2014.
Nesme, J. et al. 2016. Back to the future of soil metagenomics. Frontiers in Microbiology. 7:1–5.
Nguyen, H.T.T., V.Q. Le, A.A. Hansen, J.L. Nielsen and P.H. Nielsen. 2011. High diversity and abundance of putative polyphosphate-accumulating *Tetrasphaera*-related bacteria in activated sludge systems. FEMS Microbiology Ecology. 76(2):256–267.
Nguyen, H.T.T., J.L. Nielsen and P.H. Nielsen. 2012. "Candidatus Halomonas phosphatis", a novel polyphosphate-accumulating organism in full-scale enhanced biological phosphorus removal plants. Environmental Microbiology. 14(10):2826–2837.
Nobu, M.K., H. Tamaki, K. Kubota and W.T. Liu. 2014. Metagenomic characterization of "Candidatus Defluviicoccus tetraformis strain TFO71", a tetrad-forming organism, predominant in an anaerobic-aerobic membrane bioreactor with deteriorated biological phosphorus removal. Environmental Microbiology. 16(9):2739–2751.
Nocek, B., S. Kochinyan, M. Proudfoot, G. Brown, E. Evdokimova, J. Osipiuk, A.M. Edwards, A. Savchenko, A. Joachimiak and A.F. Yakunin. 2008. Polyphosphate-dependent synthesis of ATP and ADP by the family-2 polyphosphate kinases in bacteria. Proceedings of the National Academy of Sciences. 105(46):17730–17735.
Noe, G.B., L.J. Scinto, J. Taylor, D.L. Childers and R.D. Jones. 2003. Phosphorus cycling and partitioning in an oligotrophic everglades wetlands ecosystem: a radioisotope tracing study. Freshwater Biology. 48(2003):1993–2008.
Obersteiner, M., J. Peñuelas, P. Ciais, M. van der Velde and I. a Janssens. 2013. The phosphorus trilemma. Nature Geoscience. 6(11):897–898.
Oehmen, A., P.C. Lemos, G. Carvalho, Z. Yuan, J. Keller, L.L. Blackall and M.A.M. Reis. 2007. Advances in enhanced biological phosphorus removal: From micro to macro scale. Water Research. 41(11):2271–2300.
Ohtake, H., A. Kuroda, M. Chandrasekaran, H. Wu and S. Tanaka. 2001. Molecular genetics of bacterial polyphosphate accumulation to better understand the mechanism underlying biological phosphorus removal. In M. Healy, D. L. Wise, & M. Moo-Young, eds. Environmental Monitoring and Biodiagnostics of Hazardous Contaminants. Dordrecht, Netherlands: Kluvwer Academic Publishers, :181–196.
Olander, L.P. and P.M. Vitousek. 2000. Regulation of soil phosphatase and chitinase activity by N and P availability. Biogeochemistry. 49(2):175–190.
Oliverio, A.M., M.A. Bradford and N. Fierer. 2016. Identifying the microbial taxa which consistently respond to soil warming across time and space. Global Change Biology. 23:2117–2129.
Ouyang, Y. and X. Li. 2013. Recent research progress on soil microbial responses to drying–rewetting cycles. Acta Ecologica Sinica. 33(1):1–6.
Oyserman, B.O., J.M. Martirano, S. Wipperfurth, B.R. Owen, D.R. Noguera and K.D. McMahon. 2017. Community Assembly and Ecology of Activated Sludge under Photosynthetic Feast–Famine Conditions. Environmental Science & Technology. :acs.est.6b03976.
Oyserman, B.O., F. Moya, C.E. Lawson, A.L. Garcia, M. Vogt, M. Hefferenen, D.R. Noguera and K.D. McMahon. 2016. Ancestral genome reconstruction identifies the evolutionary

basis for trait acquisition in polyphosphate accumulating bacteria. The ISME Journal. 10(12):2931–2945.
Pallen, M.J. 2016. Microbial bioinformatics 2020. Microbial Biotechnology. 9(5):681–686.
Pellerin, B.A., B.A. Stauffer, D.A. Young, D.J. Sullivan, S.B. Bricker, M.R. Walbridge, G.A. Clyde and D.M. Shaw. 2016. Emerging tools for continuous nutrient monitoring networks: sensors advancing science and water resources protection. Journal of the American Water Resources Association. 52(4):993–1008.
Peralta, A.L., S. Ludmer, J.W. Matthews and A.D. Kent. 2014. aBacterial community response to changes in soil redox potential along a moisture gradient in restored wetlands. Ecological Engineering. 73:246–253.
Perez-Riverol, Y., L. Gatto, R. Wang, T. Sachsenberg, J. Uszkoreit, F. da Veiga Leprevost, C. Fufezan, T. Ternent, S.J. Eglen, D.S. Katz, T.J. Pollard, A. Konovalov, R.M. Flight, K. Blin, J.A. Vizcaíno. 2016. Ten simple rules for taking advantage of Git and GitHub. PLOS Computational Biology. 12(7):e1004947.
Peterson, S.B., F. Warnecke, J. Madejska, K.D. McMahon and P. Hugenholtz. 2008. Environmental distribution and population biology of *Candidatus* Accumulibacter, a primary agent of biological phosphorus removal. Environmental Microbiology. 10(10):2692–703.
Pett-Ridge, J. and M.K. Firestone. 2005. Redox fluctuation structures microbial communities in a wet tropical soil redox fluctuation structures microbial communities in a wet tropical soil. Applied and Environmental Microbiology. 71(11):6998–7007.
Poisot, T., B. Péquin and D. Gravel. 2013. High-Throughput Sequencing: a roadmap toward community ecology. Ecology and Evolution. 3(4):1125–1139.
Powell, J.R., A. Welsh, S. Hallin and S.D. Allison. 2015. Microbial functional diversity enhances predictive models linking environmental parameters to ecosystem properties. Ecology. 96(7):1985–1993.
Powers, S.M., T.W. Bruulsema, T.P. Burt, N.I. Chan, J.J. Elser, P.M. Haygarth, N.J.K. Howden, H.P. Jarvie, Y. Lyu, H.M. Peterson, A.N. Sharpley, J. Shen, F. Worrall and F. Zhang. 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nature Geoscience. 9(5):353–356.
Prestigiacomo, A.R., S.W. Effler, R.K. Gelda, D.A. Matthews, M.T. Auer, B.E. Downer, A. Kuczynski and M.T. Walter. 2016. Apportionment of bioavailable phosphorus loads entering Cayuga Lake, New York. Journal of the American Water Resources Association. 52(1):31–47.
Price, K.J. and H.J. Carrick. 2014. Quantitative evaluation of spatiotemporal phosphorus fluxes in stream biofilms. Freshwater Science. 33(1):99–111.
Price, K.J. and H.J. Carrick. 2013. Effects of physical disturbance on phosphorus uptake in temperate stream biofilms. Inland Waters. 3:321–330.
Prosser, J.I. 2013. Think before you sequence. Nature. 494:41.
Prosser, J.I., B.J.M. Bohannan, T.P. Curtis, R.J. Ellis, M.K. Firestone, R.P. Freckleton, J.L. Green, L.E. Green, K. Killham, J.J. Lennon, A.M. Osborn, M. Solan, der G. Van and J.P.W. Young. 2007. The role of ecological theory in microbial ecology. Nat Rev Micro. 5:384–392.
Rangarajan, E.S., G. Nadeau, Y. Li, J. Wagner, M.N. Hung, J.D. Schrag, M. Cygler and A. Matte. 2006. The Structure of the exopolyphosphatase (PPX) from *Escherichia coli*

O157:H7 suggests a binding mode for long polyphosphate chains. Journal of Molecular Biology. 359(5):1249–1260.
Rao, N.N., M.R. Gómez-García and A. Kornberg. 2009. Inorganic polyphosphate: essential for growth and survival. Annual Review of Biochemistry. 78:605–647.
Read, E.K., M. Ivancic, P. Hanson, B.J. Cade-Menun and K.D. McMahon. 2014. Phosphorus speciation in a eutrophic lake by 31P NMR spectroscopy. Water Research. 62:229–240.
Reddy, K.R., R.H. Kadlec, E. Flaig and P.M. Gale. 1999. Phosphorus retention in streams and wetlands: a review. Critical Reviews in Environmental Science and Technology. 29(1):83–146.
Reed, D.C., C.K. Algar, J.A. Huber and G.J. Dick. 2014. Gene-centric approach to integrating environmental genomics and biogeochemical models. Proceedings of the National Academy of Sciences. 111(5):1879–84.
Reizer, J., A. Reizer and M.H. Saier. 1993. Exopolyphosphatase and guanine pentaphosphate phosphatase belong to the sugar kinase/actin/hsp70 superfamily. Trends in Biochemical Sciences. :247–248.
Richardson, C.J. 1985. Mechanisms Controlling Phosphorus Retention. Science. 228:1424–1427.
Rier, S.T., K.C. Kinek, S.E. Hay and S.N. Francoeur. 2016. Polyphosphate plays a vital role in the phosphorus dynamics of stream periphyton. Freshwater Science. 35(2):490–502.
Riesenfeld, C.S., P.D. Schloss and J. Handelsman. 2004. Metagenomics: genomic analysis of microbial communities. Annual Reviews of Genetics. 38:525–552.
Rittenburg, R.A., A.L. Squires, J. Boll, E.S. Brooks, Z.M. Easton and T.S. Steenhuis. 2015. Agricultural BMP effectiveness and dominant hydrological flow paths: concepts and a review. Journal of the American Water Resources Association. 51(2):305–329.
Rocca, J.D., E.K. Hall, J.T. Lennon, S.E. Evans, M.P. Waldrop, J.B. Cotner, D.R. Nemergut, E.B. Graham and M.D. Wallenstein. 2015. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. Isme Journal. 9(8):1693–1699.
Rode, M., A.J. Wade, M.J. Cohen, R.T. Hensley, M.J. Bowes, J.W. Kirchner, G.B. Arhonditsis, P. Jordan, B. Kronvang, S.J. Halliday, R.A. Skeffington, J.C. Rozemeijer, A.H. Aubert, K. Rinke and S. Jomaa. 2016. Sensors in the stream: the high-frequency wave of the present. Environmental Science and Technology. 50(19):10297–10307.
Rodríguez, H. and R. Fraga. 1999. Phosphate solubilizing bacteria and their role in plant growth promotion. Biotechnology Advances. 17(4–5):319–339.
Rousk, J., P.C. Brookes, H.C. Glanville and D.L. Jones. 2011. Lack of correlation between turnover of low-molecular-weight dissolved organic carbon and differences in microbial community composition or growth across a soil pH gradient. Applied and Environmental Microbiology. 77(8):2791–2795.
Rubio-Rincón, F.J., L. Welles, C.M. Lopez-Vazquez, M. Nierychlo, B. Abbas, M. Geleijnse, P.H. Nielsen, M.C.M. van Loosdrecht and D. Brdjanovic. 2017. Long-term effects of sulphide on the enhanced biological removal of phosphorus: The symbiotic role of Thiothrix caldifontis. Water Research. 116:53–64.
Saia, S.M., P.J. Sullivan, J.M. Regan, H.J. Carrick, A.R. Buda, N.A. Locke and M.T. Walter. 2017. Evidence for polyphosphate accumulating organism (PAO)-mediated phosphorus cycling in stream biofilms under alternating aerobic/anaerobic conditions. Freshwater Science. 36(2):284-296.

Schimel, J. and J. Gulledge. 1998. Microbial community structure and global trace gases. Global Change Biology. 4(7):745–758.
Schindler, D.W. 2012. The dilemma of controlling cultural eutrophication of lakes. Proceedings of the Royal Society B. 279(1746):4322–4333.
Schindler, D.W. 1977. Evolution of phosphorus limitation in lakes. Science. 195(4275):260–262.
Schloss, P.D. and J. Handelsman. 2006. Toward a census of bacteria in soil. PLOS Computational Biology. 2(7):786–793.
Schulz, H.N. and H.D. Schulz. 2005. Large sulfur bacteria and the formation of phosphorite. Science. 307:416–418.
Scinto, L.J. and K.R. Reddy. 2003. Biotic and abiotic uptake of phosphorus by periphyton in a subtropical freshwater wetland. Aquatic Botany. 77(3):203–222.
Seviour, R.J. and S. McIlroy. 2008. The microbiology of phosphorus removal in activated sludge processes-the current state of play. Journal of Microbiology. 46(2):115–124.
Seviour, R.J., T. Mino and M. Onuki. 2003. The microbiology of biological phosphorus removal in activated sludge systems. FEMS Microbiology Reviews. 27(1):99–127.
Seviour, R.J. and P.H. Nielsen. 2010. Microbial Ecology of Activated Sludge R. J. Seviour & P. H. Nielsen, eds., London, UK: IWA Publishing.
Shade, A. 2017. Diversity is the question, not the answer. ISME Journal. 11:1–6.
Shade, A. and T.K. Teal. 2015. Computing workflows for biologists: a roadmap. PLOS Biology. 13(11):1–10.
Sharpley, A., H.P. Jarvie, A. Buda, L. May, B. Spears and P. Kleinman. 2013. Phosphorus legacy: overcome the effects of past management practices to mitigate future water quality impairment. Journal of Environmental Quality. 42(5):1308–1326.
Sharpley, A.N., P.J.A. Kleinman, D.N. Flaten and A.R. Buda. 2011. Critical source area management of agricultural phosphorus: experiences, challenges and opportunities. Water Science and Technology. 64(4):945–952.
Sharpley, A.N. and S. Rekolainen. 1997. Phosphorus in agriculture and its environmental implications. In H. Tunney, O. T. Carton, P. C. Brookes, & A. E. Johnston, eds. Phosphorus loss from soil to water. Wallingford, UK: CAB International, :1–54.
Sharpley, a N. and R.G. Menzel. 1987. The impact of soil and fertilizer phosphorus on the environment. Advances in Agronomy. 41:297–324.
Sharrer, K.L., L.E. Christianson, C. Lepine and S.T. Summerfelt. 2016. Modeling and mitigation of denitrification “woodchip” bioreactor phosphorus releases during treatment of aquaculture wastewater. Ecological Engineering. 93:135–143.
Sherson, L.R., D.J. Van Horn, J.D. Gomez-Velez, L.J. Crossey and C.N. Dahm. 2015. Nutrient dynamics in an alpine headwater stream: use of continuous water quality sensors to examine responses to wildfire and precipitation events. Hydrological Processes. 29(14):3193–3207.
Shi, X. and A. Kornberg. 2005. Endopolyphosphatase in Saccharomyces cerevisiae undergoes post-translational activations to produce short-chain polyphosphates. FEBS Letters. 579(9):2014–2018.
Silver, W.L., a E. Lugo and M. Keller. 1999. Soil oxygen availability and biogeochemical cycling along elevation and topographic gradients in Puerto Rico. Biogeochemistry. 44:301–328.

Skennerton, C.T., J.J. Barr, F.R. Slater, P.L. Bond and G.W. Tyson. 2014. Expanding our view of genomic diversity in *Candidatus* Accumulibacter clades. Environmental Microbiology. 17(5):1574–1585.
Smil, V. 2000. Phosphorus in the environment: natural flows and human interfaces. Annual Reviews of Engery and the Environment. 25(1):53–88.
Smith, D.R., K.W. King, L. Johnson, W. Francesconi, P. Richards, D. Baker and A.N. Sharpley. 2015. Surface runoff and tile drainage transport of phosphorus in the Midwestern United States. Journal of Environmental Quality. 44(2):495–502.
Smith, M.S. and J.M. Tiedje. 1979. Phases of denitrification following oxygen depletion in soil. Soil Biology and Biochemistry. 11(3):261–267.
Soo, R.M., C.T. Skennerton, Y. Sekiguchi, M. Imelfort, S.J. Paech, P.G. Dennis, J.A. Steen, D.H. Parks, G.W. Tyson and P. Hugenholtz. 2014. An expanded genomic representation of the phylum Cyanobacteria. Genome Biology and Evolution. 6(5):1031–1045.
Stante, L., C.M. Cellamare, F. Malaspina, G. Bortone and A. Tilche. 1997. Biological phosphorus removal by pure culture of Lampropedia spp. Water Research. 31(6):1317–1324.
Stout, W.L., A.N. Sharpley and H.B. Pionke. 1998. Reducing soil phosphorus solubility with coal combustion by-products. Journal of Environment Quality. 27(1):111–118.
Streichan, M., J.R. Golicki and G. Schön. 1990. Polyphosphate-accumulating bacteria from sewage plants with different processes for biological phosphorus removal. FEMS Microbiology Ecology. 73:113–124.
Sulu-Gambari, F., D. Seitaj, F.J.R. Meysman, R. Schauer, L. Polerecky and C.P. Slomp. 2016. Cable bacteria control iron-phosphorus dynamics in sediments of a coastal hypoxic basin. Environmental Science and Technology. 50(3):1227–1233.
Tapia-Torres, Y., M.D. Rodríguez-Torres, J.J. Elser, A. Islas, V. Souza, F. García-Oliva and G. Olmedo-Álvarez. 2016. How to live with phosphorus scarcity in soil and sediment: lessons from bacteria. Applied and Environmental Microbiology. 82(15):4652–4662.
Temperton, B., J. a Gilbert, J.P. Quinn and J.W. McGrath. 2011. Novel analysis of oceanic surface water metagenomes suggests importance of polyphosphate metabolism in oligotrophic environments. PLOS One. 6(1):1–14.
Todd-Brown, K.E.O., F.M. Hopkins, S.N. Kivlin, J.M. Talbot and S.D. Allison. 2012. A framework for representing microbial decomposition in coupled climate models. Biogeochemistry. 109(1–3):19–33.
Tomer, M.D., E.J. Sadler, R.E. Lizotte, R.B. Bryant, T.L. Potter, M.T. Moore, T.L. Veith, C. Baffaut, M. a. Locke and M.R. Walbridge. 2014. A decade of conservation effects assessment research by the USDA Agricultural Research Service: progress overview and future outlook. Journal of Soil and Water Conservation. 69(5):365–373.
Trelstad, P.L., P. Purdhani, W. Geißdörfer, W. Hillen and J.D. Keasling. 1999. Polyphosphate kinase of Acinetobacter sp. strain ADP1: Purification and characterization of the enzyme and its role during changes in extracellular phosphate levels. Applied and Environmental Microbiology. 65(9):3780–3786.
Tringe, S.G., C. von Mering, A. Kobayashi, A.A. Salamov, K. Chen, H.W. Chang, M. Podar, J.M. Short, E.J. Mathur, J.C. Detter, P. Bork, P. Hugenholtz and E.M. Rubin. 2005. Comparative Metagenomics of Freshwater Microbial Communities. Science. 308(5721):554–557.

Turner, B.L. and P.M. Haygarth. 2001. Phosphorus solubilization in rewetted soils. Nature. 411:258.
Tye, A.M., B.G. Rawlins, J.C. Rushton and R. Price. 2016. Understanding the controls on sediment-P interactions and dynamics along a non-tidal river system in a rural-urban catchment: The River Nene. Applied Geochemistry. 66:219–233.
Tzeng, C.M. and A. Kornberg. 1998. Polyphosphate kinase is highly conserved in many bacterial pathogens. Molecular Microbiology. 29(1):381–382.
Uhlmann, D. and H.-D. Bauer. 1988. A remark on microorganisms in lake sediments with emphasis on polyphosphate-accumulating bacteria. Internationale Revue der gesamten Hydrobiologie und Hydrographie. 73(6):703–708.
Valdivia, M.V. 2009. The role of microbial processes in soil phosphorus dynamics. PhD Thesis, Cornell University.
Verbree, D.A., S.W. Duiker and P.J.A. Kleinman. 2010. Runoff losses of sediment and phosphorus from no-till and cultivated soils receiving dairy manure. Journal of Environment Quality. 39(5):1762–1770.
Vidon, P., C. Allan, D. Burns, T.P. Duval, N. Gurwick, S. Inamdar, R. Lowrance, J. Okay, D. Scott and S. Sebestyen. 2010. Hot spots and hot moments in riparian zones: potential for improved water quality management. Journal of the American Water Resources Association. 46(2):278–298.
Wagner, M., R. Erhart, W. Manz, R. Amann, H. Lemmer, D. Wedi and K.H. Schleifer. 1994. Development of an rRNA-targeted oligonucleotide probe specific for the genus Acinetobacter and its application for in situ monitoring in sctivated sludge. Applied and Environmental Microbiology. 60(3):792–800.
Wallenstein, M.D. and E.K. Hall. 2012. A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. Biogeochemistry. 109(1–3):35–47.
Walter, M.T., J.A. Archibald, B. Buchanan, H. Dahlke, Z.M. Easton, R.D. Marjerison, A.N. Sharma and S.B. Shaw. 2009. New paradigm for sizing riparian buffers to reduce risks of polluted storm water: practical synthesis. Journal of Irrigation and Drainage Engineering. 135(2):200–209.
Walter, M.T., M.F. Walter, E.S. Brooks, T.S. Steenhuis, J. Boll and K. Weiler. 2000. Hydrologically sensitive areas: variable source area hydrology implications for water quality risk assessment. Journal of Soil and Water Conservation. 55(3):277–284.
Wang, Q., J.A. Fish, M. Gilman, Y. Sun, C.T. Brown, J.M. Tiedje and J.R. Cole. 2015. Xander : employing a novel method for efficient gene-targeted metagenomic assembly. Microbiome. 3:32.
Watts, D.B. and H.A. Torbert. 2009. Impact of gypsum applied to grass buffer strips on reducing soluble P in surface water runoff. Journal of Environment Quality. 38(2000):1511–1517.
Wentzel, M.C., L.H. Lotter, R.E. Loewenthal and G.V.R. Marais. 1986. Metabolic behaviour of Acinetobacter spp. in enhanced biological phosphorus removal – a biochemical model,
White, D.J., M.R. Noll and J.C. Makarewicz. 2008. Does manganese influence phosphorus cycling under suboxic lake water conditions? Journal of Grest Lakes Research. 34(4):571–580.
Wieder, W.R., G.B. Bonan and S.D. Allison. 2013. Global soil carbon projections are improved by modelling microbial processes. Nature Climate Change. 3(10):909–912.

Williams, M.R., K.W. King, W. Ford, A.R. Buda and C.D. Kennedy. 2016. Effects of tillage on macropore flow and phosphorus transport to tile drains. Water Resources re. 52(4):2868–2882.
Wilson, H.F., S. Satchithanantham, A.P. Moulin and A.J. Glenn. 2016. Soil phosphorus spatial variability due to landform, tillage, and input management: a case study of small watersheds in southwestern Manitoba. Geoderma. 280:14–21.
Withers, P.J.A., R. Sylvester-Bradley, D.L. Jones, J.R. Healey and P.J. Talboys. 2014. Feed the crop not the soil: Rethinking phosphorus management in the food chain. Environmental Science and Technology. 48(12):6523–6530.
Withers, P.J. and H.P. Jarvie. 2008. Delivery and cycling of phosphorus in rivers: a review. Science of the Total Environment. 400(1–3):379–395.
Wommack, K.E., J. Bhavsar and J. Ravel. 2008. Metagenomics: read length matters. Applied and Environmental Microbiology. 74(5):1453–1463.
Wurst, H. and A. Kornberg. 1994. A soluble exopolyphosphatase of Saccharomyces cerevisiae. Journal of Biological Chemistry. 269(15):10996–11001.
Yall, L., B.H. Boughton, S.A. Roinestad and N.A. Sinclair. 1972. Logical removal of phosphates. In Logical removal of phosphates. Applications of New Concepts of Physical Chemical Waste Water Treatment. Vanderbilt University Conference: Pergamon Press.
Yevdokimov, I., A. Larionova and E. Blagodatskaya. 2016. Microbial immobilization of phosphorus in soils exposed to drying-rewetting and freeze-thawing cycles. Biology and Fertility of Soils. 52(5):685–696.
Young, E.O. and R.D. Briggs. 2008. Phosphorus concentrations in soil and subsurface water: a field study among cropland and riparian buffers. Journal of Environmental Quality. 37(1):69–78.
Young, E.O. and D.S. Ross. 2016. Total and labile phosphorus concentrations as influenced by riparian buffer soil properties. Journal of Environmental Quality. 45(1):294–304.
Young, I.M. and J.W. Crawford. 2004. Interactions and Self-Organization in the Soil-Microbe Complex. Science. 304(5677):1634–1637.
Yuan, Z., S. Pratt and D.J. Batstone. 2012. Phosphorus recovery from wastewater through microbial processes. Current Opinion in Biotechnology. 23(6):878–883.
Zago, A., S. Chugani and A.M. Chakrabarty. 1999. Cloning and characterization of polyphosphate kinase and exopolyphosphatase genes from *Pseudomonas aeruginosa* 8830. Applied and Environmental Microbiology. 65(5):2065–2071.
Zhang, A.N. et al. 2016. Development of quantitative real-time PCR assays for different clades of *Candidatus* Accumulibacter. Scientific Reports. 6:23993.
Zhang, F., L.C. Blasiak, J.O. Karolin, R.J. Powell, C.D. Geddes and R.T. Hill. 2015. Phosphorus sequestration in the form of polyphosphate by microbial symbionts in marine sponges. Proceedings of the National Academy of Sciences. (14):1–6.
Zhang, H., K. Ishige and A. Kornberg. 2002. A polyphosphate kinase (PPK2) widely conserved in bacteria. Proceedings of the National Academy of Sciences. 99(26):16678–16683.
Zhang, H., Y. Sekiguchi, S. Hanada, P. Hugenholtz, H. Kim, Y. Kamagata and K. Nakamura. 2003. *Gemmatimonas aurantiaca* gen. nov., sp. nov., a Gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl. nov. International Journal of Systematic and Evolutionary Microbiology. 53(4):1155–1163.

Zhu, Y., W. Huang, S.S.K. Lee and W. Xu. 2005. Crystal structure of a polyphosphate kinase and its implications for polyphosphate synthesis. EMBO Reports. 6(7):681–687.
Zilles, J.L., C.H. Hung and D.R. Noguera. 2002a. Presence of Rhodocyclus in a full-scale wastewater treatment plant and their participation in enhanced biological phosphorus removal. Water Science and Technology. 46(1–2):123–128.
Zilles, J.L., J. Peccia, M. Kim, C. Hung and D.R. Noguera. 2002b. Involvement of Rhodocyclus -related organisms in phosphorus removal in full-scale wastewater treatment plants. Applied and Environmental Microbiology. 68(6):2763–2769.
Zimmerman, N., J. Izard, C. Klatt, J. Zhou and E. Aronson. 2014. The unseen world: environmental microbial sequencing and identification methods for ecologists. Frontiers in Ecology and the Environment. 12(4):224–231.

CHAPTER 2

EVIDENCE FOR POLYPHOSPHATE ACCUMULATING ORGANISMS (PAO)-MEDIATED PHOSHPHORUS CYCLING IN STREAM BIOFILMS UNDER ALTERNATING AEROBIC/ANAEROBIC CONDITIONS[1]

Sheila M. Saia, Patrick J. Sullivan, John M. Regan, Hunter J. Carrick, Anthony R. Buda, Nicholas A. Locke, M. Todd Walter

ABSTRACT

Phosphorus (P) availability often limits primary production in freshwater ecosystems and excessive P inputs promote accelerated eutrophication. Microbial mechanisms may control O_2-dependent uptake/release of P in stream sediments and biofilms, but specific organisms responsible for these cycles have not been identified. Polyphosphate accumulating organisms (PAOs) are purposely enriched in treatment plants to remove P from wastewater. PAOs release P under anaerobic conditions and take it up under aerobic conditions. We hypothesized that alternating aerobic/anaerobic conditions promote patterns of P uptake/release like PAOs in wastewater treatment. We collected intact, native stream biofilms and subjected them to laboratory treatments to impose conditions similar to what may occur because of diel oxygenic and respiratory cycles: 1) continuous sparging with air and 2) alternate sparging with air or anaerobic gas (20:80% by volume CO_2:N_2). We monitored phosphate (PO_4^{3-}), Ca, Mg, total Mn, K, Fe^{2+}, and total S concentrations in the water during the experiment and total P (TP) and

[1] Published in: *Freshwater Science*. 2017, DOI: 10.1086/691439.

polyphosphate (polyP) concentrations in the biofilms at the start and end of the experiment. We used microscopy and polymerase chain reaction (PCR) to quantify the percentage of cells with stored intracellular polyP and to test for known PAO genes, respectively. The water had significantly greater dissolved PO_4^{3-} concentrations during anaerobic than during aerobic conditions. Ca, K, Mg, and total Mn concentrations mimicked PO_4^{3-} concentrations over time, but Fe^{2+} and total S concentrations did not. Precipitation of Ca and Mg and reductive dissolution of Mn may have influenced P cycling. Percent microbially stored intracellular polyP was nearly 3× greater in aerobic than anaerobic conditions. We did not find previously reported PAO genes in our biofilms, indicating the presence of novel polyP accumulators. Combined biotic and abiotic processes may be important in controlling short-term P cycling in stream biofilms.

INTRODUCTION

Eutrophication of streams resulting from excess P has been an issue of world-wide importance for the past 50 y (Dubrovsky et al. 2010, Jarvie et al. 2013). P dynamics in streams are controlled by a combination of chemical, physical, and biological processes (Reddy et al. 1999, Dodds 2003, Withers and Jarvie 2008, Cohen et al. 2013), but discrepancies remain between upstream nutrient reductions and downstream responses. Thus, further description and quantification of key mechanisms governing P dynamics is needed (Davelaar 1993). In most stream environments, biotic processes constitute a larger percentage of P uptake—from 66% to 91% of the total—than do abiotic processes (Reddy et al. 1999, Cohen et al. 2013). When P is taken up, from 10 to >50% of the P in stream biofilms can be stored as intracellular polyphosphate (polyP; Taylor 2016, Rier et al. 2016). Stream biofilm assemblages also alter pH and dissolved O_2 (DO) concentrations, which influence the chemical reactivity of P with Ca, Fe, clays, and other metal

oxides (Carlton and Wetzel 1988, Hongve 1997, Dodds 2003, Withers and Jarvie 2008, Tye et al. 2016).

However, explanations and modeling of P dynamics in streams have largely (if not solely) drawn on chemical and physical controls (Davelaar 1993, Hupfer and Lewandowski 2008). Previous investigators hypothesized that alternating redox conditions in sediments may favor organisms capable of storing polyP and these organisms may affect water-column P concentrations (Davelaar 1993, Reddy et al. 1999, Hupfer et al. 2007, 2008, Hupfer and Lewandowski 2008, Diaz et al. 2012, McMahon and Reed 2013). Observations from wastewater treatment may offer important insight into the influence of biofilms on P cycling in streams (Davelaar 1993, Reddy et al. 1999, Hupfer et al. 2007, McMahon and Reed 2013) because polyphosphate accumulating organisms (PAOs) have been used as a key element of enhanced biological P removal (EBPR) to treat wastewater (Seviour et al. 2003). In EBPR wastewater treatment, condensation of intracellular polyP leads to removal of PO_4^{3-} by PAOs in the water column (Hesselmann et al. 1999, Seviour et al. 2003). Hesselmann et al. (1999) were the first to identify the model PAO species *Candidatus* Accumulibacter phosphatis (CAP), from a sequencing batch reactor seeded with EBPR sludge. The metabolism of CAP and of PAOs in general depends on the alternating aerobic/anaerobic conditions of the EBPR process (Hesselmann et al. 1999).

During anaerobic conditions, PAOs, such as CAP, use energy derived from hydrolysis of stored polyP to take up energy-rich volatile fatty acids, like acetate, and store them as polymers, such as polyhydroxyalkanoates (Hesselmann et al. 1999, Seviour et al. 2003). Hydrolysis of polyP leads

to release of PO_4^{3-} from the cell into the water column. During aerobic conditions, PAOs use energy from polyhydroxyalkanoate oxidation to replenish polyP stores, resulting in luxury uptake of PO_4^{3-} from the surrounding water. Their ability to store electron donors intracellularly during anaerobic periods enables PAOs to thrive over other microorganisms (Seviour et al. 2003). Diverse assemblages of PAOs have been found in EBPR sludges (Nakamura et al. 1995, Liu et al. 2001, Zilles et al. 2002, Kong et al. 2005, 2007, Nguyen et al. 2011, 2012, Albertsen et al. 2012, Kristiansen et al. 2013), indicating this P uptake/release metabolism may be used by diverse microorganisms.

EBPR PAOs have been identified in stream and lake sediments (Kunin et al. 2008, Peterson et al. 2008, Martins et al. 2011), and polyP has been identified in sediment-dwelling organisms (Uhlmann and Bauer 1988, Hupfer et al. 1995, 2004, Schulz and Schulz 2005, Sulu-Gambari et al. 2016) and in stream biofilms (Rier et al. 2016). Several investigators have shown the effect of changing environmental conditions on P uptake and release in the water column. Stream P concentrations decreased in the daytime and increased at night (Cohen et al. 2013, Sherson et al. 2015). Furthermore, water-column P concentrations overlying epipelic algae and sediments under aerobic conditions were lower than water-column P concentrations under anaerobic conditions (Carlton and Wetzel 1988, Gächter et al. 1988).

A number of investigators have used polyP detection techniques to study P storage dynamics on long time scales. PolyP was found in surface sediment extracts taken monthly and was associated with intracellular polyP granules but was not found at depth (Hupfer et al. 1995). In a survey of

22 European lakes, polyP was found in surface sediment extracts but not at depth (Hupfer et al. 2004). In another lake, a shift from aerobic to anaerobic conditions at the sediment water interface was accompanied by a decrease in polyP concentrations from monthly sediment extracts (Amirbahman et al. 2013). Water-column samples from a coastal inlet showed a correlation between polyP disappearance and the emergence of hypoxia (Diaz et al. 2012). In streams, biofilm sampled 1 d after P pulses associated with storm runoff had more stored polyP than biofilms sampled before the pulse (Rier et al. 2016). To our knowledge, no study of stream biofilms has linked water-column P patterns with polyP storage during short-term (i.e., hourly) alternating aerobic/anaerobic conditions.

Stream biofilms are an interesting system in which to study how coupled biological, chemical, and physical processes—including the potential role of PAOs—influence stream P cycling under changing environmental conditions. Our objective was to assess whether alternating aerobic/anaerobic periods were coupled with water-column PO_4^{3-} concentration and intracellular polyP storage in stream biofilms. We hypothesized that alternating aerobic/anaerobic conditions in stream biofilms lead to patterns in P uptake/release, respectively, from organisms in the biofilm. We tested our hypothesis by monitoring the concentration of PO_4^{3-} and other analytes in water samples taken from sealed tubs over time in the laboratory under alternating aerobic/anaerobic conditions. We also took biofilm samples at the start and end of the experiment and analyzed them for the percentage of intracellular polyP granules, presence of model PAO genes, and TP and polyP concentrations.

METHODS

Site Description

Our study site was a 2nd-order stream, Cascadilla Creek, near Ithaca, New York (Figure 2.1A). The stream is underlain by Upper Devonian shale (King and Beikman 1974), and bed material consists of primarily medium- to large-sized cobbles. The Cascadilla Creek Watershed (CCW) is a subwatershed of the Cayuga Lake Watershed and larger Owasco River Basin. The watershed has a drainage area of 3400 ha, and the predominant land cover upstream of the sample site is deciduous/mixed forest and agricultural land (Jin et al. 2013). Other than a baseline sample taken at the start of our experiment (i.e., 0 h), nutrient concentrations have not been characterized for this ungauged watershed. However, information is available for larger surrounding watersheds with similar land-cover distributions—the Fall Creek and Six Mile Creek Watersheds (Table 2.1). For context, from late 2002 to early 2015 Fall Creek and Six Mile Creek had average sample TP concentrations of 0.18 and 0.06 mg L^{-1}, respectively (USGS 2012, CSI 2015). Average sample total Kjeldahl N (TKN) concentrations for Fall Creek and Six Mile Creek during the same period were 0.80 and 0.34 mg L^{-1}, respectively (USGS 2012, CSI 2015).

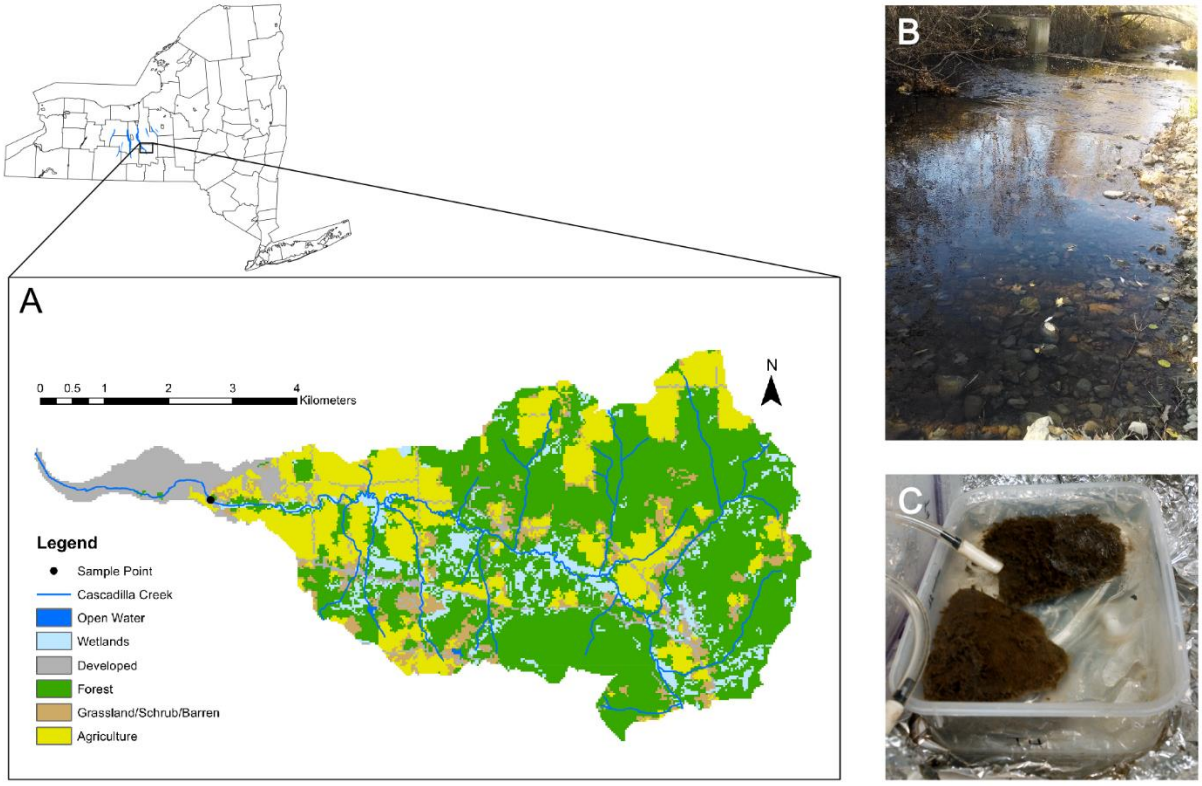

Figure 2. 1 A. Location of the study site in the Cascadilla Creek Watershed in Tompkins County, New York. B. Photograph of Cascadilla Creek. C.—Stream cobbles in sealed containers with native stream water used in the experiment.

Table 2. 1 Land cover (LC) percentages for the ungauged study watershed, Cascadilla Creek, and nearby gauged watersheds, Six Mile Creek and Fall Creek.

Description	Cascadilla Creek LC (%)	Six Mile Creek LC (%)	Fall Creek LC (%)
Developed	8.5	6.2	6.6
Forest	54.5	57.7	33.3
Agriculture	22.1	21.2	44.3
Wetlands	8.6	6.7	7.2
Open Water	0.1	0.3	0.4
Grassland/Shrubs	6.2	7.9	8.2

Biofilm Collection and Experimental Design

We collected 7 cobbles from the stream. We kept 3 of them unsubmerged. Upon returning to the laboratory, we immediately processed unsubmerged biofilms for polyP and TP extractions, microscopy, and PCR. We kept the remaining 4 cobbles (average exposed benthic biofilm surface area excluding the bottom ≈ 40 cm2) submerged in 2 sealed tubs with stream water in the laboratory (Figures 2.1B and 2.1C). We stirred the water in the tubs continuously with a stir bar and sparged the water with air to establish baseline conditions before the start of the experiment.

The emphasis of this experiment was to evaluate short-term P dynamics over time under varying O_2 levels in a controlled laboratory environment rather than a more global (i.e., spatial) assessment of P dynamics in streams. Measuring all parameters at each sampling point was time consuming. Therefore, we used only 1 tub per treatment. To offset this limitation, we increased the number of samples taken and duration of the experiment to include 2 aerobic/anaerobic cycles, which served as a replication in time. We did not include a negative control tub with bare cobbles in this experiment.

Water in the treatment 1 (T1) tub was sparged alternately with air for 12 h (aerobic period: 0900–2100 h) with a mixed anaerobic gas (20:80% by volume CO_2:N_2) for 12 h (anaerobic period: 2100–0900 h). This cycle was repeated twice. Thus, the total experimental duration was ~48 h. Water in the treatment 2 (T2) tub was continuously sparged with air for the entire experiment. We illuminated both tubs with a 12 W light-emitting diode (LED) growing light (BloomBoss Spot, Framingham, Massachusetts) from 0900-2100 h and covered them from 2100-

0900 h to mimic day and night conditions, respectively, and maintain algal growth. The average photosynthetically active radiation (PAR) emitted by the growing light was within the 0.3 to 50 mol m–2 d–1 range reported by other investigators (Mulholland et al. 2001, Julian et al. 2011). We added a C source composed of 20 mg acetate L^{-1} (sodium acetate salt) to both tubs at 2100 h each evening of the experiment. We used acetate because it is considered to be the preferred C source of bacteria in the wastewater treatment process. The acetate concentration we used was within the range observed in streams (~2–22 mg acetate L^{-1}; Johnson et al. 2009). The experiment ended at 0900 h when T1 had completed the 2nd anaerobic period and T2 was aerobic (see Figure 2.2 for a schematic of the experimental design).

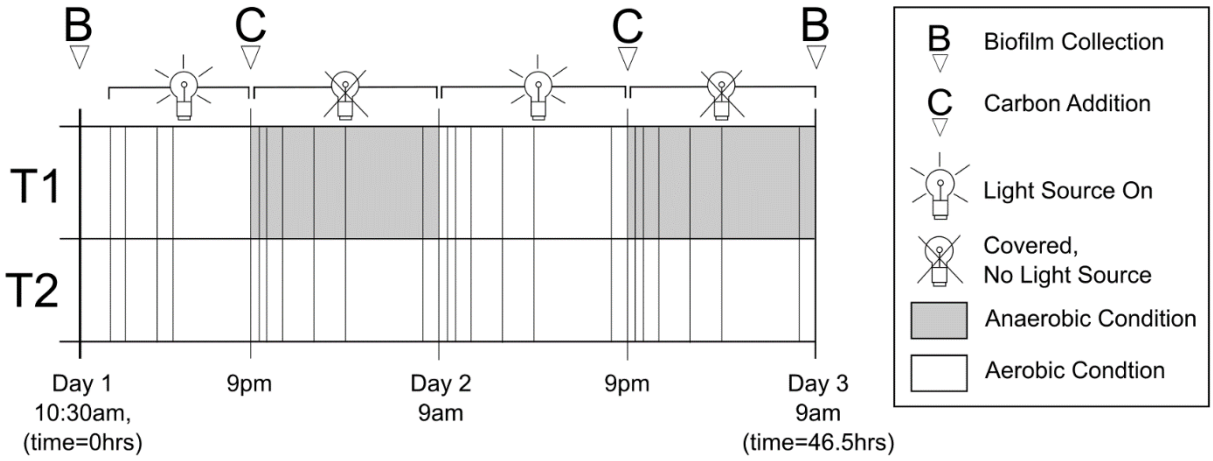

Figure 2. 2 Experimental design schematic for treatment 1 (T1) and treatment 2 (T2). Environmental measurements and water samples collection times are shown as vertical lines. See the methods section for further description.

Water Sampling and Analysis

We measured environmental variables including pH (AR50; Fisher Scientific Accumet Research, Waltham, Massachusetts), DO (YSI 550A; Yellow Springs Instruments, Yellow Springs, Ohio), temperature (AR50, YSI 550A), and light intensity (Hobo Pendant Temperature/Light Data Logger UA-002-XX; Onset Corporation, Bourne, Massachusetts) in both tubs over the duration of the experiment. We took water samples from each tub at regular intervals (i.e., at 30 min, 1, 2.5, 5, 7, 11, and 12 h) over the duration of the experiment (vertical lines in Figure 2.2). We collected each water sample with a 10-mL syringe and filtered the sample through a 0.45-μm filter (Supor Membrane Disc Filter, 25-mm diameter; Pall Life Sciences, Port Washington, New York). Immediately after filtration, we used a 23-gauge needle (BD 305165; Becton–Dickinson, Franklin Lakes, New Jersey) to transfer 1 mL of the filtered sample to an evacuated 10-mL glass vial, filled the vial with N2 gas, and stored the sample for future Fe^{2+} analysis. To halt potential microbial cycling of nutrients, we added 10 μL of 30% H2SO4 to the remaining 9 mL of filtered sample before storing it at 4°C until analysis (within 1 wk). We maintained the water level in the tubs at 1 L by adding 10 mL of distilled water after each sampling event. We had replaced ~25% of the water by the end of the experiment.

We analyzed water samples in duplicate (i.e., n = 2 laboratory replicates/sample) for dissolved PO_4^{3-} (as PO_4^{3-}-P) with the Mb-blue method (USEPA 1978) on an automated wet-chemistry analyzer (FS3000; Xylem Analytics O.I. Analytical, Beverly, Massachusetts). We averaged laboratory replicates and used these averages for all statistical analyses. The standard deviations of these laboratory replicates were <0.001 mg PO_4^{3-}-P L^{-1} (nonnormalized concentration) for all samples. We also analyzed water samples for dissolved concentrations of Ca, K, Mg, total Mn,

and total S using an Inductively Coupled Plasma-Mass Spectrometer (ICP-MS; Thermo Jarrell Ash, Franklin Massachusetts) and corrected them for S added during acidification. We analyzed evacuated samples for dissolved Fe^{2+} with the ferrozine assay (Lovley and Phillips 1986) adapted for a plate reader. We added 4 μL of filtered sample and 196 μL of 2 mM ferrozine (pH 7.0) made up in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer solution to each well of a 96-well plate (Corning 3595; Corning, Corning, New York) and analyzed with a plate reader (Tecan Infinite M200 Pro, Tecan, Männedorf, Switzerland) at 562 nm. All analyte concentrations discussed henceforth refer to the dissolved form.

Biofilm Extractions and Microbial Preservation

We scraped biofilms from cobbles at the start and end of the experiment and saved them for analyses including polyP and TP extractions, cell counts, and PCR (Figure 2.2). We normalized water-quality variables against the mass of wet biofilm scraped from each tub and biofilm extractions by total surface area of cobbles in each tub. Normalization was necessary because cobbles were not exactly the same size and stream biofilm cover was not always uniform. We measured cobble surface area with the aid of ImageJ software (version 1.48; Rasband 2014). PolyP and TP extracts were restricted to n = 2 replicates/treatment because of the limited quantity of biofilm scraped from the cobbles. We used 1 g (n = 2 replicates/treatment) of wet scrapings to estimate biofilm polyP concentration according to the hot-water extraction method presented by Eixler et al. (2005) and 2 g (n = 2 replicates/treatment) of wet scrapings/tub to estimate biofilm TP using the standard ammonium persulfate digestion (USEPA 1978). After digestion, we filtered biofilm polyP and TP extractions through a 0.45-μm filter and analyzed the

PO_4^{3-} (as PO_4^{3-}-P) concentration of each with standard Mb-blue methods (USEPA 1978) on with an automated wet-chemistry analyzer (FS3000).

We fixed 2 portions per treatment of ~0.5 mL of wet biofilm immediately after scraping. To do this, we incubated the 0.5 mL of biofilm with 1.5 mL of freshly prepared 4% paraformaldehyde in a 1× phosphate-buffered saline (PBS) solution for 1.5 h, washed the sample twice with 1× PBS by spinning the sample at 9000 g for 2 min and removing supernatant, and stored it in 1:1 1× PBS–95% ethanol at –20°C until cell counts were made. We stored the remaining biofilm unfixed at –20°C for DNA extraction and PCR analysis.

Cell Counts

We homogenized preserved biofilm samples (n = 2 replicates/treatment) further by repeatedly flushing the sample through a 26-gauge needle (30×/sample). We added 50 µL of 10-µg/mL 4',6-diamindino-2-phenylindole (DAPI) to 50 µL of homogenized sample to obtain a final DAPI concentration of 5 µg mL^{-1} as suggested by Hung et al. (2002) for polychromatic polyP staining. We stored stained samples in the dark at 4°C for 3 h. After the staining period, we centrifuged samples at 9,000 g for 2 min and replaced the supernatant with 1× PBS. We diluted samples 400× with 1× PBS, vortexed them briefly, and filtered 1 mL of each dispersed, diluted sample onto a 0.2–µm black filter (GTBP02500 Isopore Membrane Filter, 25 mm diameter; EMD Millipore, Darmstadt, Germany). We placed the black filter on a glass microscope slide with 15 µL of 0.1% p-phenylenediamine solution (anti-fade) in 1:1 PBS–glycerol under the filter and between the filter and the coverslip. We stored slides in the dark at 4°C until cell counts could be

made on an epifluorescence microscope (BX53F; Olympus, Tokyo, Japan) at 100× magnification with immersion oil. We counted DAPI-DNA stained cells on 10 randomly chosen fields/sample with a standard DAPI filter set (excitation 387/11, dichroic 409, emission 447/60) and counted DAPI–polyP stained cells on the same 10 fields with a modified DAPI filter set as described by Hung et al. (2002) but adapted for a color camera (excitation 360/70, dichroic 400, emission 420 long pass). Cells with stored polyP granules appeared light blue on the DAPI–DNA filter set and yellow under the DAPI–polyP filter set. We used cell counts to obtain measurements of the percentage of cells with stored polyP granules. We used 10 views/sample and 2 samples/treatment, resulting in 20 views/treatment.

Polymerase Chain Reaction (PCR) Analysis

We extracted DNA from each unfixed biofilm sample (n = 2 replicates/treatment) with a MoBio PowerSoil DNA Isolation Kit (12888, Qiagen MoBio, Germantown, Maryland) and tested samples for known PAO genes with PCR primers. We used the 518f/846f and 763f/1170r primer pairs and PCR conditions suggested by He et al. (2007) to test for the presence of CAP 16S rRNA and CAP type I ppk1 genes, respectively. We obtained positive controls from the anoxic zone of the University Area Joint Authority wastewater treatment plant in State College, Pennsylvania, which uses the EBPR process.

Statistical Analysis

We conducted all statistical analyses in R (version 3.0.2; R Project for Statistical Computing, Vienna, Austria). We used a generalized additive model (GAM) to test whether the experimental

treatment had an effect on PO_4^{3-} concentration as a function of time. We applied the GAM using the mgcv package (Wood 2006, 2015, Zuur 2009). The general form of the model is given as:

$$Y \sim s(t, \text{by } T1) + s(t, \text{by } T2) + \varepsilon \qquad \text{(Equation 2.1)}$$

where Y represents the response variable, t represents time, the function s( ) represents the additive response smoothed over time conditioned on treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only), and ε is the mean zero Gaussian random error. We used time as the main GAM predictor. Our analysis focused on changes in PO_4^{3-} concentration, but the response of other inorganic molecules also was of interest. Thus, we used the GAM to assess whether the treatment affected Ca, K, Mg, total Mn, Fe^{2+}, and total S concentrations over time. We used linear models to quantify correlations of PO_4^{3-} concentrations with other inorganic molecules. We compared the quality of these linear models using the Akaike information criteria (AIC). A lower AIC value indicates a better model fit. We used a 1-way analysis of variance (ANOVA) accompanied by Tukey pairwise comparison tests to identify statistically significant treatment groups (start, end of T1, end of T2) for both polyP and TP biofilm extracts. We used a 2-sample t-test to compare the percentage of cells with polyP granules at the end of T1 to the percentage at the end of T2.

