The Pennsylvania State University

The Graduate School

Department of Civil and Environmental Engineering

CLASSIFICATION OF POLYPHOSPHATE-ACCUMULATING BACTERIA IN BENTHIC BIOFILMS

A Thesis in

Environmental Engineering

by

Nicholas Locke

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2015

The thesis of Nicholas Locke was reviewed and approved* by the following:

John Regan
Professor of Environmental Engineering
Thesis Advisor

William Burgos
Professor of Environmental Engineering
Chair of Civil and Environmental Engineering Graduate Programs

Anthony Buda
Adjunct Assistant Professor of Ecosystem Science and Management

*Signatures are on file in the Graduate School

ABSTRACT

Polyphosphate accumulating organisms (PAOs) are microorganisms known to store excess phosphorus (P) as polyphosphate (poly-P) in environments subject to alternating aerobic and anaerobic conditions. There has been considerable research on PAOs in biological wastewater treatment systems, but very little investigation of these microbes in freshwater systems. We hypothesize that putative PAOs in benthic biofilms of shallow streams where daily light cycles induce alternating aerobic and anaerobic conditions are similar to those found in EBPR. To test this hypothesis, cells with poly-P inclusions were isolated, classified, and described. Eight benthic biofilms taken from a first-order stream in Mahantango Creek Watershed (Klingerstown, PA) represented high and low P loadings from a series of four flumes and were found to contain 0.39 - 6.19% PAOs. A second set of eight benthic biofilms from locations selected by Carrick and Price (2011) were from third-order streams in Pennsylvania and contained 11-48% putative PAOs based on flow cytometry particle counts. Putative PAOs of the second sample set represented 61 unique operational taxonomic units with taxonomic classification including members of *Actinomycetales*, *Bacillales*, *Ignavibacteriales*, *Sphingomonadales*, *Burkholderiales*, and *Rhodocyclales*. Streams with similar putative PAO species richness were found to have similar trophic states. Canonical correlation analysis showed that some PAO groups were directly related to pH, temperature, conductivity, or stream cover. Four probes targeting the predominant PAO groups of *Ignavibacterium*, *Thauera*, *Comamonadaceae*, and *Pseudomonas* allowed the colocalization of poly-P and group identity. Members of *Thauera* and *Comamonadaceae* were confirmed as PAOs. These results greatly expand our understanding of in-stream PAO-facilitated P cycling and will inform P fate and transport models to better simulate in-stream P cycling in lotic systems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost I would like to acknowledge my supervisor, John Regan, for providing me a tremendous opportunity to perform high impact research in the field of environmental engineering. Jay has provided me with the tools necessary to perform productive, meaningful, and profound research in engineering and science. The skills I have learned as a graduate student will be carried with me throughout my future endeavors as an engineer.

My collaborators provided me with biofilm samples, knowledge, and data to make my research more meaningful. I thank Dr. Hunter Carrick at Central Michigan University for providing me with biofilm samples and stream data. I also thank his former student, Keith Price and his current student Shayna Taylor, for their work in sampling and analyzing biofilm data. Dr. Todd Walter and his student, Sheila Saia, at Cornell University performed additional research that provided insights to my own work, and make my results more meaningful. Dr. Anthony Buda at the USDA research lab in University Park, PA has provided the entire team with a tremendous amount of data from the FD-36 experimental watershed and additional information regarding polyphosphate accumulators.

The Huck Institute of the Life Sciences at Penn State (University Park, PA) provided me with flow cytometry and cell sorting capabilities, a critical component of my experiments. I thank the operators and technicians at this facility for their time, guidance, and expertise that has made my experiments and subsequent analyses a much smoother process.

The Research and Testing Laboratory (Lubbock, TX) performed the next-generation sequencing. They also performed a significant amount of the sequencing analysis, which in turn saved me a lot of time and helped me learn the steps of proper high-throughput data analysis.

Chapter 1

Introduction

In recent decades, anthropogenic nutrient inputs to streams, lakes, and marine environments have resulted water quality deterioration. Eutrophication is a natural process in which the productivity of a water resource slowly increases over time with nutrient inputs (Wetzel 2001). However, anthropogenic phosphorus (P) and nitrogen (N) inputs cause eutrophication to occur at a rate faster than ecosystems can adapt (Carpenter, Caraco et al. 1998). In freshwater systems, P is often the nutrient controlling primary productivity (Schindler 1977). Point source and non-point source (NPS) P inputs are both a concern for ongoing eutrophication of aquatic systems. Point-source pollution is often predictable and easily addressed using chemical and/or biological removal in wastewater treatment plants (WWTPs). However, NPS pollution is often complex and addressed through mitigation strategies that require a thorough understanding of P transport to be effective.

Phosphorus, after it is applied as fertilizer, is often immobilized by soils and/or taken up for biological growth. The labile P pool for most soils typically contains 50-75% inorganic P (Sharpley, Tiessen et al. 1987). Immobilized P builds up in soils, contributing to the phenomena called “legacy P” and “environmental P sinks” which can effectively mask anticipated improvements in water supply by providing a continuing supply of P to watershed runoff (He, Zhang et al. 2006, Sharpley, Jarvie et al. 2013). In addition to environmental P sinks, hydrologic events are also known to alter P availability as a result of microbial activity responding to varying wetting/drying conditions. These biogeochemical “hotspots” have been described as areas that show disproportionately high reaction rates relative to the surrounding area (McClain, Boyer et al. 2003). The connection between hydrologic models

and microbial P metabolisms is an area that remains poorly understood despite evidence that biological uptake and release of P occurs and may play a significant role in P cycling (Fitter 2005, Hupfer, Gloess et al. 2007, Hupfer and Lewandowski 2008).

The amount of P that is incorporated into biomass is a relatively predictable mechanism of the P cycle. However, there are occasions where certain bacteria will take up "luxury P" in excess of expected cell requirements (Khoshmanesh, Sharma et al. 2001). Such microbes, called polyphosphate accumulating organisms (PAOs), are critical in enhanced biological P removal (EBPR) from wastewater. Bacteria with polyphosphate inclusions have also been identified in lake sediments, although their relationship to PAOs in EBPR remains unknown (Khoshmanesh, Sharma et al. 2001, Hupfer, Gloess et al. 2007). If microbes with the PAO metabolism exist in stream beds, they have the potential to serve as another environmental P sink and act as a biogeochemical hotspot. Therefore, the mechanisms controlling excessive biological P uptake in the environment must be understood to model their contributions to the P cycle.

For polyphosphate accumulating organisms to remove P in EBPR, activated sludge cycles through alternating carbon-deficient aerobic and carbon-rich anaerobic phases. The cycling confers a selective advantage to PAOs, which anaerobically use poly-P as an energy source to accumulate organic substrates and then aerobically oxidize these organics for cell growth and poly-P accumulation. Interest in innovative nutrient removal technologies, such as EBPR, has increased as cases of freshwater eutrophication become more common.. However, P runoff from nonpoint sources (NPSs) is a significant contributor to freshwater deterioration as well (Hatt, Fletcher et al. 2004, He, Zhang et al. 2006, Acevedo, Oehmen et al. 2012). While EBPR cannot be used to remove P from NPSs, cyclic aerobic/anaerobic conditions have been identified in natural settings and may select for PAOs that are typically known to exist in wastewater treatment plants.

Evidence suggests that benthic biofilms attached to lake and stream sediments are subject to alternating anaerobic and aerobic cycles. In daylight, oxygenic phototrophs in biofilms at the surface of lake sediments can metabolize light and carbon dioxide, producing oxygen that can be used by PAOs for the aerobic phase of their metabolism. At night, benthic biofilms cease oxygenic photosynthesis, and net respiration results in a shift to anaerobic conditions where PAOs would be expected to release phosphate. This diurnally induced aerobic and anaerobic cycle is similar to what is found in engineered EBPR systems. Both processes cycle on a periodic basis and have high carbon availability in the anaerobic phase. In lake sediments, it has been found that P influx is associated with aerobic conditions during the day, while at night P release is associated with anaerobic conditions (Carlton and Wetzel 1988). With our knowledge of EBPR, it is possible that PAOs are responsible for this diurnal P flux phenomenon in streams.

There has already been a significant amount of research characterizing PAOs in EBPR activated sludge. Using similar techniques involving full-cycle rRNA analysis, which identifies microbial cells without cultivation, PAOs in benthic biofilms can be taxonomically classified. This research investigated putative PAO bacterial community composition in benthic biofilms and compared it with those found in EBPR to interpret the potential P accumulation pathways of PAOs. The hypotheses of this study were that PAOs are naturally selected for in benthic biofilms and are evolutionarily similar to PAOs from EBPR activated sludge. Experimental success will increase scientific understanding of P movement through watersheds, which in turn will lead to strategies to improve P mitigation and reduce rates of eutrophication.

Chapter 2

Literature Review

2.1 Enhanced Biological Phosphorus Removal

Current understanding of PAOs and the associated biological P accumulation metabolism relies on biochemical characteristics of the EBPR wastewater treatment process. Biological phosphorous removal is accomplished by circulating activated sludge through anaerobic and aerobic phases and wasting a fraction of activated sludge following aerobic treatment. The conventional system is referred to as an anaerobic/oxic (A/O) process, having only one anaerobic and one aerobic phase (Tchobanoglous, Burton et al. 2002) (Figure 2-1). This base system is often modified by adding an anoxic tank between the anaerobic and aerobic phases to achieve denitrification, referred to as an A^2O process. Biological accumulation of P in EBPR has been an effective and economical alternative to chemical precipitation for removing P from wastewater. While conventional waste activated sludge typically contains 2-3% P by mass, waste activated sludge from EBPR processes typically contains 4-5% P by mass and decreases total effluent P to < 1 mg/L (Seviour, Mino et al. 2003). Under an A^2O arrangement, treatment facilities can achieve 94% removal of nitrogen and < 1 mg/L effluent P in one process (Barnard 1975). Removal efficiencies are sensitive to the influent BOD/P ratio, which typically must be 15-20 g/g for an A/O process and 20-25 g/g for an A^2O process (Grady, Daigger et al. 2011).

Polyphosphate-accumulating organisms are unique in their ability to couple the aerobic creation of polyphosphate (poly-P) chains from inorganic ortho-P monomers with the anaerobic creation of polyhydroxyalkanoates (PHAs), with the net effect of storing P beyond

typical cell requirements. The EBPR process was widely implemented before PAO metabolism was fully understood. As a result, EBPR exhibited failures and was commonly supplemented with chemical P precipitation. The PAO metabolism was initially modeled from chemical profiles of conventional EBPR processes (Figure 2-1). Components of EBPR critical to the PAO metabolism include activated sludge recycle, an anaerobic phase preceding the aerobic phase, and substrate input into the anaerobic phase. Overall, anaerobic volatile fatty acid (VFA) uptake serves as the carbon source, aerobic PHA breakdown serves as the electron donor, and oxygen serves as the terminal electron acceptor (Figure 2-2). In an A^2O system, specialized denitrifying PAOs (DPAOs) use nitrate as the electron acceptor in a secondary anaerobic phase receiving return activated sludge from the aerobic phase (Zeng, Saunders et al. 2003). DPAOs exist in A/O systems, but their presence is undesirable because denitrifying bacteria decrease the supply of organic matter (Kortstee, Appeldoorn et al. 1994).

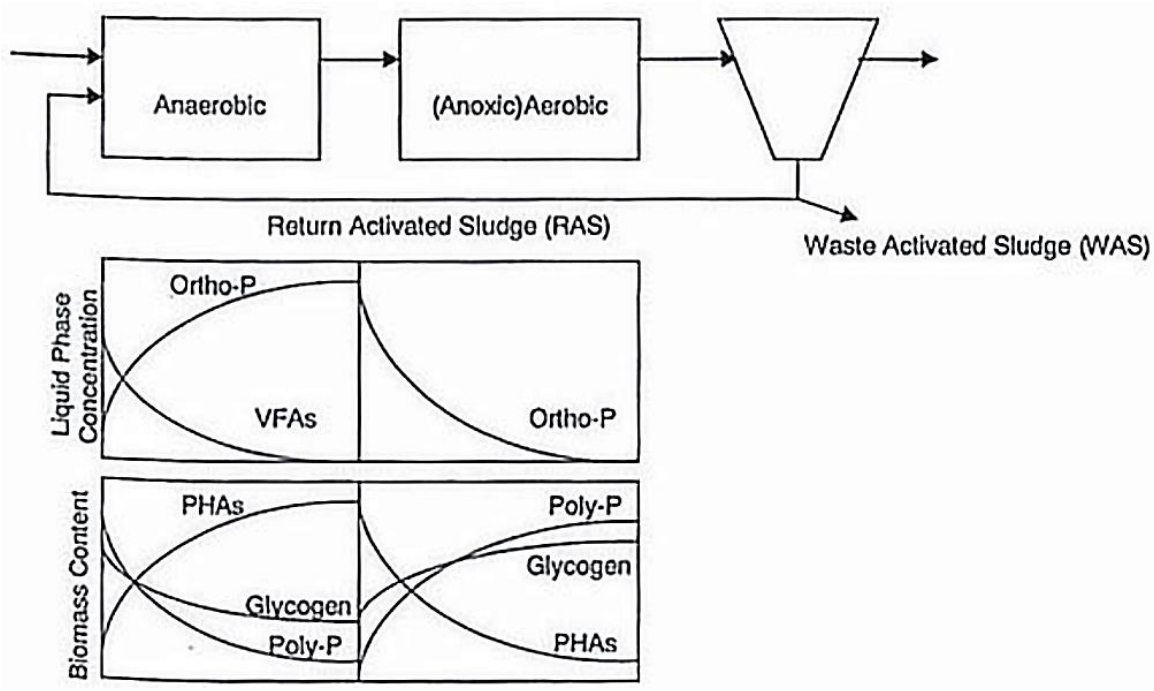

Figure 2-1. A diagram and typical chemical profiles of the conventional EBPR process (Wisconsin Department of Natural Resources Wastewater Operator Study Guide).

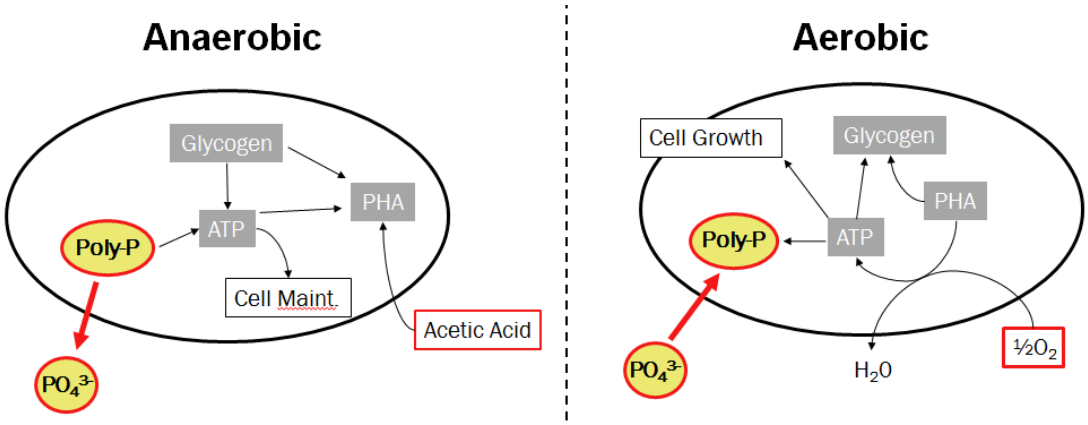

Figure 2-2. The polyphosphate accumulating metabolism (Modified from Comeau, Hall et al. 1986).

In the anaerobic phase, the hydrolysis of poly-P and glycogen generates ATP molecules that enable the uptake of carbon substrates, such as acetic acid. To maintain a balance of charge, acetic acid is deprotonated and both the proton and phosphate molecules are released into the environment. Acetate is converted to PHA through fermentation after glycogen undergoes glycolysis. Cells intracellularly store PHAs until the aerobic cycle and use ATP generated in glycolysis for cell maintenance (Comeau, Hall et al. 1986)

The aerobic phase follows the anaerobic phase. At this point, the wastewater has low BOD content (as a result of anaerobic VFA sequestration) and high P content. In the aerobic phase, PAO metabolism shifts to regenerate poly-P and glycogen stores. Electron acceptors (typically oxygen and/or nitrate) enable respiration to commence, and the resultant energy (i.e., ATP) is used to form poly-P from phosphate ions and form glycogen and cell mass from PHAs. Phosphate ions are linked together to form large poly-P chains using ATP generated from the electron transport chain. PHAs serve as the carbon source for aerobic respiration, which includes glycolysis followed by the citric acid cycle.

In competition with the polyphosphate accumulating metabolism, researchers have discovered a group of microorganisms, called glycogen accumulating organisms (GAOs), that compete with PAOs for carbon substrate in the EBPR process (Crocetti, Banfield et al.

2002). GAOs have a metabolism similar to that of PAOs, but instead of accumulating P as poly-P in the aerobic phase for energy production in the anaerobic phase, GAOs rely solely on glycogen stores to provide the energy necessary for survival. Some EBPR reactor failures have been attributed to a microbial community shift favoring GAOs and low concentrations of PAOs. This shift in microbial community populations is thought to be caused by factors such as carbon substrate type and pH. PAOs were found to take up propionate at 76.6% the rate of acetate uptake while the uptake rate of propionate by GAOs was 5.4% that of their acetate uptake rate, suggesting that adding propionate as a supplemental carbon source can reduce EBPR failures (Oehmen, Yuan et al. 2004). If the pH of EBPR is held constant, GAOs become preferred over PAOs in the activated sludge. Reactors with a pH held constant at 7 in both aerobic and anaerobic phases were found to be less efficient for P removal than reactors with uncontrolled pH, which typically measured 7.5 in the anaerobic phase and 8.6 in the aerobic phase (Serafim, Lemos et al. 2002).

2.2 Techniques to Study PAOs

No isolated cultures are observed to have all of the characteristics expected from the PAO metabolism. Therefore, the full-cycle rRNA approach (Amann, Ludwig et al. 1995) has been used to characterize and classify putative PAOs in EBPR activated sludge. In the full-cycle rRNA approach, mixed-culture samples (i.e. activated sludge) are enriched for PAOs, enrichment cultures are sequenced using 16S rRNA methods, and sequences are used to design PAO-targeted oligonucleotide probes. Fluorescent in-situ hybridization of probes to the original samples confirms that the identified species are cells that contain poly-P. As molecular identification techniques have matured over the years, so have the techniques used to physically enrich samples for PAOs, sequence putative PAO cells, and analyze the sequencing data to develop probes.

2.2.1 Poly-P detection and PAO enrichment in mixed cultures

Molecules of poly-P are commonly stained using the 4',6-diamidino-2-phenylindole (DAPI) staining technique because of its characteristic yellow fluorescence when attached to poly-P (Allan and Miller 1980, Tijssen, Beekes et al. 1982). DAPI is typically used as a nucleic acid stain based on its fluorescence maximum of 456 nm when bound to DNA and excited with ultraviolet light. However poly-P shifts the fluorescence emission maximum of DAPI to 526 nm. Therefore, DAPI serves as a single stain to visualize cells containing poly-P and distinguish them from other cells in a mixed culture. Crossover of the DAPI-DNA and DAPI-P emission spectra requires the poly-P emission to be detected at approximately 575 nm and the DNA emission to be detected at approximately 456 nm to allow separation of the two signals (Streichan, Golecki et al. 1990, Kawaharasaki, Tanaka et al. 1999, Hung, Peccia et al. 2002). Alternatively, to minimize this crossover, a violet excitation can be used instead of the typical ultraviolet excitation to reduce the DAPI-DNA emission and maintain a relatively strong DAPI-poly-P emission (Aschar-Sobbi, Abramov et al. 2008). This technique was used to visualize and quantify low concentrations of poly-P in the mitochondrion of neurons.

