BIOMARKERS IN THREE BANGLADESH AQUIFER SITES

BIOMARKERS OF BIOGEOCHEMICAL CARBON CYCLING AT THREE AQUIFER SITES IN BANGLADESH

By Reisa Joy Gascon San Pedro, BSc

A thesis submitted to the School of Graduate Studies in Partial Completion of the Requirements for the Degree of Master of Science

McMaster University MASTER OF SCIENCE (2019) Hamilton, Ontario

(Earth and Environmental Science)

TITLE: Biomarkers of Biogeochemical Carbon Cycling at Three Aquifer Sites in Bangladesh
AUTHOR: Reisa Joy Gascon San Pedro, B.Sc. SUPERVISOR: Professor Gregory F. Slater, NUMBER OF PAGES xi, 125.

Abstract

The role of aquifer microorganisms in controlling arsenic contamination of Bangladesh aquifers via oxidation of organic carbon coupled with reduction of sedimentary iron oxyhydroxides and concomitant arsenic dissolution is generally accepted. What remains to be ascertained is the *in situ* biogeochemical mechanisms of cycling different carbon sources and directly relating indigenous microbiota to arsenic release.

Using biomarker fingerprint approaches, this dissertation expanded the presently growing research in the biogeochemical carbon cycling controlling arsenic contamination in Bangladesh aquifers. Comprehensive profiles of microbial cell membrane components (PLFA and sterols) at three different aquifers tested the regional distribution of aquifer microbial community abundance, structure, and organic input potential across Araihazar. The highly variable bulk viable microbial biomass observed across these three sites confer both regional-scale and localized heterogeneous distributions of in-aquifer microbial communities which control carbon cycling in the aquifer. The lack of correlation between PLFA biomarkers and dissolved arsenic challenges the assumption that greater extent of microbial community metabolism results in an increase in arsenic in groundwater. Natural abundance radiocarbon isotope $\Delta^{14}C$ analysis of cell membrane PLFA and available carbon pools (SOC, DOC, DIC) confirmed that young organic carbon substrates are being cycled at two of the three sites investigated here. This corroborates previous reports at nearby sites (Site B and F) thereby contributing to a well-constrained carbon source which actively support microbial metabolism over a regional scale. Sterol biomarker distributions were characterized to determine potential sources of organic input into the aquifer. In particular, the importance of raw human and/or animal sewage waste as a source of labile carbon was assessed by measuring the faecal biomarker Coprostanol and comparing its abundance to other sources of biogenic sterols using sewage input proxies (Sewage Contamination Index, Coprostanol/Cholesterol ratio). This was motivated by previous findings which correlated sewage contamination with dissolved arsenic at depth at nearby sites. While sewage contamination was low in the shallow aquifers at these sites, it is more likely that plant organic matter supported the elevated microbial abundance at shallow depths. On the other hand, evidence presented in this project suggests that sewage contamination intrudes into deeper aquifers (e.g. buried Pleistocene) and contributes to the vulnerability of previous pristine aquifers to future arsenic contamination.

Acknowledgments

"Most things disappoint till you look deeper" *-Graham Greene*

This thesis would not be possible without supervision from Dr. Greg Slater who generously provided a platform upon which my academic creativity and enthusiasm can flourish. On more occasions than I can count, he has helped me stay focused and understand the insight from the evidence. His confidence in my capabilities helped me achieve research milestones I could not have imagined for myself.

My infinite gratitude also goes to Jennie Kirby for her tireless assistance with wet lab procedures and patience in training me as a novice analytical chemist; and Dr. Allyson Brady without whom seemingly endless Excel sheets of raw numbers would be meaningless. My thanks also to Dr. Kelly Whaley-Martin for laying out the integral foundations of this project in the Slater lab and extending her hand to me in friendship and intellectual guidance. I also thank the rest of the Environmental Organic Geochemistry group: Mohammed, Sian, Robbyn, Chris, Mike, Julia, Johnny, Rebecca, Jacob. Their enthusiasm for scientific discovery colored the lab, our shared offices, and often The Phoenix and made me so thankful to be surrounded by such bright minds.

I am also grateful to Dr. Alexander van Geen of Columbia University and Dr. Brian J. Mailloux of Barnard College who helped put the environmental microbiology focus of this research project into worldly context. This last year, I was fortunate to join them in Bangladesh for a field sampling expedition and while I gained invaluable field experience, I was also able to place insightful socio-economic perspectives into my remote research.

Finally, as ever, I am so thankful for my loving and understanding parents, Boniloisa and Reynaldo San Pedro, who fostered my enthusiasm for science even from a young age. No two people have understood my motivations more and cheered me on when I decided to pursue a considerably unconventional career path two years ago. My sister, Reiann San Pedro, for her delightful sense of humor; she keeps me grounded and reminds me 'there is a world outside of your laboratory'. Thank you from the bottom of my heart.

Table of Contents

List of Figures

List of Abbreviations and symbols

PLFA – phospholipid fatty acids
FAME – fatty acid methyl ester
DOC – dissolved organic carbon
DIC – dissolved inorganic carbon
SOC – sedimentary organic carbon
DCM – dichloromethane
MeOH – methane
GC-MS – gas chromatograph-mass spectrometer
WHO – World Health Organizations
Cl/Br – Chloride/Bromide ratio
Fe – iron
As – arsenic

Chapter 1. Introduction

Arsenic (As) contamination of drinking water in Bangladesh and many parts of Southeast Asia continues to be a regional public health crisis since groundwater wells in rural Bangladesh were discovered to have concentrations that exceed the Bangladeshi standard of 50µg As/L (BGS & DPHE, 2001; Centeno et al., 2006; Chakraborti et al., 2015; Smedley & Kinniburgh, 2002; van Geen et al. 2003). Chronic exposure to high concentrations of arsenic can lead to poisoning (known as Arsenicosis) which can manifest in various health problems such as skin lesions, cardiovascular disease, and cancer (Hossain, Hasibuzzaman & Himeno, 2019). The source of arsenic is largely geogenic, primarily sorbed onto sedimentary Fe oxides of alluvial and deltaic formations that make up the Bengal basin (Acharyya, Lahiri, Raymahashay, & Bhowmik, 2000; Shamsudduha, Uddin, Saunders, & Lee, 2008). The dissolution of arsenic into groundwater enabled its widespread occurrence and transport as dissolved ions, As(III) or As(V); As(III) being more toxic than As(V) (Akter & Naidu, 2006).

The primary release mechanism responsible for the accumulation of high concentrations of dissolved arsenic in groundwater is triggered by microbially mediated reductive dissolution of arsenic-bearing sedimentary Fe(oxy)hydroxides (equation [1] below). This biogeochemical process is driven by organic matter in the aquifer. For example, labile organic matter dissolved by surface water and enters the aquifer contributes to stimulate microbial metabolism driving the increase in dissolved arsenic (Harvey et al., 2006; Mailloux et al., 2013; Whaley-Martin et al., 2016). However, heterogeneities in the availability of carbon sources, subsurface biogeochemical cycling, and sedimentary arsenic concentration result in high spatial variability of arsenic contamination in different aquifers across Bangladesh. Therefore, this dissertation

expands on previous work which characterized indicators of *in situ* biogeochemical cycling driven by organic inputs in the aquifer (Mailloux et al., 2013; Whaley-Martin et al., 2017, 2016). Using sedimentary biomarkers (phospholipid fatty acids, sterols) and natural abundance radiocarbon isotope $\Delta^{14}C$ analysis, this dissertation elucidates carbon sources that may be driving the biogeochemical cycling that contributes to the spatial variability of arsenic contaminated aquifers. Constraining this variability becomes more important as the demand for clean low-arsenic groundwater increases across the country (Van Geen et al., 2014).

1.1 Redox zonation in the aquifer environment

To provide the energy to fuel microbial metabolism in the natural environment, microbial respirers catalyze redox (reduction-oxidation) reactions (Jin & Bethke, 2003). These are conducted by simultaneously reducing an electron acceptor and oxidizing an electron donor, two half-reactions. Each half reaction provides free energy available to microorganisms in the environment; the net usable energy (i.e. Gibbs free energy, ΔG^o) can determine which terminal electron accepting process (TEAP) dominates (Bethke, Sanford, Kirk, Jin, & Flynn, 2011). In general, the electron acceptors used for these TEAP include O_2, NO_3^-, Mn(IV), Fe(III), SO_4^{2-}, and available reduced carbon often as by-products of anoxic fermentation to produce methane (Bethke et al., 2011; McMahon, Chapelle, & Bradley, 2011). The reduction of these electron acceptors develop into a hierarchical redox ladder (Figure 1) (Champ, Gulens, & Jackson, 1979; McMahon et al., 2011). In the pristine aquifer environment, the "highest rung" of this thermodynamic ladder begins with aerobic respiration which dominates in oxic environments followed by nitrate reduction (ΔG^o_{TEAP} = -550 kJ/mol NO_3^- for denitrification) common in slightly anoxic environments. The energetic yields produced during Mn(IV) (-417 to -

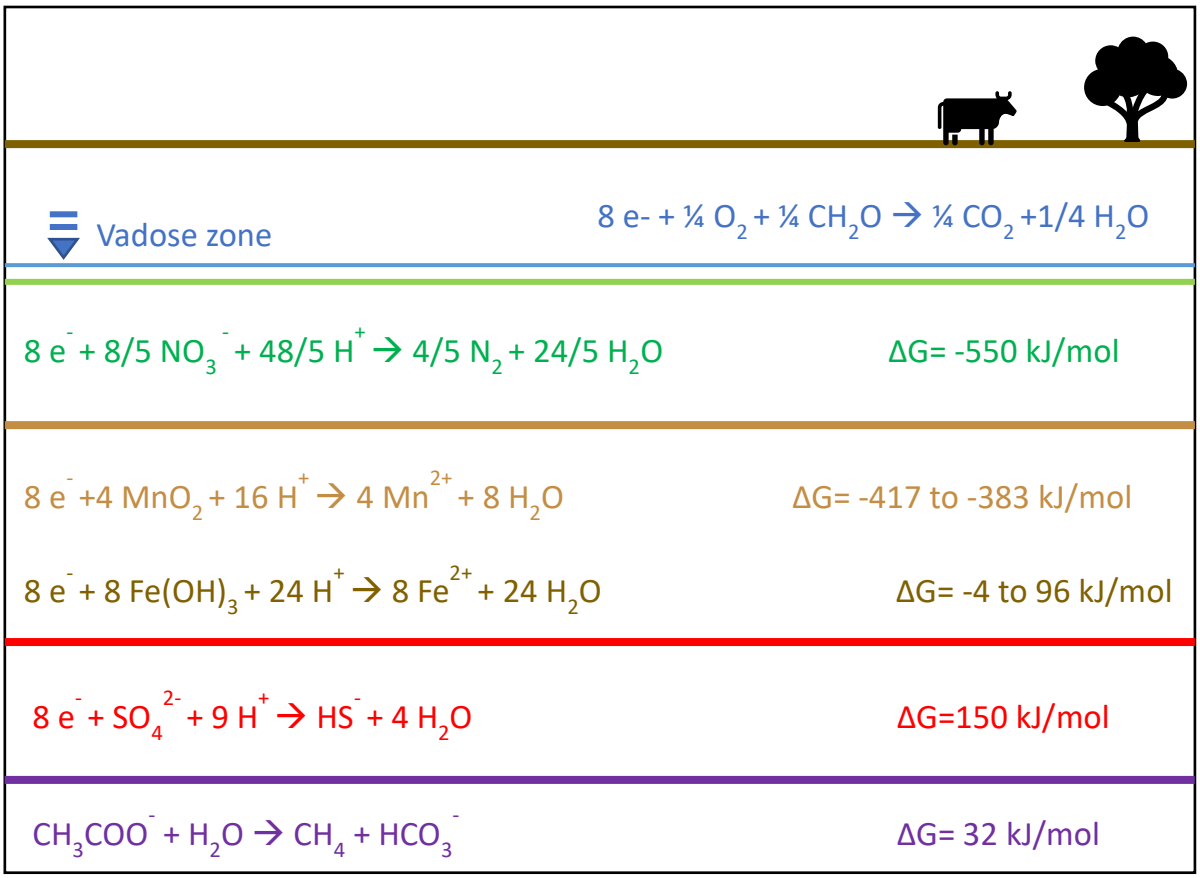


Figure 1. *'The Redox Ladder': terminal electron accepting reactions in order from most (top, aerobic) to least (bottom, methanogenic) favourable. Gibb's free energies from Bethke et al. (2011).*

383 kJ/mol MnO_2) and Fe(III) reduction can vary depending on which mineral oxide is present in the environment (Bethke et al., 2011). For example, the reduction of Goethite, an Fe oxide, produces ΔG^{o}_{TEAP} = 155 kJ/mol FeOOH (Bigham et al., 1996) while reduction of Magnetite produces ΔG^{o}_{TEAP} = 231 kJ/mol Fe_3O_4 (Bethke et al., 2011; Delany & Lundeen, 1989). Sulfate reduction follows this, producing ΔG^{o}_{TEAP} = 150 kJ/mol SO_4^{2-} (Bethke et al., 2011; Delany & Lundeen, 1989). Finally, the least thermodynamically favourable reaction is methanogenesis wherein the fermentation of substrates such as CH_3COOH and subsequent reduction of its by-product, CO_2, produces CH_4. Alternatively, the direct reduction of CO_2 can also produce CH_4. The last four electron accepting processes occur in anoxic environments considered slightly to

highly reducing. This sequence provides a general guideline to determine the prevalence of functional microbial groups in the environment. For example, because aerobic respirers can produce greater free energy yields by reducing O_2 than metal-reducing bacteria that reduce Fe(III), they can produce greater microbial yields and are thus more competitive in oxic environments. In reality, the reduction of any combination of these electron acceptors may be simultaneously present in the environment and may be enabled by mutualistic relationships between microorganisms that catalyse different electron accepting reactions (Bethke et al., 2011). For example, Fe(III) and SO_4^{2-} reduction can simultaneously occur with the oxidation of organic carbon in anoxic aquifers (Postma & Jakobsen, 1996). The presence or prevalence of each of these TEA may vary in the aquifer environment. For example, oxygen – and thus aerobic respiration – may be abundant within the vadose zone but is rapidly consumed below the water table below which anoxic conditions promote more reducing conditions. Similarly, Fe(III) oxidation would only occur where Fe(III)-bearing minerals or Fe(III) oxide-coated sediments are present (McMahon et al., 2011).

1.2 Sources and distribution of arsenic in Bangladesh aquifers

The Bengal basin, which includes Bangladesh and West Bengal, is made up of Quaternary fluvio-deltaic deposits; it is also known as the Ganges-Brahmaputra-Meghna plain after the three major rivers that course through the region (Alam, Alam, Curray, Chowdhury, & Gani, 2003). The naturally occurring, or geogenic, arsenic was deposited during the formation of the sedimentary aquifers in the Cenozoic (i.e. orogenic deposits from the nearby Himalayas). (Alam, 1989; Ravenscroft, 1999; Ravenscroft, Burgess, Ahmed, Burren, & Perrin, 2005). This arsenic is considered immobilized as sorbed constituents on sediment grains such as

Fe(oxy)hydroxide-coated sand (Smedley & Kinniburgh, 2002). Solid arsenic concentrations in Bangladesh subsurface have been reported to range from 1.0 to 14.7 mg/kg (Smedley & Kinniburgh, 2002)

Regionally, at the sub-basin scale across Bangladesh, the shallow Holocene-age aquifers (<100 m) host most of the wells that are contaminated with arsenic (>10 or >50 µg/L As). Discontinuous silt-clay stratigraphic layers are also common in shallow sandy Holocene aquifers and imparts hydrologic anisotropy, resulting in the heterogenous distribution of dissolved arsenic concentrations at the village scale (M. A. Hoque, Hoque, & Ahmed, 2007; M. Hoque, Burgess, Shamsudduha, & Ahmed, 2011; M. Hoque et al., 2009). These shallow Holocene aquifers are considered hydraulically separated from the deeper Pleistocene-age aquifers (>100 m) (Mihajlov et al., 2016). Deeper Pleistocene aquifers have generally low arsenic (<10 µg/L As) which represent potentially sustainable sources of uncontaminated groundwater (Acharyya et al., 2000; BGH & DPHE, 2001; M. Hoque et al., 2011; Ravenscroft, 1999; Shamsudduha et al., 2008; van Geen et al., 2008). High-As Holocene aquifers and low-As Pleistocene aquifers are further distinguished based on aquifer reducing conditions. In general, Holocene aquifers are highly reducing and are distinguished by the grey sediment color whereas Pleistocene aquifers have orange-colored sediments emblematic of the oxidizing conditions in these aquifers (Biswas et al., 2012). This difference in the extent of reduction of these aquifers is visually apparent due to the staining of sediments with orange-colored Fe oxides and has been quantified using diffuse spectral reflectance analysis (Horneman et al., 2004). The highly reducing conditions of grey Holocene aquifers enable the reductive dissolution of sedimentary iron oxides and concomitant arsenic release consistent with elevated dissolved arsenic in these

aquifers. On the other hand, the low arsenic concentrations in the deeper orange Pleistocene aquifers indicate that sedimentary iron oxides have not yet been reduced (dissolved) and thus, arsenic remains sorbed (immobilized).

1.3 Reductive dissolution as a primary arsenic release mechanism in Bangladesh

The primary release mechanism of arsenic into groundwater is triggered by reductive dissolution of sedimentary Fe(oxy)hydroxides and desorption of arsenic previously sorbed onto sediment surfaces (equation [1]) (Acharyya et al., 2000; BGH & DPHE, 2001; Héry et al., 2010; Islam et al., 2004; Nickson et al., 1998; Shamsudduha et al., 2008; Y. Zheng et al., 2004). Driven by reduced organic carbon (CH_2O term in equation [1]) , *in situ* bacterial communities (e.g. iron reducing bacteria) catalyze this reductive dissolution and use the free energy released to fuel metabolic processes (Dhar et al., 2011; Héry et al., 2010; Islam et al., 2004; Legg et al., 2012; Shi, Squier, Zachara, & Fredrickson, 2007).

$$CH_2O + FeOOH_{(s)} + H_2O \rightarrow Fe(II)_{(aq)} + CO_2 + 2H_2O + H^+ \qquad [1]$$

The reduction of Fe (oxy)hydroxides in the sedimentary aquifers of Bangladesh is strongly controlled by the availability of electron donors; even small concentrations of reduced carbon in a low organic matter aquifer can support metal-reducing bacterial communities (Héry et al., 2010). While equation [1] demonstrates the utilization of reduced organic carbon as a substrate, other metabolic processes demand the utilization of inorganic substrates. For example, dissolve inorganic carbon (DIC) which often includes CO_2, HCO_3^-, H_2CO_3 is an oxidized carbon pool present in the aquifer in different forms: as by-products of organic carbon respiration, introduced from the surface as part of groundwater recharge, or dissolved from

equilibrium exchange with sedimentary carbonates in the aquifer geology. In the surface environment, the most common example of inorganic carbon fixation is the photosynthetic metabolism of plants; in the aquifer, inorganic carbon is fixed by autotrophic microorganisms such as hydrogenotrophic methanogens which use H_2 as an electron acceptor and CO_2 as an electron donor to produce CH_4. These examples show that microbially mediated redox reactions strongly influence the cycling of carbon – whether organic or inorganic – in the aquifer environment.

1.4 Biomarkers as indicators of microbial community abundance and sewage

Molecular biomarkers are biosynthesized organic compounds that can be related to specific biological sources or processes and are commonly used to trace biological activity in the environment (Eglinton & Calvin, 1967; Eglinton, Scott, Belsky, Burlingame, & Calvin, 1964; Simoneit, 2004). Biomarkers have been used extensively to delineate biological cycling of carbon in the environment over geological timescales. For example, measuring radiocarbon (^{14}C) isotopic compositions of microbial lipid biomarkers and available carbon pools in the environment have confirmed the incorporation of 'fossil' carbon (e.g. shale) presumed to be recalcitrant (Petsch, Eglinton, & Edwards, 2001). Microbial biomarkers have also been used to delineate the important carbon substrates which support microbial communities in environmental systems where carbon pools contain a mixture of different substrates (Pearson, McNichol, Benitez-Nelson, Hayes, & Eglinton, 2001; V. G. Shah et al., 2007; Whaley-Martin et al., 2016). In environmental impact studies, biomarker analysis has been applied to assess the impacts of aggressive crop farming (e.g. monoculture plots, fungicide applications) (Zelles, 1999; Zelles, Bai, Beck, & Beese, 1992; Zelles et al., 1994) and analysing sewage contamination

from livestock farming and urban effluents (Carreira, Wagener, & Readman, 2004; Chalaux, Takada, & Bayona, 1995; Reeves & Patton, 2005). This research project focuses on two types of biomarkers – phospholipid fatty acids and sterols – to assess microbial abundance estimates, sewage input potential, and carbon source cycling in the anoxic aquifer environment. Molecular markers can also be subjected to a variety of analytical methods to glean more specific insights into their provenance. For example, measuring the radiocarbon isotopic content of archaeal and bacterial cell membrane components can elucidate the carbon sources metabolized by these microorganisms and conferring their dominance in an environmental system (Ahad, Burns, Mancini, & Slater, 2010; Bradford, Ziolkowski, Goad, Warren, & Slater, 2017; Brady, Goordial, Sun, Whyte, & Slater, 2018; Cowie, Greenberg, & Slater, 2010; Mahmoudi, Porter, et al., 2013; Morrill, Szponar, Johnston, Marvin, & Slater, 2014; Pearson et al., 2001; Petsch et al., 2001; S. R. Shah, Mollenhauer, Ohkouchi, Eglinton, & Pearson, 2008; Slater, White, Eglinton, & Reddy, 2005; Whaley-Martin et al., 2016; Ziolkowski, Wierzchos, Davila, & Slater, 2013).

1.5 Using Phospholipid Fatty Acids to estimate microbial community abundance

A widely used type of biomarker are molecular phospholipid fatty acids (PLFA). Phospholipids help maintain the structural integrity of bacterial cells and are thus, an integral component of microbial cell membrane (Figure 2). The cell membrane regulates the passage of water, nutrients, and solutes while maintaining homeostatic conditions ideal for microbial operations (Silhavy, Kahne, & Walker, 2010; Zhang & Rock, 2008). Phospholipids are made up of a glycerol moiety with a phosphate head group and two fatty acid chains (the PLFA) bonded via ester links. They rapidly biodegrade upon cell lysis (i.e. cell death) which provides information about the *in situ* microbial communities viable at the time of environmental sample collection

(Guckert, Antworth, Nichols, & White, 1985; Vestal & White, 1989). The utility of molecular PLFA as bacterial community tracers is attributed to their source-specific qualities. Certain PLFA compounds have been used as unique tracers for the presence of specific or related bacterial groups in the environment. For example, the detection of the PLFA compound 10me16:0 (16-carbon saturated and branched) is used to identify the sulfate-reducing bacteria (SRB) genus *desulfobacter.* However, 10me16:0 has also been detected in the iron-reducing bacteria *Geobacter metallireducens*, an organism closely related to SRB, but which uses Fe(III) instead of SO_4^{2-} as an electron acceptor to oxidized reduced organic carbon (Lovley et al., 1993; C. L. Zhang et al., 2003). On the other hand, the PLFA i17:1ω7c (17-carbon monounsaturated and branched) has been used to trace the SRB genus *desulfovibrio* (Coleman et al., 1993; Vestal & White, 1989). On a microbial community-scale, PLFA biomarkers can be quantified in bulk format instead of the species-specific identification of individual PLFA compounds. Using a generic cell conversion factor, the total or bulk cell abundance of a the *in situ* viable or extant microbial community in any environmental system can be estimated (Frostegård, Tunlid, & Bååth, 2011; Green & Scow, 2000). The biosynthesis of new PLFA and/or molecular modification of existing PLFA is strongly controlled by environmental conditions such as temperature, pH, oxygen, moisture content, and nutrient availability (Garwin, Klages, & Cronan, 1980; Lechevalier, 1977; Parsons & Rock, 2013; Zelles, 1999; Zhang & Rock, 2008). For example, changes in the relative proportion of key groups (e.g. saturated to unsaturated PLFA), changes from *cis*- to *trans*- configurations of unsaturated PLFA, and the creation of cyclopropyl PLFA have been used to assess bacterial response to stressors such as a decrease in temperature, pH, or desiccation (Figure 2) (Kieft, Ringelberg, & White, 1994; Zhang & Rock, 2008). (Lechevalier,

1977) detailed an increase in unsaturated (containing double or multiple C-C bonds) and subsequent decrease in saturated (exclusively single C-C bonds) PLFA as well as shortening of PLFA chains to maintain cell membrane fluidity when ambient temperature decreases.

Practical applications of PLFA profiling have been previously demonstrated to successfully characterize carbon cycling as mediated by microbial communities in different environmental systems. The *in situ* biodegradation of petroleum hydrocarbons which were presumed to be a recalcitrant organic carbon substrate were confirmed using PLFA profiling at impacted sites (e.g. polycyclic aromatic hydrocarbons contamination, crude oil spills) (Mahmoudi, Fulthorpe, Burns, Mancini, & Slater, 2013; Mahmoudi, Porter, et al., 2013; Morrill et al., 2014; Rooney-Varga, Anderson, Fraga, Ringelberg, & Lovley, 1999). From a resource industry (e.g. mining, petroleum refining) perspective, natural attenuation of key contaminants in reclamation sites such as end-pit lakes have been described by *in situ* PLFA distribution in order to assess their long-term sustainability (Ahad & Pakdel, 2013; Bradford et al., 2017). In Araihazar aquifers, PLFA profiles and available carbon pools have been previously characterized in conjunction with radiocarbon-based analysis to identify the predominant source of carbon supporting in-aquifer microbial community metabolism. These studies demonstrate the usefulness of PLFA biomarkers to understand biogeochemical carbon cycling controlled by microbial communities in the environment (Coleman, Hedrick, Lovley, White, & Pye, 1993; Guckert et al., 1985; Ludvigsen, Albrechtsen, Ringelberg, Ekelund, & Christensen, 1999).

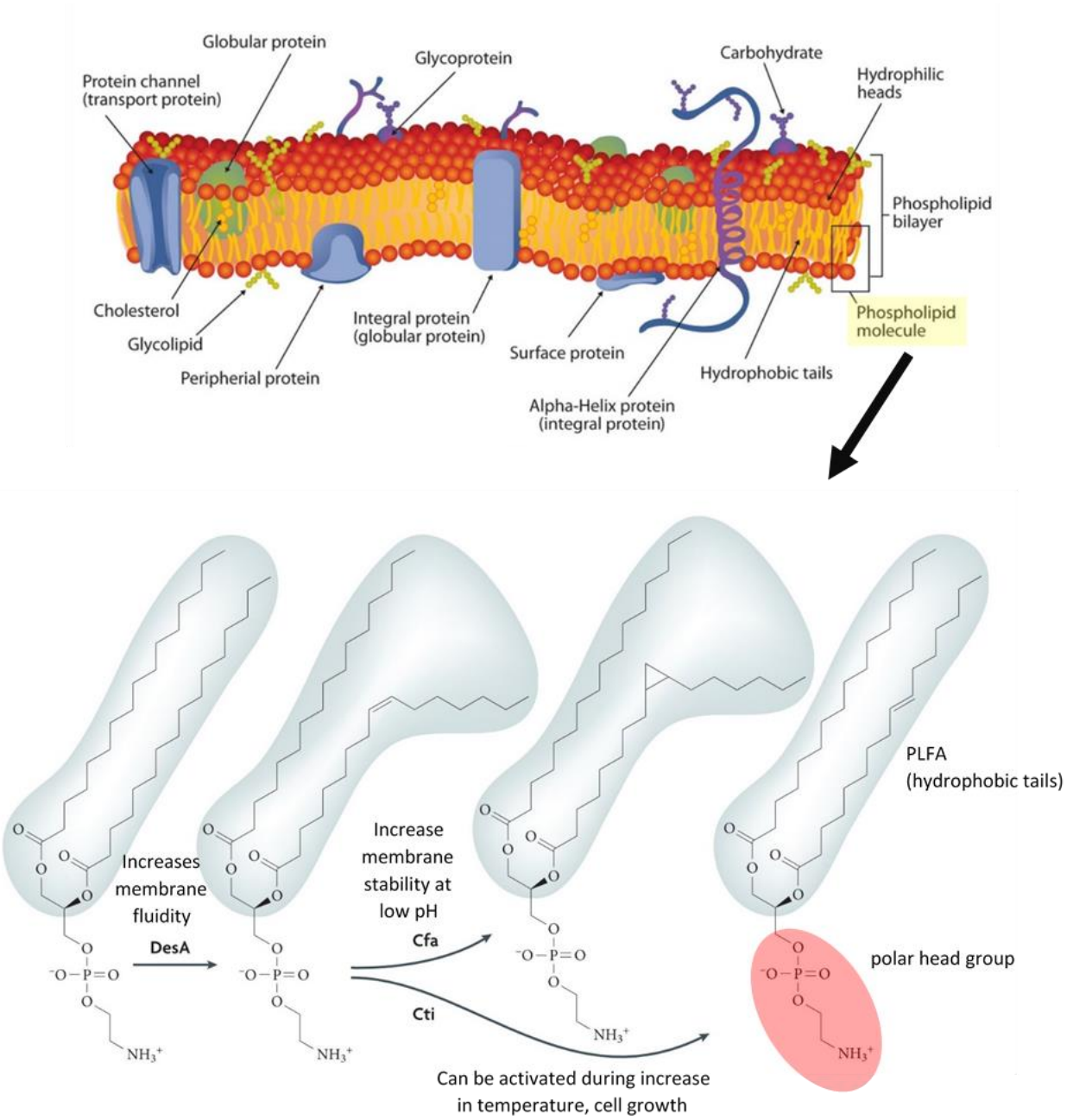


Figure 2. *Schematic model of a cell membrane (top) containing protein and lipid components integral for microbial operations; diagram modified from Chemistry LibreText, University of California Davis. Molecular structures of phospholipids and four examples of structural changes which occur in response to stressors; "DesA", "Cfa", and "Cti" represent enzymatic alterations to existing PLFA via addition of double bonds (cis, trans) and/or cyclopropyl groups; diagram modified from Zhang et al. (2008).*

1.6 Assessing sewage input potential via organic sterol profile in the environment

Sterols are a diverse group of organic compounds biologically produced in the natural environment from microbial, plant, and animal activity. Sterols (commonly with 27 to 29 carbons) are abundant in eukaryotic cells and are similar to PLFA in that they're also integral cell membrane components that help regulate membrane fluidity (Hartmann, 1998; Peters, Walters, Moldowan, 2007; Volkman et al., 1987). The provenance of sterols can be traced to specific biological activity, thus making them useful 'fingerprints' to study biogeochemical cycling of organic matter in diverse environmental systems. For example, Coprostanol is produced from the biohydrogenation of Cholesterol as mediated by enteric bacteria in the

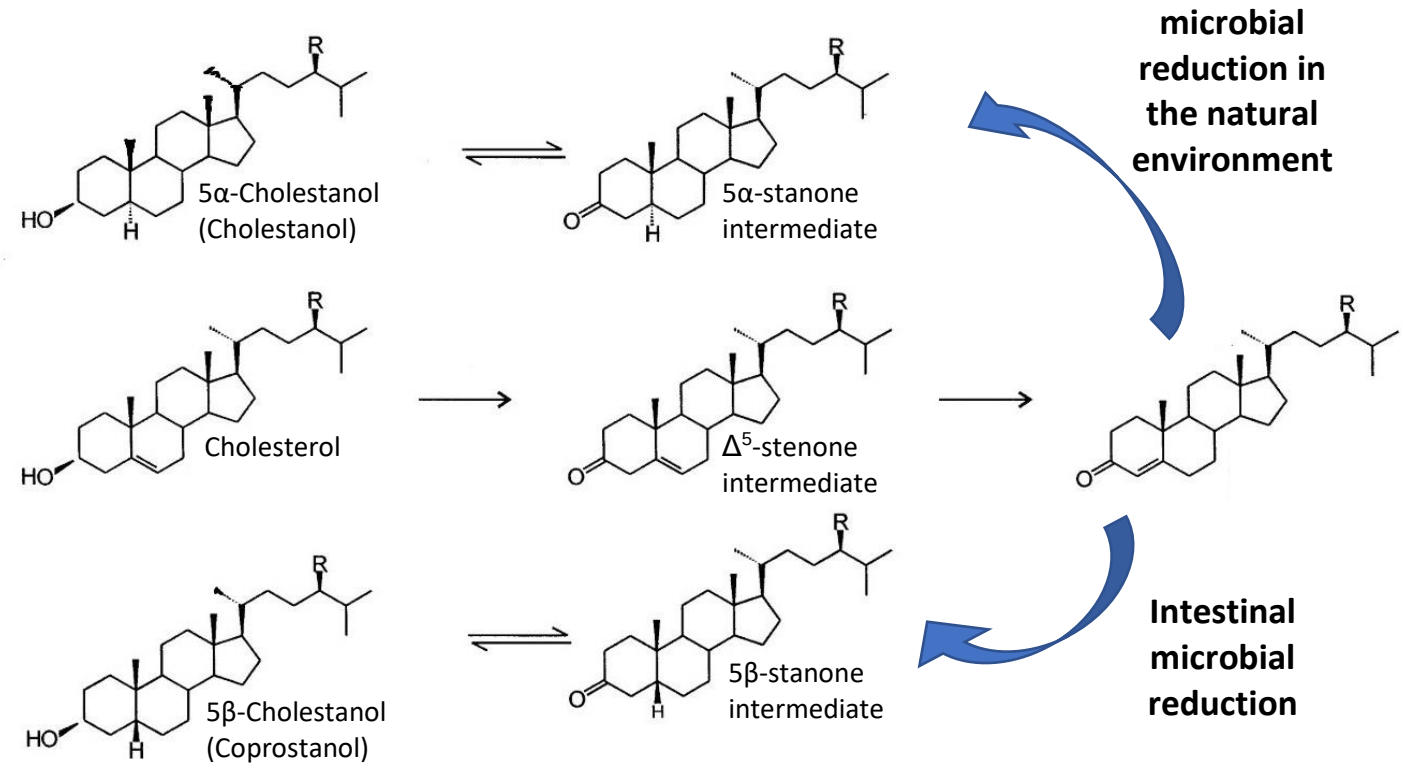


Figure 3. *Diagram showing chemical structures three common organic sterols and their degradation pathways. Cholesterol undergoes biohydrogenation in the mammalian gut and anoxic aquifer to produce the two isomers, Coprostanol and Cholestanol. Diagram modified from Bull et al. (2002).*

digestive tract of upper mammals including humans and cattle (Bull, Matthew J. Lockheart, Mohamed M. Elhmmali, David J. Roberts, & Richard P. Evershed, 2002; Grimalt, Fernández, Bayona, & Albaigés, 1990; Jardé, Gruau, Mansuy-Huault, Peu, & Martinez, 2007; Martins, Seyffert, Braun, & Fillmann, 2011; Rosenfeld, Fukushima, Hellman, & Gallagher, 1954; V. G. Shah et al., 2007). Coprostanol can make up as much as 60% of the total sterol content of human faeces (Bull et al., 2002; Isobe et al., 2002; Leeming, Ball, Ashbolt, & Nichols, 1996; S. R. Shah et al., 2008). This perhaps helps explain why absolute Coprostanol concentrations have been reported to be within the µg/g to mg/g scale at sites contaminated by sewage (Carreira et al., 2004; Dsikowitzky et al., 2017; He, Zhang, Tang, Cui, & Sun, 2018; Isobe et al., 2002; Kolpin et al., 2000; Speranza, Colombo, Skorupka, & Colombo, 2018; Vane et al., 2010).

Because Coprostanol can also be synthesized by non-enteric bacteria in anoxic sediments (Huang & Meinschein, 1979; Nishimura & Koyama, 1977), its detection in environmental samples does not unambiguously indicate inputs of human and animal sewage (Figure 3). However, because it still makes up majority of organic sterols in feces, Coprostanol remains a useful fingerprint for sewage. In order to assess the importance of sewage inputs in the aquifer, the Sewage Contamination Index (SCI) is calculated. The SCI estimates the relative abundance of sewage derived Coprostanol to other naturally occurring biogenic sterols (equation [2]). In the SCI, [5β-coprostanol] represents the absolute concentration of the sewage-derived Coprostanol while [5α-cholestanol] represents the absolute concentration of the isomer of non-enteric Cholestanol (Grimalt et al., 1990).

$$\frac{[5\beta-coprostanol]}{[5\beta-coprostanol]+[5\alpha-cholestanol]} \quad [2]$$

A calculated SCI of ≥0.7 indicates that an important component of the sterol pool is attributed to animal/human waste input thus confirming significant sewage contamination (Dsikowitzky et al., 2017; Grimalt et al., 1990; He et al., 2018; Reeves & Patton, 2005; V. G. Shah et al., 2007). By using ratios to interpret the relative importance of sewage-derived coprostanol relative to the organic sterol pool, researchers have overcome the ambiguity of the source of coprostanol from its absolute concentration (Martins et al., 2011; Reeves & Patton, 2005; Speranza et al., 2018; Vane et al., 2010; Whaley-Martin et al., 2017). Previously, the SCI has been used in Araihazar, Bangladesh to demonstrate positive correlation between increasing sewage contamination and arsenic concentration at depth (Whaley-Martin et al., 2017).

