

Sulfate Reducing Communities in Aquifer Systems can be Reliably
Stimulated by Addition of Complex Nutrients

A Dissertation Presented for the
Doctorate of Philosophy Degree
The University of Tennessee, Knoxville

Matthew Bryan Scholz
August 2012

Copyright © 2012 Matthew Bryan Scholz
All rights reserved.

Dedication

First and foremost, I have to thank my wife, Jeanna Scholz for her support, care and help during this process. I am grateful to her for all the opportunities we have found during our time together.

For my mother, father, grandparents and family: I am, and always will be, thankful and proud to be considered a part of all of your lives. In no small part, it would not have been possible for me to achieve this goal without all of your love and support.

Finally, to my daughter: Believe me, if I can do this, you can do anything. I love you, Stella Mai Carol Scholz.

Acknowledgements

It is with gratitude that I thank my advisor, Dr. Gary Sayler for his support during my career, as well as my committee, Dr. Steve Wilhelm, Dr. Alison Buchan and Dr. Mark Radosevich for all their support, patience and advice.

Thanks to Dr. Vernon McIntosh Jr. who helped me frame several important points in this dissertation.

I also need to thank Hebe Dionisi, who initiated this project, Dr. Larry McKay, who was instrumental in gathering samples in Bangladesh and Dr. Alice Layton who aided in the experimental design and other vital steps to gather information from my samples. Also thanks to Tommy Mead and Dr. Jennifer DeBruyn, who helped me with protocols, design, performing protocols, and generally keeping me sane.

Finally, everyone at the Center for Environmental Biotechnology at the University of Tennessee: Thank you for your help and support.

Abstract

The dissertation presented below is the summation of research into the potential roles of microbial communities associated with aquifers of Bangladesh contaminated with naturally occurring arsenic, including microcosm experiments to assess the role of nutrient supplementation on both the solubility of arsenic and associated bacterial community shifts. Nutrient supplementation microcosm experiments show that any supplementation can stimulate growth of sulfate reducing microbes (SRM), which can also be associated with removal of arsenic and other minerals. The addition of both a complex carbon and sulfate source shows prolonged removal of these elements from the soluble phase. Increased SRM numbers were maintained through 96 days of incubation.

A bioinformatic investigation of the identified subsystems encoded by all sequenced and annotated bacteria capable of carrying out the most conserved steps in sulfate reduction was performed. These analyses indicated that there are a number of SRM capable of directly reducing complex carbon sources, both in syntrophic communities, as well as without additional aid from the environment.

Sulfate reducing microbes are present, detectable and easily stimulated to grow in aquifer sediment. These communities of SRM are able to create conditions capable of removing arsenic from the soluble phase. The rate of growth of SRM and ability to maintain this immobilization supports the theory that SRM detected in the environment are capable of growth on complex nutrients, and require additional nutrients to successfully remediate arsenic for long periods of time.

Table of Contents

Introduction and General Information	1
Water Contamination: Microbial, Chemical and Biogeochemical	1
Arsenic in Bangladesh	2
Arsenic Contamination Led to Arsenic Exposure	3
The Importance of Groundwater to Bangladesh	6
Challenges to Arsenic Remediation.....	8
The Need for <i>in situ</i> Remediation	10
Hypotheses	10
Literature Review.....	15
Arsenic in Aquifers: Potential Sources, Cycling and Mechanisms	16
Potential Role of Iron Reducing Bacteria in Arsenic Solubility	17
Respiratory Arsenic Transformation: A Mechanism for Increased Arsenic Solubility	18
Increased Arsenic Solubility: Oxidation of Reduced Minerals	18
Immobilization or Mineralization of Arsenic by Sulfate Reduction	18
Potential Arsenic Cycling	19
Sulfate Reducing Bacteria: Mechanisms of Action	20
Impact and Significance.....	20
Overview	30
Step 1: $\text{SO}_4^{3-} + \text{ATP} \rightarrow \text{APS} + \text{PP}_i$	32
Step 2: $\text{APS} \rightarrow \text{ADP} + \text{SO}_3^{2-}$	33
Step 3: $\text{SO}_3^{2-} + 6 \text{H}^+ + 6 \text{e}^- \rightarrow \text{S}^{2+} + 3\text{H}_2\text{O}$	33
Environmental Engineering: Community Shifts as a Result of Amendment	34
Sulfide Interaction with Metals.....	38
Arsenic: Toxicity, Microbial Resistance, and Respiratory Pathways.....	40
Arsenic Resistance	41
Molecular Analysis and Detection of Sulfate Reducing Microorganisms Approach and Issues.....	44
Functional Gene Detection vs. Taxonomy.....	46
Sampling: Bias and Extraction.....	47
qPCR as Molecular Characterization.....	50
SRM Detection and Quantification.....	53
Molecular Techniques.....	53
Microcosms: Construction, Analysis and Sampling Microcosm Design, Construction and Analysis.....	55
Aquifer Sampling (Water Filtration and Sediment Collection).....	58
Well Water: Well Preparation.....	58
Well Water Filtration	59
Sediment Sampling	60
Materials and Methods.....	68
Experimental Design.....	68
Methods.....	72
Sample gathering and processing.....	72

Construction and monitoring of initial microcosms	73
Sediment Samples Gathering.....	74
Microcosm Construction.....	74
DNA extraction.....	75
PCR analysis of DNA samples	76
Amplicon Sequencing, Analysis, and Construction of Phylogenetic Trees	78
DNA Collection and T-RFLP Analysis	78
Chemical Analyses.....	79
qPCR Analysis.....	79
Pangenomic Analysis of Sulfate Reducing Microbes.....	81
Pan and Core Genome Calculations	83
Pangenome Analysis.....	84
Analysis of Annotated Genomes Using SEED/RAST Annotations	85
Subsystems Analysis.....	86
Results and Discussion	88
Arsenic Concentrations of Filter Samples	88
<i>dsrAB</i> Detection and Sequencing.....	89
<i>arrA</i> , <i>dsr</i> and 16S rDNA Detection and Quantification	89
16S rDNA Detection and Sequencing	89
Quantification of <i>dsrB</i> and 16S rDNA.....	90
Microcosm Detail.....	91
T-RFLP Analysis	92
T-RFLP Patterns	92
Molecular and Elemental Analysis	94
Gene Copy Number Changes in Samples as a Function of Time.....	95
16S rDNA Quantification and Variation Between Samples.....	96
<i>apr</i> Quantification and Variation Between Samples	97
<i>dsrB</i> Variation Between Samples	97
Chemical Variation Between Samples.....	98
Comparison of Chemical Analysis to Molecular Analysis.....	100
Pangenomic Analysis of SRM.....	102
Threaded Ortholuge Analysis	102
Analysis of SEED Annotated Genomes and Selection of Genomes	103
Subsystems Analysis.....	105
Discussion and Conclusions	118
Filter Analyses	118
Microcosm Analysis	123
Elemental Analysis of Microcosms	123
Variation in Microcosms.....	126
Population Shifts Due to Amendment in Microcosms	127
Directions for Future Research	128
Analysis of Annotated Genomes	129
List of References	132
Appendix.....	140

Vita.....	155
-----------	-----

List of Tables

Table 2-1. Selected calculated reaction energy for SRM.....	21
Table 2-2. Families of bacteria and archaea that contain confirmed sulfate reducing bacteria.	25
Table 2-3. List of primers and T(m) used in the course of these studies.....	39
Table 4-1. Arsenic concentrations of well sites from which filtered water samples were taken.	88
Table 4-2. Mean counts of qPCR detection levels of 16S rDNA, <i>dsrB</i> and <i>Apr</i> genes.....	94
Table 4-3. Correlations between measured values at 19 ft (6.3 m).....	96
Table 4-4. Median and Deviations of functional genes associated with dissimilatory sulfate reduction, and phosphorous and arsenic concentrations in the aqueous phase of all samples, divided by depth and timepoint.	101
Table 4-5. Analysis of SEED annotations of annotated SRM, grouped by taxonomic categories.....	109

Appendix:

Table A-1. Organisms with an annotated copy of <i>dsrA</i> , <i>dsrB</i> or <i>dsrC</i> as predicted by SEED	140
Table A-2. Organisms selected for pangenome subsystem/gene analysis.....	147
Table A-3. Pearson's Rho correlations between measured variables, part 1, continued in Table A-2.....	151
Table A-4. Correlations between measured elements and copy number of genes of interest.....	152

List of Figures

Figure 1-1. Aquifer structure of Bangladesh.	11
Figure 1-2. Arsenic contamination in Bangladesh.....	12
Figure 1-3. Arsenic mobilization methods via microbial activity	13
Figure 1-4. Arsenic fate and transport.....	14
Figure 2-1. Proposed pathway for Sulfate reduction in <i>Desulfovibrio</i> sp.....	62
Figure 2-2. Sulfur can be transferred between biotic/antibiotic as well as anaerobic/anaerobic spheres by a multitude of biological, photrophic and chemical processes.	63
Figure 2-3. Phylogenetic trees showing the maximum-likelihood phylogenetic trees for <i>sat apr</i> (adenosyl 5'-phosphosulfurase), <i>dsr</i> subunit A and B.....	64
Figure 2-4. Mechanism of attack of OH radical towards Guanine, generating DNA damage and potentially leading to genotoxicity	65
Figure 2-5. Figure depicting mode of entries of As(III) and (V) to eukaryotic cells, and their mechanisms of action against various proteins, as well as removal	66
Figure 2-6. Arsenic fate and transport in prokaryotic (top) and eukaryotic cells (bottom).	67
Figure 3-1. Flow chart of steps taken to study SRM in all samples.....	71
Figure 4-1. Averaged absolute quantification of <i>DsrB</i> and 16S rDNA genes from filtered water samples.	110
Figure 4-2. PCA plots of T-RFLP analysis of microcosms constructed.....	111
Figure 4-3. Ratios of <i>dsrB</i> to 16S rDNA copy number of 6.3 m site.	112
Figure 4-4. Box and whisker plots detailing average values for gene copies....	113
Figure 4-5. Bar plots of chemical data from all depths and supplement.....	114
Figure 4-6. Bootstrapped phylogenetic dendrogram of <i>DsrAB</i> genes isolated from filter samples from 2 wells (37 and 68) in Bangladesh.....	115
Figure 4-7. PCA plot generated from matrix of organisms and presence or absence of metabolic subsystems.	116
Figure 4-8. PCA plot similar to Figure 4-7, with overlapping groupings	117
Appendix:	
Figure A-1. 16S phylogenetic tree of Sequences isolated from filters.....	153
Figure A-2. PCA plot of SEED annotated subsystems for SRM grouped by isolation source.....	154

Introduction and General Information

Water Contamination: Microbial, Chemical and Biogeochemical

The role of bacterial transformations in altering the solubility, toxicity and transport of mercury, chromium, uranium, copper and arsenic, has garnered attention in recent years. These minerals have gained relevance due to their use in industrial, defense and commercial applications, coupled with potential releases into the environment. Bacterial metabolism and transformation of metals, organometallics, metalloids and minerals is well recognized, and has been studied from the aspects of bacterial physiology, toxicology and biotechnology. These transformations can range from an increase solubility or toxicity, to having little effect on the geochemistry, or reducing the impact of the mineral on the environment. These effects vary primarily depending on the minerals, geochemistry, and bacteria involved. Bacterially mediated transformations include metabolic electron transfer, cytosolic transformation to reduce cytotoxic or genotoxic effects, or use as an electron donor. The end result of transformations range from precipitation or complexation of minerals, to increased solubility and toxicity, and can vary greatly for any individual element depending on the factors above.

Anthropogenic arsenic contamination is a problem in many areas. However, arsenic has also gained global attention as a result of naturally occurring arsenic in geo-hydrological strata that provide potable water in areas

with limited alternative resources. Recent studies highlighting the pathway of human arsenic exposure through drinking water have highlighted the importance of understanding the arsenic cycle, including role of microorganisms, and highlighted the need for a thorough understanding of the role of microbial transformations on all aspects of arsenic solubility. Central to these questions are a lack of knowledge as to the abundance, distribution and physiological activity of organisms capable of transforming arsenic, sulfate, or other minerals in aquifers.

The solubility and toxicity of arsenic in natural water can be mediated by a number of biogeochemical processes and transformations. These transformations can alter the solubility of arsenic or otherwise reduce the potential for human exposure. Understanding of these processes has led to a number of attempts at human intervention to reduce arsenic concentrations: through manipulations of electron donors or acceptors such as acetate or sulfate (SO_4^{2-}). These attempts are viewed as more economically viable measures for arsenic remediation than traditional filtration based methods.

Arsenic in Bangladesh

The relatively recent discovery of arsenic contaminated well water in Bangladesh is an urgent health and safety matter, affecting millions of Bangladeshi citizens; however, understanding the sources and determining methods to remove inorganic arsenic contaminants from the system is much

more problematic. It is possible to treat water to remove soluble contaminants, however this process can be expensive, particularly in the case of elemental contamination (usually requiring distillation or osmosis based filtration). In the case of arsenic, the current standard for safe levels of arsenic in drinking water is in the range of 10 parts per billion (ppb), necessitating near total removal of arsenic from any water found to be contaminated. Bangladesh's groundwater contamination with arsenic affects millions of households in areas where detection is difficult or impossible, and filtration is prohibitively expensive. Additionally, there are no truly viable alternatives to use of groundwater, due to the extreme risks associated with drinking surface water, which can be contaminated with any number of disease causing agents, including *V.cholera*, *Shigella* species, *E.coli*, *Campylobacter jejuni*, as well as viral diarrheal diseases (Baqui, Yunus et al. 1991).

Arsenic Contamination Led to Arsenic Exposure

The discovery of arsenic contamination was not made as a result of testing or other routine measure, but as a response to an apparent epidemic of outbreak of diseases classified as arsenicosis. The symptoms of chronic, low level arsenic exposure also include so called Blackfoot disease (a blackening and blistering of the skin at the soles of the feet), loss of skin pigmentation (sometimes leading to cancers), as well as kidney ailments, other cancers, or more generalized symptoms. These symptoms were determined to be a result of

consumption of sub-lethal doses of arsenic over long periods of time. It was further determined that hand-dug tube wells created for the purpose of protecting the people of Bangladesh from surface water bearing disease causing pathogens was the source of this arsenic.

The effects of arsenic on mammals and other animals are dependent on multiple factors, including concentration, exposure time and redox state of the arsenic. The mammalian cellular response to inorganic arsenic exposure is typically a methylation or dimethylation of the arsenate or arsenite, followed by pumping and eventual filtration by the kidneys (a typical response to exposure to toxic metals/metalloids). Methylation of arsenic allows it to first be preferentially bound and pumped from individual cells, and then be filtered and removed by the kidneys. Removal of arsenic from the body is slow, however, and can result in long term toxicity. As a result arsenic is also excreted in hair and nails during growth, however this rate is much lower. Both methods (urinary and nail growth) are typically viewed as a biomarker for arsenic exposure rather than a viable method for detoxification (Button, Jenkin et al. 2009; Rivera-Nunez, Meliker et al. 2010; He 2011). Methylation of arsenic can only occur after arsenic has been detected inside a cell, gene regulation has occurred and the series of proteins responsible for arsenic methylation and export have been synthesized.

Methylated arsenic can be recognized and appears to be filtered by the kidneys, and excreted through urine. Filtration appears to occur at a constant rate, independent of concentration in the body. Additionally, work by Thomas et

al. 2001 has indicated that methylated and dimethylated arsenic compounds are more toxic than their inorganic counterparts, indicating that this process is not a pure mode of detoxification.

The exact mechanism(s) of arsenic toxicity are not fully understood. It appears that high levels of cytosolic arsenic can interfere with mRNA expression, as well as causing direct DNA damage via oxidative stress (Thomas, Styblo et al. 2001). This would indicate that arsenic exposure could have negative, long and short term effects regardless of detoxification by the body, as a result of exposure. There are several minerals that appear to increase the rate of detoxification, including selenium, which acts to mitigate its uptake into cells, but does not increase detoxification. Recent studies have shown that inorganic arsenic is also excreted through urine at a rate directly related to the concentration found in blood (Jomova, Jenisova et al. 2011). This rate is also slow, and does not appear to contribute to detoxification.

Arsenic is naturally occurring in anaerobic settings in three states, elemental arsenic, and the oxidized forms, As(III)O_3^{3-} , arsenite, and As(V)O_4^{3-} , arsenate. Arsenate is known to act as a phosphate mimic, additionally, it has been shown to have higher toxicity than its more oxidized form arsenate. The mixing of arsenate and arsenite as well as the elevated metallic content of the water in Bangladesh may also contribute to unique toxicology in the system.

The Importance of Groundwater to Bangladesh

The primary concern in many third world countries is water. In countries such as Bangladesh, this concern is not availability: surface water in Bangladesh is sourced from the Bengal river (nearly the entire country is located in the river delta, an effluvial plain of sandy sediment, which will be discussed below). This water first travels through India before arriving in Bangladesh, and is most likely contaminated with high titers of disease causing bacteria and viruses. The linkage between consumption of surface water in Bangladesh and increased infant mortality, or reduced lifespan in adults is reviewed and discussed in Caldwell et al. (Caldwell, Caldwell et al. 2003). Toxicity is of importance as there are no municipal water treatment or distribution plants outside of the city of Dhaka. Washing and bathing are done in ponds, and drinking water is transported by individual households, meaning that each household is responsible for treating its own water for drinking. In many cases it is not feasible to boil all water before consumption, leading to the aforementioned infant and adult mortality from waterborne pathogens.

In an attempt to alleviate this catastrophic lack of potable water, the British Corps of Engineers took advantage of the sandy nature of the soil in Bangladesh and several low-tech engineering tools to develop a low cost, simple method for installing hand-drilled wells across the plain. This method involves the simple use of a lever system combined with water to allow a pipe to be pounded into wet

sand or mud until the desired depth is reached. This allowed the BCE to install wells across Bangladesh, converting an estimated 95% of the population to drinking well water. It has recently come to light that an estimated 20% of these wells are contaminated with arsenic above the current WHO drinking water standards of 10 ppb. Additionally, the relative availability of materials has led to thousands of household owned wells, along with deeper community wells, each well having a varying depth and potentially varying level of arsenic.

While the specific mechanisms responsible for generating mobilized arsenic are still under investigation, there is consensus that arsenic bearing minerals from the aquifer are the original source of the arsenic. These minerals are traced upstream to the Himalaya mountains, and are abundant in the Holocene aquifer in Bangladesh (Saunders, Pritchett et al. 1997; Ahmed, Bhattacharya et al. 2004; Mitamura, Masuda et al. 2008; Saunders, Lee et al. 2008; Uddin, Shamsudduha et al. 2011)

There are several important factors to be considered when analyzing groundwater, particularly in light of recent work by Dhar et al. (Dhar, Zheng et al. 2008), detailing nutrient flux over time between differing depths of aquifer. First, the water table of Bangladesh is highly variable, reaching the surface and potentially flooding, during the monsoon, and receding by tens of feet during the dry season. This means that shallower wells may dry up during the dry season, arguably the point at which they would be most needed.

Second, in many locations in Bangladesh there is a thin clay layer at a depth of between 20-60 feet, that effectively creates two aquifer systems as illustrated by Mitamura et al. (Figure 1-1) (Mitamura, Masuda et al. 2008). The pore size and flow rate of these aquifers appear to be different, and this difference may be partially responsible for the variation of arsenic levels between the upper and lower aquifers, referred to as the shallow and deep aquifers (Weinman, Goodbred et al. 2008; Michael and Voss 2009). The recharge rate of the shallow aquifer is rapid, and allows a large amount of draw at relatively shallow depths. This is coupled with the fact that the majority of mineral sediment in the aquifer is a product of erosion from the Himalayan mountains, and is therefore mineral rich sediment, containing high levels of metals and metalloids, typically in an oxidized state (referred to as an alluvial aquifer) (Ahmed, Bhattacharya et al. 2004; Uddin, Shamsudduha et al. 2011).

The high recharge rate, coupled with the localization of the majority of wells to the shallow aquifer, as well as work indicating that the majority of contaminate wells drilled into the shallow aquifer, has led to questions as to the role of well use as a driving factor of arsenic contamination.

Challenges to Arsenic Remediation

The average household income for a rural Bangladeshi family is estimated at ~\$30 US annually. It is typically estimated that a hand-dug well in Bangladesh costs ~ \$1 US per foot of depth, meaning that drilling a well would cost the

equivalent of one year of income to reach a depth expected to maintain access to water during the dry months. This means that the majority of wells are drilled to a depth less than 40 feet.

Large scale sampling of wells across Bangladesh has illustrated that there are geographical hotspots associated with high arsenic concentrations (Figure 1-2). However, there is also a strong correlation between well depth and elevated arsenic concentration (van Geen, Ahmed et al. 2011). There is evidence suggesting that the shallow aquifer is contaminated with arsenic, while the deeper wells, and possibly deep aquifer appears less contaminated. Additionally, it is possible that some deeper wells have cracked their casings, allowing water from shallower depths to mix in the well-bore. This would decrease the number of wells drilled to the deep aquifer that are contaminated, and strengthening the inverted association between depth and arsenic contamination. While the source of arsenic in these wells is undetermined, these associations indicate a strong need to generate deeper wells to reach groundwater that is not associated with arsenic contamination.

It has also been suggested that alteration of groundwater chemistry to produce large amounts of stabilized mineral near contaminated wells can act to stably remove contaminants from groundwater, such as arsenic (Saunders, Lee et al. 2005).

The Need for *in situ* Remediation

The relatively low expense of drilling shallow wells, coupled with the high expense and health risks of alternative water sources makes the most effective method for mitigating risks associated with drinking water in Bangladesh *in situ* remediation of arsenic in the shallow aquifers, or at the individual pump sites. This requires both an engineering study, and would benefit greatly from a study of the effects these approaches have on the microbial community.

Hypotheses

This study was undertaken with the goal of determining the roles and effects of arsenic and sulfate reducing microorganisms in subsurface aquifers, particularly with regards to their role in arsenic solubility. Specific hypotheses being tested are:

- 1) Arsenic resistance genes are positively correlated with arsenic concentrations
- 2) Supplementation with a carbon source will increase microbial density in groundwater samples
- 3) Supplementation with sulfate will increase the absolute number of sulfate reducing microorganisms in the sample
- 4) There is an inverse correlation between the abundance of sulfate reducing microorganisms and arsenic concentrations in both natural and stimulated groundwater

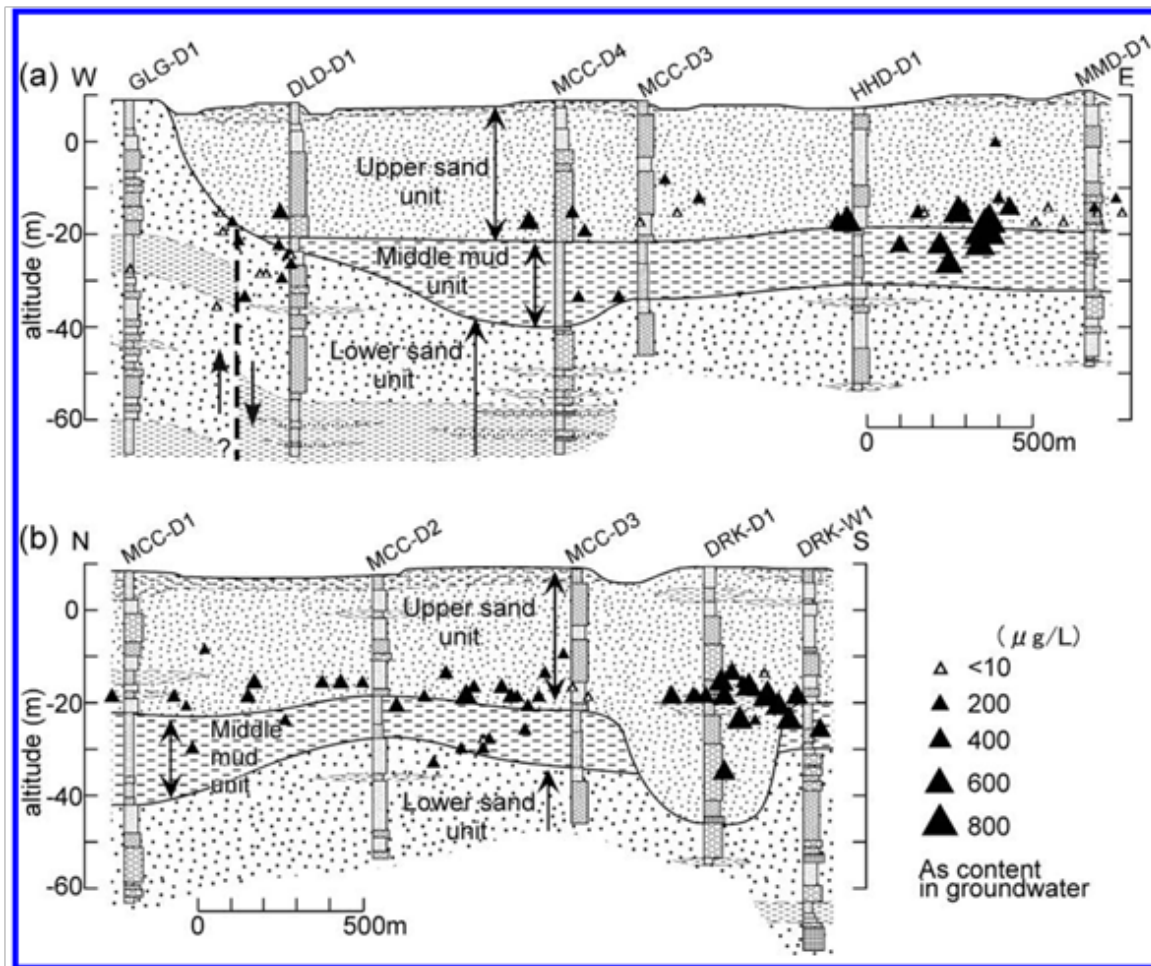


Figure 1-1. Aquifer structure of Bangladesh. Taken from Mitamura *et al.*, 2008. Aquifer layout of Bangladesh shallow and deep aquifers. There are 2 layers of sandy aquifer displayed in both locations, separated by a mud/clay layer, in these cases 20-40 M below the surface, generating the shallow and deep aquifer. This is indicative of geology of the majority of Bangladesh. Additionally, as is illustrated from triangles in chart, the majority of the wells dug are 20 M deep or less. The size of the triangle indicates the concentration of arsenic at that site. Nearly all highly contaminated wells are found in the upper, or shallow, aquifer.

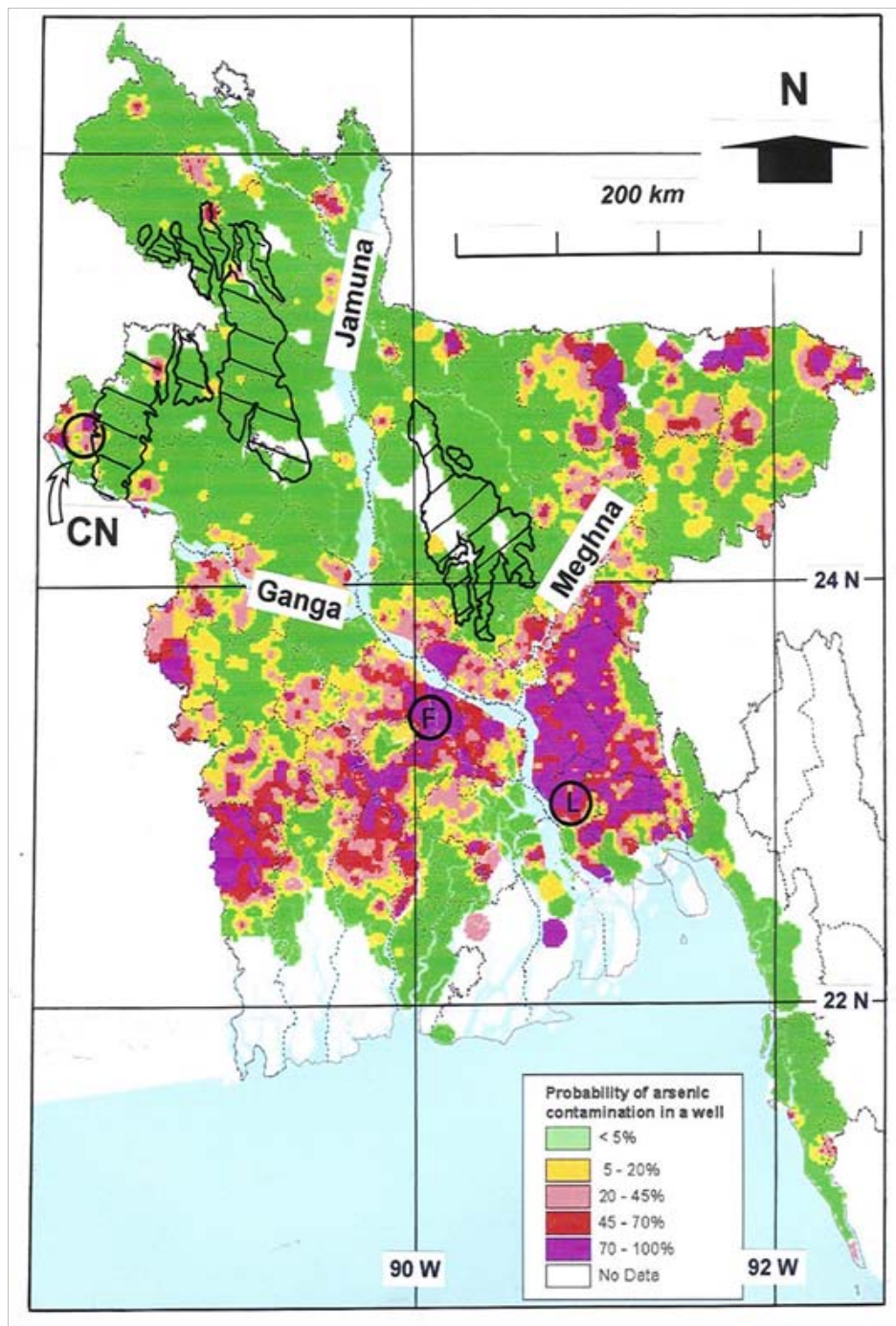


Figure 1-2. Arsenic contamination in Bangladesh. From MacArthur et al. (MacArthur 2011). A heat map of Bangladesh with the % chance a well in any given region will contain arsenic at levels above the WHO limit of 10 ppb Arsenic.

