December 2018

River Bank Inducement Influence on a Shallow Groundwater Microbial Community and Its Effects on Aquifer Reactivity

Natalie June Gayner
University of Wisconsin-Milwaukee

RIVERBANK INDUCEMENT INFLUENCE ON A SHALLOW GROUNDWATER MICROBIAL COMMUNITY AND ITS EFFECTS ON AQUIFER REACTIVITY

by

Natalie Gayner

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Freshwater Sciences & Technology

at

The University of Wisconsin-Milwaukee

December 2018

ABSTRACT

RIVER BANK INDUCEMENT INFLUENCE ON A SHALLOW GROUNDWATER MICROBIAL COMMUNITY AND ITS AFFECT ON AQUIFER REACTIVITY

by

Natalie Gayner

The University of Wisconsin-Milwaukee, 2018
Under the Supervision of Professor Ryan J. Newton, PhD

Placing groundwater wells next to riverbanks to draw in surface water, known as riverbank inducement (RBI), is common and proposed as a promising and sustainable practice for municipal and public water production across the globe. However, these systems require further investigation to determine risks associated with river infiltration especially with rivers containing wastewater treatment plant (WWTP) effluent. Since microbes drive biogeochemical transformations in groundwater and largely affect water quality, it is important to understand how the microbial communities in drinking water wells are affected by river infiltration. This study investigated if, and to what extent, the microbial community in a shallow groundwater aquifer in southeastern Wisconsin is affected by river infiltration. The study area includes an active RBI well, a previously active RBI well, a pristine background well, and the Fox River in Waukesha, WI. After targeting both DNA and RNA for V4 16S rRNA gene sequencing, the results show the microbial community compositions of the groundwater sites significantly differ from each other and from the Fox River. Microbial community compositions correlated with Total Dissolved Phosphorus (TDP) and Total Nitrogen (TN). Amplicon sequence variants

(ASVs) associated with river bacteria were found in all groundwater wells, however, these taxa were always more abundant in the active RBI well with similar distribution patterns to the river. The aquifer microbial community composition was over 50% Unclassified organisms. Some ASVs showed evidence of intron splicing in the 16S rRNA gene, a rarely recorded feature in bacteria. The aquifer microbial communities also contained common subsurface organisms and recently discovered CPR and DPANN superphyla organisms. The taxa affiliations suggest heterotrophic, fermentative, and symbiotic lifestyles, and suggest anaerobic metabolisms related to nitrate and sulfate reduction. Microbial affiliation results are consistent with free energy flux predictions for the groundwater wells. Lab experiments indicated the water itself may be C limited and that additional nitrate from river infiltration may initially accumulate in the system, which could impact required water treatment processes.

Dedicated to

Karlann “Grandma-Goodie-Legs” Coughlin
who received her M.S. in Biochemistry in 1946 from The University of Michigan

TABLE OF CONTENTS

ABSTRACT II
TABLE OF CONTENTS VI
LIST OF FIGURES VIII
LIST OF TABLES XI
LIST OF ABBREVIATIONS XIII
ACKNOWLEDGEMENTS XIV
CHAPTER 1 – INTRODUCTION 1
1.2 – BACKGROUND AND SIGNIFICANCE 3
1.2.1 GROUNDWATER USE IN THE UNITED STATES 3
1.2.2 GROUNDWATER IN SOUTHEAST WISCONSIN 3
1.2.3 RIVER BANK INDUCEMENT WELLS IN WAUKESHA, WI 4
1.2.4 MICROORGANISMS IN GROUNDWATER 5
1.3 STUDY SYSTEM 7
1.4 PREVIOUS RESEARCH 8
CHAPTER 2 – RIVER INFILTRATION EFFECTS ON THE SHALLOW GROUNDWATER AQUIFER MICROBIAL COMMUNITY 12
2.1 BACKGROUND 12
2.2 METHODS 15
2.2.1 SAMPLE COLLECTION 15
2.2.2 LAB METHODS 16
2.2.3 COMMUNITY COMPOSITION SEQUENCE DATA PROCESSING 19
2.3 ANALYSIS, RESULTS & DISCUSSION 21
2.3.1 CHEMICAL & THERMODYNAMIC RESULTS 21
2.3.1 FOX RIVER AND GROUNDWATER MICROBIAL COMMUNITY COMPOSITIONS 24
2.3.2 GROUNDWATER MICROBIAL COMMUNITY COMPARISON 27
2.3.3 GROUNDWATER MICROBIAL COMMUNITY TAXA AFFILIATIONS 32
UNCLASSIFIED ORGANISMS DOMINATE THE GROUNDWATER DATASET 32
POTENTIAL PHYSIOLOGICAL IMPLICATIONS OF DOMINANT MICROBIAL AFFILIATIONS 36
TAXA COMPARISONS BETWEEN SITES 38
2.3.3 16S RIBOSOMAL RNA:DNA ACTIVITY 46
CHAPTER 3 –NUTRIENT EFFECTS 53
3.1 BACKGROUND 53
3.2 METHODS 55
3.3 RESULTS & DISCUSSION 58
CONCLUSIONS 63
CHAPTER 4 – FUTURE RECOMMENDATIONS 66
4.1 RATIONALE 66
4.2 RECOMMENDATIONS 69

REFERENCES 70
APPENDICES 77
APPENDIX A: SAMPLING SITE COORDINATES 78
APPENDIX B: DETAILED LAB METHOD AND PROTOCOLS 79
DNA/RNA EXTRACTION FOR 16S RRNA GENE SEQUENCING 79
DNASE TREATMENT 79
REVERSE TRANSCRIPTION PCR (RT-PCR) 80
V4 16S RRNA GENE POLYMERASE CHAIN REACTION (PCR) 80
SEQUENCING 81
DAPI SAMPLE PROCESSING INSTRUCTIONS FOR GROUNDWATER SAMPLES 82
GDNA EXTRACTION FOR FUTURE SHOTGUN METAGENOMIC SEQUENCING 84
APPENDIX C: NUTRIENT EXPERIMENT RAW DATA RESULTS 88
NITRATE RESULTS 88
TOTAL DISSOLVED PHOSPHORUS RESULTS 89
AMMONIUM RESULTS 90
SOLUBLE REACTIVE PHOSPHORUS RESULTS 91
NUTRIENT EXPERIMENT CELL COUNTS 93
APPENDIX D: W12 UNIQUE DNA AND RNA TAXA SPECIALISTS 94
APPENDIX E: W13 UNIQUE DNA AND RNA TAXA SPECIALISTS 104
APPENDIX F: TOP AVERAGE RATIOS USED IN VIOLIN RATIO ANALYIS ACROSS WELLS (FIG. 13, PG 44) 110
APPENDIX G: RIVER INFILTRATION INDICATOR TAXA 114
APPENDIX H: HEATMAP CLOSE UP AVERAGE MOST ABUNDANT TAXA OF INTEREST 116
APPENDIX I: MOCK COMMUNITY DATASHEET 117
APPENDIX J: SUPPLEMENTARY 16S RIBOSOMAL RNA:DNA ACTIVITY RATIOS 119
APPENDIX K: MEASURING SOLUBLE REACTIVE PHOSPHOROUS PROTOCOL 121

LIST OF FIGURES

FIGURE 1: GEOGRAPHIC LOCATIONS OF THE STUDY SITE IN WAUKESHA, WI ALONG THE FOX RIVER. WASTEWATER TREATMENT PLANT (WWTP) EFFLUENT FLOWS INTO THE FOX RIVER UPSTREAM OF THE WELLS. W12 ACTIVELY DRAWS IN RIVER WATER AS AN RBI WELL. W11 PREVIOUSLY WAS AN RBI WELL BUT NO LONGER ACTIVELY PUMPS AT RATE WHICH DRAWS IN RIVER WATER. BACKGROUND WELL W13 IS CONSIDERED PRISTINE. 7

FIGURE 3: SUCRALOSE CONCENTRATIONS (NG/L) IN THE FOX RIVER, W11, W12, AND BACKGROUND WELL W13 (LEFT TO RIGHT). SINGLE MEASUREMENTS FROM1 LITER SINGLE SAMPLES FROM THE SPRING OF 2015 WERE PERFORMED AT UW-STEVEN'S POINT WATER AND ENVIRONMENTAL ANALYSIS LAB (WEAL). THE DETECTION LIMIT WAS 25 NG/L. ADAPTED FROM FIELDS-SOMMERS (2015). 9

FIGURE 3 MAJOR ION CONCENTRATIONS IN THE WELLS. ION CONCENTRATIONS OVER TIME WITH PUMPAGE IN RBI WELLS (TOP LEFT WK12, RIGHT WK11) AND PRISTINE WELL (BOTTOM WK13). 10

FIGURE 4 HEATMAP ILLUSTRATING THE RELATIVE ABUNDANCE OF BACTERIAL FAMILIES (ONLY FAMILIES PRESENT AT >2% OF THE COMMUNITY COMPOSITION IN AT LEAST ONE SAMPLE DEPICTED). PRELIMINARY 16S RRNA GENE COMMUNITY COMPOSITION RELATIVE ABUNDANCES OF ONE SAMPLE EACH REVEAL THE DOMINANT FAMILIES BETWEEN WWTP, FOX RIVER, AND RBI WELL W12 VARY ACROSS SITES. 11

FIGURE 5: NUTRIENT AND ION CHEMISTRY ORDINATION FROM NMDS AND EUCLIDEAN DISTANCE MATRIX FROM NUTRIENT AND ION DATA FOR THE THREE WELLS. THE LENGTH AND ANGLE OF THE ARROW CORRESPONDS TO DATE CORRELATION. 23

FIGURE 6: NMDS ORDINATION OF FOX RIVER AND GROUNDWATER MICROBIAL COMMUNITY SAMPLES.....25

FIGURE 7: RIVER INFILTRATION TAXA COMPARISON. THE BOX PLOTS SHOW THE TOTAL AVERAGE ABUNDANT FOX RIVER SEQUENCES FOR EACH SUB-CATEGORY FOR EACH LOCATION. THE FOX RIVER DATA IS ON THE LEFT AND THE WELLS ARE ON THE RIGHT WITH EACH FILTER SIZE (0.1 AND 0.2 μM INDICATED BY 1 AND 2 RESPECTIVELY) AND NUCLEIC ACID FRACTION (RNA AND DNA INDICATED BY R AND D). THE BOX SHOWS THE 25% TO 75% PERCENTILES (THE MIDDLE 50% CALLED THE INTER-QUARTILE RANGE OR IQR) OF THE AVERAGE ABUNDANCES. THE HORIZONTAL LINE THROUGH THE IQR BOX INDICATES THE MEDIAN. THE WHISKERS SHOW THE LOWEST AND HIGHEST VALUES NO FURTHER THAN 1.5X IQR AWAY FROM THE IQR, AND ALL OTHER POINTS

ABOVE/BELOW THE IQR ARE CONSIDERED OUTLIERS. NOTE THE Y-SCALES INDICATING SUM ABUNDANCE ARE NOT THE SAME, SO THE FOX RIVER HAS A HIGHER ABUNDANCE OF THESE INDICATOR SEQUENCES. 27

FIGURE 8: GROUNDWATER MICROBIAL COMMUNITY DENDROGRAM. NMDS AND BRAY-CURTIS DISSIMILARITY WERE USED TO GENERATE A DENDROGRAM DEMONSTRATING THE DIFFERENCES ACROSS GROUNDWATER MICROBIAL COMMUNITY SAMPLES. THE GROUNDWATER MICROBIAL COMMUNITIES CLUSTER FIRST BY WELL LOCATION IN THAT W12 IS SIGINIFICANTLY DIFFERENT FROM THE OTHER TWO WELLS. FILTER SIZE FRACTION THEN CLUSTER TOGETHER, THEN W13 AND W11 CLUSTER SEPARATELY, AND THEN RNA AND DNA CLUSTER TOGETHER 29

FIGURE 9: ORDINATION OF DNA (LEFT) AND RNA (RIGHT) GROUNDWATER MICROBIAL COMMUNITIES. NON-METRIC MULTIDIMENSIONAL SCALING (NMDS) BASED ON BRAY-CURTIS DISTANCES. COMMUNITIES FROM 0.1 μM SAMPLES ARE SHOWN IN TRIANGLES AND COMMUNITIES FROM 0.2 μM SAMPLES ARE SHOWING IN CIRCLES. WELL 11 IS SHOWN IN GREEN, W12 IS SHOWN IN RED, AND W13 IS SHOWN IN BLUE. SIGNIFICANT ENVIRONMENTAL PARAMETERS (TOTAL NITROGEN MG/L AND TOTAL DISSOLVED PHOSPHORUS μG/L) WERE PLOTTED TO SHOW THEIR CORRELATION IN THE DATA. CORRELATIONS BETWEEN THE ENVIRONMENTAL PARAMETERS AND NMDS ARE SHOWN IN LENGTH AND DIRECTION OF THE ARROWS. 30

FIGURE 11: MOST ABUNDANT ASVS ACROSS ALL WELLS. TAXON AFFILIATIONS ARE LISTED ON THE LEFT. DATA IS PARTITIONED BASED ON DNA (LEFT HALF OF THE HEATMAP) AND RNA (RIGHT HALF OF THE HEATMAP), FILTER SIZE WITH NUCLEIC ACID TYPE (0.1 μM LEFT, 0.2 μM RIGHT WITHIN DNA AND RNA), AND WITHIN THAT BY SITE (W11, W12, AND W13). THE LIGHT YELLOW INDICATES AN AVERAGE ABUNDANCE OF 0 MEANING THE ASV WAS NOT PRESENT FOR THAT TYPE OF SAMPLE. DARK BLUE INDICATES THE HIGHEST RELATIVE ABUNDANCE. THE FOURTH ROOT WAS USED TO DISPLAY THE RELATIVE ABUNDANCES TO SHOW THE PATTERNS MORE DISTINCTLY. SOME OF THE MOST ABUNDANT GROUNDWATER ASVS ARE SPECIFIC TO LOCATION SUGGESTING THEY ARE UNIQUE OR SPECIALIZED TO THAT LOCATION. SOME ASVS THAT ARE UNCLASSIFIED, I.E. 128, 18, 56, DO NOT APPEAR IN THE DNA BUT ARE OF THE MOST ABUNDANT RNA ASVS WHICH SUPPORTS THE THEORY OF INTRON SPLICING OCCURRING IN THE 16S RRNA GENE 39

FIGURE 12:GROUNDWATER DNA AND RNA MULTINOMIAL SPECIES CLASSIFICATION METHOD (CLAM) FOR W13 AND W12. THE BLUE TRIANGLES INDICATE RARE SEQUENCES THAT WERE TOO RARE TO CONSIDER IN THE CLASSIFICATION ANALYSIS, THE BLACK CIRCLES INDICATE GENERALISTS SPECIES/SEQUENCES THAT OCCUR BETWEEN BOTH SAMPLES, THE RED SQUARES INDICATE SPECIALISTS FOR THE W12 ENVIRONMENT, AND THE

green diamonds indicate specialist species for W13. The numbers of specialists for each well from each sub-dataset (DNA 0.2um, DNA 0.1um, RNA 0.2 μM and RNA 0.1um) were determined to be W13: (top left to right starting with DNA 0.2) 87, 82, 74, 78, and W12: 132, 128, 119, 112.....43

Figure 13:Top Average RNA:DNA Ratio ASVs from each Well. Violin plots comparing the top 5 average RNA:DNA ratio ASV's from each well resulting in 13 ASVs. Violin plots, similar to box plots, show the distribution of data values, with the mode average as the thickest part of the plot. Centered data points were plotted to show sample distribution. W11 is show in green, W12 in red, and W13 in blue. Each plot displays an individual ASV and the ratios for each well and each has its own scaled ratio y-axis. 120

Figure 14: CLAM Ratio Specialists for W12 and W13.....47

Figure 15: Selection of Most Abundant ASVs Across All Wells. Taxon affiliations are listed on the left. Data is partitioned based on DNA (left half of the heatmap) and RNA (right half of the heatmap), filter size with nucleic acid type (0.1 μM left, 0.2 μM right within DNA and RNA), and within that by site (W11, W12, and W13). The light yellow indicates an average abundance of 0 meaning the ASV was not present for that type of sample. Dark blue indicates the highest relative abundance. The fourth root was used to display the relative abundances to show the patterns more distinctly. Some of the most abundant groundwater ASVs are specific to location suggesting (ASVs: 163, 19, 35) they are unique or specialized to that location. Other ASVs show higher abundances comparatively across all wells (ASV1: Unclassified, ASV 2: Sulfurifustis) suggesting the aquifer shares specific microbes between the wells. Some ASVs that are Unclassified, (ASVs 128, 18, 56) do not appear in the DNA but are of the most abundant RNA ASVs which supports the theory of intron splicing occurring in the 16S rRNA gene. 116

LIST OF TABLES

TABLE 1: PCR CONDITIONS. THE NORMAL PCR CONDITIONS ARE LISTED AS WELL AS THE RECONDITIONED PCR CONDITIONS WHEN ONE NORMAL PCR WAS NOT SUFFICIENT FOR AMPLIFICATION....17

TABLE 2: THERMOCYLCER PCR CONDITIONS FOR V4 16S RRNA GENE AMPLIFICATION 17

TABLE 3: COMPOSITE WATER QUALITY DATA FOR THE THREE SHALLOW GROUNDWATER SITES. 21

TABLE 4: BALANCED BIOGEOCHEMICAL REACTIONS....22

TABLE 5: CHEMICAL DATA SIGNIFICANCE IN GROUNDWATER MICROBIAL COMMUNITY DISTRIBUTION....32

TABLE 6: TOP 50 COMBINED GROUNDWATER TAXA AFFILIATIONS. KNOWN TAXA AFFILIATIONS FROM ASV TAXON CLASSIFICATION WERE CONDENSED TO DISPLAY THE PERCENTAGE OF THAT TAXON IN THE DATASET. 33

TABLE 7 THE SPECIALIST TAXA DETERMINED FOR W13 OCCURRING IN ALL SUBCATEGORIES OF DNA, RNA, 0.1 AND 0.2 μM FROM CLAM 44

TABLE 8: THE SPECIALIST TAXA DETERMINED FOR W12 OCCURRING IN ALL SUBCATEGORIES OF DNA, RNA, 0.1 AND 0.2 μM FROM CLAM 45

TABLE 9: W13 AND W12 SPECIALISTS FROM CLAM ANALYSIS ON RNA:DNA RATIOS 47

TABLE 10: AMMONIUM AVERAGE RESULTS FROM NUTRIENT EXPERIMENTS 59

TABLE 11: NITRATE AVERAGE RESULTS FROM NUTRIENT EXPERIMENTS 60

TABLE 12: TOTAL DISSOLVED PHOSPHORUS AVERAGE RESULTS FROM NUTRIENT EXPERIMENT. THE DETECTION LIMIT WAS 0.5 μG/L. 60

TABLE 13: SOLUBLE REACTIVE PHOSPHORUS AVERAGE RESULTS FROM NUTRIENT EXPERIMENTS. THE DETECTION LIMIT WAS 0.5 μG/L. 60

TABLE 14: N:P AVERAGE RATIO RESULTS FROM NUTRIENT EXPERIMENTS. THE INITIAL MEASUREMENTS WERE BASED OFF OF TDP μMOL/L FROM THE BOOTSMA LAB'S MEASUREMENTS AND TN μMOL/L FROM THE TOC-TN L MACHINE IN THE ANALYTICS LAB. THE FINAL RATIOS WERE DERIVED FROM TDP μMOL/L. 60

TABLE 15: COORDINATES IN DECIMAL DEGREE FOR DESCRIBE SAMPLE SITES 78

TABLE 16: RAW DATA NITRATE RESULTS FROM NUTRIENT EXPERIMENTS....88

TABLE 17: RAW DATA TOTAL DISSOLVED PHOSPHORUS RESULTS FROM NUTRIENT EXPERIMENTS. 89

TABLE 18: RAW DATA AMMONIUM RESULTS FROM NUTRIENT EXPERIMENTS. ALL FINAL VALUES WERE BELOW DETECTION LIMIT. 90
TABLE 19: RAW DATA SOLUBLE REACTIVE PHOSPHORUS RESULTS. DETECTION LIMIT FOR PHOSPHORUS WAS 0.5 µG/L. 91
TABLE 20: NITROGEN TO PHOSPHORUS RATIOS RAW DATA FOR NUTRIENT EXPERIMENTS. THE RATIOS ARE BASED ON MOLAR RATIOS. THE INITIAL N:P RATIOS WERE BASED OFF OF TDP (BOOTSMA LAB) AND TN (TOC-TN L MACHINE) VALUES. THE FINAL N:P RATIOS ARE FROM FINAL TDP (BOOTSMA LAB) AND FINAL NITRATE-N (IC). THE TOC-TN L MACHINE WAS OUT OF ORDER AT THE END OF THE EXPERIMENTS. 92
TABLE 21:RAW DATA CELL COUNTS FROM NUTRIENT ADDITION EXPERIMENTS. VALUES OF NA MEAN THERE WERE NOT CELL COUNTS AVAILABLE FOR THOSE SAMPLES. NOTE NL 640 AND NL 634 WERE NOT INCLUDED IN THE FINAL AVERAGE RESULTS DUE TO SAMPLE DISCREPANCIES. 93
TABLE 22: UNIQUE W12 SPECIALISTS FROM DNA AND RNA CLAM RESULTS 94
TABLE 23: UNIQUE W13 SPECIALISTS FROM DNA AND RNA CLAM RESULTS 104
TABLE 24: TOP AVERAGE RNA:DNA RATIOS FROM EACH WELL 110
TABLE 25: FOX RIVER INFILTRATION INDICATOR ASV TAXA 114

LIST OF ABBREVIATIONS

ASVs	Amplicon sequence variants
bp	base pair
CAFOs	Concentrated Animal Feeding Operations
CO_2	Carbon dioxide
CPR	Candidate Phyla Radiation
DO	Dissolved oxygen
DNA	Deoxyribonucleic acid
DPANN phylum	Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, Nanohaloarchaea
EPA	Environmental Protection Agency
N	Nitrogen
NMDS	Non-metric Multidimensional Scaling
P	Phosphorus
qPCR	Quantitative Polymerase Chain Reaction
RBI well	River Bank Inducement well
RNA	Ribonucleic acid
RPM	Rotations per minute
rRNA	Ribosomal ribonucleic acid
SRP	Soluble reactive phosphorus
TOC	Total organic carbon
TDP	Total dissolved phosphorus
WDNR	Wisconsin Department of Natural Resources
WPDES	Wisconsin Pollutant Discharge Elimination Permit
WWTP	Wastewater Treatment Plant

ACKNOWLEDGEMENTS

I would like to extend a tremendous amount of gratitude towards my advisor Dr. Ryan J. Newton, whose guidance, encouragement, and expertise guided me through my journey and made the outcome of this project possible. Thank you to my committee members Dr. Tim Grundl and Dr. Matthew Smith. I'd also like to specifically acknowledge Dr. Matthew Smith for constructing the filtration system for sampling in this project. I'd like to give a great thanks to the Great Lakes Genomics Center for library preparation and illumina MiSeq sequencing. A huge appreciation goes out to Rae-Ann MacLellan-Hurd for help with nutrient measurements and to Wilson Tarpey for help with countless hours of cell counts. The support from Lou LaMartina, Natalie Rumble and other members of the Newton and McLellan Labs helped me through this experience with words of encouragement, sympathetic grumbling, bowling, laser tag, and many laughs.

CHAPTER 1 – INTRODUCTION

Groundwater aquifers are an important agricultural, industrial and domestic drinking water source, and in the U.S., they account for 25% of all freshwater used. The deep aquifer in southeastern Wisconsin has supplied residents, like those in Waukesha, with water for the last 100 years, but it is contaminated with naturally occurring radium. Over pumping has depleted the deep aquifer and impacted radium concentrations, sometimes reaching to more than three times EPA radium limits. Waukesha currently treats this water by partial removal of radium and blending deep aquifer water with radium free groundwater from a shallow aquifer. However, given their close connection to the surface, shallow groundwater aquifers are often altered by anthropogenic activities in the watershed and are also susceptible to depletion from over pumping. To counter depletion, Waukesha placed two shallow groundwater wells near the Fox River to induce flow from the river into the wells, hence they are called river bank inducement (RBI) wells. However, the Fox River receives wastewater treatment plant (WWTP) effluent upstream of the wells, and these RBI wells now receive river water containing WWTP effluent with associated ions, nutrients, and chemical constituents. Since microorganisms catalyze chemical transformations and largely influence groundwater chemistry, it is important to understand how the additional input from the Fox River affects these shallow groundwater aquifer microbial communities and thus water chemistry.

The purpose of this study was to determine if the microbial community of a shallow groundwater aquifer relates to altered water composition in river-impacted RBI wells as compared to a well that does not induce river infiltration. This study aims to investigate the first piece of a larger question of identifying athropogenically driven changes to the metagenome of a shallow groundwater aquifer and its effects on aquifer reactivity. Determining the metagenome

includes both 1) identifying the microbial community composition by targeting the specific 16S ribosomal ribonucleic acid (rRNA) gene and 2) recovering and analyzing all genetic material to determine the genetic functional potential in the groundwater environments. This study fully addresses part (1) of this larger question by characterizing changes in the microbial community composition and organism affiliated functional potential related to altered nutrient and ion compositions between pristine and river-impacted RBI wells. Part (2) was not analyzed for this study due to developing and new practices related to groundwater studies. Methods for processing samples (i.e. sample collection, genomic DNA (gDNA) extraction, gDNA concentration, and gDNA purification, etc.) for metagenomic analysis were still under development and optimization due to low microbial biomass recovery from relatively high volumes of groundwater. However, the metagenomic genetic material was recovered and extracted for future processing, applications, and analysis. Newly established techniques were applied, and microbial 16S amplicon sequence data was paired with geochemical data to address the aims of this study. Furthermore, the influence of increased nitrate with varying phosphate concentrations was examined in laboratory experiments to further characterize the microbial response to altered nutrient conditions in previously unaffected groundwater. It is most probable that nitrogen in the form of nitrate would be the primary macronutrient entering the groundwater well sites due to the geology and chemistry of the river and RBI wells (Grundl, personal communication 2018). The information acquired in this study expands the existing knowledge of groundwater microorganisms, biogeochemical processes, and may help inform decisions for water treatment processing.

1.2 – Background and Significance

1.2.1 Groundwater Use in the United States

Groundwater aquifers, a significant water source in the United States, supplied approximately 43.6 million Americans with self-supplied domestic freshwater in 2010. Groundwater withdrawals made up 22% of all US water withdrawals in 2010, and of that, 96% came from freshwater aquifers (USGS, 2017). Out of all water withdrawals in the US in 2010, groundwater made up:

- 19% of industrial (2,950 million gallons/day)
- 37% of public supply (15,700 million gallons/day)
- 47% of all irrigation (49,500 million gallons/day)
- 60% of all livestock (1,200 million gallons/day)
- and 73% of mining (3,900 million gallons/day) (Maupin & Kenny, 2010).

1.2.2 Groundwater in Southeast Wisconsin

In Wisconsin, almost two thirds of people get their drinking water from groundwater (DNR, 2017), but groundwater sources in Wisconsin risk the potential of anthropogenic and natural contamination, as well as depletion from over-use. The radioactive element radium is found naturally in the bedrock of eastern Wisconsin and contaminates water in Wisconsin's deep unconfined aquifer. There are two classifications of groundwater aquifers—confined and unconfined. Confined aquifers, deeper than unconfined, lay between two layers of impermeable rocks (Smith et al., 2012). Unconfined aquifers, the focus of this study, are closer to the surface than confined aquifers with permeable sediments allowing infiltration and seepage from the surface and can be easily impacted by anthropogenic contamination (Smith et al., 2012).

Radium consumption leads to accumulation in bones and tissues resulting in cancer, teeth deterioration, and anemia. Yet, water demands continue to rise with increasing populations which results in deeper drilling of wells and inadequate regulation on over pumping (McCoy, 2016). This practice has depleted high capacity groundwater well supplies and impacted radium concentrations up to three times EPA limits in some cases (CH2M HILL, 2013; Grundl and Cape, 2006; McCoy, 2016).

Communities like Waukesha, WI, the study area of interest, have battled radium contamination in their water for decades and need alternatives for safe and clean drinking water especially as the population expands. Waukesha treats radium contamination by partial removal, which can be costly, and by blending this deep contaminated water with shallow radium free groundwater (CH2M HILL, 2013). However, shallow groundwater in unconfined aquifers risks external contamination (Smith et al., 2012) and depletion from over-pumping.

1.2.3 River Bank Inducement Wells in Waukesha, WI

Placing groundwater wells next to riverbanks has been a common practice since the 1870's in Europe and for the last 60 years in the United States for industrial and public water supplies (Ray et al., 2002; Shamrukh and Abdel-Wahab, 2008). This practice allows for riverbank filtration into pumping wells by inducing water flow from the river (Ray, 2008; Ray et al., 2002). This method has been proposed as a promising and sustainable technology for municipal and public water production across the globe (Ray, 2008; Shamrukh and Abdel-Wahab, 2008). It has largely been assumed that this practice will filter out pollutants and contaminants for municipal and drinking water wells. However, it has been shown that river filtration does not always produce well water that conforms to drinking water standards (Ray et al., 2002; Singh et al., 2010). Also, effluent-dominated streams that have strong hydrologic

connections between surface and shallow-groundwater can transport contaminants into shallow groundwater, and these systems require further investigation to determine risks associated to contaminant infiltration (Bradley et al., 2014; Ray, 2008; Weiss et al., 2005).

Shallow wells have been drilled in Waukesha near the Fox River to induce flow from the river. This practice was meant to increase water yields from the wells and reduce depletion by quickly recharging the wells with river water as shown in (Fields-Sommers, 2015). Typically, groundwater flows into surface water bodies, but if the water table is lowered by pumping, flow is reversed and surface water is drawn into groundwater. These wells, termed River Bank Inducement (RBI) wells, are susceptible to external pollutants entering from surface water. Two shallow groundwater wells placed next to the Fox River in Waukesha now receive river water input containing upstream wastewater treatment plant (WWTP) effluent (shown in Fig. 1). It is unclear how this WWTP effluent and river water infiltration will impact microbial communities and thus water quality in these municipal wells.

1.2.4 Microorganisms in Groundwater

Microbes carry out the majority of chemical reactions in nature and largely drive the geochemical reactions in subsurface systems (Kirchman, 2012). Microorganisms greatly influence elemental and nutrient cycling in groundwater by turning over matter and energy because of the low oxygen levels, lack of sunlight, and lack of external energy sources (Smith et al., 2012, Kirchman, 2012). These organisms and transformations can impact water quality in surface water, groundwater and even concentrations of trace gases in the atmosphere (Long et al., 2016).