The DAPI staining approach is commonly used to quantify poly-P, a technique that has been applied to environmental and activated sludge samples. Results from direct cell counts have shown the fraction of PAOs making up the total cells in EBPR activated sludge ranging from 9 - 36% (Zilles, Hung et al. 2002, He, Gall et al. 2007). These values are consistent with counts of PAOs from electron transmission microscopy, which accounted for 30-50% of the total cells (Streichan, Golecki et al. 1990). In addition to direct cell counts, relative quantities of poly-P can be measured by observing increases in the fluorescence emission spectrum resulting from higher poly-P concentrations and/or longer poly-P molecule length (Aschar-Sobbi, Abramov et al. 2008, Diaz and Ingall 2010). This technique

has been applied to phytoplankton samples to measure total intracellular poly-P across a marine transition zone (Martin and Van Mooy 2013).

The DAPI staining approach enables not only quantification of PAOs, but also separation of PAOs from other cells in a mixed culture. This separation is achieved using cell sorting associated with flow cytometry technologies (Amann, Binder et al. 1990). This method has been used to quantify and produce physically enriched cultures of PAOs from EBPR activated sludge (Hung, Peccia et al. 2002, Zilles, Peccia et al. 2002). Density gradient centrifugation was also suggested as a physical enrichment method, based on the higher density of cells containing poly-P inclusions (Hung, Peccia et al. 2002). Density gradient centrifugation resulted in enrichment cultures containing up to 48% PAOs, while flow cytometry resulted in enrichment cultures of up to 72% PAOs. Flow cytometry is the most effective physical enrichment method and is commonly used for physical enrichment on environmental samples. In addition to physical enrichment, PAOs are often selectively enriched in sequencing batch reactors.

2.2.2 Sequencing Approaches

High-throughput DNA sequencing technologies such as 454 pyrosequencing (Life Technologies), MiSeq (Illumina, Inc.), and Ion Torrent (DNA Electronics) have the ability to generate considerable nucleic acid sequence datasets for use in classifying cells of a mixed culture. These “Next-Generation Sequencing” (NGS) technologies can sequence tens of thousands of DNA amplicons (<600 bp) simultaneously. Sequencing highly variable regions (called “V” regions) of the 16S rRNA gene achieves species-level identification and can allow the determination of evolutionary relationships (Woese 1987). These variable regions are targeted in NGS, although poor phylogenetic depth is a drawback in this type of analysis since only a small portion of the 16S rRNA gene is sequenced (Yarza, Yilmaz et al. 2014).

Commonly used primer pairs include 349f/515r, which targets a 185 bp region including the 66 bp V3 region of the 16S rRNA gene and achieves 91.2% coverage of known *Bacteria* sequences; and 564f/785r, which targets a 239 bp region including the 106 bp V4 region and covers 89% of *Bacteria* (Klindworth, Pruesse et al. 2013). Analyses targeting the V3-V4 region typically return higher coverage of bacteria phyla, while analyses targeting the V1-V2 region typically return lower phyla coverage with better species resolution (Yarza, Yilmaz et al. 2014). Despite the drawbacks associated with the limited sequence length capabilities of NGS, it has become a powerful tool to analyze changes in microbial communities of environmental samples due to the sheer number of sequences it generates.

2.2.3 Fluorescent in-situ Hybridization

Isolated bacteria exhibiting all the characteristics of PAOs have yet to be discovered. Therefore, PAOs must be studied using culture-independent techniques that target cells without cultivation. Fluorescent in-situ hybridization (FISH) is commonly used to identify PAOs in activated sludge and involves the design of an oligonucleotide probe that complements a unique rRNA sequence region (typically in the 16S rRNA) of the desired target organisms (Amann, Ludwig et al. 1995). Probes are fluorescently labeled, hybridized with the rRNA of target cells, and then washed from the sample preparation. Due to the abundance of ribosomes in the cytoplasm, the hybridized probe confers fluorescence to the target cells. By combining FISH with DAPI staining, it is possible to simultaneously visualize cells containing poly-P inclusions and cells carrying the PAO-targeted probe fluorescence to confirm that the selected group accumulates poly-P according to the expected metabolism.

The success of FISH depends on accessibility of the targeted rRNA region within the ribosome matrix and the proper hybridization conditions. Inaccessible regions can reduce FISH fluorescence to 1% of observed maxima (Fuchs, Wallner et al. 1998). Helper probes,

which are designed to target sequences adjacent to the probe target and disrupt secondary rRNA structure, can promote hybridization to inaccessible targets and increase fluorescence by up to 56.7% (Fuchs, Glockner et al. 2000). In addition, increasing the hybridization time to > 12 hrs can allow the hybridization reaction to approach equilibrium, which can improve fluorescence for inaccessible regions having poor reaction kinetics (Yilmaz, Okten et al. 2006). The hybridization stringency also needs to be optimized to prevent the probe from hybridizing to non-target sequences that are similar to the target sequence. This is accomplished by adding formamide to the buffer, which serves to destabilize double-stranded nucleic acids. With proper optimization experiments, the ideal formamide concentration can be determined that discriminates target from non-target cells.

2.3 PAO Classification Review

Acinetobacter spp. were the first to be isolated from EBPR activated sludge and observed to accumulate large amounts of P (Fuhs and Chen 1975). When culture-independent techniques to study individual populations in mixed cultures were limited, members of *Acinetobacter* were thought to be responsible for P removal as their cell dry mass would contain 7-10% P (Deinema, van Loosdrecht et al. 1985). However, *Acinetobacter* isolates did not accumulate acetate under anaerobic conditions in a chemostat mimicking EBPR conditions (Tandoi, Majone et al. 1998). In addition, pure culture performance tests suggested *Acinetobacter* was not present in high enough concentration to be responsible for P removal in EBPR (Crocetti, Hugenholtz et al. 2000). Studies have shown levels of *Acinetobacter* as low as 8% on average in EBPR activated sludge through FISH techniques (Wagner, Amann et al. 1994). As a result, relying on bacterial isolates from EBPR activated sludge was insufficient to fully characterize PAOs in EBPR activated sludge.

Using techniques to study individual cells without cultivation, members of *Betaproteobacteria*, *Actinobacter,* and *Rhodocyclus* were identified as PAOs responsible for EBPR. Clusters of PAOs classified under *Betaproteobacteria* became known as *Candidatus* Accumulibacter phosphatis (referred to as *Accumulibacter*) (Crocetti, Hugenholtz et al. 2000). The *Accumulibacter* group was highly concentrated in “hyper-P-removing” sludges where phosphate made up 15% of the dry weight of the biomass. Development of 16S rRNA targeted oligonucleotide probes enabled easy identification of *Accumulibacter* species in subsequent EBPR studies (probes PAO462, PAO651, and PAO846, Table 2-1). Later investigations confirmed the involvement of *Rhodocyclus*-related bacteria (a sub-group of *Betaproteobacteria*) in EBPR, and the *Accumulibacter* targeted probes were improved to cover additional *Rhodocyclus* relatives observed as PAOs (probes PAO462b, PAO846b, Table 2-1) (Zilles, Peccia et al. 2002). Both lab bench and full-scale EBPR experiments have shown that *Rhodocyclus*-related PAOs are selected for when wastewater is subjected to anaerobic/aerobic cycles, and are common among EBPR processes worldwide (Zilles, Peccia et al. 2002).

Additional PAO groups have been identified in EBPR, although these groups likely occupy different metabolic niches that contribute to the overall stability in EBPR. *Tetrasphaera*-related bacteria have been identified as candidate PAOs that are able to ferment glucose and denitrify (Kristiansen, Nguyen et al. 2013). Uncultured groups of *Halomonas*-related bacteria referred to as *Candidatus* Halomonas phosphatis have been found in full-scale EBPR in concentrations similar to that of *Accumulibacter* (Nguyen, Nielsen et al. 2012). The *Halomonas*-PAOs used low molecular weight substrates, have the ability to consume ethanol, and are unable to denitrify. Probe Hal180 was developed to identify *Halomonas*-PAOs. Members of *Rhodocyclales* closely related to the *Decholoromonas* genus were found to be more numerous than *Accumulibacter* species under low oxygen conditions

Table 2-1. Oligonucleotide probes used to identify PAOs in EBPR.

Name	**Sequence (5'-3')**	**Target**	**Reference**
RHC175	TGCTCACAGAATATGCGG	*Rhodocyclus* cluster	Hesselmann et al. 1999
RHC439	CNATTTCTTCCCCGCCGA	*Rhodocyclus* cluster	Hesselmann et al. 1999
PAO462	CCGTCATCTACWCAGGGTATTAAC	Candidatus *Accumulibacter*	Crocetti et al. 2000
PAO651	CCCTCTGCCAAACTCCAG	Candidatus *Accumulibacter*	Crocetti et al. 2000
PAO846	GTTAGCTACGGCACTAAAAGG	Candidatus *Accumulibacter*	Crocetti et al. 2000
actino_1011	TTGCGGGGCACCCATCTCT	*Tetrasphaera elongata* clones	Liu et al. 2001
PAO462b	CCGTCATCTRCWCAGGGTATTAAC	*R. tenuis* subgroup	Zilles et al. 2002
PAO846b	GTTAGCTACGGYACTAAAAGG	*Rhodocyclus Tenuis* subgroup	Zilles et al. 2002
Actino-221a	CGCAGGTCCATCCCAGAC	*Actinobacteria*-PAOs	Kong et al. 2005
Actino-658a	TCCGGTCTCCCCTACCAT	*Actinobacteria*-PAOs	Kong et al. 2005
Hal180	CCTGCTTTCTTCCTCAGAACGT	Candidatus *Halomonas phosphatis*	Nguyen et al. 2012
Tet1-266	CCCGTCGTCGCCTGTAGC	*Tetrasphaera*-PAO	Nguyen et al. 2011
Tet2-892	TAGTTAGCCTTGCGGCCG	*Tetrasphaera*-PAO	Nguyen et al. 2011
Tet2-174	GCTCCGTCTCGTATCCGG	*Tetrasphaera*-PAO	Nguyen et al. 2011
Tet3-654	GGTCTCCCCTACCATACT	*Tetrasphaera*-PAO	Nguyen et al. 2011

Table 2-2. Current primer sets used to identify PAOs in environmental samples.

Name	Sequence	Target group	Reference
518f	CCAGCAGCCGCGGTAAT	*Candidatus* Accumulibacter	He et al. 2007
PAO-846r	GTTAGCTACGGCACTAAAAGG	*Candidatus* Accumulibacter	He et al. 2007

(Zilles, Peccia et al. 2002). However, additional groups related to *Decholoromonas* were observed to anaerobically accumulate carbon substrate but not aerobically accumulate poly-P (Ahn, Schroeder et al. 2007). It was suggested that these groups only perform the expected PAO metabolism when certain conditions are met (Seviour and Nielsen 2010).

Cells having the glycogen accumulating metabolism have also been classified. GAO members of *Gammaproteobacteria* are referred to as "*Candidatus* Competibacter phosphatis" (Crocetti, Banfield et al. 2002). GAOs have also been found to be closely related to the *Alphaproteobacteria* groups *Sphingomonadales* and *Defluviicoccus* (Oehmen, Lemos et al. 2007).

With both PAOs and GAOs taxonomically classified from EBPR systems, high-throughput sequencing approaches have been used to identify conditions that highly enrich for PAOs in EBPR and repress GAOs (Weissbrodt, Schneiter et al. 2013). Similar studies have confirmed PAO preference for propionate as a VFA and have found that some PAO groups such as *Tetrasphaera spp*. and *Dechloromonas* are positively correlated with *Competibacter* groups (Tu and Schuler 2013).

2.5 Indirect Evidence of PAOs in Benthic Biofilms

Despite methods for identification and characterization of bacterial PAOs in EBPR activated sludge being thoroughly developed, there is very little research characterizing PAOs in natural settings and no research assessing their role in P cycling. Studies of microbial communities associated with the rhizosphere show a wide variety of microbial phosphate accumulators and phosphate solubilizers (Khan, Zaidi et al. 2007). In addition, PAOs in aquatic sediments have been investigated for their ability to influence P sedimentation (Hupfer, Gloess et al. 2007). Polyphosphate accumulators are known to exist in the environment, but in many cases their relationship to EBPR PAOs remains unknown. Bacterial PAOs have yet to be studied in benthic stream biofilms especially in systems where P inputs can be high and critical to eutrophication problems. For benthic biofilms, night/day cycles induce anaerobic/aerobic microenvironments, respectively, in a cycle similar to EBPR. As a result, it is likely similar microorganisms exist in these natural environments, and it is critical to characterize those communities to gain a better understanding of P fate and transport.

Hydrologic events are known to change nutrient availability as a result of microbial activity responding to varying wetting/drying conditions. Biogeochemical “hotspots” have been described as areas that show disproportionately high reaction rates relative to the surrounding area (McClain, Boyer et al. 2003). For example, hot moments are found to occur in environments subject to long dry periods between rain events. Connecting PAO metabolism with the hot moment concept, PAOs might take up and store P during a dry period (aerobic conditions), and release P during a wetting event (anaerobic conditions) which generates a hot moment where P becomes available for plant uptake and metal complexation. Similarly in benthic biofilms, hot moments can occur at the sediment-water interface in lakes and rivers. In fact, cyclic P uptake and release associated with light and

dark periods, respectively, have been shown in the uppermost layer of lake sediments (Carlton and Wetzel 1988). In addition, oxygen consumption was correlated with diel light cycles as well with higher DO concentrations observed during the light phase and lower concentrations observed during the dark phase (Figure 2-3). The dark phase (anaerobic conditions) may serve as a hot moment where P is released by benthic microorganisms and becomes available.

In a study by Price and Carrick (unpublished), a strong linear correlation was reported between the $\log_{10}$ of total phosphorus concentrations and the $\log_{10}$ of total chlorophyll concentration (a proxy of stream productivity) in benthic biofilms. Under P-limiting conditions, the biofilms were observed to accumulate more poly-P. In addition, Price and Carrick (2014) showed that biofilms will accumulate large amounts of P after initial loading events and maintain this P pool indefinitely. These preliminary data show evidence that microorganisms in biofilms accumulate P in preparation for starvation events, a characteristic that is expected of PAOs.

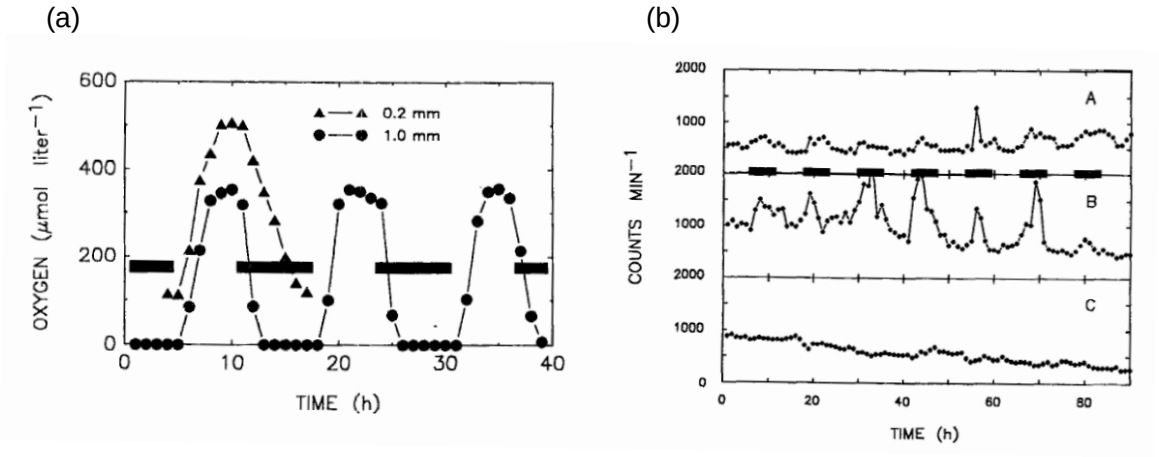

Figure 2-3. Concentrations of oxygen (a) and phosphorus (b) from benthic lake sediments (Carlton and Wetzel 1988).

Chapter 3

Investigation of Polyphosphate Accumulating Bacteria

Analyses were carried out on two different sets of benthic biofilms. The first set (PA streams survey, Table 3-1) contained eight samples that were previously studied to explore benthic P-uptake in response to P loading (Price and Carrick 2014). This sample subset was derived from 43 stream biofilms that were studied in experiments evaluating the benthic chlorophyll-a production with respect to nutrient loading (Carrick, Scanlan et al. 2009). The eight biofilms used for this study are representative of a range of geochemical conditions in Pennsylvania and the greater Mid-Atlantic coast region (Figure 3-1). In addition to the stream samples, an activated sludge sample obtained from the University Area Joint Authority Wastewater Treatment Facility (State College, PA) anoxic tank was analyzed as a positive control with known PAOs.

The second set of benthic biofilms (USDA FD-36,Table 3-2) was part of an ongoing experiment exploring the biofilm mass and stream trophic state response to varying P loadings. The sample set contained eight benthic biofilms from the USDA FD-36 experimental watershed (Mahantango Creek Watershed, PA) grown for two weeks in July 2014 on porous porcelain tiles attached to vials containing media releasing P of various concentrations. The eight biofilm samples represent four different locations along a stream in the watershed (Figure 3-2), immediately downstream of monitoring flumes (F1, F2, F3, and F4), and high (P2) and low (P4) P loadings at each location. Phosphorus loadings were designed as in situ enrichment systems (ISESs) at each flume. Biofilms were grown on porous porcelain tiles secured to vials containing either 2.5 µM (loading P2) or 5 µM (loading P4) of P as NaH_2PO_4.

Table 3-1. This investigation included eight samples from previously well-characterized streams. Data points are single point measurements at the time of biofilm collection.

Sample	Areal Chl-a (mg/m^2)	Temp. (°C)	Conduc-tivity (µS/cm)	DO* (mg/L)	Depth** (m)	pH	Cover (%)
Penns Creek	36.4	8.4	268	12.4	0.215	5.5	52.7
Cooks Creek	49.1	13	215	0.40	0.057	7.0	27.2
East Hickory Creek	55.4	7.2	30	12.8	0.109	5.4	55.3
Tionesta Creek	62.3	7.2	42	13.7	0.254	5.2	0.16
Tunkhannock Creek	178	7.3	108	0.67	0.054	5.7	24.9
Red Clay Creek	493	14	320	4.58	0.045	6.3	16.5
Spring Creek	264	9.3	277	10.0	0.086	5.5	21.2
Cowanesque River	672	10	141	-0.19	0.063	6.8	19.1

*Sensors for DO found to be malfunctioning. Values may thus be incorrect.

Table 3-2. The total P of the eight biofilm samples from an ongoing study at Mahantango Creek USDA FD-36 watershed.

Sample	Total P (mg/m^2)
Flume 1, P load 2	5.45
Flume 1, P load 4	10.3
Flume 2, P load 2	4.93
Flume 2, P load 4	8.11
Flume 3, P load 2	8.53
Flume 3, P load 4	20.8
Flume 4, P load 2	3.69
Flume 4, P load 4	8.77

The overall experimental design followed a modified full-cycle rRNA analysis (Amann, Ludwig et al. 1995). In full-cycle rRNA analysis environmental samples are first enriched for target cells then sequenced and classified based on rRNA gene sequences. Oligonucleotide probes targeting rRNA sequences were then developed from the sequencing data to identify the target cells in the original sample and in future experiments through FISH. As related to the experiments in this thesis, PAOs in the original biofilm samples were physically separated from non-PAOs using flow cytometry and simultaneous cell sorting (referred to as fluorescence activated cell sorting, “FACS”). This technique served as a basis to screen samples for poly-P accumulating cells and produce mixed cultures containing mostly PAOs. DNA of the separated putative PAO cells was extracted, and high-throughput sequencing revealed the bacterial species most closely related to the sorted PAOs. Due to low cell counts in the collected fraction of sorted cells and subsequent low DNA concentrations, only the bacterial fraction of the sorted cells was targeted during sequencing. Oligonucleotide probes target predominant PAO groups were developed from the sequencing data and fluorescent in situ hybridization experiments were used to confirm these groups are PAOs in the benthic biofilms.

