# AN ECOLOGICAL APPROACH TO GEOGENIC ARSENIC CONTAMINATION IN BANGLADESH

by

TERESA MARIE LEGG

B.S., University of Oregon, 2003

M.A., University of Colorado, 2011

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Ecology and Evolutionary Biology

2013

*This thesis entitled:*

***An Ecological Approach to Geogenic Arsenic Contamination in Bangladesh,***

*written by Teresa M. Legg*

*has been approved for the Department of Ecology and Evolutionary Biology*

______________________________________

Dr. Diana Nemergut

______________________________________

Dr. Diane McKnight

Date:_______________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards Of scholarly work in the above mentioned discipline.

Legg, Teresa Marie (Ph.D., Ecology and Evolutionary Biology)

**An Ecological Approach to Geogenic Arsenic Contamination in Bangladesh**

Thesis directed by Associate Professor Diana Nemergut

## ABSTRACT

Bacterial communities can exert significant influence on the biogeochemical cycling of groundwater arsenic (As). This has globally important implications since As in drinking water affects the health of over 100 million people worldwide, including in the Ganges-Brahmaputra Delta region of Bangladesh where geogenic arsenic in groundwater can reach concentrations of more than 10 times the World Health Organization's limit. The overall goal of this dissertation research was to examine the relative effects of sediment geochemistry versus DOM chemistry on the structure of bacterial communities across groundwater gradients in an aquifer affected by geogenic arsenic contamination in Araihazar, Bangladesh. To this end, I employed interdisciplinary methods including analytical chemistry, 16S rRNA gene sequencing to characterize the composition and diversity of bacterial communities, and fluorescence spectroscopy and PARAFAC modeling in order to describe the chemistry and source of DOM in high As aquifers. My results indicate that differences in sediment grain size and geochemistry between sites significantly influenced the structure of bacterial communities, and the relative abundances of Deltaproteobacteria and Chloroflexi. However, my research also revealed that the supply and chemistry of groundwater DOM might have a greater effect on the composition of bacterial communities in aquifer sediments than site-specific differences in sediment geochemistry. While higher DOM concentrations in incubation experiments corresponded to a

greater proportion of Deltaproteobacteria, Chloroflexi comprised a greater proportion of bacterial communities in environmental samples. Also, this research provided evidence suggesting that Chloroflexi, Epsilonproteobacteria and Deltaproteobacteria may influence groundwater As cycling through the respiratory reduction of quinone moieties in the high As aquifer environment. The combination of fluorescence spectroscopy and PARAFAC modeling, and 16S rRNA gene pyrosequencing was useful in deciphering novel relationships between bacterial taxa and DOM chemistry in the environment. Future research on geogenic As contamination should investigate the metabolic activities of Deltaproteobacteria, Chloroflexi, Epsilonproteobacteria under environmental conditions in high As aquifer environments.

**ACKNOWLEDGEMENTS**

This dissertation could not have been completed without the incredible help and support of many people. First and foremost, I would like to thank my advisor, Diana Nemergut, for her intellectual, emotional, and financial support for me, as well as for my dissertation research, over the last five years. I have learned so much from Diana, and since becoming her advisee I have transformed from a hydrologist to a microbial ecologist. I am so thankful for the ways in which Diana has encouraged me to take on a number of challenging, but enriching professional experiences, from attending and presenting at international conferences to learning novel laboratory techniques to taking difficult courses, which have made me into the scientist and person I am today. Her advising has allowed me to feel confident to embark on the next phase of my career, teaching and leading a research program with undergraduate students at St. Edward's University. The level of attention and care that Diana shows her students is truly a model for how I hope to support my own future advisees.

I would also like to thank Diane McKnight for her help and advising during this research. As a result of Diane's influence, I now have a profound appreciation for the importance of dissolved organic matter for understanding ecosystem processes, in addition to my appreciation for how difficult it is to characterize dissolved organic matter chemistry. Also, I admire Diane for her commitment to conducting research that she recognizes as environmentally relevant and important, even if it is not widely recognized as "hot" or fundable at the time. Diane's innovative approach to research has really influenced the way in which I think about my own research, and I am so grateful for her advising over the past five years.

The other members of my dissertation committee, Dr. Steve Schmidt, Dr. Richard Smith and Dr. William Lewis have also provided me with invaluable advice and input during my dissertation research. I would like to thank Steve Schmidt for his technical and intellectual input on the field and the laboratory components of my dissertation research. Richard Smith has been very generous with his time over the past five years, and has talked through the merits of various experimental designs and field research plans. In addition, Dr. Smith allowed me to use the anaerobic chamber in his laboratory, which was critical to completing the third chapter of this dissertation. Dr. William Lewis provided excellent feedback on the experimental design of my dissertation research, as well as professional and editorial advice. Dr Alexis Templeton also provided significant support and assistance of this work, and I learned an incredible amount about geochemistry from our conversations.

I received help and support during my dissertation from many other collaborators, colleagues and friends. I worked with Bailey Simone, Natalie Mladenov, Yan Zheng, and Jessica Ebert on the field research and data analysis for this work. Also, Joey Knelman, Lee Stanish, Kelly Ramirez, Jon Leff, Dan Knights and Antonio Gonzales provided invaluable help with the analysis of 16S rRNA gene sequences, as well as the statistical analyses. I was incredibly lucky to interact with an exceptional group of faculty, graduate students, post-docs and research technicians at INSTAAR, especially those in the North hallway of RL-1. Other friends in my science community including Lisa Mayhew, Elizabeth Swanner, Katherine Wright, Jennifer Underwood, Kate Campbell and Deborah Repert helped provide moral and technical support throughout my dissertation research.

I belong to a family of educators and academics, and I am very grateful for their love, and their support for my academic aspirations. My grandparents have provided such an exceptional

example of the value of education to provide positive change in people’s lives, and I aspire to have the same dedication to teaching, research and learning as my grandfather. In addition, my parents, sister and brother-in-law have always been there for me, encouraging me every step of the way. I am so lucky to be surrounded by people who understand and appreciate the ups and downs of the dissertation process, and who understand why I have chosen this professional path.

It is impossible for me to find enough words to thank my husband Greg for his unwavering love and support. I have learned so much from him about how to strike a balance between family life and professional life, and I am much happier, balanced and successful person as a result. I don’t think I would be finishing my degree now without the support from Greg.

**TABLE OF CONTENTS**

**CHAPTER 1: Introduction……………………………………………………………………1**

**1.1 Geogenic arsenic contamination…………………………………………...………1**

**1.2 The environmental chemistry of arsenic..……….……………….…………………..1**

**1.3 The origin of arsenic in Southeast Asia………...…………………………………....2**

**1.4 Hydrological processes influencing groundwater arsenic mobilization……………...3**

**1.5 Biogeochemical cycling of groundwater arsenic………………………………………4**

**1.6 The chemistry and source of dissolved organic matter………………………………..5**

**1.7 The microbial ecology of high arsenic aquifers……………………………………...7**

**1.8 Research Objectives…………………………………………………………………...8**

**CHAPTER 2: Carbon, metals and grain size correlate with bacterial community structure in sediments of a high As aquifer…………………………….…………….....12**

**2.1 Introduction…………………….……………………….……………………………12**

**2.2 Materials and Methods……………………………….……….........................15**

2.2.1 Site Description**.........................………………………..…………………15**

2.2.2 Sample Collection and Preparation**...……………………….................................17**

2.2.3 Sediment Chemical and Grain Size Analyses**…….....……...…................................18**

2.2.4 Phylogenetic Data Analysis**………………...………………………………………19**

2.2.5 Statistical Analysis……....……………………………………………………………21

**2.3 Results**……………………………………....……………………………………..…22

2.3.1 Sediment Grain Size and Chemistry..………….…………….................................22

2.3.2 Bacterial Community Characteristics………………………………….......…...27

**2.4 Discussion**……………………………………………………………….....……......33

**2.5 Conclusions**…………………………………………………….............................38

**CHAPTER 3: Labile and redox-active dissolved organic matter influence bacterial community structure and groundwater arsenic mobilization**……………….39

**3.1 Introduction**……………………………………………………..........................39

**3.2 Materials and Methods**………………………………………………………………43

3.2.1 Site Description………………………………………………………………….43

3.2.2 Sample Collection…………………………………………………………….…47

3.2.3 Incubation Experiment Preparation……………………………………………..47

3.2.4 Geochemical Analyses…………………………………………………………..49

3.2.5 Dissolved Organic Matter Characterization…………………………………..….50

3.2.6 Bacterial Community Analysis…………………………………………………51

3.2.7 Statistical Analyses……………………………………………………………..…53

**3.3 Results**………………………………………………………………………………...54

3.3.1 DOM chemistry in pre-incubation samples……………………………………….....54

3.3.2 Sediment geochemistry in post-incubation samples………………………………55

3.3.3 Groundwater geochemistry in post-incubation samples…………………………...57

3.3.4 DOM incubation treatments significantly influenced bacterial community structure……………………………………………………………………………61

3.3.5 Relationships between bacterial community structure and fluorescent DOM constituents………………………………………………………………………...67

**3.4 Discussion……………………………………………………………………............69**

**CHAPTER 4: Novel relationships between dissolved organic matter chemistry and bacterial community composition in a high arsenic aquifer: Implications for groundwater arsenic cycling in Bangladesh……...........................................75**

**4.1 Introduction…………………………………………………….........................75**

**4.2 Materials and Methods…………………………………………………………81**

4.2.1 Site Description……………………………………………………………….81

4.2.2 Sample Collection………………………………………………………….…84

4.2.3 Dissolved Organic Matter Analyses………………………………………..…84

4.2.4 Bacterial Community Analysis……………………………………………..…85

4.2.5 Statistical Analyses……………………………………………………….....…88

**4.3 Results…………………………………………………………………………...89**

4.3.1 DOM chemistry……………………………..……………………………......89

4.3.2 Bacterial Community Characteristics………………..…………………………93

4.3.3 DOM – Bacterial Relationships……………………….……………………...98

**4.4 Discussion……………………………………………………………….........106**

**CHAPTER 5: Conclusions…………………………...…………………………………113**

**5.1 Research Summary…………………………………………………….……...113**

**5.2 Future Research Directions……………………………………………….….116**

**REFERENCES……………………………………………………………………….…..120**

**APPENDIX 1………………………………………………………………………..150**

## LIST OF TABLES

**Table 2.1:** The correlation coefficients and significance values for the pairwise correlations between environmental and bacterial community data……………………….**26**

**Table 3.1:** Groundwater chemistry in pre-incubation groundwater from the four sampling locations……………………………………………………………………….**45**

**Table 3.2:** Aquifer sediment chemistry in pre-incubation samples from the four sampling locations……………………………………………………………………….**45**

**Table 3.3:** Dissolved organic matter chemistry in pre-incubation groundwater samples from the four sampling locations…………………………………………………...**46**

**Table 3.4:** The 5 components of the PARAFAC model used in this study(Ebert, 2012)…...**46**

**Table 3.5:** The percent carbon and nitrogen in post-incubation aquifer sediments……....**56**

**Table 3.6:** The iron and arsenic concentrations in post-incubation aquifer sediments……..**58**

**Table 3.7:** The phylogenetic diversity within pre- and post-incubation bacterial communities…………………………………………………………………......**65**

**Table 3.8:** The mean and standard deviations for UniFrac distances between post-incubation bacterial communities in samples under different DOM treatments…………….**65**

**Table 3.9:** PARAFAC model loadings in control and labile DOM-amended post-incubation groundwater……………………………………………………………….….**68**

**Table 3.10:** The FI and SUVA of control and labile DOM-amended post-incubation groundwater……………………………………………………………………**68**

**Table 4.1:** Water chemistry for groundwater collected at each sampling location…………**83**

**Table 4.2:** A description of the 13 components of the Cory & McKnight (2005) PARAFAC model………………………………………………………………………….**86**

**Table 4.3:** The fluorescence index and amino acid- like fluorescence in groundwater collected from each sampling location……………………………………….**91**

**Table 4.4:** PARAFAC model loadings results for groundwater collected at each sampling location……………………………………………………………………**92**

**Table 4.5:** Linear relationships and Spearman rank-order correlation results for the relationship between groundwater arsenic concentrations and DOC, and PARAFAC model loadings……………………………………………………**94**

**Table 4.6:** The relative abundances of bacterial phyla in aquifer sediment samples………..**96**

**Table 4.7:** The relative abundances of classes within the Proteobacteria in aquifer sediment samples……………………………………………………………………..**97**

**Table 4.8:** Spearman rank-order correlation results testing the relationship between PARAFAC model loadings in groundwater and the relative abundances of Chloroflexi, Dehalococcoidetes and Epsilonproteobacteria in aquifer sediment samples………………………………………………………………………**104**

**Table 4.9:** Spearman rank-order correlation results testing the relationship between the C:N in groundwater and the relative abundances of Epsilonproteobacteria and Gammaproteobacteria in aquifer sediment samples……………………………..**105**

## LIST OF FIGURES

**Figure 2.1:** The study area, Site K, in Araihazar, Bangladesh……………..........................**16**

**Figure 2.2:** Grain size distribution and soil moisture of sediments collected at the three sampling locations along the groundwater flowpath..........................................**23**

**Figure 2.3:** Sediment chemistry at each sampling location along the groundwater flowpath........................................................................................................**25**

**Figure 2.4:** Bacterial community structure is significantly at the three sampling locations...**28**

**Figure 2.5:** A Venn diagram showing the number of OTUs that are shared between sampling locations, and unique to a single sampling location...........................................**30**

**Figure 2.6:** A non-metric multidimensional scaling (NMDS) ordination showing that bacterial communities cluster based on well location, and that sediment geochemistry is related to the composition of bacterial communities.................**32**

**Figure 3.1:** The sampling locations at Site K and Site B, Araihazar, Bangladesh................**44**

**Figure 3.2:** Groundwater iron concentrations in post-incubation samples............................**59**

**Figure 3.3:** Groundwater arsenic concentrations in post-incubation samples.......................**60**

**Figure 3.4:** The relative abundances of bacterial taxa in pre-incubation and post-incubation samples from each sampling site.........................................................................**62**

**Figure 3.5:** A NMDS ordination showing that post-incubation bacterial communities cluster based on DOM treatment……………………………………………………...**66**

**Figure 4.1:** The sampling locations at Site K and Site B in Chapter 4.................................**82**

**Figure 4.2:** A schematic explaining the Mantel test method…………………………….....**90**

**Figure 4.3:** The phylogenetic diversity in bacterial communities in each aquifer sediment sample...............................................................................................................**95**

**Figure 4.4:** There was a significant relationship between bacterial community structure and tryptophan-like fluorescence...................................................................................**99**

**Figure 4.5:** There was a significant relationship between bacterial community structure and oxidized quinone fluorescence...........................................................................**100**

**Figure 4.6:** There was a significant relationship between bacterial community structure and reduced quinone fluorescence............................................................................**101**

**Figure 4.7:** There was a significant relationship between bacterial community structure and two unidentified fluorophores in the Cory & McKnight (2005) PARAFAC model. ......................................................................................................................**102**

# CHAPTER 1: Introduction

## 1.1 Geogenic arsenic contamination

Throughout the deltaic regions of Southeast Asia the primary source of drinking water is groundwater because surface waters contain high levels of human pathogens (Nordstrom, 2002) due to runoff from agricultural fields and household waste. In fact, since the 1960's residents in rural Bangladesh and West Bengal have drilled shallow tube wells in order to access an alternative, presumably safer, drinking water source (Mukherjee & Bhattacharya, 2001). Unfortunately, residents began to develop a variety of ailments, including skin lesions, which were eventually traced to arsenic (As) poisoning (Chakraborti et al., 2002; Smith et al., 2000). Today, more than 100 million people in the Ganges-Meghna-Brahmaputra Delta (GMBD) region in Bangladesh and West Bengal, the Red River Delta (RRD) in Vietnam, and the Mekong River Delta (MKD) in Vietnam and Cambodia consume groundwater containing unsafe levels of As, according to the World Health Organization regulatory guideline (10 μg/l) (Berg et al., 2007; Nickson et al., 2011). The regular consumption of unsafe levels of As causes serious illnesses such as lung, bladder and skin cancer, vascular disease and diabetes (Anawar et al., 2002; Hopenhayn, 2006), and negatively affects the intellectual development of children (Wasserman et al., 2004). Therefore understanding the processes that promote As mobilization in groundwater aquifers is vital to managing water supplies and protecting human health.

## 1.2 The environmental chemistry of arsenic

Arsenic is a metalloid found in the Earth's crust, as well as throughout the atmosphere, biosphere and hydrosphere (Cullen & Reimer, 1989). The natural sources of arsenic to soil, groundwater and surface water include geothermal fluids and the weathering products of arsenic-rich minerals such as arsenopyrite. The average natural arsenic concentrations in soils are

between five and ten parts per million (equivalent to 5 to 10 μg/l) whereas soils derived from arsenic-rich parent material can contain up to 45 parts per million of arsenic (Smedley & Kinniburgh, 2002). Natural groundwater and surface water arsenic concentrations are generally an order of magnitude lower than those found in soils.

The chemical behavior of arsenic (As) in the environment is primarily influenced by pH and the oxidation-reduction (redox) state of the environment (Cullen & Reimer, 1989). The predominant oxidation states of As, As(V) and As(III), form a variety of oxyanions which significantly influence the mobility of As in aquatic environments (O'Day, 2006). Arsenate ($H_3AsO_4$, an As(V) oxyanion) is the most common form of As in neutral pH environments. Arsenate is generally immobilized under oxidizing conditions because it has a high sorption affinity for Fe- oxide minerals (Palmer et al., 2006). Under reducing conditions, for example anaerobic groundwater, arsenate has a slightly lower sorption affinity for Fe-oxides and can be more readily desorbed from mineral surfaces. At neutral pH, arsenious acid ($H_3AsO_3$, an As(III) oxyanion) is generally only observed under reducing conditions. Arsenious acid is more mobile than arsenate since it is stable and neutrally-charged, and as a result, does not have a high sorption affinity for Fe- oxides (O'Day, 2006).

### 1.3 The origin of arsenic in Southeast Asia

The As within the shallow (<100 m) unconfined aquifers of Southeast Asia originates from As- rich metal-sulfide minerals in the Himalayas, such as arsenopyrite, which were weathered and transported through fluvial networks during the Holocene glacial retreat (~10,000 ya) (Stanger, 2005). The weathered material has been deposited within the GMBD, the RRD and MKD throughout the millennia, and today, the As is sorbed to iron (Fe)-oxide minerals which coat deltaic aquifer sediments (Charlet & Polya, 2006). Research throughout the past decade has

demonstrated that a suite of hydrologic and biogeochemical processes contributes to the release of As from Fe-oxide mineral surfaces in high As aquifers. In general, the scientific research on groundwater As mobilization has been conducted in the GMBD since geogenic As contamination in Southeast Asia was first recognized there (Nickson et al., 1998). However research from the RRD and MKD has also enhanced our understanding of groundwater As mobilization in this region.

### 1.4 Hydrologic processes influencing groundwater arsenic mobilization

The deltaic aquifers in Southeast Asia are recharged from regional and local sources. Regional groundwater recharge rates are primarily influenced by topography, and there is very little relief in the regions surrounding these large river deltas. For example, hydraulic gradients within the recharge zone of the GMBD are $\sim 10^{-4}$ (Burgess et al., 2002; Harvey, 2002) and thus regional recharge rates are relatively slow (Ahmed et al., 2004), with estimates ranging from 400 – 1800 cm/yr (Ravenscroft et al., 2004; van Geen et al., 2008). Local recharge rates vary between 0.05 to 1m/yr within the GMBD (Stute et al., 2007; van Geen et al., 2006a), and the high degree of variability is likely due to differing rates of groundwater extraction.

The primary local sources of groundwater recharge are adjacent rivers, man-made ponds, and seasonal floods from monsoonal rains (Ahmed et al., 2004). In general, As concentrations in surface water recharge sources, which are oxygenated, are below 10 μg/l. However, there is a strong relationship between groundwater age and groundwater As concentrations, suggesting that there is an accumulation of As with lateral groundwater flow (Stute et al., 2007), and a transport of As-laden groundwater to depth (Polizzotto et al., 2008). It appears that the factors influencing lateral and vertical flow and recharge rates, such as sediment grain size and human activities such as irrigation pumping, are also related to groundwater As concentrations in the Southeast

Asia deltaic aquifers. For example, groundwater flow rates are slower in fine-grained, silty sediments, which are less permeable than sandy sediments, and pore water As concentrations are also higher within fine-grained sediments. By contrast, groundwater pumping for irrigation increases the rate of vertical groundwater flow, and promotes the flow of organic-rich pond water into the aquifer, which may increase rates of groundwater As mobilization (Harvey et al., 2002; Neumann et al., 2010). Research suggests that patchiness in sediment grain size, and hydrologic perturbations due to irrigation pumping, are partially responsible for the considerable spatial heterogeneity in groundwater As concentrations observed in high As aquifers in Southeast Asia.

**1.5 Biogeochemical cycling of groundwater arsenic**

A suite of biogeochemical processes promotes groundwater As mobilization in the deltaic aquifers in Southeast Asia. The current paradigm is that the native Fe-reducing bacteria in aquifer sediments cause the reductive dissolution of Fe-oxide minerals, which results in the release of sorbed As under anaerobic, reducing conditions (Fendorf et al., 2010). In addition, research throughout the past decade has demonstrated that the rates of bacterial Fe-reduction and groundwater As mobilization are largely influenced by the availability of organic matter (Islam et al., 2004; McArthur et al., 2004; Neumann et al., 2010; Pederick et al., 2007; Rowland et al., 2007).

First, heterotrophic bacteria in the near-surface depths of the aquifer use organic matter sources for carbon and nitrogen to synthesize biomass, and for electron donors during the respiratory reduction of oxygen. Heterotrophic activity rapidly consumes oxygen and results in anaerobic, reducing conditions in the saturated zone of the aquifer. Under reducing conditions, Fe-reducing bacteria pair the oxidation of labile fractions of dissolved organic matter (DOM),

such as low-molecular weight organic acids, to the reduction of Fe(III) within Fe-oxide minerals (Islam et al., 2004; Lovley, 1991). Also, it is likely that native bacteria utilize labile DOM as an electron donor during dissimilatory As(V)-reduction in high As aquifer in Southeast Asia (Lear et al., 2007).

More recently, research has indicated that bacteria may use more recalcitrant fractions of DOM while mediating groundwater As cycling. In anaerobic aquatic environments, a diverse group of bacteria generate energy by pairing the oxidation of labile DOM to the reduction of redox-active quinone moieties within humic sustances (Cervantes et al., 2002; Coates et al., 1998; Lovley et al., 1996; Scott et al., 1998). It is likely that quinone-reduction is phylogenetically widespread amongst bacteria in neutral pH, anaerobic aquatic sediments because other electron acceptors are insoluble (Lovley et al., 1996; Kappler et al., 2004), and quinone-reduction may yield more energy than the reduction of metal-oxide minerals (Straub et al., 2005). Bacterial quinone reduction promotes an electron cascade in which reduced quinone moieties abiotically transfer electrons to Fe(III), and cause the reductive dissolution of Fe-oxide minerals (Lovley et al., 1998). Actually, evidence from laboratory experiments suggest that abiotic Fe(III)-reduction via reduced quinones is significantly faster than bacterial Fe-reduction (Jiang & Kappler, 2008). Thus, in the high As aquifers of Southeast Asia, bacterial quinone reduction could be responsible for accelerating rates of Fe-oxide dissolution and subsequently the release of As into groundwater. In fact, reduced quinone moieties have been discovered in high As groundwater in the GMBD (Mladenov et al., 2010).

### 1.6 The chemistry and source of dissolved organic matter

DOM is a chemically heterogeneous mixture of constituents with a wide range of chemical and ecological properties. The break-down of plant material and microbial biomass represent the

primary sources of DOM in aquatic environments. However, recent research suggests that DOM compounds excreted by bacteria during respiration may also constitute a significant proportion of the total DOM pool (Guillemette & Del Giorgio, 2012; McKnight et al., 2001; Stedmon & Markager, 2005). The labile fraction of DOM, which is rapidly degraded by bacteria in the environment, is composed of low molecular-weight compounds such as carbohydrates, proteins, amino acids, and lipids (Berggren et al., 2010; Coffin, 1989; Keil & Kirchman, 1991). By contrast, humic substances, which comprise between 60% and 80% of DOM, are relatively resistant to degradation by bacteria (Mulholland, 2003), although some estuarine and lake bacteria can utilize humic substances for carbon and nitrogen sources (Esham et al., 2000; Hutalle-Schmelzer et al., 2010).

In the GMBD, local groundwater recharge from ponds, rivers and seasonal flooding provides fresh sources of DOM to aquifers. Research indicates that DOM from these sources is primarily labile (Mladenov et al., 2010; Neumann et al., 2010). Latrine waste represents another source of labile DOM to shallow depths within aquifers (McArthur et al., 2004), and can also introduce human-associated bacteria including intestinal pathogens into groundwater (Knappett et al., 2011). The release of sedimentary organic matter during the reductive dissolution of Fe-oxides also contributes to the groundwater DOM pool. Sedimentary organic matter released from aquifer sediments near the ground surface is generally fairly labile (Polizzotto et al., 2008). Hence, it is likely that DOM from local surface water recharge sources and near-surface sediments provides carbon and nitrogen sources for synthesizing biomass, and electron donors for respiration. By contrast, organic matter released from sediments deeper within aquifers is more aromatic and contains humic substances (McArthur et al., 2004; Mladenov et al., 2010), and as a result, may primarily serve as an electron donor for respiratory quinone reduction.

## 1.7 The microbial ecology of high arsenic aquifers

The composition and activity of bacterial communities in high As aquifers appears to be influenced by the availability of DOM. Previous research indicates that bacterial communities in aquifer sediments are diverse, and a majority of the community is comprised of oligotrophic taxa that have not been previously identified as Fe- or As-reducers. For example, Betaproteobacteria were the dominant members of sediments collected from aquifers in the GMBD (Héry et al., 2010), and the MKD (Lear et al., 2007; Rowland et al., 2007), while Gammaproteobacteria, Alphaproteobacteria, Bacteriodetes, and Firmicutes were also relatively abundant in bacterial communities from a number of aquifers affected by geogenic As contamination in Southeast Asia (Gault et al., 2005; Islam et al., 2004; Lear et al., 2007; Rowland et al., 2007). In environmental samples from high As aquifers, Deltaproteobacteria have been comparatively less abundant than other taxa even though there are many known Fe- and As-reducers within this bacterial class (Gault et al., 2005; Héry et al., 2010; Pederick et al., 2007; Rowland et al., 2007). However, results from incubation experiments indicate that the abundance and activity of Deltaproteobacteria in the environment may be hindered by a low supply of DOM in high As aquifers.

Multiple experiments have shown that increases in labile DOM concentrations cause significant shifts in the composition of aquifer sediment bacterial communities, and an increase in the relative abundance of Deltaproteobacteria. For example, dramatic increases in the proportion of taxa closely related to known Fe- and As-reducers in the *Geobacter* genus, and corresponding increases in Fe-reduction and As mobilization rates, were observed in incubations amended with sodium acetate, a proxy for labile carbon (Héry et al., 2010; Islam et al., 2004; Pederick et al., 2007; Rowland et al., 2007). To a lesser extent, acetate additions seemed to

stimulate Clostridia related to known fermenting bacteria (Héry et al., 2010), and Gammaproteobacteria in the *Pseudomonas* genus (Lear et al., 2007; Rowland et al., 2007). Also, the relative abundances of Deltaproteobacteria and Gammaproteobacteria were higher in incubations amended with labile carbon and a model quinone compound than in incubations amended with just acetate (Rowland et al., 2007). Thus, results from these experiments suggest that the chemistry and availability of DOM influences the abundance of bacterial taxa such as Deltaproteobacteria, in addition to other subphyla within the Proteobacteria, that may promote groundwater As mobilization in high As aquifers.

**1.8 Research Objectives**

Previous research findings have led to a number of questions regarding the role of bacterial communities in mediating the biogeochemical cycling of groundwater As in Southeast Asia. First, it is not clear what proportion of the diverse bacteria recovered in aquifer sediments (Gault et al., 2005; Héry et al., 2010; Pederick et al., 2007; Rowland et al., 2007) are involved in promoting groundwater As mobilization since a majority of the taxa are not related to known respiratory Fe-, As- or quinone-reducers. In addition, although research has shown that spatial gradients in sediment geochemistry affect abiotic As mobilization processes (van Geen et al., 2003; van Geen et al., 2006b; Zheng et al., 2005), the influence of these gradients on the composition and activity of bacterial communities in high As aquifers is not understood. Finally, additional research is needed to understand how different chemical fractions of DOM, such as labile DOM from local pond recharge (Neumann et al., 2010) or redox- active, quinone-containing DOM from sediments (Mladenov et al., 2010), affect the composition and activity of bacterial communities across groundwater As concentrations. Identifying the source of DOM

that fuels bacterial-mediated As mobilization in groundwater could be critical to developing mitigation strategies for geogenic As contamination throughout Southeast Asia.

The primary goals of this dissertation research were to address these outstanding issues using interdisciplinary methods including analytical chemistry, 16S rRNA gene sequencing (Pace, 1997) to characterize the composition and diversity of bacterial communities, and fluorescence spectroscopy and PARAFAC modeling (Cory & McKnight, 2005) for describing the chemistry and source of DOM in high As aquifers. To this end, the research presented in Chapter 2 investigated patterns in bacterial community composition in environmental samples collected across geochemical gradients in an aquifer with elevated groundwater As concentrations in Araihazar, Bangladesh. This work found that differences in sediment grain size, and sediment C, Mn and Fe concentrations significantly influenced the structure of bacterial communities, including the relative abundances of Deltaproteobacteria and Chloroflexi. The fact that sediment grain size, C, Mn and Fe also influence the mobility of groundwater arsenic supports a strong link between geochemical and biological processes in high As aquifer environments in Southeast Asia.

I followed up on these research findings by conducting a full factorial incubation experiment designed to evaluate the relative effects of DOM chemistry versus sediment geochemistry on bacterial community structure and Fe- and As mobilization rates. As presented in Chapter 3, I incubated aquifer sediment and groundwater collected from four sites that represented the range of geochemical conditions observed within my study sites in Araihazar, Bangladesh with and without labile and redox-active DOM. From this research I discovered that the structure of bacterial communities from different sampling sites converged based on DOM chemistry, suggesting that DOM chemistry had a greater influence on the composition of native aquifer

bacteria than sediment chemistry. Also, the supply of both labile and redox- active DOM corresponded to significantly higher relative abundances of Deltaproteobacteria and Fe and As mobilization rates, indicating that these taxa may accelerate As mobilization through respiratory quinone reduction. Significant relationships between relative abundances of Chloroflexi and Deltaproteobacteria and the availability of specific redox-active fluorophores in post-incubation groundwater provided additional evidence for bacterial quinone reduction in high As groundwater.

Next, I was interested in comparing results from the incubation experiment described in Chapter 3 to relationships between bacterial taxa and DOM chemistry in the aquifer environment. Thus, in environmental samples collected from established study sites in Araihazar, Bangladesh I characterized the chemistry and source of groundwater DOM using a combination of fluorescence spectroscopy and PARAFAC modeling, and I employed 16S rRNA gene sequencing in order to characterize the structure of bacterial communities. I hypothesized that Deltaproteobacteria and Chloroflexi performed respiratory quinone reduction in the aquifer, and thus their relative abundances would be significantly related to the availability of labile DOM from pond recharge and quinone DOM originating from the release of sedimentary DOM. By contrast, other taxa may primarily influence As cycling by using these labile and quinone-containing DOM sources for carbon and nitrogen to synthesize biomass. In Chapter 4, I present results which suggest that labile DOM and quinone moieties in humic DOM are more important in shaping bacterial community structure than the source of DOM in high As aquifers. Also, it appears that Chloroflexi and Espilonproteobacteria may be more involved in accelerating groundwater As mobilization through respiratory quinone-reduction than Deltaproteobacteria.

In total, this dissertation research has demonstrated that interdisciplinary research approaches, such as combining analytical geochemistry, 16S rRNA gene sequencing, and fluorescence spectroscopy with PARAFAC modeling, can provide new insights into the factors promoting geogenic As contamination in aquifers throughout Southeast Asia. In addition, this work has demonstrated that the comparison of experimental results to results from environmental sampling efforts is critical for accurately characterizing biogeochemical processes. Future research on geogenic As contamination should focus on describing the metabolic activities of Deltaproteobacteria, Chloroflexi, Epsilonproteobacteria under environmental conditions in high As aquifer environments.

## CHAPTER 2: Carbon, metals and grain size correlate with bacterial community structure in sediments of a high arsenic aquifer

Legg, T.M., Y. Zheng, B. Simone, K.A. Radloff, N. Mladenov, A. González, D. Knights, H. Siu, M.M. Rahman, K.M. Ahmed, D.M. McKnight, D.R. Nemergut. 2012. *Frontiers in Microbiology* 3(82): 1-15.

### 2.1 Introduction

Throughout the last decade considerable research effort has focused on characterizing the mechanisms leading to elevated groundwater arsenic (As) concentrations in aquifers throughout South Asia. These studies have demonstrated that groundwater As mobility is affected by a number of factors including redox conditions (Zheng et al., 2004), sediment properties (Winkel et al., 2008), hydrology (Polizzotto et al., 2005), organic matter quality (Neumann et al., 2010; Mladenov et al., 2010), and microbial activity (Dhar et al., 2011; Islam et al., 2004). In the Ganges-Brahmaputra Delta (GBD) region of Bangladesh, where groundwater As concentrations are on average about 10 times higher than the World Health Organization's drinking water guideline value, groundwater As concentrations are heterogeneous (van Geen et al., 2003) and often decoupled from bulk sediment As concentrations (Dhar et al., 2008). In the GBD, areas with high groundwater As concentrations typically share similar sediment characteristics, hydrology and organic matter chemistry. Sediment layers within aquifers containing higher proportions of fine-grained material such as silts generally feature higher groundwater As concentrations than sands (van Geen et al., 2006b), which, by contrast, tend to have lower organic matter concentrations. Silty layers have a lower permeability, thereby favoring the accumulation of dissolved As in the groundwater (van Geen et al., 2006a). In addition, fine-grained sediments often exhibit higher concentrations of sediment organic matter and metals

including iron (Fe), manganese (Mn) and As (McArthur et al., 2004), which may fuel relevant microbial processes.

Indeed, it is clear that microbial processes are also important in regulating As mobilization in the GBD (Radloff et al., 2008; Islam et al., 2004). Groundwater As mobilization may be enhanced in fine sediments, as microorganisms pair the oxidation of organic carbon (C) to the reductive dissolution of Fe oxides (Lovley & Phillips, 1988), which liberates As from Fe oxide mineral surfaces (Tufano & Fendorf, 2008; McArthur et al., 2001; Cummings et al., 1999). To a lesser extent, the desorption of As from the microbial reduction of Mn oxides may also promote elevated As concentrations (Luna et al., 2009; Inskeep et al., 2002), and the release of reduced Mn into groundwater often occurs with the release of As (van Geen et al., 2006a; Zheng et al., 2004). In addition, microorganisms reduce humic substances (HS) (Mladenov et al., 2010) which promotes an electron cascade resulting in Fe oxide reduction and As desorption (Jiang & Kappler, 2008; Kappler et al., 2004), as well as the reduction of As(V) to the more mobile As(III) (Jiang et al., 2009; ). Microorganisms can also enhance As mobility in anaerobic aquifers by mediating redox reactions with As. Specifically, the detoxification pathway encoded by the *ars* operon (Macy et al., 2000; Rosen et al., 1991; Sun et al., 2004) and dissimilatory As(V) reduction (Saltikov & Newman, 2003) results in the reduction of As(V) to the more mobile As(III). Also, microorganisms transform As species during methylation (Mukhopadhyay et al., 2002) although evidence of methylation has not been observed in the GBD environment (Islam et al., 2004). By contrast, microbial sulfate reduction can decrease the mobility of As in groundwater (Kirk et al., 2004; Saalfield & Bostick, 2009). Sulfide co-precipitates with As(III) and Fe to form As trisulfide (Newman et al., 1997) under anoxic, reducing conditions (Rittle et al., 1995). Given the diverse array of microbial metabolisms with potential affects on As cycling,

a better understanding of the patterns in microbial community structure across sediment characteristics could help elucidate the roles of specific taxa involved in biogeochemical processes that affect As mobility in these anaerobic aquifers.

Recent research investigating the role of microbial communities in groundwater As cycling has used molecular phylogenetic tools to characterize the taxonomic composition of microbial communities in the GBD groundwater environment. The analysis of SSU rRNA gene sequences from microcosm experiments with As-rich aquifer sediment show the addition of a labile C source promotes bacterial community shifts characterized by an increased proportion of Deltaproteobacteria, and corresponding increases in Fe-reduction and As mobilization (Islam et al., 2004; Lear et al., 2007; Rowland et al., 2007), as well as bacterial sulfate reduction (Héry et al., 2010). These microcosm results indicate that patterns in bacterial community composition are related to changes in As cycling in groundwater environments. However, questions remain about the types of microorganisms found in ambient conditions in the environment, and how their distribution varies with native geochemical conditions.