In recent years, the use of genome sequencing technologies has discovered previously unknown microorganisms from subsurface and groundwater environments has greatly expanded

the tree of life (Hug, Baker, et al., 2016). However, biogeochemical models lack representation from the subsurface and relatively little is known about groundwater microorganisms (Hug et al., 2016; Long et al., 2016). Many of these novel and recently discovered microbes show limited metabolic capabilities suggesting many groundwater organisms perform "metabolic handoffs" where most organisms do not singly contain machinery necessary to carry out multiple sequential redox transformations. These microbes have ultra-small cell and genome sizes (Anantharaman et al., 2016; Castelle et al., 2015; Long et al., 2016). It appears these microbes drive biogeochemical cycling through symbiotic relationships. Different microbes can be selected to implement certain redox pathways as environmental conditions alter, and thus, steps in major biogeochemical cycles can also be altered (Anantharaman et al., 2016). Since resource input is limited and the apparent syntrophic nature of these organisms, any change can impact or shift microbial composition and metabolism (Hemme et al., 2015) and thus affect water quality. It's unclear if and how the biogeochemical cycles and microbial groups will be altered in these shallow aquifer drinking water wells from river water infiltration containing WWTP effluent.

The overall objective of this study is to determine if the microbial community composition, organism affiliated functional potential, and aquifer reactivity of previously pristine groundwater is altered by river infiltration containing WWTP effluent, and then to characterize any change related to river water infiltration. Newly established metagenomic techniques were utilized, and microbial data was paired with geochemical data to address these questions to expand current knowledge related to groundwater systems. It is necessary to understand anthropogenic impact on aquifer systems to secure sustainable and high quality drinking water resources. It is important to understand the effects of river water infiltration into RBI wells on the native microbial communities and implications for water quality especially since RBI wells

have been proposed as promising and sustainable water supplies worldwide. This study can serve as a model system for RBI impacts on water quality in groundwater wells.

1.3 Study System

There are three groundwater wells in a shallow sand and gravel aquifer located near the Fox River in Waukesha, Wisconsin. Up to 40% of the Fox River's flow comes from upstream wastewater treatment plant (WWTP) effluent (Fields-Sommers, 2015). Two RBI wells are 225 feet and 83 feet from the Fox River and the third well is approximately 1,500 feet from the river. The wells are screened between 60-150 feet (18-45 meters) below surface. The Wisconsin Department of Natural Resources (WDNR) unique well numbers are RL255 (W11) and RL256 (W12) for the two RBI wells, and WK947 (W13) for the background well. Throughout the rest of the text, the wells will be referred to by their Waukesha numbers—W11, W12, W13. Previous research indicated the two RBI (W12 and W11) wells are infiltrated by river water while the third well (W13), is not impacted by river water. RBI well W12 actively pumps and thus actively draws in river water and is currently impacted. RBI well W11 has

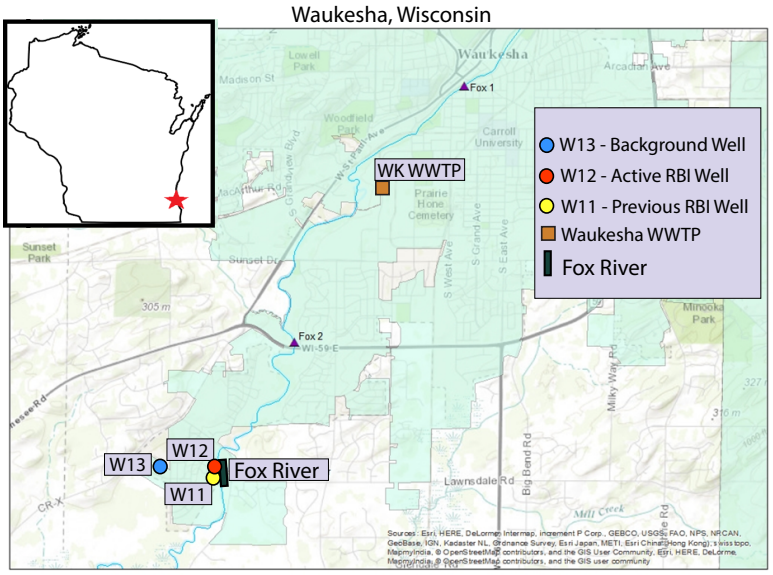


Figure 1: Geographic locations of the study site in Waukesha, WI along the Fox River. Wastewater treatment plant (WWTP) effluent flows into the Fox River upstream of the wells. W12 actively draws in river water as an RBI well. W11 previously was an RBI well but no longer actively pumps at rate which draws in river water. Background well W13 is considered pristine.

recently reduced its pumping, potentially drawing in less Fox River water, but was previously impacted when it was actively pumping. The background well W13 actively pumps, but is considered pristine compared to the other two sites with no hydrologic connection to the river (Fields-Sommers, 2015). The layout and geography of the well sites are shown in Fig. 1.

1.4 Previous Research

The goal of previous research was to determine existing and potential influences of river bank inducement, recharge mechanisms of the well field, and to discriminate the sources of sodium and chloride entering the well field. Previous research indicates that the two river bank inducement (RBI) wells W11 and W12 pump up to 40-60% Fox River water (Fields-Sommers, 2015). Artificial sweeteners, which are highly concentrated in WWTP effluent, were determined to be the most mobile of emerging contaminants and found to be a reliable tracer and indicator of river water infiltration in the RBI wells W11 and W12. Specifically, sucralose had substantially higher concentrations in W12 and W11 than pristine W13 with even higher concentrations in the Fox River and the highest concentrations in WWTP samples, demonstrating river infiltration in the two RBI wells (Fields-Sommers, 2015). These values are displayed Fig. 2.

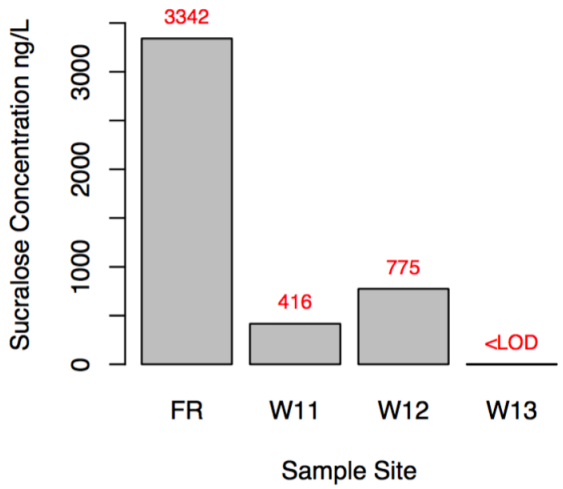


Figure 2: Sucralose Concentrations (ng/L) in the Fox River, W11, W12, and background well W13 (left to right). Single measurements from1 liter single samples from the spring of 2015 were performed at UW-Steven's Point Water and Environmental Analysis Lab (WEAL). The detection limit was 25 ng/L and methods used are described previously (McGinley et al., 2015). Adapted from Fields-Sommers (2015).

Also, increasing concentrations of sodium and chloride in W11 and W12 over time indicated infiltration from the Fox River, which is sodium and chloride rich from WWTP effluent. Chloride and sodium levels in the pristine well remained constant and lower over time than in the RBI wells W11 and W12 which continued to rise as pumpage decreased over time. Figure 3 displays concentrations of ions (left axis) and pumpage trends (right axis) over time in years for each well.

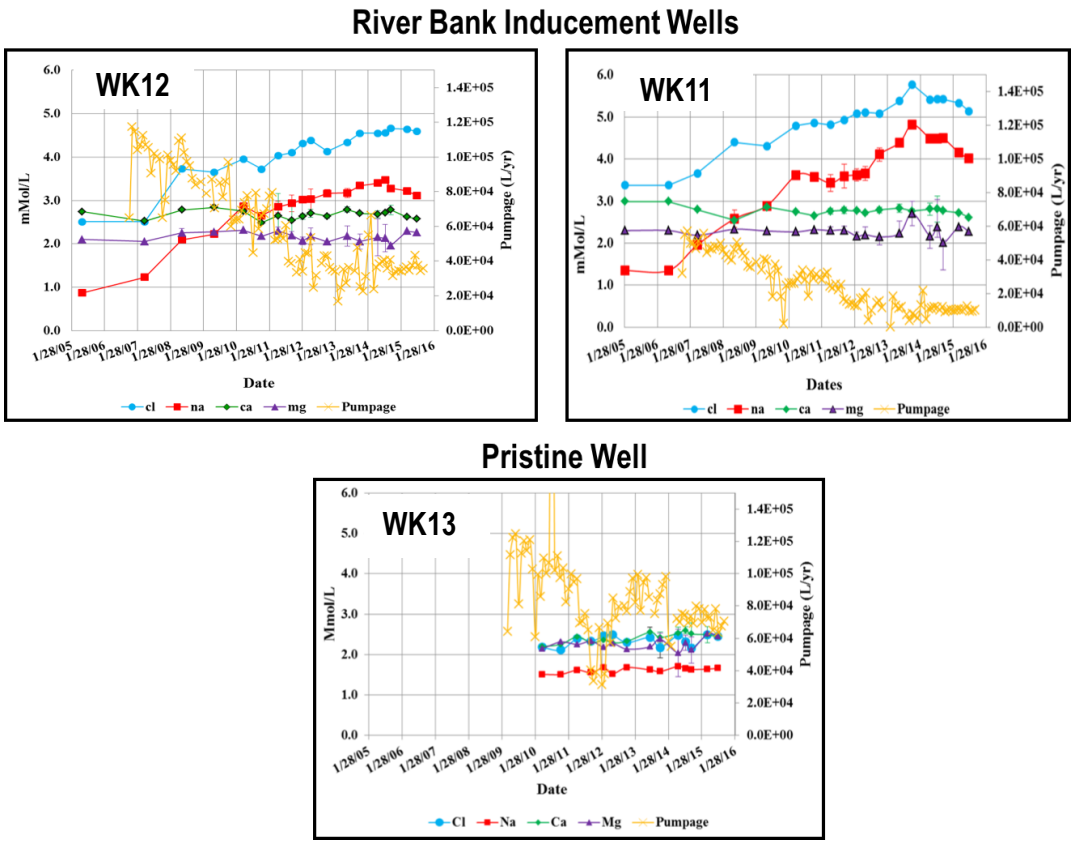


Figure 3 Major Ion Concentrations in the Wells. Ion concentrations over time with pumpage in RBI wells (top left WK12, right WK11) and pristine well (bottom WK13).

Bacterial collection and preliminary analyses were performed by the McLellan Lab at the School of Freshwater Sciences at UW-Milwaukee and Laura Fields-Sommers (2015). 16S rRNA gene sequencing indicated distinct microbial community differences between the WWTP, Fox River, and RBI well W12. This can be seen in Figure 4 which is a heatmap of the most abundant taxa from a WWTP sample, Fox River sample, and RBI well W12 sample. The dark purple shows a relative abundance of 0 and shows that the RBI well microbial community is distinct from the WWTP and river sample. Also, no fecal tracer bacteria—*Bacteroidales, E. coli, Enterococci, Lachnospiraceae, and Ruminant*—were found in the RBI well after qPCR analysis (Fields-Sommers, 2015). This preliminary analysis suggested that microorganisms are not moving through the soil matrix from the Fox River into the shallow groundwater. Only river water and its mobile chemical constituents would be entering the RBI wells affecting microbial communities.

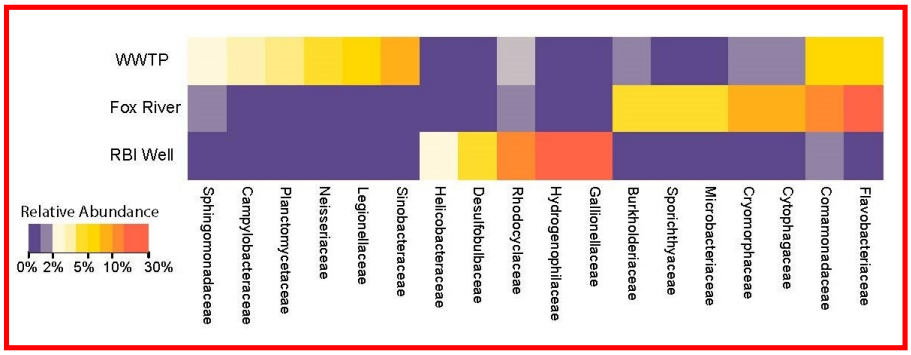


Figure 4 Heatmap illustrating the relative abundance of bacterial families (only families present at >2% of the community composition in at least one sample depicted). Preliminary 16S rRNA gene community composition relative abundances of one sample each reveal the dominant families between WWTP, Fox River, and RBI well W12 vary across sites.

CHAPTER 2 – RIVER INFILTRATION EFFECTS ON THE SHALLOW GROUNDWATER AQUIFER MICROBIAL COMMUNITY

Specific Aim 1: Identify and characterize microbial communities and geochemical reactivity in pristine and river-infiltrated portions of a shallow sand and gravel groundwater aquifer to determine differences related to infiltration.

1a) Collect and sequence microbial RNA and DNA and analyze the V4 hypervariable region of the 16S ribosomal RNA (rRNA) gene to identify and compare microbial community compositions at non-river impacted groundwater, river-infiltrated groundwater, and river sites.

1b) Determine and compare 16S ribosomal RNA:DNA ratios to infer protein synthesis potential for specific community members and between samples and sites.

1c) Collect and analyze geochemical data in each well to identify and determine thermodynamically favorable reactions from free energy yield calculations.

Specific Aim 2: Characterize microbial community response to altered nutrient additions with varying Nitrogen:Phosphorus ratios.

2a) Based on the geochemical compostion of the wells from Specific Aim 1c, perform bottle experiments with varying N:P ratios to determine the microbial response to increased nitrate concentrations.

2.1 Background

Microbial community composition data typically indicates which microorganisms are present and in what relative abundances. This data essentially describes “who” is there, “who”

may be impacting ecosystem-level element and nutrient cycles the system, and how environmental conditions may impact community structure. Anoxic environments are unique and contain highly specialized microbial communities (Vigneron et al., 2018). Some organisms, like certain orders of *Proteobacteria* such as *Burkholderiales*, have been associated with nutrient poor conditions in post-WWTP sediments (Atashgahi et al., 2015). However, if nutrient conditions become altered, the community composition and function of the ecosystem could be altered could change. Studies using multivariate analyses to study the relationship between environmental parameters and microbial communities in a hyporheic zone show that microbial community composition correlates to nutrient loads and/or oxygen concentrations (i.e. organic carbon, nitrogen, DOC, TOC, TN, and Oxygen) (Atashgahi et al., 2015).

Subsurface microorganisms utilize, as well as generate, biogeochemical gradients. Through genome resolution, specific microbial community members have been identified and associated with specific transformations in nutrient and biogeochemical cycles like in carbon, nitrogen, and sulfur cycles in the terrestrial subsurface (Brown et al., 2015; Long et al., 2016). Environmental and nutrient alterations from the river to these low oxygen, low nutrient groundwater ecosystems in any capacity may significantly alter the microbial community composition, biodiversity, and resulting ecosystem-relevant functions carried out by a new microbial community state.

Community composition can be determined using microbial 16S ribosomal RNA (rRNA) gene sequencing (Pace et al., 1986). Ribosomal RNA mediates protein synthesis as part of the ribosome. rRNA is found in all known living organisms and has remained highly conserved throughout evolutionary history. In microorganisms, there are variable regions specific to different taxa within the conserved regions of the rRNA gene, making it a good molecular

marker to target and identify phylogeny and taxonomy. The hypervariable V4 region was targeted to capture both archaeal and bacterial microorganisms (Parada et al., 2016; Walters et al., 2016). By combining community composition and geochemical data from the pristine and impacted well sites (Specific Aim 2b) we were capable of identifying parameters that impact microbial communities in these groundwater wells.

Recently, a large set of previously unknown microorganisms have been discovered from groundwater, and this discovery greatly expanded the tree of life (Hug, Baker, et al., 2016). Most of these previously unknown organisms were identified using 16S rRNA gene sequencing and shotgun metagenomic genome sequencing. These organisms, largely derived from anoxic subsurface samples, make up the archaeal DPANN superphylum and bacterial Candidate Phyla Radiation (CPR) in the new tree of life (Castelle et al., 2015; Eme & Doolittle, 2015; Hug, Baker, et al., 2016; Liu et al., 2018; Rinke et al., 2013). Candidate phyla that lack isolated representatives are expected to contribute to, and moderate nutrient cycling. It has been noted, the rRNA genes of many CPR organisms contain self-splicing introns and encode proteins. Since this is a rarely recorded bacterial characteristic, most of these organisms would not be detected through common cultivation-independent methods (Brown et al., 2015), such as standard methods of only targeting the 16S rRNA gene (i.e. the DNA encoding for the transcribed ribosomal RNA, in 16S rRNA gene sequencing for microbial community composition surveys). For this reason, this study targeted both the microbial ribosomal RNA (RNA) and the microbial ribosomal RNA gene (DNA) by simultaneously extracting RNA and DNA to generate 16S rRNA and 16S rRNA gene sequence data.

Given that many groundwater microorganisms have unusual 16S rRNA gene sequences, the approach of targeting both the DNA and RNA captures more microbial community members

than solely targeting the DNA alone. Also, 16S ribosomal RNA:DNA ratios can be used to estimate and compare protein synthesis potentials (PSP) of specific taxa across temporal, sampling, and location differences (Denef et al., 2016). In theory, comparing any change in 16S rRNA copies to rRNA gene copies (16S rRNA:DNA ratios) within a specific organism/taxa, indicates how a specific taxa's PSP changes under different conditions, since 16S rRNA is part of the ribosome, which mediates protein synthesis.

2.2 Methods

2.2.1 Sample Collection

Water samples were obtained through Waukesha Water municipal pump houses. Wells were pumped prior to arrival and each system was first flushed for approximately 5-10 minutes before sample collection in order to obtain groundwater that was not previously remaining in the pipes. Water was collected in autoclaved 1 and 2 liter Nalgene containers that were first rinsed with sample water and transferred to the filtration system. Samples were filtered on site and flash frozen in liquid nitrogen due to the fast degradation/alteration rate of RNA. Two peristaltic pumps were used to filter 2-3 L of groundwater in replicate from each of the three aquifer well sites. The sample volume varied depending on filtration time. The filtration time was capped after approximately 30 minutes to minimize RNA degradation and alteration. Samples were filtered sequentially through an in-line filtration system on polyether sulfone (PES) Millipore 47 mm diameter filters with uniform pore sizes of 3 μm, 0.2 μm, and 0.1 μm. Approximately 300 ml of Fox River water was filtered sequentially through 3 μm, 0.2 μm, and 0.1 μm filters as well. The 3 μm filter was used as a pre-filter to reduce clogging from particles to increase water flux

and microbial capture. The 3 µm pre-filters also removed particle attached microorganisms and the 0.2 µm and 0.1 µm filters were used for data analysis on the free-living microorganisms. 1-10 ml of 3 µm filtrate and 0.2 µm filtrate were collected and fixed with 4% or 21 % formaldehyde (final concentration 1-2% formaldehyde) for cell enumeration using DAPI fluorescent stain and microscopy (Porter and Feig, 1980). Filters were stored in sterile pre-labeled screw cap 2 ml tubes and placed in liquid nitrogen immediately after filtration in the field. Upon returning to the lab at the School of Freshwater Sciences, samples were placed and stored in the -80 °C freezer until processing for nucleic acid extraction.

2.2.2 Lab Methods

Simultaneous DNA and RNA extraction was performed with Qiagen's AllPrep Powerviral DNA/RNA kit with modified prerequisite and elution steps (located in Appendix B). Promega's RQ1 RNase-Free DNase (Cat #M6101) and GoScript™ Reverse Transcription System were used to treat RNA samples and reverse transcribe RNA to complementary DNA (cDNA). The reverse primer 806Rb for the v4 16S rRNA gene region was used in the cDNA synthesis (806Rb – GGACTACNVGGGTWTCTAAT) (Apprill et al., 2015). Polymerase Chain Reaction (PCR) was used to target and amplify the V4 16S rRNA gene region in the DNA and cDNA samples using 515Fb (Parada et al., 2016) and 806Rb primers with Invitrogen's™ Platinum™ Taq DNA Polymerase. Samples were run in triplicate PCR reactions and later pooled before sequencing. 5 µl of one PCR reaction out of the three triplicates for each sample was screened using gel electrophoresis to verify amplification and DNA fragment size. A modified reconditioned/nested PCR protocol was used when one normal PCR cycle (25 µl reaction volume, 1 µl template, 30 cycles) was not sufficient for sample amplification. In the

reconditioned PCR, two consecutive PCR's were carried out. The first PCR had a smaller reaction volume with a shorter cycle period but still 1 µl of template (15 µl reaction volume, 10 cycles, 1 µl template). Then, 1 µl of the reconditioned PCR was used as template in the full PCR (25 µl reaction volume, 30 cycles, 1 µl of template). A negative control was also run in all thermocycling runs. Reaction components, volumes, concentrations are described in the table below.

Table 1: PCR Conditions. The normal PCR conditions are listed as well as the reconditioned PCR conditions when one normal PCR was not sufficient for amplification

Master Mix of PCR Components	**Working Concentration**	**Normal PCR**	**Reconditioned PCR**
Reaction Volume	-	25	15
PCR Cycles	-	30	10
10x Buffer for Platinum Taq	10x	2.5 µL	1.5 µL
F Primer	5 uM stock	1 µL	0.6 µL
R Primer	5 uM stock	1 µL	0.6 µL
50 mM $MgSO_4$	50 uM	1 µL	0.6 µL
10 mM dNTP Mix	10mM	0.5 µL	0.3 µL
Platinum Taq Polymerase	5 U/µL	0.1 µL	0.06 µL

Table 2: Thermocylcer PCR Conditions for v4 16S rRNA Gene Amplification

PCR Thermocycler Conditions			
1	Initial denaturation	94°C	5 minutes
2	Denature	94°C	30 seconds
3	Anneal	50°C	45 seconds
4	Extend	72°C	60 seconds
Repeat 2 – 4 30x, or 10x for first reconditioned PCR step			
5	Final Extension	72°C	2 minutes
6	Hold	10°C	Hold

Agencourt AMPure XP magnetic bead kit protocol was used to clean and purify pooled triplicate PCR products before samples were library prepped. The protocol was followed per manufacturer's instructions with the exception that bead volume was reduced from 1.8 µl

AMPure XP beads per 1 µl of sample to 0.8 µl AMPure XP beads per 1 µl of sample (i.e. 56 µl of AMPure XP beads were used with 70 µl PCR samples). Samples were sequenced using the Illumina MiSeq 2 x 250 bp chemistry at the Great Lakes Genomics Center (GLGC). Two plates were run for sequencing. The first plate consisted solely of groundwater samples as well as a mock community (Appendix I) and a blank. The second plate contained all 22 of the river samples (11 RNA and 11 DNA) with 10 groundwater samples (5 RNA and 5 DNA). The HM-782D Microbial Mock Community B from BEI Resources containing genomic DNA from 20 known bacterial strains consisting of equimolar rRNA operon counts (Appendix I) was sequenced for quality control on both sequence plates. A blank, comprising of extraction and PCR blanks, was also sequenced for quality control (McCarthy et al., 2015).

Geochemical data was collected and analyzed by Madeline J. Salo from the Department of Geosciences, UW-Milwaukee. Parameters that change rapidly were measured in the field. These include DO, electrical conductivity, temperature, alkalinity, ferrous iron, and pH (Salo, 2018). Major ions were analyzed utilizing an ion chromatograph for anions and atomic absorption spectrometer for cations including: chloride, nitrate, phosphate, sulfate, calcium, sodium, magnesium, and potassium (Salo, 2018). Nutrients including nitrogen species (nitrate, nitrite and ammonium), total dissolved phosphorus, and H_2, CH_4, and CO_2 were analyzed as described in Salo (2018). Thermodynamically relevant reactions and free energy yield calculations were also performed by Madeline J. Salo. Geochemical data was paired as environmental conditions to predict and indicate trends and variance in the microbial community composition data.

2.2.3 Community Composition Sequence Data Processing

16S rRNA gene sequence data was processed in-house, with Mothur (Schloss et al., 2009), and DADA2 (Callahan et al., 2016). Low quality sequences, according to illumina standards, were filtered out. Illumina primers were removed utilizing cutadapt (Martin, 2011). DADA2 was used to merge reads, denoise sequence reads, and remove chimeras to create an amplicon sequence variant (ASV) table. Mothur was used to remove primers from merged reads that were binned incorrectly as Forward and Reverse, and these were added to the existing ASV count table. Mothur was used to remove sequences that were 5% shorter or longer than the median length of all sequences, which was 253 bp. Read lengths of 240-266 bps were kept. These reads were used to produce the final ASV count table. Sequence data was classified using SILVA v132. Taxonomy was added to sequences with the following parameters in Silva v132: minimum identity fight query sequence 0.9, reject sequences below identity 70%, Ref NR, SILVA taxonomy, serch-kmer-candidates 1000, ica-quorum 0.8, search-kmer-lne 10, search-kmer-mm 0. Once taxonomy was assigned, ASVs associated with the Mock community showing up in the samples were removed. The negative control was used to remove any ASVs in well samples with a lower mean count as compared to the negative control. Any taxa classified as mitochondria or chloroplast were removed, however there were none in this case. Since two sequence runs were performed, the second sequence run consisting of 10 groundwater samples and all of the Fox River samples was processed as described above. At the end of processing both sequence data sets, the ASV and count tables were merged into one matrix.

Data was analyzed and visualized using R (R Development Core Team, 2016). After performing sequence data processing and rigorous quality control using DADA2 with an error rate of 0%, the sequence dataset (RNA and DNA, 0.2um and 0.1um fractions, from W11, W12,

W13, and Fox River sites) included 51,331 unique amplicon sequence variants (ASVs), or taxa. Further processing in R removed sequences occurring at a relative abundance less than 0.01% in each sample (a relative abundance equal to 0.0001). The threshold of 0.01% was chosen to be stringent enough to remove cross contamination sequences (i.e. "sequence walking/migrating" in the Illumina flow cell during sequencing), but to also allow for rare community members to be included. Furthermore, the taxa that are in low abundances have little influence to overall community patterns.

Subsampling all data to the lowest sample can also be done to rarefy data for comparison (Weiss et al., 2017). However, known data is removed arbitrarily and this method can still cause noise in the dataset rather than eliminate it (Willis, 2017). Also, a specific count number could have been chosen as a threshold. This practice would allow the cutoff to have a different weight/significance per sample depending on the sequencing depth of each sample. This would cause differences across samples (i.e. a cutoff of a sequence count of 25 in one sample could be 25/25,000 or 0.1%, vs. 25/250,000 or 0.01% in another sample).

For this dataset, the highest total number of sequence counts for a given sample was 219,404 (0.1um W13 DNA 7/20/17) and the lowest total number of sequence counts for a sample was 27,232 (0.2 μm Fox River RNA 7/18/17). Given the lowest sample count in this dataset, the lowest cutoff could be 1/27,232 (0.003%). In order to be more stringent and to remove noise related to whether sample presence, the 0.01% threshold was chosen. This was also done under the assumption that contamination scales with sequencing depth. In this way, the dataset was cut down from 51,3331 to 21,910 unique amplicon sequence variants for analysis.

2.3 Analysis, Results & Discussion

2.3.1 Chemical & Thermodynamic Results

Chemical parameters and average measured values for the period Nov. 2016 through Jan. 2018 are shown in Table 3.

Table 3: Composite Water Quality Data for the Three Shallow Groundwater Sites.

Parameter	Units	WDNR Unique Well Number		
		RL255 W11	RL256 W12	WK947 W13
Temperature	°C	10.42 ± 1.28	10.61 ± 0.21	10.5 ± 0.14
pH		6.98 ± 0.13	6.99 ± 0.18	7.06 ± 0.62
Calcium	mg/L	93.12 ± 25.56	90.48 ± 19.63	83.71 ± 12.56
Chloride	mg/L	218.48 ± 57.18	201.28 ± 59.36	97.24 ± 33.32
Magnesium	mg/L	54.93 ± 1.69	53.32 ± 2.22	56.71 ± 3.97
Potassium	mg/L	3.3 ± 0.76	3.16 ± 0.45	2.56 ± 0.48
Sodium	mg/L	101.43 ± 4.8	81.1 ± 3.35	39.79 ± 1.46
Dissolved oxygen	mg/L	0.18 ± 0.31	0.15 ± 0.16	0.14 ± 0.17
Ferrous Iron	mg/L	0.1	0.1	0.1
Ammonium	mg/L	0.001 ± 0.002	0.07 ± 0.01	0.03 ± 0.01
Nitrate	mg/L	1.49 ± 1.07	0.3 ± 0.7	1.74 ± 1.15
Nitrite	mg/L	0.05 ± 0.02	0.003 ± 0.002	0.04 ± 0.01
Sulfate	mg/L	64.11 ± 19.56	68.17 ± 14.55	96.9 ± 11.06
Sulfide	mg/L	0.1	0.1	0.1
Total dissolved phosphorus	mg/L	0.002 ± 0.001	0.004 ± 0.002	0.003 ± 0.001
Dissolved organic carbon	mg/L	0.49 ± 0.32	0.93 ± 0.27	0.93 ± 0.37
Bicarbonate	mg/L	420.68 ± 164.44	462.25 ± 112.14	411.26 ± 104.72
Hydrogen	µmol/L	0.002 ± 0.0004	0.005 ± 0.003	0.004 ± 0.0004
Methane	µmol/L	0.007 ± 0.002	0.417 ± 0.19	0.043 ± 0.022

Free energy calculations were performed using 23 biogeochemical reactions to assess the potential metabolic pathways being carried out by the microbial consortia. The reactions include the groundwater constituents used in this study and are commonly driven by microorganisms in groundwater systems (Davidson et al., 2011; Lisle, 2014). Reactions are listed in Table 4.