RESULTS

Phosphate (PO_4^{3-}) Cycling

We induced cyclic redox conditions in T1 that produced a pattern of PO_4^{3-} release and uptake that was not seen in T2. The pattern in T1 differed significantly from the null model ($p < 0.001$; Figure 2.3, Table 2.2). During anaerobic conditions, PO_4^{3-} concentrations in T1 peaked at 0.009

and 0.005 mg PO_4^{3-}-P L^{-1} g^{-1} wet biofilm for cycles 1 and 2, respectively. During aerobic conditions, PO_4^{3-} concentrations in T1 decreased to ~0.002 mg PO_4^{3-}-P L^{-1} g^{-1} wet biofilm and converged with PO_4^{3-} concentrations in T2 (Figure 2.3). In the continuously aerobic control (T2), PO_4^{3-} concentration was ~0.002 mg PO_4^{3-}-P L^{-1} g^{-1} wet biofilm for the duration of the experiment (Figure 2.3). This pattern did not differ from the null model (p = 0.28; Table 2.2).

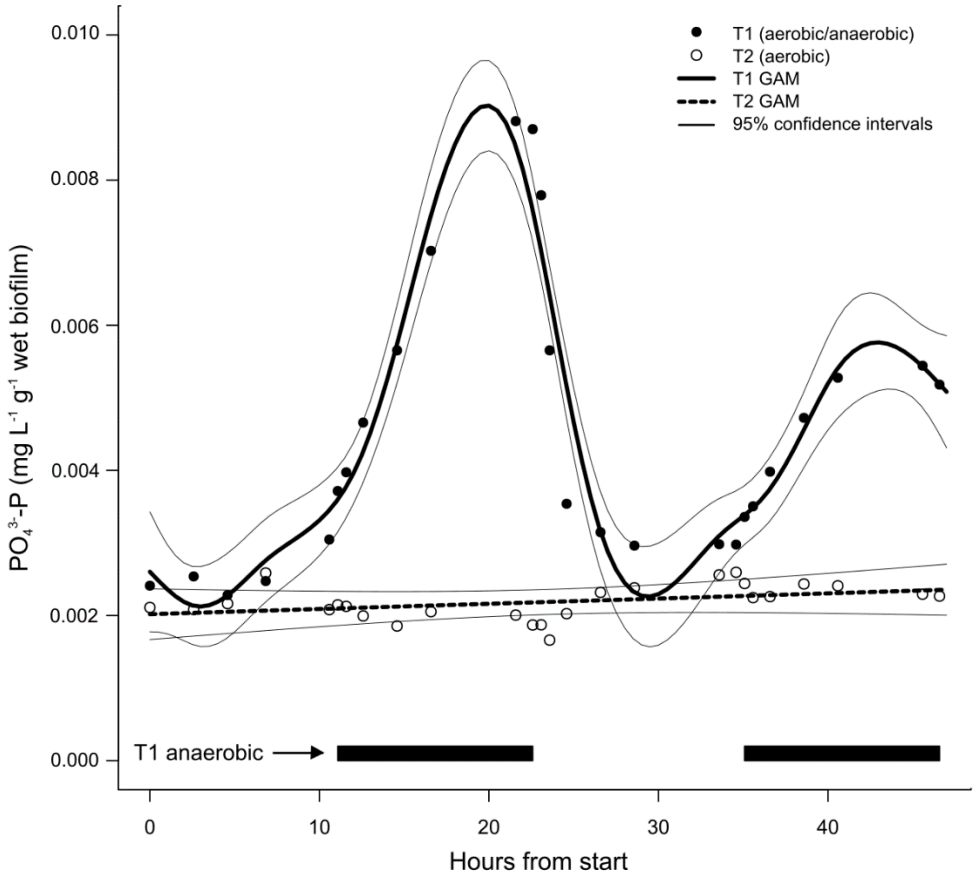

Figure 2. 3 Observed (points) and modeled (lines) PO_4^{3-}-P concentrations in the surrounding water normalized by mass of wet biofilm in the tub as a function of time for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). Black bar indicates periods when T1 was anaerobic. Unmarked periods for T1 were aerobic. Samples at 0 h were taken in the stream and the remaining samples were taken in the laboratory.

Table 2. 2 Generalized additive model results for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). SE = standard error, TS = total S, TMn = total Mn.

	PO_4^{3-}-P ~ time	Ca ~ time	K ~ time	Mg ~ time	TMn ~ time	Fe^{2+} ~ time	TS ~ time	pH ~ time
Comparison to null model								
Estimate (intercept)	2.19×10^{-3}	3.2	0.35	0.86	–0.14	0.11	18.52	4.66
SE	8.67×10^{-5}	0.16	0.04	0.08	0.03	0.02	0.85	0.59
t	25.23	20.11	9.37	10.15	–5.43	4.63	21.88	7.85
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
T1 smooth term								
F	77.41	116.76	29.23	7.84	148.22	2.84	3.78	9
p	<0.001	<0.001	<0.001	<0.001	<0.001	0.08	0.03	<0.001
T2 smooth term								
F	1.21	10.96	4.7	34.22	14.76	4.79	0.58	17.4
p	0.28	<0.001	0.002	<0.001	<0.001	0.04	0.45	<0.001
Deviance explained (%)	95	98.1	97.3	94.3	97.5	38	14.5	76.6

Biogeochemical Feedbacks

We assessed the uptake and release of other inorganic analytes over time to examine alternative controls on PO_4^{3-} cycling (Table 2.2, Figures 2.4A–F). In T1, among the analytes studied, only Fe^{2+} and total S concentrations (alternating anaerobic/aerobic conditions) did not exhibit discernable cycles of uptake and release (Figures 2.4A and 2.4B). The GAM confidence intervals for the 2 treatments overlapped for both Fe^{2+} and total S, and the model explained only 14.5 and 38.0% of the deviance for Fe^{2+} and total S, respectively (Table 2.2). Total Mn concentrations mimicked PO_4^{3-} concentrations in T1 (Figure 2.4C), and patterns in both treatments were significantly different from the null model (Table 2.2). In T1, peak total Mn concentrations during the 1st and 2nd anaerobic cycles increased from 0.05 to 0.15 mg total Mn L^{-1} g^{-1} wet biofilm, respectively (Figure 2.4C). In T1, Ca, K, and—to a lesser extent—Mg concentrations showed clear patterns of uptake/release under aerobic/anaerobic conditions, respectively (Figures 2.4D–F). In both treatments, patterns for Ca, K, and Mg differed significantly from the null model (Table 2.2). During anaerobic periods in T1, peak concentrations of Ca and K were consistent in time and equal to ~8.5 and ~0.55 mg L^{-1} g^{-1} wet biofilm, respectively, (Figures 2.4D and 2.4E). During aerobic conditions, Ca concentrations in T1 converged with concentrations in T2 (~4 mg L^{-1} g^{-1} wet biofilm) and K concentrations in T1 approached concentrations in T2 (~0.3 mg L^{-1} g^{-1} wet biofilm). Similar to PO_4^{3-} the peak Mg concentration in T1 during the 1st and 2nd anaerobic cycles decreased from 1.4 to 1.2 mg L^{-1} g^{-1} wet biofilm, respectively (Figure 2.4F). However, in contrast to PO_4^{3-}, the Mg concentration decreased consistently over time in T2. This decrease suggests that the lower Mg peak during the 2nd anaerobic period of T1 may be a result of dilution effects, whereas the lower PO_4^{3-} peak during T1 is not (Figures 2.3 and 2.4F). We used linear models to further explore the importance of

measured inorganic compounds in relation to PO_4^{3-} cycling (Table 2.3). According to linear model results, concentrations of Ca, K, and Mg were significantly related to PO_4^{3-} concentration (Table 2.3). The PO_4^{3-} vs Ca model had the most explanatory power (lowest AIC = –271.69; Table 2.3), followed by the PO_4^{3-} vs K, total Mn, and Mg models in order of decreasing explanatory power.

Table 2. 3 Linear model results for PO_4^{3-}-P vs concentrations of Ca, K, Mg, total Mn (TMn), Fe^{2+}, and total S (TS) measured in the surrounding water for treatment 1 (T1; alternating anaerobic/aerobic conditions). SE = standard error, AIC = Akaike Information Criterion.

	Estimate	SE	*t*	*p*	AIC
PO_4^{3-}-P ~ Ca model					–271.69
Intercept	-5.98×10^{-4}	8.52×10^{-4}	–0.70	0.49	
Slope	7.82×10^{-4}	1.27×10^{-4}	6.18	<0.001	
PO_4^{3-}-P ~ K model					–267.41
Intercept	-5.32×10^{-3}	1.84×10^{-3}	–2.89	0.008	
Slope	2.03×10^{-2}	3.79×10^{-3}	5.36	<0.001	
PO_4^{3-}-P ~ Mg model					–251.34
Intercept	-5.16×10^{-3}	4.60×10^{-3}	–1.12	0.27	
Slope	7.73×10^{-3}	3.68×10^{-3}	2.1	0.05	
PO_4^{3-}-P ~ TMn model					–256.88
Intercept	3.80×10^{-3}	3.71×10^{-4}	10.25	<0.001	
Slope	2.67×10^{-2}	8.01×10^{-3}	3.34	2.73×10^{-3}	
PO_4^{3-}-P ~ Fe^{2+} model					–75.16
Intercept	5.54×10^{-4}	8.21×10^{-4}	6.75	<0.001	
Slope	-5.53×10^{-2}	3.13×10^{-2}	–1.77	0.13	
PO_4^{3-}-P ~ TS model					–250.43
Intercept	9.32×10^{-4}	1.93×10^{-3}	0.48	0.63	
Slope	1.63×10^{-4}	8.80×10^{-5}	1.85	0.07	

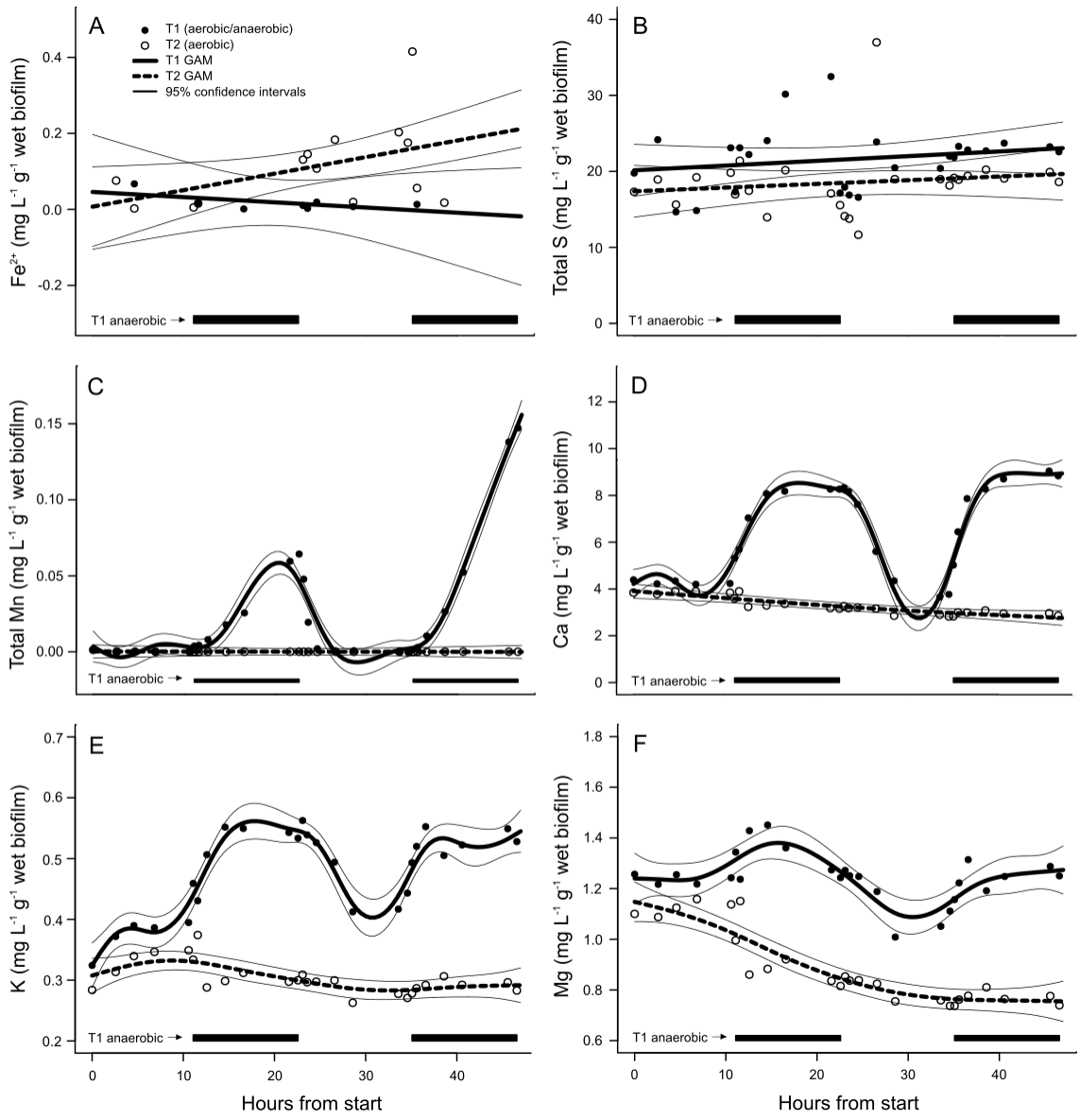

Figure 2. 4 Observed (points) and modeled (lines) concentrations of Fe^{2+} (A), total S (B), total Mn (C), Ca (D), K (E), and Mg (F) in the surrounding water normalized by mass of wet biofilm in the tub as a function of time for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only). Black bar indicates periods when T1 was anaerobic. Unmarked periods for T1 were aerobic. Samples at 0 h were taken in the stream and the remaining samples were taken in the laboratory.

The Role of PolyP Storage in Stream Biofilms

At the end of the experiment, the percentage of biofilm cells with stored polyP visualized with epifluorescence microscopy was significantly higher for aerobic (T2) than anaerobic (T1) treatments ($p = 0.004$; Figure 2.5A). The average percentages of cells with stored polyP were 6.4 and 18.9% for T1 and T2, respectively. A sample microscopy image taken of the T2 biofilm sample shows a typical cell (white arrow in Figure 2.5B) with stored polyP (white arrow in Figure 2.5C) similar to cells identified in other studies (Hung et al. 2002, Eixler et al. 2005) based on this technique. TP and polyP extraction findings did not agree with microscopy results. We observed no statistically significant differences between TP and polyP extractions from T1 and T2 at the end of the experiment ($p > 0.05$, Table 2.4). Last, we did not detect the presence of commonly targeted CAP genes (16S rRNA and type I ppk1 genes) in biofilm samples with PCR, but we did detect these genes in our positive control EBPR sludge sample.

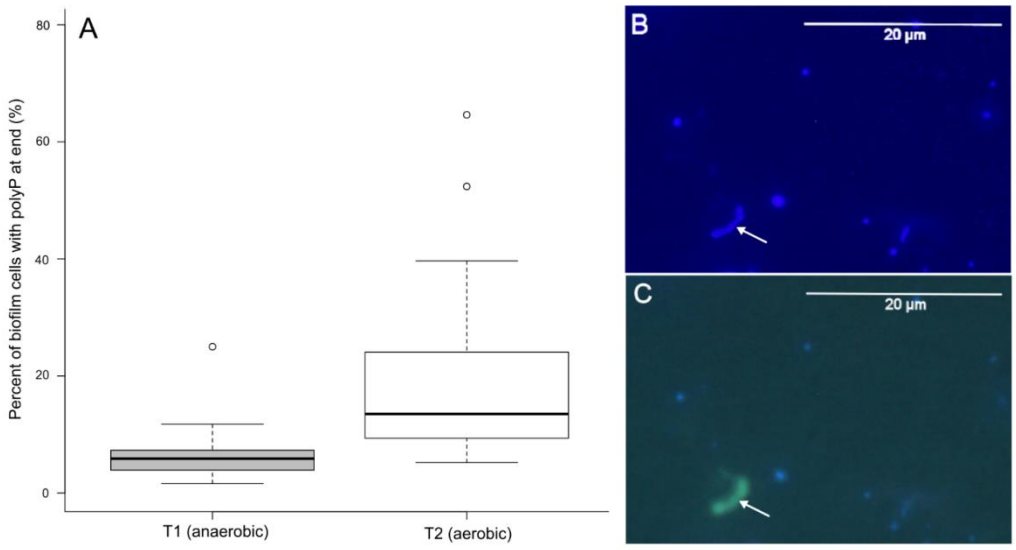

Figure 2. 5 Percentage of biofilm cells (*n* = 20 fields/treatment) with stored polyphosphate (polyP) at the end of the experiment in treatment 1 (anaerobic) and treatment 2 (aerobic). B.—Microscope image from a T2 biofilm sample at the end of the experiment and visualized with a 4',6-diamindino-2-phenylindole (DAPI)–DNA filter where light blue indicates the presence of DNA. C.—Microscope image taken a T2 biofilm sample at the end of the experiment and visualized with a DAPI–polyP filter where yellow indicates the presence of polyP. Arrow indicates an example cell with stored polyP granules.

Table 2. 4 Mean (standard error, *n* = 2) concentration of polyphosphate (polyP) and total P (TP) per unit mean biofilm surface area at the start and end of the experiment for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only).

Form of P	Start	T1 end (anaerobic)	T2 end (aerobic)
polyP (mg P/m^2)	1.77 (0.043)	1.42 (0.193)	1.27 (0.540)
TP (mg P/m^2)	19.43 (0.466)	11.03 (1.005)	15.63 (4.95)

DISCUSSION

Phosphate (PO_4^{3-}) Cycling

P patterns—higher PO_4^{3-} concentrations under anaerobic conditions than during aerobic conditions—support our hypothesis and provide evidence for PAO-mediated P cycling in stream biofilms. However, we cannot discount the impact of metal precipitation and reductive dissolution with P and discuss this in further detail below.

The reduced PO_4^{3-} peak in the 2nd anaerobic cycle in T1 (35–48 h; Figure 2.3) may have been caused by: 1) stresses on the biofilm ecosystem resulting in reduced abundance of metabolically active bacteria over time, 2) reduced availability of PO_4^{3-} as tub water was diluted with distilled water, or 3) differences in O_2 availability at the biofilm–water column interface that could not be detected based on water-column DO concentrations. Similar reductions in P flux over time were observed in biofilms taken from 8 Pennsylvania streams (Price and Carrick 2014).

Light should not have affected PO_4^{3-} concentration because light availability and timing were controlled variables during this study. Indeed, Carlton and Wetzel (1988) tested both short-term indirect (light) and direct (O_2) controls on PO_4^{3-} and found similar results. We discuss the role of precipitation and reductive dissolution of other analytes with PO_4^{3-} in further detail below, but PO_4^{3-} results tended to support our hypothesis and are consistent with the findings of other investigators (Carlton and Wetzel 1988, Drake et al. 2012, Matheson et al. 2012, Cohen et al. 2013).

Biogeochemical Feedbacks

The increase in Fe^{2+} indicated by the GAM during the 2nd half of the experiment (25–35 h; Figure 2.4A) was unexpected and does not follow known Fe redox patterns (i.e., a decrease in Fe^{2+} concentration under aerobic conditions). We suggest the increasing slope of the regression from T1 may be an artifact of sample variability combined with the high Fe^{2+} concentration observed at ~35 h (0.04 mg Fe^{2+} L^{-1} g^{-1} wet biofilm; Figure 2.4A). To test this hypothesis, we ran a GAM where Fe^{2+} concentration over time was described using a model with the slope = 0 and intercept = mean normalized Fe^{2+} concentration = 0.077 and found this null model to be statistically significant (data not shown, $p = 0.002$). Thus, the Fe^{2+} concentration can be represented by a constant value (= mean normalized Fe^{2+} concentration = 0.077) over time. We found no significant linear relationship between PO_4^{3-} and Fe^{2+} or TS, which further supports the disconnect between these analytes and PO_4^{3-} cycling (Table 2.3). Last, reductive dissolution of Fe may be important in stream or terrestrial systems where sediments or soils have high Fe concentrations (Dillon and Molot 1997, Dodds 2003, Chacon et al. 2006). However, the concentrations of Fe^{2+} were low in this experiment compared to other studies (e.g., Chacon et al. 2006) and further indicates the limited role of reductive dissolution of Fe. Thus, we suspect the reductive dissolution of Fe and S did not influence PO_4^{3-} -P patterns over time in our experiment.

In contrast to the lower peaks of PO_4^{3-} and Mg during the 2nd anaerobic cycle of T1 (Figure 2.3, Figure 2.4F), the higher 2nd peak of total Mn (40–48 h; Figure 2.4C) was unexpected, especially given the comparable pH (Figure 2.6) and DO concentrations (data not shown) between the 1st and 2nd anaerobic cycles. Thus, we hypothesize this increase was caused by release of Mn not previously exchanged with the water column; i.e., abiotic liberation of Mn from an anion (e.g.,

NO3–) other than PO_4^{3-} or biotic release of Mn. We also observed a significant linear relationship between PO_4^{3-} and total Mn concentrations ($p < 0.001$; Table 2.3). PO_4^{3-} precipitation with Mn is not thought to be as common as PO_4^{3-} precipitation with Fe, but the former can influence P cycling in sediments (White et al. 2008, Tye et al. 2016, Sulu-Gambari et al. 2016). Therefore, changes in PO_4^{3-} concentration over time may have been the result of PO_4^{3-} precipitation with Mn as $MnHPO_4$ or complexation with Mn-oxyhydroxides under aerobic conditions.

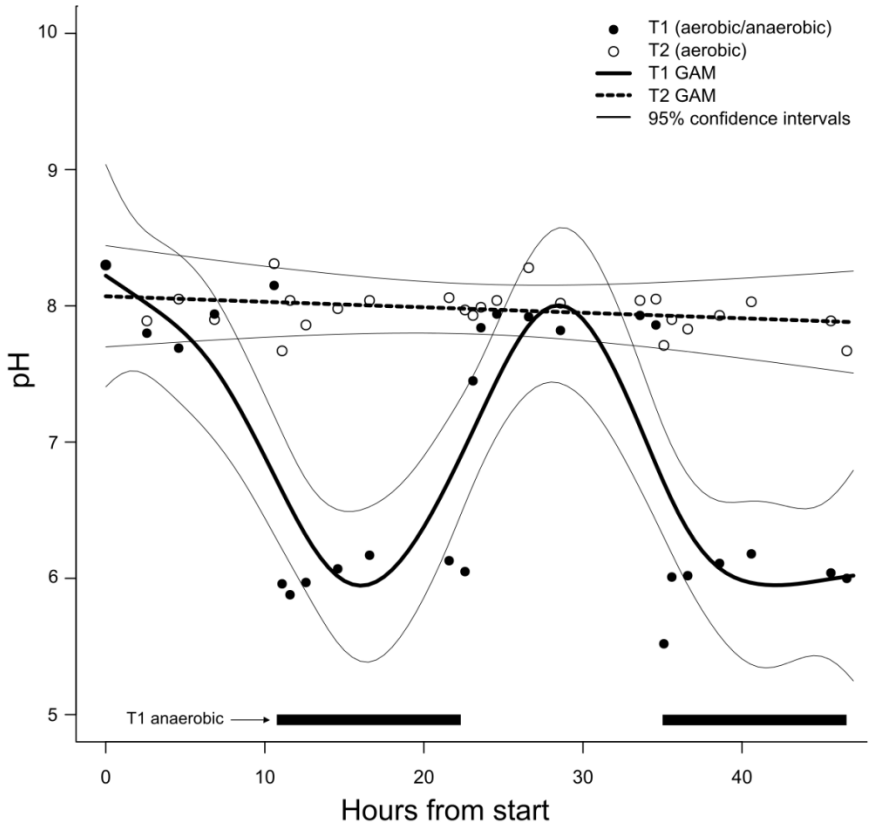

Figure 2. 6 Observed (points) and modeled (lines) pH of the surrounding water as a function of time for treatment 1 (T1; alternating anaerobic/aerobic conditions) and treatment 2 (T2; aerobic conditions only).

Addressing the role of Fe, S, and Mn interactions with PO_4^{3-} under changing redox conditions is important because reductive dissolution of these elements can mask PO_4^{3-} uptake and release attributed to PAOs present in the resident biofilms. For example, when DO concentrations are low, Fe^{3+} is reduced to Fe^{2+} and the PO_4^{3-} bound to Fe^{3+} is released into solution (Carlton and Wetzel 1988, Hupfer and Lewandowski 2008, Withers and Jarvie 2008). The process of reductive dissolution of Fe is mediated abiotically and biotically (Doong and Schink 2002). S can serve as an ultimate control on metal availability as well. When Fe^{2+} is bound to sulfide (S2–) under reduced conditions, it cannot precipitate with PO_4^{3-} (Dodds 2003). The reductive dissolution of Mn with PO_4^{3-} may also occur under changing redox conditions (Hongve 1997, White et al. 2008, Tye et al. 2016). When DO concentrations are low, dissolved forms of Mn such as Mn^{2+}, Mn^{3+}, or $Mn(HCO_3)$– dominate (Hem 1963). When DO concentrations are high, reduced Mn is oxidized to Mn^{4+} and forms precipitates, such as MnCO3, $MnHPO_4$, and MnO_2 (Hem 1963, Hongve 1997). Researchers have hypothesized the coupling of Fe, S, Mn, and P cycling by bacteria under changing redox conditions (Davelaar 1993, Diaz et al. 2012). In a study testing this hypothesis, cable bacteria were more abundant in oxic compared to anoxic sediments and stored polyP in oxic conditions (Sulu-Gambari et al. 2016). These authors also linked P dynamics to Fe, S, and Mn cycling. Specifically, Sulu-Gambari et al. (2016) suggested that the presence of oxic conditions at the sediment–water interface allows cable bacteria to use O_2 at the surface to oxidize S (originally FeS) in deeper sediments, which leads to the upward diffusion of Fe^{2+}. As it moves upward, Fe^{2+} is oxidized by Mn oxides at the surface to form Fe oxides that ultimately bind PO_4^{3-} (Sulu-Gambari et al. 2016). Thus, a combination of biotic and abiotic processes may control P cycling in environments undergoing changes in O_2 availability.

Our findings suggest a couple main points concerning abiotic controls on P cycling. First, the strong relationship between Ca and PO_4^{3-} may be important in this particular study system. Cohen et al. (2013) better explained diel patterns in water-column P concentrations when they accounted for Ca– PO_4 precipitation. They further demonstrated that 66 and 34% of P cycling in the Ichetucknee River were the results of biological and geochemical processes, respectively. The potential chemical controls on Ca– PO_4 precipitation include changes in pH because at high pH, Ca precipitates with PO_4^{3-} (Dodds 2003). In our experiment, pH was higher during aerobic and anaerobic conditions (Figure 2.6), which may have led to the precipitation of PO_4^{3-} with Ca. In addition to Ca, the precipitation of PO_4^{3-} with Mg may have occurred during aerobic conditions because of the associated pH increase. We did not measure Al, but we hypothesize that it behaves similarly to Ca and may have precipitated with PO_4^{3-} at higher pH as has been show in soils (Bloom 1981, Gerke 2010).

Second, these findings suggest the importance of biotic controls on PO_4^{3-} cycling. We observed patterns of uptake and release of K in T1 (Figure 2.4E) that mimicked patterns of PO_4^{3-} uptake and release. The precipitation of K with PO_4^{3-} was unlikely due to the tendency for K to stay dissolved in water, so the patterns of K uptake and release observed in T1 may be attributed to microbial cellular processes within the stream biofilm. Furthermore, it is not uncommon for K, Ca, Mn, and Mg to be chelators of stored polyP in microorganisms in wastewater sludge and elsewhere (Kornberg 1995, Pattarkine and Randall 1999, Schönborn et al. 2001). Therefore, the patterns of various analytes over time indicate that a combination of abiotic and biotic processes may have influenced stream biofilm P cycling in our experiment.

Besides the presence of metals, pH may also influence P cycling. For example, photosynthetically active stream biofilms can increase pH up to 1 unit (Dodds 2003). Carlton and Wetzel (1988) observed pH values at the sediment–water interface that increased from 7 to as high as 9 when they exposed sediments to light, whereas pH ranged from 7 to 8 during dark conditions (Carlton and Wetzel 1988). We did not test the effect of light but manipulated O_2 levels by sparging alternately with air and a mixed anaerobic gas composed of 20:80% by volume CO_2:N_2 (commonly used in anaerobic culturing experiments; e.g., Ingvorsen et al. 1984). This gas forced the tubs to pH of ~6 during anaerobic and 8 during aerobic periods (Figure 2.6) and may have affected the relative importance of abiotic and biotic controls on PO_4^{3-} cycling. We found a positive linear relationship between DO and pH ($p < 0.001$; data not shown). To test the effect of pH on PO_4^{3-} uptake and release, we added a linear pH term to the existing PO_4^{3-} GAM. Adding a term for pH improved the ability of the GAM to predict PO_4^{3-} concentrations over time ($p < 0.001$), but a change in pH from 6 to 8 decreased PO_4^{3-} concentration by only a small amount (0.0004 mg L^{-1} g^{-1} wet biofilm). In summary, the mixed anaerobic gas influenced pH, and thus, the relative importance of abiotic and biotic controls on PO_4^{3-} cycling in our experiment. However, PO_4^{3-} patterns over time probably were not controlled by abiotic processes alone.

The Role of Polyphosphate Storage in Stream Biofilms

Based on previous studies of EBPR wastewater treatment plant sludges (Hesselmann et al. 1999) and sediments (Carlton and Wetzel 1988), we anticipated higher polyP storage in biofilm organisms exposed to continuous aeration, such as those in T2. Our microscopy techniques yielded data that supported this expectation, but more traditional TP and polyP extraction

findings did not agree with microscopy results. The ratio of polyP to TP concentration measured by extraction (Table 2.4) agreed with ratios of polyP and TP biofilm extractions from other studies where polyP can make up anywhere from 10 to 50% of TP (Eixler et al. 2005, Rier et al. 2016, Taylor 2016). polyP estimates >10% may indicate biological processes (Hupfer and Lewandowski 2008). We suggest the difference in findings between extraction- and microscopy-based polyP measurements may have been an artifact of the limited number of samples and the spatio-temporal resolution targeted by these techniques. Microscopy-based polyP measurements represent intracellular polyP storage by microorganisms at the instant the stream biofilm sample is collected and fixed (Hung et al. 2002, Eixler et al. 2005, Aschar-Sobbi et al. 2008). In contrast, extraction-based polyP measurement does not distinguish between intracellular and extracellular polyP. Thus, it probably represents a more integrated polyP response in space and time. Moreover, microscopy-based polyP methods may capture the short-term polyP dynamics, such those targeted in this experiment, more accurately than extraction because stream biofilms can take up P on the order of minutes to hours (Carlton and Wetzel 1988, Cohen et al. 2013, Price and Carrick 2014).

We did not detect the presence of commonly targeted CAP genes (16S rRNA and type I ppk1 genes) in biofilm samples with PCR, but we did detect these genes in our positive control EBPR sludge sample. Our findings suggest that microorganisms other than CAP were present in the stream biofilm samples and exhibited PAO-like behavior similar to what has been observed in EBPR. Findings by Locke (2015) support our suggestion that CAP genes, while abundant in EBPR, are less common in stream biofilms. Locke (2015) found that <10% of bacteria accumulating polyP in Pennsylvania stream biofilms were members of the order Rhodocyclales,

which includes CAP. Accumulated polyP is used by diverse organisms (i.e., eukaryotes, archaea, and bacteria) as an energy source, P reservoir, and for many other functions (Kornberg 1995, Rao et al. 2009, Seviour et al. 2003). To the best of our knowledge, no degenerate ppk primers are available that could detect ppk genes in non-CAP PAOs. Despite not finding CAP-specific genes, our microscopy results support our hypothesis by providing in situ evidence of polyP storage by biofilm microorganisms under aerobic conditions. The emphasis of our experiment was to evaluate short-term temporal changes in P dynamics given varying O_2 levels rather than a spatial assessment of P dynamics in streams. Moreover, abiotic processes, such as precipitation and reductive dissolution of PO_4^{3-} with analytes, such as Ca, Mg, and Mn may have influenced patterns of PO_4^{3-} uptake and release in the stream biofilms over time in our experiment.

CONCLUSIONS

We tested the hypothesis that alternating aerobic/anaerobic water-column conditions produce patterns of dissolved PO_4^{3-} uptake/release from stream biofilms, respectively, that are similar to those attributed to PAOs in wastewater treatment. Dissolved PO_4^{3-} concentrations over time under alternating aerobic/anaerobic conditions support this hypothesis. In addition, significantly higher percentage of cells with intracellular polyP granules during aerobic conditions support the biotic mechanisms we proposed. However, results of other analyte concentrations over time point to complexities in coupled controls on P cycling in stream biofilms. Dissolved Ca, K, Mg, and total Mn uptake and release mimicked PO_4^{3-} cycling, which could be a result of cation accumulation by PAOs or precipitation of Ca-PO_4 and Mg-PO_4 and reductive dissolution of Mn-PO_4. The presence of intracellular polyP granules despite an absence of CAP genes points to our limited understanding of PAO diversity in the natural environment. Our results suggest that PAO

metabolism may be an important, but overlooked, component of in-stream P cycling and provide a biotic mechanism for short-term variation in water-quality patterns observed in stream ecosystems.

ACKNOWLEDGEMENTS

Author contributions: SMS, PJS, JMR, HJC, ARB, and MTW designed the study. SMS collected the data. SMS and PJS analyzed the data. SMS, PJS, JMR, HJC, ARB, NAL, and MTW interpreted the data. SMS drafted the manuscript. PJS, JMR, HJC, ARB, NAL, and MTW provided critical revision.

We thank Margaret Flemming, Peri Gerson, and Michelle Szeto for help with sampling, Erin Eggleston and Ruth Richardson for help with microscopy, Erin Menzies for help with land-cover descriptions and Community Science Institute data summaries, Shree Giri for help with ICP-MS sample analysis, Peter Marchetto and Keiran Catalina for help converting logger values to PAR equivalents, and the stream biofilms for their participation. SMS was supported by the Cornell University College of Agriculture and Life Sciences Land Grant Fellowship and US Environmental Protection Agency Science to Achieve Results Fellowship. This project was supported by funding from the US Department of Agriculture (UDSA) grant 2014-67019-21636. This article is contribution number 80 of the Institute for Great Lakes Research at Central Michigan University. Data and R scripts associated with this work are available on GitHub at https://github.com/sheilasaia/paper-p-cycling-in-stream-biofilms/tree/v1.0 and can also be

accessed via Zenodo (DOI:10.5281/zenodo.242599,

https://zenodo.org/badge/latestdoi/37694319).

REFERENCES

Albertsen, M., L. B. S. Hansen, A. M. Saunders, P. H. Nielsen, and K. L. Nielsen. 2012. A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. ISME Journal 6:1094–1106.

Amirbahman, A., B. A. Lake, and S. A. Norton. 2013. Seasonal phosphorus dynamics in the surficial sediment of two shallow temporal lakes: a solid-phase and pore-water study. Hydrobiologia 701:65–77.

Aschar-Sobbi, R., A. Y. Abramov, C. Diao, M. E. Kargacin, G. J. Kargacin, R. J. Frech, and E. Pavlov. 2008. High sensitivity, quantitative measurements of polyphosphate using a new DAPI-based approach. Journal of Fluorescence 18:859–866.

Bloom, P. R. 1981. Phosphorus adsorption by an aluminum-peat complex. Soil Science Society of America Journal 45:267–272.

Carlton, R. G., and R. G. Wetzel. 1988. Phosphorus flux from lake sediments: effect of epipelic algal oxygen production. Limnology and Oceanography 33:562–570.

Chacon, N., W. L. Silver, E. A. Dubinsky, and D. F. Cusack. 2006. Iron reduction and soil phosphorus solubilitzation in humid tropical forest soils: the role of labile carbon pools and an electron shuttle compound. Biogeochemistry 78:67–84.

Cohen, M. J., M. J. Kurz, J. B. Heffernan, J. B. Martin, R. L. Douglass, C. R. Foster, and R. G. Thomas. 2013. Diel phosphorous variation and the stoichiometry of ecosystem metabolism in a large spring-fed river. Ecological Monographs 83:155–176.

CSI (Community Science Institute). 2015. Fall Creek and Six Mile Creek databases. Community Science Institute, Ithaca, New York (Available from: http://www.communityscience.org/database)

Davelaar, D. 1993. Ecological significance of bacterial polyphosphate metabolism in sediments. Hydrobiologia 253(1–3):179–192.

Diaz, J. M., E. D. Ingall, S. D. Snow, C. R. Benitez-Nelson, M. Taillefert, and J. A. Brandes. 2012. Potential role of inorganic polyphosphate in the cycling of phosphorus within the hypoxic water column of Effingham Inlet, British Columbia. Global Biogeochemical Cycles 26:1–13.

Dillon, P. J., and L. A. Molot. 1997. Effect of landscape form on export of dissolved organic carbon, iron, and phosphorus from forested stream catchments. Water Resources Research 33:2591–2600.

Dodds, W. K. 2003. The role of periphyton in phosphorus retention in shallow freshwater aquatic systems. Journal of Phycology 39:840–849.

Doong, R.-A., and B. Schink. 2002. Cysteine-mediated reductive dissolution of poorly crystalline iron (III) oxides by *Geobacter sulfurreducens*. Environmental Science and Technology 36:2939–2945.

Drake, W. M., J. T. Scott, M. Evans-White, B. Haggard, A. Sharpley, C. W. Rogers, and E. M. Grantz. 2012. The effect of periphyton stoichiometry and light on biological phosphorus immobilization and release in streams. Limnology 13:97–106.

Dubrovsky, N. M., K. R. Burow, G. M. Clark, J. M. Gronberg, P. A. Hamilton, K. J. Hitt, D. K. Mueller, M. D. Munn, B. T. Nolan, L. J. Puckett, M. G. Rupert, T. M. Short, N. E. Spahr, L. A. Sprague, and W. G. Wilber. 2010. The quality of our Nation's waters - nutrients in the Nation's streams and groundwater, 1992 – 2004: US Geological Survey Circular 1350. US Geological Survey, Reston, Virginia.

Eixler, S., U. Selig, and U. Karsten. 2005. Extraction and detection methods for polyphosphate storage in autotrophic planktonic organisms. Hydrobiologia 533(1–3):135–143.
Gächter, R., J. S. Meyer, and A. Mares. 1988. Contribution of bacteria to release and fixation of phosphorus in lake sediments. Limnology and Oceanography 33(6, part 2):1542–1558.
Gerke, J. 2010. Humic (organic matter)-Al(Fe)-phosphate complexes: an underestimated phosphate form in soils and source of plant-available phosphate. Soil Science 175:417–425.
He, S., D. L. Gall, and K. D. McMahon. 2007. *Candidatus* Accumulibacter population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes. Applied and Environmental Microbiology 73:5865–5874.
Hem, J. D. 1963. Chemical equilibria and rates of manganese oxidation: chemistry of manganese in natural water. Water-Supply Paper 1667-A. US Geological Survey, Washington, DC.
Hesselmann, R. P., C. Werlen, D. Hahn, J. R. van der Meer, and A. J. B. Zehnder. 1999. Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. Systematic and Applied Microbiology 22:454–465.
Hongve, D. 1997. Cycling of iron, manganese, and phosphate in a meromictic lake. Limnology and Oceanography 42:635–647.
Hung, C., J. Peccia, J. L. Zilles, and D. R. Noguera. 2002. Physical enrichment accumulating activated polyphosphate in organisms sludge. Water Environment Research 74:354–361.
Hupfer, M., R. Gächter, and H. Ruegger. 1995. Polyphosphate in lake sediments ^{31}P NMR spectroscopy as a tool for its identification. Limnology and Oceanography 40:610–617.
Hupfer, M., S. Glöss, and H. P. Grossart. 2007. Polyphosphate-accumulating microorganisms in aquatic sediments. Aquatic Microbial Ecology 47:299–311.
Hupfer, M., S. Glöss, P. Schmieder, and H. P. Grossart. 2008. Methods for detection and quantification of polyphosphate and polyphosphate accumulating microorganisms in aquatic sediments. International Review of Hydrobiology 93:1–30.
Hupfer, M., and J. Lewandowski. 2008. Oxygen controls on phosphorus release in lake sediments—a long-lasting paradigm in limnology. International Reviews in Hydrobiology 90(4–5):415–432.
Hupfer, M., B. Rübe, and P. Schmieder. 2004. Origin and diagenesis of polyphosphate in lake sediments: a 31P-NMR study. Limnology and Oceanography 49:1–10.
Ingvorsen, K., A. J. B. Zehnder, and B. B. Jorgensen. 1984. Kinetics of sulfate and acetate uptake by *Desulfobacter postgatei*. Applied and Environmental Microbiology 47:403–408.
Jarvie, H. P., A. N. Sharpley, P. J. A. Withers, J. T. Scott, B. E. Haggard, and C. Neal. 2013. Phosphorus mitigation to control river eutrophication: murky waters, inconvenient truths, and "postnormal" science. Journal of Environmental Quality 42:295–304.
Jin, S., L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian. 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment 132:159–175.
Johnson, L. T., J. L. Tank, and C. P. Arango. 2009. The effect of land use on dissolved organic carbon and nitrogen uptake in streams. Freshwater Biology 54:2335–2350.
Julian, J. P., S. Z. Seegert, S. M. Powers, E. H. Stanley, and M. W. Doyle. 2010. Light as a first-order control on ecosystem structure in a temperate stream. Ecohydrology 4:422–434.
King, P. B., and H. K. Beikman. 1974. Explanatory text to accompany the geologic map of the

United States. US Government Printing Office, Washington, DC.
Kornberg, A. 1995. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. Journal of Bacteriology 177:491–496.
Kong, Y., J. L. Nielsen, and P. H. Nielsen. 2005. Identity and ecophysiology of uncultured Actinobacterial polyphosphate-accumulating organisms in full-scale enhanced biological phosphorus removal plants. Applied and Environmental Microbiology 71:4076–4085.
Kong, Y., Y. Xia, J. L. Nielsen, and P. H. Nielsen. 2007. Structure and function of the microbial community in a full-scale enhanced biological phosphorus removal plant. Microbiology 153:4061–4073.
Kristiansen, R., H. T. T. Nguyen, A. M. Saunders, J. L. Nielsen, R. Wimmer, V. Q. Le, S. J. McIlroy, S. Petrovski, R. J. Seviour, A. Calteau, K. L. Nielsen, and P. H. Nielsen. 2013. A metabolic model for members of the genus *Tetrasphaera* involved in enhanced biological phosphorus removal. ISME Journal 7:543–554.
Kunin, V., S. He, F. Warnecke, S. B. Peterson, H. Garcia Martin, M. Haynes, and P. Hugenholtz. 2008. A bacterial metapopulation adapts locally to phage predation despite global dispersal. Genome Research 18:293–297.
Liu, W. T., A. T. Nielsen, J. H. Wu, C. S. Tsai, Y. Matsuo, and S. Molin. 2001. In situ identification of polyphosphate- and polyhydroxyalkanoate-accumulating traits for microbial populations in a biological phosphorus removal process. Environmental Microbiology 3:110–122.
Locke, N. 2015. Classification of polyphosphate-accumulating bacteria in benthic biofilms. MS Thesis, Pennsylvania State University, State College, Pennsylvania.
Lovley, D. R., and E. J. Phillips. 1986. Organic matter mineralization with reduction of ferric iron in anaerobic sediments. Applied and Environmental Microbiology 51:683–689.
Martins, G., A. Terada, D. C. Ribeiro, A. M. Corral, A. G. Brito, B. F. Smets, and R. Nogueira. 2011. Structure \nd activity of lacustrine sediment bacteria involved in nutrient and iron cycles. FEMS Microbial Ecology 77:666–679.
Matheson, F. E., J. M. Quinn, and M. L. Martin. 2012. Effects of irradiance on diel and seasonal patterns of nutrient uptake by stream periphyton. Freshwater Biology 57:1671–1630.
McMahon, K. D., and E. K. Reed. 2013. Microbial contributions to phosphorous cycling in eutrophic lakes and wastewater. Annual Reviews of Microbiology 67:199–219.
Mulholland, P. J., C. S. Fellows, J. L. Tank, N. B. Grimm, J. R. Webster, S. K. Hamilton, E. Martí, L. Ashkenas, W. B. Bowden, W. K. Dodds, W. H. McDowell, M. J. Paul, and B. J. Peterson. 2001. Inter-biome comparison of factors controlling stream metabolism. Freshwater Biology 46:1503–1517.
Nakamura, K., A. Hiraishi, Y. Yoshimi, M. Kawaharasaki, K. Masuda, and Y. Kamagata. 1995. *Microlunatus phosphovoms* gen. nov., sp. nov., a new Gram-positive polyphosphate-accumulating bacterium isolated from activated sludge. International Journal of Systematic Bacteriology 45:17–22.
Nguyen, H. T. T., V. Q. Le, A. A. Hansen, J. L. Nielsen, and P. H. Nielsen. 2011. High diversity and abundance of putative polyphosphate-accumulating *Tetrasphaera*-related bacteria in activated sludge systems. FEMS Microbiology Ecology 76:256–267.
Nguyen, H. T. T., J. L. Nielsen, and P. H. Nielsen. 2012. "*Candidatus* Halomonas phosphatis", a novel polyphosphate-accumulating organism in full-scale enhanced biological phosphorus removal plants. Environmental Microbiology 14:2826–2837.
Pattarkine, V., and C. Randall. 1999. The requirement of metal cations for enhanced biological

phosphorus removal by activated sludge. Water Science and Technology 40:159–165.
Peterson, S. B., F. Warnecke, J. Madejska, K. D. McMahon, and P. Hugenholtz. 2008. Environmental distribution and population biology of *Candidatus* Accumulibacter, a primary agent of biological phosphorus removal. Environmental Microbiology 10:2692–2703.
Price, K. J., and H. J. Carrick. 2014. Quantitative evaluation of spatiotemportal phosphorus fluxes in stream biofilms. Freshwater Science 33:99–111.
Rasband, W. S. 2014. ImageJ. US National Institutes of Health, Bethesda, Maryland. (Available from: http://imagej.nih.gov/ij/)
Reddy, K. R., R. H. Kadlec, E. Flaig, and P. M. Gale. 1999. Phosphorus retention in streams and wetlands: a review. Critical Reviews in Environmental Science and Technology 29:83–146.
Rier, S. T., K. C. Kinek, S. E. Hay, and S. N. Francoeur. 2016. Polyphosphate plays a vital role in the phosphorus dynamics of stream periphyton. Freshwater Science 35:490–502.
Rao, N. N., M. R. Gomez-Garcia, and A. Kornberg. 2009. Inorganic polyphosphate: essential for growth and survival. Annual Reviews of Biochemistry 78:605–647.
Schönborn, C., H. D. Bauer, and I. Röske. 2001. Stability of enhanced biological phosphorus removal and composition of polyphosphate granules. Water Research 35:3190–3196.
Schulz, H. N., and H. D. Schulz. 2005. Large sulfur bacteria and the formation of phosphorite. Science 307:416–418.
Seviour, R. J., T. Mino, and M. Onuki. 2003. The microbiology of biological phosphorus removal in activated sludge systems. FEMS Microbiology Reviews 27:99–127.
Sherson, L. R., D. J. Van Horn, J. D. Gomez-Velez, L. J. Crossey, and C. N. Dahm. 2015. Nutrient dynamics in an alpine headwater stream: use of continuous water quality sensors to examine responses to wildfire and precipitation events. Hydrological Processes 29:3193–3207.
Sulu-Gambari, F., D. Seitaj, F. J. R. Meysman, R. Schauer, L. Polerecky, and C. P. Slomp. 2016. Cable bacteria control iron-phosphorus dynamics in sediments of a coastal hypoxic basin. Environmental Science and Technology 50:1227–1233.
Taylor, S. 2016. Polyphosphate accumulation in stream biofilms evaluated among sites of various land use using an *in-situ* enrichment experiment in Pennsylvania, USA. MS Thesis, Central Michigan University, Mount Pleasant, Michigan.
Tye, A. M., B. G. Rawlins, J. C. Rushton, and R. Price. 2016. Understanding the controls on sediment-P interactions and dynamics along a non-tidal river system in a rural-urban catchment: the River Nene. Applied Geochemistry 66:219–233.
Uhlmann, D., and H. D. Bauer. 1988. A remark on microorganisms in lake sediments with emphasis on polyphosphate-accumulating bacteria. Internationale Revue der gesamten Hydrobiologie und Hydrographie 73:703–708.
USEPA (US Environmental Protection Agency). 1978. Method 365.3. Phosphorous, all forms (colorimetric, ascorbic acid, two reagent). US Environmental Protection Agency, Washington, DC.
USGS (US Geological Survey). 2012. National Water Information System data available on the world wide web. USGS Water Data for the Nation. US Geological Survey, Reston, Virginia. (Available from: http://waterdata.usgs.gov/nwis/)
White, D. J., M. R. Noll, and J. C. Makarewicz. 2008. Does manganese influence phosphorus cycling under suboxic lake water conditions?. Journal of Great Lakes Research 34:571–

580.
Withers, P. J., and H. P. Jarvie. 2008. Delivery and cycling of phosphorus in rivers: a review. Science of the Total Environment 400(1–3):379–395.
Wood, S. N. 2006. Generalized additive models: an introduction with R. CRC Press, Boca Raton, FL.
Wood, S. N. 2015. mgcv: mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation. R package version 1.8-6. R Project for Statistical Computing, Vienna, Austria.
Zilles, J. L., C. H. Hung, and D. R. Noguera. 2002. Presence of *Rhodocyclus* in a full-scale wastewater treatment plant and their participation in enhanced biological phosphorus removal. Water Science and Technology 46(1–2):123–128.
Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York.

CHAPTER 3

THE IMPACT OF HYDROLOGY AND SOIL PHOSPHORUS CONCENTRATIONS ON THE RELATIVE ABUNDANCE OF POLYPHOSPHATE CYCLING GENES

ABSTRACT

Hydrology influence phosphorus (P) transport and cycling and shapes microbial processes. Researchers have suggested that hydrologically driven soil wetting/drying cycles may promote the growth of polyphosphate accumulating organisms (PAOs) and that PAOs may influence P cycling. Model PAOs have been identified in soils and a few studies have explored their ecological role in other environments (e.g., lake sediments). However, no studies have characterized PAO polyphosphate cycling genes (i.e., *ppk1*, *ppk2,* and *ppx*) and their correlation with soil P pools along a soil wetness index (SWI) gradient. We hypothesized (1) mobile forms of soil P would be low and (2) polyP functional genes would be more abundant in wetter sites (i.e., a high SWI). To test these hypotheses, we collected soil samples along a SWI gradient over time from sites in New York (NY) and Pennsylvania (PA) and analyzed each for three soil P pools (from most to least mobile): $CaCl_2$ extracted P ($CaCl_2$-P), oxalate extracted P (Ox-P), and total P. We carried out shotgun metagenomics sequencing on DNA extracted from soil samples, characterized dominant microbial taxa harboring *ppk1*, *ppk2*, and *ppx* genes, and determined the relative abundance of these genes along SWI and associated P gradients. SWI was a significant negative predictor of soil moisture and $CaCl_2$-P concentrations but not a significant predictor of *ppk1*, *ppk2*, or *ppx* relative abundances in NY and PA. Consistent with our hypothesis, we observed an overall negative trend between the relative abundance of all three functional genes versus $CaCl_2$-P concentration in NY. However, this negative trend was only statistically

significant in the case of *ppk2*-harboring contigs indicating that our results only weakly support the role of putative PAOs (i.e., contigs with either *ppk1* and *ppx* or *ppk2* and *ppx*) in P cycling in NY. In PA, we unexpectedly observed no significant relationship between the relative abundance of all three polyP functional genes and $CaCl_2$-P concentration. It is possible other P controls (e.g., landscape management) may have masked the role of PAO-mediated P cycling along SWI gradients included in NY and PA. We identified contigs harboring both *ppk1* and *ppx* from similar bacterial classes as known PAOs (e.g., *Betaproteobacteria*) as well as many understudied phyla (e.g., *Alphaproteobacteria* and *Melainabacteria*). These contigs may represent putative PAOs but additional research is needed to verify the PAO phenotype. While evidence for PAO-mediated P cycling in soils along a SWI gradient is limited in this study, these results indicate that SWI is a useful management tool for predicting soil $CaCl_2$-P pools. Our results also highlight the ubiquity of polyP functional genes in soils as well as the need for more research on the relative importance of abiotic and biotic controls on P transport and cycling in the environment.