Thus, the goal of this work was to advance our understanding of microbial communities in As-rich groundwater sediments by combining high-throughput gene sequencing methods with environmental chemistry and statistical analysis. I was especially interested in the investigation of relationships between environmental chemistry and microbial community structure at the level of the entire aquifer since previous research at my study site has shown that groundwater arsenic increases with aquifer depth and location along the flowpath (Mladenov et al., 2010; Radloff et al., 2008), and that dissolved organic matter chemistry changes with depth (Mladenov et al., 2010). I used the natural environmental gradients created by the groundwater flowpath and depth within the aquifer to demonstrate that patterns in bacterial community composition are correlated

to sediment chemistry across a groundwater As concentration gradient in a GBD aquifer. The results from this work show that bacterial community structure is significantly different between locations in the groundwater aquifer. In addition, I demonstrate that variations in sediment grain size, as well as sediment C, Mn and Fe concentrations correspond to variations in bacterial community structure.

## 2.2 Materials and Methods

### 2.2.1 *Site Description*

This research was conducted at Site K (Radloff et al., 2008) in Araihazar, Bangladesh, approximately 30 km northeast of Dhaka (23˚ 47’ 34” N, 90˚ 37’ 48” E). The regional climate is monsoonal, and receives more than 50% of the annual precipitation (average of 2354 mm) between June and September (Immerzeel, 2008). Consequently, like similar sites in Araihazar (Stute et al., 2007), stream and groundwater levels at Site K vary seasonally and peak during July and August when most of the study site is flooded. Site K is located in a rural area within the floodplain of the Old Brahmaputra River, an abandoned river channel that has been filled through sedimentation and reduced to a small stream (Figure 2.1). Previous research at Site K has extensively characterized the hydrology and groundwater geochemistry (Radloff et al., 2008; Radloff, 2010). Like elsewhere in the GBD, shallow groundwater As concentrations are spatially variable, and can exceed 400 µg/L (Radloff et al., 2008). I examined groundwater and sediment samples from three monitoring nests along a groundwater flowpath at the study site. Well nest K240 is located 240 m from the river within a village, whereas well nests K150 and K60, 150 and 60 m from the river respectively, are located within cultivated rice fields.

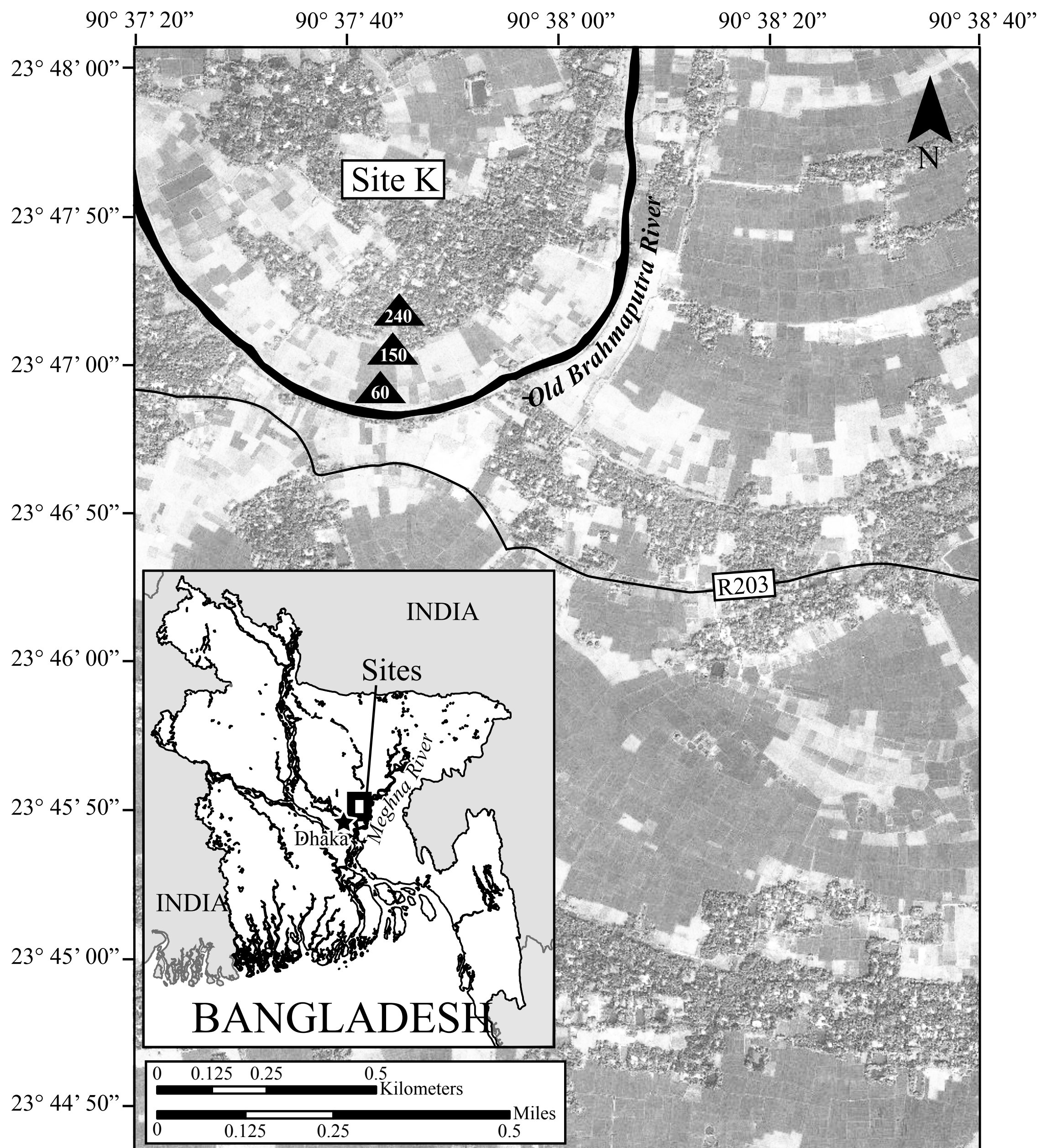



Figure 2.1. The study area, Site K, in Araihazar, Bangladesh.

Based on groundwater age estimates, the mean direction of groundwater flow at the study site is from the village to the Old Brahmaputra River (Radloff, 2010). Thus, K240 is located near the beginning of the groundwater flowpath, K150 is located in the middle, and K60 is located at the end of the flowpath, directly upgradient of the river (Figure 2.1). During the wet season (approximately May-October), all three well nests at Site K are completely saturated due to groundwater table rise, with the exception of depths between 0 and 0.5 m within the village. In February, at the height of the dry season (Stute et al., 2007), depths between 0 and 3 m below the ground surface are unsaturated at all sampling locations at Site K.

The groundwater chemistry at Site K changes with location along the flowpath and depth within the aquifer (Radloff, 2010). Like other sites in Araihazar (Dhar et al., 2008), there are not marked seasonal changes in groundwater As concentrations (Radloff, 2010). Dissolved groundwater Fe and As concentrations are lowest at shallow depths at K240, and generally increase with depth at each well location (Radloff, 2010). In addition, groundwater As concentrations increase along the groundwater flowpath, and peak at K60 at 15 m (429 μg/L). Groundwater Mn, by contrast, does not show patterns with depth or position along the flowpath (Radloff, 2010).

2.2.2 *Sample Collection and Preparation*

I used the natural environmental gradients created by the groundwater flowpath and depth within the aquifer to examine the relationship between sediment chemistry and bacterial community composition. I collected sediment cores from three different sites within the aquifer, K240, K150 and K60 (Figure 2.1), adjacent to the monitoring well nests in July 2008. At each location I collected cores from seven aquifer depths: 1.52 m, 3.05 m, 4.57 m, 6.10 m, 7.62 m, 10.67 m, and 15.24 m. I rinsed the sediment core liners with 100% ethanol prior to placing them

in the coring device in order to minimize contamination. Also, I excised the top and bottom of each core with an ethanol-sterilized saw to remove sample that had contact with drilling fluids. I preserved the sediment cores in airtight mylar bags with oxygen-absorbing packets and placed them on ice in the field for approximately four hours; thereafter the cores were stored at -80°C. In the laboratory I opened each sediment core with a sterilized dremel tool, placed the sediment in a sterile bag, and then homogenized the sediment by hand from the 21 different cores (seven depths at three sites). During sieving, I did not observe plant roots or invertebrates such as worms.

2.2.3 *Sediment Chemical and Grain Size Analyses*

I measured pH with an Accumet® AB15 pH meter (Fisher Scientific, Inc., Waltham, Massachusetts, USA) in each of the homogenized sediment core samples after adding one gram of the wet sediment to one milliliter of de-ionized water (Carter, 1993). In order to quantify the percentage of C and nitrogen (N) in each sample I first dried 5 grams of homogenized sediment at 70°C for 48 hours; dried sediments were then ground to a fine powder (Cleveland et al., 2006). I used a Thermo Scientific FlashEA 1112 Elemental Analyzer (Thermo Fisher Scientific, Inc.) with high temperature (950°C) dry combustion to measure the percentage C and N in each sample (Matejovic, 1997). I estimated soil moisture in sediment samples by dividing the difference between the mass of the wet sediment and the mass of dried sediment by the mass of the wet sediment. The Laboratory for Environmental and Geological Studies (LEGS) at the University of Colorado (http://www.colorado.edu/geolsci/legs/indexa.html) conducted the analysis for determining the concentrations of Mn, As, and Fe in each oven-dried, homogenized sediment sample. Sediment concentrations of Mn, As and Fe were determined using a protocol modified from Farrell et al. (1980). Briefly, 5 ml of a 7:3 mixture of hydrochloric acid and

hydrofluoric acid and 2 ml of nitric acid were added to sediment samples in digestion tubes. Tubes were then heated to 95° C in a digestion block for 2 hours. Next, samples were cooled and the volume of each sample was increased to 50 ml with a 1.5% by weight boric acid solution. Samples were reheated to 95° C for about 15 minutes, and then cooled to room temperature again. Metals concentrations were analyzed in the cooled solutions on a SCIEX inductively coupled plasma mass spectrometer, (Elan DRC-e, Perkin Elmer, Waltham, MA) using an Indium internal standard.

A second set of aquifer sediment samples collected at K240, K150 and K60 from depths between 1 m and 16 m was used to investigate grain size. A modification of the USGS East Coast Sediment Analysis Procedures protocol was used for this analysis (Poppe et al., 2000). Sediments were freeze-dried for 48 to 72 hours and then oven-dried at 60°C for 48 hours. In order to disaggregate the sediment, samples were suspended in distilled water for 24 hours. Then each wet sub-sample was passed serially through 150 μm and 63 μm sieves. Sediment fractions were dried and weighed and grain size distributions were reported as percentage sand (i.e., greater than 150 μm in diameter), silty-sand (between 150 μm and 63 μm in diameter), and silt (63 μm in diameter and smaller) (Wentworth, 1922). In the statistical analyses I included data from only one grain size class, percentage silt, since percentage sand, percentage silty-sand and percentage silt were strongly autocorrelated.

2.2.4 *Phylogenetic Data Analysis*

DNA was extracted from homogenized sediment samples using the Mo Bio PowerSoil™ DNA Isolation Kit following the manufacturer's suggested protocol (Mo Bio Laboratories, Inc., Carlsbad, CA, USA). Bacterial 16S rRNA genes were PCR-amplified from the genomic DNA of the 21 sediment samples for pyrosequencing (Margulies et al., 2005) analysis. A universal

bacterial primer set described in Hamady et al. (2008) that included the highly conserved bacterial primers 27F (5'-GCCTTGCCAGCCCGCTCAGTCAGAGTTTGATCCTGGCTCAG-3') and 338R, with a unique, error-correcting barcode that identifies the PCR product in each sample (5'-GCCTCCCTCGCGCCATCAGNNNNNNNNCATGCTGCCTCCCGTAGGAGT-3') was used (Fierer et al., 2008). Each reaction contained 3 µL of genomic DNA, 30 µM (final concentrations) forward and reverse primers, and 22.5 µL of Platinum SuperMix (Invitrogen, Carlsbad, CA, USA). Reaction conditions were performed as described by Fierer et al. (2008). PCRs were performed for each of the 21 samples in triplicate, and then the products were pooled from each sample for downstream processing. No template controls were included to ensure that sample DNA was not contaminated with foreign DNA. PCR products were cleaned with the Mo Bio UltraClean-htp PCR Clean-up Kit (Mo Bio Laboratories, Inc.) according to the manufacturer's recommended protocol, and then pooled in equal concentrations. The University of South Carolina Environmental Genomics Core Facility performed the sequencing of our 16S rRNA gene amplicons on a Roche FLX 454 pyrosequencing machine.

First, I used QIIME (Caporaso et al., 2010a) to perform a sequence quality filter to the original 16S rRNA gene sequence dataset based on the sequence quality log file. This quality filter eliminated sequences that were shorter than 200 nucleotides in length, in addition to those with 1 or more ambiguous bases, and/or had received a quality score of less than 25. After this sequence quality filter, pyrosequencing yielded 31,517 quality short-read (average length of 231 nucleotides) 16S rRNA gene sequences total, and an average of 1500 sequences per sample (with a standard deviation of 189). I used QIIME to conduct all of the following phylogenetic analyses of the 16S rRNA sequences (Caporaso et al., 2010a). I defined bacterial operational taxonomic units (OTUs) at 97% identity with the uclust (Edgar, 2010) and the cd-hit algorithm (Li and

Godzik, 2006). As a source for comparison, I also defined bacterial OTUs at 90%, 95% and 99% identity with the cd-hit algorithm. I conducted all of the subsequently described analyses on each of these OTU tables in order to identify any discrepancies based on the OTU definition and patterns between bacterial community structure and chemistry (5 tables total). Next, I filtered our dataset to eliminate OTUs represented by only one 16S rRNA gene sequence (singletons), as well as OTUs present in only one sample (Zhou et al., 2011). The number of sequences present in each sample after applying filtering is included in the Appendix (Table A.1). Then, I aligned the 16S rRNA gene sequences using the PyNAST alignment algorithm (Caporaso et al., 2010b) with the Greengenes database (DeSantis et al., 2006). In QIIME, I used the RDP Classifier (Wang et al., 2007) to assign the taxonomic classification to each OTU using the Greengenes database (DeSantis et al., 2006). To create a phylogeny, I implemented the FastTree algorithm (Price et al., 2009). I performed rarefaction analysis, and calculated collector's curves (Schloss & Handelsman, 2004) for many different alpha diversity metrics including the Chao1 richness estimator (Chao, 1984) and Shannon diversity index (Weaver & Shannon, 1949). To investigate patterns in beta diversity, I calculated the pairwise distances between bacterial communities with the UniFrac distance metric (Lozupone & Knight, 2005). Sequences and sediment chemistry parameters were deposited in the MG-RAST database (Meyer et al., 2008) under accession number qiime:130 according to MIMARKS standards (Yilmaz et al., 2011).

2.2.5 *Statistical Analysis*

I used univariate and multivariate statistical techniques in order to elucidate relationships between the natural gradients within the aquifer, such as depth and well location, and the environmental chemistry and the bacterial community. I applied log transformations to percentage sediment C concentration, and square-root transformations to sediment As and Fe

concentrations because the raw data for these variables had non-normal distributions (Gotelli & Ellison, 2004). Then I performed linear correlation analyses using MATLAB® 7.9.0 (2009b) and the Pearson's correlation coefficient (Zar, 1999). I performed permutational multivariate ANOVA tests using the adonis function in the vegan package in R (Oksanen, 2007) in order to evaluate the role of depth and well location in structuring the bacterial community, as characterized by both unweighted UniFrac distances and proportions of bacterial taxa, within the aquifer. Then, to examine the effects of sediment chemical parameters on structuring bacterial communities I performed Mantel tests on the bacterial community structure (i.e., the UniFrac distance matrix) and sediment chemistry data.

Next, I performed non-metric multidimensional scaling (NMDS), an unconstrained ordination technique, with the metaMDS function in the vegan package in R (Oksanen, 2007) on the unweighted UniFrac distance matrix in order to further examine patterns in bacterial community structure and environmental parameters. Then I used the envfit function in the vegan package in R (Oksanen, 2007) to fit vectors of the environmental parameters that were significantly related to bacterial community structure in the Mantel tests (square-root transformed Fe, log transformed %C, Mn, and %silt) to the NMDS ordination (Oksanen, 2007).

## 2.3 Results

### 2.3.1 *Sediment Grain Size and Chemistry*

The percentage of the sediment classified as sand ranged from 2.5 to 96 while the percentage silt ranged between 1.4 and 76.8 (Figure 2.2). The range in percentage silty-sand was from 1.6% to 77.4%. The grain size distribution of the sediment samples varied based on the sample depth.

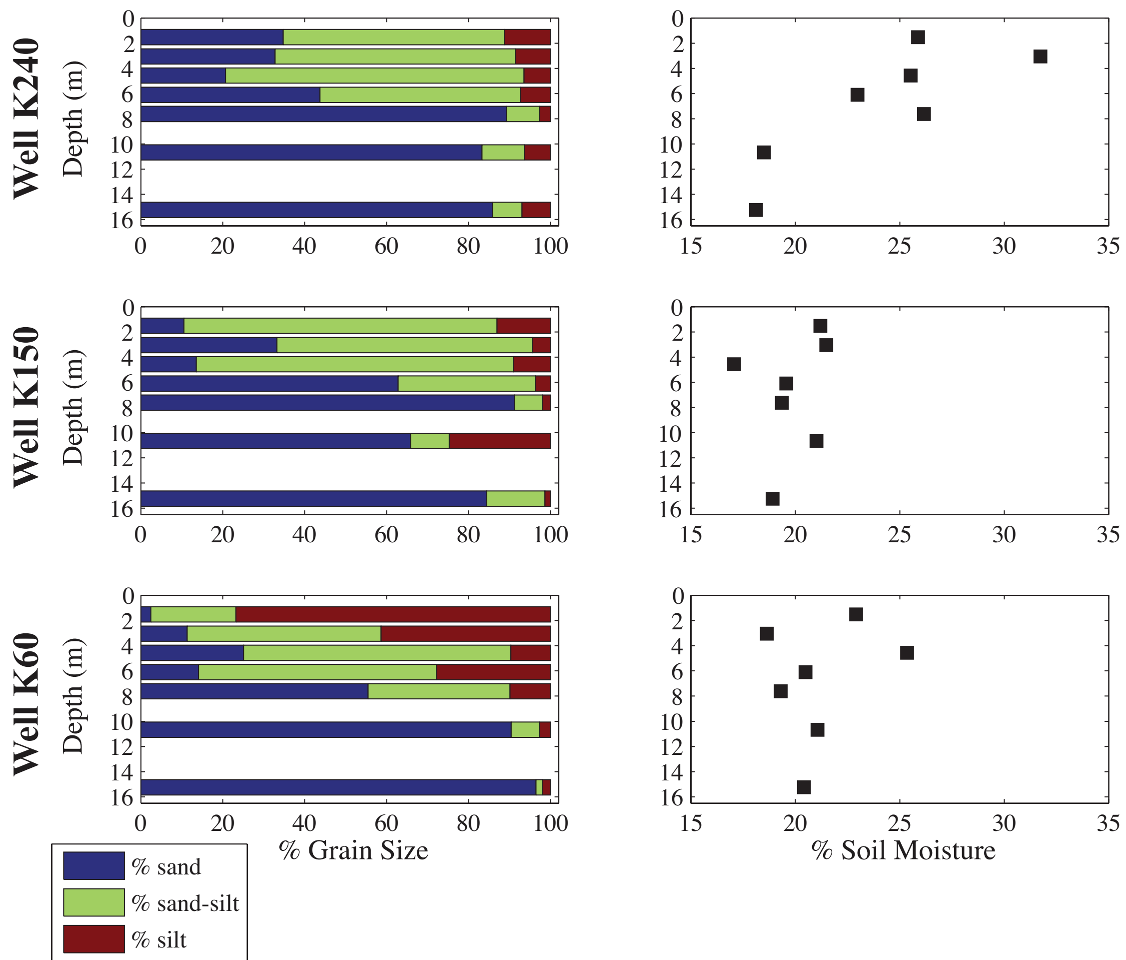



Figure 2.2. Grain size distribution and soil moisture of sediments collected at depths of 1.52, 3.05, 4.57, 6.10, 7.62, 10.67, and 15.24 m at the three sampling locations along the groundwater flowpath at Site K.

Silt-sized grains predominated in the sediments collected at shallow depths (<7 m) within the aquifer, whereas sediments collected at deeper depths (>7 m) were primarily composed of sands. Sediment C ranged from 0.04 to 0.67% across all samples (Figure 2.3), while sediment N was at or below the detection limit of the analytical method (0.01%) for 15 of the 21 samples. The samples were all circumneutral and pH ranged from 6.9 to 7.8. Sediment Mn concentrations ranged from ~140-1100 ppm, Fe from 10-50 ppt and As from 5-39 ppm (Figure 2.3). Sediment Mn concentration was the only chemical parameter that showed statistically significant differences between well sites, and was nearly twice as high at K60 than at K150 or K240 (ANOVA, $p<0.05$). Depth was significantly negatively correlated to percentage silt and sediment Mn and Fe concentrations, whereas percent silt, C, Mn, and Fe were positively correlated with one another. Sediment pH was not correlated with any other chemical parameter measured in this study (Table 1).

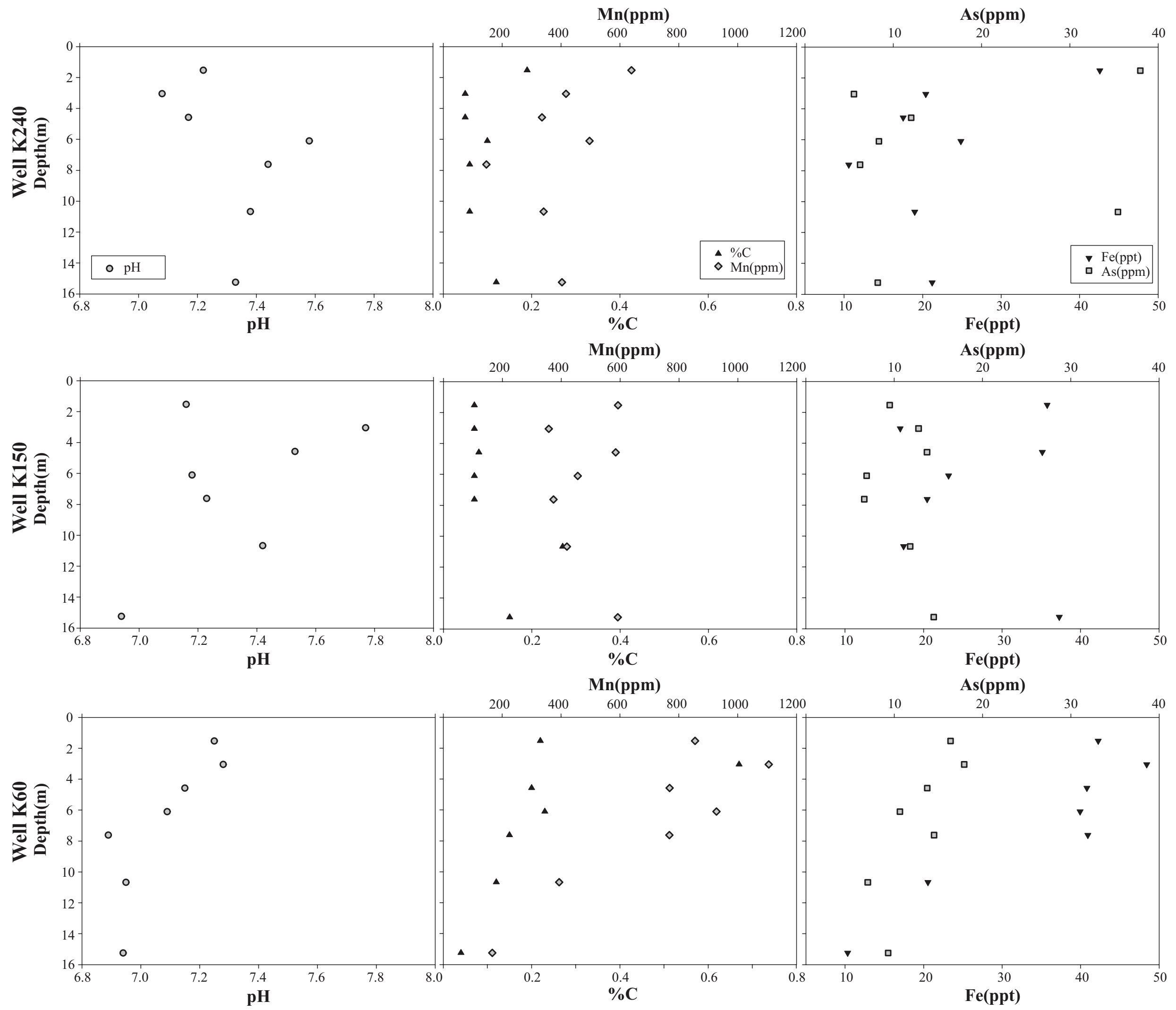



Figure 2.3. Sediment chemistry at each sampling location along the groundwater flowpath.

Table 2.1. The correlation coefficients (R values) for the Pearson's correlations that are presented in this table correspond to the pairwise correlations between environmental and bacterial community diversity data. Significance values (p-values) ≤0.05 for the correlations are shown in parentheses; 'NS' indicates non-significant correlations. I transformed data columns if the raw data did not follow a normal distribution. The transformation method I used is indicated in the row and column labels.

| ***Correlation Coefficients (R)*** | **% soil moisture** | **pH** | **Log (%C)** | **Mn (ppm)** | **√As (ppm)** | **√Fe (ppm)** | **% Silt** | **Unifrac** |
|---|---|---|---|---|---|---|---|---|
| **% soil moisture** | | -0.05 (NS) | -0.21 (NS) | -0.17 (NS) | -0.13 (NS) | -0.16 (NS) | -0.01 (NS) | 0.003 (NS) |
| **pH** | | | -0.05 (NS) | -0.18 (NS) | 0.09 (NS) | -0.22 (NS) | 0.04 (NS) | -0.14 (NS) |
| **Log (%C)** | | | | 0.82 (6E-6) | 0.32 (NS) | 0.71 (3E-4) | 0.62 (0.002) | 0.39 (0.005) |
| **Mn (ppm)** | | | | | 0.32 (NS) | 0.93 (6E-10) | 0.65 (0.001) | 0.56 (0.001) |
| **√As (ppm)** | | | | | | 0.42 (NS) | 0.22 (NS) | -0.08 (NS) |
| **√Fe (ppm)** | | | | | | | 0.51 (0.02) | 0.27 (0.003) |
| **% Silt** | | | | | | | | 0.44 (0.007) |
| **Unifrac** | | | | | | | | |

2.3.2 *Bacterial Community Characteristics*

Rarefaction analysis demonstrated that there is a large variation in the total number of OTUs between the samples. Collector's curves for the Chao1 richness estimator and Shannon diversity index show that the overall diversity does approach an asymptote in a majority of the samples (Appendix, Figure A.1), suggesting that the sequence coverage was sufficient to capture the diversity of the bacterial communities. Interestingly, it appears that the alpha diversity of samples is related to the well location: the number of OTUs per sample is highest at Well K240 and lowest at well K60.

Although 35 phyla were present in the 21 samples, only 4 phyla comprised more than 5% of the community in every sample (Figure 2.4). Proteobacteria comprised approximately 28% of the average community across all samples, whereas Chloroflexi and Acidobacteria each comprised approximately 11% of the community. The proportion of Firmicutes was approximately 5%, while other phyla represented much smaller proportions of the bacterial communities. Acidobacteria and Firmicutes were highly variable, with ranges in proportions of two orders of magnitude across all of the bacterial communities sampled, whereas the proportions of Chloroflexi and Proteobacteria varied by roughly one order of magnitude across the 21 samples. The sub-phyla Alphaproteobacteria, Betaproteobacteria, Deltaproteobacteria and Gammaproteobacteria composed an average of 7%, 7%, 8% and 5% of the overall bacterial community, respectively.

Bacterial community composition was not significantly related to depth (Table 1). Instead, bacterial communities clustered by well site; a permutational multivariate ANOVA using the adonis function in the vegan package (Oksanen, 2007) revealed that well site accounted for a significant amount of the variation in community composition ($R^2 = 0.23$; $p < 0.001$).

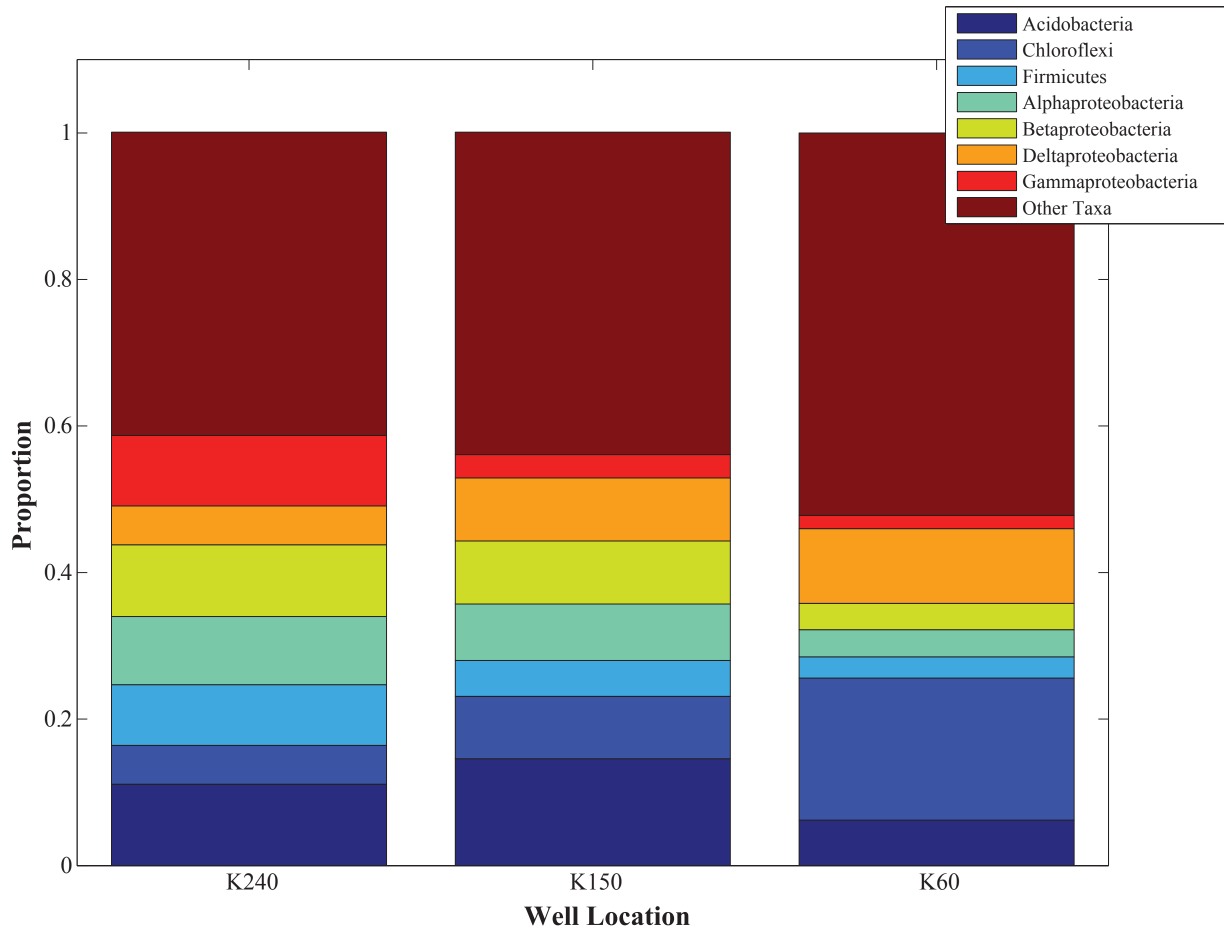



Figure 2.4. The bacterial communities, as characterized by the UniFrac metric, are significantly distinct at the three well locations at our study site (PERMANOVA, $R^2 = 0.23$; $p \leq 0.001$). The bar graphs show the relative abundances of dominant (>5% total abundance) bacterial taxa at the three locations ($n = 7$ depths for each well) and in all wells combined. Although there are gradients in sediment texture and chemistry within the aquifer, only Mn was significantly different between well (ANOVA, $p < 0.05$).

This relationship was significant ($p < 0.05$) for unweighted UniFrac distance matrices calculated from the pre-filtered and filtered datasets, and OTU tables calculated from different clustering methods (uclust and cd-hit) and different identity thresholds (90%, 95%, 97%, and 99%). Next, I investigated how the proportions of bacterial taxa contributed to these observed differences. The proportions of the most common (at least 5% of the total community) bacterial phyla and sub-phyla along the groundwater flowpath were significantly different between well locations ($p < 0.05$). Specifically, Deltaproteobacteria and Chloroflexi were found in higher relative abundance while Betaproteobacteria, Alphaproteobacteria, Gammaproteobacteria and Firmicutes were in lower proportions in K60 (near the river) than K150 (middle of the flowpath) or K240 (village site, Figure 2.4). The proportion of Acidobacteria was highest at K150, and lowest at K60.

While roughly 30% of the OTUs in the bacterial communities at well K240 and well K60 were unique to that location, 17% of the bacterial communities at K150 were comprised of unique OTUs (Figure 2.5). Bacteriodetes, Alphaproteobacteria, Acidobacteria and Betaproteobacteria were the most frequently observed OTUs unique to well K240. Acidobacteria dominated the unique OTUs at K150. The OTUs that were unique to K60, the site closest to the river, were primarily Chloroflexi (30%), and 50% of these Chloroflexi OTUs were classified in the Dehalococcoidetes class. All three of the well locations shared 294 OTUs; a majority of those OTUs (22%) were classified within the Acidobacteria. The bacterial communities at K150 and K60 had the highest number of shared OTUs, which largely belonged to the Chloroflexi, Acidobacteria, and Deltaproteobacteria (comprising 16%, 15% and 11% of the shared OTUs, respectively), the taxa that comprised the largest proportion of the 16S rRNA gene dataset at Site K. Only two taxonomic groups, Acidobacteria and Alphaproteobacteria each comprised greater than 10% of the OTUs shared between K240 and K150.

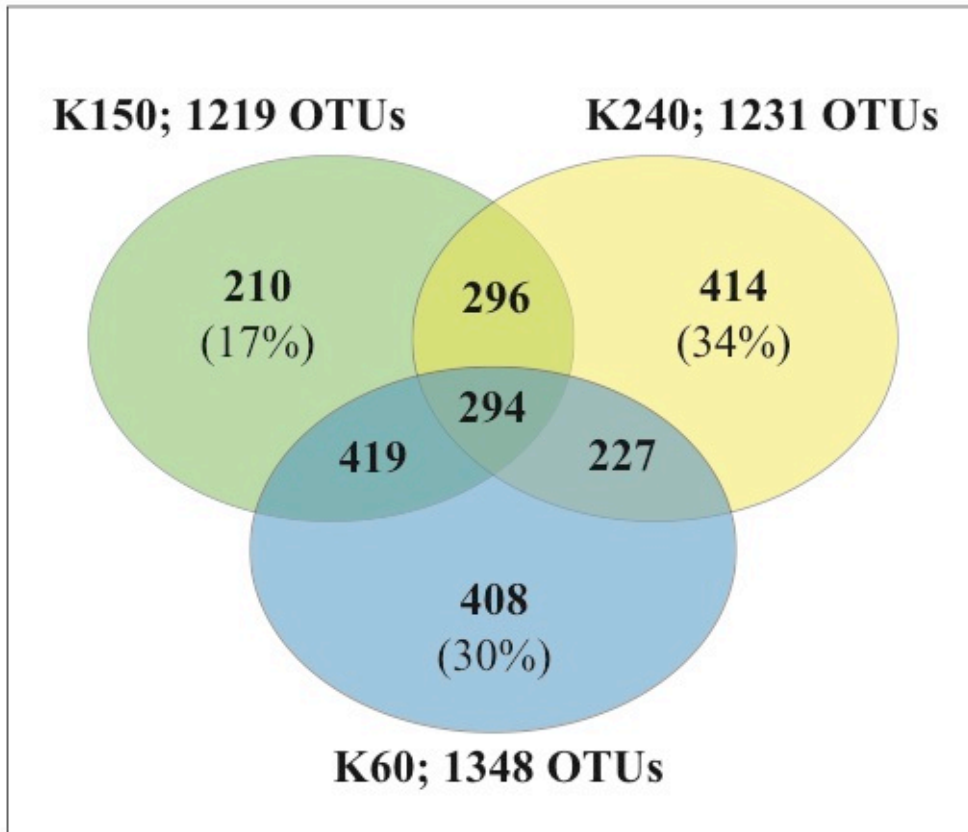



Figure 2.5. As shown in the Venn diagram, a greater percentage of the OTUs present within the bacterial communities at well locations K240 (village) and K60 (river) were unique than at well K150 (mid flow path). In addition, K150 and K60 shared a greater percentage of OTUs (approximately 30%).

By contrast, the OTUs common to K240 and K60 were more taxonomically distributed: OTUs in the Betaproteobacteria, Alphaproteobacteria, Acidobacteria, Firmicutes, and Bacteriodetes each comprised greater than 10% of the shared OTUs between K240 and K60.

Mantel tests revealed that bacterial community composition, as characterized by the unweighted UniFrac metric (Lozupone & Knight, 2005), was significantly related to percentage silt, and sediment C, Mn, and Fe concentrations, but not sediment As or pH (Table 1). I performed the Mantel tests using unweighted UniFrac distance matrices calculated from the pre-filtered and filtered datasets, as well as OTU tables calculated from different clustering methods (uclust and cd-hit), different identity thresholds (90%, 95%, 97%, and 99%), and different numbers of sequences per sample (Appendix, Table A.2). I found significant relationships between bacterial community structure and percentage silt, and sediment C, Mn, and Fe concentrations for each of these UniFrac distance matrices (Appendix, Table A.2) suggesting that the correlations observed are robust to issues related to sampling and OTU definition.