Table 4: Balanced Biogeochemical Reactions

Reaction Number	Reaction
1	$CH_4 + SO_4^{2-} \rightarrow H_2O + HCO_3^- + HS^-$
2	$Acetate + NO_3^- + H_2O \rightarrow 2HCO_3^- + NH_3$
3	$4H_2 + 1.6NO_3^- + 1.6H^+ \rightarrow 0.8N_2 + 4.8H_2O$
4	$Acetate + 1.6NO_3^- + 0.6H^+ \rightarrow 2HCO_3^- + 0.8H_2O + 0.8N_2$
5	$4H_2 + NO_3^- + H^+ \rightarrow NH_3 + 3H_2O$
6	$Acetate + SO_4^{2-} \rightarrow 2HCO_3^- + HS^-$
7	$4H_2 + H^+ + SO_4^{2-} \rightarrow HS^- + 4H_2O$
8	$4Acetate + 4H_2O \rightarrow 4CH_4 + 4HCO_3^-$
9	$4H_2 + H^+ + HCO_3^- \rightarrow CH_4 + 3H_2O$
10	$4H_2 + H^+ + 2HCO_3^- \rightarrow Acetate + 4H_20$
11	$Acetate + 8Fe(OH)_3 + 15H^+ \rightarrow 8Fe_2^+ + 20H_2O + 2HCO_3^-$
12	$HS^- + 8Fe(OH)_3 + 15H^+ \rightarrow SO_4^{2-} + 8Fe^{2+} + 20H_2O$
13	$4H_2 + 2O_2 \rightarrow 4H_2O$
14	$Acetate + 2O_2 \rightarrow 2HCO_3^- + H^+$
15	$CH_4 + 2O_2 \rightarrow HCO_3^- + H^+ + H_2O$
16	$HS^- + 2O_2 \rightarrow SO_4^{2-} + H^+$
17	$(^4/_3)NH_3 + 2O_2 \rightarrow (^4/_3)NO_2^- + (^4/_3)H^+ + (^4/_3)H_2O$
18	$H_2S + 4NO_3^- \rightarrow SO_4^{2-} + 4NO_2^- + 2H^+$
19	$3H_2S + 4NO_2^- + 2H^+ + 4H_2O \rightarrow 3SO_4^{2-} + 4NH_4^+$
20	$(^4/_3)NH_4^+ + 2O_2 \rightarrow (^4/_3)NO_2^- + (^8/_3)H^+ + (^4/_3)H_2O$
21	$4NO_2^- + 2O_2 \rightarrow 4NO_3^-$
22	$8Fe^{2+} + 2O_2 + 20H_2O \rightarrow 8Fe(OH_3) + 16H^+$
23	$NO_3^- + 4H_2 + H^+ \rightarrow NH_3 + 3H_2O$

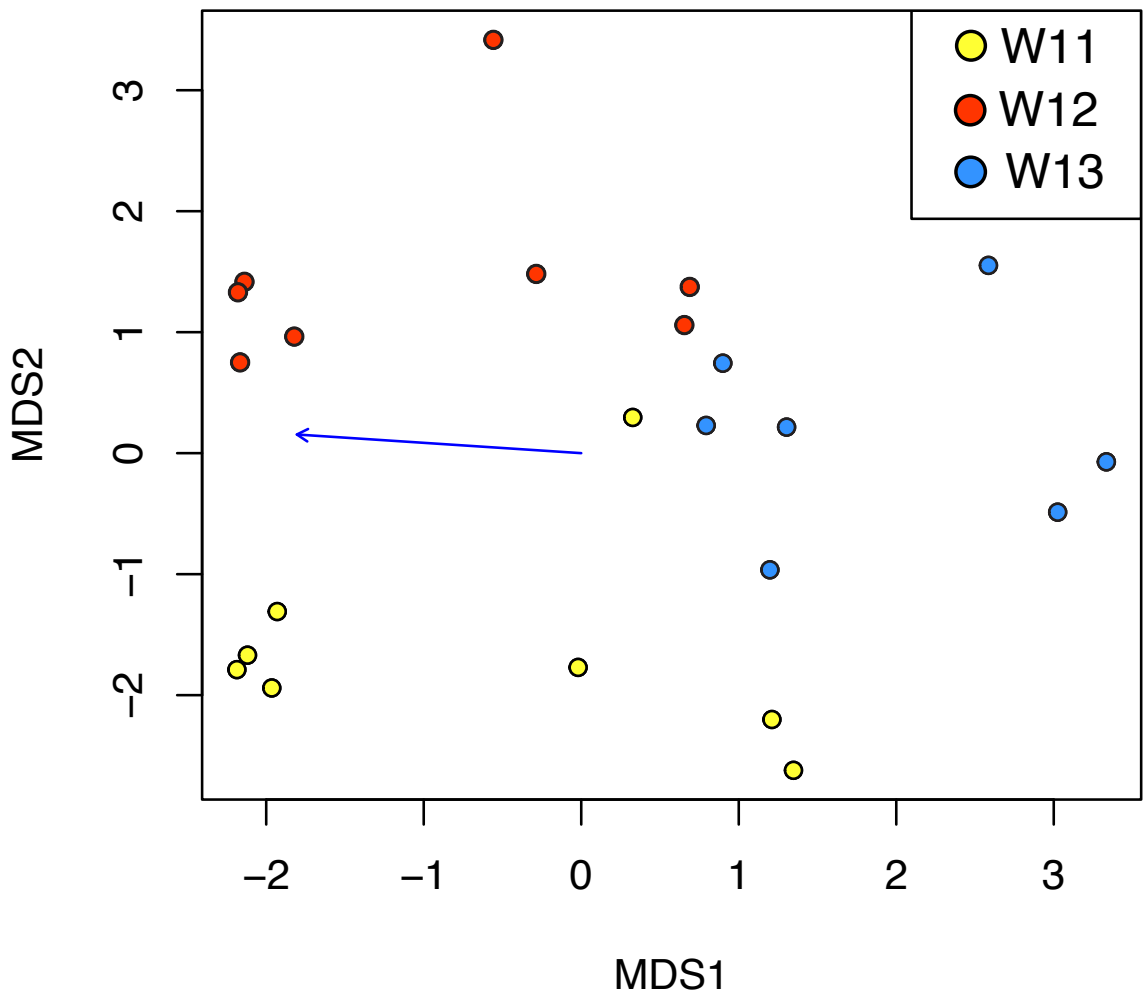


Figure 5: Nutrient and ion chemistry ordination from NMDS and Euclidean distance matrix from nutrient and ion data for the three wells. The length and angle of the arrow corresponds to date correlation. The groundwater samples displayed depict ion and nutrient chemistry data as a whole. Sample well sites are displayed with each color and were determined to be significantly different from one another based solely on chemistry data. The arrow indicates time had a correlation to the differences in the nutrient and ion composition of the groundwater samples.

To understand how the wells compared to one another based on chemistry as a whole (Fig. 5), a Euclidean distance matrix was developed for the nutrient and ion data (nitrate, nitrite, ammonia, TDP, calcium, sodium, magnesium, potassium, chloride, and sulfate [mg/L]) from all three groundwater sites (W11, W12, W13) that had corresponding samples from 2017-2018. The data was normalized by Z-score. Non-metric Multidimensional Scaling was used to develop an ordination of these chemistry samples using the function metaMDS() in the vegan package

(Oksanen et al., 2013) in R. Based on the nutrient and ion data, it was determined that samples cluster significantly by site (PERMANOVA $p = 0.001$) and by sample date (PERMANOVA $p = 0.004$), i.e. location and temporal effects were significant. Overall, it was shown that based solely on chemistry ion and nutrient data as a whole, the well sites differed significantly from each other (Fig. 5), even though the values do not appear to vary vastly between sites (Table 3). Also, a temporal component significantly correlated with the differences in well sites based solely on nutrient and ion chemistry data as a whole.

2.3.1 Fox River and Groundwater Microbial Community Compositions

The complete dataset of all groundwater and Fox River samples (W11, W12, W13, and FR) included 21,910 unique ASVs and was used to compare the microbial communities. A community distance matrix was developed using Bray-Curtis dissimilarity in the vegan package in R (Oksanen et al., 2013). Non-metric Multidimensional Scaling was used to develop an ordination of all microbial communities using the function metaMDS() in the vegan package. Essentially complex data with many dimensions is condensed down into 2D space so that it can be visualized and interpreted in a meaningful way. The Fox River and groundwater samples clearly cluster independently from each other as shown in Fig. 6. The microbial communities of the groundwater wells and the Fox River are distinct from one another and cluster by site significantly (PERMANOVA $p = 0.001$).

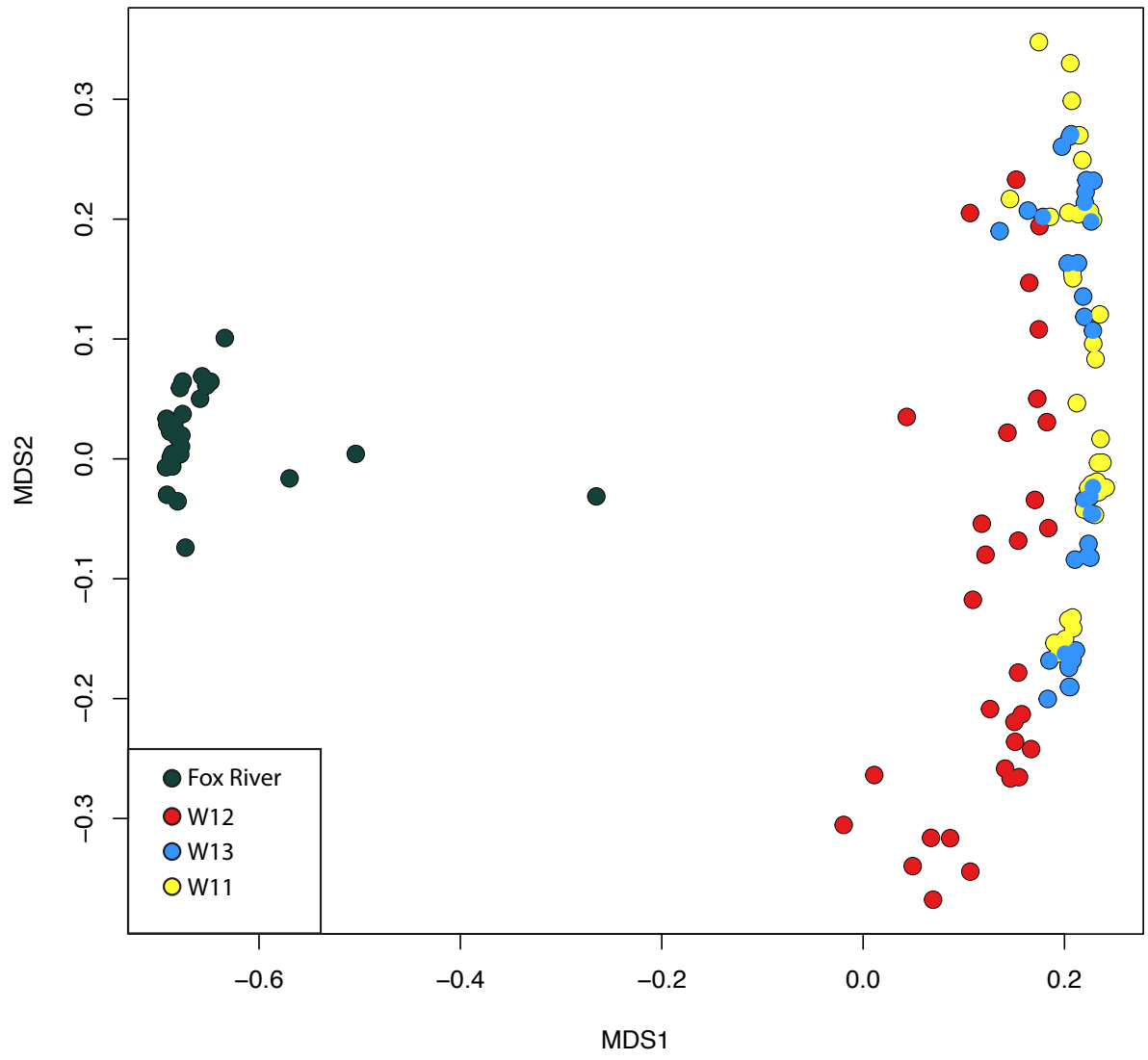


Figure 6: NMDS Ordination of Fox River and Groundwater Microbial Community Samples. Each dot depicts the microbial community of a sample as whole. The colors indicate the location of the sample. The samples from the Fox River and groundwater sites significantly differed from each other shown by the large split on the x-axis.

Although the microbial communities of the groundwater and Fox River are significantly different, some microbial community members were present in both the river and groundwater sites (Fig. 7). Previous preliminary 16S rRNA gene sequencing found no traces of fecal bacteria in the RBI well W12 suggesting that microorganisms were not entering the RBI well from the river (Fields-Sommers, 2015). This may not translate for all microorganisms. Although fecal

tracer bacteria were not found in the RBI well previously, this could mean that they may not be present, or they are at low enough abundances to fall below the detection limit.

An analysis of the most abundant river taxa was performed to determine similarities and differences between the river and wells, shown in Figure 7. The most abundant river ASVs were classified as typical fresh surface-water microorganisms, and these ASVs were used as tracers of the Fox River water in the wells. A complete list of these taxa is included in Appendix G. Some ASVs were not found in any of the groundwater wells, but other ASVs appeared in at least one well and/or across all sites. All ASVs were in much higher abundances in the Fox River—approximately 3,000x higher on average with a range of approximately 7x (W12) to 34,000x (W13), if present at all (Figure 7).

Given that W12 is actively drawing in river water, it was our hypothesis that taxa/ASVs more commonly associated with the Fox River would be more abundant in W12. The infiltration results supported this hypothesis. Although all wells contained Fox River indicator ASVs, W11 contained the least with low abundances, and W12 contained the highest abundances of Fox River indicator ASVs across the wells. W12 also followed more similar trends to the Fox River distributions, at lower abundances, than W11 and W13, suggesting that Fox River taxa may be transferring into W12 due to riverbank inducement.

W13 also contained river indicator ASVs in higher abundances than W11 but did not show similar trends to the Fox River. One explanation could be that these data suggest that the aquifer system has more of a connection to the Fox River than previously thought. W13 was considered to have no apparent hydrologic connection to the river, however, these data suggest that river microorganisms may enter the aquifer and be present in locations not experiencing river inducement. Another explanation could be that the river taxa are not traveling through the

soil matrix, but the ASVs of two considerably different microbes have similar sequences. This would occur if the evolution of the microbe diverged from the evolution of the 16S rRNA gene due to the slow rate of change in the 16S rRNA gene (Kirchman, 2012).

Overall, this study found evidence of the ASVs associated with river bacteria in all of the groundwater wells with higher abundances in W12, an active riverbank inducement well. The fate of river infiltrating microbes is unclear in these groundwater wells and whether they impact the typical activities carried out by groundwater microbes in the wells.

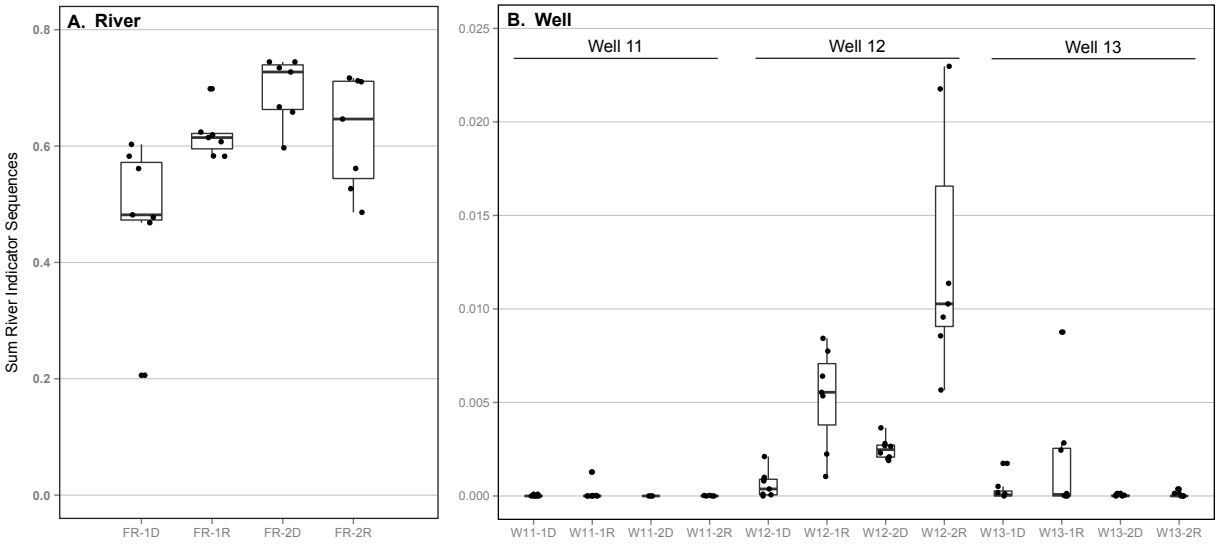


Figure 7: River Infiltration Taxa Comparison. The box plots show the total average abundant Fox River sequences for each sub-category for each location. The Fox River data is on the left and the wells are on the right with each filter size (0.1 and 0.2 µm indicated by 1 and 2 respectively) and nucleic acid fraction (RNA and DNA indicated by R and D). The boxes in the plots show the 25% to 75% percentiles (or the middle 50% ,called the inter-quartile range or IQR) of the average abundances. The horizontal line through the IQR box indicates the median. The whiskers show the lowest and highest values no further than 1.5x IQR away from the IQR, and all other points above/below the IQR are considered outliers. All data points are plotted, however. Note the y-scales indicating sum abundance are not the same, so the Fox River has a higher abundance of these indicator sequences.W12 had higher amounts of river indicator ASVs in each fraction of data as compared to the other two wells.

2.3.2 Groundwater Microbial Community Comparison

The dataset consisting of just the groundwater samples (W11, W12, and W13) was used to compare groundwater microbial communities. In Figure 8, a dendrogram was generated for the entire groundwater dataset to show similarities and dissimilarity relationships between the

groundwater samples. The dendrogram (Fig. 8) was generated using hierarchical clustering of pairwise dissimilarity between samples using Bray-Curtis dissimilarity and the functions vegdist() and hclust() in the vegan package in R (Oksanen et al., 2013). Essentially the "height" at which the branches merge at each node is relative to their similarity.

In this dataset of the groundwater sites, the first branch split (right to left) and therefore the largest factor contributing to the variation in the dataset, or driving a difference in the dataset, is site location—W12 is significantly different from W13 and W11 (PERMANOVA p= 0.001). The next factor driving the second largest difference in the dataset is the size fraction (PERMANOVA P = 0.001). The 0.2 μm and 0.1 μm communities differ within the dataset. The next factor driving a difference is location of the sites between W13 and W11. Finally, the last factor significantly contributing to a difference in the dataset is observed in the RNA and DNA fractions (PERMANOVA p = 0.001). There was no temporal significance in the distribution of the groundwater microbial community data shown in Fig. 8 (PERMANOVA p = 0.871).

Essentially, these data indicate that the overall community in W12, the RBI well actively drawing in river water, differs from W13 and W11, the pristine well and former RBI well that no longer pumps at a rate that draws in river water. This result could suggest that the former RBI well that no longer actively pumps and draws in river water is returning to a state similar to the non-river infiltrated groundwater, assuming it was previously affected like W12 is now. The data also show that cell size is a significant differentiator in the community and is a bigger factor in explaining the community variation than the location of W13 and W11. This means, the 0.2 μm communities in W13 are more similar to the 0.2 μm communities in W11 than to the 0.1 μm communities in W13—the same location—and vice versa.

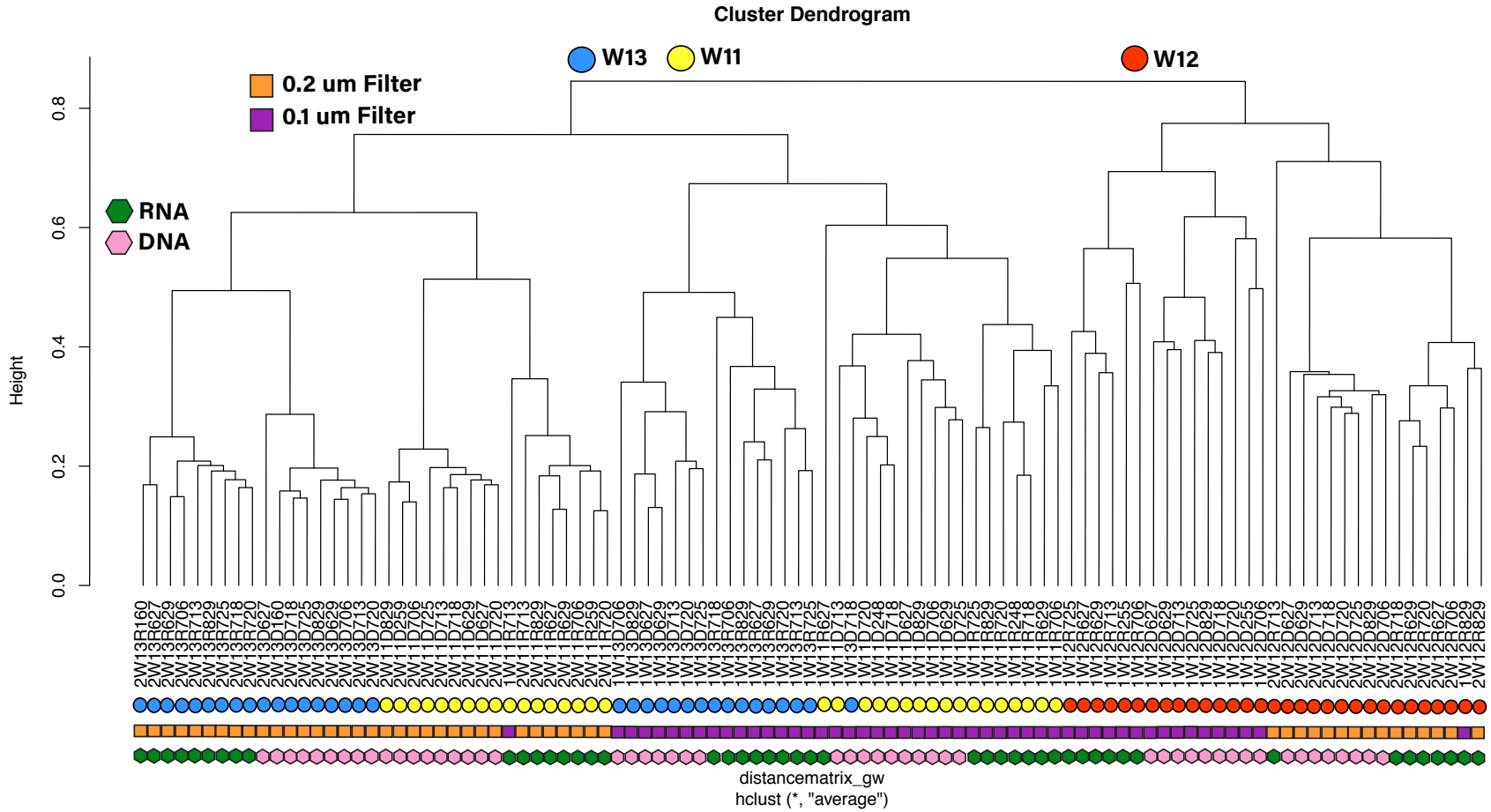


Figure 8: Groundwater Microbial Community Dendrogram. NMDS and Bray-Curtis dissimilarity were used to generate a dendrogram demonstrating the differences across groundwater microbial community samples. The groundwater microbial communities cluster first by well location in that W12 is siginificantly different from the other two wells. Filter size fraction then cluster together, then W13 and W11 cluster separately, and then RNA and DNA cluster together

These community differences can also be seen in Fig. 9. NMDS was utilized on separate RNA and DNA Bray-Curtis dissimilarity community matrices to develop ordinations of the microbial communities for the RNA and DNA groundwater communities separately to further distinguish these differences (displayed in Figure 9). In both the RNA and DNA ordinations, the groundwater sites distinctly and significantly cluster from each other. Also, the 0.1 µm and 0.2 µm communities cluster significantly. It can be noted, that the 0.2 µm communities cluster more tightly than the 0.1 µm communities, which indicates there is more variability in the 0.1 µm communities. This result is expected in that the 0.2 µm filter randomly allows 0.1 µm organisms to pass through due to filter clogging and inconsistency of always allowing the same 0.1 µm microorganisms to pass onto the 0.1 µm filter.

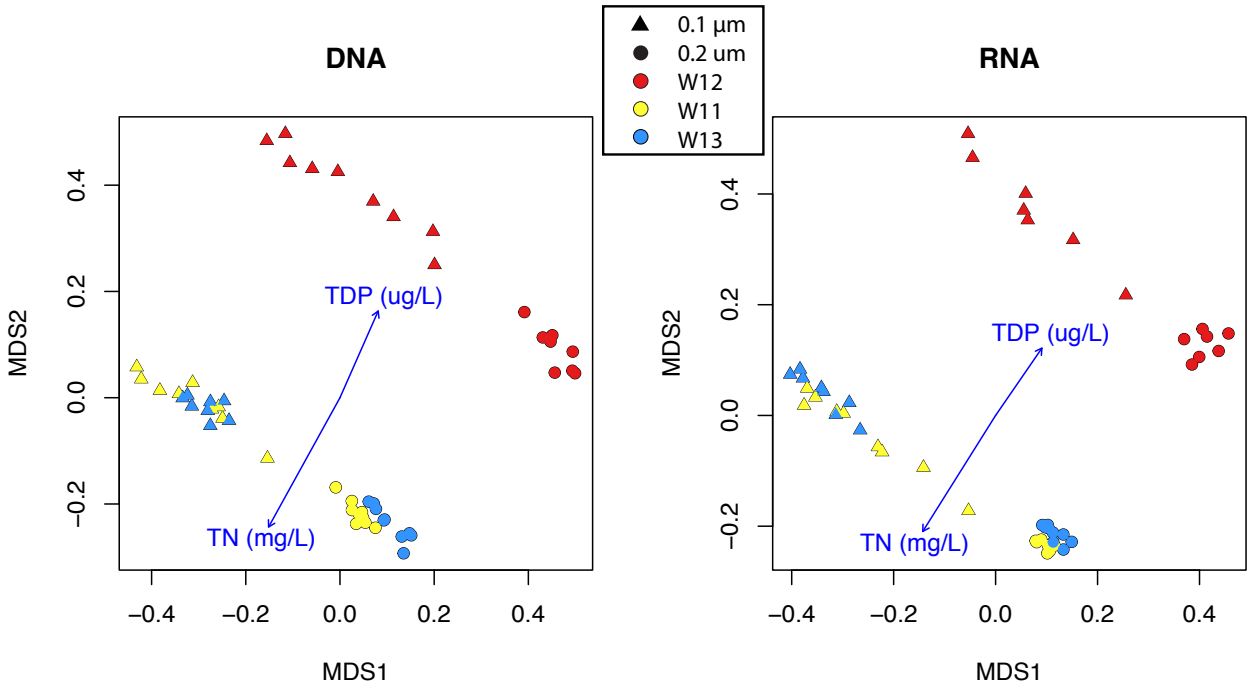


Figure 9: Ordination of DNA (left) and RNA (right) Groundwater Microbial Communities. Non-metric multidimensional scaling (NMDS) was based on Bray-Curtis distances. Communities from 0.1 µm samples are shown in triangles and communities from 0.2 µm samples are showing in circles. W11 is shown in green, W12 is shown in red, and W13 is shown in blue. Significant environmental parameters with a p value ≤ 0.01 (Total Nitrogen mg/L and Total Dissolved Phosphorus µg/L) were plotted to show their correlation in the data. Correlations between the environmental parameters and NMDS are shown in length and direction of the arrows. Nitrogen correlates to samples from W11 and W13 and phosphorus correlates to W12 samples.

Selected environmental chemical data was plotted onto the ordinations using envfit(), part of the vegan package in R (Oksanen et al., 2013). This function plots environmental conditions associated with the ordination samples and illustrates trends in those environmental samples. Direction and length demonstrate correlations between the data points and the environmental variable. Nitrogen correlates significantly with W11 and W13 (PERMANOVA $p = 0.001$ for both DNA and RNA communities) as TDP correlates significantly with W12 microbial community samples (PERMANOVA $p = 0.002$ for DNA and $p = 0.004$ for RNA communities).

Nitrate appears to be a key difference between the well locations. The chemical data in Table 3 showed that W12 measured close to the detection limit for nitrate while the other two wells had higher levels of nitrate. This could indicate there is no nitrate present in W12, or that it is being used by the microorganisms present. Nitrate is an important electron acceptor in groundwater, so it is hypothesized that the nitrate pool is reduced and/or incorporated into biomass and thus completely drawn down in W12.

Environmentally significant differences related to microbial community distribution were determined through statistical analyses. Statistical tests (PERMANOVA) were conducted on the four sub-datasets to parse out the differences in microbial community distributions associated with the environmental geochemical factors without interference from the other strong factors of nucleic acid type and filter size. Specifically nitrogen (total nitrogen, nitrate (mg/L), nitrite (mg/L), and ammonia (mg/L)) significantly (p-values ≤ 0.01) explains variation in the microbial communities across all wells (W11, W12, W13) and all sub-datasets (DNA 0.2 um, DNA 0.1 um, RNA 0.2 um, and RNA 0.1um). Below, Table 5 lists all p-values associated with metadata tested on the datasets. In this table, only p-values ≤ 0.01 are highlighted as the significant threshold.

Table 5: Chemical Data Significance in Groundwater Microbial Community Distribution.

Parameter	DNA 01 um	DNA 02 um	RNA 01 um	RNA 02 um
Date Number	0.968	0.922	0.911	0.943
Site	0.001 ***	0.001 ***	0.001 ***	0.001 ***
Nitrate mg/L	0.001 ***	0.001 ***	0.001 ***	0.001 ***
Nitrite mg/L	0.001 ***	0.001 ***	0.001 ***	0.001 ***
Ammonia mg/L	0.001 ***	0.001 ***	0.001 ***	0.001 ***
TDP µg/L	0.012	0.014	0.103	0.017
TDP mg/L	0.014	0.029	0.181	0.039
TN mg/L	0.001 ***	0.001 ***	0.001 ***	0.001 ***
Calcium mg/L	0.904	0.971	0.809	0.204
Potassium mg/L	0.055	0.055	0.047	0.029
Sodium mg/L	0.027	0.051	0.032	0.025
Magnesium mg/L	0.641	0.487	0.681	0.869
Sulfate mg/L	0.155	0.092	0.103	0.024
Chloride mg/L	0.115	0.065	0.09	0.006 **
DOC mg/L	0.031	0.19	0.073	0.117
Signif. codes: '*' 0.001 '**' 0.01**				

2.3.3 Groundwater Microbial Community Taxa Affiliations

Unclassified Organisms Dominate the Groundwater Dataset

Taxonomic affiliations were identified for each amplicon sequence variant with SILVA (Quast et al., 2013) as described in the methods. After identification, taxonomic affiliations were pooled to determine the total contribution of that known phylogeny in the dataset as a whole and are displayed in Table 6. Top taxa were also investigated across sites, filter, and nucleic acid fractions and compared (Fig. 11 Heatmap). Over 50% of the ASVs (Table 6, Fig. 10) in the dataset were not matched to any specific taxon and were therefore determined to be Unclassified organisms. This identification does not represent one specific taxon, but rather many different Unclassified taxa/ASVs.

Table 6: Top 50 Combined Groundwater Taxa Affiliations. Known taxa affiliations from ASV taxon classification were condensed to display the percentage of that taxon in the dataset.

