MICROBIALLY MEDIATED TRANSFORMATION OF ARSENIC AND SULFUR COMPOUNDS IN MONO LAKE, CALIFORNIA, USA

by

CHRISTIAN FREDERICK EDWARDSON IV

(Under the Direction of James T. Hollibaugh)

ABSTRACT

The transformation of arsenic, sulfur, and thioarsenic compounds was examined in alkaline, hypersaline Mono Lake, California, USA. The microbial communities were surveyed at five depths along a redox gradient at the deepest station in the lake using both 16S rRNA tag pyrosequencing and metatranscriptomics approaches. Most of the abundant OTUs observed in the 16S amplicon survey had been found previously in Mono Lake, thus indicating the microbial community remained relatively stable through periods of meromixis and mixing. Similar taxa were found through metatranscriptomics, but abundances varied, especially with Bacteroidetes (relatively less abundant) and Proteobacteria (relatively more abundant).

The metatranscriptome samples were examined for arsenic- and sulfur-transforming bacteria by identifying transcript hits to enzymes catalyzing arsenic and sulfur redox transformations. Arsenite oxidase transcripts were dominated by the alternative (*arxA*) rather than the canonical (*aioA*) arsenite oxidase and were affiliated with organisms not previously identified as being involved in arsenite oxidation (*Thioalkalivibrio* and *Halomonas*). Arsenate reductase (*arrA*) transcripts were dominated by *Deltaproteobacteria* and *Firmicutes*. Transcripts of important sulfur cycle genes (*soxB*, *aprA*, *dsrA*) were dominated by sulfur-oxidizing

Gammaproteobacteria and *Deltaproteobacteria* at 15 and 18 m, with a transition to sulfate reducing *Deltaproteobacteria* at 18-31 m. These results highlight the shift from arsenite and sulfide oxidation in the oxycline (15-18 m) to arsenate and sulfate reduction in the anoxic bottom waters (18-31 m), as has been detected by previous rate measurements.

A phototrophic enrichment culture that was dominated by the purple sulfur bacteria *Ectothiorhodospira* sp. was obtained by amending Mono Lake water with thioarsenic compounds. The culture was able to convert thioarsenic compounds, arsenite, thiosulfate and monothioarsenate to obtain energy for photosynthetic growth. Monothioarsenate was used as the sole electron donor for anoxygenic photosynthesis, and the transformation was light-dependent. Additional pure cultures of members of the *Ectothiorhodospiraceae* were tested and most were able to convert monothioarsenate to arsenate. These findings highlight the important interactions between sulfur and arsenic in hypersaline, alkaline environments where arsenic is prevalent. The links between the bacteria that oxidize sulfur and arsenite and bacteria that reduce sulfate and arsenate are clearer.

INDEX WORDS: Haloalkaliphiles; Metatranscriptomics; 16S rRNA; Thioarsenic; *Thioalkalivibrio*; *Ectothiorhodospira*; Biogeochemistry; Extremophiles; Purple sulfur bacteria; Mono Lake; Soda lake; Arsenic; Sulfur; Redox chemistry; Microbial ecology

MICROBIALLY MEDIATED TRANSFORMATION OF ARSENIC AND SULFUR COMPOUNDS IN MONO LAKE, CALIFORNIA, USA

by

CHRISTIAN FREDERICK EDWARDSON IV

B.S, University of Minnesota, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

MICROBIALLY MEDIATED TRANSFORMATION OF ARSENIC AND SULFUR COMPOUNDS IN MONO LAKE, CALIFORNIA, USA

by

CHRISTIAN FREDERICK EDWARDSON IV

Major Professor:	James T. Hollibaugh
Committee:	Robert Maier
	Mary Ann Moran
	William B. Whitman

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

DEDICATION

To my family and friends for their love and support.

ACKNOWLEDGEMENTS

First and foremost, this would not have been possible without the guidance and support of my advisor Tim Hollibaugh. His knowledge on all things microbial ecology has shaped my understanding of science and has given me a greater appreciation for the microbial world. I also have to acknowledge all of the members of my committee for pushing me in the right direction and helping me focus. To all of the members past and present in the Hollibaugh lab – all of the shared joys and frustrations over the years got me through each day. Most importantly, I have made some incredible friends here who have helped me push through all of the ups and downs and have been there for me. I could not have done this without them. There are numerous other people, places, and things that have made this possible. I will be forever grateful to them for the memories.

TABLE OF CONTENTS

Page

APPENDICES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Literature Review

Arsenic and Sulfur Biogeochemistry

Understanding the biogeochemistry of arsenic (As) is important to the toxicity of inorganic arsenic compounds and their presence in terrestrial and aquatic environments, from both natural (e.g. geothermal) and anthropogenic sources (e.g. mining). Dissolved inorganic arsenic is found in the environment primarily as the As^{+5} oxyanion arsenate (As(V), e.g. $HAsO_4^{2-}$) and the As^{+3} oxyanion arsenite (As(III), e.g. $H_2AsO_3^-$). Detoxification transformations of arsenic by microorganisms have been well studied, and include methylation, and non-energy yielding oxidation and reduction (Mukhopadhyay *et al.* 2002). Energy-yielding redox transformations of these compounds by microorganisms have only been identified in the last 20 years (Ahmann *et al.* 1994, Santini *et al.* 2000). Numerous reviews of arsenic biogeochemistry have been written, covering topics such as quantification of different forms of arsenic in the biosphere (Campbell and Nordstrom 2014, Cullen and Reimer 1989, Matschullat 2000, Smedley and Kinniburgh 2002) and arsenic redox transformations (Bissen and Frimmel 2003, Lievremont *et al.* 2009, Oremland and Stolz 2003, Sharma and Sohn 2009). In addition, arsenic redox chemistry has been argued to be important to the origin and evolution of life on earth (Duval *et al.* 2008, Kulp *et al.* 2008, Oremland *et al.* 2009, Schoepp-Cothenet *et al.* 2012, van Lis *et al.* 2013).

Arsenic-sulfur compounds are also present in the environment. Historically, much of the research on As-S compounds has been done on arsenic bearing minerals (arsenopyrite, orpiment, and realgar) due their prevalence in the geosphere (Lengke *et al.* 2009, Smedley and Kinniburgh 2002). Recent studies have found that dissolved arsenic can be present as stable thioarsenic compounds in some environments, especially in sulfidic waters (Hollibaugh *et al.* 2005, Wallschlager and Stadey 2007). The detection of these compounds was made possible due to methodological innovations, namely the application of ion chromatography with inductively-coupled plasma mass spectrometry (IC-ICP-MS) (Planer-Friedrich *et al.* 2007, Wallschlager and Stadey 2007, Wilkin *et al.* 2003). Thioarsenic compounds have been detected in sulfidic and neutral-to-alkaline waters (Wallschlager and Stadey 2007), including hot springs in Yellowstone National Park (Planer-Friedrich *et al.* 2007, Planer-Friedrich *et al.* 2009), groundwater (Stucker *et al.* 2013, Wallschlager and Stadey 2007), and in Mono Lake, California (Hollibaugh *et al.* 2005) and Big Soda Lake, NV (Planer-Friedrich, unpublished data). Synthesized thioarsenic standards analyzed by IC-ICP-MS were determined to be thioarsenates, with As in the +5 oxidation state (Stauder *et al.* 2005). X-ray absorption spectroscopy (XAS) has been used to differentiate between thioarsenates and thioarsenites in the same solution (Beak *et al.* 2008, Suess *et al.* 2009) and to determine that thioarsenite is formed in anoxic solutions of sulfide and arsenite (Planer-Friedrich *et al.* 2010). XAS is not sensitive enough for the analysis of environmental samples (Suess *et al.* 2009), and thioarsenites are rapidly converted to thioarsenates upon exposure to oxygen during IC-ICP-MS analysis (Planer-Friedrich *et al.* 2010). As a consequence, the redox state of the arsenic atom in environmental thioarsenic compounds is still unknown. However, since the compounds that are analyzed are believed to be thioarsenate, this is the term that is generally used, although it has been postulated that both thioarsenates and

thioarsenites could be stable in certain conditions (Helz and Tossell 2008).

Microbial Transformations of Sulfur and Arsenic

The known roles of prokaryotes in redox transformations of arsenic include dissimilatory arsenate respiration (reduction) and arsenite oxidation. These processes can yield energy or be resistance mechanisms (Oremland and Stolz 2003). Dissimilatory arsenate reducing/respiring prokaryotes are either heterotrophs that can grow on a variety of organic carbon sources or chemolithoautotrophs. Although these microorganisms grow using arsenate, many are able to use other electron acceptors such as nitrate or oxygen (Oremland and Stolz 2003, Oremland *et al.* 2009). Microbial oxidation of arsenite to arsenate is widely distributed, both in heterotrophs (as a detoxification mechanism) and in chemolithoautotrophs that use arsenite as an electron donor. Oxidation of arsenite is coupled to a variety of terminal electron acceptors including oxygen, nitrate, and Fe(III) (Oremland and Stolz 2003). In addition, anoxygenic photosynthesis has been shown to be driven by arsenite oxidation in a variety of purple sulfur bacteria (Budinoff and Hollibaugh 2008, Kulp *et al.* 2008). The role of microbes in the biological transformation of thioarsenic compounds has only been studied recently in experiments with Mono Lake water (Fisher *et al.* 2008) and in geothermal drainages in Yellowstone National Park (Planer-Friedrich *et al.* 2009). In experiments with both Mono Lake water and enrichment cultures grown in minimal medium, thioarsenic compounds were formed from arsenite and sulfide, they were stable for at least 24 hours, and then were converted to arsenate or a mixture of arsenate and arsenite. Little to no arsenate was formed in sterile or killed controls (Fisher *et al.* 2008). Transformation of thioarsenic to arsenite occurred up to 500 times faster in Yellowstone geothermal drainages than in abiotic laboratory solutions, likely due to the influence of microbial activity (Planer-Friedrich *et al.* 2009). Studies have concluded that a mix of biotic

transformations mediated by *Thermocrinis* spp., and abiotic oxidation was responsible for thioarsenic transformations in this environment (Härtig and Planer-Friedrich 2012). Pure culture work confirmed that *Thermocrinis ruber* is capable of monothioarsenate transformation via oxidation of the thiol group, in addition to arsenite oxidation (Härtig *et al.* 2014). Finally, recent work has also shown that the sulfide-oxidizing, arsenate-reducing Deltaproteobacteria strain MLMS-1 (Hoeft *et al.* 2004) is capable of disproportionation of monothioarsenate to arsenite and S(0) (Planer-Friedrich *et al.* 2015). The complexities of thioarsenic transformation in natural waters are highlighted by the fact that arsenite and arsenate are both present during thioarsenic transformation in some laboratory experiments (Fisher *et al.* 2008). In addition, the natural distribution of arsenate, arsenite, and thioarsenic compounds in Mono Lake is dynamic spatially (by depth) and temporally (Hollibaugh *et al.* 2005). This complicates the interpretation of the mechanisms of formation and decomposition based on chemical analyses alone. In addition, the role of microorganisms in this process remains to be determined. It is also unclear if this process is mediated by a consortium of microorganisms or by a single species. Further, the genes, gene products, and pathways involved in this process are unknown, thus more experiments with pure cultures and molecular analysis is warranted.

Prokaryotic transformations of arsenic have been studied extensively in Mono Lake, CA, due to its naturally high (200 μM) levels of dissolved inorganic arsenic. Mono Lake is a closed basin lake located in the western portion of the Great Basin, east of the Sierra Nevada Mountains. The lake is hypersaline (approximately 90 g L^{-1}) and alkaline (pH 9.8). The high arsenic content is due to geothermal activity and input from hot springs. The lake is seasonally stratified, with typically one mixing event per year (monomixis), but has undergone periods of extended meromixis (prolonged stratification) in the recent past (Melack and Jellison 1998). The anoxic

hypolimnion contains high concentrations of reduced compounds such as arsenite, sulfide, and thioarsenic compounds, while the aerobic epilimnion contains the oxidized forms of these compounds: arsenate and sulfate (Hollibaugh *et al.* 2005, Oremland *et al.* 2004). Arsenate reduction was measured in Mono Lake and was found to occur in both the water column (Oremland *et al.* 2000) and sediments (Kulp *et al.* 2006). In addition, sulfate reduction is an important process in the anoxic bottom waters and sediment of the lake (Kulp *et al.* 2006, Oremland *et al.* 2000). Mono Lake is dominated by a strain of the picoeukaryote *Picocystis*, which contributes up to 50% of primary production in the lake (Roesler *et al.* 2002). The dynamics of *Picocystis* and of bacterioplankton are strongly influenced by the stratification and mixing regimes of the lake, as well as by grazing by the brine shrimp, *Artemia monica* (Jellison and Melack 1993). Other studies of the microbial ecology of Mono Lake have focused on important processes mediated by microbes including ammonia and methane oxidation (Carini and Joye 2008, Joye *et al.* 1999, Lin *et al.* 2005, Ward *et al.* 2000), nitrogen fixation (Oremland 1990, Steward *et al.* 2004), chitin degradation (LeCleir *et al.* 2004, LeCleir *et al.* 2007), selenium transformations (Blum *et al.* 1998, Fisher and Hollibaugh 2008), and viral dynamics (Jiang *et al.* 2004). The most well studied organisms from Mono Lake have been those that use inorganic arsenic and sulfur compounds to gain energy for growth (Table 1.1).

Mono Lake Microbial Ecology

The composition of the Mono Lake microbial community was first analyzed using PCR-DGGE (Hollibaugh *et al.* 2001), subsequently augmented by sequencing amplicons of the 16S rRNA gene (Humayoun *et al.* 2003). Molecular characterization of both arsenic and sulfur cycling pathways in microorganisms from the lake used PCR amplification and sequencing of marker genes. Arsenate respiratory reductase (*arrA*) has been used as a marker for arsenate

reduction in the environment and in cultures (Malasarn *et al.* 2004), and studies have shown amplification of this gene from Mono Lake (Hollibaugh *et al.* 2006, Kulp *et al.* 2006). The canonical aerobic arsenite oxidase gene (*aioA*) (Lett *et al.* 2012) was not detected in arsenite-oxidizing cultures from Mono Lake (Hoeft *et al.* 2007, Kulp *et al.* 2008) using a variety of primers. However, recent studies have shown that the arsenate reductase enzyme (ArrA) can work in reverse (Richey *et al.* 2009) and an alternative arsenite oxidase gene that is more similar to *arrA* than to *aioA* (identified as *arxA*) has been found in cultures and the environment (Zargar *et al.* 2010, Zargar *et al.* 2012). The enzymes responsible for arsenic transformations (AioA, ArrA, ArxA) are all members of the DMSO family of molybdopterin oxidoreductases, which are part of a larger class of enzymes known as the complex iron sulfur molybdoenzymes (CISM) (McEwan *et al.* 2002, Rothery *et al.* 2008, Schoepp-Cothenet *et al.* 2012, van Lis *et al.* 2013). These enzymes have been postulated to be very ancient, possibly pre-dating last universal common ancestors (LUCA) (Nitschke *et al.* 2013).

Microbial Sulfur Cycling

Compared with arsenic redox cycling, sulfur redox cycling by prokaryotes is more complicated, but has received more attention. Oxidation of reduced sulfur compounds – including sulfide, thiosulfate, elemental sulfur and polysulfides – is a process that is mediated by a wide variety of organisms and molecular mechanisms. The major groups of sulfur-oxidizing bacteria (SOB) include the Green Sulfur Bacteria (*Chlorobiaceae*), Purple Sulfur Bacteria, and Purple Non-Sulfur Bacteria. Mechanisms of sulfur oxidation include sulfide:quinone reductase (SQR), the reverse-acting sulfite reductase (Dsr) system, and the Sox system (Ghosh and Dam 2009). Key enzymes in the inorganic sulfur redox cycle involve oxidation of thiosulfate by the Sox system (SoxAXKBCDYZ), sulfide oxidation by sulfide:quinone reductase (SQR), transformation

of sulfide to sulfate or sulfate to sulfide through dissimilatory sulfite reductase (DsrAB and 13 other proteins), adenosine-5'-phosphosulfate reductase (AprBA), and sulfate adenylyl transferase (Sat) (Dahl 2008, Friedrich *et al.* 2001, Frigaard and Dahl 2008, Muyzer and Stams 2008). The genes encoding these enzymes have been used as molecular markers (Loy *et al.* 2009, Meyer *et al.* 2007, Meyer and Kuever 2007, Mori *et al.* 2010, Muller *et al.* 2015). The detection of sulfate-reducing bacteria in Mono Lake using *aprA* (*apsA*) and *dsrAB* was unsuccessful in most cases (Scholten *et al.* 2005), but divergence of the microbial community in Mono Lake from the published primer sequences likely the affected ability to detect these genes, as newly designed primers performed better with Mono Lake enrichment cultures (Meyer and Kuever 2007). Detection of sulfur-oxidizing bacteria using *soxB* (Tourova *et al.* 2013) and detection of sulfate-reducing bacteria using *dsr* (Foti *et al.* 2008) has provided insights into sulfur cycling in other soda lakes.

Omics Tools for Microbial Community Analysis

Next-generation high throughput sequencing platforms (454 pyrosequencing and Illumina sequencing-by-synthesis) have become more prevalent than PCR-clone-sequence marker gene surveys and have been used in numerous studies including 16S rRNA amplicon surveys, especially in microbiome research (Caporaso *et al.* 2011), metagenomics (Dick *et al.* 2009, Sharon *et al.* 2013), and metatranscriptomics (Gifford *et al.* 2011, Hollibaugh *et al.* 2011, Moran *et al.* 2013). Previous studies of this type have looked at communities in habitats similar to Mono Lake, including meromictic soda lakes (Antony *et al.* 2013, Hawley and Hess 2014), submarine alkaline springs (Glaring *et al.* 2015), hypersaline mats (Harris *et al.* 2013), and hypersaline Lake Tyrell (Narasingarao *et al.* 2012). Arsenic and sulfur cycling have also been examined using these methods (Bertin *et al.* 2011, Plewniak *et al.* 2013, Rascovan *et al.* 2015), but generally

have been focused on metagenomics rather than metatranscriptomics, the latter of which allows identification of the active members of microbial communities.

Objectives

The objectives of this dissertation are to combine molecular, chemical, and culture-based methods to explore microbial sulfur and arsenic cycling in Mono Lake, California. In Chapter 2, a big-picture overview of the composition of the microbial community of Mono Lake, CA is presented. Samples were taken at five different depths, covering the range of redox conditions in the stratified lake. The vertical distribution of microorganisms was determined using both 16S rRNA tag pyrosequencing and from taxonomy of mRNA transcripts. The most abundant 16S rRNA OTUs (>1% relative abundance) were compared with reference databases to determine the most abundant taxonomic groups. In addition, the relative abundance and phylogenetic composition of these sequences was compared with results from a previous study of the Mono Lake microbial community. Abundant (>1% relative abundance), transcriptionally active taxa were identified and compared at a broader taxonomic level (phylum and class) with OTUs defined by 16S rRNA sequences. The potential role of the different groups of bacteria in the environment is discussed and the results are compared with other studies made of other environments that are similar to Mono Lake.

In Chapter 3, the arsenic and sulfur cycles in Mono Lake were examined more closely using metatranscriptomics. The dataset from Chapter 2 was used to explore the distribution of key transcripts for arsenite oxidation, arsenate reduction, sulfur/sulfide oxidation and sulfate reduction. A custom database approach was used to identify different molybdopterin oxidoreductase transcripts (which include arsenite oxidase and arsenate reductase). Another custom database was used to identify transcripts of key marker genes in the sulfur cycle

(adenosine-5'-phosphate reductase, *aprA*; dissimilatory sulfite reductase, *dsrA*; and thiosulfate hydrolase, *soxB*). In addition, reads were assembled to recover full length transcripts. The protein translation of these full length transcripts, as well as reference sequences that had abundant (>1% relative abundance) hits were used to create a phylogeny of each of the arsenic and sulfur cycling proteins. The abundances of hits to each reference sequence in the phylogenies were compared, highlighting the shift in sulfur and arsenic redox cycling with depth. A bacterium with abundant hits to both arsenic and sulfur transcripts (*Thioalkalivibrio*) was investigated in further detail. The difference in individual transcript abundance (relative to overall transcript abundance at each depth) was investigated to determine transcripts that were more abundant at one depth versus others. These results provide insight into specific aspects of the arsenic and sulfur cycle in Mono Lake and provide further evidence that sulfur and arsenic cycling are connected in this environment.

In Chapter 4, arsenic, sulfur, and thioarsenic transformation by a specific group of bacteria (Ectothiorhodospiraceae) is examined. An anoxygenic photosynthetic enrichment culture was obtained from Mono Lake water using a mixture of thioarsenic compounds as the sole electron donor. This enrichment culture was tested for its ability to grow on mixtures of thioarsenic compounds, and more specifically for its ability to use the pure thioarsenic compound monothioarsenate. The bacteria in the enrichment culture were identified using 16S ribosomal RNA gene sequencing. The culture was dominated by an *Ectothiorhodospira* sp., a purple sulfur bacterium. Once the dominant bacterium in the enrichment culture was identified, pure cultures related to this group were tested for their ability to grow on monothioarsenate. The ability of these anoxygenic phototrophs to use monothioarsenate as the sole electron donor was found to be widespread in the Ectothiorhodospiraceae, with few exceptions. The ability of bacteria to

transform thioarsenates directly via anoxygenic photosynthesis adds a new perspective to the arsenic and sulfur cycles.

References

Abin CA, Hollibaugh JT (2014). Dissimilatory antimonate reduction and production of antimony trioxide microcrystals by a novel microorganism. *Environ Sci Technol* **48:** 681-688.

Ahmann D, Roberts AL, Krumholz LR, Morel FMM (1994). Microbe grows by reducing arsenic. *Nature* **371:** 750-750.

Antony CP, Kumaresan D, Hunger S, Drake HL, Murrell JC, Shouche YS (2013). Microbiology of Lonar Lake and other soda lakes. *ISME J* **7:** 468-476.

Baesman SM, Stolz JF, Kulp TR, Oremland RS (2009). Enrichment and isolation of *Bacillus beveridgei* sp. nov., a facultative anaerobic haloalkaliphile from Mono Lake, California, that respires oxyanions of tellurium, selenium, and arsenic. *Extremophiles* **13:** 695-705.

Beak DG, Wilkin RT, Ford RG, Kelly SD (2008). Examination of arsenic speciation in sulfidic solutions using X-ray absorption spectroscopy. *Environ Sci Technol* **42:** 1643-1650.

Bertin PN, Heinrich-Salmeron A, Pelletier E, Goulhen-Chollet F, Arsene-Ploetze F, Gallien S *et al.* (2011). Metabolic diversity among main microorganisms inside an arsenic-rich ecosystem revealed by meta- and proteo-genomics. *ISME J* **5:** 1735-1747.

Bissen M, Frimmel FH (2003). Arsenic - a review. - Part 1: Occurrence, toxicity, speciation, mobility. *Acta Hydroch Hydrob* **31:** 9-18.

Blum JS, Bindi AB, Buzzelli J, Stolz JF, Oremland RS (1998). *Bacillus arsenicoselenatis*, sp. nov., and *Bacillus selenitireducens*, sp. nov.: two haloalkaliphiles from Mono Lake, California that respire oxyanions of selenium and arsenic. *Arch Microbiol* **171:** 19-30.

Budinoff CR, Hollibaugh JT (2008). Arsenite-dependent photoautotrophy by an *Ectothiorhodospira*-dominated consortium. *ISME J* **2:** 340-343.

Campbell KM, Nordstrom DK (2014). Arsenic speciation and sorption in natural environments. *Rev Mineral Geochem* **79:** 185-216.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* **108:** 4516-4522.

Carini SA, Joye SB (2008). Nitrification in Mono Lake, California: Activity and community composition during contrasting hydrological regimes. *Limnol Oceanogr* **53:** 2546-2557.

Cullen WR, Reimer KJ (1989). Arsenic speciation in the environment. *Chem Rev* **89:** 713-764.

Dahl C (2008). Inorganic sulfur compounds as electron donors in purple sulfur bacteria. In: Hell R, Dahl C, Knaff D, Leustek T (eds). *Sulfur Metabolism in Phototrophic Organisms*. Springer: Netherlands. pp 289-317.

Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP *et al.* (2009). Community-wide analysis of microbial genome sequence signatures. *Genome biology* **10:** R85.

Duval S, Ducluzeau AL, Nitschke W, Schoepp-Cothenet B (2008). Enzyme phylogenies as markers for the oxidation state of the environment: the case of respiratory arsenate reductase and related enzymes. *BMC evolutionary biology* **8:** 206.

Fisher JC, Hollibaugh JT (2008). Selenate-dependent anaerobic arsenite oxidation by a bacterium from Mono Lake, California. *Appl Environ Microbiol* **74:** 2588-2594.

Fisher JC, Wallschlager D, Planer-Friedrich B, Hollibaugh JT (2008). A new role for sulfur in arsenic cycling. *Environ Sci Technol* **42:** 81-85.

Foti M, Sorokin D, Zacharova E, Pimenov N, Kuenen JG, Muyzer G (2008). Bacterial diversity and activity along a salinity gradient in soda lakes of the Kulunda Steppe (Altai, Russia). *Extremophiles* **12:** 133-145.

Friedrich CG, Rother D, Bardischewsky F, Quentmeier A, Fischer J (2001). Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? *Appl Environ Microbiol* **67:** 2873-2882.

Frigaard N-U, Dahl C (2008). Sulfur metabolism in phototrophic sulfur bacteria. *Advances in Microbial Physiology*. pp 103-200.

Ghosh W, Dam B (2009). Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. *FEMS Microbiol Rev* **33:** 999-1043.

Gifford SM, Sharma S, Rinta-Kanto JM, Moran MA (2011). Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J* **5:** 461-472.

Glaring MA, Vester JK, Lylloff JE, Abu Al-Soud W, Sørensen SJ, Stougaard P (2015). Microbial diversity in a permanently cold and alkaline environment in Greenland. *PLoS ONE* **10:** e0124863.

Harris JK, Caporaso JG, Walker JJ, Spear JR, Gold NJ, Robertson CE *et al.* (2013). Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. *ISME J* **7:** 50-60.

Härtig C, Planer-Friedrich B (2012). Thioarsenate transformation by filamentous microbial mats thriving in an alkaline, sulfidic hot spring. *Environ Sci Technol* **46:** 4348-4356.

Härtig C, Lohmayer R, Kolb S, Horn MA, Inskeep WP, Planer-Friedrich B (2014). Chemolithotrophic growth of the aerobic hyperthermophilic bacterium *Thermocrinis ruber* OC 14/7/2 on monothioarsenate and arsenite. *FEMS Microbiol Ecol* **90:** 747-760.

Hawley ER, Hess M (2014). Metagenome sequencing of the prokaryotic microbiota of the hypersaline and meromictic Soap Lake, Washington. *Genome announcements* **2:** e01212-01213.

Helz G, Tossell J (2008). Thermodynamic model for arsenic speciation in sulfidic waters: A novel use of *ab initio* computations. *Geochim Cosmochim Acta* **72:** 4457-4468.

Hoeft SE, Kulp TR, Stolz JF, Hollibaugh JT, Oremland RS (2004). Dissimilatory arsenate reduction with sulfide as electron donor: Experiments with mono lake water and isolation of Strain MLMS-1, a chemoautotrophic arsenate respirer. *Appl Environ Microbiol* **70:** 2741-2747.

Hoeft SE, Blum JS, Stolz JF, Tabita FR, Witte B, King GM *et al.* (2007). *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. *Int J Syst Evol Microbiol* **57:** 504-512.

Hollibaugh J, Carini S, Gurleyuk H, Jellison R, Joye S, Lecleir G *et al.* (2005). Arsenic speciation in Mono Lake, California: Response to seasonal stratification and anoxia. *Geochim Cosmochim Acta* **69:** 1925-1937.

Hollibaugh JT, Wong PS, Bano N, Pak SK, Prager EM, Orrego C (2001). Stratification of microbial assemblages in Mono Lake, California, and response to a mixing event. *Hydrobiologia* **466:** 45-60.

Hollibaugh JT, Budinoff C, Hollibaugh RA, Ransom B, Bano N (2006). Sulfide oxidation coupled to arsenate reduction by a diverse microbial community in a soda lake. *Appl Environ Microbiol* **72:** 2043-2049.

Hollibaugh JT, Gifford S, Sharma S, Bano N, Moran MA (2011). Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J* **5:** 866-878.

Humayoun SB, Bano N, Hollibaugh JT (2003). Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl Environ Microbiol* **69:** 1030-1042.

Jellison R, Melack JM (1993). Algal photsynthetic activity and its response to meromixis in hypersaline Mono Lake, California. *Limnol Oceanogr* **38:** 818-837.

Jiang S, Steward G, Jellison R, Chu W, Choi S (2004). Abundance, distribution, and diversity of viruses in alkaline, hypersaline Mono Lake, California. *Microb Ecol* **47:** 9-17.

Joye SB, Connell TL, Miller LG, Oremland RS, Jellison RS (1999). Oxidation of ammonia and methane in an alkaline, saline lake. *Limnol Oceanogr* **44:** 178-188.

Kulp TR, Hoeft SE, Miller LG, Saltikov C, Murphy JN, Han S *et al.* (2006). Dissimilatory arsenate and sulfate reduction in sediments of two hypersaline, arsenic-rich soda lakes: Mono and Searles lakes, California. *Appl Environ Microbiol* **72:** 6514-6526.

Kulp TR, Hoeft SE, Asao M, Madigan MT, Hollibaugh JT, Fisher JC *et al.* (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science* **321:** 967-970.

LeCleir GR, Buchan A, Hollibaugh JT (2004). Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions. *Appl Environ Microbiol* **70:** 6977-6983.

LeCleir GR, Buchan A, Maurer J, Moran MA, Hollibaugh JT (2007). Comparison of chitinolytic enzymes from an alkaline, hypersaline lake and an estuary. *Envrion Microbiol* **9:** 197-205.

Lengke MF, Sanpawanitchakit C, Tempel RN (2009). The oxidation and dissolution of arsenic-bearing sulfides. *Can Mineral* **47:** 593-613.

Lett MC, Muller D, Lievremont D, Silver S, Santini J (2012). Unified nomenclature for genes involved in prokaryotic aerobic arsenite oxidation. *J Bacteriol* **194:** 207-208.

Lievremont D, Bertin PN, Lett MC (2009). Arsenic in contaminated waters: biogeochemical cycle, microbial metabolism and biotreatment processes. *Biochimie* **91:** 1229-1237.

Lin J-L, Joye SB, Scholten JCM, Schäfer H, McDonald IR, Murrell JC (2005). Analysis of methane monooxygenase genes in Mono Lake suggests that increased methane oxidation activity may correlate with a change in methanotroph community structure. *Appl Environ Microbiol* **71:** 6458-6462.

Loy A, Duller S, Baranyi C, Mussmann M, Ott J, Sharon I *et al.* (2009). Reverse dissimilatory sulfite reductase as phylogenetic marker for a subgroup of sulfur-oxidizing prokaryotes. *Environ Microbiol* **11:** 289-299.

Malasarn D, Saltikov W, Campbell KM, Santini JM, Hering JG, Newman DK (2004). *arrA* is a reliable marker for As(V) respiration. *Science* **306:** 455-455.

Matschullat J (2000). Arsenic in the geosphere - a review. *The Science of the total environment* **249:** 297-312.

McEwan AG, Ridge JP, McDevitt CA, Hugenholtz P (2002). The DMSO reductase family of microbial molybdenum enzymes; molecular properties and role in the dissimilatory reduction of toxic elements. *Geomicrobiol J* **19:** 3-21.

Melack J, Jellison R (1998). Limnological conditions in Mono Lake: contrasting monomixis and meromixis in the 1990s. *Hydrobiologia* **384:** 21-39.

Meyer B, Imhoff JF, Kuever J (2007). Molecular analysis of the distribution and phylogeny of the *soxB* gene among sulfur-oxidizing bacteria - evolution of the Sox sulfur oxidation enzyme system. *Environ Microbiol* **9:** 2957-2977.

Meyer B, Kuever J (2007). Molecular analysis of the diversity of sulfate-reducing and sulfur-oxidizing prokaryotes in the environment, using *aprA* as functional marker gene. *Appl Environ Microbiol* **73:** 7664-7679.

Moran MA, Satinsky B, Gifford SM, Luo H, Rivers A, Chan LK *et al.* (2013). Sizing up metatranscriptomics. *ISME J* **7:** 237-243.

Mori Y, Purdy KJ, Oakley BB, Kondo R (2010). Comprehensive detection of phototrophic sulfur bacteria using PCR primers that target reverse dissimilatory sulfite reductase gene. *Microbes Environ* **25:** 190-196.

Mukhopadhyay R, Rosen BP, Phung LT, Silver S (2002). Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiol Rev* **26:** 311-325.

Muller AL, Kjeldsen KU, Rattei T, Pester M, Loy A (2015). Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J* **9:** 1152-1165.

Muyzer G, Stams AJ (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nature reviews Microbiology* **6:** 441-454.

Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ *et al.* (2012). De novo metagenomic assembly reveals abundant novel major lineage of archaea in hypersaline microbial communities. *ISME J* **6:** 81-93.

Nitschke W, McGlynn SE, Milner-White EJ, Russell MJ (2013). On the antiquity of metalloenzymes and their substrates in bioenergetics. *BBA-Bioenergetics* **1827:** 871-881.

Oremland RS (1990). Nitrogen fixation dynamics of two diazotrophic communities in Mono Lake, California. *Appl Environ Microbiol* **56:** 614-622.

Oremland RS, Dowdle PR, Hoeft S, Sharp JO, Schaefer JK, Miller LG *et al.* (2000). Bacterial dissimilatory reduction of arsenate and sulfate in meromictic Mono Lake, California. *Geochim Cosmochim Acta* **64:** 3073-3084.

Oremland RS, Hoeft SE, Santini JM, Bano N, Hollibaugh RA, Hollibaugh JT (2002). Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, Strain MLHE-1. *Appl Environ Microbiol* **68:** 4795-4802.

Oremland RS, Stolz JF (2003). The ecology of arsenic. *Science* **300:** 939-944.

Oremland RS, Stolz JF, Hollibaugh JT (2004). The microbial arsenic cycle in Mono Lake, California. *FEMS Microbiol Ecol* **48:** 15-27.

Oremland RS, Saltikov CW, Wolfe-Simon F, Stolz JF (2009). Arsenic in the evolution of Earth

and extraterrestrial ecosystems. *Geomicrobiol J* **26:** 522-536.

Pikuta EV, Hoover RB, Bej AK, Marsic D, Whitman WB, Cleland D *et al.* (2003). *Desulfonatronum thiodismutans* sp. nov., a novel alkaliphilic, sulfate-reducing bacterium capable of lithoautotrophic growth. *International Journal of Systematic and Evolutionary Microbiology* **53:** 1327-1332.

Planer-Friedrich B, London J, McCleskey RB, Nordstrom DK, Wallschlager D (2007). Thioarsenates in geothermal waters of Yellowstone National Park: Determination, preservation, and geochemical importance. *Environ Sci Technol* **41:** 5245-5251.

Planer-Friedrich B, Fisher J, Hollibaugh J, Suess E, Wallschlager D (2009). Oxidative transformation of trithioarsenate along alkaline geothermal drainages—abiotic versus microbially mediated processes. *Geomicrobiol J* **26:** 339-350.

Planer-Friedrich B, Suess E, Scheinost AC, Wallschlager D (2010). Arsenic speciation in sulfidic waters: reconciling contradictory spectroscopic and chromatographic evidence. *Anal Chem* **82:** 10228-10235.

Planer-Friedrich B, Hartig C, Lohmayer R, Suess E, McCann SH, Oremland R (2015). Anaerobic chemolithotrophic growth of the haloalkaliphilic bacterium Strain MLMS-1 by disproportionation of monothioarsenate. *Environ Sci Technol* **49:** 6554-6563.

Plewniak F, Koechler S, Navet B, Dugat-Bony E, Bouchez O, Peyret P *et al.* (2013). Metagenomic insights into microbial metabolism affecting arsenic dispersion in Mediterranean marine sediments. *Molecular ecology* **22:** 4870-4883.

Rascovan N, Maldonado J, Vazquez MP, Eugenia Farias M (2015). Metagenomic study of red biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. *ISME J* **e-pub ahead of print 3 July 2015:** doi: 10.1038/ismej.2015.1109.

Richey C, Chovanec P, Hoeft SE, Oremland RS, Basu P, Stolz JF (2009). Respiratory arsenate reductase as a bidirectional enzyme. *Biochem Bioph Res Co* **382:** 298-302.

Roesler CS, Culbertson CW, Etheridge SM, Goericke R, Kiene RP, Miller LG *et al.* (2002). Distribution, production, and ecophysiology of *Picocystis* strain ML in Mono Lake, California. *Limnol Oceanogr* **47:** 440-452.

Rothery RA, Workun GJ, Weiner JH (2008). The prokaryotic complex iron–sulfur

molybdoenzyme family. *BBA-Biomembranes* **1778:** 1897-1929.

Santini JM, Sly LI, Schnagl RD, Macy JM (2000). A new chemolithoautotrophic arsenite-oxidizing bacterium isolated from a gold mine: Phylogenetic, physiological, and preliminary biochemical studies. *Appl Environ Microbiol* **66:** 92-97.

Schoepp-Cothenet B, van Lis R, Philippot P, Magalon A, Russell MJ, Nitschke W (2012). The ineluctable requirement for the trans-iron elements molybdenum and/or tungsten in the origin of life. *Sci Rep* **2:** 263.

Scholten JCM, Joye SB, Hollibaugh JT, Murrell JC (2005). Molecular analysis of the sulfate reducing and archaeal community in a meromictic soda lake (Mono Lake, California) by targeting 16S rRNA, *mcrA*, *apsA*, and *dsrAB* Genes. *Microb Ecol* **50:** 29-39.

Sharma VK, Sohn M (2009). Aquatic arsenic: toxicity, speciation, transformations, and remediation. *Environment international* **35:** 743-759.

Sharon I, Morowitz MJ, Thomas BC, Costello EK, Relman DA, Banfield JF (2013). Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res* **23:** 111-120.

Smedley PL, Kinniburgh DG (2002). A review of the source, behaviour and distribution of arsenic in natural waters. *Appl Geochem* **17:** 517-568.

Sorokin DY, Gorlenko VM, Tourova TP, Tsapin A, Nealson KH, Kuenen GJ (2002). *Thioalkalimicrobium cyclicum* sp. nov. and *Thioalkalivibrio jannaschii* sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing bacteria from hypersaline alkaline Mono Lake (California). *Int J Syst Evol Microbiol* **52:** 913-920.

Stauder S, Raue B, Sacher F (2005). Thioarsenates in sulfidic waters. *Environ Sci Technol* **39:** 5933-5939.

Steward GF, Zehr JP, Jellison R, Montoya JP, Hollibaugh JT (2004). Vertical distribution of nitrogen-fixing phylotypes in a meromictic, hypersaline lake. *Microb Ecol* **47:** 30-40.

Stucker VK, Williams KH, Robbins MJ, Ranville JF (2013). Arsenic geochemistry in a biostimulated aquifer: An aqueous speciation study. *Environ Toxicol Chem* **32:** 1216-1223.

Suess E, Scheinost AC, Bostick BC, Merkel BJ, Wallschlaeger D, Planer-Friedrich B (2009). Discrimination of thioarsenites and thioarsenates by x-ray absorption spectroscopy. *Anal Chem* **81:** 8318-8326.

Tourova TP, Slobodova NV, Bumazhkin BK, Kolganova TV, Muyzer G, Sorokin DY (2013). Analysis of community composition of sulfur-oxidizing bacteria in hypersaline and soda lakes using *soxB* as a functional molecular marker. *FEMS Microbiol Ecol* **84:** 280-289.

van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B (2013). Arsenics as bioenergetic substrates. *Biochimica et biophysica acta* **1827:** 176-188.

Wallschlager D, Stadey CJ (2007). Determination of (oxy)thioarsenates in sulfidic waters. *Anal Chem* **79:** 3873-3880.

Ward BB, Martino DP, Diaz MC, Joye SB (2000). Analysis of ammonia-oxidizing bacteria from hypersaline Mono Lake, California, on the basis of 16S rRNA sequences. *Appl Environ Microbiol* **66:** 2873-2881.

Wilkin RT, Wallschlager D, Ford RG (2003). Speciation of arsenic in sulfidic waters. *Geochem Trans* **4:** 1-7.

Zargar K, Hoeft S, Oremland R, Saltikov CW (2010). Identification of a novel arsenite oxidase gene, *arxA*, in the haloalkaliphilic, arsenite-oxidizing bacterium *Alkalilimnicola ehrlichii* strain MLHE-1. *J Bacteriol* **192:** 3755-3762.

Zargar K, Conrad A, Bernick DL, Lowe TM, Stolc V, Hoeft S *et al.* (2012). ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. *Environ Microbiol* **14:** 1635-1645.

Table 1.1. Isolates and enrichment cultures from Mono Lake, CA that use arsenic and sulfur compounds (in bold)

Organism	Electron Donor	Electron Acceptor	Reference
Bacillus arsenicoselenatis	Lactate, malate, fructose, starch, citrate	Se(VI), **As(V)**, Fe(III), nitrate, fumarate	(Blum *et al.* 1998)
Bacillus selenitireducens	Lactate, glucose, pyruvate	Se(IV), **As(V)**, nitrate, nitrite, fumarate	(Blum *et al.* 1998)
Bacillus beveridgei	Organic carbon	Te(VI), Te(IV), Se(VI), Se(IV), **As(V)**, nitrate, nitrite, fumarate	(Baesman *et al.* 2009)
Deltaproteobacteria strain MLMS-1	**Sulfide**	**As(V)**	(Hoeft *et al.* 2004)
Ectothiorhodospira sp. strain PHS-1	**As(III)**	CO_2	(Kulp *et al.* 2008)
Ectothiorhodospira-dominated community	**As(III)**	CO_2	(Budinoff and Hollibaugh 2008)
Alkalilimnicola ehrlichii MLHE-1	**As(III)**, H_2, sulfide, acetate	Nitrate, O_2	(Hoeft *et al.* 2007, Oremland *et al.* 2002)
Desulfobulbus-dominated enrichment culture	**Sulfide**	**As(V)**	(Hollibaugh *et al.* 2006)
Thioalkalivibrio jannaschii	**Sulfide**	O_2	(Sorokin *et al.* 2002)
Thioalkalimicrobium cyclicum	**Sulfide**	O_2	(Sorokin *et al.* 2002)
Bacillales strain MLFW-2	Lactate	**As(V),** Sb(VI)	(Abin and Hollibaugh 2014), (Abin, unpublished results)
Desulfonatronum thiodismutans	H_2, formate, ethanol	**Sulfate, sulfite, thiosulfate**	(Pikuta *et al.* 2003)

CHAPTER 2

ANALYSIS OF THE MICROBIAL COMMUNITIES IN ALKALINE, HYPERSALINE, MONO LAKE, CALIFORNIA, USA THROUGH METATRANSCRIPTOMCS AND HIGH-THROUGHPUT 16S RIBOSOMAL RNA SEQUENCING[1]

[1]Edwardson C.F. and J.T. Hollibaugh. To be submitted to *Environmental Microbiology*.

Abstract

The vertical distribution of microorganisms in alkaline, hypersaline, Mono Lake, California, USA, was analyzed using 16S rRNA tag sequencing and taxonomy of metatranscriptomes. Samples were collected at five depths, capturing the major physicochemical zones in the lake during the first year of meromixis. Besides the picoeukaryotic algae *Picocystis*, the lake was dominated by bacteria from the phyla Proteobacteria, Firmicutes, and Bacteroidetes. Most (80%) of the abundant (>1% relative abundance) OTUs observed in the 16S amplicon survey had been found previously in Mono Lake. Metatranscriptomic taxonomic bins at the genus level were dominated by the Gammaproteobacteria *Thioalkalivibrio* (4-13%) and *Thioalkalimicrobium* (0-14%) and Firmicutes *Dethiobacter* (0-5%) and *Clostridium* (1-4%). Overall, these results showed that the Mono Lake microbial community remained stable through long-term meromixis and annual mixing events over 12 years and provided further insight into the communities that are transcriptionally active.

Introduction

Mono Lake, located to the east of the Sierra Nevada Mountains in the Great Basin in California, USA, has been studied extensively due to its unusual chemistry and biology (reviewed in (Oremland *et al.* 2004)). Briefly, Mono Lake is a closed-basin, hypersaline (~90 g/L), and alkaline (pH ~9.8) lake with geothermal inputs leading to an elevated arsenic concentration (~200 μM). The lake is generally monomictic, but undergoes periods of meromixis (Melack and Jellison 1998), leading to elevated levels of sulfide in the monimolimnion, which in turn leads to formation of thioarsenic compounds (Hollibaugh *et al.* 2005). The biogeochemistry of the lake has been well studied (Oremland *et al.* 2004), with contributions focused on the role of microorganisms in arsenic geochemistry. These studies have used isolates (Blum *et al.* 1998, Budinoff and Hollibaugh 2008, Fisher and Hollibaugh 2008, Hoeft *et al.* 2007, Sorokin *et al.* 2002), enrichment cultures (Edwardson *et al.* 2014, Fisher *et al.* 2008, Hollibaugh *et al.* 2006), and 16S rRNA and functional gene clone libraries (Giri *et al.* 2004, Humayoun *et al.* 2003, Scholten *et al.* 2005). Other aspects of the biogeochemistry of Mono Lake that have been studied include ammonia and methane oxidation (Joye *et al.* 1999, Lin *et al.* 2005, Ward *et al.* 2000), sulfur cycling (Hollibaugh *et al.* 2006, Scholten *et al.* 2005), chitin degradation (LeCleir *et al.* 2004), and carbon fixation (Giri *et al.* 2004). The composition of microbial communities from extreme environments has been studied using both next generation sequencing platforms (Harris *et al.* 2013, Lanzén *et al.* 2013, Podell *et al.* 2014, Rascovan *et al.* 2015, Schneider *et al.* 2013, Yelton *et al.* 2013) and clone libraries (Antony *et al.* 2013, Foti *et al.* 2008, Mesbah *et al.* 2007) despite limited sequencing depth for the latter. In addition to determining the taxonomic composition of the microbial community in an environment based on marker genes (e.g. 16S rRNA), determining the transcriptional (mRNA) abundance of the microbial communities allows

identification of organisms and pathways that may be more biogeochemically active. One of the first environmental metatranscriptomics studies (Poretsky *et al.* 2005) cloned and sequenced cDNA from Mono Lake. However, technological limitations restricted sequencing to 60 reads. In this study, we sampled five depths along a depth and redox gradient and sequenced amplified 16S rRNA genes and mRNA transcripts from each depth. The samples were collected during summer stratification following a winter when the lake did not mix fully, signaling the onset of meromixis. We were interested in the microbial communities present along a redox and depth gradient in the lake at the beginning of stratification and how they compared to the microbial communities found through clone library analysis determined after long-term meromixis (5 years) (Humayoun *et al.* 2003) and analysis of transcriptionally abundant (>1% relative abundance) microorganisms and their potential contributions to biogeochemical cycling in the lake.

Materials and Methods

Field Site and Sampling

Water samples were collected from Mono Lake in July 2012 at Station 6 (37°57.85' N, 119°01.32' W, surface elevation: 1945 m), the deepest station in the lake as described previously (Hollibaugh *et al.* 2005). Vertical profiles of temperature, pressure, conductivity, photosynthetically active radiation, turbidity, *in vivo* fluorescence and oxygen were obtained with a SeaBird SBE 19 Seacat CTD equipped with a C-Star transmissometer and WETstar fluorometer (Wet Labs), a Licor PAR sensor, and a SBE43 oxygen meter (SeaBird). Water samples for chemical analysis were collected as described previously (Hollibaugh *et al.* 2005). Samples for sulfide analysis were taken as described previously (Miller *et al.* 1993). Briefly, water was collected from discrete depths and injected into 8 mL, septa-capped vials containing

250 µL of 15% solution of zinc acetate and analyzed spectrophotometrically (Cline 1969). Samples for total arsenic and thioarsenic speciation were collected and analyzed by ICP-MS and IC-ICP-MS as described previously (Planer-Friedrich *et al.* 2007).

Nucleic Acid Sampling and Processing

Water samples for nucleic acid extraction were collected in duplicate from depths of 10, 15, 18, 25 and 31 m, chosen to capture microbial communities from the major redox zones of the lake. Water for DNA analysis was collected using a 5 L Niskin bottle. The sample was drained into two acid cleaned and sample-rinsed 1L polycarbonate bottles that was allowed to overflow approximately 1 volume before being capped without headspace using a septum cap. Samples were kept in a cooler with ice packs and filtered within 8 hours of collection. Water was filtered through Sterivex-GV 0.22 µm pore-size cartridge filters (Millipore) using a peristaltic pump. Lysis buffer (1.8 ml: 0.75M sucrose, 40 mM EDTA, 50 mM TRIS; pH 8.3) was added, and samples were frozen until extraction. DNA was extracted from the filters as described previously using a lysozyme-proteinase K followed by phenol-chloroform extraction (Ferrari and Hollibaugh 1999). RNA was collected by pumping water from each depth through Masterflex PharMed tubing (Cole-Parmer), which has low light and oxygen permeability, directly onto 142 mm diameter, 0.2µm pore-size Supor membrane filters (polyethersulfone, Pall Life Sciences). Filters were folded and placed in 15 mL polypropylene tubes then flash frozen in liquid nitrogen. Samples remained frozen until processed. RNA was extracted using a modification of a method described previously (Gifford *et al.* 2011). Briefly, the frozen filter was removed from the tube, transferred to a sterile Whirl-Pak bag, shattered with a mallet, and transferred immediately to a 50 mL centrifuge tube containing 8 mL RLT lysis buffer (Qiagen RNeasy Kit) and 3 g of low-binding 200 µm zirconium beads (OPS Diagnostics). Two internal standards (916 and 970 nt)

were added (3 x 10^{10} copies of each) to the tube, (Gifford *et al.* 2011, Satinsky *et al.* 2013, Satinsky *et al.* 2014), which was then vortexed at maximum speed for 10 minutes. Large particles were removed by low-speed centrifugation, the lysate was homogenized by drawing through a 21g syringe ~5-10 times, and RNA was extracted using the RNEasy Mini Kit (Qiagen). DNA was removed with Turbo DNA-free Kit (Life Technologies). Total bacterial abundance was determined by 16S rRNA gene quantitative PCR as previously described (Kalanetra *et al.* 2009).

16S Tag Pyrosequencing

The V4-V5 region of the 16S ribosomal RNA gene was amplified using 515F (Caporaso *et al.* 2011) and 907R (Armitage *et al.* 2012). The forward primer also contained the Roche LIB-L Adapter A and a 10 nt MID tag on the 5' end. The reverse primer contained the Roche LIB-L Adapter B on the 5' end. PCR was performed in triplicate for each sample with Q5® High-Fidelity polymerase (NEB). Each 25 µL reaction contained the following: Q5 Reaction Buffer (5 µL of 5X), 200 nM of each primer, 200 µM each dNTP and 0.02U/µL of Q5 polymerase, and 1 µL (12 – 61 ng) DNA template. The PCR program used was: 98 °C for 30 s, 25 cycles of 98 °C for 10 s, 60 °C for 10 s and 72 °C for 10 s, with a final step of 72 °C for 2 minutes. Samples were run on a 1% agarose gel and single bands of the expected length were cut and extracted using a QiaQuick Spin kit (Qiagen). The triplicates were pooled and an additional cleanup step was performed with a QiaQuick Spin kit to concentrate them. A final cleanup was performed using the AmpureXP kit (Beckman Genomics). Products were quantified using the PicoGreen kit (Life Technologies) and pooled at an equivalent weight. Each sample was sequenced at the Georgia Genomics Facility using 454 Titanium chemistry on approximately 1/60^{th} of a plate.

Metatranscriptomics Processing

Ribosomal RNA was depleted from the RNA pool as described previously (Stewart *et al.* 2010) using subtractive hybridization probes synthesized from Mono Lake DNA collected as described above. Probes were synthesized from PCR products amplified using the primers described (Stewart *et al.* 2010) but with 25 µL reactions of Q5® High-Fidelity polymerase (NEB) following the manufacturers recommended reaction conditions and a modified amplification protocol: 98 °C for 30 s, 30 cycles of 98 °C for 10s, annealing (Eub16S: 63 °C, Eub23S: 55 °C, Arch16S: 67 °C, Arch23S: 64 °C, Euk18S: 65 °C, Euk28S: 61 °C) for 30 s, 72 °C for 40, 60 or 80 s (16/18S, 23S and 28S reactions respectively), and 72 °C for 2 minutes. One PCR reaction was performed for each DNA replicate from each depth (n=10) and all were pooled and purified with a QiaQuick PCR Cleanup kit (Qiagen), and concentrated using a SpeedVac 120 (Savant), then a second round of purification was performed with the AmpureXP kit (Beckman Genomics) to remove remaining primers and primer dimers. The Arch16S PCR did not yield a usable product and thus was excluded from probe synthesis. Ribosomal RNA-depleted samples were amplified using the MessageAmpII-Bacteria kit (Ambion). Double stranded cDNA was prepared using a combination of first and second strand kits: SuperScript III First Strand Synthesis (Life Technologies) primed with random hexamers, and NEBNext mRNA second Strand synthesis module (NEB). Double stranded cDNA was purified with a PureLink PCR cleanup kit (Life Technologies) followed by ethanol precipitation. The cDNA was sheared (Covaris instrument) to a targeted ~225 bp insert size, and libraries were prepared using the TruSeq DNA (Illumina) kit with custom indices developed by the Georgia Genomics Facility at the University of Georgia. Samples were pooled and sequenced on one lane of an Illumina

HiSeq2500 in Rapid Run mode with the 150PE (300 cycle) kit at HudsonAlpha Genomic Services Lab (Huntsville, AL).

Bioinformatics – 16S rRNA Sequences

Sequences were processed using QIIME versions 1.5 (initial sample processing and de-noising) and 1.8 (OTU picking and taxonomic assignment) (Caporaso *et al.* 2010). Samples were split and filtered using default quality control settings, additionally with the removal of all reads with ambiguous bases (Huse *et al.* 2007) and all reads longer than 500 bp. Reads were de-noised using Denoiser (Reeder and Knight 2010) and checked for chimeras using USEARCH 6.1 *de novo* chimera picking (UCHIME) (Edgar 2010, Edgar *et al.* 2011). OTUs were picked using the open reference method with USEARCH 6.1 at a 97% identity cutoff. The SILVA rRNA database, release 111 (Quast *et al.* 2013), was used for reference-based OTU picking and for taxonomy assignment using UCLUST (Edgar 2010). Additional taxonomy assignment was performed for representative OTU sequences that were "unassigned." These OTUs were assigned using SINA (Pruesse *et al.* 2012) against SILVA release 119 or the RDP classifier (Wang *et al.* 2007). The QIIME taxonomic assignments were compared to the SINA taxonomic assignments against SILVA release 119 to improve taxonomy assignments for a number of sequences. Chloroplast, mitochondria, and singleton OTUs were removed. Representative sequences of each OTU were used to search against the NCBI nr database using BLASTN (Altschul *et al.* 1990). Sequences from a previous Mono Lake microbial diversity study (Humayoun *et al.* 2003) (and unpublished sequences from other Mono Lake stations and depths from the same study) were downloaded from NCBI Genbank (n=274). All sequences that spanned the region amplified by the 16S rRNA primers (515F/907R) were aligned in Geneious and trimmed to the length of the amplicons (~375 nt, n=174). These sequences were processed in

QIIME as described above (except singletons were retained (n=48)) to determine OTUs at 97% identity and define taxonomy in the same manner as the amplicon sequences. The number of clones belonging to each OTU was manually determined for each depth at Station 6 (2 m, 17.5 m, 23 m and 31 m); including clones not deposited in GenBank (see Humayoun et al. 2003). Relative abundances of each OTU were determined for each depth. Alpha and beta diversity analyses were performed on the 454 amplicons using the R package *phyloseq* (McMurdie and Holmes 2013).

Bioinformatics - Metatranscriptomics

FASTQ files were paired using PEAR (version 0.9.2) (Zhang *et al.* 2014) with a minimum overlap of 1 and no statistical tests. PRINSEQ (version 0.20.3) (Schmieder and Edwards 2011) was used to trim and perform quality control using the following parameters: 10-90% GC content, minimum length 35bp, mean quality score of 20, trim from 3' and 5' ends with a sliding scale window of 3 and step of 1 with a minimum mean quality score of 20, trim >5 uncalled bases from ends, trim >5 A/T tails, and only allow 1 uncalled base (any reads with uncalled bases were later removed). RiboPicker (0.4.3) (Schmieder *et al.* 2012) was used to identify rRNA reads. The default parameters (50% coverage, 75% identity, 30 base align length, BWA-SW Z-best value of 3) and the non-redundant rRNA database were used. FASTQ files were converted to FASTA files using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) default settings, which removes any reads that contain uncalled bases. A custom BLAST database was created using BLAST+ (Camacho *et al.* 2009), using sequences of the cDNA internal standards provided by B. Satinsky and M.A. Moran (Satinsky *et al.* 2014). A BLASTN (Altschul *et al.* 1990) search against the custom database was used to count reads assigned to internal standards. Recovery calculations were performed based on counts of hits with a bit score

>50 relative to the number of internal standard sequences added to each sample (Satinsky *et al.* 2013). These sequences were removed from the input FASTA files using a QIIME script (filter_fasta.py) (Caporaso *et al.* 2010). A local RapSearch2 (Zhao *et al.* 2012) database of all protein sequences in the RefSeq (Tatusova *et al.* 2014) database (release 64) was downloaded from NCBI. Rapsearch2 was used to annotate metatranscriptome reads using an e-value cutoff of 10^{-5}, only keeping the top hit. Further processing removed all hits with a bit score <40 (Gifford *et al.* 2011). Custom scripts were used to fully annotate and assign taxonomy to hits. Absolute abundances (transcripts/L) were calculated by multiplying read counts for each library by the multiplication factor determined from internal standard recovery divided by sample volume filtered. Relative abundance at each depth was determined by taking the average of the absolute abundance for the two transcript libraries per depth and dividing by the sum of these averages. In order to focus confidently on protein-encoding transcripts, RefSeq hits to proteins with annotations that contained the terms 'hypothetical' and 'putative' were removed. It should be noted that this likely removed hits to some transcripts that encode valid proteins whose function is either not known or annotation is incorrect, but it also removed incorrect annotations of non-protein encoding transcripts (Tripp *et al.* 2011). Disproportionally abundant hits to "cell wall hydrolases" were present in the dataset. In one case (library 31A), the top Bacteria hit was 11% of all bacterial transcripts, whereas the top non-hypothetical hit was only 2%. Analysis of each of the hits which made up >1% of hypothetical/cell wall hydrolases was performed through a TBLASTN search against the nr/nt nucleotide database with the best hit RefSeq protein sequence. This analysis (data not shown) revealed misannotated ribosomal RNAs, small RNAs or ribozymes, as had also been found previously (Tripp *et al.* 2011). Thus, reads were removed with the following annotations that were deemed to be inaccurate: "hydrolase" and anything with

"predicted protein," "uncharacterized protein" and "cell wall-associated hydrolase" in the annotation.