3.1 Sampling

The eight PA stream samples were sampled from three rocks within riffle sections for each of the eight streams (Figure 3-1). The biofilms growing on each rock were removed by repeated scrubbing with a modified grout brush, rinsed with distilled water, and transferal to a plastic container. A subsample for each stream, used for this study, was preserved in ~3.5 % paraformaldehyde and stored at -20 °C. At the time of sampling, the stream ambient data for productivity, temperature, conductivity, dissolved oxygen (DO), depth, and pH were measured. Biofilm P uptake, alkaline phosphatase activity (APA), nitrate reductase activity,

poly-P, particulate P, primary production, areal C, and areal N were measured for each biofilm sample (Price and Carrick 2014).

Stream biofilm samples from the USDA FD-36 experimental watershed were collected by removal of whole tiles from the ISES vials and storage in Whirl-Pak bags. An aliquot of biofilm from each tile was scraped and transferred to a centrifuge tube containing stream water. Upon returning to the lab, samples were frozen at -20°C to reduce the risk of changes in the microbial community structure.

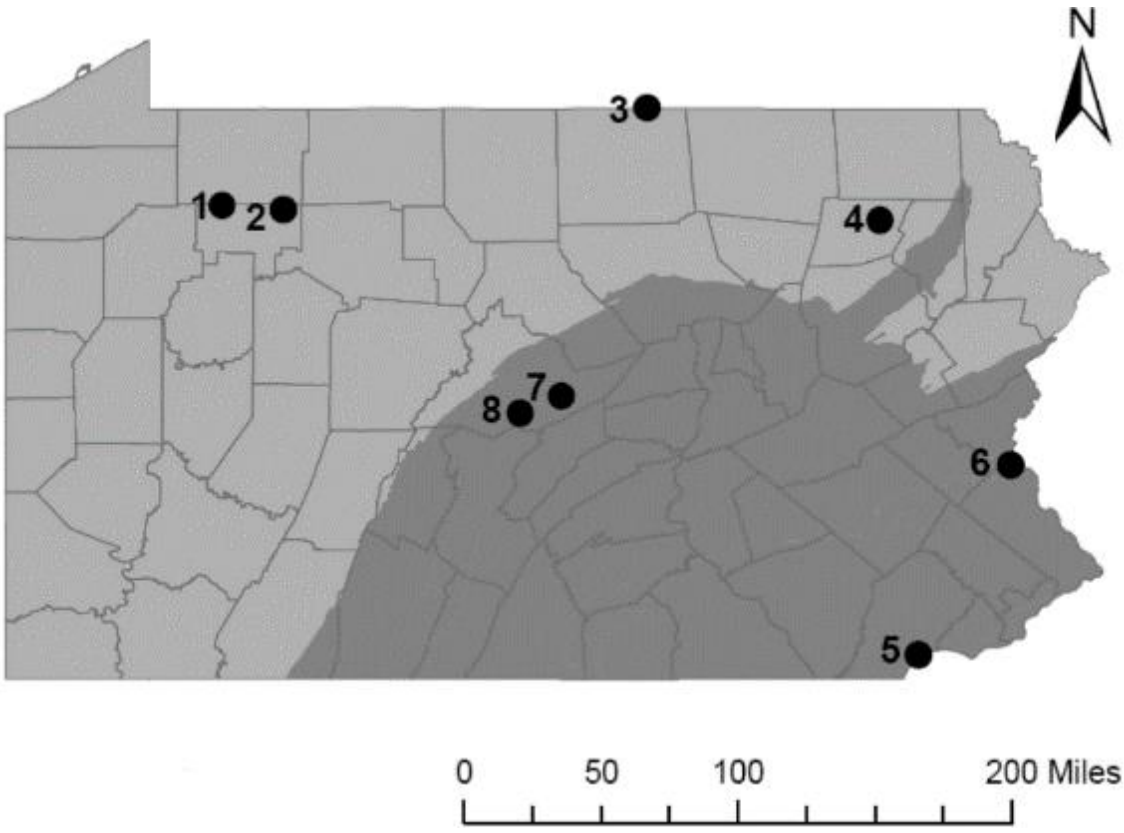

Figure 3-1. PA stream samples are from (1) East Hickory Creek, (2) Tionesta Creek, (3) Cowanesque River, (4) Tunkhannock Creek, (5) Red Clay Creek, (6) Cooks Creek, (7) Penns Creek and (8) Spring Creek. Image adapted from Price and Carrick (2014).

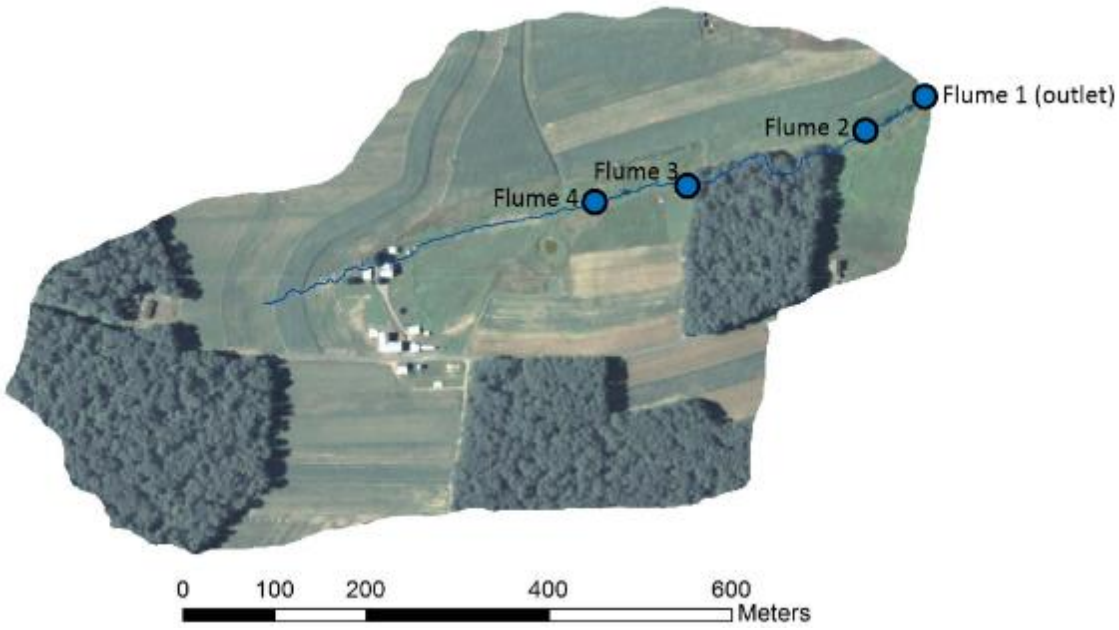

Figure 3-2. The FD-36 experimental watershed aerial image with flume locations. Water flow moves in the direction of flume 4 to flume 1.

Cell fixation using 3-4% formaldehyde prior to downstream analyses is commonly used to preserve cell morphology, permeabilize cell walls for probe penetration, and inhibit enzyme activity that can degrade target molecules (Seviour and Nielsen 2010). All biofilm samples were preserved by fixation in 3 % formaldehyde solution and long-term storage in 1:1 (v/v) ethanol to PBS solution at -20°C. Cells were recovered for analysis by pelleting and re-suspension in 1x PBS solution (pH 7.2).

3.2 Flow Cytometry and Cell Sorting

All samples were physically enriched for PAOs using flow cytometry. Cells were stained in 5 μg/mL DAPI in the dark for three hours. To disrupt the biofilm and individualize cells, PA stream samples were passed through a 21 gauge needle fifty times and samples from USDA FD-36 were exposed to ultrasonic radiation at 80 W for 2 sets of 45 second

pulses with a 15 second pause between sets. Disruption approaches differed because additional testing showed that ultrasonication produced better disruption results (data not shown). Samples were stored at 4°C in the dark until analysis. Immediately prior to processing, all samples were passed through a 35 µm cell strainer to prevent clogging of the flow cytometer orifice.

Samples were analyzed at the Penn State Huck Institute of the Life Sciences Microscopy and Flow Cytometry Facility. PA stream samples were processed using a BD Influx cell sorter (BD Biosciences) and USDA FD-36 samples were processed using a MoFlo Astrios cell sorter (Beckman Coulter). Sample sets were processed on different flow cytometers because of changes to on-campus machinery. BD Influx data was collected using a violet (405nm) excitation laser and a 550/50 nm emission sensor. MoFlo Astrios data was collected using a violet (405nm) excitation laser and a 546/20 nm emission sensor. Both experimental sets had forward scatter and side scatter measurements collected.

Sorting experiments were set up using an unstained sample set as the negative control, enabling compensation for autofluorescence. The DAPI-stained cells were compared to the negative control and served as the samples to be physically enriched for PAOs. Stained cells showing a significant increase in fluorescence relative to unstained cells were targeted for sorting. Sorting gates maintained the same general shape, although sample-dependent adjustments were made to compensate for autofluorescence differences among samples. Sorted cells were dispensed into 15 mL centrifuge tubes containing 1x PBS at 4°C. The sorting process continued until no sample remained or a sufficient number of cells was collected (~500,000 particles).

3.3 Microscopy

Phase contrast and epifluorescent microscopy were performed using a Zeiss Axiophot microscope with halogen and mercury-based lamps. A filter slider contained sets for fluorescent tags Cy3, FITC, and DAPI emissions for DNA and poly-P (Appendix A). The microscope was equipped with an AxioCam MRm camera for capturing black and white photos.

Suspended cell cultures were prepared for microscopy by allowing samples to air-dry to slides. Cells were fixed to the slide by a 50-80-100% ethanol dehydration gradient for 2 min. in each step. After the final dehydration step, slides were air-dried. DAPI stain was applied at a concentration of 1 μg/mL and cells were stained for 1 hour in the dark at room temperature. A rinse with 4°C distilled water removed the DAPI solution, and slides were allowed to air dry.

3.4 DNA Extraction

The sorted fraction of all eight stream samples and one activated sludge sample taken from the EBPR anoxic tank of the University Area Joint Authority Wastewater Treatment Plant (State College, PA) went through DNA extraction. In addition, an aliquot of the original biofilm sample of the six samples that were successfully sequenced (see 4.2) went through DNA extraction. The MoBio PowerSoil DNA Isolation and Purification kit was used to extract DNA following the procedures outlined in the manufacturer's instructions. The initial sample transferred to the provided lysis tube was 0.250 mL of culture suspended in 1x PBS. The final elution step was performed with 100 μL of solution C6 (10 mM tris) and the extracted DNA was stored at -20°C in the provided 2 mL collection tubes. Concentration and

quality of DNA was measured using a NanoDrop 2000c following the manufacturer's instructions.

3.5 Metagenomic Sequencing

The DNA of putative PAOs from the PA stream biofilm samples was sequenced by the Research and Testing Laboratory (Lubbock, TX) on the MiSeq platform (Illumina). The V1 region of the 16S rRNA gene was amplified using the 28f/519r primer set (Table 3-3). Amplifications were performed as 25 μL reactions with Qiagen HotStarTaq master mix, 1 μL each primer at 5 μM (final conc. 0.2 μM), and 1 μL of template DNA. Reactions occurred on an ABI Veriti thermocycler (Applied Biosystems, Carlsbad, CA) under the following thermal profile: Initial denaturation at 95°C for 5 min., then 35 cycles of denaturation at 94°C for 30 sec., annealing at 54°C for 40 sec., and elongation at 72°C for 1 min. followed by final extension at 72°C for 10 min. and a 4°C hold. Amplicons were sequenced using a 2x300 bp paired-end read.

Sequencing data were processed largely by a pipeline developed by the Research and Testing Laboratory (Lubbock, TX). Summarizing their methodology, reads went through the major stages of denoising, chimera checking, and microbial diversity analysis. Analysis included quality checking, sequence clustering, taxonomic identification, and microbial diversity analysis. The FASTQ formatted read files were merged using the PEAR Illumina paired-end read merger (Zhang, Kobert et al. 2014). Prefix dereplication and clustering at 4% divergence were performed using their respective USEARCH algorithms (Edgar 2010). OTU selection was performed using the UPARSE algorithm (Edgar 2013) and chimera detection was performed with UCHIME in de novo mode (Edgar, Haas et al. 2011). Each read was mapped to its corresponding nonchimeric cluster using the USEARCH global alignment algorithm (Edgar 2010). Each read base is corrected using the consensus sequence of each

Table 3-3. Primers used for metagenomic analysis on the Illumina MiSeq platform.

Name	Sequence	Target group
28f	GAGTTTGATCNTGGCTCAG	Most *bacteria*
519r	GTNTTACNGCGGCKGCTG	
28f	GAGTTTGATCNTGGCTCAG	Most *bacteria*
388r	TGCTGCCTCCCGTAGGAGT	

cluster as a guide and the corrected sequences are then written to an output FASTA file. The output FASTA file was demultiplexed and clustered into OTUs using the UPARSE algorithm (Edgar 2013). The centroid sequence from each cluster was run against the USERACH global alignment algorithm (Edgar 2010) and taxonomic identification was performed by a program developed by the Research and Testing Laboratory (Lubbock, TX). Taxonomic levels where there was not at least a 51% consensus were marked “unknown”. Final analysis files showed the taxonomic classification and number of hits for sequences in each sample.

3.6 Sequencing Data Analysis

Output MiSeq analysis files were used to perform statistical analyses. Heatmaps with associated dendrograms were generated using the “Heatplus” R package (Ploner 2014). The OTU FASTA files were aligned using the MUSCLE algorithm (Edgar 2004). The alignment file was uploaded to MEGA to manually select oligonucleotide probes targeting populations found in all six benthic biofilms. Probes were checked for specificity using RDP SeqMatch (Cole, Wang et al. 2014) and the properties were calculated using OligoCalc. The alignment file was also run through FastTree to generate a phylogenetic tree (Price, Dehal et al. 2010).

Statistical analysis began with performing PCA analysis of the ambient stream conditions, courtesy of Hunter Carrick of Central Michigan University. The microbial diversity data were correlated with stream conditions that best explained variance in the physical and chemical monitoring data among the eight PA streams. Microbial diversity and stream data were correlated with canonical correspondence analysis (CCA) using the “vegan” package within the statistical program R. During CCA analysis, the number of hits each taxonomic classification had in a given sample was disregarded, and only the presence of that group in a given sample was considered. The number of hits of a given group is considered misleading due to the amplification selectivity during the PCR phase of high throughout sequencing (Shokralla, Spall et al. 2012). Therefore, the count of a given sequence after high throughput sequencing is unreliable; abundance of PAOs is confirmed by FISH.

3.7 Probe Design and Fluorescent In Situ Hybridization

Three bacterial isolate cultures were ordered from DSMZ (Brunswick, Germany): *Ignavibacterium album* (DSM no. 19864), *Thauera chlorbenzoica* (18012), and *Diaphorobacter nitroreducens* (15985). Cells arrived either as actively growing cultures or lyophilized cells were grown according to DSMZ protocols. Cultures were grown for 2-3 weeks prior to fixation. Stocks of each culture were made in 20% glycerol and stored at -80°C. *Pseudomoas aeruginosa* was regrown from a frozen stock in lysogeny broth at 37°C for 2 days prior to fixation. All four isolates were identified though Sanger sequencing of the 16S rRNA using the 27f forward primer (Genomics Core Facility, Penn State University).

Isolate cell cultures were fixed in 3% paraformaldehyde solution for 3 hours at room temperature. Following fixation, cells were rinsed through pelleting and resuspension in 1x PBS. After the second round of rinsing, pellets were resuspended in a 1:1 ethanol to 1x PBS solution and samples were stored at -20°C until analysis.

Oligonucleotide probes were designed to be approximately 20 base pairs in length, have a GC content between 40 and 60 percent, and to have no more than three G and/or C nucleotides in the last five base pairs of the probes according to probe design guidelines (Aquino de Muro 2005). Probe characteristics were checked for secondary structures using the IDT OligoAnalyzer 3.1 (IDT Technologies, Coralville, IA). Fluorescently labelled probes targeting putative PAOs, most bacteria, and helper regions were ordered from IDT Technologies (Coralville, Iowa) (Table 3-4). Putative PAO probes were labeled at the 5' end with the Cy3 fluorophore. EUB338 probes were tagged at the 3' end with the 6-FAM fluorophore. EUB338, EUB338-II and EUB 338-III were combined on a 1:1:1 ratio to create an 'EUBmix' probe. Hybridizations were performed at 46°C for 1.5 hours in buffer containing 0.9 M NaCl, 5 uM probe, and formamide at the appropriate stringency (60% for EUB probes). Subsequent hybridizations were performed starting with the highest stringency condition and proceeding in order to the lowest. Probes developed from this study were optimized through a series of FISH experiments at 5% formamide increments starting at 0% formamide. Greyscale images were taken using an AxioCam MR with at a set exposure time for each fluorochrome. Images were analyzed in ImageJ by measuring the changes in grayscale pixel intensity with increasing formamide concentration for each probe. Fluorescence intensity during image analysis was determined by measuring the grayscale pixel intensity of five randomly selected cells and subtracting from it the background intensity. All measurements were corrected for autofluorescence.

Optimized probes underwent FISH experiments with the original biofilm samples from East Hickory Creek and Cowanesque River. Biofilm samples underwent dual FISH with universal EUB bacterial probes and PAO probes. Samples were then stained with DAPI at a concentration of 1 μg/mL for one hour. Slides were rinsed with distilled water, allowed to air dry, and mounted in 2.5% DABCO. Microscopy was performed as previously described. For

each probe, fluorescent images of FITC, Cy3, DAPI-P, and DAPI-DNA were obtained for eight different frames.

Table 3-4. 16S rRNA oligonucleotide probes used in this study.

Name	**Sequence**	**Target group**	**Reference**
EUB338	GCTGCCTCCCGTAGGAGT	Most *bacteria*	Amann et al. 1990
EUB338-II	GCAGCCACCCGTAGGTGT	*Planctomycetales*	Daims et al. 1999
EUB338-III	GCTGCCACCCGTAGGTGT	*Verrucomicrobiales*	Daims et al. 1999
IGN-PAO	TTACTCGGGATATTGCTACC	*Ignavibacterium album*	This study
RHO-PAO	GACATGGGCACGTTCCGATG	*Thauera sp.*	This study
BURK-PAO	AGAGGCTTTTCGTTCCGTAC	Uncultured *Comamonadaceae*	This study
PSE-PAO	CCCTTCCTCCCAACTTAAAG	Uncultured *Pseudomonas*	This study

Chapter 4

Results

4.1 Flow Cytometry

The negative control sample set produced relative 550/25 nm emission fluorescence maximums of $10^{1.3}$ for particles with the lowest relative forward scatter and $10^{4.5}$ for particles with the highest relative forward scatter. The experimental samples showed a 10-fold fluorescence increase under the 550/25 nm emission filter after (Figure 4-1). Most of the fluorescence increase was noticed on particles with relative forward scatter values lower than 10^2; samples with increasing forward scatter had smaller increases in fluorescence following DAPI staining. The 460/25 emission filter measured little to no increase in flourescence in the experimental set relative to the negative control set, showing that any crossover between the DAPI-DNA signal and the DAPI-poly-P signal is negligible (Figure 4-1).

Samples from PA streams contained 11.0 - 38.4% putative PAOs (Figure 4-2). The number of cells sorted from a given sample ranged from 10^3 – 10^6 cells total. The sorting gate omitted >99.5% of cells in the fluorescence range of the negative control, with most cases achieving >99.9% (Figure 4-1).

Samples obtained from the USDA FD-36 watershed contained 3.34% - 6.19% PAOs for the low P loading condition and 0.39% - 3.03% PAOs for the high P loading condition (Figure 4-2). The number of putative PAOs sorted from samples ranged from 5×10^3 to 5×10^4 cells. The sorting gate omitted >99.9% of cells in the fluorescence range of the negative control (Figure 4-1).