	Taxonomy	% Across Dataset
1	Unclassified(100);	55.13%
2	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Acidiferrobacterales(100); Acidiferrobacteraceae(100); Sulfurifustis(100);	6.50%
3	Bacteria(100); Nitrospirae(100); 4-29-1(100);	6.15%
4	Archaea(100); Nanoarchaeaeota(100); Woesearchaeia(100);	4.52%
5	Bacteria(100); Proteobacteria(100); Deltaproteobacteria(100); DTB120(100);	2.91%
6	Bacteria(100); Proteobacteria(100); Deltaproteobacteria(100); MBNT15(100);	1.62%
7	Bacteria(100); Omnitrophicaeota(100);	1.56%
8	Bacteria(100); Nitrospirae(100); HDB-SIOI1093(100);	1.12%
9	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Gallionellaceae(100);	0.97%
10	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Rhodocyclaceae(100);	0.69%
11	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Sulfuricellaceae(100); Ferritrophicum(100);	0.55%
12	Bacteria(100); Bacteroidetes(100); Ignavibacteria(100); Kryptoniales(100); BSV26(100);	0.52%
13	Bacteria(100); Elusimicrobia(100); Elusimicrobia(100); Lineage IV(100);	0.50%
14	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100);	0.50%
15	Bacteria(100); Epsilonbacteraeota(100); Campylobacteria(100); Campylobacterales(100); Thiovulaceae(100); Sulfuricurvum(100);	0.50%
16	Bacteria(100); Patescibacteria(100); Parcubacteria(100); Candidatus Nomurabacteria(100);	0.45%
17	Bacteria(100); Gemmatimonadetes(100); Gemmatimonadetes(100); Gemmatimonadales(100); Gemmatimonadaceae(100); uncultured(100);	0.43%
18	Bacteria(100); Planctomycetes(100); Brocadiae(100); Brocadiales(100); Brocadiaceae(100);	0.34%
19	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Legionellales(100); Legionellaceae(100); Legionella(100);	0.32%
20	Bacteria(100); Nitrospirae(100); Thermodesulfovibrionia(100); uncultured(100);	0.32%
21	Bacteria(100); Patescibacteria(100); Parcubacteria(100); UBA9983(100);	0.31%

22	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Nitrosomonadaceae(100); GOUTA6(100);	0.29%
23	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Rhodocyclaceae(100); Thauera(100);	0.29%
24	Bacteria(100); Omnitrophicaeota(100); Omnitrophia(100); Omnitrophales(100); Omnitrophaceae(100); Candidatus Omnitrophus(100);	0.25%
25	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Gallionellaceae(100); Gallionella(100);	0.25%
26	Bacteria(100); Epsilonbacteraeota(100); Campylobacteria(100); Campylobacterales(100); Thiovulaceae(100); Sulfurimonas(100);	0.23%
27	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Hydrogenophilaceae(100);	0.23%
28	Archaea(100); Diapherotrites(100); Micrarchaeia(100);	0.22%
29	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Nitrosomonadaceae(100); MND1(100);	0.21%
30	Bacteria(100); Patescibacteria(100); Parcubacteria(100); Candidatus Azambacteria(100);	0.21%
31	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Gallionellaceae(100); Candidatus Nitrotoga(100);	0.21%
32	Bacteria(100); Proteobacteria(100); Deltaproteobacteria(100); Myxococcales(100);	0.20%
33	Bacteria(100); Patescibacteria(100); Microgenomatia(100); Candidatus Curtissbacteria(100);	0.18%
34	Bacteria(100); Proteobacteria(100); Deltaproteobacteria(100); Desulfobacterales(100); Desulfobulbaceae(100); Desulfocapsa(100);	0.18%
35	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Gallionellaceae(100); Sideroxydans(100);	0.18%
36	Bacteria(100); Latescibacteria(100);	0.17%
37	Bacteria(100); Actinobacteria(100); Thermoleophilia(100); Gaiellales(100);	0.17%
38	Bacteria(100); Actinobacteria(100); MB-A2-108(100);	0.17%
39	Bacteria(100); Proteobacteria(100); Deltaproteobacteria(100); Myxococcales(100); bacteriap25(100);	0.16%
40	Bacteria(100); Patescibacteria(100); Microgenomatia(100); Candidatus Woesebacteria(100);	0.16%
41	Bacteria(100); Patescibacteria(100); Microgenomatia(100); Candidatus Woykebacteria(100);	0.16%
42	Bacteria(100); Proteobacteria(100); Deltaproteobacteria(100); Oligoflexales(100); 0319-6G20(100);	0.16%
43	Bacteria(100); Bacteroidetes(100); Bacteroidia(100); Sphingobacteriales(100); AKYH767(100);	0.16%
44	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); Hydrogenophilaceae(100); uncultured(100);	0.16%

45	Archaea(100); Diapherotrites(100); Iainarchaeia(100); Iainarchaeales(100);	0.16%
46	Bacteria(100); Patescibacteria(100); Microgenomatia(100); Candidatus Amesbacteria(100);	0.16%
47	Bacteria(100); Planctomycetes(100); Phycisphaerae(100); CCM11a(100);	0.15%
48	Bacteria(100); Acidobacteria(100); Subgroup 6(100);	0.15%
49	Bacteria(100); Chloroflexi(100); KD4-96(100);	0.14%
50	Bacteria(100); Proteobacteria(100); Gammaproteobacteria(100); Betaproteobacteriales(100); TRA3-20(100);	0.14%

Figure 10 further demonstrates the large fraction of Unclassified ASVs found in the groundwater dataset. All groundwater samples from this study (divided into DNA, RNA, and filter size) have a proportionately higher amount of Unclassified ASVs compared to river datasets. Figure 10 shows the average relative abundance of Unclassified ASVs from the sub-groups of each dataset.

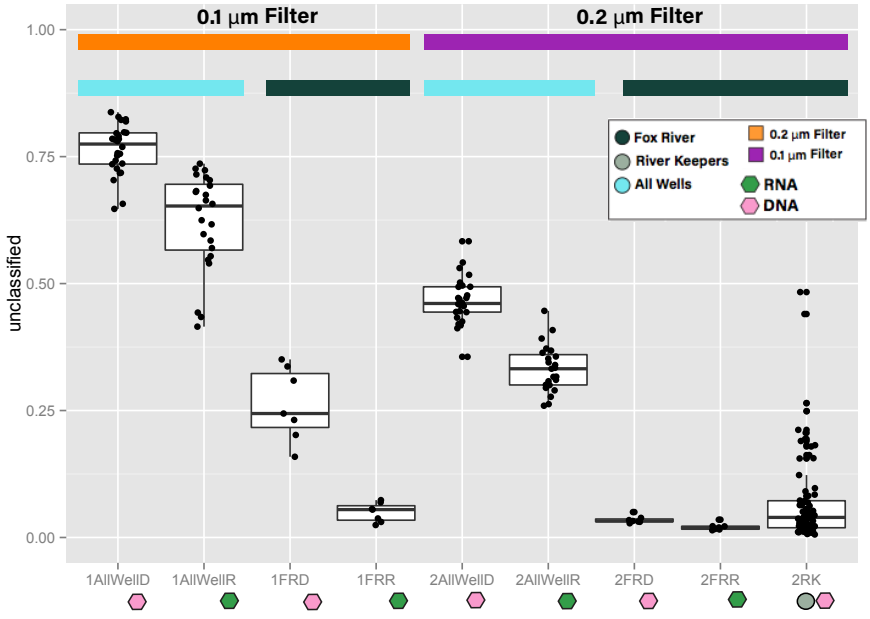


Figure 10: Proportion of Unclassified ASVs Compared Between Groundwater and River Environment Datasets. The groundwater samples are displayed below the light blue bar. River samples are displayed below the dark green bar. Filter fraction is split with 0.1 µm on the left and 0.2 µm on the right. River Keeper data is displayed on the right and consists of data from 16 Milwaukee Area rivers. All other river data is from the Fox River dataset from this study. The Groundwater dataset has higher proportions of Unclassified ASVs compared to the rivers. The 0.1 µm fraction always show higher proportions of Unclassified ASVs than their counterpart. The River Keeper data only includes 0.2 µm community data because 0.1 µm practices are not prevalent in the field currently. Likely smaller organisms have evaded detection previously comparatively to 0.2 µm members so more Unclassified ASVs are found on the 0.1µm filters.

Potential Physiological Implications of Dominant Microbial Affiliations

Although most of the groundwater ASV taxa affiliations are Unclassified, many of the classified taxa associate with known groundwater microbes and recently discovered CPR and DPANN organisms. Many CPR and DPANN organisms have mostly been detected through genomic and phylogenetic analyses from mostly oxygen-limited or anaerobic environments (Castelle et al., 2018). Here, some dominant groundwater ASV taxonomic affiliations and potential physiological impacts will be discussed.

Woesarchaeota, a relatively large proportion of this microbial dataset (~ 4%), are part of the recently discovered DPANN superphylum. Genomic analyses and metabolic reconstructions suggest these are anaerobic heterotrophic organisms that may have symbiotic relationships with other microbes, specifically methanogens (Liu et al., 2018). Some *Woesarchaeota* are thought to be fermenters with iron or methane metabolisms. Others have unknown metabolisms and are thought to be obligate symbionts (Castelle et al., 2018).

Candidatus Parcubacteria, part of the recently discovered CPR superphylum, are suggested to have a variety of metabolisms and many may be fermenters that can produce acetate, ethanol, lactate, and hydrogen. A number of genomes contain nitrite reductases (*nirK* and *nirB*) which can utilize nitrite to produce nitric oxide (*nirK*) and ammonium (*nirB*). Other genomes suggest nitrate reduction and anaerobic respiration (Castelle et al., 2018). *Candidatus* Peregrinebacteria are also part of the CPR superphylum and expected to be non-respiring anaerobes that likely ferment. *Candidatus* Micrarchaeota are predicted to have fermentation and heterotrophic O_2 respiratory capabilities.

Melainabacteria, also found in this dataset but not displayed in Table 6, have previously been recovered from aquifer sediment and appear to be anaerobic obligate fermenters and

diverged from cyanobacteria ~2.5 billion years ago (Di Rienzi et al., 2013; Soo et al., 2017). They are thought to be able to use a wide variety of carbon sources and use a FeFe hydrogenase for H_2 production, but likely need syntrophic partners to sustain low partial hydrogen pressures such as symbiotic acetogens or methanogens (Di Rienzi et al., 2013). Also, some aquifer related *Melainabacteria* show nitrogen fixation capabilities with nitrogen fixation genes (nifE, nifV, nifS, nifU, nifB, and nifB/X) and nitrogenase complexes. These microbes are believed to likely play a key part in carbon cycling in subsurface systems.

Other taxa found in the dataset appear to have strong associations with sulfur mechanisms. *Desulfocapsa* and *Desulfobacca* have been associated with sulfate reduction (Göker et al., 2011; Oude Elferink et al., 1999; Tonolla et al., 2000). *Desulfobacca* also have been described converting acetate through methanogenesis even in the presence of high amounts of sulfate (Oude Elferink et al., 1999). *Desulfomonile* has been characterized as an obligate anaerobe capable of growth through the use of fumarate, sulfate, sulfite, thiosulfate, and nitrate as electron acceptors (Sun et al., 2001). *Sulfurifustis* from the *Acidiferrobacteraceae* family are known sulfur oxidizers (Kojima et al., 2015).

Other taxa in this groundwater microbial dataset consist of previously described and studied organisms associated with specific capabilities. *Gallionellaceae* have previously been described in anaerobic groundwater environments containing ferrous iron that comes in contact with an oxygenated environment. *Hydrogenophilaceae* is a family with known members that are obligately thermophilic autotrophs that utilize molecular hydrogen while others are obligately respiratory and use oxygen or nitrate (Boden et al., 2017). Some *Nitrospirae* organisms are known to have dissimilatory sulfate reduction nitrite oxidizing capabilities. Nitrate reduction has been characterized for members of the *Bukholderiaceae* family. *Rhodocyclaceae* have been

described as fixing nitrogen (aerobically), degrading carbon compounds utilizing oxygen, nitrate, chlorate, perchlorate, selenite, etc., and as anaerobes, chemoautotrophs, and methylotrophs oxidizing sulfur and producing propionic acid from fermentation (Oren, 2014). The taxa associated with the groundwater microbial communities in this dataset suggest the potential for fermentation, methanogenesis, nitrate reduction, sulfate reduction, nitrite oxidation, iron oxidation, sulfur oxidation in the groundwater wells. Overall, many of the dominant taxa affiliations correspond to anaerobic metabolisms involving nitrate and sulfate reduction. The affiliations also suggest fermentative lifestyles as well as sulfur, H_2, nitrite, and iron oxidation capabilities.

Taxa Comparisons Between Sites

Top taxa were also investigated across sites, filter, and nucleic acid fractions and compared (Fig. 11 Heatmap). The ten average most abundant ASV's were determined from each sub-category of the dataset (well site, nucleic acid type, and filter fraction). This could potentially produce a total of 120 unique ASVs (top 10 ASVs x 2 nucleic acid types x 3 well sites x 2 filter sizes). However, a total of 56 unique ASVs were determined to be within the top ASVs across the dataset. A heatmap was generated using the relative abundances for these ASVs across samples, displayed in Fig. 11, which shows ASV relationships between RNA and DNA (activity), presence, and abundance across sites and filter sizes.

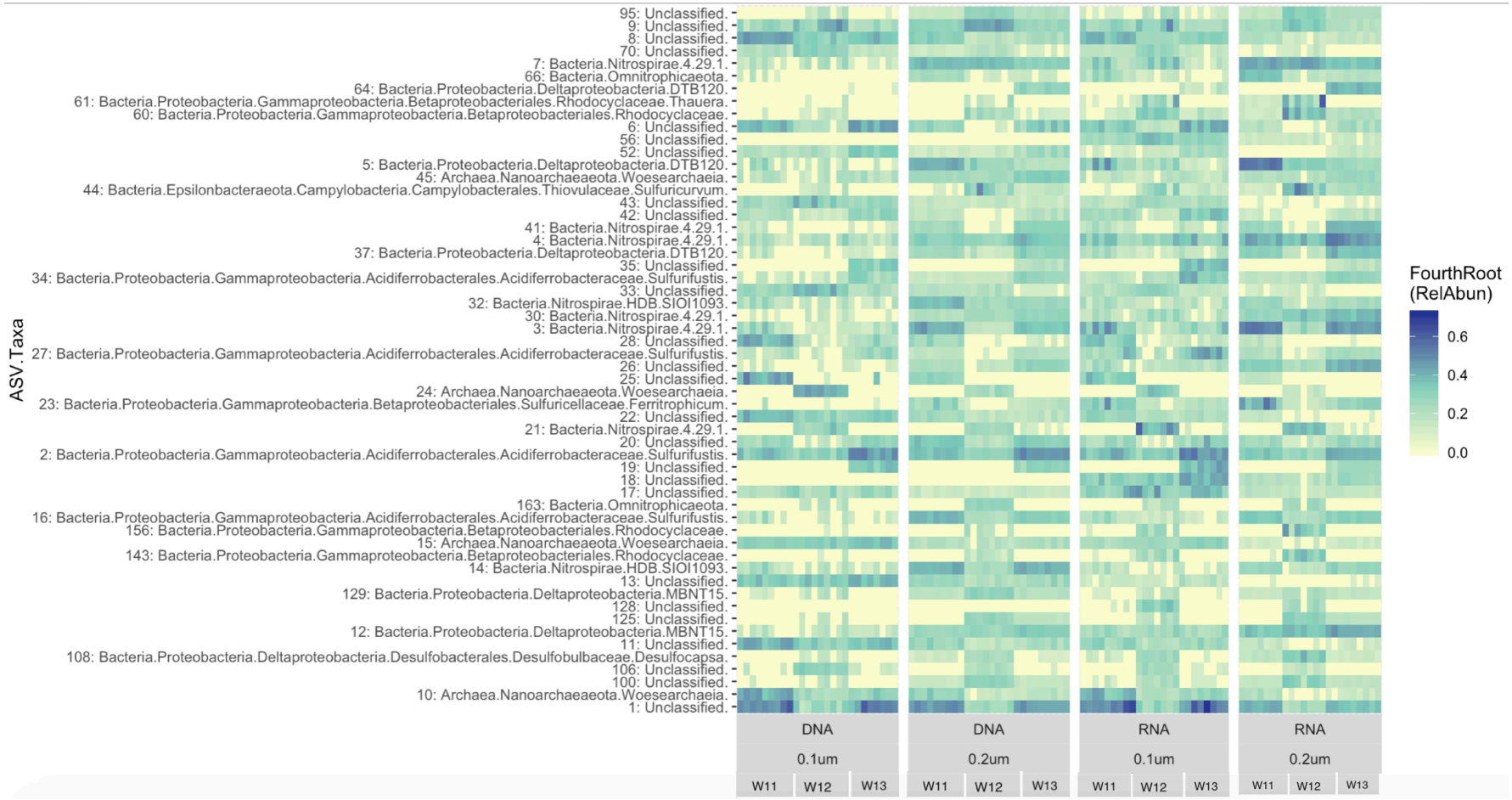


Figure 11: Most Abundant ASVs Across All Wells. Taxon affiliations are listed on the left. Data is partitioned based on DNA (left half of the heatmap) and RNA (right half of the heatmap), filter size with nucleic acid type (0.1 µm left, 0.2 µm right within DNA and RNA), and within that by site (W11, W12, and W13). The light yellow indicates an average abundance of 0 meaning the ASV was not present for that type of sample. Dark blue indicates the highest relative abundance. The fourth root was used to display the relative abundances to show the patterns more distinctly. Some of the most abundant groundwater ASVs are specific to location suggesting they are unique or specialized to that location. Some ASVs that are Unclassified, i.e. 128, 18, 56, do not appear in the DNA but are of the most abundant RNA ASVs, which supports the theory of intron splicing occurring in the 16S rRNA gene.

The data of the heatmap (Fig. 11) show unique relationships between the samples. Some ASV's are shared across samples, yet some ASVs are only found in specific samples and well locations suggesting these organisms may be unique or more specialized to their respective environments. For example, ASVs found only in W13 in the heatmap include:

- ASV 64: *Bacteria; Proteobacteria; Deltaproteobacteria; DTB120*,
- ASV 35: Unclassified, and
- ASV 19: Unclassified.

ASVs in the heatmap only found in W12 include:

- ASV 143: *Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales*; *Rhodocyclaceae*,
- ASV 24: *Archaea; Nanoarchaeaeota; Woesearchaeia* and
- ASV 163: *Bacteria; Omnitrophicaeota*.

Furthermore, ASV 61 and 60 (both *Rhodocyclaceae*) appear to be preferential to W12 and more active since the RNA fraction is more abundant than its DNA fraction. Some anaerobic *Rhodocyclaceae* have been described as fermenters of organics and as sulfur-oxidizing chemoautotrophs, methylotrophs, and fermenters (Oren, 2014). ASV 108: *Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae; Desulfocapsa* may be more abundant in W12. This family has been associated with sulfate reduction.

Similarly, ASVs 3, 4, and 41 which are all *Bacteria; Nitrospirae; 4.29.1* appear to be more abundant and more active in W13 than in W12, but still found between both sites. *Nitrospirae* have been associated with sulfate reduction and nitrite oxidation. ASV 23: *Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Sulfuricellaceae;*

Ferritrophicum also appears to be more abundant and potentially more active, especially in the 0.1 μm fraction, in W13 than W12.

Other ASVs are found throughout the sites and samples, suggesting other organisms are present throughout the aquifer. For example, Unclassified ASV 1 is present in all subsets of the data at a relatively high abundance across all samples. Among Unclassified, some of the other shared ASVs relate to *Woesearchaeia*, *Nitrospirae*, *Sulfuricurvum*, *Sulfurifustis*, *Sulfuricelleceae*, and *Desulfocapsa*.

Finally, the data in the heatmap also show an interesting and unique observation/discovery from a biological and evolutionary perspective. Some ASVs (128: Unclassified, 18: Unclassified, and 56: Unclassified) show an abundance of 0 across the DNA fraction. However, these ASVs were of the highest relative abundances in the RNA data. Since RNA is transcribed from DNA, it's likely not possible that there is no DNA encoding this subsequent RNA. Most likely, these organisms contain introns within their 16S rRNA gene which would not be captured from standard 16S rRNA gene surveys. This supports recent hypotheses and discoveries that some of these unclassified and recently discovered DPANN and CPR phyla, which vastly expanded the tree of life, may have introns and intron splicing within their 16S rRNA genes which is a rarely described feature in Bacteria and Archaea. This finding also suggests that typical 16S rRNA gene surveys only targeting the DNA fails to capture all 16S rRNA phylogenetic markers from an environment.

Unique ASVs related to W12 and W13 were further investigated using Multinomial Species Classification Method (CLAM) from the vegan package in R. CLAM was performed to identify unique ASVs distinctly affiliated with W13 or W12. CLAM is a statistical approach for classifying generalist and specialists from two distinct habitats (Chazdon et al., 2011). ASVs

were separated into four groups dependent on distribution within samples (displayed in Fig. 12). The method distinguished specialist ASVs favored for the W13 habitat, specialist ASVs favored for the W12 habitat, ASVs with apparent equal distributions (generalists), and ASVs too rare to consider for categorization. Stringent conditions were used to discriminate specialists to each habitat of W13 and W12. The conditions included a specialization threshold of 5/6 which essentially indicates a specialist organism/sequence has to occur at least 5x more in one well versus the other, an alpha of 0.0001/(total number of sequences in that dataset), and a coverage limit of 1. The number of ASV specialists for each well from each sub-dataset was determined to be: W13—DNA 0.2um: 87, DNA 0.1 um: 82, RNA 0.2 um: 74, RNA 0.1 um: 78, and W12—DNA 0.2um: 132, DNA 0.1 um: 128, RNA 0.2 um: 119, RNA 0.1 um: 112.

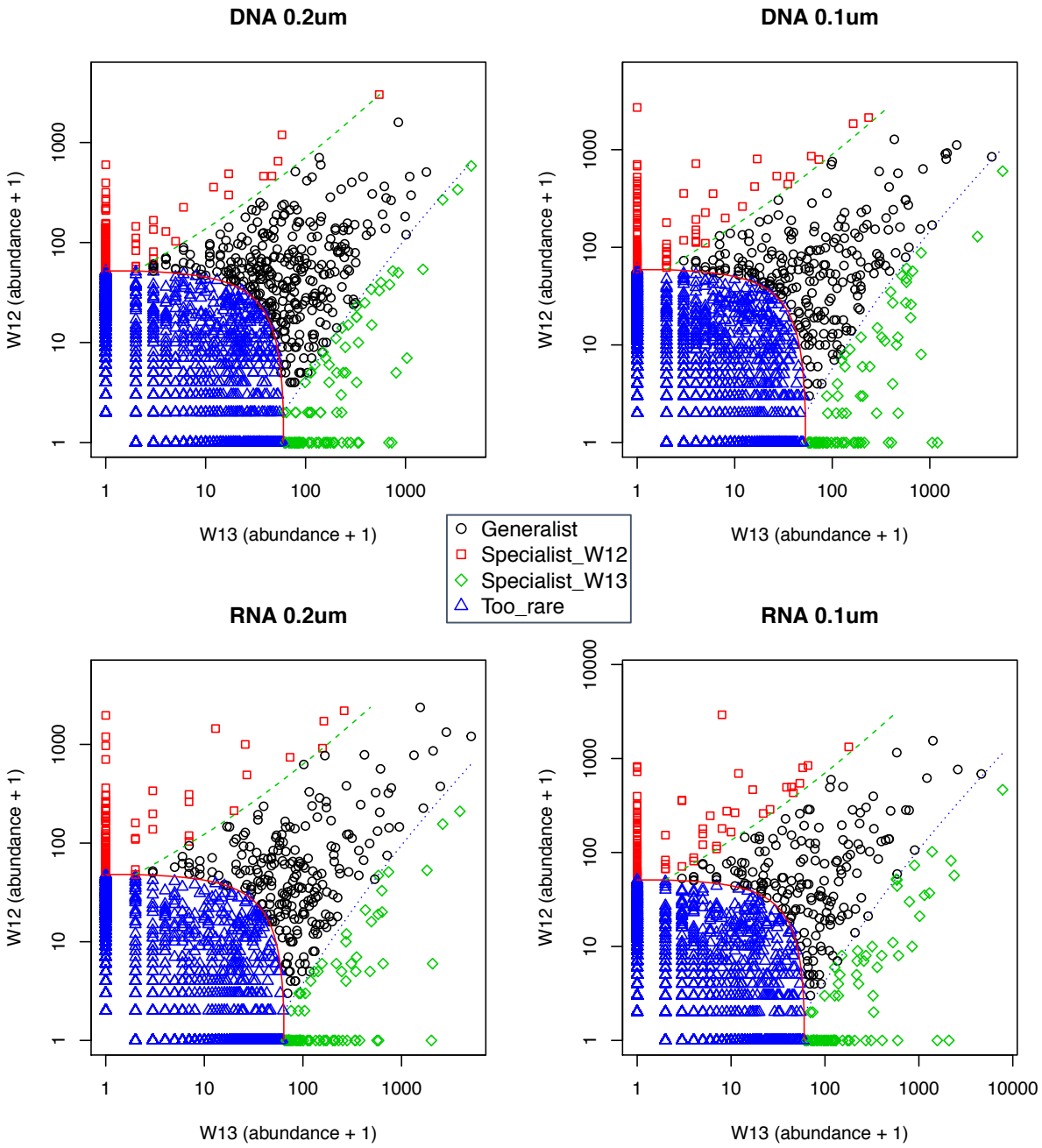


Figure 12: Groundwater DNA and RNA Multinomial Species Classification Method (CLAM) for W13 and W12. The blue triangles indicate rare sequences that were too rare to consider in the classification analysis, the black circles indicate generalists species/sequences that occur between both samples at similar abundaces, the red squares indicate specialists for the W12 environment, and the green diamonds indicate specialist species for W13. The numbers of specialists for each well from each sub-dataset (DNA 0.2um, DNA 0.1um, RNA 0.2 µm and RNA 0.1um) were determined to be W13: (top left to right starting with DNA 0.2) 87, 82, 74, 78, and W12: 132, 128, 119, 112.

The ASVs and associated SILVA taxonomic affiliations were identified for the specialists for both W13 and W12. A complete list of the unique ASVs and associated taxa can

be found in Appendix D and E. Only the specialists overlapping in all four datasets are displayed below. From this analysis, W13 appears to contain specialists potentially related to nitrite reduction, fermentation (*Parcubacteria*), iron oxidation (*Gallionellaceae*), sulfur oxidation (*Sulfurifustis*) across all four subsets of the dataset. W12 appears to contain some specialists that could potentially be related to sulfate reduction and nitrite oxidation (*Nitrospirae*), fermentation, and iron or methane metabolisms (*Woesarchaeota*).

Table 7: The specialist taxa determined for W13 occurring between all subcategories of DNA, RNA, 0.1 and 0.2 µm from CLAM. A complete list of all specialists from each subcategory can be found in Appendix E.

ASV ID	Unique W13 Taxa
6	Unclassified;
19	Unclassified;
35	Unclassified
42	Unclassified;
47	Bacteria; Patescibacteria; Parcubacteria; Candidatus; Nomurabacteria;
48	Unclassified;
88	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
89	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
102	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
104	Unclassified;
126	Bacteria; Patescibacteria; Parcubacteria; UBA9983;
232	Bacteria; Proteobacteria; Alphaproteobacteria; Acetobacterales; Acetobacterales; Incertae; Sedis; uncultured;
288	Bacteria; Patescibacteria; Parcubacteria; UBA9983;
368	Unclassified;

Table 8: The specialist taxa determined for W12 occurring between all subcategories of DNA, RNA, 0.1 and 0.2 µm from CLAM. A complete list of all specialists from each subcategory can be found in Appendix D.

ASV ID	Unique W12 Taxa
9	Unclassified;
21	Bacteria; Nitrospirae; 4; 29; 1;
53	Archaea; Nanoarchaeaeota; Woesearchaeia;
106	Unclassified;
147	Unclassified;
163	Bacteria; Omnitrophicaeota;
317	Unclassified;
330	Unclassified;

2.3.3 16S ribosomal RNA:DNA Activity

Furthermore, Multinomial Species Classification Method (CLAM) (Chazdon et al., 2011) was used to determine active specialists between W12 and W13. The Generalist ASVs from the DNA CLAM tests (Figure 12 DNA 0.1 μm and DNA 0.2 um) were used to determine differentially active ASVs between W12 and W13 among the present ASVs in both W12 and W13. All values of 0 were set to 0.00005 (half as much as the initial threshold cutoff of 0.01% described on page 20) in the DNA and RNA prior to taking the 16S ribosomal RNA:DNA ratio in order to obtain a non-zero ratio value. Values were multiplied by 1,000 and rounded in order to obtain integers for all values to use in the CLAM method. Stringent conditions were used to discriminate active specialists from each habitat of W13 and W12. The conditions included a specialization threshold of 3/4 which essentially indicates a specialist organism/sequence has to be at least 3x more active in one well versus the other, an alpha of 0.0001/(total number of sequences in that dataset), and a coverage limit of 1. W13 resulted in 29 active specialists. W12 resulted in 20 active specialists (Table 9).

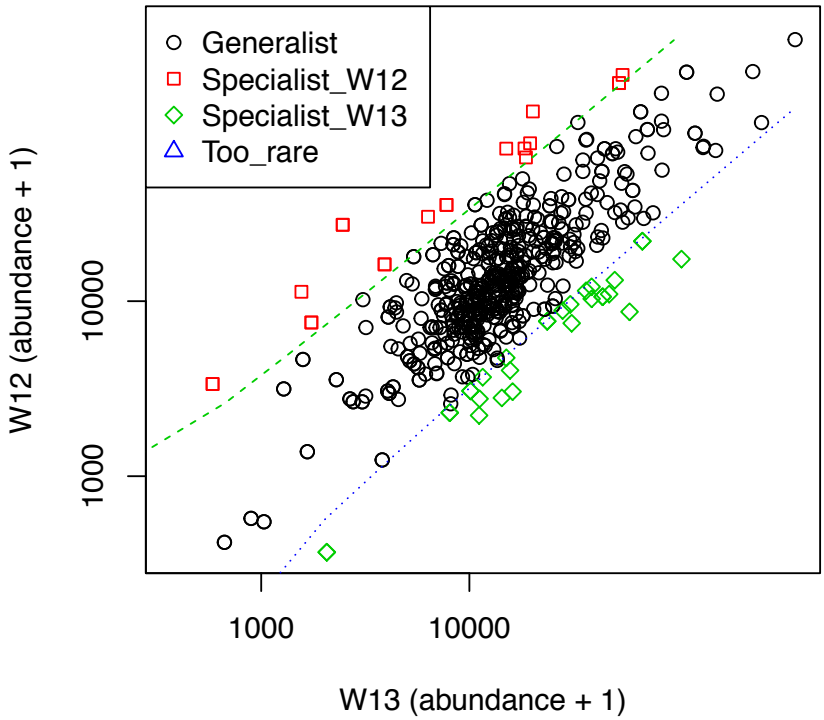


Figure 13: CLAM 16S ribosomal RNA:DNA Ratio Specialists for W12 and W13. 16S ribosomal RNA:DNA ratios indicate activity for a specific ASV between the two sites, W12 and W13, because 16S rRNA is part of the ribosome that mediates protein synthesis and the DNA is the 16S rRNA gene encoding for the RNA. In theory, the more 16S rRNA present indicates more ribosomes which would indicate a more active cell as compared to less ribosomes and less 16S rRNA from the same ASV. The black circles indicate generalists (16S ribosomal RNA:DNA ratios that occur between both samples at similar values), the red squares indicate specialist ASVs for the W12 environment with ratios at least 3 times higher than in W13, and the green diamonds indicate specialist species for W13 with ratios at least 3 times higher than ratios in W12 for that given ASV.