The results from the NMDS analysis, depicted in an ordination plot (Figure 2.6), demonstrate the relationship between bacterial community structure and environmental parameters. Each point on the ordination represents a bacterial community from a specific sample location and depth. The distances between samples (points) in the ordination indicate their level of similarity, as characterized by unweighted UniFrac distances. In order to investigate the validity of the NMDS ordination analysis I performed a stressplot, which showed that the UniFrac distances and the ordination distances were highly correlated ($R^2 = 0.98$).

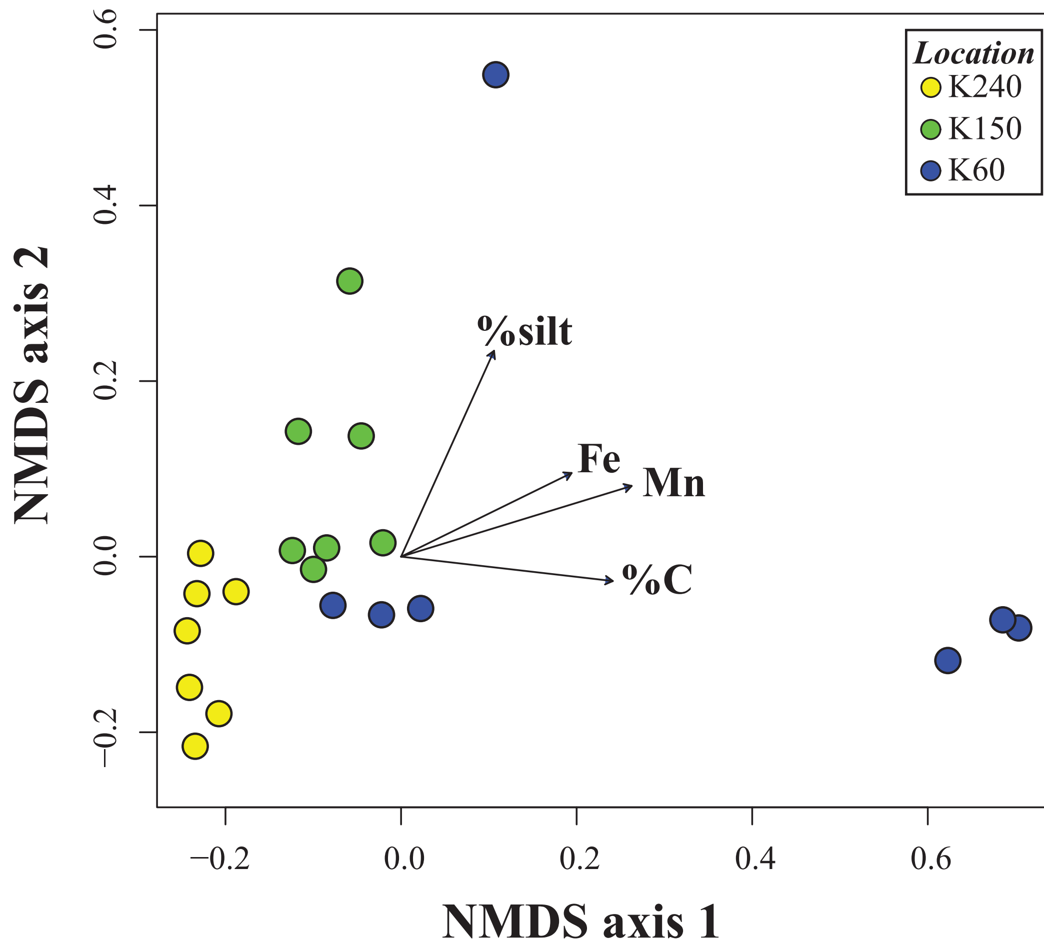



Figure 2.6. Thus non-metric multi-dimensional scaling (NMDS) ordination shows that the bacterial communities cluster based on well location. Each point on the ordination is a different bacterial community (n = 21), and the location of the points is based on UniFrac distances. The points are colored based on well location. In addition, the vectors on the ordination represent the gradient in the chemical variables for all the samples in the dataset.

The NMDS ordination (Figure 2.6) demonstrates that bacterial communities, in general, cluster based on well location, which supports results from a permutational multivariate ANOVA. Also, the ordination shows that bacterial communities at K60 are more different from one another than at the other well locations. Whereas bacterial communities from the deepest depths at K60 cluster with communities from K150, bacterial communities at depths of 3.05m, 4.57m and 6.1m at K60 form a distinct cluster between 0.6 and 0.7 on NMDS axis 1 (Figure 2.6). Similarly, the bacterial community from the shallowest depth at K60, 1.52m, is positioned at roughly 0.55 on NMDS axis 2, far from the other communities on the ordination (Figure 2.6). This suggests that the bacterial community structure at 1.52m at K60 is distinct from bacterial communities at other locations at the study site. The direction and position of the environmental parameter vectors on the ordination, calculated with the envfit function in the vegan package (Oksanen, 2007), represent the gradient in each parameter (i.e., from lowest to highest concentration) as well as the strength of the correlation between the environmental parameter and the ordination (i.e., bacterial community structure, as characterized by UniFrac distances). The vectors provide a graphical representation to identify relationships between environmental gradients and patterns in bacterial community structure.

**2.4 Discussion**

Bacterial community structure at Site K is significantly related to well location, grain size and chemical differences in groundwater sediments, including percentage silt, and sediment C, Mn, and Fe concentrations (Table 1). Taxa such as Alphaproteobacteria, Betaproteobacteria, and Acidobacteria were more abundant at the village site (K240), in aquifer sediments with higher concentrations of sand and lower concentrations of C and metal. In addition, the OTUs that were

present only at well location K240 belonged to these taxa. By contrast, OTUs classified as Deltaproteobacteria and Chloroflexi comprised a larger proportion of the communities in silty sediments with higher concentrations of C, Fe, and Mn. (Figure 2.4). These results suggest that the considerable heterogeneity in sediment chemistry at Site K drives significant differences in bacterial community structure across the three well locations.

The dominant members (greater than 5% of the community) of the bacterial communities at Site K, Acidobacteria, Chloroflexi, Firmicutes, Alphaproteobacteria, Betaproteobacteria, Deltaproteobacteria and Gammaproteobacteria (Figure 2.4), are abundant in many soil and sediment environments. Recent 16S rRNA gene sequence-based analyses have found that Acidobacteria and Proteobacteria are the dominant members of soil bacterial communities across ecosystem types (Fierer et al., 2009). A global survey of 21 16S rRNA gene sequence libraries found that while Proteobacteria and Acidobacteria comprised roughly 40% and 20% of the bacterial communities respectively, Chloroflexi and Firmicutes were also relatively abundant (greater than 5% of the community), in a range of soil environments (Janssen, 2006). Recent research shows that although these groups are dominant across soil types, the relative proportion of the bacterial community that each of these dominant groups comprises is influenced by factors such as pH, depth within the soil profile, the degree of soil saturation, and anaerobiosis (Fierer et al., 2009; Hansel et al., 2008; Jones et al., 2009). In our study, I found that shifts in bacterial community structure were related to changes in the sediment grain size distribution and changes in sediment C, Mn and Fe rather than soil moisture or depth (Table 1).

Our finding that bacterial community structure is significantly related to sediment grain size is supported by other research that has shown that sediment grain size influences microbial biomass and bacterial community structure (Sessitsch et al., 2001), and enzyme kinetics (Grandy

et al., 2008). Silts typically have higher concentrations of organic matter (Sparks, 2003), especially aromatic carbon compounds and humic acids (Guggenberger et al., 1995), and metals (Murray et al., 1999; Thorne & Nickless, 1981). Sandy sediments, characterized by grain sizes larger than 150 μm, usually contain lower organic carbon and metal concentrations (Sparks, 2003). In the sandy sediments with high Si concentrations at Site K, phyla such as Firmicutes and Alphaproteobacteria are the dominant groups in the bacterial community. By contrast, Chloroflexi and Deltaproteobacteria were the dominant members of the bacterial community in silty sediments with high C and metal concentrations. Heavy metal concentrations have been shown to correlate with bacterial community structure and function in both soil and groundwater environments (Stefanowicz et al., 2008).

Deltaproteobacteria are abundant across different soil and sediment environments (Spain et al., 2009). There is evidence that Deltaproteobacteria are more abundant in anaerobic soils (Hansel et al., 2008), perhaps because members of Deltaproteobacteria can use a variety of electron acceptors. For example, organisms within the *Geobacteraceae* family can use labile C to reduce Fe, Mn and HS (Lovley & Phillips, 1988; Lovley et al., 1996). There is ample evidence that Fe-, Mn-, and HS- reducing Deltaproteobacteria are prevalent in anaerobic freshwater sediments (Coates et al., 2002; Lovley et al., 2004), and previous research has documented that the relative abundance and diversity of members of *Geobacteraceae* corresponds to Fe and Mn concentrations in groundwater environments (Luna et al., 2009).

Although the overall community composition was not found to be correlated with sediment As (Table 1), it is important to note that there is a poor relationship between sediment and groundwater As concentrations in this environment (Radloff et al., 2008). However bacterial community structure is significantly correlated to sediment characteristics, such as percent silt,

sediment C and sediment metal concentrations in the sediment, which are related to groundwater As concentrations (van Geen et al., 2006a). Thus, the relationship between the proportion of Deltaproteobacteria in the bacterial community and percent silt and concentrations of C and Fe in the aquifer sediment may have important implications for understanding how the native microbial community influences groundwater As mobility at Site K. Fe-reducing Deltaproteobacteria could promote the mobilization of As by mediating the reductive dissolution of Fe- oxides, which results in the desorption of As from the Fe-oxide surface (Jiang et al., 2009; McArthur et al., 2001). In fact, results from GBD sediment microcosm experiments demonstrated that labile carbon additions promoted higher rates of Fe(III)-reduction and As mobilization, as well as increases in the relative abundance of Deltaproteobacteria (Islam et al., 2004). Additional evidence from sediment microcosm experiments suggests that microbial Fe-reduction is enhanced by the presence of redox-active HS in aquifer sediments, which can shuttle electrons to Fe(III), promoting Fe-oxide reduction and As desorption (Mladenov et al., 2010). Thus, Fe-reducing and HS-reducing Deltaproteobacteria could play a significant role in groundwater As mobilization.

Deltaproteobacteria could also influence groundwater As mobilization more directly. First, as mentioned above, although the overall community composition was not related to sediment As, it was correlated to the percent silt in the sediment. Fine-grained silts are less permeable than sands and often have higher dissolved As concentrations. Thus Deltaproteobacteria in the silty sediments at Site K may be more adapted to high groundwater As concentrations. Members of the genera *Desulfovibrio* (Li & Krumholz, 2007), *Desulfomicrobium* (Macy et al., 2000), and *Geobacter* (Methe et al., 2003) within the Deltaproteobacteria are known to reduce As(V) to As(III) through a detoxification pathway encoded by the *ars* operon (Rosen et al., 1991).

Groundwater As concentrations at Site K are highest at well K60 between 7 and 15 m, and as a result these conditions may favor bacteria that can detoxify As. Thus, the potential for Deltaproteobacteria to mediate the As detoxification, Fe-reduction, Mn-reduction and HS-reduction pathways could explain their higher abundance at K60 (Figure 2.4).

Chloroflexi at Site K are also significantly more abundant in silty sediments with higher concentrations of sediment C, Fe and Mn. Approximately 50% of the Chloroflexi OTUs at Site K belonged to the halo-respiring Dehalococcoidetes class, and these OTUs were only present in the C- and metal- rich silty sediments at well K60. Halorespiring Chloroflexi have also been discovered in other pristine freshwater environments (Loffler et al., 2000), and recent research suggests that halorespiring bacteria could also use other respiratory pathways including Se(VI)-reduction, As(V)-reduction, Fe(III)-reduction, Mn(IV)-reduction, as well as the oxidation and reduction of a model compound for quinone-containing HS (Luijten et al., 2004). Also, two members of the Dehalococcoides genus have the As resistance gene, *arsC*, within their genome (Kube et al., 2005; Seshadri et al., 2005). Consequently, Chloroflexi may have an advantage at well K60 over other bacteria that do not have such adaptations to the local groundwater geochemical conditions.

Based on the phylogenetic affiliations of these taxa, it is possible that these results indicate a shift to more Fe-, Mn-, and humic substance- reducers in the silts with high C and metal concentrations (Lovley & Phillips, 1988; Lovley et al., 1996; Luijten et al., 2004). However, research is needed to further elucidate the specific roles of taxa such as Deltaproteobacteria and Chloroflexi in the groundwater environment at Site K. For example, our study suggests that sediment C is important in structuring bacterial communities and thus, further work on the source and chemical characteristics of the sediment C may provide insight into the dominant

processes underlying the relationship between C and bacterial community composition. For example, if the C is serving largely as a nutrient source, I may conclude that observed community shifts are an example of the copiotrophic-oligotrophic continuum described by Fierer et al. (2007). By contrast, more recalcitrant, redox-active C sources such as HS may be important for electron shuttling that promotes Fe- reduction and As mobilization (Mladenov et al., 2010).

## 2.5 Conclusions

Results from this research have led to a more complete understanding of the bacterial community structure within GBD aquifer sediments. It is well-documented that sediment grain size, C, Mn and Fe influence the mobility of groundwater arsenic, and it is intriguing that these constituents also structure the bacterial community. This work has also demonstrated the importance of deeper 16S rRNA gene sequencing in identifying environmentally relevant patterns in bacterial community structure across groundwater As gradients.

# CHAPTER 3: Labile and redox-active dissolved organic matter influence bacterial community structure and groundwater arsenic mobilization

## 3.1 Introduction

Elucidating the mechanisms promoting spatial variability in groundwater arsenic (As) across relatively small scales is key to developing mitigation strategies for the geogenic As contamination crisis in Southeast Asia (Fendorf et al., 2010). An extensive body of research has shown that sediment geochemistry, labile dissolved organic matter (DOM) chemistry and availability, and the activity of native bacteria in aquifer sediments are the most important factors in promoting groundwater As mobilization. While it is clear that ferric iron (Fe(III))-reducing bacteria in aquifer sediments cause the release of As sorbed to iron (Fe)-oxide minerals (Héry et al., 2010; Islam et al., 2004; Radloff et al., 2008), as originally proposed by Nickson, et al. (1998), there is conflicting evidence regarding whether sediment geochemistry (Horneman et al., 2004; van Geen et al., 2004; Zheng et al., 2005) or DOM chemistry (Harvey et al., 2002; McArthur et al., 2004; Mladenov et al., 2010; Neumann et al., 2010) are more important for biological groundwater As mobilization rates.

Throughout Southeast Asia differences in the grain size of sediments and in the chemistry of the Fe-oxide minerals and As in aquifer sediments are partially responsible for the high degree of variability in groundwater As concentrations observed across relatively small spatial scales (Winkel et al., 2008; Horneman et al., 2004; Polizzotto et al., 2008). For example, dynamism in the redox state of Fe on mineral surfaces results in the uneven sorption affinity and distribution of As on Fe-oxide minerals in aquifer sediments (van Geen et al., 2004; Zheng et al., 2005). As well, fine-grained, silty aquifer sediments generally have higher concentrations of organic matter, Fe-oxide minerals, and sorbed As than sands (McArthur et al., 2004; van Geen et al.,

2006a). Sedimentary organic carbon concentrations can vary by an order of magnitude or more across gradients as small as 3 meters (Legg, et al. 2012; Rowland, et al. 2007), and organic matter in aquifer sediments may represent an important source of carbon to fuel the activity of heterotrophic Fe(III)-reducing bacteria in aquifers (McArthur et al., 2004; Mladenov et al., 2010). Evidence suggests that the sediment grain size, and carbon (C) and Fe concentrations significantly affect the composition of bacterial communities, and taxa related to heterotrophic Fe(III)-reducing bacteria are more abundant in silty sediments (Legg et al., 2012). Thus, sediment geochemistry may affect groundwater As mobilization by influencing the composition and activity of bacterial communities in aquifer sediments.

The availability and chemistry of DOM also influence groundwater As mobilization in high As aquifers. Local aquifer recharge from adjacent rivers and ponds, and seasonal flooding during the annual monsoon rains, represent a major source of DOM to aquifers (Ahmed et al., 2004). As well, extensive groundwater extraction for irrigation in cultivated areas throughout rural Bangladesh increases rates of aquifer recharge from eutrophic ponds, thereby transporting more labile DOM into the subsurface (Harvey et al., 2002; Neumann et al., 2010). Understanding the source and supply of labile DOM is critical to predicting rates of groundwater As mobilization because Fe-reducing bacteria use labile constituents as electron donors during reductive dissolution of Fe-oxides, which results in the release of As (Héry et al., 2010; Islam et al., 2004; Rowland et al., 2007). Also, heterotrophic bacteria in aquifers rely on labile DOM for sources of C and nitrogen (N) to synthesize biomass (Lear et al., 2007).

Bacteria in high As aquifers may also utilize redox-active fractions of DOM to mediate groundwater As cycling. Humic substances are a redox-active component of DOM that is relatively resistant to degradation in most freshwater environments (Lovley et al., 1996). A

phylogenetically diverse suite of bacteria can use humic substances as an electron acceptors in anaerobic aquatic environments (Cervantes et al., 2002; Coates et al., 1998). In fact, in some environments these taxa may be more abundant than Fe-reducing bacteria (Kappler et al., 2004). Quinones are the moieties within humic substances that serve as electron acceptors for bacterial humic substance-reduction (Scott et al., 1998), and reduced quinones abiotically transfer electrons to Fe-oxide minerals, thereby reducing Fe(III) up to seven times faster than bacteria (Jiang & Kappler, 2008). Also, quinone moieties represent perennial catalysts of Fe-oxide reduction because they are resistant to degradation and are continually recycled through biotic and abiotic oxidation and reduction reactions.

The recent discovery of reduced quinone fluorescence within high As groundwater (Mladenov et al., 2010) suggests that bacterial quinone-reduction could affect rates of As mobilization in the aquifers of Southeast Asia. It is likely that quinone-containing humic substances are introduced into groundwater DOM during the reductive dissolution of Fe-oxide minerals, which causes the release of organic matter from aquifer sediments (Mladenov et al., 2010). Sedimentary organic matter in the GMBD is primarily derived from the degradation of plant material, and is comprised of peat deposited during sediment burial (McArthur et al., 2004; Quicksall et al., 2008). Also, it is likely that humic substances are retained in the sediment due to sorption on Fe-oxides (Mladenov et al., 2008). Thus, there are endogenous sources of quinone moieties in the high As aquifers of Southeast Asia.

The relative importance of labile and redox-active DOM fractions in the biogeochemical cycling of As is unresolved. Incubation experiments with Bangladesh aquifer sediments have clearly demonstrated that labile DOM additions stimulate up to a three-fold increase in Fe(III)-reduction and As mobilization (Islam et al., 2004; Radloff et al., 2008). Also, DOM additions

stimulate shifts in the composition of the native bacterial communities, and an increase in the relative abundance of Deltaproteobacteria (Héry et al., 2010; Islam et al., 2004). Labile DOM also promoted bacterial community shifts and increases in groundwater As concentrations in Cambodian aquifer sediments (Lear et al., 2007; Pederick et al., 2007). However the addition of a model quinone compound resulted in even higher rates of Fe(III)-reduction and As mobilization, and a bacterial community shift towards Deltaproteobacteria (Rowland et al., 2007). Results from these experiments suggest that Deltaproteobacteria may use respiratory quinone reduction, in addition to Fe(III)-reduction, to mediate groundwater As mobilization in aquifers throughout Southeast Asia.

Thus, the goal of this research was to examine the relative effects of DOM chemistry vs. sediment geochemistry on bacterial community structure and Fe-and As mobilization rates in an aquifer affected by geogenic arsenic contamination in Araihazar, Bangladesh. To this end, I designed a full factorial experiment in which aquifer sediment and groundwater from four sites with contrasting geochemical conditions was incubated with and without labile and redox-active DOM. In addition, I employed fluorescence spectroscopy and a five-component PARAFAC model to characterize labile and redox-active DOM constituents within pre- and post-incubation groundwater. I found that the composition of the pre-incubation bacterial communities from different sites was distinct; however the composition of post-incubation bacterial communities converged according to the type of DOM treatment. Also, the addition of both labile and redox-active DOM favored bacteria related to quinone-reducing Deltaproteobacteria and promoted the highest rates of Fe and As mobilization. Finally, I discovered that the relative abundances of Chloroflexi and Deltaproteobacteria were significantly related to the availability of specific redox-active fluorophores in post-incubation groundwater.

## 3.2 Materials and Methods

### 3.2.1 *Site Description*

The study sites, Site K (Legg et al., 2012; Radloff et al., 2008) and Site B (Mladenov et al., 2010; Zheng et al., 2005), are located in Araihazar, Bangladesh, approximately 30 km northeast of Dhaka (Figure 3.1). This research focused on four locations/depths that represent a range of groundwater (Table 3.1) and aquifer sediment (Table 3.2) geochemical conditions, as well as a range of DOM chemical characteristics (Table 3.3). At Site K, I focused on a shallow (8m) depth well nest K240, which is located at the beginning of the groundwater flowpath, and a deep (14m) depth at the end of the groundwater flowpath (well K60). At Site B, I concentrated on groundwater at an intermediate (11m) and a deeper (14m) depth at a single well nest location. Of these four sampling locations, the highest groundwater Fe and As concentrations are observed at B 14m, whereas K240 8m has the lowest groundwater Fe and As concentrations (Table 3.1). Also, at the Site B sampling locations, sedimentary Fe and As concentrations, as well as the concentrations of carbon (C) and nitrogen (N) in aquifer sediment, are higher than at the two sampling locations at Site K (Table 3.2).

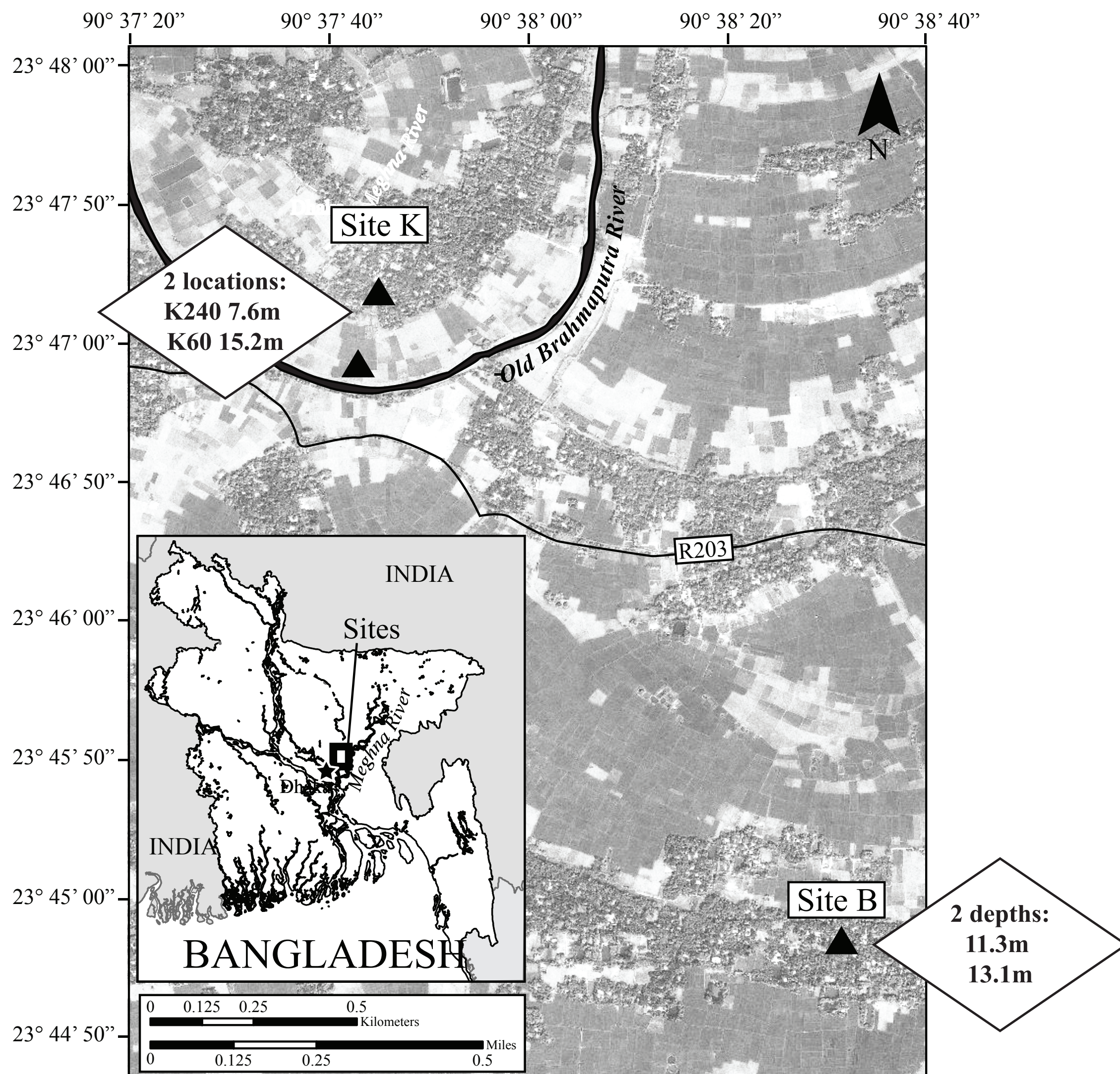



Figure 3.1. Aquifer sediment and groundwater was collected from Site K and Site B in order to characterize pre-incubation conditions in the aquifer, and to obtain material for preparing incubation experiments.

Table 3.1. The chemistry of pre- incubation groundwater collected at the four sampling location, including dissolved organic carbon (DOC), total dissolved nitrogen (TDN), the DOC:TDN ratio, total dissolved iron (Fe), and total dissolved arsenic (As).

| **SITE** | **DOC (mg/L)** | **TDN (mg/L)** | **DOC:TDN** | **Fe (ppm)** | **As (ppb)** |
|---|---|---|---|---|---|
| **B 11m** | 2.6 | 1.4 | 1.9 | 1.6 | 132 |
| **B 14m** | 3.2 | 2.3 | 1.4 | 4.3 | 256 |
| **K240 8m** | 0.8 | 0.2 | 4.8 | 1.1 | 1 |
| **K60 15m** | 1.4 | 0.8 | 1.8 | 1.6 | 115 |

Table 3.2. The total concentrations of Fe and As, and the percent carbon (C) and nitrogen (N), in pre- incubation aquifer sediment at each sampling location. The sediments from K240 8m and K60 14m had sediment nitrogen concentrations that were below the detection limit (b.d.l.) of the instrument.

| **SITE** | **Fe (ppm)** | **As (ppb)** | **%C** | **%N** |
|---|---|---|---|---|
| **B 11m** | 34376 | 12583 | 0.15 | 0.02 |
| **B 14m** | 37873 | 8749 | 0.14 | 0.01 |
| **K240 8m** | 11933 | 5483 | 0.10 | b.d.l. |
| **K60 15m** | 19670 | 7428 | 0.07 | b.d.l. |

Table 3.3. The dissolved organic matter (DOM) chemistry of pre- incubation groundwater collected at each sampling location. The fluorescence index (FI) indicates whether the source of DOM is terrestrially- or microbially- derived. The specific ultra-violet absorbance ($SUVA_{254}$) indicates aromaticity of DOM. The fluorophore types for each of the PARAFAC model components (C1 – C5) are listed in Table 3.4.

| SITE | FI | $SUVA_{254}$ | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|---|---|
| **B 11m** | 1.60 | 3.84 | 0.176 | 0.894 | 0.025 | 0.134 | 0.036 |
| **B 14m** | 1.53 | 3.67 | 0.214 | 1.114 | 0.033 | 0.138 | 0.052 |
| **K240 8m** | 1.56 | 5.36 | 0.003 | 0.109 | 0.002 | 0.079 | 0.009 |
| **K60 15m** | 1.43 | 14.26 | 0.064 | 0.368 | 0.000 | 0.126 | 0.008 |

Table 3.4. The 5 components of the PARAFAC model used in this study (Ebert 2012). For each component, the fluorophore type has been inferred based on comparisons between their EEMs and the EEMs of components within the Cory and McKnight (2005) PARAFAC model, as well as other similar fluorescent DOM PARAFAC models.

| Component | Fluorophore Type | Abbreviation |
|---|---|---|
| **1** | Microbially- associated reduced quinone | RQ |
| **2** | Sedimentary humic substance | HS |
| **3** | Microbially- associated semiquinone | SQ |
| **4** | Unidentified, Bangladesh fluorophore | UB |
| **5** | Amino acid- like | AA |

3.2.2 *Sample Collection*

In October 2009, groundwater and aquifer sediment samples were collected from the four sampling locations in order to acquire material for preparing incubation experiments, and to characterize the pre-incubation, ambient conditions in the aquifer. Groundwater was pumped from K240 at 8m, K60 at 14m, and B at depths of 11m and 14m, and samples were collected in burned, sterilized amber serum vials while maintaining anaerobic conditions, and were sealed with butyl stoppers and an aluminum crimp seal. These unfiltered groundwater samples were stored in the laboratory at 4$^{\circ}$C for up to 24 hours before they were processed during incubation experiment preparation. After collecting groundwater samples, aquifer sediments were drilled in approximately 0.5m sections from each of the four sampling locations/depths. While collecting aquifer sediment samples we exercised a number of precautions described previously (Legg et al., 2012) in order to minimize contamination and maintain anaerobic conditions. Sediment cores were stored in mylar bags with oxygen-absorbing packets on ice in the field, and then in the laboratory at 4$^{\circ}$C for up to 24 hours until they were processed for incubation experiment preparation.

3.2.3 *Incubation Experiment Preparation*

Incubation experiments were prepared using groundwater and aquifer sediment samples collected from our four sampling sites. First, a subset of groundwater samples from at each location was processed in order to capture pre-incubation conditions in the aquifer. To this end, 20 ml of groundwater from each site were filtered with glass fiber filters (GF/F, 0.7 μm) under an $N_2$ atmosphere in a Captair® Pyramid glove bag, and acidified to pH 2 with 1 M HCl to prevent the precipitation of DOM-Fe complexes. An additional 20 ml of groundwater from each site were sterile-filtered with cellulose acetate membrane filters (0.2 μm; ADVANTEC MFS®,

Dublin, CA) that were pre-rinsed with deionized water. These pre-incubation groundwater samples were stored anaerobically at 4°C in burned amber serum vials sealed with butyl stoppers and an aluminum crimp seal. Prior to preparing the incubation vials, each aquifer sediment core was homogenized by hand in a sterilized bag under an $N_2$ atmosphere in a Captair® Pyramid glove bag. A pre-incubation aliquot of the homogenized sediment from each core was placed in sterile bag, and then stored in a mylar bag with oxygen-absorbing packets on dry ice for up to one week. Thereafter the pre-incubation sediment samples were stored at −80°C in the laboratory.

Under an $N_2$ atmosphere in the Captair® Pyramid glove bag, approximately 3.5 g of the homogenized sediment and 18 ml of unfiltered groundwater from each sampling site were added to 12 different 20 ml amber serum incubation vials (previously burned and sterilized; 48 vials total). Three incubation vials from each site, hereafter called 'control incubations' (CL), were set aside and sealed with butyl stoppers and an aluminum crimp seal. Then, three of the remaining incubation vials from each site were subjected to a labile DOM treatment (LOM). To this end, sodium acetate was dissolved in groundwater from each site, and 18 ml of a 2 mM sodium acetate solution (final concentration) was added to triplicate vials. Next, a redox-active DOM treatment (ROM) was prepared. First, anthraquinone-2, 6-disulfonate (AQDS) was dissolved in groundwater from each site, and 18 ml of the AQDS solution (5 mM final concentration) was added to each of the triplicate incubation vials from each site. Finally, a labile and redox-active DOM treatment (LOM+ROM) was prepared by dissolving both sodium acetate and AQDS in groundwater from each site, and then adding 18 ml of this solution to the triplicate incubation vials from each site. In LOM+ROM incubations, the final concentration of sodium acetate was 2 mM and the final concentration of AQDS was 5 mM.

Incubation vials were mixed thoroughly, and sealed with butyl stoppers and aluminum crimp seals under the $N_2$ atmosphere in the glove bag. After 90 days, the post-incubation aquifer sediment was separated from the groundwater, and processed for downstream analyses under anaerobic conditions. An aliquot of post-incubation groundwater from each incubation vial was filtered through glass fiber filters (GF/F, 0.7 μm) and acidified to a pH of 2 with 1 M HCl, while another aliquot was sterile-filtered with pre-rinsed ADVANTEC MFS cellulose acetate membrane filters (0.2 μm) and remained un-acidified. These post-incubation groundwater samples were stored anaerobically at 4°C in burned amber serum vials sealed with butyl stoppers and aluminum crimp seals. An aliquot of post-incubation aquifer sediment was stored in 2 ml sterile Eppendorf® tubes (Hauppauge, NY) at −80°C in the laboratory, and was later used for bacterial community analysis. An additional aliquot of post-incubation aquifer sediment was stored in anaerobically sealed vials at 4°C for geochemical analyses.

3.2.4 *Geochemical Analyses*

Aliquots of pre- and post-incubation aquifer sediment samples were dried at 70℃ for 48 h, and then ground into a fine powder with a mortar and pestle. A Thermo Scientific FlashEA 1112 Elemental Analyzer (Thermo Fisher Scientific, Inc.,) with high temperature (950℃) dry combustion was used to measure the percentage C and N in each sediment sample (Matejovic, 1997). The Laboratory for Environmental and Geological Studies (LEGS) at the University of Colorado (http://www.colorado.edu/geolsci/legs/indexa.html) conducted chemical digestion of sediments with a protocol modified from Farrell et al. (1980), as described in Legg et al. (2012). Next, concentrations of Fe, As, Mn and Si in digestion solutions on a SCIEX inductively coupled plasma mass spectrometer (ICP-MS; Elan DRC-e, Perkin Elmer, Waltham, MA, USA) were determined using an Indium internal standard. Concentrations of Fe and As in unacidified,

filtered (0.2 µm) pre- and post-incubation groundwater samples were determined by ICP-MS at the LEGS lab (SCIEX; Elan DRC-e, Perkin Elmer, Waltham, MA, USA).

3.2.5 *Dissolved Organic Matter Characterization*

DOM chemistry was characterized in aliquots of pre-incubation and CL and LOM post-incubation groundwater samples filtered through 0.7 µm glass fiber filters (GF/F) and acidified to a pH of 2 with 1 M HCl. Since the ROM and LOM+ROM post-incubation samples had very high levels of dissolved ferrous Fe, it was not possible to accurately measure many of the DOM variables in these samples. To maintain the redox potential of groundwater samples during the DOM analyses, samples were processed under an $N_2$ atmosphere. The Kiowa Environmental Chemistry Laboratory at the University of Colorado (http://snobear.colorado.edu/Seiboldc/kiowa.html) measured concentrations of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) using a Shimadzu TOC-V CSN Total Organic Carbon Analyzer. Then, as described in Mladenov et al. (2010), the ultraviolet-visible (UV-vis) absorbance was characterized on an Agilent 8453 spectrophotometer in a quartz cell with a 1cm path-length. The UV absorbance at 254 nm was normalized to the DOC concentration in order to calculate the specific UV absorbance ($SUVA_{254}$), which is an indication of the aromaticity of the dissolved organic matter (Weishaar et al., 2003).

The filtered (GF/F, 0.7 µm), acidified pre- and CL and LOM post-incubation groundwater samples were transferred to a sealed quartz cuvette under an $N_2$ atmosphere, and then the fluorescence spectra of DOM in the samples was measured using a Jobin-Yvon-Horiba/Spex Fluoromax-3 spectrofluorometer. Three-dimensional fluorescence excitation-emission matrices (EEMs) for each sample were generated in S/R mode at excitations of 240-500 nm at 10 nm increments, and emissions of between 350-550 nm at 2 nm increments were generated. Initially,

the 13-component Cory & McKnight (2005) PARAFAC model was applied to the EEMs, however there were distinct peaks in the residuals for post-incubation samples, indicating that prominent fluorophores within the DOM were not fitted with this model. As a result, a five-component PARAFAC model was developed which provided a better fit to the CL and LOM post-incubation samples (Table 3.4; Ebert, 2012).

3.2.6 *Bacterial community analysis*

DNA was extracted from pre-and post-incubation aquifer sediment samples using the MO BIO PowerSoil® DNA Isolation Kit following the manufacturer's suggested protocol (MO BIO Laboratories, Inc., Carlsbad, CA, USA). The V1–V2 region of the 16S rRNA gene was PCR-amplified from genomic DNA using the highly conserved, universal bacterial primer set as described in Fierer et al. (2008), Hamady et al. (2008), and Nemergut et al. (2010), with a modification to the PCR amplification step (Stanish et al., 2012). The forward primer included the 454 Roche Titanium primer B, a 2 base pair adapter, and the highly conserved bacterial primer 27F. The reverse primer was comprised of the 454 Roche Titanium primer A, a unique, 12 base pair error-correcting Golay barcode, a 2 base pair adapter, and the highly conserved bacterial primer 338R. Each PCR was performed in triplicate, with a total reaction volume of 25 µl. The PCR mixture consisted of 10 µl of 5-PRIME hot master mix (5-PRIME, Gaithersburg, MD, USA), 10 µl of sterile water, 1 µl of the reverse primer (0.4 µM final concentration), 2 µl of the forward primer (0.4 µM final concentration), and 2 µl of genomic DNA. Samples were initially denatured for 3 min at 94°C followed by 25 cycles at 94°C for 45 s, 50°C for 30 s, 72°C for 90 s and a final elongation step at 70°C for 10 min (Stanish et al., 2012). The PCR product was quantified from each reaction, and the product from the triplicate reactions was pooled. Then the pooled PCR products were purified using the UltraClean PCR Clean-up kit (MoBio,

Carlsbad, CA) according to the manufacturer's protocol. 16S rRNA gene amplicons were sequenced using Titanium chemistry (454 Life Sciences, Bradford, CT, USA) on a GS FLX platform at the Environmental Genomics Core Facility (Engencore, University of South Carolina, USA).