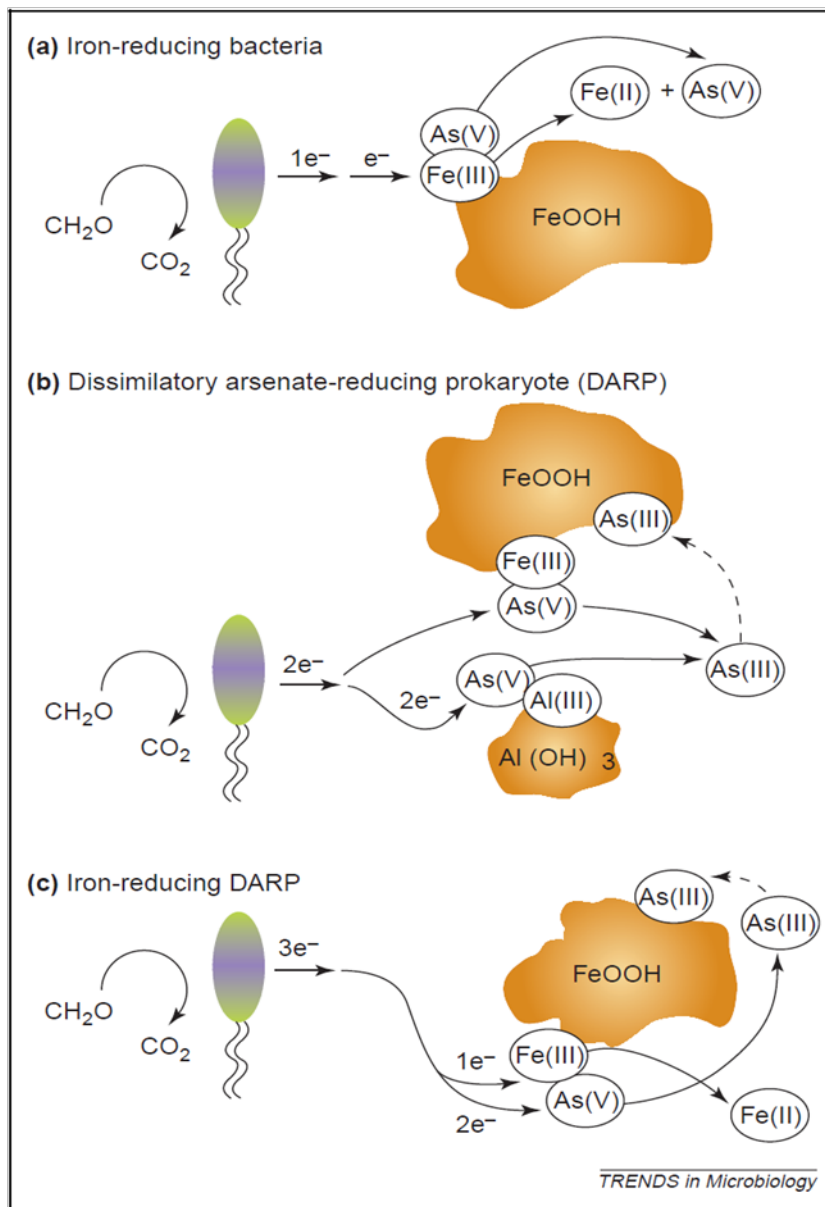


Figure 1-3. Arsenic mobilization methods via microbial activity . From Oremland and Stolz (2005). The chart shows three methods of potential release of arsenic from sediments into the soluble phase from the activity of different microbes, either Fe reducing microbes release arsenic from ferric hydroxide matrices as the iron matrix is dissolved (A), arsenic is reduced by direct arsenate reducing prokaryotes (B), or arsenate and iron are reduced by the same organism (C).

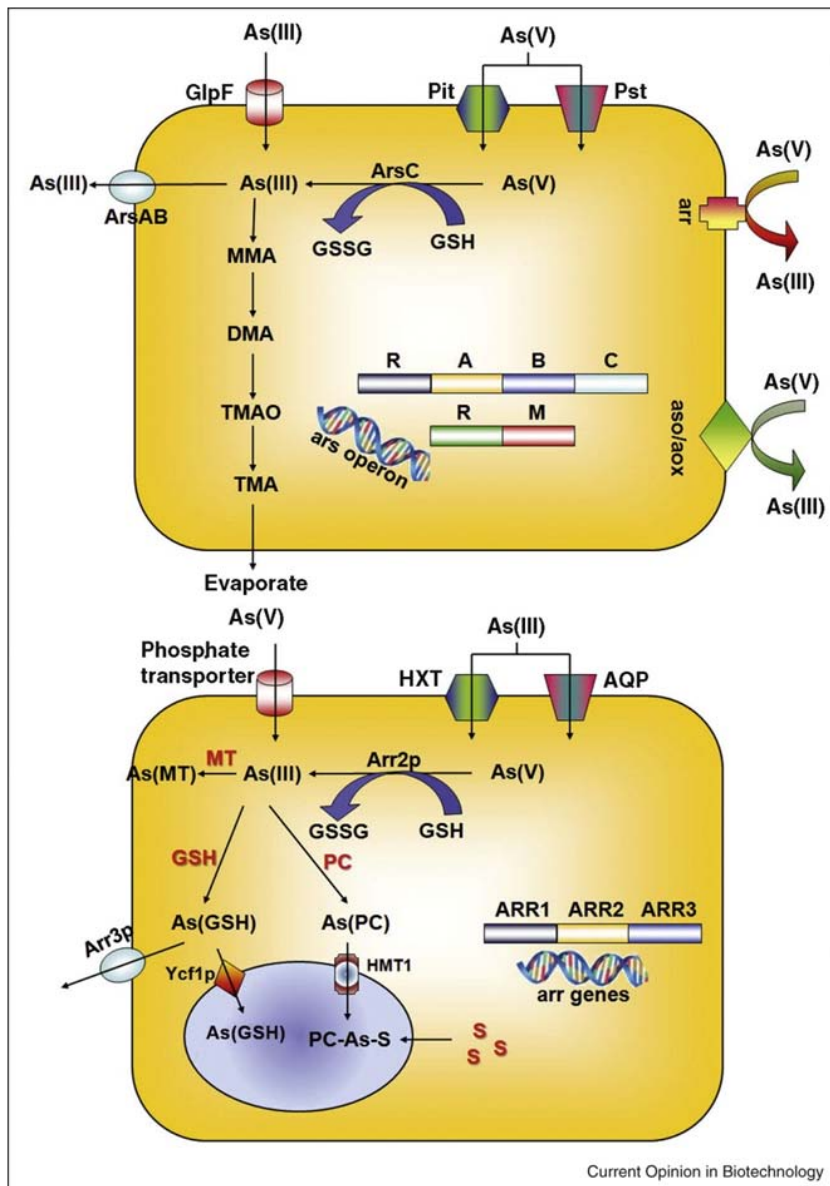


Figure 1-4. Arsenic fate and transport in prokaryotic (top) and eukaryotic cells (bottom). From Tsai *et al.*, 2009. Arsenate can be reduced to arsenite via the *arr* pathway (respiratory), the *ars* pathway (non-respiratory, intracellular), or the (*aox/aso*) pathway (non respiratory, extra cellular). Additionally, some cells can methylate arsenic to form tri-methyl arsenic (TMA) which becomes airborne. Eukaryotic cells can transport GSH (thiol group linked) arsenic out of the cell through the *arr3p* pathway, additionally, bacteria can methylate arsenic to form MMA, which diffuses out of the cell membrane.

Literature Review

Overview

The purpose of this work is to describe the role of sulfate reducing microbes in aquifers both in response to the addition of nutrients expected to selectively encourage their growth. Additionally, this work will describe the effectiveness of artificially altering microbial communities to biologically alter the chemistry of the sediments and associated pore waters to remove arsenic and other inorganic contaminants from the aqueous phase (bioremediation). This chapter will give all appropriate background for framing investigations and approaching this problem. First, a comprehensive understanding of sulfur metabolism, and the challenges and requirements for the growth of sulfate reducing microbes is required. The chemistry of sulfate reduction and its potential role in immobilization of metals and metalloids, such as arsenic are also vital. An understanding of what roles microbes are capable of utilizing alternative electron acceptors is also necessary: both for a firm understanding of the interplay between SRM and other organisms in the environment, as well as to understand the natural biogeochemical controls on SRM growth in anaerobic aquifers.

Also important is an understanding of the effects and value of *in situ* supplementation as a mechanism for encouraging growth of microbes, in specific sulfate reducing microbes is also discussed, coupled with limitations in current

research, and the value of further refining of techniques used in these studies. Additionally, the different types of arsenic resistance in bacteria (reduction of arsenate to arsenite, and selective transport, or extracellular reduction for respiration or resistance) are of importance to understand the scope of the initial contamination issue, as well as the role microbes play in altering arsenic to a more soluble and potentially more dangerous state.

Third, a detailed discussion of the merits and drawbacks of the available experimental techniques and molecular methods to model *in situ* remediation and to determine the effects of these efforts on the microbial community is included. It is both important to understand these factors, and to review all available methods to determine if the most effective experimental designs are used to answer the hypotheses outlined. Discussion of the various techniques required for these studies is discussed. There are also descriptions of the details of the methodology utilized in Chapter 3.

Arsenic in Aquifers: Potential Sources, Cycling and Mechanisms

There are several theories as to the source of arsenic contamination, which may be mutually exclusive, or may all play a role in the process. These can be divided into three main categories: chemical/mineralogical, anaerobically microbially produced, or aerobically microbially produced. There are several chemical or mineralogical effects that may play a role in arsenic mobility. These theories are based on the understanding that arsenic is generally distributed

among iron rich sediment, and would be freed during dissolution of iron rich mineral. The possibility of this iron rich mineral being a highly reduced mineral (such as pyrite), would lead to the existence of aerobic or oxidizing conditions causing this mineral to oxidize and release both iron and arsenic (Islam, Gault et al. 2004; Oremland and Stolz 2005). Alternatively, oxidized mineral (e.g. FeO_3) could also be reduced by highly reducing conditions that may occur in an anaerobic aquifer with highly reductive conditions, although this is less likely. If the initial conditions in the aquifer are oxidized sediment, the more likely scenario is microbial mediated reduction of iron, by pathways suggested by Oremland and Stolz (Figure 2-1) (Oremland and Stolz 2005).

Potential Role of Iron Reducing Bacteria in Arsenic Solubility

It has been illustrated that iron can be used for anaerobic respiration, using ferrous iron as a terminal electron acceptor (TEA). This results in highly soluble ferric iron ions, and also serves to release other elements contained in the iron matrix to the aqueous phase. This mechanism has been proposed to be dominant in at least the lower aquifer, as the levels of aqueous iron are extremely high in this location, coupled with a very low redox potential, indicating highly reductive potential.

Respiratory Arsenic Transformation:**A Mechanism for Increased Arsenic Solubility**

There are a number of genetic mechanisms in place that allow some bacteria to use oxidized arsenic as a terminal electron acceptor. This mechanism could target arsenic rich sediment to reduce arsenate, increasing its solubility and allowing it to dissolve from matrices. This process is not mutually exclusive to iron reduction, and may play a complimentary role.

Increased Arsenic Solubility: Oxidation of Reduced Minerals

In oxidative conditions, there is a newly discovered subclass of delta proteobacteria that are capable of utilizing sulfides as an electron source (Wirsén, Sievert et al. 2002; Ghosh and Dam 2009). Oxidation of sulfidic mineral could also be performed chemically with only mildly oxidizing conditions required in groundwater to oxidize sulfidic mineral to soluble metals and metalloids and sulfate.

Immobilization or Mineralization of Arsenic by Sulfate Reduction

While all forms of arsenic oxides are soluble, they can also be easily incorporated into many metallic minerals. Arsenic can be adsorbed into the matrix of iron precipitates, including iron oxides and iron sulfides, such as pyrite. These two mechanisms have been illustrated to play a major role in arsenic immobilization in natural and human impacted aquifers (Saunders, Pritchett et al.

1997; Pichler, Veizer et al. 1999; Saunders, Lee et al. 2005; Saunders, Lee et al. 2008). Soluble, reduced iron reacts readily with oxygen to form iron oxides. These oxides are much less soluble, and readily adsorb or incorporate other minerals, including arsenic, serving as a potential arsenic sink. Alternatively, sulfides in solution react with all available metals to oxidize both iron and arsenic to form insoluble precipitates. These precipitates can form a highly reactive adsorptive barrier to remove many metals from solutions, including additional iron and arsenic. This mechanism has been shown to occur as the result of biogenic sulfide formation from sulfate reducing microbes, which anaerobically utilizes sulfate rather than metals or oxygen as a terminal electron acceptor, resulting in the formation of sulfide ions or hydrogen sulfide. This reactive ion can rapidly react with any metals in the surrounding water, forming insoluble precipitates.

Potential Arsenic Cycling

It is possible that all of these processes occur, to form a cycle of oxidation/reduction of iron and arsenic transported between oxygenated and anaerobic segments of the aquifer (Figure 2-2). Additionally, a similar cycling of sulfur can occur in the same system, oxidizing and reducing sediments to form a loop of arsenic and sulfides between oxygenated and anoxic zones. While the figure describes a pelagic system, the introduction of potentially non-airtight wells into an anaerobic aquifer may produce similar gradients of oxygen for similar activities to occur.

Sulfate Reducing Bacteria: Mechanisms of Action

Impact and Significance

The process of anaerobic respiratory or dissimilatory sulfate reduction has been well studied. This process utilizes oxidized sulfate (SO_4^{2-}) or bisulfate (HSO_4^-) as a terminal electron acceptor for anaerobic respiratory processes (Figure 2-2) with sulfide as the final product. The scientific study of microbial sulfide production ranges from the study of corrosion and its prevention in wastewater treatment (Little, Lee et al. 2006; Barton and Fauque 2009; Sheoran, Sheoran et al. 2010; Johnston 2011), to the discovery of syntrophic bacteria (Muller, Worm et al. 2010), which require communities of organisms to digest end products of reactions to shift the thermodynamics to a positive energetic state, to its utility in the mineralization of toxic metals (Saunders, Lee et al. 2005).

The negative impacts of sulfate reduction on sewage treatment, and water transportation, have driven the majority of work on sulfate reducing microorganisms (SRM). Thermodynamically, sulfate is less energetically suitable as a terminal electron acceptor than most inorganic anaerobic electron acceptors (Table 2-1), due to the low Gibb's free energy yield per reaction to completely reduce sulfate to sulfide.

There is also strong evidence that sulfate reduction is negatively impacted *in situ* in microbial communities by the addition of nitrate, even with sufficient sulfate in solution (He and Zhou ; Richardson 2000; Gregoire and Soetaert 2010). In these cases, with the addition of nitrate, sulfide generation appeared to

come to a halt. The authors speculated that the higher energy yield of nitrate as compared to sulfate may lead to nitrate reducing bacteria out competing the slower growing sulfate reducers; alternatively, it has been shown that many sulfate reducing microorganisms (SRM) also carry genes implicated in the nitrate

Table 2-1. Selected calculated reaction energy for SRM. Acetate yields the least free energy per molecule, while pyruvate yields the greatest amount.

<i>Electron Source</i>	<i>Reaction</i>	<i>dG</i>
<i>Hydrogen :</i>	$4H_2 + SO_4^{2-} \rightarrow 4H_2O + S_2^-$	-123.9
<i>Acetate :</i>	$CH_3COO^- + SO_4^{2-} \rightarrow H_2O + CO_2 + HCO_3^- + S_2^-$	-12.4
<i>Formate</i>	$4HCOO^- + SO_4^{2-} \rightarrow 4HCO_3^-$	-182.6
<i>Pyruvate:</i>	$4CH_3COCOO^- + SO_4^{2-} \rightarrow 4CH_3COO^- + 4CO_2 + S_2^-$	-331.0
<i>Lactate :</i>	$2CH_3CHOHCOO^- + SO_4^{2-} \rightarrow 2CH_3COO^- + 2CO_2 + 2H_2O + S_2^-$	-178.0
<i>Malate :</i>	$2(OOCCH_2CHOHCOO)_2^- + SO_4^{2-} \rightarrow 2CH_3COO^- + 2CO_2 + 2HCO_3^- + S_2^-$	-180.9
<i>Fumarate :</i>	$2(OOCCHCHCOO)_2^- + SO_4^{2-} + 2H_2O \rightarrow 2CH_3COO^- + 2CO_2 + 2HCO_3^- + S_2^-$	-190.1
<i>Succinate :</i>	$4(OOCCH_2CH_2COO)_2^- + 3SO_4^{2-} \rightarrow 4CH_3COO^- + 4CO_2 + 4HCO_3^- + 3S_2^-$	-150.4

reduction pathway , and at least one SRM, *Desulfovibrio desulfuricans*, is capable of carrying out nitrate reduction as well as sulfate reduction (Marietou, Richardson et al. 2005). This complementary gene set could also explain the lack of sulfide generation under nitrate rich conditions, as dually capable SRM may up-regulate nitrate reducing genes and down-regulate sulfate reduction during periods of nitrate exposure.

Sulfate reducing microbes have a number of advantages that seem to improve their viability, while allowing them to utilize less energetic substrates. Unlike other obligate anaerobes, some SRM have been shown to survive exposure to oxygen, with additional evidence pointing to sulfide generation being used as a method to remove oxygen from the environment. It would appear that both of these processes are the result of the extremely reactive nature of sulfide with oxygen. Recent research has also indicated that some SRM can also perform sulfate reduction under oxygenated conditions (Brioukhanov and Netrusov 2007). This is not true of all SRM, but is a potential reason for maintenance of a gene cassette that is less energetically favorable under most conditions. Alternatively, it has not been definitively shown whether these organisms are capable of using multiple electron transport chain (ETC) pathways for ATP generation simultaneously. For example, it is commonly accepted that there is little to no benefit to up-regulation of both a nitrate reduction and a sulfate reduction pathway if both nitrate and sulfate are present in an environment, due to the low energy yield of sulfate reduction. This has not been tested empirically

in any studies of SRM that also have demonstrated capacity for alternative ETC pathways.

The sulfide ion, hydrogen sulfide, and many sulfidic minerals are rapidly oxidized in the presence of elemental oxygen to sulfate, along with any other counter ions present being oxidized according to their reaction pathways. This means that sulfate reducing microbes are protected to a degree from exposure to oxygen by their mode of respiration. This cycling from sulfide to sulfate would increase the concentration of available substrate for further respiration by SRM. This serves the dual purpose of protecting the microbes from oxidative stress, and presenting them with additional TEAs for continued respiratory activity.

Dissimilatory sulfate reduction also plays a major role in the cycling of sulfur, particularly in aquatic systems. This comes as a result of the transport of sulfide and sulfate between anoxic and aerobic zones. Sulfide can be chemically oxidized to sulfate without biological intervention. Alternatively, sulfide can be used in aerobic conditions as an electron donor by sulfide oxidizing microorganisms (SOM) (Loy, Duller et al. 2009), allowing an interchange between sulfide oxidizing microorganisms which use sulfide as an electron donor for energy generation and SRM in anoxic water (Harada, Yoshida et al. 2009). While there are a large number and wide diversity of identified SOM, it appears that at least one class of SOM has evolved as a direct result of divergent evolution of *dsrAB*. The new gene, labeled reverse dissimilatory sulfite

reductase *rdsr* has high sequence identity to *dsr*, but appears to reverse the reaction, harvesting 6 electrons from sulfide for aerobic respiration.

While SRM are assumed to be obligate anaerobes, SOM are only capable of oxidizing sulfur when oxygen is used as a terminal electron acceptor. There are a diverse number of both SRM and SOM. Families of bacteria and archaea with identified sulfate reducing microbes are listed in Table 2-2. The complete list can be found in table Supplementary Table A-1 and Table A-2. The end result is the exchange of sulfur between anoxic and aerobic zones in aquatic systems, allowing sulfur to cycle (Figure 2-2). Additionally, when metals, particularly soluble, oxidized metals are exposed to sulfide, the metal is reduced to form a covalently bound sulfidic mineral, such as FeS, FeS₂, MgS, etc., which are highly insoluble. This leads to the net removal of both metals and sulfur from the environment. This approach has recently been applied in microcosm experiments as well as field scale applications (Lee, Saunders et al. 2005; Hiibel, Pereyra et al. 2008) to remove metals from solutions. This technique has been shown to be an effective method for mineralization of many metals (Saunders, Lee et al. 2005).

Table 2-2 Families of bacteria and archaea that contain confirmed sulfate reducing bacteria.

Archea	Bacteria		
Metanobacteriales	Bacteroidales	Thermoaerobacterales	Geobacter
Methanococcales	Actinomycetales	Peptococcaceae	Desulfovibrio
Methanosarcinaceae	Dehalococcoidetes	Clostridium	Enterobacteriales
Thermoproteaceae	Thermaceae	Aquificales	Alteromonadales
Methanomicrobales	Halanaerobales	Desulfobacterales	Haemophilus

SRM Phylogeny

Phylogenetically, SRM are widespread across the bacterial kingdom, including both Gram negative and gram positive organisms, as well as several archeal species. Examination of the 16s ribosomal subunit DNA of cultured sulfate reducing microorganisms shows family and genus level distributions that are consistent with a deep rooted differentiation (Figure 2-5). Alternatively, examination of conserved functional genes shows maintenance of the genes responsible for conserved proteins, as well as the potential for horizontal gene transfer (Friedrich 2002; Wagner, Loy et al. 2005). The conservation of genes in the sulfate reduction pathway is important for a better understanding of sulfate reduction, both environmentally, and on a molecular level. The isolation and study of SRM coupled with surveys of environmental or other complex systems in which SRM play a role allows us to better predict and understand the role these organisms play both in sulfur cycling and precipitation or corrosion of metals.

The two genes studied for phylogenetic diversity are the APS reductase (*apr*), and dissimilatory sulfite reductase (*dsr*) gene cassette. The *dsr* cassette

typically consists of A and B subunits responsible for reducing sulfite to sulfide, as well as several complexes for transport and electron transport chain activities (*dsrMJKOP*). These two gene sets have been observed to be reasonably well conserved across SRM. They catalyze two of the irreversible steps for the sulfate reduction process. *dsr* has been studied in greater detail, with primers designed, and re-designed to amplify a 2.1 Kb portion of the *dsrAB* genes (Friedrich 2002; Wagner, Loy et al. 2005). More recently, a less well conserved portion of the *dsrB* gene has been used to amplify a ~150 bp region of the gene (Geets, Borremans et al. 2006), for use in DGGE (Miletto, Bodelier et al. 2007) as well as one example of use for Real-Time PCR (Geets, Borremans et al. 2006). *dsr* is an excellent marker for sulfate reduction, because it is required, moderately well conserved and has been shown to correlate closely to 16s phylogeny. Less studied, but still of value is the APS reductase, both for study and for phylogenetic typing. *Apr* does not correlate as closely to 16s phylogeny, indicating a degree of horizontal gene transfer, or the potential of SRM encoding other mechanisms for converting sulfate to sulfite (Friedrich 2002).

Both genes have been investigated for use in detection of SRM from the environment, including amplification and sequencing of amplicons to generate site specific analyses (Nakagawa, Nakagawa et al. 2004; Smith, Kostka et al. 2004; Boothman, Hockin et al. 2006; Miletto, Loy et al. 2008). From the available evidence, it appears that *dsr* is a better target for this process due to the more conserved nature of the gene, combined with the difficulties in designing

conserved primers for *apr*. PCR primers have also been designed for sub-families of SRM using 16S markers (Muhling, Woolven-Allen et al. 2008), but these are less utilized, and require a degree of *a priori* knowledge about the sample before utilization.

SRM Evolution and Isotope Preferences

It is highly likely that sulfate reduction evolved pre-photosynthesis. The low energy yields and the extreme sensitivity to oxygen both suggest that this mechanism of respiration is millions of years old. There have been several in depth studies investigating phylogenetics of the *dsrAB* gene cassette, which have shown broad distribution, indicating both a deep rooted sulfate reducing capacity, and a broad range of organisms with this capacity (Wagner, Roger et al. 1998; Geets, Borremans et al. 2005). This supports the hypothesis that sulfate reduction evolved billions of years ago. Additional evidence comes from geochemical analysis. As a general rule, it has been observed that enzymes are isotope specific (Detmers, Bruchert et al. 2001; Johnston 2011). In the case of SRM, this leads to the convenient fact that the isotope that SRM preferentially reduce, is the lighter stable isotope ^{32}S , which exists at an approximate ratio of 95/5 in favor of ^{34}S . Because ^{32}S is a stable isotope, determination of the $^{32}\text{S}/^{34}\text{S}$ ratios can be used to determine if sulfur bearing minerals were generated as a result SRM activity. This has allowed dating of the first SRM mediated mineralization of sulfur to 2.3 billion years ago (Wagner, Roger et al. 1998). Even

with the availability of a large number of identified, sequenced SRM, there is little to no work in the field of comparative genomics of the sequenced strains of identified SRM.

Investigation of the genetic backgrounds of identified SRM will greatly increase our understanding of both the evolutionary background for the known differences between *dsrAB*, as well as an understanding of what accessory genes are conserved among SRM, or classes of SRM. This is a major need in the field of study, with the amount of information available.

Isolation and Study of SRM

The growth of SRM is relatively simple, and is the primary method for quantification and detection in any cases via MPN quantification. However, it is sometimes imprecise, and takes a long time, due to the slow growth rate of most SRM (8 hour doubling time for *Desulfovibrio vulgaris*). MPN detection of SRM is fairly accurate, as any growth of an SRM can be detected by the formation of sulfidic mineral, which induces a color change, and a subsequent precipitation of metals from solution.

The currently isolated sulfate reducers have required minimal medium and trace metals for growth. They are typically isolated from enrichment cultures, usually grown in Hungate tubes, over several cycles of sub culturing. SRM are typically also slow growing bacteria, due to the low energy yield for sulfate reduction (See Sulfate Reduction Pathway and Energetics section).

Desulfovibrio vulgaris has an average doubling time of 6-10 hours in defined media, and this is near the average growth time for most organisms of this type. Of currently isolated species, there are 141 sequenced strains that have an annotated *dsr* gene or genes. Table A-1 and Table A-2 list all sequenced bacteria with known or suspected sulfate reduction pathways, based on annotations.

There are few SRM with developed genetic systems, making culture based work less appealing. The majority of work done with SRM is done from mixed cultures, to determine MPN, reduction capacity, or effects of nutrient supplementation on communities containing SRM. This is an area of need for future work in the field.

The advent of shotgun sequencing for closing of genomes has led to an explosion in the number of sequenced microbes, and a large number of SRM have been sequenced in recent years. Using annotated genomes, and pangenome analysis, some of the relationships between these organisms can be elucidated, to show the level of relatedness between all isolates.

It is also important to note that many anaerobic isolates were isolated for reasons other than sulfate reduction, but appear to be capable of dissimilatory sulfate reduction. This would indicate that SRM do not grow exclusively as sulfate reducers, but have a range of metabolic capacities.

Sulfate Reduction Pathway and Energetics

Overview

The complete reduction of sulfate to sulfide is a multi-step process, first the sulfate is reduced to sulfite (2 electrons), followed by formation of adenosine 5'-phosphosulfate (APS), and the subsequent transfer of 8 electrons and the formation of HS^- , alternatively this process can form elemental sulfur with the transfer of 6 electrons, and lower energy yield. The mechanism of ATP generation from these steps is unclear, but expected to follow other respiratory ETC based mechanisms, including pumping of H^+ to form gradients, and ATPase based proton transporters. The detailed stages of this respiratory pathway are explained below.

Thermodynamically, sulfate reduction creates a low energy yield under both standard conditions, and conditions assumed to be present in anaerobic microorganisms. In this situation, standard conditions are less likely, as sulfite should be found in excess and sulfide is pumped away from the reaction sites by transferases, yielding a slightly more favorable thermodynamic situation. The linkage between sulfate and acetate or lactate is the equivalent of 2 ATP formation events. Hydrogen yields slightly less. As discussed earlier, of the other common inorganic nutrients, nitrate yields much higher energy per reaction than any sulfate based reactions. This leads to the suggestion that sulfate reduction is a niche community, rather than a dominant community member. However, additional information indicates that many SRM also harbor genes from the nitrate reduction pathway. If true, then sulfate reduction may be merely an

additional form of maintenance. However, generally accepted theories of genetic drift and gene maintenance would indicate that unless these organisms are regularly exposed to conditions under which sulfate reduction is required for energy or maintenance of the cells, the genes would be lost.

While sulfate reduction is a low energy yield process, the niches occupied by SRM are frequently utilizing low energy yield electron donors. This makes SRM a niche organism with a wide range of potential energetic substrates, leading to it being a diverse, globally distributed organism. The majority of methanotrophs and many hydrogenic microbes are also SRM. In the case of hydrogenic microbes, this may also play a role in mitigating sulfide generation by forming hydrogen sulfide gas (H_2S), which would more rapidly diffuse out of the immediate area of the microbe. While methanotrophs have lower energy yield than organisms that completely reduce carbon to CO_2 , methane is not immediately soluble in solution, and should therefore be more rapidly diffused away from the microbe as well, in both cases shifting the thermodynamics of the energetic reaction towards higher energy yields, or more precisely, preventing the end products of respiration from reducing the thermodynamic yields.

As mentioned previously, SRM are also frequently partners in syntrophic growth. The syntrophic form of commensalism is the consumption of electron donors by a combination of two organisms, each incapable of growth on the substrate without the other. Syntrophy was discovered in sulfate reducing microorganism cocultures that were able to consume methane or other low

energy yield substrates by cooperative degradation of intermediate compounds (Schink 1997). To accomplish this, one organism (typically a SRM) first oxidizes the substrate, and produces H_2 . By itself, this reaction is thermodynamically unfavorable, which would leave these organisms in an energy deficit. The second syntrophic organism acts to alleviate this energy burden by immediately harnessing the generated gas, and metabolizing it, resulting in a net negative yield Gibbs free energy yield for both organisms. These organisms are of interest from both an evolutionary and a biochemical standpoint, as they illustrated a novel role for bacterial cocultures, and could be considered the beginnings of multi-cellular life. Beyond the scientific value in the discovery of syntrophy, this is an indication that SRM can interact closely with other members of a microbial community, and leads to the potential for horizontal gene transfer events, as well as gene loss due to shared metabolic capacities.

Step 1: $SO_4^{3-} + ATP \rightarrow APS + PP_i$

This process can occur through multiple pathways, but ends by creating covalently bound 5'APS from ATP and SO_4^{2-} , yielding APS and PP_i . This is an endergonic reaction, requiring the input of two ATP equivalents. This reaction is typically catalyzed by sulfate adenylyltransferase (*sat*), and is essentially irreversible. To overcome the energy input required, the subsequent reduction of APS to sulfite and subsequent 8 electron transfer to convert sulfite to HS^- must yield at least three ATP.

Step 2: APS $\rightarrow$ ADP + SO₃²⁻

The reduction of APS to sulfite, is catalyzed by an intracellular enzyme, APS reductase (A and B), which induces substrate level phosphorylation from the APS to an ADP, returning one equivalent of ATP to the system. This reaction is also the endpoint for two electrons, reducing sulfate to sulfite. The mechanism of this process has been explored, and it has been suggested that this process is not directly linked to energetic substrate usage (Barton and Hamilton 2007; Barton and Fauque 2009). As an example, when SRM utilize H₂ as an energy source, APS reduction is not affected, indicating a non-direct linkage between the two processes.

Step 3: SO₃²⁻ + 6 H⁺ + 6 e⁻ $\rightarrow$ S²⁺ + 3H₂O

Sulfite and bisulfite is directly reduced by the Dissimilatory (Bi)sulfite reductase pathway. This is a point at which an alternative electron acceptor, bisulfate, can enter the pathway, allowing for a higher energy yield. This pathway is encoded by a series of proteins, which are responsible for transferring electrons from the cytoplasm to the periplasmic space and the subsequent transport of the sulfide out of the cell. Additionally, protons are consumed to produce water, further improving the energy yield (which is typical of ETC). The mechanism of energy generation from this process is not completely understood, but it is assumed that cytochromes on the cell surface aid in the electron

transport chain and H^+ transfer across the cell membrane. Additional investigation into the number of H^+ required per ATP generated, as this number has also been shown to vary between two and five (Barton and Hamilton 2007; Barton and Fauque 2009). The *dsr* complex is a multi-protein complex, with *dsrAB* appearing to be conserved in all species. Several species also encode for accessory proteins (*dsrMJKOP*) which appear to relate to electron transfer, or improving efficiency of sulfate reduction. These species include both bacteria and archaea.

Detail associated with these reactions can be seen in detail in Figure 2-1, including expected cellular location of each reaction.

Environmental Engineering: Community Shifts as a Result of Amendment

The volume and nature of aquifers makes pump and filter or treat methods impractical and potentially impossible. Alternatively, *in situ* remediation is a highly effective method for treatment of organic contaminants, as they can be degraded into less dangerous or noxious forms. In the case of metal or elemental contaminants, there is no easy method for either *in situ* treatment currently available. As the solubility of many metals and metalloids is dependent on their redox state, there have been attempts to utilize metal reducing or oxidizing microbes to immobilize these contaminants (Weber, Achenbach et al. 2006; Kumar, Singh et al. 2007; Shi, Squier et al. 2007). Tests on *in situ* treatment with supplements to encourage the growth of metal reducing

microorganisms have been performed on uranium, zinc with limited success. In these cases, the desired result is a bloom of metal reducing microbes, which then reduce the metal and result in insoluble materials precipitating from the aquifer.

In many of these cases, simple nutrients are used to encourage growth of the desired organisms, such as acetate, lactate or other 3-4 carbon sugars. This is done to limit the general growth of undesired groups of organisms, and prevent them from out competing the desired species (Moon, McGuinness et al. 2010). Recent investigations into these shifts have shown that along with the desired organisms, other, unrelated populations are encouraged as well. However, these additions do serve to reduce the number of fermentative microbes that are encouraged, as these supplementation sources are the products of fermentation.