Table 9: W13 and W12 Specialists from CLAM Analysis on 16S ribosomal RNA:DNA Ratios

ASV and Taxon	Classes
32: Bacteria; Nitrospirae; HDB-SIOI1093;	Specialist_W13
68: Unclassified;	Specialist_W13
82: Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;	Specialist_W13
109: Bacteria; Patescibacteria; Microgenomatia; Candidatus Amesbacteria;	Specialist_W13
171: Unclassified;	Specialist_W13
186: Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;	Specialist_W13
239: Unclassified;	Specialist_W13
285: Unclassified;	Specialist_W13
293: Unclassified;	Specialist_W13
344: Unclassified;	Specialist_W13

351: Unclassified;	Specialist_W13
650: Unclassified;	Specialist_W13
771: Unclassified;	Specialist_W13
867: Unclassified;	Specialist_W13
829: Unclassified;	Specialist_W13
963: Unclassified;	Specialist_W13
980: Unclassified;	Specialist_W13
1128: Unclassified;	Specialist_W13
32: Bacteria; Nitrospirae; HDB-SIOI1093;	Specialist_W13
109: Bacteria; Patescibacteria; Microgenomatia; Candidatus Amesbacteria;	Specialist_W13
145: Unclassified;	Specialist_W13
171: Unclassified;	Specialist_W13
182: Unclassified;	Specialist_W13
186: Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;	Specialist_W13
292: Bacteria; Omnitrophicaeota;	Specialist_W13
293: Unclassified;	Specialist_W13
441: Bacteria; Rokubacteria; NC10; Rokubacteriales;	Specialist_W13
771: Unclassified;	Specialist_W13
872: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Rhodoferax;	Specialist_W13
7: Bacteria; Nitrospirae; 4-29-1;	Specialist_W12
133: Unclassified;	Specialist_W12
178: Unclassified;	Specialist_W12
210: Unclassified;	Specialist_W12
360: Unclassified;	Specialist_W12
475: Unclassified;	Specialist_W12
497: Unclassified;	Specialist_W12
654: Archaea; Nanoarchaeaeota; Woesearchaeia;	Specialist_W12
7: Bacteria; Nitrospirae; 4-29-1;	Specialist_W12
125: Unclassified;	Specialist_W12
136: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; TRA3-20;	Specialist_W12
133: Unclassified;	Specialist_W12
178: Unclassified;	Specialist_W12
210: Unclassified;	Specialist_W12
263: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Sulfuricellaceae; Sulfuricella;	Specialist_W12
301: Unclassified;	Specialist_W12
338: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;	Specialist_W12
360: Unclassified;	Specialist_W12

475: Unclassified;	Specialist_W12
523: Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Xanthobacteraceae; Pseudolabrys;	Specialist_W12

To compare and compile specialist taxa between sites, taxa affiliations were compiled from the specialist analyses previously described. Specialist presence or absence in W12 and W13 are indicated in Table 10. It appears some taxa affiliations are present between W13 and W12. Some families are known to be associated with certain functions. However, taxa definitions and names can be extremely broad with specific strains differing greatly. Also, some specialist taxa appear to be favored in certain environments. Specifically, the specialists determined by ratio analysis, suggest those taxa/ASVs are more active in one location as compared to the other. For example, *Woesearchaeia, Gallionellaceae,* and *Rhodocyclaceae* were ratio specialists in W12 and *Sulfurifustis*, *Omnitrophicaeota*, and *Nitrospirae HDB-SIO1093* were ratio specialists in W13.

Table 10: Compilation of Groundwater Specialist Taxa Affiliations Compared Between W12 and W13. This table shows some of the ASV taxa affiliations as described previously and determined to be specialist in either W12 or W13 through various analyses (Figures 11, 12, 13, Appendix J). The blue indicates presence in W13 and the red indicates presence in W12. Some taxa affiliations are found between sites while others are more specific for each site.

Taxon	Mean % in all wells	Affiliation	Specialists from Abundance of DNA or RNA (Fig.12)		Specialists from RNA:DNA Ratios (Fig. 13)		Highest Ratios (Appendix J)	
			W13	W12	W13	W12	W13	W12
Unclassified	55.10%	??????????						
Sulfurifustis	6.5%	Sulfur-oxidizer Chemolithoautotroph growing by oxidation of inorganic sulfur						
Nitrospirae 4-29-1	6.10%	Likely nitrite oxidizer						

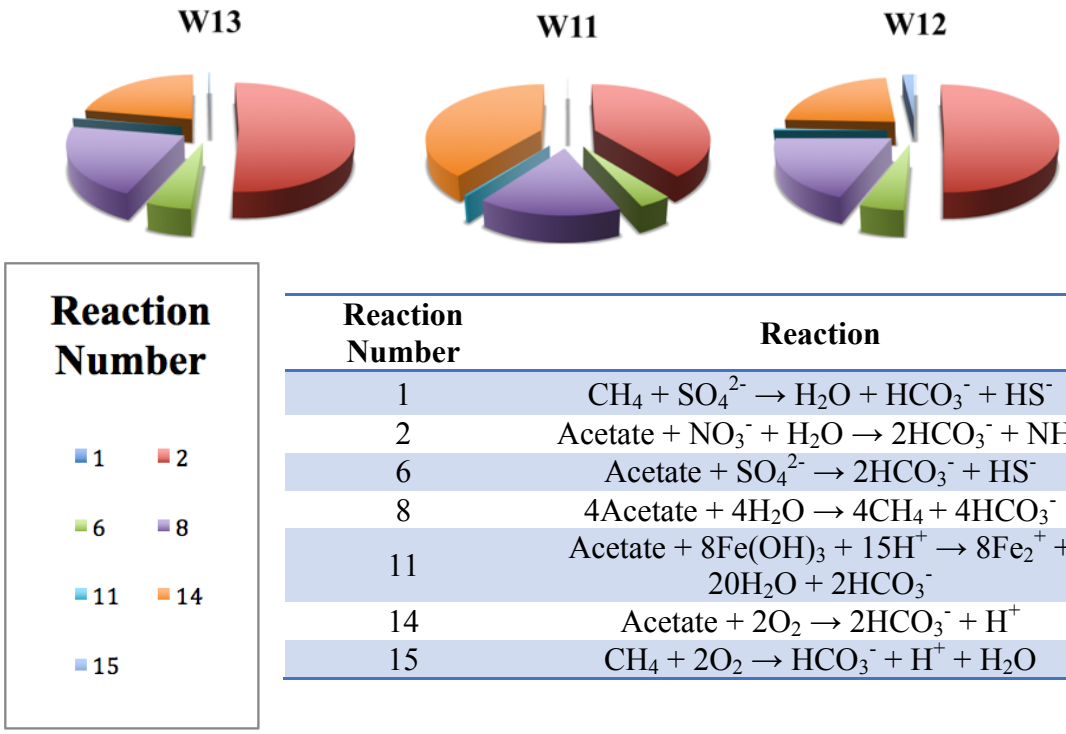
Woesearchaeia	4.50%	Anaerobic lifestyle-fermentation, likely syntrophic, unknown metabolisms but some likely iron or methane metabolisms							
Omnitrophicaeota	1.60%	Found in Anoxic environments, unknown metabolism							
Nitrospirae HDB-SIO1093	1.10%	Likely nitrite oxidizer							
Rhodocyclaceae	0.70%	Common in anoxic environments – members are known to be chemoautotrophs, methylotrophs, fermenters							
Hydrogenophilaceae	0.20%	Typically have chemolithoautotrophic lifestyles – sulfur or H_2 oxidizers							
Gallionellaceae	1.00%	Typically chemolithoautotrophs as iron & sulfur-oxidizers							
Nomurabacteria	0.50%	*Parcubacteria* superphylum (CPR) – unknown metabolism							
Sulfurimonas	0.20%	Sulfur-oxidizing bacteria, most common in marine environments							
Parcubacteria-UBA9983	0.30%	Typically larger bacteria (captured on 0.2 um), diverse and varied metabolic capabilities							
Thermodesulfovibrionia	0.30%	Typically sulfate-reduction and a thermophile							
Nitrotoga	0.20%	Nitrite-oxidizer							
Gallionella	0.30%	Typically chemolithoautotrophs as iron & sulfur-oxidizers							

Thauera	0.30%	Typically hydrocarbon, phenol degrader – often found in conditions where complex industrial contaminants are being degraded						
Omnitrophus	0.30%	Found in Anoxic environments, unknown metabolism						
Delta – DTB120	2.90%	Found in Anoxic environments – specifically when CH_4 is present, unknown metabolism						
Delta – MBNT15	1.60%	Found in Anoxic environments, unknown metabolism						
Ferritrophicum	0.60%	Typically Fe(II)-oxidizers						
Kryptoniales-BSV26	0.50%	Unknown – typically associated with geothermal springs						
Sulfuricurvum	0.50%	Sulfur-oxidizer with nitrate/nitric oxide, and possibly H_2 oxidation.						
Brocadiaceae	0.30%	Anammox group						
Thermodesulfovibrionia	0.30%	Typically sulfate-reduction and a thermophile						
Nitrosomonadaceae-GOUTA6	0.30%	Typically an ammonia-oxidizing group that also fixes CO_2						
Nitrosomonadaceae-MND1	0.20%	Typically an ammonia-oxidizing group that also fixes CO_2						
Myxococcales	0.20%	Common soil bacterium and known for spore forming capability						

Overall, known taxa affiliations of the groundwater dataset indicate fermentation, methanogenesis, nitrate reduction, sulfate reduction, nitrite oxidation, iron oxidation, and sulfur oxidation in the groundwater wells. The data also indicates anaerobic metabolisms involving

nitrate and sulfate reduction with fermentation, sulfur oxidation, H_2 oxidation, nitrite oxidation, and iron oxidation capabilities. This is also consistent with thermodynamic free energy flux results (Salo, 2018) from the groundwater wells.

The pie charts in Figure 14 show the favorable heterotrophic and fermentation biogeochemical reactions in the three shallow groundwater wells. Among the heterotrophic reactions, nitrate is the primary electron acceptor. RBI well W12 and pristine W13 appear to have the same distributions of free energy flux through all reactions. The 4 dominant reactions include heterotrophic metabolisms (14, 2, 6), nitrate reduction (2), sulfate reduction (6), and fermentation (8).



Reaction Number	Reaction
1	$CH_4 + SO_4^{2-} \rightarrow H_2O + HCO_3^- + HS^-$
2	$Acetate + NO_3^- + H_2O \rightarrow 2HCO_3^- + NH_3$
6	$Acetate + SO_4^{2-} \rightarrow 2HCO_3^- + HS^-$
8	$4Acetate + 4H_2O \rightarrow 4CH_4 + 4HCO_3^-$
11	$Acetate + 8Fe(OH)_3 + 15H^+ \rightarrow 8Fe_2^+ + 20H_2O + 2HCO_3^-$
14	$Acetate + 2O_2 \rightarrow 2HCO_3^- + H^+$
15	$CH_4 + 2O_2 \rightarrow HCO_3^- + H^+ + H_2O$

Figure 14: Thermodynamic Free Energy Flux Distributions for Heterotrophic and Fermentation Metabolisms in the Groundwater Wells. The reaction equations are from Table 4. The 4 dominant reactions include reactions 2, 8, 14, and 6 which are heterotrophic metabolisms (14, 2, 6), nitrate reduction (reaction 2, red), sulfate reduction (reaction 6, green), and fermentation (reaction 8, purple). Adapted from (Salo, 2018).

In general, major and similar biogeochemical process capabilities appear to be present across all wells based on free energy flux calculations (Fig. 14) and from microbial taxa affiliation analyses, suggesting that similar functions are carried about between sites. However, these functions may be performed by different organisms specialized for their environment. The analyses indicated W12 contains specialists related to chemoautotrophs, methylotrophs, ferementers, sulfur-oxidizers (*Rhodocyclaceae*), sulfate reduction and nitrite oxidation (*Nitrospirae*), iron oxidation (*Gallionellaceae*), fermentation, and iron or methane metabolisms (*Woesarchaeota*). The analyses indicated the pristine background well W13 contains specialists potentially related to nitrite reduction, fermentation (*Parcubacteria*), iron oxidation (*Gallionellaceae*), sulfur oxidation (*Sulfurifustis*).

CHAPTER 3 –NUTRIENT EFFECTS

3.1 Background

Previous research found the RBI wells experienced river infiltration and pumped up to 40-60% river water (Fields-Sommers, 2015). Based on this previous research, it was expected the RBI wells have altered nutrient concentrations such as nitrogen and phosphorus from river water containing WWTP effluent. It is important to understand the microbial response and potential water quality implications from the WWTP effluent and river water infiltration.

Groundwater sources, in general, are becoming increasingly more contaminated due to anthropogenic activities such as WWTP processes and effluents, septic system leaks, agricultural fertilizer runoff, and manure runoff. For example, in northeastern Wisconsin, Kewaunee County faces extreme contamination in shallow groundwater consisting of microbial pathogens, viruses, and high nitrate levels from manure and human waste (Leland, 2017). Some residents are forced

to drink solely bottled water (Leland, 2017). Out of the 131 wells sampled, 40 wells had cow manure infiltration and 29 had human wastewater contamination from sources like livestock farms and septic systems (Zimmerman, 2017). Kewaunee County is known for having more cows than people, but is not the only place with a large amount of Concentrated Animal Feeding Operations (CAFOs). Wisconsin had a total of 289 Wisconsin Pollution Discharge Elimination System (WPDES) permits for CAFOs as of January 2017 (Wisconsin Department of Natural Resources, 2017). Contamination from agricultural/livestock practices and wastewater treatment is not isolated to Kewaunee Count. However, it may provide a good demonstration as to why it is essential to understand anthropogenic influences on nutrients in groundwater.

Large alterations of bioavailable nutrients can lead to drastic changes in nutrient cycling in aquatic ecosystems. Nitrogen and phosphorous are important major biogenic elements (in the list of: C, N, P, S, Si) necessary for microorganisms. Nutrients often are the limiting factor for productivity and the limiting nutrient is determined by the ratio. Redfield Ratio, (C:N:P) 106:16:1, is used to describe nutrient ratios in healthy aquatic systems needed for phytoplankton growth. It has been noted that microorganisms typically have more nitrogen and phosphorus content relative to their size than macroscopic organisms. Heterotrophic freshwater organisms typically have more N and more P than algae (Kirchman, 2012). This is attributed to N contributing to all proteins and nucleic acids, and P contributing to nucleic acids and phospholipids. Protein makes up the majority of microbial cells and is typically 55% of dry weight of bacteria. Hence, N is important for microbial growth as it is necessary for all proteins (Kirchman, 2012). It is important to understand how additional nitrogen, i.e. nitrate, may impact cell growth in the aquifer.

Major sources of nitrates include fertilizer runoff, septic tank leaks, sewage, and natural deposit from erosion (US EPA, 2018). Nitrate can mobilize into the subsurface and groundwater (Rivett et al., 2008). Increased nitrogen in groundwater can lead to: anoxia/hypoxia environments, altered biodiversity that could lead to a change in overall ecosystem function, changes in food-web structure, general habitat degradation, increase parasitic and harmful diseases, and increased acidification in freshwater ecosystems (Bernhard, 2010). High nitrate levels, over 10 mg/L (US EPA, 2018), can lead to serious illness or even death in infants younger than 6 months, including shortness of breath and blue-baby syndrome (WI DNR Bureau of Drinking Water & Groundwater, 2017).

Furthermore, nitrate is an important electron acceptor in groundwater affecting redox conditions. Native microbes facilitate most redox reactions in groundwater, but are typically specialized in their utilization of electron acceptors (Dinicola, 2006). Since microorganisms are the key drivers of the biogeochemical transformations in the nutrient cycles in groundwater, it's important to further characterize and understand their response to altered nutrient concentrations. In terms of water quality and health, treatment processes may need to be adjusted if the system is overcome with additional nutrients, electron donors/acceptors, and if the microbial communities can no longer biogeochemically process these compounds effectively or if they produce toxic or corrosive compounds (i.e. nitrate accumulation, hydrogen sulfide, nitrogen oxides, etc.) (Iribar et al., 2008; Vigneron et al., 2018).

3.2 Methods

Pristine groundwater from W13 was collected, and nitrate and phosphate were added to lab experiment bottles to understand microbial responses to altered nutrient conditions, i.e.

increased nitrate concentrations. 20 groundwater samples were collected in volumes of 1 L on 10/17/18 at well site W13. Samples were transferred to maintain low oxygen conditions by ensuring there was no head space and using rubber butyl stoppers with screw caps. All bottles, stoppers, and caps were acid washed for at least 4 hours in 5% HCl, rinsed with Type 1 (MilliQ) water three times, and rinsed with sample water 3 times prior to sample collection. All bottle surfaces were covered with aluminum foil prior to sampling to ensure a dark environment. Upon returning to SFS, samples were purged in lab for 1-4 hours with N_2 gas passing through the butyl stoppers with precision glide needles attached to 0.2 μm filters and an outlet needle to relieve pressure. It was verified visually that the needles were submerged in the water and actively bubbling and sparging the sample.

Initial nutrient measurements were performed on separately collected samples at the School of Freshwater Sciences (the Analytical Lab with the aid of Patrick Anderson and in Dr. Harvey Bootsma's Lab by Rae-Ann MacLellan-Hurd). The Auto Analyzer 3 HR SEAL was used to measure ammonium with the molybdate blue method, the Ion Chromatograph was used to measure nitrate, and SHIMADZU TOC-TN-L was used to measure TN. TDP was measured using photo-oxidation to break down dissolved organic phosphorus compounds into orthophosphate, which was then measured using the molybdate method as described in Appendix K. SRP was measured on filtered water with the molybdate method. All nutrient samples were pre-filtered with a 0.2 μm filter prior to analysis. Experiments were run in four replicates per condition for W13 water.

Based on the initial measurements just described, 16 1L bottle experiments were set up (10/18/18) at various N:P ratios: 4 bottles in replicate as controls, 4 bottles in replicate as an N:P of 500, 4 bottles in replicate as an N:P 50, 4 bottles in replicate as an N:P of 5. 4 initial 1 L bottle

samples were used for initial cell counts and filtering. Samples were spiked with nitrogen in the form of nitrate (from a stock of 1000ppm $NaNO_3$) up to 6 mg/L total nitrate in experiment bottles to mimic previously measured Fox River nitrate concentrations. In this way, the experiments could serve as an indicator of how the groundwater communities may respond if the groundwater nitrate concentrations ever reached those of the Fox River, the most likely nutrient constituent to mobilize from the river to aquifer. Samples were spiked with phosphorus in the form of phosphate (from stocks of 1000 ppm, 60 ppm, and 2.5 ppm NaH_2PO_4) to set the N:P ratios to 500, 50, and 5 to mimic WWTP effluent, the Fox River, and manure N:P ratios respectively. Sodium was chosen as the cation for both nutrient spikes to maintain consistency between the two factors and because sodium is already at relatively high levels in the groundwater and it would largely not affect sodium concentrations, unlike the cation potassium.

Upon spiking and setting up the experiments, initial samples (4 1 L bottles not treated with nutrient additons) were used to take initial cell counts using DAPI fluorescent stain and microscopy. After spiking, samples were placed in an incubator at 10° C to mimic groundwater temperature conditions and were gently mixing at the lowest setting of 50 RPM. Experiments were left to incubate for ~7.5 days.

After the incubations, final nutrient measurements were taken similarly to initial measurements as described previously. Approximately 50 mls of each sample was obtained and filtered through a 0.2 μm filter for nutrient measurements. The remaining volume was filtered sequentially through a 3 μm pre-filter onto a 0.1 μm filter in the inline filtration system as described previously. Samples were stored in the -80 °C freezer immediately after filtration. Filtrate from the 3 μm filter was obtained during filtration for cell enumeration with DAPI fluorescent stain.

The purpose of the experiment was to observe the effects of increased nitrate concentrations. N:P ratios were varied to mimic potential sources of contamination such as manure, WWTP effluent, and river contamination, but also to ensure the system was not P limited in order to see a response in increased nitrate levels. The nitrogen was held constant as to only adjust one parameter—phosphorus. Also, phosphorus in the form of phosphate is more likely to adsorb in the soil matrix or be used prior to reaching the aquifer, where nitrate is more likely to migrate into the groundwater. Samples were collected, sparged, and spiked in a random order.

3.3 Results & Discussion

The bottle experiments aimed to determine the effects of increased nitrate in the pristine groundwater well W13. Nitrate was spiked to the same concentration (6 mg/L) in all treatments except the control, and phosphate was spiked at various concentrations to ensure that the systems were not P limited in order to see a response in increased nitrate. The N:P ratios were chosen to mimic environmentally relevant ratios related to infiltration including manure, the Fox River, and WWTP effluent N:P ratios. These ratios can vary but were chosen as follows: a manure N:P ratio of 5:1, a Fox River N:P ratio of 50:1, and a WWTP effluent N:P ratio of 500:1. Manure ratios can range from 2:1 to 4:1(M. B. Vanotti et al., 2003; USDA, 2001), the Fox River N:P ratio was measured at approximately 48:1 (Salo, 2018), and WWTP effluent N:P ratios are highly variable depending on location ranging from around 5 to 900 (Waukesha City, 2013; Ferrell et al., 2014; Foster et al., 2016; Lambert et al., 2017; Peterson et al., 2012; Spiteri et al., 2007; Welskel and Howes, 1992). The pristine W13 N:P ratio was measured at approximately 200:1 at the time of the experiments.

Overall, there were minimal microbial effects from increased nitrate in these bottle experiments. Results are displayed in the tables below (Tables 11-16). The biological response measured in these experiments were cell counts and they were not significantly different between treatments and the Controls. The difference in nitrate (initial – final nitrate concentrations) did not significantly change between treatments. This result was not expected as the hypothesis was that with increased nitrate, the microbes would utilize more nitrate.

There were small changes in the removal of some of the nutrient pools. The TDP µg/L difference, or removal and loss of TDP, significantly differed from the difference in the Control results (post hoc Tukey test, $p = 0$). The percentage of the nitrate nutrient pool that was removed differed significantly in the 5 and 50 N:P ratio experiments as compared to the Control, but the difference in nitrate removal was not significantly different between treatments and the Control. The percentage of the SRP nutrient pool differed in the 50 N:P ratio compared to the Control, and the DOP nutrient pool percentage differed in the N:P ratio 5 compared to the Control. Cell counts were not significantly different between the N:P ratios and the Control. All final ammonium concentrations were below the detection limit. Only initial Total Nitrogen measurements were able to be measured because the TOC-TN-L machine used to measure TN was under repair at the end of the experiments.

Table 11: Ammonium Average Results from Nutrient Experiments

N:P Ratio	Initial [Ammonium] mg/L	Final [Ammonium] mg/L
Control	0	NA
5	0.0095	NA
50	0.0095	NA
500	0.0095	NA

Table 12: Nitrate Average Results from Nutrient Experiments

N:P Ratio	Initial [Nitrate] mg/L	Final [Nitrate] mg/L	STDEV	Difference [Nitrate] mg/L	Percent Remaining
Control	2	0.91	0.30	1.18	43%
5	6	4.88	0.52	1.12	81%
50	6	3.42	1.95	2.59	57%
500	6	4.08	1.43	1.92	68%

Table 13: Total Dissolved Phosphorus Average Results from Nutrient Experiment. The detection limit was 0.5 µg/L.

N:P Ratio	Initial [TDP] µg/L	Final [TDP] µg/L	STDEV	Difference [TDP] µg/L	Percent Remaining
Control	4	0.70	0.42	3.0	18%
5	537.7	247.8	20.62	247.8	46%
50	53.7	49.0	3.6	4.78	91%
500	5.0	0.7	0.47	1.0	13%

Table 14: Soluble Reactive Phosphorus Average Results from Nutrient Experiments. The detection limit was 0.5 µg/L.

N:P Ratio	Initial [SRP] µg/L	Final [SRP] µg/L	STDEV	Difference [SRP] µg/L	Percent Remaining
Control	3.0	<0.5	0.2	2.6	14%
5	537.5	160.8	9.8	376.7	30%
50	53.0	47.2	4.9	6.0	88%
500	4.9	<0.5	<0.5	4.6	7%

Table 15: N:P Average Ratio Results from Nutrient Experiments. The initial measurements were based off of TDP µmol/L from the Bootsma Lab's measurements and TN µmol/L from the TOC-TN L machine in the Analytics Lab. The final ratios were derived from TDP µmol/L and Nitrate concentrations.

N:P Ratio	Initial N:P Measured	Final N:P Measured	STDEV	Difference µmol/L
500	500.0	893.3	667.4	-393.3
Control	184.7	5334.7	4949.1	-5150.0
5	5.0	7.1	4.2	-2.1
50	50.0	42.0	15.9	8.0

Table 16: Cell Count Results from DAPI Fluorescent Stain and Microscopy for Nutrient Experiments.

Sample NL Number	N:P Ratio	Initial Cell Counts cells/ml	Average Cell Count cells/ml
647	Control	2.10E+05	2.57E+05
636	5	2.10E+05	6.35E+05
638	50	2.10E+05	7.13E+05
642	500	2.10E+05	7.82E+04

Since the nitrate removal difference (initial – difference nitrate concentration) was similar across treatments and the control, the experiments could have been electron donor limited, i.e. carbon (C) limited. We were working under the assumption C was available in the form of DOC or CO_2. However, referring back to Table 3, the DOC was relatively low in W13 (DOC from Table 3: 0.93 ± 0.37 mg/L in W13). Then, referring to reaction 2 in Table 4 (Acetate + NO_3^- + $H_2O \rightarrow 2HCO_3^-$ + NH_3) assuming it is a predominant metabolism occurring, which thermodynamic calculations suggest, then the lab bottle experiments would be carbon limited and only allow up to ~2.4 mg/L of nitrate to be utilized as shown in the equations:

$$DOC \Rightarrow Acetate: \frac{0.93\ mg\ DOC}{L} * \frac{1\ mol\ C}{12\ g\ C * \frac{2C}{Acetate}} * \frac{1g}{1000mg} = \frac{3.875E^{-05}mol}{L} Acetate$$

$$NO_3: \frac{6\ mg}{L} * \frac{1\ mol\ NO_3}{62\ g\ NO_3} * \frac{1\ g}{1000\ mg} = \frac{9.7E^{-05}mol}{L}\ NO_3$$

Here, Acetate would be limiting because it is less than nitrate based on molar coefficients from reaction 2 in Table 4 (Acetate + NO_3^- + $H_2O \rightarrow 2HCO_3^-$ + NH_3).

$$\frac{3.87\ E^{-05}mol}{L} Acetate * \frac{1molNO_3}{1molAcetate} * \frac{62gNO_3}{molNO_3} = 0.0024gNO_3 = 2.4\frac{mg}{L} NO_3$$

This could explain the consistent difference in nitrate loss between all bottle experiments. Assuming acetate was the predominant carbon source and equal to previous DOC measurements, only approximately 2.4 mg/L of nitrate could be utilized regardless of nitrogen or phosphate availability due to a C limitation. If experiments were repeated, DOC should also be measured. Also, more time may be necessary in order to see a full response from the microbial community in the pristine groundwater W13 system.

However, C can be available in a vast variety of useful forms for microorganisms and some microbes are more specialized for certain sources. If we selected a specific carbon source, this could select for specific microbial community members and processes. We intended to mitigate multiple variable effects. We aimed to observe increased nitrate effects in the non-river infiltrated well by only increasing the nitrate concentration and varying the N:P ratio to allow a response to occur without P being a limiting factor. However, if most carbon was not dissolved in the water, but is in the sediment or is particle attached, then the water incubations could have eliminated primary C sources and thus developed a C limitation.

Although there was little response in these experiments likely due to Carbon limitation, this may not be representative of natural conditions if increased nitrate became prevalent. There are likely more C sources in the sediment of the groundwater which the microbes could utilize. With this in mind, W13 could also not be primed for nitrate use as well as W12. Nitrate levels had been measured in W13 at relatively higher concentrations compared to W12. This could indicate that W12 microbes are able to utilize nitrate and therefore it is removed from the W12 environment. This is probable in the sense that W12 receives infiltration from the river and would likely have incoming nitrate from the river and WWTP effluent. W13 may contain

microbes not primed for as much nitrate utilization, and a week-long experiment may not have allowed the community to shift or adapt to this altered environment.

With that said, additional nitrate entering the wells could initially result in an accumulation of nitrate in the wells. This experiment showed that environmentally relevant concentrations of nitrate and N:P ratios could affect the wells and lead to nitrate accumulation, which can affect water management practices. The background well W13 may not be primed to handle additional nitrate initially. However, overtime, the aquifer microbial community may be able to shift and adjust to altered conditions as in W12, assuming nitrate is entering W12 and being utilized immediately by the present microbial consortia.

If these experiments were repeated, more information could be determined about this aquifer system and potential effects of nitrate accumulation in the wells. DOC should be measured and if the experiments are C limited, experiments should be run again with a carbon source addition such as acetate. Also, performing experiments using both W13 and W12 water could indicate if the microbial communities are primed for nitrate utilization in W12 and not W13.

CONCLUSIONS

Overall, microbial community compositions were clearly distinct and significantly different between sites. The Fox River and groundwater microbial community compositions were significantly different. The groundwater sites differed significantly both in the ion and nutrient chemistry composition as a whole and in microbial community composition. TDP and TN were significant drivers in community differences in the groundwater microbial communities. Overall, the chemistry compositions and microbial community compositions between groundwater sites were significantly different.

This study found evidence of indicator river bacteria ASVs in all of the groundwater wells, however these ASVs were always more abundant in the RBI well W12 with similar distribution trends as the Fox River. These data suggests that river microbes may be entering the aquifer through riverbank inducement, and can potentially then transfer through the aquifer system. However, the river bacteria indicator ASVs' distribution in the background well W13 were not similar to the patterns of distribution in the river, indicating that potentially these ASVs could be from different organisms entirely but their 16S rRNA genes diverged and evolved similarly.

Based off V4 16S rRNA gene sequence data from microbial RNA and DNA, most groundwater microbial community members were Unclassified and those consistent with recently discovered microorganisms from the CPR and DPANN superphyla. Having such a large proportion (~55%) of the groundwater microbial dataset as Unclassified is not common for most microbial environments, indicating a vast amount of knowledge is still to be learned from these largely unexplored groundwater systems. This study also found evidence of the rarely recorded bacterial and archaeal feature of intron splicing in Unclassified groundwater microbial 16S rRNA genes from RNA sequences that had absent corresponding DNA sequences, which is a relatively novel discovery in the field in recent years (Brown et al., 2015; Castelle et al., 2018). This finding suggests that typical 16S rRNA gene surveys only targeting the DNA fail to capture all 16S rRNA phylogenetic markers from an environment.