Bacterial 16S rRNA gene sequence data were processed in the QIIME pipeline (Caporaso et al., 2010a). First, the sequence data were denoised (Reeder & Knight, 2010), and low-quality sequences were removed, including sequences that were shorter than 200 nucleotides in length, and/or with one or more ambiguous bases, and/or with quality scores of less than 25. After applying these filters, there were an average of 1436 16S rRNA gene sequences per sample (standard deviation of 503). Bacterial operational taxonomic units (OTUs) were defined at 97% sequence identity with the uclust algorithm (Edgar, 2010), and singleton OTUs (represented by one sequence), as well as OTUs present in only one sample were removed from the dataset (Zhou et al., 2011). The remaining 16S rRNA gene sequences were aligned using the PyNAST alignment algorithm (Caporaso et al., 2010b) with the Greengenes database (DeSantis et al., 2006). The taxonomic classification of each OTU was determined using the RDP Classifier (Wang et al., 2007), and then a phylogeny of the sequences was created using the FastTree algorithm (Price et al., 2009). The OTU taxonomic classification data and the phylogeny were used to calculate the phylogenetic diversity within each sample (PD; Faith, 1992), and UniFrac distances between samples (beta diversity; Lozupone & Knight, 2005). The unweighted UniFrac distances, which do not include information about the abundances of OTUs, were used in all of the downstream statistical analyses. The 16S rRNA gene sequences and accompanying metadata will be deposited in the MG-RAST database (Meyer et al., 2008) upon publication.

3.2.7 *Statistical analyses*

Statistical approaches were used to evaluate whether the labile and redox-active DOM treatments induced significant differences in aquifer chemistry and bacterial community structure during the incubation experiment. Outlier values, which fell outside of two standard deviations of the mean (Gotelli & Ellison, 2004), were identified and removed prior to performing statistical analyses. The C, N, Fe and As concentrations in post-incubation aquifer sediment samples, and Fe and As concentrations in post-incubation groundwater samples were normalized to the sediment weight within the incubation vial prior to performing statistical analyses. Also, a log-transformation was applied to the post-incubation groundwater Fe, sediment Fe, and sediment As concentration data since the raw data for each variable had a non-normal distribution (Gotelli & Ellison, 2004). Then, significant differences in post-incubation aquifer sediment C, Fe and As concentrations and groundwater Fe and As concentrations by DOM treatment and sampling site were evaluated using two-way ANOVA tests and Tukey's HSD tests in R (R Core Team, 2012). Also, biplots were used to investigate the relationship between sediment C and Fe concentrations and groundwater As concentrations in post-incubation samples. Statistical analyses were not performed on sediment N concentrations because samples from K240 8m and K60 14m had N concentrations below the instrument detection limit (0.01%).

A two-way ANOVA followed by a Tukey's HSD test performed in R (R Core Team, 2012) was used to investigate the influence of the incubation treatments ($n$ = 4) and the sampling site ($n$ = 4) on the phylogenetic diversity (PD) within post-incubation sediment bacterial communities. Also, a Spearman rank-order correlation test was used to evaluate the relationship between PD and the relative abundances of bacterial taxa in post-incubation bacterial communities. To evaluate whether bacterial community structure, as characterized by the unweighted UniFrac

metric, was significantly different between different DOM treatments and sampling sites, a permutational multivariate ANOVA was performed using the Adonis function in the vegan package in R (Oksanen et al., 2012). To visualize differences between pre- and post-incubation bacterial community structure, in addition to differences between communities subjected to different DOM treatments, a non-metric multidimensional scaling (NMDS) analysis was performed on the unweighted UniFrac distance matrix using the metaMDS function in the vegan package in R (Oksanen et al., 2012). A one-way ANOVA and a Tukey's HSD test were performed in order to test if there were significant differences in the relative abundances of specific taxa (>1% average relative abundance) based on DOM treatment.

The relationships between bacterial community diversity and DOM chemistry in CL and LOM post-incubation samples were evaluated. First, the five-component PARAFAC model loadings data for each CL and LOM incubation sample were used to calculate a Euclidean distance matrix. Then, a Mantel test was employed in the vegan package of R (Oksanen et al., 2012) to test the relationship between DOM chemistry, as characterized by the Euclidean distance of the PARAFAC model loadings, and bacterial community structure (unweighted UniFrac distances). This analysis was followed by Spearman rank-order correlation tests to identify the taxa (>1% of the bacterial community) within the bacterial community whose abundances were significantly related to the intensity of the five fluorophores identified in the PARAFAC model (Table 3.4).

## 3.3 Results

### 3.3.1 *DOM chemistry in pre-incubation samples*

Sampling site B 14m, where the highest groundwater As concentrations are observed, had the highest groundwater DOC and TDN concentrations of the four sampling sites, but the lowest DOC:TDN (Table 3.1). By contrast, the lowest groundwater As site, K240 8m, had the lowest groundwater DOC, TDN, and the highest DOC:TDN (Table 3.1). Although B 11m and K60 14m were similar with respect to groundwater DOC:TDN, the two sites had contrasting fluorescent DOM (FDOM) chemical characteristics (Table 3.3). The DOM in groundwater collected at B 11m had a higher fluorescence index (FI), indicating a more microbial signature, whereas at K60 14m the DOM had a plant-derived signature (Table 3.3). In addition, the groundwater DOM at K60 14m had a higher aromaticity than B 11m, as indicated by its higher specific UV absorbance ($SUVA_{254}$; Table 3.3). In general, groundwater at Site B had a greater amount of DOM fluorophores than either of the sampling locations at Site K (Table 3.3). For example, the amino acid-like (AA) fluorescence was more than an order of magnitude higher at Site B 11m and 14m than at K240 8m or K60 14m (Table 3.3). Also, under ambient conditions in the aquifer the DOM at 11m and 14m at Site had more reduced quinone fluorophores than the sampling locations at Site K.

3.3.2 *Sediment geochemistry in post-incubation samples*

The percent C in post-incubation aquifer sediments was significantly different between sampling sites (two-way ANOVA, $F(3,30) = 41.9$, $p < 1E\text{-}10$; Table 3.5), and between the DOM treatments ($F(3,30) = 6.1$, $p = 0.003$; Table 3.5); however the interaction between sampling site and treatment was not statistically significant ($F(9,30) = 1.7$; $p > 0.1$). ROM post-incubation sediments had significantly higher sediment C carbon concentrations than the CL (Tukey's HSD $p = 0.01$) or LOM (Tukey's HSD $p = 0.006$) post-incubation sediments (Table 3.5).

Table 3.5. A) The percent carbon (C) and B) percent nitrogen (N) in post- incubation aquifer sediment, normalized to sediment weight in each incubation vial, under the control (CL), the labile DOM (LOM), the redox- active DOM (ROM), and the labile DOM and redox- active DOM (LOM+ROM) treatments. Percent C and N are reported as means of the three experimental replicates within a treatment/site ± one standard deviation of the mean.

A) A two –way ANOVA test and a Tukey's HSD test showed that there were significant differences in percent C between DOM treatments, as shown by the italicized letters underneath the treatment names. Also, these tests revealed that percent sediment C was significantly different between some of the sites, as indicated by the symbols next to the site names.

B) The sediments from K240 8m and K60 14m had sediment nitrogen concentrations that were below the detection limit (b.d.l.; 0.01%) of the instrument, and thus no statistical analyses were performed on this data.

**A) *Percent C in post- incubation aquifer sediment***

| SITE | CL | LOM | ROM | LOM + ROM |
|---|---|---|---|---|
| | ***a*** | ***a*** | ***b*** | ***ab*** |
| **B 11m*** | 0.05 ± 0.02 | 0.07 ± 0.02 | 0.07 ± 0.01 | 0.06 ± 0.02 |
| **B 14m⌘** | 0.07 ± 0.02 | 0.05 ± 0.01 | 0.09 ± 0.02 | 0.09 ± 0.01 |
| **K240 8m§** | 0.02 ± 0 | 0.02 ± 0.01 | 0.04 ± 0.02 | 0.03 ± 0.005 |
| **K60 14m§** | 0.02 ± 0 | 0.03 ± 0.002 | 0.04 ± 0.006 | 0.03 ± 0.004 |

**B) *Percent N in post- incubation aquifer sediment***

| SITE | CL | LOM | ROM | LOM + ROM |
|---|---|---|---|---|
| **B 11m** | 0.006 ± 0.001 | 0.007 ± 0.001 | 0.007 ± 6E-4 | 0.006 ± 0.001 |
| **B 14m** | 0.007 ± 0.001 | 0.005 ± 3E-4 | 0.005 ± 5E-4 | 0.005 ± 0.001 |
| **K240 8m** | 0 | 0 | 0 | 0 |
| **K60 14m** | 0 | 0 | 0 | 0 |

Also, the post-incubation sediments from different sampling sites had significantly different percent C ($p < 0.05$; Table 3.5), except for K240 8m and K60 14m, which had the lowest sediment C concentrations ($p > 0.6$; Table 3.5). Similarly, there were significant differences in post-incubation sediment Fe concentrations between the sampling sites overall ($F(3,30) = 37.1$, $p < 1E\text{-}9$; Table 3.6), with significantly higher sediment Fe concentrations in post-incubation samples from B 14m than the other sites ($p < 0.01$; Table 3.6). Although there were significantly different sediment Fe concentrations in the post-incubation samples based on DOM treatment ($F(3,30) = 3$, $p = 0.05$; Table 3.6), the intra-site differences in sediment Fe concentrations were greater than the intra-treatment differences in post-incubation samples. By contrast, there were no significant differences in sediment As concentrations between sampling sites or OM treatments in post-incubation sediments ($p > 0.1$; Table 3.6). Biplots indicated that there was a positive linear relationship between sediment C concentrations and groundwater As concentrations, as well as between sediment Fe concentrations and groundwater As concentrations in post-incubation samples.

### 3.3.3 *Groundwater geochemistry in post-incubation samples*

Groundwater Fe and As concentrations in post-incubation samples were significantly influenced by sampling site and DOM treatment, as demonstrated by two-way ANOVA tests (Figures 3.2 and 3.3). There were significant differences in post-incubation groundwater Fe concentrations based on sampling site ($F(3,30) = 38.2$, $p < 3E\text{-}10$), DOM treatment ($F(3,30) = 96.4$, $p < 2E\text{-}15$), and the interaction between site and treatment ($F(9,30) = 6.9$, $p < 3E\text{-}5$). Post-incubation groundwater Fe concentrations at K240 8m were significantly higher than the other sampling sites (Tukey's HSD, $p < 1E\text{-}7$; Figure 3.2).

Table 3.6. A) The total Fe concentrations (ppm), and B) total As concentrations (ppb), in post-incubation aquifer sediment under each DOM treatment. Sediment Fe and As concentrations are reported as means of the three experimental replicates within a treatment/site ± one standard deviation of the mean.

A) A two –way ANOVA test and a Tukey's HSD test showed that there were significant differences in sediment Fe concentrations between DOM treatments, as shown by the italicized letters underneath the treatment names. Also, these tests revealed that percent sediment C was significantly different between some of the sites, as indicated by the symbols next to the site names.

B) There were no significant differences in sediment As concentrations between sampling sites or DOM treatments.

**A)** ***Fe concentrations (ppm) in aquifer sediment***

| **SITE** | **CL**<br>***ab*** | **LOM**<br>***a*** | **ROM**<br>***b*** | **LOM + ROM**<br>***ab*** |
|---|---|---|---|---|
| **B 11m*** | 9612 ± 1189 | 11176 ± 1777 | 8536 ± 1891 | 8649 ± 1631 |
| **B 14m⌘** | 12920 ± 1394 | 12627 ± 5053 | 12984 ± 4103 | 15099 ± 3461 |
| **K240 8m§** | 7033 ± 1267 | 7495 ± 3097 | 4401 ± 574 | 5678 ± 1306 |
| **K60 14m§** | 4772 ± 36 | 6398 ± 984 | 4375 ± 787 | 5130 ± 1393 |

**B)** ***As concentrations (ppb) in aquifer sediment***

| **SITE** | **CL** | **LOM** | **ROM** | **LOM + ROM** |
|---|---|---|---|---|
| **B 11m** | 10509 ± 7435 | 16968 ± 4680 | 24522 ± 9365 | 15359 ± 7916 |
| **B 14m** | 25206 ± 10193 | 20626 ± 27670 | 14474 ± 5907 | 45884 ± 53594 |
| **K240 8m** | 9940 ± 7044 | 10676 ± 4712 | 49132 ± 56666 | 8222 ± 2583 |
| **K60 14m** | 5232 ± 2123 | 4896 ± 1058 | 16290 ± 15643 | 10512 ± 376 |

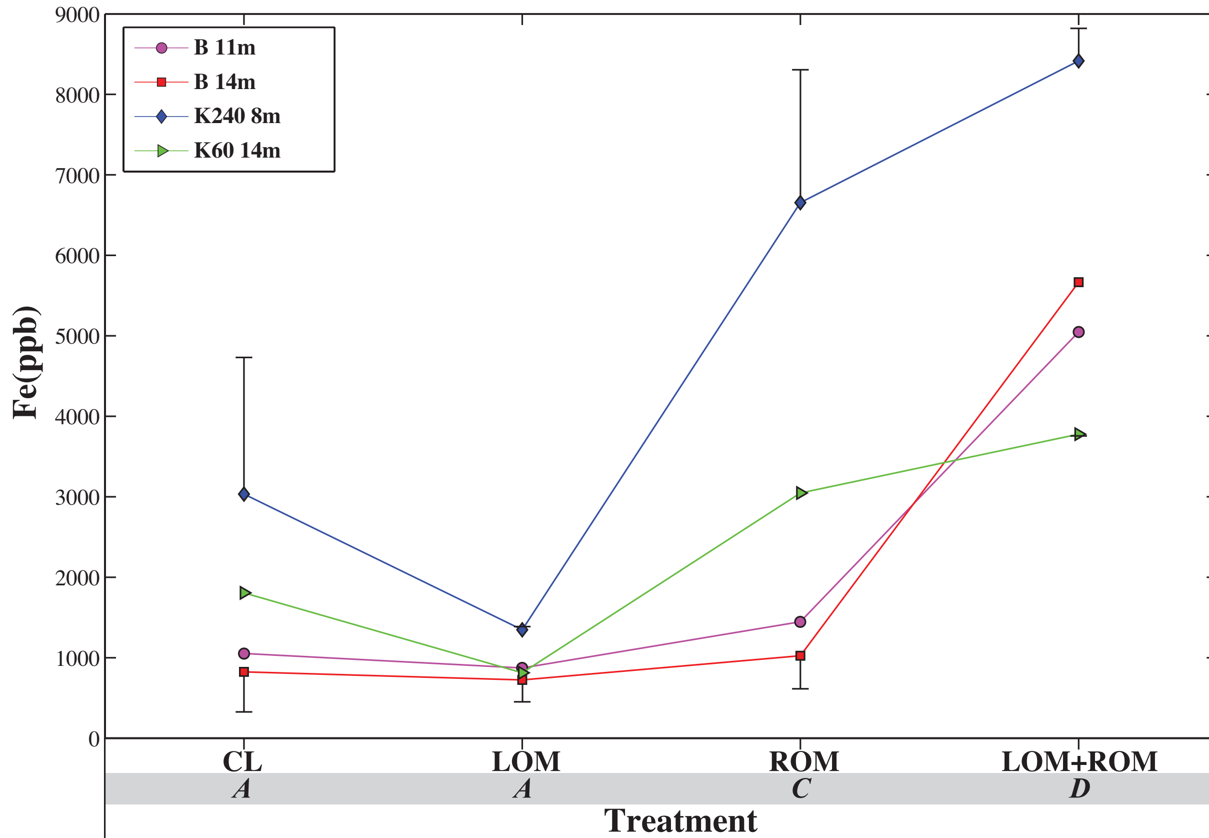



Figure 3.2. Average post-incubation groundwater Fe concentrations for triplicate technical replicates (same site and treatment). Error bars represent average post-incubation Fe concentrations ± two standard deviations. Groundwater Fe concentrations in CL and LOM post-incubation samples were significantly lower than concentrations in ROM or LOM+ROM post-incubation samples. The LOM+ROM post-incubation samples had significantly higher groundwater Fe concentrations than in the other treatments.

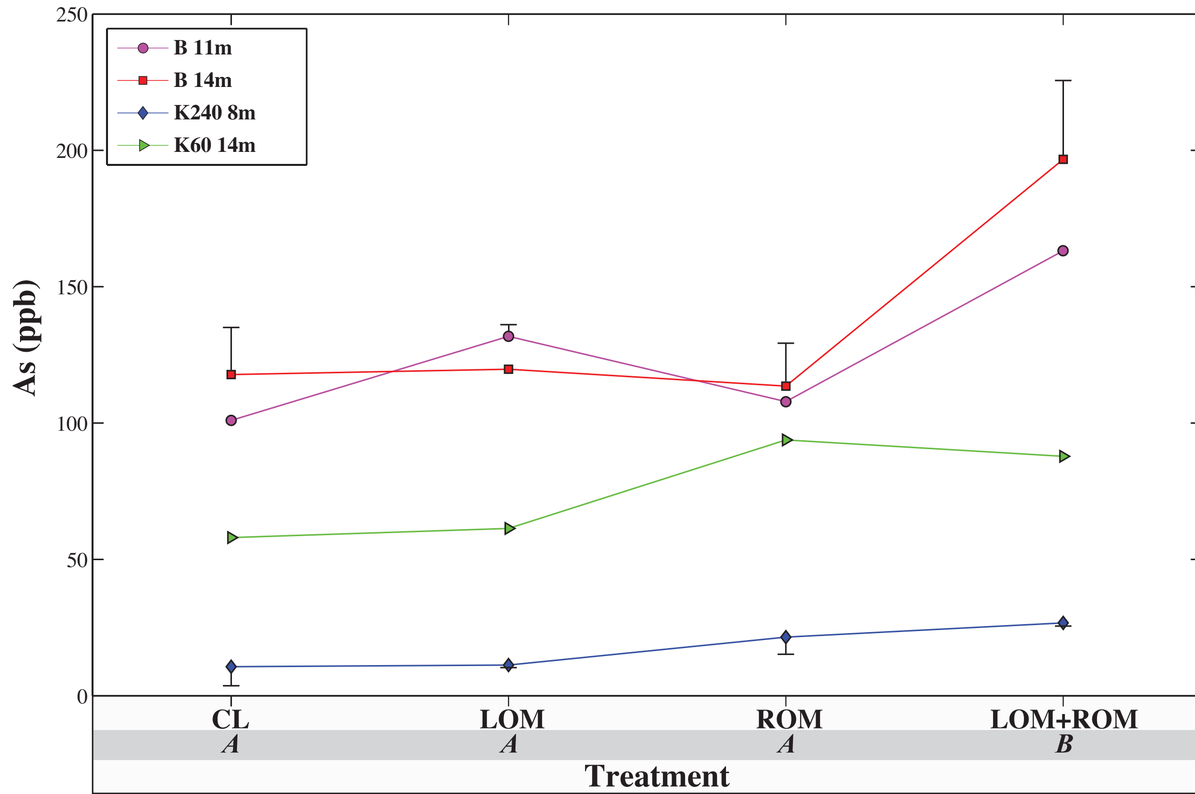



Figure 3.3. Average post-incubation groundwater As concentrations for triplicate technical replicates (same site and treatment). Error bars represent average post-incubation As concentrations ± two standard deviations. Groundwater As concentrations were significantly higher in LOM+ROM post-incubation samples than in the other treatments.

Also, post-incubation groundwater Fe concentrations were significantly higher in LOM+ROM post-incubation samples than the other treatments (Tukey's HSD, $p < 7E\text{-}4$; Figure 3.2). Likewise, post-incubation groundwater As concentrations were significantly different between DOM treatments ($F(3,30) = 33.7$, $p \leq 1E\text{-}9$), with significantly higher As concentrations observed in the LOM+ROM post-incubation samples ($p \leq 2E\text{-}6$; Figure 3.3). There was a highly statistically significant effect of sampling site on groundwater As concentrations ($F(3,30) = 233.7$, $p < 2E\text{-}16$). Also, the interaction between site and DOM treatment significantly influenced post-incubation groundwater As concentrations ($F(9,30) = 7$, $p < 3E\text{-}5$).

3.3.4 *DOM incubation treatments significantly influenced bacterial community structure*

The DOM incubation treatments significantly influenced the composition of post-incubation bacterial communities ($F(3,32) = 5.4$, $p \leq 0.001$). The average proportion of Betaproteobacteria and Gammaproteobacteria in pre-incubation aquifer sediments was 20%, yet these taxa comprised a much lower proportion of post-incubation bacterial communities (Figure 3.4). The average relative abundance of Betaproteobacteria in CL post-incubation samples varied between 2 and 6%, whereas Gammaproteobacteria comprised an average of 1-5% of CL post-incubation samples. In LOM post-incubation samples, the average relative abundance of Betaproteobacteria ranged from 2% to 12%, while the average relative abundance of Gammaproteobacteria ranged from 0% to 27% of bacterial communities. Betaproteobacteria comprised between 4% and 21% of ROM post-incubation bacterial communities, and 1% to 9% of bacterial communities in LOM+ROM post-incubation samples. In ROM post-incubation samples, the average relative abundance of Gammaproteobacteria varied between 1% and 27%, whereas their average relative abundance ranged from 0% to 20% in LOM+ROM post-incubation samples.

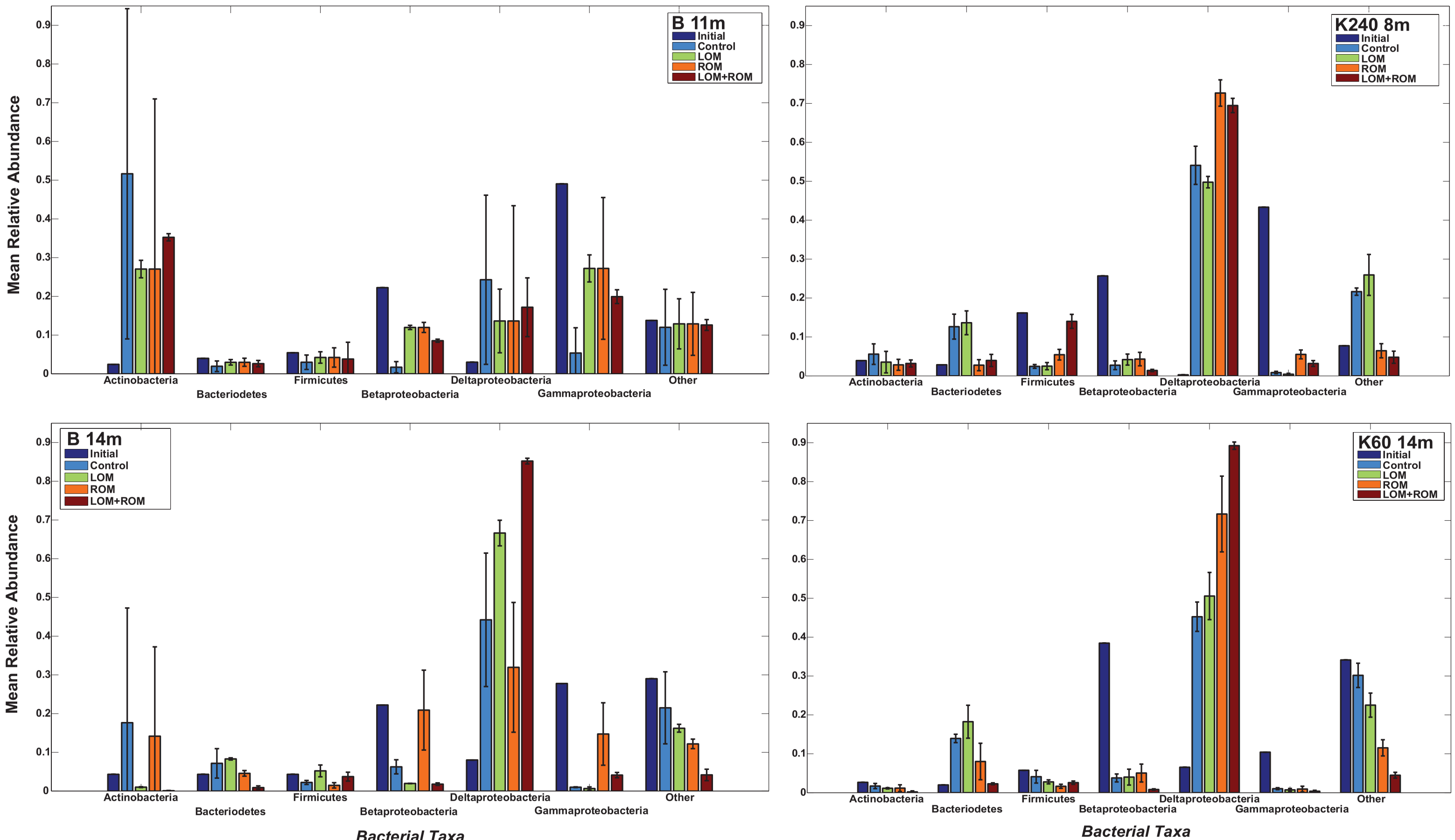



Figure 3.4. The relative abundances of bacterial taxa in pre-incubation (initial) and post-incubation samples. For post-incubation samples, colored bars represent the average relative abundances in triplicate technical replicates (same site and treatment) while error bars represent the average relative abundances ± two standard deviations.

The average relative abundance of Chloroflexi in pre-incubation samples ranged from 0.1% to 12.5%, yet in CL post-incubation samples Chloroflexi only comprised 0.5% to 1.5% of the bacterial community. The average relative abundance of Chloroflexi in LOM, ROM and LOM+ROM post-incubation samples varied from 0.5% to 1%. By contrast, the average relative abundance of Deltaproteobacteria was considerably lower in pre-incubation samples than in post-incubation samples (Figure 3.4). Deltaproteobacteria comprised between 0% and 8% of pre-incubation bacterial communities; however, their average relative abundance in CL post-incubation samples varied between 24% and 54%. In LOM and ROM post-incubation samples, the range in the average relative abundance of Deltaproteobacteria was from 14% to ~70%. The average relative abundance of Deltaproteobacteria in LOM+ROM post-incubation samples ranged between 17% and 90%.

The relative abundance of Deltaproteobacteria in post-incubation samples was significantly different between DOM treatments (two-way ANOVA, $F(3,32) = 25$, $p < 2E\text{-}8$), but not between sampling sites ($F(3,32) = 2.5$, $p > 0.05$). The average relative abundance of Deltaproteobacteria (~80% of the bacterial community) was significantly higher in the LOM+ROM post-incubation samples than in the other DOM treatments (Tukey's HSD, $p < 1E\text{-}4$). In LOM+ROM post-incubation samples, OTUs classified within the Desulfuromonadales order of the Deltaproteobacteria were significantly higher than in any other treatments (p-value $\leq 0.0003$). In ROM post-incubation sediments, Desulfuromonadales comprised an average of 95% of deltaproteobacterial OTUs, while in CL and LOM post-incubation samples 80% of Deltaproteobacteria were Desulfuromonadales OTUs. By contrast, Desulfobacterales, the second most abundant Deltaproteobacteria order, comprised approximately 6% of CL and LOM post-

incubation bacterial communities, and less than 1% in the ROM and LOM+ROM post-incubation samples.

Bacterial community diversity was significantly influenced by DOM incubation treatments. The phylogenetic diversity (PD) within bacterial communities was substantially higher in pre-incubation aquifer sediment samples than in post-incubation samples (Table 3.7). The PD of post-incubation bacterial communities was significantly different between DOM treatments (two-way ANOVA, $F(3,32) = 27.7$, $p < 6E\text{-}9$; Table 3.7). The PD of bacterial communities in ROM and LOM+ROM post-incubation sediments was significantly lower than the PD of CL and LOM bacterial communities (Table 3.7). Also, there was a significant negative relationship between bacterial community PD and the relative abundance of Deltaproteobacteria in post-incubation samples (Rho = -0.68, $p \leq 5E\text{-}7$). In addition, a permutational MANOVA test using the adonis function (Oksanen et al., 2012) showed that the beta diversity of post-incubation bacterial communities, as characterized by UniFrac distances, was significantly different between DOM treatments ($F(3,32) = 5.4$, $p \leq 0.001$; Figure 3.5), sampling sites ($F(3,32) = 3.8$, $p \leq 0.001$), and the interaction between DOM treatment and sampling site ($F(9,32) = 1.6$, $p \leq 0.001$). The UniFrac distances between the initial bacterial communities from the four different sampling sites (mean = 0.78) was greater than the UniFrac distances between bacterial communities within the same treatment group (mean = 0.59, standard deviation = 0.05; Figure 3.5). Also, the average UniFrac distance between the CL and LOM post-incubation bacterial communities was smaller than the distances between the communities from the other DOM treatments (Table 3.8), indicating that the structures of CL and LOM post-incubation bacterial communities were the most similar.

Table 3.7. The phylogenetic diversity (PD) of bacterial communities in the initial aquifer sediments (pre-incubation) and in post-incubation samples. A two-way ANOVA test and a Tukey's HSD test showed that there were significant differences in PD between DOM treatments, as shown by the italicized letters underneath the treatment names (initial samples were not included in the statistical analyses). There were no significant differences in PD between sampling sites.

| **SITE** | **Initial** | **CL**<br>***a*** | **LOM**<br>***a*** | **ROM**<br>***b*** | **LOM+ROM**<br>***c*** |
|---|---|---|---|---|---|
| **B 11m** | 34 | 18 ± 5 | 18 ± 2 | 12 ± 4 | 10 ± 1 |
| **B 14m** | 13 | 15 ± 2 | 15 ± 2 | 15 ± 4 | 8 ± 0.5 |
| **K240 8m** | 32 | 18 ± 1 | 17 ± 2 | 13 ± 3 | 12 ± 1 |
| **K60 15m** | 49 | 19 ± 1 | 17 ± 1 | 11 ± 1 | 8 ± 2 |

Table 3.8. The mean and standard deviations for UniFrac distances between post- incubation bacterial communities in samples under different DOM treatments.

| **Mean ± St. Dev.** | **CL** | **LOM** | **ROM** | **LOM+ROM** |
|---|---|---|---|---|
| **CL** | ---- | | | |
| **LOM** | 0.62 ± 0.048 | ---- | | |
| **ROM** | 0.69 ± 0.036 | 0.69 ± 0.031 | ---- | |
| **LOM+ROM** | 0.72 ± 0.037 | 0.70 ± 0.041 | 0.65 ± 0.054 | ---- |

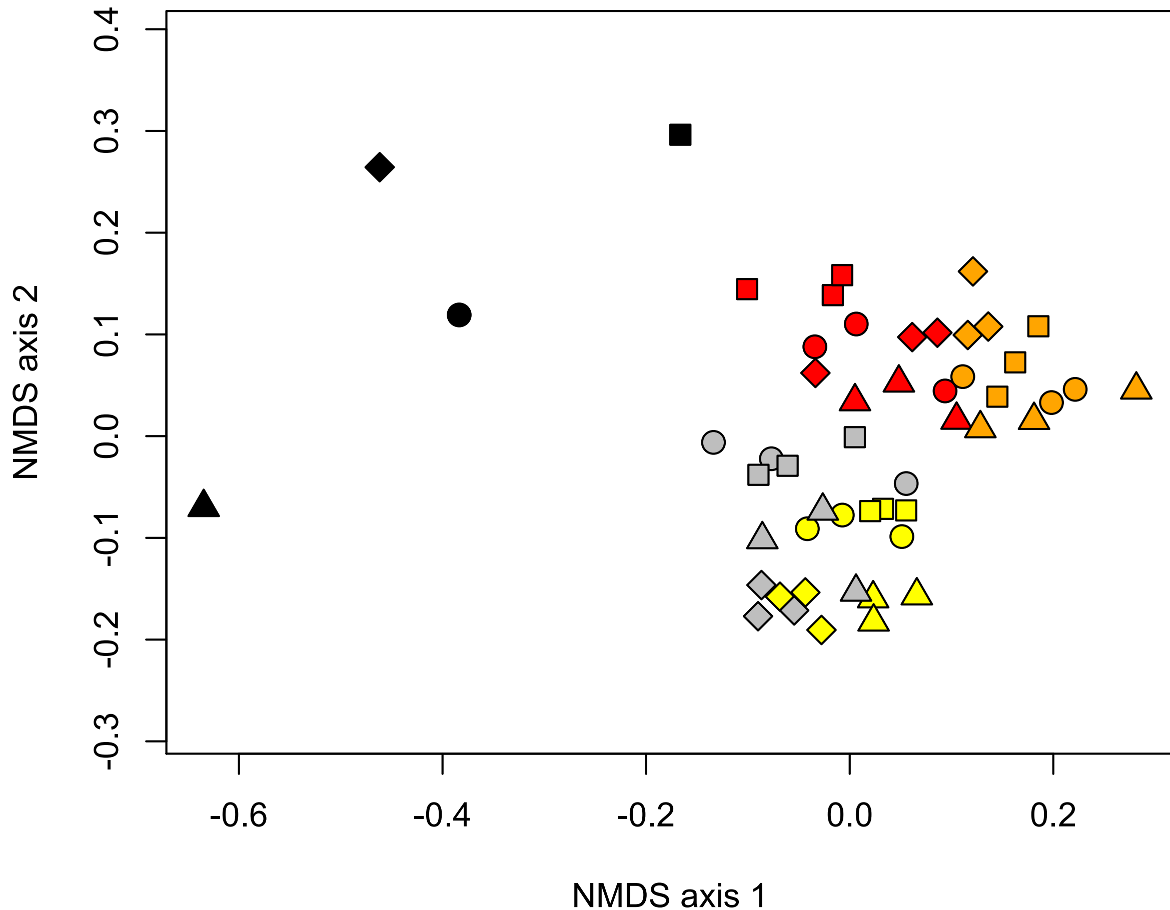



Figure 3.5. This non-metric multidimensional scaling (NMDS) ordination shows that bacterial communities cluster based on DOM treatment. In addition, the pre-incubation bacterial communities from the four sampling sites are more different from one another than the post-incubation bacterial communities within a DOM treatment. Each point on the ordination represents a different bacterial community ($n$ = 52). The black points represent bacterial communities in pre-incubation samples ($n$ = 4). The circle-shaped points represent bacterial communities from Site B 11m, while the square-shaped points represent bacterial communities from Site B 14m. The diamond-shaped points represent bacterial communities from Site K240 8m, and the triangle-shaped points represent bacterial communities from Site K60 14m. Bacterial communities in CL post-incubation samples are represented by grey points ($n$ = 12), whereas yellow points represent LOM post-incubation bacterial communities ($n$ = 12). The red points represent ROM post-incubation bacterial communities ($n$ = 12), and orange points represent LOM+ROM post-incubation bacterial communities ($n$ = 12).

Finally, ROM and LOM+ROM post-incubation bacterial communities were more similar to one another than to the other post-incubation bacterial communities (Table 3.8).

3.3.5 *Relationships between bacterial community structure and fluorescent DOM constituents*

DOM chemistry, as characterized by the five PARAFAC model loadings, varied between sampling sites (Table 3.9). Bacterial beta diversity, as characterized by unweighted UniFrac distances, was significantly related to the Euclidean distance matrix of the five PARAFAC model components in CL and LOM post-incubation samples (Mantel test, $r_s = 0.38$; $p \leq 0.001$). In CL and LOM post-incubation groundwater, the relative abundance of Chloroflexi was positively correlated to RQ fluorescence (Rho = 0.34, $p = 0.001$), HS fluorescence (Rho = 0.30, $p = 0.007$), and the fluorescence of UB, an unknown fluorophore that was unique to these incubation samples (Rho = 0.27, $p = 0.01$). In addition, the relative abundance of Dehalococcoidetes, a class within Chloroflexi, was significantly correlated with the intensity of the RQ (Rho = 0.52; $p = 0.001$) and the HS (LC2; Rho = 0.38; $p = 0.01$) fluorophores in post-incubation groundwater. The relative abundance of Deltaproteobacteria was significantly correlated with the SQ (Rho = 0.45; $p \leq 0.04$), but none of the other fluorescent DOM components. The relative abundances of other abundant taxa (>1%) were not significantly related to any of the five PARAFAC model components in the post-incubation samples.

In addition, the source (FI) and aromaticity ($SUVA_{254}$) of groundwater DOM in CL and LOM post-incubation samples (Table 3.10) were significantly related to the beta diversity of bacterial communities. There was a significant relationship between the unweighted UniFrac distance matrix and the FI (Mantel test, $r_s = 0.38$; $p = 0.001$) and the SUVA (Mantel test, $r_s = 0.27$; $p = 0.002$) of post-incubation groundwater.