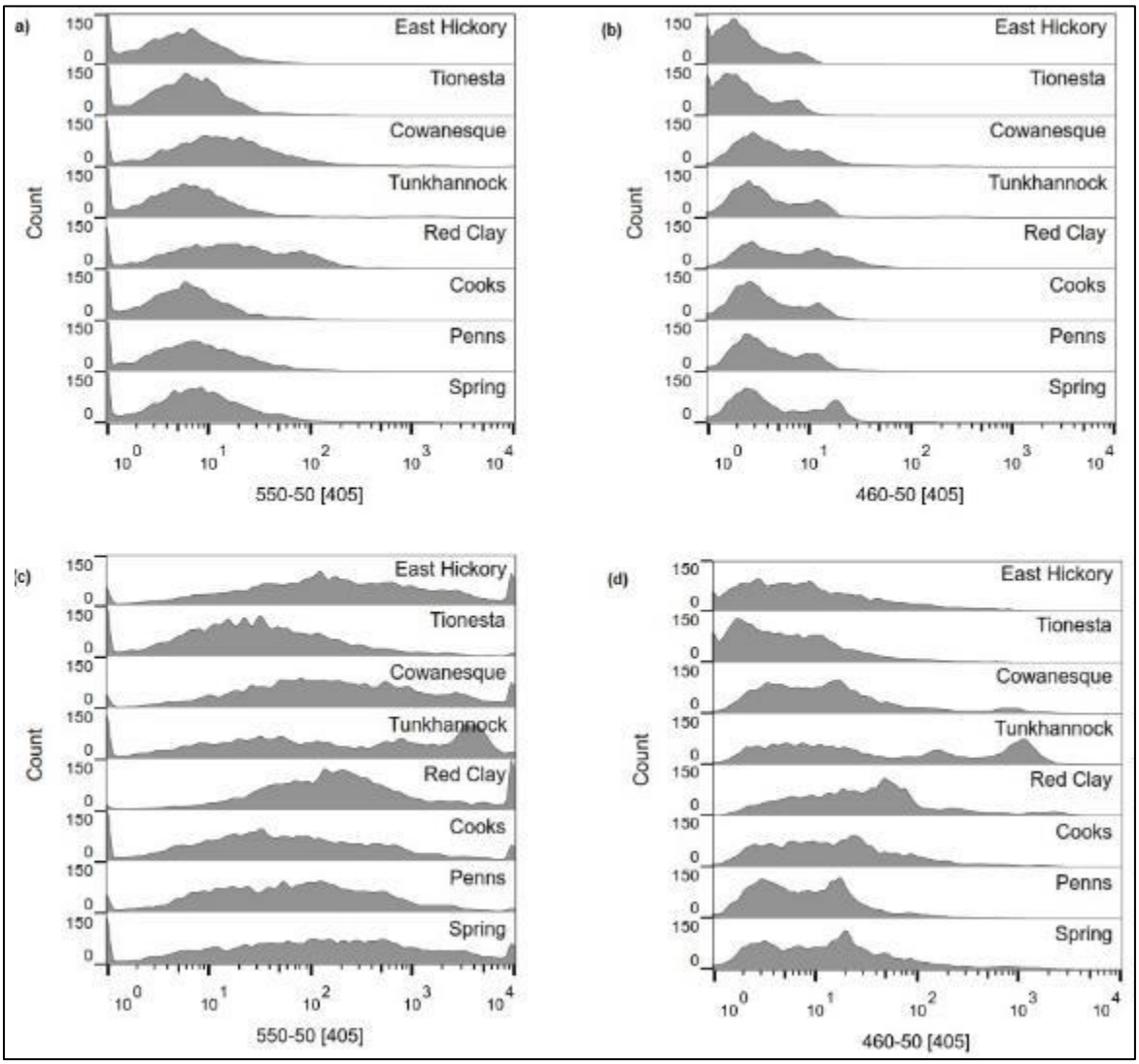

Figure 4-1. PA stream biofilm emission fluorescence measured for particles without DAPI staining in figures (a) and (b) and with DAPI staining in figures (c) and (d). Data shown is normalized to 5000 particles.

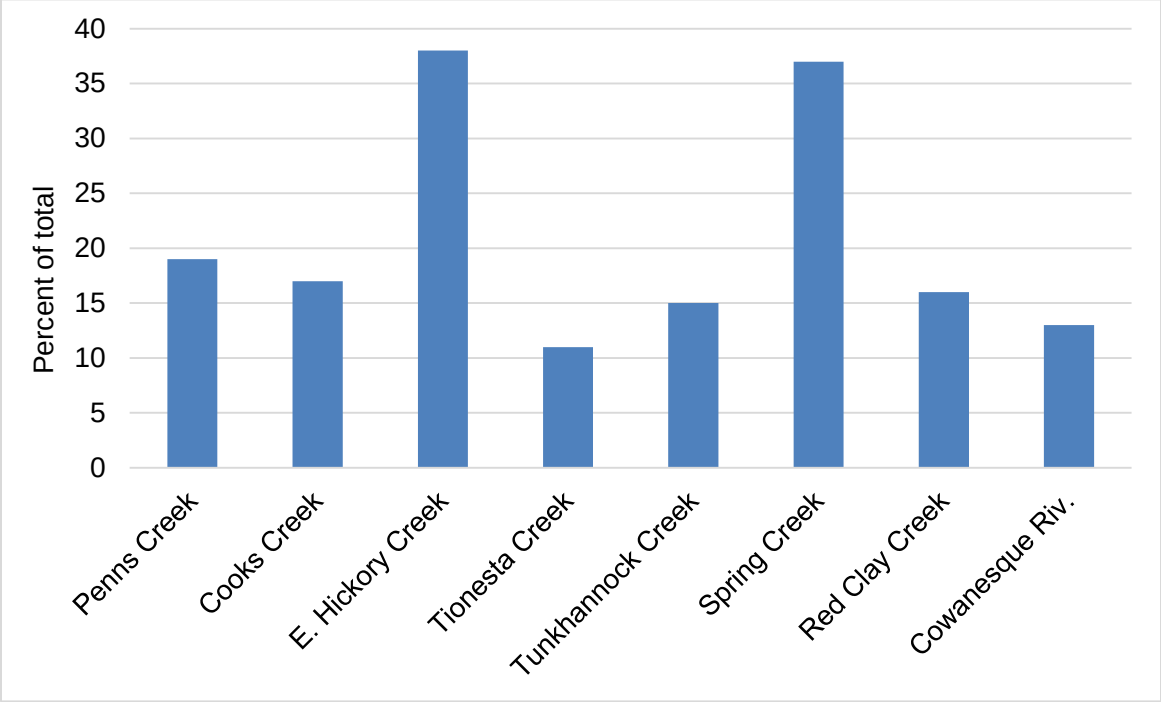

Figure 4-2. Percentage of total cells found to be putative PAOs in PA streams.

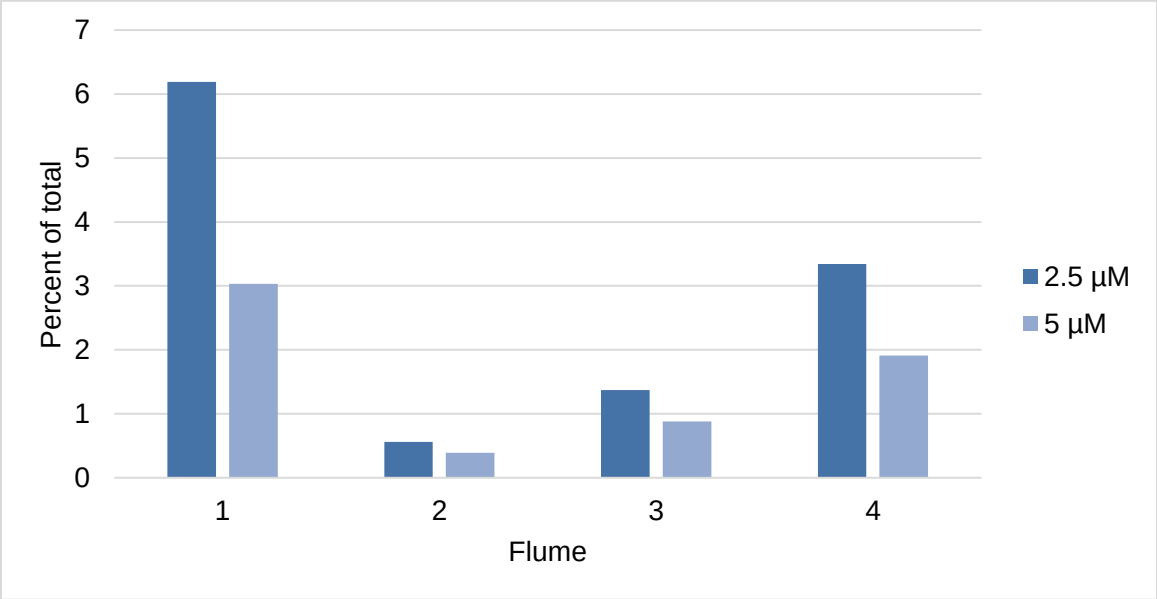

Figure 4-3. Percentage of total cells found to be putative PAOs in USDA FD-36 experimental watershed for 2.5 µM and 5 µM P loadings.

4.2 Metagenomic Sequencing

DNA extracts from the putative PAO fraction had < 0.1µg/mL DNA as measured by a Thermo Fisher Scientific NanoDrop 2000c (Appendix D). The concentrations were too low to measure on the NanoDrop 2000c. Only six of the nine samples were successfully sequenced. These samples were from the streams East Hickory Creek, Cowanesque River, Red Clay Creek, Cooks Creek, Penns Creek, and Spring Creek. These six streams still provided coverage of a wide range of geochemical conditions. In addition, these streams included 2 oligotrophic (Cooks and Penns), 1 mesotrophic (East Hickory), and 3 eutrophic states (Cowanesque, Red Clay, and Spring) using chlotophyll-a measurements as a measure of trophic state (Price and Carrick 2014).

Sequencing analysis for the PA stream biofilms identified 61 OTUs as candidate PAO groups, 2 of which had no hits in the bacterial database (Appendix E). Groups present in all six benthic biofilms are *Ignavibacterium album*, *Thauera sp.* and species related to the *Comamonadaceae* genus and the *Pseudomonas* genus (Figure 4-4). There were no hits for *Candidatus* accumulibacter.

Sequencing analysis of the USDA FD-36 biofilms identified 232 OTUs and candidate PAO groups, 1 of which had no hits in the bacterial database. Putative PAOs predominately include members of bacterial orders *Cytophagales*, *Actinomycetales*, and *Sphingomonadales* (Figure 4-5). One sample, obtained from the low P loading of flume 3, returned hits for *Candidatus* accumulibacter.

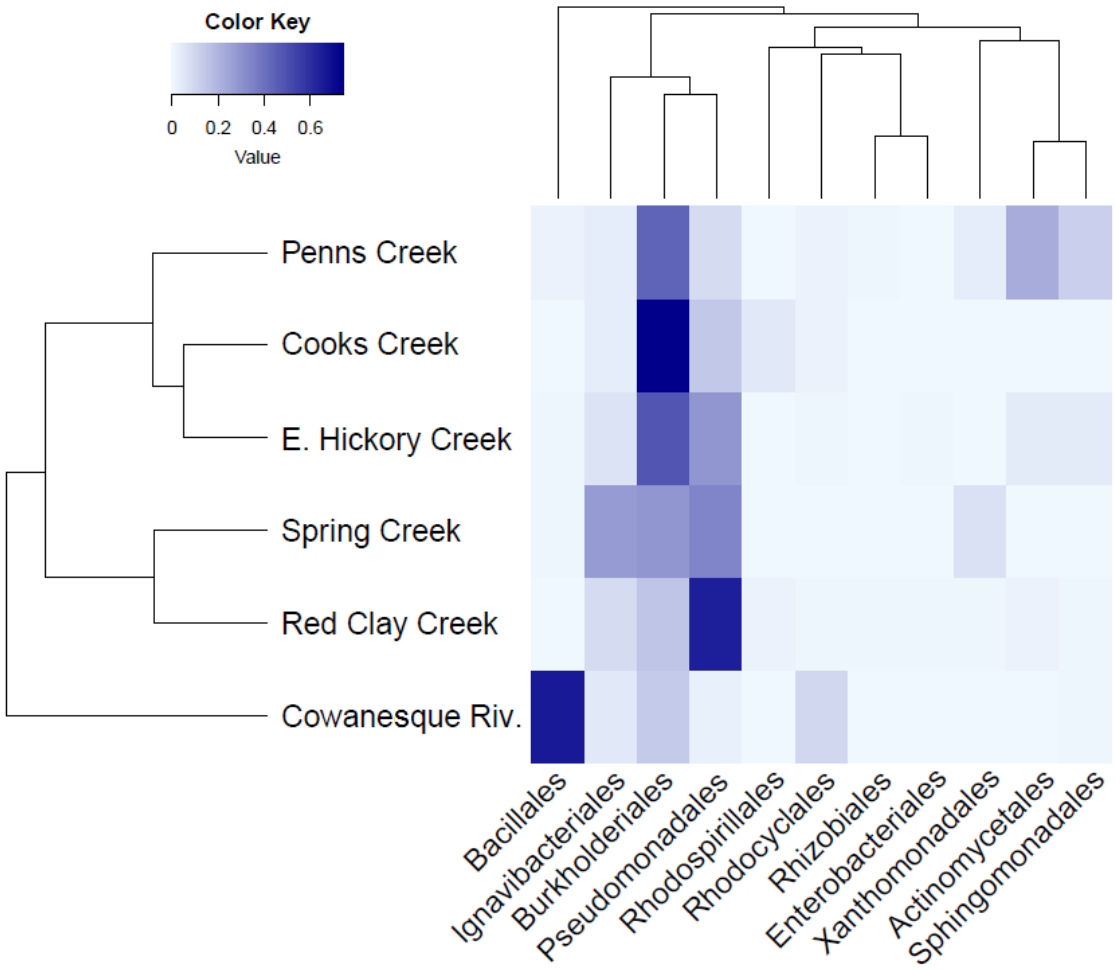

Figure 4-4. A heatmap showing the relative abundance of bacterial orders in PA streams. Dendrograms group samples based on similar diversity (y-axis) and group orders based on appearance together (x-axis).

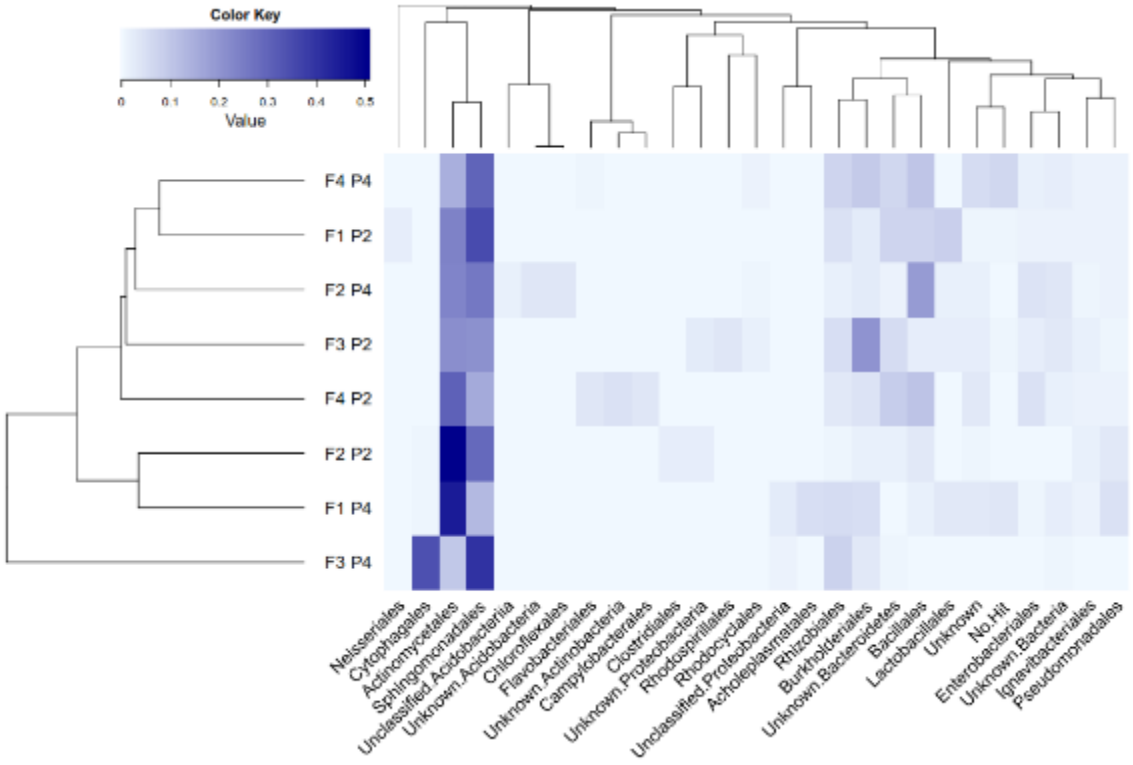

Figure 4-5. A heatmap showing the relative abundance of bacterial orders for USDA FD-36 samples. Dendrograms group samples based on similar diversity (y-axis) and group orders based on appearance together (x-axis).

4.4 Statistical Analyses

Principal components analysis (PCA) of ambient stream data into linear components of temperature/stream cover and pH/conductivity explained 87% of the variation. Temperature and stream cover (shading) were found to be inversely related (as stream cover increased, temperature decreased) as were conductivity and pH (as conductivity increased, pH decreased). These variables were used in a canonical correlation analysis (CCA) to explore the environmental conditions associated with PAO species of all six streams (Figure 4-6).

According to CCA analysis, predominant PAO groups were found to be closely related to temperature, stream cover, pH or conductivity.. PAO groups related to *Bacillales*

and *Rhodocyclales* were directly related to the stream pH, preferring a pH close to that found in Cowanesque stream (pH = 6.82 at the time of biofilm collection). PAO groups related to *Ignavibacterium* were directly related to stream temperature and PAO groups related to *Burkholderiales*, *Sphingomonadales*, and *Actinomycetales* were directly related to stream cover. Less predominant PAO groups within *Enterobacteriales* and *Xanthomonadales* were directly related to conductivity.