Phylogenetic Analysis

The phylogeny of the 16S rRNA sequences that were obtained from 454 sequencing was compared with the longer sequences obtained previously by Sanger sequencing (Humayoun *et al.* 2003). All OTUs from 454 sequencing with >1% relative abundance (n=35) were aligned with OTUs from Sanger sequencing with >1% relative abundance (n=60) and with 16S rRNA genes retrieved from genome sequences for taxa represented at >1% relative abundance in the metatranscriptomics dataset, as well as other reference 16S rRNA sequences. All sequences were trimmed to the length of the 454 reads (~375bp) and aligned in using the SINA aligner (Pruesse *et al.* 2012). Alignments were imported into Geneious 8 (Biomatters, LTD), and neighbor joining consensus trees (Jukes-Cantor distance) with 100 bootstrap replicates were built, with *Halobacterium salinarum* as outgroup. Three separate trees (Proteobacteria, Firmicutes, and other phyla) were constructed in this manner.

Results

Chemical Characteristics of Mono Lake

Mono Lake station 6, the deepest station and site of many of the previous microbiological studies of the lake, was sampled at five discrete depths based on the chemical profile of the lake at the time of sampling (Figure 2.1, Table 2.S1). The epilimnion is characterized by the highest temperatures (>15°C), highest light, highest dissolved oxygen, and is subject to intense grazing by brine shrimp, *Artemia monica* (Jellison and Melack 1993). The base of the epilimnion (10 m), the base of the oxycline (15 m), and near the base of the thermocline (18 m) was sampled. The dissolved oxygen level was 0.83 mg/L at 15 m, and decreased to the instrument's limit of

detection at 15.8 m (0.68 mg/L), thus the 15 m sample is considered to be suboxic. The anoxic hypolimnion was sampled at 25 m and 31 m.

Community Profiling by 16S rRNA Analysis

Between 4,137 and 16,913 high quality reads per depth were obtained (Table 2.1). Chloroplast sequences accounted for 5% to 57% of the reads (Table 2.1) and made up an increasing proportion of the reads as depth increased. Chloroplast sequences were dominated (>99%) by one OTU, which was 99% similar to the 16S rRNA gene sequence from *Picocystis salinarum* CCMP1897 chloroplasts and identical to a plastid sequence previously retrieved from Mono Lake (Humayoun *et al.* 2003).

A combined total of 243 OTUs in all samples from all depths was found. The distribution by depth of all 243 of these OTUs, with read counts, relative abundance, and full taxonomy is presented in Tables 2.S2 and 2.S3. The use of a variety of alpha diversity metrics revealed that alpha diversity was lowest at 10 m and highest at 18 m or 25 m, depending on the metric used (Figure 2.S1). Unweighted UniFrac analysis of beta diversity (Hamady *et al.* 2009) showed that the communities were structured by depth, with the 10 m and 31 m samples most dissimilar from the others, and the 15, 18 and 25 m samples similar to each other (Figure 2.S1). Weighted UniFrac, which takes into account relative abundances of OTUs, revealed a different relationship, with the 25 m and 31 m communities most similar to each other, and the community in the 10 m sample most dissimilar to the other communities (Figure 2.S1). The phyla Proteobacteria (26-39%), Bacteroidetes (28-39%), Firmicutes (1-26%), and Actinobacteria (4-15%) were abundant, and only one other phylum (Spirochaetes, 1-4%) made up more than 1% at any depth (Figure 2.2).

Forty-one of the 243 observed OTUs were present at all five depths. These core OTUs make up 61-74% of the overall microbial community at all depths. The most abundant (OTU 8, 5.4 – 19.8% relative abundance) was classified as a Bacteroidetes (class Cytophagia) related to clone ML602J-37 retrieved from Mono Lake in a previous study (Humayoun *et al.* 2003). Another Cytophagia (represented by clone ML310M-34) was also present in all depths (1.1-7.2%) but was more abundant in samples from below the oxycline. Five additional abundant (>1% relative abundance) core OTUs included two Actinobacteria (Microbacteriaceae, 1.7-4%; *Nitriliruptor*, 1.9-8.7%), an Alphaproteobacteria (unclassified Rhodobacteraceae, 4.1-6.7%), a Gammaproteobacteria (*Marinicella*, 1.3-5.6%), and *Spirochaeta* (1.2-2.5%).

One of the OTUs present at all depths and the most abundant at 31 m could not be classified by QIIME and had only 8 full length hits at 73-74% identity to the NCBI nr database using BLASTN. This OTU appeared to be a chimeric sequence or sequencing error, with a 54 nt region having 94% identity to clones in the nr database (Alphaproteobacteria), followed by a 72 nt region with no hits to the nr database, and a short region of 242 nt that was only 79-80% identical to clones in the nr database (Firmicutes). The presence of the same OTU in all samples indicates that it was not likely a PCR artifact, and full length 16S rRNA sequence analysis would be required to accurately classify this OTU.

Twenty-nine OTUs were present at all depths except 31 meters. The most abundant of these were members of the Gammaproteobacteria (*Spiribacter* and *Thiomicrospira*). The most abundant OTUs that were only found in samples of anoxic water (15-31 m) were related to Bacteroidales (the ML635J-40 aquatic group) previously identified in Mono Lake clone libraries (Humayoun *et al.* 2003). Thirty-five OTUs were present at >1% relative abundance at any depth and some of these (n=20) were present at all depths (Table 2.2). The closest representative

sequences in the NCBI nr database for all but four of these abundant OTUs (ML2012 OTU 80, 150, 151, and 166) were retrieved from Mono Lake or other soda lakes. In one case (OTU 80), the top hit to a sequence from Mono Lake was only slightly less similar than the top hit (99.5 vs 99.7%), from a deep subsurface bore-hole enrichment experiment. OTU 151 was only 89% similar to a 16S rRNA gene sequence from a Mongolian soda lake isolate (*Dethiobacter alkaliphilus* AHT1). Many of the other top hits were to 16S rRNA gene sequences from soda lakes (including Mono Lake) isolates. Additionally, 57% of the sequences from the closest cultured representatives were less than 95% similar to these OTUs, suggesting the presence of organisms representing previously undescribed genera (and in some cases higher taxonomic levels) in Mono Lake.

Transcriptionally Active Taxa Determined by Metatranscriptomics

In addition to the abundance of OTUs present in the 16S rRNA analysis, the transcriptionally active microbial community based on mRNA taxonomic affiliation was examined by sequencing duplicate cDNA libraries of mRNA collected in parallel with the DNA at each depth. Between 8 and 18 million paired reads per library were obtained, with an average read length of ~240 nt (Table 2.S4). On average, 68.4% of the metatranscriptome sequences were most closely related to Bacteria, 30.4% to Eukarya, 0.6% to Archaea, and 0.8% to Viruses. For the Bacteria, 45-59% of hits to RefSeq were hypothetical proteins or misannotations and subsequently removed from analysis (see Materials and Methods) (Table 2.S5). The phylum-level composition of the transcriptionally active microbial community was compared to the composition reflected by 16S rRNA genes in Figure 2.2. Metatranscriptome reads from Proteobacteria accounted for 43-61% of the transcriptionally active taxa in libraries from all depths. Firmicutes (7-32%) and Bacteroidetes (5-14%) were the next most transcriptionally

active phyla. Bacteroidetes transcripts were more abundant at 10 m than at other depths, and Firmicutes transcripts were abundant in the anoxic depths (15-31 m). Actinobacteria transcripts were abundant (14%) at 10 m but less abundant in the anoxic depths.

Because internal standards were used, overall bacterial transcript abundance (transcripts/L) could be compared among all depths. Transcript abundance ranged from ~0.25 to ~1.5 trillion transcripts per liter on average at each depth, increasing with depth (Figure 2.S3). Closer examination of the differences between the 16S OTUs and the transcriptome genome bin taxonomy relative abundance (Figure 2.3) show the largest differences within the Bacteroidetes and Proteobacteria phyla. In general, the Bacteroidetes are relatively more abundant in the 16S OTU dataset than in the metatranscriptome dataset. The opposite is true for Proteobacteria. The biggest differences at the class level are for the Cytophagia and Bacteroidia classes in the Bacteroidetes (more abundant in 16S OTUs) and in the Delta- and Gammaproteobacteria (more abundant in mRNA transcript bins).

The contribution of individual genera that contribute to >1% of the transcript pool at each depth was determined (Table 2.3). The genus-level bins that contained the most transcripts in samples from 10 m include *Spiribacter* (4%) and *Thioalkalivibrio* (4%), two Gammaproteobacteria in the Ectothiorhodospiraceae family. *Spiribacter* has been found previously in moderately halophilic environments and is a strict aerobe, which could indicate the reason for its presence only in oxic (10 m) and suboxic (15 m) depths in Mono Lake. *Thioalkalivibrio* has been isolated previously from soda lakes, including Mono Lake (Sorokin *et al.* 2001, Sorokin *et al.* 2002). Other transcriptionally abundant genera at 10 m include *Cyanobium* (3%) and *Synechococcus* (2%). A *Cyanobium* strain was previously isolated and characterized from Mono Lake, but cell counts showed that it was more abundant at aphotic

depths than at the surface (Budinoff and Hollibaugh 2007). However, the finding of more transcripts at 10 m would reinforce the hypothesis that the abundance of *Cyanobium* cells in the aphotic zone was due to inactive or sinking cells. Three phyla were only abundant at 10 m: Actinobacteria, Bacteroidetes and Verrucomicrobia. The top genomic bins for these phyla are not known to be haloalkaliphilic, but indicate the closest sequenced relative for these phyla.

Thioalkalivibrio was also the most transcriptionally abundant genus at all depths below 15 m and another Gammaproteobacteria haloalkaliphile (Sorokin *et al.* 2001), *Thioalkalimicrobium,* was the most abundant at 15 m (14% of the libraries), and combined with *Thioalkalivibrio*, these two Gammaproteobacteria accounted for 27% of the transcripts in these libraries. *Clostridium* species also appeared more abundant at 15 m and below, likely due to their anaerobic lifestyle, and at 18 m an increase of another genus in the order Clostridiales, *Dethiobacter*, was observed. This thiosulfate-reducing haloalkaliphile (Sorokin *et al.* 2008) was the third most abundant genus behind the chemolithotrophic sulfur oxidizers (*Thioalkalivibrio* and *Thioalkalimicrobium*, both 9%) abundant at 18 m. An increase in sulfate-reducing genera of the Deltaproteobacteria from 18 to 31 m was also detected. Obligate anaerobic Firmicutes and Spirochetes taxa increased in abundance at 25 and 31 m, and a decrease in the abundance of sulfur-oxidizing bacteria was seen at these depths as well. One interesting observation was the increasing presence of *Trichodesmium* transcripts (1 to 4% relative abundance) with depth. To analyze further, these transcripts were assembled using the Geneious assembler with default settings. The consensus sequences were searched with BLASTN against the Mono Lake *Picocystis* genome (C. Saltikov, unpublished). Approximately 90% of the hits were >97% identical and 25% (278/1180) were 100% identical to *Picocystis*. Further, these hits were to a single *Picocystis* consensus sequence of 87,248 bp. This sequence is 86% similar (nucleotide

identity) to the *Picocystis salinarum* chloroplast from San Francisco Bay (Lemieux *et al.* 2014), indicating these hits were likely *Picocystis* chloroplast sequences. Interestingly, a *Trichodesmium*-related Oscillatoriales species, *Phormidium*, which was identified in deep lake nitrogen-fixing aggregates in Mono Lake (Oremland 1990), was not detected in these samples.

Phylogenetic Analysis

The phylogenetic relationships and relative abundances of the 16S rRNA OTUs determined by 454 sequencing were compared to the 16S rRNA OTUs determined by Sanger clone library (Humayoun *et al.* 2003), with representative 16S rRNA sequences from both the metatranscriptome taxonomic genome bins (MTBs) as well as best hits to the NCBI nr database (Figures 2.4, 2.5, 2.6). Eleven of the 454 OTUs representative sequences were identical to the Humayoun clone OTU representative sequences. In general, the relative abundances were similar between the two datasets but in a few cases there were large discrepancies. The relative abundance of Actinobacteria OTU 83 was only 3% at 18 m, whereas the most closely related clone OTU made up almost 40% of all clone sequences at 18 m (17.5 m in the Humayoun dataset). In addition, there was a group of Firmicutes clone OTUs at 25 and 31 m with no similar 454 OTU. This group also has no closely related reference 16S rRNA sequence, with the closest being *Dethiobacter alkaliphilus* at ~90% identity. In fact all of the abundant Firmicutes (and Tenericutes, which grouped phylogenetically within the Firmicutes) 454 OTUs were only 87-96% similar to any 16S sequence in the NCBI 16S database. For the Proteobacteria (Figure 2.4, Table 2.2), only one 454 amplicon OTU was not closely related to a previously obtained clone. This OTU (ML2012 OTU 80) was most closely related to *Desulfonatronum thiodismutans* and *Desulfonatronum lacustre* isolates from Mono Lake and Lake Kivu, a Russian soda lake, respectively (Pikuta *et al.* 2003). Some of the OTUs that were obtained were most closely related

to clones from a sulfide-oxidizing, arsenate-reducing enrichment culture obtained from Mono Lake water (Hollibaugh *et al.* 2006) rather than to sequences retrieved directly from a water column sample. The other major groups identified consisted of Verrucomicrobia, Actinobacteria, Bacteroidetes (Flavobacteria, Chitinophaga), Spirochetes, Cyanobacteria, Planctomycetes, and an unclassified OTU (OTU 151) (Figure 2.6). None of the OTUs were related to the *Cyanobacteria* or *Chlorobium* MTBs that were present in the metatranscriptome. This could be due to bias in the primers used to amplify 16S rRNA genes.

Discussion

Unlike the previous study (Humayoun *et al.* 2003) a set of OTUs that spanned all depths in the lake was observed. In most cases, the relative abundance of these OTUs did not vary greatly between depths. In some cases the relative abundance of these OTUs was greater than the relative abundance of related MTBs. This could indicate presence of inactive or dormant cells (e.g. spores). Seven OTUs made up 23-45% of the community at all depths. These include two Cytophagia OTUs (12-21% relative abundance) most closely related (90-93% identity) to *Gracilimonas* (facultative aerobe) and *Balneola* (aerobic) species (Choi *et al.* 2009, Urios *et al.* 2008). An Alphaproteobacteria (Rhodobacteraceae) OTU (4-7%) was 99% similar to purple non-sulfur *Roseinatronobacter monicus* isolated from Mono Lake (Boldareva *et al.* 2007). This species is an obligate aerobe. Related *Rhodobaca* species are able to grow under anaerobic conditions if illuminated, thus the OTUs persistence just below the oxycline might be a consequence of photoheterotrophic "maintenance" or, alternatively, the organism might be associated with microoxic conditions created by the abundant oxygenic *Picocystis* found at this depth. However, it is unlikely that adequate irradiance is available in the highly reducing water below the chemocline. Recovery of this OTU from those depths may be due to the presence of

stationary or dead cells ("preserved nucleic acids"). An OTU related to *Nitriliruptor* species (Actinobacteria) contributed to 4-7% at all depths. The most closely related strain is the aerobic *Nitriliruptor alkaliphilus*, a haloalkaliphile isolated from soda lakes in isobutylnitrile-degrading enrichment cultures (Sorokin *et al.* 2009). Since the metatranscriptome was analyzed, the genome of *Nitriliruptor alkaliphilus* was released, and it is likely that if this genome had been included in the analysis there would be hits corresponding to this genome.

Due to the abundance and diversity of reduced sulfur compounds, one of the major questions in soda lake biogeochemistry is sulfur-cycling dynamics (Sorokin *et al.* 2011). Sulfur cycling microbial taxa in soda lakes are also among the most well characterized in terms of cultured isolates (Sorokin *et al.* 2013). As expected, OTUs representing many sulfur-oxidizing bacteria in both the 454 OTUs and MTBs at 15-31 meters with *Thioalkalimicrobium* (OTU 41) and *Thioalkalivibrio* (OTUs 10 and 50) made up 5-12% of the 454 OTUs and 24-12% of the MTBs. Below 15 m, *Thioalkalimicrobium* decreased but *Thioalkalivibrio* increased with depth. Both taxa should have decreased with depth, as these are aerobic sulfur oxidizers. Although some *Thioalkalivibrio* have the ability to use alternative electron acceptors such as nitrate (Sorokin *et al.* 2011), transcriptomic evidence (e.g. nitrate reductase) that would indicate that this was occurring was not found (see Chapter 3). The abundance of these two organisms as the likely dominant sulfur-oxidizing bacteria was not surprising, as *Thioalkalivibrio jannaschii* and *Thioalkalimicrobium cyclicum* were both isolated from Mono Lake (Sorokin 2002). A number of different sulfate-reducing bacteria appear between at 18 and 31 m (various Deltaproteobacteria and OTU 20). This was expected as it is known that sulfate-reducing activity is higher at depth in the lake, particularly when it is strongly stratified (Oremland *et al.* 2000). In addition, *Dethiobacter alkaliphilus*, the most abundant MTB in the Firmicutes phylum, is a sulfate-

reducing bacterium (Sorokin *et al.* 2008). Arsenic cycling, along with sulfur cycling is of interest in Mono Lake in particular due to findings from a number of previous studies. Further detail into the active microbial taxa related to sulfur and arsenic cycling are discussed in Chapter 3.

Previous work in Mono Lake looked at nitrogen cycling (Carini and Joye 2008) and observed peak ammonia oxidation rates at 12-14 m along with the presence of ammonia-oxidizing bacteria. In both the tag pyrosequencing and metatranscriptome datasets ammonia-oxidizing bacteria and archaea were at low abundance (<1% relative abundance). Moreover, little evidence was seen for the presence of ammonia-oxidizing archaea in the metatranscriptome, along with very low abundance of ammonia-oxidizing bacteria such as *Nitrosomonas*, and no OTUs representing known nitrifying bacteria in the 454 libraries. One possible explanation is the presence of methanotrophs (e.g. *Methylomicrobium* and *Methyloglobus*). It is known that some methane-oxidizers can also oxidize ammonia (Nyerges and Stein 2009).

Diversity relative to previous Mono Lake study and other soda lakes

Other recent studies have found similar bacteria and community structures in soda lakes, and closely related environments. Sequences where the closest hit was a Mono Lake clone sequence have also been recovered from other soda lakes (Lanzén *et al.* 2013, Wani *et al.* 2006). Clone sequences L1228J-1 (Firmicutes) and ML635J-40 (Bacteroidetes) were found in the interior of ikaites columns (pH >10, but permanently cold) found in the Ikka Fjord in Greenland. Ikaites are similar to calcium carbonate tufa that forms in alkaline environments. In addition, *Rhodobaca*, *Desulfonatronum*, *Thioalkalimicrobium*, *Thioalkalivibrio*, and *Tindallia* species were also found in these ikaites, which is an environment similar to classical soda lake environments (Glaring *et al.* 2015). The alpha diversity was similar to that found in the Humayoun *et al.* study, where community diversity increased with depth. However, a decrease in

community diversity at 31 m was seen. This may be due to the low sampling effort at 31 m relative to the other depths (Table 2.1). Similar changes in community diversity with depth were found in Lake Kivu, Africa, with diversity being greater in the anoxic regions of the lake (İnceoğlu *et al.* 2015). The microbial communities found in Soap Lake (Washington) also have a composition similar to Mono Lake and other soda lakes, with greater diversity in the deep sulfidic region (Dimitriu *et al.* 2008). A number of clone sequences that were abundant in the Humayoun *et al.* study were not represented by any 454 OTUs, but in general the Mono Lake microbial communities seem to have remained stable after mixing events. Differences could also be attributed to sequencing depth and primer biases.

Acknowledgements

We would like to thank Meredith Ross, Christopher Abin, Ron Oremland, Larry Miller, Sierra Nevada Aquatic Research Lab and Tom Crowe of Mono Lake Boat Tours for field support. We thank Britta Planer-Friedrich for analyzing samples for arsenic and thioarsenic speciation. We thank Brandon Satinsky for internal mRNA standards, Shalabh Sharma, and the Georgia Advanced Computing Resource Center at UGA for bioinformatics support and scripts.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215:** 403-410.

Antony CP, Shimpi GG, Cockell CS, Patole MS, Shouche YS (2013). Molecular characterization of prokaryotic communities associated with Lonar Crater basalts. *Geomicrobiol J* **31:** 519-528.

Armitage DW, Gallagher KL, Youngblut ND, Buckley DH, Zinder SH (2012). Millimeter-scale patterns of phylogenetic and trait diversity in a salt marsh microbial mat. *Frontiers in microbiology* **3**.

Blum JS, Bindi AB, Buzzelli J, Stolz JF, Oremland RS (1998). *Bacillus arsenicoselenatis*, sp. nov., and *Bacillus selenitireducens*, sp. nov.: two haloalkaliphiles from Mono Lake, California that respire oxyanions of selenium and arsenic. *Arch Microbiol* **171:** 19-30.

Boldareva EN, Bryantseva IA, Tsapin A, Nelson K, Sorokin DY, Tourova TP *et al.* (2007). The new alkaliphilic bacteriochlorophyll a-containing bacterium *Roseinatronobacter monicus* sp. nov. from the hypersaline soda Mono Lake (California, United States). *Microbiology* **76:** 82-92.

Budinoff CR, Hollibaugh JT (2007). Ecophysiology of a mono lake picocyanobacterium. *Limnol Oceanogr* **52:** 2484-2495.

Budinoff CR, Hollibaugh JT (2008). Arsenite-dependent photoautotrophy by an *Ectothiorhodospira*-dominated consortium. *ISME J* **2:** 340-343.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K *et al.* (2009). BLAST+: architecture and applications. *BMC Bioinformatics* **10:** 421.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7:** 335-336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* **108:** 4516-4522.

Carini SA, Joye SB (2008). Nitrification in Mono Lake, California: Activity and community composition during contrasting hydrological regimes. *Limnol Oceanogr* **53:** 2546-2557.

Choi DH, Zhang GI, Noh JH, Kim W-S, Cho BC (2009). *Gracilimonas tropica* gen. nov., sp. nov., isolated from a *Synechococcus* culture. *Int J Syst Evol Microbiol* **59:** 1167-1172.

Cline JD (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnol Oceanogr* **14:** 454-458.

Dimitriu PA, Pinkart HC, Peyton BM, Mormile MR (2008). Spatial and temporal patterns in the microbial diversity of a meromictic soda lake in Washington state. *Appl Environ Microbiol* **74:** 4877-4888.

Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460-2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Edwardson CF, Planer-Friedrich B, Hollibaugh JT (2014). Transformation of monothioarsenate by haloalkaliphilic, anoxygenic photosynthetic purple sulfur bacteria. *FEMS Microbiol Ecol* **90:** 858-868.

Ferrari VC, Hollibaugh JT (1999). Distribution of microbial assemblages in the Central Arctic Ocean Basin studied by PCR/DGGE: analysis of a large data set. *Hydrobiologia* **401:** 55-68.

Fisher JC, Hollibaugh JT (2008). Selenate-dependent anaerobic arsenite oxidation by a bacterium from Mono Lake, California. *Appl Environ Microbiol* **74:** 2588-2594.

Fisher JC, Wallschlager D, Planer-Friedrich B, Hollibaugh JT (2008). A new role for sulfur in arsenic cycling. *Environ Sci Technol* **42:** 81-85.

Foti M, Sorokin D, Zacharova E, Pimenov N, Kuenen JG, Muyzer G (2008). Bacterial diversity and activity along a salinity gradient in soda lakes of the Kulunda Steppe (Altai, Russia). *Extremophiles* **12:** 133-145.

Gifford SM, Sharma S, Rinta-Kanto JM, Moran MA (2011). Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J* **5:** 461-472.

Giri BJ, Bano N, Hollibaugh JT (2004). Distribution of RuBisCO genotypes along a redox gradient in Mono Lake, California. *Appl Environ Microbiol* **70:** 3443-3448.

Glaring MA, Vester JK, Lylloff JE, Abu Al-Soud W, Sørensen SJ, Stougaard P (2015). Microbial diversity in a permanently cold and alkaline environment in Greenland. *PLoS ONE* **10:** e0124863.

Hamady M, Lozupone C, Knight R (2009). Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J* **4:** 17-27.

Harris JK, Caporaso JG, Walker JJ, Spear JR, Gold NJ, Robertson CE *et al.* (2013). Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. *ISME J* **7:** 50-60.

Hoeft SE, Blum JS, Stolz JF, Tabita FR, Witte B, King GM *et al.* (2007). *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. *Int J Syst Evol Microbiol* **57:** 504-512.

Hollibaugh J, Carini S, Gurleyuk H, Jellison R, Joye S, Lecleir G *et al.* (2005). Arsenic speciation in Mono Lake, California: Response to seasonal stratification and anoxia. *Geochim Cosmochim Acta* **69:** 1925-1937.

Hollibaugh JT, Budinoff C, Hollibaugh RA, Ransom B, Bano N (2006). Sulfide oxidation coupled to arsenate reduction by a diverse microbial community in a soda lake. *Appl Environ Microbiol* **72:** 2043-2049.

Humayoun SB, Bano N, Hollibaugh JT (2003). Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl Environ Microbiol* **69:** 1030-1042.

Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome biology* **8:** R143.

İnceoğlu Ö, Llirós M, Crowe S, García-Armisen T, Morana C, Darchambeau F *et al.* (2015). Vertical distribution of functional potential and active microbial communities in meromictic Lake Kivu. *Microb Ecol* **70:** 596-611.

Jellison R, Melack JM (1993). Algal photsynthetic activity and its response to meromixis in hypersaline Mono Lake, California. *Limnol Oceanogr* **38:** 818-837.

Joye SB, Connell TL, Miller LG, Oremland RS, Jellison RS (1999). Oxidation of ammonia and methane in an alkaline, saline lake. *Limnol Oceanogr* **44:** 178-188.

Kalanetra KM, Bano N, Hollibaugh JT (2009). Ammonia-oxidizing archaea in the Arctic Ocean and Antarctic coastal waters. *Environ Microbiol* **11:** 2434-2445.

Lanzén A, Simachew A, Gessesse A, Chmolowska D, Jonassen I, Øvreås L (2013). Surprising prokaryotic and eukaryotic diversity, community structure and biogeography of Ethiopian soda lakes. *PLoS ONE* **8:** e72577.

LeCleir GR, Buchan A, Hollibaugh JT (2004). Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions. *Appl Environ Microbiol* **70:** 6977-6983.

Lemieux C, Otis C, Turmel M (2014). Six newly sequenced chloroplast genomes from prasinophyte green algae provide insights into the relationships among prasinophyte lineages and the diversity of streamlined genome architecture in picoplanktonic species. *BMC Genomics* **15:** 857.

Lin J-L, Joye SB, Scholten JCM, Schäfer H, McDonald IR, Murrell JC (2005). Analysis of methane monooxygenase genes in Mono Lake suggests that increased methane oxidation activity may correlate with a change in methanotroph community structure. *Appl Environ Microbiol* **71:** 6458-6462.

McMurdie PJ, Holmes S (2013). *phyloseq*: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8:** e61217.

Melack J, Jellison R (1998). Limnological conditions in Mono Lake: contrasting monomixis and meromixis in the 1990s. *Hydrobiologia* **384:** 21-39.

Mesbah N, Abou-El-Ela S, Wiegel J (2007). Novel and unexpected prokaryotic diversity in water and sediments of the alkaline, hypersaline lakes of the Wadi An Natrun, Egypt. *Microb Ecol* **54:** 598-617.

Miller LG, Jellison R, Oremland RS, Culbertson CW (1993). Meromixis in hypersaline Mono Lake, California. 3. Biogeochemical response to stratification and overturn. *Limnol Oceanogr* **38:** 1040-1051.

Nyerges G, Stein LY (2009). Ammonia cometabolism and product inhibition vary considerably among species of methanotrophic bacteria. *FEMS Microbiol Lett* **297:** 131-136.

Oremland RS (1990). Nitrogen fixation dynamics of two diazotrophic communities in Mono Lake, California. *Appl Environ Microbiol* **56:** 614-622.

Oremland RS, Dowdle PR, Hoeft S, Sharp JO, Schaefer JK, Miller LG *et al.* (2000). Bacterial dissimilatory reduction of arsenate and sulfate in meromictic Mono Lake, California. *Geochim Cosmochim Acta* **64:** 3073-3084.

Oremland RS, Stolz JF, Hollibaugh JT (2004). The microbial arsenic cycle in Mono Lake, California. *FEMS Microbiol Ecol* **48:** 15-27.

Pikuta EV, Hoover RB, Bej AK, Marsic D, Whitman WB, Cleland D *et al.* (2003). *Desulfonatronum thiodismutans* sp. nov., a novel alkaliphilic, sulfate-reducing bacterium capable of lithoautotrophic growth. *International Journal of Systematic and Evolutionary Microbiology* **53:** 1327-1332.

Planer-Friedrich B, London J, McCleskey RB, Nordstrom DK, Wallschlager D (2007). Thioarsenates in geothermal waters of Yellowstone National Park: Determination, preservation, and geochemical importance. *Environ Sci Technol* **41:** 5245-5251.

Podell S, Emerson JB, Jones CM, Ugalde JA, Welch S, Heidelberg KB *et al.* (2014). Seasonal fluctuations in ionic concentrations drive microbial succession in a hypersaline lake community. *ISME J* **8:** 979-990.

Poretsky RS, Bano N, Buchan A, LeCleir G, Kleikemper J, Pickering M *et al.* (2005). Analysis of microbial gene transcripts in environmental samples. *Appl Environ Microbiol* **71:** 4121-4126.

Pruesse E, Peplies J, Glockner FO (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28:** 1823-1829.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41:** D590-D596.

Rascovan N, Maldonado J, Vazquez MP, Eugenia Farias M (2015). Metagenomic study of red biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. *ISME J* **e-pub ahead of print 3 July 2015:** doi: 10.1038/ismej.2015.1109.

Reeder J, Knight R (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nat Meth* **7:** 668-669.

Satinsky BM, Gifford SM, Crump BC, Moran MA (2013). Use of internal standards for quantitative metatranscriptome and metagenome analysis. In: DeLong EF (ed). *Methods in Enzymology*. Academic Press. pp 237-250.

Satinsky BM, Crump BC, Smith CB, Sharma S, Zielinski BL, Doherty M *et al.* (2014). Microspatial gene expression patterns in the Amazon River plume. *Proc Natl Acad Sci USA* **111:** 11085-11090.

Schmieder R, Edwards R (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27:** 863-864.

Schmieder R, Lim YW, Edwards R (2012). Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* **28:** 433-435.

Schneider D, Arp G, Reimer A, Reitner J, Daniel R (2013). Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the Kiritimati Atoll, Central Pacific. *PLoS ONE* **8:** e66662.

Scholten JCM, Joye SB, Hollibaugh JT, Murrell JC (2005). Molecular analysis of the sulfate reducing and archaeal community in a meromictic soda lake (Mono Lake, California) by targeting 16S rRNA, *mcrA*, *apsA*, and *dsrAB* Genes. *Microb Ecol* **50:** 29-39.

Sorokin DY, Lysenko AM, Mityushina LL, Tourova TP, Jones BE, Rainey FA *et al.* (2001). *Thioalkalimicrobium aerophilum* gen. nov., sp. nov. and *Thioalkalimicrobium sibericum* sp. nov., and *Thioalkalivibrio versutus* gen. nov., sp. nov., *Thioalkalivibrio nitratis* sp.nov., novel and *Thioalkalivibrio denitrificans* sp. nov., novel obligately alkaliphilic and obligately chemolithoautotrophic sulfur-oxidizing bacteria from soda lakes. *Int J Syst Evol Microbiol* **51:** 565-580.

Sorokin DY (2002). Thioalkalimicrobium cyclicum sp. nov. and Thioalkalivibrio jannaschii sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing

bacteria from hypersaline alkaline Mono Lake (California). *International Journal of Systematic and Evolutionary Microbiology* **52:** 913-920.

Sorokin DY, Gorlenko VM, Tourova TP, Tsapin A, Nealson KH, Kuenen GJ (2002). *Thioalkalimicrobium cyclicum* sp. nov. and *Thioalkalivibrio jannaschii* sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing bacteria from hypersaline alkaline Mono Lake (California). *Int J Syst Evol Microbiol* **52:** 913-920.

Sorokin DY, Tourova TP, Mußmann M, Muyzer G (2008). *Dethiobacter alkaliphilus* gen. nov. sp. nov., and *Desulfurivibrio alkaliphilus* gen. nov. sp. nov.: two novel representatives of reductive sulfur cycle from soda lakes. *Extremophiles* **12:** 431-439.

Sorokin DY, van Pelt S, Tourova TP, Evtushenko LI (2009). *Nitriliruptor alkaliphilus* gen. nov., sp. nov., a deep-lineage haloalkaliphilic actinobacterium from soda lakes capable of growth on aliphatic nitriles, and proposal of *Nitriliruptoraceae* fam. nov. and *Nitriliruptorales* ord. nov. *Int J Syst Evol Microbiol* **59:** 248-253.

Sorokin DY, Kuenen JG, Muyzer G (2011). The microbial sulfur cycle at extremely haloalkaline conditions of soda lakes. *Frontiers in microbiology* **2:** 44.

Sorokin DY, Banciu H, Robertson LA, Kuenen JG, Muntyan MS, Muyzer G (2013). Halophilic and haloalkaliphilic sulfur-oxidizing bacteria. In: Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F (eds). *The Prokaryotes*. Springer Berlin Heidelberg. pp 529-554.

Stewart FJ, Ottesen EA, DeLong EF (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* **4:** 896-907.

Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I (2014). RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res* **42:** D553-D559.

Tripp HJ, Hewson I, Boyarsky S, Stuart JM, Zehr JP (2011). Misannotations of rRNA can now generate 90% false positive protein matches in metatranscriptomic studies. *Nucleic Acids Res* **39:** 8792-8802.

Urios L, Intertaglia L, Lesongeur F, Lebaron P (2008). *Balneola alkaliphila* sp. nov., a marine bacterium isolated from the Mediterranean Sea. *Int J Syst Evol Microbiol* **58:** 1288-1291.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73:** 5261-5267.

Wani AA, Surakasi VP, Siddharth J, Raghavan RG, Patole MS, Ranade D *et al.* (2006). Molecular analyses of microbial diversity associated with the Lonar soda lake in India: An impact crater in a basalt area. *Res Microbiol* **157:** 928-937.

Ward BB, Martino DP, Diaz MC, Joye SB (2000). Analysis of ammonia-oxidizing bacteria from hypersaline Mono Lake, California, on the basis of 16S rRNA sequences. *Appl Environ Microbiol* **66:** 2873-2881.

Yelton A, Comolli L, Justice N, Castelle C, Denef V, Thomas B *et al.* (2013). Comparative genomics in acid mine drainage biofilm communities reveals metabolic and structural differentiation of co-occurring archaea. *BMC Genomics* **14:** 1-15.

Zhang J, Kobert K, Flouri T, Stamatakis A (2014). PEAR: a fast and accurate Illumina paired-end read merger. *Bioinformatics* **30:** 614-620.

Zhao Y, Tang H, Ye Y (2012). RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* **28:** 125-126.

Table 2.1. Summary statistics for 16S rRNA tag pyrosequencing.

Read Statistics	Depth (m)				
	10	**15**	**18**	**25**	**31**
Reads Passing QC	16,913	8,433	9,978	8,003	4,137
Reads after Singletons, Chloroplast, Mitochondria Removed	15,555	6,180	6,261	3,811	1,718
Percent of Chloroplast Contribution	5	26	36	51	57
OTUs Observed (243 Total)	101	174	196	184	126

Table 2.2. Most abundant OTUs and top hits to NCBI databases. A database hit to a sequence retrieved from Mono Lake is only listed when it is not the top hit from the nr database (name includes "ML"). In other cases there were no similar database hits from Mono Lake. NR hits include sequence source location when retrieved from a hypersaline/alkaline environment. NCBI 16S hits include isolate source in parentheses where available.

OTU ID	RA (%) by Depth (m)					Phylum; Class; Genus (Unless Noted)	Mono Lake; NCBI nr; NCBI 16S	Accession	% ID
	10	15	18	25	31				
83	4	3	3	2	2	*Actinobacteria; Actinobacteria; Microbacteriaceae (F)*	*Actinobacterium* clone ML602M-15 *Microbacteriaceae bacterium* isolate DGGE gel band 16 *Pontimonas salivibrio* CL-TW6	AF448180 KP301133 NR_109611	99.5 99.5 98
3	9	4	3	2	2	*Actinobacteria; Nitriliruptoria; Nitriliruptor*	N/A *Actinobacterium* ML602J-33 *Nitriliruptor alkaliphilus* ANL-iso2	N/A AF486499 NR_044203	N/A 100 94
162	1	0	0	0	0	*Actinobacteria; Nitriliruptoria; Nitriliruptor*	Uncultured *Actinobacterium* clone ML615J-23 bacterium clone AN0C1CD05 (alkaline/saline soil) *Nitriliruptor alkaliphilus* ANL-iso2	AF448193 JQ426330 NR_044203	96 97.1 95
22	0	1	2	3	4	*Bacteroidetes; Bacteroidia; ML635J-40_aquatic_group*	N/A *Bacteroidetes* bacterium clone ML623J-22 *Parabacteroides goldsteinii* JCM 13446	N/A AF507858 NR_113076	N/A 100 87
214	0	0	1	1	1	*Bacteroidetes; Bacteroidia; ML635J-40_aquatic_group*	N/A *Sphingobacteria* bacterium clone A831 (Anderson Lake, OR) *Alistipes senegalensis* JC50	N/A EU283540 NR_118219	N/A 94.9 87
6	1	5	5	5	7	*Bacteroidetes; Cytophagia; ML310M-34*	Uncultured CFB group bacterium clone ML310M-18 DGGE gel band T1-1_Bac_b (Kulunda Steppe) *Balneola alkaliphila* CM41_14b	AF449772 EF622438 NR_044367	100 100 90
8	20	13	9	7	5	*Bacteroidetes; Cytophagia; ML602J-37*	Uncultured CFB group bacterium clone ML602M-3 clone 100307_0m_11H (Walker Lake, NV) *Gracilimonas mengyeensis* YIM J14	AF449782 KC358314 NR_109748	100 100 93
173	0	1	1	1	1	*Bacteroidetes; Cytophagia; ML602J-37*	*Cytophagia* bacterium enrichment culture clone EC8-10-1_F9 Bacterium YC-ZSS-LKJ111 (Yuncheng salt lake, China) *Gracilimonas tropica* CL-CB462	KM070869 KP174593 NR_044361	96.6 97.1 91
26	2	2	1	2	1	*Bacteroidetes; Flavobacteria; Brumimicrobium (F: Cryomorphaceae)*	N/A *Bacteroidetes* bacterium clone ML617.5J-5 *Brumimicrobium mesophilum* YH207	N/A AF507867 NR_115845	N/A 97.8 95
90	1	1	1	1	0	*Bacteroidetes; Flavobacteria; Psychroflexus (F: Flavobacteriaceae)*	Uncultured CFB group bacterium clone ML110M-1 *Psychroflexus lacisalsi* H7 (Hunazoko-Ike, Antarctica) *Psychroflexus lacisalsi* H7	AF452596 NR_112778 NR_112778	97.3 98.1 98
9	0	1	1	1	1	*Bacteroidetes; Sphingobacteriia; Saprospiraceae (F)*	N/A CFB group bacterium clone ML1218M-14 *Lewinella nigricans* ATCC 23147	N/A AF452599 NR_115013	N/A 98.9 90
7	6	3	2	1	1	*Bacteroidetes; Unlassified Bacteroidetes; ML602M-17*	N/A CFB group bacterium clone ML110J-35 *Owenweeksia hongkongensis* UST20020801	N/A AF452592 NR_040990	N/A 100 90
155	0	0	1	1	1	*Firmicutes; Bacilli; Paenibacillaceae (F)*	N/A low G+C Gram-positive bacterium clone ML-AsS-18	N/A DQ206409	N/A 98.8

							Desulfuribacillus alkaliarsenatis AHT28	NR_126221	94
36	0	2	6	5	4	*Firmicutes; Clostridia; Dethiobacter (F: Syntrophomonadaceae)*	N/A low G+C Gram-positive bacterium clone ML-AsS-4 *Dethiobacter alkaliphilus* AHT 1	N/A DQ206410 NR_044205	N/A 99.4 92
166	0	0	0	0	2	*Firmicutes; Clostridia; OPB54*	N/A Bacterium clone DGGE Band-4 *Dethiobacter alkaliphilus* AHT 1	N/A AB525845 NR_044205	N/A 95.7 89
150	0	5	6	7	7	*Firmicutes; Clostridia; Ruminococcaceae (F)*	N/A Bacterium isolate d21l12b41 *Anaerobacterium chartisolvens* T-1-35	N/A FR687166 NR_125464	N/A 96.5 96
28	0	2	5	6	0	*Firmicutes; Clostridia; Syntrophomonadaceae (F)*	N/A low G+C Gram-positive bacterium clone ML623J-3 *Dethiobacter alkaliphilus* AHT 1	N/A AF507876 NR_044205	N/A 99.2 91
194	0	0	0	0	5	*Firmicutes; Clostridia; Syntrophomonadaceae (F)*	N/A low G+C Gram-positive bacterium clone ML623J-3 *Dethiobacter alkaliphilus* AHT 1	N/A AF507876 NR_044205	N/A 98.7 90
64	0	0	1	1	0	*Planctomycetes; Phycisphaerae; ML-A-10*	N/A *Planctomycete* clone ML-A-10 *Phycisphaera mikurensis* NBRC 102666	N/A DQ206406 NR_074491	N/A 99.2 83
81	4	7	5	4	5	*Proteobacteria; Alphaproteobacteria; Rhodobacteraceae (F)*	Uncultured *alpha proteobacterium* clone ML623J-49 Clone S22B33 (Badain Jaran, China) *Roseinatronobacter monicus* ROS 35 (Mono Lake)	AF507835 KJ733773 NR_043914	100 100 99
95	3	1	1	1	0	*Proteobacteria; Betaproteobacteria; GKS98_freshwater_group (F: Alcaligenaceae)*	Uncultured *beta proteobacterium* clone ML617.5J-48 100307_24m_05A (Walker Lake, NV) *Kerstersia gyiorum* LMG 5906	AF507838 KC358294 NR_025669	100 100 98
80	0	0	0	0	1	*Proteobacteria; Deltaproteobacteria; Desulfonatronum (F: Desulfonatronaceae)*	*Delta proteobacterium* ML-1 Clone 1301APY_F05 *Desulfonatronum thiodismutans* MLF1 (Mono Lake)	HQ595725 NR_041848 NR_025163	99.5 99.7 99
20	0	1	2	2	3	*Proteobacteria; Deltaproteobacteria; Desulfurivibrio (F: Desulfobulbaceae)*	N/A *Delta proteobacterium* clone ML-A-1 *Desulfurivibrio alkaliphilus* AHT2	N/A DQ206405 NR_074971	N/A 100 96
11	6	3	2	1	2	*Proteobacteria; Gammaproteobacteria; Marinicella*	N/A *Gamma proteobacterium* clone ML623J-52 *Ectothiorhodospira salini* JA430	N/A AF507824 NR_104503	N/A 100 92
13	2	2	1	1	0	*Proteobacteria; Gammaproteobacteria; ML617.5J-3 (O: Oceanospirillales)*	N/A *Gamma proteobacterium* clone ML110J-5 *Halochromatium salexigens* DSM 4395	N/A AF452603 NR_036810	N/A 100 92
12	14	8	4	3	0	*Proteobacteria; Gammaproteobacteria; Spiribacter (F: Ectothiorhodospiraceae)*	N/A *Gamma proteobacterium* clone ML110J-38 *Arhodomonas recens* RS91	N/A AF452602 NR_118045	N/A 100 95
169	0	0	0	0	2	*Proteobacteria; Gammaproteobacteria; Spiribacter (F: Ectothiorhodospiraceae)*	N/A *Gamma proteobacterium* clone ML110J-38 *Spiribacter salinus* M19-40	N/A AF452602	N/A 99.7 94
10	0	3	2	4	5	*Proteobacteria; Gammaproteobacteria; Thioalkalivibrio (F: Ectothiorhodospiraceae)*	N/A *Gamma proteobacterium* clone ML110M-3 *Thioalkalivibrio nitratireducens* DSM 14787	N/A AF452604 NR_102486	N/A 100 98
50	1	0	0	0	0	*Proteobacteria; Gammaproteobacteria; Thioalkalivibrio (F: Ectothiorhodospiraceae)*	N/A *Gamma proteobacterium* clone ML615J-17 *Thioalkalivibrio versutus* DSM 13738	N/A AF453541 NR_117125	N/A 100 99

41	0	9	7	3	0	*Proteobacteria; Gammaproteobacteria; Thiomicrospira (F: Piscirickettsiaceae)*	*Gamma proteobacterium* clone ML617.5J-13 *Gamma proteobacterium* ASL8.1 (Soap Lake , WA) *Thioalkalimicrobium cyclicum* ALM1 (Mono Lake)	AF507825 DQ900622 NR_074720	99.5 99.7 99.5
176	0	0	0	0	2	*Proteobacteria; Gammaproteobacteria; Thiomicrospira (F: Piscirickettsiaceae)*	Uncultured *gamma proteobacterium* clone ML617.5J-13 *Gamma proteobacterium* ASL8.1 (Soap Lake, WA) *Thioalkalimicrobium cyclicum* ALM1 (Mono Lake)	AF507825 DQ900622 NR_074720	98.7 98.9 98
21	2	1	2	2	3	*Spirochaetes; Spirochaetes; Spirochaeta*	*Spirochaetaceae* bacterium clone ML623J-21 Clone 091808_0m_08C (Walker Lake, NV) *Spirochaeta alkalica* Z-7491	AF507856 KC358215 NR_026301	100 100 97
16	13	3	2	1	1	*Tenericutes; Mollicutes; NB1-n*	N/A *Bacillus sp.* clone ML615J-28 *Bacillus berkeleyi* KMM 6244	N/A AF454301 NR_109459	N/A 100 87
158	0	0	0	1	0	*Tenericutes; Mollicutes; RF9*	N/A *Firmicutes* bacterium A1874 (Anderson Lake, OR) *Acholeplasma morum* 72-043	N/A EU283544 NR_042959	N/A 97 86
151	0	2	4	6	8	*Unclassified Bacteria; Unlassified Bacteria; Unlassified Bacteria*	N/A *Alpha proteobacterium* clone Ola1.G05.invm13r *Piscibacillus salipiscarius* RBU1-1	N/A AB691215 NR_041260	N/A 85.3 80

RA%=relative abundance (rounded), %ID = % nucleotide identity

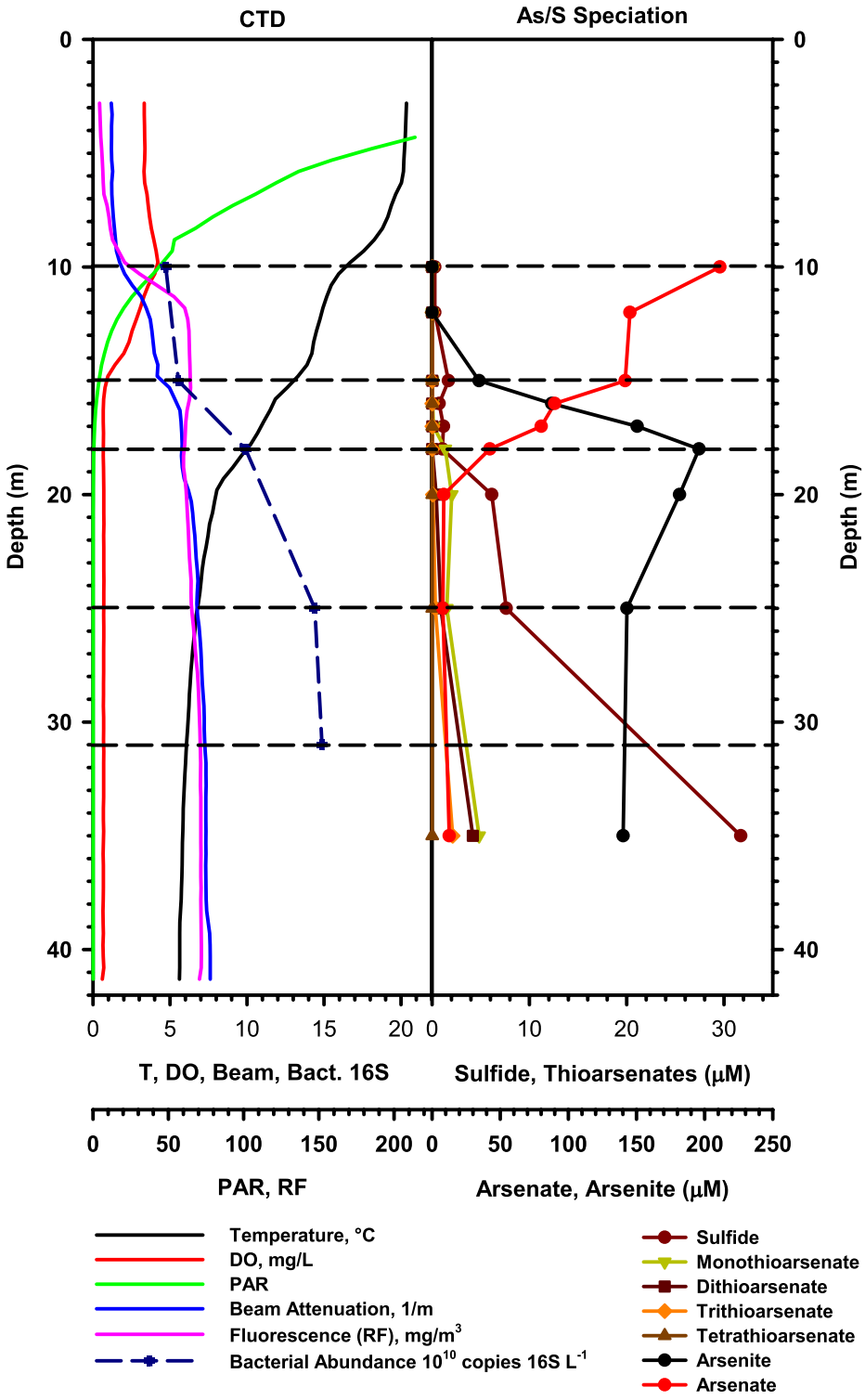
Table 2.3. Genera >1% relative abundance (rounded) at any depth with classification and top genome bins.

Genus	10 m	15 m	18 m	25 m	31 m	Phylum	Class; Order; Family	Dominant Bin (or # of bins)
Ilumatobacter	1%	0%	0%	0%	0%	Actinobacteria	Actinobacteria; Acidimicrobiales; Acidimicrobiaceae	*Ilumatobacter coccineum* YM16-304
Leifsonia	1%	0%	0%	0%	0%	Actinobacteria	Actinobacteria; Actinomycetales; Microbacteriaceae	5
Streptomyces	1%	0%	0%	0%	0%	Actinobacteria	Actinobacteria; Actinomycetales; Streptomycetaceae	>10
Anaerophaga	0%	0%	0%	0%	1%	Bacteroidetes	Bacteroidia; Bacteroidales; Marinilabiliaceae	*Anaerophaga thermohalophila* *Anaerophaga* sp. HS1
Fluviicola	1%	0%	0%	0%	0%	Bacteroidetes	Flavobacteriia; Flavobacteriales; Cryomorphaceae	*Fluviicola taffensis* DSM 16823
Owenweeksia	1%	0%	0%	0%	0%	Bacteroidetes	Flavobacteriia; Flavobacteriales; Cryomorphaceae	*Owenweeksia hongkongensis* DSM 17368
Flavobacterium	1%	0%	0%	0%	0%	Bacteroidetes	Flavobacteriia; Flavobacteriales; Flavobacteriaceae	>10
Psychroflexus	1%	0%	0%	0%	0%	Bacteroidetes	Flavobacteriia; Flavobacteriales; Flavobacteriaceae	*Psychroflexus gondawanenis* *Psychroflexus torquis* *Psychroflexus tropicus*
Pelodictyon	0%	0%	0%	0%	1%	Chlorobi	Chlorobia; Chlorobiales; Chlorobiaceae	*Chlorobium luteolum* DSM 273 *Pelodictyon phaeoclathratiforme* BU-1
Cyanobium	3%	0%	0%	0%	0%	Cyanobacteria	unclassified; Chroococcales; unclassified	*Cyanobium* sp. PCC 7001 Cyanobium gracile PCC 6307
Synechococcus	2%	0%	0%	0%	0%	Cyanobacteria	unclassified; Chroococcales; unclassified	>10
Trichodesmium	1%	1%	1%	3%	4%	Cyanobacteria	unclassified; Oscillatoriales; unclassified	*Trichodesmium erythraeum* IMS101
Bacillus	1%	1%	2%	1%	1%	Firmicutes	Bacilli; Bacillales; Bacillaceae	>10 including *Bacillus thuringiensis* *Bacillus selenitireducens* MLS10 *Bacillus smithii* *Bacillus cellulosilyticus*
Paenibacillus	0%	0%	1%	1%	1%	Firmicutes	Bacilli; Bacillales; Paenibacillaceae	>10
Staphylococcus	0%	0%	0%	0%	1%	Firmicutes	Bacilli; Bacillales; Staphylococcaceae	*Staphylococcus hominis* and others
Enterococcus	1%	1%	1%	0%	0%	Firmicutes	Bacilli; Lactobacillales; Enterococcaceae	Enterococcus faecalis *Enterococcus faecium* and others
Alkaliphilus	0%	0%	1%	1%	1%	Firmicutes	Clostridia; Clostridiales; Clostridiaceae	*Alkaliphilus metalliredigens* QYMF *Alkaliphilus oremlandii* OhILAs
Clostridium	1%	4%	5%	4%	4%	Firmicutes	Clostridia; Clostridiales; Clostridiaceae	*Clostridium difficile* *Clostridium thermocellum* *Clostridium clariflavum* *Clostridium termitidis* *Clostridium papyrosolvens*
Desulfosporosinus	0%	0%	1%	1%	1%	Firmicutes	Clostridia; Clostridiales; Peptococcaceae	*Desulfosporosinus orientis* DSM 765 and 4 others
Desulfotomaculum	0%	1%	2%	2%	2%	Firmicutes	Clostridia; Clostridiales; Peptococcaceae	7
Acetivibrio	0%	1%	0%	0%	0%	Firmicutes	Clostridia; Clostridiales; Ruminococcaceae	*Acetivibrio cellulolyticus*

Dethiobacter	0%	2%	5%	5%	5%	Firmicutes	Clostridia; Clostridiales; Syntrophomonadaceae	*Dethiobacter alkaliphilus*
Halanaerobium	0%	0%	1%	1%	1%	Firmicutes	Clostridia; Halanaerobiales; Halanaerobiaceae	*Halanaerobium hydrogeniformans* *Halanaerobium saccharolyticum* *Halanaerobium prevalens* DSM 2228
Halothermothrix	0%	0%	1%	1%	1%	Firmicutes	Clostridia; Halanaerobiales; Halanaerobiaceae	*Halothermothrix orenii* H168
Acetohalobium	0%	0%	1%	1%	1%	Firmicutes	Clostridia; Halanaerobiales; Halobacteroidaceae	*Acetohalobium arabaticum* DSM 5501
Natranaerobius	0%	0%	1%	1%	1%	Firmicutes	Clostridia; Natranaerobiales; Natranaerobiaceae	*Natranaerobius thermophilus*
Rhodopirellula	1%	0%	0%	0%	0%	Planctomycetes	Planctomycetia; Planctomycetales; Planctomycetaceae	6
Paracoccus	0%	1%	0%	0%	0%	Proteobacteria	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	5
Rhodobacter	1%	1%	1%	1%	1%	Proteobacteria	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Rhodobacter sphaeroides* *Rhodobacter capsulatus* *Rhodobacter* sp. CACIA14H1
Roseobacter	1%	1%	0%	0%	0%	Proteobacteria	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	Roseobacter sp. AzwK-3b and 6 other species
Roseovarius	1%	1%	0%	0%	0%	Proteobacteria	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	*Roseovarius* sp. 217 *Roseovarius* sp. TM1035 *Roseovarius nubinhibens*
Ruegeria	0%	1%	0%	0%	0%	Proteobacteria	Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae	7
Desulfatibacillum	0%	0%	0%	1%	1%	Proteobacteria	Deltaproteobacteria; Desulfobacterales; Desulfobacteraceae	*Desulfatibacillum alkenivorans* AK-01
Desulfococcus	0%	0%	0%	1%	2%	Proteobacteria	Deltaproteobacteria; Desulfobacterales; Desulfobacteraceae	*Desulfatibacillum multivorans* *Desulfatibacillum oleovorans* Hxd3
Desulfurivibrio	0%	1%	2%	0%	0%	Proteobacteria	Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae	*Desulfurivibrio alkaliphilus* AHT2
Desulfonatronospira	0%	0%	0%	1%	1%	Proteobacteria	Deltaproteobacteria; Desulfovibrionales; Desulfohalobiaceae	*Desulfonatronospira thiodismutans*
Desulfovibrio	0%	0%	1%	1%	2%	Proteobacteria	Deltaproteobacteria; Desulfovibrionales; Desulfovibrionaceae	>10
Geobacter	0%	0%	1%	0%	1%	Proteobacteria	Deltaproteobacteria; Desulfuromonadales; Geobacteraceae	8
delta proteobacterium	0%	1%	2%	1%	1%	Proteobacteria	Deltaproteobacteria; unclassified Deltaproteobacteria; unclassified Deltaproteobacteria	*deltaproteobacterium* MLMS-1 *deltaproteobacterium* NaphS2
Marinobacter	1%	0%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Alteromonadales; Alteromonadaceae	10
Alkalilimnicola	1%	0%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae	*Alkalilimnicola ehrlichii* MLHE-1
Spiribacter	4%	1%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae	*Spiribacter salinus* M19-40

Thioalkalivibrio	4%	13%	9%	12%	9%	Proteobacteria	Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae	*Thioalkalivibrio nitratireducens* *Thioalkalivibrio sulfidophilus* HL-EbGr7 *Thioalkalivibrio* sp. K90mix *Thioalkalivibrio thiocyanodenitrificans* and others
Escherichia	2%	2%	2%	1%	1%	Proteobacteria	Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae	*Escherichia coli* strains
Methylomicrobium	0%	1%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Methylococcales; Methylococcaceae	*Methylomicrobium alcaliphilum* 20Z *Methylomicrobium buryatense* *Methylomicrobium album*
Halomonas	2%	1%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Oceanospirillales; Halomonadaceae	14
Pseudomonas	1%	1%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae	>10
Thioalkalimicrobium	0%	14%	9%	6%	5%	Proteobacteria	Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae	*Thioalkalimicrobium cyclicum* ALM1
Thiomicrospira	0%	4%	3%	2%	1%	Proteobacteria	Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae	*Thiomicrospira crunogena* XCL-2 *Thiomicrospira arctica* *Thiomicrospira halophila*
gamma proteobacterium	1%	0%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria	9
Vibrio	1%	1%	0%	0%	0%	Proteobacteria	Gammaproteobacteria; Vibrionales; Vibrionaceae	*Vibrio parahaemolyticus* and others
Sphaerochaeta	0%	0%	1%	1%	1%	Spirochaetes	Spirochaetia; Spirochaetales; Spirochaetaceae	*Sphaerochaeta pleomorpha* str. Grapes
Spirochaeta	1%	0%	1%	1%	2%	Spirochaetes	Spirochaetia; Spirochaetales; Spirochaetaceae	*Spirochaeta africana* DSM 8902 *Spirochaeta alkalica* *Spirochaeta smaragdinae* DSM 11293
Opitutaceae bacterium	2%	0%	0%	0%	0%	Verrucomicrobia	Opitutae; Opitutales; Opitutaceae	*Opitutaceae* bacterium TAV1
Opitutus	2%	0%	0%	0%	0%	Verrucomicrobia	Opitutae; Opitutales; Opitutaceae	*Opitutus terrae* PB90-1
Coraliomargarita	2%	1%	0%	0%	0%	Verrucomicrobia	Opitutae; Puniceicoccales; Puniceicoccaceae	*Coraliomargarita akajimensis* DSM 45221
Verrucomicrobiae bacterium	2%	0%	0%	0%	0%	Verrucomicrobia	Verrucomicrobiae; Verrucomicrobiales; unclassified Verrucomicrobiales	*Verrucomicrobiae* bacterium DG1235

Figure 2.1. CTD and chemical profiles of Mono Lake Station 6. Sampling depths are indicated by horizontal dashed lines. Abbreviations: As, arsenic; S, sulfur; DO, dissolved oxygen; PAR, photosynthetically active radiation; RF, relative fluorescence; 16S, 16S ribosomal rRNA gene; Bact., Bacteria.