KEYWORDS

topographic index, soil moisture, soil, phosphorus, polyphosphate accumulating organisms, polyphosphate, microbial processes, polyphosphate kinase, exopolyphosphatase

INTRODUCTION

Hydrological processes are an important driver of various controls on phosphorus (P) transport and cycling in the environment. Overland flow transports particle bound P (Hart et al. 2004;

Rittenburg et al. 2015) and has been a key motivation for the adoption of agriculture management strategies such as no tillage (Schelde et al. 2006; Kleinman et al. 2011a), cover crops (Kleinman et al. 2005; Kovar et al. 2011; Bergström et al. 2015), and other erosion control measures (Rittenburg et al. 2015; Bergström et al. 2015). Subsurface flows are also an important P transport mechanism. Therefore, management practices that only focused on reducing particle bound P may promote the transport of unbound P forms such as phosphate (Geohring et al. 2001; Kleinman et al. 2009; Verbree et al. 2010; Kleinman et al. 2011b; King et al. 2015; Christianson et al. 2016; Jarvie et al. 2017). Additionally, hydrological processes influence chemical controls on P through changes in soil redox conditions; soil saturation leads to reduced/anaerobic conditions while soil drying leads to oxidized/aerobic conditions (Silver et al. 1999; Pett-Ridge and Firestone 2005; Chacon et al. 2008; Burgin and Groffman 2012; Peralta et al. 2014). These redox changes—also referred to as wetting/drying cycles—control the precipitation and dissolution of P from soil organic matter (Guppy et al. 2005; Gerke 2010b) and soil minerals such as Ca, Fe-oxides, and Al-oxides (Chacon et al. 2006; Chacon et al. 2008; Gerke 2010b). Finally, hydrologically mediated wetting/drying cycles shape the structure and function of microbial communities (Pett-Ridge and Firestone 2005; DeAngelis et al. 2010; Ouyang and Li 2013; Peralta et al. 2014). For example, bacterial communities originating from wetland soils undergoing frequent wetting/drying cycles were more diverse and capable of greater community composition changes compared to bacterial communities with stable soil moisture patterns (Peralta et al. 2014). With respect to P, wetting/drying cycles impact soil microbial P pools (Grierson et al. 1998; Turner and Haygarth 2001; Turner et al. 2003; Blackwell et al. 2010; Dinh et al. 2016). Specifically, the difference in water-extractable P before and after drying was proportional to microbial P (Turner and Haygarth 2001). Mechanistically, wetting/drying cycles

serve as an environmental stress for soil microbes (Schimel et al. 1999; Ouyang and Li 2013). Drought leads to water-deficit stress in microbes and soil rewetting has been shown to lead to cell lysis and a subsequent release of P (Turner et al. 2003).

Another potential microbial mechanism controlling P availability that has received far less attention is the impact of changes in soil moisture on a group of organisms known as polyphosphate accumulating organisms (PAOs). PAOs remove P in specialized waste water treatments plants (WWTPs) and are phenotypically characterized by their ability to (1) take up and store phosphate as polyphosphate (polyP) during aerobic conditions and (2) release phosphate by breaking down stored polyP during anaerobic conditions (Seviour et al. 2003; Oehmen et al. 2007; Seviour and McIlroy 2008; Seviour and Nielsen 2010). PolyP is a polymer composed of three or more phosphate molecules (Kornberg 1995). Many diverse organisms including bacteria, fungi, plants, animals utilize polyP (Kornberg 1995; Zhang et al. 2002; Schulz and Schulz 2005; Rao et al. 2009) but their ability to use the energy from polyP hydrolysis to take up C (e.g., acetate) during anaerobic periods enables PAOs to outcompete other organisms that cannot take up C under anaerobic conditions (Gebremariam et al. 2011; Yuan et al. 2012).

PolyP synthesis and degradation in PAOs is thought to be regulated by *ppk* and *ppx* genes, respectively (Seviour et al. 2003; Skennerton et al. 2014). *ppk* encodes the polyphosphate kinase (PPK) enzyme that uses the energy released from converting ATP to ADP to form an intracellular polyP chain from phosphate molecules (Ahn and Kornberg 1990; Akiyama et al.

1992; Zhang et al. 2002; Rao et al. 2009; Kawakoshi et al. 2012). There are two known *ppk*'s: *ppk1* and *ppk2*. They have similar functions albeit the protein coded by *ppk2* (PPK2) catalyzes the making of polyP using both GTP and ATP and has a higher affinity for Mn^{2+}; whereas PPK1 has a higher affinity for Mg^{2+}(Zhang et al. 2002; Rao et al. 2009). A number of organisms have both *ppk*'s while some only have one (Zhang et al. 2002; Rao et al. 2009). Since the nucleotide sequences were first discovered, *ppk1* and *ppk2* homologs have been found in a variety of organisms (Zhang et al. 2002; Rao et al. 2009; Alcántara et al. 2014). *ppk* is especially well studied in the context of model PAOs in specialized WWTPs (e.g., He et al. 2007; Zhang et al. 2016) and is thought to be a good marker for bacterial strain diversity (Tzeng and Kornberg 1998). Additionally, it is thought to be more conserved than 16S rRNA for the model PAO, known by the provisional name as *Candidatus* Accumulibacter phosphatis (CAP; Kunin et al. 2008). *ppx* encodes the exopolyphosphatase (PPX) enzyme that preferentially removes phosphate from the terminal ends of a polyP chain (Akiyama et al. 1993; Keasling et al. 1993; Zago et al. 1999; Rangarajan et al. 2006; Rao et al. 2009). Another *ppx*, known as *ppx/gppA* codes for the specific PPX enzyme known as guanosine pentaphosphate phosphohydrolase (PPX/GppA). PPX/GppA has a similar function to PPX and can also hydrolyze the stress response protein pppGpp to ppGpp (Keasling et al. 1993; Reizer et al. 1993; Zago et al. 1999; Rao et al. 2009). *ppx* has been mainly studied in pure-culture (e.g., Zago et al. 1999; Alcántara et al. 2014) with the exception of one marine metagenome study (Temperton et al. 2011).

Given the diversity of organisms utilizing polyP metabolism and harboring *ppk* and *ppx* genes, it is possible that PAO-like metabolism may be widespread in the natural environment. Additionally, given the low nutrient density of soil (Young and Crawford 2004), it is possible

that organisms with ability to make and break polyP, would be better suited for P poor soil environments compared to those that cannot. Thus, researchers have hypothesized that PAOs are present in the environment (e.g., soils, sediments, and waterbodies) and influence P cycling under naturally occurring alternating anaerobic/aerobic conditions (Gächter et al. 1988; Davelaar 1993; Reddy et al. 1999; Schulz and Schulz 2005; Pett-Ridge and Firestone 2005; Hupfer et al. 2007; Hupfer et al. 2008; Hupfer and Lewandowski 2008; Peterson et al. 2008; Diaz et al. 2012; McMahon and Read 2013; McParland et al. 2015). In terms of the application of this hypothesis to soils, wetting and drying events may lead to subsequent anaerobic and aerobic pore space conditions, respectively, that select for and drive PAO-mediated P cycling (Pett-Ridge and Firestone 2005). Only a few studies have explicitly attempted to explore the presence of PAOs in soils (Kunin et al. 2008; Peterson et al. 2008; DebRoy et al. 2013; Li et al. 2013) and we know of none exploring their role in P cycling or looking for evidence of PAOs beyond the model PAO, CAP.

Exploring the role of PAO-mediated P cycling in soils begins with identification of soil moisture gradients across the landscape. In the field of hydrology, topographic wetness indexes have been used to integrate soil moisture over time; they are temporally static and provide a means to determine what parts of the landscape are more or less likely to be saturated (Beven and Kirkby 1979; Walter et al. 2002; Lyon et al. 2004). For example, the soil wetness index (SWI) incorporates both topography and soil properties and is a good predictor of soil moisture patterns that vary in space and time (Buchanan et al. 2014; Hofmeister et al. 2016). Topographic indexes have also been used to predict soil P availability (Marjerison et al. 2011; Thomas et al. 2016;

Wilson et al. 2016) but no studies have looked at their ability to predict microbial mechanisms controlling P cycling and transport in the landscape (Vidon et al. 2010; Bernhardt et al. 2017).

Therefore, the main objectives of this study were to (1) determine whether SWI could be used to predict various P pools, (2) characterize the taxa of soil microorganisms harboring P cycling functional genes, and (3) test whether the relative abundance of P cycling functional genes varied along SWI and associated P availability gradients. We hypothesized that SWI was a good predictor of mobile forms of P (i.e., phosphate) and the concentrations of these mobile P forms was inversely related to SWI. Put another way, we expected wetter parts of the landscape (i.e., large SWI) to experience flushing/loss of abiotically and biotically sourced P while drier parts of the landscape (i.e., small SWI) accumulate P. We expected to see taxa of well-studied PAOs (e.g., CAP) as well as unstudied, putative PAOs (i.e., contigs with either *ppk1* and *ppx* or *ppk2* and *ppx*). Finally, we hypothesized that the relative abundance of all three P cycling functional genes would be higher in wetter parts of the landscape because these environments would select for microorganisms capable of intracellular polyP synthesis and degradation. We expected the same trends between all three genes because putative PAOs require pairs of these genes—either *ppk1* and *ppx* or *ppk2* and *ppx*—to carry out polyP synthesis and degradation. From a P availability perspective, since we expected wetter parts of the landscape to have less P, it followed that the relative abundance of P cycling genes would be higher where P was low. To test these hypotheses, we collected soil samples along a SWI gradient over the course of six months and analyzed each sample for three different P pools (in order of most to least mobile P): dilute salt extractable P ($CaCl_2$-P), oxalate extractable P (Ox-P), and total extractable P (Total-P). We extracted DNA from each soil sample and sequenced DNA using shotgun metagenomic

analysis techniques. We then assembled sequences, determined relative abundances of *ppk1*, *ppk2*, and *ppx*, and assigned taxonomic identities to assembled reads harboring these same genes.

METHODS

Soil Wetness Index Calculations

We selected sample sites along a SWI gradient, which incorporates topography and soil properties (Beven and Kirkby 1979; Lyon et al. 2004), to ensure each sampling event included a full range of soil moisture conditions. We determined SWI values for New York (NY) and Pennsylvania (PA) study sites using methods described previously by Hofmeister et al. (2016). Briefly, we calculated the SWI according to Equation 1 (Walter et al. 2002; Lyon et al. 2004):

$$SWI = \ln\left(\frac{A}{k_s D \tan(B)}\right) \qquad \text{(Equation 1)}$$

where A is the upslope area per unit contour (m), B is the local slope angle (rad), k_s is the mean saturated hydraulic conductivity (m day^{-1}), and D is the depth to the restrictive layer (m). We determined upslope area and local slope angle by LiDAR-based digital elevation maps (DEM) for NY and PA sites. We obtained measurements of k_s and D from US Department of Agriculture (USDA) SURRGO soil data (USDA 2006). We used SAGA-GIS and the RSAGA packages in R (version 3.0.1) to generate the SWIs (Brenning, 2007; R Core Team, 2015). We used ArcGIS 10 (ESRI, Redlands, CA) to calculate the final SWI for each sample site; the area weighted average SWI within a 3 m buffer around the GPS coordinate of the site. Here, we used SWI to select sample sites along a soil moisture gradient, where low SWI values indicate sites that are less likely to saturate/be dry and high SWI values indicate sites that are more likely to saturate/be wet.

New York Study Sites

New York (NY) study sites (Table 1) were located in pasture/hay fields at the Cornell University Cornell University Teaching and Research Center in Harford, NY. We sampled sites once per month from May to October 2015. Based on rainfall data collected within a mile of the study site from 1981-2010, the 30-year average annual precipitation near the study region is 103.6 cm (Arguez et al. 2010). The average annual precipitation for 2015 was 96.9 cm (Menne et al. 2012); slightly less (by 6.4%) than the 30-year average (Figure 3.1). NY study sites were managed as pasture/hay and have not been cultivated since available records end in 2003. Dairy manure is spread in the spring at an average rate of 113 kg ha -1 P (as P_2O_5) just upslope of site NY55. However, no manure is applied near any other NY sites. Soils are moderately well drained Lobdell (fine-loamy, mixed, active, mesic Fluvaquenic Eutrudepts), well drained Bath (coarse-loamy, mixed, active, mesic Typic Fragiudepts), well drained Valois (coarse-loamy, mixed, superactive, mesic Typic Dystrudepts), somewhat excessively drained to well drained Howard (loamy-skeletal, mixed, active, mesic, Glossic Hapludalfs), and somewhat excessively drained to well drained Palymyra (fine-loamy over sandy or sandy-skeletal, mixed, active, mesic, Glossic Hapludalfs) series (USDA 2006; USDA 2016). The dominant soil series of this region (i.e., Bath, Howard, and Palymyra) have a clay layer from as shallow as 28 cm to as deep as 114 cm from the surface (USDA 2006; USDA 2016), which often results in a perched water table (Hofmeister et al. 2016).

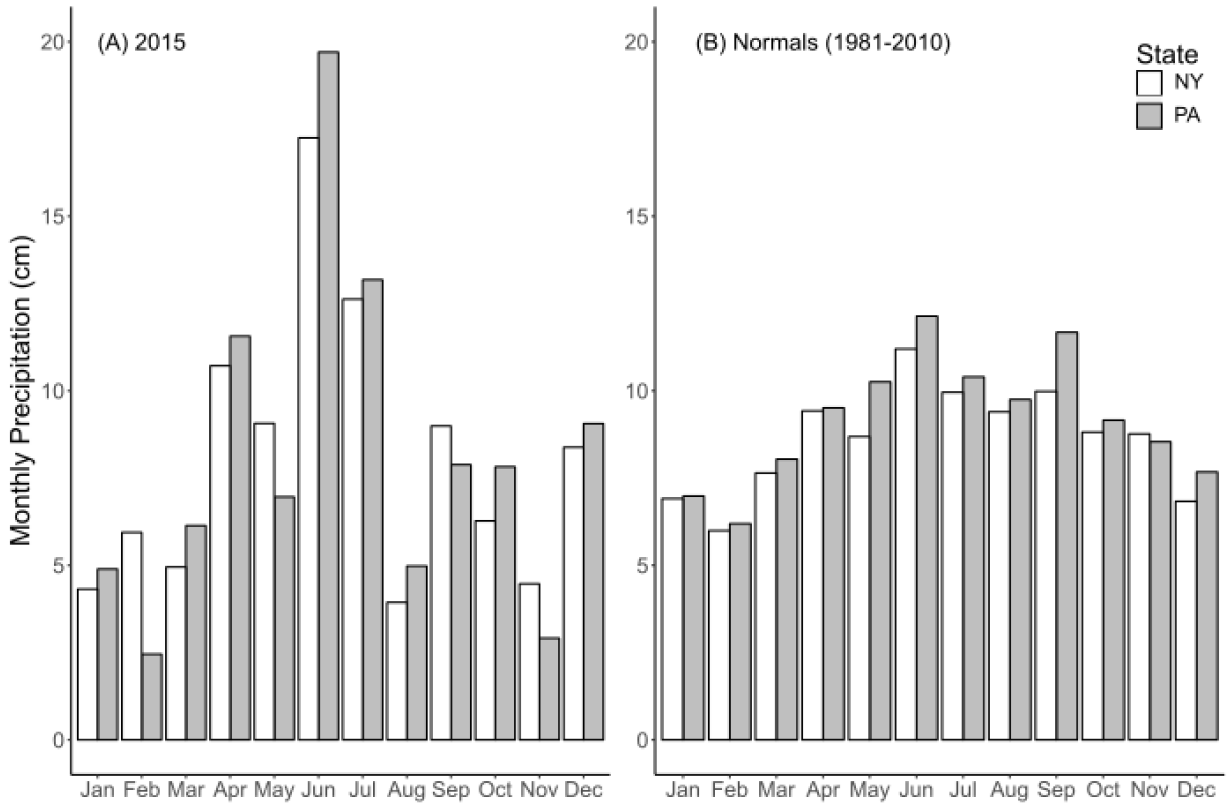

Figure 3. 1 Monthly precipitation at New York (NY) and Pennsylvania (PA) field sites for (A) 2015—the year of this study—and (B) 30-year normals.

Pennsylvania Study Sites

Pennsylvania (PA) study sites (Table 3.1) were located at the USDA Agricultural Research Service (ARS) Mattern Experimental Watershed near Klingerstown, PA. The Mattern watershed (11 ha) is part of the larger Mahantango Creek Watershed (7.3 km^2), which drains to the Chesapeake Bay. Both of these watersheds have been described previously (Buda et al. 2009; Bryant et al. 2011; Buda et al. 2013; Collick et al. 2015; Collick et al. 2016). We sampled sites once per month in May, July, and October. All sample sites were managed as a part of the USDA Conservation Reserve Program (CRP) and have not been farmed since 2008. However, sites are downslope of cropped corn-soybean fields that received an average 13.5 kg ha^{-1} P (as P_2O_5) from

beef and dairy manure in the fall and an average of 72.5 kg ha^{-1} of P (as P_2O_5) from 10-20-20:N-P-K chemical fertilizer in the spring from 2006 to 2010. Based on rainfall data collected within a mile of the PA study sites from 1981-2010, the average annual precipitation in the study region is 109.6 cm. The average annual precipitation for 2015 was 97.6 cm; slightly less (by 11.0%) than the 30-year average. Upper parts of the watershed are composed of well drained Berks (loamy-skeletal, mixed, active, mesic Typic Dystrudepts) series with an argillic (clay <30%) horizon (Buda et al. 2009). Lower parts of the watershed are composed of well drained Leck Kill (fine-loamy, mixed, semiactive, mesic Typic Hapludults) and somewhat poorly drained Albrights (fine-loamy, mixed, semiactive, mesic Aquic Fragiudalfs) series, the latter of which, is underlain by a fragipan at 60 cm (Buda et al. 2009).

Table 3. 1 Description of New York (NY) and Pennsylvania (PA) study sites. Abbreviations: soil wetness index (SWI).

Site ID	SWI	Soil Series Name	Soil Texture
NY Sites			
NY17	11.02	Erie	silt loam
NY24	4.95	Bath/Valois	channery silt loam/gravelly loam
NY25	6.90	Howard/Valois	gravelly loam
NY28	3.16	Howard/Palmyra	gravelly loam
NY29	9.42	Howard/Palmyra	gravelly loam
NY55	13.40	Lobdell	silt loam
PA Sites			
PA09	8.21	Leck Kill	channery silt loam
PA16	8.29	Leck Kill	channery silt loam
PA26	6.16	Berks	channery loam
PA45	5.82	Berks	channery loam
PA50	12.99	Albrights	silt loam
PA51	4.47	Berks	channery loam
PA54	11.17	Leck Kill	channery silt loam

Soil Sample Collection and Initial Processing

For each sample event, we collected the upper 5 cm of soil using a using a 2 cm diameter t-corer (Model HA, Oakfield Apparatus, Fond du Lac, Wisconsin). We wore nitrile gloves to avoid contamination while sampling and discarded the first core to 'rinse' our gloves and the corer. We collected an average of six cores per sample event in a 3 m radius around the sample GPS coordinate because 3 m is the average rated accuracy of our GPS unit (eTrex 20x 010-01508-00, Garmin, Olathe, Kansas). We pooled soil cores in a zip-lock bag, removed as much air as possible to prevent sample oxidation, and placed samples on ice until they could be stored in the fridge (4°C) for further processing (minimum 24 and 48 hours for NY and PA sites, respectively).

We processed samples by homogenizing the zip-lock bag by hand for approximately two minutes or until visually well mixed. We removed visibly large sticks, rocks, and leaves after pouring soil into aluminum tin wiped down with 95% ethanol. We filled sterile 15 ml centrifuge tubes with fresh soil for DNA extractions and placed them in the freezer (-20°C) until we could carry out this analysis. We covered the remaining soil and left it to air-dry. Once dry, we sieved the soil through a 2 mm sieve for subsequent analyses of three different P extractions as discussed in further detail below.

In-situ Measurements

For each soil sample, we measured soil moisture in the field using a total domain reflectometry (TDR) probe (Campbell Scientific Hydrosense II, CS658). We measured and then calibrated TDR measurements for NY and PA to gravimetric soil moisture measurements as described by

Hofmeister et al. (2016). We averaged all soil moisture readings for each site on a specific sample date and used this number henceforth for all analyses.

Phosphorus Extractions and Other Analysis

We carried out three P extractions for each soil sample (n=2): dilute-salt-extractable P ($CaCl_2$-P), oxalate-extractable P (Ox-P), and total extractable P (Total-P). In this order, these three pools represent soil P that is loosely- to strongly-bound. Specifically, the $CaCl_2$-P extraction provides a measure of soluble reactive phosphorus (SRP) that could potentially be mobilized from soil during an overland flow runoff event (Self-Davis et al. 2009). The Ox-P extraction serves as a measurement of P bound to amorphous Al/Fe-oxides in the soil (McKeague and Day 1966) and Total-P extraction provides a measure of the total soil P concentration. Therefore, of the three extractions, $CaCl_2$-P is likely more readily available to microorganisms.

To measure $CaCl_2$-P, we followed the method of Self-Davis et al. (2009). We weighed 1 g of air-dried, sieved soil, added 25 mL of 0.01M $CaCl_2$, shook samples for one hour on an end-over-end mixer, centrifuged for 10 min at 2100 g, and filtered soil extractions through a 0.45 μm filter (66191; 47 mm diameter, Pall Life Sciences, Port Washington, New York). We stored filtered extractions in the fridge (4°C) up to one week until analyzing for SRP (as P) concentration using the molybdenum blue method (USEPA, 1978) with an automated wet-chemistry analyzer (FS3000; Xylem Analytics O.I. Analytical, Beverly, Massachusetts). The quantification limit of the FS3000 machine is 0.01 mg l^{-1} and all calibration curves had an R^2 value of 0.999 or higher.

To measure Ox-P, we followed the method of McKeague and Day (1966). We weighed 0.5 g of air-dried, sieved soil, added 20 mL of oxalate solution (mixture of 0.2 M ammonium oxalate and 0.2 M oxalic acid adjusted to a pH of 3.0), shook each sample for two hours on an end-over-end mixer, centrifuged for 5 min at 2100 g, and filtered soil extractions through a 0.45 μm filter (66191; 47 mm, Pall Life Sciences, Port Washington, New York). We stored soil extractions in the dark at room temperature for up to one month until they could be analyzed for Ox-P, oxalate-extractable Fe (Ox-Fe), and oxalate-extractable Al (Ox-Al) concentrations using an Inductively Coupled Plasma-Mass Spectrometer (ICP-MS; Thermo Jarrell Ash, Franklin Massachusetts).

We followed standard protocols for analyzing the total amount of P (Total-P) in soil using concentrated HNO_3 and H_2O_2 (USEPA 3050B). We weighed out 1 g of air-dried, sieved soil, and placed soil samples in an automatic digester (Environmental Express AutoBlock, Charleston, South Carolina) set to run the USEPA 3050B method. We did not add concentrated HCl at the end of the digestion because these extractions were to be analyzed by ICP-MS (not ICP-AES). After digestion, we topped samples off to 50 mL using deionized water and filtered 20 mL into plastic scintillation vials using a 11 μm filter (1001-125; 125 mm diameter, Whatman, Maidstone, United Kingdom). We stored the capped scintillation vials at room temperature for up to one month until they could be analyzed for Total-P concentrations using an ICP-MS as discussed previously.

DNA Extractions and Sequencing Preparation

We extracted DNA in triplicate from pooled soil samples using a MoBio PowerSoil DNA Isolation Kit (12888, Qiagen MoBio, Germantown, Maryland). We pooled samples because we

wanted to focus on average site microbially-mediated P cycling patterns. Slight modifications to the MoBio DNA extraction protocol included: (1) two, 30 s homogenizations with the bead beater (10 s rest in between each homogenization), (2) elution with 50 μL (rather than 100 μL), and (3) pooling of all three elutions into one tube. We quantified DNA on an Invitrogen Qubit spectrofluorometer using a Qubit dsDNA broad range assay kit (Q32850, Thermo Fisher Scientific, Waltham, Massachusetts). We sheered DNA to an average insert size of 550 bp using a Covaris S2 adaptive focused acoustic disruptor (Covaris, Woburn, Massachusetts). We prepared libraries using an Illumina TruSeq Nano DNA HT Sample Prep Kit (FC-121-4003, Illumina, San Diego, California). We carried out shotgun metagenomic sequencing with one pooled, indexed sample on two lanes at the Cornell University Institute for Biotechnology Genomics Facility using an Illumina HiSeq2500 instrument 2x150 bp paired-end reads on Rapid Run mode.

Metagenomics Analysis

We used the workflow described in Figure 3.1 to process shotgun metagenomic sequencing data. Briefly, we quality controlled and trimmed raw reads using BBMap/BBTools (JGI 2017), assembled reads for samples from NY and PA sites separately with MEGAHIT using default settings (Li et al. 2015), called genes and translated gene sequences using Prodigal (Hyatt et al. 2010), annotated PPK1, PPK2, and PPX protein sequences using Pfam models and a HMMER reciprocal search using default settings (Finn et al. 2011), assigned taxonomy to contigs with genes of interest using Kaiju's protein-level NCBI RefSeq reference database classifier (Menzel et al. 2016), and mapped raw reads to determine relative gene abundances using bwa-mem (Li 2017). Relative gene abundances are given as the number of fragments (i.e., paired-end reads)

with a gene of interest per kilobase (kb) per million reads mapped (FPKM). Protein family models from the Pfam database (Finn et al. 2016) used in this study to identify *ppk* and *ppx* genes are shown in Table 3.2. For contigs with unique combinations of genes, we used pBLAST (NCBI 2017) to compare contig protein sequences to previously studied organisms. We selected the top three contigs by maximum relative abundance for each sample as well as the top seven contigs by mean relative abundance overall and display the z-score of these data in heatmaps. We also categorized each contig for whether it contained various combinations of *ppk1*, *ppk2*, and *ppx/gppA* (referred to henceforth as *ppx*). The data and scripts pertaining to this workflow are available on GitHub at https://github.com/sheilasaia/paper-p-cycling-in-soils.

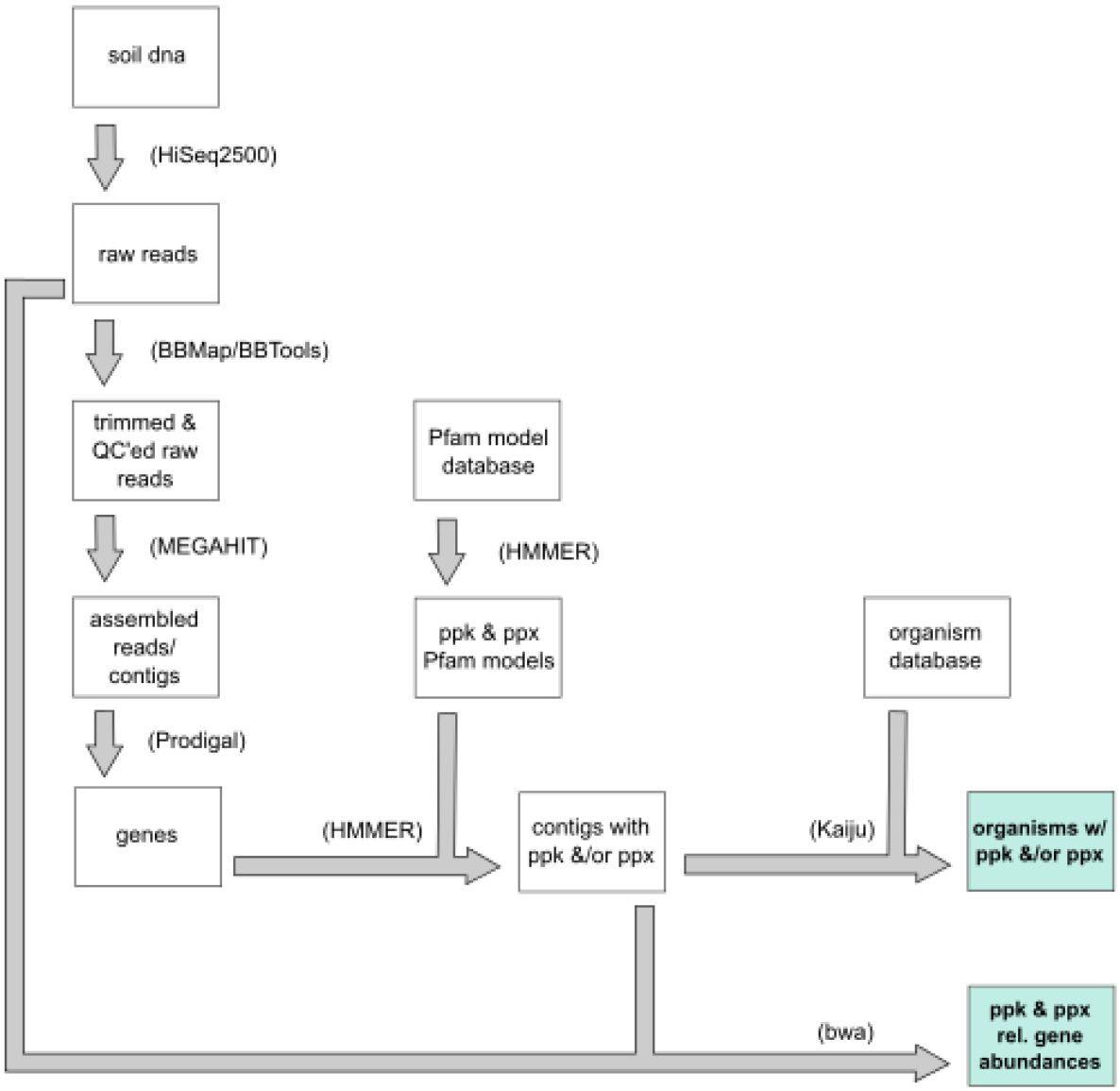

Figure 3. 2 Description of shotgun metagenomic sequencing data workflow used in this study. Final output tables used in subsequent analyses are bolded/shaded green.

Table 3. 2 Protein family (Pfam) database models used in this study.

Pfam ID	Protein	Pfam Description
PF02503.16	PPK1	polyphosphate kinase middle domain (PPK1)
PF13089.5	PPK1	polyphosphate kinase N-terminal domain (PPK1)
PF13090.5	PPK1	polyphosphate kinase C-terminal domain (PPK1)
PF03976.13	PPK2	polyphosphate kinase (PPK2)
PF02541.15	PPX	exopolyphosphatase (PPX/GppA)

Statistical Analysis

We conducted all statistical analysis in R (R Core Team 2015). Due to repeated sampling in time, we used a linear mixed effects model to test whether (1) SWI was a significant predictor of soil moisture and (2) SWI was a significant predictor of each of the three P pools (i.e., $CaCl_2$-P, Ox-P, and Total-P), and (3) SWI was a significant predictor of *ppk1*, *ppk2*, and *ppx* relative abundance. We applied the LMEM using the *nlme* package (Pinheiro et al. 2017). The general form of the model is given in Equation 1 as:

$$Y \sim \beta_0 + \beta_1 X + bZ + \varepsilon \qquad \text{(Equation 1)}$$

where Y represents the response variable, X represents the fixed predictor variable, Z represents the random effect variable, β_0 and, β_1 represent the intercept and slope, respectively, of the fixed predictor variable, b represents the mean zero Gaussian random effects parameter, and ε represents the mean zero Gaussian random error. In this study site ID was treated as a random effect. When necessary, we log transformed the response variable to ensure it was normally distributed. Tukey pairwise comparisons were made using the *multcomp* package (Hothorn et al. 2008). For all statistical tests, we used $\alpha = 0.05$ to determine statistical significance and compared the quality of LMEMs using the Akaike information criteria (AIC; Akaike 1974). A lower AIC value indicates a better model fit and AIC scores greater than two units apart are statistically different.

RESULTS & DISCUSSION

Soil Moisture and the Soil Wetness Index

When simply compared by location (i.e., NY versus PA), soil moisture measurements were not statistically significant ($p=0.061$; Figure 3.3A). However, aggregating soil moisture by location as well as sample month resulted in better predictability of soil moisture patterns; Model 3.3 has the lowest AIC score of all the models tested (Table 3.3) and model summary results show that PA trends are statistically different from NY when compared over time ($p=0.0073$; Table 3.4). PA soil moisture measurements were highest in July (Figure 3.3B) and are statistically different from measurements taken in May and October ($p<1.0E-5$). NY soil moisture trends over the course of this experiment show a wet May followed by two wet-up periods: one from June to July/August and a second from September to October (Figure 3.3B). Pairwise comparisons of NY data show that soil moisture measurements in May, June, and July were statistically the same, August measurements were most similar to June, September measurements were lower (i.e., dryer) than any other month, and October measurements were similar to June and August (Figure 3.3).

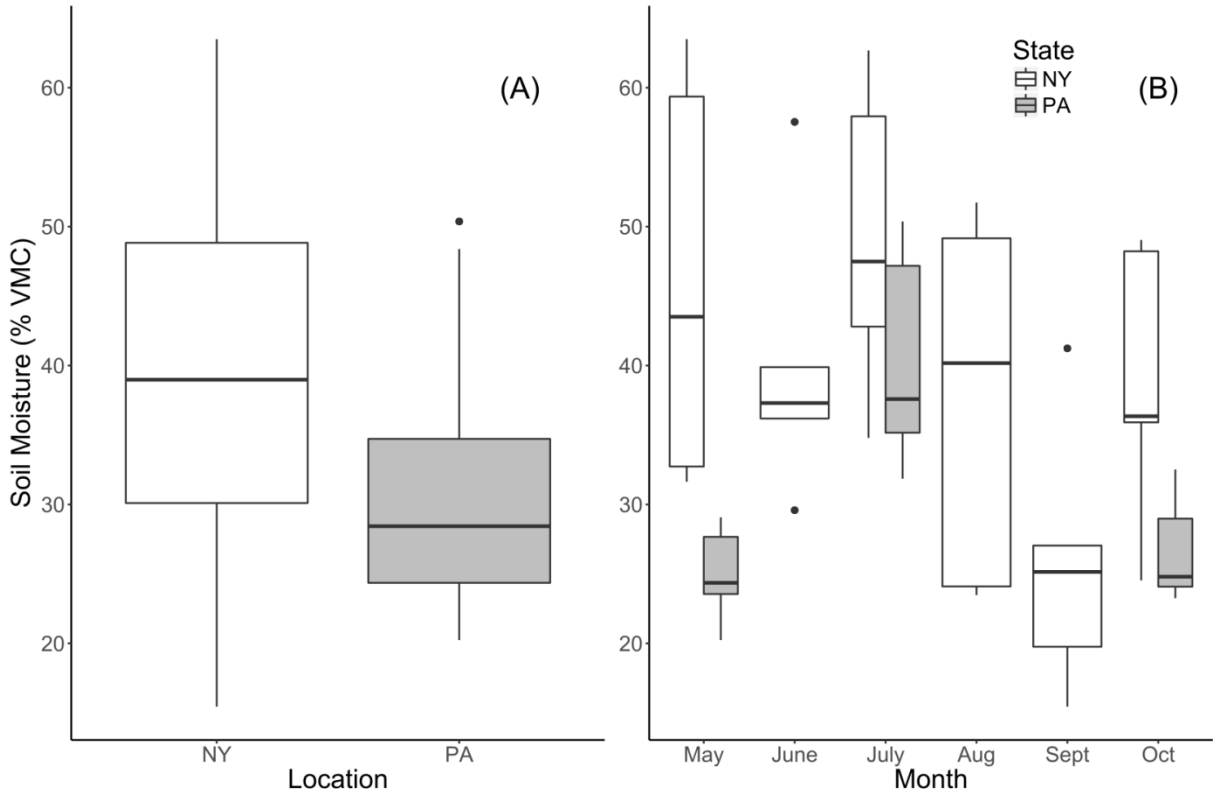

Figure 3. 3 (A) Soil moisture, measured as percent volumetric moisture content (% VMC), aggregated for all sample months versus location. (B) Soil moisture versus sample month for New York (NY) and Pennsylvania (PA).

Table 3. 3 Linear mixed effects model comparisons for soil moisture, measured as percent volumetric moisture content (% VMC). 'Location' indicates whether samples were collected in New York (NY) or Pennsylvania (PA) and 'Month' indicates the sample month. Best fit models are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC).

Model Number	Model	df	AIC
3.1	% VMC ~ Location	4	39.53
3.2	% VMC ~ Month	8	-70.08
3.3	**% VMC ~ Location + Month**	**9**	**-73.75**

Table 3. 4 Linear mixed effects model outputs for best fit model (Model 3.3 in Table 3.3) for percent volumetric moisture content (% VMC). 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df).

Model Term	Value	SE	df	t-value	p-value
Intercept*	39.51	0.09	94	43.1587	< 1.00E-05
Location (PA)	0.69	0.11	11	-3.2831	0.0073
Month (June)	1.04	0.05	94	0.72862	0.468
Month (July)	1.34	0.03	94	8.79711	< 1.00E-05
Month (August)	0.92	0.04	94	-1.845	0.068
Month (September)	0.62	0.04	94	-10.727	< 1.00E-05
Month (October)	0.96	0.03	94	-1.3184	0.1906

* Base model represents NY data in May.

SWI was a significant predictor of soil moisture in all the models tested (Table 3.5). However, the best fit model (Model 5.4 in Table 3.5) indicates the statistically significant importance of also including the location and month (Table 3.6). Comparison of observations to Model 5.4 results are shown in Figure 3.4 and indicate a positive relationship with a consistent slope. In general, this finding is consistent with previous research validating the predictability of soil moisture using SWIs (Agnew et al. 2006; Buchanan et al. 2014; Hofmeister et al. 2016). It also highlights that (1) larger SWI values correspond to wetter parts of the landscape and (2) the temporally static SWI integrates soil moisture patterns over time.

Table 3. 5 Linear mixed effects model comparisons soil moisture, measured as percent volumetric moisture content (% VMC), versus soil wetness index (SWI). 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA) and 'Month' indicates the sample month. Best fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC).

Model	Model	df	AIC
5.1	% VMC ~ SWI	4	39.103
5.2	% VMC ~ SWI + Location	5	35.275
5.3	% VMC ~ SWI + Month	9	-67.482
5.4	**% VMC ~ SWI + Location + Month**	**1**	**-78.543**
5.5	% VMC ~ SWI + Month + SWI:Month	1	-28.535
5.6	% VMC ~ SWI + Location + SWI:Location	6	42.086
5.7	% VMC ~ SWI + Location + Month + SWI:Month	1	-40.027
5.8	% VMC ~ SWI + Location + Month + SWI:Location	1	-71.838
5.9	% VMC ~ SWI + Location + Month + SWI:Month +	1	-32.752

Table 3. 6 Linear mixed effects model outputs for best fit model (Model 5.4 in Table 3.5) for percent volumetric moisture content (% VMC) versus soil wetness index (SWI). 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df).

Model Term	Value	SE	df	t-value	p-value
Intercept*	26.22	0.09	94	34.7432	< 1.00E-05
SWI	1.05	0.01	10	5.16515	0.0004
Location	0.69	0.06	10	-5.8615	0.0002
Month (June)	1.03	0.05	94	0.72124	0.4726
Month (July)	1.34	0.03	94	8.80565	< 1.00E-05
Month (August)	0.92	0.04	94	-1.8468	0.0679
Month (September)	0.62	0.04	94	-10.738	< 1.00E-05
Month (October)	0.96	0.03	94	-1.3197	0.190

* Base model represents NY data in May.

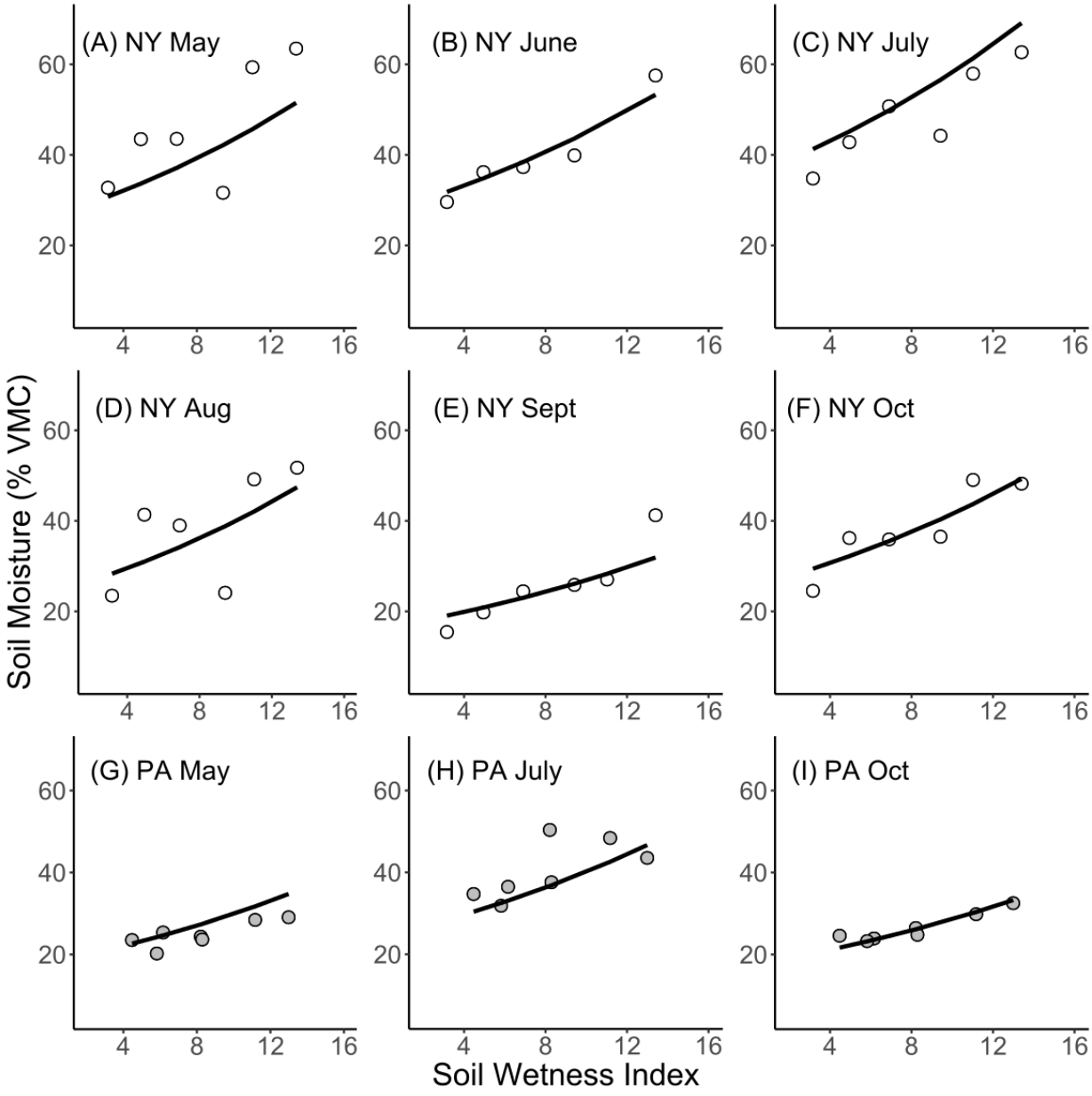

Figure 3. 4 Soil moisture, measured as percent volumetric moisture content (% VMC), versus soil wetness index for New York (NY) and Pennsylvania (PA) locations taken during each sample month. Best fit linear mixed effects model (LMEM) results (Model 5.4 in Table 3.5) shown as a black line.

These results highlight the application of SWI under very wet conditions as well as the limitations of using the SWI to predict soil moisture under dry conditions, as other studies have found (Western et al. 1999). Specifically, the April prior to the start of our study was very wet—58.1% wetter than the 30-year normal for April (Arguez et al. 2010; Menne et al. 2012). The R-squared value for data presented in Figure 3.4A for May was 0.55—indicating moderate variation in the relationship between soil moisture and SWI—but SWI was still capable of predicting the overall patterns. Variation in May soil moisture measurements from NY is also visible when looking at the spread of these same data in Figure 3.3B. When possible, we tried to avoid sampling until at least two days after a rainfall event to ensure that the water in the soil profile had a chance to redistribute. However, it is possible that soil water from rainfall events in April were still being redistributed throughout the landscape when May samples were collected in NY. With respect to dry conditions, the cumulative rainfall in August during the year of this study was 13.7% less than the 30-year normal; it is likely there was very little water in the soil profile throughout the landscape (Menne et al. 2012). While there was more variation in the relationship between soil moisture and SWI for August (R-squared value = 0.40) compared to May as mentioned above, the overall pattern of soil moisture was captured.

Soil Phosphorus and the Soil Wetness Index

Model comparisons show that aggregating $CaCl_2$-P based on location (i.e., NY versus PA) as well as location and sample month were both equality powerful; the AIC score of Model 7.1 is within two units of the AIC score of Model 7.3 (Table 3.7). However, because Model 7.1 is more parsimonious, we prefer it. Therefore, when aggregated by location, $CaCl_2$-P concentrations were significantly lower in NY compared to PA ($p=0.0012$; Table 3.8, Figure 3.5). Because the

AIC of Model 7.3 was within two units of Model 7.1, we also include $CaCl_2$-P results plotted by location as well as sample month in Figure 3.6. The smaller $CaCl_2$-P concentration in NY is evident but changes little over time (Figure 3.6B). Pairwise comparisons indicate significant differences between May and September, July and August, August and September, and September and October. The larger $CaCl_2$-P concentration in PA is evident and, similarly to NY, changes little over time (Figure 3.6B). Pairwise comparisons indicate that October $CaCl_2$-P concentrations measurements in PA are significantly different than May and July measurements. Model comparisons for Ox-P and Total-P also indicate that location is the most important predictor of these P pool trends (Table 3.7). Ox-P was statistically higher in NY compared to PA ($p=0.0013$; Table 3.8, Figure 3.5). Total-P was not statistically different between NY and PA ($p=0.5247$; Table 3.8, Figure 3.5).

Table 3. 7 Linear mixed effects model comparisons for soil extractions ($CaCl_2$-P, Ox-P, and Total-P). 'Location' indicates whether samples were collected in New York (NY) or Pennsylvania (PA) and 'Month' indicates the sample month. Best fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC).

Model Number	Model	df	AIC
$CaCl_2$-P Models			
7.1	**$CaCl_2$-P ~ Location***	**4**	**43.40**
7.2	$CaCl_2$-P ~ Month	8	50.81
7.3	**$CaCl_3$-P ~ Location + Month**	**9**	**42.45**
Ox-P Models			
7.4	**Ox-P ~ Location**	**4**	**-96.38**
7.5	Ox-P ~ Month	8	-78.47
7.6	Ox-P ~ Location + Month	9	-81.63
Total-P Models			
7.7	**Total-P ~ Location**	**4**	**-91.57**
7.8	Total-P ~ Month	8	-76.39
7.9	Total-P ~ Location + Month	9	-72.23

Table 3. 8 Best fit linear mixed effects model (LMEM) outputs (Models 7.1, 7.3, 7.4, and 7.7 in Table 3.7) for soil extractions ($CaCl_2$-P, Ox-P, and Total-P). 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df).

Model Term	Value	SE	df	t-value	p-value
$CaCl_2$-P ~ Location					
Intercept	1.66	0.16	99	3.22	0.0017
Location (PA)	2.54	0.22	11	4.35	0.0012
*$CaCl_2$-P ~ Location + Month**					
Intercept	1.62	0.16	94	2.97	0.004
Location (PA)	2.53	0.22	11	4.29	0.0013
Month (July)	1.02	0.08	94	0.23	0.8151
Month (June)	0.95	0.06	94	-0.83	0.4104
Month (May)	1.25	0.08	94	2.88	0.0049
Month (October)	0.83	0.08	94	-2.41	0.0179
Month (September)	1.14	0.06	94	2.24	0.0277
Ox-P ~ Location					
Intercept	149.61	0.13	99	37.67	< 1.00E-05
Location (PA)	0.61	0.18	11	-2.73	0.0195
Total-P ~ Location					
Intercept	304.90	0.01	99	66.77	< 1.00E-05
Location (PA)	0.93	0.12	11	-0.66	0.5247

*Note the $CaCl_2$-P ~ Location model is preferable because it is simpler and within two AIC units (Table 3.7).

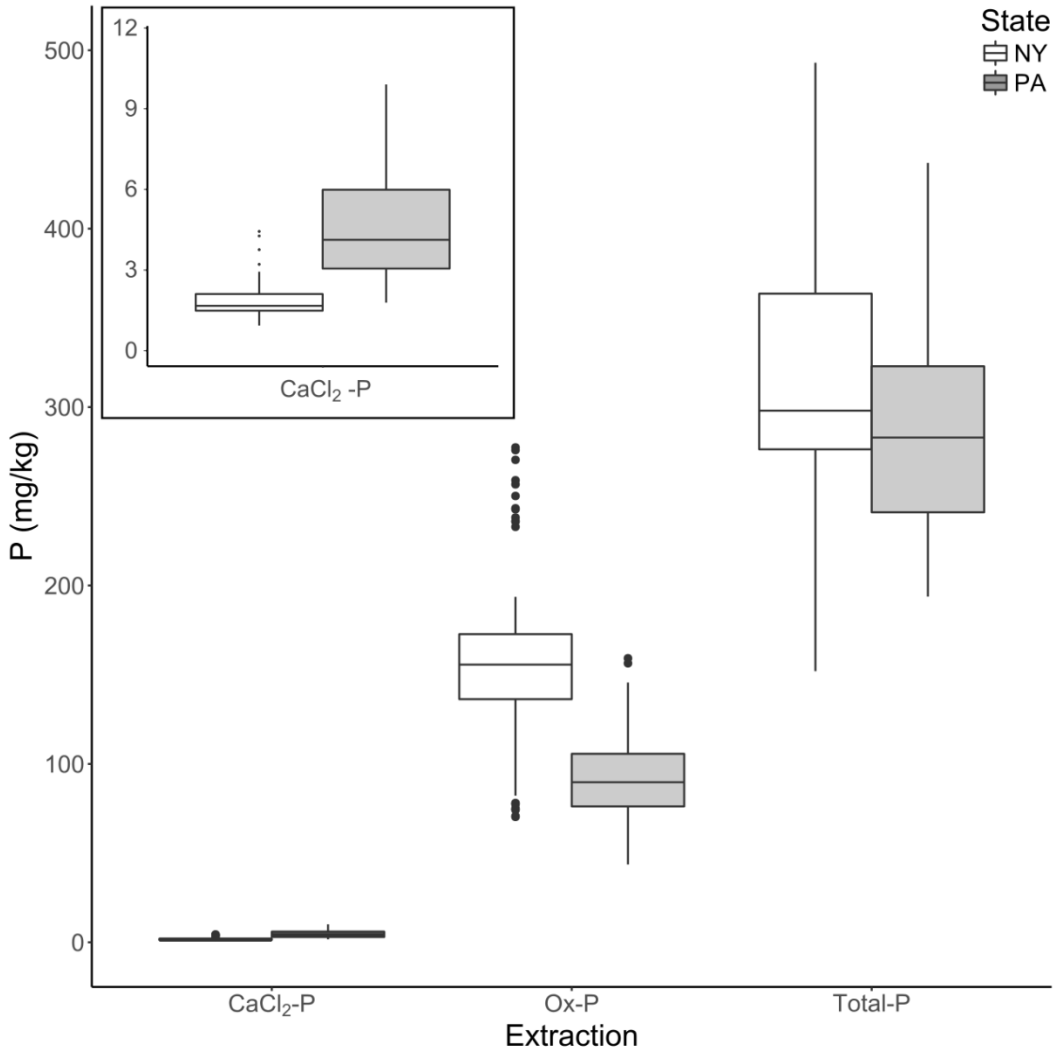

Figure 3. 5 Phosphorus concentration (mg/kg) versus extraction type for New York (NY) and Pennsylvania (PA). In-set (top left-hand corner) shows a zoomed-in view of $CaCl_2$-P extraction results.