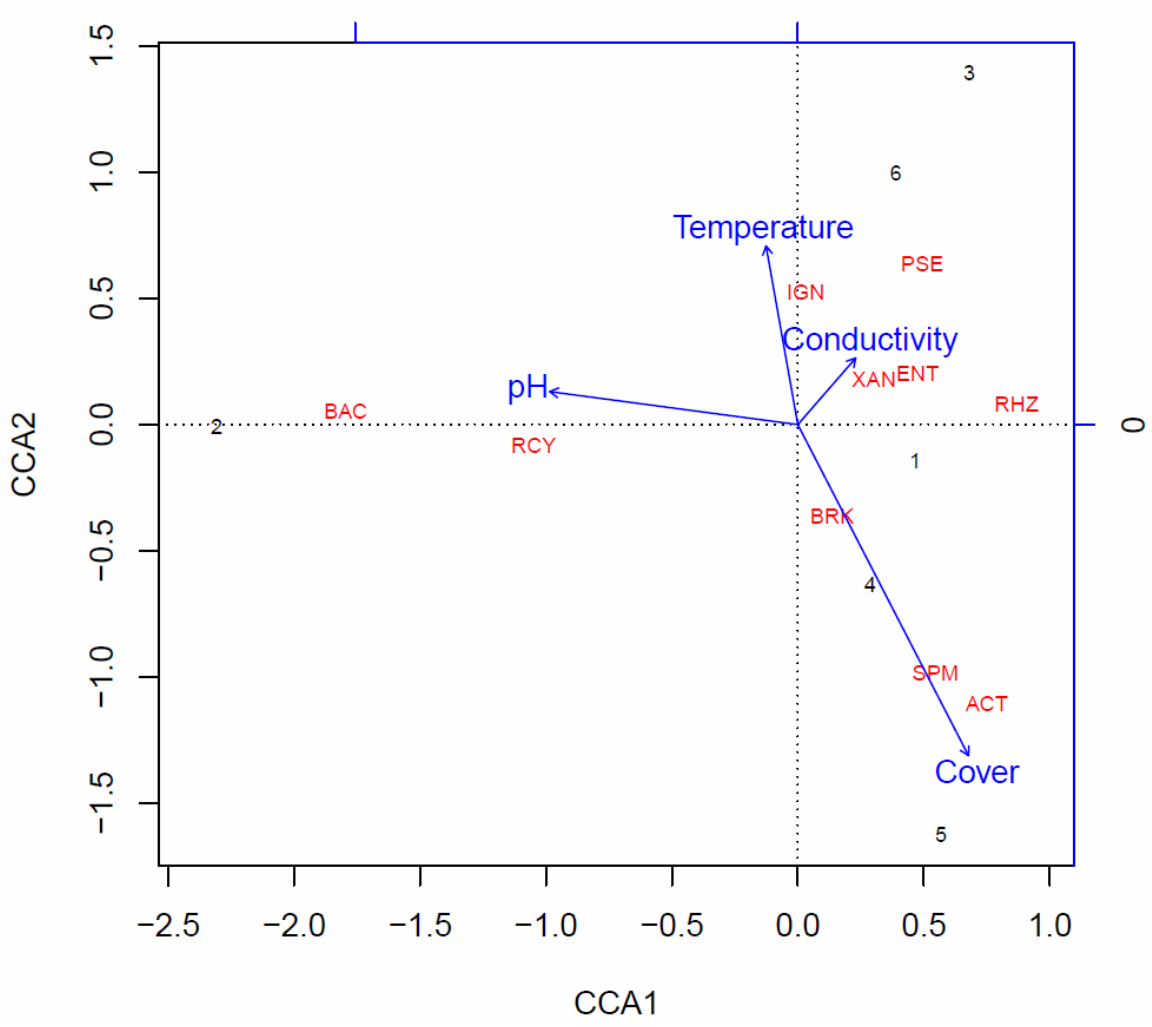

Figure 4-6. CCA analysis of environmental conditions (blue vectors) with species (red text) and streams (black numbers).

4.3 Oligonucleotide Probe Design and Optimization

Four oligonucleotide probes were developed to target predominant putative PAO groups within *Ignavibacteriales*, *Rhodocyclales*, *Burkholderiales*, and *Pseudomonadales* (Table 4-1). These groups are called “predominant” because they were found in all six streams. While some represent the highest abundance hits during high throughput sequencing, this is not necessarily why they were chosen as a focus. Isolates of *I. album*, *T. chlorobenzoica*, *D. niroreducens*, and *Pseudomonas sp.* were chosen to represent their respective PAO group and probe for FISH optimization experiments (Table 3-4). The 16S rRNA gene of these four isolates contain perfect matches to the designed probes for predominant PAO groups in the PA stream benthic biofilms. Isolates with mismatches were unavailable.

All four probes were had a strong fluorescent intensity at formamide concentrations of 0% and 5%. Probe RHO-PAO had a steady drop in fluorescence as the formamide concentration increased from 0% to 10%. At formamide concentrations of 10% and 15%, probe RHO-PAO produced a faint signal. Probe BURK-PAO had an increase in fluorescent intensity when the formamide concentration was increased to 5%. The fluorescence dropped significantly at formamide concentrations of 10 and 15%, with a faint signal at 15%.Probe PSE-PAO maintained a relatively high fluorescence for formamide concentrations between 0 and 15%.

Table 4-1. Characteristics of the designed PAO probes.

Name	Sequence	Length (bp)	GC Content (%)	Tm (°C)
IGN-PAO	TTACTCGGGATATTGCTACC	20	45	51.2
RHO-PAO	GACATGGGCACGTTCCGATG	20	60	58.8
BURK-PAO	AGAGGCTTTTCGTTCCGTAC	20	50	54.9
PSE-PAO	CCCTTCCTCCCAACTTAAAG	20	50	52.8

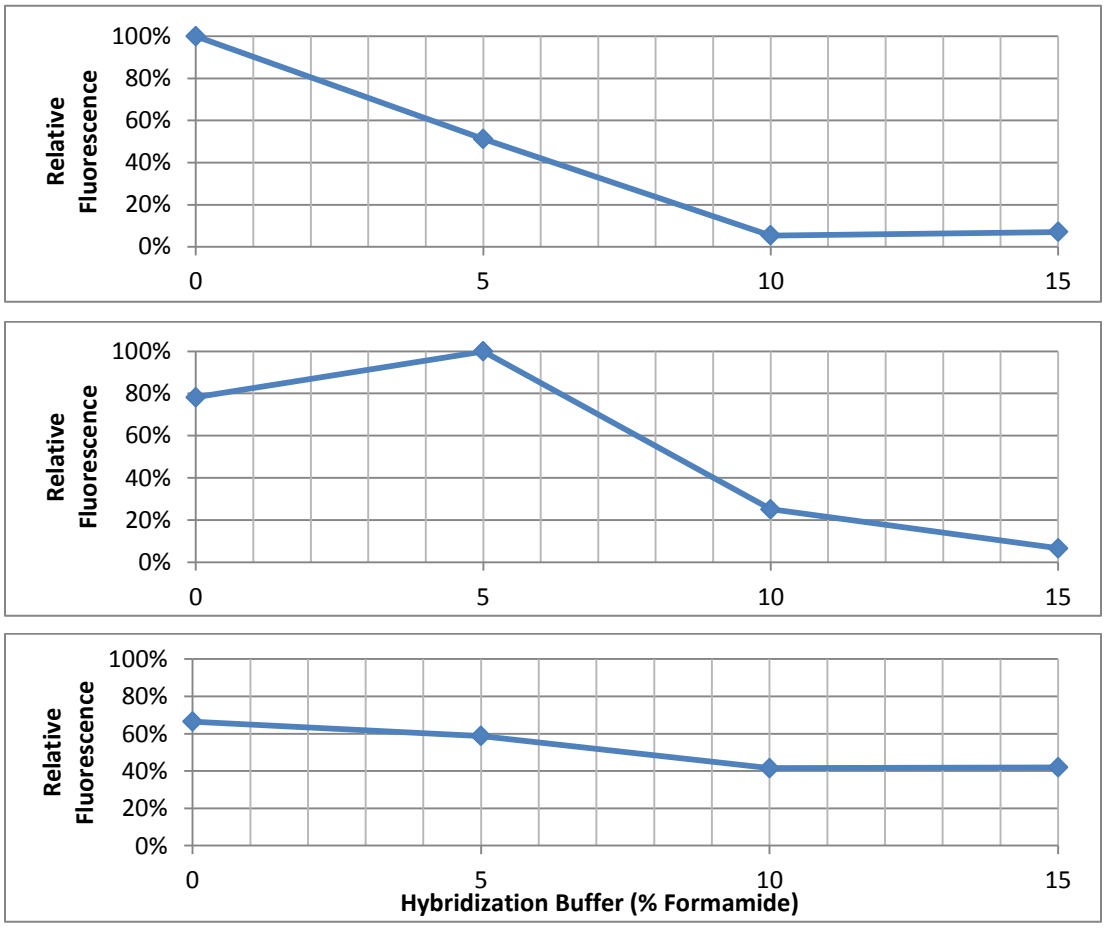

Figure 4-7.Fluorescent intensity changes at varying formamide levels for probes (a) RHO-PAO, (b) BURK-PAO and (c) PSE-PAO.

Chapter 5

Discussion

5.1 Putative PAOs of Benthic Biofilm Samples

All benthic biofilm samples contained PAOs, as indicated by the increase in yellow florescence following DAPI staining during flow cytometry. In some cases, such as the PA stream biofilm samples, PAOs can make up a significant fraction of the microbial community. In fact, the fraction of putative PAOs found in the PA stream benthic biofilms, 11% - 38%, is similar to the fraction found in EBPR (He, Gall et al. 2007). This suggests that PAOs in natural stream sediments have a significant role in P cycling. However, the USDA-FD-36 samples, which were allowed to grow for two weeks, contained 0.39% - 6.19% putative PAOs. High abundance of poly-P containing cells in benthic biofilms from stream rocks suggests that PAOs are naturally selected for over a long period of time and that their metabolic role can contribute to P cycling through water.

Startup and recovery times for an EBPR process can take months (Kortstee, Appeldoorn et al. 1994). The observation of higher abundance of putative PAOs in benthic biofilms provides evidence that like EBPR, benthic biofilms need time to acclimate. During this acclimation process, assuming the proper environmental conditions are met, PAOs slowly increase in abundance and metabolic diversity. The East Hickory Creek sample was derived from an isolated location in the Allegheny National Forest and is classified as a wilderness location. It also contains the highest percentage of PAOs among the naturally-derived samples, 38%, and is relatively diverse in comparison to the rest of the stream samples. This suggests that a wilderness environment enabled the PAO community at East Hickory Creek to mature to a level of high abundance. In contrast, biofilms from Cowanesque River contained 13% PAOs and represented primarily members of the *Bacillus* order. The

USDA FD-36 samples, which were allowed only two weeks to grow, had low percentages of the putative PAO fraction; most PAOs were from the *Actinomycetales* and *Sphingomonadales* order.

USDA FD-36 samples provide evidence that phosphorus loading plays a key role in shaping the polyphosphate accumulating community. For each flume, the sample with P loading P2 contained 1.7 times more PAOs than the respective biofilm with P loading P4. This finding is counterintuitive, as it is expected that higher P loadings would allow for more poly-P accumulation by the biomass supported by a larger PAO community.

While the level of PAO diversity varied among samples, even the least diverse samples contained enough bacterial species to suggest that a wide range of metabolic pathways for polyphosphate accumulation exist. Flow cytometry data provides supporting through plots of FSC vs. SSC measurements, a technique often used to identify cell types. For example, different types of blood cells are represented as distinct clusters on an FSC vs. SSC graph. The putative PAO fraction of the samples represents a wide range of forward scatter and side scatter measurements, with no significant clustering when FSC is plotted against SSC. Therefore, the putative PAO fraction contains a large range of cell types, indicative of representing a wide range of species.

High-throughput sequencing of the V1-V3 region of the 16S rRNA reveals that PAOs in the PA stream benthic biofilms are related to a broad range of bacterial groups. Surprisingly, no samples contained *Candidatus* Accumulibacter, indicative of additional bacterial groups performing the PAO metabolism and different pathways of biological phosphate accumulation in the environment. It remains unknown if the typical PAO metabolism modeled from EBPR is the dominant P accumulation pathway in the environment. As a result, new oligonucleotide probes were developed for naturally occurring PAOs.

Candidate PAO groups among the six PA stream samples include *Ignavibacterium album*, *Thauera sp.*, uncultured *Comamonadaceae* and uncultured *Pseudomonas* (Figure 4-4). These bacterial orders were chosen as a focal point due to their appearance in all samples and potential of containing groups with the PAO metabolism. All four groups have been previously reported to contain facultative species, in support of the currently accepted model for the polyphosphate accumulating metabolism. There are PAOs that have a denitrifying metabolism to accumulate phosphate instead of the typical aerobic metabolism; groups containing strict anaerobes can also contain PAOs. The four common groups also contain species that have been found in aquatic environments. In fact, members of *Comamonadaceae* are typically found in soils and water and contain known phytopathogens.

Ignavibacterium album was recently isolated from microbial mats living in Japanese hot water springs and is a member of the *Chlorobi* family (Iino, Mori et al. 2010). While other members of *Chlorobi* are known as green sulfur bacteria, *I. album* is non-phototrophic, capable of metabolism under both oxic and anoxic conditions, cannot fix sulfur, and cannot fix nitrogen (Liu, Frigaard et al. 2012). Its facultative metabolism and appearance in the natural environment make it a likely candidate PAO in PA streams. There is no research identifying *I. album* in EBPR activated sludge, most likely a result of limited EBPR activated sludge characterization since it was isolated. This study is the first to identify *I. album* as a putative PAO.

The *Thauera* genre contains bacteria that are gram-negative, flagellated, and rod-shaped. They typically occur in wet soils and polluted freshwater (Garrity, Brenner et al. 2005). *Thauera* sp. is the closest naturally occurring relative to *Candidatus* Accumulibacter in the PA stream benthic biofilm samples, but has not been identified as a PAO in EBPR activated sludge. In fact, members of the *Thauera* genre are typically known for their ability to break down aromatic hydrocarbons. They have not been observed to accumulate phosphorus despite their critical functional role in wastewater treatment processes where the

main pollutants are aromatic (Liu, Zhang et al. 2005). Since inorganic poly-P confers metabolic resistance under unfavorable conditions, it has been observed that bacteria accumulate poly-P when aromatics are the dominant carbon source (Chavez, Lunsdorf et al. 2004). Results suggest that *Thauera sp.* may accumulate poly-P with evidence suggesting this poly-P accumulation is connected to aromatic degradation. Their close evolutionary relationship with *Candidatus* Accumulibacter is strong evidence that they have a functional role in phosphate uptake in addition to aromatic compound degradation.

Members of the *Comamonadaceae* family are typically rod and helical shaped and include common soil and water bacteria (Willems, Deley et al. 1991). Their metabolism is strictly respiratory, being chemoorganotrophic or chemolithotrophic (Garrity, Brenner et al. 2005). While not typically found in EBPR activated sludge, members of the *Comamonadacae* group can be facultative and are often found in the environment, suggesting that they are putative PAOs. This group may have a unique polyphosphate accumulating metabolism based on their evolutionary distance from *Candidatus accumulibacter*.

Members of *Pseudomonas* are commonly found in environmental samples and cover a wide range of ecological niches (Madigan 2010). Therefore, while *Pseudomonas* is not typically considered a PAO in EBPR, there have been reports suggesting that they are PAOs in other environments. For example, Pseudomonas strain B4 has been found to accumulate poly-P during its logarithmic growth phase when PCB is the sole carbon source (Chavez, Lunsdorf et al. 2004). In addition, *Pseudomonas putida* strain CA-3 has been observed to accumulate poly-P under nitrogen limitation (Tobin, McGrath et al. 2007). The variety of carbon sources available on streambeds suggests that there are other environmental stressors triggering poly-P accumulation in *Pseudomonsas spp*.

USDA FD-36 samples were found to contain a high presence of bacteria from the *Propionibacterium* genre and *Sphingomonadales* order (Figure 4-5). It is important to note that sequencing constraints caused the USDA FD-36 sequencing to be based on the V1-V2

region of the 16S rRNA assay instead V1-V3 region used for sequencing the PA stream samples. As a result, the resolution of the USDA samples was lower and the taxonomic data is difficult to compare to that of the PA stream biofilms. However, the high quality DNA of the USDA FD-36 samples indicates that the sequencing data represents the putative PAO community well (Appendix D).

Members of *Propionibacterium* are primarily facultative and are known for their ability to metabolize propionic acid. Both characteristics have been observed in PAOs from EBPR activated sludge. Adding propionic acid as a carbon substrate to the anaerobic phase of EBPR has been found to select for PAOs (Oehmen, Saunders et al. 2006). The sequencing results in conjunction with EBPR findings suggest that there are unclassified members of *Propionibacterium* that accumulate P in excess.

5.2 Stream Conditions Associated with PAO Community Structure

It was observed that the PAO community structure was correlated with the trophic state of the stream at the time samples were collected. The PA streams represent a range of trophic states, using the chlorophyll-a concentrations of the biofilms as a measure of trophic state (36.4 mg chl-a/m^2 to 672 mg chl-a/m^2) (Figure 5-1). Samples with similar trophic state had similar taxonomic composition. For example, East Hickory Creek and Cooks Creek are both mesotrophic and contain a similar microbial community structure for putative PAOs. Cowanesque River, by far the most eutrophic stream, had a unique microbial community structure relying primarily on members of *Bacillus*. The shift in the microbial community structure suggests that P may be an indicator of stream trophic state and the PAO community can be used as additional evidence of eutrophication. As streams become more eutrophic, the PAO community of the benthic biofilms shifts from members of *Burkholderia* and *Pseudomonas* to members of only *Bacillus*. This community shift is supported by a drop

in PAO species richness according to the Chao1 and Simpson diversity measurements (Figure 5-2).

The PA stream benthic biofilms samples were analyzed for changes associated with stream ambient data (temperature, pH, conductivity, and cover). PCA analysis of the ambient stream data resulted in trends that were expected in the eight Pennsylvania streams. It was expected that temperature increases were associated with decreasing stream cover because greater exposure to sunlight generally increases stream temperature. Analyzing the PAO community shift along the principal components, specific PAO groups were found to be directly related to specific stream conditions.

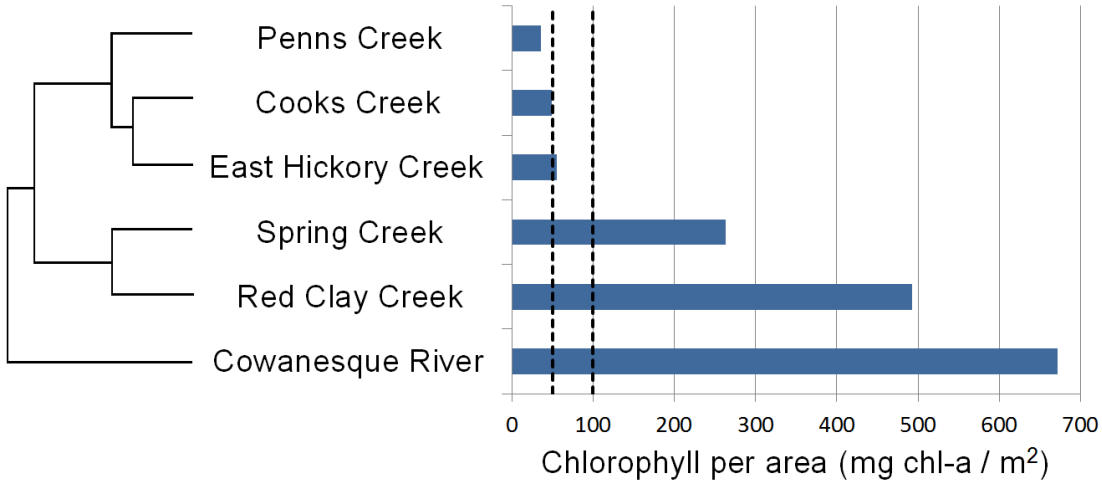

Figure 5-1. The dendrogram grouping streams based on PAO group composition superimposed on a bar chart showing stream productivity. Chlorophyll concentrations lower than 50 mg/m^2 are oligotrophic, between 50 and 100 are mesotrophic, and above 100 are eutrophic.

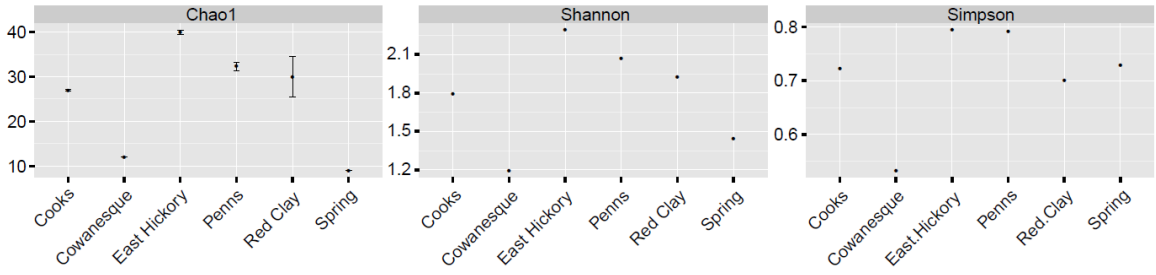

Figure 5-2. Alpha diversity measures showing the species richness of PA stream benthic biofilms.