CTD
As/S Speciation
Depth (m)
0
10
20
30
40
0 5 10 15 20
T, DO, Beam, Bact. 16S
0 10 20 30
Sulfide, Thioarsenates (µM)
0 50 100 150 200
PAR, RF
0 50 100 150 200 250
Arsenate, Arsenite (µM)
Temperature, °C
DO, mg/L
PAR
Beam Attenuation, 1/m
Fluorescence (RF), mg/m3
Bacterial Abundance 10^10 copies 16S L^-1
Sulfide
Monothioarsenate
Dithioarsenate
Trithioarsenate
Tetrathioarsenate
Arsenite
Arsenate

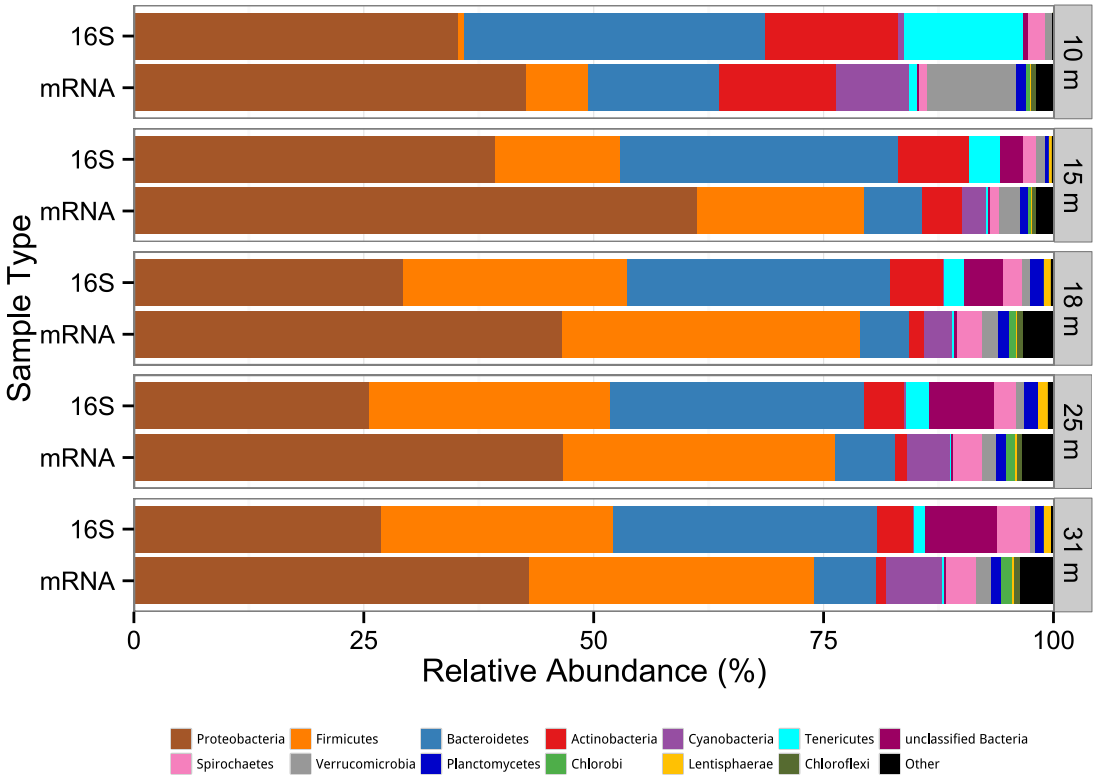


Figure 2.2. Phylum-level comparison of taxonomic affiliation of 16S rRNA gene and mRNA transcript relative abundance by depth.

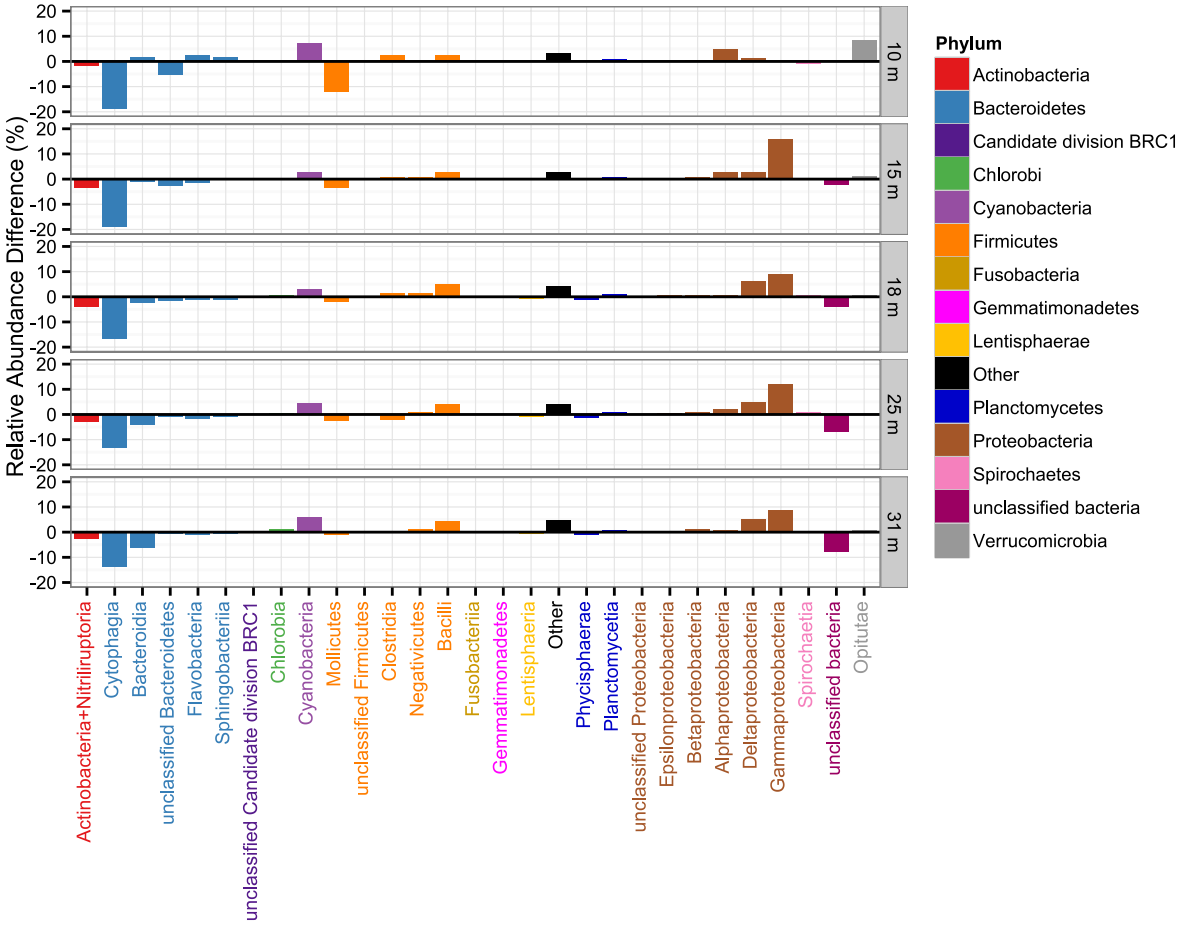


Figure 2.3. Difference in relative abundance between mRNA transcript bins and 16S rRNA OTUs by class. The relative abundance of 16S OTUs at the class level was subtracted from the relative abundance of mRNA class-level bins. Positive bars indicate greater relative abundance in class-level transcript bins. Negative bars indicate greater relative abundance in 16S OTUs at the class level. Class names are colored by phylum.

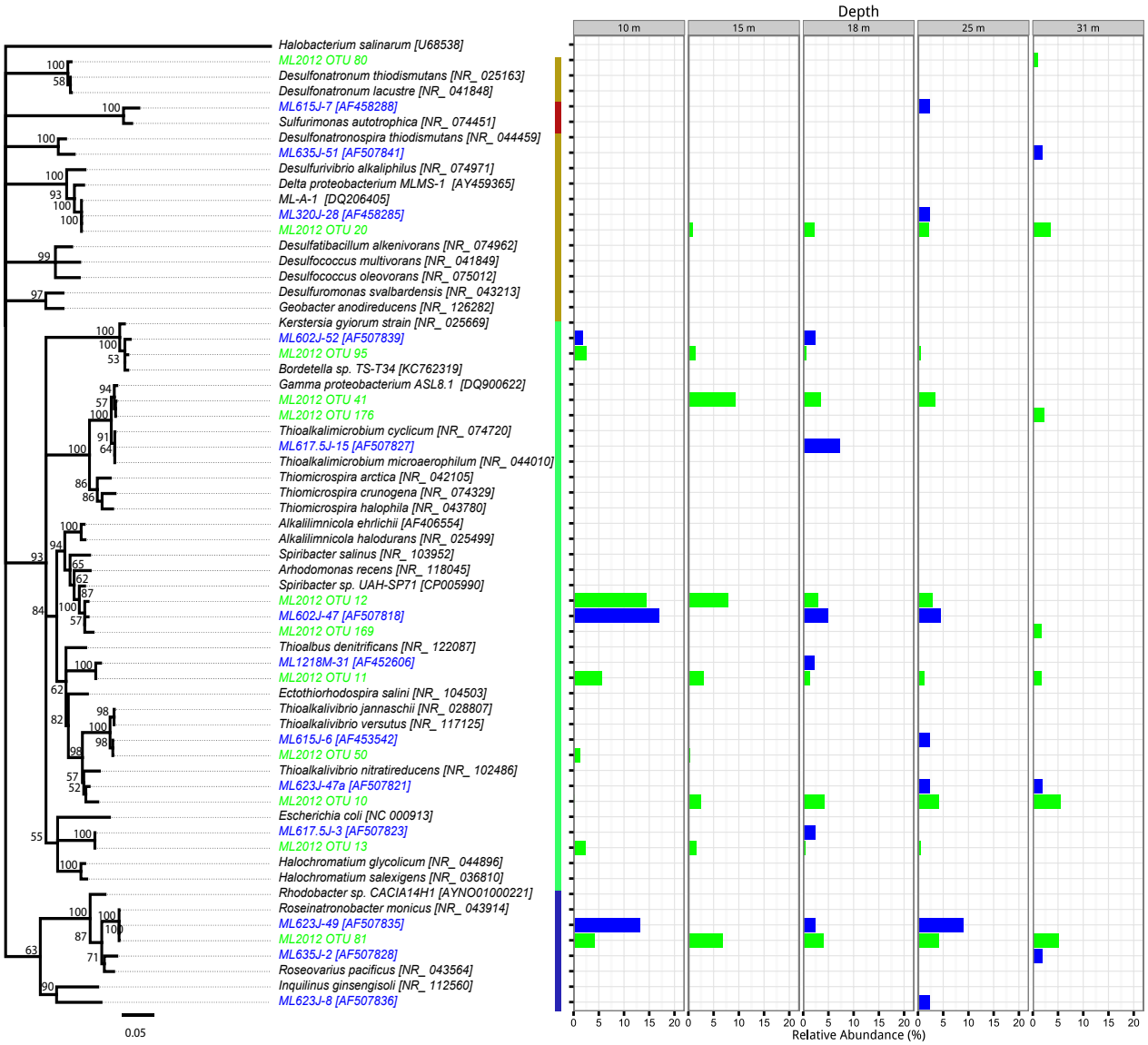


Figure 2.4. Proteobacteria abundance by depth. Green taxa are tag pyrosequencing OTU representative sequences. Blue taxa are Mono Lake clone OTU representative sequences from Humayoun *et al.*, 2003. Black sequences are metatranscriptome taxonomic bin 16S reference sequences (>1% relative abundance in metatranscriptome) and best hits to NCBI 16S rRNA database. Proteobacteria classes are colored by vertical bar as follows: Delta, gold; Epsilon, red; Gamma,green; Alpha, blue. The outgroup is *Halobacterium salinarum*.

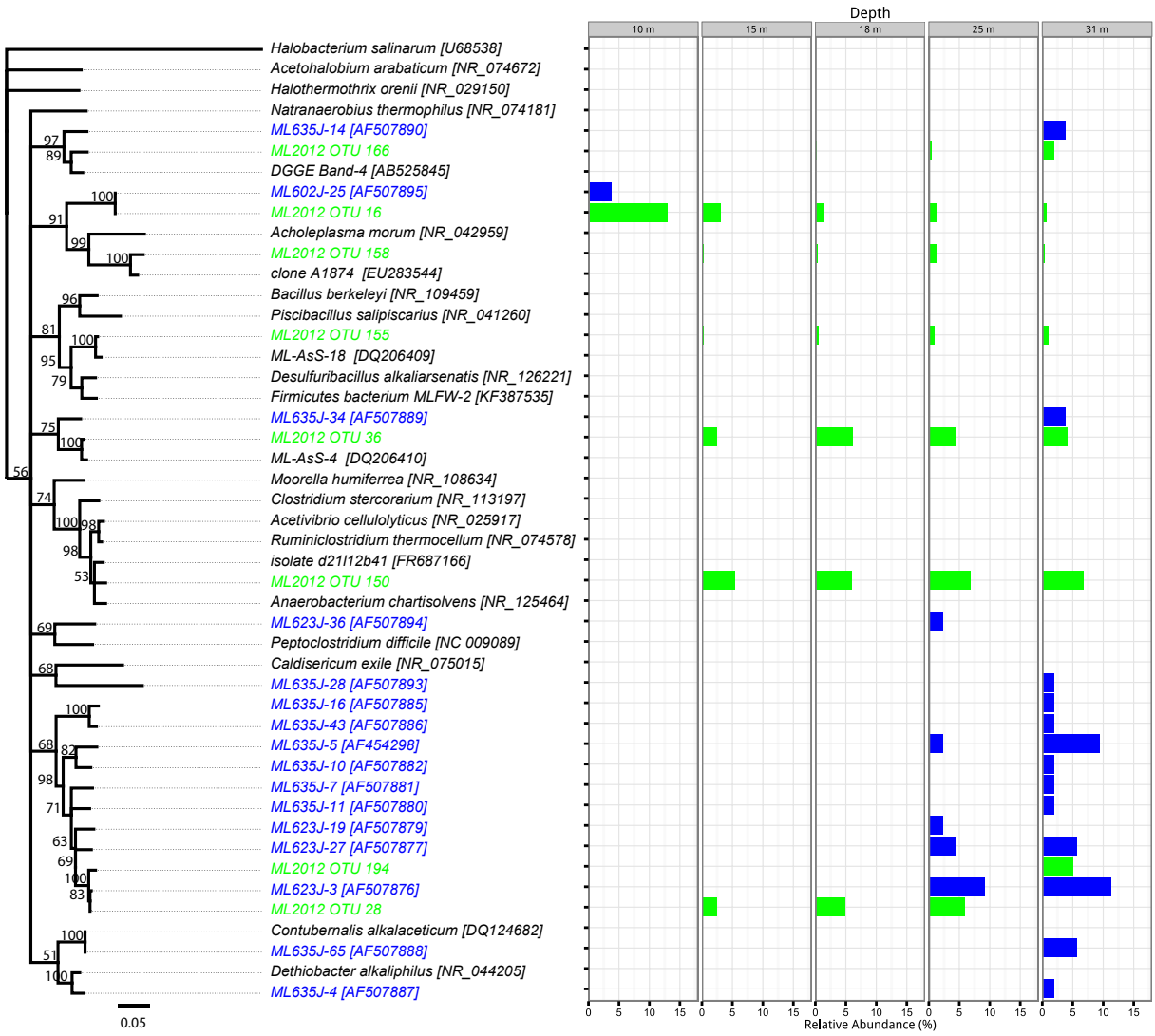


Figure 2.5. *Firmicutes* abundance by depth. Green taxa are tag pyrosequencing OTU representative sequences. Blue taxa are Mono Lake clone OTU representative sequences from Humayoun *et al.*, 2003. Black sequences are metatranscriptome taxonomic bin 16S reference sequences (>1% relative abundance in metatranscriptome) and best hits to NCBI 16S rRNA database. The outgroup is *Halobacterium salinarum*.

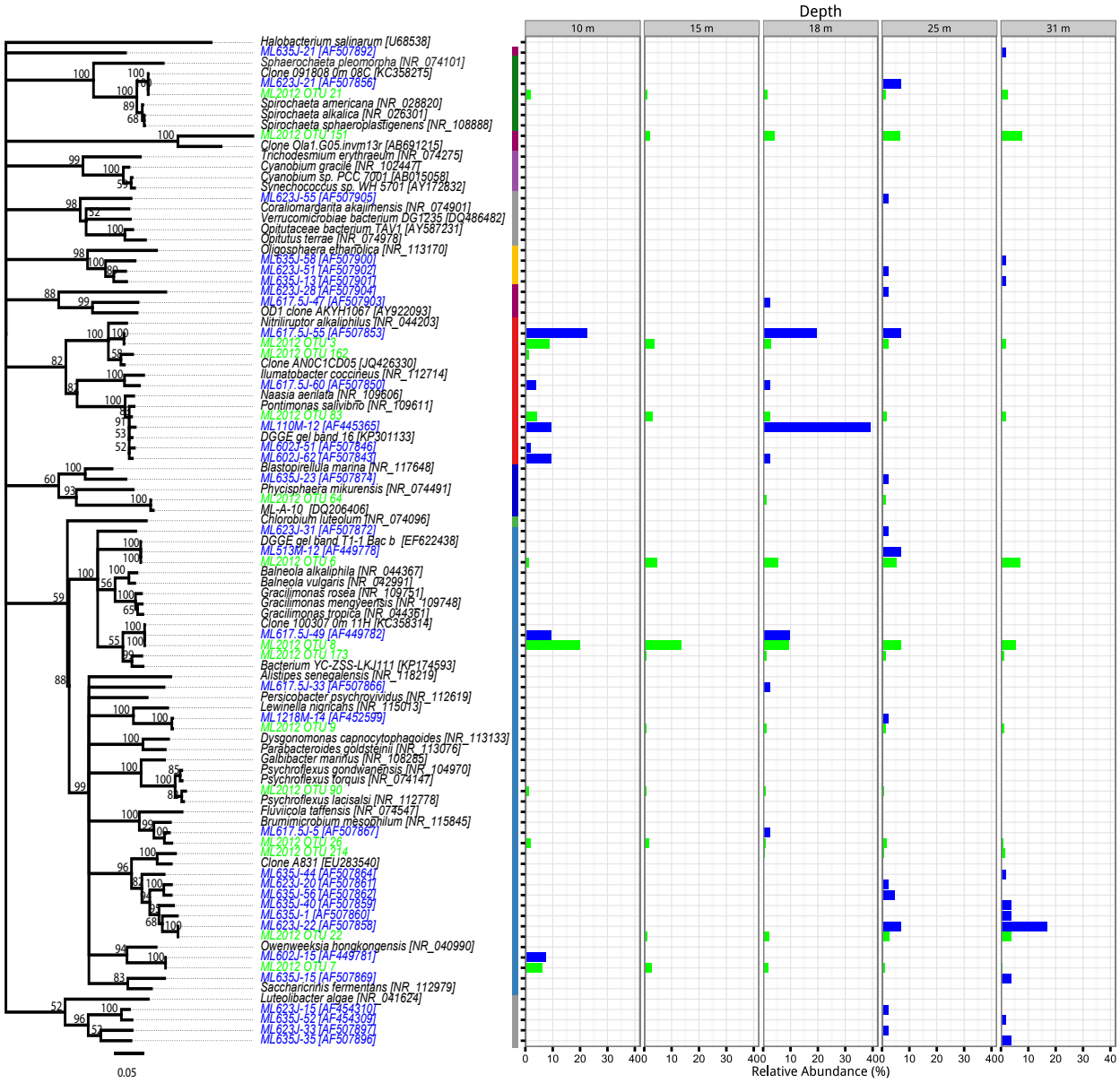


Figure 2.6. Other phyla abundance by depth. Green taxa are tag pyrosequencing OTU representative sequences. Blue taxa are Mono Lake clone OTU representative sequences from Humayoun *et al.*, 2003. Black sequences are metatranscriptome taxonomic bin 16S reference sequences (>1% relative abundance in metatranscriptome) and best hits to NCBI 16S rRNA database. Colored bars indiciate phylum. Coloring is identical to Figure 2.2. The outgroup is *Halobacterium salinarum*.

CHAPTER 3

INSIGHTS INTO THE ROLE OF SULFUR-OXIDIZING AND REDUCING PROKARYOTES IN THE SULFUR AND ARSENIC CYCLES IN ALKALINE, HYPERSALINE MONO LAKE, CALIFORNIA, USA, REVEALED THROUGH METATRANSCRIPTOMICS[1]

[1]Edwardson C.F. and J.T. Hollibaugh. To be submitted to *ISME Journal.*

Abstract

The transcriptionally active sulfur- and arsenic-cycling microbial community in Mono Lake, CA, an alkaline, hypersaline, and stratified lake was investigated. Arsenic cycling at 15 and 18 meters was dominated by arsenite-oxidizing bacteria belonging to *Thioalkalivibrio* and *Halomonas* species. A transition to arsenate-reducing belonging to both Deltaproteobacteria and Firmicutes groups was observed at 25 and 31 m. Sulfur cycling was also dominated by *Thioalkalivibrio*, in addition to *Thioalkalimicrobium*, at 15 and 18 m with a transition to sulfate-reducing bacteria belonging to the class Deltaproteobacteria at 25 and 31 m. Some of the organisms responsible for arsenate reduction were potentially also involved in sulfate reduction in the water column. Further investigation showed that *Thioalkalivibrio* transcripts related to arsenic and sulfur oxidation were more highly transcribed at 15 m relative to other depths.

Introduction

Microbial cycling of sulfur and arsenic compounds is important to human health and to the fate of these elements in the environment (Mandal and Suzuki 2002). Historically, sulfur and arsenic have been studied together due to their co-occurrence in the minerals arsenopyrite, orpiment and realgar (Lengke *et al.* 2009). In addition, soluble arsenic-sulfur oxyanions, known as thioarsenic compounds, have been identified in the environment and are thought to play an important role in the arsenic and sulfur cycles, especially in sulfidic and alkaline waters (Hollibaugh *et al.* 2005, Planer-Friedrich *et al.* 2007, Wilkin *et al.* 2003). The diversity, ecology and physiology of microorganisms catalyzing the transformation of sulfur compounds such as sulfide, elemental sulfur, sulfate, and thiosulfate, are well known (Kelly *et al.* 1997, Muyzer and Stams 2008). The microbial cycling of arsenic is also well studied, with both oxidative and reductive transformations characterized (Oremland *et al.* 2004, Oremland *et al.* 2009). More recently, microbial interactions in the sulfur and arsenic cycle have begun to be appreciated (Edwardson *et al.* 2014, Fisher *et al.* 2008b, Hoeft *et al.* 2004). Thioarsenic compounds have now been demonstrated to support microbial growth (Edwardson *et al.* 2014, Fisher *et al.* 2008b, Härtig *et al.* 2014, Planer-Friedrich *et al.* 2009, Planer-Friedrich *et al.* 2015). Many environments containing both arsenic and sulfur compounds have been investigated with regards to the microbial communities involved, including alkaline thermal springs (Planer-Friedrich *et al.* 2009) and soda lakes (Hollibaugh *et al.* 2005). One such environment is Mono Lake, a hypersaline soda lake in northern California, USA. Mono Lake contains elevated levels of arsenic (200 μM) and sulfide is prevalent, especially in anoxic parts of the lake, leading to measurable concentrations of thioarsenic compounds. These compounds are stable due to the elevated pH of the lake (Hollibaugh *et al.* 2005, Planer-Friedrich *et al.* 2010).

Most of the studies involving the microbial transformations of arsenic and sulfur have involved enrichments or pure cultures (Edwardson *et al.* 2014, Hoeft *et al.* 2004) or by amplifying functional genes from the environment (Hollibaugh *et al.* 2006). The enzymes involved in dissimilatory transformation of arsenic are well established. Arsenate respiratory reductase (ArrA) (Saltikov and Newman 2003), the aerobic arsenite oxidase (AioA) (Lett *et al.* 2012), and the more recently identified alternative arsenite oxidase (ArxA) are members of the complex iron-sulfur molybdoenzyme (CISM) family (van Lis *et al.* 2013). Also known as molybdopterin oxidoreductases due to the presence of a Mo-containing pterin cofactor, these enzymes are part of the dimethylsulfoxide (DMSO) reductase family of enzymes that also includes various oxidoreductases, such as formate dehydrogenase and respiratory nitrate reductase (McEwan *et al.* 2002, Rothery *et al.* 2008). Molecular markers for both oxidative and reductive microbial sulfur transformations include the thiosulfate/sulfur oxidation Sox enzyme system (Meyer *et al.* 2007), the reversible dissimilatory sulfite reductase (*dsrAB*) (Muller *et al.* 2015), and adenosine-5'-phosphosulfate (APS) reductase (*aprBA*). APS reductase catalyzes the transformation of APS to sulfite, which is a key step in both sulfate reduction and sulfite oxidation to sulfate (Meyer and Kuever 2007c).

The study of environmental transcripts to infer the potentially active microbial communities and their roles in biogeochemical cycling in various environments is almost ubiquitous, with insights into microbial carbon cycling in the ocean (Poretsky *et al.* 2010, Satinsky *et al.* 2014), the nitrogen cycle (Hilton *et al.* 2015, Hollibaugh *et al.* 2014), and sulfur cycle in numerous environments (Canfield *et al.* 2010, Chen *et al.* 2015, Spain *et al.* 2015, Stewart *et al.* 2011, Vila-Costa *et al.* 2010). Here, high throughput sequencing of environmental mRNA transcripts (metatranscriptomics) was used to gain further insights into the microbial

communities responsible for arsenic and sulfur cycling in Mono Lake, CA. Transcripts of key enzymes mediating the oxidative and reductive microbial sulfur and arsenic cycle were identified. The focus of this chapter was to determine the abundance of active sulfur and arsenic cycling prokaryotes and examine how the abundance of these microorganisms varied with depth. In addition, further insights into potential novel arsenic and sulfur cycling microorganisms were examined.

Materials and Methods

Field Site, Sampling and Nucleic Acid Processing

Water was collected from five depths at Mono Lake and RNA was collected and processed as described previously (Chapter 2). Briefly, water column profiles of environmental variables (conductivity, temperature, depth, dissolved oxygen and chlorophyll a concentrations and beam attenuation) were obtained using a SBE19 CTD (SeaBird Electronics) (Hollibaugh *et al.* 2005) and used the data to determine ideal sampling depths at that covered a range of redox conditions. Samples for sulfide analysis were taken as described previously (Miller *et al.* 1993). Briefly, water was collected from discrete depths and injected into 8 mL, septa-capped vials containing 250 µL of 15% solution of zinc acetate and analyzed spectrophotometrically (Cline 1969). Samples for total arsenic and thioarsenic speciation were collected and analyzed by ICP-MS and IC-ICP-MS as described previously (Planer-Friedrich *et al.* 2007). Water (~0.5-2 L) from 10, 15, 18, 25 and 31 m was filtered through 0.2 µm pore size, 142 mm, Supor (Pall Life Sciences) membrane filters, which were immediately flash frozen in liquid nitrogen. Total RNA was extracted from two filters collected from each depth and mRNA was processed as described previously (Gifford *et al.* 2011, Stewart *et al.* 2010), with internal standards added (Satinsky *et*

al. 2013), converted to cDNA and sequenced on an Illumina HiSeq instrument to obtain 150 bp paired-end reads.

Bioinformatics – Metatranscriptomics

Initial processing of the dataset and taxonomic binning was described previously (Chapter 2). Briefly, reads were quality filtered with PrinSeq (Schmieder and Edwards 2011) and paired with PEAR (Zhang *et al.* 2014), discarding reads that did not overlap. Average paired read size was ~240 bp. Ribosomal RNA and internal standard sequences were found using RiboPicker (Schmieder *et al.* 2012) , and internal standard sequences were counted using BLASTN. These sequences were removed, and remaining reads aligned with reference protein sequences (NCBI RefSeq database) using Rapsearch2 (Zhao *et al.* 2012). For hits to Bacteria, the full taxonomy (NCBI-based) and counts of these hits were determined with custom scripts.

Molybdopterin Oxidoreductase Identification

A custom database of 110 representative molybdopterin oxidoreductase (also known as Complex Iron-Sulfur Molybdoenzymes (CISM)) catalytic subunit amino acid sequences was prepared with sequences from various references (Denton *et al.* 2013, Grimaldi *et al.* 2013, Rothery *et al.* 2008, Schoepp-Cothenet *et al.* 2012) (Table 3.S1). The sequences were downloaded from NCBI and edited in Geneious (Biomatters LTD). BLASTX (Altschul *et al.* 1990) was used to perform a search of all reads from each depth against this database. All hits with a bit score greater than 40 were retained (Gifford *et al.* 2011). The query reads that had hits to the custom database were matched with their respective RapSearch2 hit to the RefSeq database. If there was no RapSearch2 hit, BLASTX was used to search against the RefSeq database. Full taxonomy was obtained with custom scripts.

Sulfur Cycling Identification

Representative amino acid sequences for proteins involved in oxidative and reductive microbial sulfur cycling (Table 3.S2) were obtained using a text search of the UniRef90 database (Suzek *et al.* 2015). Additional sequences for some proteins (AprBA, DsrAB, Sat, SorAB, SoxABCDXYZ) (Frigaard and Dahl 2008, Loy *et al.* 2009, Meyer *et al.* 2007, Meyer and Kuever 2007a, Meyer and Kuever 2007b, Mori *et al.* 2010) were obtained from GenBank and RefSeq. Duplicates were removed in Geneious, and sequences were re-clustered at 90% identity with CD-HIT (Li and Godzik 2006). All reads from each depth were searched against this database using BLASTX, and then processed as described above for the CISM database.

Metatranscriptome Assembly and Annotation

An additional quality control step was performed on the raw reads to ensure better sequence assembly. TrimGalore (www.bioinformatics.babraham.ac.uk/projects/trim_galore/) was used to remove Illumina adapter sequences. PEAR was used with default settings and MAP statistical test enabled to pair overlapping reads (Zhang *et al.* 2014). All reads in each library, including unpaired reads, were assembled using Trinity version r20140717 (Celaj *et al.* 2014, Grabherr *et al.* 2011). Bowtie2 (Langmead and Salzberg 2012) was used to map reads from each library to the resulting contigs. The contigs were annotated with Prokka (Seemann 2014) and also searched against the custom CISM and sulfur databases to confirm the presence of a full length ORF of interest using the BLAST functions in Geneious. Full length ORFs (identified as PROKKA_XXXX) that had hits to proteins of interest in the database were retained for further analysis. Additional ORFs (identified as MANUAL_XXXX) were discovered using Geneious when a contig had a BLASTX hit to the custom databases but no Prokka annotated ORF. To

reduce the number of contig sequences to analyze, all protein sequence translations of the ORFs were clustered with CD-HIT at a 90% identity cutoff.

Phylogenetic Analysis

The full length amino acid sequences corresponding to the "best hits" to the custom databases were obtained from the NCBI RefSeq database. The counts of each hit were used to calculate transcripts per liter for each library, and an average was taken for each depth. Only "best hits" greater than 1% relative abundance for a given gene at any depth were considered for phylogenetic analysis. The full-length protein translations of the assembled transcripts were also included in phylogenetical analysis. Transcript abundances for the assembled protein translations were obtained by taking the total number of hits that mapped to the full contig containing the transcript of interest and diving by the length of the contig and multiplying by the length of the transcript of interest. This hit count was then converted to transcripts/L as described above. Protein sequences were aligned in Geneious with MAFFT (Katoh and Standley 2013). The best evolutionary model was determined using Prottest 3.4 (Darriba *et al.* 2011), and that model was used to determine phylogeny using the PhyML Geneious plugin (Guindon and Gascuel 2003). Outgroups were chosen as noted in each figure and the maximum likelihood consensus tree from 100 bootstraps was used. Initial alignments and trees were analyzed for paralogs by visual inspection, and "best hits" to proteins other than those being analyzed (paralogs) were removed from analysis. Additional reference sequences from the custom databases were added to trees for sulfur cycle proteins based on a 64% identity cluster value using CD-HIT. This cutoff was chosen based on its previous use in an analysis of DsrAB (Muller *et al.* 2015). This percent identity roughly corresponds to sequences grouping at the genus level (Luo *et al.* 2014).

Thioalkalivibrio Analysis

The most abundant genus (*Thioalkalivibrio*) and genome bin (*Thioalkalivibrio nitratireducens*) were analyzed further to look at differential transcript abundance between depths. All *Thioalkalivibrio* reads were filtered out of the data set using the QIIME script filter_fasta.py. BLASTX was used to align these hits to all *Thioalkalivibrio nitratireducens* proteins in custom a database to ensure correct comparison between libraries/depths. Raw counts of hits to each protein in the *T. nitratireducens* genome were collected for each library. The Bioconductor/R package DESeq2 (Love *et al.* 2014) was used to determine $\log_2$ fold changes in transcript abundance at each depth, with a Benjamini-Hochberg false discovery rate (adjusted *p*-value) of 0.05 or lower. The log2fold changes are relative to the 15 m samples, chosen mainly because of the greatest abundance of *Thioalkalivibrio* at that depth.

Results and Discussion

Chemical Characteristics of Sampling Depths

A basic description of the limnology at the time of sampling at Mono Lake (Station 6, July 2012) has been described previously (Chapter 2), and the CTD and depth profile is included in Chapter 2 (Figure 2.1). Briefly, samples were obtained from 5 depths (10 m, 15 m, 18 m, 25 m and 31 m) along a redox gradient. The concentration of sulfide, arsenate, arsenite, and thioarsenic compounds varied with depth (Figure 2.1). The 10 m sample was from the base of the epilimnion, dominated by the oxyanion arsenate (99.9% of arsenic, 211 μM). The 15 m sample was from the base of the oxycline, and the 18 m sample was from at the base of the thermocline. Both the 15 and 18 m samples were from the redox transition zone where arsenic speciation shifts from arsenate (15 m: 80.3%, 142 μM) to arsenite (18 m: 95%, 196 μM). Trace quantities of thioarsenates began to appear at these depths. The anoxic hypolimnion was sampled at 25 m

and 31 m. Sulfide appears at these depths, with a decrease in arsenite concentration likely due to formation of thioarsenates. Sulfide and thioarsenic concentrations were much lower than recorded during long term meromixis (Hollibaugh *et al.* 2005). Although not measured at the time of sampling, sulfate concentration remains generally constant in the lake at ~100-150 mM throughout the water column (Oremland *et al.* 2000).

Overview of Metatranscriptome

The taxonomic affiliations of metatranscriptomic reads were analyzed previously (Chapter 2). Reads were aligned with the NCBI RefSeq database and grouped according to taxonomy. Counts at different taxonomic levels were performed, and absolute (transcripts/L) and relative (proportion of total transcripts in each library or depth) abundances were calculated. Genera that made up >1% relative abundance at any depth are plotted in Figure 3.1. The most abundant genera belonged to the Proteobacteria and Firmicutes phyla; especially at 15-31 m. Transcripts belonging to *Thioalkalivibrio* and *Thioalkalimicrobium* were especially abundant. These two members of the Gammaproteobacteria order Chromatiales are sulfur-oxidizing bacteria, and representative species *Thioalkalivibrio jannaschii* and *Thioalkalimicrobium cyclicum* have been isolated from Mono Lake (Sorokin *et al.* 2002). Identification of these top genera as containing taxa that are known to be able to perform arsenic and sulfur redox transformations was performed based on literature searches (Amend *et al.* 2014, Canfield *et al.* 2010, Sorokin *et al.* 2015). This search revealed that almost half (27/57) of these genera contain taxa with known capability for dissimilatory sulfur and arsenic metabolism, highlighting the importance of sulfur and arsenic oxidizing bacteria in the lake.

Arsenic Redox Activity

The abundance of transcripts hits to and the phylogeny of the catalytic subunit of both the dissimilatory arsenate reductase (ArrA) (Saltikov and Newman 2003) and arsenite oxidase (AioA, ArxA) (Lett *et al.* 2012, Zargar *et al.* 2012) was used to identify the bacteria involved in arsenic redox transformations at different depths in the lake. Because of the ubiquity and ancient origin of these enzymes and their varying functions (Schoepp-Cothenet *et al.* 2012), many of the database entries and/or genomic annotations of these enzymes are incorrect, such as known arsenate reductases that are annotated as "formate dehydrogenase." In addition, many of the annotations are non-specific with respect to the substrate of the enzyme, such as "molybdopterin oxidoreductase" or "dehydrogenase". Therefore, investigation of a specific subset of these enzymes required more than textual parsing of hits to the RefSeq database. This initial search of reads against a custom database (Table 3.S1) was used to place each read into the most likely CISM protein bin. In the next step, reads with hits to the CISM database were matched to their respective RefSeq database hit. Relative to the 15-31 m samples, the CISM transcript abundance at 10 m was low (Figure 3.2) so the focus was only on the other depths. At 15 m anaerobic arsenite oxidase (*arxA*) was the most abundant CISM transcript. The abundance of *arxA* was lower at 18 – 31 m, with an increase in abundance of arsenate reductase, implying a potential shift from arsenite oxidation to arsenate reduction as the dominant respiratory arsenic transformation. The canonical arsenite oxidase (*aioA*) was detected at very low abundance at all depths. In addition to arsenic redox enzymes, the most abundant CISM transcripts belonged to different types of formate dehydrogenases (*fdhA, fdhN, fdnG*). These enzymes catalyze the conversion of formate to CO_2 and occur in a wide variety of organisms (Grimaldi *et al.* 2013). Analysis of the top RefSeq hits identified as formate dehydrogenase in the CISM custom

database revealed two general types of formate dehydrogenase (data not shown). One is membrane bound and known to be involved with generating a proton motive force through electron transport through nitrate reductase (Jormakka *et al.* 2002), and the other is a formate dehydrogenase known to be associated with the formate hydrogen lyase complex (McDowall *et al.* 2014). Further evaluation is needed to identify their role in Mono Lake biogeochemistry.

A closer examination of the taxonomic affiliation of *arxA* and *arrA* transcripts (Figure 3.3) show that the *arxA* transcripts were dominated by *Thioalkalivibrio* and *Halomonas*, with transcripts mostly affiliated with *Thioalkalivibrio nitratireducens*, *Halomonas boliviensis,* and *Halomonas* sp. A3H3. There is no direct evidence that *Thioalkalivibrio* is able to oxidize arsenite, but two other relatives in the *Ectothiorhodospiraceae* family, *Ectothiorhodospira* sp. PHS-1 (Kulp *et al.* 2008) and *Alkalilimnicola ehrlichii* MLHE-1 (Hoeft *et al.* 2007) oxidize arsenite and transcribe the *arxA* gene (Zargar *et al.* 2010, Zargar *et al.* 2012). In addition, a Mono Lake enrichment culture that oxidized arsenite via the formation of thioarsenic compounds contained mainly *Thioalkalivibrio jannaschii* (Fisher *et al.* 2008b). *T. jannaschii* was able to transformation thioarsenic compounds, but growth was not observed in the process (Edwardson *et al.* 2014). *Halomonas* species from Big Soda Lake, NV (a lake with chemistry similar to Mono Lake) have been shown recently to oxidize arsenite and to contain *arxA* (A. Conrad, unpublished thesis). In addition, an arsenite-oxidizing, *arxA*-containing *Halomonas*-related strain (ANAO-440) was isolated from a Mongolian soda lake (Hamamura *et al.* 2014). Thus, *Halomonas* and *Thioalkalivibrio* strains may be more important members of the arsenite oxidizing community in Mono Lake may than previously understood.

The taxonomic affiliation of the *arrA* transcripts was more diverse, with representatives from at least 7 phyla. Dominant bacterial genera include *Desulfurispirillum*, *Desulfitobacterium*,

Desulfosporsinus, *Aeromonas*, *Ferrimonas*, and *Thioalkalivibrio*. Strains in all of those genera have been shown either directly (Nakagawa *et al.* 2006, Niggemyer *et al.* 2001, Pérez-Jiménez *et al.* 2005, Rauschenbach *et al.* 2012) or indirectly (Pepi *et al.* 2007) to be able to reduce arsenate. Six different *arrA* hits had greater than 10% relative abundance at 15-31m: *Thioalkalivibrio nitratireducens*, *Desulfurispirillum indicum*, *Desulfovibrio alkalitolerans*, *Natronobacterium gregoryi*, *Alkaliphilus oremlandii*, and *Natranaerobius thermophilus* (Figure 3.3). Only *Alkaliphilus oremlandii* and *Desulfurispirillum indicum* have previously been shown to reduce arsenate (Fisher *et al.* 2008a, Rauschenbach *et al.* 2012). *Natronobacterium gregoryi* has recently been hypothesized to reduce arsenate based on the similarity of its arsenate respiratory reductase operon to an arsenate-reducing haloarchaeal biofilm community (Rascovan *et al.* 2015). An additional haloarchaeal arsenate reductase was also detected (*Halobiforma* sp.). None of the other organisms to which these transcripts hit, besides the aforementioned, have been tested for their ability to transform arsenic, to our knowledge. *Thioalkalivibrio nitratireducens* appears to contain an arsenate reductase in addition to an arsenite oxidase, but since it is not known whether the organism is able to oxidize arsenite or reduce arsenate, the actual role of this organism in the arsenic cycle is less clear. It could be argued that since it is transcribing more *arxA* than *arrA*, it is more likely oxidizing arsenite, but alternatively it could switch between oxidation of arsenite oxidation and arsenate reduction depending on redox conditions. The fact that it is present at all depths introduces the question of what electron acceptor would be used by the *T. nitratireducens* in the anoxic depths of Mono Lake. Most of the characterized *Thioalkalivibrio* strains are aerobic, with some able to use nitrogen compounds (NO_3, NO_2, N_2O) as alternative electron acceptors (Sorokin *et al.* 2013). The characterized strain of *Thioalkalivibrio nitratireducens* is able to reduce nitrate (but not nitrite) (Sorokin *et al.* 2003),

however very few (<10) periplasmic nitrate reductase transcripts were detected relative to the total number of transcript hits to *T. nitratireducens* (>100000) . In addition, this strain also contains an octaheme nitrite reductase that functions *in vitro* but has unclear function *in vivo* (Tikhonova *et al.* 2006), and approximately 50 transcripts per library hit to this protein. The only other annotated nitrate-associated transcript that was detected is a nitrate/nitrite sensor.

The majority of transcripts assembling into full length ArrA and ArxA sequences belonged to one ArrA and one ArxA consensus sequence, clustered at 90% identity (Figure 3.5). A search of the consensus sequences against the NCBI database showed that the ArrA sequence (PROKKA_00186) was most closely related to *Halarsenibacter silvermanii*, a strain isolated from nearby Searles Lake (Blum *et al.* 2009), and *Natranaerobius thermophilus*, a haloalkalithermophile isolated from a soda lake in Wadi An Natrun, Egypt (Mesbah *et al.* 2007). The ArxA sequence (PROKKA_00030) was most closely related to Oceanospirillales bacteria including *Nitrincola lacisaponensis* and *Halomonas* strains. Neither of these sequences clustered closely with reference sequences in the phylogenetic trees, indicating the potential for abundant arsenic oxidizing and reducing organisms that have yet to be isolated from Mono Lake, despite numerous characterized and sequenced strains of arsenic respiring organisms from that environment. Alternatively, misassembly of transcripts could contribute to a chimeric sequence and the more dissimilar it is to cultured strains could be an artifact. The *arrA* and *arxA* transcripts do not appear hit to one dominant RefSeq sequence. This could be due to a number of factors. First, in performing the homology search using Rapsearch2, only the top hit was retained. However, when the sequences are highly dissimilar from database sequences there tends to be a number of "top hits" with similar but slightly lower bit scores. Also, with short reads, one read from an environmental transcript might align to one database sequence better than another read

from the same transcript, leading to differential binning of that transcript. Therefore, consensus sequences from correct assemblies may actually be a better representation of the majority environmental sequence. Alternatively, there may be a diverse arsenite-oxidizing and arsenate-reducing community in Mono Lake. One way to improve on this analysis would be to perform a metagenome assembly and align the metatranscriptome reads to assembled metagenome scaffolds.

Sulfur Redox Cycling

Microbial oxidation of sulfur compounds (sulfide, sulfite, thiosulfate, and elemental sulfur) is performed by both phototrophic and lithotrophic bacteria (Friedrich *et al.* 2005, Frigaard and Dahl 2008). Oxidative and reductive transformation of sulfur compounds is mediated by a large number of enzymes, but recent molecular surveys have focused on *soxB*, *aprBA*, and *dsrAB* (Meyer *et al.* 2007, Meyer and Kuever 2007c, Muller *et al.* 2015). Metatranscriptome reads from these sulfur cycling genes were examined for the presence of *soxB* (thiosulfate oxidation), *aprA* (indirect sulfite to sulfate through APS or reverse) and *dsrA* (sulfite to sulfide or the reverse) to identify the organisms most likely involved in sulfur transformations and to determine how they varied by depth (Figure 3.4). At 15 m, transcripts of key sulfur genes were dominated by *Gammaproteobacteria*, with *aprA* and *dsrA* transcripts dominated by *Thioalkalivibrio*. *SoxB* transcripts were dominated by *Thioalkalimicrobium*, although *Thioalkalivibrio* was also abundant. The relative abundance of the *soxB* transcripts, compared with *dsrA* and *aprA*, decreased with depth. Deltaproteobacteria *aprA* transcripts were present at 15 m, but they increased at 18 m, dominated by Deltaproteobacteria strain MLMS-1. At 25 and 31 m, the *aprA* transcripts were dominated by the Deltaproteobacteria *Desulfatibacillum*, *Desulfococcus*, *Desulfonatronospira*, and *Desulfovibrio*. These organisms also contributed to the

dsrA transcript pool; however, the *dsrA* transcript pool also contained reads affiliated with the sulfate-reducing Clostridia, including *Desulfotomaculum*, *Natranaerobius*, *Desulfurispora* and *Dethiobacter*. The proportion of the *dsrA* transcripts affiliated with unclassified Thermoplasmatales archaea increased with depth. *DsrA* transcripts from 15 m contained representatives of Clostridia and Deltaproteobacteria as well as some Gammaproteobacteria, but as depth increased, the contribution of Gammaproteobacteria decreased. More careful phylogenetic analysis revealed that the custom database search recruited paralogous sequences of a 4Fe-4S ferredoxin of unknown function, but possibly assimilatory sulfite reductase. So, although the Firmicutes contribute to sulfur cycling in the lake, it is unclear if they are gaining energy from the reduction of sulfur compounds or using the dissimilatory sulfite reductase enzyme complex. Phylogenetic analysis also revealed the Thermoplasmatales sequences to be paralogous as well. Of the sequences that contribute to >1% of the *dsrA* transcripts at every depth, all are members of either the oxidative ("reverse") group and are dominated by *Thioalkalivibrio* sequences, or are members of the Deltaproteobacteria families *Desulfovibrionales* and *Desulfobacterales*. Two assembled AprA consensus sequences were abundant at 15 and 18 m. PROKKA_00232, related to *Desulfobulbus propionicus* and strain MLMS-1, and PROKKA_00196, related to *Thioalkalivibrio nitratireducens*. This transcript was also abundant at 25 m, as were 5 transcripts related to sulfate-reducing Deltaproteobacteria (Figure 3.6), highlighting the transition between sulfide oxidation and sulfate reduction between 18 and 25 m. An almost identical trend is seen with DsrA (Figure 3.7). A general trend is a shift from *Desulfurivibrio alkaliphilus* and strain MLMS-1 at 15 m and 18 m to the Desulfovibrionales group (*Desulfovibrio* and *Desulfonatronospira thiodismutans*). The Desulfobulbaceae (including MLMS-1 and *Desulfurivibrio*) are known sulfide oxidizers (Hoeft

et al. 2004, Hollibaugh *et al.* 2006, Pfeffer *et al.* 2012). Thus, it appears that sulfide oxidation occurs at 15-18 m, with a switch to sulfate reduction at 25 and 31m. This is consistent with previous rate measurements (Oremland *et al.* 2000). That study also measured arsenate reduction rates in the water column and hypothesized that both arsenate and sulfite reduction were occurring simultaneously or that there was a potential shift in SRB activity between sulfite and arsenate reduction. This is supported by the present study due to the recovery of transcripts affiliated with arsenate reduction and dissimilatory sulfite reduction from both *Desulfovibrio* and *Desulfonatronospira*. Thus, sulfate reduction in Mono Lake appears to be dominated by Deltaproteobacteria members of the order Desulfovibrionales. This work is also in agreement with previous studies of sulfur-reducing bacteria in Mono Lake (Scholten *et al.* 2005) that identified abundant sulfate-reducing bacteria in the anoxic depths of the lake through 16S rRNA genes but were not able to identify functional marker genes (*dsrAB*, *aprA*), likely due to inadequate primer sequences or low abundance of extracted DNA. SoxB is solely an indicator for sulfur (thiosulfate) oxidizing bacteria, whereas the other two enzymes are found in sulfate-reducing bacteria as well. The use of the *soxB* gene sequence as a molecular marker in previous studies has shown that it is present in soda lakes other than Mono Lake (Tourova *et al.* 2013). However, PCR analysis did not verify the presence of *Thioalkalimicrobium* in those environments, either due to issues with the primer sequences or to its absence or low abundance. *Thioalkalimicrobium soxB* was abundant at all depths below 15 m, while *Thioalkalivibrio soxB* decreased with depth, even though *Thioalkalivibrio* transcripts were more abundant than *Thioalkalimicrobium* at 31m. This could indicate a potential shift in metabolism for *Thioalkalivibrio*, whereas *Thioalkalimicrobium* does not appear to be able to oxidize other sulfur compounds, based on lack of transcript hits to *aprA* or *dsrA*. Also, the *Thioalkalimicrobium*

cyclicum genome does not contain the Apr or Dsr pathway, despite the ability of *T. cyclicum* to use sulfur compounds other than thiosulfate in pure culture (Sorokin *et al.* 2002).

Thioalkalivibrio transcription analysis

Additional analyses of the top genome bin, *Thioalkalivibrio nitratireducens*, were performed due to its presence in both arsenic and sulfur cycling genes. All genus-level *Thioalkalivibrio* transcripts hits from the 10-31 m samples were realigned to the *T. nitratireducens* (*Tnat*) genome by BLASTX against all *Tnat* protein sequences. The assumption was that other sequences that binned at the genus level to *Thioalkalivibrio* strains would have homologs in the *Tnat* genome. The goal was to capture the functional diversity of the *Thioalkalivibrio* genus, since it is likely that the abundant organism in Mono Lake is a different strain of *Thioalkalivibrio*. Only 86% of all *Thioalkalivibrio* transcripts had hits in the *Tnat* genome. The average inferred amino-acid percent identity of the transcripts to the *Tnat* proteins was 78%, which corresponds to approximately the genus level. Differential expression analysis was performed to look for transcripts that might be transcribed more at one depth than another. The *Thioalkalivibrio* transcription pattern at 10, 18, 25 and 31 m was compared to transcription at 15 m, arbitrarily chosen as a reference depth for comparison purposes. The top 10 significant (adjusted p-value <0.05) differentially transcribed genes are listed in Table 3.S3 for each depth compared to 15 m. The transcription patterns of the known sulfur and arsenic genes in the *Tnat* genome were analyzed (Table 3.S4). Transcripts with hits to sulfite and arsenite oxidation genes (*dsr* and *arx* operons) were among the transcripts with the highest negative log-fold change at 10 m vs 15 m. The genes that were the most highly transcribed at 15 m (relative to the other depths) include members of the dissimilatory sulfite reductase operon and proteins involved in electron transport and respiration, indicating respiration of sulfite at 15 m. In addition, differential

transcription analysis of inferred arsenate reductase and arsenite oxidase genes show that arsenite oxidase and arsenate reductase operons are more highly transcribed at 15 m relative to other depths. This reinforces the data from the transcript counts indicating that arsenite oxidation and sulfite oxidation by *Thioalkalivibrio* are important at 15 m. A number of genes that were transcribed more strongly at 18-31 m than at 15 m were annotated as hypothetical proteins. Predictions about their functions are available through conserved domain analysis (see Table 3.S3). One such transcript encodes a putative DUF302 domain-containing protein. While the function of this protein is not known, recent studies have hypothesized that it is contained in an operon encoded by a plasmid in a *Synechocystis* strain that is induced by the presence of both arsenite and sulfide (Nagy *et al.* 2014). Thus, a role in arsenic and/or sulfur cycling would not be surprising. In addition to arsenic and sulfur cycle genes being highly differentially transcribed, a large number of genes that were more highly transcribed at 15 m relative to other depths were related to growth (ribosome and ATP related), respiration and electron transport (ubiquinone and cytochrome), and carbon assimilation (RuBisCO). One hypothesis is that *Thioalkalivibrio* is more actively growing and/or respiring at 15 m than at other depths, but the cells are present and transcription continues to occur. This would also help explain the fact that *Thioalkalivibrio* would not have an electron acceptor (oxygen, nitrate or potentially arsenate) readily available at 18-31 m. Thus, *Thioalkalivibrio* appears to be a key player in both the sulfur and arsenic cycles, especially at 15 m, but possibly at 18-31 m as well.

Acknowledgements

We thank Meredith Ross, Christopher Abin, Ron Oremland and Larry Miller at USGS, Sierra Nevada Aquatic Research Lab and Tom Crowe of Mono Lake Boat Tours for field support. We thank Britta Planer-Friedrich for analyzing samples for arsenic and thioarsenic speciation. We thank Brandon Satinsky for internal mRNA standards, Shalabh Sharma, and the Georgia Advanced Computing Resource Center at UGA for bioinformatics support and scripts.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215:** 403-410.

Amend JP, Saltikov C, Lu G-S, Hernandez J (2014). Microbial arsenic metabolism and reaction energetics. *Rev Mineral Geochem* **79:** 391-433.

Blum JS, Han S, Lanoil B, Saltikov C, Witte B, Tabita FR *et al.* (2009). Ecophysiology of "*Halarsenatibacter silvermanii*" Strain SLAS-1T, gen. nov., sp. nov., a facultative chemoautotrophic arsenate respirer from salt-saturated Searles Lake, California. *Appl Environ Microbiol* **75:** 1950-1960.

Canfield DE, Stewart FJ, Thamdrup B, De Brabandere L, Dalsgaard T, Delong EF *et al.* (2010). A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Science* **330:** 1375-1378.

Celaj A, Markle J, Danska J, Parkinson J (2014). Comparison of assembly algorithms for improving rate of metatranscriptomic functional annotation. *Microbiome* **2:** 39.

Chen L-x, Hu M, Huang L-n, Hua Z-s, Kuang J-l, Li S-j *et al.* (2015). Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J* **9:** 1579-1592.

Cline JD (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnol Oceanogr* **14:** 454-458.

Darriba D, Taboada GL, Doallo R, Posada D (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27:** 1164-1165.

Denton K, Atkinson M, Borenstein S, Carlson A, Carroll T, Cullity K *et al.* (2013). Identification of a possible respiratory arsenate reductase in *Denitrovibrio acetiphilus*, a member of the phylum Deferribacteres. *Arch Microbiol* **195:** 661-670.

Edwardson CF, Planer-Friedrich B, Hollibaugh JT (2014). Transformation of monothioarsenate by haloalkaliphilic, anoxygenic photosynthetic purple sulfur bacteria. *FEMS Microbiol Ecol* **90:** 858-868.