Coupled with the non-specific nature of supplementation, other potential hazards present themselves with the enrichment of small populations of organisms in the form of potential inhibition. Enrichment of any species can again trigger phage/microbe interplay, as lytic phage interactions with their target bacteria typically follow predator/prey relationships. This relationship would typically be a rapid growth of 'weeds', rapidly growing microbes that utilize the carbon source, followed by a spike of lytic phage, resulting in fluctuations of the rapidly growing organisms, and higher levels of community variation.

To encourage the growth of the organisms of interest, addition of excess electron acceptor can also be performed. Inorganic nutrient supplementation can

result in growth of large numbers of the corresponding reducing species (e.g. sulfate reducing or nitrate reducing microorganisms with addition of sulfate or nitrate respectively).

This reaction is not due to the increased availability of substrate, but the higher concentration of a potential electron acceptor. However, in the process these organisms will produce their corresponding byproducts in higher quantities (ammonia or sulfides) which, if left in solution, negatively impact the energy gain from subsequent respiratory actions. As has been previously noted, nitrate in solution is capable of drastically decreasing or completely halting sulfate reduction, either due to thermodynamic considerations, or as a result of gene regulation, or both.

Current Issues in the Field

Dissimilatory respiratory sulfate reduction is the least energetically favorable respiratory reaction known to be utilized by microbes. The existence of this pathway is of interest for its value in the cycling of inorganic sulfur as well as from an evolutionary standpoint. The existence of respiratory sulfate reduction, as well as syntrophic organisms, supports evolutionary theories, and indicates that niches will be filled. Accurate detection is still questionable, due to the variation of the genes, or other factors (Christophersen, Morrison et al. 2011).

On a molecular detection level, it is difficult to distinguish between organisms containing dissimilatory sulfate reduction genes and those capable of

expressing them. Alternatively, the discovery of sulfide oxidizing enzymes that are similar in sequence to *dsrAB*, may also be detected. It is currently unclear how many environmental organisms are capable of sulfate reduction even with the presence of a sulfate reducing pathway. However, of the 141 sequenced organisms with annotated copies of *dsrB*, less than 75% were isolated due to their sulfate reducing capacity, and do not appear to have evidence of sulfate reducing capacity to date.

From isolated species, such as *D. vulgaris*, there are examples of organisms containing mechanisms for sulfate reduction alongside gene cassettes implicated in nitrate reduction, metal reduction (e.g. *geobacter* species), among other potential options. The conservation of multiple low energetic yield respiration pathways is curious, as all of these processes are multi-enzyme pathways, which would point to a high energetic cost of maintenance. It is possible that these genes are maintained from an evolutionary standpoint due to fluctuations in nutrient levels in anaerobic conditions. Alternatively, the slow replication rate may improve gene maintenance from the necessity to compete with more quickly growing microorganisms (Wagner, Loy et al. 2005).

Organisms have also been found that couple the low energetic yield of sulfate reduction with degradation of low energy organic substrates (such as long chain hydrocarbons) in the so-called syntrophic organisms. These organisms are incapable of complete oxidation of the carbons, but, in the presence of simple heterotrophs, are capable of partial degradation. These organisms may be able

to grow on simple, high energy substrates, but incapable of growth on more complex, low yield products, without the presence of a second “helper” organism. Without the syntrophic interaction, it would be impossible for SRM to survive in these environments.

Sulfide Interaction with Metals

The chemical effect of sulfide interactions with iron and other commonly used metals is well documented. Chemically, there is little difference between sulfide (S^{2-}) and hydrogen sulfide gas (H_2S). In the presence of oxidized or elemental metals and metalloids, H_2S quickly reacts by reducing the metals and forming an ionic sulfidic complex. In anaerobic, or dry environments, this complex will typically form a sulfidic mineral. However in the presence of oxygen and water, sulfides can be quickly decomposed and oxidized, forming sulfate (SO_4^{2-}) and elemental or oxidized metal. This is the principle cause of sulfidic weathering, whereby metal pipes or other structures are eroded by the presence of sulfide.

Table 2-3. List of primers and T(m) used in the course of these studies.

<i>Primer Name</i>	<i>Sequence</i>	<i>T(m)</i>
<i>dsr1F</i>	ACSCACTGGAAGCACG	52.1
<i>dsr1Fa</i>	ACCCAYTGGAACACG	56.9
<i>dsr1Fb</i>	GGCCACTGGAAGCACG	56.9
<i>dsr1Fc</i>	ACCCATTGGAACATCG	49.2
<i>dsr1Fd</i>	ACTCACTGGAAGCACG	51.7
<i>dsr4r</i>	GTGTAGCAGTTACCGCA	52.8
<i>dsr4ra</i>	GTGTAACAGTTTCCACA	47.9
<i>dsr4Rb</i>	GTGTAACAGTTACCGCA	50.4
<i>dsr4Rc</i>	GTGTAGCAGTTKCCGCA	50.4
<i>dsr4Rd</i>	GTGTAGCAGTTACCACA	50.4
<i>dsr4Re</i>	GTGTAACAGTTACCACA	47.9
<i>dsr2060f</i>	CAACATCGTYCAYACCCAGGG	48.0
<i>dsr2060fA</i>	CAACATGGTYCAYACCCAGGG	48.0
<i>aprF</i>	TGGCAGATCATGWTYAAYGG	51.1
<i>aprR</i>	GGGCCGTAACCRTCYTTRAA	51.5
<i>1055f</i>	ATGGCTGTCGTCAGC	50.6
<i>1392r</i>	ACGGGCGGTGTGTAC	53.3
<i>E. coli F</i>	CATGCCGCGTGTATGAAGAA	57.3
<i>E. coli R</i>	CGGGTAACGTCAATGAGCAAA	57.9
<i>BsubR</i>	CTCAGGTCGGCTACGCATCG	63.5
<i>BsubF</i>	TACCGGATGGTTGTTTGAACCGCATGGT	63.9
<i>arsF</i>	TCGCGTAATACGCTGGAGAT	57.3
<i>arsR</i>	ACTTTCTCGCCGTCTTCCTT	57.3
<i>acr3Pf</i>	CTATGTCAGAAAGATCAAAAAAGTG	55.9
<i>acr3Pr</i>	TTGTTCCATATATAATATGGTTTA	62.7

Arsenic: Toxicity, Microbial Resistance, and Respiratory Pathways

Arsenic toxicity

On a microbial level, arsenic toxicity is highly variable depending on the redox or chemical state of the arsenic. Arsenate (AsO_4^{2-}) is the most directly toxic to prokaryotes, as it acts as a phosphate mimic, therefore being easily transported into cells, and interacting with phosphate utilization pathways in the cell. The most rapid of these effects is affecting the electron transport chain (ETC) in respiration, blocking regeneration of ATP. This energy balance issue is rapidly destabilizing to cells, and without proper resistance mechanisms, microbes die upon exposure to arsenic. Alternatively, elemental arsenic and the less oxidized arsenite is less likely to cause immediate harm to microbial life, as these elements do not mimic compounds widely targeted for cellular uptake and metabolism.

In eukaryotes, arsenic has been suggested to have several modes of activity, including as a carcinogen, having negative impacts on cardiovascular structure and function, diabetes, as well as causing skin lesions. It has also been linked to negative impacts on neurological development in children (Jomova, Jenisova et al. 2011). Genotoxicity appears to be a function of free radical activation, and subsequent oxidative stress to DNA. This linkage is currently under investigation, and in several cases has yielded inconclusive or conflicting results. The evidence of genotoxicity supports the hypothesis of hydroxyl radicals causing direct damage to guanine, leading to DNA lesions (Figure 2-4).

It has also been suggested that As(III) is more genotoxic than As(V), and there is evidence to support that it also causes the release of Fe from storage proteins, potentially increasing the rate of hydroxyl radical generation (Jomova, Jenisova et al. 2011). Generally, arsenic has been implicated in causing cancers of the skin, kidney, liver and bladder.

Arsenic Resistance

While As(V) is a phosphate mimic, inorganic As(III) is able to diffuse across cell membranes, potentially through aquaporins (Ventura-Lima, Bogo et al. 2011). While As(IV) is typically reduced in eukaryotic cells as well, As(III) efflux is linked to ATP binding cassette (ABC) efflux pumps, typically those classified as multi-drug resistance (Ventura-Lima, Bogo et al. 2011), which seem to correlate to survival of arsenite exposure. However, the most effective and important mechanism for arsenic removal from the cytoplasm remains methylation. A complete diagram of arsenic entry, effects and removal from the cell can be seen in Figure 2-5.

The combination of microbial arsenic reduction from As(V) to As(III) for the purposes of recognition and removal from the cell, and the increased genotoxicity and deleterious effects of As(III) on eukaryotic survival, leads to the survival of microbes, combined with the exposure of eukaryotes to the more toxic form of the metalloid.

Arsenic in natural environments is typically found in two, soluble forms: Arsenate (AsO_4^{3-}) and arsenite (AsO_3^{2-}). While arsenite is more lethal to animals, arsenate is a phosphate mimic, and readily taken up by microbes. Cytosomal arsenate can be recognized by arsenic resistance genes. In *E.coli*, this is encoded by the arsenic resistance cassette (*ars*) (Carlin, Shi et al. 1995), which contains genes for a cytosolic arsenate reductase (*arsA*). The arsenite transporter, *ArsC*, has been studied as a molecular marker for arsenic resistance in the environment, due to its wide distribution (Xu, Zhou et al. 1998; Rosen 1999; Mukhopadhyay, Rosen et al. 2002).

Previously identified arsenite transporter genes have also been found to be environmentally relevant for the identification of arsenic resistant bacteria (Achour, Bauda et al. 2007). *Acr3p* has been identified in yeast, as well as higher prokaryotes. It is unclear in current literature whether there is any ecological bias towards either of these resistance cassettes. This is an important question, as the current de-facto standard detection technique for arsenic resistance is the *ars* cassette, rather than *Acr3p*. If *Acr3p* is abundant in environmental samples, current detection methods are less effective in detection. This is important, ecologically speaking, as arsenic resistance genes are utilized as markers for elevated arsenic concentrations (Ford, Jay et al. 2005).

The mode of action of these resistance cassettes appears to be the reduction of arsenate to arsenite in a non-energetic electron transfer, and a subsequent selective pumping of arsenite out of the cell. Alternatively, there are

several surface proteins that can reduce arsenate to arsenite, or further, particularly the arsenic respiratory reductase (arr) and the aox cassettes. These cassettes also serve the purpose of preventing arsenic from entering the cell, giving organisms with these genes a selective advantage in conditions with high quantities of arsenate (Tsai, Singh et al. 2009). This strategy is well suited to microbes, which can survive chronic exposure to arsenite, unlike higher life forms. A complete figure of the modes of arsenic reduction and transport in prokaryotes is included in Figure 2-6

This dissertation aims to explore the effects of nutrient addition, including complex carbon sources and inorganic sulfate on microbial communities, and explore the potential interrelationships between sulfur, carbon, arsenic and other nutrients in anaerobic communities.

Molecular Analysis and Detection of Sulfate Reducing Microorganisms

Approach and Issues

The use of molecular methods to characterize a metagenomic sample is a necessity. While 16S rRNA gene characterizations are possible, it has been shown that this method can over or under estimate populations of interest (Christophersen, Morrison et al. 2011). When interrogating metagenomic samples, quantification of particular groups of microbes is most definitive when the functional genes capable of carrying out the biochemical task in question are targeted.

Molecular characterization of microbes has become the gold standard of identification. For bacteria, the small ribosomal subunit is a non-protein coding rRNA serves as the evolutionary marker to determine heritage. This is due to its absolute necessity for bacterial life (without it, there would be no protein synthesis), and its regions of highly conserved and highly variable regions. Highly conserved regions are due to the active sites of the ribosome, while variable regions are generally found at the surface of the folded rRNA. This can lead to variable, but conserved alterations in variable regions that are hereditary. Due to this fact, 16S rDNA markers are considered to be the defining characteristic defining bacterial lineage. There is known issue as to the relationship between classical biochemical and taxonomic characterization techniques and molecular techniques, leading to the new phylogenetic classification based on 16S rDNA sequence (Busse, Denner et al. 1996;

Vandamme, Pot et al. 1996). This is partially based on the potential for horizontal gene transfer, as well as the determination of widespread examples of convergent evolution, in which divergent species evolve the capacity to accomplish the same or similar biochemical reactions via different evolutionary paths.

Biochemical assays are the original method of characterizing microbes, to assay for their biochemical attributes, and use a dichotomous tree to find the correct classification. However, discovery of the ability of many microbes to obtain genes encoding for biochemical pathways by events such as horizontal gene transfer, or prophage encoded genes has thrown doubt onto the validity of biochemical tests to correctly classify microbes. Additionally, biochemical characterization can be a slow process, whereas DNA typing can take as little as 48 hours.

However, even 16S rDNA typing has apparent flaws. For example, *Escherichia coli* strains have been shown to have different metabolic characteristics, while maintaining 16S rDNA homology to each other. With the rapid growth of bacteria (minutes to days), selective pressure and random mutation appear to lead to rapid strain/species differentiation. This has led to questions as to the validity of 16S rDNA, or any conserved, so called housekeeping gene as an identification method, due to the high sequence variation in non-conserved genes, including those encoding for biochemical pathways. Additionally, the inability to culture and sequence 99%+ of the

population makes the classification by 16S rDNA characterization somewhat less functionally applicable. For a more thorough review of the issues, including apparent conservation of genes across wide separations of the tree of life, and gene loss within species, there are several general reviews (Koonin and Wolf 2008; Ward, Cohan et al. 2008), additionally, there are too many comparative genomics papers discussing intra and interspecies variations to appropriately cite.

Functional Gene Detection vs. Taxonomy

In the field of environmental microbiology, this has led to use of two opposing cataloguing methods: functional gene detection, and species cataloguing. The premise of the former is that the most important factor to classifying any soil is the enzymatic processes it is capable of carrying out, a result of the combination of functional genes present in the sample, regardless of source. In principle, this would illustrate the contribution of microbes from a given sample, which could be extrapolated to determine microbial characteristics of entire groups of soil. However, current annotation pipelines rely on homology to previously characterized proteins, which may be inexact, and will be unable to identify sequences encoding for novel proteins. This method is becoming more feasible with the advances in sequencing technology, but is also limited by detection of the most abundant sequences. Alternatively, a greater amount of information can be gathered by use of 16S rDNA detection methods, e.g. 454

Pyrotag sequencing. This methodology identifies organisms by use of conserved regions of the 16S rDNA in the genomes of organisms in the sample. While this does not directly inform questions of function in soil, the continually growing knowledge of potential activities of isolated and classified organisms can be used to extrapolate potential microbial activities (e.g. an abundance of methanotrophic organisms indicate anaerobic conditions, with carbon degradation, etc).

However, the lack of in depth sequencing information, coupled with potential PCR bias, leads to a lack of depth of understanding with this methodology.

While neither method is perfect, they are the best available methods for complete cataloguing of a microbial community. The studies contained herein utilize a 16S rDNA cataloguing attempt, and use of several conserved genes for functional characterization of samples.

Sampling: Bias and Extraction

The study of community microbiology in environmental samples is by no means a perfected science. There are multiple issues to consider in molecular biology in particular. Primary among these is sampling. Environmental samples have a high degree of variation in sampling, especially those with particulates (soil and sediment). This can be attributed to the highly heterogeneous nature of samples, as well as several other factors. There are several methodological suggestions to attempt to minimize this variation, most having to do with increasing the number of technical replicates. At this point, however, the

biological diversity inherent to environmental samples enters into the equation. Even with the most sophisticated sequencing technologies currently available, the biological diversity of a single soil sample cannot be adequately exhausted (Agrawal and Lal 2009). High throughput sequencing has also led to the discovery of the so called “rare biosphere”, a group of low abundance species that may help explain community structure variation and global biodiversity (Sogin, Morrison et al. 2006). More targeted molecular queries can be made of samples, by use of targeted primers and quantitative PCR, or other quantitative techniques. This leads to functional gene capacity quantification in a sample.

Second, sample bias also results from the type of sample, and the type of DNA extraction that is performed. Sample bias is typically a result of matrix effects, whereby the extraction of DNA is hindered by the structure of the sample. This is typically seen in soil extraction, where negatively charged particles in the soil may attract DNA, causing it to partition with the sample fraction, rather than with the released DNA. Additional matrix effects may include proximity of microbes to particles, which would shield them from mechanical lysis, or chemical effects from the matrix affecting either lysis, or extraction of DNA from the matrix. Finally, there are a number of known contaminants that can be extracted with DNA from matrix that affect downstream applications, particularly PCR inhibitors, such as Ca, Mg, or humic acids that enter the extracted phase.

Extraction protocols also play a part in potential biasing of the sample. First, lysis of microbes in the sample must be performed, which can be achieved

chemically, enzymatically, mechanically, or with some combination of the above. Chemical and enzymatic lysis are commonly performed for monoculture, as these are known factors to lyse and break down cell walls. The addition of lysozyme to chemical lysis techniques will frequently occur as a secondary lysis step for gram positive bacteria to improve the efficiency of the reaction. This method is typically eschewed for metagenomic extractions, as both matrix effects decrease the effectiveness of the extraction and the unknown distribution of microbes makes it difficult to determine extraction efficiency in these cases. Mechanical lysis is performed by a bead beating technique, in a chemical solution to both stabilize DNA and to encourage lysis of the bacteria. Mechanical lysis has several advantages, chief among them the much higher likelihood of lysis. However, the chief disadvantage of mechanical lysis is also a result of the vigorous motion of the extraction protocol. The difficulty is that bead beating and vigorous movement in general, will cause shattering of the extracted DNA to smaller and smaller pieces, meaning that the more time and energy devoted to ensuring lysis of all microbes in a sample, the shorter the average length of extracted DNA. This results in potentially less than ideal lysis protocols in favor of maintaining DNA strands long enough to detect and amplify genes, requiring longer stretches of DNA.

Once DNA is freed from the bacteria as a result of lysis, it must be separated from non-nucleotide detritus. There are several steps outlined for this, and multiple commercial kits have been designed specifically for DNA extraction

from environmental samples. Commercial kits have the advantage of being rapid, defined protocols, and are expected to be highly repeatable. However, kit based extractions are almost exclusively reliant on a filtration step in which DNA is chemically forced to precipitate on a silica matrix in a filter (vacuum or via centrifugation), followed by rinsing with ethanol to remove as much contaminant as possible. It has been demonstrated that failure of DNA to bind to the silicate column during filtration may be responsible for up to 40% of the loss of DNA during extraction (Feinstein, Sul et al. 2009). However, the alternative to this step is both time consuming and generates a large amount of waste (chloroform, isoamyl alcohol and phenol). Phenol/chloroform extractions extract the DNA in a hot phenol mix rather than salt matrix, and then mixes with chloroform (or 24:1 chloroform:isoamyl alcohol). DNA partitions to the chloroform phase, while the majority of hydrophobic contaminants (including humics) are retained in phenol. However, it has been suggested that phenol itself may transport with extracted DNA, requiring additional measures to be taken to remove phenol. This extraction method, due has much higher yields than kit based methods. However, the issue of phenol contamination is serious, and may affect downstream applications to a much greater capacity than kit extracted DNA.

qPCR as Molecular Characterization

The general term real time PCR is used to describe the use of PCR based methods to detect relative differences between DNA abundances between

samples. Real-time PCR can be generally described as the use of fluorescent, DNA binding dyes to detect the cycle of PCR that amplified enough DNA for detection (called the cycle of threshold, or $C(t)$). As PCR is an exponential amplification method, this means that a change of 1 $C(t)$ between samples is approximately equal to a change by a factor of 2, two cycles indicates a four-fold difference in the quantity in samples, and on and on. However, this method is limited to comparative analyses between samples, and cannot be used for direct quantification, or for comparisons to products from other primer sets. Enter quantitative PCR, or qPCR. Quantitative PCR is the use of a set of known standards that are amplified by the primers used by the primers, to generate a standard curve of the relationship between $C(t)$ and actual quantity. This allows each sample to be compared to the standard curve and a real, quantitative number of copies of the targeted gene present to be generated. From the perspective of sample to sample comparisons, there is no mathematical difference between methods. However, qPCR allows the comparisons of genes to each other within a single sample, or ratios to be constructed to compare between samples. This becomes useful in environmental samples with high degrees of uncertainty.

To counter the sheer number of organisms, and the complications associated with this, it is possible to design specific experiments to interrogate metagenomes by detection and quantification of specific genes. This is done with design of specific primers that are both specific to the gene of interest, and

that amplify genes of interest. This technique is more rapid, can be performed with a minimal amount of equipment, and will answer specific questions about samples. However, qPCR design, in particular, requires short amplicons (typically < 250 BP), and cannot be guaranteed to amplify all sequences that may be available in the sample.

In the case of metagenome comparison, as with any comparative analysis, it is necessary to find a mechanism to normalize between samples. This is accomplished by use of housekeeping genes in the case of RNA, or by 16S rDNA in monoculture. In metagenomes, there are two possible methods of normalization: normalization to 16S rDNA, which results in relative abundance, or normalization to an external spike of DNA, which normalizes to gram of sediment extracted. Ideally, both can be performed at the same time. As always, there are drawbacks to each method. 16S rDNA normalization assumes that there is an average number of 16S rRNA subunits encoded in all genomes present in the sample. This has been demonstrated not to be the case, with rapidly dividing bacteria containing multiple copies of their ribosomes, which would be explainable by their greater need for protein synthesis and higher levels of nutrient uptake. More fastidious microorganisms maintain one or two copies, reflecting their slower growth and lower need for rapid response of translation. However, it may also be that community shifts between samples may also increase or decrease organisms that have alternate numbers of 16S rDNA copies on the genome, potentially altering the expected ratios.

Normalization to a spiked organism, on the other hand, has the negative impact of altering the 16S rDNA numbers, making normalization to 16S rDNA more complicated. Additionally, there is the question of method of spike. Typically, a known number of organisms or genomes is spiked into the sample. Spiking genomes normalizes to the sample effects on extraction, whereas whole organism spikes normalizes to lysis efficiency as well as sample effects. With complex samples, either of these effects may play a minimal to large role in the extraction efficiencies. It is generally assumed that these two steps are mutually exclusive, and that to receive the benefit of both analyses, both a spiked and unspiked extraction must be performed on each sample. It is also important to note that a typical kit or phenol/chloroform extraction method requires between .25 and 1 gram of sample per extraction, which would then lead to a sample size requirement of at least 3 grams per sample for molecular analysis. However, sample is not always simple to obtain, which may lead to the necessity of generating microcosms, rather than native sampling.

SRM Detection and Quantification

Molecular Techniques

The major molecular techniques for detection of DNA contained by SRM are typically contained to detection of 16s rDNA or conserved functional genes, including *dsrAB* or *Apr* as markers for dissimilatory sulfate reduction, and arsenic resistance genes. *dsrAB* are highly conserved, and believed to be the only

enzyme complex capable of converting sulfite to elemental sulfur, or sulfide. The *Apr* gene is involved in the conversion of sulfate to sulfite. All three of these can be utilized for PCR, qPCR for detection and/or quantification. Each has limitations and benefits, and must be approached appropriately. Recent advances in sequencing have made high volume sequencing possible, however, in unimpacted natural settings, it has also been shown that even with 16s rDNA library construction, high throughput sequencing options, such as 454 sequencing, have greatly improved detection and throughput, to the point that major taxa are now being detected (Keijser, Zaura et al. 2008; Polymenakou, Lampadariou et al. 2009; Ishak, Plowes et al. 2011; Jaenicke, Ander et al. 2011). This only leaves detection of the so called “rare biosphere” which requires new methods to remove highly abundant species from analysis, making complete cataloguing of environmental samples still several years away.

qPCR based detection of SRM has been typically performed using site-specific primers, first sequencing several isolates from an environmental sample, and constructing 16s rDNA primers or *dsr* specific primers to those samples, and querying the environment. This method has been applied in oil fields as well as soil samples to catalogue the capacity of the sample to perform sulfate reduction. While useful, it requires several stages of labor, first isolating large inserts of the traditionally amplified *dsrAB* cassette, sequencing, and primer design from the conserved regions found in the newly sequenced amplicons. This method suffers from primer bias of the original *dsrAB* primer set discussed below, but

should, in theory, overcome the biases of primers designed from currently available sequences.

Location specific primer design is a valuable tool, but cannot be applied in a high throughput environment. Additionally, it is difficult, if not impossible, to compare results between sites under conditions such as these. Finally, there is the chance that the *dsrAB* primer set may be nonspecific enough to miss some species, due to specific amplification biases, or . Alternatively, several generic real-time primers have been designed for *dsr* (both A and B) (Geets, Borrernans et al. 2006; Ben-Dov, Brenner et al. 2007; Kondo, Shigematsu et al. 2008), as well as general PCR primers for an ~2.1 kb region of *dsrAB* (Table 2-3). These primers have the advantage of wide applicability, with a single primer set. It is important to note that they may not pick up all strains of SRM containing a copy of *dsrA* or *B*. However, the wide range of applicability and their use in both qPCR (Agrawal and Lal 2009) and DGGE allow their use with a degree of confidence that the results are accurate (Miletto, Bodelier et al. 2007; Miletto, Loy et al. 2008).

Microcosms: Construction, Analysis and Sampling

Microcosm Design, Construction and Analysis

While chemical and mineral analyses of microcosms typically yield similar results between replicates and mimic the behavior of natural samples (Illman and Alvarez 2009), microcosms appear to generate a great deal more microbial

community structure variability (Bombach, Richnow et al. 2010). This may be due to the snapshot nature of molecular microbial sampling, which would result from a highly dynamic host-prey or competition based community. This would be the result of many predator/prey interrelationships. As every organism has a different replication time, and different predators, each organism would have a different oscillation pattern (so-called “red queen dynamics,”) which could result in drastically different communities in a relatively short period of time in separated microcosms.

The concept of red queen dynamics is similar to that of evolutionary theory, namely that there is constant pressure from outside sources. The principle of “having to run just to stay where you are,” in this case, encompasses the idea that organisms that continue to reproduce without (presumably slight) phenotypic changes will quickly be overwhelmed by the myriad of rapidly evolving organisms constantly competing for the same evolutionary niche, or infected and lysed by the rapidly evolving phage population. This has been demonstrated computationally for multi-organism systems (Khibnik and Kondrashov 1997; Dercole, Ferriere et al. 2010). Additionally, the high diversity and heterogeneity of sediment and soil communities, coupled with the continual evolution of phage, make it an ideal environment for red queen dynamics to rear its head.

Alternatively, minor population differences between replicates may result from simple out-competition by a species better adapted to the slightly altered

conditions of the microcosm. This is a simple ecological result of outcompetition by a species given a slight advantage. Alternatively, abundance of an organism in a single sample may play a role by giving organisms the advantage of numbers as a result of separating communities from each other. There is no way to adequately address these topics with our current tools and understanding of environmental microbiology. This leads to the common assumption that whenever possible, microcosms are to be eschewed in favor of natural sampling.

There are cases, however, where it is impossible to perform *in situ* experimentation in environmental microbiology. As an example, testing the effects of antibiotics or toxic chemicals on soil or aquifers is not possible due to EPA regulations and common sense. In these cases there are two potential mechanisms to mitigate the effects of microcosm variation, for the sake of increased statistical power or decreased variation. First, the use of a single, large sample (mesocosm or macrocosm) rather than many small samples (microcosms) can be used. These large experiments can range from grams to kilograms. To allow for sampling under these conditions, larger samples such as these are constructed in containers that allow sterile sampling of both water and removal of sample for microbiological characterization.

As an example, Williams et al. utilized an upflow supplementation mesocosm to encourage the growth of SRM in an *in situ* experiment to determine the mineralization of heavy metals (Williams, Ntarlagiannis et al. 2005). This allows for samples to be taken from a single source, thereby minimizing noise

and uncontrolled oscillation associated with smaller microcosms. However, it is currently unclear that this method is indicative of actual microbial community behavior.

Aquifer Sampling (Water Filtration and Sediment Collection)

There are two methodologies that must be discussed here, sampling well water and the biases and benefits thereof, and collection of aquifer sediment for the purposes of microcosm construction.

Well Water: Well Preparation

There are several effects to be taken into account when sampling from wells including sterility (bore clearing), extraction efficiency, water quality and the effects of oxidation on potentially strict anaerobes during filtration. Sterility of the well is not a question of sterile groundwater, but rather if the well has no leaks in the bore, and if the mechanism used to remove water from the well is sterile so as not to introduce foreign bacteria into the sample. As with any contamination issue, contamination during sampling may result in skewing of either ratios or detection of the target organisms. In the case of cracked well bores (a more common occurrence in Bangladesh, where the wells are drilled with PVC pipes and hammered into the ground), this contamination would alter the perceived community structure by mixing communities from the target depth with communities at the depth of a crack. The common rule of thumb with water sampling for the purpose of filtration or chemical analysis from a well is to clear

3-5X the volume of the bore before sampling. This removes some degree of bias from any potential leaks, but also generates enough mixing to hopefully result in sampling of water directly from the aquifer, rather than the sheltered region within the bore. This clearing also removes most effects of oxygenation as a result of the well itself. It is possible that this effect is not entirely mitigated, as pumping also results in mixing of the water, but it is most likely beneficial.

Well Water Filtration

Once the well bore has been cleared, it is necessary to isolate bacteria from the water. This is typically done by filtration. There are two main methods of filtration currently in use in the field: direct filtration and tangential flow filtration. Direct filtration has the advantage of being low technology, and has low exposure times. In a direct filtration filter, the water is immediately passed through a .2 micron or smaller filter to catch bacteria and large particles directly. While this is a rapid filtration, dependent on the pressure of the pump and the strength of the filtration membrane, it is important to note the relatively lower biomass in groundwater as compared to surface water. Where a typical lake sample may have much higher concentrations of organisms per mL, the amount in groundwater, due to lower nutrients combined with typically anaerobic conditions, is typically lower. This means that filtration of liters of groundwater is necessary where mL of surface water would typically suffice.

The mechanism of filtration is also important. While direct filtration is rapid, even the most effective filters (e.g. Sterivex filters) have a limited ability to

filter organisms before they clog. The alternative method frequently utilized is tangential flow filtration. This methodology has the advantage of huge increases in bacterial concentrations. However, the process of tangential flow is time consuming, requiring multiple rounds of concentration to achieve higher concentrations of bacteria in the filtered fraction. This delay, coupled with changes from anaerobic to aerobic conditions, can potentially lyse bacteria before filtration, altering the perceived community structure, due to lysis, or potential replication on the filter. The changing states of aerobicity, light, and temperature may all combine to alter any sampling that can be done by these methods. However, they are still the best direct measure of planktonic cells in a water sample.

Sediment Sampling

The process of gathering sample from aquifer sediment involves drilling, sampling, and removal of the sample. This can be done in a number of ways, including continuous, semi-continuous, and mixed extraction. In a general sense, all that is required is to dig a hole, and extract sediment with as little contamination of materials, biological contaminants, or conditions different from the sample itself. In the case of anaerobic sediments, this is made more complicated by the need to prevent as much exposure to oxygen as possible during and after the extraction as possible. A high technology solution is to use continuous, large bore equipment to extract a large amount of sediment and keep it both free of outside contaminants and in the same conditions as in its

native state by removing it and immediately capping it. This can be done in a semi-continuous manner, by extracting sediment incrementally, and capping each small segment individually. In this way, discrete measurements can be taken at each depth, and these samples can be easily transported and handled.

In the more exact case of Bangladesh, due to lack of drilling crews, the previously mentioned hand-drilling equipment is used to drill to the desired depth, and semi-continuous cores are taken at 1-foot intervals. There is a higher degree of risk of contamination, due to the use of wet sand and water to lubricate the drilling process. However, the tight packing and positive flow of the aquifer helps to minimize these concerns.