With *Rhodocyclales* members being the closest relatives to *Candidatus accumulibacter*, it is not surprising to find *Rhodocyclales* members to be directly related to the stream pH. EBPR is sensitive to changes in pH, and shifts in pH can cause GAOs to be the dominant community in EBPR systems (Serafim, Lemos et al. 2002). The pH of EBPR is typically 7.5 to 8.0, higher than the pH of any of the streams. This may explain why members of *Rhodocyclus* were in low abundance in comparison to other predominant PAO groups. The direct relationship between *Rhodocyclales* PAOs and stream pH suggests that some naturally occurring PAO groups are sensitive to changes in pH.

Little is known about *I. album*, but its preference for warm environments may explain its relationship to temperature in the benthic biofilms of PA streams. Since *I. album* is typically grown at 45°C and was isolated from Japanese hot springs, it would be expected that the abundance of *I. album* increases as temperature increases and the results support these expectations.

5.3 Additional PAO-targeted Oligonucleotide Probes

Probes RHO-PAO, BURK-PAO, and PSE-PAO were designed to target the *Thauera* genre, the *Comamonadaceae* family, and the *Pseudomonas* genre, respectively (Figure 5-3). The optimal probe hybridization condition was approximated using isolates containing perfect matches to the probe because additional isolates containing one mismatch to the probes of interest were unavailable. The optimal hybridization condition was approximated as the highest formamide concentration that was observed to hybridize the probe to cultures containing a perfect sequence match. It was expected that the 0% formamide hybridization condition would set the maximum fluorescence intensity for each probe and that the fluorescent intensity would remain strong until the formamide concentration was increased

beyond the optimal hybridization condition, at which point the fluorescent intensity would quickly decrease and remain low.

Isolates of *I. album* produced poor DNA quality following extraction and disintegrated during the fixation process despite successful growth in published growing medium (Iino et al. 2010). This was unexpected as the genome of I. album was previously published (Liu, Frigaard et al. 2012). Failure of DNA extraction may be due to using only 0.25 g of cells where previously 7 g of cells was harvested for sequencing (Liu, Frigaard et al. 2012). The reason for *I. album* cells failing during fixation and in FISH experiments remains unknown. No optimization of the IGN-PAO probe was achieved.

The RHO-PAO probe followed trends expected during hybridization optimization experiments. Maximum fluorescence was observed when the hybridization buffer contained 0% formamide. The fluorescent intensity dropped from 51% to 5% of the maximum under hybridization conditions of 5% and 10% formamide, respectively. This indicates that the approximate optimal hybridization condition for FISH is 5%.

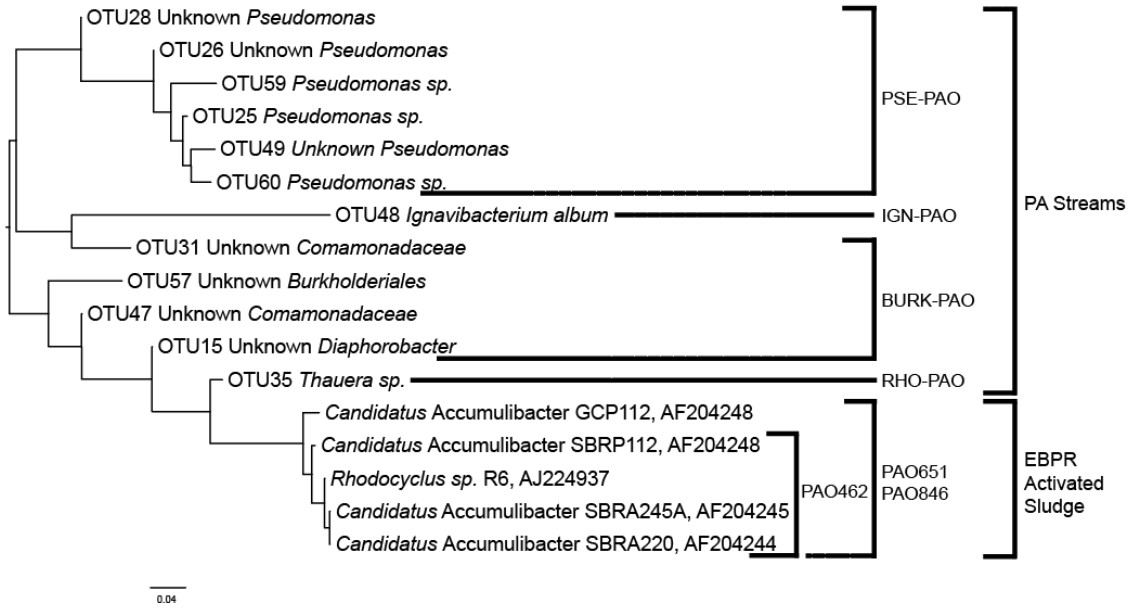

Figure 5-3. A phylogenetic tree showing the evolutionary relationships among the PAO oligonucleotide probes developed in this study and *Candidatus* Accumulibacter.

The BURK-PAO probe maximum fluorescence was observed at 5% formamide. While this result was unexpected, the fluorescent intensity at 0% formamide was 78% of the maximum, which still produced strong cell fluorescence. As expected, the fluorescent intensity dropped off at higher formamide concentrations. The fluorescence was 25% and 6% of the maximum at 5% and 10% formamide, respectively. Fluorescent intensity measurements from ImageJ are supported by visual observations (Figure 5-4). Therefore, the optimal hybridization condition was approximated as 5% formamide.

The PSE-PAO probe did not follow expectations during optimization experiments. The maximum fluorescence was observed at 20% formamide hybridization. The fluorescence at 0% formamide was only 66% of the maximum. It is unknown why the PSE-PAO probe produced erroneous results and it was difficult to determine the approximate optimal hybridization condition. It is likely that further increases to the formamide concentrations in the hybridization buffer would result in a sharp decrease in the fluorescent intensity at some point, but the inconsistencies of the fluorescent intensities at lower formamide concentrations make the data unreliable.

Successful hybridization of the BURK-PAO and RHO-PAO probes to the original biofilms samples confirms that the identified groups act as polyphosphate accumulators in the benthic biofilms of PA streams. Cells were observed to contain the key signals of nucleic acids, poly-P, the PAO probe and the bacterial probes (Figures 5-7, 5-8). Therefore, it is confirmed that members *Burkholderiales* and *Rhodocyclales* in the PA stream biofilms species contain polyphosphate.

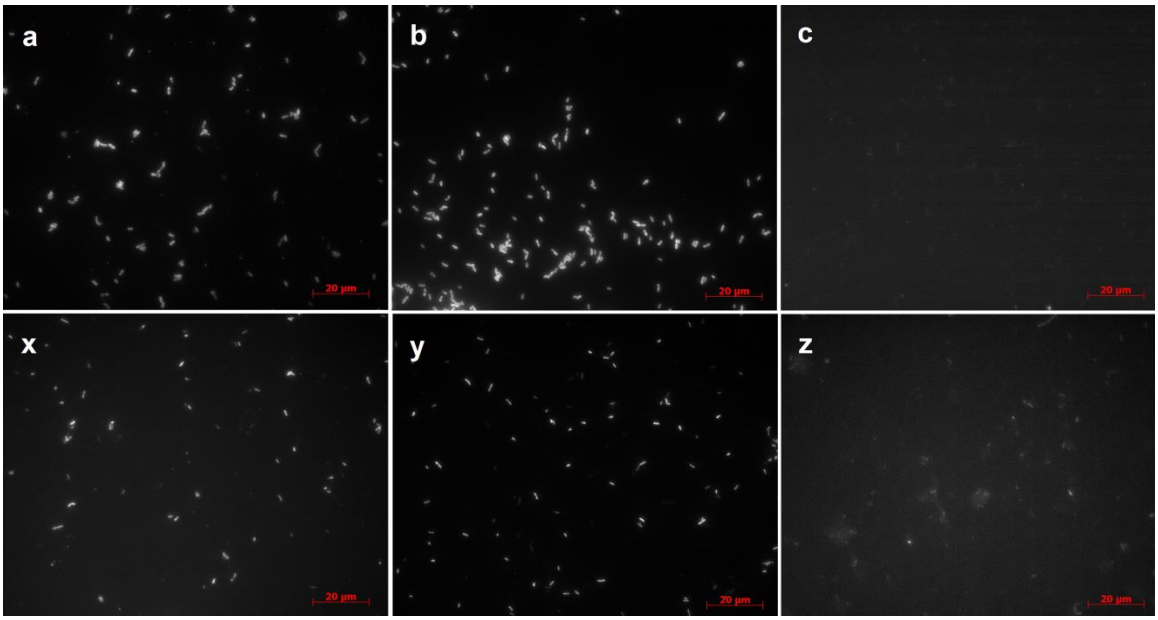

Figure 5-4. Images of probe optimization experiments. a, b, and c show RHO-PAO hybridized to *T. chlorobenzoica* at 0%, 5%, and 10% formamamide, respectively. x, y, and z show BURK-PAO hybridized to *D. nitroreducens* at 0%, 5%, and 10% formamide, respectively.

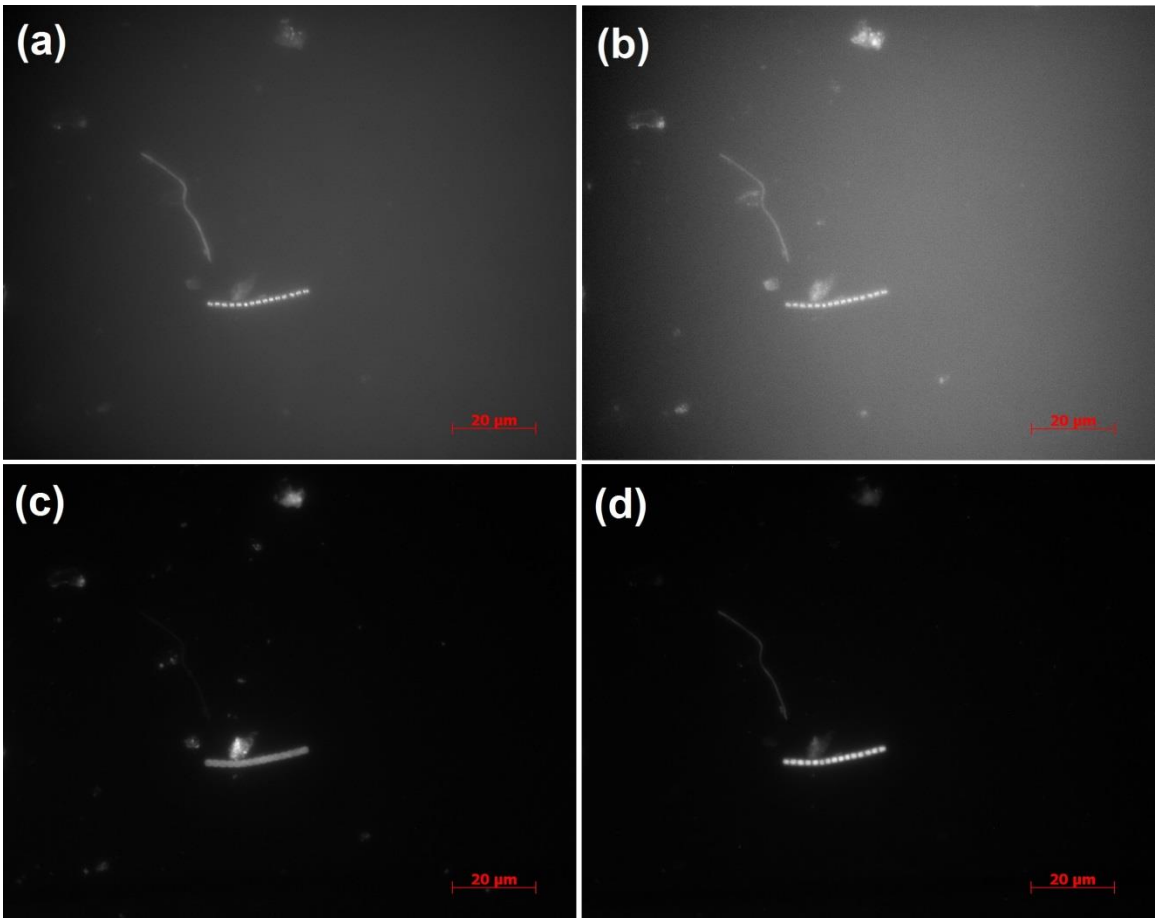

Figure 5-5. PAOs in East Hickory stream identified by FISH with probe Burk-PAO. Images show (a) nucleic acids, (b) poly-P granules, (c) Burk-PAO probe, and (d) EUB-338 mix probe.

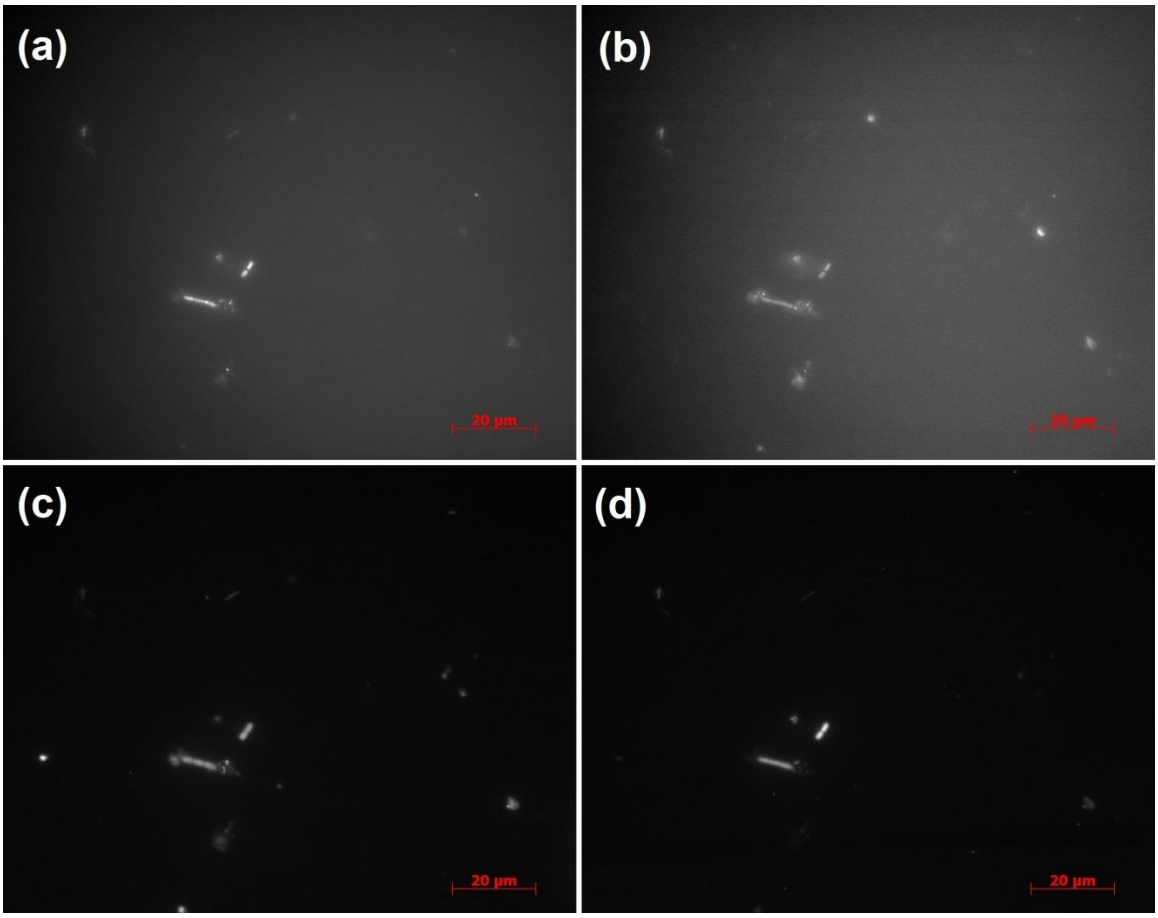

Figure 5-6. PAOs in East Hickory stream identified by FISH with probe RHO-PAO. Images show (a) nucleic acids, (b) poly-P granules, (c) Burk-PAO probe, and (d) EUB-338 mix probe.

Chapter 6

Impacts and Future Work

This research represents an initial step toward connecting biological metabolisms with P mobility and cycling in the environment. Taxonomic classification of PAOs in benthic stream biofilms was not previously accomplished until now. In addition, no previous research has heretofore attempted to compare putative PAO species in the natural environment with those found in EBPR activated sludge. Now that we know PAOs encompass a wide range of species not typically found in EBPR, research can begin to focus on the different metabolic pathways of polyphosphate accumulation and the environmental conditions that select for them.

The next step is to explore environmental conditions that generate alternating aerobic/anaerobic microenvironments and see if the putative PAO taxonomy moves towards species typically found in EBPR. If so, it is possible that these areas are biologically immobilizing a significant portion of the P pool under alternating wet and dry conditions in agricultural landscapes and over the course of diel light cycles in stream environments. Research in this area is already underway at Cornell University and the research presented herein has supported the hypotheses being tested there. This form of P immobilization can be incorporated into P models for an improved understanding of P cycling in stream systems and a better representation of these processes in in-stream P simulation models.

The connection of biological P metabolism to other forms of P immobilization in the environment can be explored as a result of this research. Ferric iron complexation with phosphate ions as a result of oxidation/reduction reactions is another important P immobilization pathway that can have both chemical and biological connections with PAOs. Saturated soil conditions shift the master variable for chemical reactions from pH to redox potential, resulting in increased rates of metal oxide dissolution (McBride 1994). Redox

potential of a soil is inversely related to P release and dissolution of P in reducing environments is related to iron oxide reduction (Patrick and Khalid 1974, Baldwin, Mitchell et al. 2000). As microbial populations, such as *Geobacter*, reduce iron to its ferrous oxidation state, phosphate ions become available for uptake by plants or sorption to sediment. Increased rates of metal oxide reduction under saturated conditions are typically not observed until oxygen is depleted. Anaerobic conditions cause a shift in microbial metabolisms to reduce ferric iron, and release phosphate ions into the environment. It is likely PAOs compete with iron reducers under aerobic conditions, where available phosphate can complex with ferric iron or be taken up by PAOs as polyphosphate (poly-P).

In addition to competition with iron reducers, there is evidence that PAOs compete with plants in the rhizosphere for phosphate availability. Complexation of phosphate with iron occurs under aerobic conditions. Plants release carboxylic acids to drop the pH, promoting the detachment and dissolution of metal oxides and increasing phosphorus availability (Bolan, Naidu et al. 1994). It has not yet been acknowledged if carboxylic acids can release phosphate stored by PAOs in aerobic environments. It is important to understand the fate of phosphate during aerobic/anaerobic cycles as microbial community structure can have significant impact on phosphate availability for plants.

While this project has focused on the connection of PAOs to agricultural P runoff, the data presented can also be used to understand PAOs in locations beyond the those that were considered in this research. As mentioned in the discussion section, there are already observations that polyphosphate accumulation is associated with the degradation of aromatic compounds. It is certainly of interest to explore how alternative pathways of biological P accumulation remove aromatic compounds in the context of water remediation. Members of *Thauera* and *Pseudomonas* have now been target as groups that can perform both aromatic degradation and polyphosphate accumulation. For *Pseudomonas*, the poly-P accumulation is associated with the breakdown of aromatics. This concept can also be taken out of the

restrictive wastewater treatment perspective and applied to environments with high concentrations of aromatics as a result of pollution to improve remediation practices.

REFERENCES

Acevedo, B., A. Oehmen, G. Carvalho, A. Seco, L. Borras and R. Barat (2012). "Metabolic shift of polyphosphate-accumulating organisms with different levels of polyphosphate storage." Water Res **46**(6): 1889-1900.

Ahn, J., S. Schroeder, M. Beer, S. McIlroy, R. C. Bayly, J. W. May, G. Vasiliadis and R. J. Seviour (2007). "Ecology of the microbial community removing phosphate from wastewater under continuously aerobic conditions in a sequencing batch reactor." Applied and Environmental Microbiology **73**(7): 2257-2270.