Fisher E, Dawson AM, Polshyna G, Lisak J, Crable B, Perera E *et al.* (2008a). Transformation of inorganic and organic arsenic by *Alkaliphilus oremlandii* sp. nov. strain OhILAs. *Ann New York Acad Sci* **1125:** 230-241.

Fisher JC, Wallschlager D, Planer-Friedrich B, Hollibaugh JT (2008b). A new role for sulfur in arsenic cycling. *Environ Sci Technol* **42:** 81-85.

Friedrich CG, Bardischewsky F, Rother D, Quentmeier A, Fischer J (2005). Prokaryotic sulfur oxidation. *Curr Opin Microbiol* **8:** 253-259.

Frigaard N-U, Dahl C (2008). Sulfur metabolism in phototrophic sulfur bacteria. *Advances in Microbial Physiology*. pp 103-200.

Gifford SM, Sharma S, Rinta-Kanto JM, Moran MA (2011). Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J* **5:** 461-472.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **29:** 644-652.

Grimaldi S, Schoepp-Cothenet B, Ceccaldi P, Guigliarelli B, Magalon A (2013). The prokaryotic Mo/W-bisPGD enzymes family: a catalytic workhorse in bioenergetic. *Biochimica et biophysica acta* **1827:** 1048-1085.

Guindon S, Gascuel O (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52:** 696-704.

Hamamura N, Itai T, Liu Y, Reysenbach A-L, Damdinsuren N, Inskeep WP (2014). Identification of anaerobic arsenite-oxidizing and arsenate-reducing bacteria associated with an alkaline saline lake in Khovsgol, Mongolia. *Environmental microbiology reports* **6:** 476-482.

Härtig C, Lohmayer R, Kolb S, Horn MA, Inskeep WP, Planer-Friedrich B (2014). Chemolithotrophic growth of the aerobic hyperthermophilic bacterium *Thermocrinis ruber* OC 14/7/2 on monothioarsenate and arsenite. *FEMS Microbiol Ecol* **90:** 747-760.

Hilton JA, Satinsky BM, Doherty M, Zielinski B, Zehr JP (2015). Metatranscriptomics of N2-fixing cyanobacteria in the Amazon River plume. *ISME J* **9:** 1557-1569.

Hoeft SE, Kulp TR, Stolz JF, Hollibaugh JT, Oremland RS (2004). Dissimilatory arsenate reduction with sulfide as electron donor: Experiments with mono lake water and isolation of Strain MLMS-1, a chemoautotrophic arsenate respirer. *Appl Environ Microbiol* **70:** 2741-2747.

Hoeft SE, Blum JS, Stolz JF, Tabita FR, Witte B, King GM *et al.* (2007). *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. *Int J Syst Evol Microbiol* **57:** 504-512.

Hollibaugh J, Carini S, Gurleyuk H, Jellison R, Joye S, Lecleir G *et al.* (2005). Arsenic speciation in Mono Lake, California: Response to seasonal stratification and anoxia. *Geochim Cosmochim Acta* **69:** 1925-1937.

Hollibaugh JT, Budinoff C, Hollibaugh RA, Ransom B, Bano N (2006). Sulfide oxidation coupled to arsenate reduction by a diverse microbial community in a soda lake. *Appl Environ Microbiol* **72:** 2043-2049.

Hollibaugh JT, Gifford SM, Moran MA, Ross MJ, Sharma S, Tolar BB (2014). Seasonal variation in the metratranscriptomes of a Thaumarchaeota population from SE USA coastal waters. *ISME J* **8:** 685-698.

Jormakka M, Tornroth S, Byrne B, Iwata S (2002). Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. *Science* **295:** 1863-1868.

Katoh K, Standley DM (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30:** 772-780.

Kelly DP, Shergill JK, Lu WP, Wood AP (1997). Oxidative metabolism of inorganic sulfur compounds by bacteria. *Antonie Van Leeuwenhoek* **71:** 95-107.

Kulp TR, Hoeft SE, Asao M, Madigan MT, Hollibaugh JT, Fisher JC *et al.* (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science* **321:** 967-970.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9:** 357-359.

Lengke MF, Sanpawanitchakit C, Tempel RN (2009). The oxidation and dissolution of arsenic-bearing sulfides. *Can Mineral* **47:** 593-613.

Lett MC, Muller D, Lievremont D, Silver S, Santini J (2012). Unified nomenclature for genes involved in prokaryotic aerobic arsenite oxidation. *J Bacteriol* **194:** 207-208.

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

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with *DESeq2*. *Genome biology* **15:** 550.

Loy A, Duller S, Baranyi C, Mussmann M, Ott J, Sharon I *et al.* (2009). Reverse dissimilatory sulfite reductase as phylogenetic marker for a subgroup of sulfur-oxidizing prokaryotes. *Environ Microbiol* **11:** 289-299.

Luo C, Rodriguez-R LM, Konstantinidis KT (2014). MyTaxa: an advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Res* **42:** e73.

Mandal BK, Suzuki KT (2002). Arsenic round the world: a review. *Talanta* **58:** 201-235.

McDowall JS, Murphy BJ, Haumann M, Palmer T, Armstrong FA, Sargent F (2014). Bacterial formate hydrogenlyase complex. *Proc Natl Acad Sci USA* **111:** E3948-E3956.

McEwan AG, Ridge JP, McDevitt CA, Hugenholtz P (2002). The DMSO reductase family of microbial molybdenum enzymes; molecular properties and role in the dissimilatory reduction of toxic elements. *Geomicrobiol J* **19:** 3-21.

Mesbah NM, Hedrick DB, Peacock AD, Rohde M, Wiegel J (2007). *Natranaerobius thermophilus* gen. nov., sp. nov., a halophilic, alkalithermophilic bacterium from soda lakes of the Wadi An Natrun, Egypt, and proposal of *Natranaerobiaceae* fam. nov. and *Natranaerobiales* ord. nov. *Int J Syst Evol Microbiol* **57:** 2507-2512.

Meyer B, Imhoff JF, Kuever J (2007). Molecular analysis of the distribution and phylogeny of the *soxB* gene among sulfur-oxidizing bacteria - evolution of the Sox sulfur oxidation enzyme system. *Environ Microbiol* **9:** 2957-2977.

Meyer B, Kuever J (2007a). Molecular analysis of the distribution and phylogeny of dissimilatory adenosine-5'-phosphosulfate reductase-encoding genes (*aprBA*) among sulfur-oxidizing prokaryotes. *Microbiology* **153:** 3478-3498.

Meyer B, Kuever J (2007b). Phylogeny of the alpha and beta subunits of the dissimilatory adenosine-5'-phosphosulfate (APS) reductase from sulfate-reducing prokaryotes--origin and evolution of the dissimilatory sulfate-reduction pathway. *Microbiology* **153:** 2026-2044.

Meyer B, Kuever J (2007c). Molecular analysis of the diversity of sulfate-reducing and sulfur-oxidizing prokaryotes in the environment, using *aprA* as functional marker gene. *Appl Environ Microbiol* **73:** 7664-7679.

Miller LG, Jellison R, Oremland RS, Culbertson CW (1993). Meromixis in hypersaline Mono Lake, California. 3. Biogeochemical response to stratification and overturn. *Limnol Oceanogr* **38:** 1040-1051.

Mori Y, Purdy KJ, Oakley BB, Kondo R (2010). Comprehensive detection of phototrophic sulfur bacteria using PCR primers that target reverse dissimilatory sulfite reductase gene. *Microbes Environ* **25:** 190-196.

Muller AL, Kjeldsen KU, Rattei T, Pester M, Loy A (2015). Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J* **9:** 1152-1165.

Muyzer G, Stams AJ (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nature reviews Microbiology* **6:** 441-454.

Nagy CI, Vass I, Rákhely G, Vass IZ, Tóth A, Duzs Á *et al.* (2014). Coregulated genes link sulfide:quinone oxidoreductase and arsenic metabolism in *Synechocystis* sp. strain PCC6803. *J Bacteriol* **196:** 3430-3440.

Nakagawa T, Iino T, Suzuki K-i, Harayama S (2006). *Ferrimonas futtsuensis* sp. nov. and *Ferrimonas kyonanensis* sp. nov., selenate-reducing bacteria belonging to the Gammaproteobacteria isolated from Tokyo Bay. *Int J Syst Evol Microbiol* **56:** 2639-2645.

Niggemyer A, Spring S, Stackebrandt E, Rosenzweig RF (2001). Isolation and characterization of a novel As(V)-reducing bacterium: implications for arsenic mobilization and the genus Desulfitobacterium. *Appl Environ Microbiol* **67:** 5568-5580.

Oremland RS, Dowdle PR, Hoeft S, Sharp JO, Schaefer JK, Miller LG *et al.* (2000). Bacterial dissimilatory reduction of arsenate and sulfate in meromictic Mono Lake, California. *Geochim Cosmochim Acta* **64:** 3073-3084.

Oremland RS, Stolz JF, Hollibaugh JT (2004). The microbial arsenic cycle in Mono Lake, California. *FEMS Microbiol Ecol* **48:** 15-27.

Oremland RS, Saltikov CW, Wolfe-Simon F, Stolz JF (2009). Arsenic in the evolution of Earth and extraterrestrial ecosystems. *Geomicrobiol J* **26:** 522-536.

Pepi M, Volterrani M, Renzi M, Marvasi M, Gasperini S, Franchi E *et al.* (2007). Arsenic-resistant bacteria isolated from contaminated sediments of the Orbetello Lagoon, Italy, and their characterization. *J Appl Microbiol* **103:** 2299-2308.

Pérez-Jiménez JR, DeFraia C, Young LY (2005). Arsenate respiratory reductase gene (*arrA*) for *Desulfosporosinus* sp. strain Y5. *Biochem Biophys Res Commun* **338:** 825-829.

Pfeffer C, Larsen S, Song J, Dong M, Besenbacher F, Meyer RL *et al.* (2012). Filamentous bacteria transport electrons over centimetre distances. *Nature* **491:** 218-221.

Planer-Friedrich B, London J, McCleskey RB, Nordstrom DK, Wallschlager D (2007). Thioarsenates in geothermal waters of Yellowstone National Park: Determination, preservation, and geochemical importance. *Environ Sci Technol* **41:** 5245-5251.

Planer-Friedrich B, Fisher J, Hollibaugh J, Suess E, Wallschlager D (2009). Oxidative transformation of trithioarsenate along alkaline geothermal drainages—abiotic versus microbially mediated processes. *Geomicrobiol J* **26:** 339-350.

Planer-Friedrich B, Suess E, Scheinost AC, Wallschlager D (2010). Arsenic speciation in sulfidic waters: reconciling contradictory spectroscopic and chromatographic evidence. *Anal Chem* **82:** 10228-10235.

Planer-Friedrich B, Hartig C, Lohmayer R, Suess E, McCann SH, Oremland R (2015). Anaerobic chemolithotrophic growth of the haloalkaliphilic bacterium Strain MLMS-1 by disproportionation of monothioarsenate. *Environ Sci Technol* **49:** 6554-6563.

Poretsky RS, Sun S, Mou X, Moran MA (2010). Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ Microbiol* **12:** 616-627.

Rascovan N, Maldonado J, Vazquez MP, Eugenia Farias M (2015). Metagenomic study of red biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. *ISME J* **e-pub ahead of print 3 July 2015:** doi: 10.1038/ismej.2015.1109.

Rauschenbach I, Bini E, Häggblom MM, Yee N (2012). Physiological response of *Desulfurispirillum indicum* S5 to arsenate and nitrate as terminal electron acceptors. *FEMS Microbiology Ecology* **81:** 156-162.

Rothery RA, Workun GJ, Weiner JH (2008). The prokaryotic complex iron–sulfur molybdoenzyme family. *BBA-Biomembranes* **1778:** 1897-1929.

Saltikov CW, Newman DK (2003). Genetic identification of a respiratory arsenate reductase. *Proc Natl Acad Sci USA* **100:** 10983-10988.

Satinsky BM, Gifford SM, Crump BC, Moran MA (2013). Use of internal standards for quantitative metatranscriptome and metagenome analysis. In: DeLong EF (ed). *Methods in Enzymology*. Academic Press. pp 237-250.

Satinsky BM, Crump BC, Smith CB, Sharma S, Zielinski BL, Doherty M *et al.* (2014). Microspatial gene expression patterns in the Amazon River plume. *Proc Natl Acad Sci USA* **111:** 11085-11090.

Schmieder R, Edwards R (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27:** 863-864.

Schmieder R, Lim YW, Edwards R (2012). Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* **28:** 433-435.

Schoepp-Cothenet B, van Lis R, Philippot P, Magalon A, Russell MJ, Nitschke W (2012). The ineluctable requirement for the trans-iron elements molybdenum and/or tungsten in the origin of life. *Sci Rep* **2:** 263.

Scholten JCM, Joye SB, Hollibaugh JT, Murrell JC (2005). Molecular analysis of the sulfate reducing and archaeal community in a meromictic soda lake (Mono Lake, California) by targeting 16S rRNA, *mcrA*, *apsA*, and *dsrAB* Genes. *Microb Ecol* **50:** 29-39.

Seemann T (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:** 2068-2069.

Sorokin DY, Gorlenko VM, Tourova TP, Tsapin A, Nealson KH, Kuenen GJ (2002). *Thioalkalimicrobium cyclicum* sp. nov. and *Thioalkalivibrio jannaschii* sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing bacteria from hypersaline alkaline Mono Lake (California). *Int J Syst Evol Microbiol* **52:** 913-920.

Sorokin DY, Tourova T, apos, P. y, Sjollema KA, Kuenen JG (2003). *Thialkalivibrio nitratireducens* sp. nov., a nitrate-reducing member of an autotrophic denitrifying consortium from a soda lake. *Int J Syst Evol Microbiol* **53:** 1779-1783.

Sorokin DY, Banciu H, Robertson LA, Kuenen JG, Muntyan MS, Muyzer G (2013). Halophilic and haloalkaliphilic sulfur-oxidizing bacteria. In: Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F (eds). *The Prokaryotes*. Springer Berlin Heidelberg. pp 529-554.

Sorokin DY, Banciu HL, Muyzer G (2015). Functional microbiology of soda lakes. *Current Opinion in Microbiology* **25:** 88-96.

Spain AM, Elshahed MS, Najar FZ, Krumholz LR (2015). Metatranscriptomic analysis of a high-sulfide aquatic spring reveals insights into sulfur cycling and unexpected aerobic metabolism. *PeerJ* **3:** e1259.

Stewart F, Dmytrenko O, DeLong E, Cavanaugh C (2011). Metatranscriptomic analysis of sulfur oxidation genes in the endosymbiont of *Solemya velum*. *Frontiers in microbiology* **2:** 134.

Stewart FJ, Ottesen EA, DeLong EF (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* **4:** 896-907.

Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH, Consortium tU (2015). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31:** 926-932.

Tikhonova TV, Slutsky A, Antipov AN, Boyko KM, Polyakov KM, Sorokin DY *et al.* (2006). Molecular and catalytic properties of a novel cytochrome c nitrite reductase from nitrate-reducing haloalkaliphilic sulfur-oxidizing bacterium *Thioalkalivibrio nitratireducens*. *BBA-Proteins Proteom* **1764:** 715-723.

Tourova TP, Slobodova NV, Bumazhkin BK, Kolganova TV, Muyzer G, Sorokin DY (2013). Analysis of community composition of sulfur-oxidizing bacteria in hypersaline and soda lakes using soxB as a functional molecular marker. *FEMS Microbiol Ecol* **84:** 280-289.

van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B (2013). Arsenics as bioenergetic substrates. *Biochimica et biophysica acta* **1827:** 176-188.

Vila-Costa M, Rinta-Kanto JM, Sun S, Sharma S, Poretsky R, Moran MA (2010). Transcriptomic analysis of a marine bacterial community enriched with dimethylsulfoniopropionate. *ISME J* **4:** 1410-1420.

Wilkin RT, Wallschlager D, Ford RG (2003). Speciation of arsenic in sulfidic waters. *Geochem Trans* **4:** 1-7.

Zargar K, Hoeft S, Oremland R, Saltikov CW (2010). Identification of a novel arsenite oxidase gene, *arxA*, in the haloalkaliphilic, arsenite-oxidizing bacterium *Alkalilimnicola ehrlichii* strain MLHE-1. *J Bacteriol* **192:** 3755-3762.

Zargar K, Conrad A, Bernick DL, Lowe TM, Stolc V, Hoeft S *et al.* (2012). ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. *Environ Microbiol* **14:** 1635-1645.

Zhang J, Kobert K, Flouri T, Stamatakis A (2014). PEAR: a fast and accurate Illumina paired-end read merger. *Bioinformatics* **30:** 614-620.

Zhao Y, Tang H, Ye Y (2012). RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* **28:** 125-126.

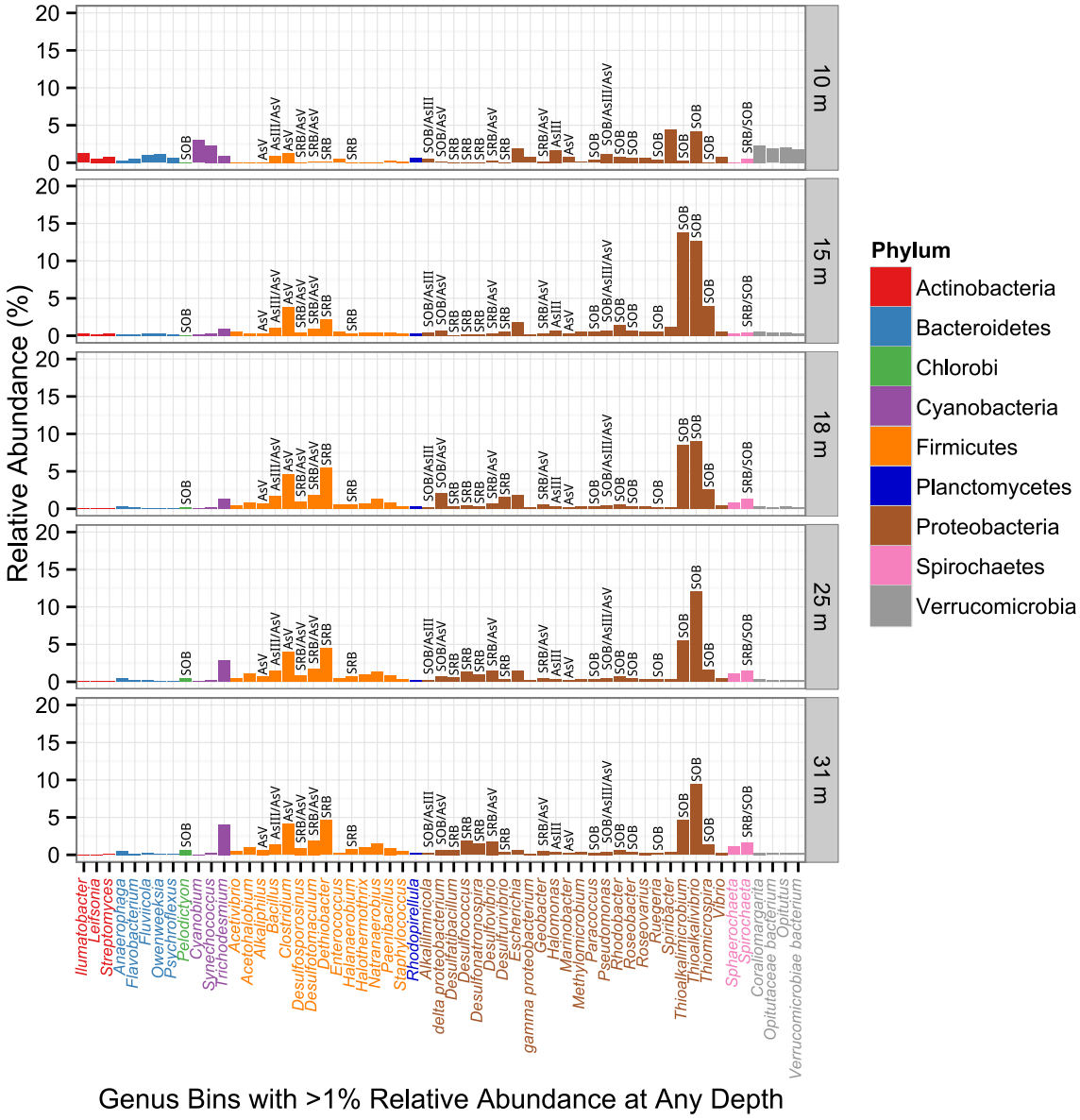


Figure 3.1. Relative abundance of genus-level transcript bins by depth. Relative abundance was calculated using the average of the absolute abundance in both transcript libraries at each depth. Genera are colored by phylum, and the text above the abundance bars indicate cultured taxa in the genera transform arsenic and sulfur compounds as follows: SOB, sulfur-oxidizing bacteria; SRB, sulfate-reducing bacteria; As(III), arsenite oxidation; As(V), arsenate reduction.

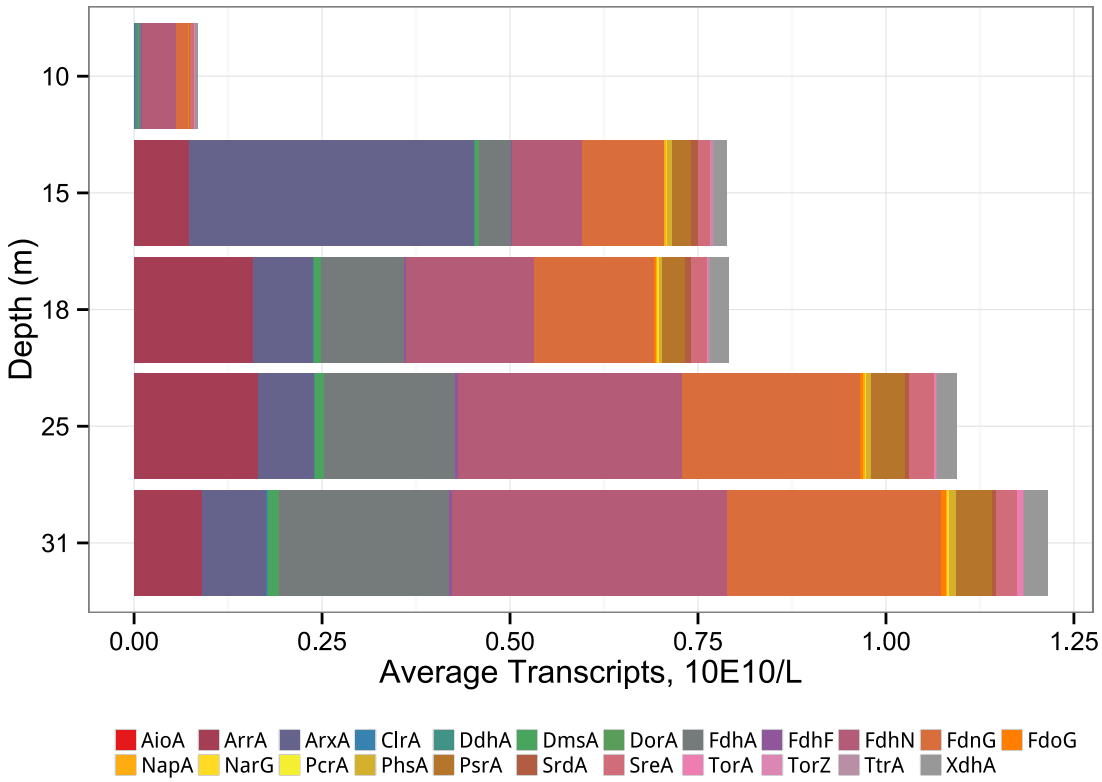


Figure 3.2. Abundance of CISM transcripts by depth. See Table 3.S1 for abbreviation definitions.

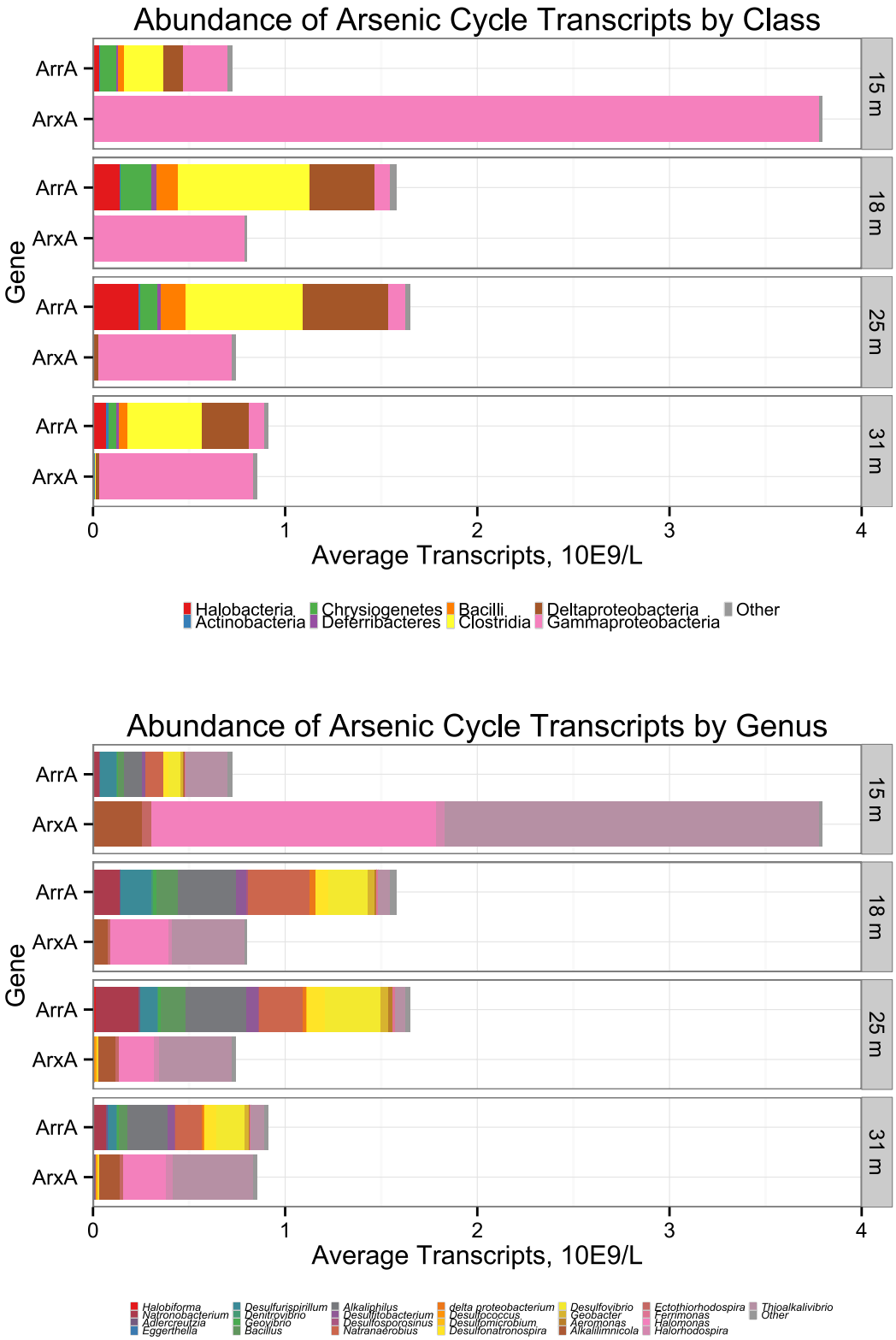


Figure 3.3. Abundance of arsenic transcripts by class and genus.

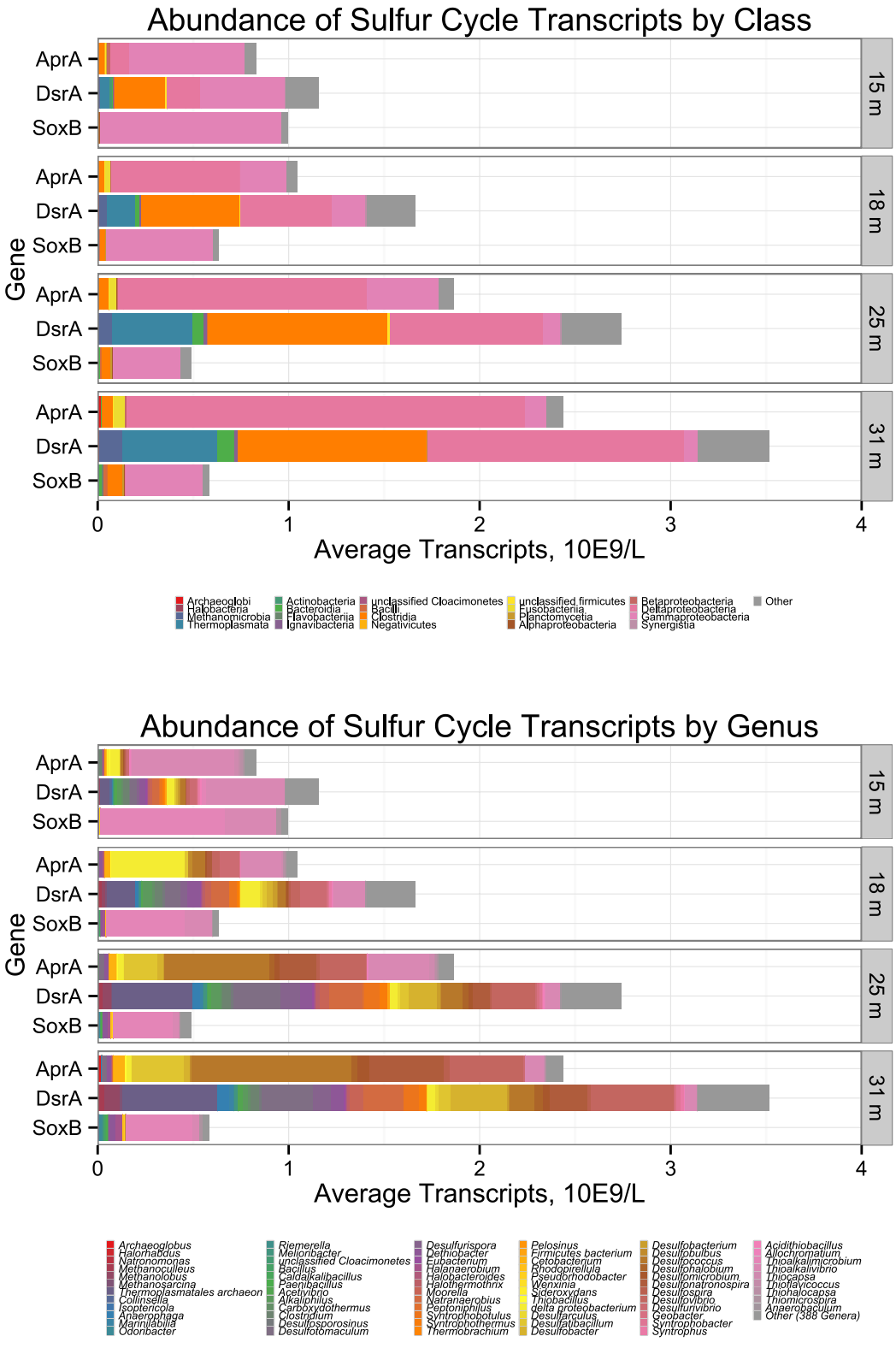


Figure 3.4. Abundance of sulfur transcripts by class and genus.

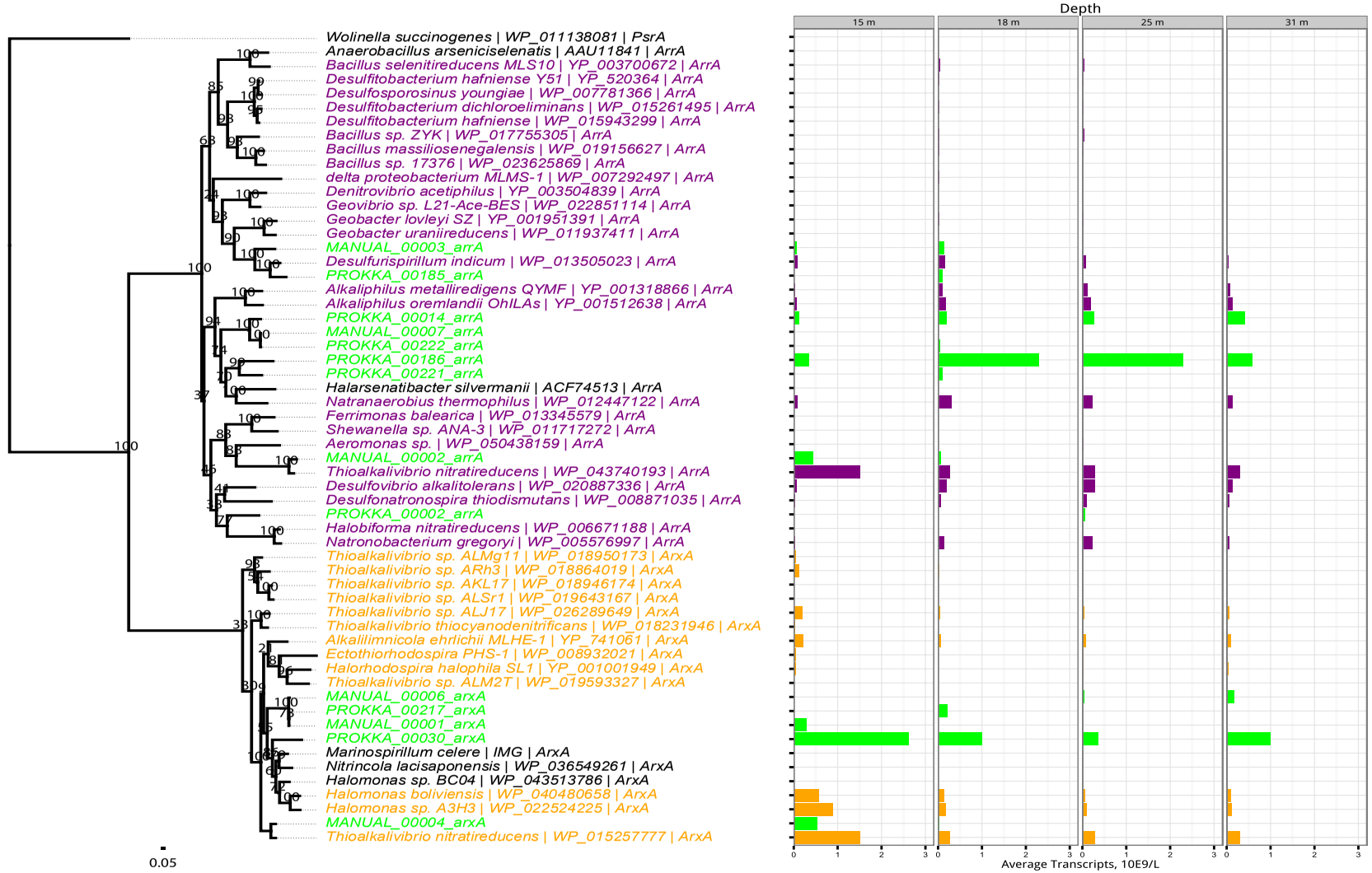


Figure 3.5. Phylogeny and abundance of arsenate reductase (ArrA*)* and arsenite oxidase (ArxA) by depth. Transcript hits (ArrA, purple; ArxA, orange) are indicated by organism and RefSeq accession number. Assembled transcripts are colored green. *Wolinella succinogenes* polysulfidereductase (PsrA) is the outgroup.

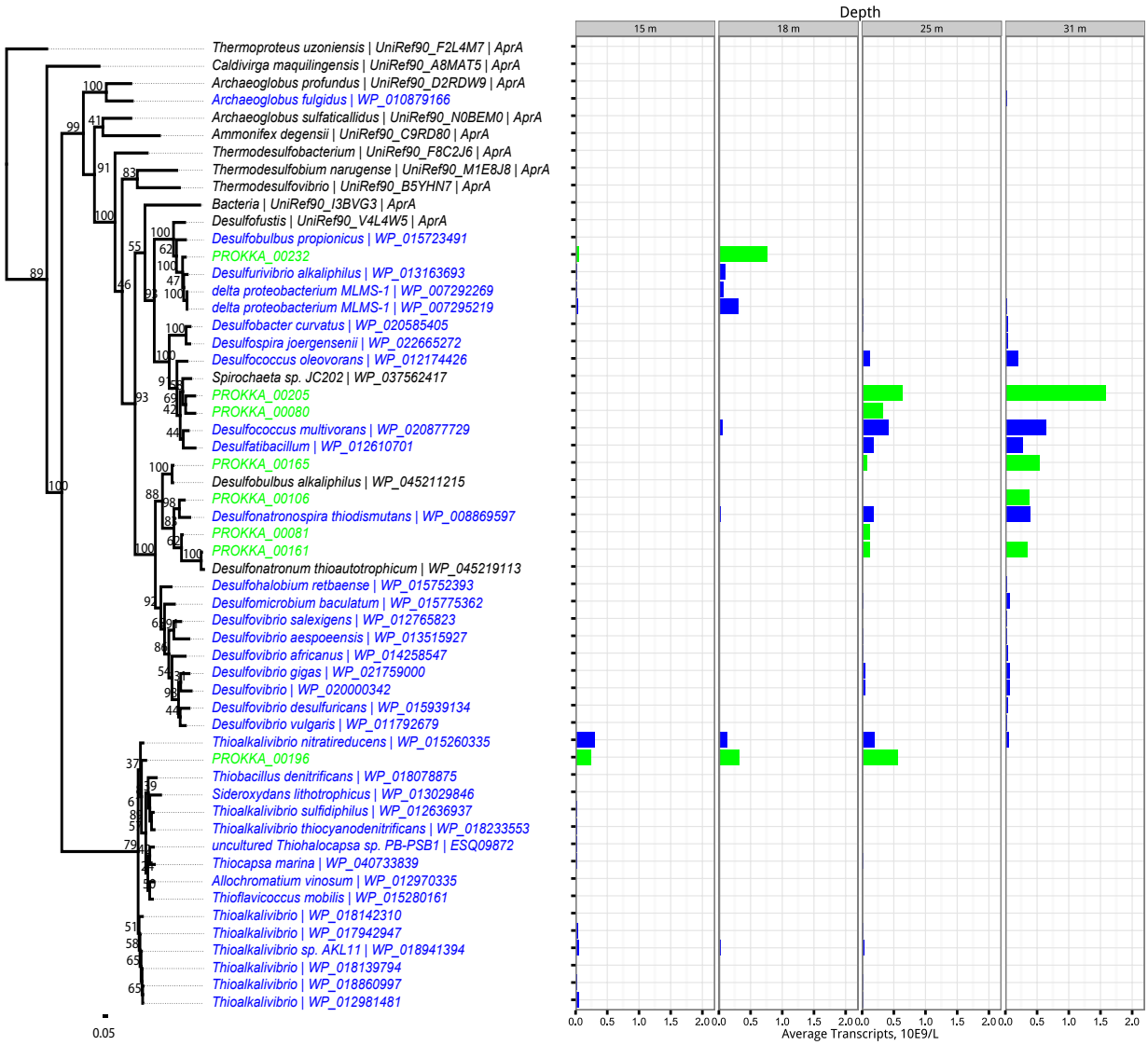


Figure 3.6. Phylogeny and abundance of adenosine-5'-phosposulfate reductase (AprA) by depth. Transcript hits (blue) are indicated by organism and RefSeq accession number. Assembled transcripts are colored green. Nodes labeled with "*AprA*" are sequences used in the reference database. *Thermoproteus uzoniensis* is the outgroup.

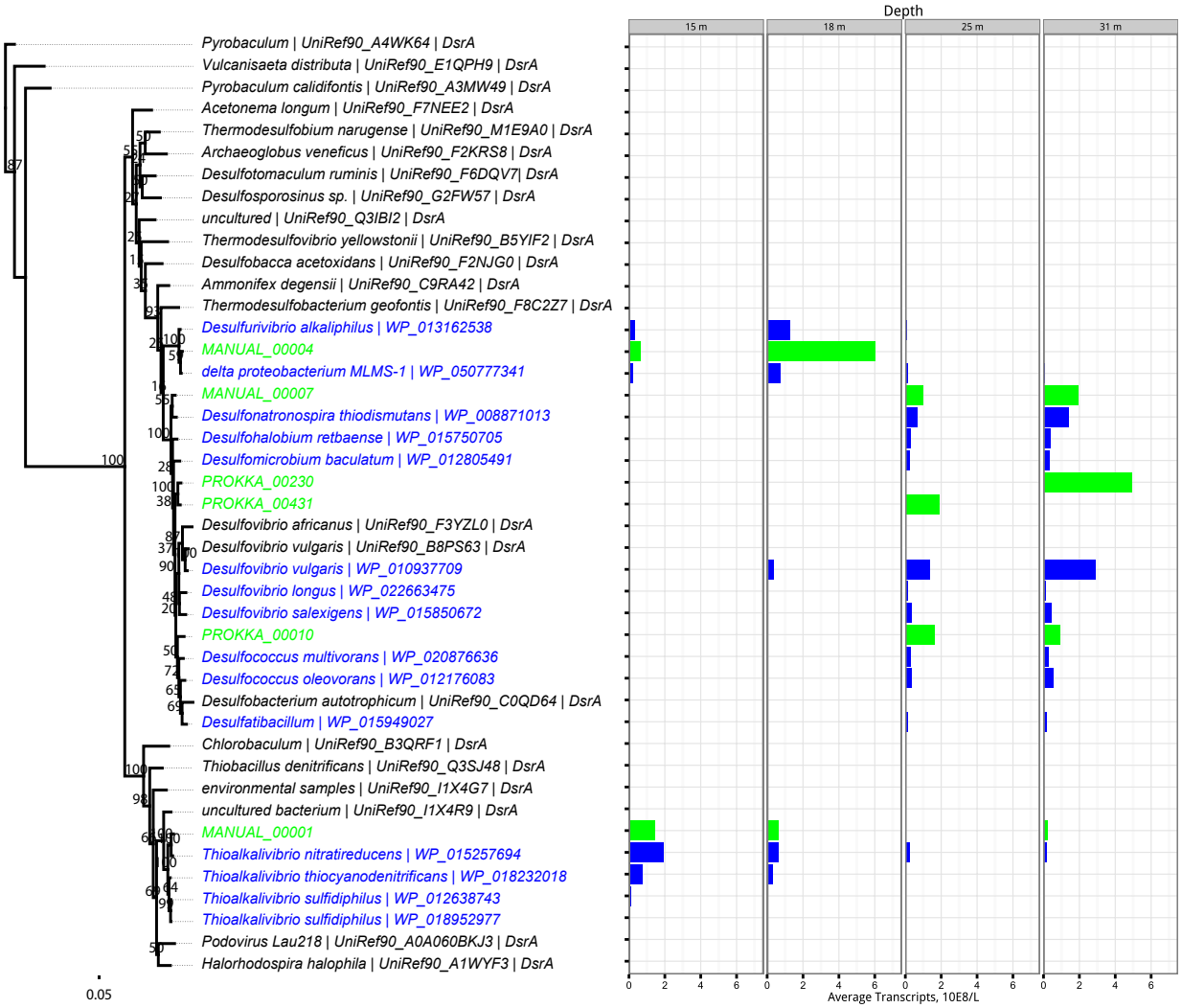


Figure 3.7. Phylogeny and abundance of dissimilatory sulfite reductase (DsrA) by depth. Transcript hits (blue) are indicated by organism and RefSeq accession number. Assembled transcripts are colored green. Nodes labeled with "*DsrA*" are sequences used in the reference database. *Pyrobaculum* is the outgroup.

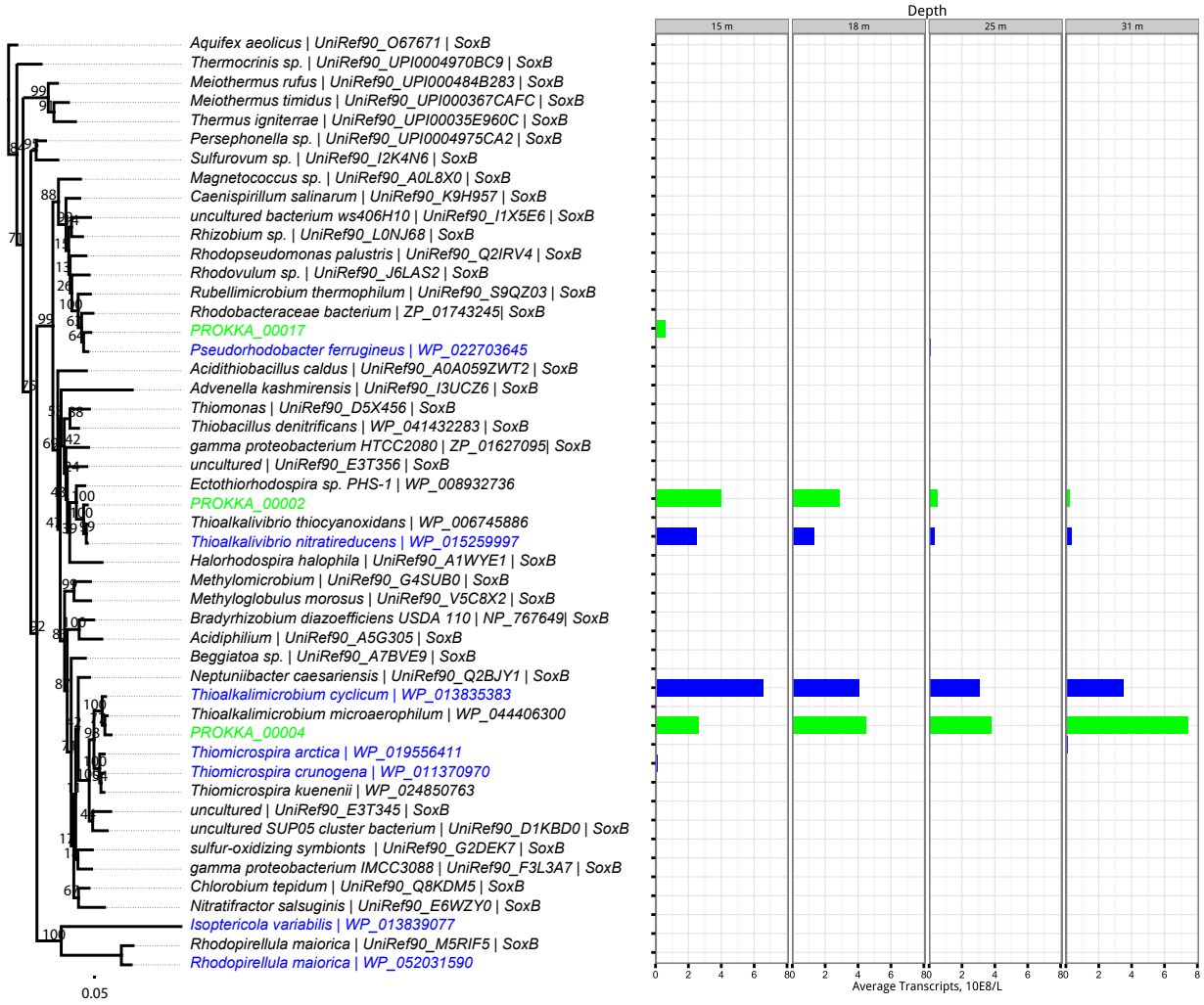


Figure 3.8. Phylogeny and abundance of SoxB by depth. Transcript hits (blue) are indicated by organism and RefSeq accession number. Assembled transcripts are colored green. Nodes labeled with "*SoxB*" are sequences used in the reference database. *Aquifex aeolicus* is the outgroup.

CHAPTER 4

TRANSFORMATION OF MONOTHIOARSENATE BY HALOALKALIPHILIC, ANOXYGENIC PHOTOSYNTHETIC PURPLE SULFUR BACTERIA[1]

[1]Edwardson C.F., B. Planer-Friedrich, and J.T. Hollibaugh. 2014. *FEMS Microbiology Ecology*. 90: 858-868. Reprinted here with permission of publisher.

Abstract

Thioarsenates are the dominant arsenic species in arsenic-rich, alkaline and sulfidic waters, but bacterial interactions with these compounds have only recently been examined. Previous studies have shown that microorganisms play a role in the transformation of monothioarsenate to arsenate, including use of monothioarsenate as a chemolithotrophic electron donor coupled with oxygen as an electron acceptor. We obtained enrichment cultures from two saline, alkaline lakes (Mono Lake, CA and Big Soda Lake, NV) that are able to use monothioarsenate as the sole electron donor for anoxygenic photosynthesis. These anoxic cultures were able to convert a 1 mM mixture of thioarsenates completely to arsenate in approximately 13 days and 4 mM monothioarsenate to arsenate in approximately 17 days. This conversion was light dependent, thus monothioarsenate can be used as the sole electron donor for anoxygenic photosynthesis. Both of the Mono Lake and Big Soda Lake enrichment cultures were dominated by an organism closely related to *Ectothiorhodospira* species. We tested additional strains of purple sulfur bacteria and found widespread ability to use monothioarsenate as an electron donor. The ability of bacteria to transform thioarsenates directly via anoxygenic photosynthesis adds a new perspective to the well-studied arsenic and sulfur cycles.

Introduction

Understanding arsenic biogeochemistry is important due to its toxicity and widespread distribution in groundwater as a result of geothermal processes and anthropogenic pollution. Biologically mediated transformations of the arsenic oxyanions arsenate (As(V)) and arsenite (As(III)), and of organoarsenicals have been well documented (Sharma and Sohn 2009). In addition, recent reports have revealed the presence of thioarsenates in the environment, especially in alkaline and sulfidic waters (Härtig and Planer-Friedrich 2012, Planer-Friedrich *et al.* 2007, Planer-Friedrich *et al.* 2009, Stauder *et al.* 2005) including the hypersaline and alkaline Mono Lake, CA, USA (Hollibaugh *et al.* 2005). In Mono Lake, periods of meromixis lead to high sulfide concentrations below the oxycline and to the accumulation of thioarsenates, which can contribute as much as half of the arsenic in the water column (Hollibaugh *et al.* 2005, Wallschlager and Stadey 2007). Microbial transformations of arsenic compounds in Mono Lake have been examined extensively (Oremland *et al.* 2004, Oremland *et al.* 2009), generally focusing on arsenite oxidation and arsenate reduction as energy-yielding reactions for bacteria (Hoeft *et al.* 2004, Hollibaugh *et al.* 2006, Oremland *et al.* 2002). Recent studies have indicated a potential biological role in the transformation of thioarsenates (Fisher *et al.* 2008, Härtig and Planer-Friedrich 2012, Planer-Friedrich *et al.* 2009). Experiments in which arsenite and sulfide were added to Mono Lake water under oxidizing conditions revealed the formation of thioarsenates, followed by the production of arsenate. Abiotic controls showed no appreciable accumulation of arsenate (Fisher *et al.* 2008). Other work under oxidizing conditions in drainage channels from sulfidic, alkaline hot springs showed that trithioarsenate, the dominant arsenic species at the source, was biologically transformed to arsenate along the channel with intermediate accumulation of arsenite and monothioarsenate (Planer-Friedrich *et al.* 2009).

Incubations of filter-sterilized geothermal water with and without the addition of filamentous microbial streamers (dominated by *Thermocrinis ruber*) showed that abiotic trithioarsenate transformation to arsenate was coupled to biotic sulfide oxidation, with an intermediate accumulation of arsenite. In sulfide-free medium, direct microbial transformation of monothioarsenate to arsenite or arsenate was shown (Härtig and Planer-Friedrich 2012). A recent experiment using a pure culture confirmed that *T. ruber* is capable of aerobic growth using monothioarsenate as an electron donor, producing arsenate and sulfate (Härtig *et al.* 2014). The mechanism for the conversion of thioarsenates to arsenate and/or arsenite have been hypothesized to involve either direct biological conversion of thioarsenate to arsenate and sulfate (equation 1) by sulfur-oxidizing bacteria; abiotic conversion to arsenite and elemental sulfur (equation 2) with subsequent biological oxidation to arsenate (equation 3) and sulfate (equation 4); or abiotic decomposition (desulfidation) (equation 5) with subsequent biological oxidation of free sulfide (equation 6) driving the chemical equilibrium toward the products arsenate and finally sulfate (Fisher *et al.* 2008, Härtig and Planer-Friedrich 2012).

$$HAs^{V}O_3S^{-II\,2-} + H_2O + 2O_2 \rightarrow HAs^{V}O_4^{2-} + S^{VI}O_4^{2-} + 2H^+ \quad (1)$$

$$HAs^{V}O_3S^{-II\,2-} + 2H^+ \rightarrow H_3As^{III}O_3 + S^0 \quad (2)$$

$$2H_3As^{III}O_3 + O_2 \rightarrow 2HAs^{V}O_4^{2-} + 4H^+ \quad (3)$$

$$2S^0 + 3O_2 + 2H_2O \rightarrow 2S^{VI}O_4^{2-} + 4H^+ \quad (4)$$

$$HAs^{V}O_3S^{-II\,2-} + OH^- \rightarrow HAs^{V}O_4^{2-} + HS^{-II\,-} \quad (5)$$

$$HS^{-II\,-} + 2O_2 \rightarrow S^{VI}O_4^{2-} + H^+ \quad (6)$$

Previous work with enrichment cultures from Mono Lake dominated by the chemolithotrophic sulfur-oxidizing bacterium *Thioalkalivibrio* sp. demonstrated the conversion of thioarsenates to arsenate (Fisher *et al.* 2008). This haloalkaliphilic member of the

Gammaproteobacteria family Ectothiorhodospiraceae has been isolated from diverse soda lakes and bioreactors (Sorokin *et al.* 2008, Sorokin *et al.* 2013). Other members of the Ectothiorhodospiraceae include the phototrophic purple sulfur bacteria (Imhoff 2006) and *Alkalilimnicola ehrlichii* MLHE-1, an arsenite-oxidizing chemolithotroph isolated from Mono Lake (Hoeft *et al.* 2007). Since phototrophic purple sulfur bacteria have previously been shown to oxidize both reduced sulfur compounds (Dahl 2008) and arsenite (Budinoff and Hollibaugh 2008, Kulp *et al.* 2008) we hypothesized that they might play a role in thioarsenate transformation. To test this hypothesis we established anaerobic enrichment cultures of anoxygenic phototrophs able to use thioarsenates for growth, using inocula from Mono Lake and Big Soda Lake, a meromictic, alkaline lake in Fallon, NV. In addition, we tested pure cultures of phototrophic purple sulfur bacteria for their ability to use thioarsenates for growth.

Materials and Methods

Field Site and Sampling

Anoxic water was collected from Station 3 (37°58.97’ N, 119°05.35’ W) (Hollibaugh *et al.* 2001) at a depth of 25 m in Mono Lake, CA, USA on 29 June 2011 and from Big Soda Lake, NV, USA (39° 31.66’ N, 118° 52.62’ W) at a depth of 20 m on 20 September 2011. The chemistry of these lakes has been described previously (Cloern *et al.* 1983, Hollibaugh *et al.* 2005). Vertical profiles of temperature, pressure, conductivity, photosynthetically active radiation (PAR), beam attenuation, in vivo fluorescence and oxygen were obtained with a Sea-Bird SBE 19 Seacat CTD equipped with a C-Star transmissometer and WETstar fluorometer (Wet Labs), PAR sensor (Licor), and dissolved oxygen meter (SBE43, Sea-Bird). Vertical profiles of arsenic and sulfur speciation in Big Soda Lake were performed by collecting samples

from discrete depths and analyzed as described below. See Figure 4.S1 for CTD and chemical depth profiles. Water was collected from a 5L Niskin bottle into polycarbonate bottles flushed with 3 volumes of flowing water and capped with no head space as described previously (Hollibaugh *et al.* 2005). Bottles were transported and stored in the dark at 4°C.

Chemicals and Chemical Analysis

All stock solutions were made using anoxic ultrapure water in an anaerobic chamber (5% H_2/95% N_2 atmosphere, Coy Laboratories). A mixture of thioarsenates was prepared by combining anoxic sodium sulfide ($Na_2S{\cdot}9H_2O$) and sodium arsenite ($NaAsO_2$) stock solutions at a 4:1 S:As molar ratio, and was used in experiments at a final concentration equivalent to 4 mM sulfide and 1 mM arsenite, unless otherwise noted. This mixture reacts spontaneously to form thioarsenic compounds, mostly trithioarsenite, which is extremely reactive and oxidizes to tri- and tetrathioarsenate. Thioarsenite oxidation can be induced by oxygen but also by elemental sulfur (e.g. found as polysulfide impurities in a sulfide standard). Tri- and tetrathioarsenate are both unstable and eventually decompose to a mixture of arsenite, arsenate, and di- and monothioarsenate (Planer-Friedrich *et al.* 2010). The thioarsenate mixture was prepared and added to medium immediately prior to inoculating cultures in all growth experiments. Monothioarsenate was synthesized as either $Na_3AsO_3S{\cdot}7H_2O$ (Suess *et al.* 2009), or $Na_3AsO_3S{\cdot}12H_2O$ (Brauer 1963). Samples for chemical analysis were collected using sterile needles and syringes, and filtered through 0.2 µm PTFE syringe filters in the anaerobic chamber. Samples for thioarsenate speciation were collected in 2 mL cryovials and immediately flash frozen in liquid nitrogen and stored at -80°C prior to analysis, when they were thawed in an anaerobic chamber. This method has been previously shown to preserve thioarsenic species (Planer-Friedrich *et al.* 2010). Total arsenic and speciation of arsenite, arsenate, and

thioarsenates was performed by ICP-MS and IC-ICP-MS, respectively, as described previously (Planer-Friedrich *et al.* 2007). Arsenic speciation in experiments with monothioarsenate amendments was analyzed by HPLC as described previously (Hoeft *et al.* 2004) using a Waters 626 HPLC with a 0.008 M H_2SO_4 eluent (pH ~2). This protocol enables detection of arsenate, monothioarsenate, and arsenite (retention times of 12.8, 14, and 17.5 minutes, respectively) but not of the more thiolated arsenic species. Detection was by absorbance at 200 nm. Others (Hoeft *et al.* 2004) have measured arsenate and arsenite at 210 nm, but we obtained better sensitivity at 200 nm. Samples for sulfide analysis were collected by 0.2 µm filtration and precipitation with zinc acetate. Free sulfide was measured spectrophotometrically by the methylene blue method (Cline 1969).