Table 3.9. The loadings for each component in the PARAFAC model in A) Control (CL), and B) labile DOM (LOM), post- incubation groundwater samples. PARAFAC model loadings are reported as means of the three experimental replicates within a treatment/site ± one standard deviation of the mean.

| **A) *Post- incubation groundwater in CL incubations*** | | | | | |
|---|---|---|---|---|---|
| **SITE** | **RQ** | **HS** | **SQ** | **UB** | **AA** |
| **B 11.3m** | 0.17 ± 0.01 | 1.32 ± 0.08 | 0 | 0.83 ± 0.10 | 0.02 ± 0.02 |
| **B 13.1m** | 0.37 ± 0.07 | 4.26 ± 1.33 | 0 | 1.51 ± 0.23 | 0 |
| **K240 7m** | 0.05 ± 0.01 | 0.34 ± 0.06 | 0.001 ± 0.001 | 0.59 ± 0.08 | 0 |
| **K60 12.8m** | 0.05 ± 0.01 | 0.34 ± 0.06 | 0.001 ± 0.001 | 0.59 ± 0.08 | 0 |
| **B) *Post- incubation groundwater in LOM incubations*** | | | | | |
| **SITE** | **RQ** | **HS** | **SQ** | **UB** | **AA** |
| **B 11.3m** | 0.19 ± 0.03 | 1.60 ± 0.16 | 0.006 ± 0.007 | 0.47 ± 0.09 | 0.03 ± 0.04 |
| **B 13.1m** | 0.17 ± 0.02 | 1.32 ± 0.10 | 0.01 ± 0.01 | 0.96 ± 0.21 | 0.02 ± 0.02 |
| **K240 7m** | 0.06 ± 0.01 | 0.43 ± 0.02 | 0.001 ± 0.002 | 0.95 ± 0.09 | 0 |
| **K60 12.8m** | 0.11 ± 0.006 | 0.70 ± 0.10 | 0 | 1.13 ± 0.45 | 0 |

Table 3.10. The fluorescence index (FI) and specific UV absorbance ($SUVA_{254}$) of post-incubation groundwater. The FI and SUVA values are reported as means of the three experimental replicates within a treatment/site ± one standard deviation of the mean.

| | **FI** | | **SUVA** | |
|---|---|---|---|---|
| **SITE** | **CL** | **LOM** | **CL** | **LOM** |
| **B 11.3m** | 1.46 ± 0.01 | 1.46 ± 0.01 | 3.65 ± 2.11 | 1.56 ± 0.26 |
| **B 13.1m** | 1.45 ± 0.02 | 1.47 ± 0.05 | 1.45 ± 0.02 | 1.02 ± 0.56 |
| **K240 7m** | 1.28 ± 0.02 | 1.24 ± 0.01 | 0.71 ± 0.49 | 0.27 ± 0.27 |
| **K60 12.8m** | 1.11 ± 0.25 | 1.33 ± 0.01 | 2.09 ± 1.38 | 0.78 ± 0.19 |

The FI was positively correlated with the relative abundance of Chloroflexi in CL and LOM post-incubation samples (Rho = 0.22; $p = 0.02$). Also, Gammaproteobacteria and Bacteriodetes were more abundant in CL and LOM post-incubation groundwater with higher SUVA values (Rho = 0.25; $p = 0.03$).

### 3.4 Discussion

This experiment examined the relative effects of DOM chemistry vs. sediment geochemistry on bacterial community structure and Fe and As mobilization rates in an aquifer affected by geogenic arsenic contamination in Araihazar, Bangladesh. Specifically, this research investigated whether labile DOM (LOM) and redox-active DOM (ROM) treatments had a greater influence on bacterial community structure and Fe and As mobilization than site-specific geochemical differences. My results indicate that although site had a greater effect on sediment C, Fe and groundwater As in post-incubation samples, the type of DOM treatment had a greater effect on post-incubation groundwater Fe concentrations and bacterial community composition and diversity. Additionally, the ROM and LOM+ROM treatments promoted a decrease in bacterial community diversity, which coincided with a marked increase in the relative abundance of Deltaproteobacteria and groundwater Fe and As concentrations. Thus DOM additions promoted significant shifts in bacterial community structure that may result in an acceleration of the reductive dissolution of Fe-oxide minerals and the subsequent release of As into groundwater.

This incubation experiment confirmed previous research findings that show that the geochemistry of aquifer sediments display considerable variability, even across small spatial scales (Horneman et al., 2004; Radloff et al., 2008; van Geen et al., 2004). The variability between sites observed in post-incubation sediment geochemistry was consistent with the

patterns observed in pre-incubation sediments. For example, sediment C and Fe concentrations were significantly higher at B 11m and B 14m than at K240 8m or K60 14m in pre-incubation and post-incubation aquifer sediments (Tables 3.5 and 3.6). Nevertheless, the magnitude of variability within technical replicates (i.e., sediment samples within the same homogenized sediment core) was remarkable. Sediment C varied by an average of 40% among post-incubation technical replicates, whereas the average variation in sediment Fe and As concentrations within technical replicates was 78% and 120%, respectively. Hence, the geochemical shifts induced by DOM treatments in post-incubation sediment chemistry were muted by the significant spatial heterogeneity in sediment geochemistry.

Previous research has shown that differences between sites in the geochemical characteristics of aquifer sediments, including the grain size and the C and Fe content, is partially responsible for the spatial variability in groundwater As mobilization observed in incubation experiments (Radloff et al., 2008; van Geen et al., 2004), as well as under ambient conditions in Bangladesh aquifers (Dowling et al., 2002; Zheng et al., 2005). My results suggested that site had a greater influence than DOM treatment on post-incubation sediment C and Fe concentrations, as well as groundwater As concentrations, providing additional support for this concept. Groundwater As concentrations in post-incubation samples were positively correlated to sediment C and Fe concentrations, but not sediment As concentrations. The decoupling between sediment As and groundwater As has been observed previously (Dhar et al., 2008), and is likely due to the fact that only a fraction of sediment As is mobilizable in the aquifer environment (Zheng et al., 2005).

Sediment geochemistry also significantly influences the structure of bacterial communities in aquifer sediments (Legg et al., 2012), which represent another key control factor on groundwater

As mobilization. Since bacteria in the high As aquifer environment rely on organic matter for carbon to build biomass and electron donors for respiration, and on Fe(III) as an electron acceptor for dissimilatory Fe-reduction (McArthur et al., 2004; Rowland et al., 2007), it is not surprising that sediment C and Fe influence the composition and diversity of bacterial communities. Although the initial bacterial communities at each of the sampling sites were distinct, the DOM treatments significantly affected the structure of post-incubation bacterial communities (Figure 3.5). In general, all of the incubation treatments catalyzed decreases in bacterial community diversity across sampling sites (Table 3.7). In CL post-incubation samples, it is likely that an increase in the groundwater to sediment ratio from ~1:8 in the environment (Islam et al., 2004) to 5:1 in the incubation experiments substantially increased the DOM concentrations since there are generally higher concentrations of organic matter in groundwater than in aquifer sediments at our sites. Also, high organic C concentrations could have been maintained during the incubation through the release of sedimentary organic matter, as indicated by the lower FI in post-incubation samples (Table 3.10) compared to the pre-incubation groundwater (Table 3.3).

The structure of bacterial communities in CL post-incubation samples was most similar to that of LOM post-incubation samples (Table 3.8; Figure 3.5), suggesting that the native groundwater DOM was relatively labile. Thus, the artificially high labile DOM concentrations in CL and LOM samples may have provided a competitive advantage to copiotrophic taxa while negatively affecting more oligotrophic taxa (Fierer et al., 2007; Hu et al., 1999), which comprise the majority of low-nutrient aquifer sediment communities (Chapelle, 2001; Legg et al., 2012). In fact, organisms related to copiotrophic, human-associated Bacteriodetes (Krieg et al., 2011), which could have been transported to aquifer sediments through latrine and pond recharge

(Knappett et al., 2011), were more abundant in CL and LOM post-incubation bacterial communities than in the initial aquifer sediments. In addition, the average relative abundance of Deltaproteobacteria was an order of magnitude greater in CL and LOM post-incubation bacterial communities than in pre-incubation communities (45% vs. 4.5%), which corroborates prior research suggesting that these taxa respond to organic matter availability in the high aquifer environment (Legg et al., 2012). Together, these results demonstrate that native bacterial communities are sensitive to shifts in organic matter concentrations in the high As aquifer environment.

Likewise, the diversity and composition of post-incubation bacterial communities were sensitive to the supply of redox-active DOM. The ROM and LOM+ROM post-incubation bacterial communities were significantly less diverse (Table 3.7), and more similar to one another (Table 3.8; Figure 3.5), than the CL and LOM post-incubation communities. The lower PD of ROM and LOM+ROM post-incubation bacterial communities coincided with a marked increase in the relative abundance of Deltaproteobacteria (Figure 3.4), especially in LOM+ROM communities where Desulfuromonadales OTUs closely related to *Geobacter* species comprised nearly 80% of the bacterial community. Additionally, the relative abundance of Deltaproteobacteria in CL and LOM post-incubation samples was positively related to the availability of the SQ fluorophore, a reduced quinone moiety associated with microbial activity (Ebert, 2012). Laboratory studies have demonstrated that multiple *Geobacter* species can respire by pairing the oxidation of labile DOM to the reduction of redox-active quinone moieties within humic substances (Coates et al., 1998; Lovley et al., 1998, 1996; Luijten et al., 2004), and that bacterial quinone reduction may accelerate the reductive dissolution of Fe-oxide minerals, and thus the mobilization of As, in aquifers (Jiang & Kappler, 2008; Mladenov et al., 2010). Results

from this experiment showed that Fe and As mobilization rates within each sampling site were highest in LOM+ROM post-incubation samples (Figures 3.2 and 3.3) where Deltaproteobacteria is the predominant taxon. This incubation experiment provides further evidence that Deltaproteobacteria may be the primary mediator of groundwater As mobilization in aquifers affected by geogenic As contamination (Héry et al., 2010; Islam et al., 2004; Lear et al., 2007; Rowland et al., 2007).

However, other bacterial taxa may also perform dissimilatory quinone reduction in the aquifer environment. Although Chloroflexi comprised a relatively small proportion of post-incubation bacterial communities (~1%), there was a positive correlation between the relative abundance of Chloroflexi and the intensity of RQ, a reduced quinone moiety (Ebert, 2012), as well as a humic substance-like moiety (HS; Ebert, 2012) in CL and LOM post-incubation samples. It is possible that Chloroflexi also produce fluorescent DOM, since they were more abundant in samples with a microbial signature (high FI; Rho = 0.22, $p = 0.02$) and with a greater intensity of UB (Rho = 0.27, $p = 0.01$), a fluorophore produced during the incubations (Ebert, 2012). Consequently, Chloroflexi could influence groundwater As cycling in the aquifer environment by producing additional sources of DOM and by reducing quinone moieties, thereby fueling the bacterially-mediated reductive dissolution of Fe-oxide minerals.

Despite the spatial variability in sediment geochemistry, DOM chemistry had a profound effect on bacterial community structure and the mobilization of Fe and As. This research confirms that native bacteria in high As aquifers posses the metabolic potential to accelerate groundwater As mobilization through respiratory quinone reduction. Yet, this experiment favored copiotrophic taxa since DOM concentrations in the incubation experiments were two to three orders of magnitude greater than DOM concentrations in the aquifer environment.

Research investigating the distribution and metabolic activity of Deltaproteobacteria, as well as oligotrophic taxa such as the Chloroflexi, under ambient conditions could elucidate the relative importance of these taxa in the biogeochemical cycling of As in the aquifer environment.

# CHAPTER 4: Novel relationships between dissolved organic matter chemistry and bacterial community composition in a high arsenic aquifer: Implications for groundwater arsenic cycling in Bangladesh

## 4.1 Introduction

Dissolved organic matter (DOM) chemistry and concentration control ecosystem processes such as nutrient mineralization rates (Bernhardt & Likens, 2002; Judd et al., 2006; Stelzer et al., 2003), light attenuation and primary productivity (Jackson & Hecky, 1980), and the mobility of contaminants (Bauer & Blodau, 2006; Kalbitz & Wennrich, 1998; Ravichandran, 2004) in aquatic ecosystems. DOM originates from the degradation of plant material and microbial biomass, and it encompasses an array of constituents with unique chemical and ecological properties. The activity and composition of microbial communities are tightly coupled with DOM chemistry (Eiler et al., 2003; Judd et al., 2006; Pinhassi et al., 1999) since microorganisms are both consumers and producers of DOM (Guillemette & Del Giorgio, 2012; Kawasaki & Benner, 2006; Ogawa et al., 2001). As DOM consumers, microorganisms use various DOM fractions to mediate the mobility and redox state of nutrients, metals and contaminants in river, lake and groundwater environments.

DOM is an important source of carbon (C) and nitrogen (N) in aquatic environments. Heterotrophic bacteria scavenge labile constituents from DOM such as amino acids, carboxylic acids and carbohydrates (Berggren et al., 2010; Coffin, 1989; Keil & Kirchman, 1991) to acquire the building blocks for synthesizing biomass. Also, some aquatic bacteria use extracellular enzymes to break down larger polysaccharides for C and N acquisition (Arnosti, 2002). A diverse group of bacterial taxa in the Firmicutes, Bacteriodetes, Gammaproteobacteria and Alphaproteobacteria can use more recalcitrant fractions of natural humic substances (HS) as

carbon and nitrogen sources (Esham et al., 2000; Hutalle-Schmelzer et al., 2010). In many environments, bacteria produce sources of DOM, for example, protein-like and humic-like products during growth in lakes (Cammack et al., 2004; Guillemette & Del Giorgio, 2012) and streams (Williams et al., 2010). In some freshwater environments, microbially-derived DOM may constitute a significant fraction of the carbon and nitrogen sources used by bacterial communities.

Bacterial communities also utilize many fractions of DOM during dissimilatory processes. Bacteria pair the oxidization of proteins, low molecular weight organic acids, and other classes of labile DOM to the reduction of oxygen (Amon & Benner, 1996), as well as a range of other electron acceptors (Lovley & Coates, 2000), during respiration. In fact, bacteria support growth by pairing the oxidation of labile DOM to the reduction of redox-active moieties within the humic fraction of DOM under anaerobic conditions (Lovley et al., 1996; Coates et al., 1998). In addition, it appears that taxa within the Proteobacteria can support growth by pairing the oxidation of humic substances to the reduction of electron acceptors such as nitrate (Coates et al., 2002). Previous work has shown that quinone moieties, which are not easily degraded and can alternately cycle between a range of oxidized and reduced states (Cory & McKnight, 2005), are the predominant electron donors and electron acceptors within humic substances (Scott et al., 1998). In anaerobic aquatic sediments, microbially-reduced quinones abiotically donate electrons to oxidized species within metal- oxide minerals (Lovley et al., 1998), thereby shuttling electrons between bacteria and insoluble electron acceptors. Indeed, it is likely that bacterial quinone respiration is ubiquitous in many anaerobic sediments because other electron acceptors are insoluble at neutral pH (Lovley et al., 1996; Kappler et al., 2004), and in some cases the energy yield from quinone-reduction may be higher than from the reduction of crystalline iron (Fe)-

oxide minerals (Straub et al., 2005). More recently, research has found that halorespiring bacteria, including *Dehalococcoides spp.* within a trichloroethylene (TCE)-degrading community, may perform dissimilatory quinone reduction (Zhang et al., 2011; Luijten et al., 2004).

Given the tight links between DOM and bacterial metabolisms, it is not surprising that DOM chemistry and source influences the structure of bacterial communities. In marine environments, researchers have extensively documented how spatial and temporal differences in DOM chemistry promote compositional and functional shifts within bacterial communities (Azam & Malfatti, 2007; McCarren et al., 2010; Mou et al., 2008). Less is known about DOM-bacteria dynamics in freshwater ecosystems, although the use of molecular techniques to characterize bacterial community structure, such as 16S rRNA gene sequencing (Pace, 1997), has facilitated new discoveries. For example, the average molecular size of constituents within DOM exerts a significant influence on the structure of bacterial communities in freshwater environments (Docherty et al., 2006; Lin et al., 2012). In fact, DOM-manipulation experiments have shown that the composition of labile DOM-amended bacterial communities is distinct from bacterial communities amended with high molecular weight DOM (Foreman & Covert, 2002; Findlay et al., 2003). It is likely that the source of DOM is an important factor in bacterial community structure since microbially-derived DOM supports bacterial communities that are taxonomically- and functionally-distinct from communities utilizing plant-derived DOM in riverine (Judd et al., 2006; Kirchman et al., 2004; Koetsier et al., 1997) and lake (Jones et al., 2009) ecosystems. In contrast to other freshwater ecosystems, little is known about the relationships between the chemistry and source of DOM and bacterial community structure across groundwater environments, although evidence suggests that DOM concentration may influence bacterial

community activity (Foulquier et al., 2011) and composition (Shirokova & Ferris, 2013) in aquifers.

Fluorescence spectroscopy paired with parallel factor analysis (PARAFAC) represents a reliable tool for characterizing the lability, the redox- reactivity, the redox-state and source of DOM (DOM) in aquatic environments. Although not all constituents within DOM fluoresce, this technique has been effective in identifying the relative contributions of labile and redox- active moieties in aquatic environments. Using this approach, researchers extract the unique spectra of DOM fluorophores from three- dimensional fluorescence excitation and emission matrices (EEMs), and then relate the distinct fluorophores to known DOM constituents, such as tryptophan and oxidized quinone moieties. For example, the Cory and McKnight (2005) PARAFAC model identified 13 components, including two amino acid-like and seven quinone-like fluorophores, in 355 water samples collected from a range of freshwater ecosystems.

Research throughout the past decade has demonstrated the utility of the EEM-PARAFAC approach to reveal details about DOM chemistry in pristine (Cawley et al., 2012; Miller et al., 2009; Pautler et al., 2012), human-influenced (Cawley et al., 2012; Tedetti et al., 2012), and engineered (Beggs et al., 2009; Murphy et al., 2011; Yu et al., 2010) ecosystems. For example, the percentage of tryptophan-like (C8 in Cory & McKnight, 2005) and tyrosine-like fluorescence (C13 in Cory & McKnight, 2005) is correlated with total hydrolysable amino acids concentrations (Yamashita & Tanoue, 2003) and biodegradable dissolved organic carbon (BDOC; (Fellman et al., 2008), and consequently these components can be used as proxies for labile DOM. The potential redox reactivity of DOM can be represented by the relative abundance (percentage) of quinone moieties (Cory & McKnight, 2005), while the redox index, a ratio of oxidized to reduced quinones, describes the redox state of DOM (Miller et al., 2006).

The fluorescence index (FI) (McKnight et al., 2001), a ratio of emission intensity at 470nm to 520nm at an excitation wavelength of 370nm (Cory & McKnight, 2005), provides an indication of the source of DOM in aquatic environments. For instance, the FI has been used to capture seasonal and spatial shifts in DOM source from a plant to a microbial source (Hood et al., 2003; Jaffé et al., 2008), and to demonstrate the role of sedimentary DOM in groundwater arsenic (As) mobilization in Southeast Asia (Mladenov et al., 2010).

Hence, fluorescence spectroscopy and EEM-PARAFAC produce information that can resolve the relative importance of different DOM fractions in the biogeochemistry of aquatic environments. The purpose of this research was to employ EEM-PARAFAC and bacterial community characterization via 16S rRNA gene pyrosequencing to improve our understanding of the biogeochemical controls on geogenic groundwater As contamination in Southeast Asia, which currently impacts over 100 million people in South and Southeast Asia (Ravenscroft et al., 2009). There is a consensus in the literature that native bacteria promote groundwater As mobilization by pairing the oxidation of labile DOM to the reductive dissolution of Fe-oxides, which desorb As from the mineral surface in the affected aquifers of Southeast Asia (Fendorf et al., 2010). Additionally, recent research has illustrated the potential for the bacterial reduction of quinone moieties within humic substances to accelerate groundwater As mobilization rates in Bangladesh and West Bengal, India (Jiang & Kappler, 2008; Mladenov et al., 2010). However, there is debate in the literature regarding the source of the labile and humic DOM to the subsurface. While some research demonstrates that microbially-derived DOM in recharge from eutrophic ponds fuels bacterial Fe- and As-reduction (Neumann et al., 2010), there is also evidence supporting the role of sedimentary DOM, which has a more plant-derived signature, in groundwater As mobilization (McArthur et al., 2004; Mladenov et al., 2010). In addition, results

from laboratory incubation experiments indicate that a relatively small group of bacterial taxa within the Deltaproteobacteria (Islam et al., 2004; Héry et al., 2010; Lear et al., 2007; Rowland et al., 2007), and to a lesser extent those within the Betaproteobacteria (Pederick et al., 2007), utilize DOM sources to mediate the biogeochemical cycling of As in anaerobic zones of Southeast Asian aquifers. However, research has also indicated that Deltaproteobacteria comprise a small percentage of the bacterial community relative to other taxa under environmental conditions (Héry et al., 2010; Rowland et al., 2007; Pederick et al., 2007; Legg et al., 2012; Islam et al., 2004), and that high As aquifer sediments harbor diverse communities of bacteria containing other taxa related to known Fe-, As- and quinone-reducers (Legg et al., 2012).

Thus, the aim of this research was to improve our understanding of the factors promoting groundwater As concentrations exceeding the WHO regulatory limit of 5 µg/L (Fawell & Mascarenhas, 2011) by characterizing the links between the chemical constituents and sources of DOM and native bacterial taxa in an affected aquifer environment. To this end, DOM chemistry and bacterial community structure were characterized in environmental samples collected from aquifers at established study sites in Araihazar, Bangladesh. I hypothesized that the chemistry and source of DOM shape the composition of bacterial communities across groundwater As concentration gradients. Some members of the bacterial community may use DOM to directly promote groundwater As mobilization by oxidizing labile DOM during Fe- and quinone-reduction, while other bacterial taxa indirectly influence As cycling as competitors for labile and quinone-containing DOM sources. This research provides evidence suggesting that labile DOM and quinone moieties in humic DOM may be more important in shaping bacterial community structure than the source of DOM in high As aquifers. In addition, my research findings suggest

that Chloroflexi and Espilonproteobacteria, in addition to Deltaproteobacteria, are important in promoting groundwater As cycling through respiratory quinone- reduction.

## 4.2 Materials and Methods

4.2.1 *Site Description*

The research was conducted in the shallow (<20 m) Holocene aquifer at two established sites, Site K (Legg et al., 2012; Radloff et al., 2008) and Site B (Mladenov et al., 2010; Stute et al., 2007), approximately 30 km northeast of Dhaka in Araihazar, Bangladesh. Site K (23° 47′ 34″ N, 90° 37′ 48″ E) is located south of a village within the floodplain of the Old Brahmaputra River, which borders the site to the east, south and west (Figure 4.1). The focus of the research at Site K was on well nests located along a groundwater flowpath from the village to the river because these wells represented the range of groundwater As concentrations observed at the site. Groundwater As concentrations generally increase with depth at each well location, and with distance along the groundwater flowpath (Table 4.1). At well nest K240, located 240 meters from the Old Brahmaputra River and within the village (Figure 4.1), groundwater As concentrations increase with depth from 2 μg $L^{-1}$ at 7.6m to approximately 90 μg $L^{-1}$ at 15.2 m (Table 4.1; Radloff et al., 2008). The highest As concentrations at Site K, approximately 400 ug $L^{-1}$, occur at K60 at the end of the groundwater flowpath (Radloff et al., 2008). At Site B (23° 46′ 51″ N, 90° 38′ 18″ E), located approximately 2 km southeast of Site K (Figure 4.1), research was conducted at single well nest where groundwater As concentrations increase from 37 μg $L^{-1}$ at 8 m depth to 555 μg $L^{-1}$ at 14 m (Table 4.1; Zheng et al., 2005).

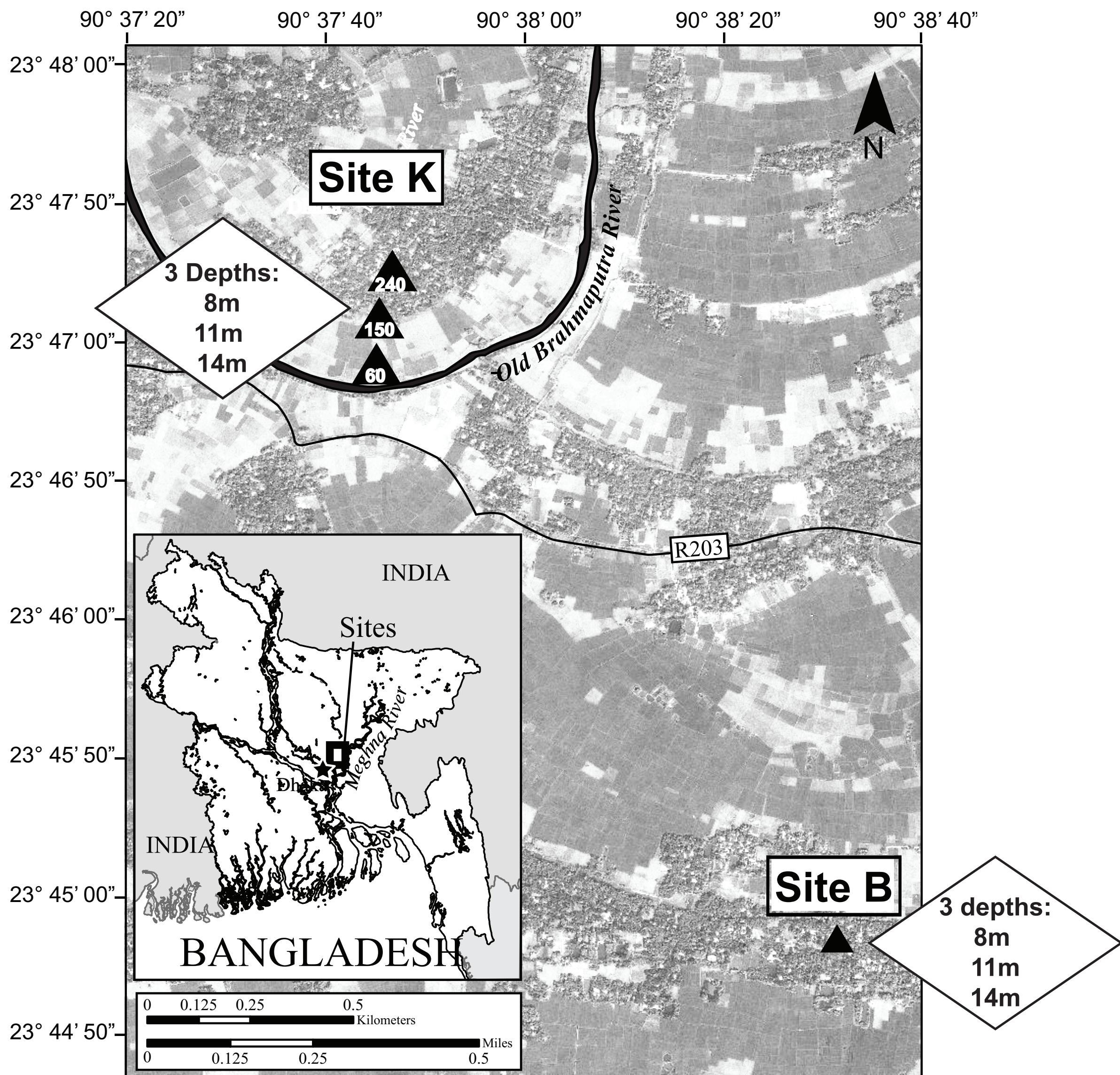



Figure 4.1. The study sites at Site K and Site B in Araihazar, Bangladesh.

Table 4.1. Water chemistry for groundwater collected at each sampling location. Data points followed by an asterisk (*) represent long-term average values (Radloff, 2010; Zheng et al., 2005), whereas the other data points were measured in October 2009. The locations/depths that are shaded were used in the incubation experiments, as well as the environmental data analysis.

| SITE | DOC (mg/L) | TDN (mg/L) | C:N | Fe (ppm) | As (ppb) |
|---|---|---|---|---|---|
| **B 8m** | 5.5 | 1.1 | 5.0 | 13.1* | 37* |
| **B 11m** | 2.6 | 1.4 | 1.9 | 1.6 / 8.3* | 132 / 366* |
| **B 14m** | 3.2 | 2.3 | 1.4 | 4.3 / 14.2* | 256 / 555* |
| **K240 8m** | 0.8 | 0.2 | 4.1 | 1.1 / 0.5* | 1 / 2* |
| **K240 11m** | 0.9 | 0.4 | 2.1 | 8.9* | 29* |
| **K240 14m** | 1.0 | 0.4 | 2.6 | 13* | 92* |
| **K150 8m** | 1.7 | 1.3 | 1.4 | 5.0* | 112* |
| **K150 11m** | 1.0 | 0.5 | 2.0 | 3.5* | 69* |
| **K150 14m** | 1.2 | 0.6 | 2.1 | 18.9* | 86* |
| **K60 8m** | 2.0 | 1.8 | 1.1 | 8.1* | 149* |
| **K60 11m** | 1.7 | 1.7 | 1.0 | 9.2* | 415* |
| **K60 14m** | 1.4 | 0.8 | 1.8 | 1.6 / 10.8* | 115 / 363* |

4.2.2 *Sample Collection*

Groundwater and sediment samples were collected at Site K and Site B in October 2009. Groundwater samples were collected at Site K from 8 m, 11 m and 14 m depth at well nests K240, K150 and K60 (located 240 m, 150 m and 60 m upgradient of the Old Brahamaputra River, respectively), and at Site B from depths of 8 m, 11 m and 14 m. Anaerobic conditions were maintained while collecting groundwater samples in amber serum vials capped with butyl stoppers. Within 24 hours of collection, groundwater samples were filtered with glass fiber (GF/F) filters (0.7 μm) under a $N_2$ atmosphere in a Captair® Pyramid glove bag, acidified to pH of 2 with 1 M HCl to prevent the precipitation of DOM-Fe complexes, and stored at 4°C in anaerobically sealed amber serum vials. Multiple sediment cores were collected from depths of 8 m, 11 m and 14 m adjacent to K240, K150, K60 and the well nest at Site B. As described in Legg et al. (2012), sample contamination was minimized by rinsing the sediment core liners with 100% ethanol prior to coring, and removing sections of sample cores that had contact with drilling fluids with an ethanol-sterilized saw. Immediately after sampling, sediment cores were stored in airtight mylar bags with oxygen-absorbing packets and placed on ice in the field. The sediment samples were stored on dry ice for approximately one week, and then at −80°C thereafter.

4.2.3 *Dissolved Organic Matter Analyses*

DOM chemistry was characterized in all groundwater samples. The redox potential of groundwater samples was maintained by preparing samples for analyses under an $N_2$ atmosphere and transferring samples to anaerobically sealed cuvettes for spectroscopic analyses. First, concentrations of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) were measured by the Kiowa Environmental Chemistry Laboratory at the University of Colorado

(http://snobear.colorado.edu/Seiboldc/kiowa.html) using a Shimadzu TOC-V CSN Total Organic Carbon Analyzer. Three-dimensional excitation-emission matrices (EEMs) were generated by running three-dimensional fluorescence scans on a Jobin-Yvon-Horiba/Spex Fluoromax-3 spectrofluorometer equipped with Datamax software in S/R mode at excitations of 240-500nm at 10nm increments, and emissions of between 350-550nm at 2nm increments. These EEMs were processed using the Cory and McKnight (2005) PARAFAC model. Using this model, it was possible to calculate the relative contributions of 13 different fluorescent DOM components (Table 4.2) and the Fluorescence Index (FI). Although all 13 components in the Cory and McKnight (2005) PARAFAC model are observed within DOM from a range of aquatic environments, the fluorescence spectra of four of the components have not yet been matched to a specific compound (e.g., oxidized quinone) are referred to as 'unidentified' fluorophores (Table 4.2).

4.2.4 *Bacterial Community Analysis*

DNA was extracted from homogenized aquifer sediment samples using the MO BIO PowerSoil® DNA Isolation Kit following the manufacturer's suggested protocol (MO BIO Laboratories, Inc., Carlsbad, CA, USA). Next, the V1–V2 region of the 16S rRNA gene was PCR-amplified from genomic DNA using the highly conserved, universal bacterial primer set as described previously (Hamady et al., 2008; Fierer et al., 2008; Nemergut et al., 2010), with a modification to the PCR amplification step (Stanish et al., 2012). Briefly, the forward primer included the 454 Roche Titanium primer B, a 2 base pair adapter, and the highly conserved bacterial primer 27F whereas the reverse primer was comprised of the 454 Roche Titanium primer A, a unique, 12 base pair error-correcting Golay barcode, a 2 base pair adapter, and the highly conserved bacterial primer 338R.

Table 4.2. The 13 components of the Cory & McKnight (2005) PARAFAC model. The 'Unidentified' components have not been matched to specific compound however, these fluorophores consistently occur in DOM samples from a range of environments. The 'microbially- associated' components were identified in an Anatarctic dataset (plant- free) and consequently these are considered to be associated with microbially- produced dissolved organic matter (Cory & McKnight 2005).

| **Component** | **Fluorophore Type** | **Abbreviation** |
|---|---|---|
| **1** | Unidentified | C1 |
| **2** | Oxidized Quinone | Q2 |
| **3** | Unidentified; Microbially- associated | C3 |
| **4** | Hydroquinone; Microbially- associated | HQ |
| **5** | Semiquinone | SQ1 |
| **6** | Unidentified; Microbially- associated | C6 |
| **7** | Semiquinone; Microbially- associated | SQ2 |
| **8** | Amino Acid- like: Tryptophan; Microbially- associated | TR |
| **9** | Semiquinone; Microbially- associated | SQ3 |
| **10** | Unidentified | C10 |
| **11** | Oxidized Quinone | Q1 |
| **12** | Oxidized Quinone; Microbially- associated | Q3 |
| **13** | Amino Acid-like: Tyrosine | TY |

Each PCR, which was performed in triplicate, consisted of 10 μl of 5-PRIME hot master mix (5-PRIME, Gaithersburg, MD, USA), 10 μl of sterile water, 1 μL of the reverse primer at 10 μM (final concentration of 0.4 μM), 2 μL of the forward primer at 5 μM (final concentration of 0.4 μM) of the forward primer, and 2 μl of genomic DNA. As described in Stanish et al. (2012), samples were initially denatured for 3 min at 94°C followed by 25 cycles at 94°C for 45 s, 50°C for 30 s, 72°C for 90 s and a final elongation step at 70°C for 10 min. PCR product from each reaction was quantified, and then the product from triplicate reactions was pooled. The pooled PCR products were purified using UltraClean PCR Clean-up kits (MoBio, Carlsbad, CA), according to the manufacturer's protocol. 16S rRNA gene amplicons were sequenced using the Titanium chemistry (454 Life Sciences, Bradford, CT, USA) on a GS FLX platform at the Environmental Genomics Core Facility (Engencore, University of South Carolina, USA).

The QIIME pipeline (Caporaso et al., 2010a) was used to process the bacterial 16S rRNA gene sequence data. An average of 1416 16S rRNA gene sequences per sample (standard deviation of 463) was obtained after denoising the sequence data (Reeder & Knight, 2010) and applying quality filters, which eliminated sequences shorter than 200 nucleotides in length, sequences with one or more ambiguous bases, and sequences with quality scores of less than 25. Bacterial operational taxonomic units (OTUs) were defined at 97% sequence identity with the uclust algorithm (Edgar, 2010), and then our dataset was filtered to eliminate OTUs represented by only one 16S rRNA gene sequence (singletons), as well as OTUs present in only one sample (Zhou et al., 2011). The PyNAST alignment algorithm (Caporaso et al., 2010b) with the Greengenes database (DeSantis et al., 2006) was used to align a representative set of 16S rRNA gene sequences that included the most abundant sequence within each OTU. Next, the RDP Classifier (Wang et al., 2007) assigned a taxonomic classification to each OTU using the

Greengenes database. The FastTree algorithm was used to create a phylogeny of the representative sequences (Price et al., 2009) which was used to calculate the phylogenetic diversity (PD; Faith, 1992) within each sample, as well as the differences between communities using the UniFrac distance metric (Lozupone & Knight, 2005). The unweighted UniFrac distance matrix, which does not consider the abundances of OTUs in calculating the distances between samples, was used for any statistical analyses investigating patterns in bacterial community beta diversity. The 16S rRNA gene sequences and accompanying environmental data have been deposited in the MG-RAST database (Meyer et al., 2008) upon publication.