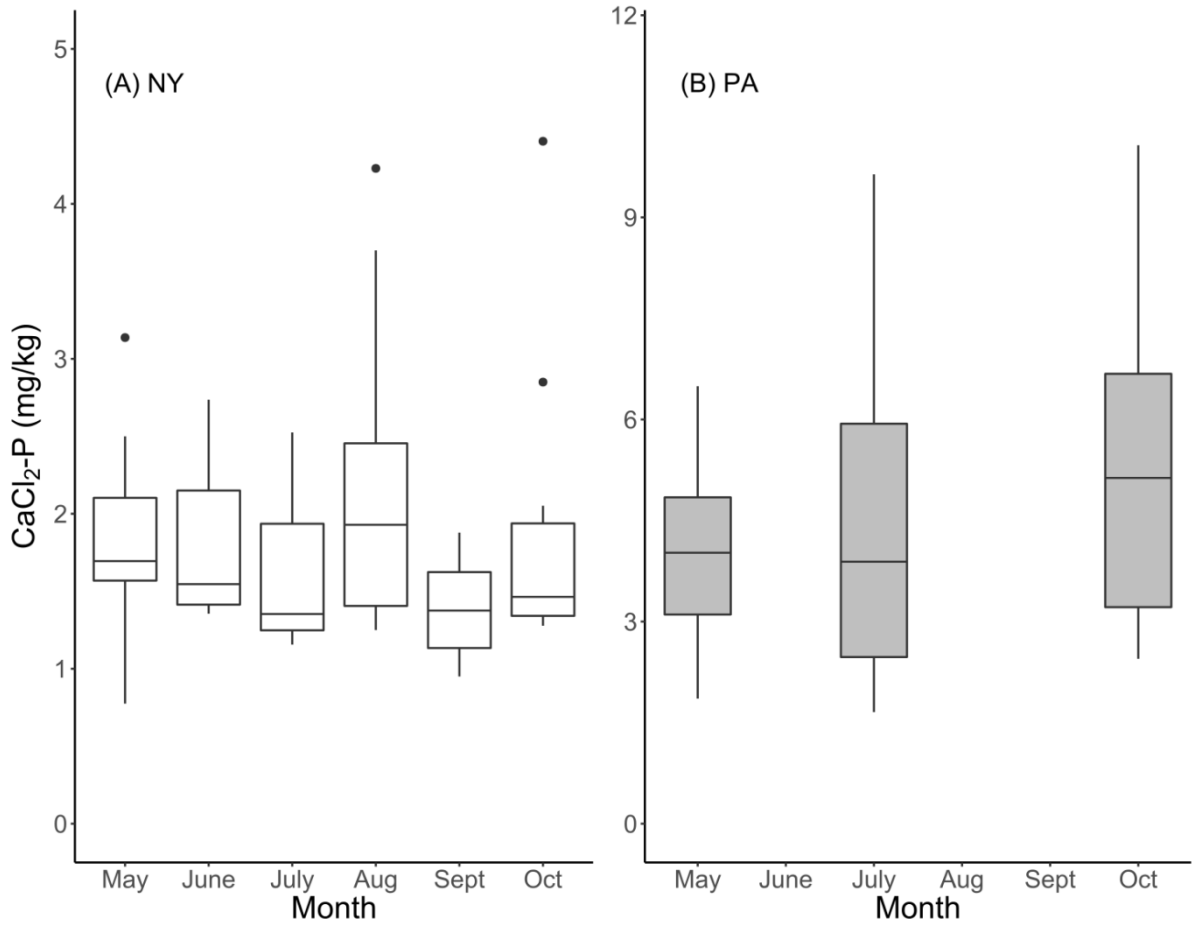

Figure 3. 6 $CaCl_2$-P concentration (mg/kg) versus sample month for (A) New York (NY) and (B) Pennsylvania (PA). Note difference in y-axis scale.

We tested several different models for each P extraction and found that both SWI and location (i.e., NY versus PA) were important predictors of $CaCl_2$-P and resulted in the model with the lowest AIC score (Model 9.2, Table 3.9). Model 9.4, which included SWI, location, as well as sample month, also had a low AIC score that was within two units of Model 9.2. However, because Model 9.2 is more parsimonious it is preferred and is used for additional analysis (i.e., Table 3.10 and Figure 3.7). Model 9.2 summaries show that SWI is a significant predictor of $CaCl_2$-P ($p=0.0349$; Table 3.10). Comparisons between Model 9.2 and observations are shown in

Figures 3.7A and 3.7B and indicate a negative slope between $CaCl_2$-P and SWI; the slope being more negative in PA compared to NY. This finding is consistent with our hypothesis that $CaCl_2$-P is more likely to be flushed out of areas with a higher SWI because these areas are more likely to generate runoff that transports mobile P forms further downslope. As noted earlier, PA sites tend to have larger $CaCl_2$-P concentrations compared to NY; thus, this larger supply may provide a reason for the difference in slope between PA and NY.

Table 3. 9 Linear mixed effects model comparisons for soil extractions ($CaCl_2$-P, Ox-P, and Total-P) versus soil wetness index (SWI). 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA) and 'Month' indicates the sample month. Best fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC).

Model Number	Model	df	AIC
$CaCl_2$-P Models			
9.1	$CaCl_2$-P ~ SWI	4	56.43
9.2	**$CaCl_2$-P ~ SWI + Location***	**5**	**45.60**
9.3	$CaCl_2$-P ~ SWI + Month	9	55.14
9.4	**$CaCl_2$-P ~ SWI + Location + Month**	**10**	**44.43**
9.5	$CaCl_2$-P ~ SWI + Month + SWI:Month	6	49.55
9.6	$CaCl_2$-P ~ SWI + Location + SWI:Location	14	83.63
9.7	$CaCl_2$-P ~ SWI + Location + Month + SWI:Month	11	48.39
9.8	$CaCl_2$-P ~ SWI + Location + Month + SWI:Location	11	48.39
Ox-P Models			
9.90	Ox-P ~ SWI	4	-87.07
9.10	**Ox-P ~ SWI + Location**	**5**	**-89.31**
9.11	Ox-P ~ SWI + Month	9	-71.82
9.12	Ox-P ~ SWI + Location + Month	10	-74.56
9.13	Ox-P ~ SWI + Month + SWI:Month	6	-36.60
9.14	Ox-P ~ SWI + Location + SWI:Location	14	-84.47
9.15	Ox-P ~ SWI + Location + Month + SWI:Month	11	-39.39
9.16	Ox-P ~ SWI + Location + Month + SWI:Location	11	-69.73
Total-P Models			
9.17	**Total-P ~ SWI**	**4**	**-87.46**
9.18	Total-P ~ SWI + Location	5	-83.49
9.19	Total-P ~ SWI + Month	9	-68.24
9.20	Total-P ~ SWI + Location + Month	10	-64.14
9.21	Total-P ~ SWI + Month + SWI:Month	6	-26.5
9.22	Total-P ~ SWI + Location + SWI:Location	14	-79.25
9.23	Total-P ~ SWI + Location + Month + SWI:Month	11	-22.38
9.24	Total-P ~ SWI + Location + Month + SWI:Location	11	-59.98

*AIC units for models 2 and 4 were not statistically significant (greater than two AIC units apart) so we select the simpler model for subsequent analyses.

Table 3.10 Best fit linear mixed effects model (LMEM) outputs (Models 9.2, 9.4, 9.10, and 9.17 in Table 3.9) for soil extractions ($CaCl_2$-P, Ox-P, and Total-P) versus soil wetness index (SWI). 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA). We ln-transformed the Value column has but kept SE in untransformed units. Abbreviations: standard error (SE), degrees of freedom (df).

Model Term	Value	SE	df	t-value	p-value
$CaCl_2$-P ~ SWI + Location					
Intercept	2.89	0.26	99	4.04	0.0001
SWI	0.93	0.03	10	-2.44	0.0349
Location (PA)	2.54	0.18	10	5.23	0.0004
*$CaCl_2$-P ~ SWI + Location + Month**					
Intercept	3.54	0.27	94	4.69	< 1.00E-05
SWI	0.93	0.03	10	-2.44	0.0347
Location (PA)	2.53	0.18	10	5.16	0.0004
Month (July)	0.76	0.08	94	-3.50	0.0007
Month (June)	0.82	0.09	94	-2.26	0.0259
Month (May)	0.80	0.08	94	-2.88	0.0049
Month (October)	0.91	0.08	94	-1.21	0.2306
Month (September)	0.67	0.08	94	-4.8	< 1.00E-05
Ox-P ~ SWI + Location					
Intercept	136.87	0.28	99	17.39	< 1.00E-05
SWI	1.01	0.03	10	0.36	0.7261
Location (PA)	0.61	0.19	10	-2.62	0.0255
Total-P ~ SWI					
Intercept	290.03	0.17	99	34.32	< 1.00E-05
SWI	1.00	0.02	11	0.06	0.9564

*Note that the *$CaCl_2$*-P ~ SWI + Location model is preferable because it is simpler and within two AIC units (Table 3.9).

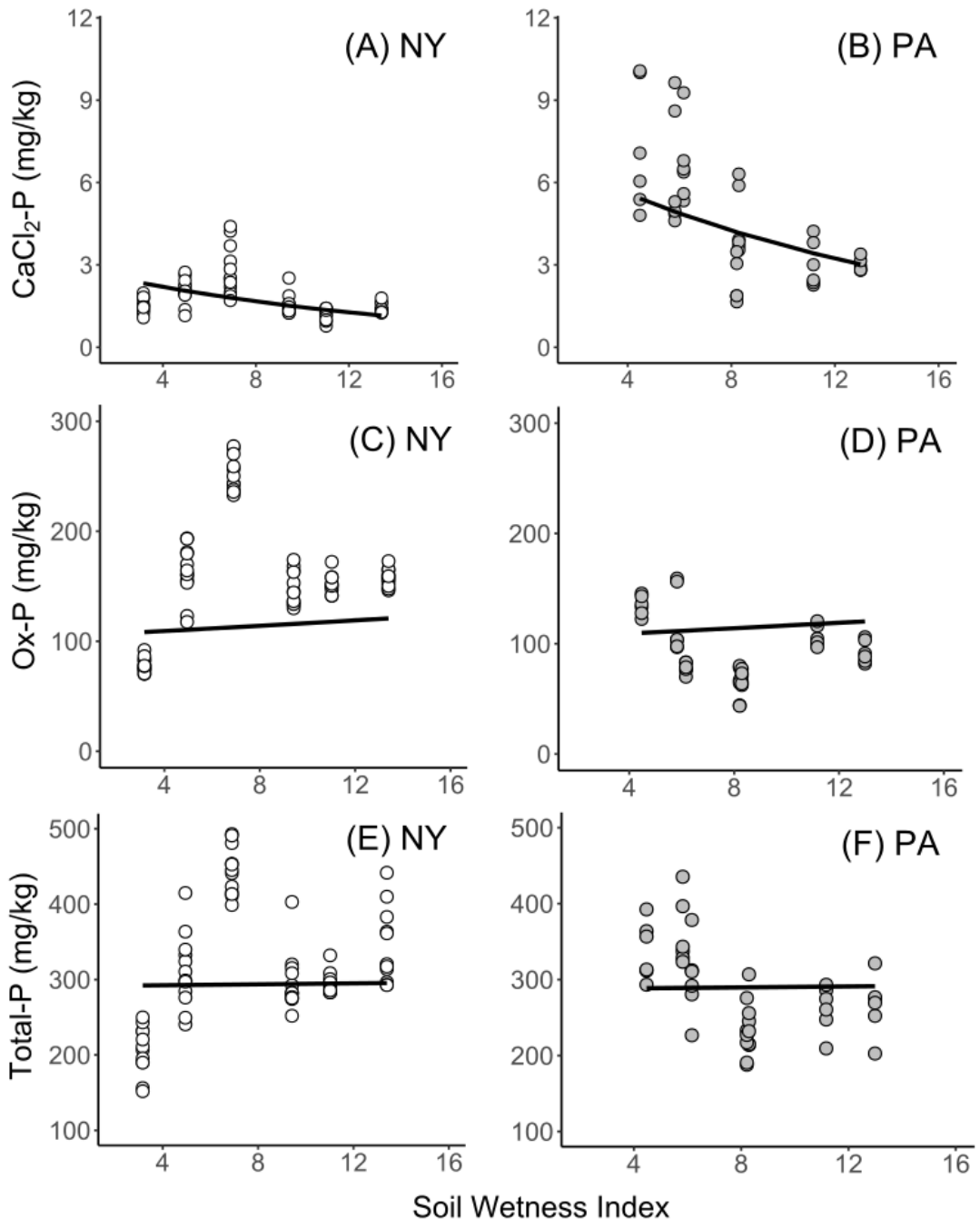

Figure 3. 7 Phosphorus concentration of $CaCl_2$-P (A and B), Ox-P (C and D), Total-P (E and F) versus soil wetness index for New York (NY; A, C, E) and Pennsylvania (PA; B, D, F). Best fit linear mixed effects model (LMEM) results (Model 9.2, Table 3.9) shown as a black line.

Our results are consistent with another study finding higher concentrations of water extractable P in upslope areas compared to wetter toeslopes (Macrae et al. 2005). This same study attributes this pattern to groundwater induced desorption processes in the toeslopes. However, our results are inconsistent with studies demonstrating higher concentrations of mobile P pools in wetter parts of the landscape (Smeck 1973; Honeycutt et al. 1990; Litaor et al. 2005; Flores-López et al. 2013; Wilson et al. 2016) likely due to the impact of sediment transport processes. We suggest these differences in finding are largely due to local landscape properties (e.g., soil and slope), land management (e.g., tillage intensity), and climate (e.g., storm size and frequency).

We also tested whether SWI could be used to predict Ox-P and Total-P pools. Comparisons of these various models demonstrate that SWI and location were the best predictors of Ox-P (Model 9.10, Table 3.9) and SWI alone is the best predictor of Total-P (Model 9.17, Table 3.9). However, closer inspection of these models indicates that SWI is not a statistically significant term for either of these P extractions. Specifically, the SWI term has a p-value equal to 0.7261 and 0.9564 for the Ox-P and Total-P models, respectively (Table 3.10). This lack of relationship between Ox-P versus SWI and Total-P versus SWI is displayed in the relatively flat slope of Figures 3.7C – 3.7F. Given that these two extractions are meant to characterize less mobile P pools, these results are not surprising; hydrological processes do not have as strong an impact on them unless the materials (e.g. organic matter or minerals) these P forms are adsorbed to are transported downslope. Especially evident in NY, there is a peak in Ox-P and Total-P concentrations at an approximate SWI value of 7 (Figures 3.7C and 3.7E) that is attributed to sample site NY25 (Table 3.1). This result was unexpected and likely due to unknown P sources.

As mentioned previously, Total-P pools were not different between PA and NY yet there were significant differences in $CaCl_2$-P and Ox-P pools between PA and NY. These findings point to either differences in (1) environmental/soil properties or (2) land management between the two locations. While dominant soils at both sites are classified as channery/gravelly silt loams, they may differ in their amount of Fe-/Al-oxides which provide a means to adsorb mobile forms of P (like $CaCl_2$-P) and contribute to higher Ox-P concentrations (McKeague and Day 1966; Kleinman and Sharpley 2002). In this study, Ox-Al and Ox-Fe concentrations were significantly higher in NY compared to PA ($p=0.0002$ and $p=0.0001$ for Ox-Al and Ox-Fe, respectively, data not shown). Thus, more Fe-/Al-oxide surfaces were available in NY to bind mobile forms of P. Nearby land management may also have contributed to some of the patterns we observed. In PA, chemical fertilizer and manure was applied upslope of the sample sites included in this study. This source of P may explain why $CaCl_2$-P concentrations are higher in PA (Figure 3.5) and there is a more negative slope between $CaCl_2$-P and SWI (Figure 3.7B) but is counter to previous research showing that topographic indexes are capable of more accurately predicting soil P under less intensive management (i.e., low nutrient input and no-tillage; Wilson et al. 2016). The higher concentrations of Total-P at the lower SWI range (Figure 3.7F) in PA also provides evidence for our assertion that sample sites in the upslope parts of the landscape, closest to the fertilized fields, may have a higher P concentrations compared to samples taken downslope. However, making this conclusion necessitates the assumption that NY sites have a history of P application prior to their status as pasture/hay (i.e., before available records end in 2003) and that historically applied P has since become bound up in the soil.

Our hypothesis assumed an inverse linear relationship between different P pools and SWI. However, our results demonstrated a non-linear (i.e., parabolic) pattern with peak P concentrations around SWI=7 (Figures 3.7A, 3.7C, and 3.7E) and trough P concentrations around SWI=8 (Figures 3.7D and 3.7F) for NY and PA, respectively. Both occur around moderate SWI values and suggest potential limitations when using the SWI to predict soil P concentrations. By definition, sites with low SWI values promote the drainage of soil water, sites with high SWI collect water from upslope areas, and sites with moderate SWI values are a combination of these two ends of the spectrum. Sites with moderate SWI values have the capacity to drain and collect water. This inherent landscape classification when using the SWI combined with its temporally static nature, appeared to limit our ability to accurately predict mid-SWI range dynamic P cycling processes in NY and PA.

Shotgun Metagenomics Assembly

For NY, we assembled a total of 5,924,925 contigs, the longest of which was 184.8 kb. There were 67 contigs greater than 50 kb in NY and they made up 0.13% of the main genome. For PA, we assembled a total of 2,398,609 contigs, the longest of which was 98.6 kb. There were 9 contigs greater than 50 kb in NY and they made up 0.04% of the main genome. A summary of the total assembly length for contigs above a certain length is shown in Table 3.11 for NY and PA samples.

Table 3. 1 1 Shotgun metagenome assembly results for samples taken in New York (NY) and Pennsylvania (PA).

Location	Minimum Contig Length	Number of Contigs
NY		
	500	2,288,892
	1,000	563,371
	10,000	2,531
PA		
	500	905,477
	1,000	186,473
	10,000	968

Microorganisms Harboring Phosphorus Functional Genes

We annotated a total of 6,493 and 2, 701 genes for NY and PA, respectively (Table 3.12). Because some contigs had more than one gene, they represented a total of 6,419 and 2,657 unique contigs for NY and PA, respectively (summation of all bars by location, Figure 3.8). As mentioned previously, we assembled sequencing data from NY and PA separately and the number of samples sequenced for NY was double the number in PA; both of these points likely explain this difference as well as the doubling effect when comparing Figure 3.8A to Figure 3.8B. However, even when taking the doubling effect into account, NY has more unique contigs (Figure 3.8).

Table 3. 1 2 Number of *ppk1*, *ppk2*, and *ppx* genes annotated in this study for soil samples taken from locations in New York (NY) and Pennsylvania (PA).

Location	Gene	Count
NY		
	ppk1	2,635
	ppk2	2,449
	ppx	1,409
	Total	6,493
PA		
	ppk1	1,252
	ppk2	847
	ppx	602
	Total	2,701

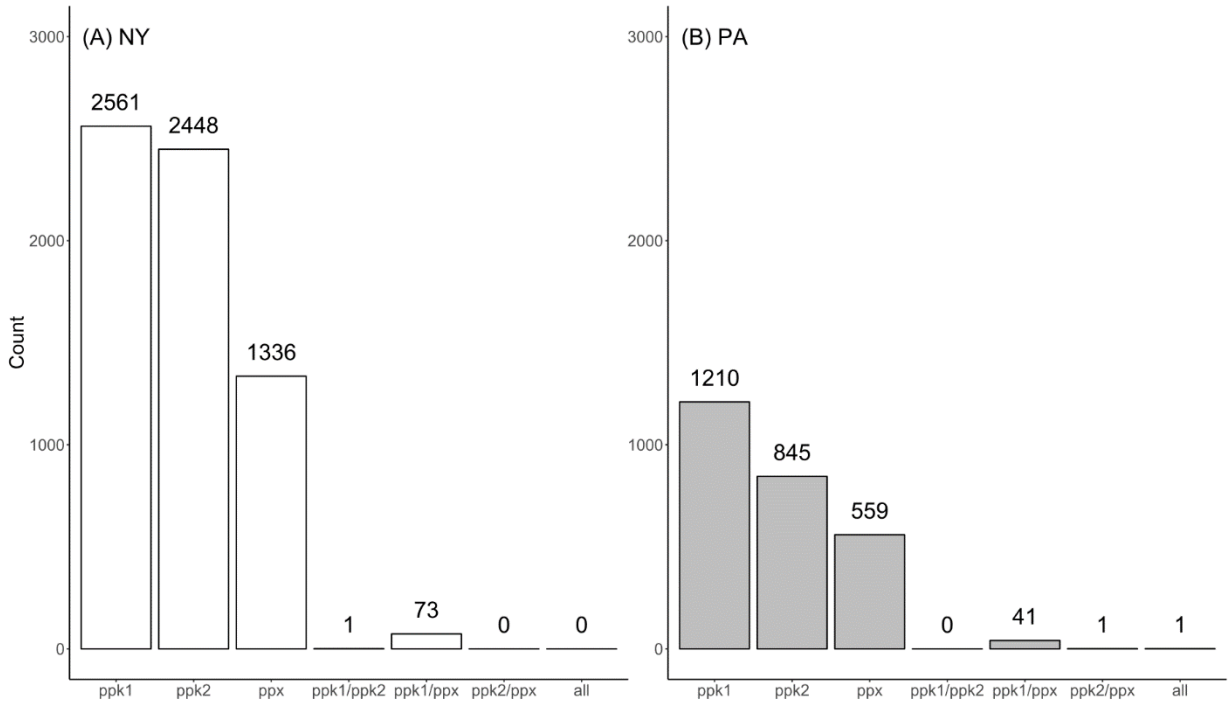

Figure 3. 8 Number of unique contigs with various gene combinations in samples from (A) New York (NY) and (B) Pennsylvania (PA). Note: the total number of samples sequenced in NY (n=30) was double the total number of samples sequenced in PA (n=15).

For both locations, unique contigs only harboring *ppk1* were most frequent followed by contigs harboring only *ppk2* and only *ppx* (Figure 3.8). The presence of contigs harboring only one gene was unexpected given that microorganisms would theoretically require either *ppk1* or *ppk2* to store intracellular polyP and *ppx* to degrade polyP. The length distribution of contigs harboring only one polyP functional gene did not appear to be especially shorter than any of the other gene combinations for NY or PA (Figure 3.9) but we cannot know for sure since the genomes for these contigs are not complete. When combined these findings indicate that either deep sequencing is required to annotate both *ppk* and *ppx* or *ppk* may have a genetic function independent of *ppx*. We suspect *ppx* genes were less frequently annotated than *ppk* genes due to either incomplete genomes, a limited focus on *ppx* in the literature/existing protein databases, or microorganism having fewer copies of *ppx* genes compared to *ppk* genes. It is understood that CAP has one copy of *ppk1* (Kunin et al. 2008) but it is possible that non-CAP PAOs have more than one copy of this gene. Only rarely did we observe contigs with both *ppk1* and *ppk2* genes as well as all three genes (Figure 3.8; Table 3.13).

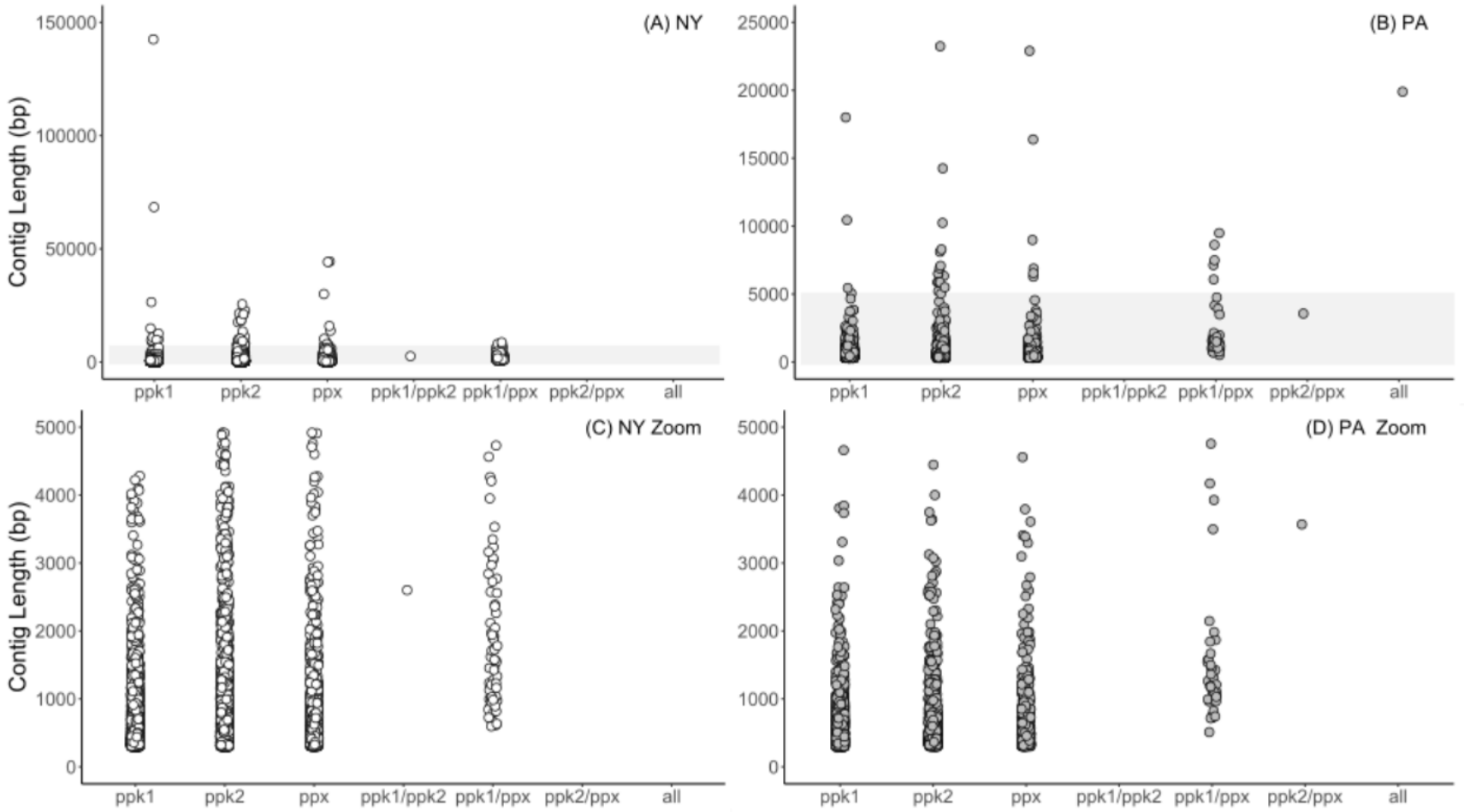

Figure 3. 9 Length distributions of contigs harboring various gene combinations from samples taken in New York (NY; A and C) and Pennsylvania (PA; B and D). Note the difference in scale between A and B. The shaded regions in plots A and B refer to the zoomed-in (0-5,000 bp) plots shown in C and D.

Table 3.13 Contigs with unique combinations of *ppk* and *ppx* genes as determined by shotgun metagenome analysis and subsequent comparison of protein sequence to existing sequences in the pBLAST database.

Contig ID	Gen	Location	Top pBLAST Hit	Top pBLAST Hit Organism	Accession Number	Identity	Environmen	Reference
k141_1087762								
	ppk1	NY	hypothetical protein A2V87_04225	*Gammaproteobacteria* bacterium RBG_16_58_17	MGQV01000063.1	85	Aquifer	Anantharaman et al. 2016
	ppk2	NY	hypothetical protein DUF344	*Chthoniobacter flavus* Ellin428 (*Verrucomicrobia*)	ABVL01000022.1	79	Soil	Kant et al. 2011
k141_1683664								
	ppk1	PA	polyphosphate kinase	*Betaproteobacteria* bacterium SG8_40	LJTS01000251.1	73	Estuary sediment	Baker et al. 2015
	ppk2	PA	Polyphosphate AMP	*Syntrophaceae* bacterium CG2_30_58_14	MNZQ01000047.1	66	Aquifer	Probst et al. 2017
	ppx	PA	exopolyphosphatase	*Betaproteobacteria* bacterium RIFCSPLOWO2_2_FULL_62_13	MERG01000201.1	69	Aquifer	Anantharaman et al. 2016
k141_975119								
	ppk2	PA	phosphate--nucleotide	*Alphaproteobacteria* bacterium 64-11	MKRH01000042.1	73	Bioreactor	Kant et al. (unpublished)
	ppx	PA	hypothetical protein AUH79_01695	*Betaproteobacteria* bacterium 13_1_40CM_4_64_4	MNFQ01000026.1	75	Soil	Butterfield et al. 2016

Regardless of whether samples were collected in NY or PA, contigs classified as putative PAOs (i.e., harboring either *ppk1* and *ppx* or *ppk2* and *ppx*) more frequently harbored a combination of *ppk1* and *ppx* genes compared to a combination of *ppk2* and *ppx* genes (Figure 3.8). The length distribution of putative PAOs does not indicate a bias toward longer contigs. This is true with the exception for a contig harboring all three polyP functional genes in PA, which was ~20 kb long (Figure 3.9B; Table 3.13). Besides assembly bias, it is possible that *ppk* does not always reside near *ppx* in a particular genome. While both organisms are not recognized as PAOs, *ppx* was found downstream of *ppk* in the *E. coli* genome—forming the *ppk* operon—but upstream of *ppk* in the *P. aeruginosa* genome (Brown and Kornberg 2008). The polyP functional gene arrangement in *P. aeruginosa* does not form an operon but co-regulation has not been ruled out (Brown and Kornberg 2008). Given our limited understanding and subsequent ability to resolve these issues, we suggest that further research is needed to (1) determine whether deeper sequencing influences the number of putative PAOs identified in soil samples, (2) improve our understanding of the organization of these genes in the genomes of different organisms, and (3) verify the phenotype of putative PAOs identified here.

In terms of overall taxonomy, the distribution of contigs harboring polyP functional genes from both locations was similar (Figure 3.10). Specifically, *Proteobacteria* (i.e., *Alphaproteobacteria)* and *Actinobacteria* were the top two most represented bacterial phyla in NY and PA (Figure 3.10). We also observed many contig assignments to the *Betaproteobacteria* class, which includes CAP—a model PAO (Hesselmann et al. 1999). Besides the top three bacterial phyla, taxonomic classifications of contigs vary slightly more for contigs harboring only one of either *ppk1*, *ppk2*, or *ppx* (Figure 3.11). For example, the *Cyanobacteria/Melainabacteria* group, which

was previously identified in full-scale EBPR reactor (Soo et al. 2014), shifts up and down in order of most abundant depending on the polyP functional gene of interest. There are generally more unclassified contigs for both NY and PA *ppx* (Figures 3.10F and 3.11C), which is not surprising given our more limited understanding of this gene compared to *ppk*. Many contigs in NY and PA were assigned to the phyla *Verrucomicrobia*, which is abundant in soil (Bergmann et al. 2011) and positively correlated with total P concentrations in sediments (Jin et al. 2017). Some contigs were also assigned to archaeal lineages (*Methanomicrobia*).

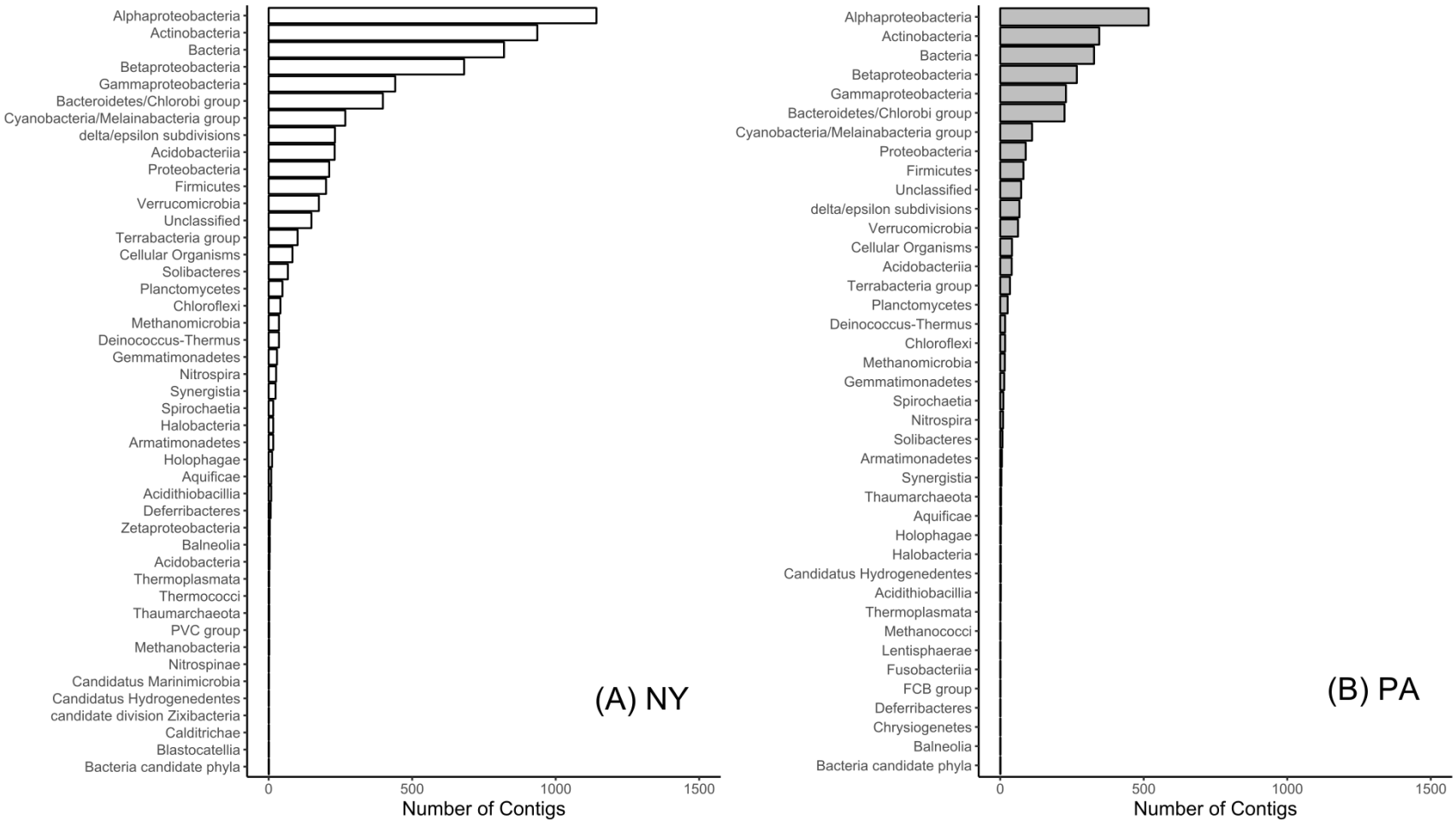

Figure 3. 1 0 Number of unique contigs classified into major bacterial and archaeal phyla/classes for (A) New York (NY) and (B) Pennsylvania (PA).

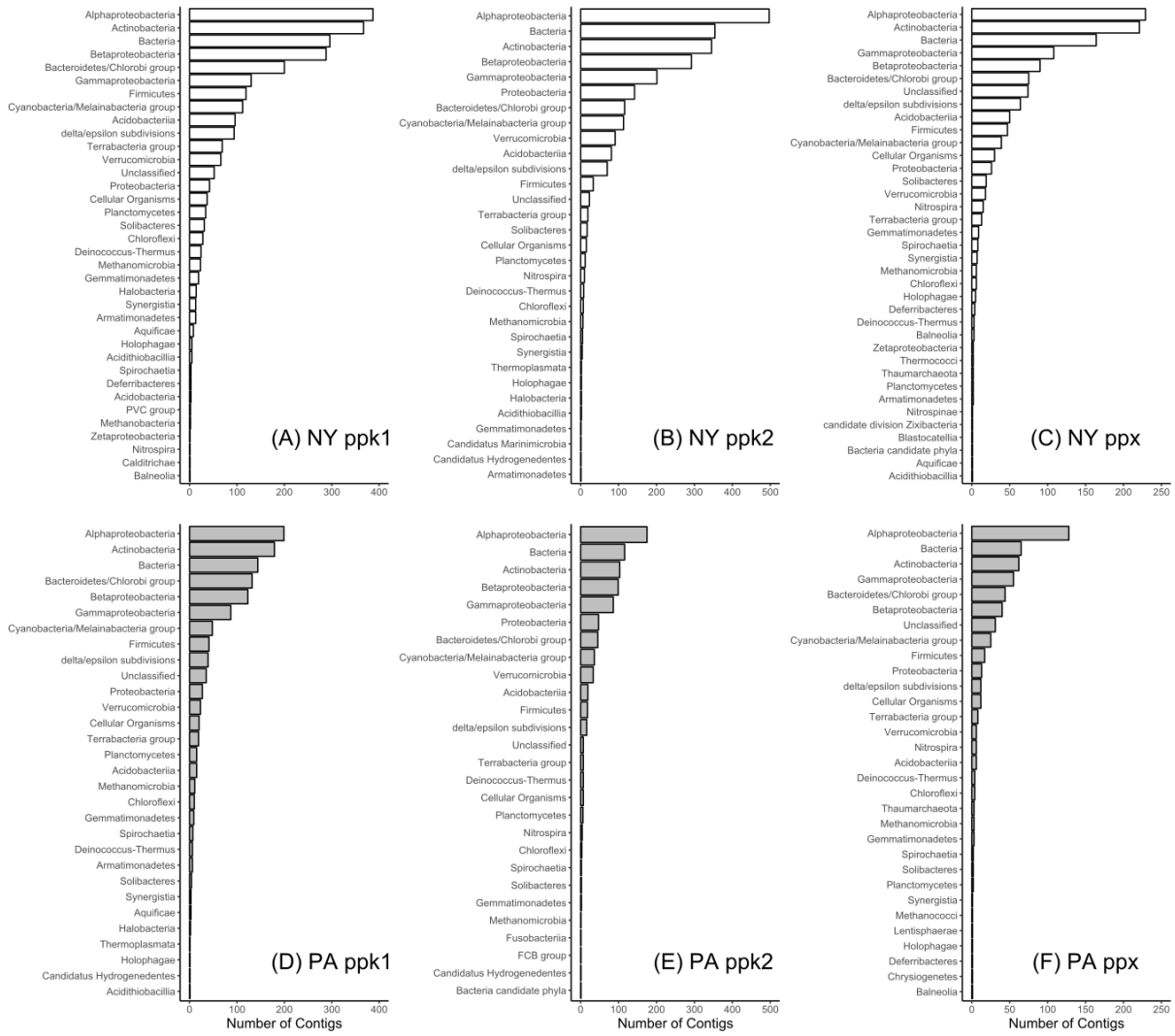

Figure 3. 1 1 Number of unique contigs classified into major bacterial and archaeal phyla/classes for New York (NY; A-C) and Pennsylvania (PA; D-E) harboring only one of either *ppk1* (A and D), *ppk2* (B and E), and *ppx* (C and F) genes.

Contigs that harbored both *ppk1* and *ppx* genes—and may represent putative PAOs—were classified as members of the *Proteobacteria* (i.e., *Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria*) phylum (Figure 3.12). Given the caveats discussed earlier associated with contig length and ability to predict putative PAOs, these findings are notable because CAP is a member of the *Betaproteobacteria* and other studies of WWTP systems have identified PAOs in the *Alphaproteobacteria* and *Gammaproteobacteria* classes (e.g., Zilles et al. 2002; Nguyen et al. 2012). We recognize these are diverse classes and suggest the need for further research to verify whether the putative PAOs identified here fit the phenotype of known PAOs. Many of the *Betaproteobacteria* with both *ppk1* and *ppx* genes were assigned to the genus *Burkholderia*, which has been found to store intracellular polyP granules and solubilize P in the environment (Locke 2015; Purahong et al. 2016). Only one contig in PA (k141_975119, Table 3.13) harbored both *ppk2* and *ppx* genes. It also appears as though it is less common to have *ppk2* and *ppx* genes than *ppk1* and *ppx* genes. We found no contigs containing all three polyP functional genes in NY but one in PA (k141_1683664, Table 3.13) that had PPK1 and PPX protein sequences that were similar to proteins belonging to *Betaproteobacteria* found in estuary sediment and aquifer environments (Baker et al. 2015; Anantharaman et al. 2016). This same contig had a PPK2 protein sequence that was related to a *Syntrophaceae* bacterium protein found in an aquifer environment (Probst et al. 2017). The length of this contig was large compared to other contigs (~20 kb; Figure 3.7B); therefore, it is possible that deeper sequencing may prove beneficial for identifying organisms with similar combinations of polyP functional genes. We found one contig in NY (k141_1087762, Table 3.13) with both *ppk1* and *ppk2* genes and none in PA. We note that the protein sequences of contigs that we identified with unique combinations of polyP functional genes were only recently characterized in the environment; that is, from 2011 to

2017 (Table 3.13). Therefore, these results indicate the potential for many newly discovered organisms with the capacity to influence P cycling in diverse environments from soils, to estuary sediments, to aquifers (and perhaps beyond).

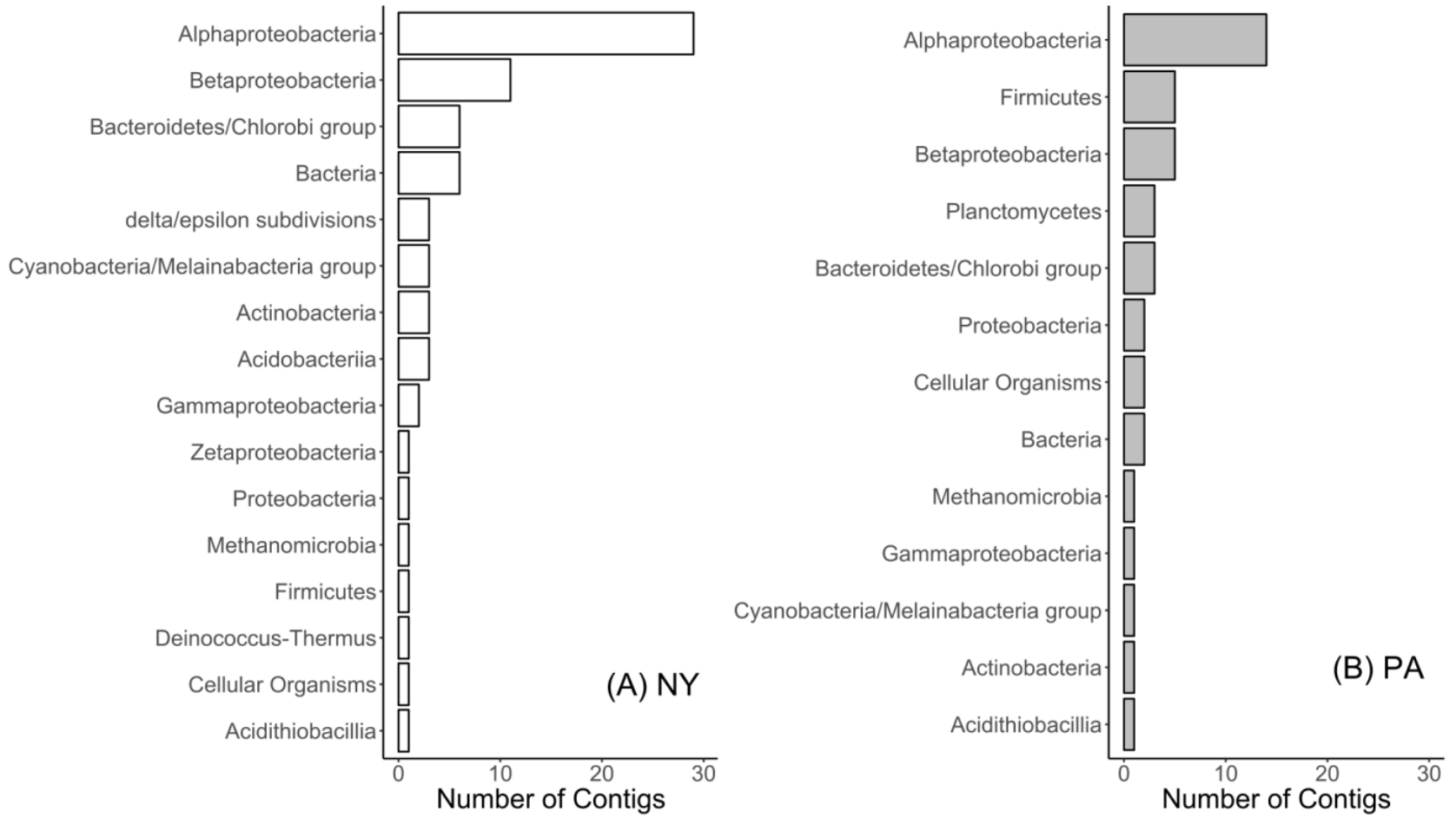

Figure 3. 1 2 Number of unique contigs harboring both *ppk1* and *ppx* genes classified into major bacterial and archaeal phyla/classes for (A) New York (NY) and (B) Pennsylvania (PA).

Hydrology, and Phosphorus Pools, and Functional Gene Relative Abundance

LMEMs relating relative abundance of *ppk1*, *ppk2*, and *ppx* to SWI, location, and the three P pools are shown in Table 3.14. Looking first at SWI, LMEM AIC summaries indicate that SWI alone had a relatively low predictive power compared to other models and was not a significant predictor of relative abundance for any of the polyP functional genes included in this study ($p>0.05$, Table 3.15). We found no observable trend when plotting relative abundance of the three polyP functional genes versus sample event ordered from lowest to highest SWI (Figure 3.13) or versus SWI when repeated measurements in time were plotted together (Figure 3.14). These results do not support our hypothesis that the relative abundance of polyP functional genes will increase for regions of the landscape with larger SWIs. They may indicate that (1) DNA is a poor indicator of PAO-mediated P cycling along the SWI gradients studied here, (2) microbially-mediated P processes may be masked by land management or other controls on P, (3) SWI characterizes soil moisture patterns at a time scale that is too large for it to be applicable to finer scale, microbial processes, or (4) *ppk1*, *ppk2*, and *ppx* are uniformly spread out in time and space (e.g., Figure 3.13F). Researchers have noted the importance of measuring environmental factors at timescales that are meaningful to biological processes (Pellerin et al. 2016; Rode et al. 2016).

Table 3.14 Relative gene abundance (RA; in units of FPKM) linear mixed effects model (LMEM) comparisons versus variables of interest for *ppk1*, *ppk2*, and *ppx*. 'Location' indicates whether samples were collected in New York (NY) or Pennsylvania (PA). Best fit models for each soil extraction are bolded. Abbreviations: soil wetness index (SWI), degrees of freedom (df), Aikaike information criteria (AIC).

Model Number	Model	df	AIC
ppk1			
11.1	RA ~ SWI	4	552.35
11.2	RA ~ Location	4	539.54
11.3	RA ~ $CaCl_2$-P	4	544.08
11.4	RA ~ $CaCl_2$-P + Location	5	532.55
11.5	**RA ~ $CaCl_2$-P + Location + $CaCl_2$-P:Location**	**6**	**518.50**
11.6	RA ~ Ox-P	4	555.01
11.7	RA ~ Ox-P + Location	5	540.78
11.8	RA ~ Ox-P + Location + Ox-P:Location	6	540.00
11.9	RA ~ Total-P	4	558.91
11.10	RA ~ Total-P + Location	5	542.05
11.11	RA ~ Total-P + Location + Total-P:Location	6	538.97
ppk2			
11.12	RA ~ SWI	4	528.77
11.13	RA ~ Location	4	525.32
11.14	RA ~ $CaCl_2$-P	4	528.60
11.15	RA ~ $CaCl_2$-P + Location	5	519.58
11.16	**RA ~ $CaCl_2$-P + Location + $CaCl_2$-P:Location**	**6**	**506.91**
11.17	RA ~ Ox-P	4	532.71
11.18	RA ~ Ox-P + Location	5	524.75
11.19	RA ~ Ox-P + Location + Ox-P:Location	6	524.07
11.20	RA ~ Total-P	4	533.36
11.21	RA ~ Total-P + Location	5	525.76
11.22	RA ~ Total-P + Location + Total-P:Location	6	525.60
ppx			
11.23	RA ~ SWI	4	472.85
11.24	RA ~ Location	4	460.47
11.25	RA ~ $CaCl_2$-P	4	471.85
11.26	RA ~ $CaCl_2$-P + Location	5	467.69
11.27	**RA ~ $CaCl_2$-P + Location + $CaCl_2$-P:Location**	**6**	**448.65**
11.28	RA ~ Ox-P	4	476.43
11.29	RA ~ Ox-P + Location	5	463.83
11.30	RA ~ Ox-P + Location + Ox-P:Location	6	464.86
11.31	RA ~ Total-P	4	479.92
11.32	RA ~ Total-P + Location	5	464.78
11.33	RA ~ Total-P + Location + Total-P:Location	6	466.96

Table 3.15 Linear mixed effects model summaries for single predictor models of relative gene abundance (RA; in units of FPKM) presented in Table 3.13 for *ppk1*, *ppk2*, and *ppx*. Abbreviations: soil wetness index (SWI), standard error (SE), and degrees of freedom (df).