Enrichment Cultures

Water collected from Mono Lake as described above was diluted 1:10 in 0.2 µm-filtered Mono Lake water amended with the 4:1 S:As thioarsenate mixture at a final concentration equivalent to 2 mM sulfide and 0.5 mM arsenite. This water was dispensed into 150 mL serum bottles and incubated anaerobically with incandescent light at room temperature (22-25 ^{o}C). Enrichments were transferred as 1:10 dilutions into the same medium monthly for three months and then transferred as either 1:10 or 1:100 dilutions into an Artificial Mono Lake Water (AMLW) medium containing the following (g L^{-1}): NaCl (60), $(NH_4)_2SO_4$ (0.5), $MgSO_4 \cdot 7H_2O$ (0.25), KCl (1.7), $Na_2B_4O_7 \cdot H_2O$ (2), Na_2SO_4 (16.5), KH_2PO_4 (0.25), K_2HPO_4 (0.5) and 900 mL ultrapure water. After the medium was autoclaved and cooled, 100 mL of a filter-sterilized solution of Na_2CO_3 (106 g L^{-1}) and $NaHCO_3$ (42 g L^{-1}) (150 mM final concentration in the AMLW medium), 10 mL L^{-1} of vitamin solution (Oremland *et al.* 1994), and 1 mL L^{-1} of unchelated trace elements SL-10 (Widdel *et al.* 1983) were added aseptically. The final pH of the

medium was 9.3-9.5. The addition of basic sulfide and arsenite stock solutions to the AMLW did not raise the pH above 9.55 because of the carbonate-bicarbonate buffer in the medium. Anoxic media was obtained by storing the loosely capped container in the anaerobic chamber for a minimum of 48 hours. Enrichment cultures were transferred every 2-4 weeks into fresh media amended with the thioarsenate mixture. Anoxic water from Big Soda Lake was amended with 0.1 mM sulfide and 0.1 mM arsenite. This enrichment culture was incubated at room temperature in the light for a month and then fed with the thioarsenate mixture. Once growth was apparent, the culture was transferred 1:1 to AMLW, and then 1:10 and 1:100 into AMLW, with the thioarsenate mixture added at each transfer. Growth was monitored by observing increases in pink to dark red pigmentation and an increase of cells monitored by epifluorescence microscopy (DAPI) (Porter and Feig 1980) or optical density measurements at 600 nm (OD_{600}) (Eppendorf BioPhotometer™).

Thioarsenate Amendment Experiments

All amendment experiments were set up and sampled in an anaerobic chamber using 150 mL serum bottles with black butyl rubber stoppers and aluminum crimp seals. All incubations were at 30 °C. Phototrophic organisms were incubated under fluorescent grow lights (three GE #49892 15W bulbs). An array of LEDs emitting at 850 nm (IR12-850 x 4, total ~2.4 W, EnvironmentalLights.com) was added to the incubator for some experiments and experiments using this setup are noted. One of our Mono Lake Enrichment cultures (EC8) was used for growth experiments. To compare growth of the enrichment culture on thioarsenates vs. sulfide alone, a late-log phase culture was diluted 1:25 into triplicate bottles of AMLW that were amended with either the thioarsenate mixture or 4 mM sulfide. Negative controls were prepared in triplicate as described above but inoculated with 0.2 µm-filtered, late-log phase culture

medium, however one replicate of the thioarsenate amended negative controls was contaminated and discarded. A late-log phase culture was diluted 1:25 into triplicate bottles of AMLW that were amended with either 4 mM monothioarsenate or 4 mM thiosulfate ($S_2O_3^{2-}$) to compare growth of the enrichment culture on monothioarsenate vs. thiosulfate. Triplicate negative controls were prepared as described above; however, one replicate of the monothioarsenate abiotic control was discarded due to contamination. We tested the dependence of growth and monothioarsenate oxidation on photoautotrophy by comparing replicates of a late-log phase culture diluted 1:25 into six bottles of AMLW. Two replicates were incubated in the dark and four replicates were incubated in light (grow lights plus IR LEDs). Growth was monitored by OD_{600} and monothioarsenate conversion to arsenate was monitored by HPLC as described above. Once growth reached mid-exponential phase, two of the light-incubated cultures were placed in the dark for the duration of the experiment.

Additional cultures were obtained from the sources listed in Table 4.S1 and grown as described there. Cultures were given an electron donor (1-5 mM monothioarsenate or 4:1 S:As thioarsenate mixture as described in Materials and Methods) at the time of inoculation. We tested the ability of these cultures to grow on thioarsenates by inoculating (1:25 dilution) fresh medium containing the electron donor being tested with a culture grown to late exponential phase, then following substrate conversion (HPLC) while monitoring growth (OD_{600}). These cultures were incubated at 30 °C in the light as described above or in the dark with nitrate (5 mM) as the electron acceptor if the organism was not phototrophic. *Thioalkalivibrio jannaschii* was incubated aerobically.

Composition of Enrichment Cultures

Two sets of 16S rRNA gene clone libraries were prepared and sequenced to determine the composition of the enrichment cultures. Genomic DNA was extracted from enrichment culture Mono Lake EC8 after 6-7 transfers during the initial selection phase. Cells from 1.9 mL of the culture were harvested by centrifugation at 12,000 RPM (10,700 x g) for 5-10 minutes. The pellet was resuspended in 1.8 mL of lysis buffer, then DNA was extracted and purified as described previously (Ferrari and Hollibaugh 1999). In addition, cells from Mono Lake (EC "ML Ecto," EC8, and EC12) and Big Soda Lake (EC13) enrichment cultures were harvested as described above to compare the composition of the cultures after 20-30 transfers. We extracted and purified DNA from these cells using the PureLink® Genomic DNA kit (Invitrogen) as per manufacturer's instructions for Gram-positive bacteria. PCR was performed in duplicate or triplicate (depending on the library) with OneTaq® DNA Polymerase (NEB) using Bacteria-specific 27f (5'-AGAGTTTGATCMTGGCTCAG-3') and universal 1492r (5'-GGTTACCTTGTTACGACTT-3') 16S rRNA gene primers (Lane 1991, Turner *et al.* 1999). The PCR program used for the initial EC8 clone libraries was as follows: initial denaturation at 94 °C for 30 s, 35 cycles of denaturation at 94°C for 30 s, annealing at 62°C for 30 s, and extension at 68°C for 30 s, with a final extension of 45 minutes (Humayoun *et al.* 2003). The PCR program used for the second set of clone libraries increased the initial denaturation to 3 minutes, decreased the annealing temperature to 48°C for decreased stringency (Frank *et al.* 2008), and decreased the final extension to only 10 minutes. PCR products were pooled and run on a 1% agarose gel. Bands corresponding to the expected product size were cut and DNA was extracted from the gel with QIAquick® Gel Extraction Spin Kit (Qiagen). The amplicons were cloned into pCR®-4-TOPO® vector using the TOPO TA Cloning® Kit for Sequencing (Invitrogen).

Colonies containing cloned inserts were picked and grown in Luria-Bertani (LB) medium with ampicillin, then inserts were sequenced by GeneWiz, Inc. Sequences were checked for chimeras with DECIPHER (Wright *et al.* 2012) and Bellerophon (Huber *et al.* 2004) and chimeric sequences were discarded. Sequences were edited and aligned, then consensus sequences were generated for alignments that were >99% similar using Geneious R7 (Biomatters, Ltd). Consensus and reference sequences were aligned using SINA (Pruesse *et al.* 2012). The SINA alignment was imported into Geneious and a neighbor-joining (Jukes-Cantor) tree was constructed with 100 bootstraps (Figure 4.S2). A BLASTN analysis (Altschul *et al.* 1990) against the NCBI nr and RefSeq databases was performed on the remaining sequences (Table 4.S2). Sequences have been deposited in GenBank under accession numbers KM070827 - KM070952.

Results

Comparison of Growth on Thioarsenates vs. Sulfide

Analysis of arsenic speciation in a Mono Lake enrichment culture (EC8) amended with a mixture of 4 mM sulfide and 1 mM arsenite (thioarsenates mixture) or 4 mM sulfide only revealed that thioarsenates were present in the medium prior to the addition of the inoculum (t = -2 hours) and that trithioarsenate and arsenate concentrations increased for the first 21 hours after inoculation (Figure 4.1a). Arsenite concentration decreased rapidly during the first 2 hours prior to adding the inoculum then more slowly over the next 211 hours. Trithioarsenate was the dominant thioarsenate compound formed, with mono-, di- and tetrathioarsenate present as minor components. Concentrations of all arsenic compounds were relatively stable between 21 and 89 hours. Trithioarsenate concentrations decreased between 89 and 211 hours, coincident with increases in mono- and dithioarsenate. This was followed by the disappearance of thioarsenates

between 211-259 hours accompanied by a large increase in arsenate concentration. All added arsenite was completely converted to arsenate (1000 ± 35 μM) over the course of the experiment. Thioarsenate species formed initially in the abiotic control as described above, but speciation changed relatively little over the course of the incubation (Figure 4.1b). Initial sulfide concentrations were lower in the thioarsenate amendment compared to the sulfide-only amendment (Figure 4.2). Sulfide concentrations began to decrease steadily after 21 hours in both the thioarsenate and sulfide-only amendments. The time course suggests that the thioarsenate-amended culture was using free sulfide in the thioarsenate mixture preferentially during the first 68 hours of the experiment, since the concentration of sulfide in the thioarsenate amendment decreased during the first 68 hours (Figure 4.3) of the incubation, whereas the concentration of thioarsenates remained constant during that time period (Figure 4.1a). Sulfide and trithioarsenate concentrations decreased after 68 hours, while di- and monothioarsenate concentrations increased until 211 hours when di- and monothioarsenate began to be converted to arsenate. Sulfide concentrations in the sulfide-only amendment changed similarly to their pattern in the thioarsenates amendment, but did not decrease as quickly between 21 and 68 hours. Sulfide concentrations decreased only slightly in abiotic controls. The growth rate of the enrichment culture was similar when amended with the thioarsenate mixture or with sulfide alone (Figure 4.3).

Comparison of Growth on Monothioarsenate vs. Thiosulfate

A Mono Lake enrichment culture (EC8) was amended with either 4 mM pure monothioarsenate or 4 mM thiosulfate in sulfide-free AMLW. The culture appeared to prefer thiosulfate to monothioarsenate, as there was a much shorter lag prior to the onset of exponential growth and the culture grew faster on thiosulfate than on monothioarsenate. The enrichment

culture grew exponentially on monothioarsenate after an initial lag of 165 hours. Monothioarsenate was completely converted to arsenate in 290 hours (total incubation of 455 hours; Figure 4.4). There was no change in monothioarsenate or arsenate concentration in the abiotic controls during the incubation. Arsenite concentrations were initially less than 65 µM in both the enrichment culture and abiotic controls and decreased to below the limit of detection over the course of the experiment.

Light-Dependence of Photoautotrophic Monothioarsenate Oxidation

We compared growth of the Mono Lake enrichment culture (EC8) in light and dark incubations to verify that the oxidation of monothioarsenate was a phototrophic process. Monothioarsenate was only oxidized when enrichment cultures were illuminated. The concentration of monothioarsenate did not change when cultures were incubated in the dark (Figure 4.5a). The concentration of monothioarsenate decreased, with concomitant production of arsenate (Figure 4.5b) and growth (Figure 4.5c), when cultures were incubated in the light. Monothioarsenate oxidation, arsenate production and growth ceased when light-incubated cultures were placed in the dark.

Monothioarsenate Oxidation by Additional Mono Lake and Big Soda Lake Enrichment Cultures

We tested additional enrichment cultures from Mono Lake (EC12) and Big Soda Lake (EC13) that had been selected for growth on thioarsenates, for their ability to grow on monothioarsenate only. We also tested the enrichment culture containing *Ectothiorhodospira* sp. ML Ecto (designated EC "ML Ecto"), established and maintained in our lab (Budinoff and Hollibaugh 2008) and originally selected for its ability to grow photoautotrophically on arsenite. All of the enrichments were able to grow photoautotrophically by oxidizing sulfide, thiosulfate, monothioarsenate and a mixture of thioarsenates (Table 4.2). All of the enrichment cultures,

including EC8, were also able to grow by oxidizing arsenite photoautotrophically, but growth was very slow (weeks to months) with a long lag before growth and oxidation began (data not shown), as first described for EC "ML Ecto" (Budinoff and Hollibaugh 2008, Kulp *et al.* 2008).

Composition of Enrichment Cultures

Ninety-four cloned amplicons from an early transfer of EC8 were sequenced yielding 83 usable sequences. Forty six (55.4%) of these shared >99 % identity to the *Ectothiorhodospira variabilis* strain WN22 (Gorlenko *et al.* 2009) 16S rRNA gene (NR_042700). The remaining sequences consisted of Firmicutes (28, 33.7%), Bacteroidetes (8, 9.6%) and Spirochetes (1, 1.2%). Their top blastn hits are listed in Table 4.S1. Additional libraries from EC8, constructed after the experiments shown in Figure 4.1 (21 transfers) and Figure 4.4 (27 transfers) were conducted, contained only sequences that were >99% identical to the *E. variabilis* WN22 16S rRNA gene. Additional libraries from other Mono Lake (EC12 and EC "ML Ecto") and Big Soda Lake (EC13) enrichment cultures also contained *Ectothiorhodospira* sp. as the dominant organism. The phylogenetic relationships of the dominant *Ectothiorhodospira* sp. in the enrichment cultures to pure cultures of members of the family Ectothiorhodospiraceae are shown in Figure 4.S2.

Growth of Pure Cultures on Monothioarsenate

We tested a number of pure cultures from the family Ectothiorhodospiraceae and one member of the Chromatiaceae (see Table 4.S1) for their ability to grow autotrophically by oxidizing monothioarsenate. This capability appears to be widespread among the Ectothiorhodospiraceae, but not universal (Table 4.2). *Halorhodospira halophila*, *Ectothiorhodospira shaposhnikovii*, *Ectothiorhodospira* sp. PHS-1 and *Ectothiorhodospira* sp. Bogoria Red were all able to grow in media containing only monothioarsenate as an electron

donor, while converting monothioarsenate to arsenate. The chemolithotrophic *Thioalkalivibrio jannaschii* oxidized monothioarsenate aerobically but did not appear to grow (no increase in OD_{600}). The chemolithotrophic *Alkalilimnicola ehrlichii* did not oxidize monothioarsenate or thiosulfate with nitrate as an electron acceptor, and *Halorhodospira abdelmalekii* was the lone phototrophic Ectothiorhodospiraceae tested that was not able to oxidize monothioarsenate or thiosulfate photoautotrophically. Additionally, we tested the well-studied purple sulfur bacterium from the family Chromatiaceae, *Allochromatium vinosum*, for its ability to oxidize thioarsenates. When amended with the thioarsenate mixture, there was a decrease in total sulfur and an increase in monothioarsenate and arsenate, and growth. However, *A. vinosum* was not able to grow on monothioarsenate or to oxidize it to arsenate (Table 4.2).

Discussion

Thioarsenates as an Electron Donor for Anoxygenic Photosynthesis

An enrichment culture from Mono Lake was shown to be able to grow on both a mixture of thioarsenates (Figure 4.1-4.3) and solely on monothioarsenate (Figure 4.4). Comparison of the enrichment culture with the abiotic control shows that decomposition of thioarsenates is driven by bacterial activity, as there is relatively little change in speciation of the thioarsenates in the abiotic control. Growth of the enrichment culture on sulfide and on the thioarsenate mixture was comparable (Figure 4.3), with a larger final OD in the sulfide-grown culture, potentially due to a difference in cell morphology or cell clumping. Sulfide concentrations were higher in enrichment cultures amended with sulfide only versus thioarsenates. The difference between treatments may reflect rapid formation of stable monothioarsenate or transient formation of insoluble elemental sulfur or arsenic-sulfur (e.g. orpiment) compounds that would have been removed by filtration prior to sulfide analysis. This experiment showed that the photoautotrophic enrichment culture

was responsible for the conversion of thioarsenates to arsenate, similarly to previous results from experiments with an aerobic, non-phototrophic enrichment culture (Fisher *et al.* 2008).

In the first set of experiments it was unclear whether the enrichment culture was oxidizing thioarsenates directly or if it was oxidizing free sulfide present in the culture, driving the equilibrium toward the abiotic desulfidation of the thioarsenates, which (except for monothioarsenate) are unstable (Härtig and Planer-Friedrich 2012, Planer-Friedrich *et al.* 2010). Thus, we examined growth of the enrichment culture on a single thioarsenate compound (monothioarsenate) and compared the time course of growth on this substrate with growth on thiosulfate because of the similarities in the structures of the two sulfur-containing compounds (Table 4.1). Our results indicate direct transformation of monothioarsenate to arsenate. Previous experiments of monothioarsenate with *T. ruber* under aerobic conditions and high temperature (80°C) had shown that abiotic decomposition of monothioarsenate was relevant (Härtig *et al.* 2014). However, abiotic decomposition of monothioarsenate did not occur under the conditions of our incubations. In addition, abiotic oxidation of monothioarsenate to arsenite and elemental sulfur (followed by the oxidation of those compounds to arsenate and sulfate) was observed in the *T. ruber* experiments. We saw no increase in arsenite concentrations in both biotic and abiotic treatments in our experiments.

Potential Mechanism of Phototrophic Growth on Thioarsenates

Our data suggest that direct oxidation of monothioarsenate may be facilitated by the thiosulfate oxidation (*sox*) pathway. This pathway has been well studied in a related member of the family Chromatiaceae, *A. vinosum*. The pathway involves oxidation of thiosulfate and attachment to the SoxYZ enzyme by SoxXAK, followed by hydrolysis and release of sulfate from the SoxYZ complex by the SoxB enzyme, and transfer of the sulfane sulfur from SoxYZ to

sulfur globules (Dahl 2008, Weissgerber *et al.* 2011). Further conversion of the sulfur globules to sulfate proceeds by a pathway discussed elsewhere (Dahl 2008). Due to the similarity of the structures of monothioarsenate and thiosulfate (Table 4.1), we hypothesized that Sox enzymes could cleave the thiol group from monothioarsenate. Among the Ectothiorhodospiraceae with sequenced genomes, *Halorhodospira halophila* (Dahl 2008), *Ectothiorhodospira* sp. PHS-1 (Zargar *et al.* 2012), *Ectothiorhodospira haloalkaliphila* (JGI), *Ectothiorhodospira shaposhnikovii* DSM 2111 (Terrence Meyer, personal communication), and a number of *Thioalkalivibrio* species (JGI), all contain *sox* genes and many have been shown here to be able to oxidize monothioarsenate. Of the strains we tested that could not oxidize monothioarsenate, *Alkalilimnicola ehrlichii* only contains *soxY* and *soxZ* genes (Hoeft *et al.* 2007). We used degenerate PCR primers (Meyer *et al.* 2007) to test for the presence of *soxB* in the enrichment cultures and the strains for which we had no genome sequences. Only *Halorhodospira abdelmalekii* did not yield a PCR product (data not shown), and this organism was unable to oxidize monothioarsenate. However, if a complete *sox* pathway is all that is required, presumably *A. vinosum* should have been able to oxidize monothioarsenate. Other factors, such as resistance to arsenic toxicity or the differences in pH and salt requirements between *A. vinosum* and the *Ectothiorhodospira* strains we tested, could play a role in their ability to oxidize monothioarsenate.

Relevance of Thioarsenate Oxidation to Mono Lake and Other Sulfidic Environments

Biological oxidation (as well as reduction, which we have not explored here) of monothioarsenate and other thioarsenates is relevant to arsenic geochemistry as it is an underexplored transformation of environmentally relevant arsenic and sulfur compounds that could affect their toxicity (Planer-Friedrich *et al.* 2008) or mobility (Stucker *et al.* 2013).

Humayoun and colleagues (Humayoun *et al.* 2003) detected many potentially sulfur-oxidizing Gammaproteobacteria related to *Thioalkalivibrio* throughout the water column of Mono Lake. The primers used in that study only cover about 25% of the diversity of the *Ectothiorhodospira* (determined by examining 16S rRNA gene reference sequences with TestPrime using the Silva SSU r117 database (Klindworth *et al.* 2013)), so it is possible that *Ectothiorhodospira* are an even larger portion of the microbial community in Mono Lake than previously indicated (Humayoun *et al.* 2003). Our work has shown that both chemolithotrophic and photoautotrophic processes may impact the concentrations of thioarsenic compounds in natural waters.

Conceptual Model of Thioarsenate Oxidation to Arsenate

Aerobic oxidation of monothioarsenate (equation 1), may occur in the oxic upper layers of Mono Lake mediated by aerobic sulfide oxidizing bacteria as previously described (Fisher *et al.* 2008). Anaerobic oxidation of monothioarsenate by purple sulfur bacteria would be more likely in deeper and more sulfidic anoxic waters. Our interpretation of the oxidation of the thioarsenate mixture by the enrichment culture is that biological oxidation of free sulfide occurs preferentially to thioarsenate oxidation. As free sulfide is oxidized biotically, the medium becomes sulfide-deficient, leading to abiotic desulfidation of tetra-, tri-, and dithioarsenates, leaving monothioarsenate which is more stable under these conditions (Härtig and Planer-Friedrich 2012, Planer-Friedrich *et al.* 2010). Monothioarsenate oxidation then occurs directly (equation 5), potentially due to direct cleavage by an enzyme in the thiosulfate oxidation pathway (*sox*) or by a similar mechanism involving an unknown enzyme or pathway. An alternative hypothesis is that monothioarsenate decomposes abiotically to arsenite and elemental sulfur (equations 2-4), as shown in experiments with *T. ruber* (Härtig *et al.* 2014) where arsenite and elemental sulfur are oxidized sequentially. More rapid growth of enrichment cultures and pure *Ectothiorhodospira*

strains on thiosulfate and monothioarsenate versus arsenite suggests that they prefer sulfur-containing compounds over arsenite (data not shown). In addition, we did not see an increase in arsenite concentration in any of our experiments. Therefore, decomposition of monothioarsenate seems less likely than direct oxidation of the sulfur group on monothioarsenate

In conclusion, we have shown that many Ectothiorhodospiraceae are capable of photoautotrophic (and perhaps chemoautotrophic) growth on thioarsenic compounds (especially monothioarsenate). This process could play an important role in the transformation of arsenic and sulfur compounds in sulfidic environments.

Acknowledgements

We would like to thank Chris Abin for synthesis of monothioarsenate and for fruitful discussions. The assistance of Robert Jellison and others at Sierra Nevada Aquatic Research Laboratory in sampling Mono Lake is greatly appreciated. We thank Ron Oremland, Chad Saltikov and others at USGS for providing field support for sampling Big Soda Lake. Ron Oremland, Christiane Dahl, Chad Saltikov and Mike Madigan supplied some of the strains used in this study. We would also like to thank the helpful comments from two anonymous reviewers that helped improved the manuscript. This research was supported by an award from the US National Science Foundation (NSF EAR 09-52271 to JTH) and an Emmy Noether grant from the German Research Foundation to Britta Planer-Friedrich (PL 302/3-1).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215:** 403-410.

Brauer G (1963). *Handbook of Preparative Inorganic Chemistry*, 2nd edn, vol. 1. Academic Press: New York.

Budinoff CR, Hollibaugh JT (2008). Arsenite-dependent photoautotrophy by an *Ectothiorhodospira*-dominated consortium. *ISME J* **2:** 340-343.

Cline JD (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnol Oceanogr* **14:** 454-458.

Cloern JE, Cole BE, Oremland RS (1983). Autotrophic processes in meromictic Big Soda Lake, Nevada. *Limnol Oceanogr* **28:** 1049-1061.

Cordero B, Gomez V, Platero-Prats AE, Reves M, Echeverria J, Cremades E *et al.* (2008). Covalent radii revisited. *Dalton Trans***:** 2832-2838.

Dahl C (2008). Inorganic sulfur compounds as electron donors in purple sulfur bacteria. In: Hell R, Dahl C, Knaff D, Leustek T (eds). *Sulfur Metabolism in Phototrophic Organisms*. Springer: Netherlands. pp 289-317.

Ferrari VC, Hollibaugh JT (1999). Distribution of microbial assemblages in the Central Arctic Ocean Basin studied by PCR/DGGE: analysis of a large data set. *Hydrobiologia* **401:** 55-68.

Fisher JC, Wallschlager D, Planer-Friedrich B, Hollibaugh JT (2008). A new role for sulfur in arsenic cycling. *Environ Sci Technol* **42:** 81-85.

Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ (2008). Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* **74:** 2461-2470.

Gorlenko VM, Bryantseva IA, Rabold S, Tourova TP, Rubtsova D, Smirnova E *et al.* (2009). *Ectothiorhodospira variabilis* sp. nov., an alkaliphilic and halophilic purple sulfur bacterium from soda lakes. *Int J Syst Evol Microbiol* **59:** 658-664.

Härtig C, Planer-Friedrich B (2012). Thioarsenate transformation by filamentous microbial mats thriving in an alkaline, sulfidic hot spring. *Environ Sci Technol* **46:** 4348-4356.

Härtig C, Lohmayer R, Kolb S, Horn MA, Inskeep WP, Planer-Friedrich B (2014). Chemolithotrophic Growth of the Aerobic Hyperthermophilic Bacterium Thermocrinis ruber OC 14/7/2 on Monothioarsenate and Arsenite. *FEMS Microbiology Ecology***:** n/a-n/a.

Hoeft SE, Kulp TR, Stolz JF, Hollibaugh JT, Oremland RS (2004). Dissimilatory arsenate reduction with sulfide as electron donor: Experiments with mono lake water and isolation of strain MLMS-1, a chemoautotrophic arsenate respirer. *Appl Environ Microbiol* **70:** 2741-2747.

Hoeft SE, Blum JS, Stolz JF, Tabita FR, Witte B, King GM *et al.* (2007). *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. *Int J Syst Evol Microbiol* **57:** 504-512.

Hollibaugh J, Carini S, Gurleyuk H, Jellison R, Joye S, Lecleir G *et al.* (2005). Arsenic speciation in Mono Lake, California: Response to seasonal stratification and anoxia. *Geochim Cosmochim Acta* **69:** 1925-1937.

Hollibaugh JT, Wong PS, Bano N, Pak SK, Prager EM, Orrego C (2001). Stratification of microbial assemblages in Mono Lake, California, and response to a mixing event. *Hydrobiologia* **466:** 45-60.

Hollibaugh JT, Budinoff C, Hollibaugh RA, Ransom B, Bano N (2006). Sulfide oxidation coupled to arsenate reduction by a diverse microbial community in a Soda Lake. *Appl Environ Microbiol* **72:** 2043-2049.

Huber T, Faulkner G, Hugenholtz P (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20:** 2317-2319.

Humayoun SB, Bano N, Hollibaugh JT (2003). Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl Environ Microbiol* **69:** 1030-1042.

Imhoff J (2006). The family Ectothiorhodospiraceae. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (eds). *The Prokaryotes*. Springer: New York. pp 874-886.

Kempa PB, Wiebcke M, Felsche J (1990). Structure of trisodium monothioarsenate dodecahydrate. *Acta Crystallogr, Sect C* **46:** 729-732.

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

Kulp TR, Hoeft SE, Asao M, Madigan MT, Hollibaugh JT, Fisher JC *et al.* (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science* **321:** 967-970.

Lane DJ (1991). 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds). *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons: New York, NY. pp 115-175.

Meyer B, Imhoff JF, Kuever J (2007). Molecular analysis of the distribution and phylogeny of the *soxB* gene among sulfur-oxidizing bacteria - evolution of the Sox sulfur oxidation enzyme system. *Environ Microbiol* **9:** 2957-2977.

Oremland RS, Blum JS, Culbertson CW, Visscher PT, Miller LG, Dowdle P *et al.* (1994). Isolation, Growth, and Metabolism of an Obligately Anaerobic, Selenate-Respiring Bacterium, Strain SES-3. *Appl Environ Microbiol* **60:** 3011-3019.

Oremland RS, Newman DK, Kail BW, Stolz JF (2002). Bacterial respiration of arsenate and its significance in the environment. In: Frankenberger WT (ed). *Environmental Chemistry of Arsenic*. Marcel Dekker, Inc: New York. pp 273-295.

Oremland RS, Stolz JF, Hollibaugh JT (2004). The microbial arsenic cycle in Mono Lake, California. *FEMS Microbiol Ecol* **48:** 15-27.

Oremland RS, Saltikov CW, Wolfe-Simon F, Stolz JF (2009). Arsenic in the evolution of Earth and extraterrestrial ecosystems. *Geomicrobiol J* **26:** 522-536.

Planer-Friedrich B, London J, McCleskey RB, Nordstrom DK, Wallschlager D (2007). Thioarsenates in geothermal waters of Yellowstone National Park: Determination, preservation, and geochemical importance. *Environmental Science & Technology* **41:** 5245-5251.

Planer-Friedrich B, Franke D, Merkel B, Wallschläger D (2008). Acute toxicity of thioarsenates to *Vibrio fischeri*. *Environ Toxicol Chem* **27:** 2027-2035.

Planer-Friedrich B, Fisher J, Hollibaugh J, Suess E, Wallschlager D (2009). Oxidative transformation of trithioarsenate along alkaline geothermal drainages—abiotic versus microbially mediated processes. *Geomicrobiol J* **26:** 339-350.

Planer-Friedrich B, Suess E, Scheinost AC, Wallschlager D (2010). Arsenic speciation in sulfidic waters: reconciling contradictory spectroscopic and chromatographic evidence. *Anal Chem* **82:** 10228-10235.

Porter KG, Feig YS (1980). Use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25:** 943-948.

Pruesse E, Peplies J, Glockner FO (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28:** 1823-1829.

Sandor E, Csordas L (1961). The crystal structure of anhydrous sodium thiosulphate. *Acta Crystallogr* **14:** 237-243.

Schneider L (2012). The role of free and As-bound sulfide during arsenite oxidation by *Alkalilimnicola ehrlichii*, strain MLHE-1. B.Sci. thesis, University of Bayreuth, Bayreuth, Germany.

Sharma VK, Sohn M (2009). Aquatic arsenic: toxicity, speciation, transformations, and remediation. *Environ Int* **35:** 743-759.

Sorokin DY, Bosch PLF, Abbas B, Janssen AJH, Muyzer G (2008). Microbiological analysis of the population of extremely haloalkaliphilic sulfur-oxidizing bacteria dominating in lab-scale sulfide-removing bioreactors. *Appl Microbiol Biotechnol* **80:** 965-975.

Sorokin DY, Banciu H, Robertson LA, Kuenen JG, Muntyan MS, Muyzer G (2013). Halophilic and haloalkaliphilic sulfur-oxidizing bacteria. In: Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F (eds). *The Prokaryotes*. Springer Berlin Heidelberg. pp 529-554.

Stauder S, Raue B, Sacher F (2005). Thioarsenates in sulfidic waters. *Environ Sci Technol* **39:** 5933-5939.

Stucker VK, Williams KH, Robbins MJ, Ranville JF (2013). Arsenic geochemistry in a biostimulated aquifer: An aqueous speciation study. *Environ Toxicol Chem* **32:** 1216-1223.

Suess E, Scheinost AC, Bostick BC, Merkel BJ, Wallschlaeger D, Planer-Friedrich B (2009). Discrimination of thioarsenites and thioarsenates by x-ray absorption spectroscopy. *Anal Chem* **81:** 8318-8326.

Turner S, Pryer KM, Miao VPW, Palmer JD (1999). Investigating Deep Phylogenetic Relationships among Cyanobacteria and Plastids by Small Subunit rRNA Sequence Analysis. *Journal of Eukaryotic Microbiology* **46:** 327-338.

van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B (2013). Arsenics as bioenergetic substrates. *Biochim Biophys Acta* **1827:** 176-188.

Wallschlager D, Stadey CJ (2007). Determination of (oxy)thioarsenates in sulfidic waters. *Anal Chem* **79:** 3873-3880.

Weissgerber T, Zigann R, Bruce D, Chang Y-J, Detter JC, Han C *et al.* (2011). Complete genome sequence of *Allochromatium vinosum* DSM 180T. *Stand Genomic Sci* **5:** 311-330.

Widdel F, Kohring G-W, Mayer F (1983). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. III. Characterization of the filamentous gliding *Desulfonema limicola* gen. nov. sp. nov., and *Desulfonema magnum* sp. nov. *Arch Microbiol* **134:** 286-294.

Wright ES, Yilmaz LS, Noguera DR (2012). DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl Environ Microbiol* **78:** 717-725.

Zargar K, Conrad A, Bernick DL, Lowe TM, Stolc V, Hoeft S *et al.* (2012). ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. *Environ Microbiol* **14:** 1635-1645.

Table 4.1. Comparison of the structures of monothioarsenate and thiosulfate.

	Atomic Distances (Å)			Bond Angles (°)		Atomic radius (pm)
Molecule	As–O or S–O (bond)	As–S or S–S (bond)	O–O or O–S	O–As or S–O	O–As or S–S	Center atom
Monothioarsenate	1.7	2.14	2.72	109	110	119
Thiosulfate	1.47	2.01	2.45	110.5	108.5	105

Distances and bond angles are from (Kempa *et al.* 1990) and (Suess *et al.* 2009) (monothioarsenate) and (Sandor and Csordas 1961) (thiosulfate). Atomic radii are from (Cordero *et al.* 2008).

Table 4.2. Oxidation and growth of Purple Sulfur Bacteria on various electron donors.

Organism/Enrichment	Thioarsenates Mixture	Monothioarsenate	Thiosulfate	Arsenite
Mono Lake EC "ML Ecto"[a]	+	+	+	+
Mono Lake EC12	+	+	+	+
Big Soda Lake EC13	+	+	+	+
***Alkalilimnicola ehrlichii* MLHE-1[b]**	-	-	+[c]	+
***Thioalkalivibrio jannaschii* DSM 14478**	+ / ND	+ / -	+	ND
***Halorhodospira abdelmalekii* DSM 2110**	-[d] / +	-	-	ND
***Halorhodospira halophila* DSM 244**	+	+	+	+[e]
***Ectothiorhodospira shaposhnikovii* DSM 2111**	+	+	+	ND
***Ectothiorhodospira* sp. ML Ecto**	+	+	+	+
***Ectothiorhodospira* sp. 'Bogoria Red'**	+	+	+	+[f]
***Ectothiorhodospira* sp. PHS-1**	+	+	ND	+
Allochromatium vinosum	-[d] / +	-	+	ND

+ indicates positive result, - indicates negative result.
Two symbols indicate different results for oxidation and growth (oxidation / growth).
ND = not determined.
a. Source of enrichment: (Budinoff and Hollibaugh 2008)
b. Data from (Hoeft *et al.* 2007) and (Schneider 2012).
c. Growth on thiosulfate was reported by (Hoeft *et al.* 2007) but we have been unable to grow this organism on thiosulfate in our lab.
d. All thioarsenate mixture was converted to monothioarsenate but not completely oxidized to arsenate.
e. Reported in (van Lis *et al.* 2013)
f. Previously reported as unable to grow on arsenite (Budinoff and Hollibaugh 2008).

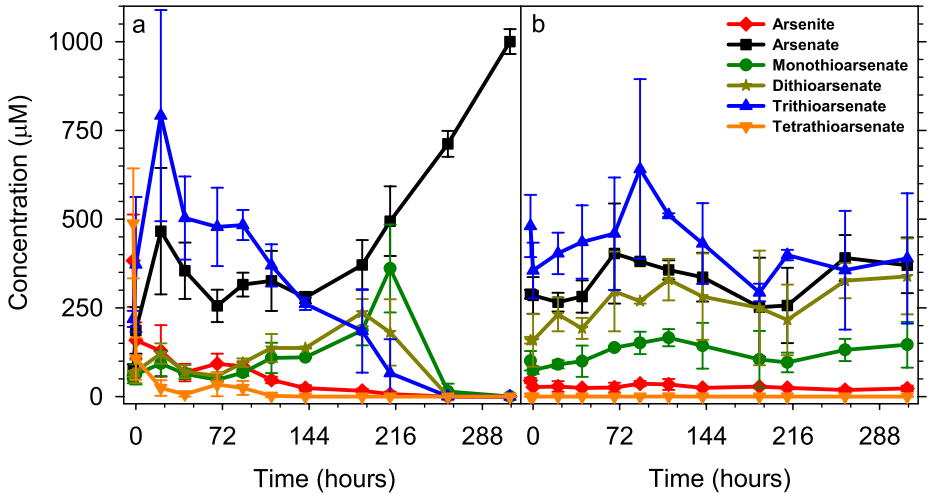


Figure 4.1. Thioarsenate speciation in (a) enrichment culture EC8 (triplicates); and (b) filtered abiotic controls (duplicates). Treatments were amended with a 4:1 S:As mixture of thioarsenates and As speciation was determined by IC-ICP-MS. Points are means of replicates with error bars indicating standard deviation for triplicates and range for duplicates.

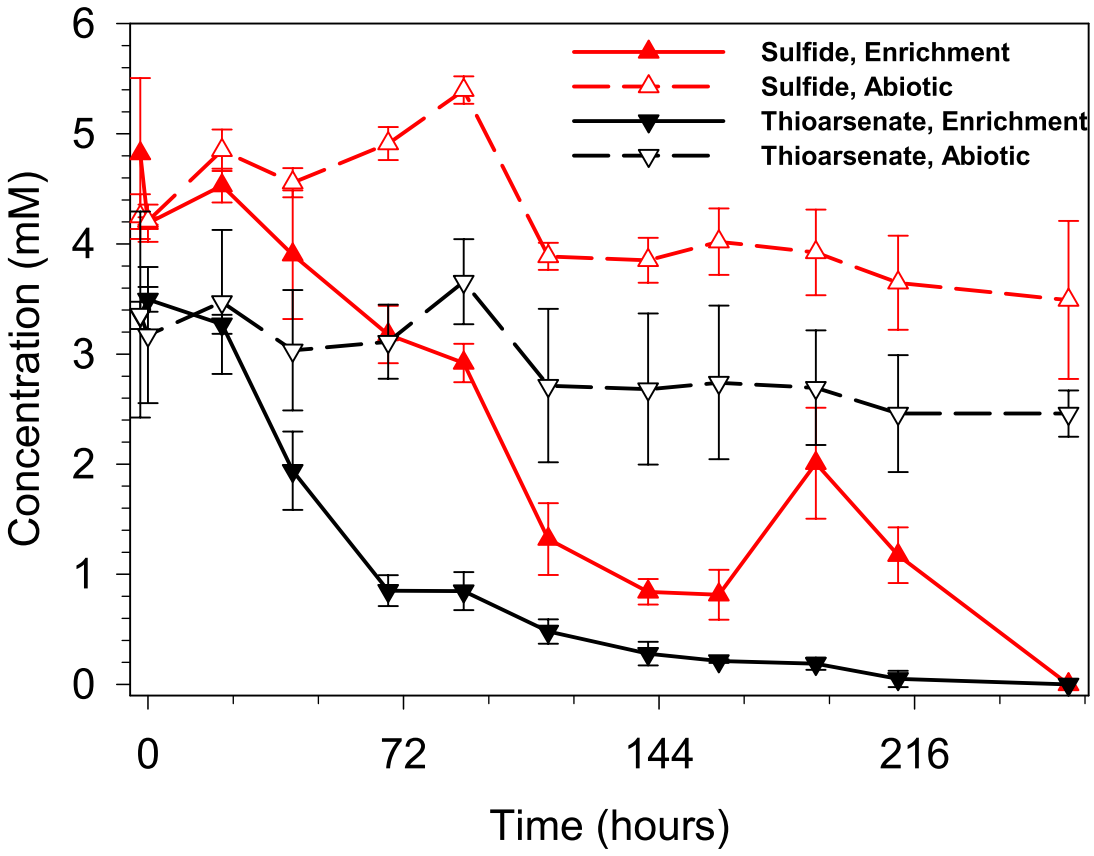


Figure 4.2. Sulfide concentrations for the enrichment culture (EC8) and abiotic controls shown in Figure 4.1. Data points are means and standard deviations, except for the thioarsenate mixture abiotic control where duplicates were analyzed and are shown as mean and range.

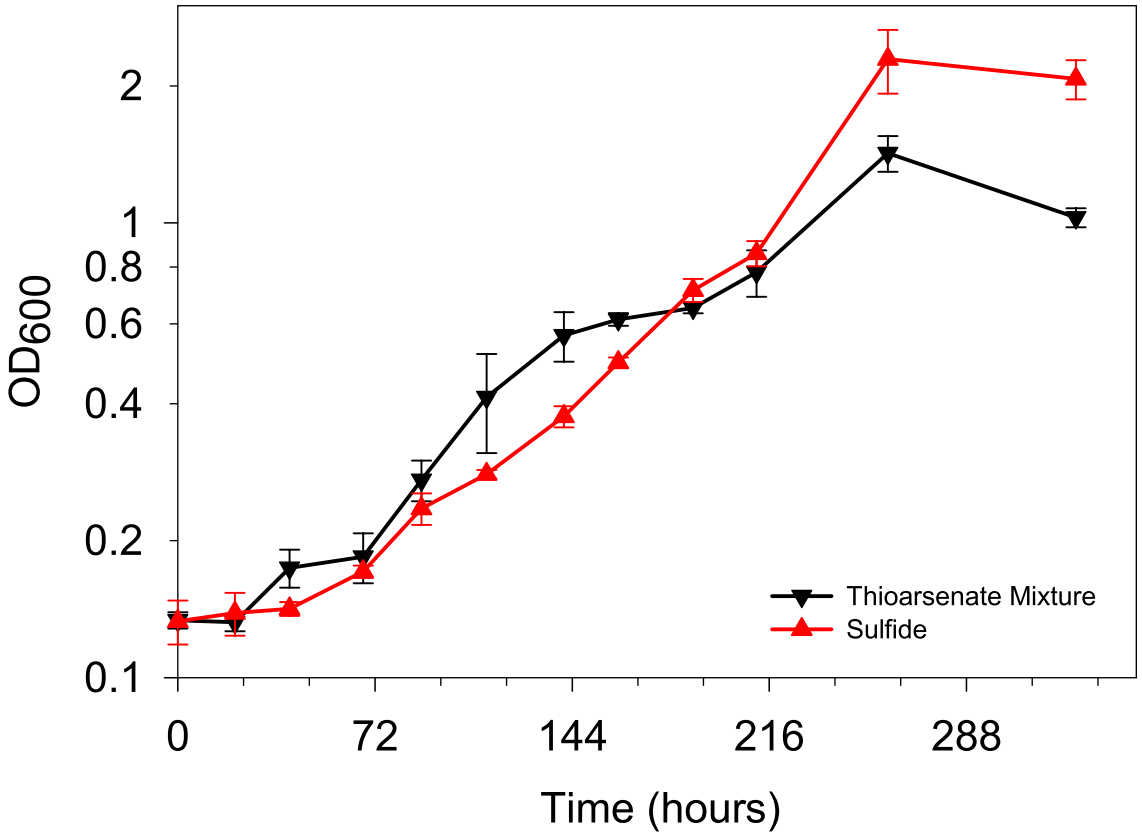


Figure 4.3. Growth of the enrichment culture (EC8) shown in Figure 4.1. Points are shown as means of triplicates with error bars indicating standard deviation. The y-axis is log scale.

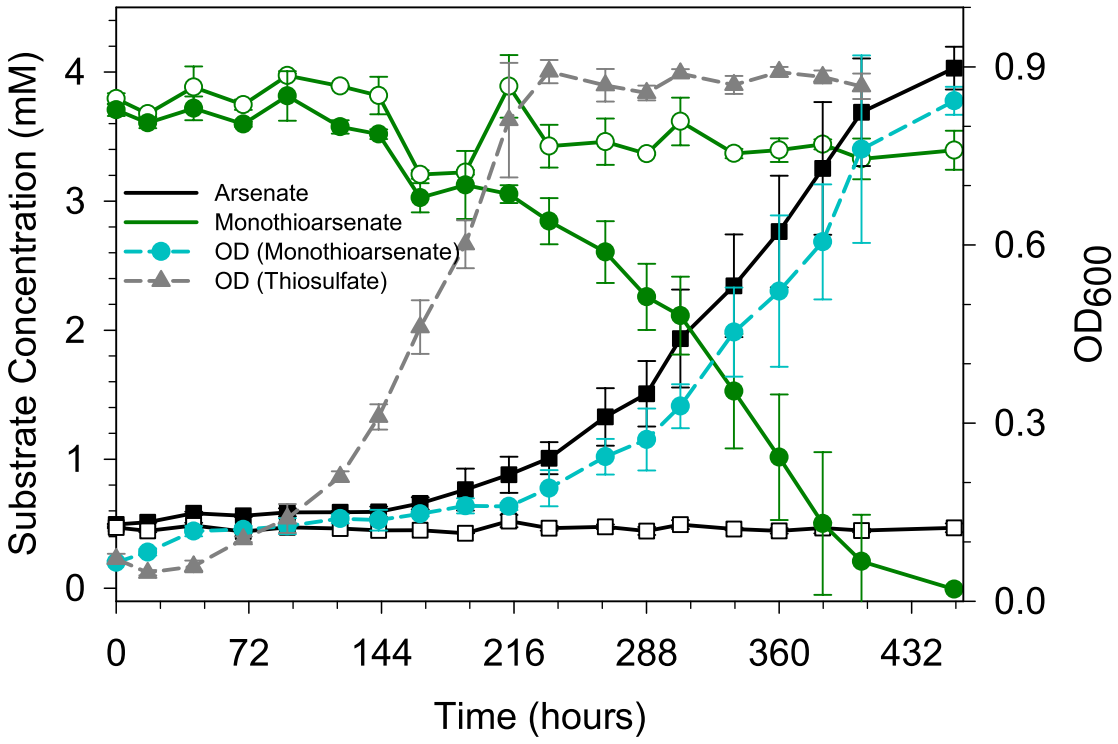


Figure 4.4. Conversion of monothioarsenate to arsenate by enrichment culture EC8 (closed symbols, solid lines) and comparison to abiotic controls (open symbols, solid line), and growth of the enrichment culture on monothioarsenate or thiosulfate (dashed lines). Experiments were run in triplicate with points shown as averages with error bars indicating standard deviation, except the monothioarsenate abiotic control in which duplicates were analyzed and error bars indicate range.

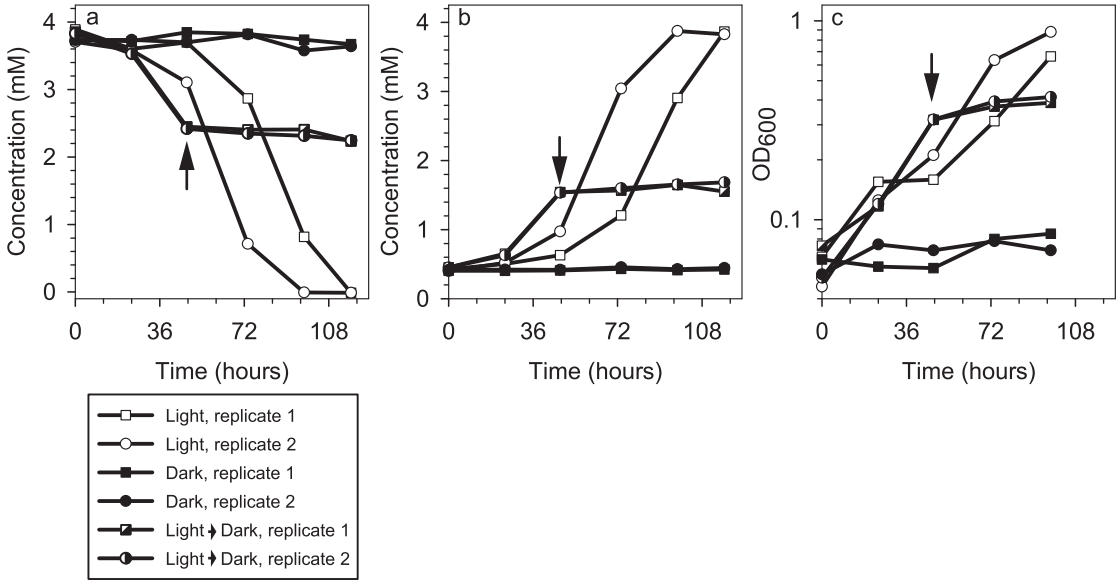


Figure 4.5. Change in monothioarsenate concentration (a), arsenate concentration (b), and growth (c) for the enrichment culture (EC8) amended with monothioarsenate (4 mM) and grown in the light, dark, or grown in light until the time point indicated by the arrow on the graph, then transferred to the dark.

CHAPTER 5

CONCLUSIONS

In the preceding chapters, molecular, chemical, and culture-based methods were used to explore the types of prokaryotes involved in arsenic and sulfur metabolism in the hypersaline, alkaline, Mono Lake, California, USA.

In Chapter 2, an overview of the microbial community of Mono Lake, CA was presented. The vertical distribution of microorganisms was determined over five depths, capturing the major physicochemical zones in the lake during the first year of meromixis. Besides the picoeukaryotic algae *Picocystis*, the lake was dominated by bacteria from the phyla Proteobacteria, Firmicutes, and Bacteroidetes. Most (80%) of the abundant (>1% relative abundance) OTUs observed in the 16S rRNA amplicon survey had been identified in a previous study in Mono Lake (Humayoun et al. 2003) and some (31%) were identical to sequenced 16S rRNA amplicons found in that study. Metatranscriptomic taxonomic bins at the genus level were dominated by the Gammaproteobacteria (*Thioalkalivibrio*, 4-13% and *Thioalkalimicrobium,* 0-14%), and Firmicutes (*Dethiobacter*, 0-5%) and *Clostridium*, 1-4%). The comparison of these results with previous studies shows that the lake community is generally stable over the long term, even with alternative regimes of stratification and mixing. These findings also reveal that many of the most abundant organisms at all depths in the lake are involved in sulfur cycling and many are probably involved in arsenic cycling as well. Although measurements of arsenic and sulfur cycling rates have been made previously, identification of the organisms involved has, up until now, been limited to surveys based on small libraries of cloned 16S rRNA amplicons and isolation of

strains. The full extent of the diversity of the biogeochemical cycles and organisms involved in microbial metabolism in Mono Lake was not determined, although the metatranscriptomes generated in this investigation can be used for future studies of important processes in biogeochemistry in the lake beyond arsenic and sulfur cycling.

In Chapter 3, the arsenic and sulfur cycles in Mono Lake were examined more closely due to the prevalence of arsenic and sulfur cycling taxa revealed by the analysis presented in Chapter 2. The metatranscriptomics dataset from Chapter 2 was used to explore key transcripts of arsenite oxidation, arsenate reduction, sulfur compound oxidation, and sulfate reduction. A custom database approach was used along with phylogenetic analysis of functional marker protein sequences and the relative abundance of transcript hits to these proteins. These were key aspects of this analysis, as database entries would not have identified many of the organisms that were found to contain and transcribe gene transcripts related to arsenic cycling. The analysis of sulfur cycling transcripts increased our knowledge of the organisms in Mono Lake that are involved in sulfur cycling. Previous studies failed to identify the taxa involved in sulfur cycling because of primer bias. Although many of the organisms identified in this analysis are likely just the closest relatives present in the database, the phylogenetic placement of the assembled transcripts gives us more confidence in the identification of arsenic and sulfur cycling taxa. Further work using metagenomics assembly and binning may reveal the genomic context of these transcripts.

In Chapter 4, an enrichment culture was used to investigate the transformation of a mixture of thioarsenic compounds. The culture was able to use monothioarsenate as the sole electron donor for anoxygenic photosynthesis. The culture was dominated by purple sulfur bacteria (*Ectothiorhodospira* sp.). Other members of the family *Ectothiorhodospiraceae* were

able to grow on monothioarsenate and thioarsenic compounds, thus this ability appears to be widespread. Thioarsenates are the dominant arsenic species in arsenic-rich, alkaline and sulfidic waters, but bacterial interactions with these compounds have only recently been examined. Although recent studies have shown the ability of bacteria to directly transform monothioarsenate (both oxidatively and reductively), this is the first study to show that it is used as an electron donor for anoxygenic photosynthesis. The molecular mechanism of this transformation was hypothesized to be mediated by sulfur-cycling related enzymes (thiosulfate hydrolase). Although this was not able to be demonstrated due to lack of genetic system for these organisms, future work should test this hypothesis once a genetic system is developed. Overall, this work builds upon previous studies of thioarsenic transformations and adds a new perspective to the arsenic and sulfur cycles.

In summary, this work has increased the understanding of prokaryotes involved in transformations of sulfur, arsenic and thioarsenic compounds. These results can be applied to global models of arsenic cycling, especially in alkaline, sulfidic waters. In addition, this work has increased the understanding of the prokaryotic diversity within a unique alkaline, hypersaline environment and revealed organisms not previously identified as being involved in arsenic and sulfur cycling.

APPENDIX A

SUPPORTING INFORMATION FOR CHAPTER 2

Table 2.S1. CTD and chemical measurements from Mono Lake Station 6, July 13, 2012.

Depth, m	Sulfide, µM	As (III), µM	As (V), µM	Mono-TAs, µM	Di-TAs, µM	Tri-TAs, µM	Tetra-TAs, µM	Sum of As Species, µM	Total As, µM	Cond., mS/cm	Temp., °C	DO, mg/L	PAR	FL, mg/m^3	Beam, 1/m	Bacteria 16S, 10^{10} copies/L
2.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	73.10	20.36	3.32	85.88	4.28	1.19	ND
3.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	73.06	20.33	3.33	128.43	4.58	1.24	ND
3.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	73.01	20.31	3.34	210.22	4.83	1.20	ND
4.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	72.96	20.27	3.35	213.99	5.15	1.20	ND
4.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	72.89	20.23	3.36	185.23	5.62	1.20	ND
5.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	72.85	20.19	3.35	158.72	6.12	1.21	ND
5.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	72.79	20.16	3.30	136.57	6.51	1.27	ND
6.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	72.15	20.03	3.35	121.72	6.74	1.21	ND
6.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	71.73	19.66	3.50	107.80	7.24	1.22	ND
7.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	71.32	19.38	3.58	92.80	9.28	1.26	ND
7.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	70.85	19.14	3.67	79.51	10.55	1.31	ND
8.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	70.13	18.80	3.79	67.91	11.41	1.39	ND
8.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	69.14	18.27	3.94	53.99	12.94	1.46	ND
9.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	68.14	17.58	4.10	52.58	16.91	1.52	ND
9.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	66.52	16.70	4.21	46.21	20.94	1.73	ND
10	0.28	0.02	211.29	0.07	0.02	0.02	0.03	211.44	252.19	ND	ND	ND	ND	ND	ND	4.72
10.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	65.51	15.96	4.01	39.97	30.82	2.05	ND
10.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	64.78	15.47	3.59	32.66	42.23	2.53	ND
11.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	64.33	15.18	3.30	25.87	53.72	3.12	ND
11.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	64.00	14.93	3.06	20.45	60.87	3.45	ND
12	0.28	0.02	145.22	0.03	0.02	0.03	0.03	145.34	246.16	ND	ND	ND	ND	ND	ND	ND
12.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	63.71	14.74	2.81	15.87	62.86	3.72	ND
12.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	63.35	14.51	2.53	12.34	63.74	3.82	ND
13.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	63.13	14.33	2.33	9.63	63.88	3.90	ND
13.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	62.89	14.21	2.00	7.52	64.03	3.98	ND
14.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	62.00	13.89	1.42	5.69	64.29	4.21	ND

14.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	60.81	13.31	0.96	4.36	64.54	4.17	ND
15	1.68	34.59	141.74	0.03	0.04	0.02	0.03	176.44	243.40	ND	ND	ND	ND	ND	ND	5.54
15.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	59.53	12.60	0.78	3.30	64.81	4.96	ND
15.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	58.91	11.84	0.69	2.39	64.10	5.31	ND
16	0.73	87.89	90.09	0.07	0.02	0.03	0.02	178.11	242.82	ND	ND	ND	ND	ND	ND	ND
16.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	58.36	11.50	0.66	1.72	62.45	5.64	ND
16.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	57.77	11.08	0.66	1.20	61.63	5.72	ND
17	1.18	150.56	80.19	0.14	0.03	0.03	0.02	230.97	237.09	ND	ND	ND	ND	ND	ND	ND
17.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	57.15	10.68	0.68	0.84	61.25	5.76	ND
17.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	56.40	10.23	0.66	0.59	60.92	5.78	ND
18	1.05	195.94	42.41	1.37	0.02	0.02	0.01	239.78	238.17	ND	ND	ND	ND	ND	ND	9.90
18.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	55.47	9.63	0.66	0.43	60.07	5.74	ND
18.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	54.65	9.08	0.68	0.32	60.10	5.80	ND
19.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	53.87	8.46	0.69	0.23	61.08	5.92	ND
19.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	53.46	8.04	0.69	0.17	61.76	6.19	ND
20	6.14	181.65	8.62	2.03	0.42	0.08	0.02	192.82	236.42	ND	ND	ND	ND	ND	ND	ND
20.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	53.31	7.89	0.70	0.12	62.27	6.38	ND
20.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	53.06	7.76	0.69	0.08	62.64	6.46	ND
21.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	52.83	7.57	0.70	0.06	63.04	6.53	ND
21.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	52.66	7.47	0.69	0.04	63.46	6.61	ND
22.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	52.44	7.35	0.69	0.03	63.70	6.65	ND
22.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	52.22	7.19	0.69	0.02	64.09	6.70	ND
23.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	52.11	7.09	0.69	0.02	64.64	6.77	ND
23.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.99	7.01	0.70	0.02	65.15	6.81	ND
24.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.84	6.93	0.70	0.01	65.07	6.77	ND
24.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.64	6.83	0.69	0.01	65.23	6.74	ND
25	7.62	143.01	7.99	1.52	0.90	0.35	0.02	153.79	237.03	ND	ND	ND	ND	ND	ND	14.40
25.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.50	6.70	0.69	0.01	65.85	6.79	ND
25.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.41	6.64	0.69	0.01	66.67	6.89	ND
26.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.28	6.56	0.69	0.01	67.29	6.94	ND
26.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.16	6.47	0.69	0.01	67.89	6.99	ND

27.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.08	6.41	0.69	0.01	68.53	7.03	ND
27.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	51.00	6.35	0.69	0.01	69.32	7.06	ND
28.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.92	6.31	0.67	0.01	69.74	7.10	ND
28.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.88	6.26	0.68	0.01	70.11	7.15	ND
29.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.83	6.23	0.67	0.00	70.43	7.20	ND
29.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.76	6.18	0.69	0.00	70.67	7.22	ND
30.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.69	6.14	0.69	0.00	70.91	7.22	ND
30.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.59	6.09	0.68	0.00	71.13	7.24	ND
31										ND	ND	ND	ND	ND	ND	14.87
31.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.56	6.04	0.70	0.00	71.31	7.28	ND
31.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.52	6.01	0.69	0.00	71.50	7.34	ND
32.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.44	5.96	0.69	0.00	71.37	7.34	ND
32.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.41	5.92	0.69	0.00	71.34	7.34	ND
33.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.37	5.90	0.68	0.00	71.56	7.33	ND
33.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.32	5.86	0.67	0.00	71.56	7.33	ND
34.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.30	5.84	0.67	0.00	71.57	7.32	ND
34.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.27	5.83	0.69	0.00	71.58	7.34	ND
35	31.72	140.24	12.62	4.83	4.19	2.12	0.02	164.02	243.16	ND	ND	ND	ND	ND	ND	ND
35.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.26	5.80	0.66	0.00	71.57	7.32	ND
35.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.25	5.80	0.65	0.00	71.62	7.33	ND
36.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.22	5.78	0.67	0.00	71.60	7.32	ND
36.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.20	5.77	0.67	0.00	71.61	7.33	ND
37.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.16	5.74	0.67	0.00	71.72	7.33	ND
37.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.12	5.72	0.66	0.00	71.61	7.34	ND
38.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.05	5.68	0.64	0.00	71.65	7.37	ND
38.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.04	5.64	0.65	0.00	71.75	7.48	ND
39.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.04	5.63	0.67	0.00	71.80	7.58	ND
39.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	50.04	5.63	0.64	0.00	71.89	7.60	ND
40.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	49.90	5.63	0.66	0.00	71.89	7.61	ND
40.8	ND	ND	ND	ND	ND	ND	ND	ND	ND	43.97	5.62	0.69	0.00	71.84	7.61	ND
41.3	ND	ND	ND	ND	ND	ND	ND	ND	ND	33.38	5.61	0.59	0.00	70.65	7.61	ND

Abbreviations: TAs = Thioarsenic, As = Arsenic, As(III) = Arsenite, As(V) = Arsenate, Cond. = Conductivity, Temp. = Temperature, DO = Dissolved Oxygen, PAR = Photosynthetically Active Radiation, FL = Fluorescence, Beam = Beam Attenuation, Bacteria 16S = 16S rRNA qPCR results

Table 2.S2. 16S rRNA amplicon sequencing OTU table with counts and relative abundances. Purple OTUs are present at all depths and >1% relative abundance, yellow OTUs are >1% relative abundance at one or more depths and blue OTUs are present at all depths but less than 1% relative abundance. RA=Relative Abundance.