The taxa associated with the groundwater microbial communities in this dataset suggest anaerobic metabolisms involving nitrate and sulfate reduction with fermentation, sulfur oxidation, H_2 oxidation, nitrite oxidation, and iron oxidation capabilities. This was consistent with thermodynamic free energy flux results from the groundwater wells. Based on free energy

flux calculations, favorable biogeochemical reactions included heterotrophic and fermentative metabolisms in the three shallow groundwater wells. Among the heterotrophic reactions, nitrate was found to be the primary electron acceptor. 3 of the 4 dominant reactions included heterotrophic metabolisms with nitrate reduction and sulfate reduction, and the fourth dominant metabolism was fermentation.

In general, major biogeochemical process capabilities appeared to be present across all wells based on free energy flux calculations and from microbial taxa affiliation analyses, suggesting that similar functions are carried about between sites. However, these functions may be performed by different organisms specialized for their environment. The analyses indicated W12 contains specialists related to chemoautotrophs, methylotrophs, ferementers, sulfur-oxidizers (*Rhodocyclaceae*), sulfate reduction and nitrite oxidation (*Nitrospirae*), iron oxidation (*Gallionellaceae*), fermentation, and iron or methane metabolisms (*Woesarchaeota*). The analyses indicated the pristine background well W13 contains specialists potentially related to nitrite reduction, fermentation (*Parcubacteria*), iron oxidation (*Gallionellaceae*), sulfur oxidation (*Sulfurifustis*).

Nutrient experiments of spiked nitrate additions in non-river infiltrated groundwater suggest that nitrate accumulation may initially occur under these circumstances. Environmentally relevant concentrations of nitrate and N:P ratios could affect the groundwater wells and lead to nitrate accumulation, which could affect water management practices. However, the nutrient experiments were carbon limited and should be performed again measuring for C and spiking with C as necessary. Nutrient experiments should also be performed again with pristine W13 groundwater and RBI well W12 groundwater to determine if W12 is primed for nitrate utilization and if W13 is not, which could result in nitrate accumulation in the wells. Further investigation

of the metabolic capabilities of the groundwater microorganisms is necessary as much is still unknown in these subsurface environments and the associated microbial consortia.

CHAPTER 4 – FUTURE RECOMMENDATIONS

This study demonstrated the vast amount of unknown microbes and their unknown functions in a shallow groundwater aquifer. This study also indicated, if disturbed or impacted, previously non-river infiltrated groundwater could initially accumulate environmentally relevant concentrations of nitrate. Accumulation of nitrate or other foreign constituents could impact water quality, water management practices and human health. It could be important to learn more about the unknown microbes in groundwater, their functional potential, and their ability to handle altered environmental conditions.

4.1 RATIONALE

Subsurface microorganisms utilize, as well as generate, biogeochemical gradients. These microbial metabolisms and transformations affect changes to important biogeochemical species, impacting things like water quality and human health (i.e. harmful levels of nitrate in groundwater). Specific microbial community members have been identified and associated with specific transformations in nutrient and biogeochemical cycles like in carbon, nitrogen, and sulfur cycles in the terrestrial subsurface through genome resolution (Brown et al., 2015; Long et al., 2016). Furthermore, due to selection under altered environmental conditions, microbial metabolic capabilities and steps in major biogeochemical cycles can be altered (Anantharaman et al., 2016). When dominant reactions are determined, they can be incorporated in reaction networks and models to better understand life strategies and overall functioning. Knowledge gained from pairing geochemical data with potential functional capabilities found in groundwater

metagenomes can aid in understanding and managing different environmental impacts and perturbations (Long et al., 2016).

Standard practices to study microorganisms long depended on lab culturing techniques. Incomplete knowledge and understanding of microbial distribution, diversity, composition, and functional capabilities were greatly due to inherent limitations of standard culturing practices, as lab conditions are not conducive for the growth and recovery of most microorganisms (Delmont & Eren, 2016). Now, the capability to recover genetic material from a natural environment for sequencing provides far broader knowledge of the diversity and composition of all the microorganisms present in a system and their functional capabilities from a specific environment (Delmont & Eren, 2016; Sharon & Banfield, 2013; Vanwonterghem, Jensen, Rabaey, & Tyson, 2016). This method, termed metagenomics, does not rely on cultivation and allows for further investigation of the physiological prediction of microbial environmental roles (Sharon & Banfield, 2013b). In short, biomass is recovered from the environment from its natural setting and the DNA is extracted and sequenced, skipping PCR bias. The subsequent data is then analyzed utilizing computational tools. In this way, this metagenomics method tackles two specifically important needs: the capability of analyzing 99% of microbes in nature which have not been cultivated and enables the study of microbes in the context of their natural environment and community (Eisen, 2011; Handelsman, Tiedje, National Research Council (US) Committee on Metagenomics: Challenges and Functional, & Applications, 2007; Sharon & Banfield, 2013a).

In recent years, sequence studies have provided insight into the magnitude and variety of little known to unknown microbes that mediate nutrient cycling in subsurface environments (Anantharaman et al., 2016). In other environments, associating 16S rRNA gene sequence data

with reference genomes has been used to assume functions for specific microorganisms. This practice is not applicable in groundwater systems because they lack reference genomes and contain microbes that evade detection through standard 16S rRNA gene surveys as mentioned previously (Brown et al., 2015; Long et al., 2016). There's a need to discover reference genomes in groundwater through metagenomics to extensively study subsurface microbial communities and subsurface biogeochemical processes to fill this gap of knowledge (Long et al., 2016). Metagenome-assembled genomes (MAGs) can be used to study microbial systems where researchers depend on functional annotations, phylogenetically conserved gene families, or distribution patterns to determine metabolic pathways, evolved relationships, and/or co-occurrence of genomic collections (Delmont & Eren, 2016). Establishing groundwater metagenomes will be essential for understating potential and possible subsurface metabolic reactions. Groundwater metagenomes will lay the backbone for metatranscriptomic and proteomic data which would show the actual metabolic functions occurring at a given time or under specific conditions (Long et al., 2016).

Recent studies also suggest the importance of metabolic byproduct exchanges between interacting organisms in groundwater. It appears many of the novel and recently discovered microbes in groundwater show limited metabolic capabilities and do not singly contain machinery to carry out sequential redox transformations (Anantharaman et al., 2016; Castelle et al., 2015a; Eme & Doolittle, 2015; Liu et al., 2018; Long et al., 2016; Probst et al., 2018; Waters et al., 2003) and have ultra-small cell and genome sizes (Luef et al., 2015). Due to the expected syntrophic behavior of many groundwater microorganisms, it is important to survey and examine all possible genes and pathways across the entire ecosystem and metagenome to fully understand all possible metabolic and biogeochemical capabilities. Then, comparing the metagenomes from

the three groundwater sites will indicate metabolic capabilities at each location and indicate any possible biogeochemical reactions or metabolic differences related to river water and WWTP effluent infiltration.

4.2 RECOMMENDATIONS

- Determine metagenomes, functional potential, and geochemical reactivity in pristine and WWTP/river-infiltrated portions of a shallow sand and gravel groundwater aquifer and determine patterns related to infiltration and geochemistry.
- Once metagenomes are determined, determine metatranscriptomes through RNAseq to further understand occurring functions in the systems. Also, pair proteomic samples with metagenomic samples to further understand occurring microbial functions in the groundwater.
- Investigate temporal affects associated with chemical data and also microbial community shifts.
- Perform nutrient experiments again using W12 water to determine any response differences specifically related to nitrogen in a system that, in theory, is primed for nitrate utilization.
- Perform nutrient experiments with W13 water again with the addition of a carbon source, such as acetate.

REFERENCES

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

Apprill, A., Mcnally, S., Parsons, R., Weber, L., 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat. Microb. Ecol. https://doi.org/10.3354/ame01753

Atashgahi, S., Aydin, R., Dimitrov, M.R., Sipkema, D., Hamonts, K., Lahti, L., Maphosa, F., Kruse, T., Saccenti, E., Springael, D., Dejonghe, W., Smidt, H., 2015. Impact of a wastewater treatment plant on microbial community composition and function in a hyporheic zone of a eutrophic river. Sci. Rep. 5. https://doi.org/10.1038/srep17284

Bernhard, A., 2010. The Nitrogen Cycle: Processes, Players, and Human Impact. Nat. Educ. Knowl. 3, 25.

Boden, R., Hutt, L.P., Rae, A.W., 2017. Reclassification of Thiobacillus aquaesulis (Wood & Kelly, 1995) as Annwoodia aquaesulis gen. nov., comb. nov., transfer of Thiobacillus (Beijerinck, 1904) from the Hydrogenophilales to the Nitrosomonadales, proposal of Hydrogenophilalia class. nov. withi. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.001927

Bradley, P.M., Barber, L.B., Duris, J.W., Foreman, W.T., Furlong, E.T., Hubbard, L.E., Hutchinson, K.J., Keefe, S.H., Kolpin, D.W., 2014. Riverbank filtration potential of pharmaceuticals in a wastewater-impacted stream. Environ. Pollut.

Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., Wilkins, M.J., Wrighton, K.C., Williams, K.H., Banfield, J.F., 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523, 208–211. https://doi.org/10.1038/nature14486

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13, 581–583. https://doi.org/10.1038/nmeth.3869

Castelle, C.J., Brown, C.T., Anantharaman, K., Probst, A.J., Huang, R.H., Banfield, J.F., 2018. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. Nat. Rev. Microbiol. https://doi.org/10.1038/s41579-018-0076-2

Castelle, C.J., Wrighton, K.C., Thomas, B.C., Hug, L.A., Brown, C.T., Wilkins, M.J., Frischkorn, K.R., Tringe, S.G., Singh, A., Markillie, L.M., Taylor, R.C., Williams, K.H., Banfield, J.F., 2015. Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. Curr. Biol.

https://doi.org/10.1016/j.cub.2015.01.014

Chazdon, R.L., Chao, A., Colwell, R.K., Lin, S.Y., Norden, N., Letcher, S.G., Clark, D.B., Finegan, B., Arroyo, J.P., 2011. A novel statistical method for classifying habitat generalists and specialists. Ecology. https://doi.org/10.1890/10-1345.1

CH2M HILL. (2013, October). Application Summary, City of Waukesha Application for a Lake Michigan Diversion with Return Flow; Volume 1 of 5 . Retrieved June 2018, from http://www.waukesha-water.com/downloads/1_City_of_Waukesha_Application__Summary.pdf

Davidson, M., Silver, B., Onstott, T., Moser, D., Gihring, T., Pratt, L., Boice, E., Lollar, B., Lippmann-Pipke, J., Pfiffner, S., Kieft, T., Seymore, W., and Ralston, C., 2011, Capture of planktonic microbial diversity in fractures by long-term monitoring of flowing boreholes, Evander Basin, South Africa: Geomicrobiology Journal, v. 28, p. 275–300.

Denef, V.J., Fujimoto, M., Berry, M.A., Schmidt, M.L., 2016. Seasonal succession leads to habitat-dependent differentiation in ribosomal RNA:DNA ratios among freshwater lake bacteria. Front. Microbiol. 7. https://doi.org/10.3389/fmicb.2016.00606

Dinicola, R. (2006). U.S. GEOLOGICAL SURVEY Scientific Investigations Report 2006-5056: Continued Biodegradation of Chloroethene Compounds in Ground Water at Operable Unit 1, Naval Undersea Warfare Center, Division Keyport, Washington: Redox Conditions in Contaminated Ground Water. Retrieved from USGS: https://pubs.usgs.gov/sir/2006/5056/

Di Rienzi, S.C., Sharon, I., Wrighton, K.C., Koren, O., Hug, L.A., Thomas, B.C., Goodrich, J.K., Bell, J.T., Spector, T.D., Banfield, J.F., Ley, R.E., 2013. The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. Elife. https://doi.org/10.7554/eLife.01102.001

DNR, 2017. Groundwater. Retrieved from Wisconsin Department of Natural Resources: https://dnr.wi.gov/topic/groundwater/

Eme, L., Ford Doolittle, W., 2015. Microbial diversity: A bonanza of phyla. Curr. Biol. 25, R227–R230. https://doi.org/10.1016/j.cub.2014.12.044

Fields-Sommers, L.R., 2015. Assessing the E ects of Riverbank Inducement on a Shallow Aquifer in Southeastern Wisconsin.

Ferrell, G.M., Yearout, M., Grimes, B., Graves, A., Fitzgerald, S., Meyer, M., 2014. Water-Quality Characteristics Indicative of Wastewater in Selected Streams in the Upper Neuse River Basin , Durham and Orange Counties , North Carolina , from 2004 to 2013.

Foster, G.M., Graham, J.L., Williams, T.J., King, L.R., 2016. Spatiotemporal Variability of Inorganic Nutrients during Wastewater Effluent Dominated Streamflow Conditions in Indian Creek , Johnson County , Kansas , 2012 – 15.

Göker, M., Teshima, H., Lapidus, A., Nolan, M., Lucas, S., Hammon, N., Deshpande, S., Cheng, J.F., Tapia, R., Han, C., Goodwin, L., Pitluck, S., Huntemann, M., Liolios, K., Ivanova, N., Pagani, I., Mavromatis, K., Ovchinikova, G., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Brambilla, E.M., Rohde, M., Spring, S., Detter, J.C., Woyke, T., Bristow, J., Eisen, J.A., Markowitz, V., Hugenholtz, P., Kyrpides, N.C., Klenk, H.P., 2011. Complete genome sequence of the acetate-degrading sulfate reducer Desulfobacca acetoxidans type strain (ASRB2T). Stand. Genomic Sci. https://doi.org/10.4056/sigs.2064705

Grundl, T. (2018). (N. Gayner, Personal Communication)

Grundl, T., Cape, M., 2006. Geochemical factors controlling radium activity in a sandstone aquifer. Ground Water 44, 518–527. https://doi.org/10.1111/j.1745-6584.2006.00162.x

Hemme, C.L., Tu, Q., Shi, Z., Qin, Y., Gao, W., Deng, Y., Van Nostrand, J.D., Wu, L., He, Z., Chain, P.S.G., Tringe, S.G., Fields, M.W., Rubin, E.M., Tiedje, J.M., Hazen, T.C., Arkin, A.P., Zhou, J., 2015. Comparative metagenomics reveals impact of contaminants on groundwater microbiomes. Front. Microbiol. 6, 1–12. https://doi.org/10.3389/fmicb.2015.01205

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D.A., Finstad, K.M., Amundson, R., Thomas, B.C., Banfield, J.F., 2016. A new view of the tree of life. Nat. Microbiol. 1. https://doi.org/10.1038/nmicrobiol.2016.48

Hug, L.A., Thomas, B.C., Sharon, I., Brown, C.T., Sharma, R., Hettich, R.L., Wilkins, M.J., Williams, K.H., Singh, A., Banfield, J.F., 2016. Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. Environ. Microbiol. 18. https://doi.org/10.1111/1462-2920.12930

Iribar, A., Sánchez-Pérez, J.M., Lyautey, E., Garabétian, F., 2008. Differentiated free-living and sediment-attached bacterial community structure inside and outside denitrification hotspots in the river-groundwater interface. Hydrobiologia 598, 109–121. https://doi.org/10.1007/s10750-007-9143-9

Kirchman, D. L. (2012). *Processes in Microbial Ecology.* Oxford, NY, USA: Oxford University Press Inc., New York.

Kojima, H., Shinohara, A., Fukui, M., 2015. Sulfurifustis variabilis gen. nov., sp. nov., a sulfur oxidizer isolated from a lake, and proposal of Acidiferrobacteraceae fam. nov. and Acidiferrobacterales ord. nov. Int. J. Syst. Evol. Microbiol.

https://doi.org/10.1099/ijsem.0.000479

Lambert, R., Opsahl, S., Musgrove, M., 2017. Water Quality, Sources of Nitrate, and Chemical Loadings in the Geronimo Creek and Plum Creek Watersheds, South-Central Texas, April 2015–March 2016.

Leland, M. (2017, May 10). *FOX 11 Investigates Kewaunee County's groundwater contamination.* Retrieved from FOX 11 News: http://fox11online.com/news/fox-11-investigates/fox11-investigates-kewaunee-countys-groundwater-contamination

Lisle, J.T. 2014. Survival of Bacterial Indicators and the Functional Diversity of Native Microbial Communities in the Floridan Aquifer System, South Florida: U.S. Geological Survey Open–File Report 2014-1011, 70 p., http://dx.doi.org/10.3133/ofr20141011

Liu, X., Li, M., Castelle, C.J., Probst, A.J., Zhou, Z., Pan, J., Liu, Y., Banfield, J.F., Gu, J.-D., 2018. Insights into the ecology, evolution, and metabolism of the widespread Woesearchaeotal lineages. Microbiome 6, 102. https://doi.org/10.1186/s40168-018-0488-2

Long, P.E., Williams, K.H., Hubbard, S.S., Banfield, J.F., 2016. Microbial Metagenomics Reveals Climate-Relevant Subsurface Biogeochemical Processes. Trends Microbiol. https://doi.org/10.1016/j.tim.2016.04.006

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. https://doi.org/10.14806/ej.17.1.200

Maupin, M.A., Kenny, J.F., 2010. Estimated Use of Water in the United States in 2010 Circular 1405, US Geological Survey. https://doi.org/http://dx.doi.org/10.3133/cir1405

McCarthy, A., Chiang, E., Schmidt, M.L., Denef, V.J., 2015. RNA preservation agents and nucleic acid extraction method bias perceived bacterial community composition. PLoS One. https://doi.org/10.1371/journal.pone.0121659

McGinley, P.M., DeVita, W. D., Nitka, A.L., 2015. Evaluating Chemical Tracers in Suburban Groundwater as Indicators of Nitrate-Nitrogen Sources. A Final Report Prepared for the Wisconsin Department of Natural Resources. College of Natural Resources University of Wisconsin-Stevens Point, Center for Watershed Science and Education.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2013. Package 'vegan.' R Packag. ver. 2.0–8 254. https://doi.org/10.4135/9781412971874.n145

Oren, A., 2014. The family Rhodocyclaceae, in: The Prokaryotes: Alphaproteobacteria and Betaproteobacteria. https://doi.org/10.1007/978-3-642-30197-1_292

Oude Elferink, S.J.W.H., Akkermans-van Vliet, W.M., Bogte, J.J., Stams, A.J.M., 1999. Desulfobacca acetoxidans gen. nov., sp. nov., a novel acetate-degrading sulfate reducer isolated from sulfidogenic granular sludge. Int. J. Syst. Bacteriol. https://doi.org/10.1099/00207713-49-2-345

Pace, N.R., Olsen, G.J., Woese, C.R., 1986. Ribosomal RNA phylogeny and the primary lines of evolutionary descent. Cell. https://doi.org/10.1016/0092-8674(86)90315-6

Parada, A.E., Needham, D.M., Fuhrman, J.A., 2016. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ. Microbiol. 18, 1403–1414. https://doi.org/10.1111/1462-2920.13023

Peterson, B.J., Stubler, A.D., Wall, C.C., Gobler, C.J., 2012. Nitrogen-rich groundwater intrusion affects productivity, but not herbivory, of the tropical seagrass thalassia testudinum. Aquat. Biol. https://doi.org/10.3354/ab00413

Porter, K.G., Feig, Y.S., 1980. The use of DAPI for identifying and counting aquatic microflora. Limnol. Oceanogr. https://doi.org/10.4319/lo.1980.25.5.0943

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41. https://doi.org/10.1093/nar/gks1219

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. https://doi.org/10.1007/978-3-540-74686-7

Ray, C., Grischek, T., Schubert, J., Wang, J.Z., Speth, T.F., 2002. A perspective of riverbank filtration. J. / Am. Water Work. Assoc. https://doi.org/10.1073/pnas.0601738103

Ray, C., 2008. Worldwide potential of riverbank filtration. Clean Technol. Environ. Policy. https://doi.org/10.1007/s10098-008-0164-5

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A., Dodsworth, J.A., Hedlund, B.P., Tsiamis, G., Sievert, S.M., Liu, W.T., Eisen, J.A., Hallam, S.J., Kyrpides, N.C., Stepanauskas, R., Rubin, E.M., Hugenholtz, P., Woyke, T., 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature. https://doi.org/10.1038/nature12352

Rivett, M.O., Buss, S.R., Morgan, P., Smith, J.W.N., Bemment, C.D., 2008. Nitrate attenuation in groundwater: A review of biogeochemical controlling processes. Water Res. https://doi.org/10.1016/j.watres.2008.07.020

Salo, M., 2018. Anthropogenically Driven Changes to Shallow Groundwater in Southeastern Wisconsin and it's Effect on the Aquifer Metagenome

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski,

R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541. https://doi.org/10.1128/AEM.01541-09

Shamrukh, M., Abdel-Wahab, A., 2008. Riverbank filtration for sustainable water supply: Application to a large-scale facility on the Nile River. Clean Technol. Environ. Policy. https://doi.org/10.1007/s10098-007-0143-2

Singh, P., Kumar, P., Mehrotra, I., Grischek, T., 2010. Impact of riverbank filtration on treatment of polluted river water. J. Environ. Manage.

Smith, R.J., Jeffries, T.C., Roudnew, B., Fitch, A.J., Seymour, J.R., Delpin, M.W., Newton, K., Brown, M.H., Mitchell, J.G., 2012. Metagenomic comparison of microbial communities inhabiting confined and unconfined aquifer ecosystems. Environ. Microbiol. 14, 240–253.

Soo, R.M., Hemp, J., Parks, D.H., Fischer, W.W., Hugenholtz, P., 2017. On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. Science (80-.). https://doi.org/10.1126/science.aal3794

Spiteri, C., Slomp, C.P., Regnier, P., Meile, C., Van Cappellen, P., 2007. Modelling the geochemical fate and transport of wastewater-derived phosphorus in contrasting groundwater systems. J. Contam. Hydrol. https://doi.org/10.1016/j.jconhyd.2007.01.002

Sun, B., Cole, J.R., Tiedje, J.M., 2001. Desulfomonile limimaris sp. nov., an anaerobic dehalogenating bacterium from marine sediments. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/00207713-51-2-365

Tonolla, M., Demarta, A., Peduzzi, S., Hahn, D., Peduzzi, R., 2000. In situ analysis of sulfate-reducing bacteria related to Desulfocapsa thiozymogenes in the chemocline of meromictic Lake Cadagno (Switzerland). Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.66.2.820-824.2000

USDA, 2001. Comprehensive nutrient managemetn planning technical guidance. Washington, D.C. USDA Natural Resources Conservation Service.

US EPA, 2018. *Ground Water and Drinking Water: National Primary Drinking Water Regulations.* Retrieved from EPA: United States Environmental Protection Agency: https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations

USGS, 2017. Groundwater use in the United States. Retrieved March 26, 2018, from The USGS Water Science School: https://water.usgs.gov/edu/wugw.html

Vanotti, M. B. , Szogi, A. A., Hunt, P. G., 2003. Extraction of Soluble Phosphorus From Swine

Wastewater. Trans. ASAE. https://doi.org/10.13031/2013.15637

Vigneron, A., Cruaud, P., Alsop, E., de Rezende, J.R., Head, I.M., Tsesmetzis, N., 2018. Beyond the tip of the iceberg; a new view of the diversity of sulfite- and sulfate-reducing microorganisms. ISME J. 1–4. https://doi.org/10.1038/s41396-018-0155-4

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., Apprill, A., Knight, R., 2016. Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems 1, e00009-15. https://doi.org/10.1128/mSystems.00009-15

Waukesha City, 2013. Wisconsin Pollutant Discharge Elimination System (WPDES) Wastewater Discharge Individual Permit Application, WPDES Permit Number 0029971-09-0, FID # 268005100. Waukesha City.

Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E.R., Knight, R., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome. https://doi.org/10.1186/s40168-017-0237-y

Weiss, W.J., Bouwer, E.J., Aboytes, R., LeChevallier, M.W., O'Melia, C.R., Le, B.T., Schwab, K.J., 2005. Riverbank filtration for control of microorganisms: Results from field monitoring. Water Res. https://doi.org/10.1016/j.watres.2005.03.018

Welskel, P.K., Howes, B.L., 1992. Differential Transport of Sewage-Derived Nitrogen and Phosphorus through a Coastal Watershed. Environ. Sci. Technol. https://doi.org/10.1021/es00026a017

WI DNR Bureau of Drinking Water & Groundwater., (2017). *Nitrate in Drinking Water.* Retrieved from Nitrate in Drinking Water DG001 2017 web: https://dnr.wi.gov/files/PDF/pubs/DG/DG0001.pdf

Willis, A., 2017. Rarefaction, alpha diversity, and statistics. bioRxiv 231878. https://doi.org/10.1101/231878

Wisconsin Department of Natural Resources, 2017. *WISCONSIN WPDES PERMITTED CONCENTRATED ANIMAL FEEDING OPERATIONS.* Retrieved from DNR WI: https://dnr.wi.gov/topic/AgBusiness/documents/cafo_statewide_map.pdf

Zimmerman, J., 2017. *Study results released on Kewaunee County water contamination.* Retrieved from ABC 2 WABY : http://www.wbay.com/content/news/Study-results-released-on-Kewaunee-County-water-contamination--427127793.html

APPENDICES

APPENDIX A: SAMPLING SITE COORDINATES

Table 17: Coordinates in Decimal Degree for Describe Sample Sites

Sampling Name	Unique Site ID	Latitude	Longitude
W11	RL255	42.959938	-88.279256
W12	RL256	42.961012	-88.279063
W13	WK947	42.961236	-88.289167
FR	Fox 2	42.977690	-88.264797
WK WWTP	Waukesha WWTP	42.998190	-88.249151

APPENDIX B: DETAILED LAB METHOD AND PROTOCOLS

DNA/RNA Extraction for 16S rRNA Gene Sequencing

The 0.2 and 0.1 μm pore size filters will be cut within their collection tubes using sterilized small dissection scissors. DNA and RNA will simultaneously be extracted from the same sample using Qiagen's AllPrep Powerviral DNA/RNA kit. Zirconian beads will be added to the sample tubes after cutting the filters into small pieces and vortexed in a bead beater for 2 minutes and 30 seconds and placed on ice for 5 minutes and repeated for a total of 2 bead beating (Smith et al., 2012) steps. The extraction will be followed according to manufacturer's instructions except for elution in three volumes up to 100ul. Extracted DNA and RNA will be in the same sample tube and stored in the -80 °C freezer.

DNase Treatment

Promega's RQ1 RNase-Free DNase (Cat #M6101) kit for DNase Treatment of RNA Samples Prior to RT-PCR was used. Promega's DNase protocol calls for 1-8ul of RNA sample in elution buffer. Since the groundwater samples are low biomass systems with low nucleic acid yields, the full 8 μl of sample was used. A total of 16 μl of RNA sample were used per sample per DNase treatment in order to have enough for two reactions (one positive with reverse transcriptase, and one negative without reverse transcriptase) in subsequent steps in the RT-PCR. 1 μl of DNase (the protocol calls for 1 unit or μl per 1 ug of RNA), and 2 μl of Buffer were used in the 16 μl sample reactions.

Reverse Transcription PCR (RT-PCR)

RNA was reverse transcribed using the Promega's GoScript™ Reverse Transcription System. The reverse primer 806Rb for the v4 16S rRNA gene region will be used in the cDNA synthesis (806Rb – GGACTACNVGGGTWTCTAAT). 8.5 µl of DNase treated RNA was used with 1.5 µl of primer for each reaction. Each sample had two reactions: one positive reaction including the reverse transcriptase and one negative reaction without the reverse transcriptase. The two reactions were used to ensure the DNase treatment worked correctly and no carryover-over of initial DNA remained in the RNA/cDNA sample for the subsequent 16S rRNA gene PCR.

v4 16S rRNA gene Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) was used to target and amplify the v4 16S rRNA gene region in the DNA and cDNA samples using Invitrogen's™ Platinum™ Taq DNA Polymerase. Forward and reverse primers 515Fb and 806Rb with Illumina adapters were used:

515Fb-illumina –

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA

806Rb-illumina –

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACNVGGGTWTCTAAT

Reactions were ran in triplicate for each sample and then pooled prior to cleanup with the AMPure Bead cleanup system. One PCR reaction out of the three triplicates for each sample will be screened using gel electrophoresis to verify amplification and DNA fragment size. A modified reconditioned/nested PCR protocol will be used when one normal PCR (25ul reaction volume, 1 µl template, 30 cycles) is not sufficient for sample amplification. In the reconditioned PCR, two

consecutive PCR's will be carried out. The first PCR will have a smaller reaction volume with and shorter cycle period but still 1 µl of template will be use (15ul reaction volume, 10 cycles, 1 µl template). Then 1 µl of the reconditioned PCR will be used as template in the full PCR (25 µl reaction volume, 30 cycles, 1 µl of template). A negative control was run each time. Reaction components, volumes, and concentrations are described in the table below.

Table 18: PCR Mastermix Conditions

Master Mix of PCR Components	Working Concentration	Normal PCR	Reconditioned PCR
Reaction Volume	-	25	15
PCR Cycles	-	30	10
10x Buffer for Platinum Taq	10x	2.5 µL	1.5 µL
F Primer	5 uM stock	1 µL	0.6 µL
R Primer	5 uM stock	1 µL	0.6 µL
50 mM $MgSO_4$	50 uM	1 µL	0.6 µL
10 mM dNTP Mix	10mM	0.5 µL	0.3 µL
Platinum Taq Polymerase	5 U/µL	0.1 µL	0.06 µL

Table 19: PCR Thermocycler Conditions

PCR Thermocycler Conditions			
1	Initial denaturation	94°C	5 minutes
2	Denature	94°C	30 seconds
3	Anneal	50°C	45 seconds
4	Extend	72°C	60 seconds
Repeat 2 – 4 30x, or 10x for first reconditioned PCR step			
5	Final Extension	72°C	2 minutes
6	Hold	10°C	Hold

Sequencing

Sequencing will be performed on the Illumina MiSeq at the Great Lakes Genomics Center (GLGC). An extraction blank will be run as a negative control and a mock community will be sequenced as a positive control for quality control and processing. Data will be processed in-

house through GLGC support (Aurash Mohaimani). Data will be further processed through Mothur (Schloss et al., 2009), DADA2 (Callahan et al., 2016), and SILVA classification (Quast et al., 2013). R and RStudio will be used to visualize and statistically analyze (Willis, 2017) processed data along with the vegan package in R(Oksanen, 2005; Oksanen et al., 2013).

DAPI Sample Processing Instructions for Groundwater Samples

Note to wear gloves throughout processing in order to avoid contaminating the sample and to avoid skin contact with formaldehyde fixative and DAPI stain.

DAPI Sample Prep

1. Measure the sample volume exactly.
    a. 1 ml or 5 mls for groundwater samples into a new labeled tube
2. Add fixative to the sample and mix (swirl, invert), allow the fixative to sit in the sample for at least 30 minutes before adding DAPI stain.
    a. 21% formaldehyde in PBS solution is 111ul fixative per 1 ml sample.
    b. 4% formaldehyde in PBS solution is 1:2 proportion,
        i. 2.5 mls fixative 4% formaldehyde in 5 mls of sample
        ii. 5 mls of fixative at 4% into 10 mls of sample
3. If the sample volume is 1 ml, add 10 µl of DAPI (1mg/ml) stain directly to sample/fixative solution in the dark. Allow the sample to incubate at room temperature in the dark for 15 mins with DAPI stain.
    a. DAPI stain is light sensitive and will fade over time when exposed to light.