Model	Model Term	Value	SE	df	t-value	p-value
ppk1						
RA ~ SWI	Intercept	621.68	106.91	34	5.82	< 1.00E-05
	SWI	2.65	12.78	8	0.21	0.8411
RA ~ $CaCl_2$-P	Intercept	512.97	44.44	33	11.54	< 1.00E-05
	$CaCl_2$-P	38.80	11.59	33	3.35	0.002
RA ~ Ox-P	Intercept	783.59	79.23	33	9.89	< 1.00E-05
	Ox-P	-1.11	0.55	33	-2.00	0.0536
RA ~ Total-P	Intercept	603.95	147.00	33	4.11	0.0002
	Total-P	0.13	0.47	33	0.27	0.7887
ppk2						
RA ~ SWI	Intercept	519.82	67.60	34	7.69	< 1.00E-05
	SWI	5.43	8.12	8	0.67	0.5224
RA ~ $CaCl_2$-P	Intercept	572.83	41.86	33	13.68	< 1.00E-05
	$CaCl_2$-P	-3.44	10.51	33	-0.33	0.7451
RA ~ Ox-P	Intercept	645.24	54.93	33	11.75	< 1.00E-05
	Ox-P	-0.64	0.38	33	-1.67	0.1046
RA ~ Total-P	Intercept	705.73	91.00	33	7.76	< 1.00E-05
	Total-P	-0.48	0.29	33	-1.64	0.1112
ppx						
RA ~ SWI	Intercept	317.05	39.31	34	8.06	< 1.00E-05
	SWI	3.74	4.70	8	0.80	0.4494
RA ~ $CaCl_2$-P	Intercept	323.03	21.07	33	15.33	< 1.00E-05
	$CaCl_2$-P	6.88	5.31	33	1.30	0.2038
RA ~ Ox-P	Intercept	397.98	31.28	33	12.72	< 1.00E-05
	Ox-P	-0.40	0.22	33	-1.85	0.0737
RA ~ Total-P	Intercept	363.32	55.60	33	6.53	< 1.00E-05
	Total-P	-0.06	0.18	33	-0.32	0.7494

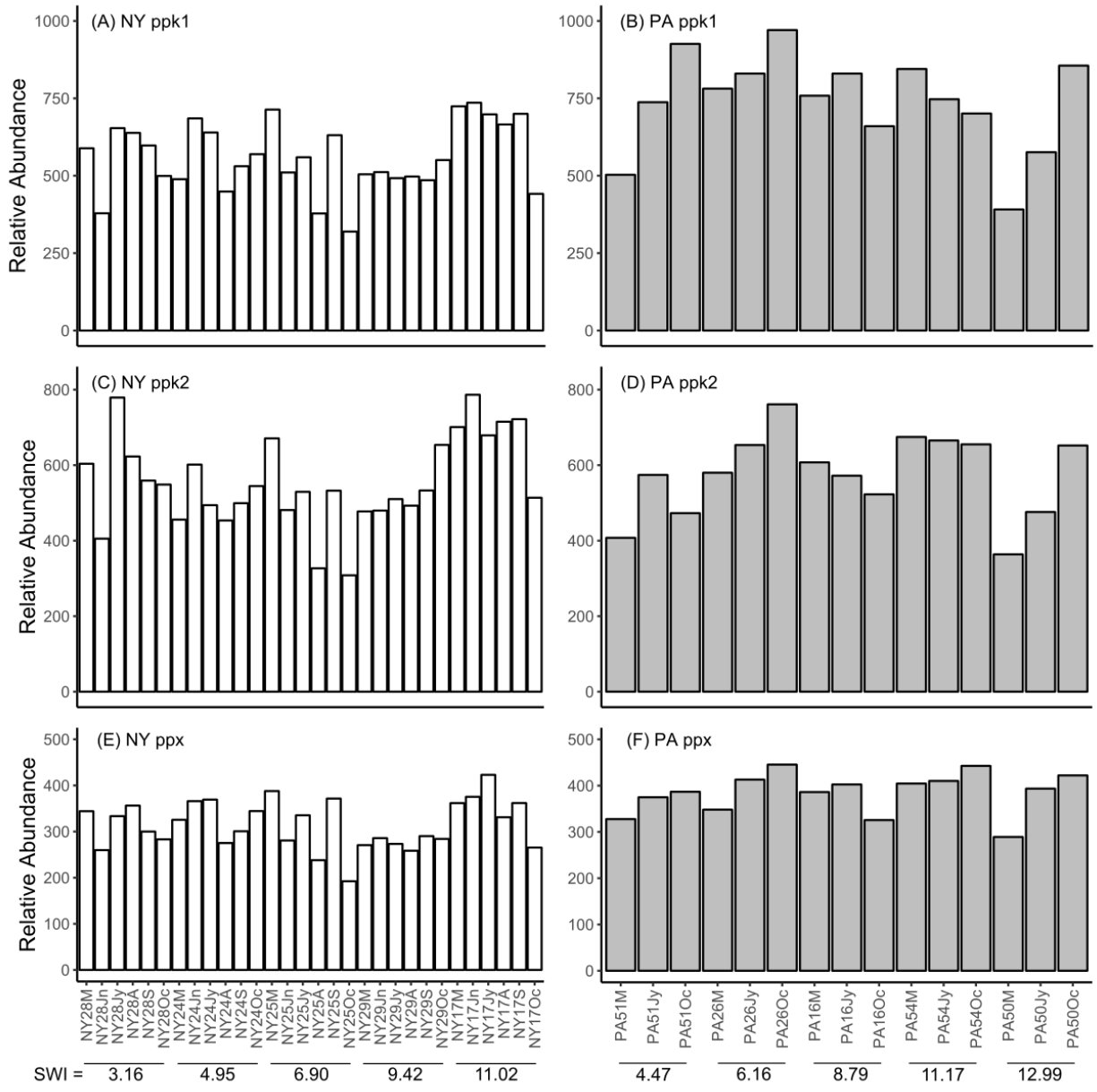

Figure 3. 1 3 Relative gene abundance (FPKM) of *ppk1* (A and B), *ppk2* (C and D), and *ppx* (E and F) versus sample event ordered from lowest to highest SWI and from May to October. New York (NY) results shown in (A), (C), and (E). Pennsylvania (PA) results shown in (B), (D), and (F).

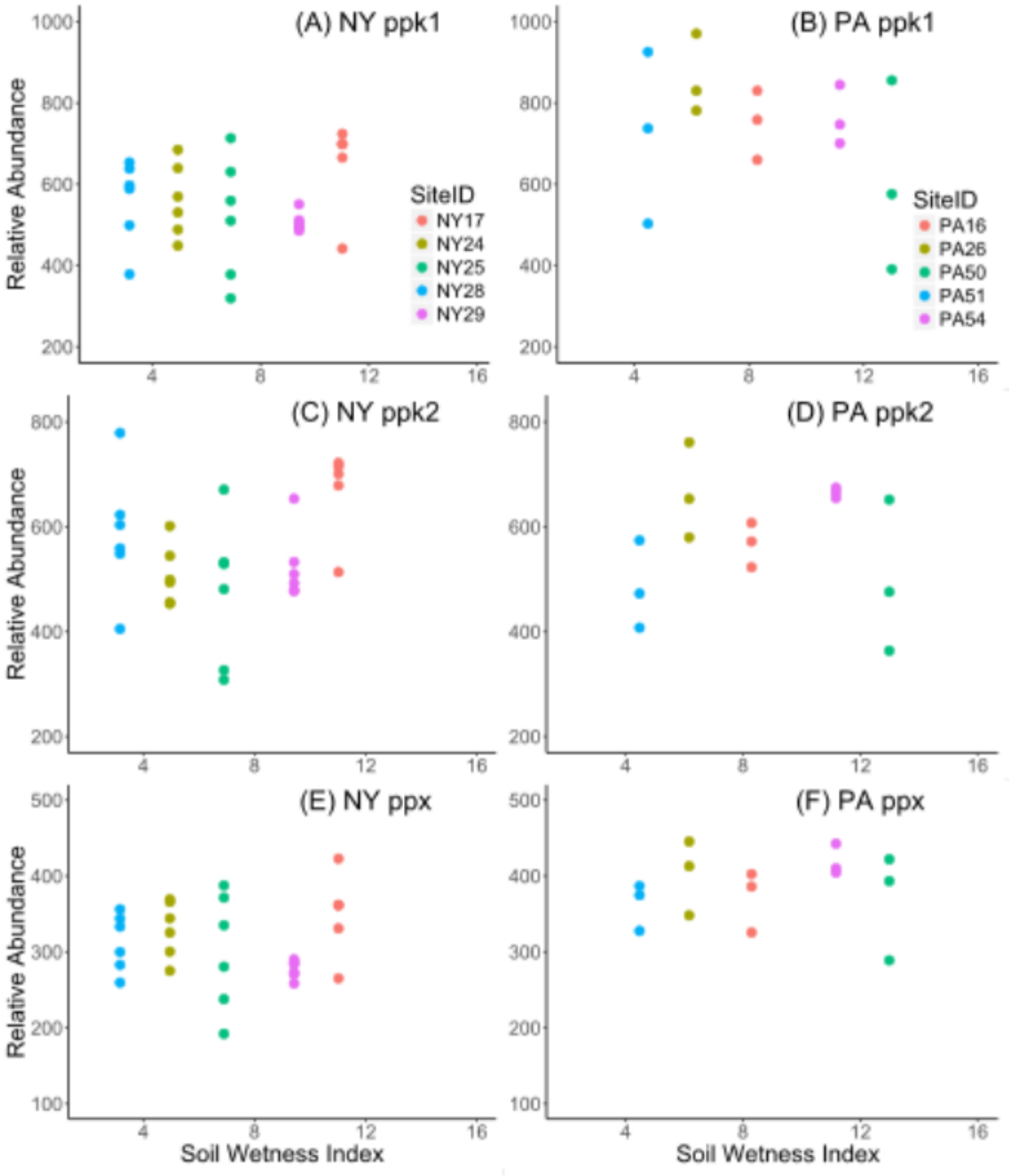

Figure 3. 1 4 Observed relative abundance (FPKM) colored by sample site versus soil wetness index for in New York (NY; A, C, and E) and Pennsylvania (PA; B, D, and F) for *ppk1* (A and B), *ppk2* (C and D), and *ppx* (E and F) genes.

Despite the lack of trend between the relative abundance of the three polyP functional genes and SWI, we observed a high variation in relative abundance at moderate SWI values in NY but not in PA (Figure 3.14). This spread is especially evident for *ppk1* and *ppx* in NY (Figures 3.14A and 3.14E). As mentioned previously, we observed peak P concentrations at moderate SWI values in NY (Figure 3.7) that may indicate the limitations of assuming a linear relationship between soil P concentrations and SWI, as discussed previously. Future studies may focus on whether regions of the landscape with moderate SWI values, rather than those with high SWI values, are more likely to select for PAOs.

Compared to SWI, location alone was a significantly better predictor (i.e., lower AIC) of relative abundance for all three polyP functional genes (Table 3.14). When data were aggregated by location, we observed significantly higher relative abundance of *ppk1* in PA compared to NY ($p=0.0037$; Model 11.2 in Table 3.14, Figure 3.15A). Similarly, the relative abundance of *ppx* was significantly higher in PA compared to NY ($p=0.0034$; Model 11.13 in Table 3.14, Figure 3.15C). However, there was no significant difference between the relative abundance of *ppk2* in NY and PA ($p=0.6351$; Model 11.24 in Table 3.14, Figure 3.15B). It is interesting to note that despite annotating fewer unique contigs in PA (Figure 3.8), the relative abundances of functional genes are similar in magnitude to NY (Figure 3.15).

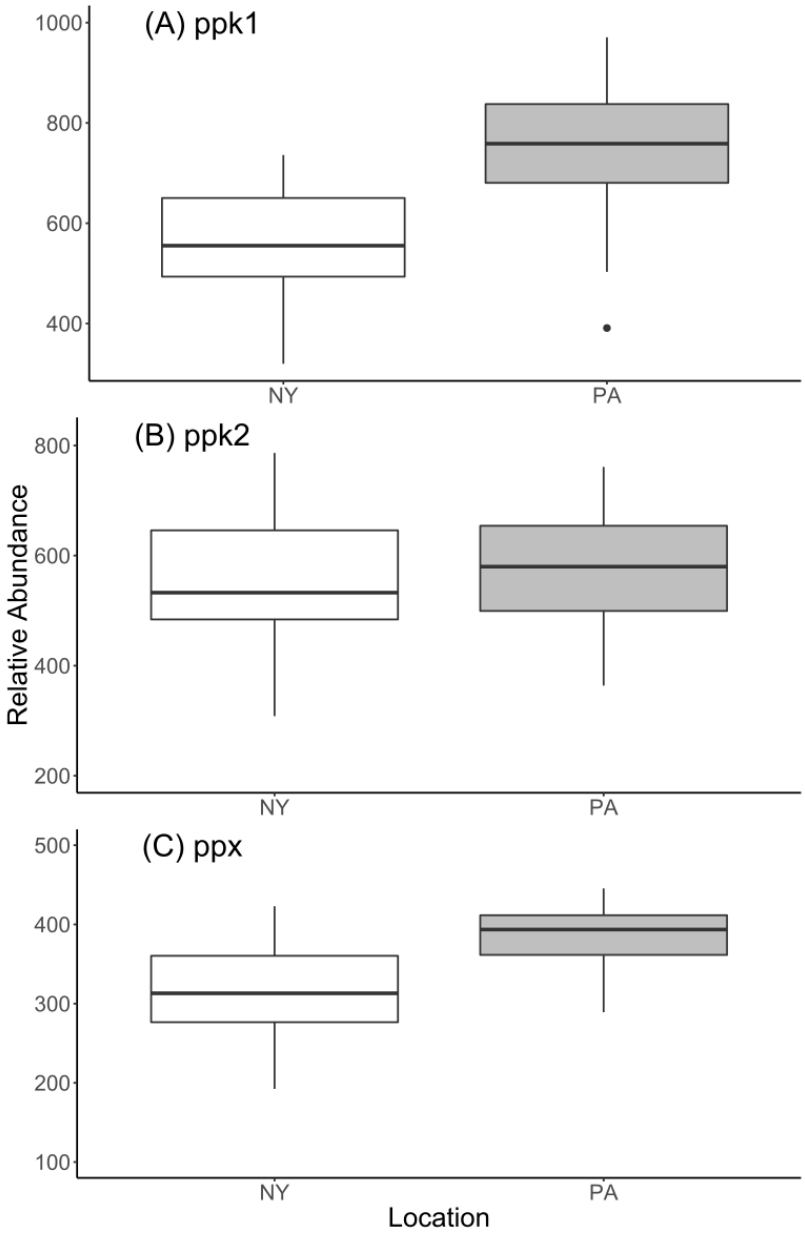

Figure 3. 1 5 Relative gene abundance (FPKM) of (A) *ppk1*, (B) *ppk2*, and (C) *ppx* aggregated by location; whether samples were taken in New York (NY) or Pennsylvania (PA).

LMEM comparisons indicate that including location and soil P pool information improve our ability to predict the relative abundances of *ppk1*, *ppk2*, and *ppx* (Table 3.14). In the case of all three genes, we found that LMEMs incorporating $CaCl_2$-P concentrations alone were more predictive (i.e., lower AIC) than SWI alone, Ox-P alone, or Total-P alone (Table 3.14). However, incorporating $CaCl_2$-P alone was only a significant predictor of *ppk1* relative gene abundance ($p=0.002$, Table 3.15). Beyond these single-term response models, the best fit LMEM for all three genes included $CaCl_2$-P concentration, location, and the interaction of $CaCl_2$-P concentration and location (Models 11.5, 11.16, and 11.27 in Table 3.14). Furthermore, the interaction terms of these best fit models indicate significantly different trends between samples taken in NY and PA for *ppk1* and *ppk2* ($p<0.05$, Table 3.16), which was likely due to differences in $CaCl_2$-P concentrations between the two sites as noted previously (Figure 3.5). Overall in NY, we observed a negative relationship between the relative abundance of all three polyP functional genes and CaCl2-P concentration (Figure 3.16) but the only significant negative trend was for *ppk2* ($p=0.0106$; Table 3.16, Figure 3.16B). In PA, we observed no significant relationship for either of the three polyP functional genes ($p=0.2000$, Table 3.16, Figure 3.16B). In the case of *ppx*, the interaction term as well as $CaCl_2$-P concentration term were not significant predictors of relative abundance ($p>0.05$, Table 3.16).

Table 3.16 Best fit linear mixed effects model (LMEM) summaries (Models 11.5, 11.16, and 11.27 in Table 3.14) for predicting relative abundance (RA; in units of FPKM) of *ppk1*, *ppk2*, and *ppx* genes. 'Location' indicates whether samples were from New York (NY) or Pennsylvania (PA). Abbreviations: standard error (SE) and degrees of freedom (df).

Model Term	Value	SE	df	t-value	p-value
ppk1: RA ~ $CaCl_2$-P + Location + $CaCl_2$-P:Location					
Intercept	671.06	65.33	32	10.27	< 1.00E-05
$CaCl_2$-P	-62.32	32.66	32	-1.91	0.0654
Location (PA)	-105.43	101.95	8	-1.03	0.3313
$CaCl_2$-P:Location (PA)	98.18	35.76	32	2.75	0.0098
ppk2: RA ~ $CaCl_2$-P + Location + $CaCl_2$-P:Location					
Intercept	702.28	62.33	32	11.27	< 1.00E-05
$CaCl_2$-P	-82.72	30.47	32	-2.71	0.0106
Location (PA)	-134.32	96.16	8	-1.40	0.2000
$CaCl_2$-P:Location (PA)	84.34	33.31	32	2.53	0.0165
ppx: RA ~ $CaCl_2$-P + Location + $CaCl_2$-P:Location					
Intercept	361.13	28.78	32	12.55	< 1.00E-05
$CaCl_2$-P	-26.04	14.24	32	-1.83	0.0769
Location (PA)	24.23	44.55	8	0.54	0.6013
$CaCl_2$-P:Location (PA)	25.93	15.57	32	1.67	0.1057

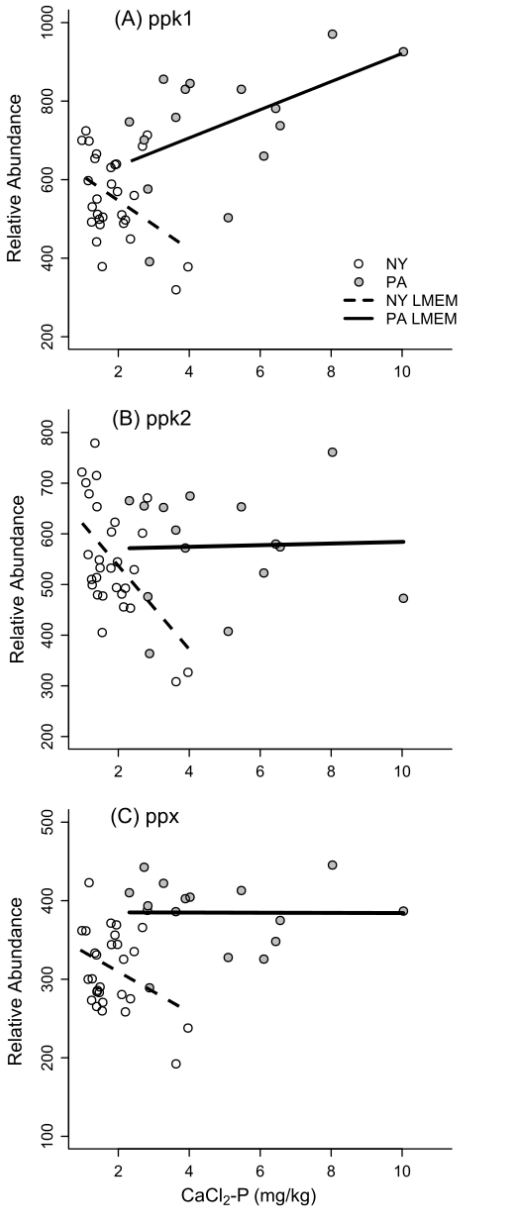

Figure 3. 1 6 Observations and best fit linear mixed effects model (LMEMs, Table 3.16) of relative abundance (FPKM) versus $CaCl_2$-P concentration for (A) *ppk1*, (B) *ppk2*, and (C) *ppx* genes. Note that the only significant ($p<0.05$) trend is for *ppk2* gene results taken from NY.

Overall, we observed that the $CaCl_2$-P pool was better, but not consistently significant, predictor of P functional gene abundance compared to SWI. We also observed that the $CaCl_2$-P pool was the best of all three P pools at predicting the relative abundance for all three genes in this study. We suggest this is because the $CaCl_2$-P pool represents the most mobile P pool compared to the Ox-P and Total-P pools and is likely more readily available to and sourced from microorganisms. The observation of a significant negative relationship between NY *ppk2* relative abundance and $CaCl_2$-P concentrations partially—we also expected SWI to influence P functional gene relative abundance—supports our hypothesis as well as the idea that P limitation might select for organisms that are more likely to store polyP. The findings of this study are consistent with another in the marine environment that showed a negative (albeit, not significant) trend in *ppk2* abundance with increasing phosphate concentrations (Temperton et al. 2011). Conversely to our findings, this same study found a significant negative trend in *ppk1* and *ppx* gene abundance with phosphate concentration. The negative relationship we expected was not significant for all cases and on some occasions indicated the opposite trend; a positive relationship between relative abundance and $CaCl_2$-P concentrations in PA (Figures 3.16A and 3.17B). The reason for this trend is unclear as it indicates the growth of opportunistic organisms that accumulate P but is not consistent with *ppx* trends in PA (Figures 3.16C and 3.17F). Given the assumption that both *ppk1* and *ppx* must be harbored to define a putative PAO, we would expect to see similar trends in relative abundance between pairs (i.e., *ppk1*/*ppx* and *ppk2*/*ppx*). We also observe a great deal of within site variation in relative abundance over time. For example, while the $CaCl_2$-P concentration of site PA50 did not vary much over the period of this study, the relative abundance of *ppk1*, *ppk2*, and *ppx* genes did (Figures 3.17B, 3.17D, and 3.17F).

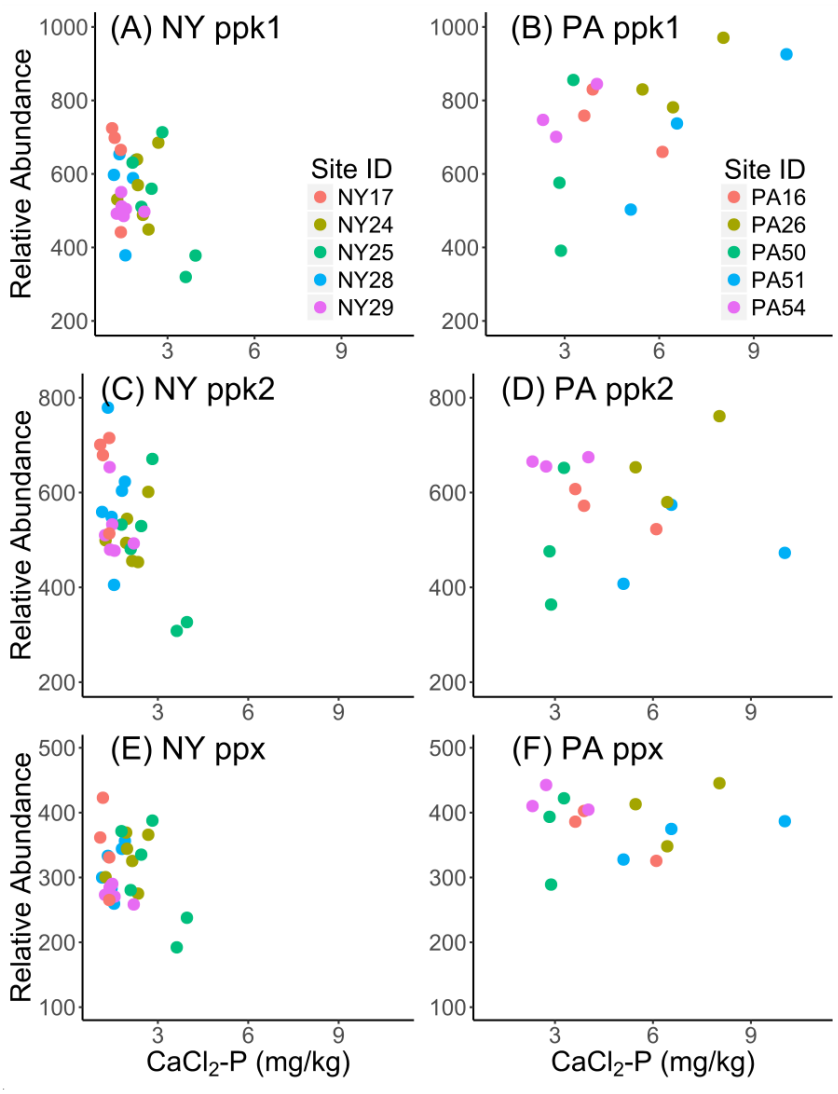

Figure 3. 1 7 Observed relative abundance (FPKM) colored by sample site versus $CaCl_2$-P concentration for in New York (NY; A, C, and E) and Pennsylvania (PA; B, D, and F) for *ppk1* (A and B), *ppk2* (C and D), and *ppx* (E and F) genes.

In addition to using LMEMs to test the impact of environmental variables such as SWI and soil P concentrations on relative abundance of polyP functional genes, we also observed some common descriptive patterns across sampling events. Specifically, we observed higher than average

relative abundances of a handful of contigs at a given sampling site that were not necessarily consistently present from one month to the next (e.g., Figures 3.18 and 3.19). There are several exceptions to this with the most notable being site NY17, which has several contigs harboring *ppk1* that remain abundant throughout the duration of this study (Figure 3.18). NY17 was consistently the wettest site in NY throughout the study period, had obvious iron redox influences (i.e., mottling), had consistently low $CaCl_2$-P concentrations (Figures 3.17A, 3.17B, and 3.17C), and had some of the highest gene abundances of samples from NY (Figures 3.13A, 3.13C, and 3.13E). Therefore, these conditions may have selected for unique organisms capable of tolerating these nutrient limited, iron-rich, continuously saturated conditions. We observed similar clustering of NY17 over time for *ppk2* and *ppx* genes as well (Figures 3.20 and 3.22). Also, specific to NY17, we observed an overwhelming abundance of one contig harboring *ppk2* (k141_4822451, top of Figure 3.20) at this site that increased from May to June, went away from July to August, re-emerged in September, and went away again in October (Figure 3.20). We suggest that this pattern may depend on soil moisture as the end of July as well as months of August and October were dry (Figure 3.1). pBLAST of the PPK2 protein sequence from this contig matched with 69% identity to a hypothetical protein from *Bosea* sp. BIWAKO-01 (*Alphaproteobacteria*). With regard to the potential for concurrent impacts of iron redox and PAOs on patterns at site NY17, we note that it is possible for abiotic iron redox chemistry to mask the role of PAOs; Fe reduction under anaerobic conditions leads to the release of phosphate while aerobic conditions leads to the precipitation of oxidized Fe with phosphate (Dodds 2003). However, one study demonstrated that microbes can mediate iron-phosphate dynamics in sediments (Sulu-Gambari et al. 2016) and another revealed the potential for Fe to stabilize intracellular polyP in estuarine diatoms (Nuester et al. 2012). To summarize, our findings

indicate the need for additional study on the coupled role of iron redox chemistry and P cycling genes in soils under alternating wetting/drying cycles.

With respect to general patterns in PA, we observed a higher than average relative abundance of a few contigs harboring *ppk2* that were consistently abundant across all samples in PA (e.g., top of grid in Figure 3.21) as well as the same for contigs harboring *ppx* in NY (Figure 3.22). Compared to NY, we observed more instances in PA where contigs harboring *ppk1*, *ppk2*, and *ppx* were up to two standard deviations below the mean (e.g., Figure 3.23).

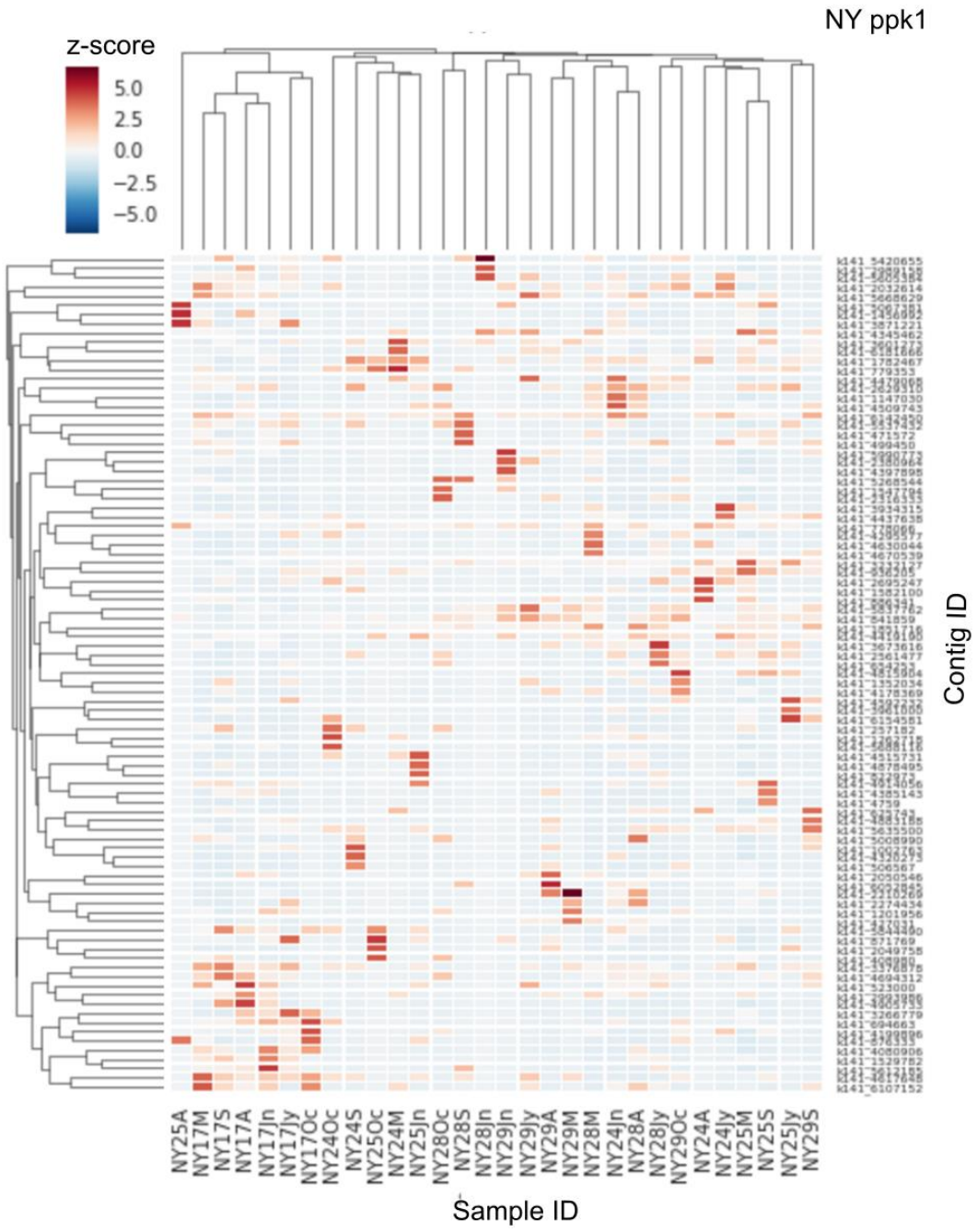

Figure 3. 1 8 Heatmap of relative abundance z-scores for *ppk1* genes found in NY samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O).

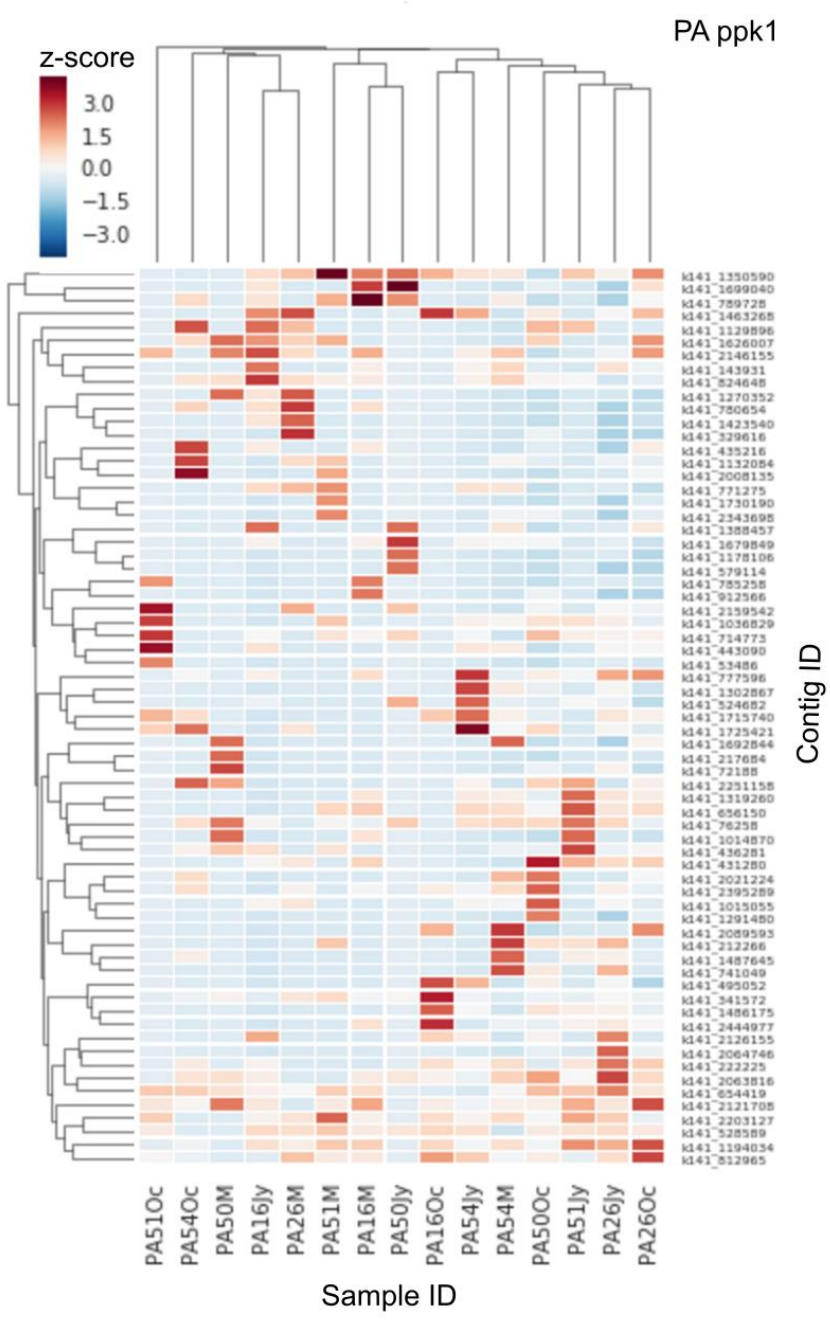

Figure 3. 1 9 Heatmap of relative abundance z-scores for *ppk1* genes found in PA samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O).

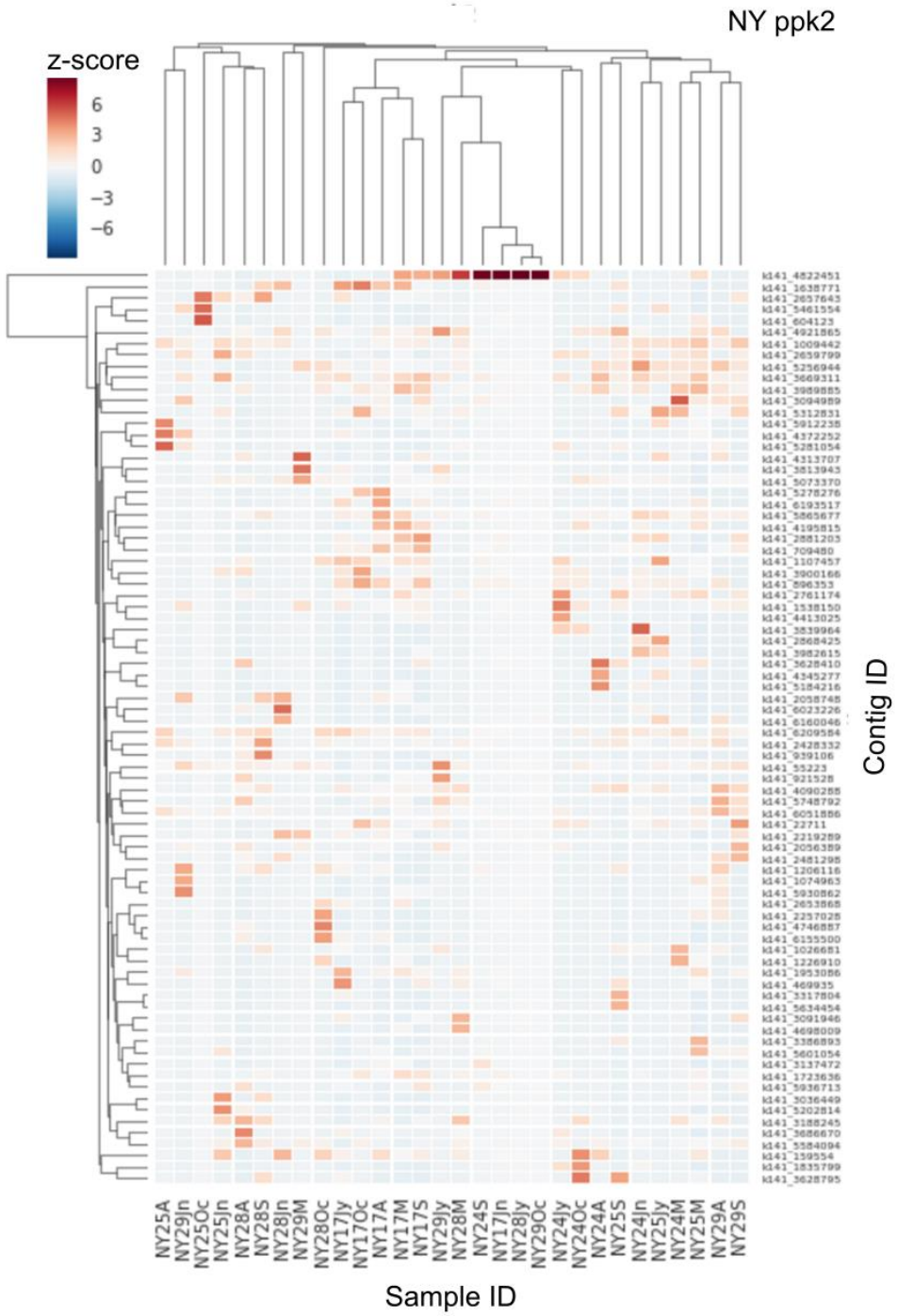

Figure 3.20 Heatmap of relative abundance z-score for *ppk2* genes found in NY samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O).

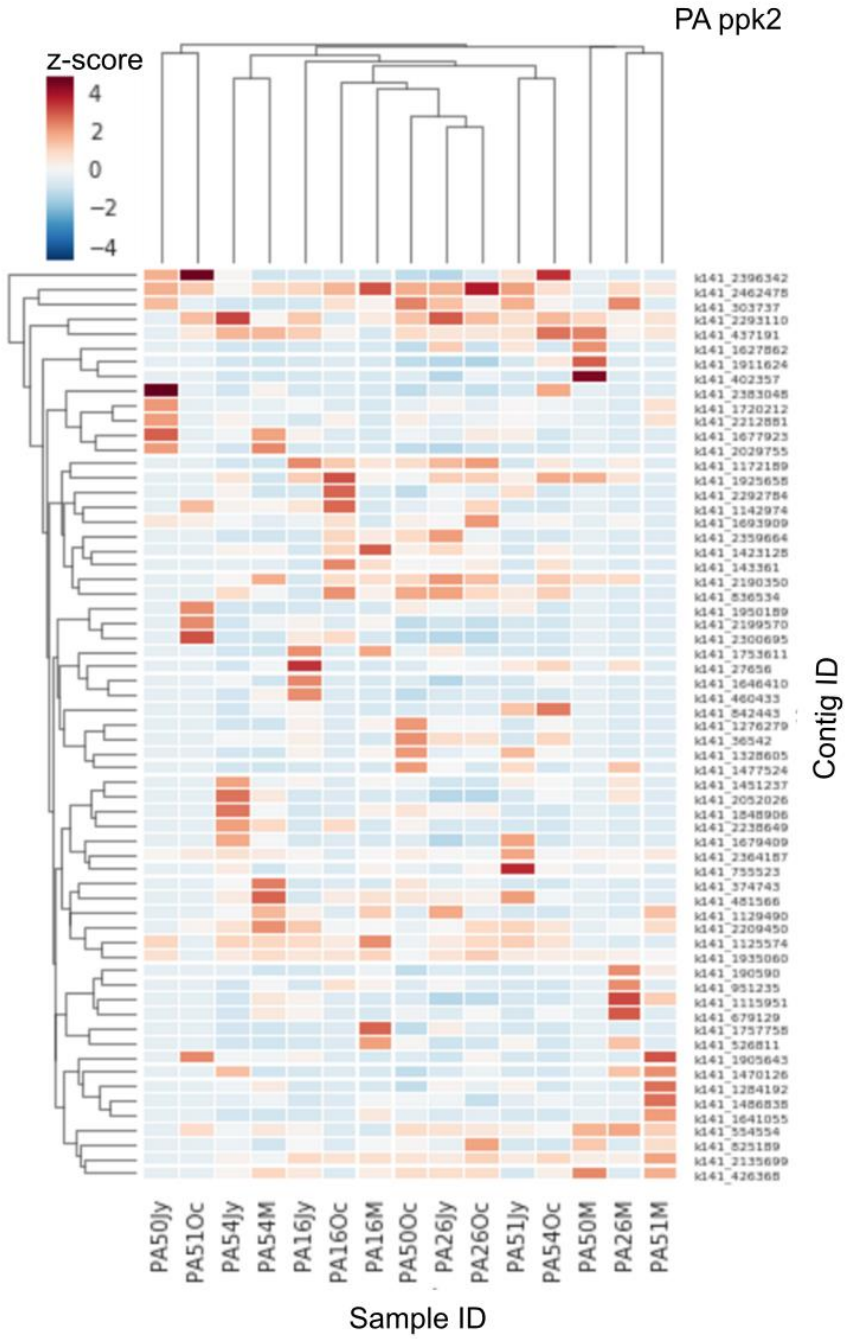

Figure 3. 2 1 Heatmap of relative abundance z-scores for *ppk2* genes found in PA samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O).

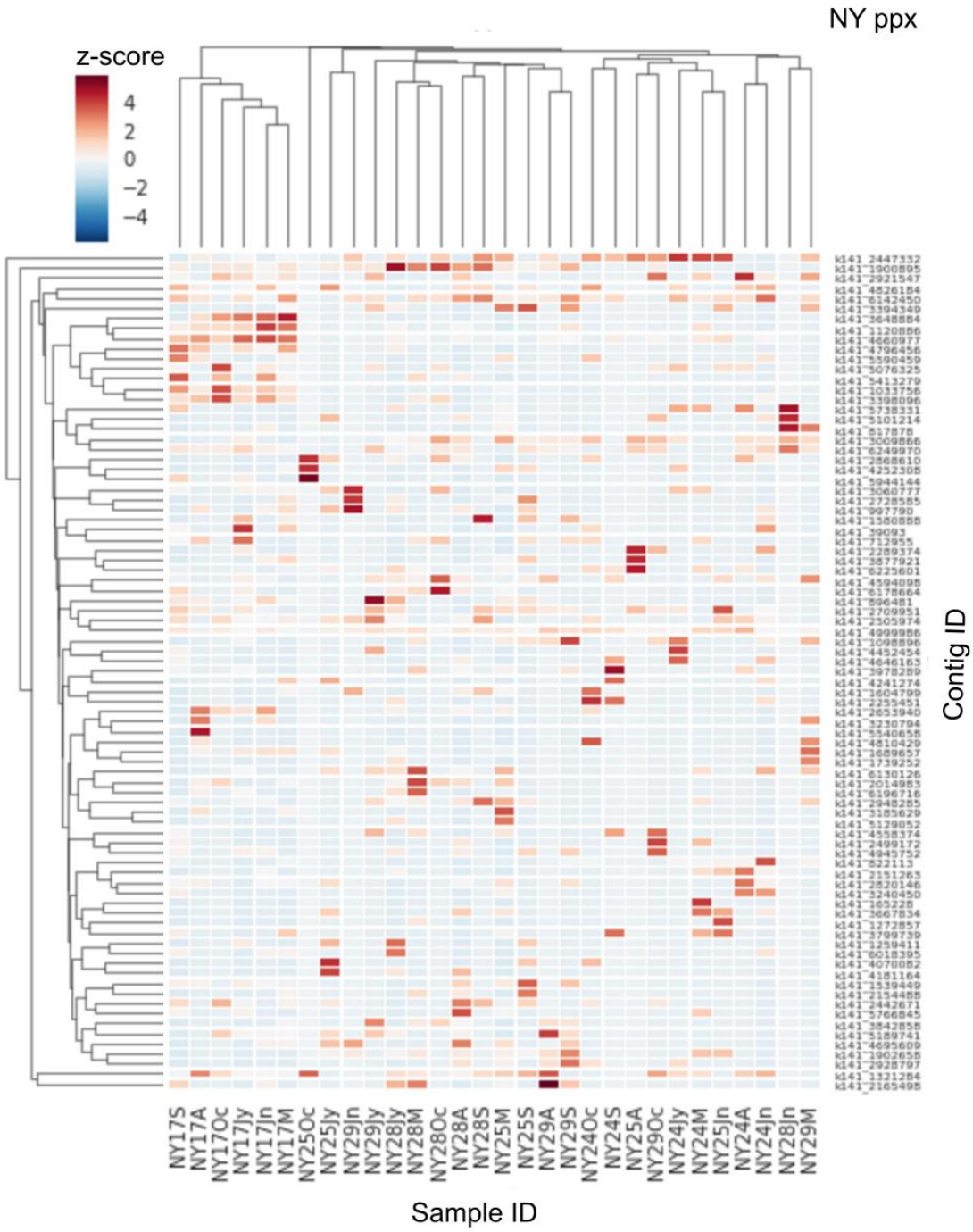

Figure 3. 2 2 Heatmap of relative abundance z-scores for *ppx* genes found in NY samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O).

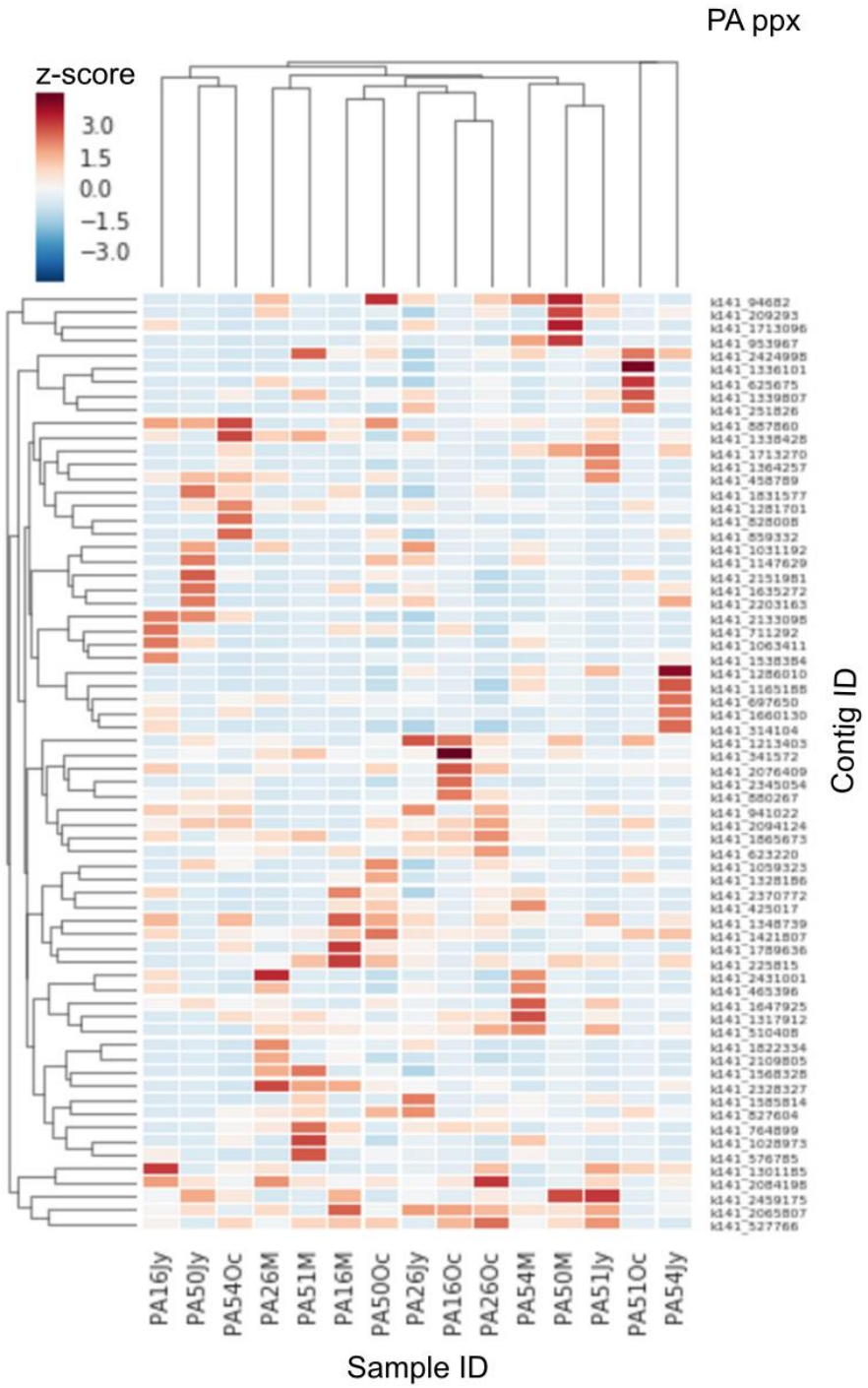

Figure 3. 2 3 Heatmap of relative abundance z-scores for *ppx* genes found in PA samples clustered by similar relative abundance profiles between sample ID and contig ID. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall. Sample ID month abbreviations: May (M), June (Jn), July (Jy), August (A), September (S), and October (O).

CONCLUSIONS

SWIs are currently used to predict soil moisture patterns and may also provide a means to predict the spatial extent of P losses as well as microbially-mediated P cycling to mitigate those losses. We found that SWI is inversely related to mobile forms of P over time—such as those obtained using $CaCl_2$-P extractions. However, despite the significant relationship between SWI and soil $CaCl_2$-P concentrations, SWI cannot be applied directly to predict the relative abundances of polyP functional genes associated. We suggest the reason for this lies in the temporally static nature of the SWI as well as landscape management and other undocumented abiotic and biotic controls. Rather, relative abundance is best predicted from $CaCl_2$-P concentrations but the inverse relationship we observed is only significant in some instances and may depend on overall P availability and location specific characteristics. Because of the lack of relationship between the SWI and PAO associated gene relative abundances as well as the limited number or contigs harboring a combination of *ppk1* and *ppx* genes, PAO-mediated P cycling does not likely play a large role in soils sampled here. Future research may focus on (1) addressing potential limitations of using the SWI as a predictor soil P at sites with moderate SWI values, (2) focusing on how P availability influences PAO-mediated P cycling in soils, and (3) addressing the relative importance of abiotic and biotic controls on P cycling in soil along soil moisture and P availability gradients. Additionally, more work is needed to verify the phenotypes of putative PAOs identified in this study (i.e., those harboring both *ppk1* and *ppx* genes).

ACKNOWLEDGMENTS

We thank Chuck Pepe-Ranney for help with sequencing analysis, Sara Fischel and the USDA-ARS, State College, PA staff for help with total P extractions, Shree Giri for help with ICP-MS

analysis and soil chemistry advice, Jennifer Mosher and Jeff Mattison at the Cornell University Institute for Biotechnology Genomics Facility for help with sample preparation and sequencing, Mallory Choudoir for help with sequencing sample preparation, and Kimberly Carrick, Nicholas Locke, Claudia Rojas, Molly Ryan, Michelle Szeto, Keiran Catalina, and Christine Georgakakos for help with sample collection and analysis. SMS was supported by the Cornell University College of Agriculture and Life Sciences Land Grant Fellowship and US Environmental Protection Agency (EPA) Science to Achieve Results (STAR) Fellowship. This publication was developed under STAR Fellowship Assistance Agreement no. FP917670-01-0 awarded by the EPA. It has not been formally reviewed by the EPA. The views expressed in this publication are solely those of SMS and EPA does not endorse any products or commercial services mentioned in this publication. This project was supported by funding from the US Department of Agriculture (UDSA) grant 2014-67019-21636. Data and R scripts associated with this work are available on GitHub at https://github.com/sheilasaia/paper-p-cycling-in-soils.

REFERENCES

Agnew, L.J., S. Lyon, P. Gérard-Marchant, V.B. Collins, A.J. Lembo, T.S. Steenhuis and M.T. Walter. 2006. Identifying hydrologically sensitive areas: Bridging the gap between science and application. Journal of Environmental Management. 78(1):63–76.

Ahn, K. and A. Kornberg. 1990. Polyphosphate kinase from *Escherichia coli*. Biochemistry. 265(20):11734–11739.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19(6):716-723.

Akiyama, M., E. Crooke and A. Kornberg. 1993. An exopolyphosphatase of *Escherichia coli*. The Journal of Biological Chemistry. 268(1):633–639.

Akiyama, M., E. Crookes and A. Kornberg. 1992. The polyphosphate kinase gene of *Escherichia coli*. The Journal of Biological Chemistry. 267(31):22556–22561.