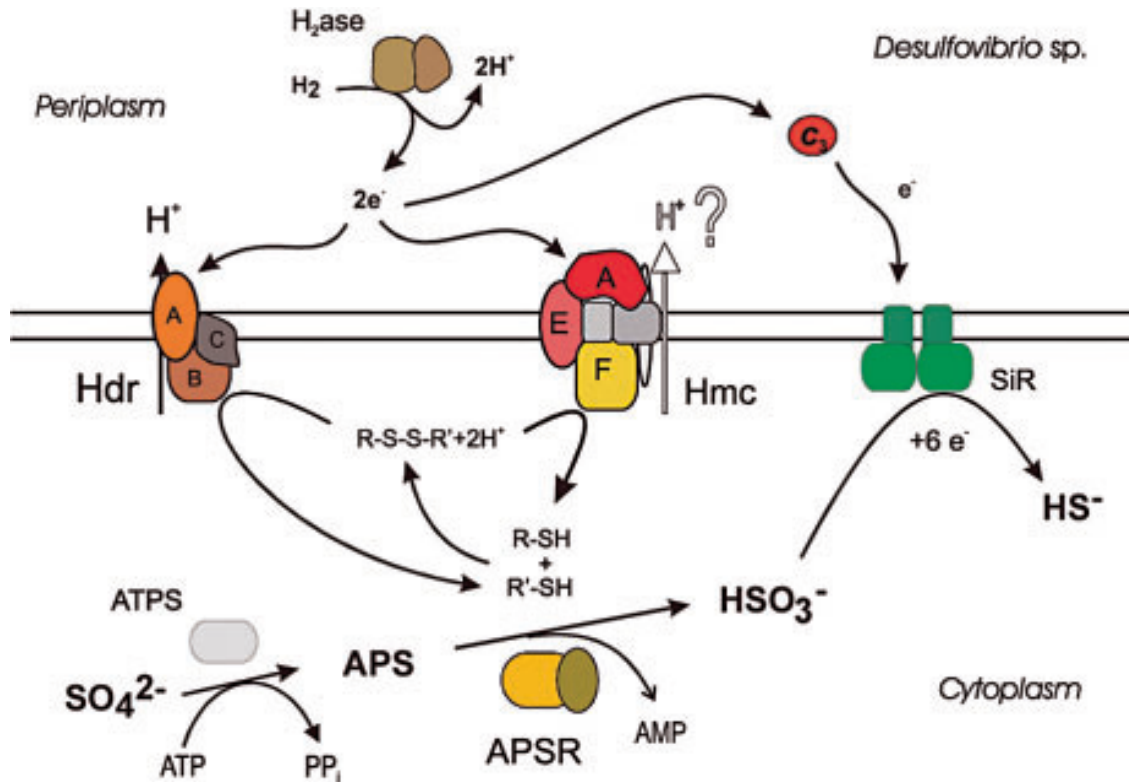


Figure 2-1. Proposed pathway for Sulfate reduction in *Desulfovibrio* sp From (Barton and Fauque 2009) Briefly sulfate is reduced to sulfite by a 2 step process generating 5' APS, followed by disproportionation to sulfite. Sulfite reductases (*dsr*) then reduce sulfite to sulfide. This generates a proton gradient in the periplasm, allowing generation of ATP. Alternative methods of sulfur reduction result in assimilatory sulfate reduction.

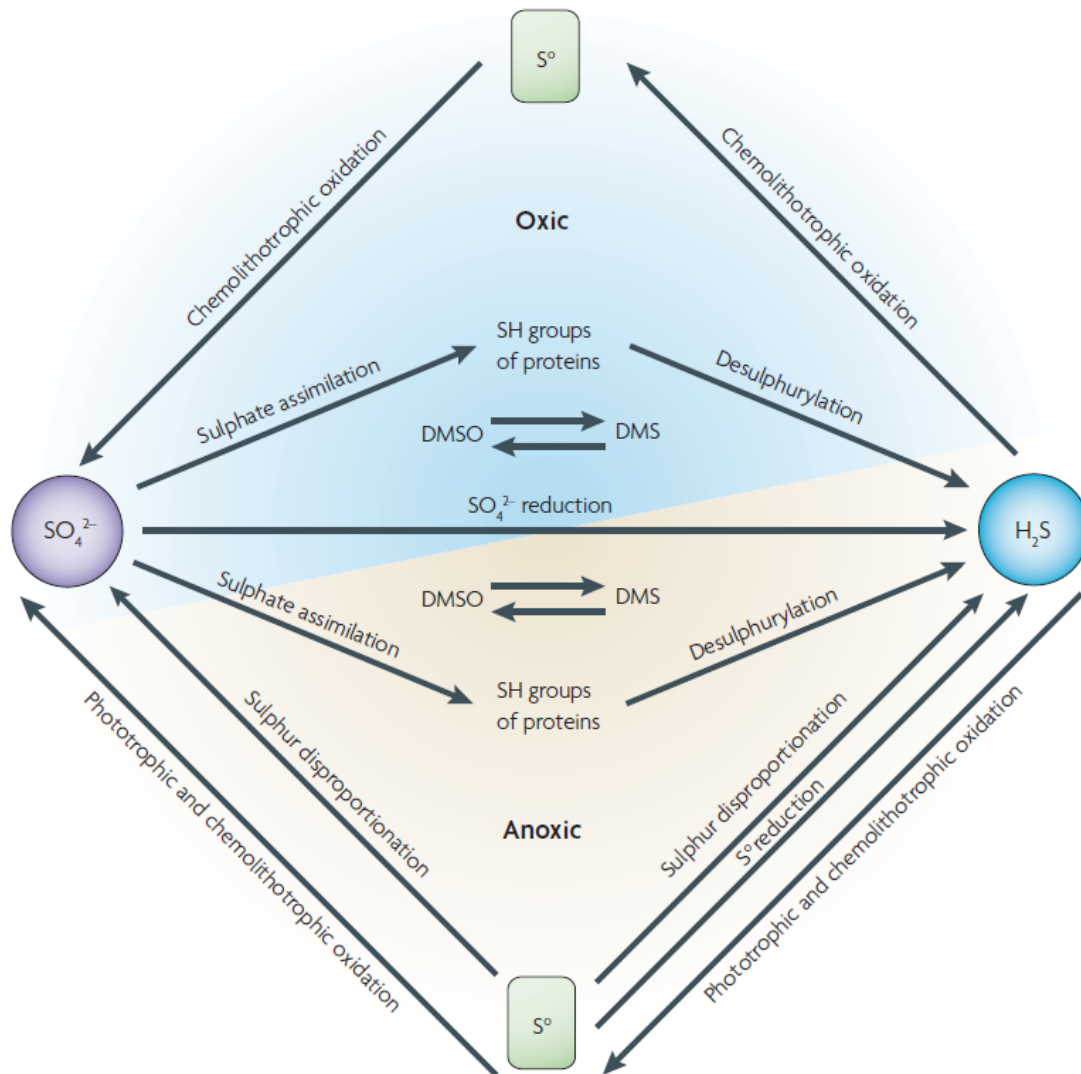


Figure 2-2. sulfur can be transferred between biotic/antibiotic as well as anaerobic/anaerobic spheres by a multitude of biological, photrophic and chemical processes. From Muyzer and Stams. This includes chemical and biological processes capable of cycling sulfate to elemental sulfur or sulfides.

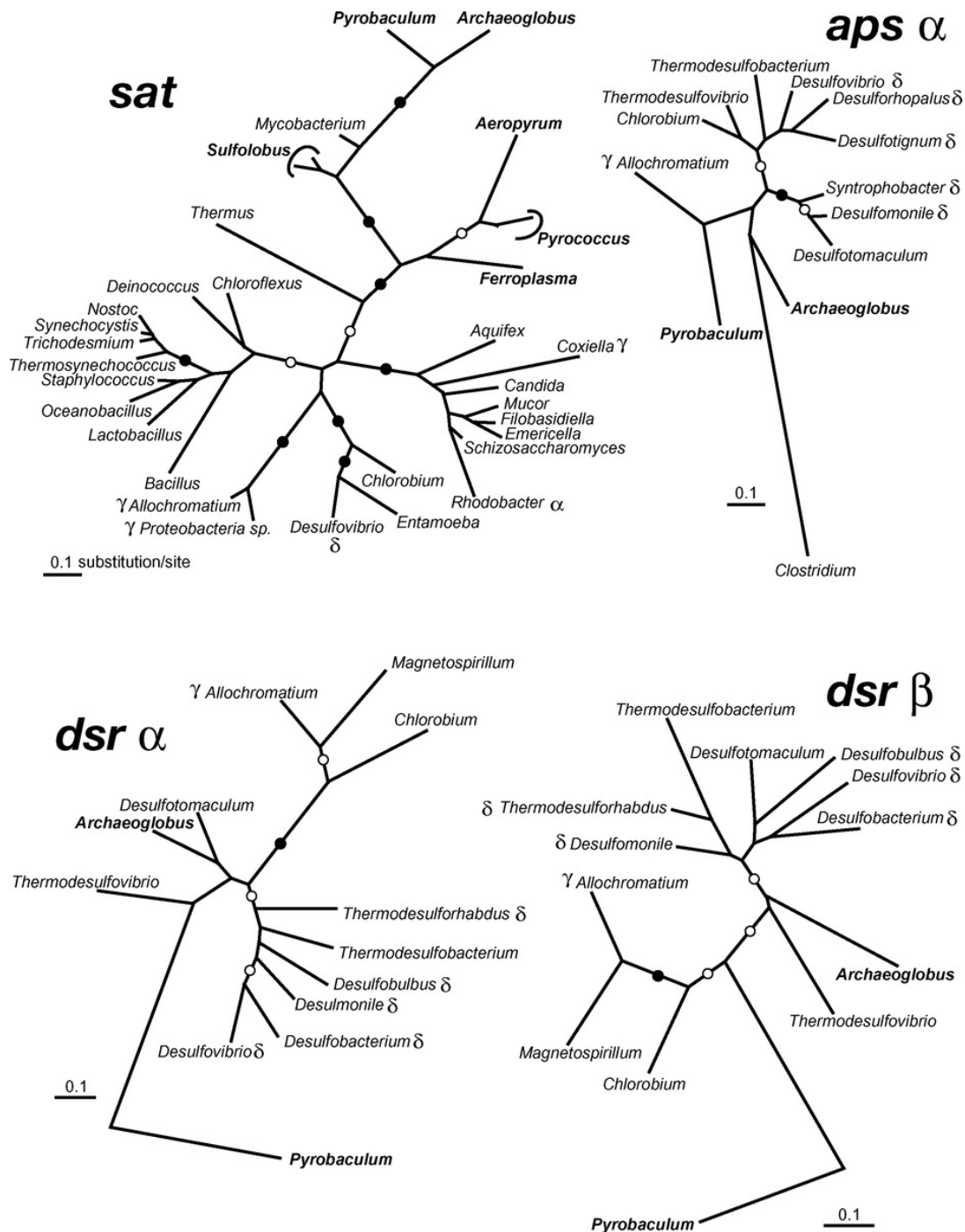


Figure 2-3 Phylogenetic trees showing the maximum-likelihood phylogenetic trees for *sat*, *apr* (adenosyl 5'-phosphosulfurase), *dsr* subunit A and B. From Boucher et al. (Boucher, Douady et al. 2003). This series of trees illustrate that the evolutionary tree is highly similar for each gene product, indicating that genes were acquired together. Black dots indicate 95% certainty of the maximum-likelihood, white dots indicate 80% certainty.

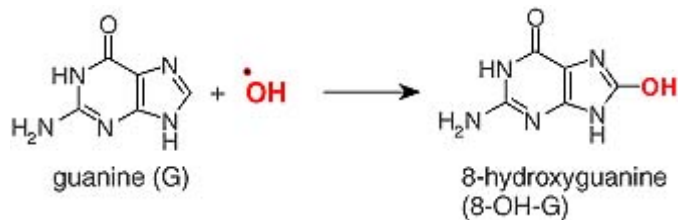


Figure 2-4. Mechanism of attack of OH radical towards Guanine, generating DNA damage and potentially leading to genotoxicity. From (Jomova, Jenisova et al. 2011). This is the likely result of arsenic entering eukaryotic cells and generating free radicals.

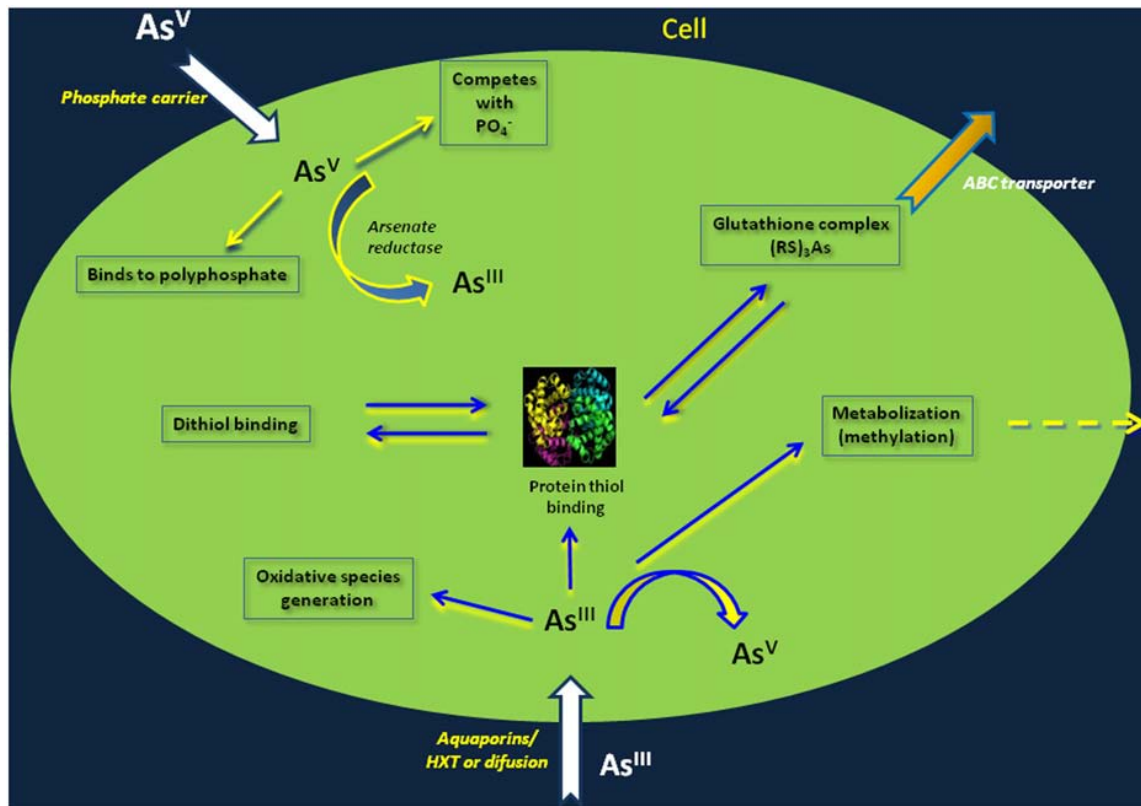


Figure 2-5 Figure depicting mode of entries of As(III) and (V) to eukaryotic cells, and their mechanisms of action against various proteins, as well as removal. From (Ventura-Lima, Bogo et al. 2011). Briefly, As(V) is reduced to As(III) in the cell, but can both compete with inorganic phosphate in energy reactions, or bind to PPI to inactivate it. The primary mode of toxicity of As(III) appears to be generation of oxidative stress and free radicals. Thiol binding and methylation are the primary points of export from the cell, mono-methyl arsenite can diffuse out of the cell, while thiol bound arsenic is pumped from the cell via ABC transporters.

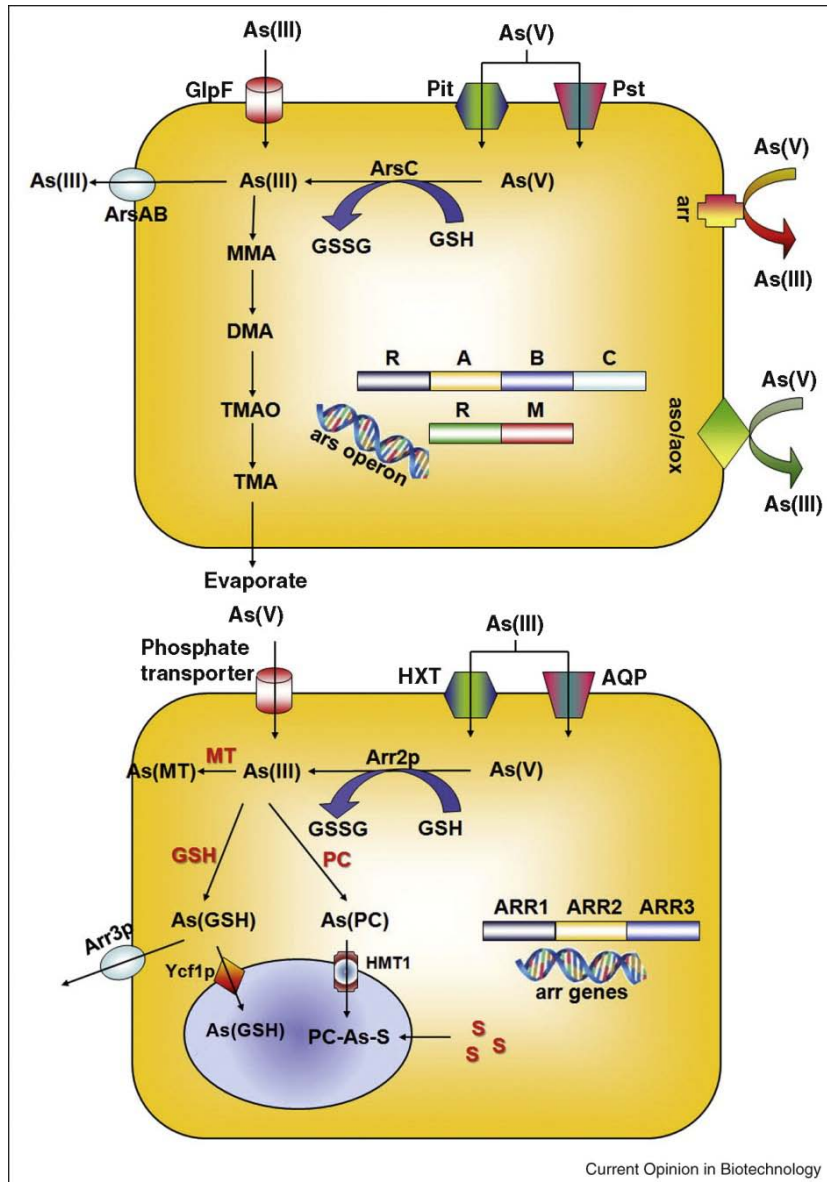


Figure 2-6. Arsenic fate and transport in prokaryotic (top) and eukaryotic cells (bottom). From Tsai et al. Arsenate can be reduced to arsenite via the *arr* pathway (respiratory), the *ars* pathway (non-respiratory, intracellular), or the (*aox/aso*) pathway (non respiratory, extra cellular). Additionally, some cells can methylate arsenic to form tri-methyl arsenic (TMA) which becomes airborne. Eukaryotic cells can transport GSH (thiol group linked) arsenic out of the cell through the Arr3p pathway, additionally, bacteria can methylate arsenic to form MMA, which diffuses out of the cell membrane.

Materials and Methods

Experimental Design

The general purpose of these experiments were to:

Aim 1) investigate the effects of arsenic on microbial communities in groundwater, including presence and phylogeny of arsenic resistance and respiratory arsenic reductases.

Aim 2) investigate the microbial and chemical effects of sulfate and organic nutrient supplementations.

Aim 3) to determine what, if any, linkages exist between naturally occurring sulfate reducing microorganism (SRM) communities and soluble arsenic levels.

Each individual aim is discussed in more detail below.

Aim 1) Investigate the effects of arsenic on microbial communities in groundwater, including presence and phylogeny of arsenic resistance and respiratory arsenic reductases.

To accomplish this goal, groundwater from several wells in the Arihaizar region of Bangladesh was collected and filtered. DNA from these filters was extracted and subjected to DNA analysis to look for presence/absence of arsenic reductases. Additional work was performed to examine the 16S rDNA distribution of individual sites with both high and low arsenic concentrations. The

working hypotheses for this aim were: a) low arsenic levels would be correlated to elevated presence of sulfate reducing microbes, b) elevated arsenic levels would positively correlate to presence of arsenic reductases and c) there would be quantifiable shifts in community structure as a function of arsenic concentrations.

Microbial community surveys have been performed with 16S rRNA surveys, along with quantification of functional genes shown to be well conserved with functions associated with both respiratory and non-respiratory arsenic reduction pathways. Quantification of 16S rRNA copy number was also performed as a proxy for microbial biomass. Relative and absolute quantifications are compared for multiple samples.

The general nature of arsenic resistance genes can also be surveyed from the pool of sequenced and annotated genomes of isolated bacteria, to indicate the background level of arsenic resistant microbes.

Aim 2) Investigate the microbial and chemical effects of sulfate and organic nutrient supplementations.

It has been illustrated that supplementation with sulfate and acetate can have beneficial effects on reducing total metal and metalloid loads from contaminated groundwater (Saunders, Lee et al. 2005). Additional work was performed in Bangladesh to assess the ability of this method to remove trace levels of naturally occurring arsenic (Uddin, Shamsudduha et al. 2011). These

methods illustrate that elemental contaminants, such as arsenic, can be removed from solution by supplementation. However, detailed studies on the connection between sulfate reducing microbes (SRM) and groundwater chemistry or contamination have not been performed. To alleviate this, we have investigated the effects of sulfate and/or molasses supplementations to groundwater on microbial community structure, as well as on the aquatic chemistry.

Coupled with this work is an ongoing investigation into the shared characteristics of organisms with identified sulfate reducing pathways. This investigation is an attempt to find more granularity from a genetic standpoint of associated genes conserved among classes of sulfate reducing bacteria.

Aim 3) Determine what, if any, linkages exist between naturally occurring sulfate reducing microorganism (SRM) communities and soluble arsenic levels.

This aim requires several analysis techniques, including analysis of community structure, phylogenetic analysis of conserved genes vital to the sulfate reduction pathway, and quantification of the same genes in each sample.

Figure 3-1 shows the general practices carried out during these experiments. Not all samples (e.g., filtered water samples) were subjected to all analysis (chemical analyses). Pangenomic analysis was carried out to attempt to lend a greater level of understanding to the conserved function of SRM and potentially explain the results as well as inform future efforts.

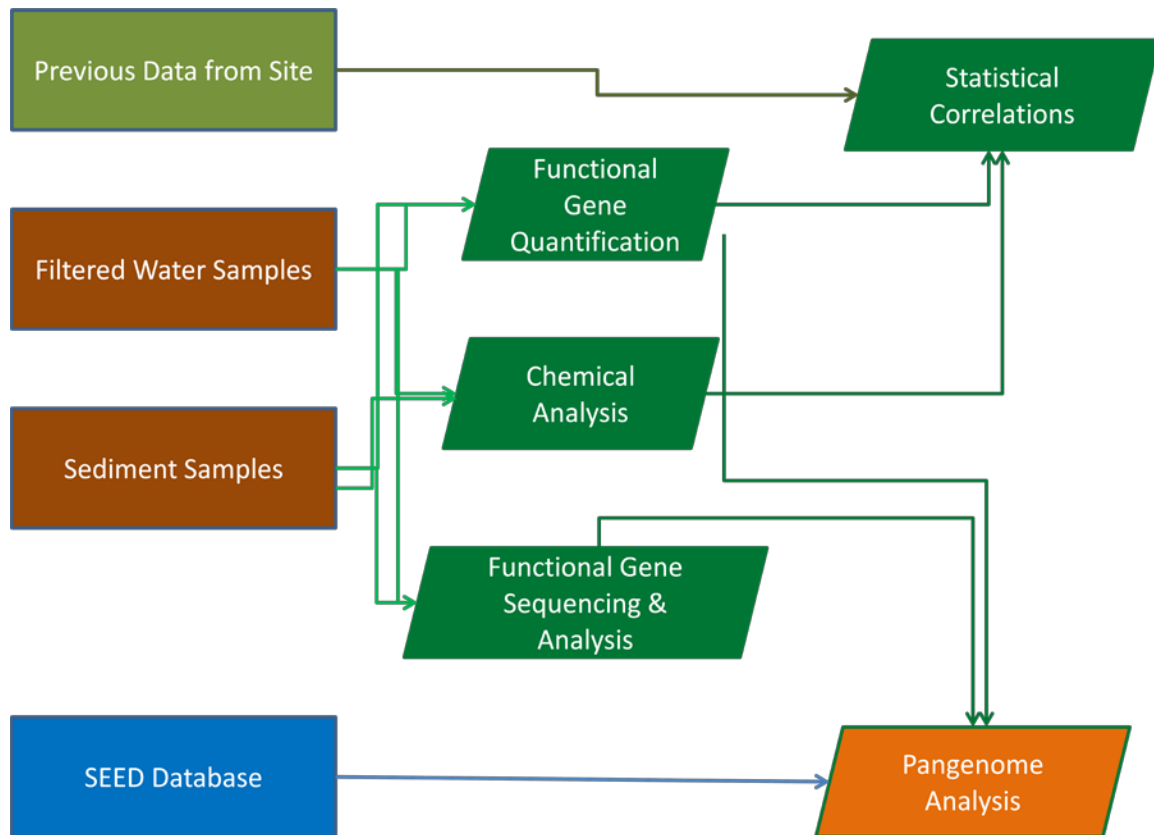


Figure 3-1. Flow chart of steps taken to study SRM in all samples

Methods

Sample gathering and processing

Water samples were obtained from individual wells by hand pumps, stored in carboys and filtered through Sterivex inline filters (Millipore, Billerica, MA). Five filter samples were obtained (labeled ML-19, ML-37, ML-39, ML-60 and ML-68) and approximately 1 L was passed through each filter. Filters were stored on dry ice until returned to the laboratory, at which point they were stored at -80°C until extraction. Due to high metal concentrations both in the original water and adhered to the filters, and interference with PCR, additional filters were not extracted. DNA was extracted using methods described in Dionisi *et al.* 2003. Briefly, filters were removed from their plastic casing, cut from the housing and placed into Qiagen FastDNA spin kit for Soil tubes (Qiagen, Valencia, CA) for extraction. DNA extraction was performed per manufacturer's protocol, with the addition of a second FastPrep bead beating step of speed 5.5X G for 30 seconds. Extracted DNA was quantified using fluorometric analysis (using Hoeffer's reagent), and stored at -20°C for analysis (Sambrook 2001).

Sediment samples for microcosms from core samples were selected from a site where the shallow aquifer showed annual cycling of sulfate concentrations, while the deep aquifer showed constant, low levels of sulfate (Dhar, Zheng et al. 2008). This site is also contaminated with arsenic, making it an ideal site to explore both the relationship between SRM and arsenic, as well as the effects of cycling of sulfate on the response to sulfate stimulation.

Rationale

Sterivex inline filters were selected due to their ability to filter large amounts of groundwater, and their ease of use, portability, the enclosed nature of the filter, as well as the existing protocol for DNA extraction. Bead beating was selected as the appropriate method for DNA extraction of these filters as it is expected that this method will lyse all bacterial cells adhered to the filter.

Construction and monitoring of initial microcosms

Sediment was transported back from Bangladesh, and used to construct preliminary microcosms with 25 mL filtered groundwater, taken from site. Five grams of sediment from Bangladesh was added to each sample under expected anaerobic conditions (5% CO₂, 5% H₂, 90% N₂), in 125 Erlenmeyer flasks and sealed with rubber stoppers. Sucrose and or sulfate were added to microcosms to generate 3 test groups and one control. Samples were sacrificed at 14, 28 and 48 days. Samples were filtered and tested for sulfate concentrations. Additional analyses of elemental content were also performed by ICP-MS performed by ACTLabs Canada. Coloration of sediment was also observed daily. Formation of black precipitates in the sediment was taken as evidence of sulfate reduction.

Rationale

This experiment was designed to mimic the addition of the principle components of molasses and sulfate. Sucrose was selected as a proxy for molasses, as it is the main sugar in molasses. The selection of time-points came as a result of observation of first appearance of precipitate in samples. Due to

the very slow doubling time of isolated SRM (8-12 Hours for *Desulfovibrio vulgaris*), time points were doubled for the two subsequent sampling points.

Sediment Samples Gathering

Sediment samples were gathered from the Araihaazar region of Bangladesh. One foot cores were sampled using the method of Van Geen *et al.*, 2008. Briefly, the technique is similar to that of digging an actual tube well, with work stopping once the desired depth for the first sample is reached, followed by sampling of 1 ft. sections by compaction. Samples were capped and stored in anaerobic cases immediately upon gathering. This resulted in 1 ft. semi-continuous core samples, each weighing ~ 500 g. Samples were stored at 4°C under anaerobic conditions until microcosm construction. Storage was intended to minimize bacterial inactivation or mortality due to freezing or long periods of storage. Samples were stored for less than 2 weeks during microcosm construction.

Microcosm Construction

Microcosms were constructed using a one foot sediment core sample (~200 g), under anaerobic conditions (83.5% N₂, 13.5% CO₂, 3% H₂). The sample was mechanically homogenized, and divided into 10 g portions. Each microcosm contained 10 g of sediment and 30 mL of synthetic groundwater. Microcosms were divided into 4 supplementation groups. Groups were supplemented to a final concentration of ppm Sulfate, 250 ppm molasses,

250 ppm sulfate + 250 ppm molasses, or a negative control containing only synthetic groundwater. Each microcosm was constructed in 50 mL conical tubes and incubated in the dark at 25°C. Triplicate microcosms were sacrificed at 14, 28 and 56 days. Water was filtered through a .22 micron syringe filter, acidified by adding 2% by volume of HNO₃ and stored at 4°C for chemical analysis by ICP-MS. Sediment was stored at -80°C for DNA extraction.

Rationale

Sampling from this site was selected due to previous work from Van Geen (Dhar, Zheng et al. 2008) indicating varying levels of sulfate concentrations over time. Additionally, the depths of the shallow aquifer and deep aquifer were known, and the general annual cycling of sulfur in these two systems was well characterized. Concentrations were calculated as an estimate of *in situ* experiments from a previously conducted experiment to add molasses and Epsom salts (MgSO₄) to a well (Uddin, Shamsudduha et al. 2011).

DNA extraction

DNA was extracted using the PowerSoil DNA kit (Mobio, Carlsbad, CA). Briefly, .5 grams of sediment were extracted per manufacturer's instructions. DNA was diluted 1:50 in Tris-HCl and stored at -20°C. Extracted DNA was used as a template for PCR (below).

PCR analysis of DNA samples

The first primers that were designed for the *dsr* gene cassettes amplified long regions of the cassette, approximately 2.1 kb.

For filter samples, PCR verification of gene presence/absence of 16S rRNA, *dsrAB*, *Apr*, *Acr3p* and *ars* (ML-19, ML-60, ML-68, ML-39, ML-37) was performed. Each sample was tested for presence/absence of each particular gene by PCR using the primers described in Table 2-3. Each sample was then run on a 1% agarose gel, stained with ethidium bromide, and photographed to determine presence/absence of detectable levels of the gene. Gene product size was determined by use of 1KB+ ladder (Fisher biosciences). All PCR reactions were subsequently subjected to TOPO 4.0 cloning, growth and plasmid isolation. One plasmid from each sample was submitted to the molecular biology resource facility for Sanger sequencing, to determine sequences of amplified gene products.

Amplicons were ligated into plasmids and transformed into competent cells using the TOPO 2.1 cloning system (Invitrogen, Carlsbad, CA), on plates containing 50 ug/mL kanamycin, and ~15 ug/mL X-Gal, for blue/white screening. Clones determined to contain an insert were isolated and sequenced by the Clemson University Genome Institute (CUGI, Clemson, SC).

Clones containing the plasmid of interest were grown and plasmids extracted using the Zyppy plasmid extraction kit (Zymo corporation, Orange, CA). Concentration of DNA was determined using a NanoDrop. Absolute copy number of plasmid per microliter H₂O was calculated, and used to generate

standard curves via serial dilutions for absolute quantification of the genes of interest during qPCR (below).