A complementary sewage input proxy is the Coprostanol/Cholesterol (C/Ch) concentration ratio; C/Ch ratios >1 indicate significant sewage contamination in the environment (Fattore, Benfenati, Marelli, Cools, & Fanelli, 1996). While Cholesterol is also abundant in sewage (Isobe et al., 2002), its subsequent enteric bacterial degradation to Coprostanol implies a lower relative concentration in environments contaminated by human and/or animal sewage. Relative concentrations of these sewage sterols are compared with plant-derived phytosterols, Campesterol, Stigmasterol, and β-Sitosterol to determine dominant sources of natural organic sterols within the Site M aquifer. In plant cells, phytosterols such as Campesterol, β-Sitosterol and Stigmasterol are used to regulate membrane fluidity – a similar biochemical function is exercised by Cholesterol in animal cells (Hartmann, 1998). For example, a study of the impact of milling on the nutrient content of rice determined that in every degree of processing, Campesterol, β-Sitosterol and Stigmasterol were consistently the most abundant phytosterols in rice (Ha et al., 2006). It is therefore reasonable to expect that at terrestrial sites

such as Bangladesh aquifers above which rice and wheat crop fields abound, the organic sterol pool is dominated by phytosterols. At previous sites in Araihazar, increasing sewage contamination correlates with increasing dissolved Arsenic concentrations at depth which suggests that sewage may be an important source of organic carbon driving microbially mediated Arsenic release in the aquifer (Whaley-Martin et al., 2017). This is conceivable given the ubiquity of domestic pit latrines and shallow ponds at the surface providing an available source of sewage into the aquifer (Knappett et al., 2011).

1.7 Sterols describe general sources of organic matter input into the aquifer

Another group of organic sterols useful as a molecular biomarker are phytosterols. Examples include Campesterol, Stigmasterol, and β-Sitosterol which are all commonly associated with higher plants (e.g. vascular plants), fungi, and zooplankton (Martins et al., 2011; Pratt, Warnken, Leeming, Arthur, & Grice, 2008; Speranza et al., 2018). The presence of these sterol biomarkers in the natural environment can elucidate sources of organic matter and can be helpful in assessing which sources dominate in a particular system. For example, the relative abundance of sewage-derived fecal sterols has been compared to terrestrial- or marine-derived phytosterols to determine impacts of anthropogenic activity on the organic matter loading in estuarine environments (e.g. growing industrial activity in metropolitan cities like Jakarta, Indonesia and the Rio de la Plata estuary off Buenos Aires, Argentina) (Dsikowitzky et al., 2017; Speranza et al., 2018). Campesterol, stigmasterol, and β-sitosterol are phytosterols abundant in rice making them particularly important molecular markers in the shallow subsurface in Bangladesh where over 70% of agricultural land is used for growing rice (Ha et al., 2006; Rahman, 2010)

Organic sterols are rapidly cycled (i.e. biodegraded) at the surface where ambient conditions are highly oxidizing (Nishimura & Koyama, 1977). However, in the highly reducing Bangladesh aquifers, anoxic conditions enable the preservation of sterols (Pratt et al., 2008). Furthermore, as hydrophobic organic compounds with high octanol-water K_{ow} partitioning coefficients (constants which determine the tendency of molecules to be dissolved in water in the environment; for coprostanol, $K_{ow} = 8.2$), sterols become preferentially sorbed onto aquifer sediments (Appelo & Postma, 2013; Froehner & Sánez, 2013; Whaley-Martin et al., 2017). Therefore, the concentration and extent of degradation of sterols at the surface (i.e. surface-derived DOC) can be expected to be different from sterol abundance and preservation within the anoxic aquifer (i.e. when sterols sorb onto sediments and become SOC) (Appelo & Postma, 2013).

1.8 Compound specific radiocarbon analysis can delineate carbon source utilization

Radiocarbon or ^{14}C is a naturally occurring radioactive isotope of carbon produced from the atmospheric breakdown of ^{14}N (Libby, 1961); ^{14}C has a half life of ~5730 years (Goodwin, 1962). The radioactivity of the ^{14}C atom means its concentration in organic samples, in addition to its known half-life, can be used to approximate the age of carbon-bearing (i.e. organic) matter in the environment. Microbial metabolism of carbon substrates in the natural environment induces kinetic isotope effects which are controlled by the mass differences between different Carbon isotopes, ^{12}C, ^{13}C, and ^{14}C. This causes the lighter isotope, ^{12}C, to be preferentially incorporated over ^{13}C or ^{14}C into biosynthesized products (e.g. molecular markers) – known as mass-dependent fractionation. The relative abundance of each isotope, denoted by the $\delta^{14}C$ notation, can be measured in a sample against a standard with a known

isotopic signature. The radiocarbon isotopic composition of organic matter can be calculated using equation [4] against a standard Oxalic Acid II (M Stuiver & Polach, 1977; Minze Stuiver et al., 1993).

$$\Delta^{14}C = \delta^{14}C - 2(\delta^{13}C + 25)\left(1 + \frac{\delta^{14}C}{1000}\right) \quad [4]$$

Equation [4] estimates the abundance of radioactive ^{14}C, eliminates the effects of mass-dependent fractionation, and corrects for the biosynthetic incorporation of atmospheric ^{14}C that was released during atomic bomb testing in the 1950s (knowns as "bomb spike") (M Stuiver & Polach, 1977). Thus, radiocarbon signatures ($\Delta^{14}C$) of molecular markers such as PLFA and carbon substrates in the environment can be compared to elucidate carbon source utilization for biosynthesis in the environment.

$$\Delta^{14}C = [F_m \times e^{\lambda(1950-Y_c)} - 1] \times 1000 \quad [5]$$

Through equation [5], the fraction of modern ^{14}C (F_m) measured from organic matter via Accelerator Mass Spectrometer in the environment in reference to the year of sample collection (Y_c) can be calculated and reported in ‰ (Woods Hole Oceanographic Institute website, 2019). Organic matter older than about 60,000 kya will have $\Delta^{14}C$ signatures close to -1000‰ while more modern (atmospheric) carbon will be positive. With respect to molecular markers, positive $\Delta^{14}C$ values indicate microbial incorporation/metabolism of organic matter containing recently fixed (modern) atmospheric ^{14}C likely produced via photosynthesis or incorporation of dissolved inorganic carbon (DIC) (Bradford et al., 2017; Cowie, Slater, Bernier, & Warren, 2009; Mahmoudi, Fulthorpe, et al., 2013; Petsch et al., 2001; Slater et al., 2005). For the purpose of natural abundance radiocarbon analysis of PLFA in this study, reagents (i.e.

Methanol) with known $\Delta^{14}C$ signatures were used for the methylation of PLFA biomarkers and conversion to FAMEs. Equation [6] is used to correct for the added carbon during methylation of PLFA to convert to fatty acid methyl esters (FAME) - where '*N*' is the average number of carbon atoms of the bulk PLFA (Brady et al., 2018; Simkus et al., 2016).

$$\Delta^{14}C_{FAME} = [(N+1) \times \Delta^{14}C_{measured} - \Delta^{14}C_{MeOH}]/N \quad [6]$$

Natural abundance radiocarbon isotope analysis is a useful tool to elucidate complex carbon cycling that is driven by biogeochemical processes in diverse environmental systems. For example, the radiocarbon signatures of integral cell biomarkers (PLFA) have been compared to available carbon substrates with highly distinct $\Delta^{14}C$ values to determine how these carbon sources drive microbial metabolism in the aquifer environment (Ahad et al., 2010; Mailloux et al., 2019, 2013; Whaley-Martin et al., 2016). With respect to site reclamation studies, radiocarbon isotope analysis has been used to evaluate and confirm the biodegradation of petroleum hydrocarbon contaminants which were previously thought to be a recalcitrant carbon source (i.e. incorporation of $\Delta^{14}C$ depleted carbon) (Ahad & Pakdel, 2013; Bradford et al., 2017; Mahmoudi et al., 2013; Slater et al., 2005). This type of analysis has also been used to demonstrate that microorganisms in inhospitable environments (e.g. deep >2 km nutrient- and carbon-limited subsurface, arid deserts) can thrive by, for example, conducting chemolithotrophy (metabolize inorganic substrates such as $H_{2(g)}$) and decreasing metabolic rates (Brady et al., 2018; Simkus et al., 2016; Slater et al., 2006; Ziolkowski et al., 2013). These studies demonstrate the utility of natural abundance ^{14}C in delineating complex biogeochemical processes mediated by indigenous microfauna and driven (or precluded) by carbon availability (or its lack thereof) in diverse environments.

1.9 Available carbon sources and microbial cycling in Bangladesh aquifers

During metabolism, microorganisms in the aquifer use free energy produced from catalyzing redox reactions like equation [1], as previously discussed. While redox zonation or the redox ladder determines which TEA are reduced to fuel metabolism, the availability and/or lability of electron donors can be a limiting factor. For example, in aquifers with little organic matter (the electron donor) and abundant electron acceptors (e.g. sedimentary Fe(III) oxides), microbial community respiration may be limited by the low organic matter content (McMahon et al., 2011). Such a system exists in Bangladesh aquifers where oligotrophic conditions (low organic content) prevail. Two overarching bulk pools of reduced carbon available in Bangladesh aquifers include solid organic matter in the aquifer sediments (sedimentary organic carbon, SOC) and dissolved organic carbon (DOC). The bulk SOC pool can include organic matter from peat preserved during burial of ancient mangroves, marshland, and swamps (Shamsudduha et al., 2008). The DOC pool can include terrestrial-derived organic matter, organic by-products of decomposition, and in-aquifer microbial cell components (Ghosh, Routh, & Bhadury, 2015; Mladenov et al., 2010). Surface sources of labile DOC can also be transported from excavated ponds (Neumann et al., 2010) and shallow pit latrines (Ahmed et al., 1994; Knappet et al., 2011; McArthur et al. 2012; Whaley-Martin et al. 2017). Allochthonous organic carbon in the aquifer can be transported from external sources through mechanical advection of groundwater while autochthonous sources develop *in situ*; autochthonous DOC can be dissolved as groundwater exchanges with SOC.

Some researchers hypothesize that the introduction of allochthonous DOC into the aquifer supplies labile carbon substrates which stimulate microbial metabolism at depth (Dhar

et al., 2011; Guo et al., 2019; Héry et al., 2010; Islam et al., 2004; Mailloux et al., 2013; Neumann, Pracht, Polizzotto, Badruzzaman, & Ali, 2014; Whaley-Martin et al., 2017, 2016). Alternatively, the addition of fresh or young labile organic matter into the subsurface can increase the bioavailability previously recalcitrant sedimentary organic matter thereby stimulating *in situ* microbial respiration (Fontaine et al., 2007; Fontaine, Mariotti, & Abbadie, 2003). The overall consequence of stimulated microbial metabolism is an increase in the reductive dissolution of sedimentary Fe(oxy)hydroxides and concomitant As release into groundwater. For example, (Mailloux et al., 2013; Whaley-Martin et al., 2016) argue that young labile organic carbon advected form the surface into the aquifer contribute to the stimulation of microbial metabolism in Bangladesh. Additionally, Whaley-Martin et al. (2017) found positive correlation between indicators of sewage contamination and dissolved As suggesting sewage inputs from surface pit latrines may contribute to As release mechanisms at depth. Similarly, using fluorescence spectroscopy, Guo et al. (2019) characterized groundwater, surface water, and sedimentary organic matter to show that labile surface carbon derived from a permanent wetland in inner Mongolia was transported into the shallow aquifer and promoted an increase in dissolved As in groundwater. This was contrasted by low dissolved As concentrations co-occurring with low organic matter content in a shallow aquifer below a recently (last 14 years) dried wetland about 4 km from the permanent wetland.

In a study of two aquifer sites in Bangladesh, researchers measured the natural abundance of ^{14}C ($\Delta^{14}C$) in microbial PLFA and DNA and compared them with $\Delta^{14}C$ signatures of bulk carbon pools including SOC, DOC, and DIC available in the aquifer. Radiocarbon $\Delta^{14}C$ values of bulk PLFA biomarkers were reported to be considerably more positive ($\Delta^{14}C_{PLFA}$ = -163‰ to

+21‰ at Site B; -167‰ to +20‰ at Site F) than bulk SOC ($\Delta^{14}C_{SOC}$ = -631 ± 54‰) which argues that the aquifer microbial communities were cycling carbon sources much younger than the bulk SOC in the aquifer. Bulk DOC ($\Delta^{14}C_{DOC}$ = -230 ± 100‰) or DIC ($\Delta^{14}C_{DIC}$ = +24 ± 30‰), which have more modern $\Delta^{14}C$ values, were proposed as alternative carbon substrates derived from the surface and transported into the aquifer (Mailloux et al., 2013; Whaley-Martin et al., 2016). Reduced organic carbon from both SOC or DOC can potentially support the heterotrophic metabolism in Bangladesh aquifers which leads to the stimulation of reductive dissolution of Fe(oxy)hydroxides coupled with As release. Previous work have shown that indigenous microorganisms within Bangladesh aquifer can respire sedimentary carbon (e.g. peat) to facilitate reductive dissolution of Fe(oxy)hydroxides and release As into groundwater (McArthur et al., 2004; Pracht, Tfaily, Ardissono, & Neumann, 2018). However, previous work has demonstrated that carbon fueling this arsenic release mechanism were predominantly young organic sources such as a modern component of the bulk DOC pool (e.g. surface ponds and pit latrines) and that old SOC (e.g. peat) remains less bioavailable (Guo et al., 2019; Harvey et al., 2002; Knappett et al., 2011; Mailloux et al., 2013; Richards et al., 2019; Whaley-Martin et al., 2017, 2016; Yu et al., 2018).

1.10 Study sites in Araihazar upazila, Bangladesh

This research project focused on three sites called Desert Island (DI), Doper Tek (DT), and Site M, in Araihazar upazila, a rural county approximately 30km east of the capital city Dhaka. While nearly two decades of extensive work in this region has addressed hydrological controls on arsenic geochemistry (Mihajlov et al., 2016; Stute et al., 2007; van Geen et al., 2008, 2003), research defining the specific role of the *in situ* biogeochemical carbon cycling controlled

Figure 3. *Study sites in Araihazar: Desert Island (DI), Doper Tek (DT), Site M, Site B, Site F all located approximately 25 km east of Dhaka. DI, DT, and Site M studied in this present research project; Site B and F assessed previously by Whaley-Martin et al. 2016, 2017; Mailloux et al. 2013).*

by indigenous microbial communities is limited. Considering previous microcosm-based research have identified specific bacteria capable of reducing sedimentary iron oxides and concomitant arsenic release (Islam et al., 2004), it is important to understand to what extent these microorganisms are driven by additions of carbon in their environment. Previous work at two other sites, Site B and F, in Araihazar showed that inputs of organic carbon derived from the surface is correlated with increased arsenic concentrations at depth and may be stimulating the microbially mediated reduction of Fe(oxy)hydroxides and Arenic release (Mailloux et al., 2013; Whaley-Martin et al., 2017, 2016). This dissertation expands on this previous work by using cell membrane PLFA and organic sterol biomarkers as well as radiocarbon isotope analysis

to quantify viable biomass and potential organic carbon inputs (e.g. sewage, plant organic matter) into the aquifers at three additional sites in Araihazar.

The aquifer at Desert Island (DI) is Holocene in age as determined from radiocarbon isotope analysis of wood fragments buried within the aquifer (6970 ± 45 ^{14}C yr at 40 m and 7180 ± ^{14}C yr at 44 m depth) (Radloff et al., 2007) while Doper Tek (DT) is Pleistocene in age based on sedimentary organic carbon (SOC) dated at 19,300 ^{14}C yr (8 m) and 18,450 ^{14}C yr (13 m). The lithostratigraphy of DI is characterized by medium to fine sand with a clay layer at 1.5-3 m while DT is dominated by fine sand with a medium sand layer at 11-12 m and clay layers at 1.5-3 m and 14 m (Table S1). The gray sediment at DI indicates reducing geochemical conditions whereas the orange sediments at DT indicates oxidizing conditions (Figure 3); this is confirmed using diffuse spectral reflectance measurements correlating the relative abundance of Fe(II)/Fe and the extent of reduction of sedimentary aquifers (Horneman et al., 2004; Radloff et al., 2007; Yan Zheng et al., 2005). DI harbours lower solid-phase and higher dissolved arsenic concentrations compared to DT (Table S1). The high solid phase As concentrations occur within the sand-dominated depths at both DI and DT and are greater than the average crustal concentration of 2 mg/kg (Yan Zheng et al., 2005). The mean groundwater arsenic concentration at DT was 7 ± 0.4 µg/L and 90 ± 5 µg/L at DI (Figure 3). Arsenic concentrations reported at Site B (100-530 ug/L) and F (5-200 ug/L) were much higher than either DI or DT (Whaley-Martin et al., 2016). DI, because of its proximity to low population denisty, is hypothesized to receive low sewage input whereas DT, located adjacent to a dense cluster of community wells and thus human population, is hypothesized to receive high sewage input (Figure 3).

1.11 Research objectives in studying microbial biomarkers and biogeochemical carbon cycling in Bangladesh aquifers

It remains widely accepted that arsenic contamination in Bangladesh aquifers is controlled by microorganisms which couple oxidation of organic carbon with reductive dissolution of arsenic-bearing (sorbed) sedimentary Fe(oxy)hydroxide. By using molecular biomarker fingerprints (PLFA, sterols), this dissertation assessed the potential relationships between viable microbial biomass, available carbon sources, and potential sewage input at three new sites, Desert Island (DI), Dopar Tek (DT), and Site M. Previous work has presented evidence that young carbon sources support microbial community metabolism and were argued to derive from the surface. Meanwhile, unsewered human and other animal waste has been presented as an example of such a surface-derived labile source of young carbon. The possibility of sewage as a potential labile carbon substrate fuelling microbial metabolism and causing an increase in arsenic in groundwater was tested at these three new sites in Araihazar.

Chapter 2 described the viable microbial community and the potential for human and animal unsewered waste as a carbon source to support microbially mediated arsenic release at the first two sites, DI and DT. The age of available carbon pools was determined using natural abundance radiocarbon isotope $\Delta^{14}C$ analysis to delineate carbon sources that support microbial metabolism. The sewage input proxies, Sewage Contamination Index and Coprostanol/Cholesterol ratio, estimated the importance of sewage by comparing Coprostanol (sewage biomarker) to other naturally derived biogenic sterols in the aquifer. These biomarker-based approaches were also applied to study a third site, Site M, in Chapter 3. Similar to DI and DT, the viable microbial abundance was quantified at Site M while PLFA distribution profiles

delineated notable shifts in the microbial community structure between these three sites. However, Site M is unique from DI and DT in that the subsurface lithology at Site M captures a deeper aquifer site than previously investigated in Araihazar. At Site M, the shallow Holocene aquifer is separated from the underlying Pleistocene by a thick organic rich clay confining unit. Previous work at Site M was motivated by the sudden increase of arsenic in a deep groundwater community well tapping a previously pristine Pleistocene aquifer; this was argued to be triggered by reactive carbon advected from the organic-rich clay confining unit due to distant municipal pumping in Dhaka. The potential for unsewered human and animal waste as another source of reactive carbon entering into deeper aquifer depths that previously reported was also examined in Chapter 3. This has important implications for the long-term sustainability of presently pristine aquifers as well as their susceptibility to arsenic contamination due to increasing anthropogenic demand (e.g. pumping). Studying these three sites at Araihazar, in addition to previous research at nearby sites (I.e. Site B and F), addressed the widespread regional distribution of biogeochemical carbon cycling controlling arsenic release in these aquifers. The in-depth characterization of aquifer viable biomass and organic inputs in this dissertation provides new insight into the impact of surface carbon sources (e.g. sewage) on arsenic contamination in the aquifer.

Bibliography

Acharyya, S. K., Lahiri, S., Raymahashay, B. C., & Bhowmik, A. (2000). Arsenic toxicity of groundwater in parts of the Bengal basin in India and Bangladesh: The role of Quaternary stratigraphy and Holocene sea-level fluctuation. *Environmental Geology*, *39*(10), 1127–1137.

Ahad, J. M. E., Burns, L., Mancini, S., & Slater, G. F. (2010). Assessing microbial uptake of petroleum hydrocarbons in groundwater systems using natural abundance radiocarbon. *Environmental Science and Technology*, *44*(13), 5092–5097.

Ahad, J. M. E., & Pakdel, H. (2013). Direct Evaluation of in Situ Biodegradation in Athabasca Oil Sands Tailings Ponds Using Natural Abundance Radiocarbon. *Environmental Science & Technology*, *47*(18), 10214–10222.

Alam, M. (1989). Geology and depositional history of Cenozoic sediments of the Bengal Basin of Bangladesh. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *69*, 125–139.

Alam, M., Alam, M. M., Curray, J. R., Chowdhury, M. L. R., & Gani, M. R. (2003). An overview of the sedimentary geology of the Bengal Basin in relation to the regional tectonic framework and basin-fill history. *Sedimentary Geology*, *155*(3–4), 179–208.

Appelo, C.A.J., & Postma, D. (2005). *Geochemistry, groundwater and pollution.* Leiden, The Netherlands: A.A. Balkema.

Bethke, C. M., Sanford, R. A., Kirk, M. F., Jin, Q., & Flynn, T. M. (2011). The thermodynamic ladder in geomicrobiology. *American Journal of Science*, *311*(3), 183–210.

BGS & DPHE (2001). Arsenic contamination of groundwater in Bangladesh. *British Geological Survey (Technical Report, WC/00/19. 4 Volumes).*

Biswas, A., Nath, B., Bhattacharya, P., Halder, D., Kundu, A. K., Mandal, U., … Jacks, G. (2012). Hydrogeochemical contrast between brown and grey sand aquifers in shallow depth of Bengal Basin: Consequences for sustainable drinking water supply. *Science of the Total Environment*, *431*, 402–412.

Bradford, L. M., Ziolkowski, L. A., Goad, C., Warren, L. A., & Slater, G. F. (2017). Elucidating carbon sources driving microbial metabolism during oil sands reclamation. *Journal of Environmental Management*, *188*, 246–254.

Brady, A. L., Goordial, J., Sun, H. J., Whyte, L. G., & Slater, G. F. (2018). Variability in carbon uptake and (re)cycling in Antarctic cryptoendolithic microbial ecosystems demonstrated through radiocarbon analysis of organic biomarkers. *Geobiology*, *16*(1), 62–79.

Bull, I. D., Matthew J. Lockheart, Mohamed M. Elhmmali, David J. Roberts, & Richard P. Evershed. (2002). The Origin of Faeces by Means of Biomarker Detection. Environ Int - Google Scholar. *Environment International*, *27*(8), 647–654.

Carreira, R. S., Wagener, A. L. R., & Readman, J. W. (2004). Sterols as markers of sewage contamination in a tropical urban estuary (Guanabara Bay, Brazil): Space-time variations.

Estuarine, Coastal and Shelf Science, *60*(4), 587–598.

Chalaux, N., Takada, H., & Bayona, J. M. (1995). Molecular markers in Tokyo Bay sediments: Sources and distribution. *Marine Environmental Research*, *40*(1), 77–92.

Champ, D. R., Gulens, J., & Jackson, R. E. (1979). Oxidation–reduction sequences in ground water flow systems. *Canadian Journal of Earth Sciences*, *16*(1), 12–23.

Coleman, M. L., Hedrick, D. B., Lovley, D. R., White, D. C., & Pye, K. (1993). Reduction of Fe(III) in sediments by sulphate-reducing bacteria. *Nature*, *361*(6411), 436–438.

Cowie, B. R., Greenberg, B. M., & Slater, G. F. (2010). Determination of microbial carbon sources and cycling during remediation of petroleum hydrocarbon impacted soil using natural abundance14C analysis of PLFA. *Environmental Science and Technology*, *44*(7), 2322–2327.

Cowie, B. R., Slater, G. F., Bernier, L., & Warren, L. A. (2009). Carbon isotope fractionation in phospholipid fatty acid biomarkers of bacteria and fungi native to an acid mine drainage lake. *Organic Geochemistry*, *40*(9), 956–962.

Dhar, R. K., Zheng, Y., Saltikov, C. W., Radloff, K. A., Mailloux, B. J., Ahmed, K. M., & Van Geen, A. (2011). Microbes Enhance Mobility of Arsenic in Pleistocene Aquifer Sand from Bangladesh. *Environ. Sci. Technol*, *45*, 31.

Dsikowitzky, L., Schäfer, L., Dwiyitno, Ariyani, F., Irianto, H. E., & Schwarzbauer, J. (2017). Evidence of massive river pollution in the tropical megacity Jakarta as indicated by faecal steroid occurrence and the seasonal flushing out into the coastal ecosystem. *Environmental Chemistry Letters*, *15*(4), 703–708.

Eglinton, G., & Calvin, M. (1967). Chemical Fossils. *Scientific American*, *216*(1), 32–43.

Eglinton, G., Scott, P. M., Belsky, T., Burlingame, A. L., & Calvin, M. (1964). Hydrocarbons of Biological Origin from a one-Billion-Year-Old Sediment. *Science*, *145*(3629), 263–264.

Fattore, E., Benfenati, E., Marelli, R., Cools, E., & Fanelli, R. (1996). Sterols in sediment samples from Venice Lagoon, Italy. *Chemosphere*, *33*(12), 2383–2393.

Fontaine, S., Barot, S., Barré, P., Bdioui, N., Mary, B., & Rumpel, C. (2007). Stability of organic carbon in deep soil layers controlled by fresh carbon supply. *Nature*, *450*(7167), 277–280.

Fontaine, S., Mariotti, A., & Abbadie, L. (2003). The priming effect of organic matter: a question of microbial competition? *Soil Biology and Biochemistry*, *35*(6), 837–843.

Frostegård, Å., Tunlid, A., & Bååth, E. (2011). Use and misuse of PLFA measurements in soils. *Soil Biology and Biochemistry*, *43*(8), 1621–1625.

Garwin, J. L., Klages, A. L., & Cronan, J. E. (1980). β-Ketoacyl-acyl carrier protein synthase II of Escherichia coli. Evidence for function in the thermal regulation of fatty acid synthesis. *Journal of Biological Chemistry*, *255*(8), 3263–3265.

Ghosh, D., Routh, J., & Bhadury, P. (2015). Characterization and microbial utilization of dissolved lipid organic fraction in arsenic impacted aquifers (India). *Journal of Hydrology*,

527, 221–233.

Goodwin, H. (1962). Half-life of Radiocarbon. *Nature*, *195*, 984.

Green, C. T., & Scow, K. M. (2000). Analysis of phospholipid fatty acids (PLFA) to characterize microbial communities in aquifers. *Hydrogeology Journal*, *8*(1), 126–141.

Grimalt, J. O., Fernández, P., Bayona, J. M., & Albaigés, J. (1990). Assessment of Fecal Sterols and Ketones as Indicators of Urban Sewage Inputs to Coastal Waters. *Environmental Science and Technology*, *24*(3), 357–363.

Guckert, J. B., Antworth, C. P., Nichols, P. D., & White, D. C. (1985). Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. *FEMS Microbiology Letters*, *31*(3), 147–158.

Guo, H., Li, X., Xiu, W., He, W., Cao, Y., Zhang, D., & Wang, A. (2019). Controls of organic matter bioreactivity on arsenic mobility in shallow aquifers of the Hetao Basin, P.R. China. *Journal of Hydrology*, *571*(November 2018), 448–459.

Ha, T. Y., Ko, S. N., Lee, S. M., Kim, H. R., Chung, S. H., Kim, S. R., … Kim, I. H. (2006). Changes in nutraceutical lipid components of rice at different degrees of milling. *European Journal of Lipid Science and Technology*, *108*(3), 175–181. https://doi.org/10.1002/ejlt.200500250

Hartmann, M. A. (1998). Plant sterols and the membrane environment. *Trends in Plant Science*, *3*(5), 170–175.

Harvey, C. F., Swartz, C. H., Badruzzaman, A. B. M., Keon-Blute, N., Yu, W., Ashraf Ali, M., … Ahmed, M. F. (2002). Arsenic Mobility and Groundwater Extraction in Bangladesh. *Science*, *298*, 1602–1606.

He, D., Zhang, K., Tang, J., Cui, X., & Sun, Y. (2018). Using fecal sterols to assess dynamics of sewage input in sediments along a human-impacted river-estuary system in eastern China. *Science of the Total Environment*, *636*, 787–797.

Héry, M., Van Dongen, B. E., Gill, F., Mondal, D., Vaughan, D. J., Pancost, R. D., … Lloyd, J. R. (2010). Arsenic release and attenuation in low organic carbon aquifer sediments from West Bengal. *Geobiology*, *8*(2), 155–168.

Hoque, M. A., Hoque, M. M., & Ahmed, K. M. (2007). Declining groundwater level and aquifer dewatering in Dhaka metropolitan area, Bangladesh: Causes and quantification. *Hydrogeology Journal*, *15*(8), 1523–1534.

Hoque, M., Burgess, W. G., Shamsudduha, M., & Ahmed, K. M. (2011). Delineating low-arsenic groundwater environments in the Bengal Aquifer System, Bangladesh. *Applied Geochemistry*, *26*(4), 614–623.

Hoque, M., Khan, A., Shamsudduha, M., Hossain, M. S., Islam, T., & Chowdhury, S. H. (2009). Near surface lithology and spatial variation of arsenic in the shallow groundwater: Southeastern Bangladesh. *Environmental Geology*, *56*(8), 1687–1695.

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.

Huang, W. Y., & Meinschein, W. G. (1979). Sterols as ecological indicators. *Geochimica et Cosmochimica Acta*, *43*(5), 739–745.

Islam, F. S., Gault, A. G., Boothman, C., Polya, D. A., Charnock, J. M., Chatterjee, D., & Lloyd, J. R. (2004). Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature*, *430*(6995), 68–71.

Isobe, K. O., Tarao, M., Zakaria, M. P., Chiem, N. H., Minh, L. Y., & Takada, H. (2002). Quantitative application of fecal sterols using gas chromatography - Mass spectrometry to investigate fecal pollution in tropical waters: Western Malaysia and Mekong Delta, Vietnam. *Environmental Science and Technology*, *36*(21), 4497–4507.

Jardé, E., Gruau, G., Mansuy-Huault, L., Peu, P., & Martinez, J. (2007). Using sterols to detect pig slurry contribution to soil organic matter. *Water, Air, and Soil Pollution*, *178*(1–4), 169–178.

Jin, Q., & Bethke, C. M. (2003). A new rate law describing microbial respiration. *Applied and Environmental Microbiology*, *69*(4), 2340–2348.

Kieft, T. L., Ringelberg, D. B., & White, D. C. (1994). Changes in Ester-Linked Phospholipid Fatty Acid Profiles of Subsurface Bacteria during Starvation and Desiccation in a Porous Medium. *Applied and Environmental Microbiology*, *60*(9), 3292–3299.

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.

Kolpin, D., Furlong, E., Meyer, M., Barber, L., Zaugg, S., Buxton, H., & Meyer, M. T. (2000). Pharmaceuticals , Hormones , and Other Organic Wastewater Contaminants in U . S . Streams , 1999-2000 : A National Reconnaissance Other Organic Wastewater Contaminants in U . S . Streams , 1999-2000 : A National Reconnaissance. *Environmental Science & Technology*, *36*(6), 1999–2000.

Lechevalier, M. P. (1977). Lipids in Bacterial Taxonomy - A Taxonomist's View. *Critical Reviews in Microbiology*, *5*(2), 109–210.

Leeming, R., Ball, A., Ashbolt, N., & Nichols, P. (1996). Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. *Water Research*, *30*(12), 2893–2900.

Legg, T. M., Zheng, V., Simone, B., Radloff, K. A., Mladenov, N., Gonzalez, A., … Nemergut, D. R. (2012). Carbon, metals, and grain size correlate with bacterial community structure in sediments of a high arsenic aquifer. *Frontiers in Microbiology*, *3*, 1–15.

Libby, W. F. (1961). Radiocarbon Dating. *American Association for the Advancement of Science*, *133*(3453), 621–629.

Ludvigsen, L., Albrechtsen, H. J., Ringelberg, D. B., Ekelund, F., & Christensen, T. H. (1999). Distribution and composition of microbial populations in a landfill leachate contaminated aquifer (Grindsted, Denmark). *Microbial Ecology*, *37*(3), 197–207.

Mahmoudi, N., Fulthorpe, R. R., Burns, L., Mancini, S., & Slater, G. F. (2013). Assessing microbial carbon sources and potential PAH degradation using natural abundance 14C analysis. *Environmental Pollution*, *175*, 125–130.

Mahmoudi, N., Porter, T. M., Zimmerman, A. R., Fulthorpe, R. R., Kasozi, G. N., Silliman, B. R., & Slater, G. F. (2013). Rapid degradation of deepwater horizon spilled oil by indigenous microbial communities in louisiana saltmarsh sediments. *Environmental Science and Technology*, *47*(23), 13303–13312.

Mailloux, B. J., Kim, C., Kichuk, T., Nguyen, K., Precht, C., Wang, S., … Buchholz, B. A. (2019). Paired RNA Radiocarbon and Sequencing Analyses Indicate the Importance of Autotrophy in a Shallow Alluvial Aquifer. *Scientific Reports*, *9*(1), 10370.

Mailloux, B. J., Trembath-Reichert, E., Cheung, J., Watson, M., Stute, M., Freyer, G. A., … van Geen, A. (2013). Advection of surface-derived organic carbon fuels microbial reduction in Bangladesh groundwater. *Proceedings of the National Academy of Sciences*, *110*(14), 5331–5335.

Martins, C. C., Seyffert, B. H., Braun, J. A. F., & Fillmann, G. (2011). Input of organic matter in a large South American tropical estuary (paranaguá estuarine system, Brazil) indicated by sedimentary sterols and multivariate statistical approach. *Journal of the Brazilian Chemical Society*, *22*(8), 1585–1594.

McArthur, J. ., Banerjee, D. ., Hudson-Edwards, K. ., Mishra, R., Purohit, R., Ravenscroft, P., … Chadha, D. . (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. *Applied Geochemistry*, *19*(8), 1255–1293.

McMahon, P. B., Chapelle, F. H., & Bradley, P. M. (2011). Evolution of redox processes in groundwater. *ACS Symposium Series*, *1071*(January), 581–597.

Mihajlov, I., Stute, M., Schlosser, P., Mailloux, B. J., Zheng, Y., Choudhury, I., … van Geen, A. (2016). Recharge of low-arsenic aquifers tapped by community wells in Araihazar, Bangladesh, inferred from environmental isotopes. *Water Resources Research*, *52*(5), 3324–3349.

Mladenov, N., Zheng, Y., Miller, M. P., Nemergut, D. R., Legg, T., Simone, B., … Mcknight, D. M. (2010). Dissolved organic matter sources and consequences for iron and arsenic mobilization in Bangladesh aquifers. *Environmental Science and Technology*, *44*(1), 123–128.

Morrill, P. L., Szponar, N., Johnston, M., Marvin, C., & Slater, G. F. (2014). Deciphering microbial

carbon substrates in PAH contaminated sediments using phospholipid fatty acids, and compound specific δ 13 C and δ 14 C. *Organic Geochemistry*, *69*, 76–87.

Neumann, R. B., Pracht, L. E., Polizzotto, M. L., Badruzzaman, A. B. M., & Ali, M. A. (2014). Biodegradable Organic Carbon in Sediments of an Arsenic-Contaminated Aquifer in Bangladesh. *Environmental Science and Technology Letters*, *1*(4), 221–225.

Nickson, R., McArthur, J., Burgess, W., Ahmed, K. M., Ravenscroft, P., & Rahman, M. (1998). Arsenic poisoning of Bangladesh groundwater [7]. *Nature*, *395*(6700), 338.

Nishimura, M., & Koyama, T. (1977). The occurrence of stanols in various living organisms and the behavior of sterols in contemporary sediments. *Geochimica et Cosmochimica Acta*, *41*(3), 379–385.