OTU ID	QIIME-Assigned OTU Name	Depth									
		10 m		15 m		18 m		25 m		31 m	
		Read Count	RA (%)	Read Count	RA (%)	Read Count	RA (%)	Read Count	RA (%)	Read Count	RA (%)
ML2012 OTU 8	AF449782	3080	19.80	831	13.45	593	9.47	264	6.93	93	5.41
ML2012 OTU 150	New.ReferenceOTU0	3	0.02	334	5.40	376	6.01	259	6.80	117	6.81
ML2012 OTU 81	HM128337	661	4.25	413	6.68	298	4.76	157	4.12	88	5.12
ML2012 OTU 6	AF449778	169	1.09	287	4.64	340	5.43	201	5.27	123	7.16
ML2012 OTU 151	New.ReferenceOTU1	63	0.41	138	2.23	248	3.96	247	6.48	130	7.57
ML2012 OTU 16	AF454301	2022	13.00	188	3.04	95	1.52	44	1.15	12	0.70
ML2012 OTU 3	AF448193	1351	8.69	224	3.62	165	2.64	86	2.26	32	1.86
ML2012 OTU 36	DQ206410	3	0.02	146	2.36	388	6.20	173	4.54	71	4.13
ML2012 OTU 10	AF452605	27	0.17	156	2.52	152	2.43	161	4.22	94	5.47
ML2012 OTU 11	AF452606	872	5.61	185	2.99	124	1.98	49	1.29	28	1.63
ML2012 OTU 83	HM129833	626	4.02	195	3.16	164	2.62	63	1.65	32	1.86
ML2012 OTU 7	AF449781	938	6.03	170	2.75	113	1.80	46	1.21	10	0.58
ML2012 OTU 21	AF507855	299	1.92	76	1.23	94	1.50	58	1.52	43	2.50
ML2012 OTU 20	AF458284	1	0.01	51	0.83	101	1.61	85	2.23	59	3.43
ML2012 OTU 26	AF507867	279	1.79	118	1.91	51	0.81	63	1.65	15	0.87
ML2012 OTU 13	AF453543	374	2.40	95	1.54	66	1.05	19	0.50	1	0.06
ML2012 OTU 173	New.ReferenceOTU3	4	0.03	48	0.78	71	1.13	53	1.39	21	1.22
ML2012 OTU 9	AF452599	9	0.06	49	0.79	72	1.15	48	1.26	20	1.16
ML2012 OTU 158	New.ReferenceOTU16	2	0.01	13	0.21	24	0.38	44	1.15	8	0.47
ML2012 OTU 162	New.ReferenceOTU2	177	1.14	29	0.47	9	0.14	1	0.03	4	0.23
ML2012 OTU 12	AF452608	2251	14.47	485	7.85	277	4.42	111	2.91	0	0.00
ML2012 OTU 41	DQ900622	23	0.15	577	9.34	417	6.66	133	3.49	0	0.00

ML2012 OTU 28	AF507875	5	0.03	148	2.39	307	4.90	227	5.96	0	0.00
ML2012 OTU 22	AF507857	0	0.00	63	1.02	129	2.06	100	2.62	63	3.67
ML2012 OTU 95	JQ231232	396	2.55	84	1.36	56	0.89	21	0.55	0	0.00
ML2012 OTU 194	New.ReferenceOTU5	0	0.00	0	0.00	0	0.00	0	0.00	86	5.01
ML2012 OTU 90	JF811035	188	1.21	58	0.94	58	0.93	30	0.79	0	0.00
ML2012 OTU 214	New.ReferenceOTU7	0	0.00	15	0.24	38	0.61	34	0.89	24	1.40
ML2012 OTU 155	New.ReferenceOTU13	0	0.00	16	0.26	36	0.57	34	0.89	18	1.05
ML2012 OTU 64	GQ135733	0	0.00	20	0.32	71	1.13	50	1.31	0	0.00
ML2012 OTU 166	New.ReferenceOTU23	0	0.00	6	0.10	12	0.19	16	0.42	33	1.92
ML2012 OTU 176	New.ReferenceOTU32	0	0.00	0	0.00	0	0.00	0	0.00	38	2.21
ML2012 OTU 50	EU709859	212	1.36	18	0.29	10	0.16	0	0.00	0	0.00
ML2012 OTU 169	New.ReferenceOTU26	0	0.00	0	0.00	0	0.00	0	0.00	29	1.69
ML2012 OTU 80	HM128188	0	0.00	0	0.00	0	0.00	0	0.00	18	1.05
ML2012 OTU 24	AF507862	1	0.01	18	0.29	31	0.50	15	0.39	17	0.99
ML2012 OTU 177	New.ReferenceOTU33	14	0.09	6	0.10	1	0.02	2	0.05	2	0.12
ML2012 OTU 187	New.ReferenceOTU43	1	0.01	12	0.19	19	0.30	6	0.16	4	0.23
ML2012 OTU 163	New.ReferenceOTU20	1	0.01	2	0.03	9	0.14	5	0.13	8	0.47
ML2012 OTU 225	New.ReferenceOTU8	55	0.35	14	0.23	13	0.21	8	0.21	2	0.12
ML2012 OTU 152	New.ReferenceOTU10	33	0.21	15	0.24	29	0.46	26	0.68	6	0.35
ML2012 OTU 15	AF454300	50	0.32	11	0.18	13	0.21	1	0.03	1	0.06
ML2012 OTU 82	HM128341	2	0.01	10	0.16	20	0.32	9	0.24	7	0.41
ML2012 OTU 153	New.ReferenceOTU11	2	0.01	24	0.39	47	0.75	25	0.66	14	0.81
ML2012 OTU 204	New.ReferenceOTU59	1	0.01	2	0.03	10	0.16	16	0.42	5	0.29
ML2012 OTU 77	HM127957	41	0.26	10	0.16	6	0.10	6	0.16	1	0.06
ML2012 OTU 39	DQ432395	13	0.08	15	0.24	8	0.13	4	0.10	4	0.23
ML2012 OTU 195	New.ReferenceOTU50	1	0.01	5	0.08	10	0.16	4	0.10	1	0.06
ML2012 OTU 127	New.CleanUp.ReferenceOTU4	1	0.01	1	0.02	6	0.10	3	0.08	3	0.17
ML2012 OTU 44	EF632658	30	0.19	19	0.31	4	0.06	3	0.08	1	0.06
ML2012 OTU 62	FJ999568	7	0.05	6	0.10	1	0.02	1	0.03	1	0.06
ML2012 OTU 165	New.ReferenceOTU22	19	0.12	19	0.31	31	0.50	22	0.58	4	0.23

ML2012 OTU 157	New.ReferenceOTU15	17	0.11	11	0.18	13	0.21	6	0.16	1	0.06
ML2012 OTU 205	New.ReferenceOTU6	50	0.32	21	0.34	8	0.13	8	0.21	3	0.17
ML2012 OTU 58	FJ153016	64	0.41	22	0.36	8	0.13	5	0.13	3	0.17
ML2012 OTU 183	New.ReferenceOTU4	100	0.64	41	0.66	37	0.59	17	0.45	6	0.35
ML2012 OTU 117	New.CleanUp.ReferenceOTU21	0	0.00	1	0.02	3	0.05	4	0.10	1	0.06
ML2012 OTU 84	HM130024	4	0.03	0	0.00	5	0.08	0	0.00	0	0.00
ML2012 OTU 109	New.CleanUp.ReferenceOTU13	2	0.01	0	0.00	0	0.00	2	0.05	0	0.00
ML2012 OTU 56	FJ152933	68	0.44	20	0.32	8	0.13	6	0.16	0	0.00
ML2012 OTU 40	DQ432396	26	0.17	9	0.15	9	0.14	3	0.08	0	0.00
ML2012 OTU 91	JN003088	2	0.01	0	0.00	1	0.02	1	0.03	0	0.00
ML2012 OTU 27	AF507869	0	0.00	0	0.00	4	0.06	3	0.08	0	0.00
ML2012 OTU 208	New.ReferenceOTU63	0	0.00	6	0.10	5	0.08	11	0.29	8	0.47
ML2012 OTU 211	New.ReferenceOTU67	0	0.00	1	0.02	1	0.02	2	0.05	0	0.00
ML2012 OTU 198	New.ReferenceOTU53	2	0.01	4	0.06	8	0.13	10	0.26	0	0.00
ML2012 OTU 219	New.ReferenceOTU74	1	0.01	4	0.06	5	0.08	6	0.16	0	0.00
ML2012 OTU 5	AF449769	0	0.00	5	0.08	18	0.29	25	0.66	0	0.00
ML2012 OTU 47	EU283512	0	0.00	2	0.03	3	0.05	5	0.13	0	0.00
ML2012 OTU 4	AF449768	0	0.00	10	0.16	15	0.24	19	0.50	12	0.70
ML2012 OTU 23	AF507860	0	0.00	0	0.00	0	0.00	0	0.00	17	0.99
ML2012 OTU 138	New.CleanUp.ReferenceOTU65	0	0.00	0	0.00	1	0.02	1	0.03	1	0.06
ML2012 OTU 218	New.ReferenceOTU73	0	0.00	1	0.02	3	0.05	0	0.00	2	0.12
ML2012 OTU 59	FJ485151	2	0.01	3	0.05	2	0.03	0	0.00	0	0.00
ML2012 OTU 223	New.ReferenceOTU78	1	0.01	1	0.02	7	0.11	0	0.00	0	0.00
ML2012 OTU 228	New.ReferenceOTU82	0	0.00	0	0.00	5	0.08	1	0.03	0	0.00
ML2012 OTU 201	New.ReferenceOTU56	0	0.00	1	0.02	3	0.05	6	0.16	0	0.00
ML2012 OTU 79	HM128153	98	0.63	39	0.63	31	0.50	10	0.26	0	0.00
ML2012 OTU 33	DQ015831	13	0.08	2	0.03	8	0.13	6	0.16	0	0.00
ML2012 OTU 54	FJ152901	52	0.33	15	0.24	13	0.21	4	0.10	0	0.00
ML2012 OTU 189	New.ReferenceOTU45	25	0.16	6	0.10	5	0.08	1	0.03	0	0.00
ML2012 OTU 238	New.ReferenceOTU93	4	0.03	4	0.06	1	0.02	0	0.00	0	0.00

ML2012 OTU 240	New.ReferenceOTU96	8	0.05	1	0.02	0	0.00	0	0.00	0	0.00
ML2012 OTU 55	FJ152914	0	0.00	0	0.00	0	0.00	0	0.00	4	0.23
ML2012 OTU 52	EU911993	0	0.00	4	0.06	16	0.26	4	0.10	7	0.41
ML2012 OTU 57	FJ152979	0	0.00	3	0.05	4	0.06	3	0.08	2	0.12
ML2012 OTU 49	EU703291	0	0.00	0	0.00	0	0.00	0	0.00	4	0.23
ML2012 OTU 123	New.CleanUp.ReferenceOTU28	1	0.01	2	0.03	0	0.00	0	0.00	0	0.00
ML2012 OTU 160	New.ReferenceOTU18	52	0.33	9	0.15	4	0.06	0	0.00	16	0.93
ML2012 OTU 25	AF507866	45	0.29	17	0.28	16	0.26	9	0.24	0	0.00
ML2012 OTU 226	New.ReferenceOTU80	0	0.00	1	0.02	1	0.02	0	0.00	0	0.00
ML2012 OTU 133	New.CleanUp.ReferenceOTU5	0	0.00	2	0.03	4	0.06	4	0.10	2	0.12
ML2012 OTU 135	New.CleanUp.ReferenceOTU54	2	0.01	1	0.02	2	0.03	0	0.00	0	0.00
ML2012 OTU 164	New.ReferenceOTU21	0	0.00	5	0.08	8	0.13	10	0.26	5	0.29
ML2012 OTU 139	New.CleanUp.ReferenceOTU66	0	0.00	0	0.00	0	0.00	2	0.05	0	0.00
ML2012 OTU 184	New.ReferenceOTU40	0	0.00	8	0.13	21	0.34	7	0.18	1	0.06
ML2012 OTU 70	GU083693	0	0.00	3	0.05	3	0.05	2	0.05	1	0.06
ML2012 OTU 125	New.CleanUp.ReferenceOTU3	0	0.00	2	0.03	6	0.10	3	0.08	2	0.12
ML2012 OTU 143	New.CleanUp.ReferenceOTU74	0	0.00	1	0.02	0	0.00	1	0.03	0	0.00
ML2012 OTU 148	New.CleanUp.ReferenceOTU94	0	0.00	2	0.03	0	0.00	0	0.00	0	0.00
ML2012 OTU 85	HM468058	0	0.00	1	0.02	8	0.13	4	0.10	1	0.06
ML2012 OTU 67	GQ377778	95	0.61	3	0.05	1	0.02	5	0.13	0	0.00
ML2012 OTU 241	New.ReferenceOTU98	1	0.01	1	0.02	4	0.06	0	0.00	2	0.12
ML2012 OTU 146	New.CleanUp.ReferenceOTU88	0	0.00	0	0.00	0	0.00	2	0.05	0	0.00
ML2012 OTU 178	New.ReferenceOTU34	43	0.28	20	0.32	11	0.18	0	0.00	0	0.00
ML2012 OTU 118	New.CleanUp.ReferenceOTU22	0	0.00	0	0.00	0	0.00	3	0.08	0	0.00
ML2012 OTU 149	New.CleanUp.ReferenceOTU95	0	0.00	0	0.00	1	0.02	2	0.05	0	0.00
ML2012 OTU 98	JQ245609	0	0.00	6	0.10	15	0.24	14	0.37	0	0.00
ML2012 OTU 31	AJ271453	0	0.00	0	0.00	0	0.00	0	0.00	17	0.99
ML2012 OTU 217	New.ReferenceOTU72	0	0.00	0	0.00	4	0.06	1	0.03	0	0.00
ML2012 OTU 2	AB294306	0	0.00	0	0.00	3	0.05	4	0.10	0	0.00
ML2012 OTU 46	EU245382	0	0.00	0	0.00	2	0.03	0	0.00	0	0.00

ML2012 OTU 221	New.ReferenceOTU76	0	0.00	1	0.02	2	0.03	2	0.05	1	0.06
ML2012 OTU 60	FJ788525	0	0.00	12	0.19	30	0.48	5	0.13	0	0.00
ML2012 OTU 126	New.CleanUp.ReferenceOTU31	0	0.00	0	0.00	0	0.00	0	0.00	2	0.12
ML2012 OTU 104	New.CleanUp.ReferenceOTU10	0	0.00	1	0.02	0	0.00	4	0.10	1	0.06
ML2012 OTU 37	DQ206423	0	0.00	2	0.03	3	0.05	4	0.10	0	0.00
ML2012 OTU 175	New.ReferenceOTU31	1	0.01	7	0.11	6	0.10	10	0.26	0	0.00
ML2012 OTU 103	New.CleanUp.ReferenceOTU1	0	0.00	0	0.00	0	0.00	0	0.00	11	0.64
ML2012 OTU 51	EU843077	0	0.00	2	0.03	0	0.00	0	0.00	0	0.00
ML2012 OTU 171	New.ReferenceOTU28	0	0.00	8	0.13	14	0.22	10	0.26	0	0.00
ML2012 OTU 142	New.CleanUp.ReferenceOTU72	0	0.00	0	0.00	0	0.00	0	0.00	2	0.12
ML2012 OTU 34	DQ124682	0	0.00	0	0.00	1	0.02	0	0.00	1	0.06
ML2012 OTU 102	New.CleanUp.ReferenceOTU0	0	0.00	10	0.16	8	0.13	10	0.26	5	0.29
ML2012 OTU 68	GQ848216	3	0.02	2	0.03	4	0.06	2	0.05	0	0.00
ML2012 OTU 29	AF507877	0	0.00	8	0.13	34	0.54	20	0.52	13	0.76
ML2012 OTU 14	AF454298	0	0.00	2	0.03	11	0.18	8	0.21	9	0.52
ML2012 OTU 200	New.ReferenceOTU55	0	0.00	5	0.08	4	0.06	7	0.18	1	0.06
ML2012 OTU 167	New.ReferenceOTU24	1	0.01	10	0.16	36	0.57	24	0.63	0	0.00
ML2012 OTU 30	AF507890	0	0.00	5	0.08	10	0.16	9	0.24	0	0.00
ML2012 OTU 235	New.ReferenceOTU9	0	0.00	5	0.08	27	0.43	26	0.68	7	0.41
ML2012 OTU 203	New.ReferenceOTU58	0	0.00	4	0.06	13	0.21	4	0.10	0	0.00
ML2012 OTU 192	New.ReferenceOTU48	0	0.00	3	0.05	10	0.16	11	0.29	0	0.00
ML2012 OTU 180	New.ReferenceOTU36	0	0.00	3	0.05	13	0.21	16	0.42	0	0.00
ML2012 OTU 108	New.CleanUp.ReferenceOTU12	0	0.00	0	0.00	0	0.00	0	0.00	3	0.17
ML2012 OTU 42	DQ988278	0	0.00	2	0.03	1	0.02	4	0.10	1	0.06
ML2012 OTU 190	New.ReferenceOTU46	0	0.00	3	0.05	3	0.05	4	0.10	0	0.00
ML2012 OTU 106	New.CleanUp.ReferenceOTU110	0	0.00	0	0.00	1	0.02	1	0.03	0	0.00
ML2012 OTU 179	New.ReferenceOTU35	0	0.00	6	0.10	22	0.35	6	0.16	2	0.12
ML2012 OTU 161	New.ReferenceOTU19	0	0.00	11	0.18	28	0.45	18	0.47	6	0.35
ML2012 OTU 237	New.ReferenceOTU92	0	0.00	0	0.00	0	0.00	2	0.05	0	0.00
ML2012 OTU 122	New.CleanUp.ReferenceOTU27	0	0.00	1	0.02	1	0.02	4	0.10	0	0.00

ML2012 OTU 105	New.CleanUp.ReferenceOTU11	0	0.00	0	0.00	0	0.00	4	0.10	0	0.00
ML2012 OTU 75	HE611101	0	0.00	0	0.00	0	0.00	1	0.03	1	0.06
ML2012 OTU 197	New.ReferenceOTU52	13	0.08	2	0.03	0	0.00	2	0.05	0	0.00
ML2012 OTU 141	New.CleanUp.ReferenceOTU70	0	0.00	0	0.00	3	0.05	0	0.00	0	0.00
ML2012 OTU 196	New.ReferenceOTU51	0	0.00	2	0.03	8	0.13	4	0.10	1	0.06
ML2012 OTU 19	AF454310	0	0.00	3	0.05	7	0.11	3	0.08	1	0.06
ML2012 OTU 18	AF454309	0	0.00	0	0.00	0	0.00	0	0.00	2	0.12
ML2012 OTU 134	New.CleanUp.ReferenceOTU53	0	0.00	0	0.00	1	0.02	2	0.05	1	0.06
ML2012 OTU 212	New.ReferenceOTU68	0	0.00	3	0.05	4	0.06	2	0.05	1	0.06
ML2012 OTU 239	New.ReferenceOTU94	0	0.00	2	0.03	2	0.03	12	0.31	2	0.12
ML2012 OTU 193	New.ReferenceOTU49	0	0.00	1	0.02	6	0.10	5	0.13	3	0.17
ML2012 OTU 168	New.ReferenceOTU25	0	0.00	7	0.11	15	0.24	12	0.31	3	0.17
ML2012 OTU 65	GQ138312	0	0.00	0	0.00	1	0.02	3	0.08	1	0.06
ML2012 OTU 35	DQ206406	0	0.00	0	0.00	0	0.00	0	0.00	12	0.70
ML2012 OTU 206	New.ReferenceOTU60	0	0.00	3	0.05	21	0.34	4	0.10	3	0.17
ML2012 OTU 78	HM128005	4	0.03	0	0.00	0	0.00	0	0.00	0	0.00
ML2012 OTU 86	HQ166789	86	0.55	28	0.45	7	0.11	2	0.05	0	0.00
ML2012 OTU 92	JN637818	1	0.01	6	0.10	1	0.02	2	0.05	0	0.00
ML2012 OTU 74	HE574924	0	0.00	0	0.00	0	0.00	0	0.00	3	0.17
ML2012 OTU 114	New.CleanUp.ReferenceOTU18	0	0.00	1	0.02	0	0.00	2	0.05	0	0.00
ML2012 OTU 124	New.CleanUp.ReferenceOTU29	0	0.00	0	0.00	1	0.02	1	0.03	0	0.00
ML2012 OTU 110	New.CleanUp.ReferenceOTU14	0	0.00	0	0.00	1	0.02	3	0.08	0	0.00
ML2012 OTU 215	New.ReferenceOTU70	0	0.00	0	0.00	3	0.05	5	0.13	1	0.06
ML2012 OTU 144	New.CleanUp.ReferenceOTU8	0	0.00	3	0.05	1	0.02	2	0.05	0	0.00
ML2012 OTU 61	FJ973586	11	0.07	2	0.03	1	0.02	0	0.00	0	0.00
ML2012 OTU 147	New.CleanUp.ReferenceOTU9	3	0.02	0	0.00	1	0.02	1	0.03	0	0.00
ML2012 OTU 76	HM038237	0	0.00	0	0.00	0	0.00	0	0.00	10	0.58
ML2012 OTU 121	New.CleanUp.ReferenceOTU26	0	0.00	1	0.02	1	0.02	4	0.10	0	0.00
ML2012 OTU 113	New.CleanUp.ReferenceOTU17	0	0.00	1	0.02	0	0.00	2	0.05	0	0.00
ML2012 OTU 72	GU289732	0	0.00	4	0.06	5	0.08	12	0.31	9	0.52

ML2012 OTU 172	New.ReferenceOTU29	0	0.00	0	0.00	7	0.11	6	0.16	10	0.58
ML2012 OTU 43	EF422413	0	0.00	1	0.02	4	0.06	1	0.03	0	0.00
ML2012 OTU 1	AAQF01000123	0	0.00	0	0.00	0	0.00	0	0.00	4	0.23
ML2012 OTU 243	Y14594	0	0.00	11	0.18	12	0.19	20	0.52	0	0.00
ML2012 OTU 97	JQ245601	0	0.00	12	0.19	27	0.43	12	0.31	1	0.06
ML2012 OTU 216	New.ReferenceOTU71	0	0.00	0	0.00	1	0.02	1	0.03	1	0.06
ML2012 OTU 207	New.ReferenceOTU61	0	0.00	0	0.00	2	0.03	2	0.05	0	0.00
ML2012 OTU 136	New.CleanUp.ReferenceOTU6	0	0.00	2	0.03	5	0.08	2	0.05	0	0.00
ML2012 OTU 186	New.ReferenceOTU42	1	0.01	4	0.06	0	0.00	0	0.00	0	0.00
ML2012 OTU 231	New.ReferenceOTU86	1	0.01	2	0.03	5	0.08	0	0.00	0	0.00
ML2012 OTU 73	GU735094	17	0.11	0	0.00	3	0.05	1	0.03	0	0.00
ML2012 OTU 119	New.CleanUp.ReferenceOTU24	0	0.00	2	0.03	0	0.00	0	0.00	0	0.00
ML2012 OTU 120	New.CleanUp.ReferenceOTU25	1	0.01	1	0.02	0	0.00	0	0.00	0	0.00
ML2012 OTU 66	GQ347369	43	0.28	46	0.74	39	0.62	30	0.79	0	0.00
ML2012 OTU 156	New.ReferenceOTU14	0	0.00	0	0.00	0	0.00	0	0.00	17	0.99
ML2012 OTU 111	New.CleanUp.ReferenceOTU15	0	0.00	2	0.03	0	0.00	2	0.05	0	0.00
ML2012 OTU 100	L35540	0	0.00	7	0.11	13	0.21	7	0.18	6	0.35
ML2012 OTU 99	JQ246431	60	0.39	13	0.21	19	0.30	11	0.29	0	0.00
ML2012 OTU 128	New.CleanUp.ReferenceOTU42	1	0.01	1	0.02	0	0.00	2	0.05	0	0.00
ML2012 OTU 101	M93356	0	0.00	0	0.00	0	0.00	0	0.00	2	0.12
ML2012 OTU 154	New.ReferenceOTU12	33	0.21	6	0.10	3	0.05	1	0.03	0	0.00
ML2012 OTU 48	EU375807	0	0.00	5	0.08	2	0.03	3	0.08	0	0.00
ML2012 OTU 45	EU180983	1	0.01	0	0.00	1	0.02	2	0.05	0	0.00
ML2012 OTU 89	HQ857609	102	0.66	32	0.52	15	0.24	7	0.18	0	0.00
ML2012 OTU 170	New.ReferenceOTU27	0	0.00	0	0.00	0	0.00	0	0.00	9	0.52
ML2012 OTU 88	HQ616312	0	0.00	3	0.05	2	0.03	0	0.00	0	0.00
ML2012 OTU 71	GU235586	0	0.00	0	0.00	3	0.05	1	0.03	0	0.00
ML2012 OTU 224	New.ReferenceOTU79	1	0.01	5	0.08	3	0.05	1	0.03	0	0.00
ML2012 OTU 32	AJ532694	1	0.01	4	0.06	2	0.03	1	0.03	0	0.00
ML2012 OTU 53	FJ152889	50	0.32	14	0.23	5	0.08	1	0.03	0	0.00

ML2012 OTU 38	DQ432391	0	0.00	1	0.02	2	0.03	0	0.00	4	0.23
ML2012 OTU 87	HQ190509	0	0.00	2	0.03	2	0.03	1	0.03	0	0.00
ML2012 OTU 209	New.ReferenceOTU64	0	0.00	1	0.02	4	0.06	3	0.08	1	0.06
ML2012 OTU 191	New.ReferenceOTU47	0	0.00	3	0.05	9	0.14	15	0.39	2	0.12
ML2012 OTU 185	New.ReferenceOTU41	0	0.00	2	0.03	3	0.05	1	0.03	1	0.06
ML2012 OTU 174	New.ReferenceOTU30	0	0.00	0	0.00	4	0.06	7	0.18	1	0.06
ML2012 OTU 63	FM957461	7	0.05	3	0.05	14	0.22	0	0.00	0	0.00
ML2012 OTU 94	JQ228578	4	0.03	1	0.02	3	0.05	0	0.00	0	0.00
ML2012 OTU 242	X74705	0	0.00	0	0.00	0	0.00	0	0.00	3	0.17
ML2012 OTU 69	GU083688	0	0.00	5	0.08	1	0.02	0	0.00	0	0.00
ML2012 OTU 236	New.ReferenceOTU90	0	0.00	0	0.00	1	0.02	0	0.00	1	0.06
ML2012 OTU 229	New.ReferenceOTU83	0	0.00	3	0.05	2	0.03	2	0.05	1	0.06
ML2012 OTU 17	AF454308	0	0.00	0	0.00	6	0.10	4	0.10	2	0.12
ML2012 OTU 159	New.ReferenceOTU17	0	0.00	7	0.11	17	0.27	18	0.47	11	0.64
ML2012 OTU 145	New.CleanUp.ReferenceOTU87	0	0.00	0	0.00	2	0.03	0	0.00	0	0.00
ML2012 OTU 137	New.CleanUp.ReferenceOTU62	0	0.00	0	0.00	0	0.00	1	0.03	1	0.06
ML2012 OTU 129	New.CleanUp.ReferenceOTU44	0	0.00	0	0.00	2	0.03	0	0.00	0	0.00
ML2012 OTU 116	New.CleanUp.ReferenceOTU20	0	0.00	1	0.02	2	0.03	1	0.03	3	0.17
ML2012 OTU 96	JQ245579	0	0.00	0	0.00	1	0.02	2	0.05	0	0.00
ML2012 OTU 213	New.ReferenceOTU69	1	0.01	0	0.00	0	0.00	1	0.03	1	0.06
ML2012 OTU 234	New.ReferenceOTU89	0	0.00	0	0.00	3	0.05	1	0.03	1	0.06
ML2012 OTU 202	New.ReferenceOTU57	0	0.00	6	0.10	16	0.26	8	0.21	0	0.00
ML2012 OTU 93	JN680639	0	0.00	1	0.02	2	0.03	2	0.05	0	0.00
ML2012 OTU 210	New.ReferenceOTU66	0	0.00	0	0.00	3	0.05	1	0.03	0	0.00
ML2012 OTU 131	New.CleanUp.ReferenceOTU48	3	0.02	0	0.00	0	0.00	0	0.00	0	0.00
ML2012 OTU 132	New.CleanUp.ReferenceOTU49	1	0.01	0	0.00	1	0.02	0	0.00	0	0.00
ML2012 OTU 232	New.ReferenceOTU87	0	0.00	1	0.02	3	0.05	5	0.13	1	0.06
ML2012 OTU 227	New.ReferenceOTU81	0	0.00	1	0.02	3	0.05	2	0.05	0	0.00
ML2012 OTU 188	New.ReferenceOTU44	0	0.00	4	0.06	1	0.02	11	0.29	1	0.06
ML2012 OTU 130	New.CleanUp.ReferenceOTU45	0	0.00	0	0.00	1	0.02	3	0.08	0	0.00

ML2012 OTU 115	New.CleanUp.ReferenceOTU19	0	0.00	0	0.00	1	0.02	1	0.03	2	0.12
ML2012 OTU 199	New.ReferenceOTU54	6	0.04	9	0.15	3	0.05	0	0.00	0	0.00
ML2012 OTU 230	New.ReferenceOTU84	0	0.00	1	0.02	0	0.00	2	0.05	0	0.00
ML2012 OTU 112	New.CleanUp.ReferenceOTU16	3	0.02	0	0.00	1	0.02	0	0.00	0	0.00
ML2012 OTU 182	New.ReferenceOTU39	0	0.00	9	0.15	3	0.05	4	0.10	0	0.00
ML2012 OTU 140	New.CleanUp.ReferenceOTU7	0	0.00	1	0.02	3	0.05	5	0.13	0	0.00
ML2012 OTU 233	New.ReferenceOTU88	2	0.01	1	0.02	1	0.02	0	0.00	1	0.06
ML2012 OTU 220	New.ReferenceOTU75	0	0.00	1	0.02	3	0.05	4	0.10	0	0.00
ML2012 OTU 107	New.CleanUp.ReferenceOTU113	0	0.00	0	0.00	1	0.02	0	0.00	2	0.12
ML2012 OTU 222	New.ReferenceOTU77	6	0.04	9	0.15	7	0.11	5	0.13	0	0.00
ML2012 OTU 181	New.ReferenceOTU38	2	0.01	6	0.10	8	0.13	11	0.29	0	0.00
Totals:		15555	100.00	6180	100.00	6261	100.00	3811	100.00	1718	100.00

Table 2.S3. 16S rRNA amplicon sequencing OTU table with SILVA taxonomy. Purple OTUs are present at all depths and >1% relative abundance, yellow OTUs are >1% relative abundance at one or more depths and blue OTUs are present at all depths but less than 1% relative abundance.

OTU ID	QIIME-Assigned OTU Name	SILVA Taxonomy
ML2012 OTU 83	HM129833	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; unclassified Microbacteriaceae*
ML2012 OTU 3	AF448193	*Bacteria; Actinobacteria; Nitriliruptoria; Nitriliruptorales; Nitriliruptoraceae; Nitriliruptor*
ML2012 OTU 162	New.ReferenceOTU2	*Bacteria; Actinobacteria; Nitriliruptoria; Nitriliruptorales; Nitriliruptoraceae; Nitriliruptor*
ML2012 OTU 8	AF449782	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 6	AF449778	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML310M-34*
ML2012 OTU 173	New.ReferenceOTU3	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 26	AF507867	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Cryomorphaceae; Brumimicrobium*
ML2012 OTU 9	AF452599	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; Saprospiraceae; uncultured Saprospiraceae*
ML2012 OTU 7	AF449781	*Bacteria; Bacteroidetes; unclassified Bacteroidetes; unclassified Bacteroidetes; unclassified Bacteroidetes; ML602M-17*
ML2012 OTU 150	New.ReferenceOTU0	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; unclassified Ruminococcacaeae*
ML2012 OTU 36	DQ206410	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Dethiobacter*
ML2012 OTU 81	HM128337	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; unclassified Rhodobacteraceae*
ML2012 OTU 20	AF458284	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae; Desulfurivibrio*
ML2012 OTU 10	AF452605	*Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae; Thioalkalivibrio*
ML2012 OTU 13	AF453543	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; unclassified Oceanospillales; ML617.5J-3*
ML2012 OTU 11	AF452606	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; Marinicella*

ML2012 OTU 21	AF507855	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 16	AF454301	*Bacteria; Tenericutes; Mollicutes; unclassified Mollicutes; unclassified Mollicutes; NB1-n*
ML2012 OTU 158	New.ReferenceOTU16	*Bacteria; Tenericutes; Mollicutes; unclassified Mollicutes; unclassified Mollicutes; RF9*
ML2012 OTU 151	New.ReferenceOTU1	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria*
ML2012 OTU 22	AF507857	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 214	New.ReferenceOTU7	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 90	JF811035	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Psychroflexus*
ML2012 OTU 155	New.ReferenceOTU13	*Bacteria; Firmicutes; Bacilli; Bacillales; Paenibacillaceae; uncultured Paenibacillaceae*
ML2012 OTU 28	AF507875	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; uncultured Syntrophomonadaceae*
ML2012 OTU 194	New.ReferenceOTU5	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; uncultured Syntrophomonadaceae*
ML2012 OTU 166	New.ReferenceOTU23	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 64	GQ135733	*Bacteria; Planctomycetes; Phycisphaerae; unclassified phycisphaerae; unclassified phycisphaerae; ML-A-10*
ML2012 OTU 95	JQ231232	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; GKS98 freshwater group*
ML2012 OTU 80	HM128188	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales; Desulfonatronaceae; Desulfonatronum*
ML2012 OTU 12	AF452608	*Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae; Spiribacter*
ML2012 OTU 50	EU709859	*Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae; Thioalkalivibrio*
ML2012 OTU 169	New.ReferenceOTU26	*Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae; Spiribacter*
ML2012 OTU 41	DQ900622	*Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Thiomicrospira*
ML2012 OTU 176	New.ReferenceOTU32	*Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Thiomicrospira*
ML2012 OTU	AF507862	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic*

24		*group*
ML2012 OTU 177	New.ReferenceOTU33	*Bacteria; Bacteroidetes; Cytophagia; Cytophagales; Cyclobacteriaceae; uncultured Cyclobacteriaceae*
ML2012 OTU 187	New.ReferenceOTU43	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; A815*
ML2012 OTU 163	New.ReferenceOTU20	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML310M-34*
ML2012 OTU 225	New.ReferenceOTU8	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Cryomorphaceae; Owenweeksia*
ML2012 OTU 152	New.ReferenceOTU10	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Psychroflexus*
ML2012 OTU 15	AF454300	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus*
ML2012 OTU 82	HM128341	*Bacteria; Firmicutes; Clostridia; Clostridiales; Peptostreptococcaceae; uncultured Peptostreptococcaceae*
ML2012 OTU 153	New.ReferenceOTU11	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Dethiobacter*
ML2012 OTU 204	New.ReferenceOTU59	*Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae; Orenia*
ML2012 OTU 77	HM127957	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Salinarimonas*
ML2012 OTU 39	DQ432395	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; uncultured Rhodobacteraceae*
ML2012 OTU 195	New.ReferenceOTU50	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales; Rhodospirillaceae; uncultured Rhodospirillaceae*
ML2012 OTU 127	New.CleanUp.ReferenceOTU4	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales; unclassified Desulfuromonadales; GR-WP33-58*
ML2012 OTU 44	EF632658	*Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; Haliea*
ML2012 OTU 62	FJ999568	*Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; Haliea*
ML2012 OTU 165	New.ReferenceOTU22	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Coxiellaceae; Coxiella*
ML2012 OTU 157	New.ReferenceOTU15	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Legionellaceae; Legionella*
ML2012 OTU 205	New.ReferenceOTU6	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; unclassified Oceanospillales; OM182 clade*
ML2012 OTU 58	FJ153016	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; Methylonatrum*

ML2012 OTU 183	New.ReferenceOTU4	*Bacteria; Verrucomicrobia; Opitutae; unclassified opitutae; unclassified opitutae; unclassified opitutae*
ML2012 OTU 117	New.CleanUp.ReferenceOTU21	*Archaea; Euryarchaeota; Halobacteria; Halobacteriales; unclassified Halobacteriales; Deep Sea Hydrothermal Vent Gp 6(DHVEG-6)*
ML2012 OTU 84	HM130024	*Bacteria; Actinobacteria; Acidimicrobiia; Acidimicrobiales; unclassified Acidimicrobiales; unclassified Acidimicrobiales*
ML2012 OTU 109	New.CleanUp.ReferenceOTU13	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; DS001*
ML2012 OTU 56	FJ152933	*Bacteria; Actinobacteria; Nitriliruptoria; Nitriliruptorales; Nitriliruptoraceae; Nitriliruptor*
ML2012 OTU 40	DQ432396	*Bacteria; Actinobacteria; Nitriliruptoria; Nitriliruptorales; Nitriliruptoraceae; Nitriliruptor*
ML2012 OTU 91	JN003088	*Bacteria; Actinobacteria; Nitriliruptoria; Nitriliruptorales; Nitriliruptoraceae; Nitriliruptor*
ML2012 OTU 27	AF507869	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Marinilabiaceae; uncultured Marinilabiaceae*
ML2012 OTU 208	New.ReferenceOTU63	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Marinilabiaceae; uncultured Marinilabiaceae*
ML2012 OTU 211	New.ReferenceOTU67	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Rikenellaceae; Blvii28 wastewater-sludge group*
ML2012 OTU 198	New.ReferenceOTU53	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 219	New.ReferenceOTU74	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 5	AF449769	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 47	EU283512	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 4	AF449768	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 23	AF507860	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 138	New.CleanUp.ReferenceOTU65	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; ML635J-40 aquatic group*
ML2012 OTU 218	New.ReferenceOTU73	*Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; unclassified Bacteroidales; unclassified Bacteroidales*
ML2012 OTU 59	FJ485151	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; CK06-06-Mud-MAS4B-21*
ML2012 OTU	New.ReferenceOTU78	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML310M-34*

223		
ML2012 OTU 228	New.ReferenceOTU82	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML310M-34*
ML2012 OTU 201	New.ReferenceOTU56	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML310M-34*
ML2012 OTU 79	HM128153	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 33	DQ015831	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 54	FJ152901	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 189	New.ReferenceOTU45	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 238	New.ReferenceOTU93	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 240	New.ReferenceOTU96	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 55	FJ152914	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 52	EU911993	*Bacteria; Bacteroidetes; Cytophagia; unclassified Cytophagia; Order III Incertae Sedis; ML602J-37*
ML2012 OTU 57	FJ152979	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Cryomorphaceae; Cryomorpha*
ML2012 OTU 49	EU703291	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Cryomorphaceae; Fluviicola*
ML2012 OTU 123	New.CleanUp.ReferenceOTU28	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Cryomorphaceae; Owenweeksia*
ML2012 OTU 160	New.ReferenceOTU18	*Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Psychroflexus*
ML2012 OTU 25	AF507866	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; Chitinophagaceae; uncultured Chitinophagaceae*
ML2012 OTU 226	New.ReferenceOTU80	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; Chitinophagaceae; uncultured Chitinophagaceae*
ML2012 OTU 133	New.CleanUp.ReferenceOTU5	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; Saprospiraceae; uncultured Saprospiraceae*
ML2012 OTU 135	New.CleanUp.ReferenceOTU54	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; unclassified Sphingobacteriales; B01R012*
ML2012 OTU 164	New.ReferenceOTU21	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; unclassified Sphingobacteriales; E6aC02*

ML2012 OTU 139	New.CleanUp.ReferenceOTU66	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; unclassified Sphingobacteriales; unclassified Sphingobacteriales*
ML2012 OTU 184	New.ReferenceOTU40	*Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales; unclassified Sphingobacteriales; WCHB1-69*
ML2012 OTU 70	GU083693	*Bacteria; Bacteroidetes; unclassified Bacteroidetes; unclassified Bacteroidetes; unclassified Bacteroidetes; BD2-2*
ML2012 OTU 125	New.CleanUp.ReferenceOTU3	*Bacteria; Bacteroidetes; unclassified Bacteroidetes; unclassified Bacteroidetes; unclassified Bacteroidetes; BD2-2*
ML2012 OTU 143	New.CleanUp.ReferenceOTU74	*Bacteria; BD1-5; unclassified BD1-5; unclassified BD1-5; unclassified BD1-5; unclassified BD1-5*
ML2012 OTU 148	New.CleanUp.ReferenceOTU94	*Bacteria; BHI80-139; unclassified BHI80-139; unclassified BHI80-139; unclassified BHI80-139; unclassified BHI80-139*
ML2012 OTU 85	HM468058	*Bacteria; Candidate division BRC1; unclassified Candidate division BRC1; unclassified Candidate division BRC1; unclassified Candidate division BRC1; unclassified Candidate division BRC1*
ML2012 OTU 67	GQ377778	*Bacteria; Cyanobacteria; unclassified Cyanobacteria; Cyanobacteria Subsection I; Cyanobacteria Family I; Synechococcus*
ML2012 OTU 241	New.ReferenceOTU98	*Bacteria; Cyanobacteria; unclassified Cyanobacteria; unclassified Cyanobacteria; unclassified Cyanobacteria; ML635J-21*
ML2012 OTU 146	New.CleanUp.ReferenceOTU88	*Bacteria; Cyanobacteria; unclassified Cyanobacteria; unclassified Cyanobacteria; unclassified Cyanobacteria; ML635J-21*
ML2012 OTU 178	New.ReferenceOTU34	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus*
ML2012 OTU 118	New.CleanUp.ReferenceOTU22	*Bacteria; Firmicutes; Clostridia; Clostridiales; Caldicoprobacteraceae; Caldicoprobacter*
ML2012 OTU 149	New.CleanUp.ReferenceOTU95	*Bacteria; Firmicutes; Clostridia; Clostridiales; Christensenellaceae; Christensenella*
ML2012 OTU 98	JQ245609	*Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Anoxynatronum*
ML2012 OTU 31	AJ271453	*Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Anoxynatronum*
ML2012 OTU 217	New.ReferenceOTU72	*Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Anoxynatronum*
ML2012 OTU 2	AB294306	*Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Anoxynatronum*
ML2012 OTU 46	EU245382	*Bacteria; Firmicutes; Clostridia; Clostridiales; Family XI; unclassified Clostridiales*
ML2012 OTU 221	New.ReferenceOTU76	*Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae; Desulfitibacter*
ML2012 OTU	FJ788525	*Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae; Desulfitispora*

60		
ML2012 OTU 126	New.CleanUp.ReferenceOTU31	*Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae; Desulfitispora*
ML2012 OTU 104	New.CleanUp.ReferenceOTU10	*Bacteria; Firmicutes; Clostridia; Clostridiales; Peptostreptococcaceae; uncultured Peptostreptococcaceae*
ML2012 OTU 37	DQ206423	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Fastidiosipila*
ML2012 OTU 175	New.ReferenceOTU31	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; unclassified Ruminococcaceae*
ML2012 OTU 103	New.CleanUp.ReferenceOTU1	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; unclassified Ruminococcaceae*
ML2012 OTU 51	EU843077	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; uncultured Ruminococcaceae*
ML2012 OTU 171	New.ReferenceOTU28	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; uncultured Ruminococcaceae*
ML2012 OTU 142	New.CleanUp.ReferenceOTU72	*Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; uncultured Ruminococcaceae*
ML2012 OTU 34	DQ124682	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Candidatus Contubernalis*
ML2012 OTU 102	New.CleanUp.ReferenceOTU0	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Dethiobacter*
ML2012 OTU 68	GQ848216	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; uncultured Syntrophomonadaceae*
ML2012 OTU 29	AF507877	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; uncultured Syntrophomonadaceae*
ML2012 OTU 14	AF454298	*Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; uncultured Syntrophomonadaceae*
ML2012 OTU 200	New.ReferenceOTU55	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; MAT-CR-H4-C10*
ML2012 OTU 167	New.ReferenceOTU24	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 30	AF507890	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 235	New.ReferenceOTU9	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 203	New.ReferenceOTU58	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 192	New.ReferenceOTU48	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*

ML2012 OTU 180	New.ReferenceOTU36	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 108	New.CleanUp.ReferenceOTU12	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; OPB54*
ML2012 OTU 42	DQ988278	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; P. palm C-A 51*
ML2012 OTU 190	New.ReferenceOTU46	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; unclassified Clostridiales*
ML2012 OTU 106	New.CleanUp.ReferenceOTU110	*Bacteria; Firmicutes; Clostridia; Clostridiales; unclassified Clostridiales; unclassified Clostridiales*
ML2012 OTU 179	New.ReferenceOTU35	*Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halanaerobiaceae; Halocella*
ML2012 OTU 161	New.ReferenceOTU19	*Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halanaerobiaceae; Halocella*
ML2012 OTU 237	New.ReferenceOTU92	*Bacteria; Firmicutes; Erysipelotrichi; Erysipelotrichales; Erysipelotrichaceae; Asteroleplasma*
ML2012 OTU 122	New.CleanUp.ReferenceOTU27	*Bacteria; Firmicutes; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes*
ML2012 OTU 105	New.CleanUp.ReferenceOTU11	*Bacteria; Firmicutes; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes*
ML2012 OTU 75	HE611101	*Bacteria; Fusobacteria; Fusobacteriia; Fusobacteriales; Leptotrichiaceae; uncultured Leptotrichiaceae*
ML2012 OTU 197	New.ReferenceOTU52	*Bacteria; Gemmatimonadetes; Gemmatimonadetes; unclassified Gemmatimonadetes; unclassified Gemmatimonadetes; BD2-11 terrestrial group*
ML2012 OTU 141	New.CleanUp.ReferenceOTU70	*Bacteria; Lentisphaerae; Lentisphaeria; unclassified Lentisphaeria; unclassified Lentisphaeria; LD1-PA34*
ML2012 OTU 196	New.ReferenceOTU51	*Bacteria; Lentisphaerae; Lentisphaeria; unclassified Lentisphaeria; unclassified Lentisphaeria; LD1-PB3*
ML2012 OTU 19	AF454310	*Bacteria; Lentisphaerae; Lentisphaeria; unclassified Lentisphaeria; unclassified Lentisphaeria; ML1228J-2*
ML2012 OTU 18	AF454309	*Bacteria; Lentisphaerae; Lentisphaeria; unclassified Lentisphaeria; unclassified Lentisphaeria; ML1228J-2*
ML2012 OTU 134	New.CleanUp.ReferenceOTU53	*Bacteria; Lentisphaerae; Lentisphaeria; unclassified Lentisphaeria; unclassified Lentisphaeria; MSBL3*
ML2012 OTU 212	New.ReferenceOTU68	*Bacteria; Lentisphaerae; Lentisphaeria; unclassified Lentisphaeria; unclassified Lentisphaeria; SS1-B-03-39*
ML2012 OTU 239	New.ReferenceOTU94	*Bacteria; Lentisphaerae; Lentisphaeria; Victivallales; Victivallaceae; uncultured Victivallaceae*
ML2012 OTU	New.ReferenceOTU49	*Bacteria; Lentisphaerae; Lentisphaeria; Victivallales; Victivallaceae; uncultured Victivallaceae*

193		
ML2012 OTU 168	New.ReferenceOTU25	*Bacteria; Lentisphaerae; Lentisphaeria; Victivallales; Victivallaceae; uncultured Victivallaceae*
ML2012 OTU 65	GQ138312	*Bacteria; Planctomycetes; Phycisphaerae; unclassified phycisphaerae; unclassified phycisphaerae; KCLunmb-38-53*
ML2012 OTU 35	DQ206406	*Bacteria; Planctomycetes; Phycisphaerae; unclassified phycisphaerae; unclassified phycisphaerae; ML-A-10*
ML2012 OTU 206	New.ReferenceOTU60	*Bacteria; Planctomycetes; Phycisphaerae; unclassified phycisphaerae; unclassified phycisphaerae; MSBL9*
ML2012 OTU 78	HM128005	*Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales; Planctomycetaceae; Rhodopirellula*
ML2012 OTU 86	HQ166789	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Roseovarius*
ML2012 OTU 92	JN637818	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Roseovarius*
ML2012 OTU 74	HE574924	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; uncultured Rhodobacteraceae*
ML2012 OTU 114	New.CleanUp.ReferenceOTU18	*Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; uncultured Rhodobacteraceae*
ML2012 OTU 124	New.CleanUp.ReferenceOTU29	*Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Rickettsiaceae; Rickettsia*
ML2012 OTU 110	New.CleanUp.ReferenceOTU14	*Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; unclassified Rickettsiales; Candidatus Odyssella*
ML2012 OTU 215	New.ReferenceOTU70	*Bacteria; Proteobacteria; Alphaproteobacteria; Unclassified alphaproteobacterium; Unclassified alphaproteobacterium; LWSR 14*
ML2012 OTU 144	New.CleanUp.ReferenceOTU8	*Bacteria; Proteobacteria; Alphaproteobacteria; Unclassified alphaproteobacterium; Unclassified alphaproteobacterium; Unclassified alphaproteobacterium*
ML2012 OTU 61	FJ973586	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; GKS98 freshwater group*
ML2012 OTU 147	New.CleanUp.ReferenceOTU9	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; GKS98 freshwater group*
ML2012 OTU 76	HM038237	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; GKS98 freshwater group*
ML2012 OTU 121	New.CleanUp.ReferenceOTU26	*Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales; Bacteriovoracaceae; Bacteriovorax*
ML2012 OTU 113	New.CleanUp.ReferenceOTU17	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobacteraceae; Desulfobotulus*
ML2012 OTU 72	GU289732	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobacteraceae; Desulfosalsimonas*

ML2012 OTU 172	New.ReferenceOTU29	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobacteraceae; uncultured Desulfobacteraceae*
ML2012 OTU 43	EF422413	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae; Desulfurivibrio*
ML2012 OTU 1	AAQF01000123	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobulbaceae; Desulfurivibrio*
ML2012 OTU 243	Y14594	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales; Desulfonatronaceae; Desulfonatronum*
ML2012 OTU 97	JQ245601	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales; Desulfuromonadaceae; Desulfuromusa*
ML2012 OTU 216	New.ReferenceOTU71	*Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales; unclassified Desulfuromonadales; GR-WP33-58*
ML2012 OTU 207	New.ReferenceOTU61	*Bacteria; Proteobacteria; Deltaproteobacteria; Syntrophobacterales; Syntrophaceae; uncultured Syntrophaceae*
ML2012 OTU 136	New.CleanUp.ReferenceOTU6	*Bacteria; Proteobacteria; Deltaproteobacteria; unclassified deltaproteobacteria; unclassified deltaproteobacteria; unclassified deltaproteobacteria*
ML2012 OTU 186	New.ReferenceOTU42	*Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; C1-B045*
ML2012 OTU 231	New.ReferenceOTU86	*Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; uncultured Alteromonadaceae*
ML2012 OTU 73	GU735094	*Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales; Ectothiorhodospiraceae; Thioalkalivibrio*
ML2012 OTU 119	New.CleanUp.ReferenceOTU24	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Coxiellaceae; Aquicella*
ML2012 OTU 120	New.CleanUp.ReferenceOTU25	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Legionellaceae; Legionella*
ML2012 OTU 66	GQ347369	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Legionellaceae; uncultured Legionellaceae*
ML2012 OTU 156	New.ReferenceOTU14	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Legionellaceae; uncultured Legionellaceae*
ML2012 OTU 111	New.CleanUp.ReferenceOTU15	*Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Legionellaceae; uncultured Legionellaceae*
ML2012 OTU 100	L35540	*Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylomicrobium*
ML2012 OTU 99	JQ246431	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Halomonadaceae; Halomonas*
ML2012 OTU 128	New.CleanUp.ReferenceOTU42	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Halomonadaceae; Halomonas*
ML2012 OTU	M93356	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Halomonadaceae; Halomonas*

101		
ML2012 OTU 154	New.ReferenceOTU12	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Halomonadaceae; unclassified Halomonadaceae*
ML2012 OTU 48	EU375807	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Oceanospirillaceae; Marinospirillum*
ML2012 OTU 45	EU180983	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Oceanospirillaceae; Nitrincola*
ML2012 OTU 89	HQ857609	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Oceanospirillaceae; Pseudospirillum*
ML2012 OTU 170	New.ReferenceOTU27	*Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Oceanospirillaceae; Pseudospirillum*
ML2012 OTU 88	HQ616312	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas*
ML2012 OTU 71	GU235586	*Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Francisellaceae; Francisella*
ML2012 OTU 224	New.ReferenceOTU79	*Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Thiomicrospira*
ML2012 OTU 32	AJ532694	*Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Thiomicrospira*
ML2012 OTU 53	FJ152889	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified Gammaproteobacteria; unclassified Gammaproteobacteria; HOC36*
ML2012 OTU 38	DQ432391	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; HOC36*
ML2012 OTU 87	HQ190509	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; Methylonatrum*
ML2012 OTU 209	New.ReferenceOTU64	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; NKB5*
ML2012 OTU 191	New.ReferenceOTU47	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; NKB5*
ML2012 OTU 185	New.ReferenceOTU41	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria*
ML2012 OTU 174	New.ReferenceOTU30	*Bacteria; Proteobacteria; Gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria; unclassified gammaproteobacteria*
ML2012 OTU 63	FM957461	*Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio*
ML2012 OTU 94	JQ228578	*Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio*
ML2012 OTU 242	X74705	*Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio*

ML2012 OTU 69	GU083688	*Bacteria; Proteobacteria; unclassified Proteobacteria; unclassified Proteobacteria; unclassified Proteobacteria; ARKICE-90*
ML2012 OTU 236	New.ReferenceOTU90	*Bacteria; SBYG-2791; unclassified SBYG-2791; unclassified SBYG-2791; unclassified SBYG-2791; unclassified SBYG-2791*
ML2012 OTU 229	New.ReferenceOTU83	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 17	AF454308	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 159	New.ReferenceOTU17	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 145	New.CleanUp.ReferenceOTU87	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 137	New.CleanUp.ReferenceOTU62	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 129	New.CleanUp.ReferenceOTU44	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 116	New.CleanUp.ReferenceOTU20	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; Spirochaeta*
ML2012 OTU 96	JQ245579	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; Spirochaetaceae; uncultured Spirochaetaceae*
ML2012 OTU 213	New.ReferenceOTU69	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; unclassified Spirochaetales; PL-11B10*
ML2012 OTU 234	New.ReferenceOTU89	*Bacteria; Spirochaetes; Spirochaetes; Spirochaetales; unclassified Spirochaetales; PL-11B10*
ML2012 OTU 202	New.ReferenceOTU57	*Bacteria; Tenericutes; Mollicutes; unclassified Mollicutes; unclassified Mollicutes; NB1-n*
ML2012 OTU 93	JN680639	*Bacteria; Tenericutes; Mollicutes; unclassified Mollicutes; unclassified Mollicutes; RF9*
ML2012 OTU 210	New.ReferenceOTU66	*Bacteria; Tenericutes; Mollicutes; unclassified Mollicutes; unclassified Mollicutes; RF9*
ML2012 OTU 131	New.CleanUp.ReferenceOTU48	*Bacteria; TM6; unclassified TM6; unclassified TM6; unclassified TM6; unclassified TM6*
ML2012 OTU 132	New.CleanUp.ReferenceOTU49	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria*
ML2012 OTU 232	New.ReferenceOTU87	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria*
ML2012 OTU 227	New.ReferenceOTU81	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria*
ML2012 OTU	New.ReferenceOTU44	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria;*

188		*unclassified bacteria*
ML2012 OTU 130	New.CleanUp.ReferenceOTU45	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria*
ML2012 OTU 115	New.CleanUp.ReferenceOTU19	*Bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria; unclassified bacteria*
ML2012 OTU 199	New.ReferenceOTU54	*Bacteria; unclassified bacteria; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes*
ML2012 OTU 230	New.ReferenceOTU84	*Bacteria; unclassified bacteria; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes; unclassified Firmicutes*
ML2012 OTU 112	New.CleanUp.ReferenceOTU16	*Bacteria; Verrucomicrobia; Opitutae; Opitutales; Opitutaceae; unclassified Opitutaceae*
ML2012 OTU 182	New.ReferenceOTU39	*Bacteria; Verrucomicrobia; Opitutae; Opitutales; Opitutaceae; unclassified Opitutaceae*
ML2012 OTU 140	New.CleanUp.ReferenceOTU7	*Bacteria; Verrucomicrobia; Opitutae; Puniceicoccales; Puniceicoccaceae; Puniceicoccus*
ML2012 OTU 233	New.ReferenceOTU88	*Bacteria; Verrucomicrobia; Opitutae; Puniceicoccales; Puniceicoccaceae; unclassified Puniceicoccaceae*
ML2012 OTU 220	New.ReferenceOTU75	*Bacteria; Verrucomicrobia; Opitutae; Puniceicoccales; Puniceicoccaceae; unclassified Puniceicoccaceae*
ML2012 OTU 107	New.CleanUp.ReferenceOTU113	*Bacteria; Verrucomicrobia; Opitutae; Puniceicoccales; Puniceicoccaceae; unclassified Puniceicoccaceae*
ML2012 OTU 222	New.ReferenceOTU77	*Bacteria; Verrucomicrobia; Opitutae; Puniceicoccales; Puniceicoccaceae; uncultured Puniceicoccaceae*
ML2012 OTU 181	New.ReferenceOTU38	*Unclassified; unclassified root; unclassified root; unclassified root; unclassified root; unclassified root*

Table 2.S4. Summary statistics for metatranscriptome sequencing.