Rationale

Selection of genes for amplification and sequencing was selected as the best molecular marker for the process of interest. Selection of genes for sulfate reduction was based on literature. For sulfate reduction, *dsrAB* is the most highly conserved gene marker of SRM. These primers are ideal for binary detection (presence/absence) and phylogenetic analyses. To improve upon the original primers, additional degenerate primers amplifying the *dsrAB* region have been designed and have been added by Wagner et al (Wagner, Loy et al. 2005). Additionally, specific primer sets for *dsr* have been designed to detect the gene at several sites. This was performed to improve sensitivity, and presumably decrease false positives. To accomplish this goal, however, isolation or shotgun sequencing of *dsr* from a site must be performed (Vladar, Rusznyak et al. 2008; Kjeldsen, Tang et al. 2009).

The addition of *Apr* is to investigate the value of an alternative method of SRM detection by use of a downstream gene in the dissimilatory sulfate reduction pathway that is less well characterized. Arsenic resistance cassettes can be classified as one of two main pathways, *ars* and *Acr3p*. Both were used to determine the general presence/absence and distribution of these genes in the samples.

Amplicon Sequencing, Analysis, and Construction of Phylogenetic Trees

Sequences were examined for homology to known gene products by BLAST comparison, using default parameters on the NCBI servers (ncbi.nlm.nih.gov). Products that displayed matches to non-specific sequence were discarded. Additionally, sequences with >20% N or other ambiguity code were also discarded. Sequences with high homology to known genes were compiled and examined for signal quality using MEGA 5 software (Kumar, Nei et al. 2008).

Sequences were analyzed using MEGA 5 to create multiple alignments by use of the Muscle algorithm. Sequences were trimmed and aligned to generate both alignment trees and bootstrap phylogenetic trees. Alignments were used to create bootstrap trees using available algorithms, including nearest neighbor, maximum parsimony, and minimum evolution. Minimum evolution trees using 5000 iterations were utilized to construct and display phylogenetic trees.

Rationale

Analysis of sequence similarity and phylogeny illustrates the novelty of *dsr* sequences from the samples. This is important to describe the samples. If sequences are shown to be similar to sequenced and classified SRM, it is possible to ascribe activity to samples.

DNA Collection and T-RFLP Analysis

DNA was collected from five time points, including one duplicate DNA extraction. Extraction was performed using the MoBio PowerSoil kit, as per

manufacturer's instructions. All T-RFLP amplification steps and analyses were performed in triplicate. Extracted DNA was amplified using AccuPrimell mastermix (Invitrogen, Carlsbad, CA) using primers 5'Hex-27f, and 807r (DeBruyn and Sayler 2009). Amplified DNA was quantified, digested with HhaI and AluI for three hours as per manufacturer's instructions. Digested fragments were purified using the Wizard DNA cleanup kit and analyzed on an ABI-Analyzer 3000. Patterns were analyzed using T-REX online software to generate PCA analyses.

Chemical Analyses

All groundwater was acidified with 2% HCl and stored anaerobically until analyzed by ICP-MS at ACT-Labs Canada.

Rationale

ICP-MS analysis is one of the few methods available with the capacity to detect sub-100 PPB concentrations of metals. As the desired concentration for these experiments was sub-25 PPB levels, ICP-MS was necessary. Additionally, ICP-MS analysis results in detection of a wide range of metals and other elements (S, C, N), which allows for more statistical test to be performed.

qPCR Analysis

All qPCR protocols were carried out with using Brilliant II qPCR mastermix with SYBR green (Invitrogen Carlsbad, CA). Reactions were carried out in 96-well BioRad qPCR plates on an Opticon instrument using Chromo4 analysis

software (BioRad). Standard curves were constructed using the plasmids isolated and verified in the previous steps at a final concentration of $25 - 2.5 \times 10^7$ copies per sample in triplicate. Absolute quantities were predicted using the log linear estimation generated from this curve.

Quantitative PCR was performed for detection and quantification of *dsrB* by use of primers *dsr2060f* and *dsr4r* (Geets, Borremans et al. 2005; Geets, Borremans et al. 2006). Brilliant II SYBR Green qPCR mastermix (Stratagene, La Jolla, CA) was used for these assays. All mastermix samples were filtered through 50 kDa size exclusion filters (Millipore), to remove any potential DNA or other contaminants in the mastermix. Cycles were performed as per Geets et al. (Geets, Borremans et al. 2005; Geets, Borremans et al. 2006): Briefly, Cycling was performed as 35 cycles of 1 minute at 94°C, 45 seconds at 55°C and 45 seconds at 72°C (Geets, Borremans et al. 2005). Detection of fluorescence was performed at 3 temperatures, and the temperature with the best R^2 value was used for quantification purposes. *Apr1f* and *Apr4r* were used to determine presence/absence of the 5'APS reductase gene. The PCR was performed as 35 cycles of 1 minute at 94°C, 45 seconds at 50°C and 45 seconds at 72°C [9]. Detection of 16S rDNA was performed using conserved regions by use of the 1055f and 1392r conserved primers (Zhang and Fang 2006). Amplification was detected by use of SYBR green. The PCR program performed was incubation for 3 min at 50°C, 10 min at 95°C and 45 cycles of repeated steps of 95°C for 30 s,

50°C for 60 s, and 72°C for 20 s. Samples were normalized to mass of sediments.

Rationale

For environmental samples, SYBR green is the preferred method for quantification, due to the potential for sequence variation in any intermediate regions. While contaminants may play a role in affecting the efficiency of both PCR and fluorescence, this effect is expected to be similar for all samples of the same type (e.g. all filter samples are expected to have similar effects, all microcosm experiments are expected to have similar effects). The processes for *dsrA*, *dsrB*, 16S rDNA, and *ars* are well defined, and explored. Use of standards and standard curves, allows for normalization between experiments, and direct comparison of samples. The introduction of additional *dsrB* primers (Wagner, Loy et al. 2005), was done to determine the effects, if any, of multiplexed primers to *dsrB* quantification.

Pangenomic Analysis of Sulfate Reducing Microbes

Several approaches were utilized to identify potential SRM. Known SRM identified explicitly by genus (*Desulfovibrio*, etc.), were selected and compared. Additionally, gene annotation information from deposited genomes was mined to select genomes annotated to encode for one or more proteins involved in dissimilatory sulfate reduction. Annotations and subsystem information of selected genomes were then downloaded and compared to each other.

Selected genomes were analyzed in two manners. First, genomes were grouped into families and subjected to analysis in a method based on orthologs (Fulton, Li et al. 2006). This method creates pairwise groups of orthologous protein sequences, based on clustering compared to an outgroup, followed by reciprocal best BLAST comparisons between the two species. The custom program is able to combine multiple pairwise comparisons into a single pangenome for all species considered. In this case, all identified gamma and delta proteobacteria were analyzed, along with *clostridia* genera, and archaea. The selected outgroup was *Mycobacterium marinum*, an obligate aerobe. Generated pangenomes were analyzed to determine how many shared protein clusters existed between any and all species within the groupings, as a measure of genome similarity.

To perform a core gene genome analysis, it is necessary to obtain genome annotations that are performed by a single source, with consistent gene calling techniques. As all selected genomes have been annotated and placed into the SEED database and stored by RAST (Aziz, Bartels et al. 2008), a second level of analysis is also possible, by using SEED subsystem style analyses to determine what metabolic subsystems are encoded by each genome.

All available bacterial and archaeal genome annotations were compared to select genomes with dissimilatory sulfate reducing capacity. These genomes

were then analyzed using the SEED FigureFAM groupings, along with subsystems groupings to find genomes with similar metabolic pathways.

Rationale

These steps are performed to give a better understanding of the level of relatedness of the class of sulfate reducing microorganisms in regards to protein level similarity as well as shared metabolic potential. Additionally, selection of the appropriate genomes was an iterative process, to determine the correct information to use for filtration. Final selection was based on *dsrAB* presence/absence, resulting in the final genome set outlined.

Pan and Core Genome Calculations

Genome annotations were downloaded from NCBI and SEED, and searched for components of dissimilatory reducing pathways, including *dsrA*, *dsrB*, *dsrC* and *dsrMJKOP*. For NCBI annotated genomes, selected genomes were compared to each other by use of several pangenomics analysis tools, particularly one based on ortholuge, written by Tracey Freitas (Los Alamos National Laboratory). Pangenomic analysis of this nature is based on gene similarity, to form groups of orthologous function, meaning that functional information is lost, but comparisons between genomes of interest can be performed.

Pangenome Analysis

The pangenome was created in two major stages. The first stage involved the identification of ortholog families using the pairwise relationships described above. The second stage is the merging of all pairwise orthologs identified into larger classes. Six tests were performed to assess the validity of these ortholog families representing the pangenome:

- 1) All orthologs identified in Ortholuge must remain in families of no less than 2 members after the merging process (pairwise limit).
- 2) Ortholog families can have no more ortholog sequences than there are genomes.
- 3) No gene can be a member in more than one family (non-redundant gene requirements).
- 4) Ortholog families may only contain 1 gene from any genome(non-redundant source).
- 5) All orthologs identified must remain in the pangenome (gene persistence).
- 6) Finally, all genes from within the organism's FASTA files must be placed uniquely into the pangenome (equal gene representation).

Rationale

This method, still under development, when it is able to be applied to the number of genomes selected, will generate a core and pangenome based solely on gene similarity matrices, rather than annotation based matrices. This distinction is important, as it allows a determination of the similarity of the whole

genomes of all selected organisms. If COG style annotations are available, granularity of the similarities can be improved by examination of individual pathways, genes, etc.

Analysis of Annotated Genomes Using SEED/RAST Annotations

To analyze shared metabolic potentials, SEED annotated genomes were compared. SEED annotated genomes are classified first as individual protein annotations, then classified into both annotation groups (e.g. *dsrA*, *Acr3p*, etc) as well as into functional classifications (e.g. respiratory sulfate reducing genes). Preliminary analysis was performed using gene annotations, followed by analysis of encoded subsystems.

Genomes containing *dsrMJKOP* were selected first, as this list was the least permissive. To add to this list, genomes with annotated *dsrA*, *dsrB*, or *dsrC* functionality were included. However, due to the high number of genomes with no known sulfate reduction capacity (including multiple *Salmonella* and *Escherichia* species and strains), additional steps were required to remove those genomes without annotated sulfate reduction pathway genes. To remove non-SRM from the list, SEED annotations of pathways were utilized as follows:

Each genome with an annotated copy of a *dsr* Gene was examined for presence/absence of sulfate reducing pathways, which are conveniently annotated as a different group than general anaerobic respiratory reductases. All genomes without at least one annotated member of a sulfate reducing pathway

were assumed to have been incorrectly annotated to contain *dsr* genes, or that the *dsr* genes in question did not result in dissimilatory sulfate reduction.

Genomes were classified by isolation location, and grouped as Freshwater, Saltwater, Sediment, Animal Wastewater, Extremophile (halophiles and thermophiles), and taxonomic details. Additional work to group these species by ecotype was also performed, with limited success, either due to lack of information as to isolation, or due to the general ubiquity of some organisms in the groups.

Rationale

Identification of SRM is difficult, both due to the number of organisms and complexity of annotation. The iterative process of selecting genomes was similar to the previously described method of selection of genomes with annotated *dsr* gene products. The widespread ecological niches for SRM means that there have been potentially millions of years of evolutionary pressure specific to their environmental niches. Grouping by taxonomy and/or ecological niche allows for finer granularity of analysis.

Subsystems Analysis

For selected genomes, the subsystems for the annotated proteins is a highly useful method to determine the potential metabolic capacity of the organisms. Genomes were analyzed for their capacity for motility, carbon utilization, and metabolite production. These data were examined and compared to the results of microcosm analysis, to determine if current knowledge of

sequenced SRM was able to support the hypothesis that SRM growth is directly or indirectly encouraged by addition of complex carbon nutrients to the microcosm.

Rationale

The use of subsystems for a top-down analysis of genomes allows for a more compact analysis for comparative genomics. Presence or absence of a particular subsystem is the first level of analysis, followed by the analysis of number of entries in the subsystem per genome. Additionally, subsystems can be more easily used in PCA or other dimension reduction statistical analyses.

Results and Discussion

Arsenic Concentrations of Filter Samples

Arsenic concentrations of wells from which filter samples were taken are displayed in Table 4-1. Briefly, arsenic concentrations ranged from 1-72 ppb in 5 wells chosen for filtration (ML-19, ML-37, ML-60 and ML-68). Additional measurements of iron, phosphorus, and sulfur were also taken. There are correlations between these concentrations. Arsenic is only correlated to manganese in these samples. Manganese also shows weak correlation to iron concentrations, which also shows significant correlation to sulfur concentrations. Results can be seen in Table A-3 and Table A-4.

Table 4-1. Arsenic concentrations of well sites from which filtered water samples were taken.

Sample Site	Arsenic Concentration (PPB)
ML-19	28.2
ML-30	42.8
ML-37	8.7
ML-60	20.4
ML-68	72.5

***dsrAB* Detection and Sequencing**

Amplicons for *dsrAB* were detected by PCR in samples ML-37 and ML-68, with ML-68 having a stronger banding pattern. Approximately 90 sequences from each of these sites were obtained, aligned, and used to construct a phylogenetic tree (Figure 4-6). This tree indicates that there are few clusters containing more than one site based on *dsr* sequence identity.

***arrA*, *dsr* and 16S rDNA Detection and Quantification**

arrA was not detected in any wells by PCR analysis. Additional attempts to generate clone libraries of PCR samples without banding patterns also failed to produce evidence of arsenic respiratory reductases in any samples. *dsrA* primers for the detection of *dsrA* by alternative primers were also unable to amplify a detectable band in any samples. *dsrB* primers were able to amplify and quantify *dsrB* from all samples. Figure 4-1 shows the quantity of 16S rRNA and *dsr* extracted from samples as determined by qPCR.

16S rDNA Detection and Sequencing

Primers for 16S rDNA amplified in all samples. The amplicons were cloned and resulting in approximately 90 16S rDNA sequences from each site. Sequences were trimmed, aligned, and used to construct phylogenetic trees (Figure A-1). This tree indicates a number of sequences similar to *E.coli*. Additionally, several sites appear to contain 16S rDNA sequences that are similar

to each other (ML-68 clusters with ML-30, etc). It is also valuable to note that there does not appear to be sequence similarity between ML-68 and ML-37 at the 16S rDNA level, indicating that community structure, as well as *dsrAB* sequences are divergent at these two sites.

Quantification of *dsrB* and 16S rDNA

dsrB and 16S rRNA amplicons were quantified for all samples as shown in Figure 4-2. Briefly, *dsrB* quantity was not an indicator of arsenic concentration, nor was 16S rDNA quantity directly related to copies of *dsrB*. Detected levels of *dsrB* were higher in both samples found to have detectable amplification of *dsrAB*. The ratio of *dsrB* copy to 16S rDNA copy number ranged from .001 to .01, indicating that in these environments, from .1% to 1% of the community is capable of dissimilatory sulfate reduction at the time of filtration.

dsrB copy number did not correlate significantly to any measured chemical factor using either Spearman ranks or Pearson tests for correlation. Similarly, there were no significant correlations between the *dsrB*/16S rDNA copy number ratio in any of the measured samples. Manganese appeared to have a weak (0.06) correlation to the ratio. This may be an artifact of correlation to total 16S rDNA copy number.

Using the Spearman ranks correlations, 16S rDNA copy number was found to correlate significantly to all chemical factors, except for arsenic (Table A-3). This would indicate that total biomass in a sample has a relationship with the

overall level of minerals in the water, but does not appear to have a linear relationship.

Microcosm Detail

It was discovered, after microcosms were sacrificed and analyzed, that the samples were not incubated under anaerobic conditions. Those samples that were supplemented with sucrose became anaerobic (as per chemical indicators), but negative controls and sulfate-only experimental conditions did not appear to remove oxygen from solution. Additionally, due to the methods of filtration of water after receipt, metal concentrations including iron and arsenic, were nearly undetectable. This made analysis of anything other than basic reaction kinetics impossible. Therefore, DNA from these samples was not extracted. However, ICP-MS analysis indicated that sulfur was removed from the aqueous phase and, in addition, analysis of the sulfate concentration by ionic analysis indicated that sulfate was indeed being consumed in all microcosms with sulfate addition. This is not a perfect indication of dissimilatory sulfate reduction, as assimilatory sulfate reduction can also occur, leading to sulfite being stored in bacterial cells, or incorporation of sulfur into thiols, cysteines, etc. However, combined with the other evidence, it is likely that SRM activity played a role in consumption of sulfate.

T-RFLP Analysis

T-RFLP Patterns

Two PCA charts were generated (Figure 4-2) to illustrate the clustering of molasses-supplemented versus molasses-free microcosms. The restriction enzyme *AluI* shows a more easily distinguished response. These responses can be measured either by simple mathematical algebraic equations PC2, appears to be sufficient to determine the difference between molasses and molasses-negative samples), or by use of clustering equations. The use of clustering is less viable, however, due to low sample number. While fragment size is the necessary variable for determination of T-RFLP patterns, the abundance measurement can be analyzed by a number of different methods. In methods similar to DGGE, the presence/absence of a band can be analyzed to generate a binary matrix. For measurements of abundance of each fragment, peak height or peak volume can also be calculated and utilized for analysis. In all cases, resulting PCA analyses resulted in similar prediction patterns.

T-RFLP analysis resulted in a very few detectable peaks in microcosms not supplemented with molasses. This was replicated across several replicates of the experiment, indicating a simple community structure, with few dominant species. The corollary conclusion to this is that the increased diversity in samples was linked to molasses addition. This diversity is not decreased with the addition of sulfate, also supporting the possibility that molasses addition and, by proxy, addition of carbon sources, function to increase diversity while addition

of electron acceptors can increase the abundance of individual taxa, but do not have major effects on the overall diversity.

The ability to utilize single PCA as an indicator for exposure to molasses in these conditions can also lead to the conclusion that there are specific biomarkers that can be used to determine the addition of carbon sources. To further analyze this, microcosms constructed with additional simple carbon sources could also be constructed and analyzed to determine if there are additional markers or if the response to all carbon sources is similar from a community structure standpoint. This analysis is important to microbial ecology from the standpoint of bioremediation: better understanding of what nutrient additions cause shifts in the community, and whether they are important to the desired effect, has the potential to reduce variability and costs of remediation efforts.

The addition of molasses as a carbon source was chosen to mimic *in situ* conditions of an experiment performed by Saunders et al., published in 2008, which indicated that arsenic was effectively mineralized upon addition of molasses and sulfate to an anaerobic aquifer in Bangladesh (Saunders, Lee et al. 2008). This experiment was limited to chemical analyses, without subsequent analysis of the microbial response. Typical measurements of the response would be performed by targeting genes of interest (sulfate reducing genes, iron reduction, 16S rDNA sequencing of SRM, etc). This study has

illustrated the value of community profiling on surveys such as this to improve understanding of the effects of *in situ* supplementation.

Molecular and Elemental Analysis

The results of qPCR and ICP-MS analysis of water and data gathered from microcosm samples were averaged, with the results shown in Table 4-2. Analysis is included below; however, it is worth noting that *dsrB* levels were elevated in all samples after time 0. Additionally, 16S rRNA levels were elevated in all samples, but average copy number was much higher in samples supplemented with molasses as a carbon source.

Table 4-2. Mean counts of qPCR detection levels of 16S rDNA, *dsrB* and *Apr* genes

Supplement	Days	Mean 16S copies/uL	Mean <i>dsrB</i> Copies/uL	Mean <i>Apr</i> Copies/uL
Molasses +Sulfate	14	2.92E+05	2.09E+04	5.84E+03
	28	1.89E+05	1.59E+04	5.77E+02
	56	1.32E+05	1.59E+04	1.75E+03
	Total	2.05E+05	1.76E+04	2.72E+03
Molasses	14	2.98E+05	1.16E+04	1.53E+03
	28	2.23E+05	2.01E+04	3.65E+03
	56	3.34E+05	2.32E+04	1.24E+03
	Total	2.85E+05	1.83E+04	2.14E+03
No Addition	14	5.94E+04	5.51E+03	5.68E+02
	28	5.40E+04	1.15E+04	1.97E+02
	56	1.56E+04	1.05E+03	2.85E+03
	Total	4.30E+04	6.03E+03	1.21E+03
Sulfate	14	6.11E+04	1.51E+04	9.18E+02
	28	6.55E+04	2.42E+04	1.16E+04
	56	6.04E+04	1.52E+04	1.32E+03
	Total	6.23E+04	1.82E+04	4.60E+03

Gene Copy Number Changes in Samples as a Function of Time

At 19 feet (6.3 m), the majority of shift as a function of time and supplement occurs between 0 and 14 days. During this time, the mean copies of *dsrB* increase in all microcosms. There are also significant differences between negative controls and microcosms supplemented with sulfate, either alone, or in combination with molasses. There is also a detectable change in *Apr* copies in all amended microcosms at 14 days. Additionally, 16S rDNA copies are also higher between these two time points (Figure 4-4). Subsequent changes in *dsrB* or *Apr* gene copy are not statistically significant.

At 28 feet, there is no such shift away from the initial values of *dsr* or *Apr*. Copies of 16S rDNA are increased, however, with addition of molasses to any sample. At this depth, at 14 days and beyond, there is a detectable difference between samples in *dsrB* copy number, with all supplemented microcosms shown to have a higher copy number of *dsrB*/g sediment than negative controls. This is true at both 14 and 28 day time points. At 56 days, average *dsrB* copy number increased in the negative control, to show no significant difference between treatments. *Apr* changes show a lower detection overall, but do indicate significant changes between negative controls, molasses, and molasses + sulfate amended microcosms at 14 and 28 days. Differences between controls and microcosms with molasses added show differences at 28 and 56 days, whereas molasses + sulfate microcosms show differences at 14 and 56 days (Figure 4-4).

16S rDNA Quantification and Variation Between Samples

Quantities of 16S rDNA were consistently higher than *dsrB* quantities ($P=0.00$). As expected, the addition of molasses resulted in a higher quantity of 16S rDNA copies/g sediment. The variation is higher in samples amended with both sulfate and molasses. Total 16S rDNA copies appear to decrease in these samples, using simple linear projections, with an R^2 value of .49. Comparisons between day 14 and Day 56 yield a significant difference using a cutoff value of 0.1 (Table 4-3).

In sum, there are individual groups (19 ft (6.3 m) Molasses + Sulfate, 25 ft (8.3 m) Molasses) which show significant differences between 14 and 56 day time points (decreased for 19 ft (6.3 m) molasses + sulfate and increased for molasses), but these shifts are less than five-fold changes. There does not appear to be any evidence that these shifts are correlated to changes in *dsrB* quantities.

Table 4-3. Correlations between measured values at 19 ft (6.3 m) . Samples of 14,28 and 56 days were pooled for this analysis. Full set of statistical data are in Appendix Table A-3 and Table A-4 Type of correlation and significant of correlation are displayed in cells.

Comparator	Phosphorous	Sulfate consumed	# Copies <i>apr</i>	Arsenic
# Copies <i>dsr</i>	Positive	Postive	Positive + 0.05	X
Sulfate Consumed	Positive	X	X	X
Iron	X	X	X	Positive

***apr* Quantification and Variation Between Samples**

apr levels were low in all samples (average below 1000 copies/g sediment). There are no distinct trends in these data. While there are several points that show discrete differences from other nutrient additions at a given time point, these show no linkage to the *dsrB* data. Additionally, at 25 feet (8.3 meters), the measured values do not differ statistically from the 0 day sample. All samples, other than 28 day negative microcosm and 14 day sulfate addition are higher than the 0 day values at 19 feet (6.3 m). This may be due to the previously noted high variability and horizontal transfer of the *apr* gene, with little linkage phylogenetically to *dsr*, leading to a large pool of *apr/apr*-like genes that are not amplified by the primers currently available. As there is not a definitive pathway described for dissimilatory sulfate reduction, it is also possible that *Apr* is not conserved in all sulfate reduction pathways. This would be possible, if any form of convergent evolution allows another APS reductase to remove a phosphate from APS to form sulfite.

***dsrB* Variation Between Samples**

Examination of microcosm samples over time yielded no statistically significant changes in *dsrB* copy number after initial shifts from 0 to 14 days at either depth ($p=1.00$) by ANOVA comparisons or Bonferroni post-hoc analysis (Appendix Table A-3 and Table A-4). Additionally, there appears to be an effect over time on both sulfur ($P=.000$) and Iron ($P=0.003$) using one-way ANOVA, but

no correlation between *dsrB* levels ($P=.982$) or arsenic concentrations ($P=.499$) as related to time points 14, 28, or 56 days. This would indicate, then, that sulfate reducing capacity of the samples are encouraged between 0 and 14 days, and remains stable over at least a span of 6 weeks when groundwater is amended with nutrients involved in sulfate reducing pathways.

There is a lack of significant difference between nutrient addition effects on *dsrB* copy number at either depth, indicating that either nutrient (molasses) or terminal electron acceptor (sulfate) are sufficient to encourage growth of SRM in these samples.

The variability of response is large between samples at individual time points, but is normal when all non-zero time points are combined. There is no significant ($P=0.05$) correlation between arsenic and *dsrB* copy number in any sample groups.

Chemical Variation Between Samples

Iron, sulfur, arsenic and phosphorus concentrations over time are illustrated in Figure 4-7. Over the course of 56 days, soluble sulfur decreased over time in all cases. Iron increases in all cases between 0 and 14 or 28 days, with all samples decreasing again by 56 days. This decrease is to levels approaching 0 ppm, except in microcosms supplemented with molasses alone. Phosphorous increased over time, for all amended microcosms at both depths. This increase is more pronounced both at 19 feet (6.3 m), as well as in samples

supplemented with magnesium sulfate. Arsenic levels varied over time. In all samples, there was a measurable decrease in arsenic over the first 14-28 days, followed by a degree of return of arsenic to soluble phase. This late increase in soluble arsenic was highest at both depths in samples amended with sulfate alone or molasses alone.

Sulfate was consumed in all cases, with a significant difference in remaining sulfur in solution between magnesium sulfate amended samples and non-amended samples at 56 days (Table 4-4), indicating that not all sulfate is consumed in non-carbon amended samples. Rates of total sulfate consumption indicate that the majority of sulfate is utilized before 28 days in the cases of sulfate addition. Molasses adds a small, but detectable, level of sulfur to the media, which is removed from soluble phase with the majority exhausted by 28 days.

The rate of increase in phosphorous availability in solution was different between supplement and depth. At 19 feet (6.3 m), the highest phosphorous concentration was found in samples amended with both molasses and sulfate. The overall increase in phosphorous was over three times the initial level measured at 0 days (Table 4-4).

There is also a significant difference in soluble arsenic concentrations between treatments. While arsenic decreases in all samples from 0 to 14 days, between 14 and 56 days, arsenic increases in sulfate supplemented microcosms. This increase was consistent between depths, resulting in arsenic levels

approximately equal to day 0 concentrations by 56 days at both depths. Sulfate samples at either depth, or combined, is significantly different ($P=0.00, 0.00, 0.03$ for Molasses, Molasses+Sulfate, and Negative control respectively) from the other treatments. There is also a slight separation ($P=0.18$) between molasses amended samples and the negative control.

Comparison of Chemical Analysis to Molecular Analysis

There are distinct linkages between chemical factors and the community response. Specifically, molasses, as expected, has a sustained increase in biomass, as indicated by a much higher 16S rDNA copy number.

Additionally, *dsrB* copy number is most closely correlated with addition of molasses, rather than sulfate. While it was expected that molasses addition, coupled with addition of magnesium sulfate to the microcosms, would result in a large increase in SRM, this shift was observed in all additions, with the most notable increases in *dsrB* copy number occurring in microcosms supplemented with either molasses or molasses and magnesium sulfate.

There is an apparent trending of SRM downward over the course of 48 days in molasses-only supplemented microcosms; however, this trend is not statistically significant.

Table 4-4. Median and Deviations of functional genes associated with dissimilatory sulfate reduction, and phosphorous and arsenic concentrations in the aqueous phase of all samples, divided by depth and timepoint.

Days/Depth (ft)		Apr Copies/ 16S copies		Phosphorus (ppm)		Arsenic (PPB)		Dsr copies/ 16S copies	
		Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
14	19	2.21E+03	3.95E+03	1.62E+00	9.95E-01	4.36E+01	4.54E+01	1.35E+04	6.78E+03
	25	4.74E+02	6.66E+02	7.14E-01	2.70E-01	1.16E+02	1.12E+02	2.74E+03	3.14E+03
	Total	1.34E+03	2.91E+03	1.17E+00	8.50E-01	7.98E+01	9.12E+01	8.13E+03	7.56E+03
28	19	3.99E+03	9.67E+03	1.50E+00	1.27E+00	6.25E+01	6.87E+01	1.79E+04	1.68E+04
	25	1.01E+03	3.15E+03	1.31E+00	1.48E+00	7.70E+01	9.24E+01	3.10E+03	4.71E+03
	Total	2.50E+03	7.20E+03	1.41E+00	1.35E+00	6.97E+01	8.00E+01	1.05E+04	1.42E+04
56	19	1.79E+03	8.32E+02	2.86E+00	2.17E+00	1.02E+02	4.67E+01	1.41E+04	1.01E+04
	25	4.76E+02	5.52E+02	9.82E-01	6.88E-01	9.08E+01	8.06E+01	2.14E+03	1.31E+03
	Total	1.13E+03	9.63E+02	1.92E+00	1.84E+00	9.66E+01	6.47E+01	8.13E+03	9.34E+03
Total	19	2.67E+03	5.96E+03	2.00E+00	1.64E+00	6.95E+01	5.86E+01	1.52E+04	1.18E+04
	25	6.53E+02	1.85E+03	1.00E+00	9.58E-01	9.46E+01	9.44E+01	2.66E+03	3.28E+03
	Total	1.66E+03	4.49E+03	1.50E+00	1.42E+00	8.21E+01	7.90E+01	8.92E+03	1.07E+04

Pangenomic Analysis of SRM

Threaded Ortholuge Analysis

From searching all finished and annotated NCBI genomes, 122 were found to be annotated to contain *dsrA* or B. These species were divided into groups by family. Again, supplementary Table A-1 and Table A-2 display the genomes selected for this study. Core genomes for each family indicate that a limited number of genes are shared. Due to the computational limitations of performing multiple core genome calculations using the ortholuge program, only groupings by 16S rDNA phylogeny were performed. This method had limited success due to the highly variable nature of gene content at the family level. While all grouping shared a number of orthologous genes, due to the lack of complete knowledge base as to annotated genes, it is difficult to determine the actual role of these shared genes. Additionally, the orthologue program, and its successors, are designed to find related genes between close relatives, and do not take into account divergent sequences with similar functions. The need to compare gene function, rather than orthologous genes, is required for a metabolic capacity study, as was originally intended in this work. To further this end, the SEED database (which serves as the backbone to the RAST annotation server) was utilized to select sequenced and annotated genomes which encoded for genes conserved in the dissimilatory reducing pathway.

Analysis of SEED Annotated Genomes and Selection of Genomes

Table A-1 shows the results of searching for *dsrA*, *dsrB*, *dsrC* within SEED annotated genomes. Genomes determined to be linked to dissimilatory sulfate reducing pathways can be seen in Table A-2. As can be seen, *dsrA*, B, and C are more widely identified than the MKJOP complex, including such organisms as *Escherichia coli*, and numerous *Clostridia* and *Vibrio* species, which have not been demonstrated to use sulfate as an electron acceptor, or any anaerobic respiratory functions. Preliminary assessment of the lists of organisms with annotated copies of *dsrA*, *dsrB*, or *dsrC* as compared to the compact list of sulfate reducing micro-organisms, indicates a number of organisms with annotated genes with no known activity of dissimilatory sulfate reduction. Using the method described above to remove non-sulfate reducing microbes from the list, 60 annotated genomes with known sulfate reducing capacity remained.

Annotation comparisons for each group were performed, and the total number of genes for each organism falling within each functional gene category was tabulated. This resulted in a matrix of 87 genomes by ~ 500 functional categories. Figure 4-8 shows the results of PCA analysis performed on this matrix with clustering performed either by phylogenetic characterization or by habitat. While the majority of conserved genes fell into conserved categories (GTPases, ABC transporters, etc.), several interesting items were observed. Archeal genomes were obviously most similar to each other, but only conserved ~20 clusters across all genomes; conversely, *deltaproteobacteria*, with many more representatives, showed the most genetic maintenance of the groups, with

> 60 groups conserved across all 30 species. Clostridia were even more highly preserved, with 80 gene clusters found to exist in all 8 bacteria within the Clostridia family with annotated dissimilatory sulfate reducing capacity.