Parsons, J. B., & Rock, C. O. (2013). Bacterial lipids: Metabolism and membrane homeostasis. *Progress in Lipid Research*.

Pearson, A., McNichol, A. P., Benitez-Nelson, B. C., Hayes, J. M., & Eglinton, T. I. (2001). Origins of lipid biomarkers in Santa Monica Basin surface sediment: A case study using compound-specific Δ14C analysis. *Geochimica et Cosmochimica Acta*, *65*(18), 3123–3137.

Petsch, S. T., Eglinton, T. I., & Edwards, K. J. (2001). 14C-dead living biomass: Evidence for microbial assimilation of ancient organic carbon during shale weathering. *Science*, *292*(5519), 1127–1131.

Postma, D., & Jakobsen, R. (1996). Redox zonation: Equilibrium constraints on the Fe(III)/SO4-reduction interface. *Geochimica et Cosmochimica Acta*, *60*(17), 3169–3175.

Pracht, L. E., Tfaily, M. M., Ardissono, R. J., & Neumann, R. B. (2018). Molecular characterization of organic matter mobilized from Bangladeshi aquifer sediment: Tracking carbon compositional change during microbial utilization. *Biogeosciences*, *15*(6), 1733–1747.

Pratt, C., Warnken, J., Leeming, R., Arthur, M. J., & Grice, D. I. (2008). Degradation and responses of coprostanol and selected sterol biomarkers in sediments to a simulated major sewage pollution event: A microcosm experiment under sub-tropical estuarine conditions. *Organic Geochemistry*, *39*(3), 353–369.

Radloff, K. A., Cheng, Z., Rahman, M. W., Ahmed, K. M., Mailloux, B. J., Juhl, A. R., … Van Geen, A. (2007). Mobilization of arsenic during one-year incubations of grey aquifer sands from Araihazar, Bangladesh. *Environmental Science and Technology*, *41*(10), 3639–3645.

Rahman, S. (2010). Six decades of agricultural land use change in Bangladesh: Effects on crop diversity, productivity, food availability and the environment, 1948-2006. *Singapore Journal of Tropical Geography*, *31*(2), 254–269.

Ravenscroft, P. (1999). Distribution of groundwater arsenic in Bangladesh related to geology. In *Groundwater Arsenic Contamination in the Bengal Delta Plain of Bangladesh* (pp. 41–56).

Ravenscroft, P., Burgess, W. G., Ahmed, K. M., Burren, M., & Perrin, J. (2005). Arsenic in groundwater of the Bengal Basin, Bangladesh: Distribution, field relations, and

hydrogeological setting. *Hydrogeology Journal*, *13*(5–6), 727–751.

Reeves, A., & Patton, D. (2005). Faecal sterols as indicators of sewage contamination in estuarine sediments of the Tay Estuary, Scotland: an extended baseline survey. *Hydrology and Earth System Sciences*, *9*(1/2), 81–94.

Richards, L. A., Magnone, D., Sültenfuß, J., Bryant, C., Boyce, A. J., van Dongen, B. E., … Polya, D. A. (2019). Dual in-aquifer and near surface processes drive arsenic mobilization in Cambodian groundwaters. *Science of The Total Environment*, *659*, 699–714.

Rooney-Varga, J. N., Anderson, R. T., Fraga, J. L., Ringelberg, D., & Lovley, D. R. (1999). Microbial communities associated with anaerobic benzene degradation in a petroleum-contaminated aquifer. *Applied and Environmental Microbiology*, *65*(7), 3056–3063.

Rosenfeld, R. S., Fukushima, D. K., Hellman, L., & Gallagher, T. F. (1954). The transformation of cholesterol to coprostanol. *Journal of Biological Chemistry*, *211*(1), 301–311.

Shah, S. R., Mollenhauer, G., Ohkouchi, N., Eglinton, T. I., & Pearson, A. (2008). Origins of archaeal tetraether lipids in sediments: Insights from radiocarbon analysis. *Geochimica et Cosmochimica Acta*, *72*(18), 4577–4594.

Shah, V. G., Hugh Dunstan, R., Geary, P. M., Coombes, P., Roberts, T. K., & Von Nagy-Felsobuki, E. (2007). Evaluating potential applications of faecal sterols in distinguishing sources of faecal contamination from mixed faecal samples. *Water Research*, *41*(16), 3691–3700.

Shamsudduha, M., Uddin, A., Saunders, J. A., & Lee, M. K. (2008). Quaternary stratigraphy, sediment characteristics and geochemistry of arsenic-contaminated alluvial aquifers in the Ganges-Brahmaputra floodplain in central Bangladesh. *Journal of Contaminant Hydrology*, *99*(1–4), 112–136.

Shi, L., Squier, T. C., Zachara, J. M., & Fredrickson, J. K. (2007, July 1). Respiration of metal (hydr)oxides by Shewanella and Geobacter: A key role for multihaem c-type cytochromes. *Molecular Microbiology*. John Wiley & Sons, Ltd (10.1111).

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology*.

Simkus, D. N., Slater, G. F., Sherwood Lollar, B., Wilkie, K., Kieft, T. L., Magnabosco, C., … Onstott, T. C. (2016). Variations in microbial carbon sources and cycling in the deep continental subsurface. *Geochimica et Cosmochimica Acta*, *173*, 264–283.

Simoneit, B. R. T. (2004). Biomarkers (molecular fossils) as geochemical indicators of life. *Advances in Space Research*, *33*(8), 1255–1261.

Slater, G. F., Lippmann-Pipke, J., Moser, D. P., Reddy, C. M., Onstott, T. C., Lacrampe-Couloume, G., & Lollar, B. S. (2006). 14C in Methane and DIC in the Deep Terrestrial Subsurface: Implications for Microbial Methanogenesis. *Geomicrobiology Journal*, *23*, 453–462.

Slater, G. F., White, H. K., Eglinton, T. I., & Reddy, C. M. (2005). Determination of microbial carbon sources in petroleum contaminated sediments using molecular 14C analysis.

Environmental Science and Technology, *39*(8), 2552–2558.

Smedley, P. ., & Kinniburgh, D. . (2002). A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry*, *17*(5), 517–568.

Speranza, E. D., Colombo, M., Skorupka, C. N., & Colombo, J. C. (2018). Early diagenetic alterations of sterol biomarkers during particle settling and burial in polluted and pristine areas of the Rio de la Plata Basin. *Organic Geochemistry*, *117*, 1–11.

Stuiver, M, & Polach, H. (1977). Discussion reporting of 14C data. *Radiocarbon*, *19*(3), 355–363.

Stuiver, Minze, Pearson, G. W., Becker, B., Vogel, J. C., Fuls, A., Visser, E., … Dehling, H. (1993). Radiocarbon. *Radiocarbon*, *35*(1), 1–255.

Stute, M., Zheng, Y., Schlosser, P., Horneman, A., Dhar, R. K., Datta, S., … Van Geen, A. (2007). Hydrological control of As concentrations in Bangladesh groundwater. *Water Resources Research*, *43*(9).

Van Geen, A., Ahmed, E. B., Pitcher, L., Mey, J. L., Ahsan, H., Graziano, J. H., & Ahmed, K. M. (2014). Comparison of two blanket surveys of arsenic in tubewells conducted 12years apart in a 25km2 area of Bangladesh. *Science of the Total Environment*, *488–489*(1), 484–492.

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 and Technology*, *42*(7), 2283–2288.

van Geen, A., Zheng, Y., Versteeg, R., Stute, M., Horneman, A., Dhar, R., … Ahmed, K. M. (2003). Spatial variability of arsenic in 6000 tube wells in a 25 km2 area of Bangladesh. *Water Resources Research*, *39*(5).

Vane, C. H., Kim, A. W., McGowan, S., Leng, M. J., Heaton, T. H. E., Kendrick, C. P., … Swann, G. E. A. (2010). Sedimentary records of sewage pollution using faecal markers in contrasting peri-urban shallow lakes. *Science of the Total Environment*, *409*(2), 345–356.

Vestal, J. R., & White, D. C. (1989). Lipid Analysis in Microbial Ecology. *BioScience*, *39*(8), 535–541.

Volkman, J. K., Farrington, J. W., & Gagosian, R. B. (1987). Marine and terrigenous lipids in coastal sediments from the Peru upwelling region at 15°S: Sterols and triterpene alcohols. *Organic Geochemistry*, *11*(6), 463–477.

Whaley-Martin, K. J., Mailloux, B. J., van Geen, A., Bostick, B. C., Ahmed, K. M., Choudhury, I., & Slater, G. F. (2017). Human and livestock waste as a reduced carbon source contributing to the release of arsenic to shallow Bangladesh groundwater. *Science of the Total Environment*, *595*, 63–71.

Whaley-Martin, K. J., Mailloux, B. J., Van Geen, A., Bostick, B. C., Silvern, R. F., Kim, C., … Slater, G. F. (2016). Stimulation of Microbially Mediated Arsenic Release in Bangladesh Aquifers

by Young Carbon Indicated by Radiocarbon Analysis of Sedimentary Bacterial Lipids. *Environmental Science and Technology*, *50*(14), 7353–7363.

Yu, K., Gan, Y., Zhou, A., Liu, C., Duan, Y., Han, L., & Zhang, Y. (2018). Organic carbon sources and controlling processes on aquifer arsenic cycling in the Jianghan Plain, central China. *Chemosphere*, *208*, 773–781.

Zelles, L. (1999). Fatty acid patterns of phospholipids and lipopolysaccharides in the characterization of microbial communities in soil: A review. *Biology and Fertility of Soils*, (29), 111–129.

Zelles, L., Bai, Q. Y., Beck, T., & Beese, F. (1992). Signature fatty acids in phospholipids and lipopolysaccharides as indicators of microbial biomass and community structure in agricultural soils. *Soil Biology and Biochemistry*, *24*(4), 317–323.

Zelles, L., Bai, Q. Y., Ma, R. X., Rackwitz, R., Winter, K., & Beese, F. (1994). Microbial biomass, metabolic activity and nutritional status determined from fatty acid patterns and poly-hydroxybutyrate in agriculturally-managed soils. *Soil Biology and Biochemistry*, *26*(4), 439–446.

Zhang, Y. M., & Rock, C. O. (2008, March 1). Membrane lipid homeostasis in bacteria. *Nature Reviews Microbiology*. Nature Publishing Group. https://doi.org/10.1038/nrmicro1839

Zheng, Y., Stute, M., van Geen, A., Gavrieli, I., Dhar, R., Simpson, H. J., … Ahmed, K. M. (2004). Redox control of arsenic mobilization in Bangladesh groundwater. *Applied Geochemistry*, *19*(2), 201–214.

Zheng, Yan, van Geen, A., Stute, M., Dhar, R., Mo, Z., Cheng, Z., … Ahmed, K. M. (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.

Ziolkowski, L. A., Wierzchos, J., Davila, A. F., & Slater, G. F. (2013). Radiocarbon Evidence of Active Endolithic Microbial Communities in the Hyperarid Core of the Atacama Desert. *Astrobiology*.

Chapter 2: Exploring regional-scale microbial carbon cycling and sewage inputs at two sites (Desert Island, Doper Tek) in Araihazar aquifers using phospholipid and sterol biomarkers

San Pedro, R.J.[1], Whaley-Martin, K.J.[2], Mailloux, B.J.[3], van Geen, A.[4], Bostick, B.C.[4], Ahmed, K. M.[5], Mozumder, R.[4], Ellis, T.[4], Slater, G.F.[1]

[1]School of Geography and Earth Sciences, McMaster University, ON, Canada
[2]Department of Civil and Mineral Engineering, University of Toronto, ON, Canada
[3]Environmental Sciences Department, Barnard College, NY, NY 10027
[4]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964
[5]Department of Geology, University of Dhaka, Dhaka 1000, Bangladesh

2.1 Abstract

Over 40 million Bangladeshi rely on groundwater with arsenic (As) concentrations often exceeding the Bangladesh water quality guideline (50 μg/L). This geogenic As is released into groundwater by dissimilatory iron-reducing microbes oxidizing surrounding organic carbon. In order to assess relationships between dissolved arsenic and microbial carbon cycling enabling arsenic release, molecular PLFA and sterol biomarkers were characterized at two aquifer sites in Araihazar, Bangladesh. Desert Island (DI) with light grey reduced Holocene sediments bearing high As concentrations (37-188 μg/L), and Doper Tek (DT) with oxidized orange clay-rich Pleistocene sediment with low As (avg 7 μg/L). Average total PLFA concentrations were used to estimate microbial community abundance which were found to range from $6.4x10^7$ to $1.8x10^8$ cells/g dry sediment at DI and from $2.1x10^7$ to $4.5x10^7$ cells/g dry sediment. The microbial abundances were at comparable between DT and nearby sites (Site B, F), although DI has a much greater biomass suggesting an overall elevated total microbial metabolism at this site. The sewage input potential was assessed using comparative ratios of sterol biomarkers. The sewage input proxies, Sewage Contamination Index and Coprostanol/Cholesterol, estimated the relative abundance of sewage derived Coprostanol to other sources of natural biogenic sterols (e.g. phytosterols from plants). Total bulk sterol concentrations were at least ten times greater at both DI (range = 35 ± 3.6 to 1300 ± 210 ng/g dry sediment) and DT (range = 86 ± 9.7 to 150.0 ± 19 ng/g dry sediment) compared to previous sites (Site B, F), although sewage input proxies indicated little sewage contamination at both DI and DT. This suggests a greater total available organic carbon input into these present sites likely sourced from plant organic matter due to the greater proportion of phytosterols in the total sterol content. The visual community groundwater well density was used as a visual proxy for human population density; previous sites where sewage contamination correlated with high arsenic had a visually greater human density compared to either DI or DT. This can potentially explain why there is comparatively little sewage contamination despite a higher total organic sterol input at DI and DT.

2.2 Introduction

The public health crisis of regional groundwater arsenic (As) contamination in Bangladesh and many parts of Southeast Asia continues to be a complex problem for national policy makers. Increasingly, groundwater aquifers across Bangladesh are found to exceed water quality guidelines set by the World Health Organization (<10µg As/L) or the Bangladesh government (<50µg As/L) (BGS & DPHE, 2001; van Geen et al., 2003; World Health Organization, 2011). A large source of this arsenic is geogenic (previously sorbed into aquifer sediments or otherwise immobilized) which is thought to be released into groundwater via the reductive dissolution of sedimentary Fe (oxy)hydroxides (Héry et al., 2010; Islam et al., 2004; Legg et al., 2012; Mailloux et al., 2013). Microorganisms such as metal reducers have been shown to be at least partially responsible for this release mechanism by coupling the reduction of electron acceptors such as solid Fe to the oxidation of reduced electron donors (e.g. organic substrates) (BGS & DPHE, 2001; Ratan K Dhar et al., 2011; Islam et al., 2004) as shown by the example in equation [1],

$$CH_2O + FeOOH_{(s)} + H_2O \rightarrow Fe(II)_{(aq)} + CO_2 + 2H_2O + H^+ \quad [1]$$

In equation [1], arsenic would be associated with the solid Fe (oxy)hydroxide and subsequently released upon the reduction and dissolution of Fe(II) into groundwater. The CH_2O in equation [1] represents any reduced organic carbon acting as an electron donor. The availability of reduced organic matter is therefore critical in driving the redox reaction.

Previous research has proposed both sedimentary organic carbon (SOC) and dissolved organic carbon (DOC) as potential sources for the organic matter driving iron reduction and

thus, As release. SOC in the shallow Holocene aquifers include peat preserved from burial and diagenesis of mangrove, marshland, and swamp plant matter (Shamsudduha et al., 2008). Indigenous microorganisms within Bangladesh aquifer have been shown to be capable of respiring sedimentary carbon (e.g. peat) to facilitate reductive dissolution of Fe(oxy)hydroxides and release As into groundwater (McArthur et al., 2004; Pracht et al., 2018). Alternatively, labile DOC can be derived from dissolution of *in situ* SOC or can be transported to depth from surface sources such as terrestrial plants and soil (Mladenov et al., 2010), excavated ponds (Neumann et al., 2010), and shallow pit latrines (Ahmed et al., 1994; Knappett et al., 2011; McArthur et al., 2012; Whaley-Martin et al., 2017). Indeed, the addition of labile carbon (e.g. DOC) has been demonstrated to stimulate microbially mediated reductive As release resulting in groundwater contamination in Bangladesh, Vietnam, and Cambodia (Ratan K Dhar et al., 2011; Harvey et al., 2002, 2006; Héry et al., 2010; Islam et al., 2004; Kocar et al., 2008; Mailloux et al., 2013; Richards et al., 2019; Whaley-Martin et al., 2016). Previous research conducted at Site B and F in Araihazar used natural abundance radiocarbon analysis ($\Delta^{14}C$) to show that the microbial communities within the aquifer systems were using young carbon sources at these sites (Mailloux et al., 2013; Whaley-Martin et al., 2016). Such young radiocarbon signatures points to organic matter which has been recently synthesized from modern carbon sources at the surface. This supports the hypothesis that addition of surface-derived organic carbon drives microbial metabolism and concurrent arsenic release at depth (Harvey et al., 2002, 2006; Mailloux et al., 2013; Richards et al., 2019; Whaley-Martin et al., 2017; Yu et al., 2018). The cycling of young carbon at these sites implies that SOC, bearing a much older radiocarbon

signature, is not a primary carbon substrate consistent with the expectation that recalcitrant sedimentary carbon (e.g. peat) are less bioavailable to microorganisms in the aquifer.

Molecular source-specific biomarkers were quantified in order to assess microbial cycling of carbon and potential for sewage input in the aquifer (Eglinton et al., 1964; Eglinton & Calvin, 1967; Simoneit, 2004). Phospholipid fatty acids (PLFA) are integral cell membrane components of microorganisms representing viable microbial biomass. Similarly, organic sterols are produced by plants and animals and are used here to assess likely sources of organic input into the aquifer. Phospholipid fatty acids (PLFA) are bonded to a polar phosphate group by an ester link and make up a significant component of viable microbial cells. Because cell membrane lipids such as PLFA rapidly degrade during lysis (i.e. cell death), their quantification in aquifer environments provides a relatively instantaneous measure of the abundance of viable or extant microbial communities operating in the aquifer (Vestal & White, 1989; D. White et al., 1996; D. C. White et al., 1979). Bulk microbial cell abundance can be estimated from PLFA concentrations using a direct cell conversion factor $2x10^4$ cells/pmol PLFA (C. T. Green & Scow, 2000; D. C. White et al., 1979). The distribution or profiles of PLFA in environmental samples can also delineate whether changes in the microbial community has occurred as a result of shifts in environmental factors such as temperature, pH, or nutrients (Coleman et al., 1993; Ludvigsen et al., 1999). Compound specific natural abundance radiocarbon isotope $\Delta^{14}C$ of organic matter in the environment can elucidate the age of carbon substrates utilized for their biosynthesis enabling researchers to trace the biogeochemical cycling of diverse carbon pools through different environmental systems (Mailloux et al., 2013; Petsch et al., 2001; Slater et al., 2005; Whaley-Martin et al., 2016).

The potential for unsewered waste input in the aquifer at DI and DT was analyzed using sedimentary organic sterol biomarkers. Specifically, Coprostanol, Cholesterol, Cholestanol, Campesterol, Stigmasterol, and β-Sitosterol were quantified to represent the non-exhaustive bulk organic sterols at DI and DT. Organic sterols in the natural environment are rapidly degraded at the surface where oxic conditions prevail but are better preserved in anoxic environments such as the saturated zone of an aquifer (Pratt et al., 2008). The preservation of organic sterols in anoxic and oligotrophic sedimentary aquifers such as those found in Araihazar, can be attributed to their high partitioning coefficients (K_{ow}) which leads to preferential sorption onto aquifer sediments and subsequent retardation with respect to groundwater flow into the aquifer (Chiou et al., 1982; Froehner & Sánez, 2013; Hansch et al., 1968; Whaley-Martin et al., 2017). Nevertheless, the source-specificity and resistance to degradation of organic sterols make them useful tools to assess important sources of organic inputs as well as sewage contamination in the environment (Bull et al., 2002; Huang & Meinschein, 1979; Leeming & Nichols, 1998; Nishimura & Koyama, 1977). Of the six sterol biomarkers measured here, Coprostanol and Cholesterol are two of the most abundant sterols found in the faeces of humans and other upper mammals such as cows (Isobe et al., 2002; S. R. Shah et al., 2008). For example, Isobe et al. (2002) report that Coprostanol and Cholesterol were present at comparable proportions (~15-16%) in the faecal sterol profiles of cows. Coprostanol, however, is not exclusively produced by upper mammals; it can also be produced by aquatic zooplankton and macrophytes (Nishimura & Koyama, 1977) and have been reported as a diagenetic by-product of the breakdown of Cholesterol in anoxic environments (G. Green et al., 1992; Nishimura, 1982). Coprostanol (5β-cholestanol) is produced by the anaerobic

degradation of Cholesterol in the digestive tract of upper mammals as mediated by enteric bacteria; Coprostanol makes up approximately 60% of the total sterol content of human faeces (Bull et al., 2002; Leeming et al., 1996; Leeming & Nichols, 1998). Previous studies have reported absolute Coprostanol concentrations in the μg/g to mg/g scale at sites contaminated by human sewage (Carreira et al., 2004; Dsikowitzky et al., 2017; He et al., 2018; Isobe et al., 2002; Kolpin et al., 2000; Speranza et al., 2018; Vane et al., 2010). Cholestanol (5α-Cholestanol), an isomer of Coprostanol, is also be produced via the biohydrogenation of Cholesterol in the natural environment. 5α-Cholestanol is reported to be the dominant isomer in environments uncontaminated by sewage because its production does not involve enteric bacteria (Bull et al., 2002; Gaskell & Eglinton, 1975; Grimalt et al., 1990; Huang & Meinschein, 1979; Nishimura & Koyama, 1977; Speranza et al., 2018). For example, Huang & Meinschein (1979) report Cholestanol relative abundances of 19.4 – 25.6% in the sediment profile of a saline lagoon environment. The other sterols used in this study – Campesterol, Stigmasterol, and β-Sitosterol – are phytosterols commonly derived from terrestrial vascular plants, fungi, algae, and zooplankton (Huang & Meinschein, 1979; Liu et al., 2016; Pearson et al., 2001). In this study, these phytosterols, along with Cholestanol, are used to represent other biogenic sources naturally occurring in pristine environments (e.g. not impacted by sewage).

Since the detection of Coprostanol and Cholesterol does not unambiguously indicate sewage contamination in the environment, supporting proxies were used to verify the relative importance of sewage and other biogenic sources (e.g. plants, fungi) into the organic carbon pool. For example, the Sewage Contamination Index (SCI) is calculated using equation [2] and estimates the relative importance of sewage-derived coprostanol to the other by-product of

Cholesterol biohydrogenation in the aquifer (Grimalt et al., 1990; He et al., 2018; Martins et al., 2011; Reeves & Patton, 2005; Vane et al., 2010; Whaley-Martin et al., 2017).

$$\frac{[5\beta-coprostanol]}{[5\beta-coprostanol]+[5\alpha-cholestanol]} \quad [2]$$

$$\frac{[coprostanol]}{[cholesterol]} \quad [3]$$

The Sewage Contamination Index (SCI) value of 0.7 represents the threshold above which contamination of human or animal waste is considered significant (i.e. sewage-derived coprostanol makes an important contribution to the bulk organic sterol pool) (Dsikowitzky et al., 2017; Grimalt et al., 1990; He et al., 2018; Reeves & Patton, 2005; V. G. Shah et al., 2007). At previous sites in Araihazar, the SCI were reported to be >0.7 at depth and were determined to correlate with dissolved arsenic concentrations implying that sewage may provide organic input driving reductive dissolution of Fe(oxy)hydroxides and arsenic release (Whaley-Martin et al., 2017). Alternatively, the Coprostanol/Cholesterol (Cop/Ch) ratio, calculated using equation [3], compares the prevalence of sewage to other natural biogenic sources of sterols (e.g. planktonic, macrophytes). Coprostanol/Cholesterol ratios <1 signify that other biogenic sources make a greater contribution to the sterol pool, whereas Cop/Ch ratios >1 indicate the dominance of sewage as a sterol source (Fattore et al., 1996). The shift in the Cop/Ch ratio has been used to investigate sewage as a component of urban effluent into surface water bodies (e.g. estuaries, rivers) (Frena et al., 2016; Matić Bujagić et al., 2016; Reeves & Patton, 2005). Put simply, the SCI and Cop/Ch are two sides of the same coin – they provide corroborating

evidence to ascertain the sources of organic sterols and helps to determine the relative importance of sewage input in the environment.

The goal of this study was to expand the assessment of carbon sources supporting microbial communities in Bangladesh aquifers by investigating two sites, Desert Island (DI) and Doper Tek (DT). Molecular biomarkers, PLFA and organic sterols, were used to quantify viable *in situ* microbial biomass and assess the potential of sewage input to drive high arsenic concentrations into the aquifer. Concurrently, natural abundance radiocarbon isotope analysis of PLFA served to determine the age of the dominant carbon source cycled in these aquifers. To parallel previous studies which found positive correlation between indicators of sewage and dissolved arsenic at depth, Coprostanol/(Cholestanol+Coprostanol) and Coprostanol/Cholesterol ratios were calculated along with proportions of phytosterols to distinguish the relative importance of different sources of organic sterols in these aquifers. Furthermore, differences in local population densities between DI and DT as well as previous sites are assessed to determine potential direct relationships between local population densities and sewage input into the aquifer.

2.3 Study Sites

The two sites studied are located in Araihazar upazila, Bangladesh: Desert Island (23.770 N, 90.609 E) and Doper Tek (23.864 N, 90.653 E) (Figure S1). Desert Island (DI) is located in the middle of an agricultural field approximately 500 m away from dense human population; it represents a low local population density proxy site. Doper Tek (DT), located at the corner of a dense human population community, is used as a high local population density proxy. DI and DT

are located ~11 km from each other and, at most, ~11 km from previously studied sites A (R K Dhar et al., 2008; Zheng et al., 2005), B, F, M, N, O, S, and T (R K Dhar et al., 2008; van Geen et al., 2003; Mailloux et al., 2013; Stute et al., 2007; Whaley-Martin et al., 2016, 2017; Zheng et al., 2005). The generally flat topography in Araihazar is characterized by low hydraulic gradients resulting in slow groundwater flow velocities (BGS and DPHE, 2001; Burgess & Ahmed, 2006). The aquifer at Desert Island (DI) is Holocene in age (6970 ± 45 ^{14}C yr at 40 m and 7180 ± ^{14}C yr at 44 m depth) (Radloff et al., 2007) while Doper Tek (DT) is Pleistocene in age based on sedimentary organic carbon (SOC) dated at 19,300 ^{14}C yr (8 m) and 18,450 ^{14}C yr (13 m) (Table S4). The lithostratigraphy of DI is characterized by medium to fine sand with a clay layer at 1.5-3 m while DT is dominated by fine sand with a medium sand layer at 11-12 m and clay layers at 1.5-3 m and 14 m (Table S1). The gray sediment at DI indicates reducing geochemical conditions whereas the orange sediments at DT indicates oxidizing conditions (Figure S1); this is confirmed using diffuse spectral reflectance measurements correlating the relative abundance of Fe(II)/Fe and the extent of reduction of sedimentary aquifers (Horneman et al., 2004; Radloff et al., 2007; Zheng et al., 2005). The DI aquifer has lower solid and higher dissolved arsenic concentrations compared to DT (Table S1). The high solid phase As concentrations occur within the sand-dominated depths at both DI and DT and are greater than the average crustal concentration of 2 mg/kg (Zheng et al., 2005). The mean groundwater arsenic concentration at DT was 7 ± 0.4 µg/L and 90 ± 5 µg/L at DI (Figure S1). Arsenic concentrations reported at Site B (100-530 ug/L) and F (5-200 ug/L) were much higher than either DI or DT (Whaley-Martin et al., 2016). DI, because of its proximity to low population denisty, is hypothesized to receive low sewage input whereas DT, located adjacent to a dense cluster of community wells and thus

human population, is hypothesized to receive high sewage input (Figure S1). The contrasting dissolved arsenic concentrations, conventional radiocarbon age $\Delta^{14}C$, geochemical reducing conditions, and local human population densities between DI and DT are the characterizing biogeochemical parameters that may be interrelated to determine arsenic cycling in these complex aquifer environments. The apparent difference in human population density as represented by the density of domestic wells (Figure S1) imply that a greater contribution of sewage waste can be reasonably expected to enter the aquifer at DT where the human population is greater compared to DI.

Two sedimentary matrix cores were drilled at DI (4.6m – 23m depth) and DT (11m – 15m depth) using a gravity corer, frozen immediately, and shipped to McMaster University. The solid and aqueous arsenic concentrations as well as relative extent of reduction at the aquifers are used to hypothesize that the microbially mediation reduction is stimulated at DI facilitated by a greater organic matter input resulting in greater dissolved arsenic. These parameters are in contrast to observations at DT where we hypothesize that the same microbially mediated arsenic release reactions are suppressed in part as a result of a more oxidized aquifer.

2.4 Methodology

2.4.1 PLFA biomarker extractions

Microbial biomarkers (PLFA and sterols) were extracted from freeze-dried aquifer sediments using a modified Bligh & Dyer solvent extraction method similar to (Whaley-Martin et al., 2016). Briefly, a mixture of 1:2:0.8 dichloromethane (DCM), methanol, and phosphate buffer mixture were used to dissolve polar PLFA and sterol compounds from sediment samples

(Guckert et al., 1985; D. C. White et al., 1979); sediment samples were extracted for 24 hrs twice. After each extraction, the supernatant was collected and filtered to remove solids. The extract was transferred into separatory funnels and Methanol and DCM were added to separate phases and the organic phase containing the lipid biomarkers were collected as the total lipid extract (TLE). Silica gel column chromatography was used to separate the extractable organic compounds contained in the TLE into three fractions based on polarity (Guckert et al., 1985): F1 – neutral lipids (DCM), F2 – uncharged polar lipids (Acetone), F3 – very polar or charged lipids (Methanol) (Vestal & White, 1989). Phospholipids elute in F3 while sterols elute in F1 and F2. To analyze the fatty acid tail ends of phospholipids, F3 was concentrated and mild alkaline (KOH) methanolysis was performed to converting PLFA to fatty acid methyl esters (FAME) amenable for gas chromatograph-mass spectrometer (GC-MS) analysis (D. White et al., 1996). A secondary silica gel chromatography was used as a "clean up" step to further separate FAME compounds from other lipid components which eluted in F3 during the primary column chromatography (glycerols, archaeal lipids). The three fractions distinguished based on polarity are F1 (4:1 Hexane:DCM), F2 (DCM), and F3 (MeOH). The FAME samples, eluted in F3, were concentrated to 1mL samples prior to GC-MS analysis.

PLFA and sterol biomarkers were quantified using an Agilent 6890N Gas Chromatograph (30m × 0.25mm DB5-ms DG, 0.25µm film thickness) coupled to a 5973 quadrupole mass spectrometer monitoring for masses (50-450 m/z for FAME and 50-500 m/z for TMS-sterols). For the PLFA analysis, 1µL sample injections were analyzed at an initial temperature of 50^{o}C held for 1 min then increased up to 130^{o}C at a rate of 20^{o}C/min then a further increase to 160^{o}C at a rate of 4^{o}C/min and finally a final ramp at a rate of 8^{o}C/min until a final maximum

temperature of 300^{o}C was achieved and held for 5 min. Fatty acid compounds were identified using two FAME mixture standards (26 component Sigma-Aldrich, and 37 component Supelco Inc.).

2.4.2 Natural abundance radiocarbon analysis of FAME

Methylated PLFA derivatized into fatty acid methyl ester (FAMEs) were concentrated in sample to 500µL and analyzed for compound specific natural abundance radiocarbon isotopic composition ($\Delta^{14}C$) at the National Ocean Sciences Accelerator Mass Spectrometer (NOSAMS) laboratory at the Woods Hole Oceanographic Institute (Woods Hole, Massachusetts, USA). Oxalic Acid II and Vienna Pee Dee Belemnite (VPDB) standards for Carbon were used and a conservative error of ±20‰ was included to accommodate errors in analyzing samples containing <150 µg Carbon using an Accelerator Mass Spectrometer (AMS) (Pearson et al., 1998, 2001).

2.4.3 Organic sterol biomarker extraction

To analyze sterols, the column chromatography fraction F2 was evaporated to dryness under $N_{2(g)}$ and redissolved in DCM then combined with F1 (Birk et al., 2012; Whaley-Martin et al., 2017). A trimethylsilyl (TMS) derivatization reaction was performed to convert intact sterols to TMS-derivatives amenable for GC-MS analysis. The samples were first concentrated to 100uL, 30uL of the derivatizing agent N,O-Bis(trimethylsilyl)trifluoro acetamide (BSTFA) (Sigma Aldrich) was then added along with 10uL of pyridine as a catalyst. Samples were heated at 60^{o}C for 30 minutes to ensure a complete derivatization reaction before analysis using GC-MS (Whaley-Martin et al., 2017).

The GC-MS program used to analyze the derivatized sterols was similar to that used to analyze PLFA. Specifically, 2µL injections where analyzed at an initial temperature of 50^{o}C held for 1 min then increased at 4^{o}C/min up to a final temperature of 300^{o}C which was held for 24 min. Sterol compounds were identified using Sigma-Aldrich sterol compound standards: 5β-cholestan-3β-ol (coprostanol), cholest-5-en-3β-ol (cholesterol), 5α-cholestan-3β-ol (cholestanol), 24-methylcholest-5-en-3β-ol (campesterol), 24-ethylcholesta-5,22(*E*)-dien-3β-ol (stigmasterol), 24-ethylcholest-5-en-3β-ol (β-sitosterol). PLFA and sterol sample compounds were identified using the manufactured standards in tandem with characteristic fragmentation patterns and mass spectral identification using the National Institute of Standards and Technology (NIST) compounds library. The limit of detection is 0.5 ng/µL – 25 ng/µL for all sterols. Because some samples were well above 25 ng/µL, concentrated samples were diluted within the limit of detection and reanalyzed in the GC-MS.

2.4.4 Dissolved groundwater arsenic measurements

Dissolved groundwater arsenic was measuring using a high resolution inductively coupled plasma mass spectrometer (HR ICP-MS) at the Lamont Doherty Earth Observatory (Columbia University) as per the methods of (Cheng et al., 2004). Lower detection limits were <0.1 µg/L.

2.5 Results

2.5.1 Microbial community abundance

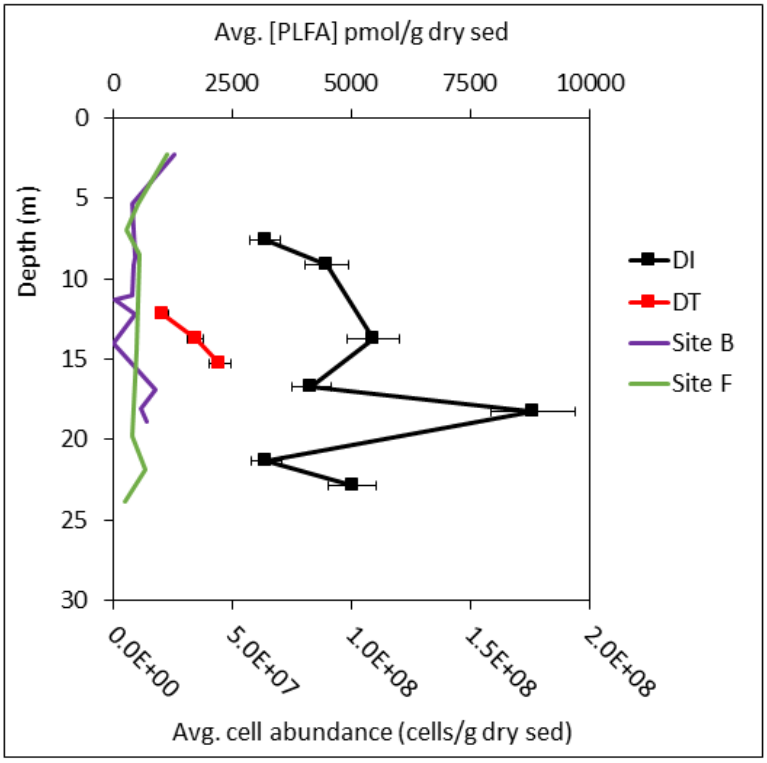


Figure 1*. Depth profile of average total PLFA concentrations (upper x-axis) and estimated microbial abundance (lower x-axis) at DI and DT compared to previous sites (B, F).*

PLFA concentrations at Doper Tek (DT) increased from 1100 pmol/g dry sediment (gds) at 11 m depth to 2200 pmol/gds at 15 m depth (mean 1700 ± 170). These are equivalent to estimated cellular abundances of 2.1×10^7 (11 m) to 4.5×10^7 (15 m) cells/gds using a cell conversion factor of 2×10^4 cells/pmol PLFA (C. T. Green & Scow, 2000) (Figure 1). While PLFA concentrations increased with depth at DT, the fact that only three samples were obtained means it is difficult to assess the overall trend. PLFA concentrations at Desert Island (DI) were higher than DT. At DI, PLFA concentrations initially increased from 7.6 m to 15 m (3200 to 5500 pmol PLFA/gds) followed by a slight decrease at 17 m (4100 pmol PLFA/gds) then a rapid

increase to the maximum PLFA concentration at 18 m (8800 pmol PLFA/gds), and followed by an immediate decrease in PLFA concentrations comparable to those between 7.6 m and 15 m (Figure 1). The PLFA range at DI is equivalent to estimated cell abundances of minimum 6.4×10^7 to a maximum 1.8×10^8 cells/gds. In general, PLFA concentrations at DI and DT were approximately 10-100 times greater than the results of previous research at Site B which ranged from 17-1300 pmol/gds and about 10 times greater than observed at Site F which ranged from 260-1100 pmol/gds (Whaley-Martin et al., 2016).