Alcántara, C., A. Blasco, M. Zúñiga and V. Monedero. 2014. Accumulation of polyphosphate in *Lactobacillus* spp. and its involvement in stress resistance. Applied and Environmental Microbiology. 80(5):1650–1659.

Anantharaman, K., C.T. Brown, L.A. Hug, I. Sharon, C.J. Castelle, A.J. Probst, B.C. Thomas, A. Singh, M.J. Wilkins, U. Karaoz, E.L. Brodie, K.H. Williams, S.S. Hubbard and J.F. Banfield. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nature Communications. 7:13219.

Arguez, A., I. Durre, S. Appelquist, M. Squires, R. Vose, X. Yin, and R. Biotta. 2010. NOAA's U.S. Climate Normals (1981-2010). Station: GHCND:USW00064758. NOAA National Centers for Environmental Information. DOI:10.7289/V5PN93JP Accessed February 1, 2017.

Baker, B.J., C.S. Lazar, A.P. Teske and G.J. Dick. 2015. Genomic resolution of linkages in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria. Microbiome. 3:14.

Bergmann, G.T., S.T. Bates, K.G. Eilers, C.L. Lauber, G. Caporaso, W.A. Walters, R. Knight and N. Fierer. 2011. The under-recognized dominance of Verrucomicrobia in soil bacterial communities. Soil Biologiy ane Biochemistry. 43(7):1450–1455.

Bergström, L., H. Kirchmann, F. Djodjic, K. Kyllmar, B. Ulén, J. Liu, H. Andersson, H. Aronsson, G. Börjesson, P. Kynkäänniemi, A. Svanbäck and A. Villa. 2015. Turnover and losses of phosphorus in Swedish agricultural soils: long-term changes, leaching trends, and mitigation measures. Journal of Environmental Quality. 44(2):512–523.

Bernhardt, E.S., J.R. Blaszczak, C.D. Ficken, M.L. Fork, K.E. Kaiser and E.C. Seybold. 2017. Control points in ecosystems: moving beyond the hot spot hot moment concept. Ecosystems. 20(2):1–18.

Beven, K.J. and M.J. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin. 24(1):43–69.

Blackwell, M.S.A., P.C. Brookes, N. de la Fuente-Martinez, H. Gordon, P.J. Murray, K.E. Snars, J.K. Williams, R. Bol and P.M. Haygarth. 2010. Phosphorus solubilization and potential transfer to surface waters from the soil microbial biomass following drying-rewetting and freezing-thawing 1st ed., Elsevier Inc.

Brenning, A. 2007. RSAGA: SAGA geoprocessing and terrain analysis in R. http://cran. r-project.org/web/packages/RSAGA/index.html.

Brown, M.R.W. and A. Kornberg. 2008. The long and short of it - polyphosphate, PPK and bacterial survival. Trends in Biochemical Sciences. 33(6):284–290.
Bryant, R.B., T.L. Veith, G.W. Feyereisen, A.R. Buda, C.D. Church, G.J. Folmar, J.P. Schmidt, C.J. Dell and P.J.A. Kleinman. 2011. U.S. Department of Agriculture Agricultural Research Service Mahantango Creek Watershed, Pennsylvania, United States: physiography and history. Water Resources Research. 47(8):1–6.
Buchanan, B.P., M. Fleming, R.L. Schneider, B.K. Richards, J. Archibald, Z. Qiu and M.T. Walter. 2014. Evaluating topographic wetness indices across central New York agricultural landscapes. Hydrology and Earth System Sciences. 18(8):3279–3299.
Buda, A.R., P.J.A. Kleinman, G.W. Feyereisen, D.A. Miller, P.G. Knight, P.J. Drohan and R.B. Bryant. 2013. Forecasting runoff from Pennsylvania landscapes. Journal of Soil and Water Conservation. 68(3):185–198.
Buda, A.R., P.J.A. Kleinman, M.S. Srinivasan, R.B. Bryant and G.W. Feyereisen. 2009. Effects of hydrology and field management on phosphorus transport in surface runoff. Journal of Environmental Quality. 38(6):2273–2284.
Burgin, A.J. and P.M. Groffman. 2012. Soil O2 controls denitrification rates and N2O yield in a riparian wetland. Journal of Geophysical Research. 117(1):1–10.
Chacon, N., N. Dezzeo, M. Rangel and S. Flores. 2008. Seasonal Changes in Soil Phosphorus Dynamics and Root Mass along a Flooded Tropical Forest Gradient in the Lower Orinoco River, Venezuela. Biogeochemistry. 87(2):157–168.
Chacon, N., S. Flores and A. Gonzalez. 2006. Implications of iron solubilization on soil phosphorus release in seasonally flooded forests of the lower Orinoco River, Venezuela. Soil Biology and Biochemistry. 38(6):1494–1499.
Christianson, L.E., R.D. Harmel, D. Smith, M.R. Williams and K. King. 2016. Assessment and synthesis of 50 years of published drainage phosphorus losses. Journal of Environment Quality. 45(5):1467–1477.
Collick, A.S., D.R. Fuka, P.J.A. Kleinman, A.R. Buda, J.L. Weld, M.J. White, T.L. Veith, R.B. Bryant, C.H. Bolster and Z.M. Easton. 2015. Predicting phosphorus dynamics in complex terrains using a variable source area hydrology model. Hydrological Processes. 29(4):588–601.
Collick, A.S., T.L. Veith, D.R. Fuka, P.J.A. Kleinman, A.R. Buda, J.L. Weld, R.B. Bryant, P.A. Vadas, M.J. White, R.D. Harmel and Z.M. Easton. 2016. Improved simulation of edaphic and manure phosphorus loss in SWAT. Journal of Environment Quality. 45(4):1215–1225.
Davelaar, D. 1993. Ecological significance of bacterial polyphosphate metabolism in sediments. Hydrobiologia. 253(1):179–192.
DeAngelis, K.M., W.L. Silver, A.W. Thompson and M.K. Firestone. 2010. Microbial communities acclimate to recurring changes in soil redox potential status. Environmental Microbiology. 12(12):3137–3149.
DebRoy, S., P. Mukherjee, S. Roy, A.R. Thakur and S. Raychaudhuri. 2013. Draft genome sequence of a phosphate-accumulating *Bacillus* sp .,. Genome Announcements. 1(1):1–2.
Diaz, J.M., E.D. Ingall, S.D. Snow, C.R. Benitez-Nelson, M. Taillefert and J.A. Brandes. 2012. Potential role of inorganic polyphosphate in the cycling of phosphorus within the hypoxic water column of Effingham Inlet, British Columbia. Global Biogeochemical Cycles. 26(2):1–14.

Dinh, M.-V., T. Schramm, M. Spohn and E. Matzner. 2016. Drying-rewetting cycles release phosphorus from forest soils. Journal of Plant Nutrition and Soil Science. 179(5):670–678.

Dodds, W.K. 2003. The role of periphyton in phosphorus retention in shallow freshwater aquatic systems. Journal of Phycology. 39(5):840–849.

Finn, R.D., J. Clements and S.R. Eddy. 2011. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Research. 39:W29–W37.

Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate, and A. Bateman. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 44(D1):D279-85. Database Issue 44:D279-D285

Flores-López, F., Z.M. Easton, L.D. Geohring, P.J. Vermeulen, V.R. Haden and T.S. Steenhuis. 2013. Factors affecting phosphorous in groundwater in an alluvial valley aquifer: Implications for best management practices. Water (Switzerland). 5(2):540–559.

Gächter, R., J.S. Meyer and A. Mares. 1988. Contribution of bacteria to release and fixation of phosphorus in lake sediments. Limnology and Oceanography. 33(6 part 2):1542–1558.

Geohring, L.D., O. V. McHugh, M.T. Walter, T.S. Steenhuis, M.S. Akhtar and M.F. Walter. 2001. Phosphorus transport into surface drains by macropores after manure applications: implications for best manure management practices. Soil Science. 166(12):896–909.

Gebremariam, S.Y., M.W Beutel, D. Christian, T.F. Hess. 2011. Research advances and challenges in the microbiology of enhanced biological phosphorus removal—a critical review. Water Environmental Research. 83(3):195– 219.

Gerke, J. 2010. Humic (organic matter)-Al(Fe)-phosphate complexes: an underestimated phosphate form in soils and source of plant-available phosphate. Soil Science. 175(9):417–425.

Grierson, P.F., N.B. Comerford and E.J. Jokela. 1998. Phosphorus mineralization kinetics and response of microbial phosphorus to drying and rewetting in a Florida Spodosol. Soil Biology and Biochemistry. 30(10–11):1323–1331.

Guppy, C.N., N.W. Menzies, P.W. Moody and F.P.C. Blamey. 2005. Competitive sorption reactions between phosphorus and organic matter in soil: a review. Australian Journal of Soil Research. 43(2):189.

Hart, M.R., B.F. Quin and M.L. Nguyen. 2004. Phosphorus runoff from agricultural land and direct fertilizer effects: a review. Journal of Environmental Quality. 33(6):1954–1972.

He, S., D.L. Gall and K.D. McMahon. 2007. *Candidatus* Accumulibacter population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes. Applied and Environmental Microbiology. 73(18):5865–5874.

Hesselmann, R.P., C. Werlen, D. Hahn, J.R. van der Meer and a J. Zehnder. 1999. Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. Systematic and Applied Microbiology. 22(3):454–465.

Hofmeister, K.L., C.B. Georgakakos and M.T. Walter. 2016. A runoff risk model based on topographic wetness indices and probability distributions of rainfall and soil moisture for central New York agricultural fields. Journal of Soil and Water Conservation. 71(4):289–300.

Honeycutt, C.W., R.D. Heil and C. V Cole. 1990. Climatic and Topographic Relations of Three Great Plains Soils : II . Carbon, Nitrogen, and Phosphorus. Soil Science Society of America Journal. 483:476-483.

Hupfer, M., S. Gloess and H. Grossart. 2007. Polyphosphate-accumulating microorganisms in aquatic sediments. Aquatic Microbial Ecology. 47(3):299–311.

Hupfer, M., S. Glöss, P. Schmieder and H.P. Grossart. 2008. Methods for detection and quantification of polyphosphate and polyphosphate accumulating microorganisms in aquatic sediments. International Review of Hydrobiology. 93(1):1–30.

Hupfer, M. and J. Lewandowski. 2008. Oxygen controls the phosphorus release from lake sediments - A long-lasting paradigm in limnology. International Review of Hydrobiology. 93(4–5):415–432.

Hyatt, D., G.-L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer and L.J. Hauser. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 11:119.

Jarvie, H.P., L.T. Johnson, A.N. Sharpley, D.R. Smith, D.B. Baker, T.W. Bruulsema and R. Confesor. 2017. Increased soluble phosphorus loads to Lake Erie: unintended consequences of conservation practices? Journal of Environment Quality. 46(1):123–132.

JGI (Joint Genome Institute). 2017. BBMap/BBTools. Available online: https://sourceforge.net/projects/bbmap/. Accessed: March 1, 2017.

Jin, X., Y. Ma, Z. Kong, W. Kou and L. Wu. 2017. The Variation of Sediment Bacterial Community in Response to Anthropogenic Disturbances of Poyang Lake, China. Wetlands. :doi:10.1007/s13157-017-0909-1.

Kant, R., M.W.J. van Passe, A. Palva, S. Lucas, A. Lapidus, T. del Rio, E. Dalin, H. Tice, D. Bruce, L. Goodwin, S. Pitluck, F.W. Larimer, M.L. Land, L. Hauser, P. Sangwan, W.M. de Vos, P.H. Janssen and H. Smidt. 2011. Genome sequence of Chthoniobacter flavus Ellin428, an aerobic heterotrophic soil bacterium. Journal of Bacteriology. 193(11):2902–2903.

Kawakoshi, A., H. Nakazawa, J. Fukada, M. Sasagawa, Y. Katano, S. Nakamura, A. Hosoyama, H. Sasaki, N. Ichikawa, S. Hanada, Y. Kamagata, K. Nakamura, S. Yamazaki and N. Fujita. 2012. Deciphering the genome of polyphosphate accumulating *Actinobacterium Microlunatus phosphovorus*. DNA Research. 19(5):383–394.

Keasling, J.D., L. Bertsch and A. Kornberg. 1993. Guanosine pentaphosphate phosphohydrolase of *Escherichia coli* is a long-chain exopolyphosphatase. Proceedings of the National Academy of Sciences. 90(15):7029–7033.

King, K.W., M.R. Williams, M.L. Macrae, N.R. Fausey, J. Frankenberger, D.R. Smith, P.J.A. Kleinman and L.C. Brown. 2015. Phosphorus transport in agricultural subsurface drainage: a review. Journal of Environmental Quality. 44(2):467–485.

Kleinman, P.J.A. and A.N. Sharpley. 2002. Estimating soil phosphorus sorption saturation from Mehlich-3 data. Communications in Soil Science and Plant Analysis. 33(11/12):1825–1839.

Kleinman, P.J.A., A.N. Sharpley, A.R. Buda, R.W. Mcdowell and A.L. Allen. 2011a. Soil controls of phosphorus in runoff: management barriers and opportunities. Canadian Journal of Soil Science. 91(3):329–338.

Kleinman, P.J.A., A.N. Sharpley, R.W. McDowell, D.N. Flaten, A.R. Buda, L. Tao, L. Bergstrom and Q. Zhu. 2011b. Managing agricultural phosphorus for water quality protection: principles for progress. Plant and Soil. 349(1–2):169–182.

Kleinman, P.J.A., A.N. Sharpley, L.S. Saporito, A.R. Buda and R.B. Bryant. 2009. Application of manure to no-till soils: phosphorus losses by sub-surface and surface pathways. Nutrient Cycling in Agroecosystems. 84(3):215–227.
Kleinman, P.J., P. Salon, N. Sharpley and L.S. Saporito. 2005. Effect of cover crops established at time of corn planting on phosphorus runoff from soils before and after dairy manure application. Journal of Soil and Water Conservation. 60(6):311–322.
Kornberg, A. 1995. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. Journal of Bacteriology. 177(3):491–496.
Kovar, J.L., T.B. Moorman, J.W. Singer, C.A. Cambardella and M.D. Tomer. 2011. Swine manure injection with low-disturbance applicator and cover crops reduce phosphorus loss. Journal of Environment Quality. 40(2):329–336.
Kunin, V., S. He, F. Warnecke, S.B. Peterson, H. Garcia Martin, M. Haynes, N. Ivanova, L.L. Blackall, M. Breitbart, F. Rohwer, K.D. McMahon and P. Hugenholtz. 2008. A bacterial metapopulation adapts locally to phage predation despite global dispersal. Genome Research. 18(2):293–297.
Li, H. 2017. bwa-mem. Available online: http://bio-bwa.sourceforge.net/. Accessed: March 1, 2017.
Li, D., C.M. Liu, R. Luo, K. Sadakane and T.W. Lam. 2015. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 31(10):1674–1676.
Li, X., H. Yuan, J. Yang and B. Li. 2013. Genome sequence of the polyphosphate-accumulating organism Arthrobacter sp. strain PAO19 isolated from maize rhizosphere soil. Genome Announcements. 1(4):11039–11040.
Litaor, M.I., T.R. Seastedt, M.D. Walker, M. Carbone and A. Townsend. 2005. The biogeochemistry of phosphorus across an alpine topographic/snow gradient. Geoderma. 124(1–2):49–61.
Lyon, S.W., T.M. Walter, P. Gerard-Marchant and T.S. Steenhuis. 2004. Using a topographic index to distribute variable source area runoff predicted with the SCS curve-number equation. Hydrological Processes. 18(15):2757–2771.
Marjerison, R.D., H. Dahlke, Z.M. Easton, S. Seifert and M.T. Walter. 2011. A Phosphorus Index transport factor based on variable source area hydrology for New York State. Journal of Soil and Water Conservation. 66(3):149–157.
McKeague, J.A. and J.H. Day. 1966. Dithionite- and oxalate-extractable Fe and Al as aids in differentiating various classes of soils. Canadian Journal of Soil Science. 46(1):13–22.
McMahon, K.D. and E.K. Read. 2013. Microbial contributions to phosphorus cycling in eutrophic lakes and wastewater. Annual review of microbiology. 67:199–219.
McParland, E., C.R. Benitez-Nelson, G.T. Taylor, R. Thunell, A. Rollings and L. Lorenzoni. 2015. Cycling of suspended particulate phosphorus in the redoxcline of the Cariaco Basin. Marine Chemistry. 176:64–74.
Menne, M.J., I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, X. Yin, S. Anthony, R. Ray, R.S. Vose, B.E. Gleason, and T.G. Houston. 2012. Global Historical Climatology Network - Daily (GHCN-Daily), Version 3. Station: GHCND:USW00064758. NOAA National Climatic Data Center. doi:10.7289/V5D21VHZ [March 1, 2017].
Menzel, P., K. L. Ng, and A. Krogh. 2016. Kaiju : Fast and sensitive taxonomic classification for metagenomics. Nature Communications. 7:11257.

NCBI (National Center for Biotechnology Information, Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2017 Mar 06]. Available from: https://www.ncbi.nlm.nih.gov/.

Nguyen, H.T.T., J.L. Nielsen and P.H. Nielsen. 2012. "Candidatus Halomonas phosphatis", a novel polyphosphate-accumulating organism in full-scale enhanced biological phosphorus removal plants. Environmental Microbiology. 14(10):2826–2837.

Nuester, J., S. Vogt and B.S. Twining. 2012. Localization of iron within centric diatoms of the genus Thalassiosira. Journal of Phycology. 48(3):626–634.

Oehmen, A., P.C. Lemos, G. Carvalho, Z. Yuan, J. Keller, L.L. Blackall and M.A.M. Reis. 2007. Advances in enhanced biological phosphorus removal: From micro to macro scale. Water Research. 41(11):2271–2300.

Ouyang, Y. and X. Li. 2013. Recent research progress on soil microbial responses to drying–rewetting cycles. Acta Ecologica Sinica. 33(1):1–6.

Pellerin, B.A., B.A. Stauffer, D.A. Young, D.J. Sullivan, S.B. Bricker, M.R. Walbridge, G.A. Clyde and D.M. Shaw. 2016. Emerging tools for continuous nutrient monitoring networks: sensors advancing science and water resources protection. Journal of the American Water Resources Association. 52(4):993–1008.

Peralta, A.L., S. Ludmer, J.W. Matthews and A.D. Kent. 2014. Bacterial community response to changes in soil redox potential along a moisture gradient in restored wetlands. Ecological Engineering. 73:246–253.

Peterson, S.B., F. Warnecke, J. Madejska, K.D. McMahon and P. Hugenholtz. 2008. Environmental distribution and population biology of *Candidatus* Accumulibacter, a primary agent of biological phosphorus removal. Environmental Microbiology. 10(10):2692–703.

Pett-Ridge, J. and M.K. Firestone. 2005. Redox fluctuation structures microbial communities in a wet tropical soil redox fluctuation structures microbial communities in a wet tropical soil. Applied and Environmental Microbiology. 71(11):6998–7007.

Pinheiro J, D. Bates, S. DebRoy, S. Sakar, and R Core Team. 2017. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-131, https://CRAN.R-project.org/package=nlme.

Probst, A.J., C.J. Castelle, A. Singh, C.T. Brown, K. Anantharaman, I. Sharon, L.A. Hug, D. Burstein, J.B. Emerson, B.C. Thomas and J.F. Banfield. 2017. Genomic resolution of a cold subsurface aquifer community provides metabolic insights for novel microbes adapted to high CO2 concentrations. Environmental Microbiology. 19:459–474.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rangarajan, E.S., G. Nadeau, Y. Li, J. Wagner, M.N. Hung, J.D. Schrag, M. Cygler and A. Matte. 2006. The Structure of the exopolyphosphatase (PPX) from *Escherichia coli* O157:H7 suggests a binding mode for long polyphosphate chains. Journal of Molecular Biology. 359(5):1249–1260.

Rao, N.N., M.R. Gómez-García and A. Kornberg. 2009. Inorganic polyphosphate: essential for growth and survival. Annual Review of Biochemistry. 78:605–647.

Reddy, K.R., R.H. Kadlec, E. Flaig and P.M. Gale. 1999. Phosphorus retention in streams and wetlands: a review. Critical Reviews in Environmental Science and Technology. 29(1):83–146.

Reizer, J., A. Reizer and M.H. Saier. 1993. Exopolyphosphatase and guanine pentaphosphate phosphatase belong to the sugar kinase/actin/hsp70 superfamily. Trends in Biochemical Sciences. :247–248.
Rittenburg, R.A., A.L. Squires, J. Boll, E.S. Brooks, Z.M. Easton and T.S. Steenhuis. 2015. Agricultural BMP effectiveness and dominant hydrological flow paths: concepts and a review. Journal of the American Water Resources Association. 51(2):305–329.
Rode, M., A.J. Wade, M.J. Cohen, R.T. Hensley, M.J. Bowes, J.W. Kirchner, G.B. Arhonditsis, P. Jordan, B. Kronvang, S.J. Halliday, R.A. Skeffington, J.C. Rozemeijer, A.H. Aubert, K. Rinke and S. Jomaa. 2016. Sensors in the stream: the high-frequency wave of the present. Environmental Science and Technology. 50(19):10297–10307.
Schelde, K., L.W. de Jonge, C. Kjaergaard, M. Laegdsmand and G.H. Rubæk. 2006. Effects of Manure Application and Plowing on Transport of Colloids and Phosphorus to Tile Drains. Vadose Zone Journal. 5(1):445.
Schimel, J.P., J.M. Gulledge, J.S. Clein-Curley, J.E. Lindstrom and J.F. Braddock. 1999. Moisture effects on microbial activity and community structure in decomposing birch litter in the Alaskan taiga. Soil Biology and Biochemistry. 31(6):831–838.
Schulz, H.N. and H.D. Schulz. 2005. Large sulfur bacteria and the formation of phosphorite. Science. 307:416–418.
Self-Davis, M.L., P.A. Moore, Jr., and B.C. Joern. 2009. Water- or Dilute Salt-Extractable Phosphorus in Soil. P. 22-24. In Kovar, J.L. And G.M. Pierzynski (eds). Methods of Phosphorus Analysis for Soils, Sediments, Residuals and Waters, 2nd Edition. SERA-17, Southern Cooperative Series Bulletin. Available online at: http://www.sera17.ext.vt.edu/Documents/P_Methods2ndEdition2009.pdf
Seviour, R.J. and S. McIlroy. 2008. The microbiology of phosphorus removal in activated sludge processes-the current state of play. Journal of Microbiology. 46(2):115–124.
Seviour, R.J., T. Mino and M. Onuki. 2003. The microbiology of biological phosphorus removal in activated sludge systems. FEMS Microbiology Reviews. 27(1):99–127.
Seviour, R.J. and P.H. Nielsen. 2010. Microbial Ecology of Activated Sludge R. J. Seviour & P. H. Nielsen, eds., London, UK: IWA Publishing.
Silver, W.L., a E. Lugo and M. Keller. 1999. Soil oxygen availability and biogeochemical cycling along elevation and topographic gradients in Puerto Rico. Biogeochemistry. 44:301–328.
Smeck, N.E. 1973. Phosphorus: an indicator of pedogenetic weathering processes. Soil Science. 115:199–206.
Soo, R.M., C.T. Skennerton, Y. Sekiguchi, M. Imelfort, S.J. Paech, P.G. Dennis, J.A. Steen, D.H. Parks, G.W. Tyson and P. Hugenholtz. 2014. An expanded genomic representation of the phylum Cyanobacteria. Genome Biology and Evolution. 6(5):1031–1045.
Sulu-Gambari, F., D. Seitaj, F.J.R. Meysman, R. Schauer, L. Polerecky and C.P. Slomp. 2016. Cable bacteria control iron-phosphorus dynamics in sediments of a coastal hypoxic basin. Environmental Science and Technology. 50(3):1227–1233.
Temperton, B., J. a Gilbert, J.P. Quinn and J.W. McGrath. 2011. Novel analysis of oceanic surface water metagenomes suggests importance of polyphosphate metabolism in oligotrophic environments. PLOS One. 6(1):1–14.
Thomas, I.A., P.E. Mellander, P.N.C. Murphy, O. Fenton, O. Shine, F. Djodjic, P. Dunlop and P. Jordan. 2016. A sub-field scale critical source area index for legacy phosphorus

management using high resolution data. Agriculture, Ecosystems and Environment. 233:238–252.
Torsten Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous Inference in General Parametric Models. Biometrical Journal 50(3):346-363.
Turner, B.L., J.P. Driessen, P.M. Haygarth and I.D. Mckelvie. 2003. Potential contribution of lysed bacterial cells to phosphorus solubilization in two rewetted Australian pasture soils. Soil Biology and Biochemistry. 35(1):187–189.
Turner, B.L. and P.M. Haygarth. 2001. Phosphorus solubilization in rewetted soils. Nature. 411:258.
USEPA (US Environmental Protection Agency). 1978. Acid Digestion of Sediments, Sludges, and Soils. US Environmental Protection Agency, Washington, DC.
USEPA (US Environmental Protection Agency). 1978. Method 365.3. Phosphorous, all forms (colorimetric, ascorbic acid, two reagent). US Environmental Protection Agency, Washington, DC.
USDA (US Department of Agriculture). 2006. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online at https://sdmdataaccess.sc.egov.usda.gov. Accessed February 1, 2015.
USDA (US Department of Agriculture). 2016. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Official Soil Series Descriptions. Available online: https://soilseries.sc.egov.usda.gov/osdname.aspx. Accessed December 1, 2016.
Verbree, D.A., S.W. Duiker and P.J.A. Kleinman. 2010. Runoff losses of sediment and phosphorus from no-till and cultivated soils receiving dairy manure. Journal of Environment Quality. 39(5):1762–1770.
Vidon, P., C. Allan, D. Burns, T.P. Duval, N. Gurwick, S. Inamdar, R. Lowrance, J. Okay, D. Scott and S. Sebestyen. 2010. Hot spots and hot moments in riparian zones: potential for improved water quality management. Journal of the American Water Resources Association. 46(2):278–298.
Walter, M.T., T.S. Steenhuis, V.K. Mehta, D. Thongs, M. Zion and E. Schneiderman. 2002. Refined conceptualization of TOPMODEL for shallow subsurface flows. Hydrological Processes. 16(10):2041-2046.
Western, A. W., R. B. Grayson, G. Bloschl, G. R. Willgoose, and T. A. McMahon. 1999. Observed spatial organization of soil moisture and its relation to terrain indices. Water Resources Research. 35(3):797-810.
Wilson, H.F., S. Satchithanantham, A.P. Moulin and A.J. Glenn. 2016. Soil phosphorus spatial variability due to landform, tillage, and input management: a case study of small watersheds in southwestern Manitoba. Geoderma. 280:14–21.
Young, I.M. and J.W. Crawford. 2004. Interactions and Self-Organization in the Soil-Microbe Complex. Science. 304(5677):1634–1637.
Yuan, Z., S. Pratt and D.J. Batstone. 2012. Phosphorus recovery from wastewater through microbial processes. Current Opinion in Biotechnology. 23(6):878–883.
Zago, A., S. Chugani and A.M. Chakrabarty. 1999. Cloning and characterization of polyphosphate kinase and exopolyphosphatase genes from *Pseudomonas aeruginosa* 8830. Applied and Environmental Microbiology. 65(5):2065–2071.

Zhang, A.N. et al. 2016. Development of quantitative real-time PCR assays for different clades of *Candidatus* Accumulibacter. Scientific Reports. 6:23993.
Zhang, H., K. Ishige and A. Kornberg. 2002. A polyphosphate kinase (PPK2) widely conserved in bacteria. Proceedings of the National Academy of Sciences. 99(26):16678–16683.
Zilles, J.L., C.H. Hung and D.R. Noguera. 2002. Presence of Rhodocyclus in a full-scale wastewater treatment plant and their participation in enhanced biological phosphorus removal. Water Science and Technology. 46(1–2):123–128.

CHAPTER 4

TEMPORAL TRENDS OF DECOMPOSING LEAF LITTER PHOSPHORUS CONCENTRATIONS AND THE RELATIVE ABUNDANCE OF POLYPHOSPHATE CYCLING GENES

ABSTRACT

Leaf litter is an important source of phosphorus (P) in forests ecosystems and a few studies suggest that microbially-mediated processes may influence P retention and cycling in these environments. One potential microbial P retention mechanism that has received very little attention involves polyphosphate accumulating organisms (PAOs), which have been shown to accumulate and liberate intracellular P as polyphosphate (polyP), depending on environmental conditions. The abundance and characterization of microorganisms harboring key genes involved in polyP metabolism of PAOs—*ppk* and *ppx*—are not well studied in the environment and may provide a new perspective for microbial controls on P cycling in forest ecosystems. We hypothesized the relative abundance of polyP functional genes (i.e., *ppk1*, *ppk2,* and *ppx*) on leaf litter increases concurrently over time with increases in leaf-associated P (leaf P). We also expected to see taxa of known and unknown PAOs inhabiting the leaf litter surface. To test these hypotheses, we placed leaf litter bags in the stream and on the forest floor, collected them over several months, and analyzed them for decomposition rates and leaf P concentration. We also extracted DNA from leaf litter samples and assembled shotgun metagenomes to characterize the microbial species harboring *ppk1*, *ppk2,* and *ppx* and determined the relative abundance of these genes. Stream leaf P concentrations increased significantly over time but relative abundances of *ppk1*, *ppk2,* and *ppx* remained constant indicating an alternative (likely non-PAO related)

mechanism explained the increases in leaf P over time. We suggest one possible explaination could be due to the buildup of sediment-bound P on the leaf litter over time. In forest floor, leaf P concentrations decreased significantly over time while the relative abundance varied. Specifically, of *ppk2* increased significantly and *ppk1* and *ppx* remained constant. We found contigs harboring both *ppk1* and *ppx* genes that belonged to similar taxa as known PAOs (e.g., *Betaproteobacteria*) as well as many understudied taxa (e.g., *Burkholderia*). These contigs may represent putative PAOs but additional work is required to verify this suggestion. While we identified putative PAOs, non-significant patterns in relative abundance over time indicate limited support for PAO-mediated P cycling associated with leaf litter. Additionally, these results demonstrate the commonality of polyP functional genes on leaf litter in streams as well as on the forest floor and provide a stepping stone for future work addressing the ecological role of PAOs in nutrient limited forest environments.

KEYWORDS

phosphorus, polyphosphate, polyphosphate accumulating organisms, leaf litter decomposition, polyphosphate kinase, exopolyphosphatase, microbial processes

INTRODUCTION

Leaf litter provides energy and nutrients, such as phosphorus (P), to forest ecosystems (Fisher and Likens 1973; Mehring et al. 2015). Leaf litter in streams can retain P via a combination of abiotic and biotic mechanisms including (1) microbial accumulation of water column P, (2) entrapment of sediment bound P in the leaf litter biofilm, (3) chemical binding of water column

P with the leaf organic matter, and (4) entrapped sediment minerals (Gregory 1978; Meyer 1980; Dodds 2003; Mehring et al. 2015). Several studies have observed accumulations of P on leaf litter in streams over time due to the combination of these processes (Gregory 1978; Meyer 1980; Mehring et al. 2015). Leaf litter also provides a source of P to organisms living on the forest floor; studies have documented the decline in leaf P concentrations on the forest floor over time (Moore et al. 2006; Ball et al. 2009; Schmidt et al. 2016; Smyth et al. 2016). The decline occurs faster when the soil below is P limited (Moore et al. 2006).

Regardless of whether leaf litter falls in the stream or on the forest floor, many studies document the role of microorganisms in leaf litter decomposition and associated P cycling (for a full review see Lladó et al. 2017). Microorganisms are thought to colonize the surface of the leaf where they can accumulate P from water as it flows past (Gregory 1978) or release acid phosphatase to increase the availability of P from leaf litter on the forest floor (Purahong et al. 2016). While there is still some debate concerning the timing and role of bacterial versus fungal communities on leaf litter decomposition, it is clear that these populations are dynamic and driven by nutrient availability and nutrient demand (Gregory 1978; Elwood et al. 1981; Webster and Benfield 1986; Gulis and Suberkropp 2003; Cheever et al. 2012; Xu et al. 2013; Fanin et al. 2016; Martínez et al. 2016; Purahong et al. 2016). Researchers have measured the bulk influence of microorganisms on decomposing leaf litter and associated P cycling via respiration experiments (Elwood et al. 1981; Martínez et al. 2016), molecular biology and microscopy based methods to determine of microbial biomass (Gulis and Suberkropp 2003; Mehring et al. 2015), radioactive P isotopes (Gregory 1978; Newbold et al. 1983), and enzyme assays (Purahong et al. 2016). For example, one study found that the bacterial richness of leaf litter colonizers was positively

correlated with acid phosphatase activity (Purahong et al. 2016). This same study also used DNA sequencing techniques to track the dynamics of bacterial and fungal communities over time. We found studies testing specific metabolic mechanisms controlling leaf litter decomposition and associated P cycling.

One potential metabolic mechanism that may enable microorganisms to thrive in P limited environments such as forest ecosystems depends on the molecule polyphosphate (polyP). PolyP is a polymer composed of three or more phosphate molecules bound together via a phosphoanhydride bond (Kornberg 1995). The ability to store polyP intracellularly has been identified in many diverse organisms including bacteria, fungi, plants, animals utilize polyP (Kornberg 1995; Zhang et al. 2002; Schulz and Schulz 2005; Rao et al. 2009). There are a handful of studies that have either identified intracellular polyP granules or characterized the role of intracellular polyP storage in freshwater sediments and biofilms (Uhlmann and Bauer 1988; Hupfer and Gächter 1995; Hupfer et al. 2004; Locke 2015; Rier et al. 2016; Saia et al. 2017) but we know of no studies that have identified intracellular polyP granules in soils or other terrestrial ecosystems (e.g., leaf litter). A few studies of soils and sediments have identified genes associated with microorganisms that are known to accumulate P as polyP (Kunin et al. 2008; Peterson et al. 2008; Martins et al. 2011). More specifically, these microorganisms—referred to as polyphosphate accumulating organisms (PAOs)—are well studied in the context of specialized wastewater treatments plants (WWTP) for their ability to accumulate phosphate and store it as intracellular polyP under aerobic conditions and to break down polyP (releasing phosphate) to fuel C uptake under anaerobic conditions (Seviour et al. 2003; Seviour and Nielsen 2010). Since the discovery of the model PAO (i.e., *Candidatus* Accumulibacter phosphatis;

Hesselmann et al. 1999), many diverse PAOs have been identified and studied in WWTPs (Nakamura et al. 1995b; Kong et al. 2007; Günther et al. 2009; Nguyen et al. 2011; Kristiansen et al. 2013) and a couple have been identified in soils (Li et al. 2013; DebRoy et al. 2013).

PolyP synthesis and degradation in PAOs is thought to be controlled by *ppk* and *ppx* genes, respectively (Seviour et al. 2003; Skennerton et al. 2014). *ppk* encodes the polyphosphate kinase (PPK) enzyme that catalyzes the formation of intracellular polyP chains from phosphate molecules (Ahn and Kornberg 1990; Akiyama et al. 1992; Zhang et al. 2002; Rao et al. 2009; Kawakoshi et al. 2012). There are two known *ppk* genes: *ppk1* and *ppk2*. They have similar functions albeit the protein coded by *ppk2* (PPK2) catalyzes the making of polyP using both GTP and ATP and has a higher affinity for Mn^{2+}, whereas PPK1 has a higher affinity for Mg^{2+} (Zhang et al. 2002; Rao et al. 2009). Organisms may have either both or one *ppk* (Zhang et al. 2002; Rao et al. 2009). *ppx* regulates the exopolyphosphatase (PPX) enzyme that preferentially removes phosphate from the terminal ends of a polyP chain (Akiyama et al. 1993; Keasling et al. 1993; Zago et al. 1999; Rangarajan et al. 2006; Rao et al. 2009). A second *ppx*, known as *ppx/gppA* regulates the PPX enzyme known as guanosine pentaphosphate phosphohydrolase (PPX/GppA). PPX/GppA has a similar function to PPX and can also hydrolyze the stress response protein pppGpp to ppGpp (Keasling et al. 1993; Reizer et al. 1993; Zago et al. 1999; Rao et al. 2009).

Since the nucleotide sequences of *ppk* and *ppx* were first discovered, they have both been found in a variety of organisms (Rao et al. 2009; Alcántara et al. 2014). However, it is commonly understood that the ability of PAOs to synthesize large amounts of polyP under aerobic

conditions and use the energy from breaking down this stored polyP under anaerobic conditions allows PAOs to outcompete other heterotrophic organisms in specialized WWTPs (Gebremariam et al. 2011). *ppk* is especially well studied in model PAOs (e.g., He et al. 2007; Zhang et al. 2016) and is thought to be a good marker for strain diversity since it better conserved than 16S rRNA (Kunin et al. 2008). However, *ppx* has only been studied in a few cultured organisms (e.g., Zago et al. 1999; Alcántara et al. 2014) with the exception of one marine metagenome study (Temperton et al. 2011). Therefore, given the importance of these genes in polyP cycling, more research is needed to characterize both in environmental systems where our knowledge of microbially-mediated controls on P cycling is limited.

The goals of this experiment were to compare decomposition rates, P accumulation and loss rates, and the abundance of polyP functional genes (i.e., *ppk* and *ppx*) of stream and forest floor leaf litter samples over time. We hypothesized decomposition rates would be faster in the stream than on the forest floor and leaf P would increase over time on the stream litter and decrease on the forest floor. With respect to biotic mechanisms, which were the main focus of this study, we expected that the abundance of polyP functional genes would increase concurrently with leaf P over time because microorganisms with the ability to store P as polyP (i.e., PAO-like microorganisms) may colonize the leaves and increase leaf P over time. This is true especially in the stream where leaf surfaces provide an anchor and the stream itself proves a continuous supply of nutrients. We expected the same trends between all three genes because putative PAOs require pairs of these genes—either *ppk1* and *ppx* or *ppk2* and *ppx*—to carry out polyP synthesis and degradation. We also hypothesized that we would find microorganisms related to known as well as putative PAOs. To test these hypotheses, we monitored the decomposition rate and leaf P

concentrations of leaf litter samples placed in one of two treatments: (1) on the forest floor or (2) in the stream. We collected bags over a period of seven months and analyzed each for decomposition rate and leaf P concentration. We also extracted DNA from leaf samples and sequenced DNA using shotgun metagenomic analysis techniques. We then assembled sequences, determined relative abundances of *ppk1*, *ppk2*, and *ppx* genes, and assigned taxonomic identities to contigs with these same genes.

METHODS

Experimental Design and Leaf Litter Collection

The study site was located at the 303 ha Mianus River Gorge Preserve (MRGP) in Bedford, NY. The main portion of this property runs north to south along the Mianus River. The Mianus River Watershed has an area of 55 km^2 and drains to the Long Island Sound (Figure 4.1). The northern portion of the watershed, where the study site is located, is primarily forested with some development while the land is mostly developed in the south (Figure 4.1). Seasonal water quality samples taken from 2011 to 2013 at a site 25 m upstream from the study site had an average soluble reactive phosphorus (SRP-P) concentration of 0.034 ppm (n=9).

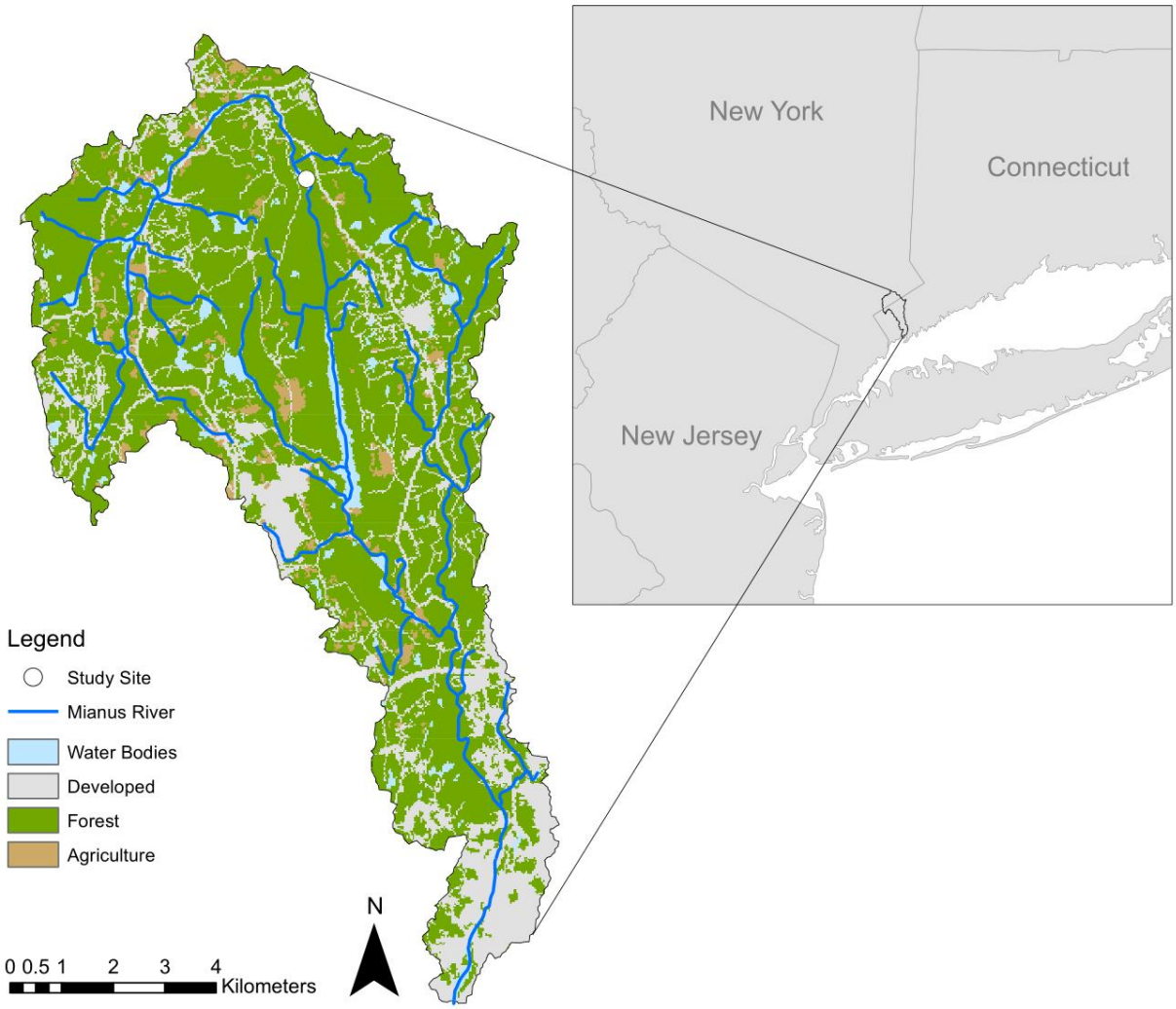

Figure 4. 1 Map of the Mianus River Watershed.

In preparation for this experiment, we cut sugar maple (*Acer saccharum*), black birch (*Betula lenta*), and American beech (*Fagus grandifolia*) leaves directly from the canopy to avoid contamination with soil microbes. We selected these tree species because they were found to be dominant riparian species in the MRGP (O'Brien et al. 2011). We brought the leaves back to the laboratory, air-dried them until dry to the touch, and dried them overnight in an oven at 60°C to ensure all moisture was removed before weighing. We weighed out approximately 1 g of each

leaf litter type and placed them in sealed plastic bags (total composite was about 3 g) until they were placed in the field on September 22, 2013. Upon placement in the field, we removed leaf litter from the plastic bags and put it into 7 x 8 inch, 500 μm pore-size, nylon mesh bags. We used fine mesh to focus mainly on microbial controls on leaf litter decomposition rather than the impact of macroinvertebrates (i.e., shredders); most adult macroinvertebrates are larger than 1 mm in size (Hauer and Lamberti 2006; Methot et al. 2012). We secured the mesh bags at a (1) forest floor site and (2) stream site. We placed the forest floor mesh bags on top of existing leaf litter and anchored stream mesh bags in a Mianus River run about 25 m downslope from the forest floor site. In this experiment, the forest floor site was meant as a control to the stream site, which would likely experience changes in oxygen availability over time. We randomly removed leaf litter bags (n=5 total replicates; n=3 for decomposition/leaf P and n=2 for DNA analysis) from both sites at approximately 0, 1, 3, 5, 7, 11, 26, and 30 weeks. We placed collected samples immediately placed in large air tight plastic bags and stored them in a 4°C and -20°C freezer until processing for decomposition/leaf P and DNA analyses, respectively. Throughout sample preparation, collection, and analysis we wore nitrile gloves to prevent microbial contamination and wiped down all equipment discussed henceforth between samples with 95% ethanol.

Leaf Litter Decomposition Analysis

We thawed stored samples at room temperature and processed them according to Meyer (1980) and Hauer and Lamberti (2006). Briefly, we removed leaf litter from the mesh bag over an 850 mm sieve and rinsed gently with distilled water to remove excess sediment. We then transferred leaves to aluminum tins and dried them in the oven for 24 hours at 105°C. We crushed dry leaves

in a food processor for approximately 30 seconds and weighted out 0.5 g into a pre-weighed aluminum tins. We put the tins into the muffle furnace at 550°C for 20 minutes, stirred them, and put them back into the muffle furnace at 550°C for 20 minutes. We allowed samples to cool and sprayed them down with distilled water to prevent loss of ash. We oven dried the ash at 60°C for 24 hours. After cooling, we weighed the ashed samples and used the difference in weight before and after combustion to calculate ash free dry mass (AFDM)—the mass of organic matter remaining—according to Hauer and Lamberti (2006).

Leaf Phosphorus Analysis

We weighed out 0.5 g of dry, crushed leaves into glass test tubes for ammonium persulfate digestion (USEPA, 1978). Briefly, 25 ml of distilled water was added to the leaf samples, followed by 1 ml of ammonium persulfate solution (0.2 g/ml concentration) and 0.5 ml of 30% H_2SO_4. Glass test tubes were autoclaved for 30 minutes at 121°C (15-20 psi). Extractions were left to cool, filtered through a 0.45um filter, and analyzed for SRP (as P) using the molybdenum blue method (USEPA, 1978) and auto-sampler (O.I. Analytical FS3000). The quantification limit of the FS3000 machine is 0.01 mg l^{-1} and all calibration curves had an R^2 value of 0.999 or higher.

DNA Extractions and Sequencing Preparation

We extracted DNA in triplicate from leaf litter samples using a MoBio PowerSoil DNA Isolation Kit (12888, Qiagen MoBio, Germantown, Maryland). Slight modifications to the protocol

included: (1) two, 30 s homogenizations with the bead beater (10 s rest in between each homogenization), (2) elution with 50 μL (rather than 100 μL), and (3) pooling of all three elutions into one tube. We quantified DNA on an Invitrogen Qubit spectrofluorometer using a Qubit dsDNA broad range assay kit (Q32850, Thermo Fisher Scientific, Waltham, Massachusetts). We sheered DNA to an average insert size of 550 bp using a Covaris S2 adaptive focused acoustic disruptor (Covaris, Woburn, Massachusetts). We prepared libraries using an Illumina TruSeq Nano DNA HT Sample Prep Kit (FC-121-4003, Illumina, San Diego, California). We carried out shotgun metagenomic sequencing with 1 pooled, indexed sample on two lanes at the Cornell University Institute for Biotechnology Genomics Facility using an Illumina HiSeq2500 instrument 2x150 bp paired-end reads on Rapid Run mode.

Metagenomics Analysis

We used the workflow described in Figure 3.1 to process shotgun metagenomic sequencing data. Briefly, we quality controlled and trimmed raw reads using BBMap/BBTools (JGI 2017), assembled reads for all samples in both treatments together with MEGAHIT using default settings (Li et al. 2015), called genes and translated gene sequences using Prodigal (Hyatt et al. 2010), annotated PPK1, PPK2, and PPX protein sequences using Pfam models with a HMMER reciprocal search using default settings (Finn et al. 2011), assigned taxonomy to contigs with genes of interest using Kaiju's protein-level NCBI RefSeq reference database classifier (Menzel et al. 2016), and mapped raw reads to determine relative gene abundances using bwa-mem (Li 2017). Relative gene abundances are given as the number of fragments (i.e., paired-end reads) with a gene of interest per kb per million reads mapped (FPKM). Protein family models from the Pfam database (Finn et al. 2016) used in this study to identify *ppk* and *ppx* genes are shown in

Table 3.2. For contigs with unique combinations of genes, we used pBLAST (NCBI 2017) to compare contig protein sequences to previously studied organisms. We selected the top three contigs by maximum relative abundance for each sample as well as the top seven contigs by mean relative abundance overall and display the z-score of these data in heatmaps. We also categorized each contig for whether it contained various combinations of *ppk1*, *ppk2*, and *ppx/gppA* (referred to henceforth as *ppx*). The data and scripts pertaining to this workflow are available on GitHub at https://github.com/sheilasaia/paper-p-leaf-litter.

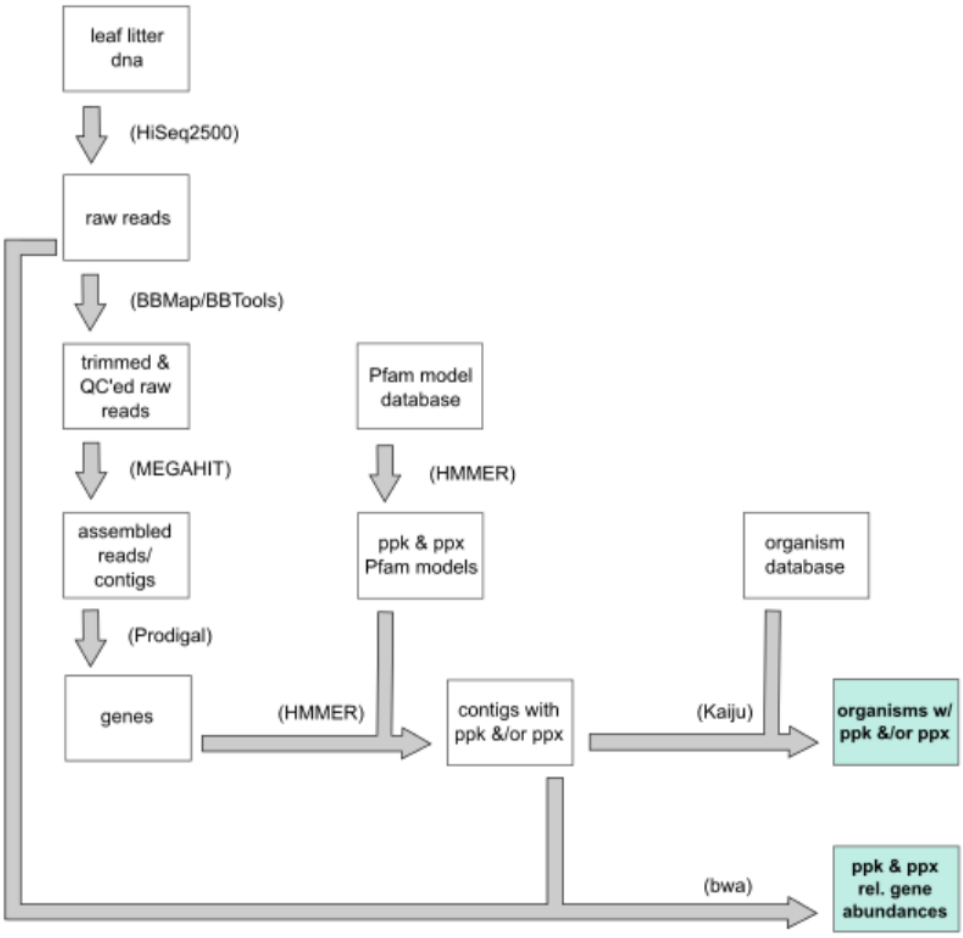

Figure 4. 2 Description of shotgun metagenomic sequencing data workflow used in this study. Final output tables used in subsequent analyses are bolded/shaded green.