What can be shown from PCA plots and analysis of gene clusters is that genomes with the highest level of similarity in functional gene clusters are better explained by phylogeny than isolation source. This may be a result of gene conservation in a hereditary fashion. However it may also be due to the ability of many genera (particularly *Desulfovibrio*) to colonize multiple habitats, making classification of organisms by their habitats difficult. Choosing alternate methods of classification or subclassifications can also result in better groupings, such as dividing extremophiles into hyperthermophiles and halophiles. Archea have the most distinct signatures of gene clustering, as a result of their evolutionary divergence. Additionally, removal of genomes with very few similarities to the remaining groups (by removing the genome responsible for the fastest decline in shared functional clusters) greatly improves the clustering of each group. It is important to note that all species identified by this manner shared sulfur-reducing pathways, as well as additional families involved in sulfur metabolism, indicating that this method of analysis is capable of selecting groups appropriately.

Subsystems Analysis

Of 60 annotated genomes, all encode for expected subsystems, including anaerobic respiratory reductases, and sulfate reduction associated complexes. In the genomes with conserved subsystems for particular regions of interest, there are several metabolic pathways that can be considered highly conserved, including pyruvate metabolism, glycogen metabolism, glycerolipid and glycerophospholipid metabolism, indicating the ability to consume secondary metabolites including pyruvate glycogens. This supports literature suggesting addition of pyruvate, citrate, or lactate as a carbon source to encourage growth of SRM, as well as other anaerobic microbes capable of some form of anaerobic respiration. Additionally, the succinate-dehydrogenase family of proteins is annotated in this set of genomes, potentially serving as the electron shuttle for anaerobic sulfate reduction.

Glycolysis and gluconeogenesis pathways are also found in more than 50 of the annotated genomes, ranging from 3 to 15 members of known glycolysis or gluconeogenesis activity. This would suggest, rather strongly, that coupled with the glycogen and glycerolipid degradation pathways, these genomes are capable of utilizing complex carbon sources as electron donors for sulfate reduction. However, only three genomes have annotated sucrose metabolism annotations. As sucrose is the major component of molasses, this suggests that SRM are not directly responding to molasses supplementation, but are growing syntrophically, or consuming metabolites from fermentation reactions by other organisms in the community.

The capacity of SRM to interact with nitrate is also investigated, due to the observed interplay of nitrate and sulfate *in situ*, namely that the presence of nitrate has been suggested to inhibit sulfate reduction. There are three major types of subsystem annotated by SEED for nitrogen compounds, nitrate stress response, nitrogen assimilation, and nitrate reduction. While there are only three (5%) genomes with annotated nitrate reduction capacity, 52 have nitrosative stress responses, which may include gene regulation, illustrating a potential pathway for nitrate to regulate sulfate reduction. More than half of the genomes analyzed have nitrogen fixation and/or ammonification pathways, indicating that multiple mechanisms for SRM to interact with nitrogen, both for nutrition, as well as chemical modification.

Nearly all genomes in question encode for a subsystem for cobalt-zinc-cadmium resistance and fluorquinone resistance (59 of 60). Additionally, more than half of selected genomes encode for arsenic resistance or zinc resistance, with some overlap, meaning that 26 genomes encode for both zinc and arsenic resistance, an additional 26 encode for only one or the other, and 8 do not encode for resistance to either zinc or arsenic. This can be coupled with indications that approximately the same number of genomes (31) are capable of selenate uptake to indicate that the majority of SRM are primed for resistance to heavy metals.

Grouping by taxonomic groups was also performed, using *deltaproteobacter*, *gammaproteobacter*, *Clostridia*, or Green/purple sulfur

bacteria as groupings. This accounted for 50 of the genomes. If this association were not related to subsystem distribution, the number of subsystems expressed in all members of an individual group can be expected to vary as an inverse function of the number of genomes in the group (i.e. the more genomes, the fewer shared subsystems). Table 4-5 shows the calculations for this variable, with a range of 141 (green/purple sulfur bacteria) to 92 (deltaproteobacteria). Some degree of this variation can be explained by random distribution; however, when environmental niches are applied, these numbers increase at a rate that is not supported by the data for all genomes.

Analysis of the Deltaproteobacteria with known SRM activity shows that even though it is the largest subset (27 genomes), there are a number of more well conserved subsystems in these systems. These include ABC transporters for branched chain amino acids, and synthesis of aromatic compounds. Of the 7 Clostridial species, the majority (4) were isolated from hydrothermal vents, or hot springs. While no trends or conserved subsystems can be observed from all Clostridia, examination of thermophiles only indicate that all encode for several interesting subsystems, including carbon starvation, hydrogenases, and heat shock associated complexes. Green/purple sulfur bacteria show a high degree of similarity between genomes, showing the most shared subsystems, even though they have the 2nd fewest members. This conservation does not appear to be strongly linked to environment, as removal of thermophiles results in only the

additional subsystem for cold shock being conserved among non-thermophilic green/purple sulfur bacteria.

Table 4-5. Analysis of SEED annotations of annotated SRM, grouped by taxonomic categories.

Classification	Clostridia	Delta proteobacteria	Gamma proteobacteria	Green Sulfur Bacteria	Miscellaneous Bacteria	Archea
# of within division Organisms	6	30	6	17	13	14
# Shared Subsystems	80	65	97	67	53	38

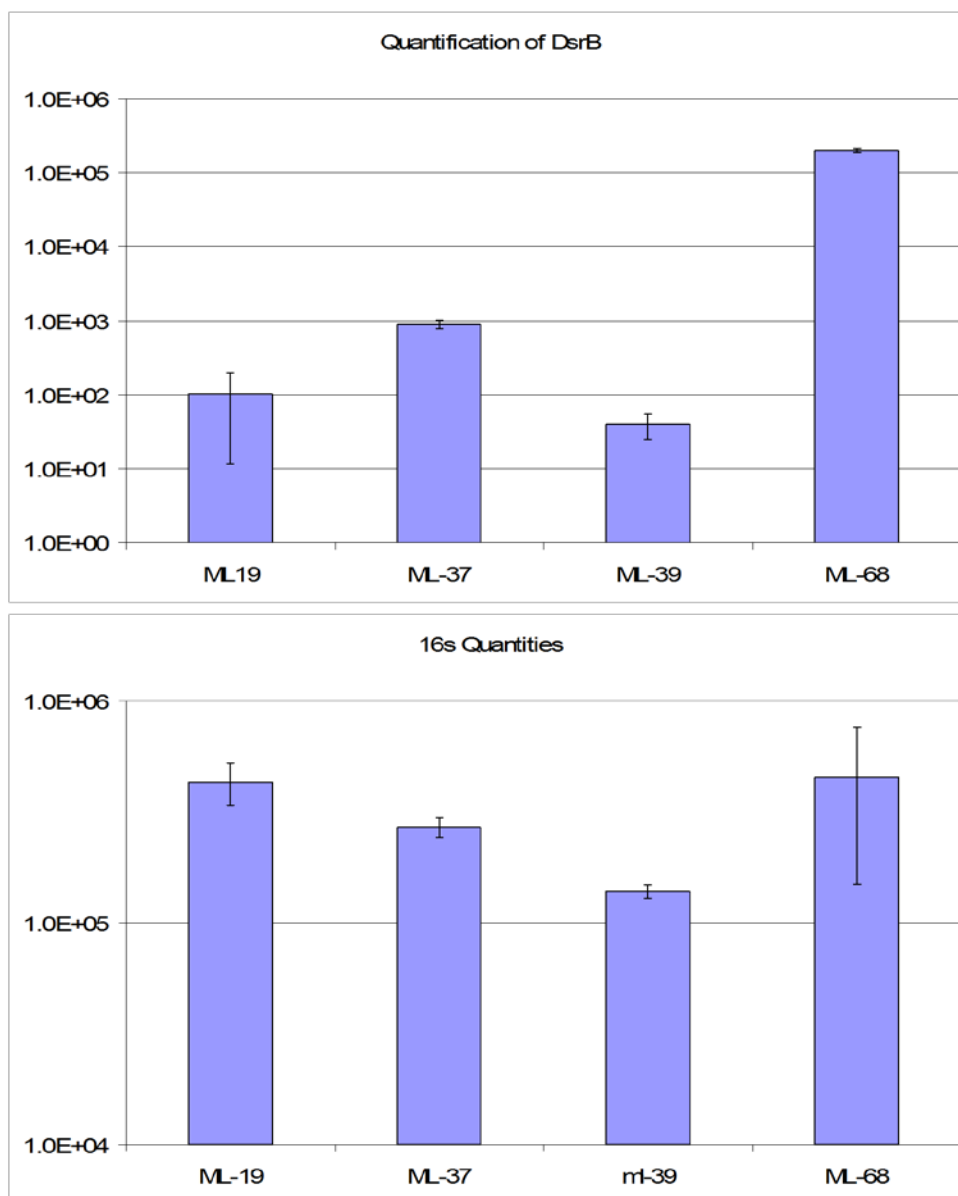


Figure 4-1. Averaged absolute quantification of *DsrB* and 16s rDNA genes from filtered water samples. Error bars represent variation in samples. There is no significant difference between quantities of *dsr* in either ML-19 or ML-39. Compared to detectable 16s rDNA quantities, it is clear that ML-68 and ML-37 have measurably higher quantities of *dsrB* gene both per uL of extracted DNA and per 16s rDNA copy than the other 2 wells. ML-60, the 5th site had exceptionally low 16s rDNA detection levels, indicating poor DNA extraction efficiency. Absolute abundances of *dsr* are approximately equal, but coupled with variable 16S rDNA quantities, indicate that ML-68 has the lowest ratio of *dsr*:16S rDNA.

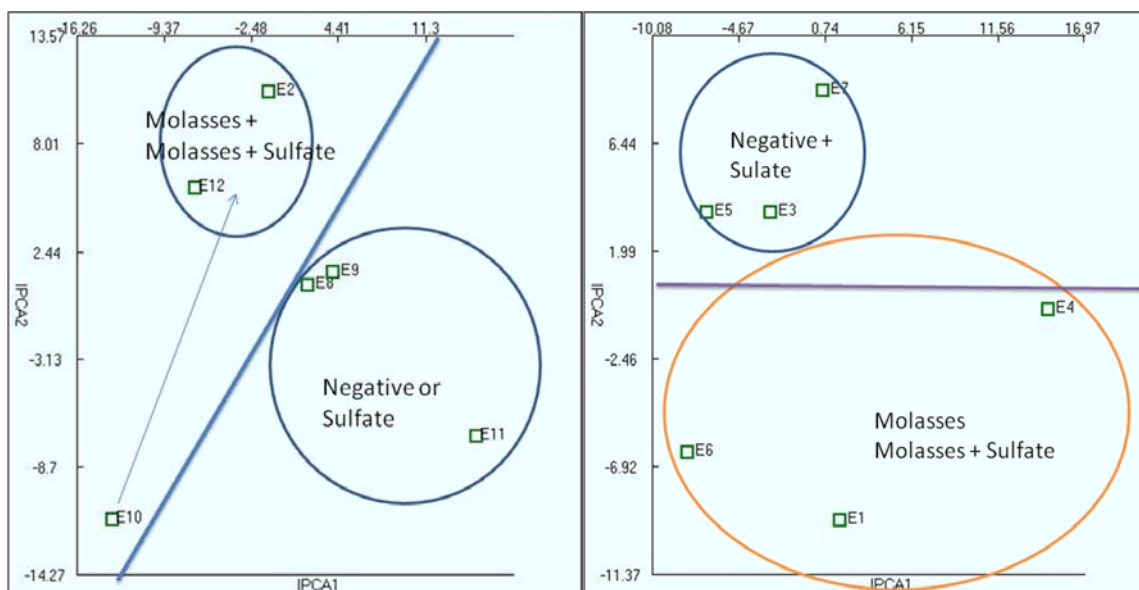


Figure 4-2. PCA plots of T-RFLP analysis of microcosms constructed using AluI digests. Loose clustering is shown between molasses supplemented and molasses negative data sets.

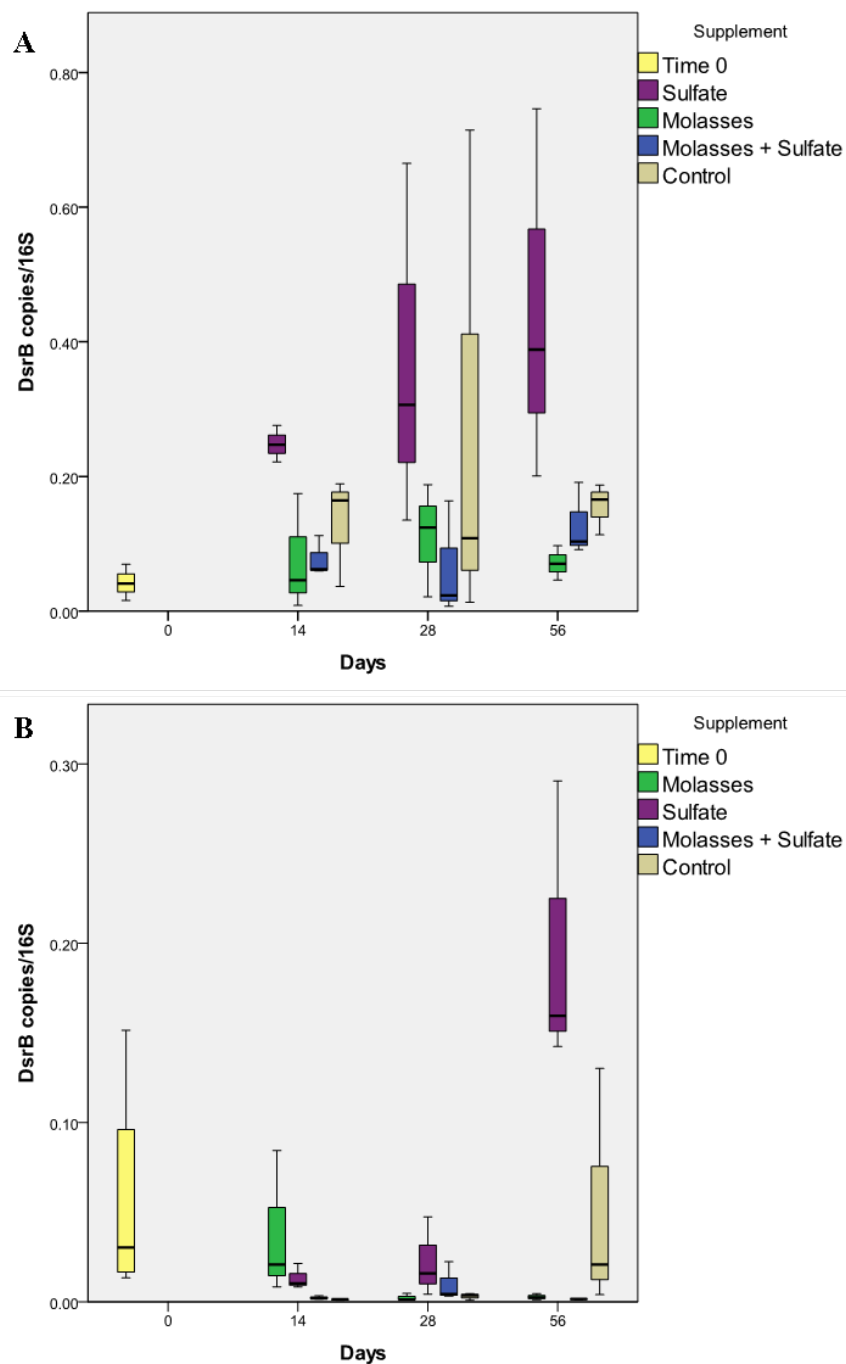


Figure 4-3. Ratios of *dsrB* to 16S rDNA copy number of 6.3 m site. Expressed in decimal form, shown by box and whisker plots calculated at 95% confidence intervals. Box shows 1 standard deviation from the mean (bar). Whiskers show standard deviation. Ratios at 19 feet (6.3 m) are shown in panel A, while 25 foot (8.3 meter) ratios are shown in panel B. Briefly, at 14 and 56 days at 19 feet (6.3 m), sulfate amended samples show an increased ratio when compared to microcosms from the same time point. At 25 foot (8.3 meter) depth, 56 day microcosms show significantly higher ratios of *DsrB* to 16S rDNA copy number.

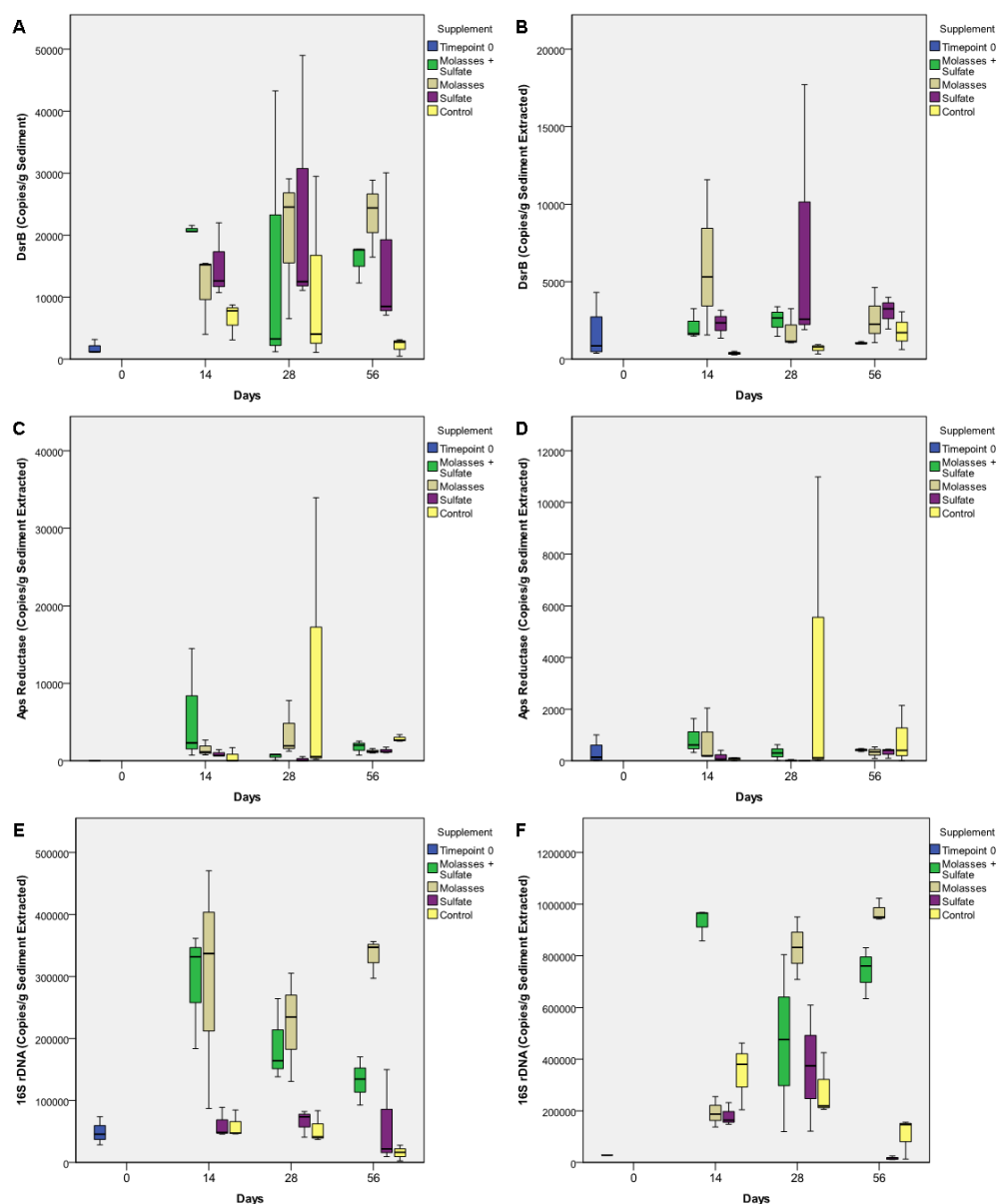


Figure 4-4. Box and whisker plots detailing average values for gene copies. Average copies/g sediment of *DsrB* (A,B), *Apr* (C,D) or 16S rDNA (E,F) at depths of 19 feet (6.3 m) (A,C, E) or 25 feet (8.3 meters) (B,D, F) respectively. shown by box and whisker plots calculated at 95% confidence intervals. Box shows 1 standard deviation from the mean (bar). Whiskers show standard deviation. Y-axis is log scale, X axis is time points of 0, 14, 28 and 56 days (not to scale). Documents show significantly higher levels of *DsrB* expression in all supplemented microcosms when compared to negative control at 56 days in the 19 ft (6.3 m) microcosm, and at 14 and 28 days in 25 ft (8.3 m) microcosm. *Apr* did not show significant differences at 19 feet (6.3 m). At 25 feet (8.3 meters), the trends do not mimic those of *DsrB*. Molasses supplementation results in consistently higher 16S rDNA copy number than negative control in all samples. Molasses plus sulfate amendment is higher at 14 and 28 days at 19 feet (6.3 m), and at 14 and 56 days at 25 feet (8.3 meters). There is a greater shift away from the 0 day levels of 16S rDNA copy number at 25 feet (8.3 meters) than at 19 feet (6.3 m).

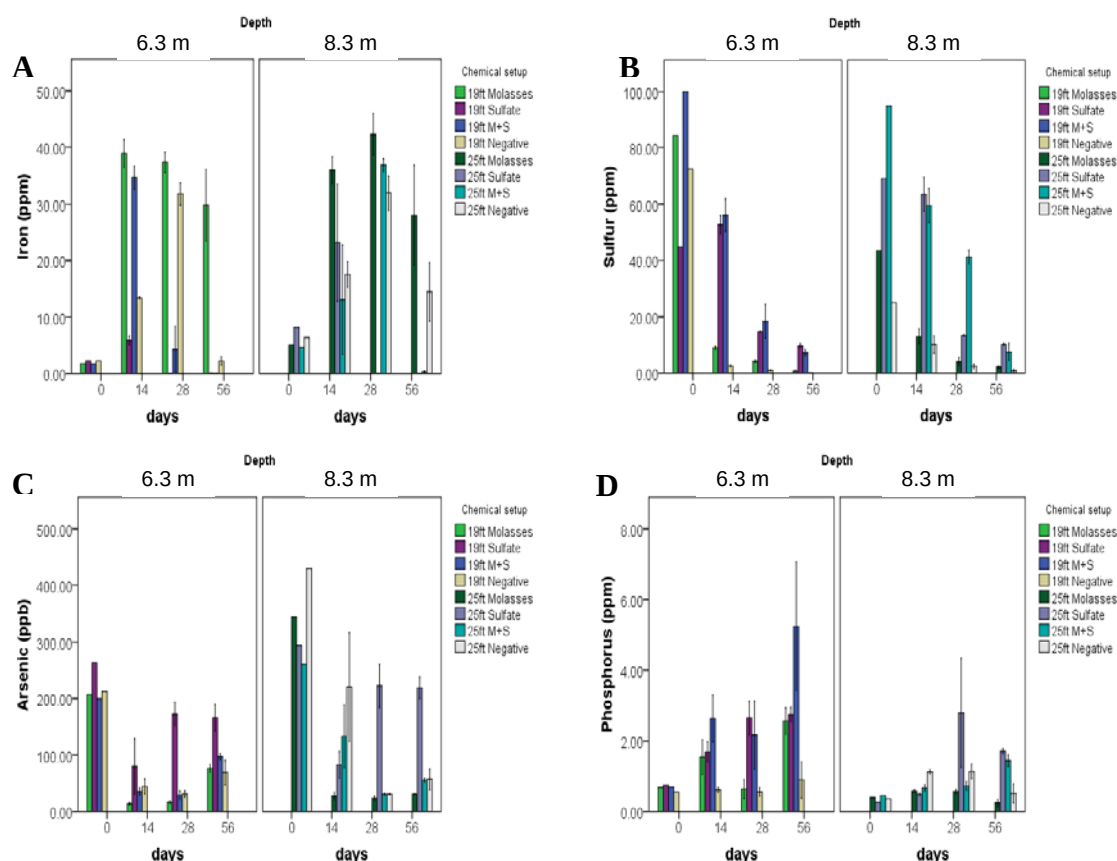


Figure 4-5. Bar plots of chemical data from all depths and supplement . Iron (A) shows an initial increase, followed by decreases over time in all samples at both depths. Shown by box and whisker plots calculated at 95% confidence intervals. Box shows 1 standard deviation from the mean (bar). Whiskers show standard deviation. Molasses supplemented microcosms show a significantly higher final concentration of iron than other amended samples, as does the negative control. Soluble Sulfur (B), used as a proxy for sulfate, shows continual decrease over time. There is a significantly higher level of sulfur at 56 days in samples amended with magnesium sulfate with or without molasses than in the molasses alone or negative control at both depths. Arsenic (C) Shows an initial decrease in soluble arsenic followed by increases in all amended microcosms. Magnesium sulfate amendment resulted in significantly higher arsenic concentrations at both depths. Phosphorus (D) shows increases at 19 feet (6.3 m) in amended samples, with the highest magnitude of change occurring in the dual addition microcosm by 56 days.

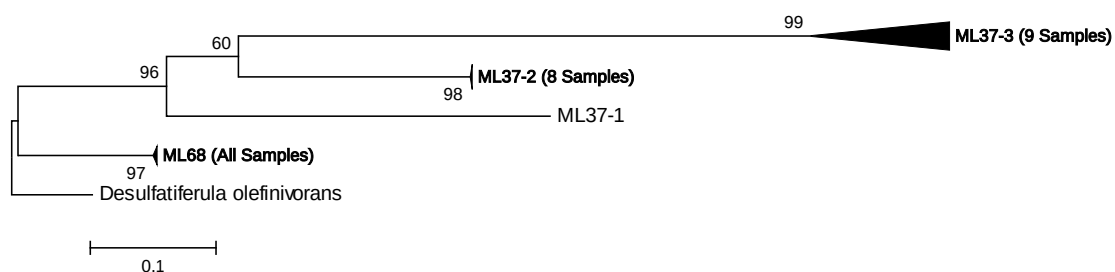


Figure 4-6. Bootstrapped phylogenetic dendrogram of *DsrAB* genes isolated from filter samples from 2 wells (37 and 68) in Bangladesh. Sample labels of N37 and N68 come from site ML-37 and ML-68 Respectively. Clustering of sequences appears to indicate similarity between samples from the same sites, but little similarity across sites. While ML68 is most closely related to *D. olefinivorans*, this distance is still large.

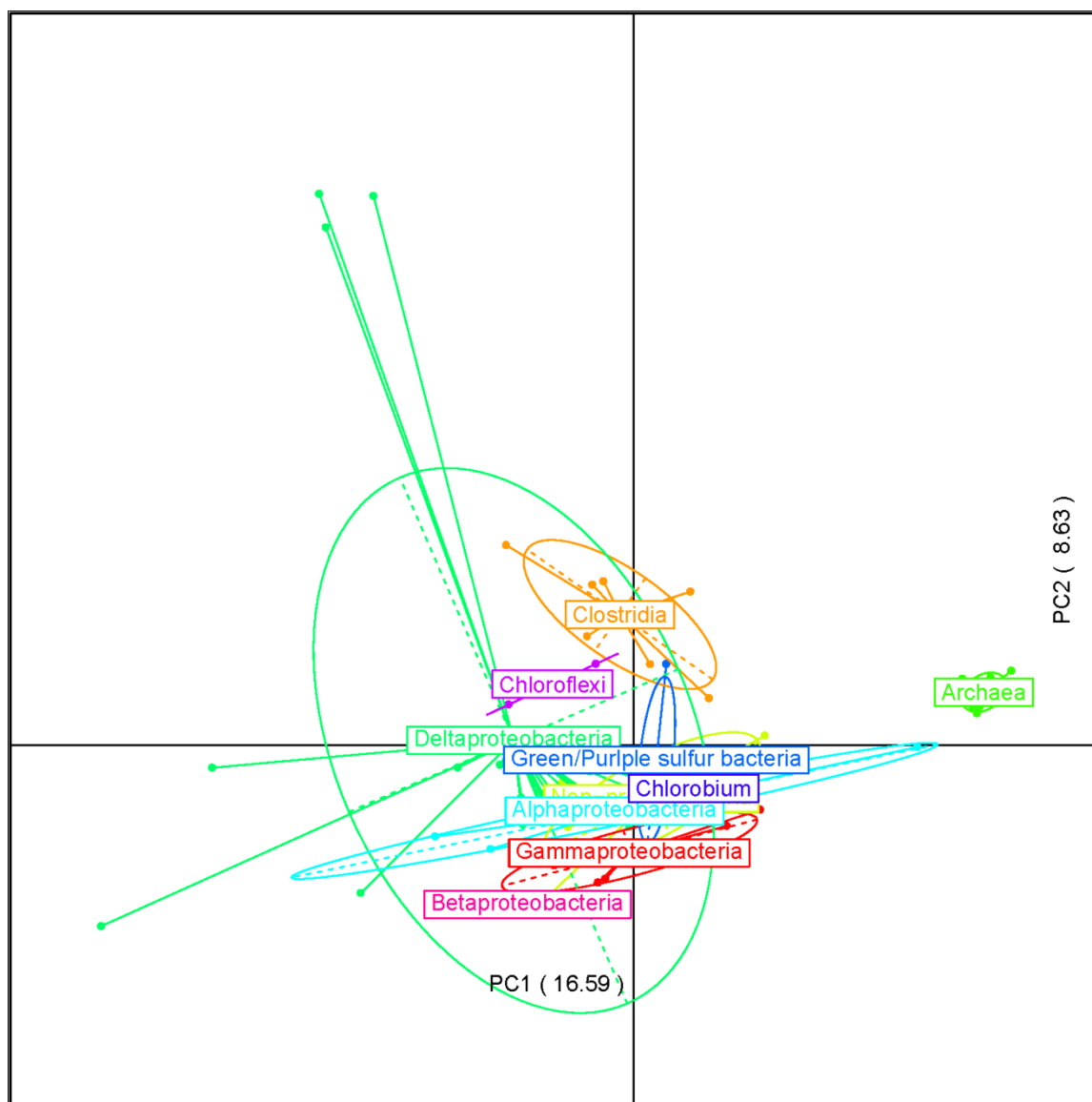


Figure 4-7. PCA plot generated from matrix of organisms and presence or absence of metabolic subsystems. Circles are calculated from groupings of the same phylogenetic classification. *Archaea* show distinct separation from other organisms, as do *Clostridia* species to a lesser degree.