4.2.5 *Statistical Analyses*

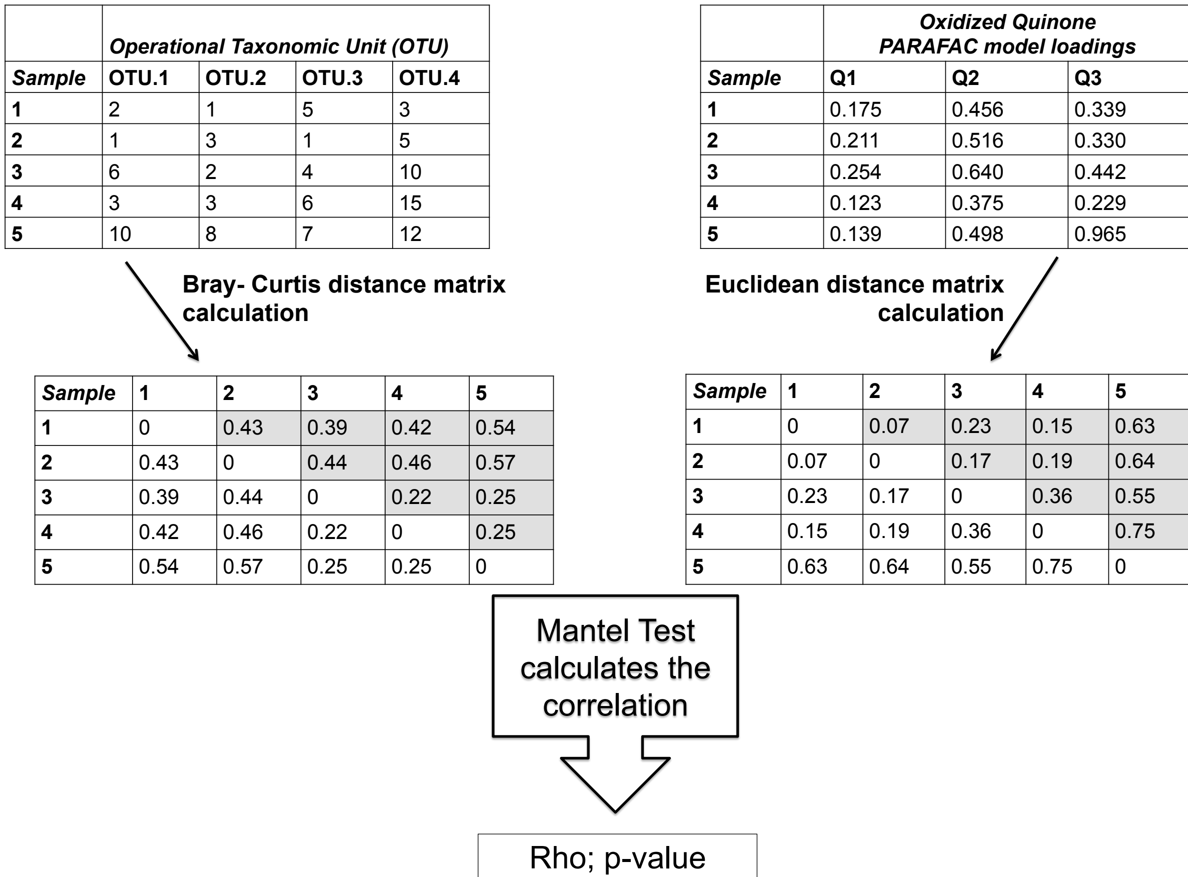
Prior to performing statistical analyses, outlier data points that fell outside of two standard deviations from the mean were removed from the dataset (Gotelli & Ellison, 2004). The relationships between DOM parameters (DOC, TDN, C:N, FI, and PARAFAC model loadings) and groundwater Fe and As concentrations were visualized in biplots. Also, relationships between bacterial community alpha diversity (phylogenetic diversity, PD) and DOM parameters were visualized using biplots. The differences in bacterial community structure between samples, as characterized by the unweighted UniFrac distance matrix, were visualized by performing a principal coordinates analysis (PCoA; Gotelli & Ellison, 2004) and plotting the scores for each sample on the first and second PCoA axes in the vegan package of R (Oksanen et al., 2012). Spearman rank- order correlations between the PCoA axis 1 scores and the 13 PARAFAC model loadings (Table 4.2) were performed to evaluate the statistical significance of relationships between bacterial community structure and DOM chemical characteristics. In addition, Spearman rank- order correlations were performed between the PCoA axis 1 scores and DOC, TDN, and C:N in order to test if the concentration and stoichiometry of DOM were significantly

related to bacterial community structure. Next the correlations between bacterial community structure and the source of DOM were evaluated using Spearman rank-order correlation tests between PCoA axis 1 scores and the FI. Then, to investigate if specific bacterial taxa were driving relationships between bacterial community structure and DOM chemistry, Spearman rank- order correlations were performed between the relative abundances of bacterial phyla comprising, on average, greater than 1% of the bacterial communities in the dataset and each of the DOM parameters. To investigate how shifts in the abundance of operational taxonomic units (OTUs) *within* each abundant phyla (>1%) were related to shifts in the DOM chemistry, Bray-Curtis distance matrices of the OTU-abundance tables within each phyla (i.e., Chloroflexi OTUs) were compared to distance matrices of PARAFAC model loadings within each chemical classification (i.e., Oxidized Quinones) using Mantel test in the vegan package of R (Figure 4.2; Oksanen et al. 2012).

## 4.3 Results

### 4.3.1 *DOM chemistry*

The concentration of DOC in groundwater ranged from 0.8 mg $L^{-1}$ to 5.5 mg $L^{-1}$ while TDN concentrations ranged from 0.2 mg $L^{-1}$ to 2.3 mg $L^{-1}$ (Table 4.1). The C:N ratio in groundwater varied between 1.0 at K60 11m and 5.0 at B 8m (Table 4.1). FI ranged from 1.05 at K150 11m to 1.60 at B 8m and 11m (Table 4.3). The FI at K150 11m was identified as an outlier value since it fell outside of two standard deviations from the mean for each variable. The percentage of amino acid- like fluorescence varied between 4% and 13% (Table 4.3). As shown in Table 4.4, the total range in PARAFAC model loadings was 0.006 – 2.05.




| | Operational Taxonomic Unit (OTU) | | | |
|---|---|---|---|---|
| **Sample** | **OTU.1** | **OTU.2** | **OTU.3** | **OTU.4** |
| **1** | 2 | 1 | 5 | 3 |
| **2** | 1 | 3 | 1 | 5 |
| **3** | 6 | 2 | 4 | 10 |
| **4** | 3 | 3 | 6 | 15 |
| **5** | 10 | 8 | 7 | 12 |

| | Oxidized Quinone PARAFAC model loadings | | |
|---|---|---|---|
| **Sample** | **Q1** | **Q2** | **Q3** |
| **1** | 0.175 | 0.456 | 0.339 |
| **2** | 0.211 | 0.516 | 0.330 |
| **3** | 0.254 | 0.640 | 0.442 |
| **4** | 0.123 | 0.375 | 0.229 |
| **5** | 0.139 | 0.498 | 0.965 |

| Sample | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **1** | 0 | 0.43 | 0.39 | 0.42 | 0.54 |
| **2** | 0.43 | 0 | 0.44 | 0.46 | 0.57 |
| **3** | 0.39 | 0.44 | 0 | 0.22 | 0.25 |
| **4** | 0.42 | 0.46 | 0.22 | 0 | 0.25 |
| **5** | 0.54 | 0.57 | 0.25 | 0.25 | 0 |

| Sample | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **1** | 0 | 0.07 | 0.23 | 0.15 | 0.63 |
| **2** | 0.07 | 0 | 0.17 | 0.19 | 0.64 |
| **3** | 0.23 | 0.17 | 0 | 0.36 | 0.55 |
| **4** | 0.15 | 0.19 | 0.36 | 0 | 0.75 |
| **5** | 0.63 | 0.64 | 0.55 | 0.75 | 0 |

Figure 4.2. A schematic describing the steps involved in performing a Mantel test evaluating the statistical relationship between the OTU composition of bacterial phyla and PARAFAC model loadings. First, Bray-Curtis distance matrices were calculated for OTU tables for each relatively abundant (>1%) phyla. Also, a Euclidean distance matrix was calculated for the PARAFAC model loadings data. Then, a Spearman rank-order correlation was performed to evaluate the relationship between the two distances matrices.

Table 4.3. Chemistry of dissolved organic matter (DOM) in groundwater collected at each sampling location in October 2009. The fluorescence index (FI) indicates whether the source of DOM is terrestrially- or microbially- derived. Amino acid (AA)- like fluorescence is a measure of the protein content, and the lability of the DOM pool for microbial assimilation. Data values followed by the pound sign (#) are outliers, and were removed from the dataset prior to performing statistical analyses.

| Sample | FI | AA-like fluorescence |
|---|---|---|
| **B 8m** | 1.6 | 5 |
| **B 11m** | 1.60 | 4 |
| **B 14m** | 1.53 | 4 |
| **K240 8m** | 1.56 | 9 |
| **K240 11m** | 1.52 | 6 |
| **K240 14m** | 1.49 | 6 |
| **K150 8m** | 1.42 | 3 |
| **K150 11m** | 1.05# | 9 |
| **K150 14m** | 1.44 | 4 |
| **K60 8m** | 1.48 | 13# |
| **K60 11m** | 1.46 | 4 |
| **K60 14m** | 1.43 | 5 |

Table 4.4. The PARAFAC model loadings data for groundwater collected in each of the sampling locations. Loadings values presented in this table have been rounded to the nearest hundredth decimal place. Data values followed by the pound sign (#) are outliers, and were removed from the dataset prior to performing statistical analyses.

| SITE | C1 | Q2 | C3 | HQ | SQ1 | C6 | SQ2 | TR | SQ3 | C10 | Q1 | Q3 | TY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **B 8m** | 0.24 | 0.46 | 0.19 | 0.45 | 0.09 | 0.11 | 0.14 | 0.07 | 0.16 | 0.08 | 0.18 | 0.34 | 0.06 |
| **B 11m** | 0.26 | 0.52 | 0.17 | 0.49 | 0.09 | 0.15 | 0.17 | 0.07 | 0.14 | 0.08 | 0.21 | 0.33 | 0.06 |
| **B 14m** | 0.34 | 0.64 | 0.22 | 0.61 | 0.12 | 0.17 | 0.18 | 0.08 | 0.17 | 0.10 | 0.25 | 0.44 | 0.06 |
| **K240 8m** | 0.02 | 0.04 | 0.03 | 0.06 | 0.01 | 0.01 | 0.01 | 0.03 | 0.02 | 0.01 | 0.01 | 0.04 | 0.00 |
| **K240 11m** | 0.04 | 0.08 | 0.04 | 0.10 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.02 | 0.04 | 0.07 | 0.00 |
| **K240 14m** | 0.07 | 0.14 | 0.07 | 0.20 | 0.04 | 0.02 | 0.04 | 0.03 | 0.05 | 0.03 | 0.05 | 0.11 | 0.02 |
| **K150 8m** | 0.17 | 0.38 | 0.09 | 0.46 | 0.10 | 0.13 | 0.10 | 0.02 | 0.08 | 0.06 | 0.12 | 0.23 | 0.03 |
| **K150 11m** | 0.07 | 0.50 | 0.26 | 2.05# | 0.22# | 0.01 | 0.05 | 0.47# | 0.22 | 0.09 | 0.14 | 0.97 | 0.00 |
| **K150 14m** | 0.08 | 0.16 | 0.06 | 0.21 | 0.05 | 0.05 | 0.05 | 0.01 | 0.05 | 0.04 | 0.07 | 0.11 | 0.03 |
| **K60 8m** | 0.24 | 0.62 | 0.17 | 0.57 | 0.12 | 0.14 | 0.13 | 0.20 | 0.10 | 0.10 | 0.22 | 0.27 | 0.19 # |
| **K60 11m** | 0.18 | 0.37 | 0.12 | 0.42 | 0.10 | 0.13 | 0.11 | 0.03 | 0.09 | 0.07 | 0.15 | 0.24 | 0.06 |
| **K60 14m** | 0.09 | 0.20 | 0.07 | 0.22 | 0.04 | 0.05 | 0.06 | 0.03 | 0.04 | 0.03 | 0.08 | 0.15 | 0.02 |

There was a positive linear relationship between oxidized quinone fluorescence and reduced quinone fluorescence ($R^2 \geq 0.91$, $p < 0.00006$) across study sites. Also, the intensities of the PARAFAC model loadings increased linearly with DOC and As concentrations, with the exception of the tryptophan- like fluorophore (Table 4.5). However, there was no discernable relationship between PARAFAC model loadings and groundwater Fe concentrations.

4.3.2 *Bacterial Community Characteristics*

There were 747 OTUs in the dataset after eliminating singletons, with an average of 183 OTUs per sample (standard deviation = 85). The alpha diversity of bacterial communities in aquifer sediments, as measured by the Phylogenetic Diversity metric (PD), ranged from 13 to 49 (Figure 4.3). Between 42% and 90% of the OTUs in aquifer sediments belonged to the Proteobacteria (mean = 70%, standard deviation = 19%; Table 4.6). Five classes within the Proteobacteria, including the Alphaproteobactria, Betaproteobacteria, Deltaproteobacteria, Epsilonproteobacteria, and Gammaproteobacteria, displayed average relative abundances greater than 1% across the aquifer sediment samples (Table 4.7). The Bacteriodetes, Chloroflexi and Firmicutes each comprised approximately 5% of the bacterial communities, and their relative abundances in samples varied by two orders of magnitude (Table 4.6). The proportions of Acidobacteria and Actinobacteria in the bacterial community were less variable, and these phyla each comprised 2% of bacterial communities sampled, on average (Table 4.6). On average, 65% of the Epsilonproteobacteria belonged to the Campylobacterales order, whereas Campylobacterales comprised 100% of the Epsilonproteobacteria OTUs in nine of the 12 aquifer sediment samples. The Dehalococcoidetes class constituted an average of 52 of the OTUs in the Chloroflexi phyla, and more than 80% of the OTUs in three aquifer sediment samples.

Table 4.5. The linear relationships observed in biplots (linear fit equation, $R^2$), and Spearman rank-order correlation results (Rho, *p*- value; shaded grey), between the intensity of the PARAFAC model loadings for each component (after removing outlier data points) and A) DOC or B) groundwater As concentrations.

| | **A) *Loadings vs. DOC concentration*** | | | | **B) *Loadings vs. As concentrations*** | | | |
|---|---|---|---|---|---|---|---|---|
| **Component** | **Equation** | **$R^2$** | **Rho** | ***p*-value** | **Equation** | **$R^2$** | **Rho** | ***p*-value** |
| C1 | 0.13x – 0.07 | 0.96 | 0.94 | 2E-05 | 1178x + 12.3 | 0.42 | 0.76 | 0.009 |
| Q2 | 0.27x – 0.13 | 0.85 | 0.94 | 2E-05 | 498x + 37.7 | 0.32 | 0.71 | 0.020 |
| C3 | 0.08x – 0.03 | 0.93 | 0.97 | 5E-07 | 1592x + 23.4 | 0.31 | 0.68 | 0.030 |
| HQ | 0.23x – 0.06 | 0.81 | 0.90 | 2E-04 | 546x + 13.1 | 0.31 | 0.74 | 0.010 |
| SQ1 | 0.05x – 0.01 | 0.73 | 0.79 | 4E-03 | 2692x + 13.4 | 0.34 | 0.78 | 0.007 |
| C6 | 0.07x – 0.04 | 0.83 | 0.85 | 8E-04 | 2005x + 27.7 | 0.43 | 0.86 | 0.002 |
| SQ2 | 0.08x – 0.04 | 0.95 | 0.96 | 2E-06 | 2050x + 8.7 | 0.41 | 0.73 | 0.020 |
| TR | 0.03x | 0.20 | 0.67 | 0.02 | 403x + 178.7 | 0.01 | 0.47 | 0.15 |
| SQ3 | 0.06x – 0.03 | 0.96 | 0.95 | 1E-05 | 1849x + 46.5 | 0.27 | 0.62 | 0.050 |
| C10 | 0.04x – 0.01 | 0.81 | 0.94 | 2E-05 | 3189x + 20.1 | 0.29 | 0.66 | 0.030 |
| Q1 | 0.11x – 0.05 | 0.90 | 0.93 | 4 E-05 | 1441x + 6.4 | 0.37 | 0.75 | 0.010 |
| Q3 | 0.16x – 0.07 | 0.98 | 0.97 | 5 E-07 | 943x + 1.3 | 0.40 | 0.68 | 0.030 |
| TY | 44.4x + 0.54 | 0.56 | 0.89 | 3E-04 | 5348x + 32.5 | 0.49 | 0.62 | 0.040 |

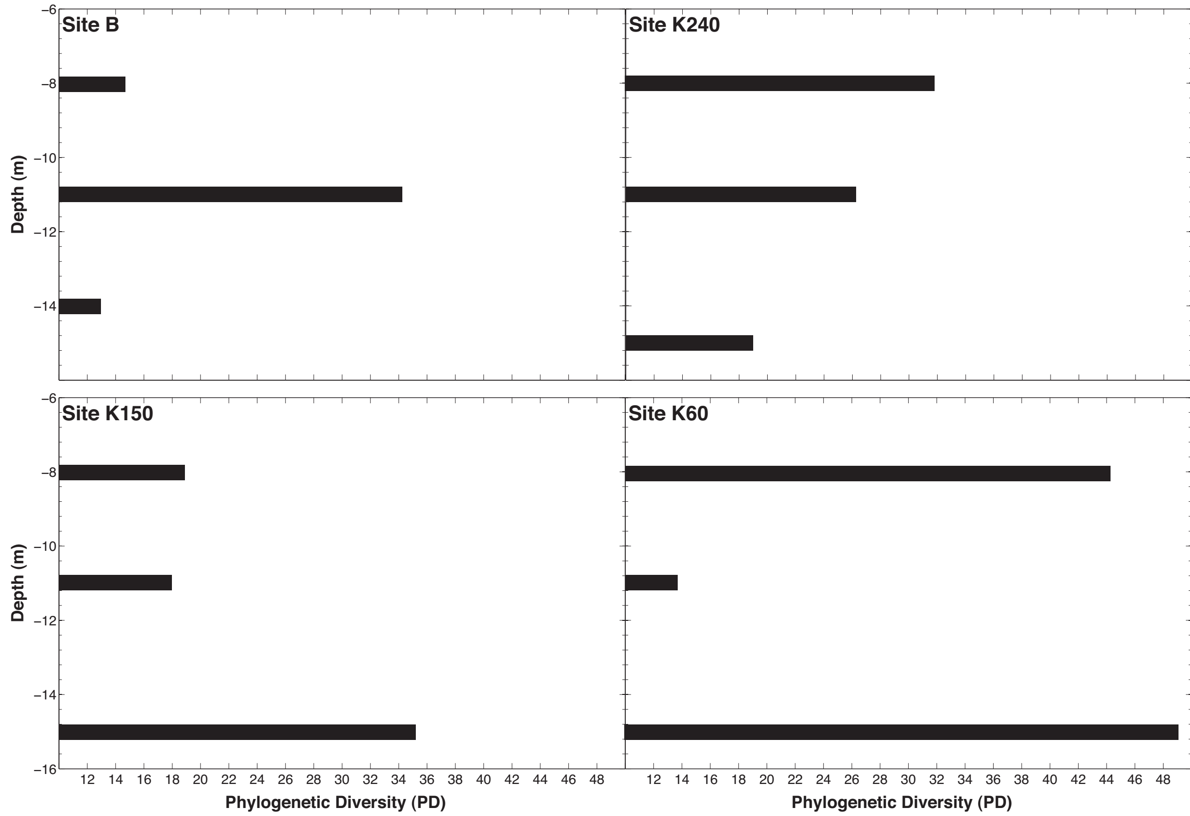



Figure 4.3. The phylogenetic diversity of bacterial communities in each aquifer sediment sample.

Table 4.6. The relative abundances (%) of phyla comprising greater than 1% of bacterial communities in aquifer sediment samples.

| SITE | Acido-bacteria | Actino-bacteria | Bacteriodetes | Chloroflexi | Firmicutes | Proteo-bacteria |
|---|---|---|---|---|---|---|
| **B 8m** | 0.5 | 0 | 2.5 | 1.2 | 1.2 | 91.4 |
| **B 11m** | 1.2 | 3.1 | 4.4 | 5.2 | 5.9 | 67.0 |
| **B 14m** | 1.5 | 5.3 | 3.9 | 12.6 | 6.3 | 53.4 |
| **K240 8m** | 1.7 | 3.6 | 16.8 | 0.5 | 20.6 | 46.2 |
| **K240 11m** | 1.6 | 2.5 | 2.4 | 0.9 | 9.7 | 79.1 |
| **K240 14m** | 1.2 | 0.9 | 5.6 | 0.6 | 4.3 | 85.3 |
| **K150 8m** | 0.5 | 1.8 | 5.4 | 0.4 | 0.4 | 89.7 |
| **K150 11m** | 0.7 | 0.5 | 4.3 | 5.4 | 0.6 | 80.4 |
| **K150 14m** | 6.6 | 3.8 | 5.2 | 1.5 | 1.2 | 73.5 |
| **K60 8m** | 5.4 | 0.62 | 4.9 | 12.8 | 4.6 | 42.2 |
| **K60 11m** | 1.2 | 0.37 | 0.6 | 0.6 | 0.9 | 89.9 |
| **K60 14m** | 2.9 | 0.20 | 2.5 | 12.0 | 6.7 | 44.3 |

Table 4.7. The relative abundances of classes within the Protoebacteria phyla that comprise, on average, greater than 1% of the bacterial communities in aquifer sediment samples.

| SITE | Alpha-proteobacteria | Beta-proteobacteria | Delta-proteobacteria | Epsilon-proteobacteria | Gamma-proteobacteria |
|---|---|---|---|---|---|
| **B 8m** | 0.2 | 26.4 | 0.3 | 3.9 | 60.7 |
| **B 11m** | 3.5 | 19.5 | 3.7 | 0.4 | 39.7 |
| **B 14m** | 4.4 | 18.0 | 7.3 | 1.0 | 21.8 |
| **K240 8m** | 3.0 | 15.4 | 1.3 | 0.6 | 25.8 |
| **K240 11m** | 2.5 | 59.6 | 1.7 | 0.2 | 15.1 |
| **K240 14m** | 0.5 | 75.3 | 2.0 | 0.4 | 7.0 |
| **K150 8m** | 1.1 | 71.5 | 0.2 | 0.1 | 16.9 |
| **K150 11m** | 0.2 | 54.8 | 1.8 | 20.2 | 3.1 |
| **K150 14m** | 3.8 | 64.4 | 1.8 | 0.1 | 3.3 |
| **K60 8m** | 1.3 | 27.9 | 5.0 | 0.7 | 6.9 |
| **K60 11m** | 1.6 | 86.0 | 1.0 | 0.5 | 0.9 |
| **K60 14m** | 4.8 | 26.1 | 6.2 | 0.1 | 6.6 |

An average of 61% of the Gammaproteobacteria were classified in the Pseudomonadales order, whereas the Alteromonadales and Aeromonadales comprised an average of 3% and 12% of the Gammaproteobacteria OTUs, respectively.

4.3.3 *DOM – Bacterial Relationships*

The significance of relationships between bacterial community diversity and composition was investigated in order to test the hypothesis that the chemistry and source of DOM shapes the composition of bacterial communities across groundwater As concentration gradients. Bacterial community alpha diversity in aquifer sediments was not significantly related to any of the DOM chemical parameters measured in the 12 groundwater samples ($p > 0.05$), however, bacterial community beta diversity was significantly related to multiple aspects of DOM chemistry. Spearman rank- order correlations between the axis 1 scores from the PCoA based on UniFrac distances and the C:N of groundwater demonstrated that bacterial community structure was significantly related to the stoichiometry of DOM in groundwater samples (Rho = 0.63, $p = 0.04$), but not the concentration of DOC or TDN ($p > 0.05$). Also, bacterial community structure was significantly related to the labile, tryptophan- like fluorophore (TR; $|Rho| = 0.62$, $p = 0.05$; Figure 4.4). Additional Spearman rank- order correlation tests identified statistically significant relationships between bacterial community structure and the loadings of three oxidized quinone ($0.73 \leq |Rho| \leq 0.78$, $p \leq 0.005$; Figure 4.5) and three reduced quinone ($0.74 \leq |Rho| \leq 0.78$, $p \leq 0.008$; Figure 4.6) fluorophores. Bacterial community structure was also significantly correlated with an unidentified microbially- associated fluorophore (C3; $|Rho| = 0.83$, $p \leq 0.001$; Figure 4.7), and an unidentified fluorophore found in a range of aquatic environments (C10; $|Rho| = 0.80$, $p \leq 0.002$; Figure 4.7).

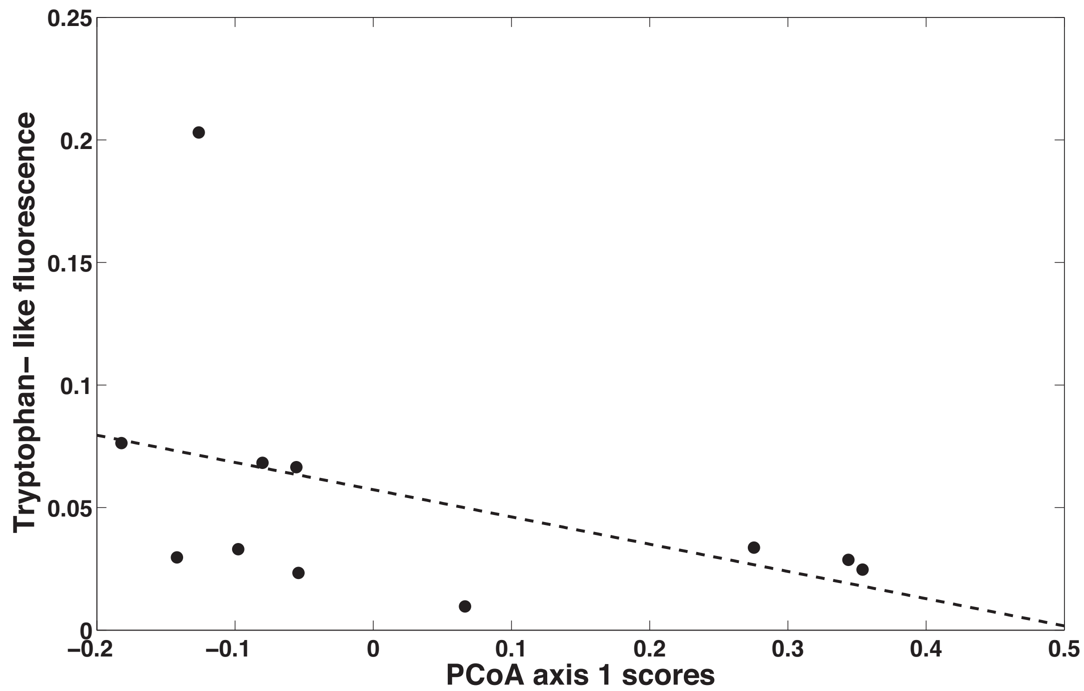



Figure 4.4. A Spearman rank-order correlation test showed that there was a statistically significant relationship between bacterial community structure (PCoA axis 1 scores) and tryptophan-like fluorescence (|Rho| = 0.62, $p = 0.05$).

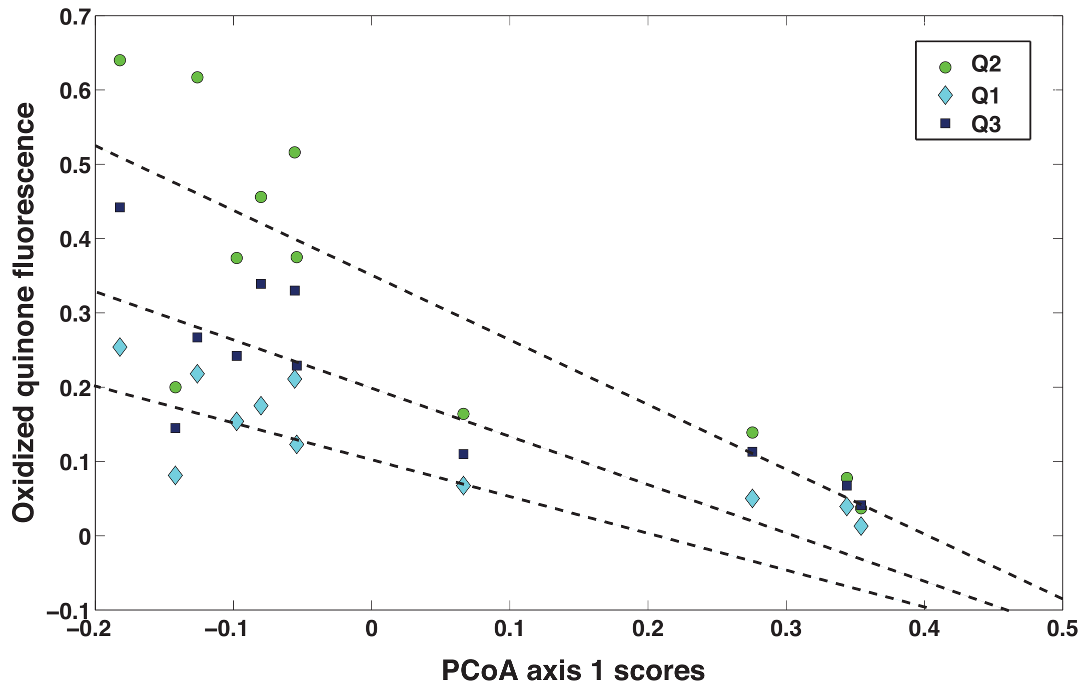



Figure 4.5. Spearman rank-order correlation tests showed that bacterial community structure (PCoA axis 1 scores) was significantly related to the three oxidized quinone fluorophores in the Cory & McKnight (2005) PARAFAC model ($0.73 \leq |\text{Rho}| \leq 0.78$, $p \leq 0.005$).

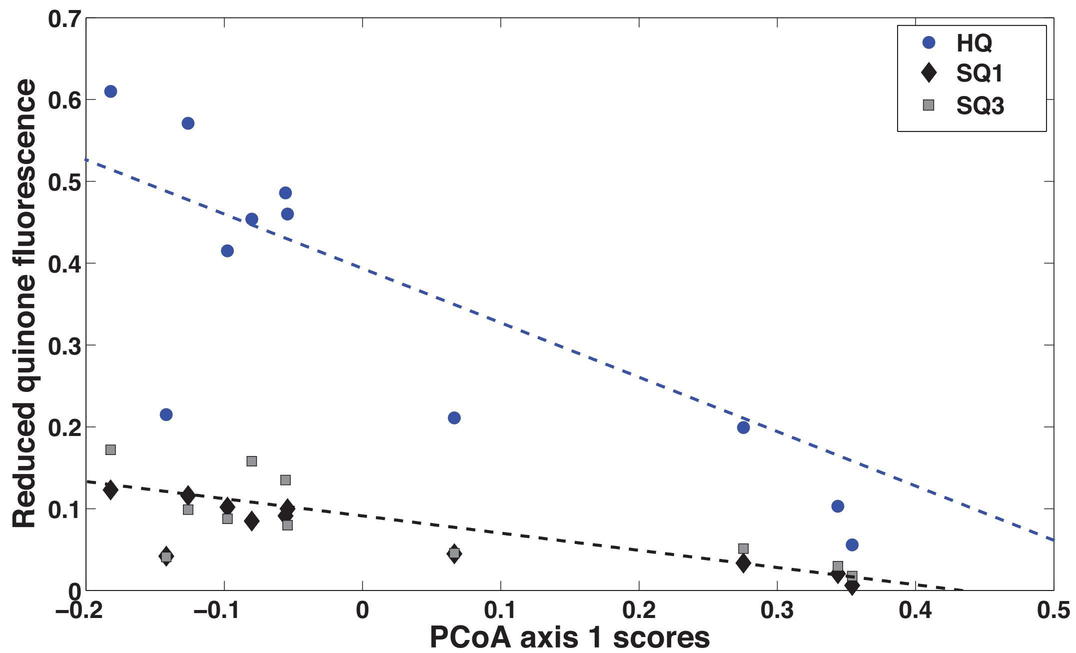



Figure 4.6. Spearman rank-order correlation tests showed that bacterial community structure (PCoA axis 1 scores) was significantly related to the three reduced quinone fluorophores in the Cory & McKnight (2005) PARAFAC model ($0.74 \leq |\text{Rho}| \leq 0.78$, $p \leq 0.008$).

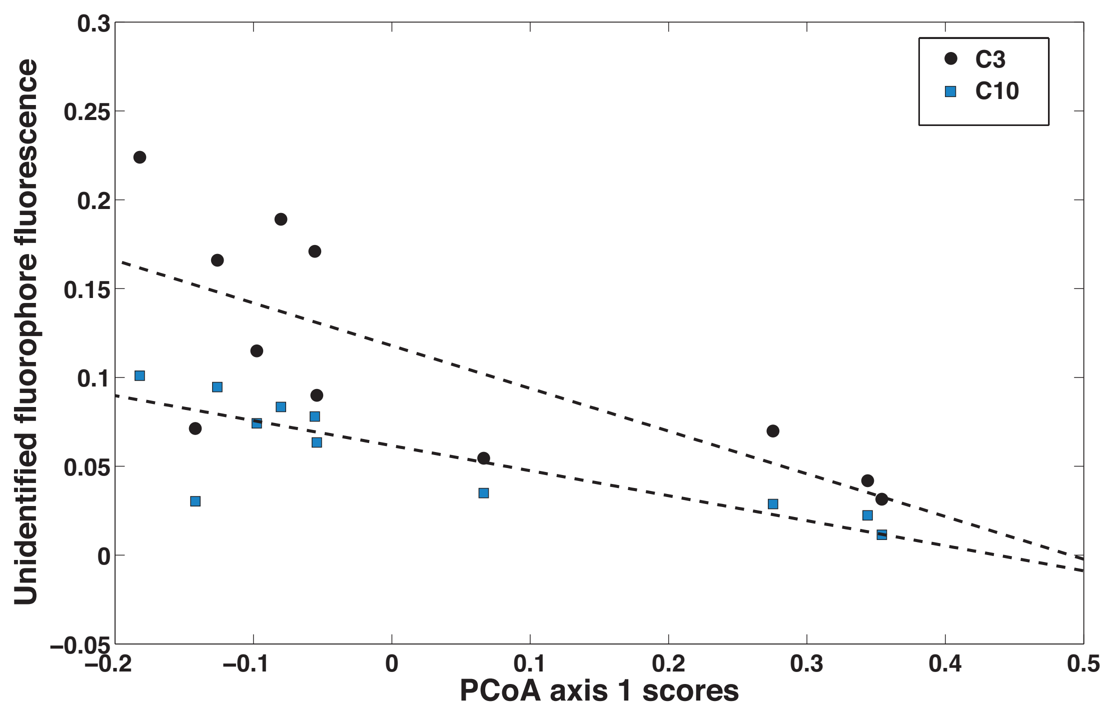



Figure 4.7. Spearman rank-order correlation tests showed that bacterial community structure (PCoA axis 1 scores) was significantly related to an unidentified microbially- associated fluorophore in the Cory & McKnight (2005) PARAFAC model (C3; |Rho| = 0.83, $p \leq 0.001$), and an unidentified fluorophore found in a range of aquatic environments (C10; |Rho| = 0.80, $p \leq 0.002$).

Additional statistical analyses revealed evidence that specific bacterial taxa were driving the observed relationships between bacterial community structure and DOM chemistry. The relative abundance of the Chloroflexi phyla was significantly correlated to the fluorescence of two oxidized quinone moieties (Q1 and Q2), a reduced quinone (HQ), a tryptophan-like moiety (TR), and an unidentified DOM fluorophore (C10) (Table 4.8). In addition, the relative abundance of Dehalococcoidetes, a class within Chloroflexi, was significantly correlated to all of the 13 PARAFAC model loadings (Table 4.8). The relative abundance of the Epsilonproteobacteria was significantly correlated to the loadings of microbially- associated fluorophores, including C3, Q3, SQ3, and TR (Table 4.8).

Mantel tests identified significant relationships between the composition of OTUs within a phyla and DOM chemistry. For example, the composition of Chloroflexi OTUs in aquifer sediment samples was significantly related to oxidized quinone ($r_s = 0.37$, $p = 0.02$), reduced quinone ($r_s = 0.33$, $p = 0.05$) and tryptophan- like ($r_s = 0.60$, $p = 0.05$) fluorescence. Also, Mantel tests detected significant relationships between the structure of OTUs in the Deltaproteobacteria and oxidized quinone ($r_s = 0.37$, $p = 0.02$), as well as reduced quinone ($r_s = 0.33$, $p = 0.03$), fluorescence. The OTU composition of the Bacteriodetes phyla was significantly related to oxidized ($r_s = 0.26$, $p = 0.03$) and reduced ($r_s = 0.23$, $p = 0.05$) quinone fluorescence, as well as the fluorescence of amino acid- like ($r_s = 0.33$, $p = 0.01$) and unidentified microbially- associated ($r_s = 0.28$, $p = 0.02$) moieties.

The stoichiometry of DOM, as described by the C:N, was significantly related to bacterial community structure, and Spearman rank- order correlation tests indicate that the relative abundances of Epsilonproteobacteria and Gammaproteobacteria are significantly correlated to this aspect of DOM chemistry (Table 4.9).

Table 4.8. The Rho and *p*- values are shown from Spearman rank- order correlations between the relative abundances of bacterial taxa in aquifer sediment samples and Cory and McKnight (2005) PARAFAC model loadings in groundwater samples. Correlation results yielding p-values greater than 0.05 are denoted with 'NS'.

| **Component** | **Chloroflexi** | **Chloroflexi class Dehalococcoidetes** | **Proteobacteria class Epsilonproteobacteria** |
|---|---|---|---|
| **C1** | 0.55; 0.07 (NS) | 0.74; 0.006 | 0.21; 0.51 (NS) |
| **Q2** | 0.67; 0.02 | 0.85; 0.0004 | 0.45; 0.15 (NS) |
| **C3** | 0.55; 0.06 (NS) | 0.75; 0.005 | 0.63; 0.03 |
| **HQ** | 0.61; 0.04 | 0.81; 0.001 | 0.46; 0.13 (NS) |
| **SQ1** | 0.52; 0.09 (NS) | 0.73; 0.007 | 0.46; 0.14 (NS) |
| **C6** | 0.43; 0.17 (NS) | 0.62; 0.03 | 0.01; 0.97 (NS) |
| **SQ2** | 0.50; 0.10 (NS) | 0.71, 0.01 | 0.22; 0.49 (NS) |
| **TR** | 0.64; 0.03 | 0.68; 0.01 | 0.83; 0.001 |
| **SQ3** | 0.48; 0.11 (NS) | 0.65; 0.02 | 0.64; 0.02 |
| **C10** | 0.66; 0.02 | 0.82; 0.001 | 0.57; 0.05 |
| **Q1** | 0.64; 0.03 | 0.83; 0.001 | 0.45; 0.15 (NS) |
| **Q3** | 0.55; 0.06 (NS) | 0.75; 0.005 | 0.63; 0.03 |
| **TY** | 0.43; 0.16 (NS) | 0.63, 0.03 | 0.21; 0.51 (NS) |

Table 4.9. The Rho and *p*- values are shown from Spearman rank- order correlations between the relative abundances of Epsilonproteobacteria and Gammaproteobacteria and the C:N of groundwater. Campylobacterales is a class within the Epsilonproteobacteria while Pseudomonadales, Aeromonadales, and Alteromonadales are classes within the Gammaproteobacteria.

| **DOM Variable** | **C:N** |
|---|---|
| **Epsilonproteobacteria** | 0.65; 0.03 |
| **Campylobacterales** | 0.73; 0.008 |
| **Gammaproteobacteria** | 0.67; 0.02 |
| **Pseudomonadales** | 0.59; 0.04 |
| **Aeromonadales** | 0.84; 0.0005 |
| **Alteromonadales** | 0.82; 0.001 |

By contrast, the source of DOM, as characterized by the FI, was not significantly related to bacterial community structure, as characterized by the UniFrac distance.