While the total PLFA concentrations were different between DI and DT, the PLFA profile distributions were similar between the sites (Figure S2). The presence and relative abundance (mol%) of each major PLFA group remained constant with depth at both DI where the average mol% were 37.9% (saturated), 32.6% (branched saturated), 19.2% (monounsaturated), 6.6% (cyclic) and 2.8% (polyunsaturated) and DT where the average mol% were 37.6% (branched saturated), 33.1% (saturated), 21.7% (monounsaturated), 2.5% (polyunsaturated), and 2.1% (cyclic) (Figure S2).

2.5.2 Radiocarbon signatures of PLFA and carbon pools

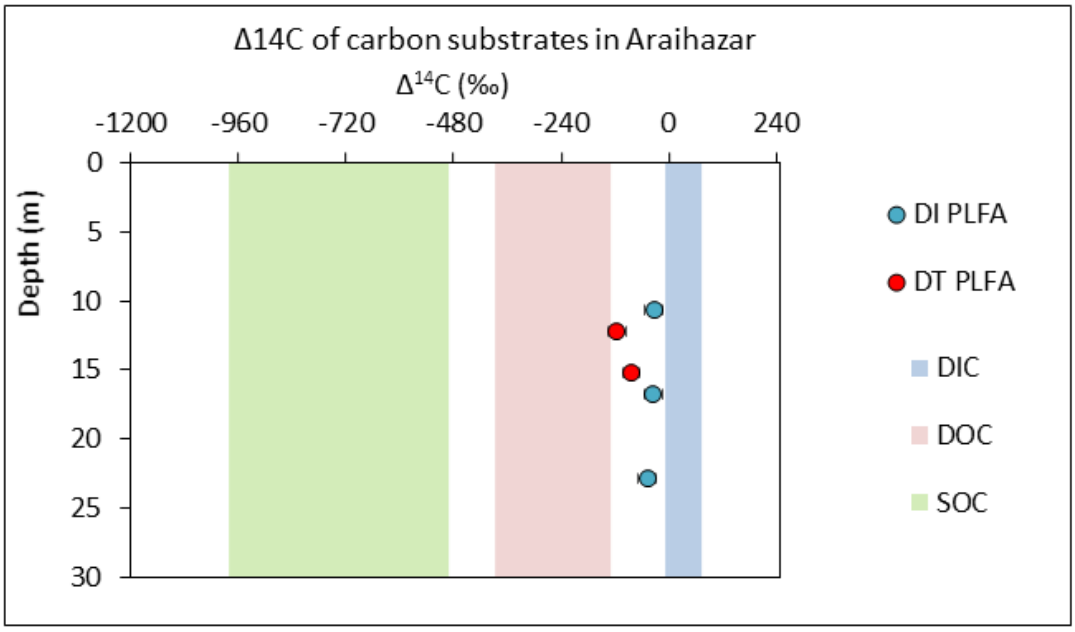


Figure 2. *Radiocarbon signature values of PLFA and representative carbon substrate pools available in Araihazar, Bangladesh: DIC, DOC, and SOC (carbon pools data modified from Mailloux et al. 2013; Whaley-Martin et al. 2016).*

By comparing the $\Delta^{14}C$ signatures of PLFA with the bulk carbon pools available in the aquifer (i.e. DIC, DOC, and SOC), the age of the dominant carbon substrate incorporated during cell membrane biosynthesis can be constrained. Analysis of natural abundance radiocarbon isotope ($\Delta^{14}C$) of PLFA demonstrated that the microbial communities in these aquifers were actively metabolizing young carbon sources. The radiocarbon signatures of bulk PLFA ($\Delta^{14}C_{PLFA}$) were slightly more enriched at DI ranging from +9 ‰ to +25 ‰ (±20‰) compared to DT which ranged from -63‰ to -30‰ (±20‰) (Figure 2, Table S4). These $\Delta^{14}C_{PLFA}$ signatures are comparable to PLFA from Site B which ranged from -163‰ to +21‰ (Whaley-Martin et al., 2016) and Site F which ranged from -167‰ to +20‰ (Whaley-Martin et al., 2016; Mailloux et al., 2013). While no site specific Δ14C data is available for DIC and DOC at DI and DT, previous research has reported ranges for the bulk carbon pools in Araihazar: dissolved inorganic carbon or DIC ($\Delta^{14}C_{DIC}$ = -8.2‰ to 75‰); dissolved organic carbon or DOC ($\Delta^{14}C_{DOC}$ = -379‰ to -131‰);

and sedimentary organic carbon or SOC ($\Delta^{14}C_{SOC}$ = -910‰ to -482‰) (Mailloux et al., 2013; Whaley-Martin et al., 2016). When the $\Delta^{14}C$ values of SOC, DOC, and DIC in Araihazar are compared to PLFA at DI and DT, the $\Delta^{14}C_{PLFA}$ at DI and DT were slightly more depleted relative to $\Delta^{14}C_{DIC}$ and slightly more enriched relative to $\Delta^{14}C_{DOC}$ (Figure 2). Bulk sedimentary organic carbon ($\Delta^{14}C_{SOC}$) at DT were -910 ± 20‰ at 7.6 m and -900 ± 20‰ at 13 m which are more depleted than $\Delta^{14}C_{SOC}$ from Site B and F (-675 ± 140‰). However, the $\Delta^{14}C_{SOC}$ values at DT are comparable to previously reports at Site M, S, and T but slightly more depleted compared to Site N and O (Whaley-Martin et al., 2016). Sediment samples from DI are presently under analysis for their $\Delta^{14}C_{SOC}$.

2.5.3 Organic sterol abundance

Total organic sterol concentrations, the sum of each individual sterol biomarker measured, ranged from 35 ± 3.6 to 1300 ± 210 ng/g dry sediment (gds) at DI and from 86 ± 9.7 to 150.0 ± 19 ng/gds at DT. These ranges are at least 10^1 greater than previous sites B (0.1 ± 0.0 to 5.2 ± 0.5 ng/gds) and F (0.1 ± 0.0 to 3.4 ± 0.6 ng/gds) (Whaley-Martin et al. 2017). In general, at both DI and DT, Coprostanol was the least abundant sterol measured followed in order of increasing abundance by Cholestanol, Cholesterol, Campesterol, β-Sitosterol, and Stigmasterol (Figure 3). The three phytosterols, Campesterol, β-Sitosterol, and Stigmasterol, were consistently the three most abundant sterols at all depths at both sites. At DI, all sterols were initially very low at the shallowest depth (7.6 m) and suddenly increased at 9.1 m (Figure 3). Sterols remained relatively constant throughout the middle of the core until another sudden

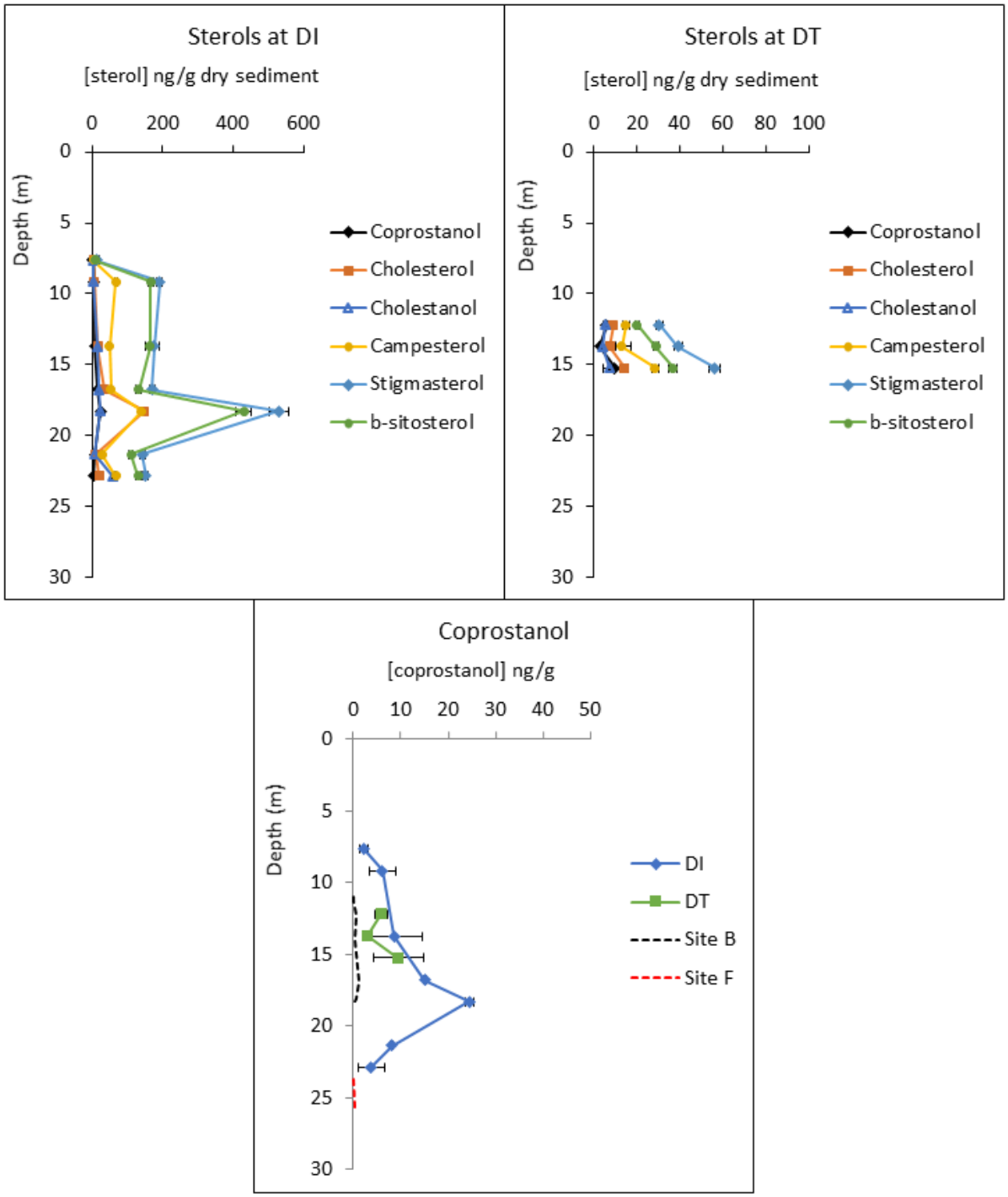


Figure 3*. Average concentrations of sterol biomarkers, Coprostanol, Cholesterol, Cholestanol, Campesterol, Stigmasterol, β-Sitosterol at Desert Island (DI, upper left) and Doper Tek (DT, upper right, x-axis scale 6x <DI). Comparison of Coprostanol concentrations (bottom) at Desert Island (DI), Doper Tek (DT), and previous sites (Site B, F).*

increase at 18 m where all sterols reach maximum concentrations. On the other hand, at DT,

Coprostanol, Cholestanol, Cholesterol, and Campesterol were effectively constant from 11 m to

12 m and were followed by a rapid increase at 15 m. By contrast, Stigmasterol and β-Sitosterol appeared to increase with depth from 11 m to 15 m. The stark elevated concentrations of all sterols at 18 m at DI and 15 m at DT reflected depths where PLFA concentrations were also elevated at these sites.

Compared to previous sites B and F, the two sterols commonly associated with sewage (Coprostanol and Cholesterol) were generally 10 times greater at DI and DT while the non-enteric by-product of Cholesterol (Cholestanol) was 10^2 greater at DI and 10^1 greater at DT. All phytosterols at DI and DT were at least 10^1 and up to 10^3 greater compared to previous sites. Coprostanol constituted a lower proportion of total sterols at DI (range = 1-6%) and DT (3-7%) compared to Site B (<LOQ-67%) and F (47-63%) (Whaley-Martin et al. 2017). Cholesterol, the other biomarker abundant in sewage, constituted 1-13% of total sterols at DI, 8-11% at DT, <LOQ-29% at Site B, and <LOQ-100% at Site F. The relative abundance of both Coprostanol and Cholesterol were comparable to Cholestanol at DI (respectively, mean 2% and 7% vs. 4%) and DT (respectively, 6% and 9% vs. 5%) but were both 10^1 higher at Sites B (respectively, 25% and 26% vs. 4%) and F (respectively, 32% and 36% vs. 6%). At DI and DT, the phytosterols (Campesterol, Stigmasterol, and β-Sitosterol) were the most abundant sterols: between 69-96% of total sterols at DI and 76-85% at DT. These were, on average, greater than previous sites: 61% at Site B and 56% at Site F. At DI, Campesterol was 9-16%, Stigmasterol was 34-47%, and β-Sitosterol was 24-38% of total sterols. At DT, Campesterol was 14-19%, Stigmasterol was 34-41%, and β-Sitosterol was 23-30% of total sterols.

2.5.4 Sterol ratios as a measure of sewage contamination

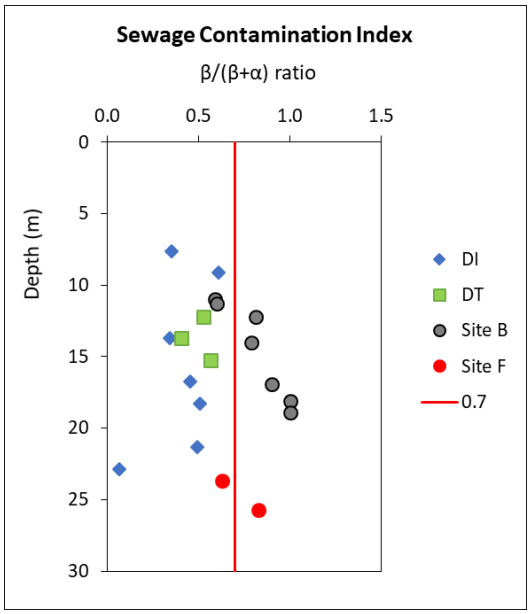


Figure 4*. Sewage Contamination Index profiles in Araihazar. Ratios >0.7 indicate significant sewage input. Site B and F from Whaley-Martin et al. 2017.*

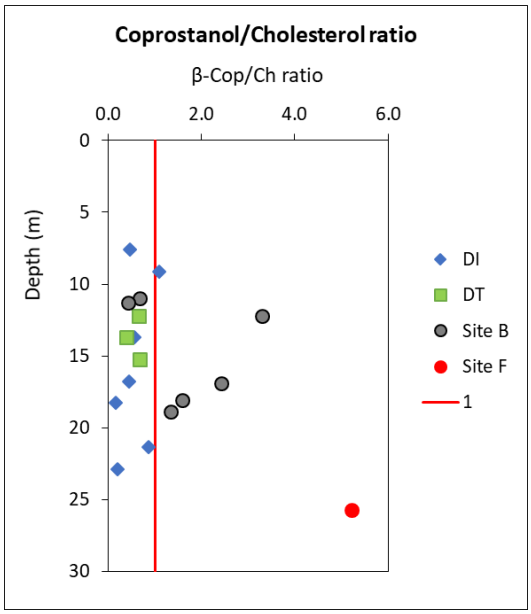


Figure 5. *Coprostanol to Cholesterol concentration ratios in Araihazar. Ratios >1 indicate significant sewage input. Site B and F data from Whaley-Martin et al. 2017.*

The Sewage Contamination Index (SCI) calculated using the Coprostanol/(Coprostanol+ Cholestanol) ratio, which estimates sewage input potential in the aquifer, ranged from 0.1 to 0.6 at DI and from 0.4 to 0.6 at DT (Figure 4). Along depth, these values are consistently lower than 0.7, the threshold value indicating significant sewage contamination in the environment. This argues that sewage was a less important source of organic sterols than other naturally derived biogenic sterols in the environment at either DI or DT. Contrastingly, reports from previous sites increasing SCI ratios with depth: from 0.6 (11m) to 1 (18.9m) at Site B, and 0.7 (23.7m) and 0.8 (25.7m) at Site F (Whaley-Martin et al., 2017). The alternative proxy for sewage input potential, Coprostanol/Cholesterol (Cop/Ch), corroborate the evidence presented by the SCI and confirms the importance of other biogenic sources of sterols other than sewage in Araihazar aquifers. Generally, Cop/Ch ratios were <1 (indicating low sewage contributions) at

both DI and DT except for a slight increase >1 at 9 m at DI (Figure 5). On the other hand, Cop/Ch ratio at both Site B and F are mostly >1 (high sewage input) despite lower absolute concentrations of the sewage biomarker, Coprostanol, compared to either DI or DT. These observations contradict the prevailing assumption that greater sewage input should translate to greater concentrations of Coprostanol, thus necessitating complementary sewage input indices such as the SCI and Cop/Ch ratio.

2.6 Discussion

2.6.1 Viable microbial biomass, carbon cycling, and arsenic release at DI and DT

The PLFA distributions at DI and DT (Figure S2) confirm that the microbial community structure present at these two aquifers were generally similar. Despite a large difference in concentrations, the relative abundance (in mol %) of major PLFA groups were not significantly different between DI and DT. To the extent that changes in PLFA distributions can explain changes in microbial community structure, this suggests that the overall microbial communities were similar and did not change significantly along depth at DI and DT. In order to accurately analyse and determine the taxonomic diversity of microbial species in these aquifers, more specific genetic-based analysis (e.g. identifying microbial DNA or RNA) is necessary. Such genetic-based analysis will enable the identification and relative abundance of specific bacteria and determine the capacity of *in situ* microbiota to enable carbon cycling and arsenic release into groundwater.

The differences in groundwater arsenic concentrations in the groundwater at DI and DT are likely a result of differences in both the initial sedimentary arsenic concentration and the

extent of microbial iron reduction triggering arsenic release. The low biomass and dissolved arsenic at DT (7 µg/L) is concurrent with orange (oxidized) sediments that indicate little arsenic release has occurred (McArthur et al., 2004). Additionally, sedimentary arsenic is higher here than at DI (DI = 3.2 mg/kg; DT = 6.7 mg/kg) and may represent a larger reservoir of arsenic immobilized due to the oxidized conditions of the aquifer. In contrast, DI has lower solid arsenic but elevated dissolved arsenic concentrations (37-188 µg/L). The grey and highly reducing aquifer at DI implies a greater extent of reductive dissolution has occurred and is likely driven by the greater microbial biomass relative to DT.

The positive $\Delta^{14}C_{PLFA}$ at DI and DT confirm that young carbon sources are actively metabolized in these aquifers consistent with reports from previous sites (Site B, F) where microbial communities incorporate predominantly DOC or DIC-aged (i.e. atmospheric $\Delta^{14}C$ values) components of DOC (Mailloux et al., 2013; Whaley-Martin et al., 2016). These results demonstrate that the microbial use of young carbon sources is occurring beyond the initial study site and may be generally true over a wider regional scale across Araihazar. This use of younger carbon is consistent with SOC being a recalcitrant source of organic carbon that is less bioavailable to the sediment-associated microbial community. These observations maintain that SOC in the aquifer does not drive microbially mediated iron reduction and arsenic release. The $\Delta^{14}C_{PLFA}$ signatures at DI plot closely to the bulk $\Delta^{14}C_{DIC}$ from sites B and F whereas $\Delta^{14}C_{PLFA}$ signatures at DT fall somewhere between bulk $\Delta^{14}C_{DIC}$ and $\Delta^{14}C_{DOC}$. While autotrophic fixation of DIC in these aquifers cannot be excluded, previous reports confirm that the microbial communities which operate in these oligotrophic aquifers in Araihazar are dominantly heterotrophic (Mailloux et al., 2013; Whaley-Martin et al., 2016). It is therefore more likely that

the dominant carbon source utilized is a modern component of DOC with $\Delta^{14}C$ signatures comparable to $\Delta^{14}C_{DIC}$. This is conceivable given the large diversity of organic compounds with different $\Delta^{14}C$ values making up the DOC pool. Previous research suggest that the modern radiocarbon signatures of these carbon sources confer surface origins; for example, as part of groundwater recharge or drawdown imposed by groundwater pumping (Harvey et al., 2002, 2006; Mailloux et al., 2013; Richards et al., 2019; Whaley-Martin et al., 2016, 2017).

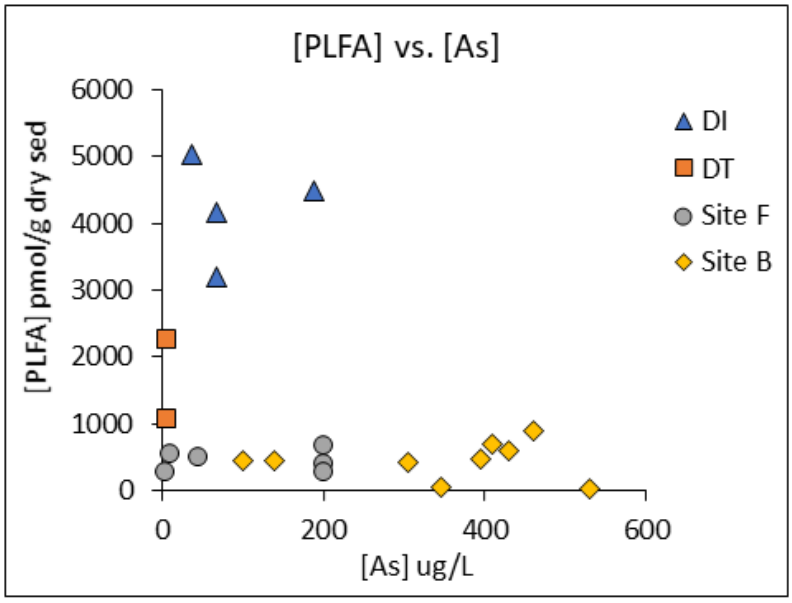


Figure 6. *Scatter plot comparing sedimentary PLFA concentration and dissolved groundwater Arsenic in Araihazar, Bangladesh showing weak correlation. Site B and F data from Whaley-Martin et al. (2016).*

Because of their high bioavailability, surface-derived carbon sources (e.g. raw domestic sewage, organic matter from ponds) have been implicated in stimulating in-aquifer microbially mediated reductive dissolution of iron oxides and concomitant arsenic release (Ghosh et al., 2015; Guo et al., 2019; Mailloux et al., 2013; Neumann et al., 2010; Richards et al., 2019; Whaley-Martin et al., 2016, 2017). However, the poor correlation between total PLFA and arsenic in Araihazar (Figure 6) contradicts the argument that a greater viable biomass, presumably from stimulated

microbial metabolism, would increase the extent of reductive dissolution of iron oxides thus, resulting in higher arsenic concentrations. The expectation of greater viable biomass resulting from greater microbial metabolism and subsequent increase in arsenic stems from bacterial culture-based experiments which showed that metal (Fe(III)) reducing bacteria found in aquifer sediments in Bangladesh can be stimulated to cause arsenic release (Ratan K Dhar et al., 2011; Islam et al., 2004; Sultana et al., 2011). This dissociation between bulk viable biomass and the microbial metabolism responsible for arsenic release may be, in part, a result of the large taxonomic diversity of microorganisms that is poorly resolved by PLFA analysis alone. It is possible that the component of the microbial community capable of reducing iron oxides and releasing arsenic into groundwater varies in proportion within different aquifers. If so, this would translate into disproportionate concentrations of arsenic which can be released into groundwater. Clearly, there is a need to characterize the importance of microbial groups present in the microbial community that are specifically capable of reductive dissolution of Fe(oxy)hydroxides and concomitant arsenic release. In this regard, more specific genetic-based methods of analyzing microbial community structures is necessary.

2.6.2 Sterol distribution and sewage input potential within Araihazar aquifers

Coprostanol was used in this study as a sewage biomarker because of its high (~60% total sterols) concentration in human and animal sewage waste (Leeming et al., 1996). It would, therefore, be reasonable to expect elevated concentrations of Coprostanol in environments contaminated by sewage. For example, a river in eastern China downstream from human settlements reported Coprostanol concentrations as high as 6.3×10^4 ng/g river sediment (He et al., 2018) while a coastal estuary downstream from the megacity Jakarta contained 3.0×10^2 to

$4.0x10^5$ ng/g river sediment (Dsikowitzky et al., 2017). Compared to these impacted sites, Coprostanol concentrations in Araihazar were at least 10^2 lower (2.2 ± 0.8 to 24.5 ± 0.9 ng/g ds at DI and DT, <LOD to 1.3 ng/gds at Site B, F). Furthermore, Coprostanol concentrations in Araihazar are lower than some pristine sites where Coprostanol has been reported to range from $<5x10^1$ to $4x10^3$ ng/g (Dsikowitzky et al., 2017; Grimalt et al., 1990; Speranza et al., 2018). Since the concentration of Coprostanol at impacted sites vary significantly, sewage input proxies such as the Sewage Contamination Index (SCI) and Coprostanol/Cholesterol (Cop/Ch) ratio reduce the ambiguity of the source of sterols in Araihazar aquifers. Therefore, despite greater absolute concentrations of Coprostanol at DI and DT, the low SCI (<0.7) and Cop/Ch (<1) attest to the low impact of sewage at these aquifers compared to previous sites B and F. The higher absolute Coprostanol concentrations at DI and DT are likely by-products of the anaerobic breakdown of Cholesterol by non-enteric bacteria in the aquifer. This is further supported by the greater proportion of the other biogenically produced sterol biomarkers such as phytosterols at DI. The source of the slightly greater Coprostanol concentration at DT compared to either Site B or F is as yet unknown considering DT is a relatively oxidizing orange aquifer. This contrasts Site B and F where key sewage biomarkers Coprostanol and Cholesterol overpower other sources of biogenic sterols. Furthermore, indicators of sewage contamination such as the SCI and Coprostanol/Cholesterol ratios are consistently greater than 0.7 and 1, respectively, along depth at Site B and F. It appears that sewage is a more important source of the organic sterols at Site B and F whereas plant-derived and other biogenic (e.g. macrophytes) are more important at DI and DT.

The depths where maximum sterol concentrations were measured at DI (18 m) and DT (15 m) are comparable to previous reports of elevated sterols at nearby sites (Whaley-Martin et al., 2017). As was previously argued, this may be a result of the accumulation of organic sterols from sewage or plants/fungi which enter the aquifer via preferential flowpaths along the annulus of drilled wells (Whaley-Martin et al., 2017). Additionally, the high K_{ow} of hydrophobic organic compounds such as sterols indicates that they preferentially partition to sediment surfaces upon entering the aquifer (Froehner and Sánez, 2013; Takada and Eganhouse, 1998) such that greater concentrations in groundwater would accumulate in the sediment matrix. Furthermore, shallow groundwater community wells abound at the surface providing abundant 'preferential flowpaths' which commonly tap depth ranges between 10 to 25 m into shallow Holocene aquifers (BGS & DPHE, 2001). However, this direct pathway created by drilled wells is absent at DI where the total and sewage sterol Coprostanol are greater than DT, Site B, or Site F. As an alternative, plant organic matter (e.g. from swamps and/or marine plankton) buried during the formation of the sedimentary aquifers may be a source of sedimentary phytosterols preserved at depth (McArthur et al., 2004; Meharg et al., 2006) which can explain the very high proportion of the phytosterols at DI and DT.

2.6.3 Correlation between PLFA and organic sterols

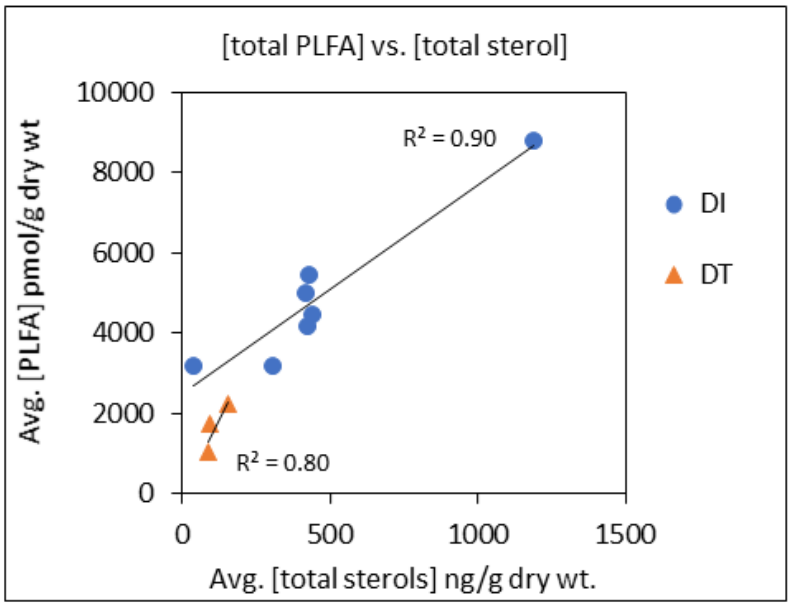


Figure 7. *Linear regression plot (right) comparing total sterols and PLFA at DI and DT with their respective correlation coefficient r-squared values.*

Positive correlation between total PLFA and total bulk sterols at both DI (r^2=0.90; Pearson r=0.95) and DT (r^2=0.80; Pearson r=0.89) suggests that active microbial metabolism is associated with the organic sterol pool (Figure 7). Because of the reduce bioavailability of hydrophobic organic sterols, this positive correlation does not indicate increased microbial metabolism due to elevated organic sterols as a carbon substrate. Instead, this supports that organic sterols may be transported simultaneously with more labile components of the DOC pool into the aquifer through, for example, groundwater recharge or drawdown from pumping.

PLFA were more positively correlated with all the phytosterols than with Coprostanol at both DI and DT. At DI, Pearson r values between PLFA and each of the sterols were 0.94 (Campesterol), 0.94 (Stigmasterol), 0.95 (β-Sitosterol), 0.78 (Coprostanol), 0.71 (Cholesterol), 0.34 (Cholestanol). On the other hand, at DT, Pearson r values between PLFA and each of the sterols were 0.76 (Campesterol), 0.97 (Stigmasterol), 1.0 (β-Sitosterol), 0.48 (Coprostanol), 0.67

(Cholesterol), 0.55 (Cholestanol). The generally more positive correlation of PLFA with each of the phytosterols compared to PLFA with the sewage biomarker is consistent with the argument that plant-derived organic matter is a more important component of the organic carbon pool than sewage. Furthermore, the low Sewage Contamination Index, and low Coprostanol/Cholesterol ratios confirmed that sewage inputs in either of these aquifers are low.

2.7 Implications of local human populations as a source of sewage

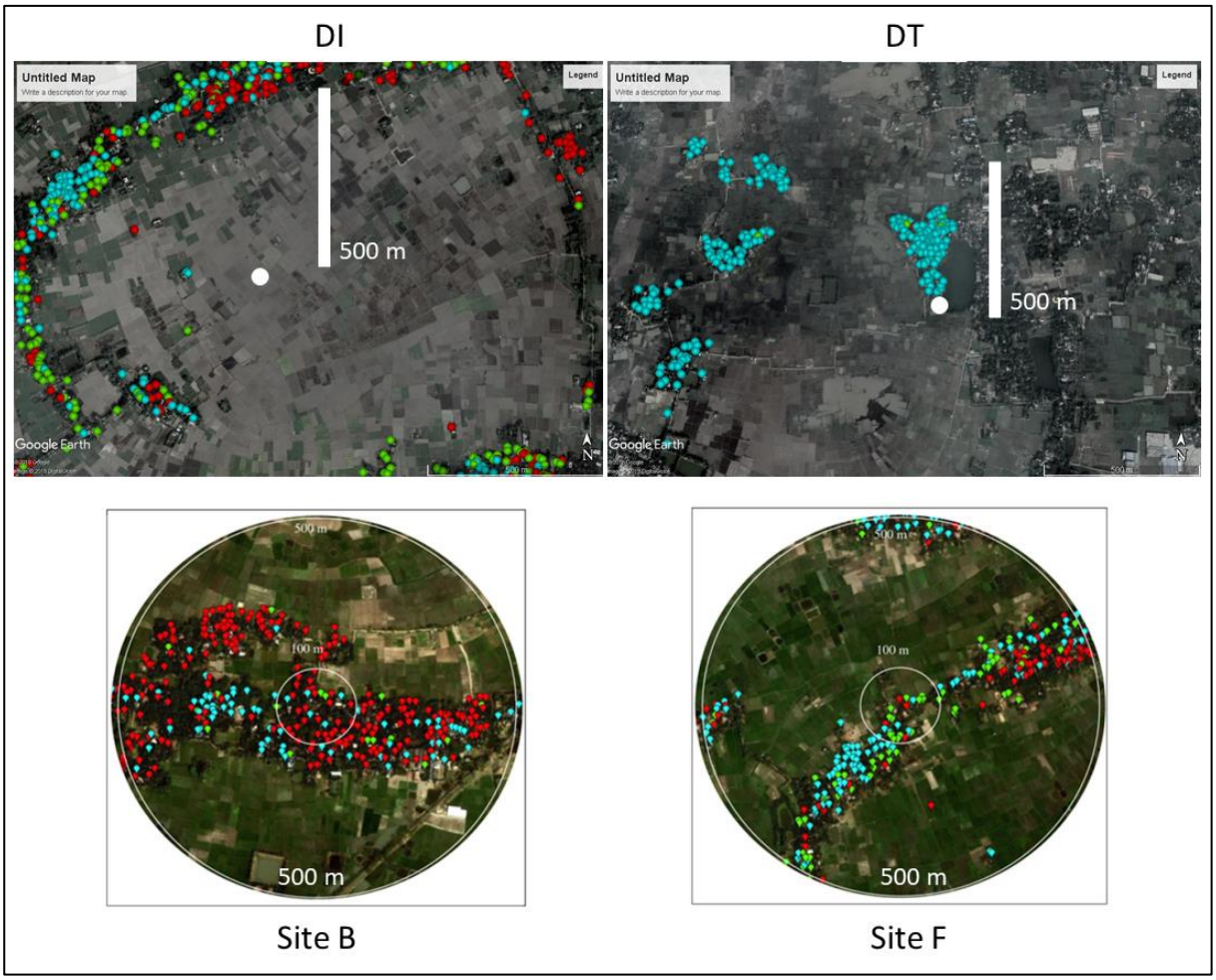


Figure 8. Abundance of community groundwater wells at four study sites in Araihazar: DI, DT, B, F. Sites B and F retrieved from Whaley-Martin et al. (2017); outer white circle = 500 m from drill site.

The density of local human populations is important to understand because humans are an obvious source of sewage in the environment (Bull et al., 2002; Frena et al., 2016; Knappett et al., 2011). As a qualitative descriptor, the density of community wells seems a reasonable proxy for local population density in rural Bangladesh. When compared in this context, Site B and F appear to have a greater community well density – and thus, a greater human population – than either Desert Island (DI) or Doper Tek (DT) (Figure 8). However, there is no evidence that the number of community wells directly reflects the human population making community well density a non-robust proxy for population density. Additionally, exploring a small number of sites (four) in a localized region relative to the country's surface area limits any regional scale arguments at this time. Nevertheless, these initial observations enable some fair speculation about sewage input potential into the aquifer.