DAPI Filtering

1. Set up vacuum filtration apparatus

a. Note: Be sure to use a secondary container/vacuum flask so no moisture passes into the vacuum system.

2. Pour sterile water into the column and turn on the vacuum to rinse the system/apparatus.
    a. Rinse up to at least the volume of your total sample (fixative + DAPI), but best practice would be to rinse the entire column fully.
3. Add the 0.1 µm (or 0.2 um) filter to the apparatus.
    a. Note: keep the filter face up by keeping it in the same direction as it comes in the box/package. Essentially whichever side of the filter is face-up in the box/packaging, keep that side face-up for filtering (cells will be on the face-up side after filtering).
    b. Use forceps from freshly prepared 70% ethanol to transfer and apply the filter.
        i. Ethanol will evaporate if it is left out. 70% ethanol has been shown to be the most effective concentration for coagulating protein all the way through entire cells to kill microorganisms.
4. Replace the filter column/clamp. Add your sample to the filter column and slowly vacuum the sample through. Turn off vacuum when it is fully filtered.
5. Add PBS up to the volume that the sample reached and filter through to rinse the sides of the filter column and bring down any cells.
6. Do not filter the PBS entirely through if DAPI stain has not been applied to the sample already (>1ml samples). Leave a small amount of PBS covering the filter so that 10 µl of the DAPI stain can be pipetted directly into the PBS covering the filter. Pipette up and down and make sure the DAPI is thoroughly spread and mixed well in the PBS covering the filter.

a. To reduce use of DAPI stain resources (i.e. scaling DAPI up by 5 to 10 x for 5 ml and 10 ml samples respectively), add the 10 µl of DAPI stain directly onto the fitler with enough PBS so that it is evenly distributed on the filter. In this way, the same volume of DAPI will be used and will stain the cells already on the filter rather than staining the cells in solution in 1 ml sample volumes.

7. Leave DAPI on the filter for at least 15 mins in the dark.
8. Filter DAPI through the filter.
9. Apply PBS to rinse the DAPI from the filter.
    a. If you just let the DAPI filter through without rinsing with PBS after, auto fluorescence from DAPI can occur, so it is important to rinse DAPI after filtering.

gDNA Extraction for Future Shotgun Metagenomic Sequencing

Modified Qiagen DNeasy PowerMax Soil Kit combined with Qiagen AllPrep Powerviral DNA/RNA DNA Extraction Protocol:

1. Turn on water bath, 60°C (place solution C1 if it has precipitated in water bath while cutting filters).
2. Thoroughly disinfect bench, working area, Bunsen burner area, pipettes, scissors, ice bucket, dissection scissors etc. (70% ethanol)
3. Get ice, turn on rotisserie incubator to 65° C
4. Obtain 4 screw top (2ml) tubes per sample and fill with zirconian beads (similar to volume of powerviral bead tubes).
5. Take out one filter from -80° C at a time, place on ice, and label new whirl pak with sample info adding current date.

6. After partial thawing of the filter (just until it's no longer rigid), using sterilized (large) scissors, cut the filter in half. Place the unused half in the new labeled whirl pak and place back on ice and into the -80° C freezer as soon as possible. Cut the half for extraction into 4 portions dividing up between 4 screw caps with zirconian beads. Keep on ice throughout protocol.
7. Using sterilized (ethanol and flame) dissection scissors, cut filter portions into tiny pieces in screw top 2 ml tubes.
    a. Steps 5-7 take ~20-30 minutes per half of a filter.
8. Add 300 µl of PowerBead Solution into each tube.
    a. This is buffer for dispersion, dissolving humic acids and nucleic acid degradation protection.
9. Add 300 µl of solution C1 (1.2 mls altogether for one full half filter sample).
    a. Lysis buffer with SDS is the disruption agent which begins to break down fatty acids and lipids in membranes of some organisms.
10. Add 6 µl of Proteinase K.
11. Bead beat for 30 seconds only!
12. Pour/dump contents of the 4 tubes into 50 ml conical tube.
13. Rinse 2 ml tubes with PowerBead solution 2-3 times with ~2 mls to transfer all contents from small screw cap tube into the conical tube. Then add PowerBead Solution to the 50 ml conical tube to bring up the volume to 15 mls.
14. Rotisserie incubator, tape tubes perpendicular to rod to increase mixing. Ensure all caps are fully close and be careful removing tape as to not catch and screw off the caps.
15. Incubate at 65° C spinning and mixing for at least 30 minutes.

16. Centrifuge at 2500x g for 3 mins room temp in the large centrifuge.
    a. Make weights/balances for large centrifuge (~50.5 g).
17. Transfer supernatant to clean collection tube.
    a. ~13 mls, may still have "soil" particles but subsequent steps get rid of it.
18. Add 5 mls of solution C2, invert twice to mix, incubate 2-8° C (fridge) for 10 mins.
19. Centrifuge 2500 x g 4 mins at room temp.
    a. This step may not pellet well.
20. Avoiding pellet, transfer supernatant to clean collection tube.

****
21. Add 4 ml of solution C3 and invert twice to mix, incubate 2-8 C for 10 mins.
22. Centrifuge tubes at 2500 x g for 4 mins at room temp.
23. Avoiding the pellet, transfer to a clean collection tube.
    a. Completely avoid all pellet.

********** repeat 21-23
24. Shake to mix C4 solution.
25. Add 30 mls of solution C4 and invert twice.
    a. Salt solution increases salt concentrations so DNA binds to silica membrane.
    b. Serological pipette or eyeball it.

Proceed to Qiagen's PowerViral Allprep DNA/RNA Extraction Kit Step 12

26. Load 625 µl of supernatant onto a Spin Filter and centrifuge at 13,000 g for 1 minute.
    a. Use two spin columns per sample.
    b. This should take ~50 loads (split between 2 columns, 25 loadings and 25 1 minute centrifugations).

27. Shake to mix solution PV5, add 600 µl of solution PV5 to the spin filter, and centrifuge at 13,000 x g for 1 minute.
28. Discard flow through and add 600 µl of solution PV4 and centrifuge at 13,000 x g for 1 minute.
    a. Pv4 and Pv5 are isopropanol and ethanol washes.
29. Discard flow through and centrifuge again at 13,000 x g for 2 minutes to remove any residual was solution.
30. Place spin filter basket into clean 2 ml collection tube.
31. Add 50 µl of solution PV6 directly onto column and let it incubate for 2-5 mins.
32. Centrifuge 13,000xg for 1 minutes.
33. Repeat 50 µl elution (31-32) and discard spin filter basket.

APPENDIX C: NUTRIENT EXPERIMENT RAW DATA RESULTS

Nitrate Results

Table 20: Raw Data Nitrate Results from Nutrient Experiments

Sample NL Number	N:P Ratio	Initial Nitrate mg/L	Final Nitrate mg/L	Difference Nitrate mg/L	Percent Remaining
646	Control	2.08	0.45	1.63	22%
647	Control	2.08	1.08	1.00	52%
648	Control	2.08	1.04	1.04	50%
649	Control	2.08	1.05	1.03	50%
634	5	6	4.11	1.89	69%
635	5	6	5.24	0.76	87%
636	5	6	5.01	0.99	84%
637	5	6	5.16	0.84	86%
638	50	6	3.37	2.63	56%
639	50	6	4.28	1.72	71%
640	50	6	5.28	0.72	88%
641	50	6	0.73	5.27	12%
642	500	6	2.43	3.57	41%
643	500	6	5.36	0.64	89%
644	500	6	3.34	2.66	56%
645	500	6	5.18	0.82	86%

Total Dissolved Phosphorus Results

Table 21: Raw Data Total Dissolved Phosphorus Results from Nutrient Experiments.

Sample NL Number	N:P Ratio	Initial TPD µg/L	Final TPD µg/L	Difference TPD µg/L	Percent Remaining
646	Control	3.93	1.01	2.92	26%
647	Control	3.93	BDL	3.72	5%
648	Control	3.93	0.57	3.36	15%
649	Control	3.93	1.11	2.82	28%
634	5	537.72	238.41	299.31	44%
635	5	537.72	236.20	301.52	44%
636	5	537.72	237.93	299.79	44%
637	5	537.72	278.72	259.00	52%
638	50	53.77	48.28	5.49	90%
639	50	53.77	51.25	2.52	95%
640	50	53.77	52.15	1.62	97%
641	50	53.77	44.28	9.49	82%
642	500	5.38	BDL	5.07	6%
643	500	5.38	1.04	4.34	19%
644	500	5.38	BDL	5.10	5%
645	500	5.38	1.16	4.22	22%

Ammonium Results

Table 22: Raw Data Ammonium Results from Nutrient Experiments. All final values were below detection limit.

Sample NL Number	N:P Ratio	Initial Ammonium mg/L	Final Ammonium	Difference Ammonium	Percent Remaining
646	Control	0.0095	NA	0	NA
647	Control	0.0095	NA	0	NA
648	Control	0.0095	NA	0	NA
649	Control	0.0095	NA	0	NA
634	5	0.0095	NA	0	NA
635	5	0.0095	NA	0	NA
636	5	0.0095	NA	0	NA
637	5	0.0095	NA	0	NA
638	50	0.0095	NA	0	NA
639	50	0.0095	NA	0	NA
640	50	0.0095	NA	0	NA
641	50	0.0095	NA	0	NA
642	500	0.0095	NA	0	NA
643	500	0.0095	NA	0	NA
644	500	0.0095	NA	0	NA
645	500	0.0095	NA	0	NA

Soluble Reactive Phosphorus Results

Table 23: Raw Data Soluble Reactive Phosphorus Results. Detection limit for Phosphorus was 0.5 µg/L.

Sample NL Number	N:P Ratio	Initial µg/L	Final µg/L	Difference µg/L	Percent Remaining
646	Control	3.0	0.6	2.4	19%
647	Control	3.0	0.4 BDL	2.7	12%
648	Control	3.0	0.5	2.5	17%
649	Control	3.0	0.2 BDL	2.8	7%
634	5	537.5	151.7	385.8	28%
635	5	537.5	155.4	382.1	29%
636	5	537.5	174.0	363.5	32%
637	5	537.5	162.2	375.3	30%
638	50	53.3	46.3	7.1	87%
639	50	53.3	49.6	3.7	93%
640	50	53.3	52.1	1.2	98%
641	50	53.3	40.7	12.6	76%
642	500	4.9	0.2 BDL	4.7	4%
643	500	4.9	0.3 BDL	4.6	6%
644	500	4.9	0.1BDL	4.8	1%
645	500	4.9	0.7	4.2	15%

Table 24: Nitrogen to Phosphorus Ratios Raw Data for Nutrient Experiments. The ratios are based on molar ratios. The initial N:P ratios were based off of TDP (Bootsma Lab) and TN (TOC-TN L machine) values. The final N:P ratios are from final TDP (Bootsma Lab) and final nitrate-N (IC). The TOC-TN L machine was out of order at the end of the experiments.

N:P Ratio	Initial N:P Measured	Final N:P Measured
Control	185	2043
Control	185	12592
Control	185	4378
Control	185	2325
5	5	7
5	5	9
5	5	11
5	5	1
50	50	25
50	50	52
50	50	32
50	50	58
500	500	723
500	500	519
500	500	1879
500	500	452

Nutrient Experiment Cell Counts

Table 25:Raw data cell counts from nutrient addition experiments. Values of NA mean there were not cell counts available for those samples. Note NL 640 and NL 634 were not included in the final average results due to sample discrepancies.

Sample NL Number	N:P Ratio	Initial Cell Counts cells/ml	Cell Counts cells/ml	Difference Cell Count cells/ml	Average Cell Count cells/ml
647	Control	2.10E+05	1.45E+05	-6.50E+04	2.57E+05
648	Control	2.10E+05	3.68E+05	1.58E+05	
634	5	2.10E+05	1.41E+06	1.20E+06	
635	5	2.10E+05	3.71E+05	1.61E+05	6.35E+05
636	5	2.10E+05	1.24E+05	-8.60E+04	
637	5	2.10E+05	NA	NA	
638	50	2.10E+05	NA	NA	
639	50	2.10E+05	1.75E+05	-3.50E+04	7.13E+05
640	50	2.10E+05	1.74E+06	1.53E+06	
641	50	2.10E+05	2.25E+05	1.50E+04	
642	500	2.10E+05	NA	NA	7.82E+04
643	500	2.10E+05	3.99E+04	-1.70E+05	
644	500	2.10E+05	6.88E+04	-1.41E+05	
645	500	2.10E+05	1.26E+05	-8.40E+04	

APPENDIX D: W12 UNIQUE DNA AND RNA TAXA SPECIALISTS

Table 26: Unique W12 Specialists from DNA and RNA CLAM Results

ASV ID	Unique W12 Taxa
9	Unclassified;
21	Bacteria; Nitrospirae; 4; 29; 1;
24	Archaea; Nanoarchaeaeota; Woesearchaeia;
44	Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Thiovulaceae Sulfuricurvum;
60	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae
53	Archaea; Nanoarchaeaeota; Woesearchaeia;
61	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae Thauera;
68	Unclassified;
100	Unclassified;
106	Unclassified;
108	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae; Desulfocapsa;
156	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae
143	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae
147	Unclassified;
144	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae; C39;
163	Bacteria; Omnitrophicaeota;
170	Unclassified;
206	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
252	Unclassified;
251	Bacteria; Armatimonadetes; Chthonomonadetes; Chthonomonadales;
231	Unclassified;
267	Unclassified;
290	Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Sedimenticolaceae; uncultured;
255	Unclassified;
277	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Hydrogenophilaceae; uncultured;
278	Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Thiovulaceae; Sulfuricurvum;
291	Unclassified;
302	Unclassified;
317	Unclassified;

318	Unclassified;
316	Unclassified;
309	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Hydrogenophilaceae;
343	Unclassified;
329	Archaea; Thaumarchaeota; Marine; Benthic; Group; A;
330	Unclassified;
407	Unclassified;
402	Unclassified;
366	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Gallionella;
395	Unclassified;
379	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Sideroxydans;
445	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; MSB; 3C8;
443	Unclassified;
436	Bacteria; Verrucomicrobia; Verrucomicrobiae; LD1; PB3;
481	Unclassified;
435	Unclassified;
439	Unclassified;
466	Unclassified;
474	Unclassified;
580	Unclassified;
567	Unclassified;
562	Unclassified;
616	Unclassified;
595	Bacteria; Nitrospirae; HDB; SIOI1093;
637	Bacteria; Elusimicrobia; Elusimicrobia; Lineage; IV;
587	Bacteria; Acidobacteria; Subgroup; 22;
652	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
615	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Gallionella;
668	Bacteria; Omnitrophicaeota;
611	Unclassified;
646	Unclassified;
675	Unclassified;
681	Unclassified;
682	Unclassified;
705	Archaea; Nanoarchaeaeota; Woesearchaeia;
731	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Nitrosomonadaceae; GOUTA6;

730	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Gallionella;
750	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
746	Bacteria; Elusimicrobia; Elusimicrobia; Lineage; IV;
776	Unclassified;
792	Unclassified;
844	Bacteria; Actinobacteria; Acidimicrobiia; IMCC26256;
828	Bacteria; Omnitrophicaeota;
840	Bacteria; Patescibacteria; Gracilibacteria; Candidatus; Peregrinibacteria;
814	Bacteria; Omnitrophicaeota;
849	Unclassified;
870	Unclassified;
890	Unclassified;
820	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
951	Bacteria; Fibrobacteres; Chitinivibrionia; uncultured;
963	Unclassified;
903	Unclassified;
916	Bacteria; Proteobacteria; Deltaproteobacteria;
955	Unclassified;
928	Unclassified;
961	Unclassified;
997	Archaea; Nanoarchaeaeota; Woesearchaeia;
1027	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;
1012	Unclassified;
1033	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobacteraceae;
1020	Bacteria; Elusimicrobia; Elusimicrobia; Lineage; IV;
1054	Unclassified;
1153	Unclassified;
1135	Bacteria; Zixibacteria;
1114	Bacteria; Acidobacteria; Subgroup; 22;
1365	Unclassified;
1137	Bacteria; Actinobacteria; MB; A2; 108;
1265	Bacteria; Planctomycetes; Phycisphaerae; MSBL9; GWC2; 45; 44;
1172	Bacteria; Acidobacteria; Subgroup; 6;
1280	Unclassified;
1260	Unclassified;
1412	Unclassified;
1239	Bacteria; Bacteroidetes; Bacteroidia; Chitinophagales; Chitinophagaceae; Sediminibacterium;
1222	Bacteria; Omnitrophicaeota;

1449	Bacteria; Omnitrophicaeota;
1317	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;
1345	Unclassified;
1313	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Syntrophus;
1275	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Syntrophus;
1576	Bacteria; Bacteroidetes; Bacteroidia; Sphingobacteriales; S15A; MN91;
1361	Archaea; Crenarchaeota; Bathyarchaeia;
1425	Unclassified;
1364	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Syntrophus;
1473	Unclassified;
1566	Bacteria; Omnitrophicaeota;
1507	Unclassified;
1500	Bacteria; Proteobacteria; Deltaproteobacteria; MBNT15;
1501	Unclassified;
1577	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Syntrophus;
1539	Archaea; Crenarchaeota; Bathyarchaeia;
1578	Unclassified;
1776	Unclassified;
1835	Bacteria; Actinobacteria; WCHB1; 81;
1592	Bacteria; Omnitrophicaeota; Omnitrophia; Omnitrophales; Omnitrophaceae; Candidatus; Omnitrophus;
1793	Bacteria; Omnitrophicaeota; Omnitrophia; Omnitrophales; Omnitrophaceae; Candidatus; Omnitrophus;
1804	Bacteria; Omnitrophicaeota;
1560	Unclassified;
1921	Unclassified;
1971	Unclassified;
1889	Unclassified;
2202	Bacteria; Schekmanbacteria;
2290	Bacteria; Omnitrophicaeota;
2198	Unclassified;
33	Unclassified;
51	Unclassified;
70	Unclassified;
95	Unclassified;
107	Unclassified;
129	Bacteria; Proteobacteria; Deltaproteobacteria; MBNT15;
127	Unclassified;

145	Unclassified;
146	Unclassified;
160	Archaea; Nanoarchaeaeota; Woesearchaeia;
183	Bacteria; Omnitrophicaeota;
173	Unclassified;
164	Unclassified;
185	Unclassified;
203	Unclassified;
199	Unclassified;
242	Unclassified;
283	Unclassified;
296	Unclassified;
314	Unclassified;
322	Unclassified;
409	Unclassified;
388	Unclassified;
414	Unclassified;
404	Unclassified;
413	Unclassified;
411	Archaea; Nanoarchaeaeota; Woesearchaeia;
437	Unclassified;
485	Archaea; Nanoarchaeaeota; Woesearchaeia;
502	Unclassified;
556	Unclassified;
577	Unclassified;
547	Unclassified;
569	Archaea; Nanoarchaeaeota; Woesearchaeia;
584	Unclassified;
596	Unclassified;
606	Archaea; Nanoarchaeaeota; Woesearchaeia;
613	Unclassified;
633	Unclassified;
655	Bacteria; Spirochaetes; Spirochaetia; Spirochaetales; Spirochaetaceae; uncultured;
660	Unclassified;
695	Unclassified;
674	Unclassified;
688	Bacteria; Firestonebacteria;
711	Unclassified;
717	Unclassified;
761	Unclassified;

747	Unclassified;
758	Unclassified;
721	Archaea; Nanoarchaeaeota; Woesearchaeia;
763	Unclassified;
753	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
817	Unclassified;
786	Unclassified;
822	Bacteria; Patescibacteria; Parcubacteria; Candidatus; Ryanbacteria;
816	Unclassified;
810	Unclassified;
830	Unclassified;
865	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Woesebacteria;
866	Unclassified;
853	Unclassified;
850	Unclassified;
855	Unclassified;
896	Unclassified;
897	Unclassified;
874	Unclassified;
936	Unclassified;
971	Unclassified;
940	Archaea; Nanoarchaeaeota; Woesearchaeia;
1007	Unclassified;
1031	Unclassified;
987	Unclassified;
974	Unclassified;
1018	Unclassified;
1041	Unclassified;
989	Unclassified;
1080	Unclassified;
1040	Bacteria; Patescibacteria; Parcubacteria; Candidatus; Yanofskybacteria;
1071	Unclassified;
1077	Unclassified;
1154	Unclassified;
1115	Unclassified;
1103	Unclassified;
1148	Unclassified;
1105	Unclassified;
1151	Unclassified;
1382	Unclassified;

1243	Unclassified;
1422	Unclassified;
1162	Unclassified;
1240	Unclassified;
1293	Unclassified;
1314	Unclassified;
1278	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Amesbacteria;
1444	Unclassified;
1286	Unclassified;
1285	Unclassified;
1318	Bacteria; Firestonebacteria;
1375	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Woesebacteria;
1528	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Woesebacteria;
1857	Unclassified;
128	Unclassified;
138	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
386	Unclassified;
428	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Hydrogenophilaceae; uncultured;
487	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae; MSBL7;
462	Unclassified;
492	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Polynucleobacter;
582	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;
543	Unclassified;
575	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Desulfomonile;
586	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Hydrogenophilaceae; uncultured;
626	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
661	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Sideroxydans;
636	Unclassified;
659	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Gallionella;
769	Unclassified;
785	Unclassified;
749	Unclassified;
832	Unclassified;
841	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae; C39;

843	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Sideroxydans;
894	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Limnobacter;
985	Unclassified;
981	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;
977	Unclassified;
979	Bacteria; Planctomycetes; BD7; 11;
991	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Xanthobacteraceae;
1101	Unclassified;
1136	Bacteria; Proteobacteria; Deltaproteobacteria; Sva0485;
1070	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales; UASB; TL25;
1241	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Desulfomonile;
1170	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
1272	Bacteria; Calditrichaeota; Calditrichia; Calditrichales; Calditrichaceae; SM23; 31;
1495	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Hydrogenophilaceae; Thiobacillus;
1514	Unclassified;
1544	Unclassified;
1552	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae; Denitratisoma;
1532	Unclassified;
1641	Unclassified;
1643	Bacteria; Rokubacteria; NC10; Methylomirabilales; Methylomirabilaceae; Candidatus; Methylomirabilis;
1655	Bacteria; Proteobacteria; Gammaproteobacteria;
1608	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Desulfobacca;
1691	Bacteria; Proteobacteria; Gammaproteobacteria; Beggiatoales; Beggiatoaceae; uncultured;
1850	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
1873	Bacteria; Lentisphaerae; Lentisphaeria; Victivallales; PRD18C08;
2084	Unclassified;
2073	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Sideroxydans;
2118	Unclassified;
2487	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Maritimimonas;
2412	Bacteria; Omnitrophicaeota;
2395	Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; Desulfomonile;
2491	Unclassified;

2561	Unclassified;
2689	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales; Eel; 36e1D6;
2714	Bacteria; Planctomycetes; Phycisphaerae; MSBL9; SG8; 4;
3613	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidetes; BD2; 2;
4697	Unclassified;
5	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
7	Bacteria; Nitrospirae; 4; 29; 1;
105	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
125	Unclassified;
136	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; TRA3; 20;
202	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
230	Bacteria; Spirochaetes; Spirochaetia; Spirochaetales; Spirochaetaceae; Treponema;
263	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Sulfuricellaceae; Sulfuricella;
301	Unclassified;
357	Unclassified;
387	Unclassified;
463	Unclassified;
433	Unclassified;
434	Unclassified;
447	Bacteria; Nitrospirae; 4; 29; 1;
506	Unclassified;
523	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Xanthobacteraceae; Pseudolabrys;
491	Bacteria; Spirochaetes; Spirochaetia; Brevinematales; Brevinemataceae; Brevinema;
521	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales; BIrii41;
666	Unclassified;
826	Unclassified;
777	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
766	Bacteria; Spirochaetes; Spirochaetia; Spirochaetales; Spirochaetaceae; Spirochaeta; 2;
788	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
1014	Unclassified;
1022	Unclassified;
1056	Unclassified;
1204	Unclassified;
1245	Unclassified;
1409	Unclassified;
1438	Archaea; Nanoarchaeaeota; Woesearchaeia;
1465	Unclassified;
1662	Unclassified;

1790	Bacteria; Firmicutes; Bacilli; Lactobacillales; Streptococcaceae; Lactococcus;

APPENDIX E: W13 UNIQUE DNA AND RNA TAXA SPECIALISTS

Table 27: Unique W13 Specialists from DNA and RNA CLAM Results

ASV ID	Unique W13 Taxa
1	Unclassified;
2	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
3	Bacteria; Nitrospirae; 4; 29; 1;
6	Unclassified;
14	Bacteria; Nitrospirae; HDB; SIOI1093;
19	Unclassified;
26	Unclassified;
34	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
37	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
35	Unclassified;
41	Bacteria; Nitrospirae; 4; 29; 1;
40	Unclassified;
42	Unclassified;
47	Bacteria; Patescibacteria; Parcubacteria; Candidatus; Nomurabacteria;
48	Unclassified;
52	Unclassified;
64	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
59	Unclassified;
75	Unclassified;
73	Unclassified;
74	Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Thiovulaceae; Sulfurimonas;
82	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
85	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Candidatus; Nitrotoga;
84	Bacteria; Proteobacteria; Deltaproteobacteria; DTB120;
78	Unclassified;
88	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
89	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
97	Unclassified;
102	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;

104	Unclassified;
117	Unclassified;
116	Unclassified;
141	Bacteria; Planctomycetes; Brocadiae; Brocadiales; Brocadiaceae;
126	Bacteria; Patescibacteria; Parcubacteria; UBA9983;
139	Unclassified;
148	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
166	Unclassified;
168	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
188	Unclassified;
167	Unclassified;
220	Unclassified;
217	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales; mle1; 27;
232	Bacteria; Proteobacteria; Alphaproteobacteria; Acetobacterales; Acetobacterales; Incertae; Sedis; uncultured;
236	Unclassified;
282	Unclassified;
288	Bacteria; Patescibacteria; Parcubacteria; UBA9983;
312	Bacteria; Planctomycetes; Brocadiae; Brocadiales; Brocadiaceae;
313	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;
334	Bacteria; Actinobacteria; Thermoleophilia; Gaiellales;
345	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
358	Unclassified;
336	Unclassified;
359	Bacteria; Latescibacteria;
367	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Nitrosomonadaceae; MND1;
393	Archaea; Diapherotrites; Micrarchaeia;
375	Unclassified;
355	Unclassified;
368	Unclassified;
444	Bacteria; Bacteroidetes; Bacteroidia; Sphingobacteriales; env; OPS; 17;
459	Bacteria; Omnitrophicaeota;
451	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
458	Bacteria; Proteobacteria; Alphaproteobacteria; Acetobacterales; Acetobacterales; Incertae; Sedis; uncultured;
539	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Nitrosomonadaceae; MND1;
554	Unclassified;
531	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;

558	Unclassified;
535	Archaea; Nanoarchaeaeota; Woesearchaeia;
570	Bacteria; Acidobacteria; Blastocatellia; ; Subgroup; 4; ; Blastocatellales; Blastocatellaceae;
563	Bacteria; Acidobacteria; Subgroup; 6;
619	Unclassified;
598	Unclassified;
608	Unclassified;
631	Bacteria; Omnitrophicaeota;
691	Bacteria; Chloroflexi; KD4; 96;
698	Bacteria; Patescibacteria; Parcubacteria; UBA9983;
737	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae; Sideroxydans;
752	Unclassified;
793	Bacteria; Planctomycetes; Brocadiae; Brocadiales; Brocadiaceae;
791	Unclassified;
877	Archaea; Nanoarchaeaeota; Woesearchaeia;
900	Unclassified;
941	Unclassified;
1006	Unclassified;
1189	Bacteria; Patescibacteria; ABY1; Candidatus; Uhrbacteria;
1231	Archaea; Nanoarchaeaeota; Woesearchaeia;
1435	Bacteria; Bacteroidetes; Bacteroidia; Sphingobacteriales;
1702	Bacteria; Latescibacteria;
20	Unclassified;
25	Unclassified;
27	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
50	Unclassified;
81	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
137	Unclassified;
151	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
142	Unclassified;
157	Unclassified;
175	Unclassified;
182	Unclassified;
212	Unclassified;
214	Unclassified;
244	Bacteria; Patescibacteria; Parcubacteria; Candidatus; Nomurabacteria;
246	Unclassified;
259	Unclassified;

260	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
274	Unclassified;
323	Bacteria; Elusimicrobia; Lineage; IIc;
353	Unclassified;
408	Unclassified;
406	Unclassified;
401	Unclassified;
432	Unclassified;
448	Unclassified;
442	Unclassified;
449	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Woykebacteria;
454	Unclassified;
513	Unclassified;
498	Unclassified;
529	Unclassified;
515	Archaea; Nanoarchaeaeota; Woesearchaeia;
549	Unclassified;
532	Unclassified;
540	Unclassified;
641	Unclassified;
693	Unclassified;
696	Unclassified;
736	Unclassified;
727	Unclassified;
744	Unclassified;
783	Unclassified;
813	Unclassified;
879	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Woykebacteria;
889	Unclassified;
885	Unclassified;
911	Unclassified;
1001	Unclassified;
1025	Bacteria; Patescibacteria; Parcubacteria; Candidatus; Vogelbacteria;
1058	Unclassified;
1083	Unclassified;
1258	Bacteria; Patescibacteria; Microgenomatia; Candidatus; Amesbacteria;
23	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Sulfuricellaceae; Ferritrophicum;
31	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;

83	Unclassified;
155	Unclassified;
218	Unclassified;
247	Unclassified;
303	Unclassified;
346	Bacteria; Nitrospirae; Thermodesulfovibrionia; uncultured;
344	Unclassified;
503	Bacteria; Bacteroidetes; Ignavibacteria; Kryptoniales; BSV26;
478	Unclassified;
510	Unclassified;
495	Unclassified;
724	Bacteria; Chloroflexi; Anaerolineae; RBG; 13; 54; 9;
703	Unclassified;
684	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae;
835	Unclassified;
880	Unclassified;
893	Unclassified;
939	Bacteria; Proteobacteria; Gammaproteobacteria; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
950	Unclassified;
972	Unclassified;
1185	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae;
1251	Unclassified;
1236	Bacteria; Planctomycetes; Brocadiae; Brocadiales; Brocadiaceae;
1140	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales; Blfdi19;
1288	Unclassified;
1279	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;
1404	Unclassified;
1377	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Gallionellaceae;
1634	Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales; Bdellovibrionaceae; Bdellovibrio;
1696	Unclassified;
1670	Bacteria; Proteobacteria; Deltaproteobacteria; MBNT15;
79	Unclassified;
171	Unclassified;
340	Unclassified;
374	Unclassified;
603	Unclassified;
643	Unclassified;
610	Unclassified;

704	Unclassified;
751	Unclassified;
839	Unclassified;
852	Unclassified;
924	Unclassified;
992	Unclassified;
1069	Unclassified;
1421	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Shewanellaceae; Shewanella;