**4.4 Discussion**

This research provides support for my hypothesis that DOM chemistry significantly influences bacterial community structure across groundwater As gradients, and suggests that DOM is important to bacterial communities as carbon and nitrogen sources, as well as a source of electron donors and acceptors for respiration in aquifer environment. These analyses have demonstrated that the concentration of groundwater DOM was not a significant factor in structuring of bacterial communities, which was surprising given the relationship previously observed between bacterial community structure and sediment C concentrations at Site K (Legg et al., 2012). Instead, my results indicate that the availability of labile, tryptophan-like fluorophores and redox- active, quinone fluorophores was closely liked to the structure of bacterial communities across groundwater As gradients. Although previous research has shown that Deltaproteobacteria were the taxa most likely to mediate groundwater As mobilization (Islam et al., 2004; Héry et al., 2010; Lear et al., 2007; Rowland et al., 2007), my research suggests that Chloroflexi and Epsilonproteobacteria may also have important roles in the biogeochemical cycling of As through respiratory quinone- reduction. Also, it appears that Chloroflexi, Epsilonproteobacteria and Bacteroidetes were influenced by the availability of labile DOM. However, the source of DOM (plant vs. microbial) does not appear to be an important control on the overall composition of bacterial taxa in aquifer sediment communities, although it may be important in understanding the mechanism through which DOM comes into the aquifer.

Numerous studies have suggested that labile DOM fuels microbially-mediated groundwater As mobilization throughout Southeast Asia (Rowland et al., 2007; Pederick et al., 2007; Neumann et al., 2010; Lear et al., 2007; Islam et al., 2004; Héry et al., 2010; Harvey et al., 2002); yet, to date, there is little known about the chemical composition and biological availability of labile DOM in the high As aquifer environment (Fendorf et al., 2010). To characterize the chemical composition of labile DOM, I measured tryptophan- and tyrosine-like fluorophores in groundwater collected across As gradients in the aquifer. Since there is no evidence that tryptophan is an electron donor for relevant bacterial respiratory pathways in anaerobic groundwater sediments, such as dissimilatory Fe- or quinone-reduction, it is likely that tryptophan is a carbon and nitrogen source for bacteria in the high As aquifer environment. Amino acids represent one type of labile DOM which can be assimilated by phylogenetically diverse bacteria (Mayali et al., 2012), and previous research indicates that amino acid- like fluorescence provides a estimate of the biodegradability of DOM in the environment (Fellman et al., 2008). Although results from previous work at Site B found a decrease in amino acid- like fluorescence with increasing groundwater As concentrations (Mladenov et al., 2010), there was no discernable relationship between tryptophan- like fluorescence and groundwater As concentrations in this dataset. By contrast, tryptophan- like fluorescence was significantly related to bacterial community structure, suggesting that bacterial communities respond to labile DOM availability.

Bacterial taxa within the Dehalococcoidetes class and the Campylobacterales order may drive the observed relationship between bacterial community structure and tryptophan-like fluorescence since they comprised a significantly higher proportion of communities in groundwater with higher tryptophan- like fluorescence intensities. In addition, there was a

significant compositional shift at the OTU-level within the Bacteriodetes phyla in concert with increases in tryptophan-like fluorescence, suggesting that some Bacteriodetes taxa may utilize tryptophan-like constituents. There is a dearth of information available on the utilization of tryptophan by bacteria in the environment, however biochemical and genomic studies have shown that tryptophan is used by taxa that synthesize Nicotinamide Adenine Dinucleotide ($NAD^+$) from tryptophan using the NAD biosynthesis II pathway (Kurnasov et al., 2003). In fact, the genomes of taxa within the Dehalococcoidetes, Campylobacterales, and Bacteriodetes contain genes encoding the NAD biosynthesis II pathway (Caspi et al., 2012). Also, higher concentrations of tryptophan may alleviate the need for members within the Dehalococcoidetes to synthesize tryptophan using the tryptophan synthesis pathway (trp; Gutiérrez-Preciado et al., 2007; West et al., 2008), which is energetically expensive. It is interesting to note that the genomes of multiple Dehalococcoidetes and Campylobacterales contain the genes encoding the arsenate detoxification pathway (arsC; Uchiyama et al., 2012), which would confer an additional competitive advantage for taxa in groundwater with high As concentrations. Thus, this research indicates that a combination of 16S rRNA gene pyrosequencing and detailed DOM chemical characterization through EEM-PARAFAC can provide novel insights into how bacterial taxa may respond to labile DOM fractions in the high As environment.

Bacterial communities may also respond to the availability of quinone moieties. Based on previous research findings, quinone moieties may promote groundwater As mobilization when bacteria within the aquifer sediments reduce quinones through dissimilatory quinone- reduction, and the reduced quinones abiotically donating electrons to Fe(III) within Fe- oxide minerals, resulting in re-oxidized quinones and the release of sorbed As (Figure 4.8; Jiang & Kappler, 2008; Scott et al., 1998; Mladenov et al., 2010; Klapper et al., 2002; Lovley et al., 1996). In fact,

I found that groundwater with higher oxidized quinone fluorescence also had greater levels of reduced quinone fluorescence, indicating that quinone moieties are alternately cycling between oxidized and reduced states in the aquifer environment. Further support that quinone moieties serve as an intermediary between bacteria and As-bearing Fe- oxide minerals in aquifer sediments lies in the fact that there was no direct relationship between bacterial community structure and groundwater As concentrations, even though quinone fluorescence was significantly related to both bacterial community structure (Figure 4.5 and 4.6), and groundwater As concentrations (Table 4.5). Bacterial quinone reduction could considerably accelerate Fe-oxide mineral dissolution (Jiang & Kappler, 2008), and thus the desorption of As. As such, evidence of bacterial quinone reduction in this high As environment is very relevant for elucidating the mechanisms causing geogenic arsenic contamination throughout Southeast Asia.

Previous laboratory experiments have found that Deltaproteobacteria, and specifically *Geobacter* spp., are the predominate taxa promoting groundwater As mobilization in Southeast Asian aquifers through the reductive dissolution of Fe-oxide minerals (Islam et al., 2004; Héry et al., 2010; Rowland et al., 2007). Although bacterial quinone-reduction appears to be phylogenetically widespread in anaerobic environments (Cervantes et al., 2002), our understanding of this metabolism is primarily based on results from incubation experiments (Rowland et al., 2007) and laboratory experiments with cultured members in the *Geobacter* and *Shewanella* (Gammaproteobacteria) genera (Scott et al., 1998; Newman & Kolter 2000; Lovley et al., 1996; Jiang & Kappler, 2008). In this study I observed OTU-level shifts in the composition of Deltaproteobacteria with increases in quinone fluorescence, however there was not an increase in *Geobacteraceae* OTUs concurrent with a greater availability of quinone fluorescence. Instead, OTUs throughout the Deltaproteobacteria phylogeny appeared to respond to quinone

fluorescence. The incongruence between my research findings and results from laboratory experiments could be due to a variety of factors, including that the groundwater at my study sites feature markedly lower DOM concentrations (1-2 orders of magnitude) than the DOM concentrations used in laboratory experiments (Islam et al., 2004; Héry et al., 2010; Rowland et al., 2007). Thus, a wider diversity of Deltaproteobacteria, which are competitive under the oligotrophic conditions typical of the aquifer, may be involved in the biogeochemical cycling of As in the environment.

This research has revealed support for novel roles for Chloroflexi and Epsilonproteobacteria, taxa that appear to be competitive for labile and quinone-containing DOM fractions under oligotrophic conditions, in groundwater As mobilization. The relative abundances of Chloroflexi and Epsilonproteobacteria were significantly related to the availability of oxidized and reduced quinone fluorophores in groundwater, in addition to tryptophan-like fluorescence (Table 4.8). Taxa in the Dehalococcoidetes are primarily driving the relationships between quinone fluorescence and the abundance and OTU- level composition of Chloroflexi. In high As aquifer sediments, Dehalococcoidetes may be quinone-reducing generalists, given that their relative abundance is significantly correlated to the loadings of each of the oxidized and reduced fluorophores in the Cory & McKnight (2005) PARAFAC model (Table 4.8). Respiratory quinone-reduction has not been previously characterized within the Dehalococcoidetes class, whose members are primarily known for respiring halogenated organic contaminants, though previous research indicates that halorespiring Firmicutes perform quinone-reduction (Cervantes et al. 2002). More recently, researchers discovered that the addition of a model quinone compound, anthraquinone-2,6-disulfonate (AQDS), stimulated the degradation of TCE by a *Dehalococcoides spp*.- containing enrichment culture, although the mechanism has not been

definitely characterized (Zhang et al., 2011). Hence, additional research is needed to elucidate the role of Dehalococcoidetes as respiratory quinone-reducing generalist taxa in high As aquifers.

By contrast, Epsilonproteobacteria may be quinone-reducing specialists. The relative abundance of Epsilonproteobacteria is significantly related to the availability of one microbially-associated oxidized quinone moiety (Q3), and the fluorescence of a microbially-associated reduced quinone (SQ3). Epsilonproteobacteria could specialize in reducing microbially-derived quinone moieties during dissimilatory quinone-reduction. This metabolism has not been previously identified for representatives within the Epsilonproteobacteria, although there are dissimilatory Fe- and As-reducing Campylobacterales that would be well-adapted to the neutral pH, high As aquifers of Southeast Asia (Stolz et al., 1999), and other known Fe-reducers can also reduce quinone moieties (Coates et al., 1998; Straub et al., 2005; Lovley & Coates, 2000; Lovley et al., 1998). An alternative hypothesis is that Epsilonproteobacteria are producing the Q3 and SQ3 fluorophores. Laboratory experiments have demonstrated that bacteria produce quinones (Newman & Kolter, 2000), and numerous environmental studies have documented the bacterial production of humic substances, which contain quinone moieties (Yamashita & Tanoue, 2004; Stedmon & Markager, 2005; Shimotori et al., 2009). Nevertheless, these results introduce the potential importance of Epsilonproteobacteria in mediating the reductive dissolution of Fe-oxide minerals, and groundwater As release, through the respiration or production of quinone moieties.

This research has identified novel relationships between bacterial community structure and DOM chemistry. Bacterial community structure is significantly influenced by the availability of labile DOM and quinone moieties, as well as the intensity of two unidentified fluorophores (C3 and C10; Figure 4.7, Table 4.8). The identification of these unidentified fluorophores, and

additional research on the interactions between chemically- distinct DOM fractions and bacterial taxa in the environment, could refine our understanding of how diverse bacterial communities mediate the groundwater As cycling. My research suggests that it is the chemistry, but not the source of DOM (plant vs. microbial, surface water vs. sedimentary), that influences bacterially-mediated groundwater As mobilization. Still, identifying the source of DOM could lead to designing effective strategies for ceasing the supply of these labile and quinone fluorophores to bacterial communities, and potentially decreasing rates of groundwater As in the environment. It is clear that investigating the biogeochemistry of groundwater As cycling under environmental conditions is essential to progressing our understanding of the factors promoting geogenic arsenic contamination in Southeast Asia.

# CHAPTER 5: Conclusions

## 5.1 Research Summary

In this dissertation I presented research that refined our understanding of how sediment geochemistry and groundwater dissolved organic matter (DOM) chemistry affect groundwater arsenic (As) mobilization through their influence on the composition of native bacterial communities in aquifer sediments. This work built upon extensive research indicating that native bacteria promote groundwater As mobilization by pairing the oxidation of labile DOM to the reductive dissolution of iron (Fe)-oxide minerals in high As aquifers throughout Southeast Asia (Dowling et al., 2002; Harvey et al., 2002; Héry et al., 2010; Islam et al., 2004; McArthur et al., 2004; Neumann et al., 2010; Nickson et al., 1998; Radloff et al., 2008; Rowland et al., 2007; and others). More recently, research indicated that bacterial quinone-reduction, and the resulting rapid electron transfer between reduced quinones and Fe-oxide minerals (Jiang & Kappler, 2008), may be occurring in high As zones of aquifers (Mladenov et al., 2010). Despite the broad scope of literature available on the factors promoting geogenic As contamination, many aspects of the biogeochemical cycling of groundwater As are still not well-understood. My dissertation research contributes new information suggesting that both sediment geochemistry and groundwater DOM chemistry influence the composition, and potentially also the activity, of bacterial communities, although the availability of specific DOM fractions may have the greatest influence on bacterial community structure and rates of bacterially-mediated groundwater As mobilization.

Previous research on the microbial ecology of high As aquifers in Southeast Asia has discovered an unexpectedly high degree of taxonomic diversity within bacterial communities from single location (Islam et al., 2004), and between bacterial communities from different

depths at a single location (Gault et al., 2005; Héry et al., 2010; Pederick et al., 2007; Rowland et al., 2007). The results presented in Chapter 2 suggest that characterizing bacterial communities from different locations within a high As aquifer is key to understanding links between the biological and geochemical processes that promote groundwater As mobilization. At my study site in Araihazar, Bangladesh, 'Site K', there were significant differences in the structure of bacterial communities between locations along the groundwater flowpath, yet it appeared that depth gradients did not have a sizeable effect on the composition of bacterial communities. Hence, my dissertation research has demonstrated the importance of designing sampling schemes that capture the heterogeneity observed across small spatial gradients in aquifers (Fendorf et al., 2010; van Geen et al., 2003; van Geen, Zheng, et al., 2006a) in order to understand the role of bacterial communities in groundwater As mobilization.

Additionally, results presented in Chapter 2 suggest that Deltaproteobacteria and Chloroflexi may be involved in groundwater As mobilization through respiratory Fe-, Mn- or quinone-reduction. This conclusion was based on the finding that Deltaprotebacteria and Chloroflexi were more abundant in fine-grained sediments with higher concentrations of carbon (C), manganese (Mn) and Fe, and previous research has identified a positive relationship between these geochemical characteristics and groundwater concentrations (McArthur et al., 2004; van Geen, et al., 2006a; van Geen, et al., 2006b). It was not surprising to discover that Deltaproteobacteria may mediate groundwater As mobilization in silty sediments since results from multiple incubation studies have highlighted the Fe-oxide reducing activity of Deltaproteobacteria in the *Geobacter* genus (Héry et al., 2010; Islam et al., 2004; Rowland et al., 2007). Bacteria in the *Geobacter* genus are well-known for their ability to respire using a variety of electron acceptors, including oxidized metals like Fe(III), Mn(IV) and uranium (VI) (Lovley et al., 1993), and

quinones within humic substances (Cervantes et al., 2003; Coates et al., 1998; Lovley et al., 1996) which are generally present in silty sediments (Guggenberger et al.,, 1994). By contrast, there has not been any prior mention of the involvement of Chloroflexi, or taxa within the Dehalococcoidetes, in mediating groundwater As mobilization in Southeast Asia. Dehalococcoidetes are primarily known for reductively dehalogenating organic contaminants such as trichloroethylene (TCE) and tetrachloroethene (PCE) (He et al., 2012; Maymó-Gatell et al., 2001), but there is evidence that other halorespiring bacteria may respire using alternative electron acceptors such as Fe(III), Mn(IV), and As(V) (Luijten et al., 2004). In addition, previous research has found that genes encoding the As detoxification pathway (*arsC*) have been detected within the genomes of *Dehalococcoides spp.* and *Geobacter spp.* (Kube et al., 2005; Methé et al., 2003; Seshadri et al., 2005), therefore indicating that these taxa may be able to tolerate the high As concentrations observed in the deltaic aquifers in Southeast Asia.

The research presented in Chapters 3 and 4 expanded upon the potential roles for Deltaproteobacteria and Chloroflexi in the biogeochemical cycling of groundwater As. In Chapter 3, I present results showing that the addition of both labile and redox- active DOM promoted significantly higher relative abundances of Deltaproteobacteria closely related to *Geobacter spp.*, and corresponding increases in the flux of Fe and As from aquifer sediments to groundwater. In addition, the relative abundance of Deltaproteobacteria was significantly correlated with the availability of a reduced quinone fluorophore in control incubations, as well as in incubations amended with labile DOM. Chloroflexi comprised a considerably lower proportion of the post-incubation bacterial communities than Deltaproteobacteria, yet their relative abundance was positively correlated to the fluorescence of a reduced quinone moiety and a humic substance moiety in post-incubation samples. Likewise, the relative abundance of

Chloroflexi was significantly related to the availability of oxidized quinone and reduced quinone fluorophores in environmental samples (Chapter 4). These results strongly support the hypothesis that Deltaproteobacteria and Chloroflexi accelerate As mobilization through respiratory quinone reduction in deltaic aquifers throughout Southeast Asia (Jiang & Kappler, 2008; Rowland et al., 2007).

In comparing the results from Chapters 3 and 4, it is clear that the relative abundances of Chloroflexi and Deltaproteobacteria are sensitive to groundwater DOM concentrations. In incubation experiments Deltaproteobacteria are generally the dominant taxa; however, they comprised a significantly smaller fraction of bacterial communities in environmental samples (Chapters 3 and 4; Héry et al., 2010; Islam et al., 2004; Rowland et al., 2007). On the other hand, Chloroflexi appear to be oligotrophic taxa that are out-competed in incubation experiments with higher DOM concentrations. The research presented in Chapter 4 suggests that in the aquifer environment Chloroflexi may be more involved in accelerating groundwater As mobilization through respiratory quinone- reduction than Deltaproteobacteria. These results show that comparing experimental results with results from environmental sampling efforts can lead to a more accurate depiction of how bacterial taxa influence groundwater As mobilization in aquifers affected by geogenic As contamination.

## 5.2 Future Research Directions

This dissertation research provided new insights into the potential for bacterial quinone reduction to accelerate groundwater As mobilization in aquifers throughout Southeast Asia. In addition, this work identified novel relationships between Chloroflexi, Epsilonproteobacteria and Deltaproteobacteria and the availability of specific DOM fractions. To build upon the findings presented in this dissertation, the priorities for future research on geogenic As contamination in

Southeast Asia include: 1) identifying the source of quinone moieties in high As aquifers; 2) performing a wider range of chemical analyses to characterize the abundance and distribution of non-fluorescent labile and redox-active DOM constituents; and 3) further investigating the metabolic activities of Chloroflexi, specifically Dehalococcoidetes, as well as Epsilonproteobacteria and Deltaproteobacteria, under environmental conditions in high As aquifers.

Previous research suggested that quinone moieties in high As groundwater originate from DOM released during the reductive dissolution of Fe-oxide minerals in aquifer sediments (Mladenov et al., 2010). However, it is also possible that DOM-rich local recharge from ponds, rivers and seasonal monsoonal floods could contain quinone-containing humic substances. Also, bacteria in aquatic environments produce quinone-like and humic substance-like moieties (Ebert, 2012; Guillemette & del Giorgio, 2012; Stedmon & Markager, 2005), and thus could represent a significant source of electron acceptors for respiratory quinone reduction in the high As aquifer environment. The first step in identifying the source(s) of quinones to high As groundwater would be a more comprehensive survey of the quinone fluorescence in recharge sources, aquifer sediments, and groundwater across As concentration gradients using fluorescence spectroscopy and the Cory and McKnight (2005) PARAFAC model. In addition, the fluorescence index (FI; McKnight et al. 2001) could be used to define the signature (plant-derived vs. microbially-derived) of DOM within ponds and rivers, sediment leachates, and groundwater collected from numerous locations throughout areas impacted by geogenic arsenic contamination.

Employing a wider range of chemical analyses to measure concentrations of non-fluorescent DOM constituents in high As groundwater could improve our understanding of the range of constituents available to bacterial taxa in the aquifer environment. I did not discover any

significant relationships between the relative abundances of bacterial taxa and the fluorescence of tryptophan- or tyrosine-like moieties (Chapter 4), the labile constituents captured by fluorescence spectroscopy and the Cory & McKnight (2005) PARAFAC model. However, bacterial communities rely on labile DOM for sources of carbon and nitrogen to synthesize biomass, and for electron donors in anaerobic respiratory pathways such as Fe- and quinone-reduction. Thus techniques such as high performance liquid chromotagraphy (HPLC) could be used to measure the concentrations of other labile moieties including organic acids. In addition, HPLC could be used to measure the concentrations of non-fluorescence redox-active moieties such as phenazines (Fernández & Pizarro, 1997). Phenazines behave similarly to quinones in the environment, and they are produced by multiple bacteria taxa including *Pseudomonas spp.* (Pierson & Pierson, 2010). Bacteria can utilize phenazines as electron acceptors for respiration, and bacterially-reduced phenazines transfer electrons to, and thus abiotically reduce, metal oxide minerals (Hernandez et al., 2004). Additional information on the DOM sources available to bacterial communities in aquifer sediments could help researchers design experiments that more closely reflect the conditions in the high As aquifer environment.

My dissertation research has demonstrated that future experiments targeted towards clarifying the respiratory pathways used by Dehalococcoidetes, Epsilonproteobacteria and Deltaproteobacteria in aquifer sediments should utilize DOM concentrations which are reflective of conditions in the high As aquifer environment. Experimental evidence confirming that Dehalococcoidetes and Epsilonproteobacteria support growth through respiratory quinone-reduction is essential to understanding the correlations I observed between the relative abundances of these taxa and quinone fluorescence (Chapters 3 and 4). Experiments comparing the rates of respiratory quinone-reduction between Chloroflexi, Epsilonproteobacteria and

Deltaproteobacteria under environmental DOM concentrations could elucidate their relative impact on the mobilization of As in aquifers affected by geogenic As contamination in Southeast Asia.

## REFERENCES


Ahmed, K. M., Bhattacharya, P., Hasan, M. A., Akhter, S. H., Alam, S. M. M., Bhuyian, M. A. H., … Sracek, O. 2004. Arsenic enrichment in groundwater of the alluvial aquifers in Bangladesh: an overview. *Applied Geochemistry 19*(2): 181–200.

Amon, R. M. W., and R. Benner. 1996. Bacterial Utilization of Different Size Classes of Dissolved Organic Matter. *Limnology and Oceanography* 41(1): 41–51.

Anawar, H. M., Akai, J., Mostofa, K. M. G., Safiullah, S., & Tareq, S. M. (2002). Arsenic poisoning in groundwater: Health risk and geochemical sources in Bangladesh. *Environment International* 27(7): 597–604.

Arnosti, C. 2002. *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. Edited by S. E. G. Findlay and R. L. Sinsabaugh. Academic Press.

Azam, F., and F. Malfatti. 2007. Microbial Structuring of Marine Ecosystems. *Nature Reviews Microbiology* 5(10): 782–791.

Bauer, M., and C. Blodau. 2006. Mobilization of Arsenic by Dissolved Organic Matter from Iron Oxides, Soils and Sediments. *Science of The Total Environment* 354(2–3): 179–190.

Beggs, K. M. H., R. S. Summers, and D. M. McKnight. 2009. Characterizing Chlorine Oxidation of Dissolved Organic Matter and Disinfection By-product Formation with Fluorescence Spectroscopy and Parallel Factor Analysis. *Journal of Geophysical Research: Biogeosciences* 114(G4): doi:10.1029/2009JG001009.

Berg, M., Stengel, C., Trang, P. T. K., Hung Viet, P., Sampson, M. L., Leng, M., … Fredericks, D. 2007. Magnitude of arsenic pollution in the Mekong and Red River Deltas — Cambodia and Vietnam. *Science of The Total Environment* 372(2–3): 413–425.

Berggren, M., H. Laudon, M. Haei, L. Ström, and M. Jansson. 2010. Efficient Aquatic Bacterial Metabolism of Dissolved Low-molecular-weight Compounds from Terrestrial Sources. *The ISME Journal* 4(3): 408–416.

Bernhardt, E. S., and G. E. Likens. 2002. Dissolved Organic Carbon Enrichment Alters Nitrogen Dynamics in a Forest Stream. *Ecology* 83(6): 1689–1700.

Burgess, W. G., Burren, M., Perrin, J., & Ahmed, K. M. 2002. Constraints on sustainable development of arsenic-bearing aquifers in southern Bangladesh. Part 1: A conceptual model of arsenic in the aquifer. *Geological Society, London, Special Publications* 193(1), 145–163.

Cammack, W. K. L., J. Kalff, Y. T. Prairie, and E. M. Smith. 2004. Fluorescent Dissolved Organic Matter in Lakes: Relationships with Heterotrophic Metabolism. *Limnology and Oceanography*: 2034–2045.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., et al. 2010a. QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.* 7: 335-336.

Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., and Knight, R. 2010b. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26: 266 -267.

Carter, M. R. 1993. *Soil Sampling and Methods of Analysis*. Boca Raton, FL: Lewis Publishers.

Caspi, R., T. Altman, K. Dreher, C. A Fulcher, P. Subhraveti, I. M. Keseler, A. Kothari, et al. 2012. The MetaCyc Database of Metabolic Pathways and Enzymes and the BioCyc Collection of Pathway/genome Databases. *Nucleic Acids Research* 40(Database issue): D742–753.

Cawley, K. M., K. D. Butler, G. R. Aiken, L. G. Larsen, T. G. Huntington, and D. M. McKnight. 2012. Identifying Fluorescent Pulp Mill Effluent in the Gulf of Maine and Its Watershed. *Marine Pollution Bulletin* 64(8): 1678–1687.

Cawley, K. M., P. Wolski, N. Mladenov, and R. Jaffé. 2012. Dissolved Organic Matter Biogeochemistry Along a Transect of the Okavango Delta, Botswana. *Wetlands* 32(3): 475–486.

Cervantes, F. J., Bok, F. A. M. de, Duong-Dac, T., Stams, A. J. M., Lettinga, G., & Field, J. A. 2002. Reduction of humic substances by halorespiring, sulphate-reducing and methanogenic microorganisms. *Environmental Microbiology* 4(1): 51–57.

Cervantes, F. J., Duong-Dac, T., Ivanova, A. E., Roest, K., Akkermans, A. D. L., Lettinga, G., & Field, J. A. 2003. Selective enrichment of Geobacter sulfurreducens from anaerobic granular sludge with quinones as terminal electron acceptors. *Biotechnology Letters* 25(1): 39–45.

Chakraborti, D., M. M., Rahman, K., Paul, U. K., Chowdhury, M. K., Sengupta, D., Lodh,… S. C., Mukherjee. 2002. Arsenic calamity in the Indian subcontinent: What lessons have been learned? *Talanta 58*(1): 3–22.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.* 11: 265-270.

Chapelle, F. 2001. *Ground-Water Microbiology and Geochemistry*. John Wiley & Sons.

Charlet, L., & Polya, D. A. 2006. Arsenic in shallow, reducing groundwaters in southern Asia: an environmental health disaster. *Elements* 2(2): 91–96.

Cleveland, C.C., Reed, S. C., and Townsend, A. R. 2006. Nutrient regulation of organic matter decomposition in a tropical rain forest. *Ecology* 87: 492-503.

Cleveland, C. C., D. R. Nemergut, S. K. Schmidt, and A. R. Townsend. 2007. Increases in Soil Respiration Following Labile Carbon Additions Linked to Rapid Shifts in Soil Microbial Community Composition. *Biogeochemistry* 82(3): 229–240.

Coates, J. D., Cole, K. A., Chakraborty, R., O'Connor, S. M., and Achenbach, L. A. 2002. Diversity and ubiquity of bacteria capable of utilizing humic substances as electron donors for anaerobic respiration. *Appl. Environ. Microbiol.* 68: 2445-2452.

Coates, J. D., Ellis, D. J., Blunt-Harris, E. L., Gaw, C. V., Roden, E. E., & Lovley, D. R. 1998. Recovery of Humic-Reducing Bacteria from a Diversity of Environments. *Applied and Environmental Microbiology* 64(4): 1504–1509.

Coffin, R. B. 1989. Bacterial Uptake of Dissolved Free and Combined Amino Acids in Estuarine Waters. *Limnology and Oceanography* 34(3): 531–542.

Cory, R. M., & McKnight, D. M. 2005. Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter. *Environmental science & technology* 39(21): 8142–8149.

Cullen, W.R., and K.J. Reimer. 1989. Arsenic speciation in the environment. *Chemical Reviews* 89(4): 713-764.

Cummings, D. E., Caccavo, Fendorf, S., and Rosenzweig, R. F. 1999. Arsenic mobilization by the dissimilatory Fe(III)-reducing bacterium *Shewanella alga* BrY. *Environ. Sci. Technol.* 33: 723-729.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72: 5069-5072.

Dhar, R. K., Zheng, Y., Saltikov, C. W., Radloff, K. A., Mailloux, B. J., Ahmed, K. M., and van Geen, A. 2011. Microbes enhance the mobility of arsenic in the Pleistocene aquifer sand from Bangladesh. *Environ. Sci. Technol.* 45: 2648-2654.

Dhar, R. K., Zheng, Y., Stute, M., van Geen, A., Cheng, Z., Shanewaz, M., Shamsudduha, M., Hoque, A., Rahman, M.W., and Ahmed, K.M. 2008. Temporal variability of groundwater chemistry in shallow and deep aquifers of Araihazar, Bangladesh. *J. Contam. Hydrol.* 99: 97-111.

Docherty, K. M., K. C. Young, P. A. Maurice, and S. D. Bridgham. 2006. Dissolved Organic Matter Concentration and Quality Influences Upon Structure and Function of Freshwater Microbial Communities. *Microbial Ecology* 52(3): 378–388.

Dowling, C. B., Poreda, R. J., Basu, A. R., Peters, S. L., & Aggarwal, P. K. 2002. Geochemical study of arsenic release mechanisms in the Bengal Basin groundwater. *Water Resources Research* 38(9): 12–1–12–18.

Ebert, J. L. 2012. *Advances in the understanding of the contributions of microbial dissolved organic matter to the fluorescence signature of natural waters using Parallel Factor Analysis* (Masters of Science). University of Colorado, Boulder, CO.

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

Esham, E. C., Ye, W., & Moran, M. A. 2000. Identification and characterization of humic substances-degrading bacterial isolates from an estuarine environment. *FEMS Microbiology Ecology* 34(2): 103–111.

Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61(1), 1–10.

Farrell, R. F., Matthes, S. A., and Mackie, A. J. 1980. *Simple, Low-Cost Method for the Dissolution of Metal and Mineral Samples in Plastic Pressure Vessels*. Washington, D.C.: Department of the Interior, Bureau of Mines.

Fawell, J. K., and R. Mascarenhas. 2011. Arsenic in Drinking-Water: Background Document for Development of WHO Guidelines for Drinking-water Quality. WHO GDWQ Background Document WHO/SDE/WSH/03.04/75/Rev/1. World Health Organization.

Fellman, J. B., D. V. D'Amore, E. Hood, and R. D. Boone. 2008. Fluorescence Characteristics and Biodegradability of Dissolved Organic Matter in Forest and Wetland Soils from Coastal Temperate Watersheds in Southeast Alaska. *Biogeochemistry* 88(2): 169–184.

Fendorf, S., H. A. Michael, & A. van Geen 2010. Spatial and temporal variations of groundwater arsenic in South and Southeast Asia. *Science* 328(5982): 1123–1127.

Fernández, R. O., & Pizarro, R. A. 1997. High-performance liquid chromatographic analysis of Pseudomonas aeruginosa phenazines. *Journal of Chromatography* 771(1–2): 99–104.

Fierer, N., Bradford, M. A., & Jackson, R. B. 2007. Toward an ecological classification of soil bacteria. *Ecology* 88(6): 1354–1364.

Fierer, N., Hamady, M., Lauber, C. L., and Knight, R. 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *P. Natl. Acad. Sci. USA* 105: 17994-17999.

Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A., and Cleveland, C. C. 2009. Global patterns in belowground communities. *Ecol. Lett.* 12: 1238–1249.

Findlay, S. E. G, R. L. Sinsabaugh, W. V. Sobczak, and M. Hoostal. 2003. Metabolic and Structural Response of Hyporheic Microbial Communities to Variations in Supply of Dissolved Organic Matter. *Limnology and Oceanography* 48(4): 1608–1617.

Foreman, C. M., and J. S. Covert. 2002. *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. Edited by S. E. G. Findlay and R. L. Sinsabaugh. Academic Press.

Foulquier, A., F. Mermillod-Blondin, F. Malard, and J. Gibert. 2011. Response of Sediment Biofilm to Increased Dissolved Organic Carbon Supply in Groundwater Artificially Recharged with Stormwater. *Journal of Soils and Sediments* 11(2): 382–393.

Gault, A. G., Islam, F. S., Polya, D. A., Charnock, J. M., Boothman, C., Chatterjee, D., & Lloyd, J. R. 2005. Microcosm depth profiles of arsenic release in a shallow aquifer, West Bengal. *Mineralogical Magazine 69*(5): 855–863.

Gotelli, N. J., & Ellison, A. M. 2004. *A Primer Of Ecological Statistics*. Sinauer Associates.

Grandy, A. S., Sinsabaugh, R. L., Neff, J. C., Stursova, M., and Zak, D. R. 2008. Nitrogen deposition effects on soil organic matter chemistry are linked to variation in enzymes, ecosystems and size fractions. *Biogeochemistry* 91: 37-49.

Guggenberger, G., Zech, W., Haumaier, L., and Christensen, B. T. 1995. Land-use effects on the composition of organic matter in particle-size separates of soils: II. CPMAS and solution 13C NMR analysis. *Eur. J. Soil Sci.* 46: 147-158.

Guillemette, F., & del Giorgio, P. A. 2012. Simultaneous consumption and production of fluorescent dissolved organic matter by lake bacterioplankton. *Environmental microbiology* 14(6): 1432–1443.

Gutiérrez-Preciado, A., C. Yanofsky, and E. Merino. 2007. Comparison of Tryptophan Biosynthetic Operon Regulation in Different Gram-positive Bacterial Species. *Trends in Genetics* 23(9): 422–427.

Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J., and Knight, R. 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat. Meth.* 5: 235-237.

Hansel, C. M., Fendorf, S., Jardine, P. M., and Francis, C. A. 2008. Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile. *Appl. Environ. Microbiol.* 74: 1620-1633.

Harvey, C. 2002. Technical Comment: Groundwater Flow in the Ganges Delta. *Science 296*: 1563A.

Harvey, C. F., C. H. Swartz, A. B. M. Badruzzaman, N. Keon-Blute, W. Yu, M. A. Ali, J. Jay, et al. 2002. Arsenic Mobility and Groundwater Extraction in Bangladesh. *Science* 298(5598): 1602–1606.

He, J., Ritalahti, K. M., Aiello, M. R., & Löffler, F. E. 2003. Complete Detoxification of Vinyl Chloride by an Anaerobic Enrichment Culture and Identification of the Reductively Dechlorinating Population as a Dehalococcoides Species. *Applied and Environmental Microbiology* 69(2): 996–1003.

Hernandez, M. E., Kappler, A., & Newman, D. K. 2004. Phenazines and Other Redox-Active Antibiotics Promote Microbial Mineral Reduction. *Applied and Environmental Microbiology* 70(2): 921–928.

Héry, M., Van Dongen, B. E., Gill, F., Mondal, D., Vaughan, D. J., Pancost, R. D., Polya, D. A., and Lloyd, J.R. 2010. Arsenic release and attenuation in low organic carbon aquifer sediments from West Bengal. *Geobiology* 8: 155–168.

Hood, E., D. M. McKnight, and M. W. Williams. 2003. Sources and Chemical Character of Dissolved Organic Carbon Across an Alpine/subalpine Ecotone, Green Lakes Valley, Colorado Front Range, United States. *Water Resources Research* 39(7): 1188.

Hopenhayn, C. 2006. Arsenic in Drinking Water: Impact on Human Health. *Elements* 2(2): 103–107.

Horneman, A., van Geen, A., Kent, D. V., Mathe, P. E., Zheng, Y., Dhar, R. K., … Ahmed, K. M. 2004. Decoupling of As and Fe release to Bangladesh groundwater under reducing conditions. Part I: Evidence from sediment profiles. *Geochimica et Cosmochimica Acta* 68(17), 3459–3473.

Hu, S. J., Van Bruggen, A. H. C., & Grünwald, N. J. 1999. Dynamics of bacterial populations in relation to carbon availability in a residue-amended soil. *Applied Soil Ecology* 13(1): 21–30.

Hutalle-Schmelzer, K. M. L., Zwirnmann, E., Krüger, A., & Grossart, H.-P. 2010. Enrichment and cultivation of pelagic bacteria from a humic lake using phenol and humic matter additions. *FEMS Microbiology Ecology* 72(1): 58–73.

Immerzeel, W. 2008. Historical trends and future predictions of climate variability in the Brahmaputra basin. *Int. J. Climatol.* 28: 243-254.

Inskeep, W.P., T.R. McDermott, and S.E. Fendorf. 2002 Arsenic (V/III) cycling in soils and natural waters: Chemical and microbiological processes. In W.T. Frankenberger, Jr. (Ed.) Environmental Chemistry of Arsenic. Springer-Verlag. p. 183-215.