If both DI and DT can be considered low local population proxies, this can explain the low impact of human and animal sewage as indicated by low SCI and Cop/Ch values. By contrast, the comparably higher population at Site B and F can provide a localized source of domestic sewage into the aquifer supporting sewage contamination previously reported at these sites (Whaley-Martin et al., 2017). This suggests that inputs of labile carbon in the aquifer may be influenced, in part, by differences in local population density. Specifically, the impact of sewage is more important in high population sites. This is conceivable given that unsanitary household pit latrines has been reported to contaminate shallow ponds used domestically for bathing and washing in rural Bangladesh (Knappett et al., 2011). On the other hand, low population sites, although uncontaminated by sewage, may receive other sources of labile carbon such as plants, fungi, or macrophytes that drive microbial metabolism as shown by the

positive correlation between bulk sterols and bulk PLFA. These other alternative labile carbon sources may help explain why, despite being located in an unpopulated agricultural field, the bulk sterol at DI was significantly greater than at DT, Site B, or F.

Overall, local human population density and sewage input into the aquifer requires further robust testing to examine the implications of sewage input as it relates to arsenic contamination in the aquifer. While improved sewage management may be expected to reduce the input of labile organic carbon at aquifers in high population regions, other natural carbon sources (e.g. plant matter) may be available and continue to promote high arsenic in otherwise pristine sites as shown by the comparably high dissolved arsenic at DI (unpopulated farm land), Site B, and F.

2.8 Conclusion

This study sought to investigate the role of viable microbial communities in controlling the carbon cycling and microbially mediated reductive arsenic release at two sites in Araihazar, Desert Island (DI) and Doper Tek (DT). These two sites are distinguished by their geochemistry: the aquifer at DI is of Holocene age with grey sediments, is highly reducing, and has high concentrations of dissolved arsenic (37-188 μg/L). On the other hand, the Pleistocene orange aquifer at DT has low dissolved arsenic (7 μg/L). The greater total concentrations of phospholipid fatty acids (PLFA) at DI indicated greater viable biomass and organic sterol content compared to DT. These same biomarkers were much lower at previous sites (Site B, F; Whaley-Martin et al., 2016; 2017), although dissolved arsenic concentrations at these sites were much higher than DI. These observations suggest that greater viable microbial biomass does not

directly result in a greater extent of reductive dissolution of sedimentary iron oxides and subsequent increase in dissolved arsenic. The disconnect between the size of the viable microbial biomass and arsenic concentration in these aquifers may be a result of differences in the proportion of the microbiota capable of reductive dissolution of iron oxides and concomitant arsenic release (e.g. iron reducing bacteria). At DI and DT, the greater bulk viable biomass may consist of fewer iron reducing bacteria compared to the microbial community at aquifers where dissolved arsenic is much greater despite a smaller microbial abundance. This indirect relationship between microbial metabolism and arsenic is further complicated by the initial sedimentary iron oxide content prior to the development of reducing conditions in the aquifer and its decoupling from arsenic in groundwater.

The positive correlation between dissolved arsenic and sewage contamination has been demonstrated at nearby sites (i.e. Site B, F) suggesting inputs of sewage containing labile carbon substrates may drive microbial metabolism leading to an increase in arsenic. In order to assess the potential for sewage input at DI and DT, the proportion of different types of organic sterols as representative indicators of carbon inputs into the aquifer were compared. Coprostanol was used as a sewage biomarker because it is produced in high concentrations (~60% of total sterols) in the gut of upper mammals such as humans and cattle. The occurrence of Coprostanol at low proportions at DI (2%) and DT (6%) and correspondingly high phytosterols (Campesterol, Stigmasterol, β-Sitosterol) (69-96% at DI; 76-85% at DT) that plant-derived organic matter make up a significantly greater component of the organic sterol pool in these aquifers. This is further confirmed by low Sewage Contamination Index (SCI) and Coprostanol/Cholesterol (Cop/Ch) ratio, sewage input proxies which compares the relative

importance of sewage derived Coprostanol to other naturally occurring biogenic sterols in the environment. Both of these proxies were calculated to be below their threshold values (SCI <0.7; Cop/Ch <1) at both DI and DT, confirming little sewage input at both sites. This is in contrast to reports from the previous sites which found increasing sewage contamination with arsenic concentrations at depth. The positive correlation between PLFA and organic sterols biomarkers at both DI and DT suggests that organic sterols in these aquifers are associated with active microbial community metabolism. The significant phytosterol concentrations and sewage input proxies that indicate little sewage input at both sites suggest the positive correlation is disproportionately influenced by plant-derived organic sources. These may come from buried organic matter such as terrestrial plants (e.g. swamps, marshland) which have accumulated at these shallow depths during the formation of the aquifer (Meharg et al., 2006; Shamsudduha et al., 2008).

Since radiocarbon analysis ($\Delta^{14}C$) indicated young potentially surface-derived carbon is actively metabolized by the microbial community, bulk organic sterol input may also be derived from similarly young sources, perhaps as part of the DOC pool transported from the surface into the aquifer. This was previously proposed to explain elevated organic sterols (especially sewage-derived sterols) at depth at nearby sites. Specifically, they argued that surface-derived organic sterols incorporated into the DOC pool accumulated in the aquifer as they are transported down the annulus of community groundwater wells. This may explain the elevated biomarker concentrations at 15 m depth at DT which could contribute to the high viable microbial biomass at this depth. However, the location of DI in the middle of farmland where community groundwater wells are absent precludes direct input of surface-derived organic

sterols via these proposed preferential pathways. In this case, other biogenic sterol sources (for example, buried plant matter) generates a much greater bulk organic sterol concentrations at depth. The visual community groundwater well density was used as a visual proxy for human population density; previous sites where sewage contamination correlated with high arsenic had a visually greater human density compared to either DI or DT. This can potentially explain why there is comparatively little sewage contamination despite a higher total organic sterol input at DI and DT. In high population regions such as Site B and F, sewage can be an important source of organic carbon which can drive arsenic release – this is conceivable given the ubiquity of open pit latrines in rural Bangladesh. This study demonstrated that higher total viable biomass as a result of greater microbial community metabolism does not directly result in an increase in arsenic concentration resulting in highly contaminated aquifer sites.

2.9 Statement of research contributions

All molecular biomarkers (PLFA and sterols) from Desert Island (DI) and Doper Tek (DT) sites were extracted, identified, and quantified by Reisa San Pedro at the Environmental Organic Geochemistry lab at McMaster University. Radiocarbon analysis of bulk PLFA samples were externally analyzed at the NOSAMS facility in Woods Hole, Massachusetts. Sediment matrix cores and groundwater (for arsenic measurements) samples were collected from Bangladesh and analyzed at the Lamont Doherty Earth Observatory by the Columbia University (Alexander van Geen, Tyler Ellis) and Barnard College (Brian Mailloux) team. Raw values of biomarker concentrations and radiocarbon data (PLFA, DNA, carbon pools) from previous sites (Site B and F) were retrieved from previous studies by Whaley-Martin et al. (2016, 2017) and Mailloux et al. (2012).

Bibliography

Ahmed, F., Clemens, J. D., Rao, M. R., & Banik, A. K. (1994). Family latrines and paediatric shigellosis in rural Bangladesh: Benefit or risk? *International Journal of Epidemiology*, *23*(4), 856–862. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed3&NEWS=N&AN=1994269289

BGS, & DPHE. (2001). Arsenic contamination of groundwater in Bangladesh. *British Geological Survey, Department of Public Health Engineering*, *1*. Retrieved from http://nora.nerc.ac.uk/id/eprint/11986

Birk, J. J., Dippold, M., Wiesenberg, G. L. B., & Glaser, B. (2012). Combined quantification of faecal sterols, stanols, stanones and bile acids in soils and terrestrial sediments by gas chromatography-mass spectrometry. *Journal of Chromatography A*, *1242*, 1–10. https://doi.org/10.1016/j.chroma.2012.04.027

Bull, I. D., Matthew J. Lockheart, Mohamed M. Elhmmali, David J. Roberts, & Richard P. Evershed. (2002). The Origin of Faeces by Means of Biomarker Detection. Environ Int - Google Scholar. *Environment International*, *27*(8), 647–654. Retrieved from https://scholar.google.de/scholar?hl=en&q=%2C+Evershed+RP+%282002%29+The+Origin+of+Faeces+by+Means+of+Biomarker+Detection.+Environ+Int&btnG=&as_sdt=1%2C5&as_sdtp=

Carreira, R. S., Wagener, A. L. R., & Readman, J. W. (2004). Sterols as markers of sewage contamination in a tropical urban estuary (Guanabara Bay, Brazil): Space-time variations. *Estuarine, Coastal and Shelf Science*, *60*(4), 587–598. https://doi.org/10.1016/j.ecss.2004.02.014

Cheng, Z., Zheng, Y., Mortlock, R., & Van Geen, A. (2004). Rapid multi-element analysis of groundwater by high-resolution inductively coupled plasma mass spectrometry. *Analytical and Bioanalytical Chemistry*, *379*(3), 512–518. https://doi.org/10.1007/s00216-004-2618-x

Chiou, C. T., Schmedding, D. W., & Manes, M. (1982). Partitioning of organic compounds in octanol-water systems. *Environmental Science & Technology*, *16*(1), 4–10. Retrieved from https://pubs.acs.org/sharingguidelines

Coleman, M. L., Hedrick, D. B., Lovley, D. R., White, D. C., & Pye, K. (1993). Reduction of Fe(III) in sediments by sulphate-reducing bacteria. *Nature*, *361*(6411), 436–438. https://doi.org/10.1038/361436a0

Dhar, R K, Zheng, Y., Stute, M., van Geen, A., Cheng, Z., Shanewaz, M., et al. (2008). Temporal variability of groundwater chemistry in shallow and deep aquifers of Araihazar, Bangladesh. *Journal of Contaminant Hydrology*, *99*(1–4), 97–111. https://doi.org/10.1016/j.jconhyd.2008.03.007

Dhar, Ratan K, Zheng, Y., Saltikov, C. W., Radloff, K. A., Mailloux, B. J., Ahmed, K. M., & Van Geen, A. (2011). Microbes Enhance Mobility of Arsenic in Pleistocene Aquifer Sand from Bangladesh. *Environ. Sci. Technol*, *45*, 31. https://doi.org/10.1021/es1022015

Dsikowitzky, L., Schäfer, L., Dwiyitno, Ariyani, F., Irianto, H. E., & Schwarzbauer, J. (2017). Evidence of massive river pollution in the tropical megacity Jakarta as indicated by faecal steroid occurrence and the seasonal flushing out into the coastal ecosystem. *Environmental Chemistry Letters*, *15*(4), 703–708. https://doi.org/10.1007/s10311-017-0641-3

Eglinton, G., & Calvin, M. (1967). Chemical Fossils. *Scientific American*, *216*(1), 32–43. https://doi.org/10.2307/24931372

Eglinton, G., Scott, P. M., Belsky, T., Burlingame, A. L., & Calvin, M. (1964). Hydrocarbons of Biological Origin from a one-Billion-Year-Old Sediment. *Science*, *145*(3629), 263–264. Retrieved from https://www-jstor-org.libaccess.lib.mcmaster.ca/stable/pdf/1713980.pdf?refreqid=excelsior%3A264061d827eea2c073504e3707e4141e

Fattore, E., Benfenati, E., Marelli, R., Cools, E., & Fanelli, R. (1996). Sterols in sediment samples from Venice Lagoon, Italy. *Chemosphere*, *33*(12), 2383–2393. https://doi.org/10.1016/S0045-6535(96)00340-2

Frena, M., Bataglion, G. A., Tonietto, A. E., Eberlin, M. N., Alexandre, M. R., & Madureira, L. A. S. (2016). Assessment of anthropogenic contamination with sterol markers in surface sediments of a tropical estuary (Itajaí-Açu, Brazil). *Science of the Total Environment*, *544*, 432–438. https://doi.org/10.1016/j.scitotenv.2015.11.137

Froehner, S., & Sánez, J. (2013). Evaluation of potential sewage contamination by fecal sterol biomarkers adsorbed in natural biofilms. *Environmental Sciences: Processes and Impacts*, *15*(11), 2080–2086. https://doi.org/10.1039/c3em00025g

Gaskell, S. J., & Eglinton, G. (1975). Rapid hydrogenation of sterols in a contemporary lacustrine sediment. *Nature*, *254*(5497), 209–211. Retrieved from https://www.nature.com/articles/254209b0.pdf

van Geen, A., Zheng, Y., Versteeg, R., Stute, M., Horneman, A., Dhar, R., et al. (2003). Spatial variability of arsenic in 6000 tube wells in a 25 km2 area of Bangladesh. *Water Resources Research*, *39*(5). https://doi.org/10.1029/2002WR001617

Ghosh, D., Routh, J., & Bhadury, P. (2015). Characterization and microbial utilization of dissolved lipid organic fraction in arsenic impacted aquifers (India). *Journal of Hydrology*, *527*, 221–233. https://doi.org/10.1016/j.jhydrol.2015.04.051

Green, C. T., & Scow, K. M. (2000). Analysis of phospholipid fatty acids (PLFA) to characterize microbial communities in aquifers. *Hydrogeology Journal*, *8*(1), 126–141. https://doi.org/10.1007/s100400050013

Green, G., Skerratt, J. H., Leeming, R., & Nichols, P. (1992). Hydrocarbon and Coprostanol Levels in Seawater, Sea-ice Algae and Sediments near Davis Station in Eastern Antarctica: *Marine Pollution Bulletin*, *25*, 293–296. Retrieved from https://pdf.sciencedirectassets.com/271825/1-s2.0-S0025326X00X03797/1-s2.0-0025326X9290685Y/main.pdf?x-amz-security-token=AgoJb3JpZ2luX2VjEEYaCXVzLWVhc3QtMSJIMEYCIQCXSiDwa%2Fd6HXN0k8rhzxBANLhLNLyb5Ed9JkQM%2FZC3nQIhAPshGqkICufpJAV1XvpUQKyJ80qBvSUQOxsxwvFRT

Grimalt, J. O., Fernández, P., Bayona, J. M., & Albaigés, J. (1990). Assessment of Fecal Sterols and Ketones as Indicators of Urban Sewage Inputs to Coastal Waters. *Environmental Science and Technology*, *24*(3), 357–363. https://doi.org/10.1021/es00073a011

Guckert, J. B., Antworth, C. P., Nichols, P. D., & White, D. C. (1985). Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. *FEMS Microbiology Letters*, *31*(3), 147–158. https://doi.org/10.1016/0378-1097(85)90016-3

Guo, H., Li, X., Xiu, W., He, W., Cao, Y., Zhang, D., & Wang, A. (2019). Controls of organic matter

bioreactivity on arsenic mobility in shallow aquifers of the Hetao Basin, P.R. China. *Journal of Hydrology*, *571*(November 2018), 448–459. https://doi.org/10.1016/j.jhydrol.2019.01.076

Hansch, C., Quinlan, J. E., & Lawrence, G. L. (1968). The linear free-energy relationship between partition coefficients and the aqueous solubility of organic liquids. *Journal of Organic Chemistry*, *33*(1), 347–350. https://doi.org/10.1016/0016-0032(63)90468-x

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

Harvey, C. F., Ashfaque, K. N., Yu, W., Badruzzaman, A. B. M., Ali, M. A., Oates, P. M., et al. (2006). Groundwater dynamics and arsenic contamination in Bangladesh. *Chemical Geology*, *228*(1-3 SPEC. ISS.), 112–136. https://doi.org/10.1016/j.chemgeo.2005.11.025

He, D., Zhang, K., Tang, J., Cui, X., & Sun, Y. (2018). Using fecal sterols to assess dynamics of sewage input in sediments along a human-impacted river-estuary system in eastern China. *Science of the Total Environment*, *636*, 787–797. https://doi.org/10.1016/j.scitotenv.2018.04.314

Héry, M., Van Dongen, B. E., Gill, F., Mondal, D., Vaughan, D. J., Pancost, R. D., et al. (2010). Arsenic release and attenuation in low organic carbon aquifer sediments from West Bengal. *Geobiology*, *8*(2), 155–168. https://doi.org/10.1111/j.1472-4669.2010.00233.x

Horneman, A., van Geen, A., Kent, D. V, Mathe, P. E., Zheng, Y., Dhar, R. K., et al. (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. https://doi.org/10.1016/j.gca.2004.01.026

Huang, W. Y., & Meinschein, W. G. (1979). Sterols as ecological indicators. *Geochimica et Cosmochimica Acta*, *43*(5), 739–745. https://doi.org/10.1016/0016-7037(79)90257-6

Islam, F. S., Gault, A. G., Boothman, C., Polya, D. A., Charnock, J. M., Chatterjee, D., & Lloyd, J. R. (2004). Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature*, *430*(6995), 68–71. https://doi.org/10.1038/nature02638

Isobe, K. O., Tarao, M., Zakaria, M. P., Chiem, N. H., Minh, L. Y., & Takada, H. (2002). Quantitative application of fecal sterols using gas chromatography - Mass spectrometry to investigate fecal pollution in tropical waters: Western Malaysia and Mekong Delta, Vietnam. *Environmental Science and Technology*, *36*(21), 4497–4507. https://doi.org/10.1021/es020556h

Knappett, P. S. K., Escamilla, V., Layton, A., McKay, L. D., Emch, M., Williams, D. E., et al. (2011). Impact of population and latrines on fecal contamination of ponds in rural Bangladesh. *Science of the Total Environment*, *409*(17), 3174–3182. https://doi.org/10.1016/j.scitotenv.2011.04.043

Kocar, B. D., Polizzotto, M. L., Benner, S. G., Ying, S. C., Ung, M., Ouch, K., et al. (2008). Integrated biogeochemical and hydrologic processes driving arsenic release from shallow sediments to groundwaters of the Mekong delta. *Applied Geochemistry*, *23*(11), 3059–3071. https://doi.org/10.1016/j.apgeochem.2008.06.026

Kolpin, D., Furlong, E., Meyer, M., Barber, L., Zaugg, S., Buxton, H., & Meyer, M. T. (2000). Pharmaceuticals , Hormones , and Other Organic Wastewater Contaminants in U . S . Streams , 1999-2000 : A National Reconnaissance Other Organic Wastewater

Contaminants in U . S . Streams , 1999-2000 : A National Reconnaissance. *Environmental Science & Technology*, *36*(6), 1999–2000.

Leeming, R., & Nichols, P. (1998). Determination of the sources and distribution of sewage and pulp fibre-derived pollution in the Derwent Estuary, Tasmania, using sterol biomarkers. *Marine Freshwater Research*, *49*, 7–17. https://doi.org/10.1071/MF99078

Leeming, R., Ball, A., Ashbolt, N., & Nichols, P. (1996). Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. *Water Research*, *30*(12), 2893–2900. https://doi.org/10.1016/S0043-1354(96)00011-5

Legg, T. M., Zheng, V., Simone, B., Radloff, K. A., Mladenov, N., Gonzalez, A., et al. (2012). Carbon, metals, and grain size correlate with bacterial community structure in sediments of a high arsenic aquifer. *Frontiers in Microbiology*, *3*, 1–15. https://doi.org/10.3389/fmicb.2012.00082

Liu, J., Yao, X., Lu, J., Qiao, X., Liu, Z., & Li, S. (2016). Distribution and factors affecting adsorption of sterols in the surface sediments of Bosten Lake and Manas Lake, Xinjiang. *Environmental Science and Pollution Research*, *23*(6), 5892–5901. https://doi.org/10.1007/s11356-015-5807-6

Ludvigsen, L., Albrechtsen, H. J., Ringelberg, D. B., Ekelund, F., & Christensen, T. H. (1999). Distribution and composition of microbial populations in a landfill leachate contaminated aquifer (Grindsted, Denmark). *Microbial Ecology*, *37*(3), 197–207. https://doi.org/10.1007/s002489900143

Mailloux, B. J., Trembath-Reichert, E., Cheung, J., Watson, M., Stute, M., Freyer, G. A., et al. (2013). Advection of surface-derived organic carbon fuels microbial reduction in Bangladesh groundwater. *Proceedings of the National Academy of Sciences*, *110*(14), 5331–5335. https://doi.org/10.1073/pnas.1213141110

Martins, C. C., Seyffert, B. H., Braun, J. A. F., & Fillmann, G. (2011). Input of organic matter in a large South American tropical estuary (paranaguá estuarine system, Brazil) indicated by sedimentary sterols and multivariate statistical approach. *Journal of the Brazilian Chemical Society*, *22*(8), 1585–1594. https://doi.org/10.1590/S0103-50532011000800023

Matić Bujagić, I., Grujić, S., Jauković, Z., & Laušević, M. (2016). Sterol ratios as a tool for sewage pollution assessment of river sediments in Serbia. *Environmental Pollution*, *213*, 76–83. https://doi.org/10.1016/j.envpol.2015.12.036

McArthur, J. ., Banerjee, D. ., Hudson-Edwards, K. ., Mishra, R., Purohit, R., Ravenscroft, P., et al. (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. *Applied Geochemistry*, *19*(8), 1255–1293. https://doi.org/10.1016/j.apgeochem.2004.02.001

McArthur, J. M., Sikdar, P. K., Hoque, M. A., & Ghosal, U. (2012). Waste-water impacts on groundwater: Cl/Br ratios and implications for arsenic pollution of groundwater in the Bengal Basin and Red River Basin, Vietnam. *Science of the Total Environment*, *437*, 390–402. https://doi.org/10.1016/j.scitotenv.2012.07.068

Meharg, A. A., Scrimgeour, C., Hossain, S. A., Fuller, K., Cruickshank, K., Williams, P. N., & Kinniburgh, D. G. (2006). Codeposition of organic carbon and arsenic in Bengal Delta aquifers. *Environmental Science and Technology*, *40*(16), 4928–4935. https://doi.org/10.1021/es060722b

Mladenov, N., Zheng, Y., Miller, M. P., Nemergut, D. R., Legg, T., Simone, B., et al. (2010).

Dissolved organic matter sources and consequences for iron and arsenic mobilization in Bangladesh aquifers. *Environmental Science and Technology*, *44*(1), 123–128. https://doi.org/10.1021/es901472g

Neumann, R. B., Ashfaque, K. N., Badruzzaman, A. B. M., Ashraf Ali, M., Shoemaker, J. K., & Harvey, C. F. (2010). Anthropogenic influences on groundwater arsenic concentrations in Bangladesh. *Nature Geoscience*, *3*(1), 46–52. https://doi.org/10.1038/ngeo685

Nishimura, M. (1982). 5β-isomers of stanols and stanones as potential markers of sedimentary organic quality and depositional paleoenvironments. *Geochimica et Cosmochimica Acta*, *46*(3), 423–432. https://doi.org/10.1016/0016-7037(82)90233-2

Nishimura, M., & Koyama, T. (1977). The occurrence of stanols in various living organisms and the behavior of sterols in contemporary sediments. *Geochimica et Cosmochimica Acta*, *41*(3), 379–385. https://doi.org/10.1016/0016-7037(77)90265-4

Pearson, A., Mcnichol, A. P., Schneider, R. J., Von Reden, K. F., & Zheng, Y. (1998). Microscale AMS 14C Measurement at NOSAMS. *Radiocarbon*, *40*(1), 61–75. https://doi.org/10.1017/S0033822200017902

Pearson, A., McNichol, A. P., Benitez-Nelson, B. C., Hayes, J. M., & Eglinton, T. I. (2001). Origins of lipid biomarkers in Santa Monica Basin surface sediment: A case study using compound-specific Δ14C analysis. *Geochimica et Cosmochimica Acta*, *65*(18), 3123–3137. https://doi.org/10.1016/S0016-7037(01)00657-3

Petsch, S. T., Eglinton, T. I., & Edwards, K. J. (2001). 14C-dead living biomass: Evidence for microbial assimilation of ancient organic carbon during shale weathering. *Science*, *292*(5519), 1127–1131. https://doi.org/10.1126/science.1058332

Pracht, L. E., Tfaily, M. M., Ardissono, R. J., & Neumann, R. B. (2018). Molecular characterization of organic matter mobilized from Bangladeshi aquifer sediment: Tracking carbon compositional change during microbial utilization. *Biogeosciences*, *15*(6), 1733–1747. https://doi.org/10.5194/bg-15-1733-2018

Pratt, C., Warnken, J., Leeming, R., Arthur, M. J., & Grice, D. I. (2008). Degradation and responses of coprostanol and selected sterol biomarkers in sediments to a simulated major sewage pollution event: A microcosm experiment under sub-tropical estuarine conditions. *Organic Geochemistry*, *39*(3), 353–369. https://doi.org/10.1016/j.orggeochem.2007.06.009

Radloff, K. A., Cheng, Z., Rahman, M. W., Ahmed, K. M., Mailloux, B. J., Juhl, A. R., et al. (2007). Mobilization of arsenic during one-year incubations of grey aquifer sands from Araihazar, Bangladesh. *Environmental Science and Technology*, *41*(10), 3639–3645. https://doi.org/10.1021/es062903j

Reeves, A., & Patton, D. (2005). Faecal sterols as indicators of sewage contamination in estuarine sediments of the Tay Estuary, Scotland: an extended baseline survey. *Hydrology and Earth System Sciences*, *9*(1/2), 81–94. https://doi.org/10.5194/hess-9-81-2005

Richards, L. A., Magnone, D., Sültenfuß, J., Bryant, C., Boyce, A. J., van Dongen, B. E., et al. (2019). Dual in-aquifer and near surface processes drive arsenic mobilization in Cambodian groundwaters. *Science of The Total Environment*, *659*, 699–714. https://doi.org/10.1016/j.scitotenv.2018.12.437

Shah, S. R., Mollenhauer, G., Ohkouchi, N., Eglinton, T. I., & Pearson, A. (2008). Origins of archaeal tetraether lipids in sediments: Insights from radiocarbon analysis. *Geochimica et*

Cosmochimica Acta, *72*(18), 4577–4594. https://doi.org/10.1016/j.gca.2008.06.021
Shah, V. G., Hugh Dunstan, R., Geary, P. M., Coombes, P., Roberts, T. K., & Von Nagy-Felsobuki, E. (2007). Evaluating potential applications of faecal sterols in distinguishing sources of faecal contamination from mixed faecal samples. *Water Research*, *41*(16), 3691–3700. https://doi.org/10.1016/j.watres.2007.04.006
Shamsudduha, M., Uddin, A., Saunders, J. A., & Lee, M. K. (2008). Quaternary stratigraphy, sediment characteristics and geochemistry of arsenic-contaminated alluvial aquifers in the Ganges-Brahmaputra floodplain in central Bangladesh. *Journal of Contaminant Hydrology*, *99*(1–4), 112–136. https://doi.org/10.1016/j.jconhyd.2008.03.010
Simoneit, B. R. T. (2004). Biomarkers (molecular fossils) as geochemical indicators of life. *Advances in Space Research*, *33*(8), 1255–1261. https://doi.org/10.1016/j.asr.2003.04.045
Slater, G. F., White, H. K., Eglinton, T. I., & Reddy, C. M. (2005). Determination of microbial carbon sources in petroleum contaminated sediments using molecular 14C analysis. *Environmental Science and Technology*, *39*(8), 2552–2558. https://doi.org/10.1021/es048669j
Speranza, E. D., Colombo, M., Skorupka, C. N., & Colombo, J. C. (2018). Early diagenetic alterations of sterol biomarkers during particle settling and burial in polluted and pristine areas of the Rio de la Plata Basin. *Organic Geochemistry*, *117*, 1–11. https://doi.org/10.1016/j.orggeochem.2017.11.013
Stute, M., Zheng, Y., Schlosser, P., Horneman, A., Dhar, R. K., Datta, S., et al. (2007). Hydrological control of As concentrations in Bangladesh groundwater. *Water Resources Research*, *43*(9). https://doi.org/10.1029/2005WR004499
Sultana, M., Härtig, C., Planer-Friedrich, B., Seifert, J., & Schlömann, M. (2011). Bacterial communities in Bangladesh aquifers differing in aqueous arsenic concentration. *Geomicrobiology Journal*, *28*(3), 198–211. https://doi.org/10.1080/01490451.2010.490078
Vane, C. H., Kim, A. W., McGowan, S., Leng, M. J., Heaton, T. H. E., Kendrick, C. P., et al. (2010). Sedimentary records of sewage pollution using faecal markers in contrasting peri-urban shallow lakes. *Science of the Total Environment*, *409*(2), 345–356. https://doi.org/10.1016/j.scitotenv.2010.09.033
Vestal, J. R., & White, D. C. (1989). Lipid Analysis in Microbial Ecology. *BioScience*, *39*(8), 535–541. https://doi.org/10.2307/1310976
Whaley-Martin, K. J., Mailloux, B. J., Van Geen, A., Bostick, B. C., Silvern, R. F., Kim, C., et al. (2016). Stimulation of Microbially Mediated Arsenic Release in Bangladesh Aquifers by Young Carbon Indicated by Radiocarbon Analysis of Sedimentary Bacterial Lipids. *Environmental Science and Technology*, *50*(14), 7353–7363. https://doi.org/10.1021/acs.est.6b00868
Whaley-Martin, K. J., Mailloux, B. J., van Geen, A., Bostick, B. C., Ahmed, K. M., Choudhury, I., & Slater, G. F. (2017). Human and livestock waste as a reduced carbon source contributing to the release of arsenic to shallow Bangladesh groundwater. *Science of the Total Environment*, *595*, 63–71. https://doi.org/10.1016/j.scitotenv.2017.03.234
White, D., Stair, J., & Ringelberg, D. (1996). Quantitative comparisons ofin situ microbial biodiversity by signature biomarker analysis. *Journal of Industrial Microbiology & Biotechnology*, *17*(3–4), 185–196. https://doi.org/10.1007/BF01574692
White, D. C., Davis, W. M., Nickels, J. S., King, J. D., & Bobbie, R. J. (1979). Determination of the

sedimentary microbial biomass by extractible lipid phosphate. *Oecologia*, *40*(1), 51–62. https://doi.org/10.1007/BF00388810

World Health Organization. (2011). *Guidelines for Drinking-water Quality*. *World Health Organization*. WHO Press. Retrieved from http://www.who.int

Yu, K., Gan, Y., Zhou, A., Liu, C., Duan, Y., Han, L., & Zhang, Y. (2018). Organic carbon sources and controlling processes on aquifer arsenic cycling in the Jianghan Plain, central China. *Chemosphere*, *208*, 773–781. https://doi.org/10.1016/j.chemosphere.2018.05.188

Zheng, Y., van Geen, A., Stute, M., Dhar, R., Mo, Z., Cheng, Z., 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. https://doi.org/10.1016/j.gca.2005.06.001

2.10 Chapter 2. Supporting Information: Exploring regional-scale microbial carbon cycling and sewage inputs at two sites (Desert Island, Doper Tek) in Araihazar aquifers using phospholipid and sterol biomarkers

List of Figures

List of Tables

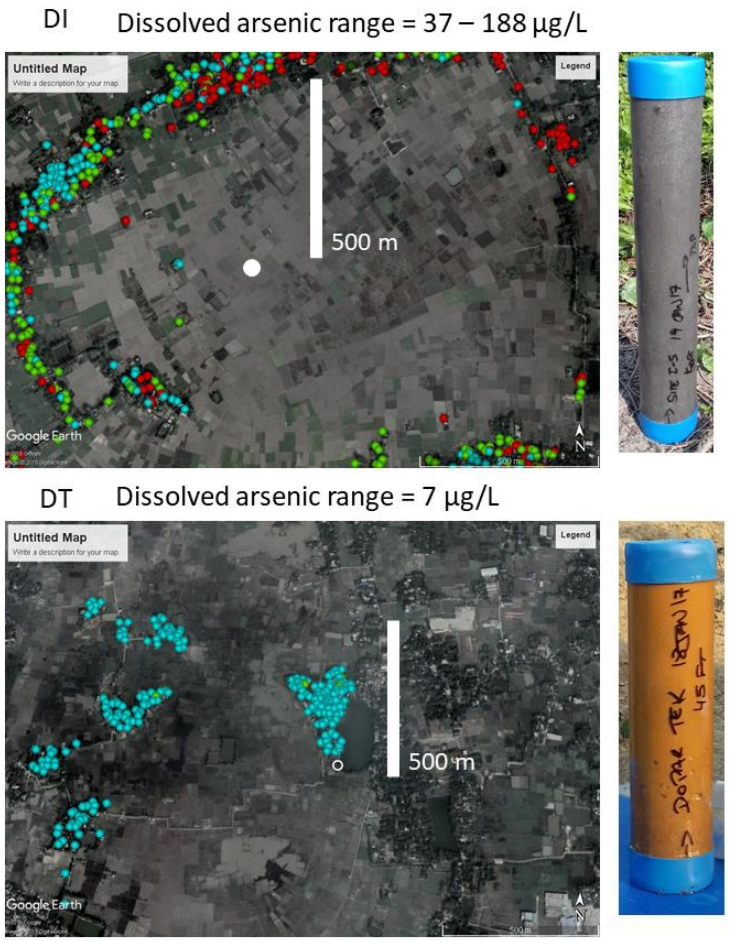


Figure S1. *Site maps (left) and sedimentary matrix cores (right) at Desert Island (top row) and Doper Tek (bottom row). Colored dots on site maps represent community wells and dissolved As measurements (cyan = 0-10 µg/L As, green = >10-50 µg/L As, red = >50 µg/L As).*

Table S1. Lithology and arsenic distribution along the Desert Island (DI) and Doper Tek (DT) core.					
Site Name	**Depth (ft)**	**Depth (m)**	**Lithology**	**Solid-phase sedimentary As (mg/kg)**	**Dissolved groundwater As (ug/L) ± 5%**
Desert Island	5	1.5	Brown clay	4.2 ± 0.6	
Desert Island	10	3.0	Brown clay	3.6 ± 0.6	
Desert Island	15	4.6	Grey Medium to fine Sand	<LOQ	
Desert Island	20	6.1	Grey Medium to fine Sand	<LOQ	68
Desert Island	25	7.6	Grey Medium to fine Sand	<LOQ	
Desert Island	30	9.1	Grey Medium to fine Sand	<LOQ	
Desert Island	35	10.7	Grey Medium to fine Sand	<LOQ	188
Desert Island	40	12.2	Grey Medium Sand	1.9 ± 0.5	
Desert Island	45	13.7	Grey Medium Sand	<LOQ	

Desert Island	50	15.24	Grey Medium to fine Sand	<LOQ	
Desert Island	55	16.8	Grey Medium Sand	<LOQ	68
Desert Island	60	18.3	Grey Medium Sand	<LOQ	
Desert Island	65	19.8	Grey Medium Sand	<LOQ	
Desert Island	70	21.3	Grey Medium to fine Sand	<LOQ	
Desert Island	75	22.9	Grey Medium to fine Sand	<LOQ	37
Doper Tek	0	0		10.4 ± 0.8	
Doper Tek	5	1.5	Brown Clay	10.2 ± 0.8	
Doper Tek	10	3.0	Brown Clay	10 ± 0.7	
Doper Tek	15	4.6	Brown Silty Clay	5 ± 0.7	
Doper Tek	20	6.1	Brown Fine Sand	4.9 ± 0.6	
Doper Tek	25	7.6	Brown Fine Sand	3.9 ± 0.6	
Doper Tek	30	9.1	Brown Fine Sand	<LOQ	
Doper Tek	35	10.7	Brown Medium Sand	<LOQ	7
Doper Tek	40	12.2	Brown Medium Sand	<LOQ	
Doper Tek	43	13.1	Brown Clay	2.6 ± 0.6	
Doper Tek	45	13.7	Brown Clay	<LOQ	
Doper Tek	50	15.2	Brown Medium to fine Sand	<LOQ	7

Table S2. Concentrations of PLFA and corresponding microbial cell abundance estimates at the two sites, Desert Island (DI) and Doper Tek (DT).

Site	Depth (ft)	Depth (m)	Avg. total pmol PLFA/g dry sed (± 10%)	Cell abundance[a] cells/g dry sed (± 10%)
DI	25	7.6	3199	6.4E+07
DI	30	9.1	4486	9.0E+07
DI	45	14	5473	1.1E+08
DI	55	16.8	4165	8.3E+07
DI	60	18.3	8812	1.8E+08
DI	70	21.3	3211	6.4E+07
DI	75	22.9	5022	1.0E+08
DT	40	12.2	1058	2.1E+07
DT	45	13.7	1731	3.5E+07
DT	50	15.2	2244	4.5E+07

[a] Microbial cell abundance estimated using generic cell conversion factor $2x10^4$ cells/pmol PLFA

Table S3. Concentrations of sedimentary sterols at Desert Island (DI) and Doper Tek (DT).