APPENDIX F: TOP AVERAGE RATIOS USED IN VIOLIN RATIO ANALYIS ACROSS WELLS (FIG. 13, PG 44)

Table 28: Top Average RNA:DNA Ratios from each Well

Sample	ASV 4: Bacteria; Nitrospirae; 4; 29; 1;	ASV 209: Bacteria; Proteobacteria; Gammaproteobacteria ; Betaproteobacteriales; Gallionellaceae;	ASV 593: Bacteria; Proteobacteria; Gammaproteobacteria ; Betaproteobacteriales; Burkholderiaceae; Massilia;	ASV 103: Bacteria; Proteobacteria; Gammaproteobacteria ; Betaproteobacteriales; Hydrogenophilaceae;	ASV 3: Bacteria; Nitrospirae; 4; 29; 1;	ASV 354: Bacteria; Nitrospirae; 4; 29; 1;	ASV 82: Bacteria; Proteobacteria; Gammaproteobacteria ; Acidiferrobacterales; Acidiferrobacteraceae; Sulfurifustis;
1W11R248	7.151074796	Inf	NA	Inf	5.412180143	Inf	Inf
1W11R627	13.93044991	Inf	95.82218913	Inf	6.505658645	Inf	NA
1W11R629	4.451353435	Inf	NA	Inf	4.124882005	Inf	Inf
1W11R706	7.385440915	Inf	NA	Inf	9.708840184	Inf	Inf
1W11R713	4.717108303	51.01800849	NA	Inf	3.273864735	Inf	Inf
1W11R718	8.37435036	Inf	NA	Inf	6.985757192	Inf	Inf
1W11R720	9.236385136	Inf	NA	Inf	8.230537959	Inf	Inf
1W11R725	15.03044829	Inf	Inf	Inf	7.010421361	NA	Inf
1W11R829	11.0962172	6.521591856	NA	Inf	17.70055219	Inf	Inf
2W11R259	2.678383826	9.291355915	NA	13.04802559	2.925580947	8.286586589	Inf
2W11R627	3.060980986	Inf	NA	14.865831	2.349438372	Inf	Inf
2W11R629	2.319927152	4.801989909	NA	10.66844569	1.779173567	7.812166523	Inf
2W11R706	2.881343381	8.228651108	NA	10.74435725	3.302665079	10.2901402	Inf
2W11R713	3.548100879	4.251041972	NA	8.345373085	3.044646086	Inf	NA
2W11R720	2.800887897	5.599493696	NA	19.17796641	2.525152451	Inf	Inf
2W11R829	3.10858451	8.61150735	NA	14.89240616	2.704483562	Inf	4.600339283
1W12R829	2.501709364	Inf	NA	10.50914292	3.625236129	Inf	0.839909434
1W12R255	Inf	NA	Inf	Inf	Inf	Inf	NA
1W12R627	7.125437667	Inf	NA	Inf	Inf	Inf	NA
1W12R629	8.105195951	NA	NA	Inf	5.187201535	Inf	Inf
2W12R629	2.18182886	Inf	NA	Inf	2.680261648	Inf	0.964410593

1W12R706	0.949450606	NA	NA	Inf	0	NA	NA
1W12R713	4.0674289	Inf	NA	12.15070862	5.881024797	Inf	Inf
1W12R725	4.177556196	NA	NA	Inf	1.881281448	Inf	NA
2W12R829	1.563813589	1.107932626	NA	9.811595657	1.782764316	3.358420772	0.432093724
2W12R627	2.341489196	2.847664687	NA	5.99080822	1.390760873	5.096795301	NA
2W12R706	1.934944291	1.438907799	NA	9.930855351	2.562027492	Inf	NA
2W12R713	7.106628811	0	NA	12.4463247	18.37078063	NA	NA
2W12R718	2.849181599	NA	NA	7.414991004	4.122102303	7.708331307	NA
2W12R720	3.352472509	NA	NA	Inf	3.991949532	Inf	NA
1W13R627	8.610416973	NA	NA	Inf	7.969515242	Inf	5.948928268
1W13R629	11.16170251	NA	NA	Inf	6.882996547	Inf	5.874952272
1W13R706	16.28428893	NA	NA	Inf	68.31440633	Inf	9.301665384
1W13R713	18.73345451	NA	NA	Inf	8.979848408	NA	8.559162372
1W13R718	4.046184257	NA	NA	Inf	0.66678961	NA	34.9133339
1W13R720	20.00538215	NA	NA	Inf	4.596836535	NA	9.773629163
1W13R725	7.468561011	Inf	NA	Inf	6.056428205	Inf	7.03702443
1W13R829	6.824961396	NA	NA	Inf	5.627744002	NA	3.785954143
2W13R160	3.470939804	Inf	NA	11.46150258	2.434607211	11.49659597	1.802290119
2W13R627	2.84132288	Inf	NA	Inf	2.903002927	21.33225399	1.046815413
2W13R629	3.45800882	Inf	NA	8.392528938	2.598490955	12.47533767	2.155746478
2W13R706	3.181557146	Inf	NA	9.984504624	2.322942497	14.78648961	2.349007935
2W13R713	2.720286187	Inf	NA	9.32078152	3.141154533	15.06454428	1.819089557
2W13R718	3.516404556	Inf	NA	10.52707809	2.612586197	16.77619161	2.378091792
2W13R720	3.253387461	Inf	NA	7.572275949	2.27286084	18.21025066	3.30073567
2W13R725	2.74710245	2.097350543	NA	7.881642512	2.469764182	15.64202899	1.873405797
2W13R829	2.932128743	NA	NA	16.99535981	2.679389154	Inf	2.495264095

Table 28: Continued : Top Average RNA:DNA Ratios from each Well

Sample	ASV 118: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales;	ASV 60: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae;	ASV 290: Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Sedimenticolaceae; uncultured;	ASV 61: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae; Thauera;	ASV 143: Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae;	ASV 44: Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Thiovulaceae; Sulfuricurvum;
1W11R248	3.822991842	Inf	NA	Inf	NA	NA
1W11R627	Inf	NA	NA	NA	NA	NA
1W11R629	Inf	NA	NA	Inf	Inf	NA
1W11R706	Inf	NA	NA	NA	NA	NA
1W11R713	Inf	0.558351595	NA	NA	NA	Inf
1W11R718	Inf	Inf	NA	NA	NA	NA
1W11R720	Inf	Inf	NA	Inf	NA	NA
1W11R725	Inf	Inf	NA	NA	NA	NA
1W11R829	Inf	Inf	NA	Inf	Inf	NA
2W11R259	20.51293627	Inf	NA	Inf	NA	Inf
2W11R627	Inf	Inf	NA	Inf	NA	Inf
2W11R629	9.837760906	Inf	NA	Inf	NA	Inf
2W11R706	11.89579301	Inf	NA	Inf	NA	Inf
2W11R713	Inf	Inf	NA	Inf	NA	1.188658481
2W11R720	Inf	Inf	NA	Inf	Inf	Inf
2W11R829	10.77315818	Inf	NA	Inf	NA	NA
1W12R829	8.645251376	10.09111233	9.788870413	28.15941467	10.25250218	6.739948558
1W12R255	NA	Inf	Inf	Inf	Inf	Inf
1W12R627	Inf	86.30132545	77.89955949	Inf	Inf	15.6899194
1W12R629	11.67519361	71.11653305	Inf	8.783152717	Inf	36.55172838
2W12R629	9.344854997	11.05425251	7.511476841	Inf	13.31264819	3.904345064
1W12R706	Inf	Inf	NA	Inf	Inf	27.6040713

1W12R713	6.605742387	33.69633329	Inf	Inf	38.40426675	13.37342144
1W12R725	Inf	23.48285294	Inf	Inf	Inf	Inf
2W12R829	10.96214108	7.890172971	8.553891596	21.76015417	6.150655386	2.28962707
2W12R627	6.231753276	6.343844604	7.006474069	20.32293589	11.55266784	1.370637715
2W12R706	5.813253363	5.712355774	5.133400797	16.81784744	6.776877074	1.302452402
2W12R713	3.136081815	2.395618053	32.58040553	8.534389315	18.70048787	3.284355114
2W12R718	6.327960066	5.066354773	10.33617153	25.19635033	10.16785235	2.482578322
2W12R720	7.634271556	11.23878674	11.17363864	Inf	11.36972086	1.804153815
1W13R627	Inf	Inf	NA	NA	NA	Inf
1W13R629	Inf	Inf	NA	NA	NA	Inf
1W13R706	Inf	Inf	NA	NA	NA	Inf
1W13R713	Inf	Inf	NA	NA	NA	Inf
1W13R718	Inf	Inf	NA	NA	NA	Inf
1W13R720	NA	Inf	NA	NA	NA	Inf
1W13R725	Inf	Inf	NA	NA	NA	Inf
1W13R829	Inf	Inf	NA	NA	NA	Inf
2W13R160	9.373332414	8.605440182	Inf	NA	NA	2.850907896
2W13R627	10.85325203	Inf	Inf	NA	NA	5.855912861
2W13R629	7.938924872	8.520731643	NA	NA	NA	5.690102375
2W13R706	22.96436063	Inf	NA	NA	NA	5.793624562
2W13R713	6.77968603	7.09670705	Inf	NA	NA	5.012753392
2W13R718	11.34656968	8.833208102	Inf	NA	NA	4.001037868
2W13R720	10.58662426	9.902207011	Inf	NA	NA	4.28149725
2W13R725	7.394143744	10.65113043	NA	NA	NA	2.933314871
2W13R829	8.877438213	9.585509058	Inf	NA	NA	8.94647512

APPENDIX G: RIVER INFILTRATION INDICATOR TAXA

Table 29: FOX River Infiltration Indicator ASV Taxa

River Infiltration Indicator Taxa
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Rhodocyclaceae; C39;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Crocinitomicaceae; Fluviicola;
Bacteria; Bacteroidetes; Bacteroidia; Cytophagales; Spirosomaceae; Pseudarcicella;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Polynucleobacter;
Bacteria; Actinobacteria; Actinobacteria; Frankiales; Sporichthyaceae; hgcI clade;
Bacteria; Bacteroidetes; Bacteroidia; Chitinophagales; Chitinophagaceae; Sediminibacterium;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Limnohabitans;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Aurantimicrobium;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Polynucleobacter;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Rhodoluna;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Limnohabitans;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Limnohabitans;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Rhodoluna;
Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Arcobacteraceae; Arcobacter;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Flavobacterium;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Candidatus Planktoluna;
Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Arcobacteraceae; Arcobacter;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Crocinitomicaceae; Fluviicola;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Flavobacterium;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Limnohabitans;
Bacteria; Actinobacteria; Actinobacteria; Frankiales; Sporichthyaceae;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae; Polynucleobacter;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Methylophilaceae;

Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiales Incertae Sedis; uncultured;
Bacteria; Bacteroidetes; Bacteroidia; Chitinophagales; Chitinophagaceae; Sediminibacterium;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Candidatus Aquiluna;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae;
Bacteria; Bacteroidetes; Bacteroidia; Chitinophagales; Chitinophagaceae; Sediminibacterium;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Crocinitomicaceae; Fluviicola;
Bacteria; Epsilonbacteraeota; Campylobacteria; Campylobacterales; Arcobacteraceae; Arcobacter;
Bacteria; Bacteroidetes; Bacteroidia; Sphingobacteriales; NS11-12 marine group;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; NS9; marine; group;
Bacteria; Actinobacteria; Actinobacteria; Frankiales; Sporichthyaceae; hgcI clade;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Crocinitomicaceae; Fluviicola;
Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae;
Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteriales; Burkholderiaceae;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Flavobacterium;
Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Flavobacterium;
Bacteria; Actinobacteria; Actinobacteria; Frankiales; Sporichthyaceae; hgcI clade;
Bacteria; Actinobacteria; Actinobacteria; Frankiales; Sporichthyaceae; hgcI clade;
Bacteria; Proteobacteria; Alphaproteobacteria; SAR11; clade; Clade; III;
Bacteria; Proteobacteria; Gammaproteobacteria; Diplorickettsiales; Diplorickettsiaceae; Rickettsiella;
Bacteria; Proteobacteria; Gammaproteobacteria; Diplorickettsiales; Diplorickettsiaceae; Rickettsiella;

APPENDIX H: HEATMAP CLOSE UP AVERAGE MOST ABUNDANT TAXA OF INTEREST

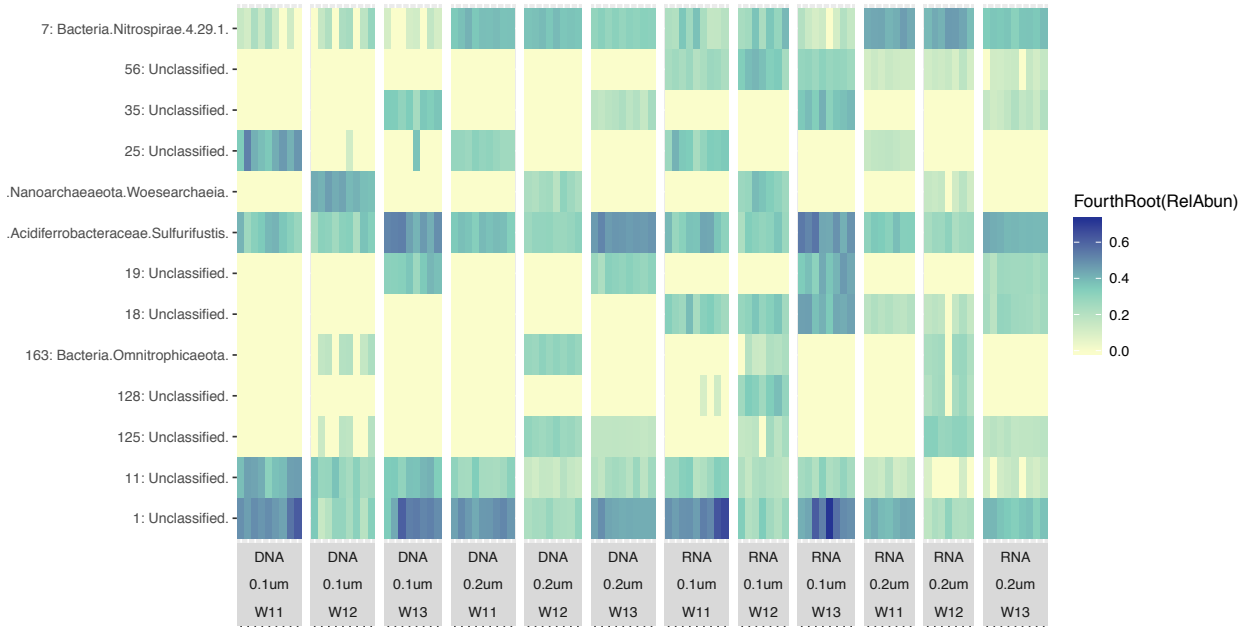


Figure 15: Selection of Most Abundant ASVs Across All Wells. Taxon affiliations are listed on the left. Data is partitioned based on DNA (left half of the heatmap) and RNA (right half of the heatmap), filter size with nucleic acid type (0.1 µm left, 0.2 µm right within DNA and RNA), and within that by site (W11, W12, and W13). The light yellow indicates an average abundance of 0 meaning the ASV was not present for that type of sample. Dark blue indicates the highest relative abundance. The fourth root was used to display the relative abundances to show the patterns more distinctly. Some of the most abundant groundwater ASVs are specific to location suggesting (ASVs: 163, 19, 35) they are unique or specialized to that location. Other ASVs show higher abundances comparatively across all wells (ASV1: Unclassified, ASV 2: Sulfurifustis) suggesting the aquifer shares specific microbes between the wells. Some ASVs that are Unclassified, (ASVs 128, 18, 56) do not appear in the DNA but are of the most abundant RNA ASVs which supports the theory of intron splicing occurring in the 16S rRNA gene.

APPENDIX I: MOCK COMMUNITY DATASHEET

b|e|i RESOURCES
SUPPORTING INFECTIOUS DISEASE RESEARCH

Product Information Sheet for HM-782D

Genomic DNA from Microbial Mock Community B (Even, Low Concentration), v5.1L, for 16S rRNA Gene Sequencing

Catalog No. HM-782D

For research use only. Not for human use.

Contributor and Manufacturer:
Sarah K. Highlander, Associate Professor, Department of Molecular Virology and Microbiology; Baylor College of Medicine, Houston, Texas

Product Description:
HM-782D contains genomic DNA from 20 bacterial strains containing equimolar (Even) ribosomal RNA operon counts (100,000 copies per organism per µL). This mock community is recommended for 16S rRNA gene sequencing by Sanger or amplicon sequencing methods. The recommended amount to use per experiment is 1 µL.[1] The bacterial strains that DNA was extracted from are listed in Table 1.

Note: The label for HM-782D is incorrect. HM-782D contains genomic DNA from microbial mock community B and not microbial mock community A.

Table 1. Microbial Mock Community B

Organism	NCBI Reference Sequence
Acinetobacter baumannii, strain 5377	NC_009085
Actinomyces odontolyticus, strain 1A.21	NZ_AAYI02000000
Bacillus cereus, strain NRS 248	NC_003909
Bacteroides vulgatus, strain ATCC® 8482™	NC_009614
Clostridium beijerinckii, strain NCIMB 8052	NC_009617
Deinococcus radiodurans, strain R1 (smooth)	NC_001263, NC_001264
Enterococcus faecalis, strain OG1RF	NC_17316
Escherichia coli, strain K12, substrain MG1655	NC_000913
Helicobacter pylori, strain 26695	NC_000915
Lactobacillus gasseri, strain 63 AM	NC_008530
Listeria monocytogenes, strain EGDe	NC_003210
Neisseria meningitidis, strain MC58	NC_003112
Propionibacterium acnes, strain KPA171202	NC_006085
Pseudomonas aeruginosa, strain PAO1-LAC	NC_002516
Rhodobacter sphaeroides, strain ATH 2.4.1	NC_007493, NC_007494
Staphylococcus aureus, strain TCH1516	NC_010079
Staphylococcus epidermidis, FDA strain PCI 1200	NC_004461
Streptococcus agalactiae, strain 2603 V/R	NC_004116
Streptococcus mutans, strain UA159	NC_004350
Streptococcus pneumoniae, strain TIGR4	NC_003028

HM-782D has been qualified for PCR applications by amplification of approximately 1500 base pairs of the 16S ribosomal RNA gene.

Material Provided:
Each vial contains approximately 20 µL of the bacterial genomic DNA mixture suspended in Tris-HCl, pH ~ 7.5. The concentration is shown on the Certificate of Analysis. The vial should be centrifuged prior to opening.

Packaging/Storage:
HM-782D was packaged aseptically in screw-capped plastic cryovials. The product is provided frozen on ice and should be stored at -20°C or colder immediately upon arrival. Freeze-thaw cycles should be minimized.

Citation:
Acknowledgment for publications should read "The following reagent was obtained through BEI Resources, NIAID, NIH as part of the Human Microbiome Project: Genomic DNA from Microbial Mock Community B (Even, Low Concentration), v5.1L, for 16S RNA Gene Sequencing, HM-782D."

Biosafety Level: 1
Appropriate safety procedures should always be used with this material. Laboratory safety is discussed in the following publication: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health. Biosafety in Microbiological and Biomedical Laboratories. 5th ed. Washington, DC: U.S. Government Printing Office, 2009; see www.cdc.gov/biosafety/publications/bmbl5/index.htm.

Disclaimers:
You are authorized to use this product for research use only. It is not intended for human use.

Use of this product is subject to the terms and conditions of the BEI Resources Material Transfer Agreement (MTA). The MTA is available on our Web site at www.beiresources.org.

While BEI Resources uses reasonable efforts to include accurate and up-to-date information on this product sheet,

BEI Resources
www.beiresources.org
E-mail: contact@beiresources.org
Tel: 800-359-7370
Fax: 703-365-2898

neither ATCC® nor the U.S. Government makes any warranties or representations as to its accuracy. Citations from scientific literature and patents are provided for informational purposes only. Neither ATCC® nor the U.S. Government warrants that such information has been confirmed to be accurate.

This product is sent with the condition that you are responsible for its safe storage, handling, use and disposal. ATCC® and the U.S. Government are not liable for any damages or injuries arising from receipt and/or use of this product. While reasonable effort is made to ensure authenticity and reliability of materials on deposit, the U.S. Government, ATCC®, their suppliers and contributors to BEI Resources are not liable for damages arising from the misidentification or misrepresentation of products.

Use Restrictions:

This material is distributed for internal research, non-commercial purposes only. This material, its product or its derivatives may not be distributed to third parties. Except as performed under a U.S. Government contract, individuals contemplating commercial use of the material, its products or its derivatives must contact the contributor to determine if a license is required. U.S. Government contractors may need a license before first commercial sale.

References:

1. S. K. Highlander, Personal Communication.

ATCC® is a trademark of the American Type Culture Collection.

Support Provided by NIAID

BEI Resources
www.beiresources.org

E-mail: contact@beiresources.org
Tel: 800-359-7370
Fax: 703-365-2898

APPENDIX J: SUPPLEMENTARY 16S RIBOSOMAL RNA:DNA ACTIVITY RATIOS

Violin plots (Fig. 15), similar to box plots that demonstrate data distribution with the thickest part of the plot showing the mode average, were developed in R to compare the ASVs that had the highest RNA:DNA ratios from each site. The ratios of all corresponding RNA to DNA samples were determined for each ASV. To determine the top 5 average ratio ASVs for each well, data values that were Infinity (dividing by 0 in the DNA) and Not a Number (0 in the RNA) were set to 0 in order to take the averages for each ASV. The 5 highest average ratio and their corresponding ASVs from each well were compiled for a final list of 13 ASVs.

The data plotted in the violin plot analysis included RNA values of 0 which would be indicative of a true ratio value of 0. Calculations returning infinite values were excluded from the analysis (i.e. division by 0 from DNA values of 0). Although no DNA values were measured for these ASVs, it is likely that these ASVs represent organisms that have the rarely recorded bacterial phenomenon of intron splicing in their 16S rRNA gene, and thus were not detected from targeting the DNA. ASVs with an abundance value of 0 in the DNA that have corresponding abundance measurements of RNA could not be used in this ratio analysis since the RNA could not be compared to a quantitative DNA abundance value.

Figure 15 includes some overlap in taxa used in the heatmap (Fig. 11) that were previously discussed. Additional taxa in the ratio violin plots include ASVs: 354, 209, 82, 118, 103, and 291. ASVs 3, 4, and 354 are related to the phylum *Nitrospirae* and have similarly active ratios between all sites. *Hydrogenophilaceae* (ASV 103) was also similarly active across all sites, which has been associated with molecular hydrogen, oxygen, or nitrate

utilization. *Sulfurifustis* (ASV 82) appears to be more active in W13 than W12. It should also be noted that ASVs had highly variable ratios for any given site.

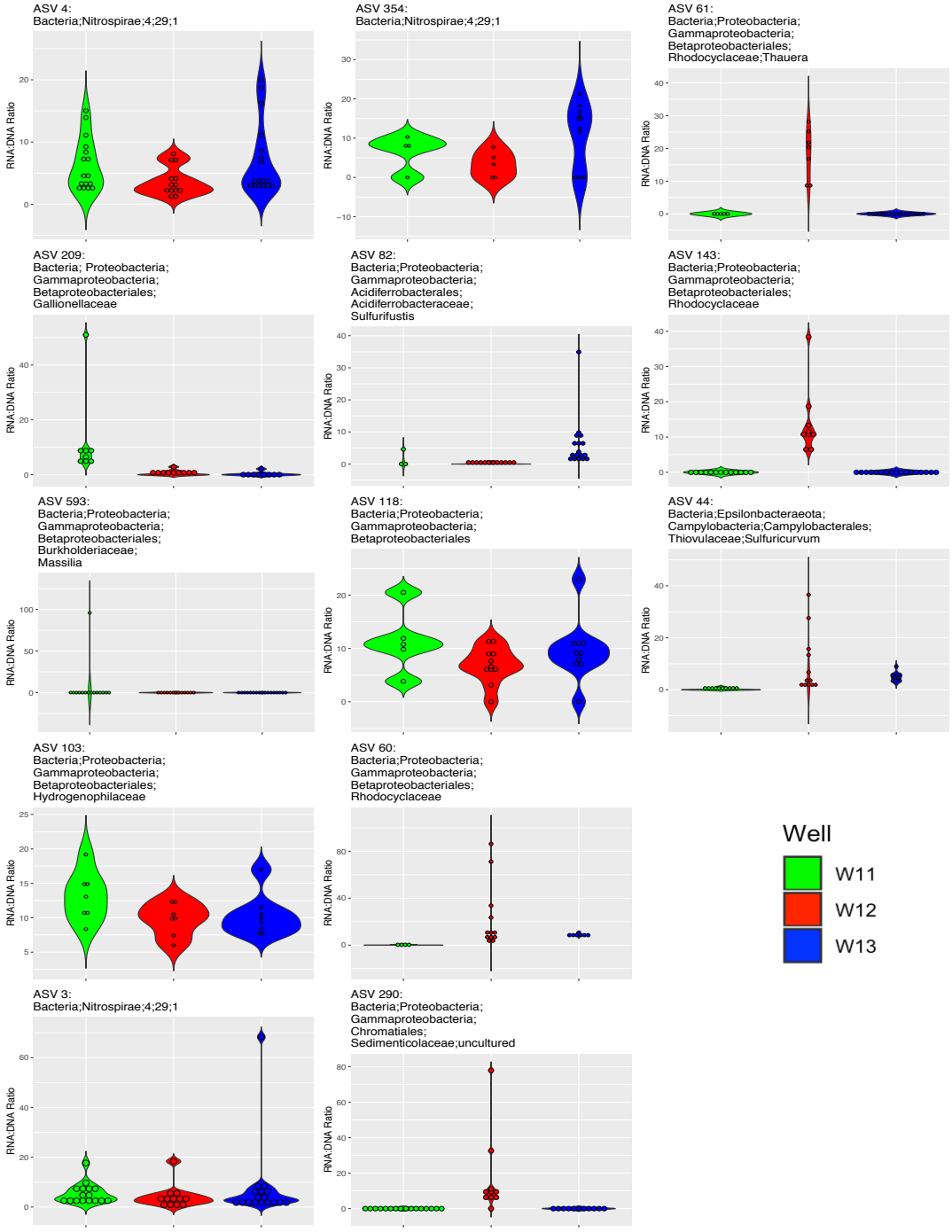


Figure 16:Top Average RNA:DNA Ratio ASVs from each Well. Violin plots comparing the top 5 average RNA:DNA ratio ASV's from each well resulting in 13 ASVs. Violin plots, similar to box plots, show the distribution of data values, with the mode average as the thickest part of the plot. Centered data points were plotted to show sample distribution. W11 is show in green, W12 in red, and W13 in blue. Each plot displays an individual ASV and the ratios for each well and each has its own scaled ratio y-axis.

APPENDIX K: MEASURING SOLUBLE REACTIVE PHOSPHOROUS PROTOCOL

SOLUBLE REACTIVE PHOSPHORUS

PRECISION

Level	-	100 µg P/liter
Standard deviation	-	1 µg P/liter

OPERATING RANGE - 5-250 µg P/liter

SAMPLE CONDITION

- Filtered through Whatman GF/F filter; stored at 5°C.
- Unacidified.
- Less than 24 hours old.
- Less than 5 µg AsO_4^{-3}/L

METHOD PRINCIPLE

Phosphate, silicate, arsenate and germanate ions react under acidic conditions with molybdate to form heteropoly acids which can be converted, by suitable reducing agents, to blue compounds of uncertain composition.

Using appropriate acid and molybdate strength, ascorbic acid as reductant and antimony as a color enhancing species, an intensely blue coloured complex is formed with PO_4^{-3} and AsO_4^{-3} having an absorbance maximum at 885 nm.

It should be noted that, while the formation of the blue complex is specific to PO_4^{-3}, some observers feel that the reaction conditions are capable of hydrolysing labile organic phosphorus compounds. This would give an overestimate of PO_4^{-3} and hence an unreliable estimate of biologically available phosphorus.

REAGENTS

A. Acid molybdate-antimony

- Distilled water: 500 ml
- Ammonium paramolybdate $(NH_4)_6MO_7O_{24}\cdot 4H_2O$: 7.5 g
- Antimony potassium tartrate: 0.14 g (140 mg)
- Sulphuric acid (S.G. = 1.84): 88 ml
- Mix in order given, cool and make to 1000 ml.
  Keep in dark glass bottle.

B. Ascorbic acid

- Dissolve 2.5 g of L-ascorbic acid in 100 ml of distilled water. This reagent is stable for a few days if kept refrigerated.

C. Mixed molybdate for natural colour determination.

- Mix 4 parts of reagent **A** with 1 part of distilled water.

D. Mixed molybdate for orthophosphate determination

- Mix 4 parts of reagent **A** with 1 part of reagent **B**. Stable for 1 day.

E. Phosphate standard stock solution.

- Dissolve 0.2197 g potassium dihydrogen phosphate (KH_2PO_4) in 100 ml of water and make to 1000 ml with water saturated with chloroform. 1.00 ml of stock solution contains 50.0 µg PO_4^{3-}, so the P concentration in this stock solution is 50,000 µg/L, or 50 mg/L.

PROCEDURE

1. Prepare in duplicate 100 ml of orthophosphate standards having a concentration approximating that of the samples to be analyzed. (See separate instructions for preparation of standards).

2. Place 25 ml of each standard and 25 ml of distilled water into 125 ml Erlenmeyer flasks. Also place 25 ml of distilled water into 125 ml Erlenmeyer flask. Add 5.0 ml of mixed molybdate reagent **D** to each flask.

3. Place 25 ml of sample into a 125 ml Erlenmeyer flask. Add 5.0 ml of mixed molybdate reagent **C**.

4. Place 25 ml of sample into a second 125 ml Erlenmeyer flask. Add 5.0 ml of mixed molybdate reagent **D**.

5. After 5 minutes and within 3 hours measure:

E_{Bl} = Absorbance of distilled water + reagents
E_o = Absorbance of sample without reductant
E_1 = Absorbance of standards or sample with reductant.

Absorbance is measured at 885 nm using a 10.00 cm path length.

Note: The above analysis can be carried out with samples of any volume, as long as the volume of reagent added is 1/5 of the sample volume.

CALCULATIONS

1. Unit extinction factor for PO_4-P is calculated as follows:

$$F = \frac{\text{Standard concentration (µg P/liter)}}{E_1 \text{ (standard)} - E_{B1}}$$

This should be around 170 ug PO_4-P/liter for a 10 cm path length.

Concentration of soluble reactive phosphorus [SRP] in sample is calculated as:

$$[SRP] = F \times (E_1 \text{ sample} - (E_o + E_{B1}))$$

SELECTED REFERENCE

Murphy, J. and J.P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chim. Acta 27: 31-36.

Stainton, M. P. 1980. Erros in molybdenum blue methods for determining orthophosphate in freshwater. Can. J. Fish. Aqaut. Sci. 37: 472-478.