Islam, F. S., Gault, A. G., Boothman, C., Polya, D. A., Charnock, J. M., Chatterjee, D., and Lloyd, J. R. 2004. Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* 430: 68-71.

Jackson, T. A., and R. E. Hecky. 1980. Depression of Primary Productivity by Humic Matter in Lake and Reservoir Waters of the Boreal Forest Zone. *Canadian Journal of Fisheries and Aquatic Sciences* 37(12): 2300–2317.

Jaffé, R., D. McKnight, N. Maie, R. Cory, W. H. McDowell, and J. L. Campbell. 2008. Spatial and Temporal Variations in DOM Composition in Ecosystems: The Importance of Long-term Monitoring of Optical Properties. *Journal of Geophysical Research: Biogeosciences* 113(G4): doi:10.1029/2008JG000683.

Janssen, P. H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA Genes. *Appl. Environ. Microbiol.* 72: 1719-1728.

Jiang, J. and Kappler, A. 2008. Kinetics of microbial and chemical reduction of humic substances: implications for electron shuttling. *Environ. Sci. Technol.* 42: 3563-3569.

Jiang, J., Bauer, I., Paul, A., and Kappler, A. 2009. Arsenic redox changes by microbially and chemically formed semiquinone radicals and hydroquinones in a humic substance model quinone. *Environ. Sci. Technol.* 43: 3639-3645.

Jones, R. T., Robeson, M. S., Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. 2009. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J.* 3: 442-453.

Jones, S. E., R. J. Newton, and K. D. McMahon. 2009. Evidence for Structuring of Bacterial Community Composition by Organic Carbon Source in Temperate Lakes. *Environmental Microbiology* 11(9): 2463–2472.

Judd, K. E., Byron C. C., and G. W. Kling. 2006. Variation in Dissolved Organic Matter Controls Bacterial Production and Community Composition. *Ecology* 87(8): 2068–2079.

Kalbitz, K, and R Wennrich. 1998. Mobilization of Heavy Metals and Arsenic in Polluted Wetland Soils and Its Dependence on Dissolved Organic Matter. *Science of The Total Environment* 209(1): 27–39.

Kappler, A., Benz, M., Schink, B., & Brune, A. 2004. Electron shuttling via humic acids in microbial iron(III) reduction in a freshwater sediment. *FEMS microbiology ecology* 47(1): 85–92.

Kawasaki, Nobu, and Ronald Benner. 2006. Bacterial Release of Dissolved Organic Matter During Cell Growth and Decline: Molecular Origin and Composition. *Limnology and Oceanography* 51(5): 2170–2180.

Keil, R. G., & Kirchman, D. L. 1991. Contribution of dissolved free amino acids and ammonium to the nitrogen requirements of heterotrophic bacterioplankton. *Mar. Ecol. Prog. Ser*, 73(1): 1–10.

Kirchman, D. L., A. I. Dittel, S. E. G Findlay, and D. Fischer. 2004. Changes in Bacterial Activity and Community Structure in Response to Dissolved Organic Matter in the Hudson River, New York. *Aquatic Microbial Ecology* 35(3): 243–257.

Kirk, M. F., Holm, T. R., Park, J., Jin, Q., Sanford, R. A., Fouke, B. W., and Bethke, C. M. 2004. Bacterial sulfate reduction limits natural arsenic contamination in groundwater. *Geology* 32: 953-956.

Klapper, L., D. M. McKnight, J. R. Fulton, E. L. Blunt-Harris, K. P. Nevin, D. R. Lovley, and P. G. Hatcher. 2002. Fulvic Acid Oxidation State Detection Using Fluorescence Spectroscopy. *Environmental Science & Technology* 36(14): 3170–3175.

Knappett, P. S. K., Escamilla, V., Layton, A., McKay, L. D., Emch, M., Williams, D. E., … van Geen, A. 2011. Impact of population and latrines on fecal contamination of ponds in rural Bangladesh. *Science of The Total Environment* 409(17): 3174–3182.

Koetsier, P., V. McArthur, and L. Leff. 1997. Spatial and Temporal Response of Stream Bacteria to Sources of Dissolved Organic Carbon in a Blackwater Stream System. *Freshwater Biology* 37(1): 79–89.

Krieg, N. R., Ludwig, W., Whitman, W. B., Hedlund, B. P., Paster, B. J., Staley, J. T., … Parte, A. (Eds.). 2011. Bergey's Manual of Systematic Bacteriology. In *Bergey's Manual of Systematic Bacteriology* (2nd ed., Vols. 1-5, Vol. 4). New York: Springer.

Kube, M., Beck, A., Zinder, S. H., Kuhl, H., Reinhardt, R., & Adrian, L. 2005. Genome sequence of the chlorinated compound-respiring bacterium Dehalococcoides species strain CBDB1. *Nature biotechnology* 23(10): 1269–1273.

Kurnasov, O., V. Goral, K. Colabroy, S. Gerdes, S. Anantha, A. Osterman, and T. P. Begley. 2003. NAD Biosynthesis: Identification of the Tryptophan to Quinolinate Pathway in Bacteria. *Chemistry & Biology* 10(12): 1195–1204.

Lear, G., Song, B., Gault, A. G., Polya, D. A., & Lloyd, J. R. 2007. Molecular Analysis of Arsenate-Reducing Bacteria within Cambodian Sediments following Amendment with Acetate. *Applied and Environmental Microbiology* 73(4): 1041–1048.

Legg, T. M., Y. Zheng, B. Simone, K. A. Radloff, N. Mladenov, A. González, D. Knights, et al. 2012. Carbon, Metals, and Grain Size Correlate with Bacterial Community Structure in Sediments of a High Arsenic Aquifer. *Frontiers in Microbiology* 3: doi:10.3389/fmicb.2012.00082.

Li, W., and Godzik, A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22: 1658 -1659.

Li, X., and Krumholz, L. R. 2007. Regulation of arsenate resistance in *Desulfovibrio desulfuricans* G20 by an *arsRBCC* operon and an *arsC* gene. *J. Bacteriol.* 189: 3705-3711.

Lin, X., S. Green, M. M. Tfaily, O. Prakash, K. T. Konstantinidis, J. E. Corbett, J. P. Chanton, W. T. Cooper, and J. E. Kostka. 2012. Microbial Community Structure and Activity Linked to Contrasting Biogeochemical Gradients in Bog and Fen Environments of the Glacial Lake Agassiz Peatland. *Applied and Environmental Microbiology* 78(19): 7023–7031.

Loffler, F. E., Sun, Q., Li, J., and Tiedje, J. M. 2000. 16S rRNA gene-based detection of tetrachloroethene-dechlorinating *Desulfuromonas* and *Dehalococcoides* species. *Appl. Environ. Microbiol.* 66: 1369-1374.

Lovley, D. R., and Phillips, E. J. P. 1988. Novel mode of microbial energy metabolism: organic carbon oxidation coupled to dissimilatory reduction of iron or manganese. *Appl. Environ. Microbiol.* 54: 1472-1480.

Lovley, D. R. 1991. Dissimilatory Fe(III) and Mn(IV) reduction. *Microbiological Reviews* 55(2): 259–287.

Lovley, D. R., Giovannoni, S. J., White, D. C., Champine, J. E., Phillips, E. J. P., Gorby, Y. A., & Goodwin, S. 1993. Geobacter metallireducens gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals. *Archives of Microbiology* 159(4): 336–344.

Lovley, D. R., Coates, J. D., Blunt-Harris, E. J. P., & Woodward, J. C. 1996. Humic substances as electron acceptors for microbial respiration. *Nature* 382: 445–448.

Lovley, D. R., Fraga, J. L., Blunt-Harris, E. L., Hayes, L. A., Phillips, E. J. P., & Coates, J. D. 1998. Humic substances as a mediator for microbially catalyzed metal reduction. *Acta hydrochimica et hydrobiologica* 26(3): 152–157.

Lovley, D., and J. Coates. 2000. Novel Forms of Anaerobic Respiration of Environmental Relevance. *Current Opinion in Microbiology* 3(3): 252–256.

Lovley, D. R, Holmes, D. E., and Nevin, K. P. 2004. Dissimilatory Fe(III) and Mn(IV) reduction. *Adv. Microb. Physiol.* 49: 219-286.

Lozupone, C., & Knight, R. 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* 71(12): 8228–8235.

Luijten, M. L. G. C., Weelink, S. A. B., Godschalk, B., Langenhoff, A. A. M., Van Eekert, M. H. A., Schraa, G., & Stams, A. J. M. 2004. Anaerobic reduction and oxidation of quinone moieties and the reduction of oxidized metals by halorespiring and related organisms. *FEMS Microbiology Ecology* 49(1): 145–150.

Luna, G. M., Dell'Anno, A., Corinaldesi, C., Armeni, M., and Danovaro, R. 2009. Diversity and spatial distribution of metal-reducing bacterial assemblages in groundwaters of different redox conditions. *Int. Microbiol.* 12: 153-159.

Macy, J. M., Santini, J. M., Pauling, B. V., O'Neill, A. H., and Sly, L. I. 2000. Two new arsenate/sulfate-reducing bacteria: mechanisms of arsenate reduction. *Arch. Microbiol.* 173: 49-57.

Maness, A. D., Bowman, K. S., Yan, J., Rainey, F. A., & Moe, W. M. 2012. *Dehalogenimonas spp.* can Reductively Dehalogenate High Concentrations of 1,2-Dichloroethane, 1,2-Dichloropropane, and 1,1,2-Trichloroethane. *AMB Express 2*(54): doi:10.1186/2191-0855-2-54

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380.

Matejovic, I. 1997. Determination of carbon and nitrogen in samples of various soils by the dry combustion. *Communications in Soil Science and Plant Analysis* 28(17-18): 1499–1511.

Mayali, X., P. K Weber, E. L Brodie, S. Mabery, P. D Hoeprich, and J. Pett-Ridge. 2012. High-throughput Isotopic Analysis of RNA Microarrays to Quantify Microbial Resource Use. *The ISME Journal* 6(6): 1210–1221.

Maymó-Gatell, X., Nijenhuis, I., & Zinder, S. H. 2001. Reductive Dechlorination of cis-1,2-Dichloroethene and Vinyl Chloride by *Dehalococcoides ethenogenes*. *Environmental Science & Technology* 35(3): 516–521.

McArthur, J. M., Banerjee, D. M., Hudson-Edwards, K. A., Mishra, R., Purohit, R., Ravenscroft, P., Cronin, A., Howarth, R. J., Chatterjee, A., Talukder, T., Lowry, D., Houghton, S., and Chadha, D. K. 2004. Natural organic matter in sedimentary basins and its relation to arsenic in anoxic ground water: the example of West Bengal and its worldwide implications. *Appl. Geochem.* 19: 1255–1293.

McArthur, J. M., Ravenscroft, P., Safiulla, S., and Thirlwall, M. F. 2001. Arsenic in groundwater: Testing pollution mechanisms for sedimentary aquifers in Bangladesh. *Water Resour. Res.* 37: 109–117.

McCarren, J., J. W. Becker, D. J. Repeta, Y. Shi, C. R. Young, R. R. Malmstrom, S. W. Chisholm, and E. F. DeLong. 2010. Microbial Community Transcriptomes Reveal Microbes and Metabolic Pathways Associated with Dissolved Organic Matter Turnover in the Sea. *Proceedings of the National Academy of Sciences* 107(38): 16420–16427.

McKnight, D. M., Boyer, E. W., Westerhoff, P. K., Doran, P. T., Kulbe, T., & Andersen, D. T. 2001. Spectrofluorometric Characterization of Dissolved Organic Matter for Indication of Precursor Organic Material and Aromaticity. *Limnology and Oceanography* 46(1): 38–48.

Methé, B. A., Nelson, K. E., Eisen, J. A., Paulsen, I. T., Nelson, W., Heidelberg, J. F., … Fraser, C. M. 2003. Genome of Geobacter sulfurreducens: metal reduction in subsurface environments. *Science* 302(5652): 1967–1969.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., and Edwards, R. A. 2008. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9: 386-393.

Miller, M. P., D. M. McKnight, R. M. Cory, M. W. Williams, and R. L. Runkel. 2006. Hyporheic Exchange and Fulvic Acid Redox Reactions in an Alpine Stream/Wetland Ecosystem, Colorado Front Range. *Environmental Science & Technology* 40(19): 5943–5949.

Miller, M. P., D. M. McKnight, S. C. Chapra, and M. W. Williams. 2009. A Model of Degradation and Production of Three Pools of Dissolved Organic Matter in an Alpine Lake. *Limnology and Oceanography* 54(6): 2213–2227.

Mladenov, N., Huntsman-Mapila, P., Wolski, P., Masamba, W. R. L., & McKnight, D. M. 2008. Dissolved organic matter accumulation, reactivity, and redox state in ground water of a recharge wetland. *Wetlands* 28(3): 747–759.

Mladenov, N., Zheng, Y., Miller, M. P., Nemergut, D. R., Legg, T., Simone, B., Hageman, C., Rahman, M. M., Ahmed, K. M., and McKnight, D. M. 2010. Dissolved organic matter sources and consequences for iron and arsenic mobilization in Bangladesh aquifers. *Environ. Sci. Technol.* 44: 123-128.

Mou, X., S. Sun, R. A. Edwards, R. E. Hodson, and M. Moran. 2008. Bacterial Carbon Processing by Generalist Species in the Coastal Ocean. *Nature* 451(7179): 708–711.

Mukherjee, A. B., & Bhattacharya, P. 2001. Arsenic in groundwater in the Bengal Delta Plain: slow poisoning in Bangladesh. *Environmental Reviews* 9(3): 189–220.

Mukhopadhyay, R., Rosen, B. P., Phung, L. T., and Silver, S. 2002. Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiol. Rev.* 26: 311 - 325.

Mulholland, P. J. 2003. Large- scale patterns in dissolved organic carbon concentration, flux and sources. In *Interactivity of Dissolved Organic Matter*; Findlay, S. E. G., Sinsabaugh, R. L., Eds.; Academic Press: San Diego, CA: 139-159.

Murphy, K. R., A. Hambly, S. Singh, R. K. Henderson, A. Baker, R. Stuetz, and S. J. Khan. 2011. Organic Matter Fluorescence in Municipal Water Recycling Schemes: Toward a Unified PARAFAC Model. *Environmental Science & Technology* 45(7): 2909–2916.

Murray, K. S., Cauvet, D., Lybeer, M., and Thomas, J. C. 1999. Particle size and chemical control of heavy metals in bed sediment from the Rouge River, Southeast Michigan. *Environ. Sci. Technol.* 33: 987-992.

Nemergut, D. R., C. C. Cleveland, W. R. Wieder, C. L. Washenberger, and A. R. Townsend. 2010. Plot-scale Manipulations of Organic Matter Inputs to Soils Correlate with Shifts in Microbial Community Composition in a Lowland Tropical Rain Forest. *Soil Biology and Biochemistry* 42(12): 2153–2160.

Neumann, R. B., Ashfaque, K. N., Badruzzaman, A. B. M., Ashraf Ali, M., Shoemaker, J. K., and Harvey, C. F. 2010. Anthropogenic influences on groundwater arsenic concentrations in Bangladesh. *Nature Geosci.* 3: 46-52.

Newman, D K, and R Kolter. 2000. A Role for Excreted Quinones in Extracellular Electron Transfer. *Nature* 405(6782): 94–97.

Newman, D., Beveridge, T., and Morel, F. 1997. Precipitation of arsenic trisulfide by *Desulfotomaculum auripigmentum*. *Appl. Environ. Microbiol.* 63: 2022-2028.

Nickson, R., MacArthur, J., Burgess, W., Ahmed, K. M., & Ravenscroft, M. 1998. Arsenic poisoning of Bangladesh groundwater. *Nature* 395: 338.

Nordstrom, D. K. 2002. Worldwide Occurrences of Arsenic in Ground Water. *Science* 296(5576), 2143–2145.

O’Day, P.A. 2006. Chemistry and mineralogy of arsenic, *Elements* 2: 77-83.

Ogawa, H., Y. Amagai, I. Koike, K. Kaiser, and R. Benner. 2001. “Production of Refractory Dissolved Organic Matter by Bacteria.” *Science* 292 (5518): 917–920.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, Peter R. Minchin, R. B. O’Hara, G. L. Simpson, P. Solymos, M. H. Stevens, and H. Wagner. 2012. *Vegan: Community Ecology Package* (version R package version 2.0-4). http://CRAN.R-project.org/package=vegan.

Pace, N. R. 1997. A molecular view of microbial diversity and the biosphere. *Science* 276(5313): 734–740.

Palmer, N.E., Freudenthal, J.H., R. von Wandruszka. 2006. Reduction of arsenates by humic materials, *Environmental Chemistry* 3: 131-136.

Pautler, B. G., G. C. Woods, A. Dubnick, A. J. Simpson, M. J. Sharp, S. J. Fitzsimons, and M. J. Simpson. 2012. Molecular Characterization of Dissolved Organic Matter in Glacial Ice: Coupling Natural Abundance 1H NMR and Fluorescence Spectroscopy. *Environmental Science & Technology* 46(7): 3753–3761.

Pederick, R. L., Gault, A. G., Charnock, J. M., Polya, D. A., & Lloyd, J. R. 2007. Probing the biogeochemistry of arsenic: Response of two contrasting aquifer sediments from Cambodia to stimulation by arsenate and ferric iron. *Journal of Environmental Science and Health, Part A* 42(12): 1763–1774.

Pierson, L. S., & Pierson, E. A. 2010. Metabolism and function of phenazines in bacteria: impacts on the behavior of bacteria in the environment and biotechnological processes. *Applied Microbiology and Biotechnology* 86(6): 1659–1670.

Polizzotto, M. L., Harvey, C. F, Sutton, S. R., and Fendorf, S. 2005. Processes conducive to the release and transport of arsenic into aquifers of Bangladesh. *Proc. Nat. Acad. Sci.* 102: 18819-18823.

Polizzotto, M. L., Kocar, B. D., Benner, S. G., Sampson, M., & Fendorf, S. 2008. Near-surface wetland sediments as a source of arsenic release to ground water in Asia. *Nature 454*: 505–509.

Poppe, L. J., Eliason, A. H., Fredericks, J. J., Rendigs, R. R., Blackwood, D., and Polloni, C. 2000. *USGS east-coast sediment analysis: procedures, database, and georeferenced*

*displays* (U.S. Geological Survey Open-File Report No. 00-358). Reston, Virginia: U.S. Department of the Interior, U.S. Geological Survey.

Price, M. N., Dehal, P. S., & Arkin, A. P. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular biology and evolution* 26(7), 1641–1650.

Quicksall, A. N., Bostick, B. C., & Sampson, M. 2008. Linking organic matter deposition and iron mineral transformations to groundwater arsenic levels in the Mekong delta, Cambodia. *Applied Geochemistry* 23(11): 3088–3098.

R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Radloff, K. 2010. *Geochemical and hydrologic determinants of arsenic distribution in sedimentary aquifers in Bangladesh* (Doctor of Philosophy). Columbia University, New York.

Radloff, K. A., Manning, A. R., Mailloux, B., Zheng, Y., Moshiur Rahman, M., Rezaul Huq, M., Ahmed, K. M., van Geen, A. 2008. Considerations for conducting incubations to study the mechanisms of As release in reducing groundwater aquifers. *Appl. Geochem.* 23: 3224-3235.

Ravenscroft, P., Brammer, H., & Richards, K. 2011. *Arsenic Pollution: A Global Synthesis*. John Wiley & Sons.

Ravenscroft, P., Burgess, W. G., Ahmed, K. M., Burren, M., & Perrin, J. 2004. Arsenic in groundwater of the Bengal Basin, Bangladesh: Distribution, field relations, and hydrogeological setting. *Hydrogeology Journal* 13(5-6): 727–751.

Ravenscroft, P., H. Brammer, and K. Richards. 2009. *Arsenic Pollution: A Global Synthesis*. http://onlinelibrary.wiley.com/book/10.1002/9781444308785.

Ravichandran, Mahalingam. 2004. Interactions Between Mercury and Dissolved Organic Matter—a Review. *Chemosphere* 55(3): 319–331.

Reeder, J., & Knight, R. 2010. Rapid denoising of pyrosequencing amplicon data: exploiting the rank-abundance distribution. *Nature methods* 7(9): 668–669.

Rittle, K. A., Drever, J. I., and Colberg, P. J. S. 1995. Precipitation of arsenic during bacterial sulfate reduction. *Geomicrobiol. J.* 13: 1-11.

Rosen, B. P., Weigel, U., Monticello, R. A., and Edwards, B. P. 1991. Molecular analysis of an anion pump: purification of the *ArsC* protein. *Arch. Biochem. Biophys.* 284: 381-385.

Rowland, H. A. L., Pederick, R. L., Polya, D. A., Pancost, R. D., Dongen, B. E. V., Gault, A. G., Vaughan, D. J., Bryant, C., Anderson, B., and Lloyd, J.R. 2007. The control of organic matter on microbially mediated iron reduction and arsenic release in shallow alluvial aquifers, Cambodia. *Geobiology* 5: 281-292.

Saalfield, S. L., and Bostick, B. C. 2009. Changes in iron, sulfur, and arsenic speciation associated with bacterial sulfate reduction in ferrihydrite-rich systems. *Environ. Sci. Technol.* 43: 8787-8793.

Saltikov, C. W., and Newman, D. K. 2003. Genetic identification of a respiratory arsenate reductase. *P. Natl. Acad. Sci. USA* 100: 10983 -10988.

Schloss, P. D., and Handelsman, J. 2004. Status of the microbial census. *Microbiol. Mol. Biol. Rev.* 68: 686-691.

Scott, D. T., McKnight, D. M., Blunt-Harris, E. L., Kolesar, S. E., & Lovley, D. R. 1998. Quinone Moieties Act as Electron Acceptors in the Reduction of Humic Substances by Humics-Reducing Microorganisms. *Environmental Science & Technology* 32(19), 2984–2989.

Seshadri, R., Adrian, L., Fouts, D. E., Eisen, J. A., Phillippy, A. M., Methe, B. A., Ward, N. L., et al. 2005. Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. *Science* 307: 105-108.

Sessitsch, A., Weilharter, A., Gerzabek, M. H., Kirchmann, H., and Kandeler, E. 2001. Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. *Appl. Environ. Microbiol.* 67: 4215-4224.

Shimotori, K., Y. Omori, and T. Hama. 2009. Bacterial Production of Marine Humic-like Fluorescent Dissolved Organic Matter and Its Biogeochemical Importance. *Aquatic Microbial Ecology* 58(1): 55–66.

Shirokova, V. L., and F. G. Ferris. 2013. Microbial Diversity and Biogeochemistry of a Shallow Pristine Canadian Shield Groundwater System. *Geomicrobiology Journal* 30(2): 140–149.

Smedley, P.L. and D.G. Kinniburgh. 2002. A review of the source, behaviour and distribution of arsenic in natural waters, *Applied Geochemistry* 17: 517-568.

Smith, A. H., Lingas, E. O., & Rahman, M. 2000. Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. *Bulletin of the World Health Organization* 78(9): 1093–1103.

Spain, A. M., Krumholz, L. R., and Elshahed, M. S. 2009. Abundance, composition, diversity and novelty of soil *Proteobacteria*. *ISME J.* 3: 992-1000.

Sparks, D. L. 2003. *Environmental soil chemistry*. San Diego, California: Academic Press.

Stanger, G. 2005. A palaeo-hydrogeological model for arsenic contamination in southern and south-east Asia. *Environmental geochemistry and health* 27(4): 359–368.

Stanish, L. F., S. P. O'Neill, A. Gonzalez, T. M. Legg, J. Knelman, D. M. McKnight, S. Spaulding, and D. R. Nemergut. 2012. Bacteria and Diatom Co-occurrence Patterns in Microbial Mats from Polar Desert Streams. *Environmental Microbiology*: doi:10.1111/j.1462-2920.2012.02872.x.

Stedmon, C. A, S. Markager, and R. Bro. 2003. Tracing Dissolved Organic Matter in Aquatic Environments Using a New Approach to Fluorescence Spectroscopy. *Marine Chemistry* 82(3–4): 239–254.

Stedmon, C. A., & Markager, S. 2005. Tracing the production and degradation of autochthonous fractions of dissolved organic matter by fluorescence analysis. *Limnology and Oceanography*, 1415–1426.

Stefanowicz, A. M., Niklinska, M., and Laskowski, R. 2008. Metals affect soil bacterial and fungal functional diversity differently. *Environ. Toxicol. Chem.* 27: 591-598.

Stelzer, R. S., J. Heffernan, and G. E. Likens. 2003. The Influence of Dissolved Nutrients and Particulate Organic Matter Quality on Microbial Respiration and Biomass in a Forest Stream. *Freshwater Biology* 48(11): 1925–1937.

Stolz, J. F., D. J. Ellis, J. S. Blum, D. Ahmann, D. R. Lovley, and R. S. Oremland. 1999. Sulfurospirillum Barnesii Sp. Nov. and Sulfurospirillum Arsenophilum Sp. Nov., New Members of the Sulfurospirillum Clade of the ε-Proteobacteria. *International Journal of Systematic Bacteriology* 49(3): 1177–1180.

Straub, K. L., Kappler, A., & Schink, B. 2005. Enrichment and isolation of ferric-iron- and humic- acid- reducing bacteria. *Methods in Enzymology* 397: 58–77.

Stute, M., Zheng, Y., Schlosser, P., Horneman, A., Dhar, R., Datta, S., Hoque, M., Seddique, A.A., Shamsudduha, M., Ahmed, K.M., and van Geen, A. 2007. Hydrological control of As concentrations in Bangladesh groundwater. *Water Resour. Res.* 43: W09417.

Sun, Y., Polishchuk, E. A., Radoja, U., and Cullen, W. R. 2004. Identification and quantification of *arsC* genes in environmental samples by using real-time PCR. *J. Microbiol. Meth.* 58: 335-349.

Tedetti, M., R. Longhitano, N. Garcia, C. Guigue, N. Ferretto, and M. Goutx. 2012. Fluorescence Properties of Dissolved Organic Matter in Coastal Mediterranean Waters Influenced by a Municipal Sewage Effluent (Bay of Marseilles, France). *Environmental Chemistry* 9(5): 438–449.

Thorne, L. T., and Nickless, G. 1981. The relation between heavy metals and particle size fractions within the severn estuary (U.K.) inter-tidal sediments. *Sci. Total Environ.* 19: 207-213.

Tufano, K. J., and Fendorf, S. 2008. Confounding impacts of iron reduction on arsenic retention. *Environ. Sci. Technol.* 42: 4777-4783.

Uchiyama, I., M. Mihara, H. Nishide, and H. Chiba. 2012. MBGD Update 2013: The Microbial Genome Database for Exploring the Diversity of Microbial World. *Nucleic Acids Research* 41 (D1): D631–D635.

van Geen, A., Zheng, Y., Versteeg, R., Stute, M., Horneman, A., Dhar, R. K., Steckler, M., Gelman, A., Small, C., Ahsan, H., Graziano, J., Hussein, I. and Ahmed, K. M. (2003). Spatial variability of arsenic in 6000 tube wells in a 25 $km^2$ area of Bangladesh. *Water Resour. Res.* 39: 1140-1156.

van Geen, A., Rose, J., Thoral, S., Garnier, J. M., Zheng, Y., & Bottero, J. Y. (2004). Decoupling of As and Fe release to Bangladesh groundwater under reducing conditions. Part II: Evidence from sediment incubations. *Geochimica et Cosmochimica Acta* 68(17): 3475–3486.

van Geen, A., Aziz, Z., Horneman, A., Weinman, B., Dhar, R. K., Zheng, Y., Goodbred, S., Versteeg, R., Seddique, A. A., Hoque, M. A., and Ahmed, K. M. (2006a). Preliminary evidence of a link between surface soil properties and the arsenic content of shallow groundwater in Bangladesh. *J. Geochem. Explor.* 88: 157-161.

van Geen, A., Zheng, Y., Cheng, Z., Aziz, Z., Horneman, A., Dhar, R. K., Mailloux, B., Stute, M., Weinman, B., Goodbred, S., Seddique, A. A., Hoque, M. A., and Ahmed, K. M. (2006b). A transect of groundwater and sediment properties in Araihazar, Bangladesh: Further evidence of decoupling between As and Fe mobilization. *Chem. Geol.* 228: 85-96.

van Geen, A., Zheng, Y., Goodbred, S., Horneman, A., Aziz, Z., Cheng, Z., … Ahmed, K. M. (2008). Flushing History as a Hydrogeological Control on the Regional Distribution of Arsenic in Shallow Groundwater of the Bengal Basin. *Environmental Science & Technology* 42(7): 2283–2288.

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology* 73(16): 5261–5267.

Wasserman, G. A., Liu, X., Parvez, F., Ahsan, H., Factor-Litvak, P., van Geen, A., … Graziano, J. H. 2004. Water Arsenic Exposure and Children's Intellectual Function in Araihazar, Bangladesh. *Environmental Health Perspectives* 112(13): 1329–1333.

Weaver, W. and Shannon, C. E. 1949) *Recent contributions to the mathematical theory of communication*. Urbana, Illinois: University of Illinois Press.

Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R., & Mopper, K. 2003. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environmental Science & Technology* 37(20): 4702–4708.

Wentworth, C. K. (1922). A scale of grade and class terms for clastic sediments. *J. Geol.* 30: 377-392.

West, K. A., D. R. Johnson, P. Hu, T. Z. DeSantis, E. L. Brodie, P. K. H. Lee, H. Feil, G. L. Andersen, S. H. Zinder, and L. Alvarez-Cohen. 2008. Comparative Genomics of 'Dehalococcoides Ethenogenes' 195 and an Enrichment Culture Containing Unsequenced 'Dehalococcoides' Strains. *Applied and Environmental Microbiology* 74(11): 3533–3540.

Williams, C. J., Y. Yamashita, H. F. Wilson, R. Jaffé, and M. A. Xenopoulos. 2010. Unraveling the Role of Land Use and Microbial Activity in Shaping Dissolved Organic Matter Characteristics in Stream Ecosystems. *Limnology and Oceanography* 55(3): 1159–1171.

Winkel, L., Berg, M., Amini, M., Hug, S.J., and Johnson, C. A. 2008. Predicting groundwater arsenic contamination in Southeast Asia from surface parameters. *Nature Geosci.* 1: 536-542.

Yamashita, Y., and E. Tanoue. 2003. Chemical Characterization of Protein-like Fluorophores in DOM in Relation to Aromatic Amino Acids. *Marine Chemistry* 82(3–4): 255–271.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., Gilbert, J. A., et al. 2011. Minimum information about a marker gene sequence (MIMARKS) and

minimum information about any (x) sequence (MIxS) specifications. *Nat. Biotech*. 29: 415-420.

Yu, G., P. He, and L. Shao. 2010. Novel Insights into Sludge Dewaterability by Fluorescence Excitation-emission Matrix Combined with Parallel Factor Analysis. *Water Research* 44(3): 797–806.

Zar, J. H. 1999. *Biostatistical analysis* (Fourth edition). Upper Saddle River, New Jersey: Prentice Hall.

Zhang, Ying, Miao Hu, Zhao Jiang, and Yang Liu. 2011. Effect of Anthraquinone-2,6-disulfonate on the Trichloroethene Degradation by Dehalococcoides-containing Consortium. *Journal of Hazardous Materials* 192(3): 1896–1899.

Zheng, Y., A. van Geen, M. Stute, R. Dhar, Z. Mo, Z. Cheng, A. Horneman, et al. 2005. Geochemical and Hydrogeological Contrasts Between Shallow and Deeper Aquifers in Two Villages of Araihazar, Bangladesh: Implications for Deeper Aquifers as Drinking Water Sources. *Geochimica Et Cosmochimica Acta* 69(22): 5203–5218.

Zheng, Y., Stute, M., van Geen, A., Gavrieli, I., Dhar, R. K., Simpson, H. J., Schlosser, P., and Ahmed, K. M. (2004). Redox control of arsenic mobilization in Bangladesh groundwater. *Appl. Geochem.* 19, 201-214.

Zhou, J., Wu, L., Deng, Y., Zhi, X., Jiang, Y.-H., Tu, Q., Xie, J., et al. 2011. Reproducibility and quantitation of amplicon sequencing-based detection. *ISME J* 5: 1303-1313.

# APPENDIX 1

Table A.1. The range and mean of the number of sequences in each sample after applying the following filtering regime to the sequence dataset: removing OTUs represented by a single sequence, and removing OTUs present in only one sample.

| **OTU-picking Method/Identity Threshold** | **Number of Sequences after Filtering** |
|---|---|
| uclust/97 | 590-1392; mean: 889.8 |
| cd-hit/97 | 575-1375; mean: 858.3 |
| cd-hit/90 | 994-1734; mean: 1301.1 |
| cd-hit/95 | 741-1542; mean: 1039.3 |
| cd-hit/99 | 334-1009; mean: 550.7 |

Table A.2. Results from Mantel tests evaluating the relationship between bacterial community structure (UniFrac) and environmental variables. The UniFrac distances were calculated two different OTU-clustering methods (uclust and cd-hit), and for four different OTU identity thresholds (90%, 95%, 97% and 99%) for the cd-hit method. In addition, UniFrac distances were calculated for rarified datasets within each cluster method and identity threshold (i.e., 100 sequences from the each sample for cd-hit at 97%). In the table the *r* statistics are listed, followed by the *p*- value.

| **Mantel *r* Statistic and Corresponding *p*-value**<br>Unweighted Unifrac vs. Environmental Variables<br>(*r* Statistic; *p*-value) | | | | | | | |
|---|---|---|---|---|---|---|---|
| **OTU-picking method/ OTU threshold/ No. sequences sampled** | **%Soil Moisture** | **pH** | **Log (%C)** | **Mn (ppm)** | **√As (ppm)** | **√Fe (ppm)** | **%Silt** |
| uclust/97%/100 | -0.05; 0.55 | -0.17; 0.96 | 0.38; 0.002 | 0.59; 0.001 | -0.10; 0.71 | 0.28; 0.004 | 0.48; 0.01 |
| uclust/97%/280 | -0.05; 0.55 | -0.17; 0.96 | 0.43; 0.003 | 0.59; 0.001 | -0.09; 0.65 | 0.27; 0.010 | 0.49; 0.01 |
| uclust/97%/580 | 0.003; 0.36 | -0.14; 0.91 | 0.39; 0.002 | 0.56; 0.001 | -0.08; 0.61 | 0.27; 0.005 | 0.49; 0.01 |
| cd-hit/97%/100 | -0.05; 0.53 | -0.15; 0.92 | 0.42; 0.004 | 0.59; 0.001 | -0.10; 0.71 | 0.28; 0.007 | 0.44; 0.01 |
| cd-hit/97%/290 | -0.04; 0.48 | -0.17; 0.96 | 0.40; 0.003 | 0.56; 0.001 | -0.11; 0.73 | 0.24; 0.008 | 0.41; 0.02 |
| cd-hit/97%/575 | -0.04; 0.54 | -0.12; 0.87 | 0.35; 0.008 | 0.53; 0.003 | -0.08; 0.63 | 0.25; 0.006 | 0.39; 0.02 |
| cd-hit/90%/100 | -0.07; 0.67 | -0.17; 0.97 | 0.43; 0.004 | 0.58; 0.001 | -0.15; 0.92 | 0.27; 0.006 | 0.42; 0.01 |
| cd-hit/90%/360 | -0.05; 0.55 | -0.16; 0.93 | 0.38; 0.007 | 0.57; 0.001 | -0.06; 0.55 | 0.27; 0.007 | 0.43; 0.02 |
| cd-hit/90%/630 | -0.04; 0.52 | -0.13; 0.88 | 0.38; 0.004 | 0.55; 0.001 | -0.07; 0.60 | 0.25; 0.008 | 0.45; 0.01 |
| cd-hit/90%/990 | -0.03; 0.47 | -0.11; 0.84 | 0.35; 0.009 | 0.53; 0.002 | -0.07; 0.62 | 0.26; 0.008 | 0.42; 0.01 |
| cd-hit/95%/100 | -0.01; 0.40 | -0.15; 0.92 | 0.41; 0.005 | 0.58; 0.001 | -0.16; 0.92 | 0.25; 0.006 | 0.38; 0.03 |
| cd-hit/95%/420 | -0.01; 0.40 | -0.18; 0.97 | 0.41; 0.005 | 0.56; 0.001 | -0.03; 0.43 | 0.27; 0.007 | 0.42; 0.02 |
| cd-hit/95%/740 | -0.03; 0.49 | -0.13; 0.87 | 0.39; 0.003 | 0.56; 0.001 | -0.07; 0.57 | 0.28; 0.005 | 0.42; 0.02 |
| cd-hit/99%/100 | -0.11; 0.81 | -0.19; 0.97 | 0.39; 0.005 | 0.58; 0.001 | -0.06; 0.57 | 0.29; 0.003 | 0.50; 0.004 |
| cd-hit/99%/210 | -0.06; 0.59 | -0.15; 0.91 | 0.36; 0.009 | 0.54; 0.001 | -0.09; 0.67 | 0.24; 0.011 | 0.49; 0.003 |
| cd-hit/99%/320 | -0.04; 0.48 | -0.13; 0.87 | 0.35; 0.003 | 0.55; 0.001 | -0.09; 0.68 | 0.26; 0.011 | 0.46; 0.01 |

Figure A.1. Collector's curves for the Chao 1 richness estimator and Shannon diversity index for two different OTU-picking methods. These curves show that the alpha diversity of a majority of the samples approaches an asymptote, and there is a wide range in the alpha diversity of the samples.