Site	Depth		Coprostanol (ng/g dry sed)[a]	Cholesterol (ng/g dry sed)[a]	Cholestanol (ng/g dry sed)[a]	Campesterol (ng/g dry sed)[a]	Stigmasterol (ng/g dry sed)[a]	b-sitosterol (ng/g dry sed)[a]	Sewage Contamination Index
	ft	m							
DI	25	7.6	2.2 ± 0.8	4.6 ± 2.8	4.1 ± 2.3	4.0 ± 2.3	11.8 ± 7.7	8.6 ± 4.9	0.3
DI	30	9.1	6.2 ± 2.8	5.6 ± 4.3	4.0 ± 3.1	66.9 ± 3.3	191 ± 10	167 ± 8	0.6
DI	45	13.7	8.6 ± 6.1	15.6 ± 0.8	16.5 ± 0.8	48.4 ± 3.9	175 ± 14	163 ± 12	0.3
DI	55	16.8	15.1 ± 0.2	33.5 ± 1.7	18.3 ± 0.9	53.7 ± 2.7	170 ± 9	132 ± 7	0.5
DI	60	18.3	24.5 ± 0.9	148 ± 7	23.9 ± 1.2	141 ± 7	530 ± 27	431 ± 22	0.5
DI	70	21.3	8.2 ± 0.3	9.4 ± 0.9	8.4 ± 0.8	27.3 ± 2.0	143 ± 7	110 ± 6	0.5
DI	75	22.9	3.8 ± 2.7	19.4 ± 1.0	57 ± 2.9	66.7 ± 3.3	151 ± 8	131 ± 7	0.1
DT	40	12.2	5.9 ± 1.3	9.1 ± 0.7	5.4 ± 0.5	14.9 ± 1.7	30.5 ± 1.5	20.2 ± 1.0	0.5
DT	45	13.7	2.9 ± 0.5	7.4 ± 2.8	4.2 ± 1.6	13.0 ± 4.4	39.3 ± 2.0	28.8 ± 1.4	0.4
DT	50	15.2	9.6 ± 5.2	14.4 ± 1.4	7.4 ± 0.7	28.6 ± 1.4	56.2 ± 2.8	36.9 ± 1.8	0.6

[a] Standard curve reproducibility is 10%; with the exception of samples indicated by large error bars, reproducibility of replicates were generally 5%.

Table S4. Correlation coefficients (Pearson r) values between concentrations of PLFA and each of the 6 sterol biomarkers at DI and DT.

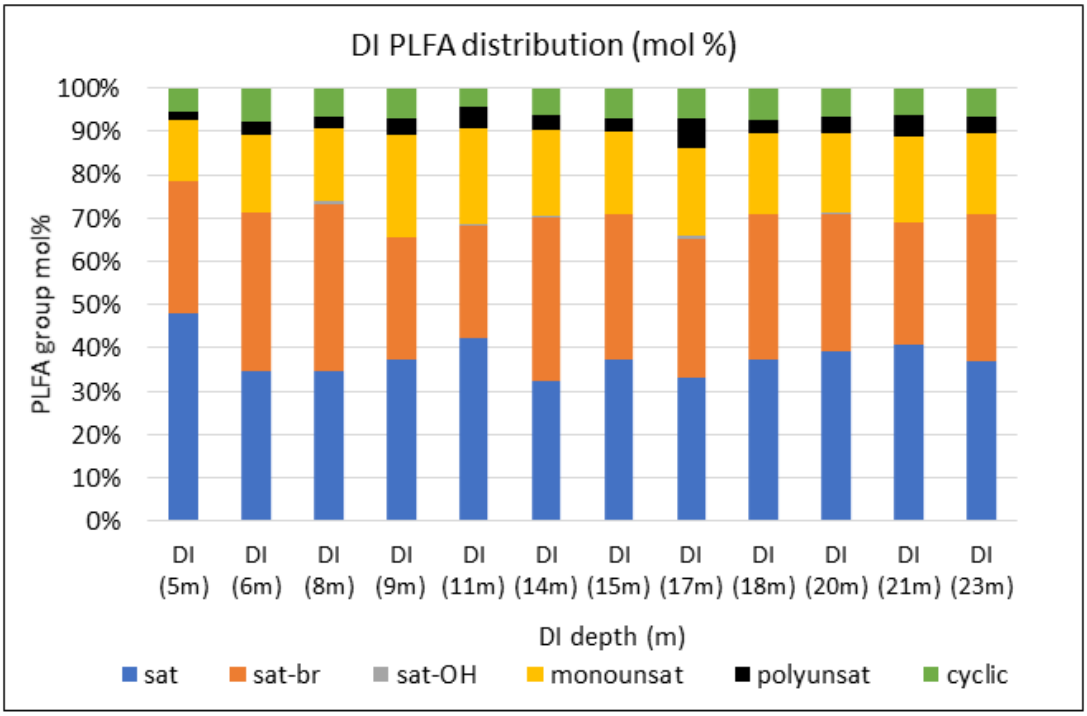
Site, n= samples	Pearson r values between PLFA and individual sterols					
	[PLFA] vs. [coprostanol]	**[PLFA] vs. [cholesterol]**	**[PLFA] vs. [cholestanol]**	**[PLFA] vs. [campesterol]**	**[PLFA] vs. [stigmasterol]**	**[PLFA] vs. [β-sitosterol]**
DI, n=7	0.78	0.71	0.34	0.94	0.94	0.95
DT, n=3	0.48	0.67	0.55	0.76	0.97	1.0

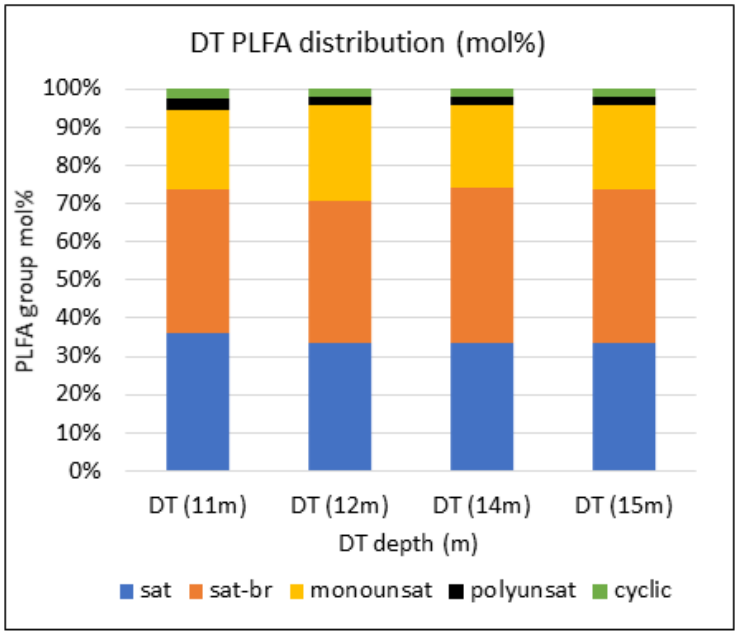
Table S5. Corrected radiocarbon signatures of sedimentary PLFA and organic carbon sources (SOC) at Desert Island (DI) and Doper Tek (DT).

Site	PLFA, Cell Abundance $\Delta^{14}C$ depth (m)	Corrected[c] $\Delta^{14}C_{PLFA}$ (‰) ±20[a]	PLFA ^{14}C age (years)	SOC $\Delta^{14}C$ depth (m)	$\Delta^{14}C_{SOC}$ (‰) ±20[a]	SOC ^{14}C age (years)
DI	10.7	25	200 (± 80)			
DI	16.8	22	220 (± 60)			
DI	22.9	9	320 (± 30)			
DI				40		6970 (± 45)[b]
DI				44		7180 (± 55)[b]
DT	12.2	-63	920 (± 100)	7.6	-910.23	19300 (±190)
DT	15.2	-30	635 (± 55)	13.1	-900.31	18450 (±200)

[a]A conservative error used for analyzing <150 µg Carbon samples using AMS (Pearson et al. 1998). [b]Reported by Radloff et al. (2007). [c] Formula used to correct for the methanolysis to convert PLFA to FAME (labelled Methanol $\Delta^{14}C$ = -997.7‰): $\Delta^{14}C=((N+1)*\Delta^{14}C_{measured} - \Delta^{14}C_{MeOH})/N$ where N=weighted average number of Carbon atoms in PLFA compounds. Formula from Brady et al. 2009, 2018.

Figure S2. Relative abundance reported as mol % of major PLFA groups measured in aquifer sediments from Desert Island (left) and Doper Tek (right).





Chapter 3: ***Using molecular PLFA and sterol biomarkers to distinguish viable microbial abundance and community structure, and sewage input potential in shallow and intermediate depth aquifers in Araihazar, Bangladesh.***

San Pedro, R.J.[1], Whaley-Martin, K.J.[2], Mailloux, B.J.[3], van Geen, A.[4], Bostick, B.C.[4], Ahmed, K.M.[5], Mozumder, R.[4], Ellis, T.[4], Slater, G.F.[1]

[1]School of Geography and Earth Sciences, McMaster University, ON, Canada
[2]Department of Civil and Mineral Engineering, University of Toronto, ON, Canada
[3]Environmental Sciences Department, Barnard College, NY, NY 10027
[4]Lamont-Doherty Earth Observatory, Columbia University, Pallisades, NY 10964
[5]Department of Geology, University of Dhaka, Dhaka 1000, Bangladesh

3.1 Abstract

The regional scale contamination of groundwater aquifers by arsenic in Bangladesh is controlled by microbially mediated dissimilatory dissolution of sedimentary iron oxides with concomitant arsenic release as driven by labile reduced carbon. In shallow Holocene aquifers in Araihazar, this labile carbon was argued to include human and/or animal sewage transported into the aquifer to stimulate microbial metabolism and increase arsenic. This hypothesis was tested in this present study using biomarker-based approaches (PLFA, sterols). The relationships between viable microbial biomass, sewage contamination, and arsenic were examined in-depth in a shallow Holocene and intermediate Pleistocene aquifer separated by an organic-rich confining clay aquitard (Site M). In order to assess the importance of sewage to drive microbial metabolism and arsenic release, organic sterols were profiles while Sewage Contamination Index and Coprostanol/Cholesterol ratio were used as aquifer sewage input proxies. The viable microbial abundance (range = 1.0x108 to 1.3x108 cells/g dry sediment), reflected by PLFA concentrations, and total organic sterols (range = 270 to 610 ng/g dry sediment) were both elevated in the shallow high arsenic Holocene aquifer. These values are comparable to a nearby similarly high arsenic site and suggests high organic carbon is available to drive microbial metabolism resulting in high biomass. An even greater viable microbial abundance was observed within the lower intermediate low arsenic Pleistocene aquifer. However, low total organic sterols (140 ng/g dry sediment) and the disappearance of phytosterols at this depth suggests that organic sterol inputs are different between lower Pleistocene and shallow Holocene aquifers. The high Coprostanol/Cholesterol ratio observed in the intermediate Pleistocene (contrary to the shallow Holocene) indicates significant sewage contamination at this depth. Sewage may be transported laterally from nearby shallow depths induced by municipal pumping from Dhaka, as previously argued at this site. As a source of labile reduced carbon stimulating microbial metabolism, this may explain the previously observed anomalous increase in arsenic immediately below the confining clay layer. This has implications for the long-term sustainability of presently pristine aquifers in Bangladesh and across Southeast Asia as anthropogenic impacts (e.g. overpumping) increase in these regions causing regional-scale hydrological changes in the aquifer.

3.2 Introduction

This study investigated microbial and carbon source biomarkers at a site where initially low Arsenic groundwater wells experienced sudden spikes in Arsenic concentrations. This study investigated both the shallow Holocene high Arsenic aquifer as well as the deeper low Arsenic Pleistocene aquifer. At the sub-basin scale (i.e. community-scale), the shallow Holocene and deep Pleistocene aquifers are hydrologically separated (Hoque, Burgess, Shamsudduha, & Ahmed, 2011; Hoque et al., 2009) and it is the shallow Holocene aquifers that generally have high Arsenic concentrations above the Bangladeshi water quality guideline (50 µg/L). Alternatively, deeper (>100 or >150 m) Pleistocene aquifers are commonly low in Arsenic leading experts to suggest that drilling deep community wells will provide a safer source of potable water (BGS & DPHE, 2001; Michael & Voss, 2008; Ravenscroft, McArthur, & Rahman, 2018). However, reports of sudden increases of Arsenic concentrations in tubewells tapping these deep Pleistocene aquifers have prompted investigations into the vulnerability of presently pristine deep aquifers as alternative groundwater sources (van Geen et al., 2007). Indeed, previous work investigated a community well located at Site M (Figure 1) that "failed" because of its sudden increase in high Arsenic levels (>50 µg/L As) from initially low (<10 µg/L As) values. Recent work proposed that this was driven by bioreactive carbon from a thick organic-rich clay layer which diffused and/or advected into the underlying Pleistocene aquifer promoting the reductive dissolution of iron oxides and concomitant Arsenic release (Mihajlov et al., in prep) (Figure 1). This hypothesis stems from previous studies which suggest that addition of labile organic carbon drives this microbially mediated Arsenic release mechanism (Dhar et al., 2011; Islam et al., 2004). Previous work has proposed that these labile carbon sources have

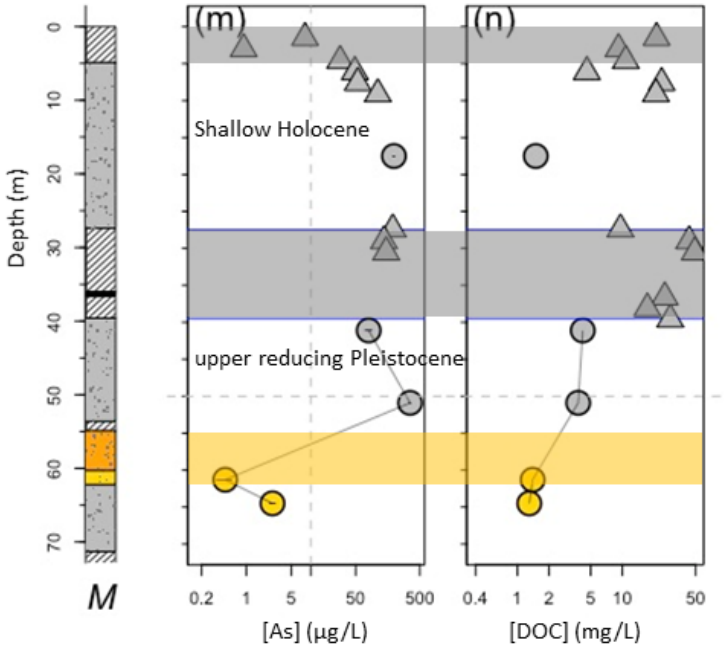


***Figure 1**. Site M aquifer lithology modified from Mozumder et al. (in prep). Grey intervals represent clay-rich layers. Orange interval represents oxidized sandy Pleistocene aquifer layer.*

surface origins and were transported into shallow aquifers to drive Arsenic contamination (Guo et al., 2019; Harvey et al., 2002; Mailloux et al., 2013; Neumann et al., 2010; Neumann, Pracht, Polizzotto, Badruzzaman, & Ali, 2014; Whaley-Martin et al., 2016). Advection of organic material from the subsurface clay would represent an additional mechanism to drive As release. Alternatively, advection of reactive carbon sourced from subsurface clay, which was argued to occur at Site M, represents an additional mechanism to drive Arsenic release at Site M, was argued to be exacerbated by municipal-scale pumping (Mihajlov et al., in prep; Mozumder et al., in prep). The impact of groundwater pumping on advective flow in aquifers from has been argued to be potentially exacerbated by distant pumping as demonstrated by the dramatic changes in aquifer hydrology immediately surrounding Dhaka (Knappett et al., 2016; Mihajlov et al., 2016)

Natural abundance radiocarbon isotope analysis has been used to trace carbon sources metabolized by microbial communities within these aquifers. The $\Delta^{14}C$ signatures of organic compounds in the environment confer an age fingerprint which elucidate the modern (i.e. atmospheric) or old (i.e. fossil) sources of carbon utilized in their production (Pearson, McNichol, Benitez-Nelson, Hayes, & Eglinton, 2001; Petsch, Eglinton, & Edwards, 2001; Shah, Mollenhauer, Ohkouchi, Eglinton, & Pearson, 2008). This method helps to constrain the carbon sources which support entire microbial communities in the environment. In Araihazar, for example, $\Delta^{14}C$ signature analysis of carbon pools in the aquifer distinguish whether the bacterial community relies predominantly on the old (negative $\Delta^{14}C$ values) clay-derived organic carbon or younger and perhaps allochthonous carbon (positive $\Delta^{14}C$ values) (Mailloux et al., 2013; Whaley-Martin et al., 2016; Chapter 2 of this thesis).

This study was part of ongoing research to assess the vulnerability of deeper Pleistocene aquifers which had been previously established to be generally lower in arsenic. In order to quantify the microbial abundance, characterize community profile/structure, and assess the potential for sewage input at Site M, two types of molecular biomarkers were used: phospholipid fatty acids (PLFA) and organic sterols. Molecular biomarkers are biosynthesized organic compounds which can be traced to specific sources or biological processes in the environment (Eglinton & Calvin, 1967; Eglinton, Scott, Belsky, Burlingame, & Calvin, 1964; Simoneit, 2004). PLFA are useful source-specific fingerprinting tools because they are important cell membrane components of bacterial and eukaryotic organisms. They rapidly degrade upon cell lysis providing an *in-situ* profile and bulk abundance of the viable biomass at the time of sample collection (D. C. White, Davis, Nickels, King, & Bobbie, 1979). Certain PLFA compounds

have also been used as diagnostic biomarkers of specific organisms in the natural environment. For example, C18:2ω6,9 have been used as biomarkers of fungal biomass in soils (Frostegård, Tunlid, & Bååth, 2011; Kaiser, Frank, Wild, Koranda, & Richter, 2010). Additionally, microorganisms change the composition or configuration of PLFA in their cell membranes in response to environmental stress such as changing moisture content, salinity, pH/alkalinity, or addition of contaminants (Green & Scow, 2000; Kieft, Ringelberg, & White, 1994). As a result, shifts in environmental conditions as it affects microbial communities can be inferred from bulk changes in the proportion of major PLFA groups (e.g. (Fierer, Schimel, & Holden, 2003; Zelles, Bai, Beck, & Beese, 1992; Zelles et al., 1994)). However, the low resolution and taxonomic specificity of PLFA distribution poorly identify species that make up microbial groups in the environment. In order to describe microbial members, more specific genetic-based analysis is necessary (e.g. (Zhang et al., 2017)). Nevertheless, PLFA are useful tracers of the structure of microbial communities in the environment which have enabled researchers to study *in situ* changes in different environments (Frostegård et al., 2011).

Previous studies have also proposed that inputs of labile carbon derived from unsewered domestic waste may be driving arsenic release at other sites in Bangladesh (McArthur, Sikdar, Hoque, & Ghosal, 2012; Whaley-Martin et al., 2017). In order to assess the potential for sewage input at depth at Site M, sedimentary sterol biomarkers were quantified and compared. Organic sterols in the natural environment are rapidly degraded at the surface where oxic conditions prevail but are better preserved in anoxic environments such as the saturated zone of an aquifer (Pratt, Warnken, Leeming, Arthur, & Grice, 2008). The preservation of organic sterols (and indeed, any hydrophobic organic compound) in anoxic and oligotrophic

sedimentary aquifers, can be attributed to their high partitioning coefficients (K_{ow}) which predicts their preferential sorption onto aquifer sediments and subsequent retardation with respect to groundwater flow into the aquifer (Chiou, Schmedding, & Manes, 1982; Froehner & Sánez, 2013; Hansch, Quinlan, & Lawrence, 1968; Whaley-Martin et al., 2017). The source-specificity and resistance to degradation of organic sterols make them useful tools to assess important sources of organic inputs in the environment (Bull, Matthew J. Lockheart, Mohamed M. Elhmmali, David J. Roberts, & Richard P. Evershed, 2002; Huang & Meinschein, 1979; Leeming & Nichols, 1998; Nishimura & Koyama, 1977). In this study, six individual sterol compounds were used to represent sources of organic input into the subsurface: Coprostanol, Cholesterol, Cholestanol, Campesterol, Stigmasterol, and β-sitosterol. Coprostanol is commonly produced from the biohydrogenation of Cholesterol as mediated by enteric bacteria in the digestive tract of upper mammals (e.g. humans, cattle) (Grimalt, Fernández, Bayona, & Albaigés, 1990; Rosenfeld, Fukushima, Hellman, & Gallagher, 1954) and has been reported to make up as much as 60% of total sterol content in human feces (Bull et al., 2002; Leeming, Ball, Ashbolt, & Nichols, 1996). However, because Coprostanol is also produced in small amounts by non-enteric microorganisms in the natural environment, sewage input proxies have been developed to determine the relative importance of sewage in comparison to other biogenic sources of sterols in nature. For example, the Sewage Contamination Index (SCI), calculated by Coprostanol/(Cholestanol+Coprostanol) ratio (Grimalt et al., 1990). Values of SCI ≥ 0.7 confer significant sewage contribution into the sterol pool relative to other naturally occurring biogenic sterols in anoxic sedimentary aquifers (Grimalt et al., 1990; Reeves & Patton, 2005; Speranza, Colombo, Skorupka, & Colombo, 2018).

Another sewage index is based on calculating the Coprostanol/Cholesterol (C/Ch) concentration ratio; C/Ch ratios >1 indicate significant sewage contamination in the environment (Fattore, Benfenati, Marelli, Cools, & Fanelli, 1996). While Cholesterol is also abundant in sewage (Isobe et al., 2002), its enteric bacterial degradation to Coprostanol implies a lower relative concentration in environments contaminated by human and/or animal sewage. Relative concentrations of these sewage sterols are compared with plant-derived phytosterols, Campesterol, Stigmasterol, and β-Sitosterol to determine dominant sources of natural organic sterols within the Site M aquifer. In plant cells, phytosterols such as Campesterol, β-Sitosterol and Stigmasterol are used to regulate membrane fluidity – a similar biochemical function is exercised by Cholesterol in animal cells (Hartmann, 1998). For example, a study of the impact of milling on the nutrient content of rice determined that in every degree of processing, Campesterol, β-Sitosterol and Stigmasterol were consistently the most abundant phytosterols in rice (Ha et al., 2006). It is therefore reasonable to expect that at terrestrial sites such as Bangladesh aquifers above which rice and wheat crop fields abound, the organic sterol pool is dominated by phytosterols. At previous sites in Araihazar, increasing sewage contamination correlates with increasing dissolved Arsenic concentrations at depth which suggests that sewage may be an important source of organic carbon driving microbially mediated Arsenic release in the aquifer (Whaley-Martin et al., 2017). This is conceivable given the ubiquity of domestic pit latrines and shallow ponds at the surface providing an available source of sewage into the aquifer (Knappett et al., 2011). By quantifying source-specific molecular biomarkers and investigating complementary geochemical and radiocarbon data, this research chapter looks at microbial community characteristics (e.g. bulk abundance and structure) as well as

sewage input to determine and constrain the extent to which microbial metabolism and carbon cycling control Arsenic release in shallow and intermediate aquifers at Site M.

3.3 Methodology

3.3.1 Study site

The study site, Site M (long 90.63235, lat 23.77612), is located in Araihazar upazila, Bangladesh about 20km east of Dhaka within a 1-2.5 km radius from previous sites DI, Site B, and Site F and 10 km southwest of DT (Figure S1). The sedimentary matrix cores analyzed for molecular biomarkers in this study were collected within 10 m of a community groundwater well, C12, installed in 2003. CW12 was one of four groundwater community wells in Araihazar which "failed" when groundwater Arsenic concentrations, which were initially <10 μg/L from July 2003 to October 2004, suddenly increased to 230 μg/L over a two month period then followed by a decrease in Arsenic eventually stabilizing at ~60 μg/L (van Geen et al., 2007). This prompted further work which proposed that impermeable clay layers provided a source of labile carbon which stimulated the reductive dissolution and concomitant As release into groundwater (Mihajlov et al., 2019, in prep). Two distinct clay layers were previously described within the grey Holocene aquifer at Site M: the first, an upper clay layer which extends from the surface to ~8 m below the surface and a second thick lower clay layer from ~26 m to ~39 m depth (Mihajlov et al., 2019 (in prep); Mozumder et al., 2019 (in prep)). Using radiocarbon isotope analyses, these researchers determined that this lower clay layer divides the shallow sandy grey Holocene aquifer from the lower Pleistocene aquifer; the Pleistocene aquifer is further divided into a shallow or upper grey-colored portion and a lower orange-colored layer which occurs from ~55 m to ~67 m depth (Mihajlov et al., 2019; Mozumder et al. 2019, in prep).

The grey color of the aquifer sediments at these depths implies the dissolution of sufficient bulk orange-colored Fe(III) oxides from sediment surfaces culminating in highly reducing conditions which has been shown using diffuse spectral reflectance (Horneman et al., 2004). Monitoring wells reported high dissolved Arsenic (20 to 250 µg/L) within these grey-colored aquifer intervals (i.e. both shallow Holocene and the upper portion of the Pleistocene aquifers). Dissolved Arsenic concentrations rapidly decrease (mean = 6 µg/L) into the lower orange-colored Pleistocene aquifer (~55 m to ~67 m depth). Four individual monitoring wells were previously installed within a 100 m radius surrounding the failed CW12 well. Taken together, these four monitoring wells, the failed CW12 well, and the drill site from which the sedimentary matrix core was collected will henceforth be referred to collectively as 'Site M' (Figure S2).

The sedimentary matrix core from Site M captured a total depth interval from 7.2 m to 42.8 m and was gray to dark grey in color throughout its length. The sediment lithology was medium to coarse sand along the core length; a very coarse sandy layer occurs at 26.8 m. Below this, a very fine-grained clay-rich layer is present at 37.9 m. Plant debris was found in the shallow grey Pleistocene aquifer at 42.8 m depth while peat fragments were present at 22.9 m and 26.8 m depth. Arsenic concentrations were reported to fluctuate but ultimately increase from 40 µg/L in 2011 to 150 µg/L in 2017 at 41 m depth just below the confining clay aquitard layer (Mihajlov et al., in prep; Mozumder et al., in prep). Solid organic content (sedimentary organic carbon, SOC) within the lower clay layer analysed by Mihajlov et al. (in prep) were as high as 7% while dissolved organic carbon (DOC) peaked at 23 mg/L, these values are considered high concentrations in these oligotrophic aquifers.

3.3.2 Collecting and quantifying PLFA biomarkers

Sediment samples were freeze-dried and stored in sterile Whirlpak bags. Approximately 100 g of sediment (dry weight) per sample were extracted using a modified Bligh & Dyer room temperature solvent extraction method. Microbial biomarkers (PLFA and sterols) were extracted from freeze-dried aquifer sediments using a modified Bligh & Dyer solvent extraction method similar to (Whaley-Martin et al., 2016). Briefly, a mixture of 1:2:0.8 dichloromethane (DCM), methanol, and phosphate buffer mixture were used to dissolve polar PLFA and sterol compounds from sediment samples (Guckert, Antworth, Nichols, & White, 1985; D. C. White et al., 1979); sediment samples were extracted for 24 hrs twice. After each extraction, the supernatant was collected and filtered to remove solids. The extract was transferred into separatory funnels and Methanol and DCM were added to separate phases and the organic phase containing the lipid biomarkers were collected as the total lipid extract (TLE). Silica gel column chromatography was used to separate the extractable organic compounds contained in the TLE into three fractions based on polarity (Guckert et al., 1985): F1 – neutral lipids (DCM), F2 – uncharged polar lipids (Acetone), F3 – very polar or charged lipids (Methanol) (Vestal & White, 1989). Phospholipids elute in F3 while sterols elute in F1 and F2. The F3 sample fraction was concentrated and mild alkaline (KOH) methanolysis was performed to cleave phosphate head groups from at their ester bonds and attaching a methyl group to the lipid fatty acid chain creating fatty acid methyl esters (FAME). These FAMEs are amenable for gas chromatograph-mass spectrometer (GC-MS) analysis (D. White, Stair, & Ringelberg, 1996). A secondary silica gel chromatography was used as a “clean up” step to further separate FAME compounds from other lipid components which eluted in F3 during the primary column chromatography

(glycerols, archaeal lipids). The three fractions distinguished based on polarity are F1 (4:1 Hexane:DCM), F2 (DCM), and F3 (MeOH). The FAME samples, eluted in F3, were concentrated to 1mL samples prior to GC-MS analysis.

PLFA and sterol biomarkers were quantified using an Agilent 6890N Gas Chromatograph (30m × 0.25mm DB5-ms DG, 0.25µm film thickness) coupled to a 5973 quadrupole mass spectrometer monitoring for masses (50-450 m/z for FAME and 50-500 m/z for TMS-sterols). For the PLFA analysis, 1µL sample injections were analyzed at an initial temperature of 50^{o}C held for 1 min then increased up to 130^{o}C at a rate of 20^{o}C/min then a further increase to 160^{o}C at a rate of 4^{o}C/min and finally a final ramp at a rate of 8^{o}C/min until a final maximum temperature of 300^{o}C was achieved and held for 5 min. Fatty acid compounds were identified using two FAME mixture standards (26 component Sigma-Aldrich, and 37 component Supelco Inc.).

3.3.3 Extraction and derivatization of organic sterol biomarkers

To analyze sterols, the column chromatography fraction F2 was evaporated to dryness under $N_{2(g)}$ and redissolved in DCM then combined with F1 (Birk, Dippold, Wiesenberg, & Glaser, 2012; Whaley-Martin et al., 2017). A trimethylsilyl (TMS) derivatization reaction was performed to convert intact sterols to TMS-derivatives amenable for GC-MS analysis. The samples were first concentrated to 100uL, 30uL of the derivatizing agent N,O-Bis(trimethylsilyl)trifluoro acetamide (BSTFA) (Sigma Aldrich) was then added along with 10uL of pyridine as a catalyst. Samples were heated at 60^{o}C for 30 minutes to ensure a complete derivatization reaction before analysis using GC-MS (Whaley-Martin et al., 2017).

The GC-MS program used to analyze the derivatized sterols was similar to that used to analyze PLFA. Specifically, 2µL injections where analyzed at an initial temperature of 50^{o}C held for 1 min then increased at 4^{o}C/min up to a final temperature of 300^{o}C which was held for 24 min. Sterol compounds were identified using Sigma-Aldrich sterol compound standards: 5β-cholestan-3β-ol (coprostanol), cholest-5-en-3β-ol (cholesterol), 5α-cholestan-3β-ol (cholestanol), 24-methylcholest-5-en-3β-ol (campesterol), 24-ethylcholesta-5,22(*E*)-dien-3β-ol (stigmasterol), 24-ethylcholest-5-en-3β-ol (β-sitosterol). PLFA and sterol sample compounds were identified using the manufactured standards in tandem with characteristic fragmentation patterns and mass spectral identification using the National Institute of Standards and Technology (NIST) compounds library. The limit of detection is 0.5 ng/µL – 25 ng/µL for all sterols. Because some samples were well above 25 ng/µL, concentrated samples were diluted within the limit of detection and reanalyzed in the GC-MS.

3.3.4 Dissolved groundwater arsenic measurements

Dissolved groundwater arsenic was measuring using a high resolution inductively coupled plasma mass spectrometer (HR ICP-MS) at the Lamont Doherty Earth Observatory (Columbia University) as per the methods of (Cheng, Zheng, Mortlock, & Van Geen, 2004). Lower detection limits were <0.1 µg/L.

3.4 Results

3.4.1 Microbial abundance and PLFA profiles at Site M

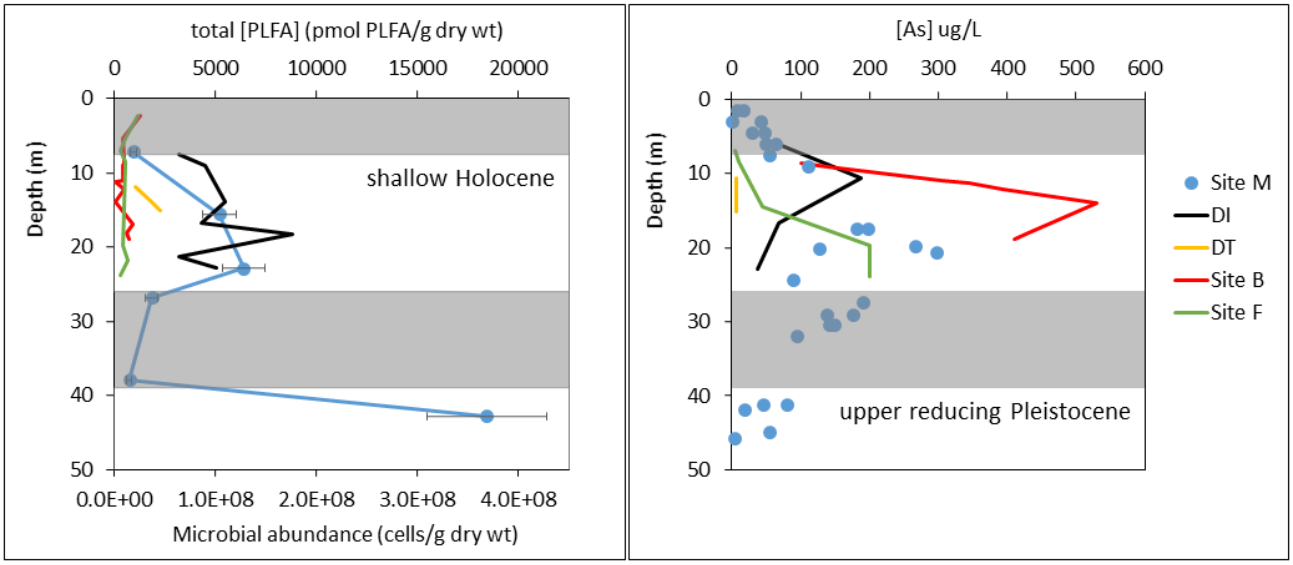


Figure 2. *PLFA concentrations, estimated cell abundance, and dissolved groundwater Arsenic (As) at five study sites in Araihazar upazila, Bangladesh. PLFA and Arsenic concentrations at Site B and F retrieved from Whaley-Martin et al. (2016). Arsenic concentrations at Site M retrieved from Mihajlov et al. (in prep) and Mozumder et al. (in prep).*

Total PLFA concentrations at Site M ranged from an average of 750 ± 120 pmol PLFA/gram dry sediment (gds) to 18000 ± 2900 pmol PLFA/gds (Figure 3). Using a generic cell conversion factor of 2×10^4 cells/pmol PLFA (Green & Scow, 2000), this is equivalent to a cell abundance range of 2×10^7 to 4×10^8 cells/gds. Microbial cell abundance was initially low at 7.2 m – comparable to nearby sites B, F, and DT – and increased with depth along the shallow sandy Holocene aquifer (Figure 3) (Whaley-Martin et al., 2016; Chapter 2, this thesis). Within the lower thick clay aquitard, a rapid decrease in microbial abundance occurs at 26.8 m which continues to the bottom of the clay layer at 37.9 m. Immediately below this clay layer (42.8 m), the maximum PLFA concentration was detected (18000 nmol PLFA/gds) equivalent to microbial cell abundance of 4×10^8 cells/gds at 42.8 m, the highest concentration reported in Araihazar.

This cell concentration is 10^1 – 10^3 greater than previous sites, B and F (Whaley-Martin et al., 2016). While elevated total or bulk PLFA correlated with arsenic in the shallow Holocene aquifer at Site M (Figure 2), other sites (e.g. Site B, F) with similarly high arsenic had much lower PLFA and thus, low viable biomass. Within the upper clay layer at Site M, both average total PLFA and dissolved arsenic were low; whereas within the lower thick clay layer, PLFA was low while arsenic was elevated. At intermediate depths (below 39 m) within the upper grey Pleistocene aquifer where maximum total viable biomass at this site occurs, dissolved arsenic decreases to values comparable to the upper clay above the Holocene aquifer.