Table 4. 1 Protein family (Pfam) database models used in this study.

Pfam ID	Protein	Pfam Description
PF02503.16	PPK1	polyphosphate kinase middle domain (PPK1)
PF13089.5	PPK1	polyphosphate kinase N-terminal domain (PPK1)
PF13090.5	PPK1	polyphosphate kinase C-terminal domain (PPK1)
PF03976.13	PPK2	polyphosphate kinase (PPK2)
PF02541.15	PPX	exopolyphosphatase (PPX/GppA)

Statistical Analysis

We conducted all statistical analysis in R (R Core Team 2015). We used a linear model (LM) as well as a generalized additive model (GAM) to test whether AFDM as well as leaf P concentrations for the two treatments differed over time. We applied the GAM using the *mgcv* package (Wood 2006, 2015, Zuur 2009). The general form of the GAM is given in Equation 1 as:

$$Y \sim s(t, \text{by stream}) + s(t, \text{by forest floor}) + \varepsilon \qquad \text{(Equation 1)}$$

where Y represents the response variable (i.e., AFDM or leaf P), t represents time, the function s( ) represents the additive response smoothed over time conditioned on either leaf litter placed in the stream or leaf litter placed on the forest floor, and ε is the mean zero Gaussian random error. For all statistical tests, we used $\alpha = 0.05$ to determine statistical significance and compared the quality of LMs and GAMs using the Akaike information criteria (AIC; Akaike 1974). A lower AIC value equates to a better model fit and AIC scores greater than two units apart are statistically different.

RESULTS & DISCUSSION

Leaf Litter Decomposition

We used LMs and GAMs to test for a relationship between leaf litter decomposition (i.e., decreasing AFDM) over time for the forest floor and stream treatments (Table 4.2). AIC scores indicate that Model 2.3 and Model 2.5 both describe the data best and are not significantly different from one another because they have equal AIC scores (Table 4.2). Therefore, we proceed in our analysis using the simpler LM (Model 2.3). Summary statistics of Model 2.3 are shown in Table 4.3 and plotted in Figure 4.3. LM results indicate a significant interaction term between the two treatments (Table 4.2), which provides support for their divergent decomposition rates (i.e., the slope of the LM between AFDM and time in Figure 4.3). The slope of the stream treatment is more negative than the slope of the forest floor treatment (Figure 4.3). When plotted, Model 2.5 (i.e., the GAM) results are similar to those of Model 2.3 (data not shown). The Model 2.5 smooth term for the forest floor leaf litter samples over time is not significantly different from a null model with zero slope ($p=0.8410$) but the smooth for the stream leaf litter samples is significant and negative ($p=< 1.00E-05$; Table 4.4). Model 2.5 explains 85.5% of the deviance in observations (Table 4.4). Therefore, we see significant decomposition rates in leaf litter left in the stream but not on the forest floor.

Table 4. 2 Linear model (LM) and generalized additive model (GAM) comparisons for ash free dry mass (AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC).

Model	Model	Model	df	AIC
2.1	LM	AFDM ~ Time	3	-154.4
2.2	LM	AFDM ~ Time + Treatment	4	-174.0
2.3	**LM**	**AFDM ~ Time + Treatment + Time: Treatment**	**5**	**-231.9**
2.4	GAM	AFDM ~ s(Time)	3	-154.4
2.5	**GAM**	**AFDM ~ s(Time, by Forest Floor) + s(Time, by**	**5**	**-231.9**

Table 4. 3 Best fit linear model (LM) outputs (Model 2.3 in Table 4.2) for ash free dry mass (AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: standard error (SE).

Model Term	Value	SE	t-value	p-value
Intercept*	4.71E-01	7.15E-03	68.85	< 1.00E-05
Time	-1.45E-05	7.20E-05	-0.20	0.8410
Treatment (Stream)	3.30E-03	1.01E-02	0.33	0.7440
Time:Treatment (Stream)	-9.90E-04	9.87E-05	-10.03	< 1.00E-05

* Base model represents forest floor treatment.

Table 4. 4 Best fit general additive model (GAM) outputs (Model 2.5 in Table 4.2) for ash free dry mass (AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: standard error (SE). The percent deviance explained by this GAM was 85.5%.

Model Term	Value	SE	t-value	p-value	F
Intercept	0.47	0.007	65.85	< 1.00E-05	
Forest Floor smooth term				0.841	0.04
Stream smooth term				< 1.00E-05	139.62

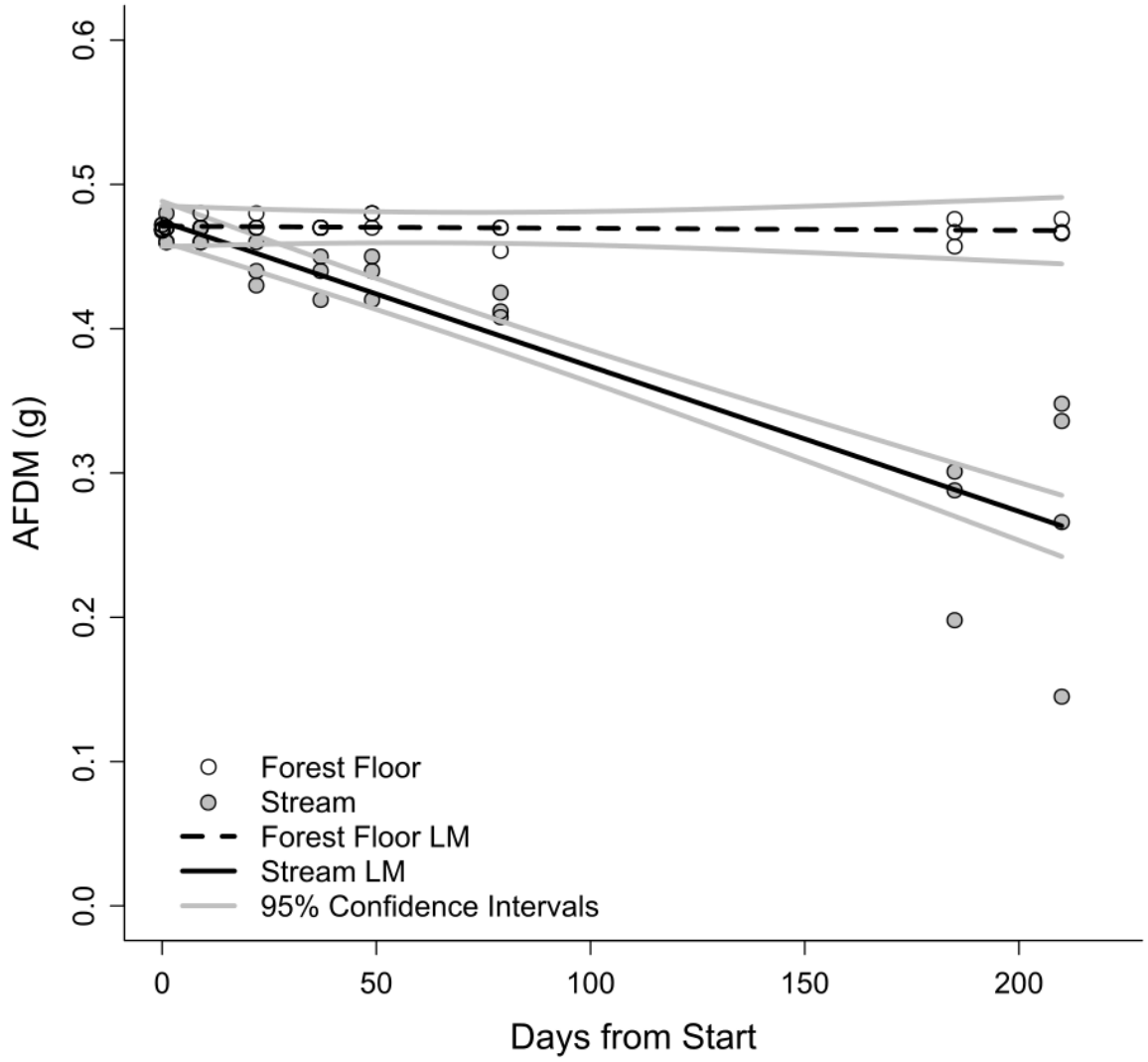

Figure 4. 3 Observations and linear model (LM) results of ash free dry mass (AFDM) versus days from start of experiment for leaf litter bags placed on the forest floor (F) and in the stream (S) treatments.

Stream site decomposition rates in this study are within the range cited in other works using similar types of leaf litter (Meyer 1980; Gulis and Suberkropp 2003). The decomposition of forest floor leaf litter in this study is much less than other studies (e.g., Cotrufo et al. 1995). There are several reasons why our results may have differed from previous studies. Focusing

first on leaf litter decomposition in streams, research conducted by Meyer (1980) took place in the undeveloped, forested Bear Brook Watershed. Bear Brook is a sub-watershed of the Hubbard Brook Watershed in New Hampshire (Fisher and Likens 1973) and studies around the same time determined that SRP concentrations in Bear Brook ranged from 0.0001 to 0.001 ppm in the summer and winter (Meyer and Likens 1979). Seasonal water quality samples collected from the Mianus River, 25 m upstream of our study site, indicate stream SRP concentrations ranged from 0.018 to 0.080 ppm (average 0.034 ppm for n=9 replicates). Therefore, P availability is higher in the Mianus River compared to Bear Brook. This is not surprising because, despite being mostly forested, the watershed just upstream of our study site is somewhat developed and may provide a source of P to the stream. Other studies support our assertion that increases in nutrient availability lead to higher rates of leaf litter decomposition in streams (Elwood et al. 1981; Gulis and Suberkropp 2003; Martínez et al. 2014). Like streams, differences in nutrient availability have been shown to influence leaf litter decomposition rates of leaf litter in forest soils (Cotrufo et al. 1995).

Second, we used a smaller leaf litter bag mesh size compared to other leaf litter studies (Meyer 1980; Cotrufo et al. 1995; Sponseller and Benfield 2001; Martínez et al. 2016) because the focus of this work was on microorganisms. This may have (1) decreased the impact of erosive processes (Meyer 1980) and (2) excluded stream invertebrates (Hauer and Lamberti 2006). Excluding these larger (>500 x 500 um) soil organisms (e.g. nematodes, anthropoids, and mollusks) likely reduced the rate of further leaf litter break down after initial microbial colonization (Webster and Benfield 1986). Finally, other environmental factors such as stream temperature and pH, litter quality, and flow velocity may have led to differences between our

findings and the findings of other studies (Cummins 1974; Federle et al. 1982; Chamier 1987; Friberg et al. 2009; Martínez et al. 2014; Martínez et al. 2016). With respect to the forest floor treatment, differences in climate (i.e., temperature and moisture) have been shown to influence leaf litter decomposition rates in soils (Gholz et al. 2000; Gonzalez and Seastedt 2001).

These results generally support our hypothesis that decomposition of leaf litter occurs faster in the stream than on the forest floor and we suggest a combination of abiotic and biotic mechanisms influenced these differences. Specifically, the consistent mechanical stress of water flowing past the leaves placed in the stream accelerated leaf litter decomposition. Precipitation through-fall or overland flow events may provide similar mechanical stress and nutrient supplies but both are likely limited due to the dense cover of primarily deciduous trees at this study site. Additionally, stream water—compared to likely intermittent overland flow—may have provided a consistent nutrient supply to microorganisms colonizing and decomposing the stream leaf litter.

Leaf Phosphorus

Model comparisons (Table 4.5) indicate the GAM—Model 5.5—had the lowest AIC score, and therefore, could better predict leaf P concentrations over time for both treatments compared to the other LMs and GAMs tested. Model 5.5 summaries indicate that both smooth terms were significant ($p<0.05$) and the percent deviance explained was 70.2% (Table 4.6). Overlap in 95% confidence intervals at ~50 days from the start of the experiment indicates that the two treatments were not significantly different during this period (Figure 4.4). However, the lack of

overlap in 95% confidence intervals during the rest of the experiment demonstrates the overall significant difference between forest floor and stream leaf litter bag treatments (Figure 4.4).

Table 4. 5 Linear model (LM) and generalized additive model (GAM) comparisons for leaf phosphorus (leaf P; ppm P/g of AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: degrees of freedom (df), Aikaike information criteria (AIC).

Model	Model	Model	df	AIC
5.1	LM	Leaf P ~ Time	3	92.0
5.2	LM	Leaf P ~ Time + Treatment	4	93.4
5.3	LM	Leaf P ~ Time + Treatment + Time: Treatment	5	60.6
5.4	GAM	Leaf P ~ s(Time)	4	91.9
5.5	**GAM**	**Leaf P ~ s(Time, by Forest Floor) + s(Time, by**	**10**	**43.3**

Table 4. 6 Best fit general additive model (GAM) output (Model 5.5 in Table 4.5) for leaf phosphorus (leaf P; ppm P/g AFDM) versus time and treatment (i.e., leaf litter placed either on the forest floor or in the stream). Abbreviations: standard error (SE). The percent deviance explained by this GAM was 70.2%.

Model Term	Value	SE	t-value	p-value	F
Intercept	1.65	0.08	19.99	< 1.00E-05	
Forest Floor smooth term				2.65E-04	9.32
Stream smooth term				< 1.00E-05	13.13

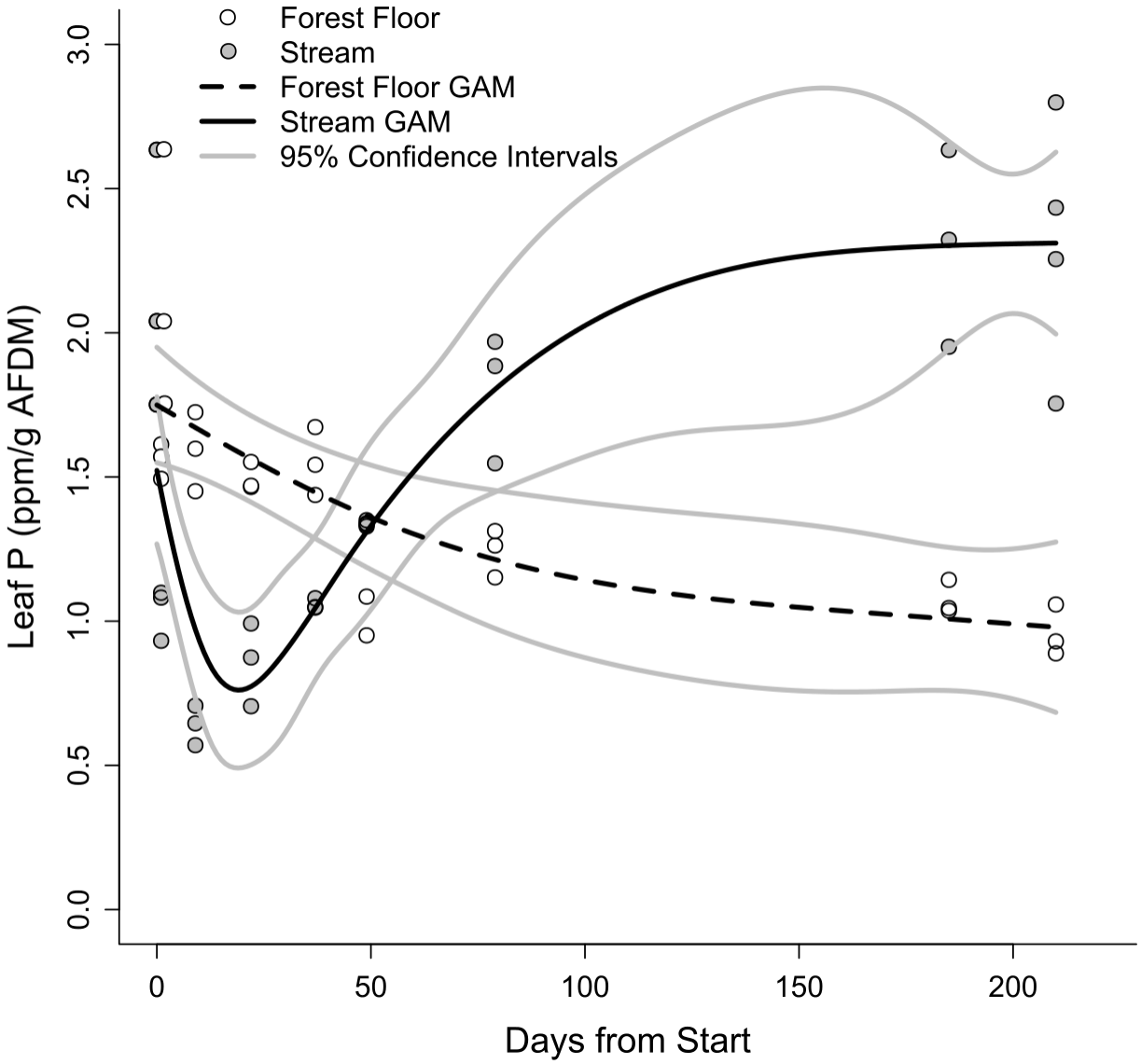

Figure 4. 4 Observations and generalized additive model (GAM) results of leaf P concentration (ppm/g AFDM) versus days from start of experiment for leaf litter bags placed on the forest floor and in the stream treatments.

Model results demonstrate the consistent wash-off/leaching of leaf P from leaf litter collected from the forest floor over the entire experiment. In contrast, leaf litter collected from the stream shows a more dynamic pattern in leaf P that can be summarized into two main stages: (1) initial

wash-off/leaching of P followed by an (2) accumulation of leaf P (Figure 4.4). The initial loss of leaf P has been previously documented (Meyer 1980) and may be due to the washing off of various salts that may have condensed on the leaf surface (Burkhardt et al. 1999) and/or atmospheric sources of P (Tamatamah et al. 2005; Camarero and Catalan 2012; Tipping et al. 2014) accumulated on the surface of the leaves. This flush may have also been due to lysis of bacteria cells upon rewetting with distilled water (Turner et al. 2003), the composition of the leaf; leaf litter chemistry has been shown to influence decomposition (Soong et al. 2015) and impact subsequent P loss (Martínez et al. 2016).

To test our assertion concerning the rapid wash-off/leaching of P from leaf litter, we placed an equal mass (air-dried) of each of the three leaf species from this experiment in air-tight plastic bags, soaked these leaves in equal volumes of distilled water for ~1 minute, 8 hours, and 24 hours, and analyzed the rinse water soluble reactive phosphorus (SRP) and leaf P concentrations as described in the methods section. Results of this laboratory study support our assertion; we see a rapid loss of P into the rinse water from the leaves that eventually levels out (Figure 4.5A). We also see a concurrent rapid decline in leaf P (Figure 4.5B). We compared the leaf P results of these three soaking periods and using ANOVA and post-hoc Tukey pair-wise comparisons and found that the ~1 minute soaking treatment was significantly different from the other treatments ($p<0.05$) and the 8 hour soaking treatment was not significantly different from the 24 hour treatment ($p=0.2565$). We were surprised that the concentration of SRP was ~ 1ppm after only ~1 minute of rinsing the leaf litter. This finding provides support for our wash-off assertion made previously. These laboratory results are consistent with research showing a rapid release of soluble nutrients from the leaves within 1-2 days of immersion in the stream as well as evidence

that the remaining soluble nutrients continue to leach out slowly after this initial period (Hauer and Lamberti 2006; Ball et al. 2009).

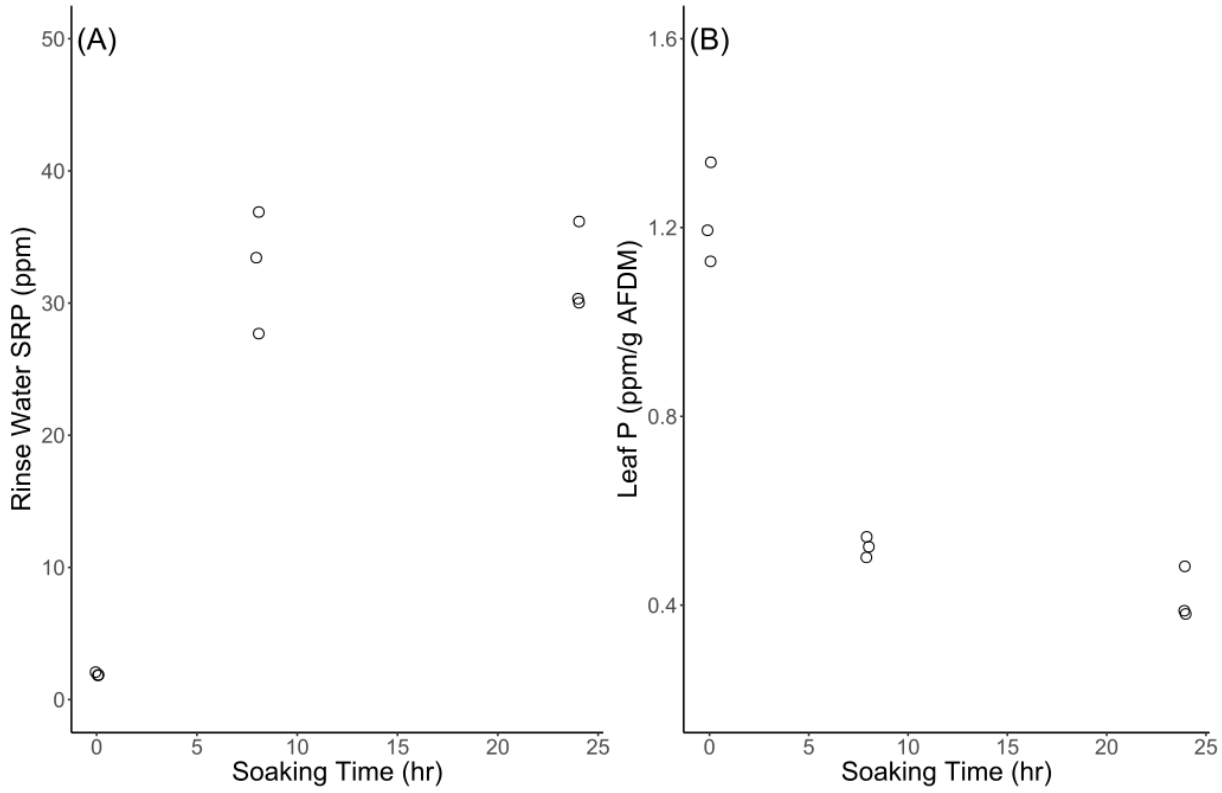

Figure 4. 5 (A) Rinse water soluble reactive phosphorus (SRP) concentration and (B) leaf P concentration versus soaking time for laboratory wash-off experiment. Samples at ~ 1 min (~0 hours) were rinsed with water and immediately drained.

With regard to the second stage in stream leaf P dynamics, the accumulation of P on leaves in streams over time has been observed by other researchers (Meyer 1980; Mehring et al. 2015). Sediment was present on the leaf litter overtime and gently washing leaves before leaf P analysis should have removed excess sediment. However, given the methods used in this study, we cannot discount the accumulation of sediment on stream leaf litter over time. Future studies may

consider using epi-fluorescence microscopy to determine microbial abundances over time (Gulis and Suberkropp 2003) and the percentage of cells on the leaf litter with intracellular polyP granules (Rier et al. 2016; Saia et al. 2017).

Shotgun Metagenomics Assembly

For the leaf litter dataset, we assembled a total of 3,090,663 contigs, the longest of which was 521.4 kb. There were 160 contigs greater than 50 kb in NY and they made up 0.6% of the main genome. A summary of the total assembly length for contigs above a certain length is shown in Table 4.7.

Table 4. 7 Shotgun metagenome assembly results for leaf litter samples included in this study (both forest floor and stream treatments were assembled together).

Minimum Contig Length	Number of Contigs
500	1,423,206
1,000	460,352
10,000	6,522

Microorganisms Harboring Phosphorus Functional Genes

We annotated a total of 2,484 and 2,919 genes for leaf litter samples collected from the forest floor and stream, respectively (Table 4.8). Because some contigs had more than one gene, they represented a total of 2,384 and 2,819 unique contigs for stream and forest floor treatments, respectively (summation of all bars by location, Figure 4.6). We suspect that the reason for

slightly more unique genes in the stream treatment is the potential for stream water to transport new organisms along with it. However, it is surprising that despite the physical (albeit hydrological connected) differences between a stream and an upslope forest floor, we assembled metagenomes with matching contigs.

Table 4. 8 Number of *ppk1*, *ppk2*, and *ppx* genes annotated in this study for leaf litter samples taken from on the forest floor and in the stream.

Location	Gene	Count
Forest Floor		
	ppk1	1,076
	ppk2	675
	ppx	733
	Total	2,484
Stream		
	ppk1	1,272
	ppk2	823
	ppx	824
	Total	2,919

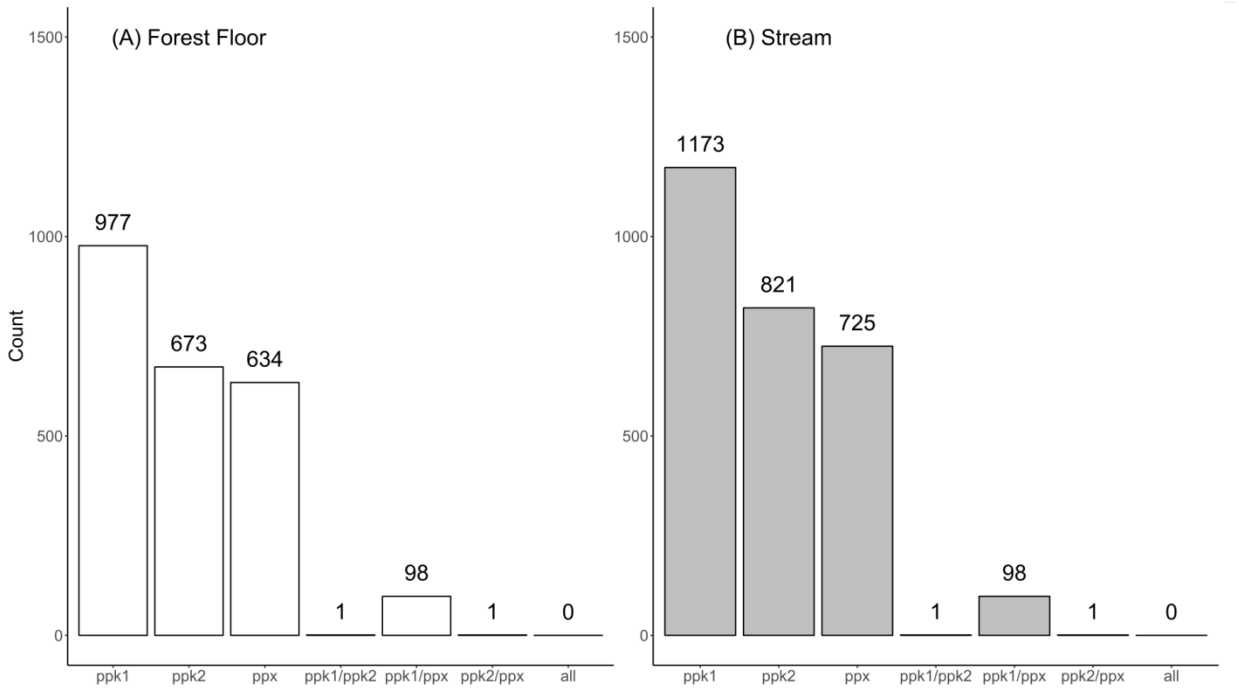

Figure 4. 6 Number of unique contigs with various gene combinations in samples from (A) forest floor and (B) stream. Note: the total number of samples sequenced is the same for both treatments (n=16).

For both locations, *ppk1* genes were most frequently harbored by contigs followed by *ppk2* and then *ppx* genes annotated (Figure 4.6). We also saw the same number of contigs that harbored both *ppk1* and *ppx* genes (Figure 4.6). Rarely, did contigs harbor both *ppk2* and *ppx* and we found no contigs in both treatments that harbored all three genes (Figure 4.6). The observation of contigs harboring only one polyP functional gene was unexpected because microorganisms would theoretically need either *ppk1* or *ppk2* to synthesize polyP as well as *ppx* to break down polyP. By comparing the length distribution of contigs harboring only one polyP functional gene to those harboring *ppk1* and *ppx* appears to be unbiased towards longer contigs below 7,500 bp

but there are several contigs harboring both *ppk1* and *ppx* that are longer than 7,500 bp (Figures 4.7C and 4.7D) This indicates that deeper sequencing may be necessary to ensure thorough classification of putative PAOs from non-PAOs. Because these genomes are incomplete, it is difficult for us to say that contigs harboring only one polyP functional gene are not putative PAOs. Another possible explanation for contigs with only one polyP functional gene is that *ppk* and *ppx* genes do not always reside near each other in a genome. While not recognized as PAOs, *E. coli* have been shown to have a *ppk* operon where *ppk* was located upstream of *ppx* (Brown and Kornberg 2008). Conversely, another study of P. aeruginosa does not have a *ppk* operon as *ppx* is upstream of *ppk* (Brown and Kornberg 2008). Therefore, further research is needed to determine the organization of these genes in putative PAO genomes.

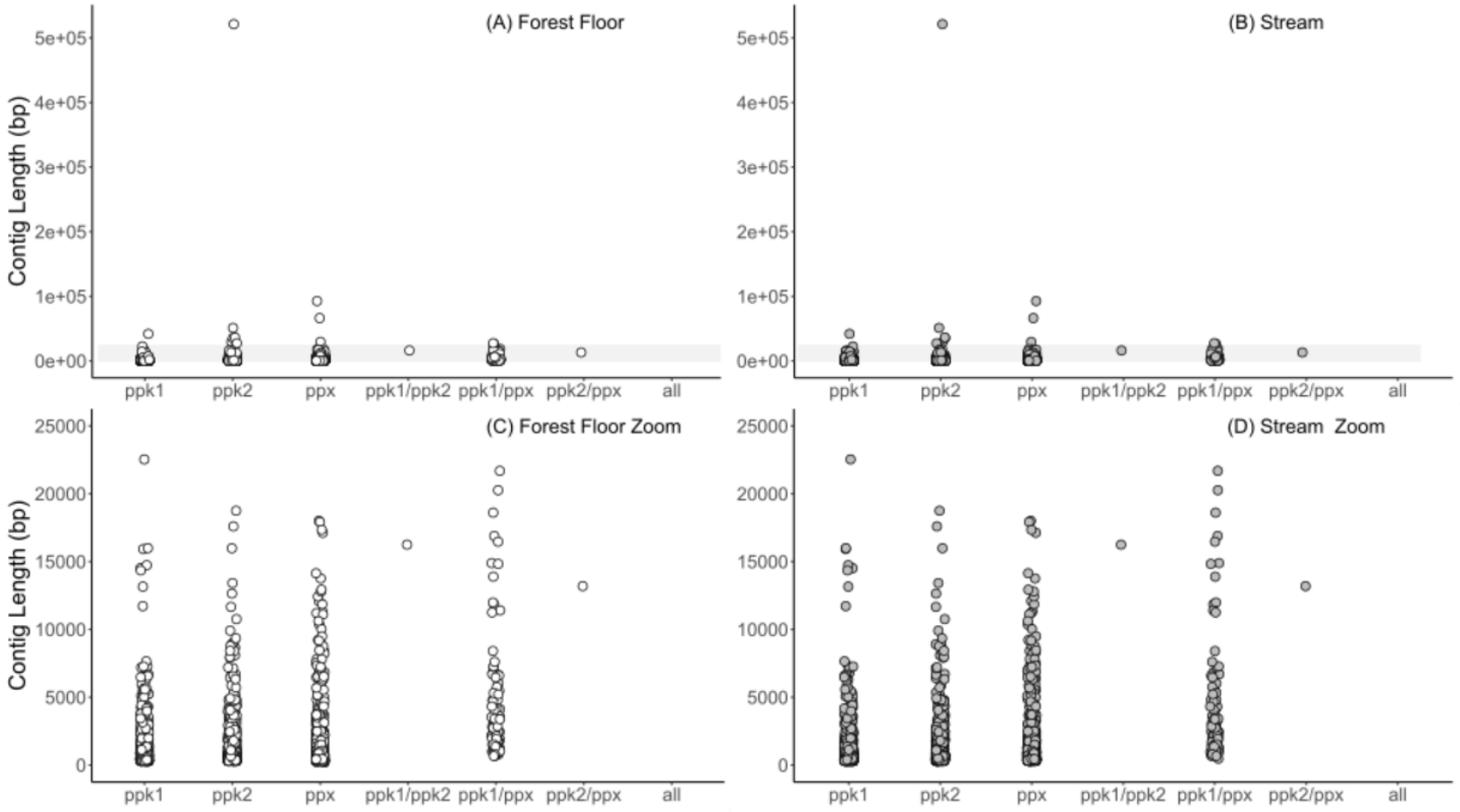

Figure 4. 7 Length distributions of contigs harboring various gene combinations from samples collected from the stream (A and C) and on the forest floor (B and D). Note the difference in scale between A and B. The shaded regions in plots A and B refer to the zoomed-in (0-25,000 bp) plots shown in C and D.

In terms of taxonomy, the distribution of contigs from both locations is similar with only differences in magnitude (Figure 4.8). *Proteobacteria* (*Alphaproteobacteria* and *Gammaproteobacteria)* are the top two most represented bacterial phyla represented in leaf litter from both treatments (Figure 4.8). We also see many contig assignments to the *Betaproteobacteria* class and the phyla of *Actinobacteria*, which are associated with known PAOs (Hesselmann et al. 1999; Nakamura et al. 1995a; Nakamura et al. 1995b; Nguyen et al. 2011). Some contigs for leaf litter samples here were assigned to the genus *Burkholderia* (*Betaproteobacteria*), which is known to solubilize phosphate and dominated the first stages of decomposition of forest floor leaf litter bacterial communities (Purahong et al. 2016). *Burkholderia* has also been found to accumulate intracellular polyP in stream biofilms (Locke 2015).

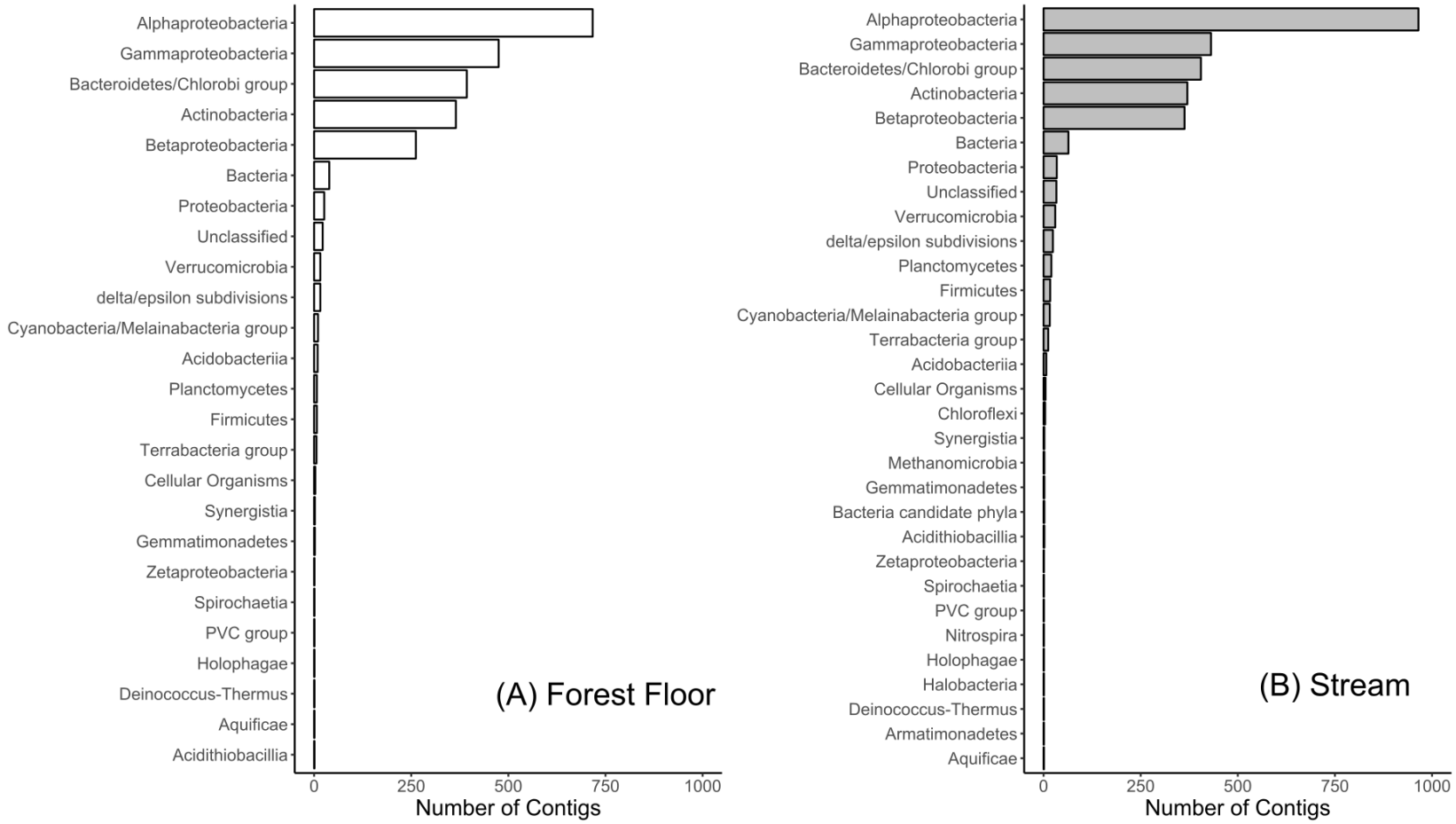

Figure 4. 8 Number of unique contigs classified into major bacterial and archaeal phyla/classes for leaf litter samples collected from the (A) forest floor and (B) stream.

Contigs that harbor either *ppk1*, *ppk2*, or *ppx* were consistently and most frequently assigned to the *Alphaproteobacteria* class (Figure 4.9). Other frequent assignments to contigs with only one of the three genes include *Bacteroidetes/Chlorobi*, *Gammaproteobacteria*, *Actinobacteria*, and *Betaproteobacteria* (Figure 4.9). More contigs are assigned to the *Betaproteobacteria* class in the stream treatment compared to the forest floor (Figure 4.9). Compared to the stream, forest floor samples have more contigs harboring only *ppk1* that were assigned to *Gammaproteobacteria* (Figures 4.9A and 4.9D). We observed contigs assigned to the phyla *Verrucomicrobia*, which are abundant in soil (Bergmann et al. 2011) and positively correlated with total P concentrations in sediments (Jin et al. 2017), contigs assigned to archaeal lineages (*Methanomicrobia*), and contigs assigned to the *Cyanobacteria/Melainabacteria* group, the latter of which was identified in full-scale EBPR reactor (Soo et al. 2014). Contigs that may potentially represent known and new PAOs because they harbor both *ppk1* and *ppx* genes are very similar between treatments (Figure 4.10), except for those classified as *Bacteroidetes/Chlorobi*, represent classes/phyla of known PAOs as mentioned previously. We found no contigs that harbored all three genes, one (k141_1220833) that harbored both *ppk1* and *ppk2*, and one (k141_1114903) that harbored both *ppk2* and *ppx* (Table 4.9). k141_1220833 and k141_1114903 were found in both treatments. Both protein sequences from contig k141_1114903 were assigned to *Sphingomonas* sp. KC8 (Table 4.9).

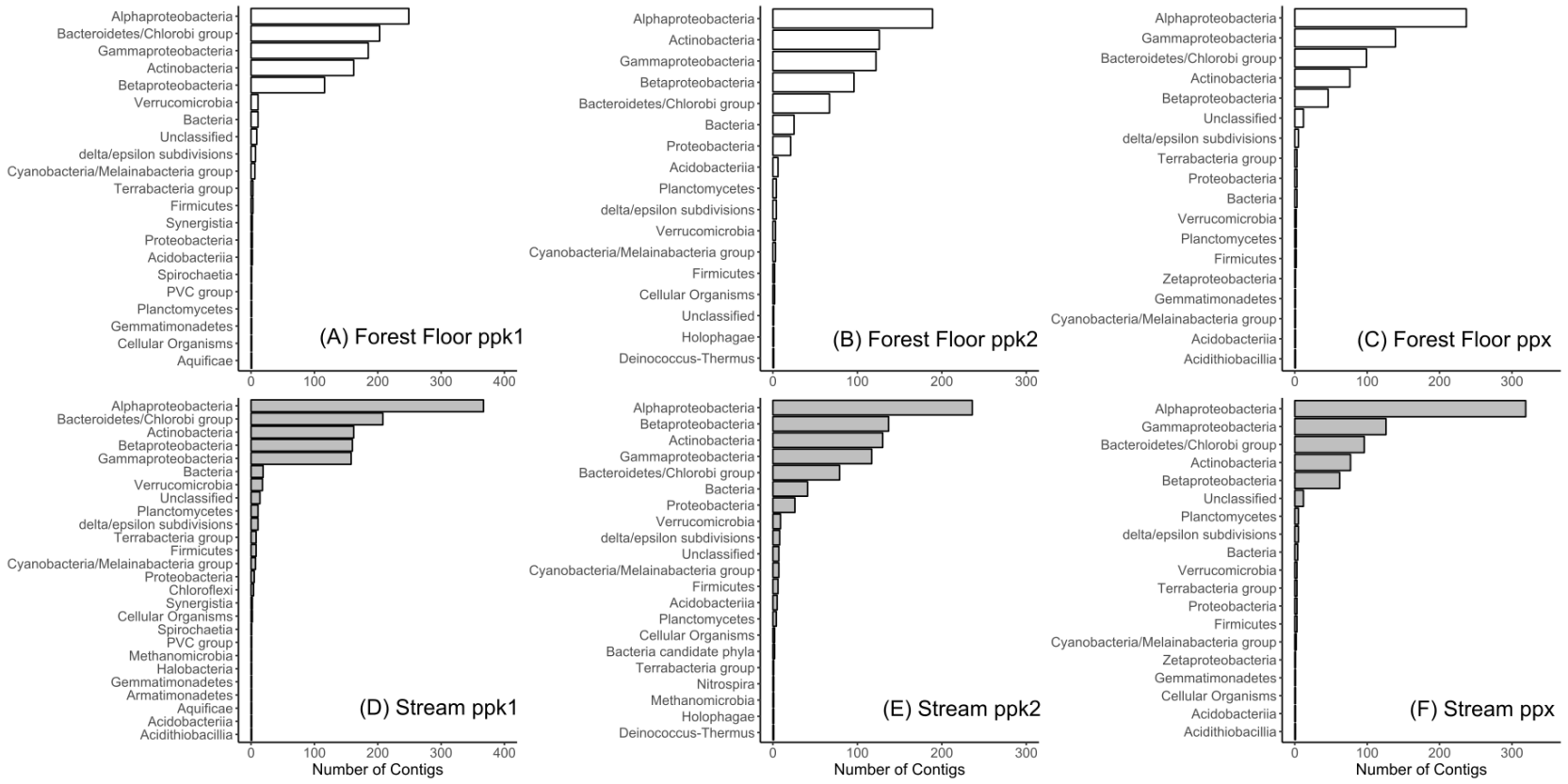

Figure 4. 9 Number of unique contigs classified into major bacterial and archaeal phyla/classes for leaf litter samples collected on the (A) forest floor (A-C) and in the stream (D-E) harboring only one of either *ppk1* (A and D), *ppk2* (B and E), and *ppx* (C and F) genes.

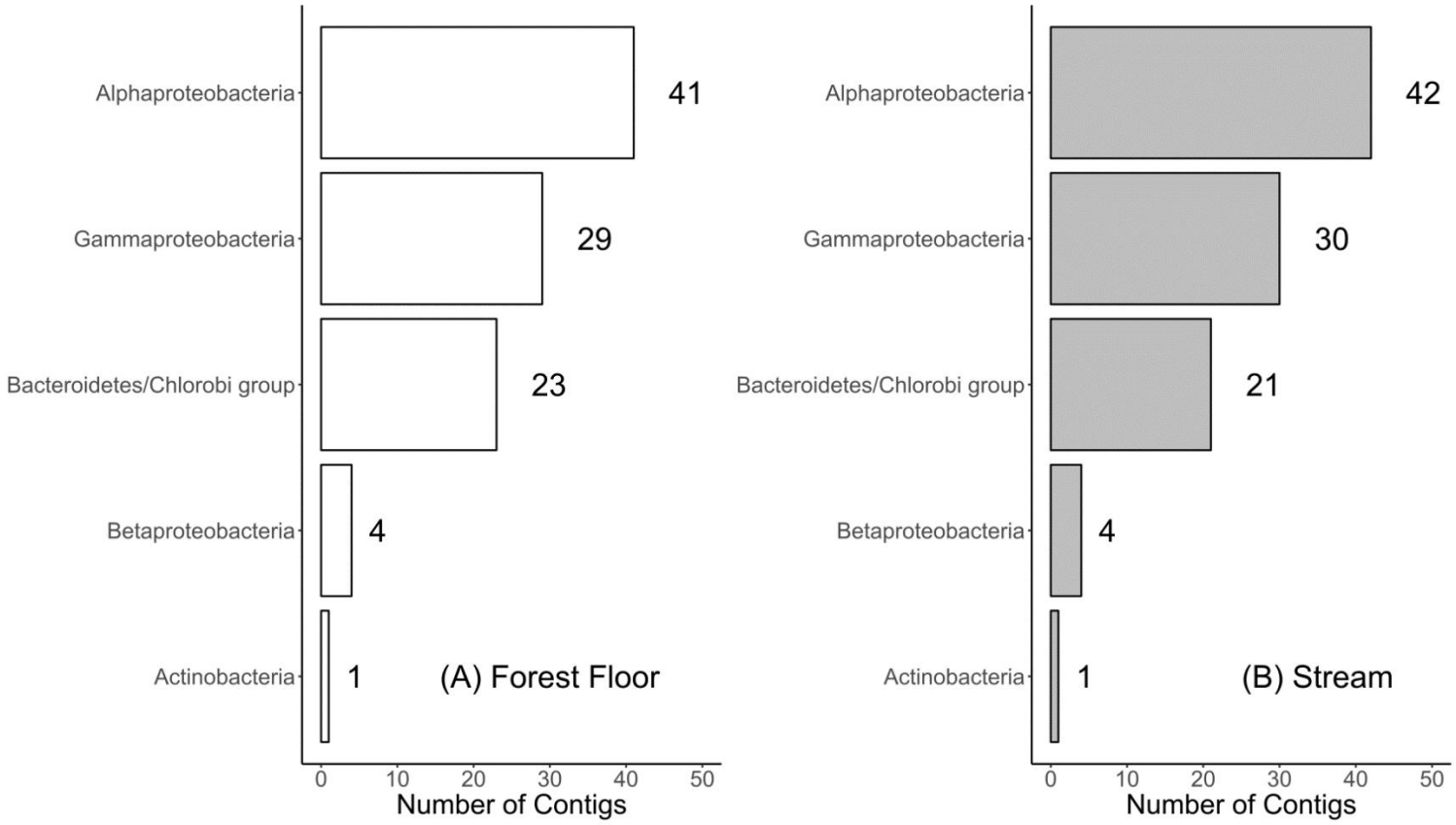

Figure 4.10 Number of unique contigs harboring both *ppk1* and *ppx* genes classified into major bacterial and archaeal phyla/classes for leaf litter samples collected from the (A) forest floor and (B) stream. Number of unique contigs per phylum/class are shown to the right of each bar.

Table 4. 9 Contigs with unique combinations of *ppk* and *ppx* genes from leaf litter samples collected from the forest floor and in the stream as determined by shotgun metagenome analysis and subsequent comparison of protein sequence to existing sequences in the pBLAST database. Abbreviations: not given (NG).

Contig ID	Gene	Top pBLAST Hit Protein Name	Top pBLAST Hit Organism	Accession Number	Identity (%)	Environment	Reference
k141_1220833							
	ppk1	polyphosphate kinase	*Ohtaekwangia koreensis*	SKC82618.1	90	NG	NG
	ppk2	polyphosphate-nucleotide phosphotransferase	Bacterium 336/3	KOY85390.1	74	NG	NG
k141_1114903							
	ppk2	hypothetical protein	*Sphingomonas* sp, KC8	WP_010125262.1	77	NG	NG
	ppx	exopolphosphatase	*Sphingomonas* sp. KC8	WP_010125257.1	69	NG	NG

Overall, these results demonstrate broader similarities between the two treatments in terms of the number of *ppk1*, *ppk2*, and *ppx* annotated as well as the type of microorganisms classified. Of the major classes and phyla represented in this dataset, many overlap with known PAOs (e.g., *Betaproteobacteria)* while others require further study (e.g., *Burkholderia* and *Verrucomicrobia*). We recognize these are diverse classes and recommend the need for further research verifying the phenotype of putative PAOs identified here. Unexpectedly, leaf litter collected from the stream had similar unique contigs harboring both *ppk1* and *ppx* compared to leaf litter collected on the forest floor, which was hydrologically up-gradient of the stream site. This is counter to a study demonstrating limited dispersal of hydrologically connected environments (Graham et al. 2017). However, hydrology and nutrient availability are important controls on microbial community structure and function (Pett-Ridge and Firestone 2005; DeAngelis et al. 2010; Ouyang and Li 2013; Peralta et al. 2014; Graham et al. 2017).

Temporal Patterns in Functional Gene Relative Abundance

Using GAMs, we found that adding in treatment (i.e., whether samples were collected from the forest floor or the stream) improved our ability to predict the relative abundance of *ppk2* and *ppx* (Table 4.10). In the case of the relative abundance of *ppk1*, Models 10.1 and 10.2 (Table 4.10) were not statistically different but we precede using Model 10.2 to further demonstrate the lack of difference between the two treatments. Namely, Model 10.2 summary results indicate that both the forest floor and stream smooth terms are not significantly different from the null model (i.e., a model with no trend) of *ppk1* relative abundance versus time ($p>0.05$; Table 4.11). Further, there is a nearly complete overlap of confidence intervals for both treatments throughout the duration of this experiment; meaning, treatments are not statistically different (Figure 4.11A).

When modeling *ppk2* relative abundance, the forest floor smooth term of Model 10.4 is significantly different from the null model of *ppk2* relative abundance versus time ($p < 1.00E-05$; Table 4.11) but the stream smooth term is not ($p=0.084$). Last, the best fit model of *ppx* relative abundance over time (Model 10.6 in Table 4.10) demonstrates a trend that is statistically different from the null model for leaf litter samples collected from the forest floor ($p=0.008$; Table 4.11) and not statistically different from a null model for samples from the stream ($p=0.676$; Table 4.11). However, the confidence intervals of the two treatments are similar except for periods at the start of this experiment and around ~50 days. Therefore, with respect to *ppx*, treatment only has a weak influence.

Table 4.10 Relative abundance (RA; in units of FPKM) generalized additive model (GAM) comparisons versus time and treatment (i.e., whether leaf litter was place in the stream or on the forest floor) for *ppk1*, *ppk2*, and *ppx* genes. fit models for each soil extraction are bolded. Abbreviations: degrees of freedom (df) and Aikaike information criteria (AIC).

Model Number	Model	df	AIC
ppk1			
10.1	RA ~ Time	4	434.8
10.2	**RA ~ Time + Treatment***	**6**	**436.3**
ppk2			
10.3	RA ~ Time	3	412.0
10.4	**RA ~ Time + Treatment**	**5**	**379.4**
ppx			
10.5	RA ~ Time	5	387.6
10.6	**RA ~ Time + Treatment**	**8**	**379.4**

* This model is not statistical different from Model 10.1 was demonstrate the treatments.

Table 4.11 Best fit generalized additive model (GAM) summaries (Models 10.2, 10.4, and 10.8 in Table 4.10) for predicting relative abundance (RA; in units of FPKM) of *ppk1*, *ppk2*, and *ppx* genes. Abbreviations: standard error (SE).