Read Statistics	Library[a]									
	10a	**10b**	**15a**	**15b**	**18a**	**18b**	**25a**	**25b**	**31a**	**31b**
Raw Unpaired Reads[b]	21.36	23.13	25.58	31.67	29.55	20.73	44.06	33.01	32.02	26.91
Paired Reads	9029458	10085082	11748031	11524766	10448205	8610511	19077062	13118718	12240380	9 116 277
Quality Trimmed	8382894	9474888	11052641	10396583	9696575	8036024	17974417	12605162	11790853	8727294
Non-rRNA (% of good reads)	4746336 (56.62)	6985620 (73.73)	6122080 (55.39)	6473827 (62.27)	6404351 (66.05)	4987006 (62.06)	9599613 (53.41)	6334622 (50.25)	6453264 (54.73)	4854897 (55.63)
Internal Standard Recovery %[c]	0.000082	0.000151	0.000074	0.000103	0.000115	0.000087	0.000077	0.000070	0.000054	0.000074
Average Read Length	251	242	232	247	244	253	246	228	226	236
Number of Hits to RefSeq (% of non-rRNA)	1189373 (25)	1881713 (27)	2385704 (39)	2627661 (41)	2556907 (40)	2188271 (44)	4003459 (42)	2228664 (36)	2496333 (39)	1894806 (39)
Domain Distribution (B/A/E/V)[d]	58/0.1/ 40.7/1.1	71.4/0.2/ 27.4/1	67.8/0.4/ 30.8/1	68.9/0.4/ 29.7/1	72.1/0.9/ 26.1/1	79.1/0.7/ 19.4/0.8	67/0.7/ 31.3/0.9	67.5/0.7/ 31.3/0.4	62/0.7/ 36.9/0.4	68.4/0.8/ 30.4/0.4

[a]number indicates depth, letter indicates replicate
[b]million
[c]number of standards counted in "good" non-rRNA reads/number of standards added (average of two standards)
[d]B=Bacteria, A=Archaea, E=Eukaryota, V=Virus

Table 2.S5. Summary of bacteria Refseq hits.

	Library[a]									
Read Statistics	**10a**	**10b**	**15a**	**15b**	**18a**	**18b**	**25a**	**25b**	**31a**	**31b**
Number of Bacteria Hits to RefSeq	689899	1343402	1617784	1809515	1843678	1730129	2682441	1504932	1547612	1295511
Putative/Hypothetical and "Bad Accessions"*	361278 (52.4)	696032 (51.8)	832047 (51.4)	806040 (44.5)	957400 (51.9)	987830 (57.1)	1547441 (57.7)	1547441 (56.2)	845854 (58.5)	2682441 (58.2)
Protein Encoding Transcripts	328621 (47.6)	647370 (48.2)	785737 (48.6)	1003475 (55.5)	886278 (48.1)	742299 (42.9)	1135000 (42.3)	659078 (43.8)	642357 (41.5)	540965 (41.8)

*see text.

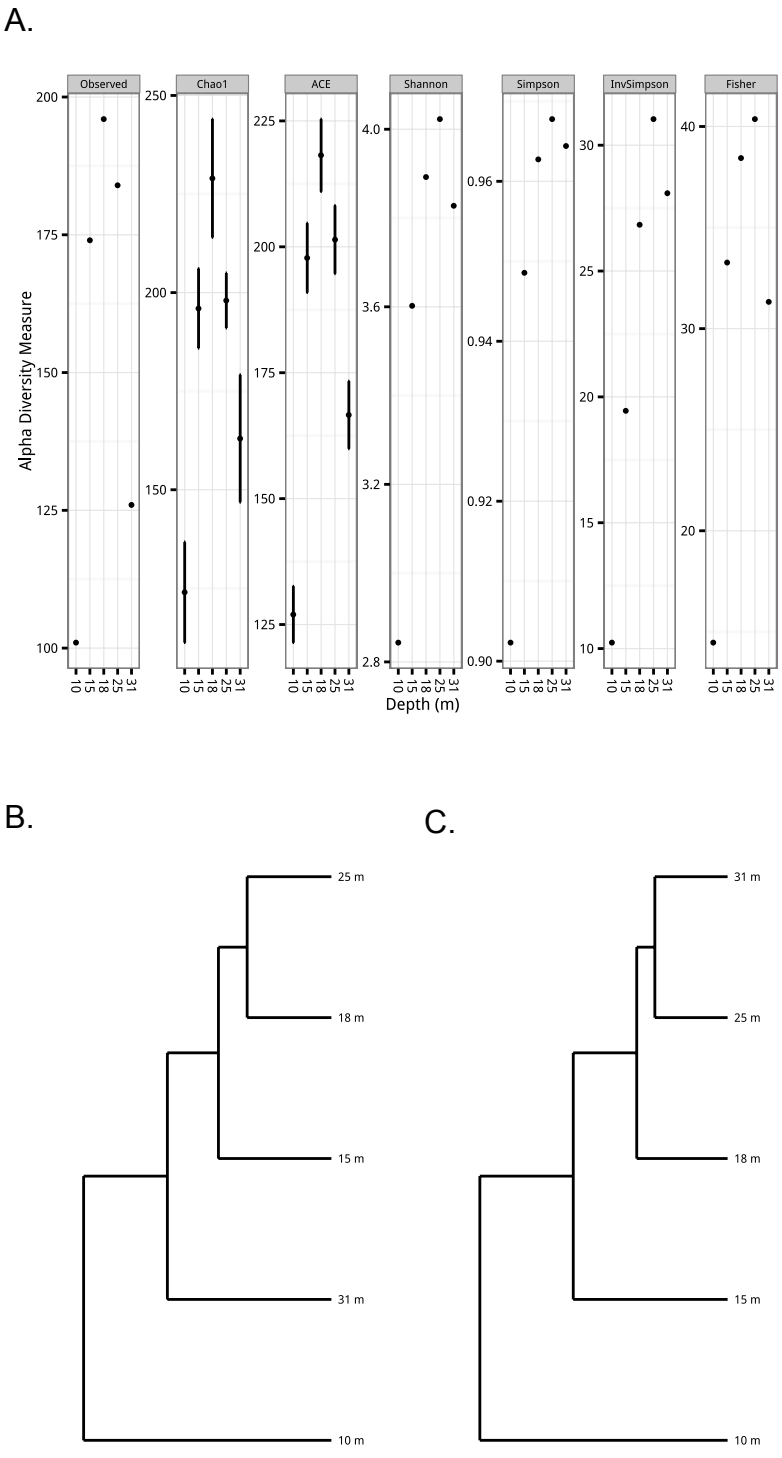


Figure 2.S1. Diversity metrics. A) Alpha diversity metrics, B) unweighted Unifrac and C) weighted Unifrac heirarchical clustering dendrograms, clustered by sample depth.

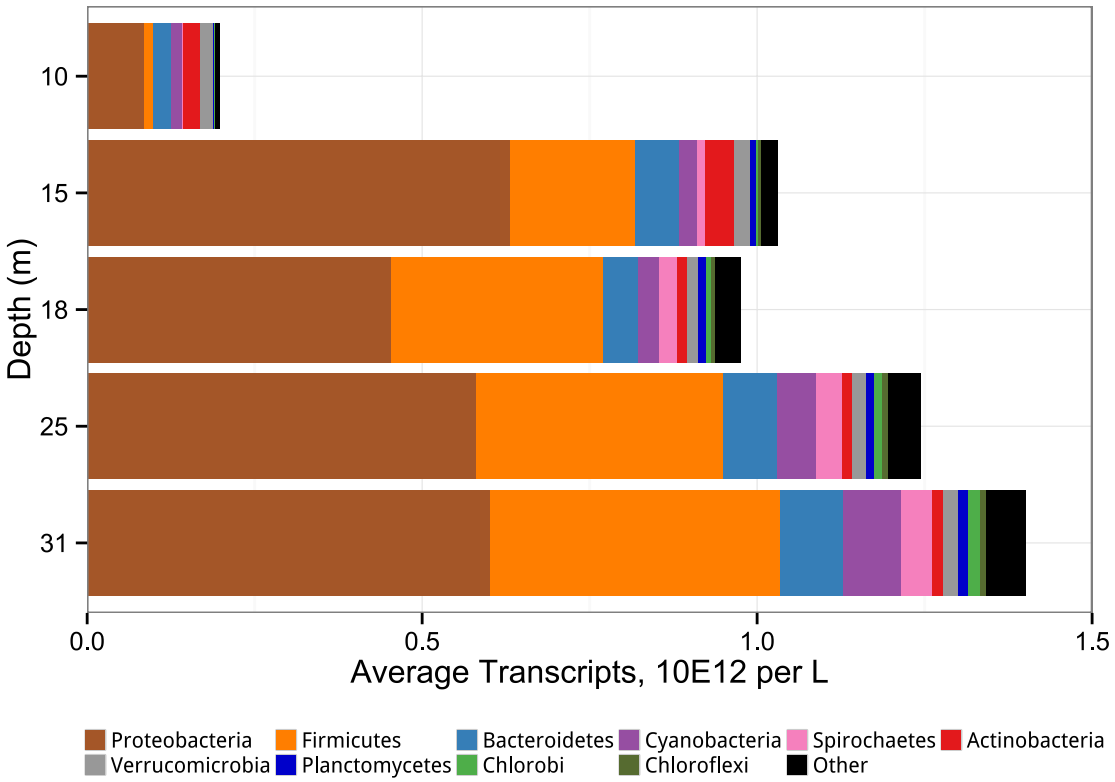


Figure 2.S2. Absolute transcript abundance by depth (phylum level).

APPENDIX B

SUPPORTING INFORMATION FOR CHAPTER 3

Table 3.S1. Molybdopterin oxidoreductase (CISM) catalytic subunit custom database information.

Protein	Protein Name	Organism	General Classification	Protein ID (NCBI or UniProtKB)	Database Accession
AioA	Arsenite Oxidase	*Marinobacter santoriniensis*	Bacteria, Proteobacteria (gamma)	WP_008937809	arsenite oxidase large subunit
AioA	Arsenite Oxidase	*Ralstonia* sp. 22	Bacteria, Proteobacteria (beta)	ACX69823	arsenite oxidase Mo-pterin subunit
AioA	Arsenite Oxidase	*arsenite-oxidising bacterium* NT-26	Bacteria, Proteobacteria (alpha)	AAR05656	AroA
AioA	Arsenite Oxidase	*Nitrobacter hamburgensis* X14	Bacteria, Proteobacteria (alpha)	YP_571843	arsenate reductase (azurin)
AioA	Arsenite Oxidase	*Thiomonas arsenitoxydans*	Bacteria, Proteobacteria (beta)	CAM58792	arsenite oxidase catalytic subunit
AioA	Arsenite Oxidase	*Aeropyrum pernix* K1	Archaea, Crenarchaeota, Thermoprotei	NP_148692	arsenite oxidase large subunit
AioA	Arsenite Oxidase	*Chloroflexus aurantiacus* J-10-fl	Bacteria, Chloroflex, Chloroflexales	YP_001634827	arsenite oxidase large subunit
AioA	Arsenite Oxidase	*Thermus thermophilus* HB8	Bacteria, Deinococcus-Thermus, Deinococci	BAD71923	arsenite oxidase, large subunit
AioA	Arsenite Oxidase	*Chlorobium limicola* DSM 245	Bacteria, Chlorobi, Chlorobia	YP_001942454	arsenite oxidase, large subunit
AioA	Arsenite Oxidase	*Sulfolobus tokodaii* str. 7	Archaea, Crenarchaeota, Thermoprotei	NP_378391	hypothetical protein ST2391
AioA	Arsenite Oxidase	*Alcaligenes faecalis*	Bacteria, Proteobacteria (beta)	Q7SIF4	RecName: Full=Arsenite oxidase subunit AioA; Short=AOI; AltName: Full=Arsenite oxidase Mo-pterin subunit
AioA	Arsenite Oxidase	*Herminiimonas arsenicoxydans*	Bacteria, Proteobacteria	WP_011870033	arsenite oxidase subunit AioA

			(beta)		
ArrA	Arsenate Reductase	*Shewanella* sp. ANA-3	Bacteria, Proteobacteria (gamma)	AAQ01672	molybdopterin oxidoreductase
ArrA	Arsenate Reductase	*Chrysiogenes arsenatis* DSM 11915	Bacteria, Chrysiogenetes, Chrysiogenates	AAU11839	arsenate respiratory reductase
ArrA	Arsenate Reductase	*Halarsenatibacter silvermanii*	Bacteria, Firmicutes, Clostridia	ACF74513	arsenate respiratory reductase catalytic subunit
ArrA	Arsenate Reductase	*Sulfurospirillum barnesii* SES-3	Bacteria, Proteobacteria (epsilon)	AAU11840	arsenate respiratory reductase, partial
ArrA	Arsenate Reductase	*Desulfosporosinus* sp. Y5	Bacteria, Firmicutes, Clostridia	ABB02056	arsenate respiratory reductase, partial
ArrA	Arsenate Reductase	*Alkaliphilus oremlandii* OhILAs	Bacteria, Firmicutes, Clostridia	YP_001512638	molybdopterin oxidoreductase
ArrA	Arsenate Reductase	*Denitrovibrio acetiphilus* DSM 12809	Bacteria, Deferribacteres, Deferribacteres	YP_003504839	molybdopterin oxidoreductase
ArrA	Arsenate Reductase	delta proteobacterium MLMS-1	Bacteria, Proteobacteria (delta)	WP_007292497	dehydrogenase
ArrA	Arsenate Reductase	*Pyrobaculum aerophilum* str. IM2	Archaea, Crenarchaeota, Thermoprotei	NP_559177	molybdopterin oxidoreductase, molybdopterin binding subunit
ArrA	Arsenate Reductase	*Bacillus selenitireducens*	Bacteria, Firmicutes	YP_003700672	molybdopterin oxidoreductase Fe4S4 region
ArxA	Anaerobic Arsenite Oxidase	*Alkalilimnicola ehrlichii* MLHE-1	Bacteria, Proteobacteria (gamma)	YP_743256	molybdopterin oxidoreductase
ArxA	Anaerobic Arsenite Oxidase	*Halorhodospira halophila* SL1	Bacteria, Proteobacteria (gamma)	YP_001001949	formate dehydrogenase
ArxA	Anaerobic Arsenite Oxidase	*Thioalkalivibrio* sp. ALM2T	Bacteria, Proteobacteria (gamma)	WP_019593327	molybdopterin oxidoreductase

ArxA	Anaerobic Arsenite Oxidase	*Alkalilimnicola ehrlichii* MLHE-1	Bacteria, Proteobacteria (gamma)	YP_741061	molybdopterin oxidoreductase
ArxA	Anaerobic Arsenite Oxidase	*Ectothiorhodospira* sp. PHS-1	Bacteria, Proteobacteria (gamma)	WP_008932021	molybdopterin oxidoreductase
ClrA	Chlorate reductase	*Halomonas jeotgali*	Bacteria, Proteobacteria (gamma)	WP_017429402	hypothetical protein
ClrA	Chlorate reductase	*Pseudomonas chloritidismutans*	Bacteria, Proteobacteria (gamma)	ADO63825	putative chlorate reductase ClrA
ClrA	Chlorate reductase	*Ideonella dechloratans*	Bacteria, Proteobacteria (beta)	P60068	RecName: Full=Chlorate reductase subunit alpha; AltName: Full=Chlorate reductase molybdenum subunit; Flags: Precursor
DdhA	Dimethylsulfide Dehydrogenase	*Sagittula stellata*	Bacteria, Proteobacteria (alpha)	WP_005861173	dimethylsulfide dehydrogenase subunit alpha
DdhA	Dimethylsulfide Dehydrogenase	*Thiorhodococcus drewsii*	Bacteria, Proteobacteria (gamma)	WP_007041785	dimethylsulfide dehydrogenase subunit alpha
DdhA	Dimethylsulfide Dehydrogenase	*Rhodovulum sulfidophilum*	Bacteria, Proteobacteria (alpha)	Q8GPG4	RecName: Full=Dimethylsulfide dehydrogenase subunit alpha; AltName: Full=DMS DH molybdenum subunit; AltName: Full=DMS DH subunit alpha; AltName: Full=Dimethyl sulfide:cytochrome c2 reductase subunit alpha; AltName: Full=Dimethylsulfide dehydrogenase molybdenum subunit; Flags: Precursor
DmsA	Membranous DMSO Reductase	*Vibrio fischeri* MJ11	Bacteria, Proteobacteria (gamma)	YP_002157663	anaerobic dimethyl sulfoxide reductase chain a
DmsA	Membranous DMSO Reductase	*Desulfovibrio vulgaris* str. 'Miyazaki F'	Bacteria, Proteobacteria (delta)	YP_002436028	anaerobic dimethyl sulfoxide reductase subunit A
DmsA	Membranous DMSO Reductase	*Propionibacterium acnes* KPA171202	Bacteria, Actinobacteria, Actinomycetales	YP_055228	anaerobic dimethyl sulfoxide reductase subunit A
DmsA	Membranous DMSO Reductase	*Bacillus azotoformans*	Bacteria, Firmicutes, Bacillales	WP_003332889	anaerobic dimethyl sulfoxide reductase, A subunit, DmsA/YnfE

DmsA	Membranous DMSO Reductase	*Salmonella enterica*	Bacteria, Proteobacteria (gamma)	WP_000850266	dimethyl sulfoxide reductase subunit A
DmsA	Membranous DMSO Reductase	*Laribacter hongkongensis* HLHK9	Bacteria, Proteobacteria (beta)	YP_002796487	DmsA
DmsA	Membranous DMSO Reductase	*Alkaliphilus metalliredigens* QYMF	Bacteria, Firmicutes, Clostridia	YP_001319191	molybdopterin oxidoreductase
DmsA	Membranous DMSO Reductase	*Desulfitobacterium hafniense*	Bacteria, Firmicutes, Clostridia	WP_026183857	DMSO reductase
DmsA	Membranous DMSO Reductase	*Wolinella succinogenes* DSM 1740	Bacteria, Proteobacteria (epsilon)	NP_907591	anaerobic dimethyl sulfoxide reductase subunit A
DmsA	Membranous DMSO Reductase	*Halobacterium* sp NRC-1	Archaea, Eyryaechaeota, Halobacteria	NP_279804	dimethylsulfoxide reductase
DorA	DMSO Reductase	*Sulfurospirillum deleyianum* DSM 6946	Bacteria, Proteobacteria (epsilon)	YP_003303806	molybdopterin guanine dinucleotide-containing S/N-oxide reductase
DorA	DMSO Reductase	*Rhodobacter sphaeroides*	Bacteria, Proteobacteria (alpha)	Q57366	RecName: Full=Dimethyl sulfoxide/trimethylamine N-oxide reductase; Short=DMSO reductase; Short=DMSOR; Short=Me2SO reductase; Short=TMAOR; Flags: Precursor
DorA	DMSO Reductase	*Rhodopseudomonas palustris* BisB18	Bacteria, Proteobacteria (alpha)	YP_531753	molybdopterin guanine dinucleotide-containing S/N-oxide reductase
FdhA	Formate Dehydrogenase Subunit Alpha	*Clostridium ljungdahlii* DSM 13528	Bacteria, Firmicutes, Clostridia	YP_003780168	formate dehydrogenase subunit alpha
FdhA	Formate Dehydrogenase Subunit Alpha	*Desulfovibrio piezophilus* C1TLV30	Bacteria, Proteobacteria (delta)	YP_007495372	Formate dehydrogenase subunit alpha
FdhA	Formate Dehydrogenase Subunit Alpha	*Methanocaldococcus jannaschii*	Archaea, Eyryaechaeota, Methanococci	P61159	RecName: Full=Formate dehydrogenase subunit alpha
FdhA	Formate Dehydrogenase Subunit Alpha	*Methanobacterium formicicum*	Archaea, Eyryaechaeota, Methanococci	P06131	RecName: Full=Formate dehydrogenase subunit alpha

FdhA	Formate Dehydrogenase Subunit Alpha	*Desulfovibrio gigas*	Bacteria, Proteobacteria (delta)	Q934F5	RecName: Full=Formate dehydrogenase subunit alpha; Short=FDH subunit alpha; AltName: Full=Formate dehydrogenase large subunit; Flags: Precursor
FdhA	Formate Dehydrogenase Subunit Alpha	*Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27774	Bacteria, Proteobacteria (delta)	ABE73760	formate dehydrogenase alpha subunit precursor
FdhF	Formate Dehydrogenase-H	*Escherichia coli*	Bacteria, Proteobacteria (gamma)	NP_418503	formate dehydrogenase-H, selenopolypeptide subunit
FdhN	Formate Dehydrogenase	*Pyrobaculum aerophilum* str. IM2	Archaea, Crenarchaeota, Thermoprotei	NP_560168	formate dehydrogenase alpha subunit
FdhN	Formate Dehydrogenase	*Anaeromyxobacter dehalogenans* 2CP-C	Bacteria, Proteobacteria (delta)	YP_466957	formate dehydrogenase alpha subunit
FdhN	Formate Dehydrogenase	*Ralstonia eutropha* H16	Bacteria, Proteobacteria (beta)	YP_840970	formate dehydrogenase alpha subunit
FdhN	Formate Dehydrogenase	*Wolinella succinogenes* DSM 1740	Bacteria, Proteobacteria (epsilon)	NP_907333	formate dehydrogenase large subunit precursor
FdhN	Formate Dehydrogenase	*Bacillus selenitireducens* MLS10	Bacteria, Firmicutes	YP_003699297	formate dehydrogenase subunit alpha
FdhN	Formate Dehydrogenase	*Aquifex aeolicus* VF5	Bacteria, Aquificales	NP_213709	formate dehydrogenase subunit alpha
FdhN	Formate Dehydrogenase	*Eggerthella lenta* DSM 2243	Bacteria, Actinobacteria	YP_003182037	molydopterin dinucleotide-binding protein
FdnG	NAD-Dependent Formate Dehydrogenase	*Hydrogenobacter thermophilus* TK-6	Bacteria, Aquificales	YP_005512343	formate dehydrogenase alpha subunit ;NAD-dependent formate dehydrogenase catalytic subunit
FdnG	NAD-Dependent Formate Dehydrogenase	*Escherichia coli* str. K-12 substr. MG1655	Bacteria, Proteobacteria (gamma)	NP_415991	formate dehydrogenase-N, alpha subunit, nitrate-inducible
FdnG	NAD-Dependent Formate Dehydrogenase	*Clostridium acidurici* 9a	Bacteria, Firmicutes	YP_006788521	NADP-dependent formate dehydrogenase FdhA

FdnG	NAD-Dependent Formate Dehydrogenase	*Enterococcus faecalis*	Bacteria, Firmicutes	WP_002384573	NAD-dependent formate dehydrogenase subunit alpha
FdnG	NAD-Dependent Formate Dehydrogenase	*Desulfosporosinus acidiphilus* SJ4	Bacteria, Firmicutes, Clostridia	YP_006465170	NAD-dependent formate dehydrogenase catalytic subunit
FdoG	Formate Dehydrogenase-O	*Oligotropha carboxidovorans* OM4	Bacteria, Proteobacteria (alpha)	YP_005951806	formate dehydrogenase-O major subunit FdoG
FdoG	Formate Dehydrogenase-O	*Escherichia coli*	Bacteria, Proteobacteria (gamma)	NP_418330	formate dehydrogenase-O, large subunit
NapA	Periplasmic Nitrate Reductase	*Bacillus azotoformans* LMG 9581	Bacteria, Firmicutes	EKN63995	periplasmic nitrate reductase
NapA	Periplasmic Nitrate Reductase	*Paracoccus pantotrophus*	Bacteria, Proteobacteria (alpha)	Q56350	RecName: Full=Periplasmic nitrate reductase; Flags: Precursor
NapA	Periplasmic Nitrate Reductase	*Shewanella oneidensis* MR-1	Bacteria, Proteobacteria (gamma)	NP_716479	periplasmic nitrate reductase molybdopterin-binding subunit NapA
NapA	Periplasmic Nitrate Reductase	*Desulfovibrio desulfuricans*	Bacteria, Proteobacteria (delta)	WP_012624252	periplasmic nitrate reductase
NarG	Membrane-Bound Nitrate Reductase	*Halomonas maura*	Bacteria, Proteobacteria (gamma)	AAT47523	NarG
NarG	Membrane-Bound Nitrate Reductase	*Herminiimonas arsenicoxydans*	Bacteria, Proteobacteria (beta)	YP_001099944	nitrate reductase 1 subunit alpha
NarG	Membrane-Bound Nitrate Reductase	*Pyrobaculum arsenaticum* DSM 13514	Archaea, Crenarchaeota, Thermoprotei	YP_001152746	nitrate reductase subunit alpha
NarG	Membrane-Bound Nitrate Reductase	*Jonesia denitrificans* DSM 20603	Bacteria, Actinobacteria, Actinobacteridae	YP_003161194	nitrate reductase subunit alpha
NarG	Membrane-Bound Nitrate Reductase	*Veillonella parvula* DSM 2008	Bacteria, Firmicutes, Negativicutes	YP_003311189	nitrate reductase subunit alpha

NarG	Membrane-Bound Nitrate Reductase	*Roseobacter denitrificans* OCh 114	Bacteria, Proteobacteria (alpha)	YP_682694	nitrate reductase subunit alpha
NarG	Membrane-Bound Nitrate Reductase	*Psychrobacter arcticus* 273-4	Bacteria, Proteobacteria (gamma)	YP_263899	respiratory nitrate reductase subunit alpha apoprotein
NarG	Membrane-Bound Nitrate Reductase	*Bacillus subtilis*	Bacteria, Firmicutes	NP_391609	nitrate reductase alpha chain
NarG	Membrane-Bound Nitrate Reductase	*Haloarcula marismortui*	Archaea, Eyryaechaeota, Halobacteria	Q9P9I5	Nitrate reductase subunit 1 (Nitrate reductase alpha chain)
NarG	Membrane-Bound Nitrate Reductase	*Thermus thermophilus*	Bacteria, Deinococcus-Thermus, Deinococci	WP_014630613	nitrate reductase
PcrA	Perchlorate Reductase	*Dechloromonas aromatica*	Bacteria, Proteobacteria (beta)	YP_285785	molybdopterin oxidoreductase:molydopterin dinucleotide-binding region
PhsA	Polysulfide/Thiosulfate Reductase	*Shewanella oneidensis* MR-1	Bacteria, Proteobacteria (gamma)	NP_719592	sulfur reductase molybdoperterin-binding subunit PhsA
PhsA	Polysulfide/Thiosulfate Reductase	*Wolinella succinogenes* DSM 1740	Bacteria, Proteobacteria (epsilon)	NP_906934	thiosulfate reductase
PhsA	Polysulfide/Thiosulfate Reductase	*Salmonella enterica*	Bacteria, Proteobacteria (gamma)	WP_000028233	thiosulfate reductase
PsrA	Polysulfide/Thiosulfate Reductase	*Geobacter lovleyi* SZ	Bacteria, Proteobacteria (delta)	YP_001951406	formate dehydrogenase
PsrA	Polysulfide/Thiosulfate Reductase	*Moorella thermoacetica* ATCC 39073	Bacteria, Firmicutes, Clostridia	YP_429324	formate dehydrogenase
PsrA	Polysulfide/Thiosulfate Reductase	*Alkalilimnicola ehrlichii* MLHE-1	Bacteria, Proteobacteria (gamma)	YP_741477	molybdopterin oxidoreductase
PsrA	Polysulfide/Thiosulfate Reductase	*Archaeoglobus fulgidus* DSM 4304	Archaea, Eyryaechaeota, Archaeoglobi	NP_071207	molybdopterin oxidoreductase, molybdopterin binding subunit

PsrA	Polysulfide/Thiosulfate Reductase	*Pyrobaculum aerophilum* str. IM2	Archaea, Crenarchaeota, Thermoprotei	NP_560307	molybdopterin oxidoreductase, molybdopterin binding subunit
PsrA	Polysulfide/Thiosulfate Reductase	*Acidiphilium cryptum* JF-5	Bacteria, Proteobacteria (alpha)	YP_001233491	nitrate reductase
PsrA	Polysulfide/Thiosulfate Reductase	*Sulfurospirillum barnesii*	Bacteria, Proteobacteria (epsilon)	WP_014768488	polysulfide reductase
PsrA	Polysulfide/Thiosulfate Reductase	*Campylobacter showae*	Bacteria, Proteobacteria (epsilon)	WP_009493170	polysulfide reductase, subunit A
PsrA	Polysulfide/Thiosulfate Reductase	*Thermus thermophilus* HB27	Bacteria, Deinococcus-Thermus, Deinococci	YP_004130	thiosulfate reductase precursor
PsrA	Polysulfide/Thiosulfate Reductase	delta proteobacterium MLMS-1	Bacteria, Proteobacteria (delta)	WP_007292283	twin-arginine translocation pathway signal protein
SerA	Selenate Reductase	*Thauera selenatis*	Bacteria, Proteobacteria (beta)	Q9S1H0	RecName: Full=Selenate reductase subunit alpha; AltName: Full=Selenate reductase molybdenum subunit; Flags: Precursor
SrdA	Selenate Reductase	*Bacillus selenatarsenatis* SF-1	Bacteria, Firmicutes	BAJ83594	putative selenate reductase subunit A
SreA	Sulfur Reductase	*Aquifex aeolicus*	Bacteria, Aquificales	WP_010880782	formate dehydrogenase
SreA	Sulfur Reductase	*Acidianus ambivalens*	Archaea, Crenarchaeota, Thermoprotei	CAC86937	sulfur reductase molybdopterin subunit
SreA	Sulfur Reductase	*Sulfolobus solfataricus*	Archaea, Crenarchaeota, Thermoprotei	WP_010923481	molybdopterin oxidoreductase
TorA	TMAO Reductase	*Sulfurospirillum deleyianum* DSM 6946	Bacteria, Proteobacteria (epsilon)	YP_003303433	molybdopterin oxidoreductase
TorA	TMAO Reductase	*Halomonas jeotgali*	Bacteria, Proteobacteria (gamma)	WP_026106068	trimethylamine N-oxide reductase I catalytic subunit

TorA	TMAO Reductase	*Shewanella massilia*	Bacteria, Proteobacteria (gamma)	O87948	RecName: Full=Trimethylamine-N-oxide reductase; Short=TMAO reductase; Short=Trimethylamine oxidase; Flags: Precursor
TorA	TMAO Reductase	*Burkholderia oklahomensis*	Bacteria, Proteobacteria (beta)	WP_010111741	trimethylamine N-oxide reductase I catalytic subunit
TorA	TMAO Reductase	*Escherichia coli*	Bacteria, Proteobacteria (gamma)	NP_415517	trimethylamine N-oxide (TMAO) reductase I, catalytic subunit
TorZ	TMAO Reductase (System III)	*Escherichia coli*	Bacteria, Proteobacteria (gamma)	YP_490134	trimethylamine N-oxide reductase system III, catalytic subunit
TtrA	Tetrathionate Reductase	*Wolinella succinogenes* DSM 1740	Bacteria, Proteobacteria (epsilon)	NP_907142	TtrA
TtrA	Tetrathionate Reductase	*Thiobacillus denitrificans* ATCC 25259	Bacteria, Proteobacteria (beta)	YP_315499	tetrathionate reductase subunit A
TtrA	Tetrathionate Reductase	*Shewanella* sp. ANA-3	Bacteria, Proteobacteria (gamma)	YP_868277	molydopterin dinucleotide-binding region
XdhA	Xanthine Dehydrogenase	*Paenibacillus mucilaginosus* K02	Bacteria, Firmicutes	YP_006190319	xanthine dehydrogenase

Table 3.S2. Sulfur cycle custom database accession list

Protein	Accession (NCBI/UniRef)
AprA	AAV31643
AprA	ABV80013
AprA	ABV80029
AprA	ABV80041
AprA	ABV80043
AprA	ABV80054
AprA	ABV80093
AprA	ABV80095
AprA	ABV80103
AprA	UniRef90_A0A011A0F3
AprA	UniRef90_A0A069RM03
AprA	UniRef90_A0A075WE48
AprA	UniRef90_A0LH40
AprA	UniRef90_A4J274
AprA	UniRef90_A4WK85
AprA	UniRef90_A5CXS9
AprA	UniRef90_A8MAT5
AprA	UniRef90_A8ZWK0
AprA	UniRef90_B1I5T0
AprA	UniRef90_B4SAM7
AprA	UniRef90_B5YHN7
AprA	UniRef90_B6WW23
AprA	UniRef90_B8FAH1
AprA	UniRef90_B8J3Y0
AprA	UniRef90_B8PS61
AprA	UniRef90_C0QHK8
AprA	UniRef90_C2CYH7
AprA	UniRef90_C4XI69
AprA	UniRef90_C6PP03
AprA	UniRef90_C7LXN7
AprA	UniRef90_C8VW02
AprA	UniRef90_C8X4M7
AprA	UniRef90_C9KJT0
AprA	UniRef90_C9RD80
AprA	UniRef90_D2RDW9
AprA	UniRef90_D3SC63
AprA	UniRef90_D5CSL2
AprA	UniRef90_D6SNE3
AprA	UniRef90_D6Z3P1

Protein	Accession (NCBI/UniRef)
AprA	UniRef90_D8F004
AprA	UniRef90_E1QFL6
AprA	UniRef90_E3HDY5
AprA	UniRef90_E8RDT8
AprA	UniRef90_F0QUP3
AprA	UniRef90_F2KPU8
AprA	UniRef90_F2L4M7
AprA	UniRef90_F2NF60
AprA	UniRef90_F3YU89
AprA	UniRef90_F6CPR9
AprA	UniRef90_F8AE53
AprA	UniRef90_F8C2J6
AprA	UniRef90_G2E1B5
AprA	UniRef90_G2FXB9
AprA	UniRef90_G7W5Z9
AprA	UniRef90_I3BVG3
AprA	UniRef90_I3YDV4
AprA	UniRef90_I4C3L4
AprA	UniRef90_J0WPC1
AprA	UniRef90_J7IY83
AprA	UniRef90_K0B3R1
AprA	UniRef90_L0R9F9
AprA	UniRef90_L7EH99
AprA	UniRef90_M1E8J8
AprA	UniRef90_M1P518
AprA	UniRef90_M1WQM0
AprA	UniRef90_N0BEM0
AprA	UniRef90_Q313I5
AprA	UniRef90_Q3A8Q8
AprA	UniRef90_Q3SGL5
AprA	UniRef90_Q3SKF6
AprA	UniRef90_Q4FMD5
AprA	UniRef90_Q5UEX5
AprA	UniRef90_Q97MT7
AprA	UniRef90_R5EAZ1
AprA	UniRef90_R5GBE3
AprA	UniRef90_R6BBC6
AprA	UniRef90_R6LVL7
AprA	UniRef90_R6VLM0
AprA	UniRef90_R7CPD7
AprA	UniRef90_S6BVC5
AprA	UniRef90_S7T3E4

AprA	UniRef90_S7UU91
AprA	UniRef90_T2G6Z9
AprA	UniRef90_U2NIQ2
AprA	UniRef90_V4L4W5
AprA	UniRef90_V4LK44
AprA	UniRef90_V9U1P3
AprA	UniRef90_W0SI19
AprA	UniRef90_W6NLB4
AprA	UniRef90_W9V999
AprA	UniRef90_X0PN84
AprB	ABV80004
AprB	ABV80012
AprB	ABV80018
AprB	ABV80020
AprB	ABV80028
AprB	ABV80034
AprB	ABV80036
AprB	ABV80040
AprB	ABV80044
AprB	ABV80051
AprB	ABV80053
AprB	ABV80055
AprB	ABV80077
AprB	ABV80079
AprB	ABV80085
AprB	ABV80089
AprB	ABV80094
AprB	ABV80096
AprB	ABV80100
AprB	ABV80105
AprB	UniRef90_A0A075HJU7
AprB	UniRef90_A0A075WHM3
AprB	UniRef90_A0LH39
AprB	UniRef90_A3MW32
AprB	UniRef90_A4J273
AprB	UniRef90_A4WK84
AprB	UniRef90_A5CXS8
AprB	UniRef90_A8MAT4
AprB	UniRef90_A8ZWJ9
AprB	UniRef90_B1I5T1
AprB	UniRef90_B4SAM8
AprB	UniRef90_B5YHN8
AprB	UniRef90_B6WW24

AprB	UniRef90_B8FAH2
AprB	UniRef90_B8GUT8
AprB	UniRef90_B8J3Y1
AprB	UniRef90_C4XI68
AprB	UniRef90_C7LXN6
AprB	UniRef90_C8VW03
AprB	UniRef90_C8X4M8
AprB	UniRef90_C9RD81
AprB	UniRef90_D2RDX0
AprB	UniRef90_D3RSA0
AprB	UniRef90_D5CSL1
AprB	UniRef90_D6SNE2
AprB	UniRef90_D6Z3P2
AprB	UniRef90_D8F003
AprB	UniRef90_E1QFL5
AprB	UniRef90_E3ISN5
AprB	UniRef90_E8RDT7
AprB	UniRef90_F0JIU7
AprB	UniRef90_F0QUP4
AprB	UniRef90_F2KPU7
AprB	UniRef90_F2L4M6
AprB	UniRef90_F2NF61
AprB	UniRef90_F3YU90
AprB	UniRef90_F6CPR8
AprB	UniRef90_F8A9X1
AprB	UniRef90_F8C2J7
AprB	UniRef90_F9UHZ8
AprB	UniRef90_G2E1B6
AprB	UniRef90_G2FE27
AprB	UniRef90_G4RNG3
AprB	UniRef90_G7VH35
AprB	UniRef90_G7W5Z8
AprB	UniRef90_I3BVG2
AprB	UniRef90_I3YDV5
AprB	UniRef90_I4C3L3
AprB	UniRef90_I4D642
AprB	UniRef90_J9YWB1
AprB	UniRef90_K0NCP7
AprB	UniRef90_L0RB73
AprB	UniRef90_M1E9F9
AprB	UniRef90_M1WM61
AprB	UniRef90_N0BP05
AprB	UniRef90_Q1NTY2

AprB	UniRef90_Q313I6
AprB	UniRef90_Q3SGL4
AprB	UniRef90_Q3SKF5
AprB	UniRef90_Q5UEX6
AprB	UniRef90_S6AHB8
AprB	UniRef90_S7UAR1
AprB	UniRef90_S7UPH6
AprB	UniRef90_S7VIS5
AprB	UniRef90_T1B7T0
AprB	UniRef90_T1C8T8
AprB	UniRef90_T2G899
AprB	UniRef90_W0DJW5
AprB	UniRef90_W0SIV0
AprB	UniRef90_W6KDK5
AprB	UniRef90_W9W058
AprB	UniRef90_X0QIE0
DsrA	ABX82416
DsrA	ABX82429
DsrA	ABX82437
DsrA	ABX82440
DsrA	BAJ17549
DsrA	UniRef90_A0A060BKJ3
DsrA	UniRef90_A0A068JI07
DsrA	UniRef90_A0A075WSN6
DsrA	UniRef90_A0L9L4
DsrA	UniRef90_A0LQK6
DsrA	UniRef90_A1BCS1
DsrA	UniRef90_A1HQ02
DsrA	UniRef90_A1WYF3
DsrA	UniRef90_A3MW49
DsrA	UniRef90_A3MW60
DsrA	UniRef90_A4J9D6
DsrA	UniRef90_A4WJC1
DsrA	UniRef90_A4WK64
DsrA	UniRef90_A5CVW6
DsrA	UniRef90_A8ZX93
DsrA	UniRef90_B1I6P0
DsrA	UniRef90_B1YBR0
DsrA	UniRef90_B3EHJ7
DsrA	UniRef90_B3ELT3
DsrA	UniRef90_B3QRF1
DsrA	UniRef90_B4S963
DsrA	UniRef90_B4SE97

DsrA	UniRef90_B5YIF2
DsrA	UniRef90_B8FDV1
DsrA	UniRef90_B8FME3
DsrA	UniRef90_B8GUE5
DsrA	UniRef90_B8IZA1
DsrA	UniRef90_B8PS63
DsrA	UniRef90_C0QD64
DsrA	UniRef90_C4XH83
DsrA	UniRef90_C7LUX2
DsrA	UniRef90_C7N3T9
DsrA	UniRef90_C8W1Y8
DsrA	UniRef90_C8WZS1
DsrA	UniRef90_C9RA42
DsrA	UniRef90_D2RFY6
DsrA	UniRef90_D5CSI3
DsrA	UniRef90_D6E7E1
DsrA	UniRef90_D6SSJ8
DsrA	UniRef90_D6Z6P8
DsrA	UniRef90_D8F2V3
DsrA	UniRef90_E1QFJ4
DsrA	UniRef90_E1QPH9
DsrA	UniRef90_E1YFB2
DsrA	UniRef90_E3HZD3
DsrA	UniRef90_E5Y7F0
DsrA	UniRef90_E6QSS7
DsrA	UniRef90_E8RE36
DsrA	UniRef90_F0JJB3
DsrA	UniRef90_F2KRS8
DsrA	UniRef90_F2NJG0
DsrA	UniRef90_F3YZL0
DsrA	UniRef90_F6B661
DsrA	UniRef90_F6CRI3
DsrA	UniRef90_F6DQV7
DsrA	UniRef90_F7NEE2
DsrA	UniRef90_F8AAL5
DsrA	UniRef90_F8C2Z7
DsrA	UniRef90_G2DAV7
DsrA	UniRef90_G2FW57
DsrA	UniRef90_G2H7S1
DsrA	UniRef90_G4RNF1
DsrA	UniRef90_G4RNF2
DsrA	UniRef90_G7W9K1
DsrA	UniRef90_G7WBP5

DsrA	UniRef90_H5WQR1
DsrA	UniRef90_H6Q9Y4
DsrA	UniRef90_I1X4G7
DsrA	UniRef90_I1X4J3
DsrA	UniRef90_I1X4R9
DsrA	UniRef90_I1X4V2
DsrA	UniRef90_I1X507
DsrA	UniRef90_I1X533
DsrA	UniRef90_I1X5A3
DsrA	UniRef90_I2PX20
DsrA	UniRef90_I3BY25
DsrA	UniRef90_I3XV44
DsrA	UniRef90_I4C0R2
DsrA	UniRef90_I4C3X6
DsrA	UniRef90_I4D5U6
DsrA	UniRef90_I4D7Z0
DsrA	UniRef90_J7J276
DsrA	UniRef90_K0ND70
DsrA	UniRef90_K2QXR7
DsrA	UniRef90_L0DW28
DsrA	UniRef90_L0F545
DsrA	UniRef90_L0GQK7
DsrA	UniRef90_L0R6S6
DsrA	UniRef90_M1E9A0
DsrA	UniRef90_M1PP41
DsrA	UniRef90_M1WMP4
DsrA	UniRef90_N0BK96
DsrA	UniRef90_O33998
DsrA	UniRef90_P45574
DsrA	UniRef90_Q0A838
DsrA	UniRef90_Q251E4
DsrA	UniRef90_Q2RI32
DsrA	UniRef90_Q2W1V4
DsrA	UniRef90_Q315R9
DsrA	UniRef90_Q3A9H7
DsrA	UniRef90_Q3AP71
DsrA	UniRef90_Q3B6V5
DsrA	UniRef90_Q3IBI2
DsrA	UniRef90_Q3SG18
DsrA	UniRef90_Q3SJ48
DsrA	UniRef90_Q3SJA5
DsrA	UniRef90_Q59109
DsrA	UniRef90_Q6AQ47

DsrA	UniRef90_Q93TS8
DsrA	UniRef90_Q9ACL3
DsrA	UniRef90_R4KMH8
DsrA	UniRef90_S6AAH8
DsrA	UniRef90_S7TLS8
DsrA	UniRef90_S7TVX7
DsrA	UniRef90_S7UKR1
DsrA	UniRef90_V4JJE7
DsrA	UniRef90_V4L2I2
DsrA	UniRef90_V6F6X7
DsrA	UniRef90_W0E527
DsrA	UniRef90_W0SHF1
DsrA	UniRef90_W6KK51
DsrA	ZP_00053120
DsrB	ABX82417
DsrB	ABX82438
DsrB	ABX82439
DsrB	BAJ17550
DsrB	BAJ17556
DsrB	UniRef90_A0L9L5
DsrB	UniRef90_A0LQK7
DsrB	UniRef90_A1AXC8
DsrB	UniRef90_A1HQ01
DsrB	UniRef90_A1RQX0
DsrB	UniRef90_A1WYF4
DsrB	UniRef90_A3MW48
DsrB	UniRef90_A3MW59
DsrB	UniRef90_A4J9D5
DsrB	UniRef90_A4SC59
DsrB	UniRef90_A4WJC2
DsrB	UniRef90_A4WK63
DsrB	UniRef90_A4WK74
DsrB	UniRef90_A8MD33
DsrB	UniRef90_A8ZX94
DsrB	UniRef90_B1I6N9
DsrB	UniRef90_B3ELT2
DsrB	UniRef90_B3QRG4
DsrB	UniRef90_B4SE96
DsrB	UniRef90_B5YIF3
DsrB	UniRef90_B8FME2
DsrB	UniRef90_B8GUE6
DsrB	UniRef90_B8IZA0
DsrB	UniRef90_C4XH84

DsrB	UniRef90_C7LUX3
DsrB	UniRef90_C8W1Y9
DsrB	UniRef90_C8WZS2
DsrB	UniRef90_C9RA41
DsrB	UniRef90_D2RFY7
DsrB	UniRef90_D3RSN2
DsrB	UniRef90_D5CSI2
DsrB	UniRef90_D6E8X4
DsrB	UniRef90_D6SSJ7
DsrB	UniRef90_D6Z6P9
DsrB	UniRef90_D8F2V2
DsrB	UniRef90_E1QFJ3
DsrB	UniRef90_E1YFB1
DsrB	UniRef90_E3HZD2
DsrB	UniRef90_E5Y7F1
DsrB	UniRef90_E6QSS8
DsrB	UniRef90_E8RE37
DsrB	UniRef90_F0JJB4
DsrB	UniRef90_F0QUQ2
DsrB	UniRef90_F2KRS7
DsrB	UniRef90_F2NJF9
DsrB	UniRef90_F3YZL1
DsrB	UniRef90_F3Z1K8
DsrB	UniRef90_F6CRI2
DsrB	UniRef90_F7NEE3
DsrB	UniRef90_F8AAL4
DsrB	UniRef90_F8C2Z6
DsrB	UniRef90_G1UUT2
DsrB	UniRef90_G2FHM6
DsrB	UniRef90_G2FW56
DsrB	UniRef90_G4RNF0
DsrB	UniRef90_G4RNF3
DsrB	UniRef90_G7WBP4
DsrB	UniRef90_H5WQR0
DsrB	UniRef90_I1X4G8
DsrB	UniRef90_I1X4J2
DsrB	UniRef90_I1X4S0
DsrB	UniRef90_I1X4V1
DsrB	UniRef90_I1X508
DsrB	UniRef90_I1X574
DsrB	UniRef90_I1X5A4
DsrB	UniRef90_I3BY26
DsrB	UniRef90_I3XV44

DsrB	UniRef90_I4C3X5
DsrB	UniRef90_I4D7Y9
DsrB	UniRef90_I5B1F1
DsrB	UniRef90_K2QXR7
DsrB	UniRef90_K6UH91
DsrB	UniRef90_L0F3N5
DsrB	UniRef90_L0R867
DsrB	UniRef90_M1E861
DsrB	UniRef90_M1P948
DsrB	UniRef90_M1WKM5
DsrB	UniRef90_N0BL33
DsrB	UniRef90_P45575
DsrB	UniRef90_Q0A837
DsrB	UniRef90_Q2RI08
DsrB	UniRef90_Q2RI33
DsrB	UniRef90_Q2W1V3
DsrB	UniRef90_Q315R8
DsrB	UniRef90_Q3AP72
DsrB	UniRef90_Q3SG19
DsrB	UniRef90_Q59110
DsrB	UniRef90_Q6AQ46
DsrB	UniRef90_R4KL11
DsrB	UniRef90_R6GEJ7
DsrB	UniRef90_R6THS5
DsrB	UniRef90_S7UGF5
DsrB	UniRef90_S7V9M2
DsrB	UniRef90_T2G8V8
DsrB	UniRef90_V6F8U6
DsrB	UniRef90_W0SE78
DsrB	UniRef90_W6KAS1
sat	AAC23622
sat	NP_661756
sat	YP_314632
sat	YP_379885
sat	YP_903355
SorA	AAK58572
SorA	AF154565
SorA	CAL62480
SorA	NP_773897
SorA	SOR_ACIAM
SorA	UniRef90_A0A010QBH0
SorA	UniRef90_A0A031JPR6
SorA	UniRef90_A0A072SZV5

SorA	UniRef90_A1VQR6
SorA	UniRef90_A1W216
SorA	UniRef90_A1WIA1
SorA	UniRef90_A4SYY7
SorA	UniRef90_A5FZZ7
SorA	UniRef90_A5V4V0
SorA	UniRef90_B6BGQ1
SorA	UniRef90_D7BCC9
SorA	UniRef90_D8PJI2
SorA	UniRef90_E0US48
SorA	UniRef90_E1SSH7
SorA	UniRef90_E3HRZ2
SorA	UniRef90_E3HWB7
SorA	UniRef90_E4TZK0
SorA	UniRef90_E4U0G5
SorA	UniRef90_E6X486
SorA	UniRef90_H0Q3T6
SorA	UniRef90_I2K7N2
SorA	UniRef90_I3UAU6
SorA	UniRef90_I3UE79
SorA	UniRef90_K0I4F5
SorA	UniRef90_L0AMI1
SorA	UniRef90_L9W9I7
SorA	UniRef90_M0L922
SorA	UniRef90_M0LKM9
SorA	UniRef90_M1XT02
SorA	UniRef90_Q01QB1
SorA	UniRef90_Q128W0
SorA	UniRef90_Q15RV6
SorA	UniRef90_Q1LDM3
SorA	UniRef90_Q1QGB1
SorA	UniRef90_Q1QI23
SorA	UniRef90_Q3SN69
SorA	UniRef90_Q46WP0
SorA	UniRef90_Q7P1Q5
SorA	UniRef90_Q8U195
SorA	UniRef90_Q8Y177
SorA	UniRef90_R4XMR5
SorA	UniRef90_T0I2I8
SorA	UniRef90_V5UEC2
SorA	YP_004737743
SorA	YP_982913
SorA	ZP_00051098

SorB	DOXA_ACIAM
SorB	DOXD_ACIAM
SorB	NP_773898
SorB	UniRef90_A0A010PCW4
SorB	UniRef90_A0A076FWK2
SorB	UniRef90_A0A076PLH1
SorB	UniRef90_A1VQR5
SorB	UniRef90_A1W215
SorB	UniRef90_A1WIA2
SorB	UniRef90_A4G7J2
SorB	UniRef90_A4SYY6
SorB	UniRef90_A5FZZ8
SorB	UniRef90_B3R8H9
SorB	UniRef90_C5T8U3
SorB	UniRef90_D3DHG5
SorB	UniRef90_D7BCD0
SorB	UniRef90_D8N691
SorB	UniRef90_D8NMH5
SorB	UniRef90_D8NWU0
SorB	UniRef90_E0US49
SorB	UniRef90_E4U0G4
SorB	UniRef90_E6PMK2
SorB	UniRef90_E6PUC6
SorB	UniRef90_F8GN54
SorB	UniRef90_G2IUB0
SorB	UniRef90_H0Q3T5
SorB	UniRef90_H0S2N5
SorB	UniRef90_H1FYW1
SorB	UniRef90_I3UE78
SorB	UniRef90_M5J3H2
SorB	UniRef90_Q128W1
SorB	UniRef90_Q15RV7
SorB	UniRef90_Q1LDM2
SorB	UniRef90_Q1QI22
SorB	UniRef90_Q3SN68
SorB	UniRef90_Q46WP1
SorB	UniRef90_Q7P1Q4
SorB	UniRef90_Q8EIW3
SorB	UniRef90_Q9LA15
SorB	UniRef90_U7NKB4
SorB	UniRef90_V8QTJ2
SorB	ZP_00051120
SorB	ZP_01044877

SoxA	CAL61370
SoxA	NP_214239
SoxA	NP_767651
SoxA	NP_769372
SoxA	NP_770154
SoxA	NP_949804
SoxA	NP_949805
SoxA	UniRef90_A0A067A3E1
SoxA	UniRef90_A0A069E0F4
SoxA	UniRef90_A0A069I3Y9
SoxA	UniRef90_A0A073A2I4
SoxA	UniRef90_A0L1A9
SoxA	UniRef90_A1SUH0
SoxA	UniRef90_B2UG47
SoxA	UniRef90_B8GN87
SoxA	UniRef90_B8GUQ0
SoxA	UniRef90_C1DX05
SoxA	UniRef90_D0J0E4
SoxA	UniRef90_D3DJG4
SoxA	UniRef90_D3DJG5
SoxA	UniRef90_D6CVG6
SoxA	UniRef90_D7A6E5
SoxA	UniRef90_E3T355
SoxA	UniRef90_E6PWC2
SoxA	UniRef90_F7QKL8
SoxA	UniRef90_G2DG34
SoxA	UniRef90_L0E1H3
SoxA	UniRef90_O33434
SoxA	UniRef90_Q08IS0
SoxA	UniRef90_Q0BZC2
SoxA	UniRef90_Q0K5U7
SoxA	UniRef90_Q1LHT4
SoxA	UniRef90_Q1W3E4
SoxA	UniRef90_Q8KDM7
SoxA	UniRef90_Q939U1
SoxA	UniRef90_Q9K4M4
SoxA	UniRef90_R7X0C6
SoxA	UniRef90_UPI0002624BF0
SoxA	UniRef90_UPI0002625345
SoxA	UniRef90_UPI0002E9D2AC
SoxA	UniRef90_UPI0003664F8C
SoxA	UniRef90_UPI00036B1448
SoxA	UniRef90_UPI0003763A9D

SoxA	UniRef90_UPI000378A4EA
SoxA	UniRef90_UPI00039DD43E
SoxA	UniRef90_UPI0003A123E7
SoxA	UniRef90_UPI0003AA4674
SoxA	UniRef90_UPI0003B67AD7
SoxA	UniRef90_UPI0003F723CF
SoxA	UniRef90_UPI0003F8947D
SoxA	UniRef90_UPI0003FB4C92
SoxA	UniRef90_UPI0003FBB608
SoxA	UniRef90_UPI000404EC9D
SoxA	UniRef90_UPI000426A502
SoxA	UniRef90_UPI000426C4E8
SoxA	UniRef90_UPI00045E7ED0
SoxA	UniRef90_UPI0004781C04
SoxA	UniRef90_UPI00047A67D1
SoxA	UniRef90_UPI000481FA01
SoxA	UniRef90_UPI0004887A23
SoxA	UniRef90_UPI00048C8096
SoxA	UniRef90_UPI00048D508B
SoxA	UniRef90_UPI00048FD07C
SoxA	UniRef90_UPI000490A5AB
SoxA	UniRef90_UPI000493E076
SoxA	UniRef90_UPI000493E69C
SoxA	UniRef90_UPI0004942F6A
SoxA	UniRef90_UPI000494E41C
SoxA	UniRef90_UPI000497B2F6
SoxA	UniRef90_UPI0004A71C50
SoxA	UniRef90_UPI0004A74927
SoxA	UniRef90_UPI0004A76731
SoxA	UniRef90_UPI0004A77426
SoxA	UniRef90_V5UEB6
SoxA	UniRef90_V8QZI7
SoxA	UniRef90_W0SI65
SoxA	YP_001003514
SoxA	YP_001021624
SoxA	YP_005023
SoxA	YP_144681
SoxA	YP_166248
SoxA	YP_286330
SoxA	YP_314322
SoxA	YP_314676
SoxA	YP_380216
SoxA	YP_390871

SoxA	YP_392779
SoxA	YP_465488
SoxA	YP_487970
SoxA	YP_487971
SoxA	YP_549441
SoxA	YP_571374
SoxA	YP_571375
SoxA	YP_578861
SoxA	YP_681833
SoxA	YP_865820
SoxA	YP_903997
SoxA	ZP_00588640
SoxA	ZP_00628737
SoxA	ZP_00912864
SoxA	ZP_00956138
SoxA	ZP_00961295
SoxA	ZP_00963530
SoxA	ZP_01014858
SoxA	ZP_01037117
SoxA	ZP_01055913
SoxA	ZP_01102558
SoxA	ZP_01144907
SoxA	ZP_01167150
SoxA	ZP_01225766
SoxA	ZP_01439477
SoxA	ZP_01549054
SoxA	ZP_01583268
SoxA	ZP_01627097
SoxA	ZP_01743246
SoxA	ZP_01748360
SoxB	AAF99435
SoxB	AAL68888
SoxB	ABR67341
SoxB	ABR67342
SoxB	ABR67343
SoxB	ABR67346
SoxB	ABR67348
SoxB	ABR67350
SoxB	ABR67351
SoxB	ABR67353
SoxB	ABR67355
SoxB	ABR67356
SoxB	ABR67358