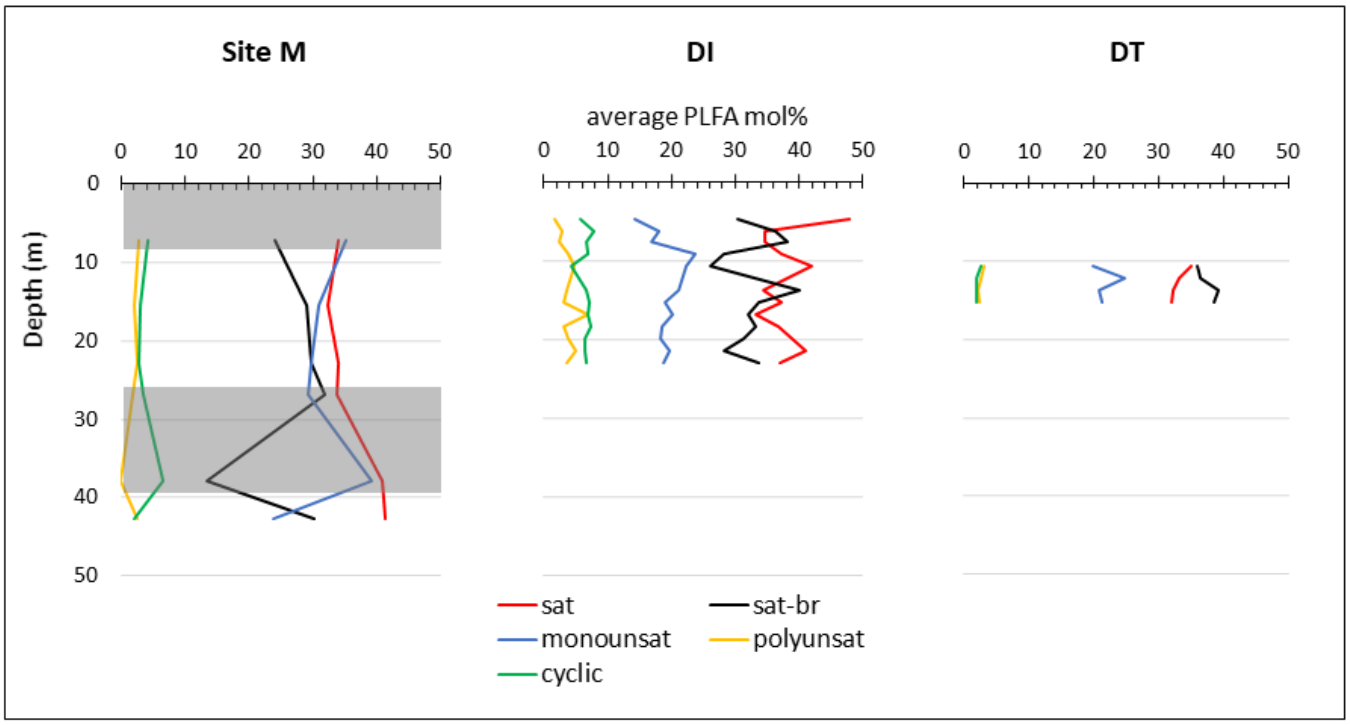


Figure 3. *Changes in relative abundance (mol%) of major PLFA groups along depth at Site M, DI, and DT. The five major PLFA groups are saturated (sat), branched saturated (sat-br), monounsaturated, polyunsaturated, and cyclic PLFA. PLFA mol% at DI and DT modified from Chapter 2, this dissertation.*

The microbial community structure within Site M was described by the relative abundance of five major PLFA groups. These were quantified by summing the mol% of individual FAME compounds into the appropriate groups: saturated straight chain, branched

saturated (i.e. straight chain with one methyl branch anywhere along the fatty acid chain), monounsaturated (i.e. straight chain PLFA with one carbon-carbon double bond anywhere along the fatty acid chain), polyunsaturated (i.e. straight chain PLFA with more than one double carbon-carbon bond anywhere along the fatty acid chain), and cyclic (i.e. straight chain PLFA with carbon rings (cyclopropyl) anywhere along the fatty acid chain). While the proportions of these PLFA groups varied along the Site M depth profile (Figure 3), their average relative abundance were 36% (saturated), 26% (branched saturated), 31% (monounsaturated), 2% (polyunsaturated), and 4% (cyclic). The relative abundance of all PLFA groups were relatively consistent along the shallow Holocene and upper grey Pleistocene at Site M. Notably, the relative abundances are significantly different within the upper and lower clay units (Figure 3). For example, at the base of the lower clay unit (37.9 m), branched saturated PLFA decreased by ~20% coincident with an increase in monounsaturated PLFA as well as a slight increase in saturated PLFA. Within the upper grey Pleistocene aquifer (42.8 m) just below this clay unit, the proportions of all PLFA groups return to values similar to the shallow Holocene aquifer.

The average proportions of saturated, polyunsaturated, and cyclic PLFA at Site M were similar to previous sites, DI (sat = 38%; polyunsat = 4%; cyc = 7%) and DT (sat = 33%; polyunsat = 3%; cyc = 2%). However, the proportions of these PLFA groups fluctuate significantly along the shallow aquifer intervals at these previous sites; whereas the same PLFA groups remain relatively constant along similar depths at Site M. Monounsaturated PLFA was a noticeably greater component of the microbial community structure at Site M (31%) than at previous sites (DI = 19%, DT = 22%). Meanwhile, branched saturated PLFA at Site M (26%) are generally lower compared to DI (33%) or DT (38%). Five individual FAME compounds make up 60% of total

FAMEs at Site M – 16:0 (22.0 ± 3.6%), i15:0 (6.9 ± 1.5%), a15:0 (5.0 ± 2.0%), total 16:1 (12 ± 6.9%), total 18:1 (13 ± 1.8%). The relative abundance of 16:1 and 18:1 were calculated by adding the mol% of every 16:1 or 18:1 FAME compound, respectively. These five FAMEs, with the addition of 18:0, were also the most abundant FAMEs at Site B and Site F (~53%) (Whaley-Martin et al. 2017, unpublished thesis chapter) as well as DI (58%) and DT (57%) (Chapter 2, this dissertation). Furthermore, the relative abundance of each of the five most abundant FAMEs were comparable between Site M and these previous sites. Their detection at previous sites were argued to indicate the presence of bacterial and microeukaryotic cells which commonly produce these FAMEs as part of cell membranes (Whaley-Martin et al. 2017, PhD thesis). This is also likely the case at Site M and suggests a similarity in the bulk microbial community structure at all the sites studied thus far in Araihazar.

3.4.2 Sedimentary organic sterols at Site M

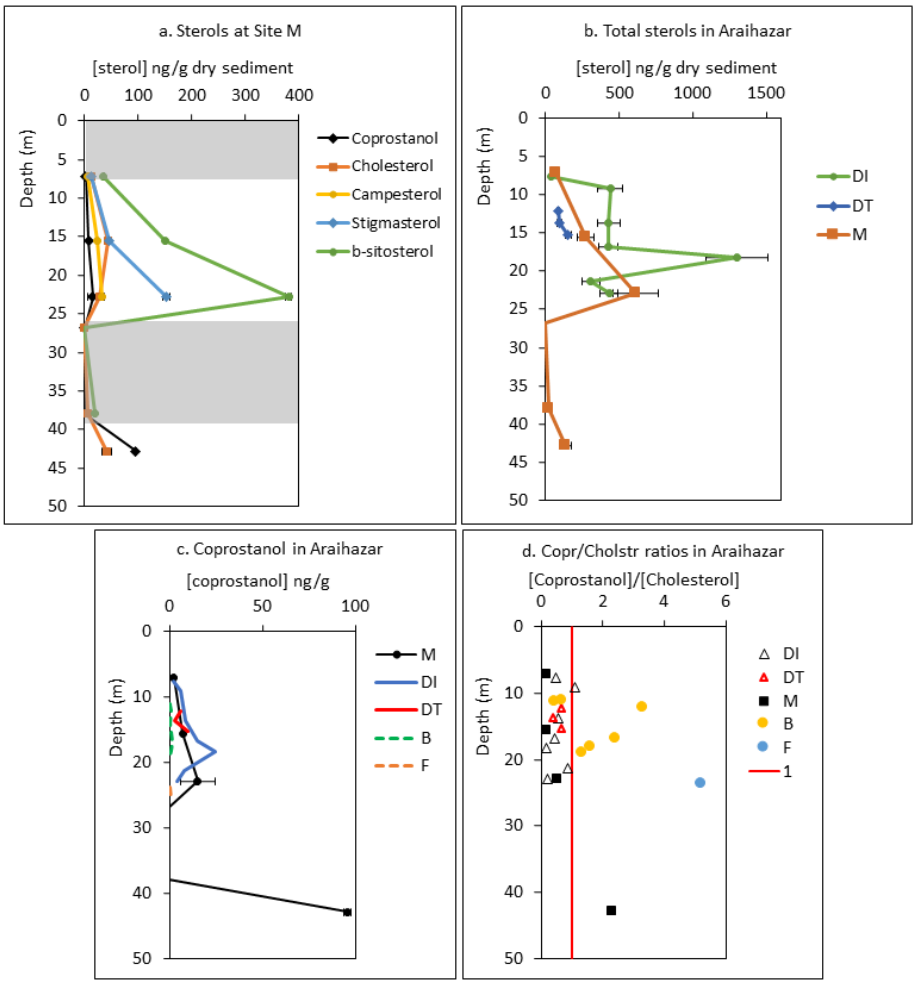


Figure 4. *a). Concentrations of individual sterol biomarkers measured at Site M and b). total sterol concentrations at DI, DT, and Site M. c). Coprostanol concentrations at the five sites studied in Araihazar: coprostanol was <LOD between 26.8 m and 37.9 m at Site M, Site B and F values were derived from Whaley-Martin et al. (2017). d). Coprostanol to cholesterol ratios calculated at the five sites in Araihazar as a proxy for sewage input into the aquifer (1 indicates threshold of sewage contamination).*

Total sterols were greater within the sandy aquifer layers (both Holocene and upper Pleistocene) compared to the clay aquitard layers at Site M. Of the three phytosterol biomarkers (Campesterol, Stigmasterol, and β-sitosterol), only Stigmasterol, and β-sitosterol, dominated the sterol profile in the shallow Holocene aquifer. Where it is present, Campesterol was usually lower than Cholesterol. This contrasts previous sites where all three phytosterols were consistently greater than the sewage related biomarkers Coprostanol and Cholesterol.

Total sedimentary organic sterols (the sum of each individual sterol biomarker measured) ranged from <LOD to 610 ± 150 ng/g dry sediment (gds) at Site M (Figure 4b). Unlike previous sites, Cholestanol was not detected at Site M. Nevertheless, these ranges are comparable to DI (35 ± 4 to 1300 ± 210 ng/gds) and DT (86 ± 10 to 150 ± 19 ng/gds) and were 10^1 to 10^2 greater than previous sites B (0.1 ± 0.0 to 5.2 ± 0.5 ng/gds) and F (0.1 ± 0.0 to 3.4 ± 0.6 ng/gds) (Whaley-Martin et al. 2017; Chapter 2, this dissertation). All organic sterols measured were very low within the lower confining clay layer: none of the sterol biomarkers measured was detected at 26.8 m while only Cholesterol (4.8 ± 0.1 ng/gds) and β-Sitosterol (19 ± 0.2 ng/gds) were detected at 37.9 m. Similar to the shallow aquifers at previous sites (DI, DT, B, F), phytosterols (Campesterol, Stigmasterol, β-sitosterol) comprised the greatest proportion of total sterols between 7.2 m to 37.9 m at Site M. Phytosterols at these depths at Site M ranged from 79% to 93% (mean 83%) which are, on average, comparable to DI (85%) and DT (80%) and higher than B (61%) and F (56%). Contrastingly, only β-sitosterol was detected at 26.8 m (within lower clay), while none of the three phytosterols were detected at 42.8 m (below clay layer).

Within the shallow Holocene at Site M, the concentration of the sewage biomarker Coprostanol ranged from 2.3 ± 0.1 to 15.0 ± 9.3 ng/gds which was comparable to similar depths at DI (2.2 ± 0.2 to 24.5 ± 2.4 ng/gds) and DT (2.9 ± 0.3 to 9.6 ± 1.0 ng/gds) (Figure 5c). It was not detected within the lower clay layer at 26.8 m and 37.9 m but suddenly peaked at 96 ± 1.6 ng/gds below this clay layer at 42.8 m. Coprostanol comprised a small proportion of total sterols (3%) within the shallow aquifer which was comparable to reports at previous similarly shallow aquifers, DI and DT. On the other hand, coincident with the rapid increase in absolute concentration, Coprostanol increased to 70% of total sterols at 42.8 m (upper grey Pleistocene)

which is comparable to maximum proportions of Coprostanol at other sites where total sterols were significantly lower, Site B (67%) and F (63%) (Whaley-Martin et al. 2017). Since Cholestanol was not detected at Site M, the Sewage Contamination Index (the ratio measuring relative abundance of sewage-derived coprostanol to cholestanol) could not be calculated. Instead, the Coprostanol/Cholesterol (Cop/Ch) ratios were calculated at depth as an alternative sewage input proxy. The Cop/Ch ratio was consistently <1 within the shallow Holocene aquifer but was not calculated within the lower clay layer where Coprostanol was below detection (Figure 5d). At 42.8 m, the Cop/Ch ratio was 2.3, well above the threshold value of 1 indicating significant sewage input. This Cop/Ch value within the lower Pleistocene at Site M is comparable to values within much of the shallow Holocene aquifers of Site B and F where significant sewage contamination has been proposed (Whaley-Martin et al. 2017).

3.5 Discussion

3.5.1 Distribution of viable biomass and sewage input potential across Araihazar

The variation in viable microbial abundance at Site M illustrated localized heterogeneity in aquifers with varying depths in Araihazar. This difference in bulk viable biomass can be interpreted to be a result of a stimulated microbial metabolism in the upper reducing Pleistocene relative to the shallow Holocene aquifer. The stimulated microbial community in the upper portion of the Pleistocene aquifer may be driven by greater input of labile carbon sources at this depth.

In the shallow Holocene aquifer at Site M, the much greater proportion of phytosterols compared to sewage-derived sterols indicated that plant-derived organic matter was the dominant organic sterol source at these shallow depths. This is consistent with findings at DI

and DT where phytosterols were the proportionally dominant sterol sources. The disappearance of the phytosterols at Site M – first, Campesterol and Stigmasterol at the base of the shallow Holocene (22.9 m), and eventually all three at the shallow grey Pleistocene (41.8 m) – coincides with the increasing dominance of sewage-derived sterols (Coprostanol, Cholesterol) with depth. This indicates that turnover occurs from predominantly plant-derived organic sterols in the shallow Holocene to sewage-derived sterols in the grey Pleistocene at intermediate depths. The complete absence of Cholestanol at Site M implies that the biohydrogenation pathways which breakdown Cholesterol to Cholestanol did not occur at any depth. This nulls the Sewage Contamination Index (SCI) necessitating the Cop/Ch ratio as an alternative sewage input proxy. Indeed, the Cop/Ch profile reinforced the depth trend suggested by the sterol distribution: the turnover from a sterol pool dominated by phytosterols to one which reflects the increase in sewage contamination at a deeper depth. While SCI could not be calculated at Site M, the elevated Cop/Ch ratio (>1) and absence of phytosterols in the upper Pleistocene indicates significant sewage contamination occurred at these depths. As compared to previous sites (Site B, F) where SCI and Cop/Ch ratios were both high (SCI >0.7, Cop/Ch >1), low Cop/Ch (<1) indicating little sewage input yet greater absolute Coprostanol concentration in the shallow aquifers of Site M, DI, and DT disputes the prevailing assumption that high Coprostanol concentration results from significant sewage contamination in the aquifer. Rather, the relative contribution of Coprostanol as compared to its Cholesterol precursor and other biogenic sterols must be considered. The observations in the upper Pleistocene at Site M suggest that sewage contamination may occur at greater depths than previously reported in Bangladesh.

3.5.2 Distribution of viable biomass and sewage input potential across Araihazar

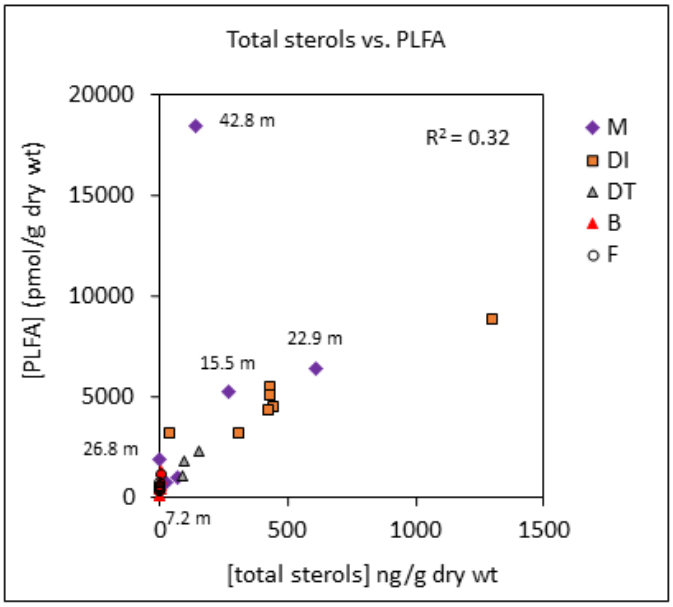


Figure 5. *PLFA and total sterol concentrations at five study sites in Araihazar. Depth labels included for data points from Site M (core range = 7.2 m to 42.8 m).*

In general, the total sterols correlate positively with total PLFA across the five sites (M, DI, DT, B, F) in Araihazar where data is available (Figure 5). This suggests an associative relationship between the viable microbial community and bulk organic sterols in the shallow aquifers in Araihazar. While microbial degradation of organic sterols cannot be discounted, their reduced bioavailability in the aquifer sediment matrix means sterols as a carbon substrate are unlikely to predominantly drive a linear relationship. However, it is possible that the microbial community utilize a labile component of the organic carbon pool that includes these organic sterols. For example, surface water recharge may contain both labile and recalcitrant (including sterols) dissolved organic matter that are collectively transported into depth (Mailloux et al., 2013; McArthur et al., 2012; Whaley-Martin et al., 2017). While the hydrophobic sterols preferentially partition to the sediment matrix, more labile components of

this DOC may be subsequently metabolised by *in situ* microbial communities. In Figure 5, the linear correlation coefficient r^2 value of 0.32 accounts for all data points including the outlier at 42.8 m, within the intermediate aquifer at Site M. Without this outlier at 42.8 m, the r^2 value is 0.86 and is visually reflected by the narrow spread of data in the graph. The data point at 42.8 m depth from Site M is the deepest point where PLFA and sterols have been quantified thus far in this region. Previous work has shown that the low-arsenic Pleistocene aquifer is considered to be hydrologically separated from the overlying shallow high-arsenic aquifer (Mihajlov et al., 2016). It is possible that as an outlier, the single data point at 42.8 m represents a glimpse of the significant differences in the bulk viable biomass and available organic source distribution between the shallow Holocene aquifers vs. deep Pleistocene aquifers in Bangladesh. This supports that the lower clay layer is an important source of reactive carbon which can stimulate microbial metabolism and drive the intermediate Pleistocene aquifer to reducing conditions as previously argued (Miahjlov et al, in prep; Mozumder et al., in prep).

3.5.3 Sewage input into the aquifer as impacted by groundwater pumping

If shallow groundwater that is laterally transported and enters the confined Pleistocene aquifer at Site M as was previously argued (Mozumder et al., in prep), this provides labile carbon to stimulate microbial metabolism at this depth. Shallow sources of sewage can come from excavated ponds with high faecal contamination (Knappett et al., 2011) or else surface water recharge which first enter shallow Holocene aquifers (Whaley-Martin et al., 2017) and then laterally transported into deeper depths via drawdown through regions where confining clay units are absent (Knappett et al., 2016; Mozumder et al., in prep). This explains the observed elevated viable biomass and Coprostanol concentration which drove the increase in

Cop/Ch ratio indicating sewage input stimulated microbial metabolism in the upper Pleistocene aquifer at Site M. An additional factor may be the concentrations of sedimentary organic carbon (SOC) within the aquifer sedimentary matrix; both PLFA and sterols are associated with SOC surfaces from organic matter which have undergone diagenesis or buried during the formation of the aquifer. Analysis of SOC concentrations is presently underway for sites B, F, DI, DT, and M.

This study demonstrated that sewage transported to depth may be an additional source of labile organic matter to potentially stimulate the microbial community metabolism in the deeper Pleistocene aquifer at Site M. However, absolute concentrations of raw sewage has not been quantified at this site and cannot be reasonably compared to the amount of reactive carbon released from the thick clay layer at Site M. Additional data such as Cl/Br ratios may be able to more specifically assess contributions of sewage into the aquifer and enable comparison with other sources of dissolved organic carbon (McArthur et al., 2012; Whaley-Martin et al., 2017).

3.5.4 Aquifer microbial community structures in Araihazar

The similarity of the relative abundance of the major PLFA groups and the proportion of the five most abundant FAMEs at Site M and previous sites (DI, DT, B, F) suggest similarities between the general microbial community structure at different sites across Araihazar. Furthermore, microbial community structures were comparable within the shallow Holocene and Pleistocene aquifers at Site M. This implies that the microbial community which drove the Holocene aquifer to highly reducing conditions with elevated arsenic is present within the intermediate Pleistocene aquifer below 39 m. This helps explain the anomalous increase in

arsenic (from 40 to 140 µg/L) which occurred over a 5-year period at 41.1 m depth immediately below the confining clay layer (Mozumder et al., in prep). If this is caused by the release of reactive carbon from the confining clay layer into the underlying Pleistocene aquifer due to pumping, the component of the microbial community capable of mediating arsenic release into groundwater (e.g. proportion of metal reducing bacteria) is stimulated by this reactive carbon. Further work will require more specific approaches such as genetic-based analyses to quantify the importance of metal reducing bacteria in the aquifer microbial community.

The variation in PLFA distribution along depth at Site M coincide with the heterogenous aquifer lithology at Site M (i.e. lithology as defined by sand-dominated aquifer vs. clay-rich aquitard). This is in contrast to previous sites which were all at shallow depths. For example, the proportions of PLFA groups fluctuated along depth at DI despite a relatively homogenous grey sandy aquifer. The most significant change at DI is the simultaneous decrease in branched saturated PLFA and increase in saturated and monounsaturated PLFA coincident with an elevation in arsenic at 10.7 m as shown in Figure 2. This change in distribution of PLFA groups at 10.7 m at DI resembled the observations within the lower clay layer (42.8 m) at Site M where arsenic was also elevated. This could suggest similar responses by the microbial community structure to changes in their environment at their respective sites. On the other hand, the thin clay layer at 13.7 m at DT did not impose notable shifts in the PLFA distribution along the narrow depth described here. However, the short PLFA profile at DT, compared to DI and Site M, is an inherently limited window within which significant changes in the microbial community structure can be characterized. Overall, the PLFA distributions at these aquifers suggests that the microbial community structure may experience localized changes along individual aquifers

(e.g. along Site M) while being relatively similar between different aquifer sites (e.g. across Araihazar). Interestingly, previous studies of sediments from DI showed a high capacity for arsenic to be released at similarly shallow depths (5 m and 12 m) (Radloff et al., 2007). However, the decoupled relationship between iron and arsenic in the aquifer has been complicated by resorption and subsequent decrease of groundwater arsenic. In a microcosm study at DI, this resorption has been shown to be linked specifically to additions of oxygen changing the reducing conditions necessary for arsenic to remain dissolved (Radloff et al., 2007). However, such a specific geochemical parameter as oxygen content to characterize the heterogeneity of the lithology at Site M was not expounded in this study. Since the variation of microbial abundance and arsenic appears to vary with an unknown yet important parameter of the subsurface lithology, this deserves further investigation to understand the potential for arsenic contamination in these aquifers.

3.5.5 Available carbon substrate and dissolved arsenic in Araihazar

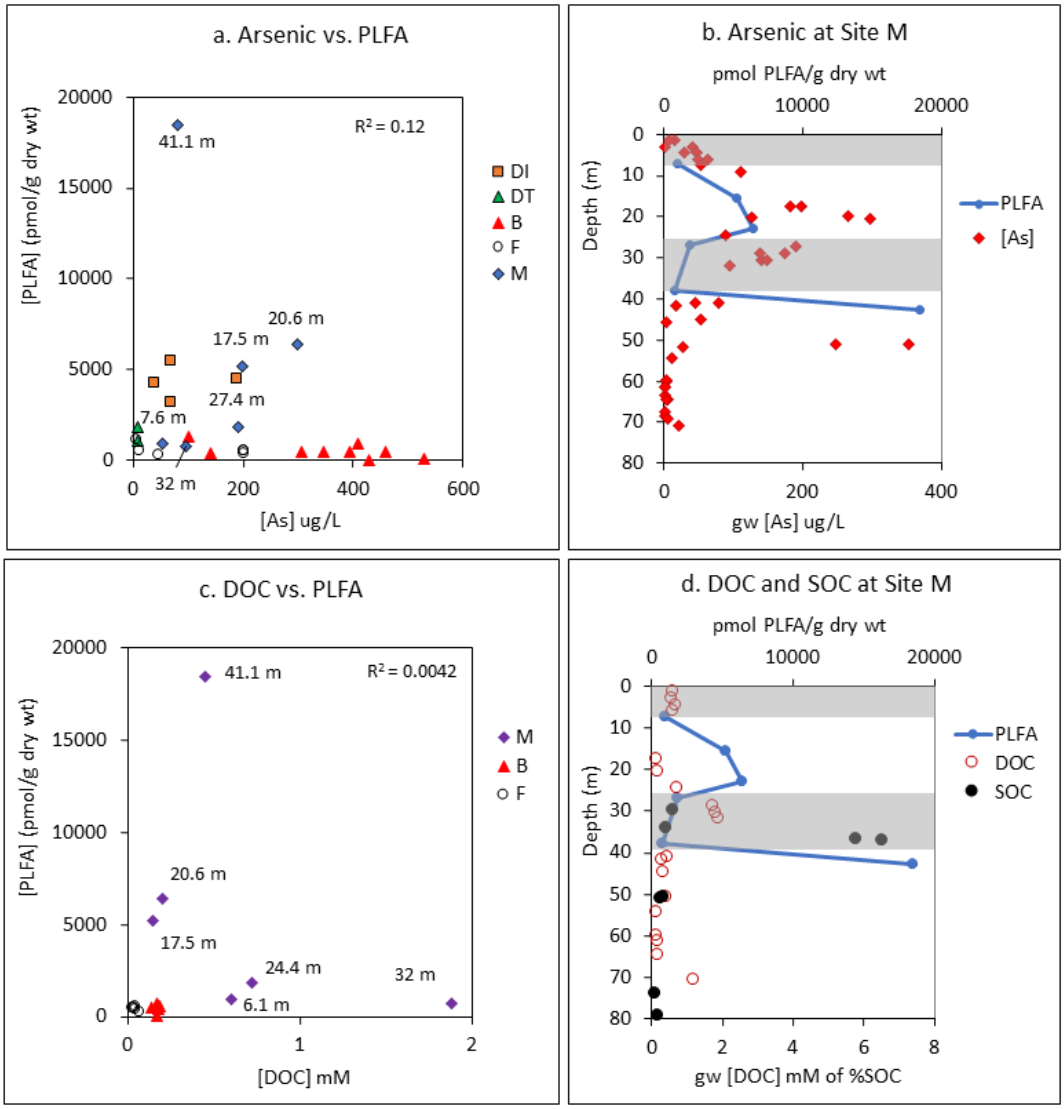


Figure 6*. Linear regression plots comparing dissolved groundwater Arsenic (As) vs. PLFA (a) and DOC vs. PLFA (c) at Site M and previous sites. Depth profiles of dissolved groundwater PLFA, As, DOC, and SOC at Site M (b and d). Dissolved As data at Site M from Mihajlov et al. (2019, in prep), at Site B and F from Whaley-Martin et al. (2016).*

Microbial PLFA is not linearly correlated with dissolved Arsenic in Araihazar (Figure 6a). At least at Site M, this poor correlation along depth is likely driven by the very low microbial biomass and high arsenic in the lower clay layer followed by a sudden and significant elevation in PLFA and rapid decrease in arsenic immediately below this clay aquitard. Upon closer inspection of the shallow Holocene aquifer at Site M, it is clear that groundwater arsenic is

concurrently elevated along with PLFA concentration. This trend is consistent with a previous site, DI, which has similarly elevated PLFA and arsenic. PLFA was also poorly correlated with bulk dissolved organic carbon (DOC) at all sites in Araihazar (Figure 6c). Previously, radiocarbon isotope analysis successfully demonstrated that microbial communities actively metabolize a young component of the DOC pool in Araihazar, as opposed to older carbon (i.e. depleted radiocarbon signatures) such as sedimentary organic carbon (SOC) (Mailloux et al., 2013; Whaley-Martin et al., 2016; Chapter 2, this dissertation). Because this trend has been demonstrated, with impressive precision, at four other sites (DI, DT, B, F), a regional-scale metabolism of young organic carbon by *in situ* microorganisms can be argued to occur across Araihazar. Further investigation will be required to confirm whether this trend is consistent at other regions in Bangladesh and perhaps Southeast Asia. Furthermore, the PLFA distributions at Site M and these previous sites suggest some similarity in their microbial community structure. It is therefore possible that a young component of the DOC supports the microbial community within the aquifer layers at Site M. While the PLFA profiles are similar between the shallow Holocene and upper portion of the Pleistocene aquifer at Site M, the previously discussed poor correlation between the biomarkers, dissolved arsenic, and DOC suggests a disconnect between the viable microbial community and increasing arsenic with increasing organic carbon substrate. As shown in Figures 5, 6a, and 6c, the data points at either 41.1 m (data points from Mihajlov et al. (in prep) and/or Mozumder et al. (in prep)) or 42.8 m depth at Site M contribute significantly to the poor correlation between PLFA and total sterols, PLFA and arsenic, and PLFA and DOC in Araihazar aquifers. Since Holocene and Pleistocene aquifers are more hydrologically connected than previously reported (Knappett et al., 2016; Mihajlov et al., 2016), it is possible

that other factors preclude the microbially mediated increase in arsenic contamination at these depths. Further investigations will be necessary to understand the specific factors which can preclude arsenic contamination in aquifers with very high viable biomass.

The poor correlation between PLFA and DOC at Site M is largely induced by the heterogeneity of total PLFA concentrations coincident with the heterogenous distribution of clay- and sand-dominated layers in the aquifer. Alternatively, the very large viable biomass in the upper portion of the Pleistocene aquifer may also be supported by another source of carbon in addition to DOC. The rapid increase of SOC at the base of the lower clay layer is close to the same depth where PLFA is very high (i.e. into the upper Pleistocene) (Figure 6d). Because of insufficient PLFA data points which coincide with SOC concentrations at depth, a correlative relationship between PLFA and SOC could not be appropriately assessed. While there is no SOC concentration value within the upper portion of the Pleistocene aquifer, the possibility of SOC as a primary carbon substrate to support the very large viable biomass at this depth cannot be discounted. To the best of our knowledge, this study is the first to quantify PLFA and sterol biomarkers and assess their relationships to organic carbon and arsenic in a geochemically reducing Pleistocene aquifer. These observations suggest that the biogeochemical factors which impact the viable microbial community (or else, is controlled by the microbial community) such as available organic carbon and dissolved arsenic may be different between the shallow Holocene and deep Pleistocene aquifers.

3.6 Implications of microbial community structure, carbon sources, and arsenic release

Future research in these aquifers will need to examine the variation of viable biomass and organic inputs as they relate to parameters that characterize the aquifer's lithology (e.g.

total organic carbon content, redox conditions). For example, the abundance of TOC may explain the correlation between PLFA and organic sterols and confirm their association with organic matter (i.e. sorbed onto sediment matrix). While this will not indicate that TOC is directly cycled by microbial communities, it will clarify differences in the organic content, and thus the significant concentration differences in the organic biomarkers (PLFA, sterols), between the sand- and clay-dominated aquifers at Site M. A TOC profile analysis of sediments from Site M is presently in progress.

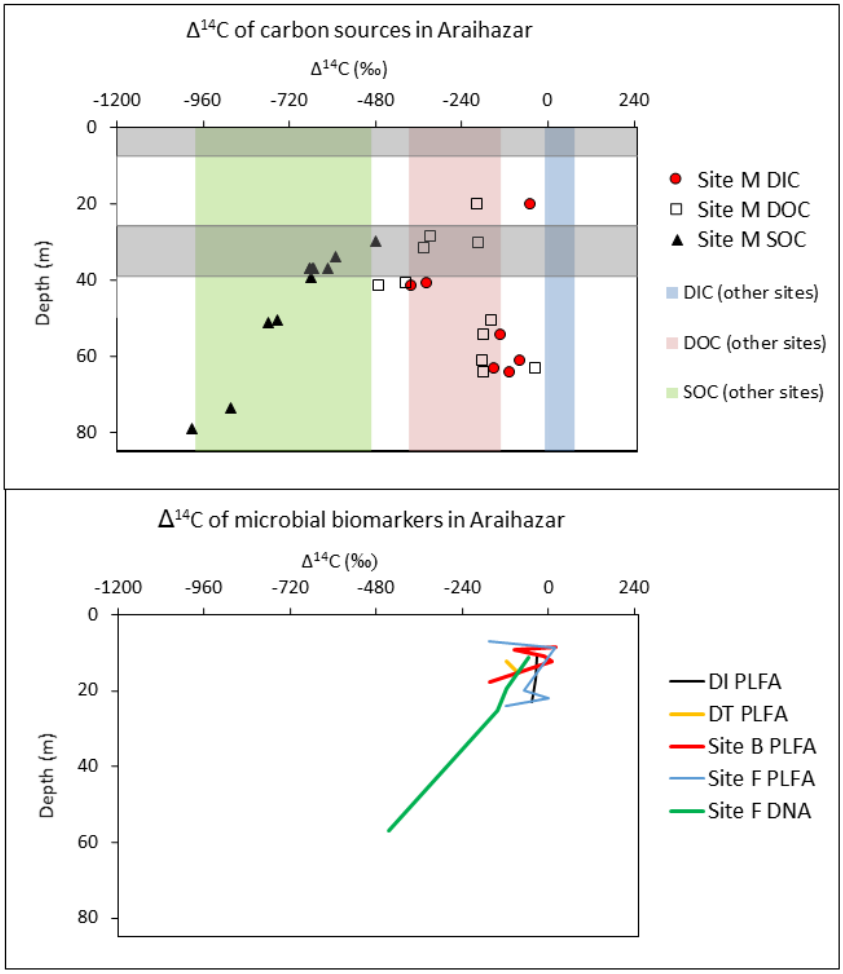


***Figure 7**. $\Delta^{14}C$ signatures of carbon sources (top) and biomarkers (bottom) in Araihazar. Gray horizontal bars = clay layers. $\Delta^{14}C$ values modified from Mihajlov et al. 2019, in prep (Site M), Whaley-Martin et al. 2016 and Mailloux et al. 2013 (Site B, F), and Chapter 2, this thesis (DI, DT).*

Previous use of natural abundance radiocarbon isotope ($\Delta^{14}C$) analysis has successfully delineated the relative age of carbon sources actively metabolized by *in situ* microbial communities in Araihazar (Mailloux et al., 2013; Whaley-Martin et al., 2016; Chapter 2, this dissertation). The well-constrained narrow range of radiocarbon signatures of microbial PLFA from previous sites ($\Delta^{14}C$ = 21‰ to -167‰) demonstrate a consistent source of carbon actively metabolized at different aquifers across Araihazar (Figure 7, bottom graph). This range is considerably narrow given the large variation captured by the radiocarbon signatures reported from three representative carbon pools in these aquifers: dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), and sedimentary organic carbon (SOC) (Figure 7, top graph). Radiocarbon analysis of PLFA from Site M will determine the age of carbon metabolized in both the Holocene and Pleistocene aquifers providing reconciling information about the importance of carbon sources derived from the confining clay unit (which will presumably have SOC radiocarbon signatures) and modern surface-derived sewage (presumably, DOC radiocarbon signatures). This work is currently in progress.

3.7 Conclusion

This study investigated the *in situ* viable microbial abundance, microbial community structure, and sewage input within an aquifer by using molecular biomarkers (phospholipid fatty acids or PLFA and sterols). The focus of this study is an aquifer at Site M where dissolved arsenic suddenly increased (to 60µg/L) in a community well after maintaining initially low concentrations (<10µg/L). This event prompted on-going work which assessed factors controlling the Arsenic cycling at depth. For this present study, the microbial abundance as described by PLFA concentrations varied significantly across a shallow Holocene aquifer and the

upper portion of a lower Pleistocene aquifer at Site M. The microbial abundance was usually greater within the sandy aquifer layers and significantly lower within the clay aquitard layers contributing to a heterogenous distribution of viable biomass with varying aquifer lithology. The significantly greater viable biomass within the upper Pleistocene aquifer compared to the shallow Holocene aquifer can be interpreted to be driven by input of reactive carbon from the overlying clay aquitard. This is consistent with a previous hypothesis which argued that reactive carbon from the thick clay aquitard diffused and/or advected (imposed by distant large-scale pumping) into the underlying Pleistocene aquifer causing an increase in the microbially mediated reductive dissolution of Fe(oxy)hydroxides and concomitant Arsenic release (Mihajlov et al., in prep; Mozumder et al., in prep). The sterol distribution at Site M points to a dual source input or organic sterol sources into these aquifers. The dominance of biogenic sources of sterols (e.g. phytosterols from plants and fungi) within the shallow Holocene aquifer at Site M is consistent with observations at other nearby shallow aquifers in Araihazar. On the other hand, the disappearance of these phytosterols and subsequent increase of quantitative indicators such as Coprostanol concentrations and the Coprostanol/Cholesterol ratios below this shallow aquifer indicate that sewage rapidly becomes an important source of organic sterols within the upper Pleistocene aquifer below the confining clay aquitard. Altogether, these observations demonstrate that the large viable biomass in the relatively deep upper portion of the Pleistocene aquifer is driven by inputs of reactive carbon from the overlying organic-rich clay aquitard as well as sewage sources within the aquifer. Along with previous studies at other sites in Araihazar, this study adds to the growing body of research which characterize the role of *in situ* microbial carbon cycling and Arsenic contamination of groundwater.