Model Term	Value	SE	t-value	p-value	F
ppk1					
Intercept	503.80	73.00	6.90	< 1.00E-05	
Forest Floor smooth term				0.133	2.06
Stream smooth term				0.654	0.21
ppk2					
Intercept	395.73	30.34	13.04	< 1.00E-05	
Forest Floor smooth term				< 1.00E-05	23.18
Stream smooth term				0.084	3.22
ppx					
Intercept	361.90	29.12	12.43	< 1.00E-05	
Forest Floor smooth term				0.008	3.91
Stream smooth term				0.676	0.18

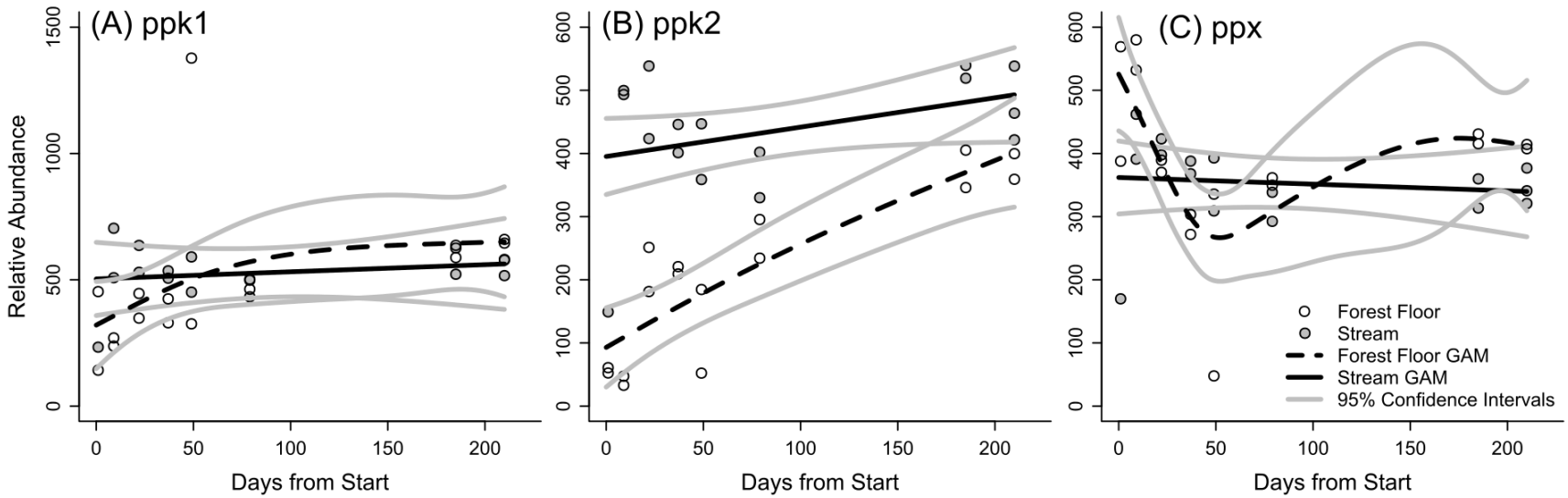

Figure 4. 1 1 Observations and generalized additive model (GAM) results (Models 8.2, 8.4, and 8.6 in Tables 4.8 and 4.9) of relative abundance (FPKM) of (A) *ppk1*, (B) *ppk2*, and (C) *ppx* genes over time for leaf litter bags collected from the forest floor and stream.

These results are not consistent with our hypothesis that the relative abundance of polyP functional genes increases proportionally with increases in leaf P. We observed an increase in stream leaf P concentrations over the course of this experiment (Figure 4.4) but the relative abundances of *ppk1*, *ppk2*, and *ppx* in this treatment stayed the same (Figure 4.11). Therefore, we suggest that the build-up of leaf P in the stream was largely sourced from sediment-bound P remained attached to the leaves, or potentially trapped in a biofilm on the leaf surface, even after being rinsed in the laboratory. In terms of the forest floor treatment, we saw a notable increase in *ppk2* over time (Figure 4.11B). This pattern of increasing *ppk2* relative abundance with decreasing P availability was noted once before (Temperton et al. 2011). Therefore, microorganisms with *ppk2* may tend to thrive in more nutrient limited environments like the forest floor as compared to the stream in this study. While the smooth term was significantly different from a null model and indicates an increase in the relative abundance of *ppx* associated with forest floor leaf litter over time, the overlap in confidence intervals between the two treatments suggests this difference was weak (Figure 4.11C). There is no consistent and significant trend in *ppk1* and *ppx* relative abundance versus time—the most likely polyP functional gene combination representing putative PAOs. Therefore, when taken in combination with the leaf P data presented earlier, these results indicate a limited role of PAOs in leaf litter associated P cycling.

It is interesting to note that the dynamics of relative abundance patterns over time are very different for forest floor samples harboring *ppk2* compared to those harboring *ppk1* or *ppx*. Specifically, we observed a consistent increase in *ppk2* relative abundance over time (Figure

4.11B) but a relatively flat trend in *ppk1* and *ppx* relative abundance with time (Figures 4.11A and 4.11C). We have no explanation for the oscillation in *ppx* relative abundance (Figure 4.11C) other than (1) the trend is being driven by an outlier at ~50 days from the start of the experiment or (2) the organisms with this gene are more sensitive to low temperatures—the start of December occurred at ~80 days from the start of the experiment.

Heatmaps of *pp1*, *ppk2*, and *ppx* relative abundance of different contigs across all samples collected highlight the impact of treatment on the leaf litter microbial community (Figures 4.12-4.14). Specifically, when clustered by similar relative abundance profiles, we see that samples are clearly grouped by treatment (see blue and green labels for forest floor and stream samples at top of grid, respectively; Figures 4.12-4.14) except for one sample from the stream (LL85) that groups with forest floor samples for all three genes in this study. Additionally, groups of anywhere from 4-7 contigs dominate the profile and are observed consistently in time. One potential reason for the switch in contigs from the start of the experiment to the end may be due to the impact of winter conditions from ~80-180 days. This switch is especially evident in the *ppk1* and *ppx* heatmaps (Figures 4.12 and 4.13).

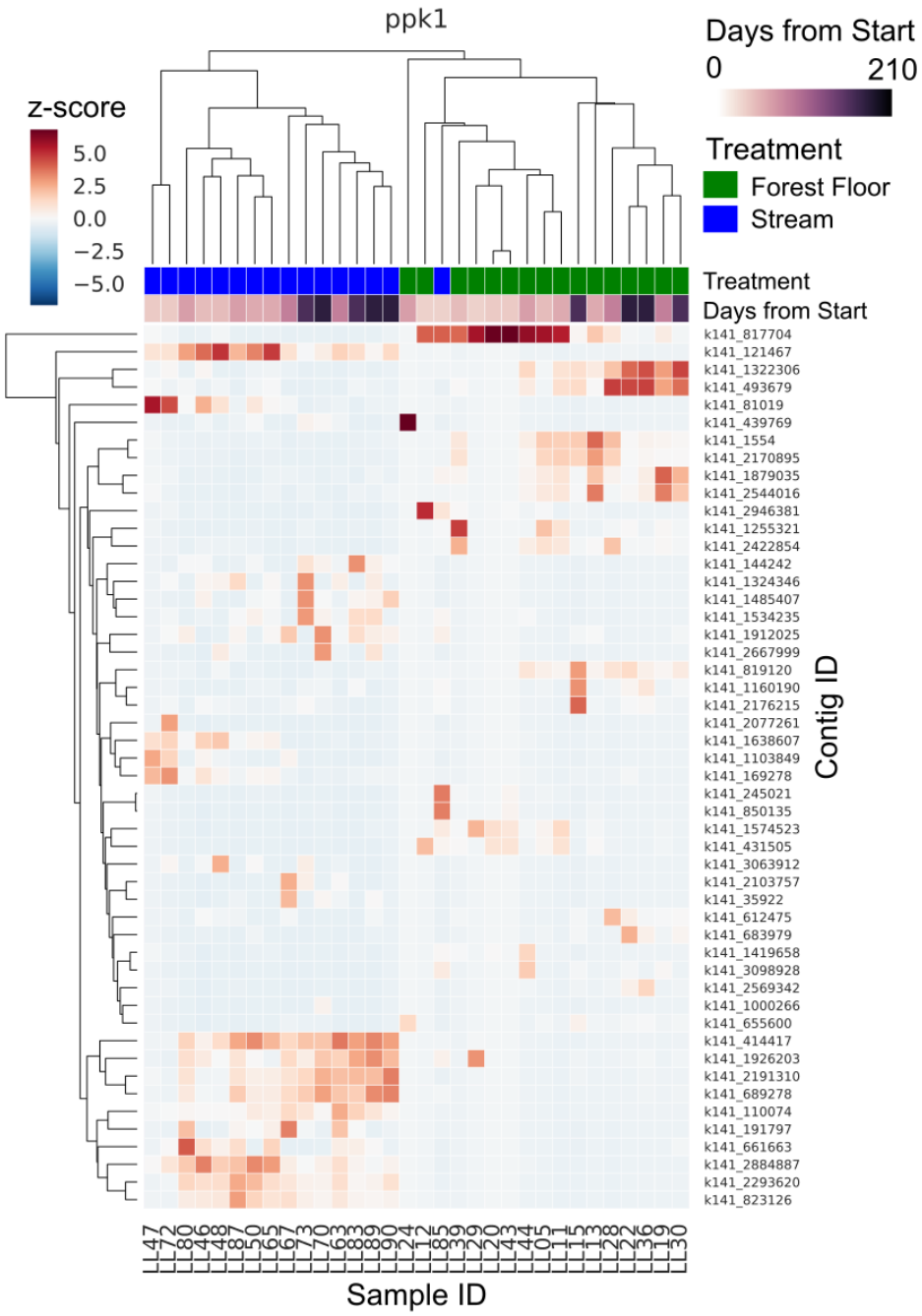

Figure 4. 1 2 Heatmap of the relative abundance z-score for *ppk1* genes clustered by similar relative abundance profiles between sample ID and contig ID. Key on top shows treatment type and collection time for each sample. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall.

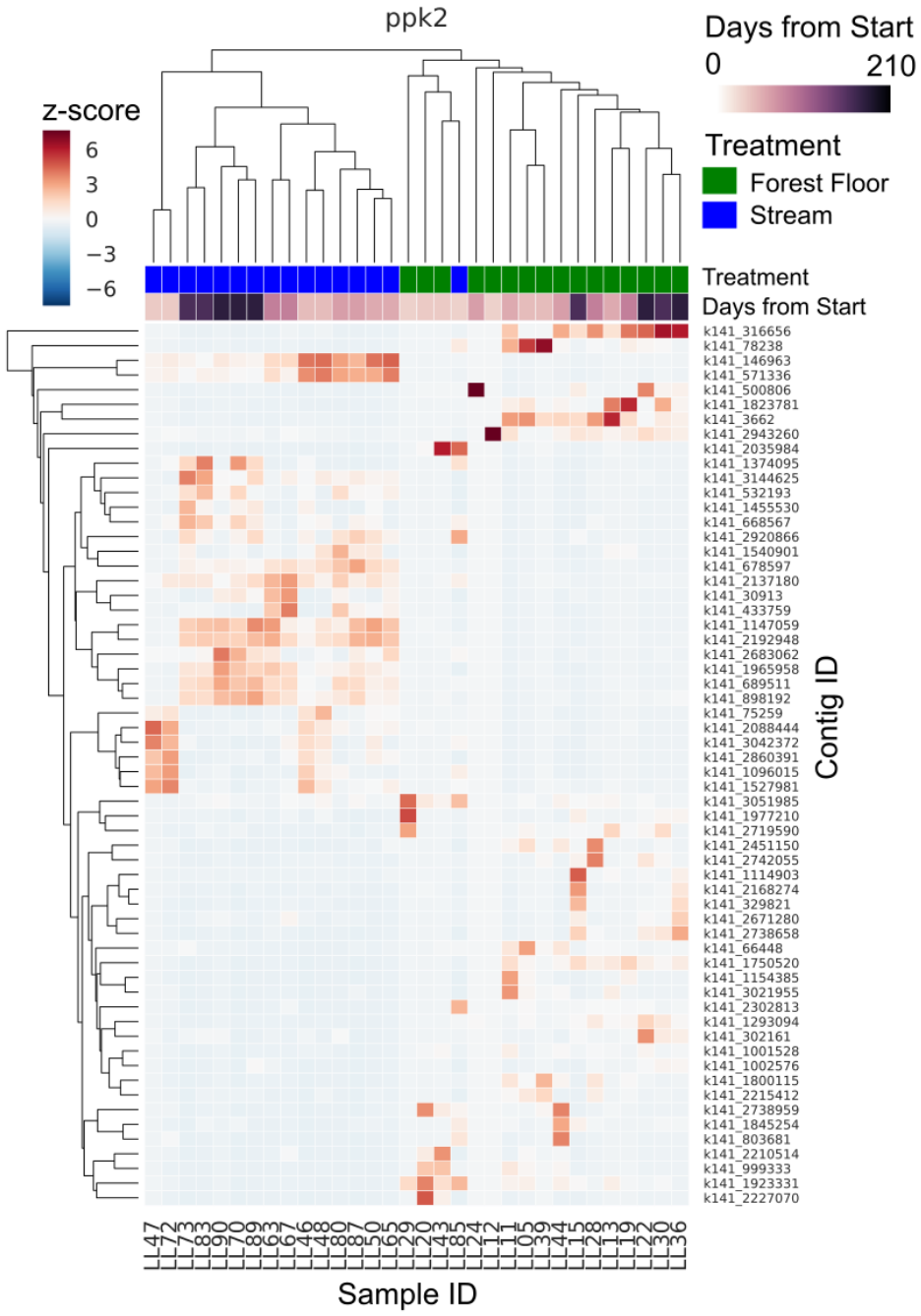

Figure 4. 1 3 Heatmap of the relative abundance z-score for *ppk2* genes clustered by similar relative abundance profiles between sample ID and contig ID. Key on top shows treatment type and collection time for each sample. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall.

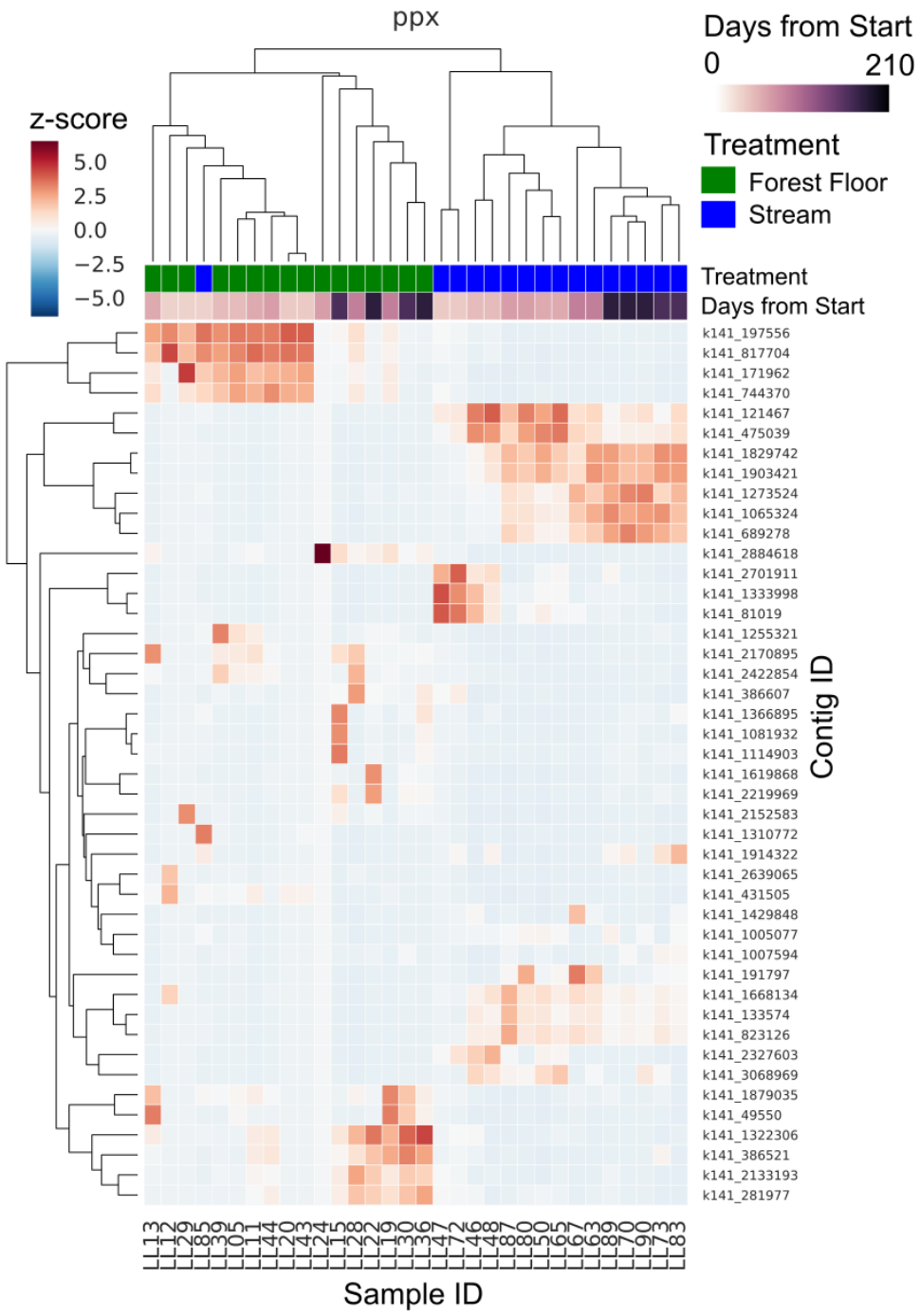

Figure 4. 1 4 Heatmap of the relative abundance z-score for *ppx* genes clustered by similar relative abundance profiles between sample ID and contig ID. Key on top shows treatment type and collection time for each sample. Includes the non-redundant top 3 contigs by maximum abundance for each sample as well as the top 7 contigs by mean abundance overall.

CONCLUSIONS

Leaf litter is an important nutrient source to terrestrial ecosystems and may influence microbially-mediated P cycling. However, with respect to the influence of PAO-mediated processes, these results suggest PAOs have a limited role in leaf litter associated P cycling. More specifically, we observed no trend in *ppk1* and a limited trend in *ppx* gene abundance—the most common gene combination for putative PAOs in this study—over time in both treatments despite changes in leaf P for both treatments. As a result, P bound to sediment may have been an important driver of the increases in leaf P observed in the stream treatment. Interestingly, we observed an increase in the relative abundance of *ppk2* in the forest floor treatment, which lost leaf P over the course of the experiment. While these results do not support our hypotheses as they relate to PAO-mediated controls on P cycling, they do highlight the need for more research (1) characterizing putative PAOs found in environmental samples as well as (2) addressing the impact of P limitation on organisms with *ppk* and *ppx* genes.

ACKNOWLEDGMENTS

We thank Chuck Pepe-Ranney for help with sequencing analysis, Rod Christie and Chris Nagy of the Mianus River Gorge Preserve in Bedford, NY for their help with experiment set-up and sample collection, Emily Stephan and Danielle Andriozzi for their help with experiment set-up, Shree Giri for help with laboratory techniques, Jennifer Mosher and Jeff Mattison at the Cornell University Institute for Biotechnology Genomics Facility for help with sample preparation and sequencing, and Mallory Choudoir for help with sequencing sample preparation. SMS was supported by a fellowship from the Mianus River Gorge Preserve, the Cornell University College

of Agriculture and Life Sciences Land Grant Fellowship, and US Environmental Protection Agency (EPA) Science to Achieve Results (STAR) Fellowship. This publication was developed under STAR Fellowship Assistance Agreement no. FP917670-01-0 awarded by the EPA. It has not been formally reviewed by the EPA. The views expressed in this publication are solely those of SMS and EPA does not endorse any products or commercial services mentioned in this publication. Data and R scripts associated with this work are available on GitHub at https://github.com/sheilasaia/paper-p-leaf-litter.

REFERENCES

Ahn, K. and A. Kornberg. 1990. Polyphosphate kinase from *Escherichia coli*. Biochemistry. 265(20):11734–11739.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19(6):716-723.

Akiyama, M., E. Crooke and A. Kornberg. 1993. An exopolyphosphatase of *Escherichia coli*. The Journal of Biological Chemistry. 268(1):633–639.

Akiyama, M., E. Crookes and A. Kornberg. 1992. The polyphosphate kinase gene of *Escherichia coli*. The Journal of Biological Chemistry. 267(31):22556–22561.

Alcántara, C., A. Blasco, M. Zúñiga and V. Monedero. 2014. Accumulation of polyphosphate in *Lactobacillus* spp. and its involvement in stress resistance. Applied and Environmental Microbiology. 80(5):1650–1659.

Ball, B.A., M.A. Bradford and M.D. Hunter. 2009. Nitrogen and phosphorus release from mixed litter layers is lower than predicted from single species decay. Ecosystems. 12(1):87–100.

Bergmann, G.T., S.T. Bates, K.G. Eilers, C.L. Lauber, G. Caporaso, W.A. Walters, R. Knight and N. Fierer. 2011. The under-recognized dominance of Verrucomicrobia in soil bacterial communities. Soil Biologiy ane Biochemistry. 43(7):1450–1455.

Brown, M.R.W. and A. Kornberg. 2008. The long and short of it - polyphosphate, PPK and bacterial survival. Trends in Biochemical Sciences. 33(6):284–290.

Burkhardt, J., H. Kaiser, H. Goldbach and L. Kappen. 1999. Measurements of electrical leaf surface conductance reveal recondensation of transpired water vapour on leaf surfaces. Plant, Cell and Environment. 22(2):189–196.

Camarero, L. and J. Catalan. 2012. Atmospheric phosphorus deposition may cause lakes to revert from phosphorus limitation back to nitrogen limitation. Nature Communications. 3:1–5.

Chamier, A.-C. 1987. Effect of pH on Microbial-Degradation of Leaf Litter in 7 Streams of the English-Lake-District. Oecologia. 71:491–500.

Cheever, B.M., E.B. Kratzer and J.R. Webster. 2012. Immobilization and mineralization of N and P by heterotrophic microbes during leaf decomposition. Freshwater Science. 31(1):133–147.
Cotrufo, F.M., P. Ineson and D.J. Roberts. 1995. Decomposition of birch leaf litters with varying C-to-N ratios. Soil Biology and Biochemistry. 27(9):1219–1221.
Cummins, K.W. 1974. Structure and function of stream ecosystems. BioScience. 24(11):631–641.
DeAngelis, K.M., W.L. Silver, A.W. Thompson and M.K. Firestone. 2010. Microbial communities acclimate to recurring changes in soil redox potential status. Environmental Microbiology. 12(12):3137–3149.
DebRoy, S., P. Mukherjee, S. Roy, A.R. Thakur and S. Raychaudhuri. 2013. Draft genome sequence of a phosphate-accumulating *Bacillus* sp .,. Genome Announcements. 1(1):1–2.
Dodds, W.K. 2003. The role of periphyton in phosphorus retention in shallow freshwater aquatic systems. Journal of Phycology. 39(5):840–849.
Elwood, J.W., J.D. Newbold, A.F. Trimble and R.W. Stark. 1981. The Limiting Role of Phosphorus in a Woodland Stream Ecosystem: Effects of P Enrichment on Leaf Decomposition and Primary Producers. Ecology. 62(1):146–158.
Fanin, N., S. Hättenschwiler, P.F. Chavez Soria and N. Fromin. 2016. (A)synchronous availabilities of N and P regulate the activity and structure of the microbial decomposer community. Frontiers in Microbiology. 6(6):1507.
Federle, T.W., V.L. McKinley and J.R. Vestal. 1982. Physical determinants of microbial colonization and decomposition of plant litter in an Arctic Lake. Microbial Ecology1. 8(2):127–138.
Finn, R.D., J. Clements and S.R. Eddy. 2011. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Research. 39:W29–W37.
Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate, and A. Bateman. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 44(D1):D279-85. Database Issue 44:D279-D285
Fisher, S.G. and G.E. Likens. 2011. Energy Flow in Bear Brook , New Hampshire : An Integrative Approach to Stream Ecosystem Metabolism. Ecological Monographs. 43(4):421–439.
Friberg, N., J.B. Dybkj??r, J.S. Olafsson, G.M. Gislason, S.E. Larsen and T.L. Lauridsen. 2009. Relationships between structure and function in streams contrasting in temperature. Freshwater Biology. 54(10):2051–2068.
Gholz, H.L., D.A. Wedin, S.M. Smitherman, M.E. Harmon and W.J. Parton. 2000. Long-term dynamics of pine and hardwood litter in contrasting environments: Toward a global model of decomposition. Global Change Biology. 6(7):751–765.
Gebremariam, S.Y., M.W Beutel, D. Christian, T.F. Hess. 2011. Research advances and challenges in the microbiology of enhanced biological phosphorus removal—a critical review. Water Environmental Research. 83(3):195– 219.
Gonzalez, G. and T. Seastedt. 2001. Soil fauna and plant litter decomposition in tropical and subalpine forests. Ecology. 82(4):955–964.

Graham, E.B., A.R. Crump, C.T. Resch, S. Fansler, E. Arntzen, D.W. Kennedy, J.K. Fredrickson and J.C. Stegen. 2017. Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. Environmental Microbiology. 19(4):1552–1567.
Gregory, S. V. 1978. Residence time studies and deciduous-coniferous comparison. Verhandlungen des Internationalen Verein Limnologie. 20:1340–1346.
Gulis, V. and K. Suberkropp. 2003. Leaf litter decomposition and microbial activity in nutrient-enriched and unaltered reaches of a headwater stream. Freshwater Biology. 48(1):123–134.
Günther, S., M. Trutnau, S. Kleinsteuber, G. Hause, T. Bley, I. Röske, H. Harms and S. Müller. 2009. Dynamics of polyphosphate-accumulating bacteria in wastewater treatment plant microbial communities detected via DAPI (4',6'-diamidino-2- phenylindole) and tetracycline labeling. Applied and Environmental Microbiology. 75(7):2111–2121.
Hauer, F.R. and G.A. Lamberti. 2006. Methods in stream ecology, Burlington, MA: Academic Press.
He, S., D.L. Gall and K.D. McMahon. 2007. *Candidatus* Accumulibacter population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes. Applied and Environmental Microbiology. 73(18):5865–5874.
Hesselmann, R.P., C. Werlen, D. Hahn, J.R. van der Meer and a J. Zehnder. 1999. Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. Systematic and Applied Microbiology. 22(3):454–465.
Hupfer, M. and R. Gächter. 1995. Polyphosphate in lake sediments: 31P NMR spectroscopy as a tool for its identification. Limnology and Oceanography. 40(3):610–617.
Hupfer, M., B. Rube and P. Schmieder. 2004. Origin and diagenesis of polyphosphate in lake sediments: a 31P-NMR study. Limnology and Oceanography. 49(1):1–10.
Hyatt, D., G.-L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer and L.J. Hauser. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 11:119.
JGI (Joint Genome Institute). 2017. BBMap/BBTools. Available online: https://sourceforge.net/projects/bbmap/. Accessed: March 1, 2017.
Jin, X., Y. Ma, Z. Kong, W. Kou and L. Wu. 2017. The Variation of Sediment Bacterial Community in Response to Anthropogenic Disturbances of Poyang Lake, China. Wetlands. :doi:10.1007/s13157-017-0909-1.
Kawakoshi, A., H. Nakazawa, J. Fukada, M. Sasagawa, Y. Katano, S. Nakamura, A. Hosoyama, H. Sasaki, N. Ichikawa, S. Hanada, Y. Kamagata, K. Nakamura, S. Yamazaki and N. Fujita. 2012. Deciphering the genome of polyphosphate accumulating *Actinobacterium Microlunatus phospho*rus. DNA Research. 19(5):383–394.
Keasling, J.D., L. Bertsch and A. Kornberg. 1993. Guanosine pentaphosphate phosphohydrolase of *Escherichia coli* is a long-chain exopolyphosphatase. Proceedings of the National Academy of Sciences. 90(15):7029–7033.
Kong, Y., Y. Xia, J.L. Nielsen and P.H. Nielsen. 2007. Structure and function of the microbial community in a full-scale enhanced biological phosphorus removal plant. Microbiology. 153(12):4061–4073.

Kornberg, A. 1995. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. Journal of Bacteriology. 177(3):491–496.
Kristiansen, R. et al. 2013. A metabolic model for members of the genus *Tetrasphaera* involved in enhanced biological phosphorus removal. The ISME Journal. 7(3):543–554.
Kunin, V., S. He, F. Warnecke, S.B. Peterson, H. Garcia Martin, M. Haynes, N. Ivanova, L.L. Blackall, M. Breitbart, F. Rohwer, K.D. McMahon and P. Hugenholtz. 2008. A bacterial metapopulation adapts locally to phage predation despite global dispersal. Genome Research. 18(2):293–297.
Li, D., C.M. Liu, R. Luo, K. Sadakane and T.W. Lam. 2015. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 31(10):1674–1676.
Li, X., H. Yuan, J. Yang and B. Li. 2013. Genome sequence of the polyphosphate-accumulating organism Arthrobacter sp. strain PAO19 isolated from maize rhizosphere soil. Genome Announcements. 1(4):11039–11040.
Li, H. 2017. bwa-mem. Available online: http://bio-bwa.sourceforge.net/. Accessed: March 1, 2017.
Lladó, S.F., R. López-Mondéjar and P. Baldrian. 2017. Forest Soil Bacteria: Diversity, Involvement in Ecosystem Processes, and Response to Global Change. Microbiology and Molecular Biology Reviews. 81(2):1–27.
Locke, N. 2015. Classification of polyphosphate-accumulating bacteria in benthic biofilms. Pennsylvania State University, State College, Pennsylvania.
Martínez, A., A. Larrañaga, J. Pérez, E. Descals and J. Pozo. 2014. Temperature affects leaf litter decomposition in low-order forest streams: Field and microcosm approaches. FEMS Microbiology Ecology. 87(1):257–267.
Martínez, A., S. Monroy, J. Pérez, A. Larrañaga, A. Basaguren, J. Molinero and J. Pozo. 2016. In-stream litter decomposition along an altitudinal gradient: does substrate quality matter? Hydrobiologia. 766(1):17–28.
Martins, G., A. Terada, D.C. Ribeiro, A.M. Corral, A.G. Brito, B.F. Smets and R. Nogueira. 2011. Structure and activity of lacustrine sediment bacteria involved in nutrient and iron cycles. FEMS Microbiology Ecology. 77(3):666–679.
Mehring, A.S., K.A. Kuehn, A. Thompson, C.M. Pringle, A.D. Rosemond, M.R. First, R.R. Lowrance and G. Vellidis. 2015. Leaf litter nutrient uptake in an intermittent blackwater river: Influence of tree species and associated biotic and abiotic drivers. Functional Ecology. 29(6):849–860.
Menzel, P., K. L. Ng, and A. Krogh. 2016. Kaiju : Fast and sensitive taxonomic classification for metagenomics. Nature Communications. 7:11257.
Methot, G., C. Hudon, P. Gagnon, B. Pinel-Alloul, A. Armellin and A.-M.T. Poirier. 2012. Macroinvertebrate size–mass relationships: how specific should they be? Freshwater Science. 31(3):750–764.
Meyer, J.L. 1980. Dynamics of Phosphorus and Organic Matter during Leaf Decomposition in a Forest Stream Dynamics of phosphorus and organic matter during leaf decomposition in a forest stream. Oikos. 34(1):44–53.
Meyer, J.L. and G.E. Likens. 1979. Transport and Transformation of Phosphorus in a Forest Stream Ecosystem. Ecology. 60(6):1255–1269.

Moore, T.R., J.A. Trofymow, C.E. Prescott, J. Fyles and B.D. Titus. 2006. Patterns of carbon, nitrogen and phosphorus dynamics in decomposing foliar litter in Canadian forests. Ecosystems. 9(1):46–62.

Nakamura, K., A. Hiraishi, Y. Yoshimi, M. Kawaharasaki, K. Masuda and Y. Kamagata. 1995a. *Microlunatus phosphovorus* gen. nov., sp. nov., a new Gram-positive polyphosphate-accumulating bacterium isolated from activated sludge. International Journal of Systematic Bacteriology. 45(1):17–22.

Nakamura, K., S. Ishikawa and M. Kawaharasaki. 1995b. Phosphate uptake and release activity in immobilized polyphosphate-accumulating bacterium *Microlunatus phosphovorus* strain NM-1. Journal of Fermentation and Bioengineering. 80(4):377–382.

NCBI (National Center for Biotechnology Information, Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2017 Mar 06]. Available from: https://www.ncbi.nlm.nih.gov/.

Newbold, J.D., J.W. Elwood, R. V. O'Neill and A.L. Sheldon. 1983. Phosphorus Dynamics in a Woodland Stream Ecosystem: A Study of Nutrient Spiralling. Ecology. 64(5):1249–1265.

Nguyen, H.T.T., V.Q. Le, A.A. Hansen, J.L. Nielsen and P.H. Nielsen. 2011. High diversity and abundance of putative polyphosphate-accumulating *Tetrasphaera*-related bacteria in activated sludge systems. FEMS Microbiology Ecology. 76(2):256–267.

O'Brien, M.J., C.E. Gomola and T.R. Horton. 2011. The effect of forest soil and community composition on ectomycorrhizal colonization and seedling growth. Plant and Soil. 341(1–2):321–331.

Ouyang, Y. and X. Li. 2013. Recent research progress on soil microbial responses to drying–rewetting cycles. Acta Ecologica Sinica. 33(1):1–6.

Peralta, A.L., S. Ludmer, J.W. Matthews and A.D. Kent. 2014. Bacterial community response to changes in soil redox potential along a moisture gradient in restored wetlands. Ecological Engineering. 73:246–253.

Peterson, S.B., F. Warnecke, J. Madejska, K.D. McMahon and P. Hugenholtz. 2008. Environmental distribution and population biology of *Candidatus* Accumulibacter, a primary agent of biological phosphorus removal. Environmental Microbiology. 10(10):2692–703.

Pett-Ridge, J. and M.K. Firestone. 2005. Redox fluctuation structures microbial communities in a wet tropical soil redox fluctuation structures microbial communities in a wet tropical soil. Applied and Environmental Microbiology. 71(11):6998–7007.

Purahong, W., T. Wubet, G. Lentendu, M. Schloter, M.J. Pecyna, D. Kapturska, M. Hofrichter, D. Krüger and F. Buscot. 2016. Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. Molecular Ecology. 25(16):4059–4074.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rangarajan, E.S., G. Nadeau, Y. Li, J. Wagner, M.N. Hung, J.D. Schrag, M. Cygler and A. Matte. 2006. The Structure of the exopolyphosphatase (PPX) from *Escherichia coli* O157:H7 suggests a binding mode for long polyphosphate chains. Journal of Molecular Biology. 359(5):1249–1260.

Rao, N.N., M.R. Gómez-García and A. Kornberg. 2009. Inorganic polyphosphate: essential for growth and survival. Annual Review of Biochemistry. 78:605–647.
Reizer, J., A. Reizer and M.H. Saier. 1993. Exopolyphosphatase and guanine pentaphosphate phosphatase belong to the sugar kinase/actin/hsp70 superfamily. Trends in Biochemical Sciences. :247–248.
Rier, S.T., K.C. Kinek, S.E. Hay and S.N. Francoeur. 2016. Polyphosphate plays a vital role in the phosphorus dynamics of stream periphyton. Freshwater Science. 35(2):490–502.
Saia, S.M., P.J. Sullivan, J.M. Regan, H.J. Carrick, A.R. Buda, N.A. Locke and M.T. Walter. 2017. Evidence for polyphosphate accumulating organism (PAO)-mediated phosphorus cycling in stream biofilms under alternating aerobic/anaerobic conditions. Freshwater Science. 36(2):284–296.
Schmidt, M., E. Veldkamp and M.D. Corre. 2016. Tree-microbial biomass competition for nutrients in a temperate deciduous forest, central Germany. Plant and Soil. 408:227–242.
Schulz, H.N. and H.D. Schulz. 2005. Large sulfur bacteria and the formation of phosphorite. Science. 307:416–418.
Seviour, R.J., T. Mino and M. Onuki. 2003. The microbiology of biological phosphorus removal in activated sludge systems. FEMS Microbiology Reviews. 27(1):99–127.
Seviour, R.J. and P.H. Nielsen. 2010. Microbial Ecology of Activated Sludge R. J. Seviour & P. H. Nielsen, eds., London, UK: IWA Publishing.
Skennerton, C.T., J.J. Barr, F.R. Slater, P.L. Bond and G.W. Tyson. 2014. Expanding our view of genomic diversity in *Candidatus* Accumulibacter clades. Environmental Microbiology. 17(5):1574–1585.
Smyth, C.E., B. Titus, J.A. Trofymow, T.R. Moore, C.M. Preston and C.E. Prescott. 2016. Patterns of carbon, nitrogen and phosphorus dynamics in decomposing wood blocks in Canadian forests. Plant and Soil. 409(459–477).
Soo, R.M., C.T. Skennerton, Y. Sekiguchi, M. Imelfort, S.J. Paech, P.G. Dennis, J.A. Steen, D.H. Parks, G.W. Tyson and P. Hugenholtz. 2014. An expanded genomic representation of the phylum Cyanobacteria. Genome Biology and Evolution. 6(5):1031–1045.
Soong, J.L., W.J. Parton, F. Calderon, E.E. Campbell and M.F. Cotrufo. 2015. A new conceptual model on the fate and controls of fresh and pyrolized plant litter decomposition. Biogeochemistry. 124(1–3):27–44.
Sponseller, A. and E.F. Benfield. 2001. Influences of land use on leaf breakdown in southern Appalachian headwater streams: a multiple-scale analysis. Journal of the North American Benthological Society. 20(1):44–59.
Tamatamah, R.A., R.E. Hecky and H.C. Duthie. 2005. The atmospheric deposition of phosphorus in Lake Victoria (East Africa). Biogeochemistry. 73(2):325–344.
Temperton, B., J. a Gilbert, J.P. Quinn and J.W. McGrath. 2011. Novel analysis of oceanic surface water metagenomes suggests importance of polyphosphate metabolism in oligotrophic environments. PLOS One. 6(1):1–14.
Tipping, E., S. Benham, J.F. Boyle, P. Crow, J. Davies, U. Fischer, H. Guyatt, R. Helliwell, L. Jackson-Blake, A.J. Lawlor, D.T. Monteith, E.C. Rowe and H. Toberman. 2014. Atmospheric deposition of phosphorus to land and freshwater. Environmental Science Processes & Impacts. 16(7):1608–1617.

Torsten Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous Inference in General Parametric Models. Biometrical Journal 50(3):346-363.
Uhlmann, D. and H.-D. Bauer. 1988. A remark on microorganisms in lake sediments with emphasis on polyphosphate-accumulating bacteria. Internationale Revue der gesamten Hydrobiologie und Hydrographie. 73(6):703–708.
USEPA (US Environmental Protection Agency). 1978. Method 365.3. Phosphorous, all forms (colorimetric, ascorbic acid, two reagent). US Environmental Protection Agency, Washington, DC.
Webster, J.R. and E.F. Benfield. 1986. Vascular plant breakdown in freshwater ecosystems. Annual Reviews in Ecological Systems. 17:567–594.
Wood, S. N. 2015. mgcv: mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation. R package version 1.8-6. R Project for Statistical Computing, Vienna, Austria.
Xu, W., L. Shi, O. Chan, J. Li, P. Casper and X. Zou. 2013. Assessing the effect of litter species on the dynamic of bacterial and fungal communities during leaf decomposition in microcosm by molecular techniques. PLOS ONE. 8(12):1–10.
Zago, A., S. Chugani and A.M. Chakrabarty. 1999. Cloning and characterization of polyphosphate kinase and exopolyphosphatase genes from *Pseudomonas aeruginosa* 8830. Applied and Environmental Microbiology. 65(5):2065–2071.
Zhang, A.N. et al. 2016. Development of quantitative real-time PCR assays for different clades of *Candidatus* Accumulibacter. Scientific Reports. 6:23993.
Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York.
Zhang, H., K. Ishige and A. Kornberg. 2002. A polyphosphate kinase (PPK2) widely conserved in bacteria. Proceedings of the National Academy of Sciences. 99(26):16678–16683.

CHAPTER 5

CONCLUSIONS

The goal of this work was to investigate the role of polyphosphate accumulating organisms (PAOs) on P cycling in the environment. After a synthesis of existing work, it was clear that only a few studies explicitly addressed the control of PAOs on P cycling outside specialized wastewater treatment plants (WWTPs). Of these, we found none that explicitly set out to test the impact of PAO-mediated P cycling under fluctuating oxygen conditions in streams and soils. Therefore, we set out to fill this knowledge gap by studying PAOs in stream biofilms and soils.

Stream biofilms took up P under aerobic conditions and released P under anaerobic conditions similarly to PAOs in WWTPs. We also observed more cells with intracellular polyphosphate (polyP) under aerobic conditions compared to anaerobic conditions, which is also consistent with the PAO phenotype. However, we were unable to determine the relative importance of PAO-mediated P cycling compared to abiotic controls on P. In our soil experiment, we saw relationships between the soil wetness index (SWI), which integrates soil moisture patterns in space and time, and mobile P pools but the predictability of these patterns did not transfer to an ability to predict the relative abundance of P cycling functional genes (i.e., *ppk* and *ppx*). Soils are complex biological environments (Dunbar et al. 2002; Gans et al. 2005; Tringe et al. 2005; Fierer and Jackson 2006; Schloss and Handelsman 2006; Hug et al. 2016; Nesme et al. 2016), and therefore, PAOs may have a limited role on P cycling considering all the other

microorganisms competing for this limited resource as well as the potential abiotic factors controlling P (e.g., land management and iron redox).

In addition to our limited understanding of PAO-mediated P cycling under changing environmental conditions, we found only one study (in marine systems) that related the abundance of *ppk* and *ppx* functional genes to environmental P concentrations. Additionally, there many studies in WWTPs have determined *ppk* relative abundance but not *ppx*. This is surprising given that a combination of both these genes is thought to be required for microorganism to be considered PAOs. More specifically, PAOs must be able to (1) store P as intracellular polyP under aerobic conditions and then (2) beak it down during anaerobic conditions; this metabolism is regulated by *ppk* and *ppx*, respectively. Furthermore, most WWTP studies that quantify *ppk* have only focused on *Candidatus* Accumulibacter phosphatis (CAP). Given the evidence that polyP functional genes were widespread in the environment, broader techniques (e.g., next generation-sequencing) are needed to characterize the diversity of PAOs as well as the relative abundance of *ppk* and *ppx* in natural systems.

By employing some of these broader techniques in this study, we identified many organisms in soil and leaf litter samples that were related to phyla/classes of known PAOs as well as new putative PAOs. We recognize these phyla/classes are very diverse but this information could be used to develop new tools such as quantitative polymerase change reaction (qPCR) primers and fluorescence in-situ hybridization (FISH) probes to verify the phenotypes and quantify the

presence of these putative PAOs in the laboratory or field. Specifically, we encourage future studies to focus on the genus *Burkholderia*, which was identified in our soil and leaf litter experiments as well as other leaf litter and stream biofilm studies (Locke 2013; Purahong et al. 2016). *Burkholderia*, was shown to store intracellular polyP granules (Locke 2013) and is also known for its ability to solubilize P (Purahong et al. 2016). These data may also be helpful in identifying putative PAOs, like *Burkholderia*, that are already studied in pure culture and may be helpful in exploring the role of PAOs in the environment, and thus, may provide a more controlled setting to test (1) the influence of environmental conditions and (2) the relative importance of PAOs versus abiotic processes on P cycling.

Last, the consistent finding by other researchers that leaf P concentrations increased in the stream over time and that these may be due to microbial accumulation, motivated us to explore the role of PAOs in this setting. We observed leaf P trends that were consistent with previous studies but found no trend in the relative abundance of *ppk* and *ppx* genes over time. This finding was unexpected and leads us to believe that PAOs were likely playing a limited role in P cycling associated with leaf litter decomposition. However, one exception emerged in the forest floor leaf litter treatments where ppk2 genes significantly increased over time despite decreased leaf P concentrations over time. We suggest this finding signals the need for more research exploring the relative contributions of abiotic and biotic controls on P cycling in various environments as well as along environmental gradients (e.g., P availability gradients).

FUTURE DIRECTIONS

This research uncovered many additional research questions and experimental considerations with regards to PAO-mediated P cycling in natural systems. We summarize future directions for each major study (i.e. stream biofilms, soils, and leaf litter).

Stream Biofilms

- Rather than forcing treatment 1 (T1) to be anaerobic an aerobic using a gas (i.e., N_2 or mixed N_2:CO_2), repeat experiment using alternating dark and light cycles to indirectly control O_2 availability over time. See Carlton and Wetzel (1988) for details on this experimental design.
- Consider continuously recording (i.e., every 5 minutes) dissolved oxygen (DO), pH, and phosphate sensors close to the biofilm and within the water column. The purpose of this is to test whether a thin O_2 rich layer develops around the surface of the biofilm and how this O_2 rich layer alters local pH and phosphate concentrations as the treatment transitions from dark to light conditions.
- Include more tub replicates and collect biofilms from streams with different P loads to study the impact of historical P availability on PAO communities and polyP accumulation efficiency within the stream biofilm.
- Compare laboratory results to field measurements of DO and phosphate concentration in the water column and near the water column-biofilm interface over 24-hour period.
- Consider repeating this experiment with alternative controls including (1) a bare rock treatment to make sure the stream substrate does not influence P cycling and (2) an abiotic

only treatment—where an antibiotic or other biocide is used to kill living organisms in the biofilm—to study the magnitude of abiotic and biotic processes on P cycling.

- Harvest the biofilm over the course of the experiment and analyze each sample using DAPI-polyP staining as well as PAO targeting FISH probes. This may also include sorting out cells that have polyP using flow cytometry and sequencing microbial communities before and after sorting like Locke (2015). Additionally, extract DNA from cells and use shotgun metagenomics to analyze for the abundance of functional genes (e.g., *ppk* and *ppx*) over time.
- Ceramic tiles offer a homogeneous surface size to grow biofilms in the stream and can be used to quantify how quickly stream biofilms grown in the stream and how they accumulate polyP over time. Also of interest is how much stream biofilm is scoured off in storm events of various sizes and what is the magnitude of P transport downstream.
- Use NanoSIMS to co-locate the presence of intracellular polyP and metals that are known to stabilize the negative charge of polyP (e.g., Ca and Mg). Estimate the impact of biotic versus abiotic processes on the uptake and release of these cations.
- Characterize non-bacterial microorganisms within the stream biofilm using traditional classification techniques (e.g., Price and Carrick 2013), amplicon sequencing, or other strategies discussed in chapter 1.
- Consider the role of other nutrients (e.g., N or Fe) in the stream biofilm ecosystem as well specific microbially-mediated processes that go along with them (e.g., denitrification and iron reduction). Use some of the approaches summarized in chapter 1 to study the microorganisms involved in the co-cycling of these nutrients and compare findings to that of PAOs to look for potential competitors and symbionts.

- Develop qPCR primers and qFISH probes from existing shotgun metagenomes for the quantification of functional genes (e.g., *ppk* and *ppx*) and PAOs, respectively. Especially for *Alphaproteobacteria* putative PAOs, which dominated metagenome hits.

Soils

- Consider gathering samples along a more detailed soil moisture/SWI gradient (i.e., gridded sample design) as well as a more detailed P limitation gradient for each of the different land cover types and consider controlling for OM and pH gradients. Consider collecting samples up to 20 cm to study whether we see similar results.
- Samples have been saved in the freezer for FISH and DAPI staining of PAOs. Analyze these to test the hypothesis that more polyP is stored under aerobic (unsaturated) conditions compared to anaerobic (saturated) conditions. Sequencing data from this experiment can be used to design FISH probes as well as qPCR probes for known EBPR PAOs and functional genes, respectively.
- There are many other types of soil P extractions that could have been carried out in this study. Carry out full Hedley consideration on select soil samples to understand impact of abiotic processes on P cycling. ^{18}O-phosphate measurements from various P pools can also be compared to determine whether P originated from biotic or abiotic processes. See Joshi et al. (2016) and Tamburini et al. (2012) as examples of this method. Additionally, measurements of microbial P pools and microbial activity (respiration) can be measured along the soil moisture gradient.

- Sequence non-grassland samples using shotgun metagenomics approaches and compare the abundance and diversity of functional genes as well as community members between managed (e.g., cropped land) and un-managed land (e.g., forest) along soil moisture/SWI gradients.
- Look for co-dependences between P, C, N, and other micro-nutrients (e.g., Fe, Mg) with changes in soil moisture/SWI.
- Carry out laboratory methods experiments to establish that $CaCl_2$-P is a better approximation of runoff P compared to water extractable P or oxalate extractable P (Ox-P) for the soils included in this study.
- Analyze microbial and fungal biomass so it may be correlated with polyP functional genes and soil P.
- Use NanoSIMS to identify the presence of metals that are known to be associated with PPK1 and PPK2 (i.e., Mg and Mn, respectively).

Leaf Litter

- Rather than mixing all leaves together in one bag, separate out different leaf types and measure P accumulation in the stream over time. Test for patterns in P accumulation with leaf litter properties (e.g., leaf litter C:N ratio). Study how the microbial communities as well as diversity and abundance of functional genes vary with different types of leaves. Also consider putting the same types of leaves in streams along a P load gradient to look at how P supply impacts litter decomposition and P accumulation on leaves in the stream.

- Use DAPI and FISH techniques to verify the presence of PAOs on the leaves over time. Develop new FISH probes as well as qPCR probes for known EBPR PAOs and functional genes, respectively
- Use a laboratory flume set-up to study the impact of flow, velocity, temperature, P availability, and O_2 availability on leaf P accumulation over time.
- Draw out forest floor study to test whether P accumulation ever begins.
- Look for co-dependences between P, C, N, and other micro-nutrients (e.g., Fe, Mg) with as P accumulates on the leaves over time.
- Calculate mass balance of P that enters and cycles through the forest/stream ecosystem as result of leaf litter additions.
- Analyze microbial and fungal biomass so it may be correlated with polyP functional genes and leaf P.
- Use NanoSIMS to identify the presence of metals that are known to be associated with PPK1 and PPK2 (i.e., Mg and Mn, respectively).

REFERENCES

Carlton, R.G. and R.G. Wetzel. 1988. Phosphorus flux from lake sediments: effect of epipelic algal oxygen production. Limnology and Oceanography. 33(4):562–570.

Dunbar, J., S.M. Barns, L.O. Ticknor and C.R. Kuske. 2002. Empirical and Theoretical Bacterial Diversity in Four Arizona Soils Empirical and Theoretical Bacterial Diversity in Four Arizona Soils. Applied and Environmental Microbiology. 68(6):3035–3045.

Li, W., S.R. Joshi, G. Hou, D.J. Burdige, D.L. Sparks and D.P. Jaisi. 2015. Characterizing phosphorus speciation of Chesapeake Bay sediments using chemical extraction, 31P-NMR, and X‑ray absorption fine structure spectroscopy. Environmental Science and Technology. 49:203–211.

Locke, N. 2015. Classification of polyphosphate-accumulating bacteria in benthic biofilms. Pennsylvania State University, State College, Pennsylvania.

Mao, Y., D.W. Graham, H. Tamaki and T. Zhang. 2015. Dominant and novel clades of *Candidatus* Accumulibacter phosphatis in 18 globally distributed full-scale wastewater treatment plants. Scientific Reports. 5:11857.

Price, K.J. and H.J. Carrick. 2013. Effects of physical disturbance on phosphorus uptake in temperate stream biofilms. Inland Waters. 3:321–330.

Purahong, W., T. Wubet, G. Lentendu, M. Schloter, M.J. Pecyna, D. Kapturska, M. Hofrichter, D. Krüger and F. Buscot. 2016. Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. Molecular Ecology. 25(16):4059–4074.

Tamburini, F., V. Pfahler, E.K. Bunemann, K. Guelland, S.M. Bernasconi and E. Frossard. 2012. Oxygen isotopes unravel the role of microorganisms in phosphate cycling in soils. Environmental Science and Technology. 46(11):5956–5962.