Allan, R. A. and J. J. Miller (1980). "Influence of S-Adenosylmethionine on Dapi-Induced Fluorescence of Polyphosphate in the Yeast Vacuole." Canadian Journal of Microbiology **26**(8): 912-920.

Amann, R. I., B. J. Binder, R. J. Olson, S. W. Chisholm, R. Devereux and D. A. Stahl (1990). "Combination of 16s Ribosomal-Rna-Targeted Oligonucleotide Probes with Flow-Cytometry for Analyzing Mixed Microbial-Populations." Applied and Environmental Microbiology **56**(6): 1919-1925.

Amann, R. I., W. Ludwig and K. H. Schleifer (1995). "Phylogenetic Identification and in-Situ Detection of Individual Microbial-Cells without Cultivation." Microbiological Reviews **59**(1): 143-169.

Aquino de Muro, M. (2005). Probe Design, Production, and Applications. Totowa, NJ, Humana Press.

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." J Fluoresc **18**(5): 859-866.

Baldwin, D. S., A. M. Mitchell and G. N. Rees (2000). "The effects of in situ drying on sediment-phosphate interactions in sediments from an old wetland." Hydrobiologia **431**(1): 3-12.

Barnard, J. L. (1975). "Biological nutrient removal without the addition of chemicals." Water Research **9**: 485-490.

Bolan, N. S., R. Naidu, S. Mahimairaja and S. Baskaran (1994). "Influence of Low-Molecular-Weight Organic-Acids on the Solubilization of Phosphates." Biology and Fertility of Soils **18**(4): 311-319.

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., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley and V. H. Smith (1998). "Nonpoint pollution of surface waters with phosphorus and nitrogen." Ecological Applications **8**(3): 559-568.

Carrick, H., A. Scanlan and R. Wagner (2009). Use of perphyton to estimate TMDL end-points. P. S. I. o. E. a. t. Environment.

Chavez, F. P., H. Lunsdorf and C. A. Jerez (2004). "Growth of polychlorinated-biphenyl-degirading bacteria in the presence of biphenyl and chlorobiphenyls generates oxidative stress and massive accumulation of inorganic polyphosphate." Applied and Environmental Microbiology **70**(5): 3064-3072.

Cole, J. R., Q. Wang, J. A. Fish, B. L. Chai, D. M. McGarrell, Y. N. Sun, C. T. Brown, A. Porras-Alfaro, C. R. Kuske and J. M. Tiedje (2014). "Ribosomal Database Project: data and tools for high throughput rRNA analysis." Nucleic Acids Research **42**(D1): D633-D642.

Comeau, Y., K. J. Hall, R. E. W. Hancock and W. K. Oldham (1986). "Biochemical-Model for Enhanced Biological Phosphorus Removal." Water Research **20**(12): 1511-1521.

Crocetti, G. R., J. F. Banfield, J. Keller, P. L. Bond and L. L. Blackall (2002). "Glycogen-accumulating organisms in laboratory-scale and full-scale wastewater treatment processes." Microbiology-Sgm **148**: 3353-3364.

Crocetti, G. R., P. Hugenholtz, P. L. Bond, A. Schuler, J. Keller, D. Jenkins 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.

Deinema, M. H., M. van Loosdrecht and A. Scholten (1985). "Some physiological characteristics of acinetobacter spp. accumulating large amounts of phosphate." Water Science and Technology **17**: 119-125.

Diaz, J. M. and E. D. Ingall (2010). "Fluorometric Quantification of Natural Inorganic Polyphosphate." Environmental Science & Technology **44**(12): 4665-4671.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Research **32**(5): 1792-1797.

Edgar, R. C. (2010). "Search and clustering orders of magnitude faster than BLAST." Bioinformatics **26**(19): 2460-2461.

Edgar, R. C. (2013). "UPARSE: highly accurate OTU sequences from microbial amplicon reads." Nature Methods **10**(10): 996-+.

Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince and R. Knight (2011). "UCHIME improves sensitivity and speed of chimera detection." Bioinformatics **27**(16): 2194-2200.

Fitter, A. H. (2005). "Darkness visible: reflections on underground ecology." Journal of Ecology(93): 231-243.

Fuchs, B. M., F. O. Glockner, J. Wulf and R. Amann (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.

Fuchs, B. M., G. Wallner, W. Beisker, I. Schwippl, W. Ludwig and R. Amann (1998). "Flow cytometric analysis of the in situ accessibility of Escherichia coli 16S rRNA for fluorescently labeled oligonucleotide probes." Applied and Environmental Microbiology **64**(12): 4973-4982.

Fuhs, G. W. and M. Chen (1975). "Microbiological Basis of Phosphate Removal in the Activated Sludge Process for the Treatment of Wastewater." Microbial Ecology **2**(2): 119-138.

Garrity, G. M., D. J. Brenner, N. R. Krieg and J. T. Staley (2005). Bergey's Manual of Systematic Bacteriology. New York, Springer.

Grady, C. P., G. T. Daigger, N. G. Love and C. D. Filipe (2011). Biological Wastewater Treatment. Boca Raton, FL, CRC Press.

Hatt, B. E., T. D. Fletcher, C. J. Walsh and S. L. Taylor (2004). "The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams." Environ Manage **34**(1): 112-124.

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, Z. L., M. K. Zhang, P. J. Stoffella, X. E. Yang and D. J. Banks (2006). "Phosphorus Concentrations and Loads in Runoff Water under Crop Production." Soil Science Society of America Journal **70**(5): 1807.

Hung, C. H., J. Peccia, J. L. Zilles and D. R. Noguera (2002). "Physical enrichment of polyphosphate-accumulating organisms in activated sludge." Water Environment Research **74**(4): 354-361.

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

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.

Iino, T., K. Mori, Y. Uchino, T. Nakagawa, S. Harayama and K. Suzuki (2010). "Ignavibacterium album gen. nov., sp nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of Ignavibacteria classis nov., for a novel lineage at the periphery of green sulfur bacteria." International Journal of Systematic and Evolutionary Microbiology **60**: 1376-1382.

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-diamidino-2-phenylindol (DAPI) at a polyphosphate-probing concentration." Water Research **33**(1): 257-265.

Khan, M. S., A. Zaidi and P. A. Wani (2007). "Role of phosphate-solubilizing microorganisms in sustainable agriculture - A review." Agronomy for Sustainable Development **27**(1): 29-43.

Khoshmanesh, A., R. Sharma and R. Beckett (2001). "Biomass of sediment bacteria by sedimentation field-flow fractionation." Journal of Environmental Engineering-Asce **127**(1): 19-25.

Klindworth, A., E. Pruesse, T. Schweer, J. Peplies, C. Quast, M. Horn and F. O. Glockner (2013). "Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies." Nucleic Acids Res **41**(1): e1.

Kortstee, G. J. J., K. J. Appeldoorn, C. F. C. Bonting, E. W. J. Vanniel and H. W. Vanveen (1994). "Biology of Polyphosphate-Accumulating Bacteria Involved in Enhanced Biological Phosphorus Removal." Fems Microbiology Reviews **15**(2-3): 137-153.

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**(3): 543-554.

Liu, B., F. Zhang, X. Feng, Y. Liu, X. Yan, X. Zhang, L. Wang and L. Zhao (2005). "Thauera and Azocarus as funtionally important genera in a denitrifying quinoline-removal bioreactor as revealed by microbial community structure comparison." FEMS Microbiol Ecol **55**: 274-286.

Liu, Z. F., N. U. Frigaard, K. Vogl, T. Iino, M. Ohkuma, J. Overmann and D. A. Bryant (2012). "Complete genome of Ignavibacterium album, a metabolically versatile, flagellated, facultative anaerobe from the phylum Chlorobi." Frontiers in Microbiology **3**.

Madigan, M. T. M., J. M.; Stahl, D.; Clark, D. P.; (2010). Brock Bioogy of Microorganisms.

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.

McBride, M. B. (1994). Environmental Chemistry of Soils. New York, NY, Oxford University Press.

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.

Nguyen, H. 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." Environ Microbiol **14**(10): 2826-2837.

Oehmen, A., P. C. Lemos, G. Carvalho, Z. G. 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.

Oehmen, A., A. M. Saunders, M. T. Vives, Z. G. Yuan and H. Keller (2006). "Competition between polyphosphate and glycogen accumulating organisms in enhanced biological phosphorus removal systems with acetate and propionate as carbon sources." Journal of Biotechnology **123**(1): 22-32.

Oehmen, A., Z. Yuan, L. L. Blackall and J. Keller (2004). "Short-term effects of carbon source on the competition of polyphosphate accumulating organisms and glycogen accumulating organisms." Water Science and Technology **50**(10): 139-144.

Patrick, W. H. and R. A. Khalid (1974). "Phosphate Release and Sorption by Soils and Sediments - Effect of Aerobic and Anaerobic Conditions." Science **186**(4158): 53-55.

Ploner, A. (2014). "Heatplus: Heatmaps with rows and/or column covariates and colored clusters." R package version 2.14.0.

Price, K. J. and H. J. Carrick (2014). "Quantitative evaluation of spatiotemporal phosphorus fluxes in stream biofilms." Freshwater Science **33**(1): 99-111.

Price, M. N., P. S. Dehal and A. P. Arkin (2010). "FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments." Plos One **5**(3).

Schindler, D. W. (1977). "Evolution of Phosphorus Limitation in Lakes." Science **195**(4275): 260-262.

Serafim, L. S., P. C. Lemos and M. A. M. Reis (2002). "Effect of pH control on EBPR stability and efficiency." Water Science and Technology **46**(4-5): 179-184.

Seviour, B. and P. Nielsen (2010). "Microbial Ecology of Activated Sludge." Microbial Ecology of Activated Sludge.

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.

Sharpley, A., H. P. Jarvie, A. Buda, L. May, B. Spears and P. Kleinman (2013). "Phosphorus Legacy: Overcoming the Effects of Past Management Practices to Mitigate Future Water Quality Impairment." Journal of Environmental Quality **42**(5): 1308-1326.

Sharpley, A. N., H. Tiessen and C. V. Cole (1987). "Soil-Phosphorus Forms Extracted by Soil Tests as a Function of Pedogenesis." Soil Science Society of America Journal **51**(2): 362-365.

Shokralla, S., J. L. Spall, J. F. Gibson and M. Hajibabaei (2012). "Next-generation sequencing technologies for environmental DNA research." Molecular Ecology **21**(8): 1794-1805.

Streichan, M., J. R. Golecki and G. Schon (1990). "Polyphosphate-Accumulating Bacteria from Sewage Plants with Different Processes for Biological Phosphorus Removal." Fems Microbiology Ecology **73**(2): 113-124.

Tandoi, V., M. Majone, J. May and R. Ramadori (1998). "The behaviour of polyphosphate accumulating Acinetobacter isolates in an anaerobic-aerobic chemostat." Water Research **32**(10): 2903-2912.

Tchobanoglous, G., F. L. Burton and H. D. Stensel (2002). Wastewater Engineering Treatment and Reuse, McGraw-Hill Education.

Tijssen, J. P. F., H. W. Beekes and J. Vansteveninck (1982). "Localization of Polyphosphates in Saccharomyces-Fragilis, as Revealed by 4',6-Diamidino-2-Phenylindole Fluorescence." Biochimica Et Biophysica Acta **721**(4): 394-398.

Tobin, K. M., J. W. McGrath, A. Mullan, J. P. Quinn and K. E. O'Connor (2007). "Polyphosphate accumulation by Pseudomonas putida CA-3 and other medium-chain-length polyhydroxyalkanoate-accumulating bacteria under aerobic growth conditions." Applied and Environmental Microbiology **73**(4): 1383-1387.

Tu, Y. J. and A. J. Schuler (2013). "Low Acetate Concentrations Favor Polyphosphate-Accumulating Organisms over Glycogen-Accumulating Organisms in Enhanced Biological Phosphorus Removal from Wastewater." Environmental Science & Technology **47**(8): 3816-3824.

Wagner, M., R. Amann, H. Lemmer, W. Manz and K. H. Schleifer (1994). "Probing Activated-Sludge with Fluorescently Labeled Ribosomal-Rna-Targeted Oligonucleotides." Water Science and Technology **29**(7): 15-23.

Weissbrodt, D. G., G. S. Schneiter, J. M. Furbringer and C. Holliger (2013). "Identification of trigger factors selecting for polyphosphate- and glycogen-accumulating organisms in

aerobic granular sludge sequencing batch reactors." Water Research **47**(19): 7006-7018.

Wetzel, R. G. (2001). Limnology: lake and river ecosystems. San Diego, Academic Press.

Willems, A., J. Deley, M. Gillis and K. Kersters (1991). "Comamonadaceae, a New Family Encompassing the Acidovorans Ribosomal-Rna Complex, Including Variovorax-Paradoxus Gen-Nov, Comb-Nov, for Alcaligenes-Paradoxus (Davis 1969)." International Journal of Systematic Bacteriology **41**(3): 445-450.

Woese, C. R. (1987). "Bacterial Evolution." Microbiological Reviews **51**(2): 221-271.

Yarza, P., P. Yilmaz, E. Pruesse, F. O. Glockner, W. Ludwig, K. H. Schleifer, W. B. Whitman, J. Euzeby, R. Amann and R. Rossello-Mora (2014). "Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences." Nature Reviews Microbiology **12**(9): 635-645.

Yilmaz, L. S., H. E. Okten and D. R. Noguera (2006). "Making all parts of the 16S rRNA of Escherichia coli accessible in situ to single DNA oligonucleotides." Applied and Environmental Microbiology **72**(1): 733-744.

Zeng, R. J., A. M. Saunders, Z. G. Yuan, L. L. Blackall and J. Keller (2003). "Identification and comparison of aerobic and denitrifying polyphosphate-accumulating organisms." Biotechnology and Bioengineering **83**(2): 140-148.

Zhang, J. J., K. Kobert, T. Flouri and A. Stamatakis (2014). "PEAR: a fast and accurate Illumina Paired-End reAd mergeR." Bioinformatics **30**(5): 614-620.

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.

Zilles, J. L., J. Peccia, M. W. Kim, C. H. Hung and D. R. Noguera (2002). "Involvement of Rhodocyclus-Related Organisms in Phosphorus Removal in Full-Scale Wastewater Treatment Plants." Applied and Environmental Microbiology **68**(6): 2763-2769.

Appendix A

Microscopy Filter Set

Table A-1. Filter sets used to visualize fluorescent signals.

Filter #	EX/EM	Wavelength	Bandwidth	Fluorochrome
D360/40 X 47899	EX	360	40	DAPI
N/A, UV	EX	360	60	DAPI
XF1073 475AF40 230 0121	EX	475	40	FITC
HQ545/30 X 47605	EX	545	30	Cy3
D460/50 M 47713	EM	460	50	DAPI-DNA
HQ575/40 M	EM	575	40	DAPI-P
D535/40 M 69144	EM	535	40	FITC
HQ610/75 M 47359	EM	610	75	Cy3

Appendix B

DNA Concentration Data

Table B-1. Concentrations of DNA from the raw biofilm samples and EBPR activated sludge from the University Area Joint Authority biological nutrient removal process.

Sample	Conc. (ng/μL)	A260	A280	260/280	260/230
East Hickory Cr.	1.6	0.031	0.012	2.59	0.51
Cowanesque R.	12.1	0.242	0.160	1.51	0.53
Red Clay Cr.	1.9	-0.037	0.035	1.06	0.41
Cooks Cr.	1.1	0.022	0.004	5.77	1.59
Penns Cr.	0.5	0.010	0.003	2.07	0.19
Spring Cr.	12.6	0.252	0.155	1.62	0.79
EBPR	14.6	0.292	0.185	1.58	0.52

Table B-2. Concentrations of DNA from the non-sorted FD-36 experimental watershed biofilm samples.

Sample	Conc. (ng/μL)	A260	A280	260/280	260/230
FD-36 F1 P2	2.6	0.052	0.036	1.44	0.35
FD-36 F1 P4	3.3	0.066	0.039	1.70	0.41
FD-36 F2 P2	0.2	0.005	0.008	0.58	0.09
FD-36 F2 P4	0.4	0.009	-0.001	-11.11	0.11
FD-36 F3 P2	0.4	0.007	0.000	91.61	0.15
FD-36 F3 P4	3.7	0.075	0.037	2.01	0.58
FD-36 F4 P2	0.7	0.015	0.002	6.93	0.30
FD-36 F4 P4	1.2	0.024	0.019	1.27	0.25

Table B-3. Concentrations of DNA from the PAO sorted fraction of the FD-36 experimental watershed biofilm samples.

Sample	Conc. (ng/μL)	A260	A280	260/280	260/230
FD-36 F1 P2	10.2	0.205	0.159	1.29	0.57
FD-36 F1 P4	1.3	0.027	0.008	3.30	0.42
FD-36 F2 P2	7.7	0.154	0.094	1.64	0.53
FD-36 F2 P4	0.8	0.015	0.008	1.96	0.48
FD-36 F3 P2	0.3	0.007	-0.004	-1.72	0.49
FD-36 F3 P4	0.9	0.018	0.013	1.34	0.27
FD-36 F4 P2	0.7	0.015	0.005	3.23	0.27
FD-36 F4 P4	0.5	0.010	0.008	1.17	0.26

Table B-4. Concentrations of DNA from the non-PAO sorted fraction of the FD-36 experimental watershed biofilm samples.

Sample	Conc. (ng/μL)	A260	A280	260/280	260/230
FD-36 F1 P2	3.5	0.071	0.038	1.87	0.51
FD-36 F1 P4	2.7	0.054	0.033	1.63	0.53
FD-36 F2 P2	0.9	0.018	0.009	2.10	0.35
FD-36 F2 P4	0.4	0.008	0.013	0.61	0.25
FD-36 F3 P2	0.4	0.007	0.008	0.94	0.21
FD-36 F3 P4	1.4	0.029	0.019	1.53	0.51
FD-36 F4 P2	4.6	0.091	0.064	1.42	0.56
FD-36 F4 P4	1.5	0.031	0.018	1.62	0.44

Appendix C

Illumina MiSeq Microbial Diversity Data

Table C-1. OTU classification and number of sequence hits for samples of the Pennsylvania streams.