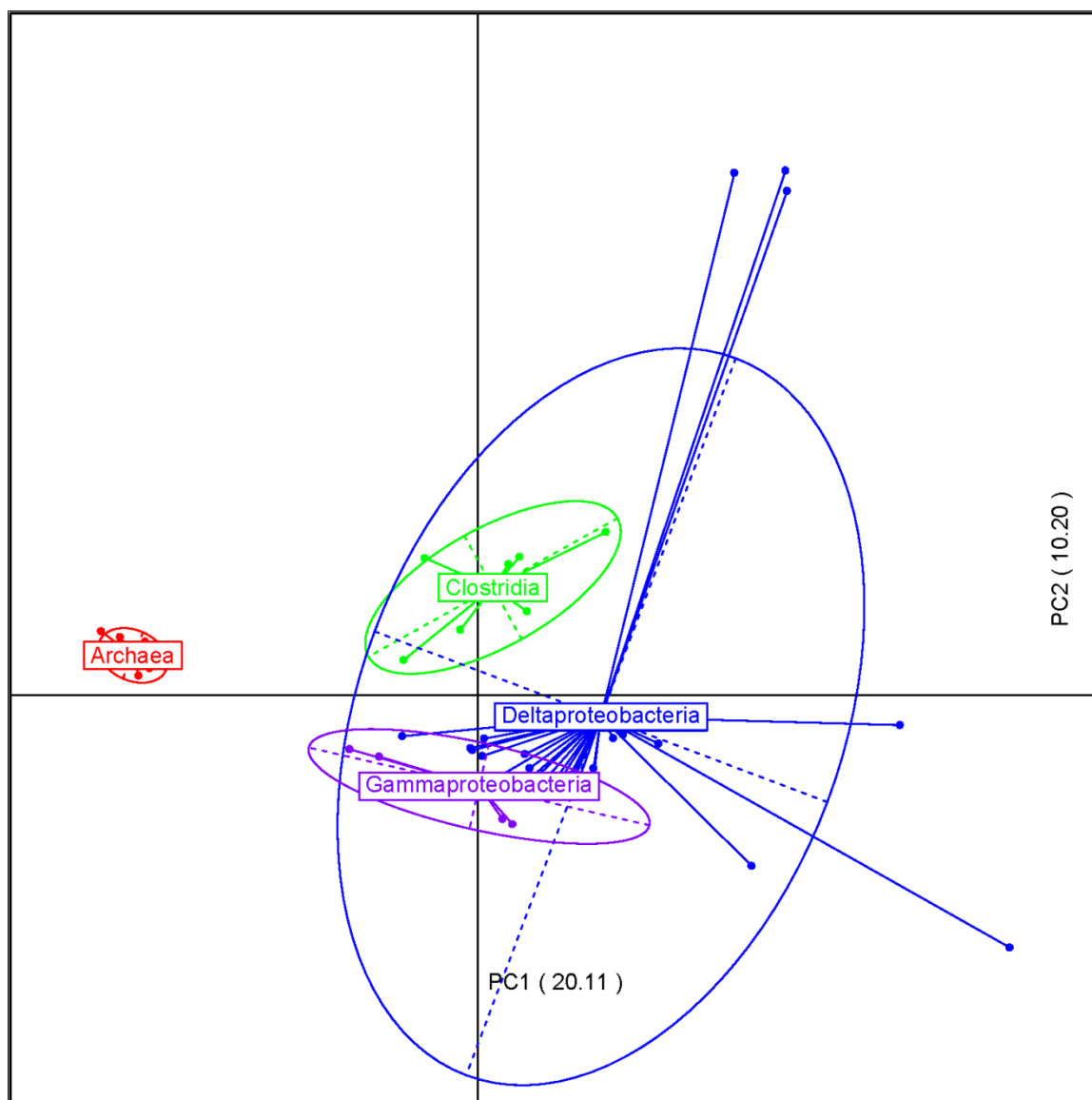


Figure 4-8. PCA plot similar to Figure 4-7, with overlapping groupings removed to illustrate that the most prominent phylogenetic groupings are distinct.

Discussion and Conclusions

While there are many factors that may affect the interplay of SRM and immobilization of metals and metalloids, microcosm experiments with oxygen contaminated soils raise the possibility that SRM may be encouraged to grow after oxygen exposure, as oxygenated water can rapidly dissolve sulfidic minerals, such as arsenopyrite, producing sulfate. The increased oxygen exposure would lead to higher levels of metals and metalloids in solution, and oxidize sulfides back to sulfite or sulfate, giving SRM an excellent set of starting materials to begin growth. If the interactions can be shown between oxygenation and sulfate reduction and can be verified, this would also potentially illustrate a strong niche and selective pressure for survival of SRM in environments with potential for exposure to oxygen.

Filter Analyses

dsrB was detectable at low levels (sub 100 copies/uL extracted sample) in all filtered water samples from Bangladesh. This indicates that *dsrB* is present at low levels across all sampling locations, supporting the theory of a wide global distribution of sulfate reducing microbes throughout anaerobic sediments. The ratio of sulfate reducing genes to 16S rDNA copies is below 1:100 in all measured cases, indicating that even in a highly sulfate reducing aquifer, the organisms still make up a small portion of the total community. The difference in detection between ML-68 (high *dsr* copy number, high arsenic) and ML-37 (low

dsr copy number, low arsenic), is pronounced, and indicates that levels of SRM are highly variable between sites. This data indicates that use of *dsrB* primers is a reliable detection technique for sulfate reducing capacity and activity in these systems, and that design of specialized primers is not required for this task. Application of these primers to future research can allow for rapid quantifications without isolation and sequencing of SRM strains.

The sequence diversity between two sites (ML-68 and ML-37) of both 16S rDNA and *dsrAB* appears to indicate that each of these sites are composed of distinct communities, with few overlapping species at the 16S rDNA or gene level. These factors combine to show a picture of a ubiquitous, low copy number organism with little direct effect on environmental conditions. There are a number of 16S rDNA sequences from the sequenced sites that are highly similar to *E.coli* indicating potential in these sites for transportation of fecal coliforms from surface ponds into sub-surface sediments (Akita, Leber et al. 2008; van Geen, Ahmed et al. 2011). Other 16S rDNA sequences were not matches to isolated organisms. 16S rDNA phylogeny also indicates that metal or metalloid concentrations, including arsenic, are not indicators of community structure. Finally, ML-37 also shows high abundance of pseudomonads (>30% of library), indicating that there is potentially a correlative link between elevated levels of arsenic, and decreased abundance of pseudomonads. Metal susceptibility is not expected to play a role as many pseudomonads are highly resistant to most metal stresses (Biro, Bayoumi et al. 1995). The relative

dominance of pseudomonads in the sample with the highest measured arsenic concentration may, in fact, be a result of elevated arsenic resistance but there is not sufficient power to determine linkages at this time.

Observation of the overall trends indicates that there are general statistical correlations between mineral factors and 16S rDNA gene copy, as these factors have a relationship when analyzed from a spearman ranks perspective. This ranking would indicate that while there is no linear relationship between these samples, higher metal concentration samples have lower overall 16S rDNA copy number. This ratio may be an indication either of water chemistry determining microbial community structure or of microbial growth limiting metal concentrations in aquifers. Alternatively, elevated microbial load in aquifers may drive immobilization of soluble metals as well as consumption of sulfur.

There is no indication of a statistically significant correlation between sulfate reducing microorganisms and any measured chemical factor. However, there is an inverse linkage between the overall 16S rDNA copy number per sample and concentrations of elemental minerals, particularly manganese, iron, and sulfur. The causes of this correlation have several potential explanations. First, it is possible that wash down of nutrients into this well site caused a bacterial bloom, leading to fermentative processes and anaerobic respiration. If this were the case, it would be expected that any oxidized minerals, particularly iron and manganese, would be easily reduced by microbial anaerobic respiration. Alternatively, oxygen entering this system could potentially oxidize sulfidic

mineral, leading to freeing of sulfate and ferric iron and mineralized manganese. Under these conditions, it would also be expected to see elevated arsenic concentrations. In either case, the increase in 16S rDNA copy number would be suspected to be from increased available electrons to encourage growth in the environment. This would then indicate that Mn, Fe, and S in solution are byproducts of rapid growth of microbes in the environment.

Alternatively, anaerobic Fe sequestration can typically be found to be in the form of pyrite due to the presence of reduced sulfides. These minerals may be dissolved during iron reduction freeing both iron and sulfides, warranting further investigation. The highly mobile state of the groundwater, coupled with the high level of uncertainty as to the sources of input for nutrients and chemicals, make *in situ* studies in Bangladesh less favorable than future microcosm work to describe the relationship between precipitated Mn, Fe, and S with microbial abundance.

This survey has indicated that sulfate reducing microbes are ubiquitous at low levels across a wide range of anaerobic aquifer sediments in Bangladesh. Neither SRM, nor 16s rDNA quantity, could be used as an indicator for chemical factors in the water. This may be a result of the mobile nature of aquifers (e.g. water movement from one location is altered by upstream elements), or may be attributable to sediment, rather than planktonic organisms. Phylogenetic analysis of *dsrAB* shows distinct lineages of SRM present at ML-37 and ML-68, potentially indicating either chemical or geographical selection for individual strains of SRM.

16S rDNA analysis at all 5 sites indicates that 16S rDNA absolute quantities are not directly correlated to any chemical factor in these samples. A shallow sequencing of 16S rDNA sequence shows high diversity at each site, but appears to indicate similarities between sites with similar water chemistry. Overall, this survey has indicated the potential for encouraging SRM growth in aquifers in Bangladesh, and the use of the ubiquitous distribution of *dsrAB* as an indicator of growth. There is no apparent natural linkage between these organisms at low concentrations and metal concentrations, or corrosion, as has been seen in water treatment and municipal water systems, where SRM are at higher concentrations. These factors all lead to the assumption that natural arsenic concentration is not affected by SRM presence/absence or quantity, either as a proportion of the community or as an absolute copy number measurement.

There is not enough statistical power to indicate strong linkages between chemical factors and associated microbes. It has illustrated the applicability of qPCR protocols coupled with normalization procedures included in this chapter. These techniques can be coupled with elemental analysis to generate statistical analyses of gene/element linkages, resulting in potential models for nutrient and elemental cycling in samples of interest. Additionally, 16S rDNA phylogeny of these samples indicates that geographic proximity is not a strong indicator of phylogenetic distribution. Additionally, the strong presence of pseudomonads in a single sample contaminated with high levels of arsenic indicates that this family

may be correlatively linked to arsenic concentrations. Further study would be necessary to determine if this similarity was due to geographical transportation, niche formations, or random chance. This study has also found at least two novel *dsrAB* genes according to the current repository at NCBI.

Microcosm Analysis

While the construction of microcosms was fatally contaminated with oxygen, there are two important factors to note. First, addition of sucrose was indeed sufficient to encourage microbial growth in sufficient numbers to restore anaerobic conditions. This is unsurprising, but the formation of sulfides (noted by odor, concentration differences, and discoloration of sample) indicates that SRM in natural environments can survive non-minimal exposure to oxygen, or are protected in sediment samples from exposure to oxygen. Secondly, this study also conclusively illustrated the removal of sulfate from solution concomitant with the addition of sulfate and sucrose, indicating that sucrose is necessary for sulfate consumption, and suggesting that this process is due to SRM, not assimilatory sulfate reduction.

Elemental Analysis of Microcosms

Sulfate is reduced in all samples, indicating the presence of sulfate reducing microorganisms as early as 14 days after incubation. Concurrent with the consumption of sulfate, detected levels of iron in solution decrease, indicating

that sulfide production is resulting in precipitation of iron as iron sulfides. While elemental analysis is not conclusive, there are only two plausible mechanisms for sulfate to be removed from the aqueous phase, either via incorporation into biomass (assimilatory sulfate reduction), or reduction to sulfide and subsequent precipitation. While there are indications of increase in biomass, the concentration drops in sulfur is more consistent with sulfidic mineral formation than sulfur incorporation into biomass. Additionally, this, combined with the fluctuations in arsenic, can be explained as a byproduct of sulfide being oxidized by reduced arsenic products chemically, resulting in an increase in soluble arsenic.

Iron appears to increase from 0 to 14 days, and subsequently decrease. This, combined with increased levels of SRM and consumption of soluble sulfate, would suggest that iron is freed during the first 14 days, and that, after 14 days, iron reacts with generated sulfides, or other chemicals, and forms insoluble minerals. This is supported by observation of dark precipitate formation in the samples (data not shown). Additionally, nutrient supplementation would be expected to result in an increase in iron reducing bacteria under anaerobic conditions, leading to reduction of any oxidized, insoluble iron. This would explain increases over the short term in soluble iron concentrations.

Arsenic levels decrease over the first 14 days in all samples, indicating arsenic desolubilization. This effect appears to be transiently linked to consumption of sulfate, with an abundance of sulfate consumption triggering a

return to solubility of arsenic. In non-amended and molasses-only supplemented microcosms, arsenic is maintained in a non soluble form, whereas sulfate alone and molasses-sulfate dual supplementation microcosms showed an increase over time in arsenic concentrations after the initial drop. Samples amended with sulfate alone result in arsenic levels at or near the initial spiked concentration of soluble arsenic (150 ppb), indicating that this treatment increases arsenic solubility when compared to a negative control or molasses additions. Sulfate consumption alone does not correlate strongly to this result. These findings are most likely due to the known effect of high concentrations of sulfides to increase the solubility of arsenic from minerals (Barton and Fauque 2009). This result is not out of the range of expected results; however, as a microcosm system, it will allow a buildup of higher concentrations of sulfides due to the static nature of the system. The reaction of sulfide with arsenite to re-oxidize sulfide to elemental sulfur or sulfite/sulfate has been described (Rochette, Bostick et al. 2000), and is a likely explanation for this reaction.

Phosphate levels do appear to be linked to sulfate additions, with phosphate levels elevated in samples amended with sulfate when compared to negative controls or molasses amended samples. This level was lower in all 25-foot samples than in 19-foot samples, but was elevated when compared to negative controls at the same depth. Phosphorous is of interest as another inorganic mineral with fluctuating levels over time at this site (Dhar, Zheng et al. 2008).

Variation in Microcosms

While there is variability inherent to microcosms, there is some evidence to support the hypothesis that molasses can serve as a carbon source along with inorganic sulfate addition to increase the population of sulfate reducing microorganisms in a community, with no other additions necessary. This work also indicates a high sample variation in microcosms in regards to molecular quantification. This intra-sample variation makes prediction and analysis of individual trends difficult. To appropriately control for this variation, either large scale mesocosms with continuous sampling or large numbers ($N > 5$) of individual microcosms are suggested to reduce variability. Mesocosms have the advantage of averaging behavior across a larger sample size. Because the samples are technically mixed, all replicates are technical replicates, meaning that chemical measurements would have little power. Additionally, this minimizes waste concerns, as a single waste stream is generated. Increasing the number of replicates poses the difficulty of gathering more sample, which is technically difficult in Bangladesh, but possible. While it is expected that variation would still be large, more replicates would allow better correlations to be drawn.

Finally, there would be great value in performing these microcosm experiments in a continuous flow, chemostat environment, with a flow rate similar to that of the aquifers in Bangladesh. However, the waste stream this would generate is large. To be exact, if the assumed aquifer refresh rate is 2 hours,

and the volume in the microcosm were 25 mL with 3 test conditions and 1 control, coupled with 5 replicates each, 1L of waste would be generated every 4 hours. Couple this calculation with a 48-day experimental set-up, and you generate approximately 200L of arsenic contaminated waste by the end of the experiment.

Population Shifts Due to Amendment in Microcosms

This investigation utilizes previously validated methods to investigate population shifts of sulfate reducing microorganisms. While there are recent investigations into quantification of SRM using *dsrB* and quantitative PCR, these investigations focused on sampling natural settings. These data indicate that sulfate is necessary for increases in SRM population, but that this sulfate can be garnered from alternative sources, such as molasses. This is in line with previous suggestions that sulfate addition is the sole requirement for encouragement of SRM *in situ* (Sheoran, Sheoran et al. 2010).

Additionally, this investigation appears to indicate that geographical location plays a major part in the lag and magnitude of an increase in SRM. Previous work at this site, indicating seasonal changes in sulfate level at the 25-foot depth, indicates that sediment gathered from the 25-foot microcosm had been exposed to sulfate. This exposure was not sufficient to allow for a rapid and robust response to the nutrient additions in this case.

The population shifts detected in sulfate-amended microcosms towards SRM indicate that simple inorganic sulfate additions are sufficient to change the dynamics of the community, and encourage sulfate reduction.

dsrB copy number was consistently higher than *Apr* copies. There are several possible explanations for this. While it is assumed that *Apr* is conserved across species, it is possible that there are alternative mechanisms for the conversion of sulfate to sulfite either biologically via another mechanism of sulfate reduction. Alternatively, there may be enough reducing power in many of these environments to chemically reduce sulfate to sulfite. It is also possible that the primers designed to amplify the APS gene are not degenerate enough in design to detect any sulfate reducing species present in this sample. If *Apr* is necessary for SRM survival, primers will need to be developed to match a wider range of SRM, enabling use of two genetic markers to indicate presence/absence and relative level of SRM presence/activity *in situ*.

Directions for Future Research

These experiments indicate that sulfate reducing microorganisms can be encouraged to grow in anaerobic sediments with the addition of sulfate, sulfate bearing nutrient such as molasses, or both, with little difference in total SRM populations. Additionally, this shift appears to occur within 14 days of initial treatments, with little change in overall concentrations over 6 weeks of incubation. Sulfate reducing microbes quantities do not appear linked to total biomass of the samples, leading to the potential that SRM activity is self-limiting,

due to exhaustion of terminal electron acceptors, or presence of sulfides in solution. This leads to a potential investigation into the community shifts during this 14 week time frame. Additional work to determine the rate of SRM community growth may indicate the expected doubling time of this organism when taken as part of the community.

Additionally, this work supports the use of biogenic sulfides to precipitate metals from soluble phases. In this case iron was precipitated into insoluble form in microcosms supplemented with sulfate, from a level of 30 ppm to nearly 0. Arsenic desolubilization appears to occur during the early stages of sulfate reducing activity, but the process reverses over time, most likely due to chemical buildup of sulfides over this time frame.

The microbial organisms responsible for these activities appear to have reached a plateau by 14 days. This, combined with the first visible changes in sediment color and structure, indicate that sulfate reducing bacteria do not produce sufficient sulfide to react with metals until they have reached their population plateau. Additional studies are needed to determine the molecular and chemical changes during the first 14 days of supplementation.

Analysis of Annotated Genomes

Use of annotations from NCBI deposited genomes to select SRM has been shown to be inaccurate, with both false positive and false negatives, resulting in a poor selection of appropriate genomes for analysis. Additionally,

the computational requirements to utilize *de novo* calculations of orthologous proteins from a list of >100 genomes has proven to be infeasible at the current time. Due to the ubiquity and taxonomic diversity of SRM, analysis of subsets of genomes (e.g. all Clostridial genomes) was shown to be impractical, resulting in few orthologues between genomes, and less annotated differences than desired. Use of SEED annotated genomes was able to identify 60 annotated SRM from a broad range of sources, and taxonomic groups. Analysis of these annotations and subsystem classifications was able to correctly identify all expected genomes of SRM in the candidate pool. Additionally, it was shown that these genomes have conserved functions for citric acid cycles, pyruvate metabolism, and glycolipid/glycophospholipid metabolisms, consistent with demonstrated SRM activity. Further analysis of shared features between genomes both as a whole, and within classifications by taxonomy and habitat show that several conserved pathways exist for particular groupings, including heat shock proteins for thermophilic organisms, cold shock proteins for non-thermophilic green or purple sulfur bacteria, and branched chain amino acid uptake for *deltaproteobacteria*.

Metabolic pathways for SRM are highly conserved for secondary metabolites such as lactate, pyruvate, as well as glycolipids and phosphoglycolipids. There is additional evidence that a majority (>30) are also capable of performing some form of glycolysis, indicating that some SRM may be able to respond directly to addition of complex carbon. However, only a select

few have annotated response mechanisms for sucrose, the main component of the nutrient of choice in this work, molasses.

The prevalence of carbon starvation subsystems in SRM also lends a degree of understanding to the results of supplementation, indicating that sulfur addition in any form (including that present in molasses) is sufficient to trigger growth of SRM. This lends a potential explanation to the evidence that sulfur additions in microcosms can trigger SRM growth, and generate conditions (elevated sulfide concentrations) to precipitate metals and metalloids.

List of References

- Achour, A. R., P. Bauda, et al. (2007). "Diversity of arsenite transporter genes from arsenic-resistant soil bacteria." Research in Microbiology **158**(2): 128-137.
- Agrawal, A. and B. Lal (2009). "Rapid detection and quantification of bisulfite reductase genes in oil field samples using real-time PCR." Fems Microbiology Ecology **69**(2): 301-312.
- Ahmed, K. M., P. Bhattacharya, et al. (2004). "Arsenic enrichment in groundwater of the alluvial aquifers in Bangladesh: an overview." Applied Geochemistry **19**(2): 181-200.
- Akita, Y., J. Leber, et al. (2008). "Spatial Analysis of the Distribution of Bacterial Pathogen Indicators Across Shallow Aquifers in Bangladesh." Epidemiology **19**(6): S205-S205.
- Aziz, R. K., D. Bartels, et al. (2008). "The RAST Server: rapid annotations using subsystems technology." BMC Genomics **9**: 75.
- Baqui, A. H., M. D. Yunus, et al. (1991). "Surveillance Of Patients Attending A Rural Diarrhea Treatment Center In Bangladesh." Tropical and Geographical Medicine **43**(1-2): 17-22.
- Barton, L. and W. A. Hamilton (2007). Sulphate-reducing bacteria : environmental and engineered systems. Cambridge ; New York, Cambridge University Press.
- Barton, L. L. and G. D. Fauque (2009). Biochemistry, Physiology and Biotechnology of Sulfate-Reducing Bacteria. Advances in Applied Microbiology, Vol 68. San Diego, Elsevier Academic Press Inc. **68**: 41-98.
- Ben-Dov, E., A. Brenner, et al. (2007). "Quantification of sulfate-reducing bacteria in industrial wastewater, by real-time polymerase chain reaction (PCR) using *dsrA* and *apsA* genes." Microbial Ecology **54**(3): 439-451.
- Biro, B., H. Bayoumi, et al. (1995). "Metal Sensitivity Of Some Symbiotic N-2-Fixing Bacteria And Pseudomonas Strains." Acta Biologica Hungarica **46**(1): 9-16.
- Bombach, P., H. H. Richnow, et al. (2010). "Current approaches for the assessment of in situ biodegradation." Applied Microbiology and Biotechnology **86**(3): 839-852.
- Boothman, C., S. Hockin, et al. (2006). "Molecular analysis of a sulphate-reducing consortium used to treat metal-containing effluents." Biometals **19**(6): 601-609.
- Boucher, Y., C. J. Douady, et al. (2003). "Lateral gene transfer and the origins of prokaryotic groups." Annual Review of Genetics **37**: 283-328.
- Brioukhanov, A. L. and A. I. Netrusov (2007). "Aerotolerance of strictly anaerobic microorganisms and factors of defense against oxidative stress: A review." Applied Biochemistry and Microbiology **43**(6): 567-582.
- Busse, H. J., E. B. M. Denner, et al. (1996). "Classification and identification of bacteria: Current approaches to an old problem. Overview of methods used in bacterial systematics." Journal of Biotechnology **47**(1): 3-38.

- Button, M., G. R. T. Jenkin, et al. (2009). "Human toenails as a biomarker of exposure to elevated environmental arsenic." Journal of Environmental Monitoring **11**(3): 610-617.
- Caldwell, B. K., J. C. Caldwell, et al. (2003). "Searching for an optimum solution to the Bangladesh arsenic crisis." Social Science & Medicine **56**(10): 2089-2096.
- Carlin, A., W. Shi, et al. (1995). "The ars operon of Escherichia coli confers arsenical and antimonial resistance." J. Bacteriol. **177**(4): 981-986.
- Christophersen, C. T., M. Morrison, et al. (2011). "Overestimation of the Abundance of Sulfate-Reducing Bacteria in Human Feces by Quantitative PCR Targeting the Desulfovibrio 16S rRNA Gene." Applied and Environmental Microbiology **77**(10): 3544-3546.
- DeBruyn, J. M. and G. S. Sayler (2009). "Microbial community structure and biodegradation activity of particle-associated bacteria in a coal tar contaminated creek." Environmental Science & Technology **43**(9): 3047-3053.
- Dercole, F., R. Ferriere, et al. (2010). "Chaotic Red Queen coevolution in three-species food chains." Proceedings of the Royal Society B-Biological Sciences **277**(1692): 2321-2330.
- Detmers, J., V. Bruchert, et al. (2001). "Diversity of sulfur isotope fractionations by sulfate-reducing prokaryotes." Applied and Environmental Microbiology **67**(2): 888-894.
- Dhar, R. K., Y. Zheng, et al. (2008). "Temporal variability of groundwater chemistry in shallow and deep aquifers of Araihasar, Bangladesh." Journal of Contaminant Hydrology **99**(1-4): 97-111.
- Feinstein, L. M., W. J. Sul, et al. (2009). "Assessment of Bias Associated with Incomplete Extraction of Microbial DNA from Soil." Applied and Environmental Microbiology **75**(16): 5428-5433.
- Ford, T., J. Jay, et al. (2005). "Use of ecotoxicological tools to evaluate the health of New Bedford Harbor sediments: A microbial biomarker approach." Environmental Health Perspectives **113**(2): 186-191.
- Friedrich, M. W. (2002). "Phylogenetic analysis reveals multiple lateral transfers of adenosine-5'-phosphosulfate reductase genes among sulfate-reducing microorganisms." Journal of Bacteriology **184**(1): 278-289.
- Fulton, D. L., Y. Y. Li, et al. (2006). "Improving the specificity of high-throughput ortholog prediction." BMC Bioinformatics **7**: 270.
- Geets, J., B. Borremans, et al. (2006). "DsrB gene-based DGGE for community and diversity surveys of sulfate-reducing bacteria." Journal of Microbiological Methods **66**(2): 194-205.
- Geets, J., B. Borremans, et al. (2005). "Molecular monitoring of SRB community structure and dynamics in batch experiments to examine the applicability of in situ precipitation of heavy metals for groundwater remediation." Journal of Soils and Sediments **5**(3): 149-163.

- Geets, J., B. Borrernans, et al. (2006). "DsrB gene-based DGGE for community and diversity surveys of sulfate-reducing bacteria." Journal of Microbiological Methods **66**(2): 194-205.
- Ghosh, W. and B. Dam (2009). "Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea." Fems Microbiology Reviews **33**(6): 999-1043.
- Gregoire, M. and K. Soetaert (2010). "Carbon, nitrogen, oxygen and sulfide budgets in the Black Sea: A biogeochemical model of the whole water column coupling the oxic and anoxic parts." Ecological Modelling **221**(19): 2287-2301.
- Harada, M., T. Yoshida, et al. (2009). "Expression of genes for sulfur oxidation in the intracellular chemoautotrophic symbiont of the deep-sea bivalve *Calyptogena okutanii*." Extremophiles **13**(6): 895-903.
- He, K. (2011). "Trace elements in nails as biomarkers in clinical research." European Journal of Clinical Investigation **41**(1): 98-102.
- He, Q. and J. Z. Zhou Bioremediation of heavy metals in soil and groundwater: impact of nitrate as an inhibitor. Geoenvironmental engineering and geotechnics: progress in modeling and applications. Proceedings of sessions of GeoShanghai 2010, Shanghai, China, 3-5 June, 2010., American Society of Civil Engineers (ASCE).
- Hiibel, S. R., L. P. Pereyra, et al. (2008). "Microbial community analysis of two field-scale sulfate-reducing bioreactors treating mine drainage." Environmental Microbiology **10**(8): 2087-2097.
- Illman, W. A. and P. J. Alvarez (2009). "Performance Assessment of Bioremediation and Natural Attenuation." Critical Reviews in Environmental Science and Technology **39**(4): 209-270.
- Ishak, H. D., R. Plowes, et al. (2011). "Bacterial Diversity in *Solenopsis invicta* and *Solenopsis geminata* Ant Colonies Characterized by 16S amplicon 454 Pyrosequencing." Microbial Ecology **61**(4): 821-831.
- Islam, F. S., A. G. Gault, et al. (2004). "Role of metal-reducing bacteria in arsenic release from Bengal delta sediments." Nature **430**(6995): 68-71.
- Jaenicke, S., C. Ander, et al. (2011). "Comparative and Joint Analysis of Two Metagenomic Datasets from a Biogas Fermenter Obtained by 454-Pyrosequencing." Plos One **6**(1).
- Johnston, D. T. (2011). "Multiple sulfur isotopes and the evolution of Earth's surface sulfur cycle." Earth-Science Reviews **106**(1-2): 161-183.
- Jomova, K., Z. Jenisova, et al. (2011). "Arsenic: toxicity, oxidative stress and human disease." Journal of Applied Toxicology **31**(2): 95-107.
- Keijser, B. J. F., E. Zaura, et al. (2008). "Pyrosequencing analysis of the Oral Microflora of healthy adults." Journal of Dental Research **87**(11): 1016-1020.
- Khibnik, A. I. and A. S. Kondrashov (1997). "Three mechanisms of Red Queen dynamics." Proceedings of the Royal Society of London Series B-Biological Sciences **264**(1384): 1049-1056.

- Kjeldsen, K. U., L. Tang, et al. (2009). "Enumeration and identification of dominant types of sulfate-reducing bacteria in pulp from a paper-recycling plant: a multiphasic approach." Fems Microbiology Ecology **69**(3): 481-493.
- Kondo, R., K. Shigematsu, et al. (2008). "Rapid enumeration of sulphate-reducing bacteria from aquatic environments using real-time PCR." Plankton & Benthos Research **3**(3): 180-183.
- Koonin, E. V. and Y. I. Wolf (2008). "Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world." Nucleic Acids Research **36**(21): 6688-6719.
- Kumar, R., S. Singh, et al. (2007). "Bioremediation of radionuclides: Emerging technologies." Omics-a Journal of Integrative Biology **11**(3): 295-304.
- Kumar, S., M. Nei, et al. (2008). "MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences." Briefings in Bioinformatics **9**(4): 299-306.
- Lee, M. K., J. A. Saunders, et al. (2005). "Geochemical modeling of arsenic speciation and mobilization: Implications for bioremediation." Advances in Arsenic Research **915**: 398-413.
- Little, B. J., J. S. Lee, et al. (2006). "Diagnosing microbiologically influenced corrosion: A state-of-the-art review." Corrosion **62**(11): 1006-1017.
- Loy, A., S. Duller, et al. (2009). "Reverse dissimilatory sulfite reductase as phylogenetic marker for a subgroup of sulfur-oxidizing prokaryotes." Environmental Microbiology **11**(2): 289-299.
- Marietou, A., D. Richardson, et al. (2005). "Nitrate reduction by *Desulfovibrio desulfuricans*: A periplasmic nitrate reductase system that lacks NapB, but includes a unique tetraheme c-type cytochrome, NapM." Fems Microbiology Letters **248**(2): 217-225.
- McArthur, J. M. R., P.; Safiullah, S.; Thirwall, M.F. (2011). "Arsenic in groundwater: testing pollution mechanisms for sedimentary aquifers in Bangladesh." Water Resources Research **37**(1): 8.
- Michael, H. A. and C. I. Voss (2009). "Estimation of regional-scale groundwater flow properties in the Bengal Basin of India and Bangladesh." Hydrogeology Journal **17**(6): 1329-1346.
- Miletto, M., P. L. E. Bodelier, et al. (2007). "Improved PCR-DGGE for high resolution diversity screening of complex sulfate-reducing prokaryotic communities in soils and sediments." Journal of Microbiological Methods **70**(1): 103-111.
- Miletto, M., A. Loy, et al. (2008). "Biogeography of sulfate-reducing prokaryotes in river floodplains." Fems Microbiology Ecology **64**(3): 395-406.
- Mitamura, M., H. Masuda, et al. (2008). "Geological structure of an arsenic-contaminated aquifer at Sonargaon, Bangladesh." Journal of Geology **116**(3): 288-302.
- Moon, H. S., L. McGuinness, et al. (2010). "Microbial reduction of uranium under iron- and sulfate-reducing conditions: Effect of amended goethite on

- microbial community composition and dynamics." Water Research **44**(14): 4015-4028.
- Muhling, M., J. Woolven-Allen, et al. (2008). "Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities." Isme Journal **2**(4): 379-392.
- Mukhopadhyay, R., B. P. Rosen, et al. (2002). "Microbial arsenic: from geocycles to genes and enzymes." Fems Microbiology Reviews **26**(3): 311-325.
- Muller, N., P. Worm, et al. (2010). "Syntrophic butyrate and propionate oxidation processes: from genomes to reaction mechanisms." Environmental Microbiology Reports **2**(4): 489-499.
- Nakagawa, T., S. Nakagawa, et al. (2004). "Phylogenetic diversity of sulfate-reducing prokaryotes in active deep-sea hydrothermal vent chimney structures." Fems Microbiology Letters **232**(2): 145-152.
- Oremland, R. S. and J. F. Stolz (2005). "Arsenic, microbes and contaminated aquifers." Trends in Microbiology **13**(2): 45-49.
- Pichler, T., J. Veizer, et al. (1999). "Natural input of arsenic into a coral reef ecosystem by hydrothermal fluids and its removal by Fe(III) oxyhydroxides." Environmental Science & Technology **33**(9): 1373-1378.
- Polymenakou, P. N., N. Lampadariou, et al. (2009). "Phylogenetic diversity of sediment bacteria from the southern Cretan margin, Eastern Mediterranean Sea." Systematic and Applied Microbiology **32**(1): 17-26.
- Richardson, D. J. (2000). "Bacterial respiration: a flexible process for a changing environment." Microbiology-Sgm **146**: 551-571.
- Rivera-Nunez, Z., J. R. Meliker, et al. (2010). "Reliability of spot urine samples in assessing arsenic exposure." International Journal of Hygiene and Environmental Health **213**(4): 259-264.
- Rochette, E. A., B. C. Bostick, et al. (2000). "Kinetics of arsenate reduction by dissolved sulfide." Environmental Science & Technology **34**(22): 4714-4720.
- Rosen, B. P. (1999). "Families of arsenic transporters." Trends in Microbiology **7**(5): 207-212.
- Sambrook, J. (2001). Molecular cloning : a laboratory manual / Joseph Sambrook, David W. Russell. Cold Spring Harbor, N.Y. :, Cold Spring Harbor Laboratory.
- Saunders, J. A., M.-K. Lee, et al. (2005). "Geochemical, microbiological, and geophysical assessments of anaerobic immobilization of heavy metals." Bioremediation Journal **9**(1): 33-48.
- Saunders, J. A., M. K. Lee, et al. (2008). "Geochemistry and mineralogy of arsenic in (natural) anaerobic groundwaters." Applied Geochemistry **23**(11): 3205-3214.
- Saunders, J. A., M. A. Pritchett, et al. (1997). "Geochemistry of biogenic pyrite and ferromanganese coatings from a small watershed: A bacterial connection?" Geomicrobiology Journal **14**(3): 203-217.