3.8 Statement of research contributions

All molecular biomarkers (PLFA and sterols) from Site M, Desert Island (DI), and Doper Tek (DT) sites were extracted, identified, and quantified by Reisa San Pedro at the Environmental Organic Geochemistry lab at McMaster University. Sediment matrix cores and groundwater (for arsenic measurements) samples were collected from Bangladesh and analyzed at the Lamont Doherty Earth Observatory by the Columbia University (Alexander van Geen, Tyler Ellis) and Barnard College (Brian Mailloux) team. Raw values of biomarker concentrations and radiocarbon data (PLFA, DNA, carbon pools) from previous sites (Site B and F) were retrieved from previous studies by Whaley-Martin et al. (2016, 2017) and Mailloux et al. (2012). Raw concentration values of groundwater arsenic and carbon sources (DIC, DOC, SOC) were retrieved from supplementary information of publications by Mihajlov et al. (2019) and Mozumder et al. (2019) in preparation for journal submission.

Bibliography

BGS, & DPHE. (2001). Arsenic contamination of groundwater in Bangladesh. *British Geological Survey, Department of Public Health Engineering*, *1*. Retrieved from http://nora.nerc.ac.uk/id/eprint/11986

Birk, J. J., Dippold, M., Wiesenberg, G. L. B., & Glaser, B. (2012). Combined quantification of faecal sterols, stanols, stanones and bile acids in soils and terrestrial sediments by gas chromatography-mass spectrometry. *Journal of Chromatography A*, *1242*, 1–10. https://doi.org/10.1016/j.chroma.2012.04.027

Bull, I. D., Matthew J. Lockheart, Mohamed M. Elhmmali, David J. Roberts, & Richard P. Evershed. (2002). The Origin of Faeces by Means of Biomarker Detection. Environ Int - Google Scholar. *Environment International*, *27*(8), 647–654. Retrieved from https://scholar.google.de/scholar?hl=en&q=%2C+Evershed+RP+%282002%29+The+Origin+of+Faeces+by+Means+of+Biomarker+Detection.+Environ+Int&btnG=&as_sdt=1%2C5&as_sdtp=

Cheng, Z., Zheng, Y., Mortlock, R., & Van Geen, A. (2004). Rapid multi-element analysis of groundwater by high-resolution inductively coupled plasma mass spectrometry. *Analytical and Bioanalytical Chemistry*, *379*(3), 512–518. https://doi.org/10.1007/s00216-004-2618-x

Chiou, C. T., Schmedding, D. W., & Manes, M. (1982). Partitioning of organic compounds in octanol-water systems. *Environmental Science & Technology*, *16*(1), 4–10. Retrieved from https://pubs.acs.org/sharingguidelines

Dhar, R. K., Zheng, Y., Saltikov, C. W., Radloff, K. A., Mailloux, B. J., Ahmed, K. M., & Van Geen, A. (2011). Microbes Enhance Mobility of Arsenic in Pleistocene Aquifer Sand from Bangladesh. *Environ. Sci. Technol*, *45*, 31. https://doi.org/10.1021/es1022015

Eglinton, G., & Calvin, M. (1967). Chemical Fossils. *Scientific American*, *216*(1), 32–43. https://doi.org/10.2307/24931372

Eglinton, G., Scott, P. M., Belsky, T., Burlingame, A. L., & Calvin, M. (1964). Hydrocarbons of Biological Origin from a one-Billion-Year-Old Sediment. *Science*, *145*(3629), 263–264. Retrieved from https://www-jstor-org.libaccess.lib.mcmaster.ca/stable/pdf/1713980.pdf?refreqid=excelsior%3A264061d827eea2c073504e3707e4141e

Fattore, E., Benfenati, E., Marelli, R., Cools, E., & Fanelli, R. (1996). Sterols in sediment samples from Venice Lagoon, Italy. *Chemosphere*, *33*(12), 2383–2393. https://doi.org/10.1016/S0045-6535(96)00340-2

Fierer, N., Schimel, J. P., & Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry*, *35*(1), 167–176. https://doi.org/10.1016/S0038-0717(02)00251-1

Froehner, S., & Sánez, J. (2013). Evaluation of potential sewage contamination by fecal sterol biomarkers adsorbed in natural biofilms. *Environmental Sciences: Processes and Impacts*, *15*(11), 2080–2086. https://doi.org/10.1039/c3em00025g

Frostegård, Å., Tunlid, A., & Bååth, E. (2011). Use and misuse of PLFA measurements in soils. *Soil Biology and Biochemistry*, *43*(8), 1621–1625. https://doi.org/10.1016/j.soilbio.2010.11.021

Green, C. T., & Scow, K. M. (2000). Analysis of phospholipid fatty acids (PLFA) to characterize

microbial communities in aquifers. *Hydrogeology Journal*, *8*(1), 126–141. https://doi.org/10.1007/s100400050013

Grimalt, J. O., Fernández, P., Bayona, J. M., & Albaigés, J. (1990). Assessment of Fecal Sterols and Ketones as Indicators of Urban Sewage Inputs to Coastal Waters. *Environmental Science and Technology*, *24*(3), 357–363. https://doi.org/10.1021/es00073a011

Guckert, J. B., Antworth, C. P., Nichols, P. D., & White, D. C. (1985). Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. *FEMS Microbiology Letters*, *31*(3), 147–158. https://doi.org/10.1016/0378-1097(85)90016-3

Guo, H., Li, X., Xiu, W., He, W., Cao, Y., Zhang, D., & Wang, A. (2019). Controls of organic matter bioreactivity on arsenic mobility in shallow aquifers of the Hetao Basin, P.R. China. *Journal of Hydrology*, *571*(November 2018), 448–459. https://doi.org/10.1016/j.jhydrol.2019.01.076

Ha, T. Y., Ko, S. N., Lee, S. M., Kim, H. R., Chung, S. H., Kim, S. R., … Kim, I. H. (2006). Changes in nutraceutical lipid components of rice at different degrees of milling. *European Journal of Lipid Science and Technology*, *108*(3), 175–181. https://doi.org/10.1002/ejlt.200500250

Hansch, C., Quinlan, J. E., & Lawrence, G. L. (1968). The linear free-energy relationship between partition coefficients and the aqueous solubility of organic liquids. *Journal of Organic Chemistry*, *33*(1), 347–350. https://doi.org/10.1016/0016-0032(63)90468-x

Hartmann, M. A. (1998). Plant sterols and the membrane environment. *Trends in Plant Science*, *3*(5), 170–175. Retrieved from https://pdf.sciencedirectassets.com/271825/1-s2.0-S0025326X00X03797/1-s2.0-0025326X9290685Y/main.pdf?x-amz-security-token=AgoJb3JpZ2luX2VjEEYaCXVzLWVhc3QtMSJIMEYCIQCXSiDwa%2Fd6HXN0k8rhzxBANLhLNLyb5Ed9JkQM%2FZC3nQIhAPshGqkICufpJAV1XvpUQKyJ80qBvSUQOxsxwvFRT

Harvey, C. F., Swartz, C. H., Badruzzaman, A. B. M., Keon-Blute, N., Yu, W., Ashraf Ali, M., … Ahmed, M. F. (2002). Arsenic Mobility and Groundwater Extraction in Bangladesh. *Science*, *298*, 1602–1606.

Hoque, M., Burgess, W. G., Shamsudduha, M., & Ahmed, K. M. (2011). Delineating low-arsenic groundwater environments in the Bengal Aquifer System, Bangladesh. *Applied Geochemistry*, *26*(4), 614–623. https://doi.org/10.1016/j.apgeochem.2011.01.018

Hoque, M., Khan, A., Shamsudduha, M., Hossain, M. S., Islam, T., & Chowdhury, S. H. (2009). Near surface lithology and spatial variation of arsenic in the shallow groundwater: Southeastern Bangladesh. *Environmental Geology*, *56*(8), 1687–1695. https://doi.org/10.1007/s00254-008-1267-3

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. https://doi.org/10.1016/j.gca.2004.01.026

Huang, W. Y., & Meinschein, W. G. (1979). Sterols as ecological indicators. *Geochimica et Cosmochimica Acta*, *43*(5), 739–745. https://doi.org/10.1016/0016-7037(79)90257-6

Islam, F. S., Gault, A. G., Boothman, C., Polya, D. A., Charnock, J. M., Chatterjee, D., & Lloyd, J. R. (2004). Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature*, *430*(6995), 68–71. https://doi.org/10.1038/nature02638

Isobe, K. O., Tarao, M., Zakaria, M. P., Chiem, N. H., Minh, L. Y., & Takada, H. (2002).

Quantitative application of fecal sterols using gas chromatography - Mass spectrometry to investigate fecal pollution in tropical waters: Western Malaysia and Mekong Delta, Vietnam. *Environmental Science and Technology*, *36*(21), 4497–4507. https://doi.org/10.1021/es020556h

Kaiser, C., Frank, A., Wild, B., Koranda, M., & Richter, A. (2010). Negligible contribution from roots to soil-borne phospholipid fatty acid fungal biomarkers 18:2ω6,9 and 18:1ω9. *Soil Biology and Biochemistry*. https://doi.org/10.1016/j.soilbio.2010.05.019

Kieft, T. L., Ringelberg, D. B., & White, D. C. (1994). Changes in Ester-Linked Phospholipid Fatty Acid Profiles of Subsurface Bacteria during Starvation and Desiccation in a Porous Medium. *Applied and Environmental Microbiology*, *60*(9), 3292–3299. Retrieved from http://aem.asm.org/

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. https://doi.org/10.1016/j.scitotenv.2011.04.043

Knappett, P. S. K., Mailloux, B. J., Choudhury, I., Khan, M. R., Michael, H. A., Barua, S., … van Geen, A. (2016). Vulnerability of low-arsenic aquifers to municipal pumping in Bangladesh. *Journal of Hydrology*, *539*, 674–686. https://doi.org/10.1016/j.jhydrol.2016.05.035

Leeming, R., Ball, A., Ashbolt, N., & Nichols, P. (1996). Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. *Water Research*, *30*(12), 2893–2900. https://doi.org/10.1016/S0043-1354(96)00011-5

Leeming, R., & Nichols, P. (1998). Determination of the sources and distribution of sewage and pulp fibre-derived pollution in the Derwent Estuary, Tasmania, using sterol biomarkers. *Marine Freshwater Research*, *49*, 7–17. https://doi.org/10.1071/MF99078

Mailloux, B. J., Trembath-Reichert, E., Cheung, J., Watson, M., Stute, M., Freyer, G. A., … van Geen, A. (2013). Advection of surface-derived organic carbon fuels microbial reduction in Bangladesh groundwater. *Proceedings of the National Academy of Sciences*, *110*(14), 5331–5335. https://doi.org/10.1073/pnas.1213141110

McArthur, J. M., Sikdar, P. K., Hoque, M. A., & Ghosal, U. (2012). Waste-water impacts on groundwater: Cl/Br ratios and implications for arsenic pollution of groundwater in the Bengal Basin and Red River Basin, Vietnam. *Science of the Total Environment*, *437*, 390–402. https://doi.org/10.1016/j.scitotenv.2012.07.068

Michael, H. A., & Voss, C. I. (2008). Evaluation of the sustainability of deep groundwater as an arsenic-safe resource in the Bengal Basin. *Proceedings of the National Academy of Sciences*, *105*(25), 8531–8536. https://doi.org/10.1073/pnas.0710477105

Mihajlov, I., Stute, M., Schlosser, P., Mailloux, B. J., Zheng, Y., Choudhury, I., … van Geen, A. (2016). Recharge of low-arsenic aquifers tapped by community wells in Araihazar, Bangladesh, inferred from environmental isotopes. *Water Resources Research*, *52*(5), 3324–3349. https://doi.org/10.1002/2015WR018224\

Mihajlov, I., Mozumder, M. R. H., Bostick, B. C., Stute, M., Mailloux, B. J., Knappett, P. S. K., … van Geen, A. (2019). Arsenic contamination of Bangladesh aquifers excacerbated by clay layers. Manuscript in preparation.

Mozumder, M. R. H., Michael, H. A., Mihajlov, I., Khan, M. R., Bostick, B. C., Mailloux, B. J., … van Geen, A. (2019). Origin of groundwater arsenic in a Pleistocene aquifer of Bangladesh

depressurized by distal municipal pumping. Manuscript in preparation.
Neumann, R. B., Ashfaque, K. N., Badruzzaman, A. B. M., Ashraf Ali, M., Shoemaker, J. K., & Harvey, C. F. (2010). Anthropogenic influences on groundwater arsenic concentrations in Bangladesh. *Nature Geoscience*, *3*(1), 46–52. https://doi.org/10.1038/ngeo685
Neumann, R. B., Pracht, L. E., Polizzotto, M. L., Badruzzaman, A. B. M., & Ali, M. A. (2014). Biodegradable Organic Carbon in Sediments of an Arsenic-Contaminated Aquifer in Bangladesh. *Environmental Science and Technology Letters*, *1*(4), 221–225. https://doi.org/10.1021/ez5000644
Nishimura, M., & Koyama, T. (1977). The occurrence of stanols in various living organisms and the behavior of sterols in contemporary sediments. *Geochimica et Cosmochimica Acta*, *41*(3), 379–385. https://doi.org/10.1016/0016-7037(77)90265-4
Pearson, A., McNichol, A. P., Benitez-Nelson, B. C., Hayes, J. M., & Eglinton, T. I. (2001). Origins of lipid biomarkers in Santa Monica Basin surface sediment: A case study using compound-specific Δ14C analysis. *Geochimica et Cosmochimica Acta*, *65*(18), 3123–3137. https://doi.org/10.1016/S0016-7037(01)00657-3
Petsch, S. T., Eglinton, T. I., & Edwards, K. J. (2001). 14C-dead living biomass: Evidence for microbial assimilation of ancient organic carbon during shale weathering. *Science*, *292*(5519), 1127–1131. https://doi.org/10.1126/science.1058332
Pratt, C., Warnken, J., Leeming, R., Arthur, M. J., & Grice, D. I. (2008). Degradation and responses of coprostanol and selected sterol biomarkers in sediments to a simulated major sewage pollution event: A microcosm experiment under sub-tropical estuarine conditions. *Organic Geochemistry*, *39*(3), 353–369. https://doi.org/10.1016/j.orggeochem.2007.06.009
Radloff, K. A., Cheng, Z., Rahman, M. W., Ahmed, K. M., Mailloux, B. J., Juhl, A. R., … Van Geen, A. (2007). Mobilization of arsenic during one-year incubations of grey aquifer sands from Araihazar, Bangladesh. *Environmental Science and Technology*, *41*(10), 3639–3645. https://doi.org/10.1021/es062903j
Ravenscroft, P., McArthur, J. M., & Rahman, M. S. (2018). Identifying multiple deep aquifers in the Bengal Basin: Implications for resource management. *Hydrological Processes*, *32*(24), 3615–3632. https://doi.org/10.1002/hyp.13267
Reeves, A., & Patton, D. (2005). Faecal sterols as indicators of sewage contamination in estuarine sediments of the Tay Estuary, Scotland: an extended baseline survey. *Hydrology and Earth System Sciences*, *9*(1/2), 81–94. https://doi.org/10.5194/hess-9-81-2005
Rosenfeld, R. S., Fukushima, D. K., Hellman, L., & Gallagher, T. F. (1954). The transformation of cholesterol to coprostanol. *Journal of Biological Chemistry*, *211*(1), 301–311.
Shah, S. R., Mollenhauer, G., Ohkouchi, N., Eglinton, T. I., & Pearson, A. (2008). Origins of archaeal tetraether lipids in sediments: Insights from radiocarbon analysis. *Geochimica et Cosmochimica Acta*, *72*(18), 4577–4594. https://doi.org/10.1016/j.gca.2008.06.021
Simoneit, B. R. T. (2004). Biomarkers (molecular fossils) as geochemical indicators of life. *Advances in Space Research*, *33*(8), 1255–1261. https://doi.org/10.1016/j.asr.2003.04.045
Speranza, E. D., Colombo, M., Skorupka, C. N., & Colombo, J. C. (2018). Early diagenetic alterations of sterol biomarkers during particle settling and burial in polluted and pristine areas of the Rio de la Plata Basin. *Organic Geochemistry*, *117*, 1–11. https://doi.org/10.1016/j.orggeochem.2017.11.013

van Geen, A., Cheng, Z., Jia, Q., Seddique, A. A., Rahman, M. W., Rahman, M. M., & Ahmed, K. M. (2007). Monitoring 51 community wells in Araihazar, Bangladesh, for up to 5 years: Implications for arsenic mitigation. *Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering*, *42*(12), 1729–1740. https://doi.org/10.1080/10934520701564236

Vestal, J. R., & White, D. C. (1989). Lipid Analysis in Microbial Ecology. *BioScience*, *39*(8), 535–541. https://doi.org/10.2307/1310976

Whaley-Martin, K. J., Mailloux, B. J., van Geen, A., Bostick, B. C., Ahmed, K. M., Choudhury, I., & Slater, G. F. (2017). Human and livestock waste as a reduced carbon source contributing to the release of arsenic to shallow Bangladesh groundwater. *Science of the Total Environment*, *595*, 63–71. https://doi.org/10.1016/j.scitotenv.2017.03.234

Whaley-Martin, K. J., Mailloux, B. J., Van Geen, A., Bostick, B. C., Silvern, R. F., Kim, C., … Slater, G. F. (2016). Stimulation of Microbially Mediated Arsenic Release in Bangladesh Aquifers by Young Carbon Indicated by Radiocarbon Analysis of Sedimentary Bacterial Lipids. *Environmental Science and Technology*, *50*(14), 7353–7363. https://doi.org/10.1021/acs.est.6b00868

White, D. C., Davis, W. M., Nickels, J. S., King, J. D., & Bobbie, R. J. (1979). Determination of the sedimentary microbial biomass by extractible lipid phosphate. *Oecologia*, *40*(1), 51–62. https://doi.org/10.1007/BF00388810

White, D., Stair, J., & Ringelberg, D. (1996). Quantitative comparisons ofin situ microbial biodiversity by signature biomarker analysis. *Journal of Industrial Microbiology & Biotechnology*, *17*(3–4), 185–196. https://doi.org/10.1007/BF01574692

Zelles, L., Bai, Q. Y., Beck, T., & Beese, F. (1992). Signature fatty acids in phospholipids and lipopolysaccharides as indicators of microbial biomass and community structure in agricultural soils. *Soil Biology and Biochemistry*, *24*(4), 317–323. https://doi.org/10.1016/0038-0717(92)90191-Y

Zelles, L., Bai, Q. Y., Ma, R. X., Rackwitz, R., Winter, K., & Beese, F. (1994). Microbial biomass, metabolic activity and nutritional status determined from fatty acid patterns and poly-hydroxybutyrate in agriculturally-managed soils. *Soil Biology and Biochemistry*, *26*(4), 439–446. https://doi.org/10.1016/0038-0717(94)90175-9

Zhang, Z., Qu, Y., Li, S., Feng, K., Wang, S., Cai, W., … Deng, Y. (2017). Soil bacterial quantification approaches coupling with relative abundances reflecting the changes of taxa. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-05260-w

3.9 Chapter 3. Supplementary Information. Using molecular PLFA and sterol biomarkers to distinguish viable microbial abundance and community structure, and sewage input potential in shallow and intermediate depth aquifers in Araihazar, Bangladesh

List of Figures

List of Tables

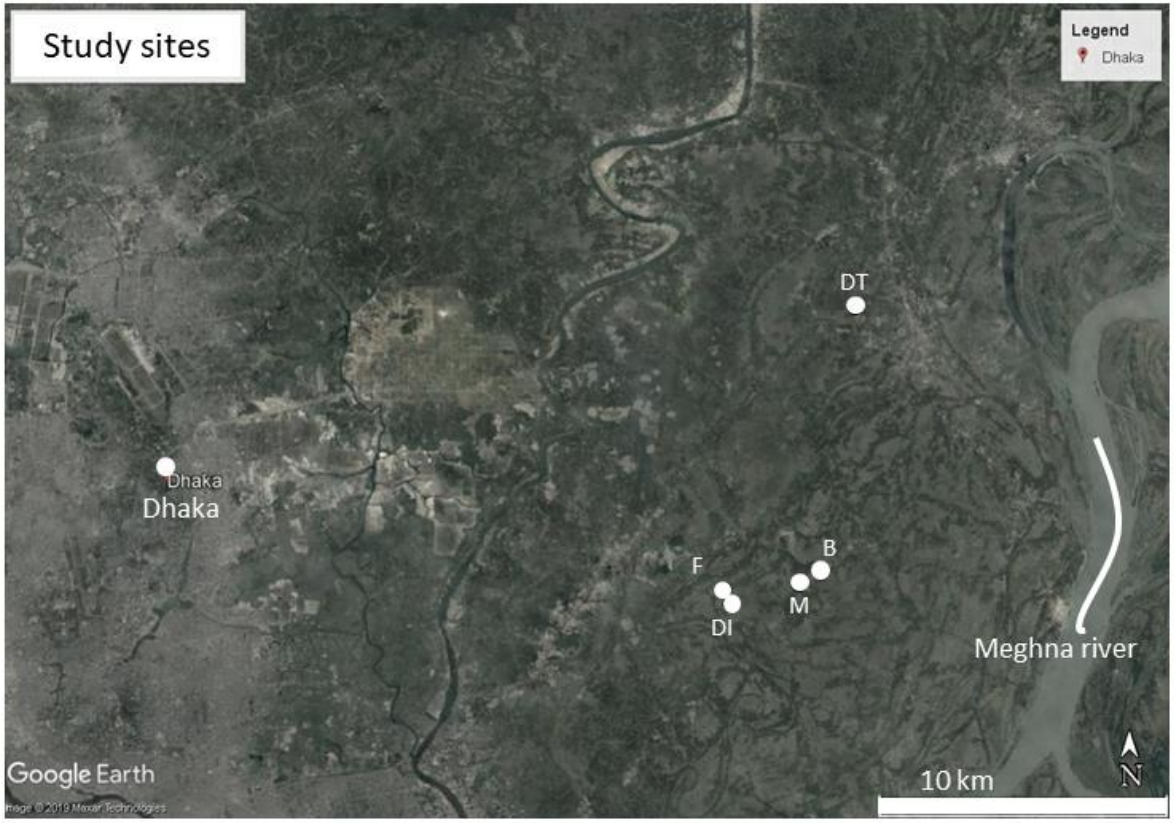


Figure S1. *Study sites examined in this dissertation. Desert Island (DI), Doper Tek (DT), Site M, Site B, Site F all located approximately 25 km east of Dhaka. DI, DT, and Site M studied in this present research project; Site B and F assessed previously by Whaley-Martin et al. 2016, 2017; Mailloux et al. 2013).*

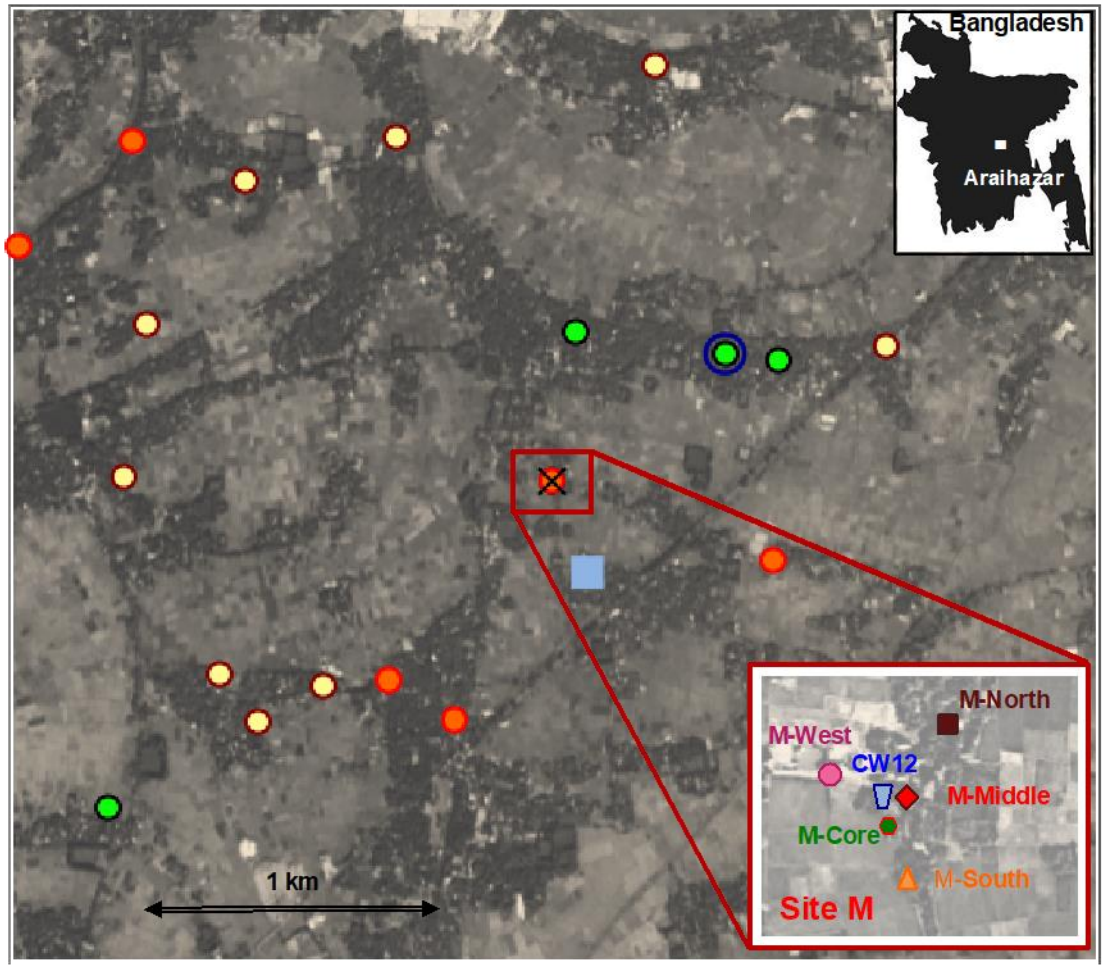


Figure S2. *Site M nest of monitoring wells, failed community well CW12, and location of core drill site. Modifief from Mihajlov et al. (2019, in prep). Well labels indicate their location relative to the failed CW12 well; the sediment matrix drill site is located near the well.*

Table S1. Concentrations of PLFA and corresponding microbial abundance estimated using the cell conversion factor . Dissolved arsenic and DOC data from Mihajlov et al. (in prep).

Depth (ft)	Depth Midpnt (m)	Sedimentary lithology	Avg. total pmol PLFA/g dry sed (± 10%)	Cell abundance[a] cells/g dry sed (± 10%)	Depth of dissolved [As] (m)	Dissolved [As] (ug/L) ± 5%	Depth of [DOC] (m)	[DOC] (mM)
23-24	7.2	Dark grey medium to coarse sand	935	1.9E+07	7.6	53.9	6.1	0.60
51	15.5	Grey coarse sand	5196	1.0E+08	17.5	197.3	17.5	0.15
75	22.9	Grey coarse sand (wood fragments/peat debris)	6404	1.3E+08	20.6	298	20.6	0.20
88	26.8	Grey very coarse sand (wood fragments/peat debris)	1851	3.7E+07	27.4	189.6	24.4	0.72
123-126	37.9	Dark grey very fine clay-rich	746	1.5E+07	32	95	32	1.9
140-141	42.8	Light grey medium sand with clay (plant debris at bottom core)	18429	3.7E+08	41.1	79.9	41.1	0.45

[a]Microbial cell abundance estimated using generic cell conversion factor $2x10^4$ cells/pmol PLFA.

Table S2. Concentrations of six sterol biomarkers along the Site M profile (SD = ±10%).

Depth (ft)	Depth Midpnt (m)	Avg. Coprostanol (ng/g)	Avg. Cholesterol (ng/g)	Avg. Cholestanol (ng/g)	Avg. Campesterol (ng/g)	Avg. Stigmasterol (ng/g)	Avg. β-Sitosterol (ng/g)	Total sterols (ng/g)
23-24	7.2	2.3	12.1	<LOQ	6.1	13.4	34.2	68.1
51	15.5	7.0	42.6	<LOQ	24.1	46.4	150.1	270.3
75	22.9	15.0	29.0	<LOQ	32.8	152.2	380.3	609.3
88	26.8	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ
123-126	37.9	<LOQ	4.8	<LOQ	<LOQ	<LOQ	19.2	23.9
140-141	42.8	95.7	41.6	<LOQ	<LOQ	<LOQ	<LOQ	137.3

Chapter 4: Conclusions and Further Research

Arsenic contamination continues to be a problem in a large number of groundwater wells across Bangladesh as well as many parts of Southeast Asia and has motivated nearly two decades of research to elucidate the widespread mechanisms of release and distribution of this toxic contaminant. The discovery of the role of microbial metabolism in releasing geogenic arsenic into groundwater underlined the importance of widespread biogeochemical carbon cycling. The dissimilatory reductive dissolution with concomitant release driven by available reduced carbon motivated studies which identified potential sources of organic carbon in the aquifer. However, little is known about the active *in situ* biogeochemical cycling of different organic carbon sources by microbial communities. This dissertation expanded the present knowledge of active in-aquifer microbial carbon cycling driving arsenic contamination in Bangladesh by using environmental biomarker-based approaches.

4.1 Research Summary

The investigation of three new sites (Desert Island DI, Doper Tek DT, Site M) in Bangladesh served to expand our understanding of the regional-scale active carbon cycling by microbial communities as it relates to arsenic contamination of aquifers. Source-specific molecular markers (i.e. Phospholipid fatty acid (PLFA) cell membrane components and organic sterols) were quantified to assess the abundance of viable *in situ* microbial biomass as well as potential sources of organic input into depth. Interpretations about sewage input as a carbon source was enabled by characterizing the organic sterol profiles in these aquifers. In particular, sewage derived Coprostanol was compared to other naturally occurring biogenic sterols in the

aquifer by using the Sewage Contamination Index and Coprostanol/Cholesterol ratio. Previous reports sewage contamination to correlate with dissolved arsenic at depth implying that inputs of sewage as a labile carbon substrate drives microbial metabolism and arsenic increase in groundwater. DI and DT, as outlined in Chapter 2, were distinguished by the extent of reduction of the aquifer highlighted by the grey and orange color of the sediments as well as their lithological age: DI is young Holocene while DT is a shallow exposed Pleistocene aquifer. The viable microbial abundance in the grey-colored high arsenic (37-188 µg/L) aquifer DI was significantly greater than at DT, the low arsenic (7 µg/L) relatively oxidizing shallow Pleistocene aquifer. This argues a greater extent of bulk microbial community metabolism in highly reducing arsenic contaminated aquifers and supports that oxidizing conditions preclude dissolution of Fe(oxy)hydroxides and/or arsenic release into groundwater. The very high total organic sterol concentration at DI compared to DT lends support to greater available organic carbon sources in this aquifer. Plant-derived phytosterols dominated the total sterol content at both DI and DT with comparatively little sewage-derived sterols (i.e. Coprostanol and Cholesterol). Furthermore, the Sewage Contamination Index (SCI) – used here to measure the relative importance of sewage compared to other sources of organic matter in the aquifer – were consistently <0.7 at both sites, indicating little sewage input into these aquifers. These observations argue that sewage-derived organic carbon does not contribute to increase microbial abundance at these present sites compared to previous sites (Site B, F).

Site M (Chapter 3) is unique from these sites because it captures a deeper aquifer environment than previously studied in Bangladesh. The motivations to study Site M stemmed from the failure of a community well which experienced a sudden increase in arsenic (60µg/L)

after initially stabilizing at low concentrations (<10µg/L). Therefore, Chapter 3 addressed concerns about the vulnerability of presently pristine deeper Pleistocene aquifers to arsenic contamination due increasing anthropogenic activity such as pumping. The viable microbial biomass at Site M described a localized heterogeneity in the size of viable microbial biomass within the subsurface: greater microbial abundance within sandy aquifer layers compared to clay-rich aquitards. This provides new insight about the distribution of viable microbial communities in deeper low arsenic aquifers previously assumed to be hydrologically separated, and thus protected, from shallow high arsenic aquifers. The maximum microbial cell abundance within the upper reducing Pleistocene aquifer may be driven by a large input of reactive carbon. The sterol profiles at DI, DT (Chapter 2), and Site M (Chapter 3) indicate that this reactive carbon may be derived from plant matter in the shallow Holocene aquifer and sewage within the intermediate depth Pleistocene aquifer. These findings corroborate previous arguments that reactive carbon from thick organic-rich clay layers provide a source of labile DOC which drive arsenic increase at Site M which was proposed to be accelerated by distant municipal-scale groundwater pumping (i.e. from Dhaka).

Radiocarbon isotope analysis ($\Delta^{14}C$) of PLFA biomarkers and bulk SOC, DOC, and DIC elucidated the relative age of carbon sources actively metabolized by the indigenous microbial community in these aquifers. Modern $\Delta^{14}C$ signatures of PLFA from both DI and DT suggested that aquifer microbial community were cycling predominantly young/modern carbon sources consistent with previous sites where a younger component of DOC was argued to be transported from the surface to the aquifer. These observations at DI and DT compared to previous sites indicate that a greater bulk microbial community metabolism may not directly

result in the increase of dissolved arsenic in the aquifer. The low impact of sewage at DI and DT but high sewage impact at previous sites suggests a regionally heterogenous organic matter available in Bangladesh aquifers. Human and/or animal sewage waste from the surface may not be a consistent carbon source which drive arsenic contamination in shallow Holocene aquifers across Bangladesh. It is possible that the differences in density of local human population in rural Bangladesh contributes to greater sewage contamination in shallow aquifers.

4.2 Implications and future work

With regards to further work, in addition to data generated in this dissertation, the radiocarbon profile of bulk carbon pools at DI and DT as well as PLFA biomarkers at Site M can be used in a two-end member model to estimate the relative importance of available carbon substrates at these present sites. This will provide further insight and confirm the microbial cycling of young or modern carbon sources over a regional scale across Bangladesh as well as elucidate whether microbial communities in deeper aquifers cycle modern carbon sources or else metabolize reduced carbon predominantly derived from organic-rich clay layers buried in the aquifer. As elucidated in this dissertation, lack of correlation between total PLFA and dissolved arsenic at depth challenges the notion that the primary release mechanism causing arsenic contamination in groundwater is mediated by the total microbial community in the aquifer. However, environmental microbial communities are highly taxonomically diverse, especially in sedimentary aquifers where available carbon substrates and geochemical conditions are heterogenous such as in Araihazar and across Bangladesh. It is possible that microbial communities in arsenic contaminated aquifers are composed of a greater proportion of bacterial groups which can couple oxidation of organic carbon with reductive dissolution of

iron oxides and concomitant arsenic release compared to low arsenic aquifers. Since PLFA profiling cannot resolve the taxonomic composition of microbial communities, future genetic-based analyses (e.g. DNA, rRNA) can specifically elucidate the taxonomic diversity of aquifer microbial communities.

Through biomarker-based approaches, this research dissertation was able to enhance regional scale understanding of biogeochemical carbon cycling in complex aquifers in Bangladesh. Using source-specific cell membrane components (PLFA) and organic sterols, viable microbial abundance and sewage/plant matter input potential was assessed at new neighbouring aquifer sites in Araihazar upazila. Profiling of microbial abundance and organic sterols confirmed the highly variable microbial biomass and potential bulk carbon sources such as sewage and plant/fungi organic matter at different sites across the region. The lack of a direct relationship between these biomarkers and groundwater arsenic demonstrates that reductive dissolution of sedimentary Fe(oxy)hydroxides resulting in arsenic release is complex and likely impacted by other aquifer parameters. The implications of the experiments in this dissertation extend to the long-term sustainability of shallow and deep aquifers as a source of low-arsenic groundwater across Bangladesh and other impacted sites in Asia.