OTU	Taxonomy (to highest resolution)	Count					
		E. H.	Cow.	R. C.	Cook	Penn	Spr.
0	*Curtobacterium sp*	22	0	0	0	4	0
1	*Alcaligenaceae*	0	0	0	0	7	0
2	*Methylobacterium*	0	0	0	0	4	0
3	*Sphingomonadaceae*	224	0	0	0	0	0
4	*Enterobacteriaceae*	37	0	0	0	0	0
5	*Propionibacterium acnes*	103	33	11	19	925	0
6	*Kozakia baliensis*	0	0	0	24	0	0
7	*Stenotrophomonas*	29	0	18	0	145	96
8	*Azospirillum*	0	0	44	0	0	0
9	*Proteobacteria*	12	0	0	0	10	0
10	*Corynebacterium*	24	0	21	5	0	0
11	*Sphingobacteriaceae*	7	0	0	0	0	0
12	No Hit	0	0	20	0	0	0
13	*Derxia sp*	93	0	15	18	10	2
14	*Escherichia coli*	28	23	29	16	25	0
15	*Diaphorobacter*	0	0	0	4	2	0
16	*Burkholderiales*	86	68	0	34	145	36
17	*Comamonadaceae*	0	0	0	25	8	0
18	*Bacteria*	141	0	0	0	0	0

19	*Rhizobiales*	0	0	22	0	39	0
20	*Sphingobium sp*	0	20	19	2	525	0
21	*Bacteria*	0	0	0	0	60	0
22	*Rhizobiales*	24	0	0	0	0	0
23	*Planctomyces sp*	40	0	0	0	0	0
24	*Rhodobacteraceae*	16	0	0	0	0	0
25	*Pseudomonas sp*	40	0	1	2	1	0
26	*Pseudomonas*	292	179	201	320	115	471
27	*Proteobacteria*	105	66	42	46	43	2
28	*Pseudomonas*	0	0	1	32	1	0
29	*Proteobacteria*	20	9	0	1	84	20
30	*Staphylococcus epidermidis*	0	96	0	0	0	0
31	*Comamonadaceae*	0	0	0	0	15	0
32	*Acidobacterium sp*	13	0	0	0	0	0
33	*Micrococcus sp*	10	0	0	0	0	0
34	*Bradyrhizobium*	0	0	8	0	6	0
35	*Thauera sp*	68	807	28	108	69	2
36	*Alcaligenes sp*	8	0	0	0	2	0
37	*Diaphorobacter*	11	0	0	28	19	0
38	No Hit	0	0	0	0	4	0
39	*Gluconacetobacter*	0	0	0	224	0	0
40	*Burkholderia sp*	140	0	13	59	17	0
41	*Tissierella sp*	22	0	0	0	0	0
42	*Schlegelella sp*	0	0	38	0	0	0
43	*Comamonadaceae*	0	0	17	0	0	0

44	*Acinetobacter sp*	16	0	34	5	0	0
45	*Massilia sp*	0	0	0	1372	0	0
46	*Methylobacillus*	11	0	0	0	0	0
47	*Comamonadaceae*	2217	1024	204	2409	1593	359
48	*Ignavibacterium album*	335	383	169	188	144	383
49	*Pseudomonas*	1	0	1	3	9	0
50	*Proteobacteria*	30	0	0	6	0	0
51	*Lactococcus lactis*	53	0	0	0	0	0
52	*Delftia*	0	0	11	0	0	0
53	*Herbaspirillum*	83	0	22	0	9	0
54	*Curtobacterium*	58	0	13	0	0	0
55	*Diaphorobacter nitroreducens*	0	0	0	0	9	0
56	*Geobacillus*	35	5283	14	21	0	0
57	*Burkholderiales*	21	0	0	0	0	0
58	*Porphyromonas sp*	27	0	0	0	0	0
59	*Pseudomonas sp*	13	0	0	191	0	0
60	*Pseudomonas sp*	1243	0	1110	206	236	0

Table C-2. OTU classification and number of sequence hits for samples of the USDA FD-36 experimental watershed.

OTU	**Taxonomy (to highest resolution)**	**F1 P2**	**F1 P4**	**F2 P2**	**F2 P4**	**F3 P2**	**F3 P4**	**F4 P2**	**F4 P4**
0	*Roseomonas sp*	0	23	0	0	0	0	0	0
1	*Pseudomonas fragi*	0	1034	0	0	0	0	0	0
2	*Sphingomonadaceae*	1191	332	1121	1630	2408	221	139	93
3	*Rhizobium*	0	0	0	0	0	669	0	0

4	*Pantoea*	8	0	0	0	0	0	0	0
5	*Bacteroidetes*	0	0	1169	0	0	0	0	0
6	*Dokdonella sp*	0	3	0	0	243	0	127	0
7	*Neisseriaceae*	0	0	0	0	0	0	16	0
8	*Intrasporangiaceae*	7	0	0	0	0	0	0	0
9	*Sugarcane phytoplasma*	0	3287	0	1	0	0	0	0
10	*Synechococcus sp*	222	0	0	0	0	0	0	0
11	*Ignavibacterium*	1098	638	1167	579	1369	235	340	205
12	*Myxococcales*	0	0	5	0	0	0	0	0
13	No Hit	0	0	0	0	0	0	0	929
14	*Curtobacterium sp*	0	0	0	0	0	487	0	0
15	*Novosphingobium*	1	0	0	1713	0	0	0	0
16	*Gordonia*	0	0	662	0	0	0	0	0
17	*Chloroflexus sp*	0	0	0	2297	0	0	0	0
18	*Streptomyces sp*	0	891	0	0	0	0	0	0
19	*Cyanobacteria*	0	0	0	0	0	0	0	502
20	*Listeria fleischmannii*	0	0	1921	0	0	0	0	0
21	*Burkholderiales*	0	0	0	0	0	2	0	0
22	*Bacteria*	770	1392	4	2104	1237	126	559	428
23	*Anaerococcus*	60	0	0	0	0	0	0	0
24	*Actinobacteria*	0	9	133	0	0	0	0	0
25	*Actinomycetales*	0	0	0	0	0	0	0	166
26	*Bacteria*	0	0	0	0	0	319	0	0
27	*Haemophilus parainfluenzae*	221	0	0	0	0	0	257	0

28	*Alloprevotella tannerae*	0	0	0	0	0	0	92	0
29	*Defluviicoccus sp*	0	0	0	0	438	0	0	0
30	*Burkholderiales*	0	0	2	0	0	0	0	0
31	*Burkholderiales*	1	0	0	0	2052	0	0	0
32	*Burkholderiales*	0	0	4	60	73	0	0	0
33	*Anaerococcus sp*	0	0	593	0	0	0	0	0
34	*Sphingomonadaceae*	8908	4121	15680	9602	7290	2979	3172	4151
35	*Steroidobacter sp*	0	0	10	0	0	7	33	0
36	*Burkholderiales*	0	0	0	0	0	183	0	0
37	*Corynebacterium*	1074	0	0	0	0	0	0	0
38	*Sphingomonas sp*	0	0	0	0	0	13927	0	0
39	*Deltaproteobacteria*	127	0	0	0	0	166	0	0
40	*Meganema sp*	0	0	1	0	91	0	0	0
41	*Rhizobiales*	0	0	0	0	103	0	18	0
42	*Corynebacterium tuberculostearicum*	0	0	0	1386	43	0	0	0
43	*Actinomycetales*	11	0	0	0	0	0	0	0
44	*Alphaproteobacteria*	0	0	0	0	0	0	0	2
45	*Nitrospira sp*	2	0	0	0	132	0	0	0
46	*Curtobacterium sp*	0	0	0	0	0	1080	0	0
47	*Sphingomonadaceae*	0	0	15	0	1	98	0	1
48	*Cyanobacteria*	139	0	910	0	0	0	0	0
49	*Burkholderiales*	0	0	0	0	507	0	0	0
50	*Rubrobacter sp*	0	0	0	0	0	0	136	0
51	*Proteobacteria*	0	0	0	0	0	0	16	0
52	*Acidobacteriia*	0	0	0	1024	0	1	0	0

53	*Cytophagales*	0	413	0	0	0	0	0	0
54	*Sphingomonas sp*	0	0	0	0	0	57	0	0
55	*Actinomyces sp*	0	0	0	0	18	0	0	0
56	*Bacteria*	97	0	0	0	0	0	0	0
57	*Neisseria*	1160	0	0	0	0	0	0	0
58	*Dokdonella sp*	0	147	0	0	0	0	0	0
59	*Clostridia*	22	0	0	0	0	0	0	0
60	*Crocinitomix sp*	0	0	0	0	0	0	0	127
61	*Alphaproteobacteria*	0	0	0	0	0	0	85	0
62	*Phycisphaerae*	0	0	0	0	0	0	11	0
63	*Actinobacteria*	0	0	0	0	0	0	1505	0
64	*Aurantimonas sp*	0	0	0	0	0	3664	0	0
65	*Burkholderiaceae*	0	0	0	0	0	234	0	0
66	*Bacteria*	0	0	0	0	390	0	0	0
67	*Sphingobacteriales*	12	0	0	0	164	0	0	0
68	*Intrasporangiaceae*	0	0	0	0	5	69	0	0
69	*Burkholderiales*	0	189	0	0	0	0	0	0
70	*Microbacterium sp*	0	0	0	0	0	319	0	0
71	*Thermobifida sp*	0	0	15	0	0	0	0	0
72	*Sphingobacteriales*	0	0	0	0	0	0	2	0
73	No Hit	209	0	54	0	0	67	0	0
74	*Bacteroidetes*	0	0	0	0	0	0	0	36
75	*Proteobacteria*	0	0	1568	0	0	0	0	0
76	*Oxalobacteraceae*	0	0	279	0	0	0	0	0
77	*Burkholderiales*	0	13	0	0	0	0	0	0

78	*Rhizobiales*	1422	0	0	2	0	0	0	0
79	*Micrococcus*	0	0	0	0	0	0	808	0
80	*Actinomycetales*	0	0	0	0	0	666	0	0
81	No Hit	0	0	0	0	146	0	0	0
82	*Methylobacterium*	0	0	0	1	0	25	0	0
83	*Curtobacterium sp*	0	0	0	0	0	100	0	0
84	No Hit	0	0	0	0	143	0	0	0
85	*Hymenobacter*	0	0	0	0	0	1771	0	0
86	*Actinomyces sp*	0	0	0	0	0	0	47	0
87	*Acidobacteriales*	0	0	0	0	183	0	0	0
88	*Acidobacteria*	0	0	0	2253	0	0	0	0
89	No Hit	0	0	36	0	0	0	0	0
90	*Betaproteobacteria*	293	2071	453	1085	1120	131	935	217
91	*Sphingomonas*	274	0	0	0	173	9	0	0
92	*Intrasporangiaceae*	0	0	0	718	0	0	0	0
93	*Bacteria*	2	1	0	0	0	0	0	0
94	*Clostridia*	0	0	0	0	593	0	0	0
95	*Rickettsiales*	0	0	0	0	0	0	18	0
96	*Pseudomonas sp*	0	0	0	94	0	1	2	1
97	*Rhizobiales*	0	0	0	0	0	0	0	1085
98	*Bacillus sp*	2863	0	0	6489	0	0	980	1752
99	*Bacteroidetes*	5486	0	0	779	3678	696	2888	1122
100	*Enterobacteriaceae*	448	442	0	1622	460	62	1299	297
101	*Acidobacterium sp*	0	0	0	0	0	0	0	26
102	*Clostridiales*	0	0	0	0	0	0	0	3

103	*Rhizobiales*	180	0	0	0	0	668	0	0
104	No Hit	0	0	0	0	2	0	0	0
105	*Rhodospirillales*	0	0	14	0	0	0	0	0
106	*Granulicatella adiacens*	0	0	0	0	1423	0	0	0
107	*Rhizobiales*	0	0	0	996	0	0	0	0
108	*Spirosoma sp*	1	0	0	1	0	21857	0	0
109	No Hit	77	2060	0	0	0	0	0	0
110	*Ochrobactrum tritici*	1	68	0	1	0	0	0	0
111	No Hit	0	0	103	0	0	0	0	0
112	*Streptococcus*	223	0	0	0	0	0	0	0
113	*Microscilla sp*	0	0	367	0	0	0	0	0
114	*Staphylococcus*	0	0	0	0	0	0	204	0
115	*Sphingomonas*	0	0	0	0	0	2364	0	0
116	*Methylobacterium*	362	0	0	0	0	4	14	0
117	*Burkholderiales*	29	2	5	666	2420	2	536	0
118	*Enterococcus cecorum*	0	0	0	0	0	0	100	0
119	*Gemella*	0	0	0	1303	0	0	0	0
120	*Pseudomonas*	339	1427	1842	740	471	149	416	0
121	*Burkholderiales*	0	0	0	2	35	0	0	0
122	*Staphylococcus*	59	717	0	0	1246	0	454	0
123	*Betaproteobacteria*	159	0	47	0	444	0	79	756
124	*Gammaproteobacteria*	5	3	6	4	198	6	1	1
125	No Hit	0	0	0	0	0	0	136	0
126	*Proteobacteria*	0	0	0	0	0	0	11	0
127	*Bacteria*	3	0	0	0	0	0	0	0

128	*Proteobacteria*	0	0	0	0	976	0	0	0
129	*Flavobacterium sp*	0	0	0	0	202	0	0	0
130	No Hit	0	77	0	0	0	0	0	0
131	*Burkholderiaceae*	8	0	0	0	0	0	0	76
132	*Bacillus*	1137	0	0	3352	461	167	945	0
133	No Hit	99	195	0	0	0	0	0	147
134	*Propionibacterium*	55	0	0	0	0	0	0	0
135	*Comamonadaceae*	71	0	0	0	0	0	0	0
136	*Actinomycetales*	57	0	0	30	3	0	0	0
137	*Actinomycetales*	2	0	0	0	282	0	0	0
138	*Alphaproteobacteria*	600	0	0	0	0	0	0	0
139	*Rhizobiales*	0	0	0	0	0	0	108	0
140	*Granulicatella*	501	0	0	0	0	0	0	0
141	*Bacteroidetes*	0	0	0	0	270	0	0	0
142	*Rothia sp*	14	30	0	0	0	0	0	0
143	*Rhizobiales*	552	8	0	0	667	0	0	8
144	*Friedmanniella*	148	0	0	0	0	65	0	0
145	*Actinomycetales*	0	0	0	65	0	0	0	0
146	*Bacteroidetes*	772	0	0	0	0	0	0	0
147	*Delftia*	0	0	0	0	1	1	0	124
148	*Oxalobacteraceae*	0	1186	0	0	0	0	0	0
149	*Leptotrichia sp*	0	0	0	0	0	0	931	0
150	*Streptococcus parasanguinis*	0	0	0	892	0	0	0	0
151	*Cytophagales*	0	0	0	0	0	30	0	0
152	*Campylobacter*	0	0	0	0	0	0	1242	0

153	*Bacteria*	0	0	0	0	530	0	0	0
154	*Pedobacter*	0	0	0	0	0	272	0	0
155	*Bosea*	235	0	0	0	778	0	0	0
156	*Pseudomonas*	0	0	0	0	0	63	0	0
157	*Neisseriales*	625	0	0	0	0	0	0	0
158	No Hit	0	0	0	0	2	0	0	0
159	*Meganema sp*	141	0	0	0	0	0	0	0
160	*Arthrospira platensis*	0	0	502	0	0	0	0	0
161	*Acidovorax*	344	257	0	369	7018	435	61	409
162	*Dechloromonas sp*	0	0	0	0	514	0	0	0
163	*Haemophilus*	0	9	0	0	0	0	0	0
164	*Dolosigranulum sp*	625	0	0	0	0	0	0	0
165	*Finegoldia sp*	0	0	873	0	0	0	0	0
166	*Rhodocyclales*	0	0	133	0	542	0	0	0
167	*Bacteria*	0	0	0	0	26	0	0	0
168	*Sphingomonas sp*	666	0	0	1	0	7189	0	0
169	*Sphingomonas*	0	0	0	0	0	9	0	0
170	*Microbacteriaceae*	0	0	0	0	0	111	0	0
171	*Burkholderiales*	0	0	0	0	0	1316	0	0
172	*Corynebacterium*	954	0	0	0	0	0	0	0
173	*Proteobacteria*	0	0	0	0	1020	0	0	0
174	*Methylobacterium sp*	0	0	0	0	0	193	0	0
175	*Burkholderiales*	34	0	0	0	443	0	0	0
176	*Aquiflexum sp*	2	0	0	0	68	0	0	0
177	*Propionibacterium granulosum*	0	0	0	0	1	325	0	0

178	*Fusobacterium*	0	0	0	0	0	0	47	0
179	*Prevotella micans*	0	0	0	0	0	0	392	0
180	*Prevotella sp*	582	0	0	0	0	0	26	0
181	*Bacteria*	50	0	0	0	0	0	0	0
182	*Hymenobacter sp*	0	0	0	0	0	157	0	0
183	*Corynebacterium*	0	0	0	1250	0	0	0	0
184	*Neisseriaceae*	0	0	0	0	61	0	0	0
185	*Methylobacterium*	0	0	0	0	0	79	0	0
186	*Streptococcus*	4631	1878	0	0	35	1	44	0
187	*Dechloromonas sp*	0	67	0	0	0	0	0	0
188	*Proteobacteria*	105	0	0	0	0	0	0	0
189	*Betaproteobacteria*	0	0	0	0	0	0	0	6
190	*Aquiflexum sp*	0	0	0	0	4	0	0	0
191	*Methylobacterium sp*	40	0	0	0	0	0	0	0
192	*Propionibacterium*	15443	25781	31127	10282	14859	2699	8621	2491
193	*Dechloromonas sp*	45	0	0	586	23	0	0	195
194	*Pseudomonas*	26	260	104	0	0	0	0	187
195	*Sphingobacteriales*	6	0	0	1	116	0	0	0
196	*Candidatus Accumulibacter sp*	27	0	0	0	391	0	0	0
197	*Flavobacterium sp*	0	0	0	0	0	0	651	0
198	*Bradyrhizobium sp*	527	3327	507	0	2221	0	873	3
199	*Escherichia coli*	444	0	9	913	1252	75	200	0
200	*Acinetobacter sp*	0	0	267	0	0	0	0	0
201	*Proteobacteria*	0	1790	13	0	0	951	21	0
202	*Acinetobacter sp*	380	0	0	0	0	0	0	0

203	*Burkholderiales*	1612	1519	736	508	1728	6	720	1000
204	*Geodermatophilus sp*	0	0	0	0	0	780	0	0
205	No Hit	0	0	0	0	0	0	0	86
206	*Sphingomonadales*	14773	3141	1358	2310	4976	570	1776	1027
207	*Phyllobacteriaceae*	0	0	0	0	0	0	0	186
208	No Hit	0	0	0	0	8	78	24	4
209	*Sphingomonas sp*	0	0	0	0	0	435	0	0
210	*Bacteria*	0	311	0	0	0	0	0	0
211	*Streptococcus*	154	0	0	0	0	0	0	0
212	*Rhodobacter*	0	0	0	0	173	0	0	0
213	*Deltaproteobacteria*	0	0	0	0	919	0	0	0
214	*Actinobacteria*	0	0	0	0	205	0	0	0
215	*Aquabacterium sp*	0	0	0	0	0	0	0	69
216	*Chitinophagaceae*	526	0	861	0	0	0	0	0
217	*Tolumonas sp*	130	0	0	0	0	0	0	0
218	*Staphylococcus sp*	1333	364	0	0	4	179	886	60
219	*Chroococcales*	0	0	0	0	0	0	0	9
220	*Rhodospirillales*	0	0	0	0	2301	1	0	0
221	*Deltaproteobacteria*	0	0	0	0	80	0	0	0
222	*Caldilinea sp*	820	0	0	0	0	0	0	0
223	*Bacteria*	649	24	0	0	0	0	0	0
224	*Sphingomonadales*	51	0	0	0	0	0	0	0
225	*Staphylococcus*	54	0	0	0	0	0	0	0
226	*Nesterenkonia sp*	0	65	0	0	0	0	0	0
227	*Xanthomonadales*	0	0	0	0	0	17	0	0

228	*Burkholderiales*	0	0	245	0	0	0	0	0
229	*Flavobacteriales*	0	0	0	0	0	0	546	0
230	*Acidobacteriia*	0	0	0	0	199	0	0	0
231	*Actinomyces sp*	0	0	0	0	0	0	300	0
232	No Hit	0	0	0	0	226	0	0	0

Appendix D

Biogeochemical Stream Data

Table D-1. Benthic biofilm stream measurements for the eight Pennsylvania streams studied.

Sample	P Uptake ((mg chl-a/m^2)$^{-1}$ min^{-1})	Total P uptake rate (nmolP/ ugchl/ hr)	APA (µg ch-a^{-1} hr^{-1})	Nitrate Reductase (nmol NO_2)	poly-P (mg P / m^2)	Particulate P (mg P / m^2)	14C Production (mgC/µgChl/d)
Penns Creek	0.49	4.57	0.27	4.44	4.59	117	0.0023
Cooks Creek	0.11	1.31	1.16	1.72	10.2	147	0.0106
East Hickory Creek	0.42	3.14	3.83	0.92	0.42	4.18	0.0016
Tionesta Creek	0.73	4.97	4.45	0.42	0.38	8.97	0.0015
Tunkhannock Creek	0.48	2.14	1.80	1.00	3.09	65.7	0.0012
Red Clay Creek	0.020	0.19	0.65	3.15	43.2	349	0.0004
Spring Creek	0.135	1.93	2.72	4.36	23.6	248	0.0005
Cowanesque River	0.067	0.32	1.34	6.32	12.5	94.9	0.0029