- Schink, B. (1997). "Energetics of syntrophic cooperation in methanogenic degradation." Microbiology and Molecular Biology Reviews **61**(2): 262-&.
- Sheoran, A. S., V. Sheoran, et al. (2010). "Bioremediation of acid-rock drainage by sulphate-reducing prokaryotes: A review." Minerals Engineering **23**(14): 1073-1100.
- Shi, L., T. C. Squier, et al. (2007). "Respiration of metal (hydr)oxides by Shewanella and Geobacter: a key role for multihaem c-type cytochromes." Molecular Microbiology **65**(1): 12-20.
- Smith, A. C., J. E. Kostka, et al. (2004). "Seasonal composition and activity of sulfate-reducing prokaryotic communities in seagrass bed sediments." Aquatic Microbial Ecology **37**(2): 183-195.
- Sogin, M. L., H. G. Morrison, et al. (2006). "Microbial diversity in the deep sea and the underexplored "rare biosphere"." Proceedings of the National Academy of Sciences of the United States of America **103**(32): 12115-12120.
- Thomas, D. J., M. Styblo, et al. (2001). "The cellular metabolism and systemic toxicity of arsenic." Toxicology and Applied Pharmacology **176**(2): 127-144.
- Tsai, S. L., S. Singh, et al. (2009). "Arsenic metabolism by microbes in nature and the impact on arsenic remediation." Current Opinion in Biotechnology **20**(6): 659-667.
- Uddin, A., M. Shamsudduha, et al. (2011). "Mineralogical profiling of alluvial sediments from arsenic-affected Ganges-Brahmaputra floodplain in central Bangladesh." Applied Geochemistry **26**(4): 470-483.
- van Geen, A., K. M. Ahmed, et al. (2011). "Fecal Contamination of Shallow Tubewells in Bangladesh Inversely Related to Arsenic." Environmental Science & Technology **45**(4): 1199-1205.
- Vandamme, P., B. Pot, et al. (1996). "Polyphasic taxonomy, a consensus approach to bacterial systematics." Microbiological Reviews **60**(2): 407-+.
- Ventura-Lima, J., M. R. Bogo, et al. (2011). "Arsenic toxicity in mammals and aquatic animals: A comparative biochemical approach." Ecotoxicology and Environmental Safety **74**(3): 211-218.
- Vladar, P., A. Rusznyak, et al. (2008). "Diversity of sulfate-reducing bacteria inhabiting the rhizosphere of Phragmites australis in lake velencei (Hungary) revealed by a combined cultivation-based and molecular approach." Microbial Ecology **56**(1): 64-75.
- Wagner, M., A. Loy, et al. (2005). Functional marker genes for identification of sulfate-reducing prokaryotes. Environmental Microbiology. San Diego, Elsevier Academic Press Inc. **397**: 469-489.
- Wagner, M., A. Loy, et al. (2005). "Functional marker genes for identification of sulfate-reducing prokaryotes." Environmental Microbiology **397**: 469-489.
- Wagner, M., A. J. Roger, et al. (1998). "Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration." Journal of Bacteriology **180**(11): 2975-2982.

- Ward, D. M., F. M. Cohan, et al. (2008). "Genomics, environmental genomics and the issue of microbial species." Heredity **100**(2): 207-219.
- Weber, K. A., L. A. Achenbach, et al. (2006). "Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction." Nature Reviews Microbiology **4**(10): 752-764.
- Weinman, B., S. L. Goodbred, et al. (2008). "Contributions of floodplain stratigraphy and evolution to the spatial patterns of groundwater arsenic in Araihasar, Bangladesh." Geological Society of America Bulletin **120**(11-12): 1567-1580.
- Williams, K. H., D. Ntarlagiannis, et al. (2005). "Geophysical imaging of stimulated microbial biomineralization." Environ Sci Technol **39**(19): 7592-7600.
- Wirsén, C. O., S. M. Sievert, et al. (2002). "Characterization of an autotrophic sulfide-oxidizing marine *Arcobacter* sp that produces filamentous sulfur." Applied and Environmental Microbiology **68**(1): 316-325.
- Xu, C., T. Q. Zhou, et al. (1998). "Metalloid resistance mechanisms in prokaryotes." Journal of Biochemistry **123**(1): 16-23.
- Zhang, T. and H. Fang (2006). "Applications of real-time polymerase chain reaction for quantification of microorganisms in environmental samples." Applied Microbiology and Biotechnology **70**(3): 281-289.

Appendix

Table A-1. Organisms with an annotated copy of *dsrA*, *dsrB* or *dsrC* as predicted by SEED These organisms largely do not contain an annotated copy of *dsrMJKOP*, and these annotations of *dsr* are generally not included in the dissimilatory sulfate reduction subsystem in SEED

SEED Annotated <i>dsrA</i>	SEED Annotated <i>dsrB</i>	SEED Annotated <i>dsrC</i>
Acetohalobium arabaticum	<i>Acetohalobium arabaticum</i>	Acetohalobium arabaticum
Acidaminococcus fermentans	<i>Acidaminococcus fermentans</i>	Acidaminococcus fermentans
Aeromonas hydrophila	<i>Aeromonas hydrophila</i>	Aeromonas hydrophila
Aeromonas veronii	<i>Aeromonas veronii</i>	Aeromonas veronii
Anaerococcus hydrogenalis	<i>Anaerococcus hydrogenalis</i>	Anaerococcus hydrogenalis
Anaerostipes caccae	<i>Anaerostipes caccae</i>	Anaerolinea thermophila
Blautia hydrogenotrophica	<i>Blautia hydrogenotrophica</i>	Anaerostipes caccae
Caldicellulosiruptor hydrothermalis	<i>Caldicellulosiruptor hydrothermalis</i>	Blautia hydrogenotrophica
Chlorobium chlorochromatii	<i>Chlorobium chlorochromatii</i>	Caldicellulosiruptor hydrothermalis
Clostridiales bacterium	<i>Citrobacter freundii</i>	Chloroherpeton thalassium
Clostridium acetobutylicum	<i>Citrobacter koseri</i>	Clostridiales bacterium
Clostridium asparagiforme	<i>Citrobacter rodentium</i>	Clostridium acetobutylicum
Clostridium bartlettii	<i>Citrobacter youngae</i>	Clostridium asparagiforme
Clostridium beijerincki	<i>Clostridiales bacterium</i>	Clostridium bartlettii
Clostridium beijerinckii	<i>Clostridium acetobutylicum</i>	Clostridium beijerincki
Clostridium bolteae	<i>Clostridium asparagiforme</i>	Clostridium beijerinckii
Clostridium botulinum	<i>Clostridium bartlettii</i>	Clostridium bolteae
Clostridium butyricum	<i>Clostridium beijerincki</i>	Clostridium botulinum
Clostridium carboxidivorans	<i>Clostridium beijerinckii</i>	Clostridium butyricum
Clostridium cellulovorans	<i>Clostridium bolteae</i>	Clostridium carboxidivorans

Table A-1

SEED Annotated <i>dsrA</i>	SEED Annotated <i>dsrB</i>	SEED Annotated <i>dsrC</i>
Clostridium difficile	<i>Clostridium botulinum</i>	Clostridium cellulovorans
Clostridium ljungdahlii	<i>Clostridium butyricum</i>	Clostridium difficile
Clostridium novyi	<i>Clostridium carboxidivorans</i>	Clostridium kluyveri
Clostridium perfringens	<i>Clostridium cellulovorans</i>	Clostridium ljungdahlii
Clostridium ramosum	<i>Clostridium difficile</i>	Clostridium novyi
Clostridium sporogenes	<i>Clostridium ljungdahlii</i>	Clostridium perfringens
Clostridium sticklandii	<i>Clostridium novyi</i>	Clostridium ramosum
Clostridium symbiosum	<i>Clostridium perfringens</i>	Clostridium sporogenes
Clostridium tetani	<i>Clostridium ramosum</i>	Clostridium sticklandii
Desulfotomaculum acetoxidans	<i>Clostridium sporogenes</i>	Clostridium symbiosum
Desulfotomaculum carboxydivorans	<i>Clostridium sticklandii</i>	Clostridium tetani
Desulfotomaculum nigrificans	<i>Clostridium symbiosum</i>	Desulfobacca acetoxidans
Desulfovibrio africanus	<i>Clostridium tetani</i>	Desulfotomaculum acetoxidans
Edwardsiella ictaluri	<i>Cronobacter sakazakii</i>	Desulfotomaculum carboxydivorans
Edwardsiella tarda	<i>Cronobacter turicensis</i>	Desulfotomaculum kuznetsovii
Epulopiscium N.t.	<i>Cronobacter turicensis</i>	Desulfotomaculum nigrificans
Erysipelotrichaceae bacterium	<i>Desulfotomaculum acetoxidans</i>	Desulfotomaculum ruminis
Escherichia albertii	<i>Desulfotomaculum carboxydivorans</i>	Desulfovibrio desulfuricans
Escherichia coli	<i>Desulfotomaculum nigrificans</i>	Edwardsiella tarda
Eubacterium limosum	<i>Dickeya dadantii</i>	Epulopiscium N.t.
Eubacterium saburreum	<i>Dickeya zeae</i>	Erysipelotrichaceae bacterium
Fusobacterium D11	<i>Edwardsiella ictaluri</i>	Escherichia albertii

Table A-1

SEED Annotated <i>dsrA</i>	SEED Annotated <i>dsrB</i>	SEED Annotated <i>dsrC</i>
Fusobacterium mortiferum	<i>Enterobacter cancerogenus</i>	Escherichia fergusonii
Fusobacterium nucleatum	<i>Enterobacter cloacae</i>	Escherichia TW09231
Fusobacterium periodonticum	<i>Enterobacter mori</i>	Eubacterium limosum
Fusobacterium ulcerans	<i>Enterobacteriaceae bacterium</i>	Eubacterium saburreum
Fusobacterium varium	<i>Erwinia billingiae</i>	Fusobacterium D11
Ilyobacter polytropus	<i>Erwinia carotovora</i>	Fusobacterium gonidiaformans
Leptotrichia buccalis	<i>Erysipelotrichaceae bacterium</i>	Fusobacterium mortiferum
Leptotrichia hofstadii	<i>Escherichia albertii</i>	Fusobacterium nucleatum
Mollicutes bacterium	<i>Escherichia coli</i>	Fusobacterium periodonticum
Moorella thermoacetica	<i>Escherichia fergusonii</i>	Fusobacterium ulcerans
Peptostreptococcus anaerobius	<i>Eubacterium limosum</i>	Fusobacterium varium
Peptostreptococcus stomatis	<i>Eubacterium saburreum</i>	Ilyobacter polytropus
Photobacterium damsela	<i>Fusobacterium gonidiaformans</i>	Intrasporangium calvum
Photobacterium leiognathi	<i>Fusobacterium mortiferum</i>	Klebsiella MS
Photobacterium profundum	<i>Fusobacterium nucleatum</i>	Koribacter versatilis
Pseudoramibacter alactolyticus	<i>Fusobacterium periodonticum</i>	Leptotrichia buccalis
Ruminococcaceae bacterium	<i>Fusobacterium ulcerans</i>	Leptotrichia hofstadii
Ruminococcus obeum	<i>Fusobacterium varium</i>	Marinomonas posidonica
Salmonella bongori	<i>Ilyobacter polytropus</i>	Methylobacterium album
Salmonella enteric	<i>Klebsiella 1</i>	Methylobacterium methanica
Salmonella enteric	<i>Klebsiella MS</i>	Mollicutes bacterium
Salmonella typhimurium	<i>Klebsiella pneumoniae</i>	Moorella thermoacetica

Table A-1

SEED Annotated <i>dsrA</i>	SEED Annotated <i>dsrB</i>	SEED Annotated <i>dsrC</i>
Sebaldella termitidis	<i>Klebsiella variicola</i>	Moraxella catarrhalis
Shigella flexneri	<i>Leptotrichia buccalis</i>	Nakamurella multipartita
Shigella sonnei	<i>Leptotrichia hofstadii</i>	Peptostreptococcus anaerobius
Shuttleworthia satellites	<i>Mollicutes bacterium</i>	Peptostreptococcus stomatis
Thermoanaerobacter brockii	<i>Moorella thermoacetica</i>	Photobacterium damsela
Thermoanaerobacter ethanolicus	<i>Pantoea aB</i>	Photobacterium leiognathi
Thermoanaerobacter italicus	<i>Pantoea ananatis</i>	Photobacterium profundum
Thermoanaerobacter mathranii	<i>Pantoea At-9b</i>	Pseudomonas fulva
Thermoanaerobacter pseudethanolicus	<i>Pantoea vagans</i>	Pseudomonas mendocina
Thermoanaerobacter wiegelii	<i>Pectobacterium atrosepticum</i>	Pseudoramibacter alactolyticus
Thermoanaerobacter X514	<i>Pectobacterium carotovorum</i>	Rheinheimera A13L
Thermoanaerobacter X561	<i>Pectobacterium wasabiae</i>	Ruminococcaceae bacterium
Thermoanaerobacterium thermosaccharolyticum	<i>Peptostreptococcus anaerobius</i>	Ruminococcus obeum
Thermoanaerobacterium xylanolyticum	<i>Peptostreptococcus stomatis</i>	Salmonella bongori
Treponema primitia	<i>Photobacterium damsela</i>	Salmonella enterica
Turcibacter HGF1	<i>Photobacterium leiognathi</i>	Salmonella typhimurium
Turcibacter PC909	<i>Photobacterium profundum</i>	Sebaldella termitidis
Vibrio harveyi	<i>Proteus penneri</i>	Serratia AS9
Vibrio rotiferianus	<i>Providencia rettgeri</i>	Shewanella baltica
	<i>Providencia stuartii</i>	Shigella boydii
	<i>Pseudoramibacter alactolyticus</i>	Shigella dysenteriae
	<i>Rahnella Y9602</i>	Shigella flexneri

Table A-1

SEED Annotated <i>dsrA</i>	SEED Annotated <i>dsrB</i>	SEED Annotated <i>dsrC</i>
	<i>Ruminococcus obeum</i>	Sideroxydans lithotrophicus
	<i>Salmonella bongori</i>	Thermoanaerobacter brockii
	<i>Salmonella enterica</i>	Thermoanaerobacter ethanolicus
	<i>Salmonella enterica</i>	Thermoanaerobacter italicus
	<i>Salmonella typhimurium</i>	Thermoanaerobacter mathranii
	<i>Sebaldella termitidis</i>	Thermoanaerobacter pseudethanolicus
	<i>Serratia marcescens</i>	Thermoanaerobacter wiegelii
	<i>Serratia odorifera</i>	Thermoanaerobacter X514
	<i>Serratia proteamaculans</i>	Thermoanaerobacter X561
	<i>Shigella boydii</i>	Thermoanaerobacterium thermosaccharolyticum
	<i>Shigella dysenteriae</i>	Thermoanaerobacterium xylanolyticum
	<i>Shigella flexneri</i>	Thermodesulfatator indicus
	<i>Shigella sonnei</i>	Thermodesulfobacterium OPB45
	<i>Shuttleworthia satelles</i>	Thermodesulfobium narugense
	<i>Sodalis glossinidius</i>	Treponema primitia
	<i>Thermoanaerobacter brockii</i>	Turicibacter HGF1
	<i>Thermoanaerobacter ethanolicus</i>	Turicibacter PC909
	<i>Thermoanaerobacter italicus</i>	Vibrio anguillarum
	<i>Thermoanaerobacter mathranii</i>	Vibrio harveyi
	<i>Thermoanaerobacter pseudethanolicus</i>	Vibrio rotiferianus
	<i>Thermoanaerobacter wiegelii</i>	

Table A-1

SEED Annotated <i>dsrA</i>	SEED Annotated <i>dsrB</i>	SEED Annotated <i>dsrC</i>
	<i>Thermoanaerobacterium xylanolyticum</i>	
	<i>Treponema primitia</i>	
	<i>Vibrio harveyi</i>	
	<i>Vibrio rotiferianus</i>	
	<i>Yersinia aldovae</i>	
	<i>Yersinia bercovieri</i>	
	<i>Yersinia enterocolitica</i>	
	<i>Yersinia frederiksenii</i>	
	<i>Yersinia intermedia</i>	
	<i>Yersinia kristensenii</i>	
	<i>Yersinia mollaretii</i>	
	<i>Yersinia pestis</i>	
	<i>Yersinia pseudotuberculosis</i>	
	<i>Yersinia rohdei</i>	
	<i>Yersinia ruckeri</i>	

Table A-2 Organisms selected for pangenome subsystem/gene analysis. SEED Annotated genomes were utilized for subsystems analysis and PCA plot, NCBI annotated genomes were used for preliminary work using Los Alamos National Laboratory's Threaded Ortholog process.

Seed Annotated Sulfate Reducing Pathway	NCBI Annotated <i>dsrA</i>
Alkalilimnicola ehrlichii MLHE-1	Acetohalobium arabaticum
Allochromatium vinosum	Acinetobacter baumannii
Ammonifex degensii KC4	Actinobacillus pleuropneumoniae
Archaeoglobus fulgidus	Alkaliphilus oremlandii
Caldivirga maquilingensis IC-167	Alteromonas macleodii
Carboxydotherrmus hydrogenoformans	Ammonifex degensii
Chlorobaculum parvum	Archaeoglobus profundus
Chlorobium chlorochromatii CaD3	Caldicellulosiruptor bescii
Chlorobium limicola	Caldicellulosiruptor hydrothermalis
Chlorobium luteolum	Caldicellulosiruptor kristjanssonii
Chlorobium phaeobacteroides	Caldicellulosiruptor kronotskyensis
Chlorobium phaeovibrioides	Caldicellulosiruptor saccharolyticus
Chlorobium tepidum TLS	Candidatus Desulforudis
Desulfatibacillum alkenivorans AK-01	Carboxydotherrmus hydrogenoformans
Desulfitobacterium hafniense DCB-2	Clostridium beijerinckii
Desulfitobacterium sp Y51	Clostridium botulinum
Desulfobacterium autotrophicum HRM2	Clostridium difficile
Desulfobulbus propionicus	Clostridium ljungdahlii
Desulfococcus oleovorans Hxd3	Clostridium novyi
Desulfohalobium retbaense	Clostridium perfringens

Table A-2

Seed Annotated Sulfate Reducing Pathway	NCBI Annotated <i>dsrA</i>
Desulfomicrobium baculatum	Clostridium phytofermentans
Desulfonatronospira thiodismutans	Clostridium tetani
Desulforudis audaxviator MP104C	Clostridium thermocellum
Desulfotalea psychrophila LSv54	Desulfarculus baarsii
Desulfotomaculum acetoxidans	Desulfatibacillum alkenivorans
Desulfotomaculum carboxydivorans	Desulfitobacterium hafniense
Desulfotomaculum kuznetsovii	Desulfobacterium autotrophicum
Desulfotomaculum nigrificans	Desulfobulbus propionicus
Desulfotomaculum reducens MI-1	Desulfococcus oleovorans
Desulfotomaculum ruminis	Desulfohalobium retbaense
Desulfovibrio aespoeensis Aspo-2	Desulfomicrobium baculatum
Desulfovibrio africanus str Walvis Bay	Desulfotalea psychrophila
Desulfovibrio desulfuricans G20	Desulfotomaculum acetoxidans
Desulfovibrio fructosovorans JJ	Desulfotomaculum reducens
Desulfovibrio magneticus RS-1	Desulfovibrio aespoeensis
Desulfovibrio piger	Desulfovibrio desulfuricans
Desulfovibrio salexigens	Desulfovibrio magneticus
Desulfovibrio sp 3 1 syn3	Desulfovibrio salexigens
Desulfovibrio sp FW1012B	Desulfovibrio vulgaris
Desulfovibrio sp ND132	Desulfurivibrio alkaliphilus
Desulfovibrio vulgaris DP4	Erwinia amylovora
Desulfurivibrio alkaliphilus AHT2	Eubacterium eligens
Desulfuromonas acetoxidans	Eubacterium rectale
Halorhodospira halophila SL1	Ferrimonas balearica

Table A-2

Seed Annotated Sulfate Reducing Pathway	NCBI Annotated <i>dsrA</i>
Magnetococcus sp MC-1	Geobacter bemidjiensis
Magnetospirillum gryphiswaldense MSR-1	Geobacter metallireducens
Magnetospirillum magneticum AMB-1	Geobacter sulfurreducens
Moorella thermoacetica	Geobacter uraniireducens
Pelodictyon luteolum	Haemophilus influenzae
Pelodictyon phaeoclathratiforme BU-1	Halorhabdus utahensis
Prosthecochloris aestuarii	Halothermothrix orenii
Prosthecochloris vibrioformis	Methanobrevibacter ruminantium
Pyrobaculum islandicum	Methanocaldococcus fervens
Ruthia magnifica str Cm Calyptogenia magnifica	Methanocaldococcus FS406
Syntrophobacter fumaroxidans MPOB	Methanocaldococcus infernus
Syntrophobotulus glycolicus	Methanocaldococcus jannaschii
Syntrophus aciditrophicus SB	Methanocaldococcus vulcanius
Thermodesulfobacterium yellowstonii	Methanococcoides burtonii
Thermosinus carboxydivorans Nor1	Methanococcus aeolicus
Thioalkalivibrio sp HL-EbGR7	Methanococcus maripaludis
Thiobacillus denitrificans	Methanococcus vanniellii
Vesicomyxosocius okutanii HA	Methanocorpusculum labreanum
	Methanoculleus marisnigri
	Methanohalobium evestigatum
	Methanohalophilus mahii
	Methanoplanus petrolearius
	Methanopyrus kandleri
	Methanosarcina acetivorans

Table A-2

Seed Annotated Sulfate Reducing Pathway	NCBI Annotated <i>dsrA</i>
	Methanosarcina barkeri
	Methanosarcina mazei
	Methanospirillum hungatei
	Methanothermobacter marburgensis
	Methanothermobacter thermautotrophicus
	Methanothermus fervidus
	Natranaerobius thermophilus
	Pantoea vagans
	Pelotomaculum thermopropionicum
	Pyrobaculum aerophilum
	Salmonella enterica
	Syntrophobacter fumaroxidans
	Syntrophobotulus glycolicus
	Thermoanaerobacter italicus
	Thermoanaerobacter mathranii
	Thermoanaerobacter pseudethanolicus
	Thermoanaerobacter tengcongensis
	Thermoanaerobacterium thermosaccharolyticum
	Thermococcus gammatolerans
	Vibrio harveyi
	Vulcanisaeta distributa

Table A-3. Pearson's Rho correlations between measured variables, part 1, continued in Table A-2

Variable	Copy number 16S	Log 16s copy number	Na (ppm)	Mg (ppm)	Si (ppm)	P (ppm)	S (ppm)	K (ppm)
Copy number 16S	1.000	1.000**	.380**	.332*	.413**	.261	.102	.695**
Log 16s copy number	1.000**	1.000	.380**	.332*	.413**	.261	.102	.695**
Na (ppm)	.380**	.380**	1.000	.607**	.091	.414**	.421**	.592**
Mg (ppm)	.332*	.332*	.607**	1.000	.069	.486**	.772**	.482**
Si (ppm)	.413**	.413**	.091	.069	1.000	.384**	-.359*	.567**
P (ppm)	.261	.261	.414**	.486**	.384**	1.000	.217	.432**
S (ppm)	.102	.102	.421**	.772**	-.359*	.217	1.000	.223
K (ppm)	.695**	.695**	.592**	.482**	.567**	.432**	.223	1.000
Ca (ppm)	.533**	.533**	.502**	.510**	.364*	.409**	.090	.542**
Mn (ppm)	.637**	.637**	.253	.138	.342*	-.179	-.051	.524**
Fe (ppm)	.464**	.464**	.133	-.257	.042	-.368*	-.205	.297*
As (ppb)	-.460**	-.460**	.063	.140	-.232	.347*	.213	-.262
Sr (ppb)	.706**	.706**	.503**	.442**	.434**	.287*	.070	.747**
Ba (ppb)	.725**	.725**	.414**	.365*	.112	.089	.218	.545**
Log APS copy number	.149	.149	.021	.067	.359*	.095	-.232	.240
Log <i>dsr</i> copy number	.487**	.487**	.128	.289*	.388**	.472**	.062	.287*

Table A-4. Correlations between measured elements and copy number of genes of interest

Variable	Ca (ppm)	Mn (ppm)	Fe (ppm)	As (ppb)	Sr (ppb)	Ba (ppb)	Log APS copy number	Log <i>dsr</i> copy number
Copy number 16S	.533**	.637**	.464**	-.460**	.706**	.725**	.149	.487**
Log 16s copy number	.533**	.637**	.464**	-.460**	.706**	.725**	.149	.487**
Na (ppm)	.502**	.253	.133	.063	.503**	.414**	.021	.128
Mg (ppm)	.510**	.138	-.257	.140	.442**	.365*	.067	.289*
Si (ppm)	.364*	.342*	.042	-.232	.434**	.112	.359*	.388**
P (ppm)	.409**	-.179	-.368*	.347*	.287*	.089	.095	.472**
S (ppm)	.090	-.051	-.205	.213	.070	.218	-.232	.062
K (ppm)	.542**	.524**	.297*	-.262	.747**	.545**	.240	.287*
Ca (ppm)	1.000	.401**	.105	-.150	.872**	.598**	.329*	.357*
Mn (ppm)	.401**	1.000	.783**	-.777**	.662**	.713**	.141	.304*
Fe (ppm)	.105	.783**	1.000	-.712**	.383**	.622**	.095	.123
As (ppb)	-.150	-.777**	-.712**	1.000	-.413**	-.496**	-.247	-.132
Sr (ppb)	.872**	.662**	.383**	-.413**	1.000	.731**	.308*	.346*
Ba (ppb)	.598**	.713**	.622**	-.496**	.731**	1.000	.176	.407**
Log APS copy number	.329*	.141	.095	-.247	.308*	.176	1.000	.214
Log <i>dsr</i> copy number	.357*	.304*	.123	-.132	.346*	.407**	.214	1.000



Figure A-1. 16S phylogenetic tree of Sequences isolated from filters. Sequences are not closely aligned with identified organisms, and are mainly discreet from each other, and separate by site, with slightly more intermingling than *dsr* phylogeny.

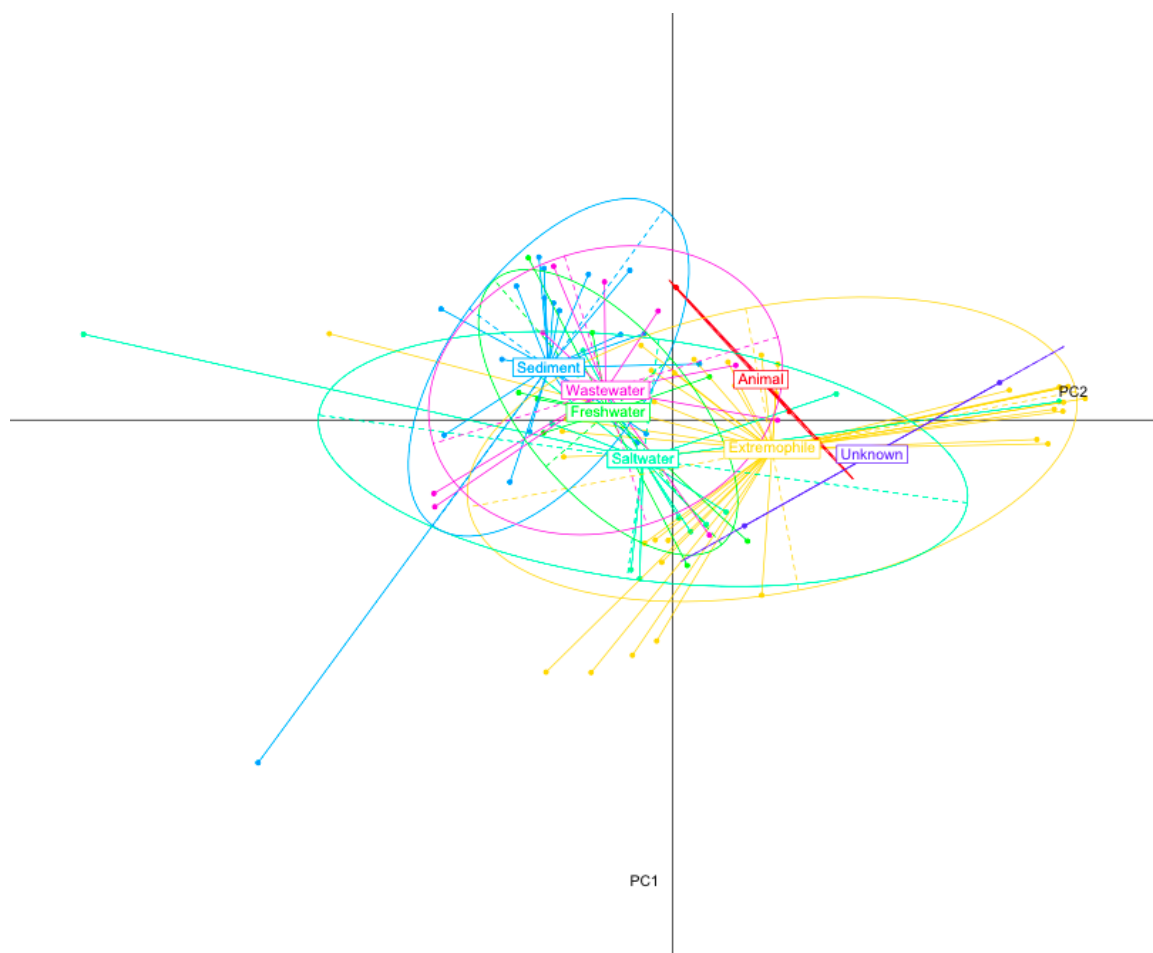


Figure A-2. PCA plot of SEED annotated subsystems for SRM grouped by isolation source. Little to no separation can be found from this method.

Vita

Matthew Scholz was born July 23, 1978. He attended high school in Albuquerque, NM. He graduated from the University of Washington, Seattle with dual degrees in biochemistry and drama. At the University of Tennessee, he explored microbial ecology and the roles and relationships of sulfate reducing bacteria on contaminated aquifer samples. The work is centered around improving quality of life for Bangladeshi citizens by reducing their chances of exposure to arsenic.

Currently, Matthew is working as a student at Los Alamos National Laboratory his current work centers around environmental microbiology and metagenome assembly. This work focuses on bioinformatics techniques to improve assembly methods and better understand characterization and assembly of complex samples.