SoxB	ABR67362
SoxB	ABR67363
SoxB	ABR67364
SoxB	ABR67365
SoxB	ABR67367
SoxB	ABR67368
SoxB	ABR67369
SoxB	ABR67371
SoxB	ABR67372
SoxB	ABR67373
SoxB	ABR67374
SoxB	ABR67375
SoxB	ABR67376
SoxB	ABR67378
SoxB	ABR67379
SoxB	ABR67381
SoxB	ABR67382
SoxB	ABR67383
SoxB	ABR67385
SoxB	ABR67386
SoxB	BAF34125
SoxB	NP_767649
SoxB	UniRef90_A0A024IG50
SoxB	UniRef90_A0A059ZWT2
SoxB	UniRef90_A0A060A1Q5
SoxB	UniRef90_A0A068X7F5
SoxB	UniRef90_A0A073A2I0
SoxB	UniRef90_A0A076JZ94
SoxB	UniRef90_A0A085F516
SoxB	UniRef90_A0L8X0
SoxB	UniRef90_A1AVJ1
SoxB	UniRef90_A1B9M4
SoxB	UniRef90_A1WYE1
SoxB	UniRef90_A2SIJ8
SoxB	UniRef90_A4A7K1
SoxB	UniRef90_A4SCF8
SoxB	UniRef90_A4U035
SoxB	UniRef90_A4WZF4
SoxB	UniRef90_A5CXN3
SoxB	UniRef90_A5E9W5
SoxB	UniRef90_A5G305
SoxB	UniRef90_A6Q623
SoxB	UniRef90_A6Q7K2

SoxB	UniRef90_A6SV46
SoxB	UniRef90_A7BVE9
SoxB	UniRef90_A8ESB5
SoxB	UniRef90_A8UQT9
SoxB	UniRef90_B1XV30
SoxB	UniRef90_B2UG44
SoxB	UniRef90_B3Q926
SoxB	UniRef90_B8J5R8
SoxB	UniRef90_D0J0E2
SoxB	UniRef90_D1KBD0
SoxB	UniRef90_D3PPJ4
SoxB	UniRef90_D3RVS5
SoxB	UniRef90_D5X456
SoxB	UniRef90_D6V8B3
SoxB	UniRef90_D7A6E2
SoxB	UniRef90_E0US21
SoxB	UniRef90_E3T345
SoxB	UniRef90_E3T356
SoxB	UniRef90_E4TZM8
SoxB	UniRef90_E6PWC4
SoxB	UniRef90_E6WZY0
SoxB	UniRef90_E8PNQ6
SoxB	UniRef90_F3L3A7
SoxB	UniRef90_F7QKM2
SoxB	UniRef90_F7ZJW5
SoxB	UniRef90_G0EZG6
SoxB	UniRef90_G0JQ03
SoxB	UniRef90_G2DEK7
SoxB	UniRef90_G4SUB0
SoxB	UniRef90_G8NAE9
SoxB	UniRef90_H9ZQC5
SoxB	UniRef90_I1X4F3
SoxB	UniRef90_I1X5E6
SoxB	UniRef90_I1X5K7
SoxB	UniRef90_I2K4N6
SoxB	UniRef90_I3BTH9
SoxB	UniRef90_I3UCZ6
SoxB	UniRef90_I7EL48
SoxB	UniRef90_I9W4T4
SoxB	UniRef90_J6LAS2
SoxB	UniRef90_J7FQ58
SoxB	UniRef90_K7R6V1
SoxB	UniRef90_K9H957

SoxB	UniRef90_L0E2P8
SoxB	UniRef90_L0NJ68
SoxB	UniRef90_L9PLX3
SoxB	UniRef90_M4Z9F8
SoxB	UniRef90_M5RIF5
SoxB	UniRef90_N6YCZ3
SoxB	UniRef90_O67671
SoxB	UniRef90_Q07M24
SoxB	UniRef90_Q12AA1
SoxB	UniRef90_Q1LHT7
SoxB	UniRef90_Q1QH98
SoxB	UniRef90_Q2BJY1
SoxB	UniRef90_Q2IRV4
SoxB	UniRef90_Q30TY6
SoxB	UniRef90_Q31FD2
SoxB	UniRef90_Q3APA0
SoxB	UniRef90_Q47BC3
SoxB	UniRef90_Q89PG5
SoxB	UniRef90_Q8KDM5
SoxB	UniRef90_Q93RP6
SoxB	UniRef90_S6AIY2
SoxB	UniRef90_S6BI80
SoxB	UniRef90_S9QBQ7
SoxB	UniRef90_S9QZ03
SoxB	UniRef90_S9SH65
SoxB	UniRef90_T0JU50
SoxB	UniRef90_U3AIF0
SoxB	UniRef90_UPI00035E960C
SoxB	UniRef90_UPI000367CAFC
SoxB	UniRef90_UPI0003B4636C
SoxB	UniRef90_UPI000484B283
SoxB	UniRef90_UPI0004970BC9
SoxB	UniRef90_UPI0004975CA2
SoxB	UniRef90_V4RMK9
SoxB	UniRef90_V5C8X2
SoxB	UniRef90_V9WLD2
SoxB	UniRef90_W0DC07
SoxB	UniRef90_W0TR16
SoxB	UniRef90_W6KJI5
SoxB	UniRef90_W8RTR7
SoxB	UniRef90_W9VAG2
SoxB	YP_166249
SoxB	YP_314321

SoxB	ZP_00051299
SoxB	ZP_00944480
SoxB	ZP_00956139
SoxB	ZP_00961294
SoxB	ZP_00963529
SoxB	ZP_01014859
SoxB	ZP_01037116
SoxB	ZP_01225765
SoxB	ZP_01439476
SoxB	ZP_01549055
SoxB	ZP_01627095
SoxB	ZP_01743245
SoxB	ZP_01748359
SoxC	AAF99436
SoxC	BAF34119
SoxC	UniRef90_A0A031GIF1
SoxC	UniRef90_A0A031GJ34
SoxC	UniRef90_A0A068TK00
SoxC	UniRef90_A0A072CFS9
SoxC	UniRef90_A0A076K0U9
SoxC	UniRef90_A0A085F517
SoxC	UniRef90_A1B9M5
SoxC	UniRef90_A2SHK3
SoxC	UniRef90_A4SX05
SoxC	UniRef90_A4WZF3
SoxC	UniRef90_A5EHW2
SoxC	UniRef90_A5EL64
SoxC	UniRef90_A5G2Z8
SoxC	UniRef90_A6Q6A0
SoxC	UniRef90_A6SV39
SoxC	UniRef90_A8ESA8
SoxC	UniRef90_B1A9Y7
SoxC	UniRef90_B3Q925
SoxC	UniRef90_B4ZYT0
SoxC	UniRef90_B6AX77
SoxC	UniRef90_B6BGQ8
SoxC	UniRef90_B6DXB1
SoxC	UniRef90_B8KX43
SoxC	UniRef90_B9VWK7
SoxC	UniRef90_C3KM89
SoxC	UniRef90_C7CBQ1
SoxC	UniRef90_D4GGW8
SoxC	UniRef90_D5X462

SoxC	UniRef90_D7A6E1
SoxC	UniRef90_E0URX4
SoxC	UniRef90_E0XCR7
SoxC	UniRef90_E6PWB8
SoxC	UniRef90_E6WZX3
SoxC	UniRef90_E8PNQ0
SoxC	UniRef90_F5RC00
SoxC	UniRef90_F7QKM3
SoxC	UniRef90_G8PPV5
SoxC	UniRef90_H1S1S8
SoxC	UniRef90_H1S535
SoxC	UniRef90_H9ZQD1
SoxC	UniRef90_I1B0R0
SoxC	UniRef90_I1X4F4
SoxC	UniRef90_I1X5E7
SoxC	UniRef90_I2IKC6
SoxC	UniRef90_I2K4I3
SoxC	UniRef90_I3UGE4
SoxC	UniRef90_I4WBQ2
SoxC	UniRef90_I9BU82
SoxC	UniRef90_K2IRL8
SoxC	UniRef90_L0NHN6
SoxC	UniRef90_Q07M25
SoxC	UniRef90_Q0JZ20
SoxC	UniRef90_Q0K5U1
SoxC	UniRef90_Q130M0
SoxC	UniRef90_Q16A44
SoxC	UniRef90_Q1YN56
SoxC	UniRef90_Q2IK73
SoxC	UniRef90_Q2PY06
SoxC	UniRef90_Q30NU7
SoxC	UniRef90_Q31JC1
SoxC	UniRef90_Q47BB7
SoxC	UniRef90_Q5ZQM9
SoxC	UniRef90_Q608I6
SoxC	UniRef90_R0EJL1
SoxC	UniRef90_R9RUA0
SoxC	UniRef90_S9QYV4
SoxC	UniRef90_V9VR91
SoxC	UniRef90_W4HNG0
SoxC	UniRef90_X2C028
SoxC	YP_001021628
SoxC	YP_166250

SoxC	YP_549445
SoxC	YP_576401
SoxC	ZP_00956140
SoxC	ZP_00961293
SoxC	ZP_01014860
SoxC	ZP_01037115
SoxC	ZP_01102563
SoxC	ZP_01167156
SoxC	ZP_01549056
SoxC	ZP_01627101
SoxC	ZP_01743243
SoxC	ZP_01748358
SoxD	AAF99437
SoxD	BAF34120
SoxD	CAA55825
SoxD	CAL61376
SoxD	NP_772760
SoxD	NP_772761
SoxD	UniRef90_A0A017HEE8
SoxD	UniRef90_A0A076K5N0
SoxD	UniRef90_A0A085BSM5
SoxD	UniRef90_A0A085F518
SoxD	UniRef90_A0A087LQX8
SoxD	UniRef90_A0KDR0
SoxD	UniRef90_A0NXV5
SoxD	UniRef90_A1B9M6
SoxD	UniRef90_A2SIK3
SoxD	UniRef90_A3JVA5
SoxD	UniRef90_A3K9X0
SoxD	UniRef90_A3SQX3
SoxD	UniRef90_A3SZG8
SoxD	UniRef90_A3VJS7
SoxD	UniRef90_A3W515
SoxD	UniRef90_A3WF82
SoxD	UniRef90_A3X8I7
SoxD	UniRef90_A4G4C9
SoxD	UniRef90_A4SX06
SoxD	UniRef90_A5EHW3
SoxD	UniRef90_A5G2Z9
SoxD	UniRef90_A6FWD0
SoxD	UniRef90_A9E7B1
SoxD	UniRef90_B3Y967
SoxD	UniRef90_B6B6H7

SoxD	UniRef90_B7QUA6
SoxD	UniRef90_B7RJ59
SoxD	UniRef90_C5ASA3
SoxD	UniRef90_D0CYH6
SoxD	UniRef90_D5X461
SoxD	UniRef90_E6PWB9
SoxD	UniRef90_F3L3A1
SoxD	UniRef90_F7QKM4
SoxD	UniRef90_G0EZH4
SoxD	UniRef90_G2HTV9
SoxD	UniRef90_G8PPV4
SoxD	UniRef90_G8PQT3
SoxD	UniRef90_H0TEH6
SoxD	UniRef90_H9ZQD2
SoxD	UniRef90_I1B0Q9
SoxD	UniRef90_I1X4F5
SoxD	UniRef90_I1X5E8
SoxD	UniRef90_I2K4I2
SoxD	UniRef90_I7EBT7
SoxD	UniRef90_J6UBC9
SoxD	UniRef90_K2K603
SoxD	UniRef90_K9GM86
SoxD	UniRef90_L0NGY6
SoxD	UniRef90_M7Z612
SoxD	UniRef90_Q07M26
SoxD	UniRef90_Q0F958
SoxD	UniRef90_Q0FI42
SoxD	UniRef90_Q0G1J3
SoxD	UniRef90_Q12A96
SoxD	UniRef90_Q130M1
SoxD	UniRef90_Q16A43
SoxD	UniRef90_Q1YN57
SoxD	UniRef90_Q2CI91
SoxD	UniRef90_Q2IK74
SoxD	UniRef90_Q2IRV6
SoxD	UniRef90_Q30NU8
SoxD	UniRef90_Q31JC0
SoxD	UniRef90_Q46W68
SoxD	UniRef90_Q47BB8
SoxD	UniRef90_Q5LUQ4
SoxD	UniRef90_Q5ZQM8
SoxD	UniRef90_Q6N1E3
SoxD	UniRef90_Q89PG3

SoxD	UniRef90_Q9LAH4
SoxD	UniRef90_S6GR29
SoxD	UniRef90_S9Q6N8
SoxD	UniRef90_S9QW81
SoxD	UniRef90_S9RG15
SoxD	UniRef90_S9SEX2
SoxD	UniRef90_U2Z8R6
SoxD	UniRef90_V4RPQ7
SoxD	UniRef90_V4TNH9
SoxD	UniRef90_W1IAE9
SoxD	UniRef90_W8S6R7
SoxD	UniRef90_W8SPY1
SoxD	YP_576402
SoxD	ZP_01102562
SoxD	ZP_01167155
SoxD	ZP_01627100
SoxX	AAF99431
SoxX	AAL68883
SoxX	BAF34124
SoxX	NP_214238
SoxX	NP_767654
SoxX	UniRef90_A0A017HCY5
SoxX	UniRef90_A0A024IFB5
SoxX	UniRef90_A0A037ZKQ3
SoxX	UniRef90_A0A058ZH00
SoxX	UniRef90_A0A059FH77
SoxX	UniRef90_A0A059FUE6
SoxX	UniRef90_A0A059ILR8
SoxX	UniRef90_A0A060A1Q0
SoxX	UniRef90_A0A061SWJ3
SoxX	UniRef90_A0A062V9X3
SoxX	UniRef90_A0A066ZSV2
SoxX	UniRef90_A0A069E0J3
SoxX	UniRef90_A0A069I633
SoxX	UniRef90_A0A072SPP3
SoxX	UniRef90_A0A073IJ63
SoxX	UniRef90_A0A085BSL9
SoxX	UniRef90_A0A085EWI5
SoxX	UniRef90_A0A085F512
SoxX	UniRef90_A0L8X4
SoxX	UniRef90_A0LE09
SoxX	UniRef90_A1AX72
SoxX	UniRef90_A1B9M0

SoxX	UniRef90_A3JVB2
SoxX	UniRef90_A3K9X6
SoxX	UniRef90_A3SQX9
SoxX	UniRef90_A3SZH5
SoxX	UniRef90_A3VJS1
SoxX	UniRef90_A3W521
SoxX	UniRef90_A3X8I1
SoxX	UniRef90_A4A7K6
SoxX	UniRef90_A4G4D4
SoxX	UniRef90_A4SCG3
SoxX	UniRef90_A4SX10
SoxX	UniRef90_A4SY09
SoxX	UniRef90_A4U039
SoxX	UniRef90_A4WZF8
SoxX	UniRef90_A4YQR0
SoxX	UniRef90_A5CW22
SoxX	UniRef90_A5EFF6
SoxX	UniRef90_A5EHV7
SoxX	UniRef90_A5G300
SoxX	UniRef90_A6FWC4
SoxX	UniRef90_A6Q627
SoxX	UniRef90_A7BMU2
SoxX	UniRef90_A8V4S5
SoxX	UniRef90_A9DG90
SoxX	UniRef90_A9E7C5
SoxX	UniRef90_B1Y5Z4
SoxX	UniRef90_B2UG46
SoxX	UniRef90_B2V885
SoxX	UniRef90_B3QNI6
SoxX	UniRef90_B4SEK9
SoxX	UniRef90_B6B6I3
SoxX	UniRef90_B6BLD5
SoxX	UniRef90_B7A5J7
SoxX	UniRef90_B7QUB2
SoxX	UniRef90_B7RJ65
SoxX	UniRef90_B8J5R7
SoxX	UniRef90_B8KK52
SoxX	UniRef90_C0QSK9
SoxX	UniRef90_C1DX01
SoxX	UniRef90_C5AS97
SoxX	UniRef90_C7CBM7
SoxX	UniRef90_D0J0E3
SoxX	UniRef90_D1KB33

SoxX	UniRef90_D3DJG6
SoxX	UniRef90_D3RVS6
SoxX	UniRef90_D5X457
SoxX	UniRef90_D6ZZE3
SoxX	UniRef90_D7A6E6
SoxX	UniRef90_D7BJA2
SoxX	UniRef90_D7BJA4
SoxX	UniRef90_E3I2K1
SoxX	UniRef90_E3T352
SoxX	UniRef90_E4TZM4
SoxX	UniRef90_E6PWC3
SoxX	UniRef90_E8PNQ5
SoxX	UniRef90_F3L3A2
SoxX	UniRef90_F6D9V8
SoxX	UniRef90_F7QKL7
SoxX	UniRef90_G2DG26
SoxX	UniRef90_G2FEE4
SoxX	UniRef90_G2HTW0
SoxX	UniRef90_G4ST10
SoxX	UniRef90_G7D5M1
SoxX	UniRef90_G8PPW0
SoxX	UniRef90_G9ZYW2
SoxX	UniRef90_H0RXF2
SoxX	UniRef90_H0SDE7
SoxX	UniRef90_H0T6J6
SoxX	UniRef90_H0TE11
SoxX	UniRef90_H0TEH0
SoxX	UniRef90_H0TTT9
SoxX	UniRef90_H9ZQC6
SoxX	UniRef90_I1B0R5
SoxX	UniRef90_I1X4E9
SoxX	UniRef90_I1X5E2
SoxX	UniRef90_I3BTG8
SoxX	UniRef90_I7EL29
SoxX	UniRef90_J5PG72
SoxX	UniRef90_K2KNX1
SoxX	UniRef90_K2M8N9
SoxX	UniRef90_K9GPT3
SoxX	UniRef90_L0DZN3
SoxX	UniRef90_L0NHP1
SoxX	UniRef90_L9PJU5
SoxX	UniRef90_M1SFR1
SoxX	UniRef90_M4ZBM7

SoxX	UniRef90_M5FCM9
SoxX	UniRef90_Q0BZB9
SoxX	UniRef90_Q0F963
SoxX	UniRef90_Q0FI48
SoxX	UniRef90_Q12AA0
SoxX	UniRef90_Q16A49
SoxX	UniRef90_Q1LHT5
SoxX	UniRef90_Q1QH93
SoxX	UniRef90_Q1YN51
SoxX	UniRef90_Q2BJX5
SoxX	UniRef90_Q30TZ0
SoxX	UniRef90_Q31I23
SoxX	UniRef90_Q3APA5
SoxX	UniRef90_Q3SF12
SoxX	UniRef90_Q3SKB7
SoxX	UniRef90_Q46W74
SoxX	UniRef90_Q47BC2
SoxX	UniRef90_Q5LUR0
SoxX	UniRef90_Q5ZQN1
SoxX	UniRef90_Q89PG9
SoxX	UniRef90_Q8KDN0
SoxX	UniRef90_S6AAT6
SoxX	UniRef90_S6CGK8
SoxX	UniRef90_S9QAN8
SoxX	UniRef90_S9QYV0
SoxX	UniRef90_S9S607
SoxX	UniRef90_U3AIF3
SoxX	UniRef90_U3QMU2
SoxX	UniRef90_U7FK99
SoxX	UniRef90_UPI0002557C77
SoxX	UniRef90_UPI0002624BF3
SoxX	UniRef90_UPI0002625346
SoxX	UniRef90_UPI00031C4E59
SoxX	UniRef90_UPI00035CC676
SoxX	UniRef90_UPI00037AE83F
SoxX	UniRef90_UPI0003B33203
SoxX	UniRef90_UPI0003B48222
SoxX	UniRef90_UPI0003F6E61F
SoxX	UniRef90_UPI0003FB94FE
SoxX	UniRef90_UPI0003FF5931
SoxX	UniRef90_UPI000402021D
SoxX	UniRef90_UPI0004047A27
SoxX	UniRef90_UPI0004114A47

SoxX	UniRef90_UPI000418F09F
SoxX	UniRef90_UPI000421118D
SoxX	UniRef90_UPI000427E63B
SoxX	UniRef90_UPI0004630B9F
SoxX	UniRef90_UPI00046491FB
SoxX	UniRef90_UPI0004649619
SoxX	UniRef90_UPI000468AFF0
SoxX	UniRef90_UPI000468B01A
SoxX	UniRef90_UPI00046A36EF
SoxX	UniRef90_UPI000474F109
SoxX	UniRef90_UPI000479E22C
SoxX	UniRef90_UPI00047C8CAD
SoxX	UniRef90_UPI00047D4666
SoxX	UniRef90_UPI00047DAE3C
SoxX	UniRef90_UPI00047E9CEA
SoxX	UniRef90_UPI00047F0B43
SoxX	UniRef90_UPI0004894EDC
SoxX	UniRef90_UPI00048B7235
SoxX	UniRef90_UPI00048CB8E7
SoxX	UniRef90_UPI00048E6357
SoxX	UniRef90_UPI00048EA2E6
SoxX	UniRef90_UPI00048FD590
SoxX	UniRef90_UPI0004947346
SoxX	UniRef90_UPI0004954A60
SoxX	UniRef90_UPI0004A27858
SoxX	UniRef90_UPI0004A759B6
SoxX	UniRef90_UPI0004DEE564
SoxX	UniRef90_V2GK44
SoxX	UniRef90_V4RW72
SoxX	UniRef90_V5UD20
SoxX	UniRef90_V8QYH5
SoxX	UniRef90_V9VR85
SoxX	UniRef90_W0DGR9
SoxX	UniRef90_W0SL49
SoxX	UniRef90_W0TR11
SoxX	UniRef90_W3RII9
SoxX	UniRef90_W6K750
SoxX	UniRef90_W8RTP1
SoxX	UniRef90_W9VKJ1
SoxX	UniRef90_X6KYI6
SoxX	UniRef90_X7F369
SoxX	YP_001003514
SoxX	YP_001021623

SoxX	YP_144684
SoxX	ZP_00051298
SoxX	ZP_01439480
SoxX	ZP_01583271
SoxX	ZP_01627096
SoxY	AAL68884
SoxY	NP_214241
SoxY	UniRef90_A0A011N2L9
SoxY	UniRef90_A0A011NUF7
SoxY	UniRef90_A0A011PYX1
SoxY	UniRef90_A0A011QRB6
SoxY	UniRef90_A0A014Q9I5
SoxY	UniRef90_A0A037ZM09
SoxY	UniRef90_A0A038FXH4
SoxY	UniRef90_A0A058ZI38
SoxY	UniRef90_A0A059FGX6
SoxY	UniRef90_A0A059FTV3
SoxY	UniRef90_A0A059ILL9
SoxY	UniRef90_A0A059ZX70
SoxY	UniRef90_A0A061SWV4
SoxY	UniRef90_A0A062VFF2
SoxY	UniRef90_A0A066ZTQ2
SoxY	UniRef90_A0A069DZZ9
SoxY	UniRef90_A0A073IW05
SoxY	UniRef90_A0A076K5N2
SoxY	UniRef90_A0A076PL66
SoxY	UniRef90_A0A085EWI4
SoxY	UniRef90_A0A085U1Y4
SoxY	UniRef90_A0A086MPU4
SoxY	UniRef90_A0L8X3
SoxY	UniRef90_A0LBZ5
SoxY	UniRef90_A0LE08
SoxY	UniRef90_A0Y7L0
SoxY	UniRef90_A1AX71
SoxY	UniRef90_A1B9M1
SoxY	UniRef90_A1BCG2
SoxY	UniRef90_A1VMV3
SoxY	UniRef90_A1WYE3
SoxY	UniRef90_A2SCZ4
SoxY	UniRef90_A2SHK2
SoxY	UniRef90_A2SHV5
SoxY	UniRef90_A2SIK2
SoxY	UniRef90_A3JVB1

SoxY	UniRef90_A3K9X5
SoxY	UniRef90_A3SQX8
SoxY	UniRef90_A3SZH4
SoxY	UniRef90_A3VJS2
SoxY	UniRef90_A3W520
SoxY	UniRef90_A3X8I0
SoxY	UniRef90_A4A7K5
SoxY	UniRef90_A4G4D1
SoxY	UniRef90_A4SCG2
SoxY	UniRef90_A4U038
SoxY	UniRef90_A4WZF7
SoxY	UniRef90_A4YQQ9
SoxY	UniRef90_A4Z2R6
SoxY	UniRef90_A5CW21
SoxY	UniRef90_A5EFF5
SoxY	UniRef90_A5EHV8
SoxY	UniRef90_A6FWC5
SoxY	UniRef90_A6FWD1
SoxY	UniRef90_A6Q0Z3
SoxY	UniRef90_A6Q626
SoxY	UniRef90_A6Q6A2
SoxY	UniRef90_A6Q7J9
SoxY	UniRef90_A9DG89
SoxY	UniRef90_A9E7C1
SoxY	UniRef90_B0UD54
SoxY	UniRef90_B0ULZ3
SoxY	UniRef90_B1XV34
SoxY	UniRef90_B1Y5Z3
SoxY	UniRef90_B1ZD61
SoxY	UniRef90_B2UG50
SoxY	UniRef90_B2V883
SoxY	UniRef90_B3EDN6
SoxY	UniRef90_B3QNI5
SoxY	UniRef90_B3QVQ9
SoxY	UniRef90_B4S942
SoxY	UniRef90_B4SAA2
SoxY	UniRef90_B6BLD4
SoxY	UniRef90_B7RJ64
SoxY	UniRef90_B8GSV6
SoxY	UniRef90_B8IQV4
SoxY	UniRef90_B8J5R3
SoxY	UniRef90_B8KK53
SoxY	UniRef90_C0QSK7

SoxY	UniRef90_D0J5V0
SoxY	UniRef90_D0KVR0
SoxY	UniRef90_D1KB32
SoxY	UniRef90_D3PPJ0
SoxY	UniRef90_D3RVA1
SoxY	UniRef90_D3SQG3
SoxY	UniRef90_D5X056
SoxY	UniRef90_D6ZZE4
SoxY	UniRef90_D7A6E4
SoxY	UniRef90_D7BJA8
SoxY	UniRef90_D8JT39
SoxY	UniRef90_D8JX76
SoxY	UniRef90_D8N4B2
SoxY	UniRef90_E0US18
SoxY	UniRef90_E2CPN8
SoxY	UniRef90_E3T353
SoxY	UniRef90_E4TZM5
SoxY	UniRef90_E6PWC0
SoxY	UniRef90_E6VNJ6
SoxY	UniRef90_E6WZX6
SoxY	UniRef90_E8PNR0
SoxY	UniRef90_F0IXH2
SoxY	UniRef90_F2J559
SoxY	UniRef90_F3L3A3
SoxY	UniRef90_F6D9V7
SoxY	UniRef90_F7QKM0
SoxY	UniRef90_F7XDM9
SoxY	UniRef90_F8J7I5
SoxY	UniRef90_F9UFE8
SoxY	UniRef90_F9ZRL9
SoxY	UniRef90_G2HTW1
SoxY	UniRef90_G4E208
SoxY	UniRef90_G4SW94
SoxY	UniRef90_G7D5M2
SoxY	UniRef90_G8NAE5
SoxY	UniRef90_G8PPV9
SoxY	UniRef90_G9AF80
SoxY	UniRef90_H0SDE8
SoxY	UniRef90_H0T6J7
SoxY	UniRef90_H0TE12
SoxY	UniRef90_H0TEH1
SoxY	UniRef90_H1G5T4
SoxY	UniRef90_I0GCH3

SoxY	UniRef90_I0HQT0
SoxY	UniRef90_I0HT98
SoxY	UniRef90_I1B0R4
SoxY	UniRef90_I1X4F0
SoxY	UniRef90_I1X5E3
SoxY	UniRef90_I1XH31
SoxY	UniRef90_I1YI19
SoxY	UniRef90_I2K4N9
SoxY	UniRef90_I2QQX6
SoxY	UniRef90_I3BR67
SoxY	UniRef90_I3BTH3
SoxY	UniRef90_I7EBQ8
SoxY	UniRef90_I9BUS6
SoxY	UniRef90_J0JHQ0
SoxY	UniRef90_J6IZW2
SoxY	UniRef90_K0C270
SoxY	UniRef90_K2J7E4
SoxY	UniRef90_K2K608
SoxY	UniRef90_K2M5E4
SoxY	UniRef90_K2M5S3
SoxY	UniRef90_K9HEN3
SoxY	UniRef90_L0DS17
SoxY	UniRef90_L0NGZ0
SoxY	UniRef90_L9PJS5
SoxY	UniRef90_M3ADQ6
SoxY	UniRef90_M4ZXP5
SoxY	UniRef90_M5FC80
SoxY	UniRef90_M7NY92
SoxY	UniRef90_N0B728
SoxY	UniRef90_Q07M22
SoxY	UniRef90_Q08IS2
SoxY	UniRef90_Q0A810
SoxY	UniRef90_Q0BZC0
SoxY	UniRef90_Q0F962
SoxY	UniRef90_Q0FI47
SoxY	UniRef90_Q0G1I8
SoxY	UniRef90_Q0K5U4
SoxY	UniRef90_Q12A97
SoxY	UniRef90_Q16A48
SoxY	UniRef90_Q1H4W5
SoxY	UniRef90_Q1QH94
SoxY	UniRef90_Q1YN52
SoxY	UniRef90_Q2IRV2

SoxY	UniRef90_Q30NU9
SoxY	UniRef90_Q30TY9
SoxY	UniRef90_Q31I24
SoxY	UniRef90_Q3APA4
SoxY	UniRef90_Q47BB9
SoxY	UniRef90_Q5SME5
SoxY	UniRef90_Q5ZQN0
SoxY	UniRef90_Q606M9
SoxY	UniRef90_Q89PG8
SoxY	UniRef90_Q89RM9
SoxY	UniRef90_Q8KDM9
SoxY	UniRef90_Q939U3
SoxY	UniRef90_Q9LCU9
SoxY	UniRef90_S6AN33
SoxY	UniRef90_S6BK39
SoxY	UniRef90_S6GM86
SoxY	UniRef90_S9Q6P7
SoxY	UniRef90_S9QW78
SoxY	UniRef90_S9SEW8
SoxY	UniRef90_T0JHL9
SoxY	UniRef90_U2FZY1
SoxY	UniRef90_U2Z8S0
SoxY	UniRef90_U7FLG9
SoxY	UniRef90_UPI0001D2F204
SoxY	UniRef90_UPI0002624BF2
SoxY	UniRef90_UPI000262534B
SoxY	UniRef90_UPI000262CA43
SoxY	UniRef90_UPI000303F7FD
SoxY	UniRef90_UPI0003471B8A
SoxY	UniRef90_UPI00035EAC6F
SoxY	UniRef90_UPI0003692C93
SoxY	UniRef90_UPI0003776C2D
SoxY	UniRef90_UPI0003794352
SoxY	UniRef90_UPI00037DC00C
SoxY	UniRef90_UPI00037EF751
SoxY	UniRef90_UPI000381A352
SoxY	UniRef90_UPI000399A29E
SoxY	UniRef90_UPI0003B3274C
SoxY	UniRef90_UPI0003B3BEF2
SoxY	UniRef90_UPI0003B488E9
SoxY	UniRef90_UPI0003B49753
SoxY	UniRef90_UPI0003D3E6EC
SoxY	UniRef90_UPI0003F8CDD6

SoxY	UniRef90_UPI000401FAF6
SoxY	UniRef90_UPI0004044E76
SoxY	UniRef90_UPI000404D468
SoxY	UniRef90_UPI0004159FDB
SoxY	UniRef90_UPI00041906BC
SoxY	UniRef90_UPI0004198596
SoxY	UniRef90_UPI00041D828D
SoxY	UniRef90_UPI00041DF2C9
SoxY	UniRef90_UPI0004242031
SoxY	UniRef90_UPI000428C6D4
SoxY	UniRef90_UPI00045EABAE
SoxY	UniRef90_UPI000463563B
SoxY	UniRef90_UPI000463915B
SoxY	UniRef90_UPI000464FF21
SoxY	UniRef90_UPI0004652082
SoxY	UniRef90_UPI00046898FE
SoxY	UniRef90_UPI000468AE21
SoxY	UniRef90_UPI000468F355
SoxY	UniRef90_UPI0004699585
SoxY	UniRef90_UPI000476F967
SoxY	UniRef90_UPI00047803F5
SoxY	UniRef90_UPI00047830A7
SoxY	UniRef90_UPI0004798D06
SoxY	UniRef90_UPI00047D0486
SoxY	UniRef90_UPI00047DD103
SoxY	UniRef90_UPI00047E679C
SoxY	UniRef90_UPI00047E8334
SoxY	UniRef90_UPI00047F3688
SoxY	UniRef90_UPI000482BF58
SoxY	UniRef90_UPI0004838221
SoxY	UniRef90_UPI000484A9D2
SoxY	UniRef90_UPI0004873CE1
SoxY	UniRef90_UPI00048743A7
SoxY	UniRef90_UPI000489AF17
SoxY	UniRef90_UPI00048D610F
SoxY	UniRef90_UPI00048E3569
SoxY	UniRef90_UPI00048E3ED9
SoxY	UniRef90_UPI00048E4C68
SoxY	UniRef90_UPI00048EB85E
SoxY	UniRef90_UPI00048EF79A
SoxY	UniRef90_UPI00048F4A6E
SoxY	UniRef90_UPI0004909348
SoxY	UniRef90_UPI000491074A

SoxY	UniRef90_UPI00049140EF
SoxY	UniRef90_UPI000493CF2B
SoxY	UniRef90_UPI0004944899
SoxY	UniRef90_UPI000494D7DB
SoxY	UniRef90_UPI0004961FC5
SoxY	UniRef90_UPI000496D483
SoxY	UniRef90_UPI000497F93C
SoxY	UniRef90_UPI0004A4BF04
SoxY	UniRef90_UPI0004A73F59
SoxY	UniRef90_V2HG49
SoxY	UniRef90_V4RPR2
SoxY	UniRef90_V4XXB6
SoxY	UniRef90_V5SG23
SoxY	UniRef90_V5UFF6
SoxY	UniRef90_V7FKT0
SoxY	UniRef90_V8QWX8
SoxY	UniRef90_V9R0A0
SoxY	UniRef90_V9VS74
SoxY	UniRef90_W0DFP8
SoxY	UniRef90_W0TSS3
SoxY	UniRef90_W1I831
SoxY	UniRef90_W1I9V2
SoxY	UniRef90_W1JTS6
SoxY	UniRef90_W3RFW4
SoxY	UniRef90_W3RHE1
SoxY	UniRef90_W4HNF5
SoxY	UniRef90_W4M2V1
SoxY	UniRef90_W6K921
SoxY	UniRef90_W6KAT5
SoxY	UniRef90_W7W9K2
SoxY	UniRef90_W7X0Z9
SoxY	UniRef90_W8KKJ7
SoxY	UniRef90_W8SPX6
SoxY	UniRef90_W9GV49
SoxY	UniRef90_W9T5H5
SoxY	UniRef90_W9V147
SoxY	UniRef90_W9V1I1
SoxY	UniRef90_X6GHK6
SoxY	UniRef90_X6KVK9
SoxY	UniRef90_X7F5A5
SoxY	YP_166246
SoxY	YP_314324
SoxY	ZP_00048026

SoxY	ZP_00944482
SoxY	ZP_01627099
SoxZ	AAL68885
SoxZ	NP_214240
SoxZ	UniRef90_A0A014NK72
SoxZ	UniRef90_A0A016XI15
SoxZ	UniRef90_A0A023WSD0
SoxZ	UniRef90_A0A024ELE4
SoxZ	UniRef90_A0A024IH18
SoxZ	UniRef90_A0A038GP89
SoxZ	UniRef90_A0A058ZJ83
SoxZ	UniRef90_A0A059FH44
SoxZ	UniRef90_A0A059FTX8
SoxZ	UniRef90_A0A059ZSV1
SoxZ	UniRef90_A0A059ZT96
SoxZ	UniRef90_A0A060B571
SoxZ	UniRef90_A0A062VHU1
SoxZ	UniRef90_A0A063Y027
SoxZ	UniRef90_A0A069E5G5
SoxZ	UniRef90_A0A069P469
SoxZ	UniRef90_A0A069PT42
SoxZ	UniRef90_A0A072SS55
SoxZ	UniRef90_A0A076K500
SoxZ	UniRef90_A0A076PPB1
SoxZ	UniRef90_A0A076PSP1
SoxZ	UniRef90_A0A081Y6N8
SoxZ	UniRef90_A0A083UHG7
SoxZ	UniRef90_A0A084CGI3
SoxZ	UniRef90_A0A085AJK3
SoxZ	UniRef90_A0A085EWI3
SoxZ	UniRef90_A0A085F515
SoxZ	UniRef90_A0A085U1Y5
SoxZ	UniRef90_A0L8X2
SoxZ	UniRef90_A0LBZ6
SoxZ	UniRef90_A0Y7K9
SoxZ	UniRef90_A1AX70
SoxZ	UniRef90_A1B9M2
SoxZ	UniRef90_A1BCG3
SoxZ	UniRef90_A1VMV4
SoxZ	UniRef90_A1WYE4
SoxZ	UniRef90_A2SHK2
SoxZ	UniRef90_A2SIK1
SoxZ	UniRef90_A4A7K4

SoxZ	UniRef90_A4G4D2
SoxZ	UniRef90_A4SCG1
SoxZ	UniRef90_A4SY07
SoxZ	UniRef90_A4U037
SoxZ	UniRef90_A4WZF6
SoxZ	UniRef90_A4Z2R7
SoxZ	UniRef90_A5CW20
SoxZ	UniRef90_A5EFF4
SoxZ	UniRef90_A5EHV9
SoxZ	UniRef90_A5G302
SoxZ	UniRef90_A6Q0Z4
SoxZ	UniRef90_A6Q6A3
SoxZ	UniRef90_A6Q7K0
SoxZ	UniRef90_A6SV43
SoxZ	UniRef90_A6UGH9
SoxZ	UniRef90_A7BZ97
SoxZ	UniRef90_A7IFH5
SoxZ	UniRef90_A7IG79
SoxZ	UniRef90_A8LNR0
SoxZ	UniRef90_B0UD53
SoxZ	UniRef90_B0UE04
SoxZ	UniRef90_B0ULZ4
SoxZ	UniRef90_B1LTL2
SoxZ	UniRef90_B1LTL3
SoxZ	UniRef90_B1XV33
SoxZ	UniRef90_B1Y090
SoxZ	UniRef90_B1Y5Z2
SoxZ	UniRef90_B1Y6K7
SoxZ	UniRef90_B1ZD62
SoxZ	UniRef90_B1ZHG3
SoxZ	UniRef90_B2UG49
SoxZ	UniRef90_B2V882
SoxZ	UniRef90_B3EDN7
SoxZ	UniRef90_B3Q927
SoxZ	UniRef90_B3QK90
SoxZ	UniRef90_B3QNI4
SoxZ	UniRef90_B3QVQ8
SoxZ	UniRef90_B4S943
SoxZ	UniRef90_B4SAA3
SoxZ	UniRef90_B4SEL1
SoxZ	UniRef90_B4U5W2
SoxZ	UniRef90_B4WZA8
SoxZ	UniRef90_B6BGQ5

SoxZ	UniRef90_B6BLD3
SoxZ	UniRef90_B6BUI0
SoxZ	UniRef90_B6JC94
SoxZ	UniRef90_B6JCA1
SoxZ	UniRef90_B8EMG4
SoxZ	UniRef90_B8GSV5
SoxZ	UniRef90_B8IQV5
SoxZ	UniRef90_B8ITV3
SoxZ	UniRef90_B8J5R4
SoxZ	UniRef90_B9R512
SoxZ	UniRef90_C0N5K9
SoxZ	UniRef90_C0N5L0
SoxZ	UniRef90_C0N6H3
SoxZ	UniRef90_C0QSK6
SoxZ	UniRef90_C1DX04
SoxZ	UniRef90_C4ZLW3
SoxZ	UniRef90_C6WT69
SoxZ	UniRef90_C6WY83
SoxZ	UniRef90_C6WZ72
SoxZ	UniRef90_C7CBM5
SoxZ	UniRef90_C8S585
SoxZ	UniRef90_D0CYI0
SoxZ	UniRef90_D0D4Q4
SoxZ	UniRef90_D0DC14
SoxZ	UniRef90_D0DC24
SoxZ	UniRef90_D0KVR1
SoxZ	UniRef90_D1KB31
SoxZ	UniRef90_D3DJG3
SoxZ	UniRef90_D3PPJ1
SoxZ	UniRef90_D3RVA2
SoxZ	UniRef90_D3SCC1
SoxZ	UniRef90_D3SQG2
SoxZ	UniRef90_D5CSJ5
SoxZ	UniRef90_D5V1N4
SoxZ	UniRef90_D5X057
SoxZ	UniRef90_D6V9F0
SoxZ	UniRef90_D6ZZE5
SoxZ	UniRef90_D7A6E3
SoxZ	UniRef90_D7A706
SoxZ	UniRef90_D7BJA6
SoxZ	UniRef90_D7BJA7
SoxZ	UniRef90_D7DL18
SoxZ	UniRef90_D7DN95

SoxZ	UniRef90_D8JTG3
SoxZ	UniRef90_D8JUZ8
SoxZ	UniRef90_D8JVS2
SoxZ	UniRef90_D8JX75
SoxZ	UniRef90_E0URX7
SoxZ	UniRef90_E0US19
SoxZ	UniRef90_E2CL87
SoxZ	UniRef90_E3CXC4
SoxZ	UniRef90_E3HWB4
SoxZ	UniRef90_E3I4S2
SoxZ	UniRef90_E3T354
SoxZ	UniRef90_E4QLU8
SoxZ	UniRef90_E4QQ55
SoxZ	UniRef90_E4TZM6
SoxZ	UniRef90_E6PWC1
SoxZ	UniRef90_E6WZX7
SoxZ	UniRef90_E7RVZ0
SoxZ	UniRef90_E8PNQ9
SoxZ	UniRef90_E8TEB2
SoxZ	UniRef90_F2KHA7
SoxZ	UniRef90_F2N5P0
SoxZ	UniRef90_F3KQZ3
SoxZ	UniRef90_F3L3A4
SoxZ	UniRef90_F3LM28
SoxZ	UniRef90_F4E032
SoxZ	UniRef90_F6D9V6
SoxZ	UniRef90_F7QKM1
SoxZ	UniRef90_F7XDN0
SoxZ	UniRef90_F8G2W4
SoxZ	UniRef90_F8JF67
SoxZ	UniRef90_F8JGS0
SoxZ	UniRef90_F9U727
SoxZ	UniRef90_F9U7E3
SoxZ	UniRef90_F9UF34
SoxZ	UniRef90_F9UFE9
SoxZ	UniRef90_F9ZV84
SoxZ	UniRef90_F9ZYJ3
SoxZ	UniRef90_G0JQ01
SoxZ	UniRef90_G2DAU2
SoxZ	UniRef90_G2DX33
SoxZ	UniRef90_G2HTW2
SoxZ	UniRef90_G3IS74
SoxZ	UniRef90_G3IY47

SoxZ	UniRef90_G4E207
SoxZ	UniRef90_G4SW93
SoxZ	UniRef90_G4T2D2
SoxZ	UniRef90_G8QKP3
SoxZ	UniRef90_G9ZYW4
SoxZ	UniRef90_H0HRE2
SoxZ	UniRef90_H0T6J8
SoxZ	UniRef90_H0TE13
SoxZ	UniRef90_H0TEH2
SoxZ	UniRef90_H1G5T3
SoxZ	UniRef90_H5WQS5
SoxZ	UniRef90_H5WUU6
SoxZ	UniRef90_H5WV54
SoxZ	UniRef90_H5YBP8
SoxZ	UniRef90_H5YL39
SoxZ	UniRef90_H8GQG4
SoxZ	UniRef90_H8Z1K6
SoxZ	UniRef90_H9ZQC2
SoxZ	UniRef90_I0GCH4
SoxZ	UniRef90_I0GGT5
SoxZ	UniRef90_I0HQT1
SoxZ	UniRef90_I0HT97
SoxZ	UniRef90_I1X4F1
SoxZ	UniRef90_I1X5E4
SoxZ	UniRef90_I1XH32
SoxZ	UniRef90_I1YI18
SoxZ	UniRef90_I2K4I0
SoxZ	UniRef90_I2K4N8
SoxZ	UniRef90_I3BR66
SoxZ	UniRef90_I3BTH0
SoxZ	UniRef90_I3CHS1
SoxZ	UniRef90_I3CIE6
SoxZ	UniRef90_I3UAU9
SoxZ	UniRef90_I3YDE8
SoxZ	UniRef90_I4WBQ5
SoxZ	UniRef90_I7DP42
SoxZ	UniRef90_I9CNH3
SoxZ	UniRef90_I9W4T0
SoxZ	UniRef90_J0BBT1
SoxZ	UniRef90_J2LF23
SoxZ	UniRef90_J6UCE8
SoxZ	UniRef90_J7QG90
SoxZ	UniRef90_K0CEB3

SoxZ	UniRef90_K2EHK9
SoxZ	UniRef90_K2KNA9
SoxZ	UniRef90_K7RJS1
SoxZ	UniRef90_K8R1Q6
SoxZ	UniRef90_K9GM94
SoxZ	UniRef90_K9NLQ5
SoxZ	UniRef90_L0DQM7
SoxZ	UniRef90_L0NHJ7
SoxZ	UniRef90_L0NIN5
SoxZ	UniRef90_L9PKB0
SoxZ	UniRef90_M1SUD1
SoxZ	UniRef90_M3ACX6
SoxZ	UniRef90_M4ZCH9
SoxZ	UniRef90_M5FBW6
SoxZ	UniRef90_M7PTJ2
SoxZ	UniRef90_N0B4W0
SoxZ	UniRef90_N0B9T2
SoxZ	UniRef90_N0BD58
SoxZ	UniRef90_N6Y5U2
SoxZ	UniRef90_N6Z231
SoxZ	UniRef90_Q07M23
SoxZ	UniRef90_Q08IS1
SoxZ	UniRef90_Q0A809
SoxZ	UniRef90_Q0BZC1
SoxZ	UniRef90_Q12A98
SoxZ	UniRef90_Q16A47
SoxZ	UniRef90_Q16BU1
SoxZ	UniRef90_Q1H4W4
SoxZ	UniRef90_Q1LHT2
SoxZ	UniRef90_Q1QH95
SoxZ	UniRef90_Q2IRV3
SoxZ	UniRef90_Q30NV0
SoxZ	UniRef90_Q30TY8
SoxZ	UniRef90_Q31I25
SoxZ	UniRef90_Q3APA3
SoxZ	UniRef90_Q47BC0
SoxZ	UniRef90_Q47BF9
SoxZ	UniRef90_Q4KEK7
SoxZ	UniRef90_Q606M7
SoxZ	UniRef90_Q89GZ8
SoxZ	UniRef90_Q89PG7
SoxZ	UniRef90_Q8KDM8
SoxZ	UniRef90_Q939U2

SoxZ	UniRef90_Q93RP7
SoxZ	UniRef90_S6AD68
SoxZ	UniRef90_S6BZD5
SoxZ	UniRef90_S6HCD0
SoxZ	UniRef90_S6HWT4
SoxZ	UniRef90_S6K860
SoxZ	UniRef90_S9R144
SoxZ	UniRef90_S9RG20
SoxZ	UniRef90_S9S3R2
SoxZ	UniRef90_T0KQT7
SoxZ	UniRef90_T0KUE9
SoxZ	UniRef90_U1GSD6
SoxZ	UniRef90_U1GWL5
SoxZ	UniRef90_U2YPX1
SoxZ	UniRef90_U4UY68
SoxZ	UniRef90_U6ZU35
SoxZ	UniRef90_U7HZE3
SoxZ	UniRef90_UPI0001D2E42C
SoxZ	UniRef90_UPI0002557C75
SoxZ	UniRef90_UPI000262534A
SoxZ	UniRef90_UPI000262CA42
SoxZ	UniRef90_UPI0002D43C00
SoxZ	UniRef90_UPI0002FC5BAD
SoxZ	UniRef90_UPI0003043DDC
SoxZ	UniRef90_UPI000316C1C9
SoxZ	UniRef90_UPI000347A4AE
SoxZ	UniRef90_UPI00036BCE3E
SoxZ	UniRef90_UPI000375F72D
SoxZ	UniRef90_UPI000376BB71
SoxZ	UniRef90_UPI00037B61A2
SoxZ	UniRef90_UPI00038205FC
SoxZ	UniRef90_UPI00039F1F0A
SoxZ	UniRef90_UPI0003A83415
SoxZ	UniRef90_UPI0003AB4228
SoxZ	UniRef90_UPI0003B342F9
SoxZ	UniRef90_UPI0003B41470
SoxZ	UniRef90_UPI0003B6E9FD
SoxZ	UniRef90_UPI0003F60C12
SoxZ	UniRef90_UPI0003F7BF12
SoxZ	UniRef90_UPI0003F9BE42
SoxZ	UniRef90_UPI0003F9FF46
SoxZ	UniRef90_UPI0003FC4C9B
SoxZ	UniRef90_UPI0003FE0C33

SoxZ	UniRef90_UPI0003FE3227
SoxZ	UniRef90_UPI0003FF5A8C
SoxZ	UniRef90_UPI000403F7D1
SoxZ	UniRef90_UPI0004046508
SoxZ	UniRef90_UPI0004059456
SoxZ	UniRef90_UPI0004062EE0
SoxZ	UniRef90_UPI0004067DEB
SoxZ	UniRef90_UPI00040FE41C
SoxZ	UniRef90_UPI0004108A98
SoxZ	UniRef90_UPI0004115B3F
SoxZ	UniRef90_UPI000416AE6C
SoxZ	UniRef90_UPI0004175FD1
SoxZ	UniRef90_UPI000417B19D
SoxZ	UniRef90_UPI0004182055
SoxZ	UniRef90_UPI00041C3D99
SoxZ	UniRef90_UPI00041D7F60
SoxZ	UniRef90_UPI000427A264
SoxZ	UniRef90_UPI00042A746A
SoxZ	UniRef90_UPI00045EA11F
SoxZ	UniRef90_UPI00045EA925
SoxZ	UniRef90_UPI0004628D24
SoxZ	UniRef90_UPI000464CABF
SoxZ	UniRef90_UPI0004674533
SoxZ	UniRef90_UPI000469F979
SoxZ	UniRef90_UPI00046EED09
SoxZ	UniRef90_UPI0004712D78
SoxZ	UniRef90_UPI00047147A5
SoxZ	UniRef90_UPI000474636C
SoxZ	UniRef90_UPI000475CCA1
SoxZ	UniRef90_UPI000475D8AC
SoxZ	UniRef90_UPI0004764188
SoxZ	UniRef90_UPI0004764F79
SoxZ	UniRef90_UPI00047685CE
SoxZ	UniRef90_UPI0004768EB6
SoxZ	UniRef90_UPI000477F027
SoxZ	UniRef90_UPI000478BA3C
SoxZ	UniRef90_UPI00047CA8CD
SoxZ	UniRef90_UPI00047D74EF
SoxZ	UniRef90_UPI0004825105
SoxZ	UniRef90_UPI000483ACAD
SoxZ	UniRef90_UPI000486F572
SoxZ	UniRef90_UPI000487E6F0
SoxZ	UniRef90_UPI0004881DA0

SoxZ	UniRef90_UPI0004887F27
SoxZ	UniRef90_UPI00048A8CDE
SoxZ	UniRef90_UPI00048AAEFB
SoxZ	UniRef90_UPI00048B0517
SoxZ	UniRef90_UPI00048B0EBD
SoxZ	UniRef90_UPI00048B8261
SoxZ	UniRef90_UPI00048BD26C
SoxZ	UniRef90_UPI00048BD286
SoxZ	UniRef90_UPI00048D80B2
SoxZ	UniRef90_UPI00048D9694
SoxZ	UniRef90_UPI00048DB779
SoxZ	UniRef90_UPI00048E7C0C
SoxZ	UniRef90_UPI000491AE87
SoxZ	UniRef90_UPI000491B33C
SoxZ	UniRef90_UPI0004945061
SoxZ	UniRef90_UPI0004946A64
SoxZ	UniRef90_UPI00049529C6
SoxZ	UniRef90_UPI0004954A99
SoxZ	UniRef90_UPI00049737F9
SoxZ	UniRef90_UPI000497F376
SoxZ	UniRef90_UPI0004A6B850
SoxZ	UniRef90_UPI0004DEF66E
SoxZ	UniRef90_UPI0004DFC4AE
SoxZ	UniRef90_V4TNI5
SoxZ	UniRef90_V5BQV4
SoxZ	UniRef90_V5SE51
SoxZ	UniRef90_V5SEW1
SoxZ	UniRef90_V5SGH0
SoxZ	UniRef90_V5SI82
SoxZ	UniRef90_V5UDU6
SoxZ	UniRef90_V6F7Y5
SoxZ	UniRef90_V6JAM9
SoxZ	UniRef90_V8QX52
SoxZ	UniRef90_V9R0A0
SoxZ	UniRef90_W0DCQ4
SoxZ	UniRef90_W0PFG0
SoxZ	UniRef90_W0SIR7
SoxZ	UniRef90_W0SK33
SoxZ	UniRef90_W0TPF2
SoxZ	UniRef90_W1I849
SoxZ	UniRef90_W1K234
SoxZ	UniRef90_W2DML3
SoxZ	UniRef90_W2DSY5

SoxZ	UniRef90_W6KA14
SoxZ	UniRef90_W6KAM3
SoxZ	UniRef90_W8KIR7
SoxZ	UniRef90_W8S6R4
SoxZ	UniRef90_W9H4S0
SoxZ	UniRef90_W9VS11
SoxZ	UniRef90_X6GLR8
SoxZ	UniRef90_X6KZR4
SoxZ	UniRef90_X6L3S8
SoxZ	UniRef90_X7EE56
SoxZ	UniRef90_X7F6W8
SoxZ	UniRef90_X7FC27
SoxZ	YP_166247
SoxZ	YP_314323
SoxZ	ZP_00056090
SoxZ	ZP_00956137
SoxZ	ZP_00963531
SoxZ	ZP_01014857
SoxZ	ZP_01037118
SoxZ	ZP_01055914
SoxZ	ZP_01167151
SoxZ	ZP_01167153
SoxZ	ZP_01225767
SoxZ	ZP_01439478
SoxZ	ZP_01549053
SoxZ	ZP_01627098
SoxZ	ZP_01743247
SoxZ	ZP_01748361
Total Protein Count	
AprA	91
AprB	92
DsrA	127
DsrB	102
Sat	5
SorA	51
SorB	43
SoxA	116
SoxB	155
SoxC	85
SoxD	90
SoxX	204
SoxY	298
SoxZ	413

Table 3.S3. List of transcripts that are significantly differentially transcribed between 15 m and other depths. Log2 fold change is relative to 15 m sample. Protein names are NCBI genome annotations. The other description column contains information taken from NCBI Conserved Domain Database (CDD), UniProt database, Pfam database, and analysis in this chapter (Arx and Dsr). NA = not applicable or available.

Depth	RefSeq Accession	Locus Tag	Log2 Fold-change	Adjusted p-value	Protein Name	Other Description
10 m	WP_015258468.1	TVNIR_RS07890	5.791	9.58E-18	hypothetical protein	Cysteine sulfinate desulfinase/cysteine desulfurase (CDD)
10 m	WP_043739787.1	TVNIR_RS14285	4.845	2.98E-06	cytochrome B561	Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA (UniProt)
10 m	WP_015257305.1	TVNIR_RS02055	4.508	3.16E-21	hypothetical protein	truncated hemoglobins (CDD)
10 m	WP_052316620.1	TVNIR_RS06290	4.468	0.001353	hypothetical protein	Protein Kinases, catalytic domain (CDD)
10 m	WP_052316608.1	TVNIR_RS03735	4.462	7.94E-25	hypothetical protein	Protein of unknown function (DUF461); Putative membrane or periplasmic protein. (CDD)
10 m	WP_052316646.1	TVNIR_RS14320	4.04	0.000296	cytochrome B559 subunit alpha	Cytochrome c oxidase subunit 2, CtaC (Uniprot)
10 m	WP_015259469.1	TVNIR_RS12875	3.999	4.66E-05	alkyl hydroperoxide reductase	NA
10 m	WP_043740715.1	TVNIR_RS14045	3.952	0.00239	hypothetical protein	Cell wall-associated hydrolase, NlpC family (CDD)
10 m	WP_015259726.1	TVNIR_RS14100	3.929	6.39E-05	pterin-4 alpha-carbinolamine dehydratase-like protein	NA
10 m	WP_006748004.1	TVNIR_RS04765	3.845	0.003118	MULTISPECIES: hypothetical protein	Uncharacterized low-complexity protein (CDD)
10 m	WP_015257778.1	TVNIR_RS04435	-3.845	8.72E-09	Ferredoxin	ArxB2
10 m	WP_015257771.1	TVNIR_RS04405	-3.866	2.81E-06	FKBP-type peptidyl-prolyl cis-trans isomerase	NA
10 m	WP_015257686.1	TVNIR_RS03975	-3.89	6.45E-05	Sulfite reduction-associated complex DsrMKJOP protein DsrK (HmeD)	DsrK
10 m	WP_015260521.1	TVNIR_RS18075	-4.119	7.19E-05	V-type ATP synthase subunit I	NA
10 m	WP_015258842.1	TVNIR_RS09675	-4.243	1.70E-06	Uptake hydrogenase large subunit	NA
10 m	WP_015257775.1	TVNIR_RS04420	-4.292	5.79E-15	polysulfide reductase NrfD	ArxC
10 m	WP_015258073.1	TVNIR_RS05905	-4.312	0.000257	succinate dehydrogenase flavoprotein subunit	NA
10 m	WP_052316670.1	TVNIR_RS04415	-4.382	6.44E-07	hypothetical protein	ArxD

10 m	WP_015257777.1	TVNIR_RS04430	-4.615	8.05E-18	formate dehydrogenase O subunit alpha	ArxA
10 m	WP_043740181.1	TVNIR_RS04425	-4.878	2.12E-11	ferredoxin (hmeA, FeS-containing hydrogenase components)	ArxB
18 m	WP_043739884.1	TVNIR_RS17020	3.751	5.53E-04	hypothetical protein	NA
18 m	WP_015256938.1	TVNIR_RS00285	3.187	1.34E-13	DUF302 domain-containing protein	NA
18 m	WP_015260523.1	TVNIR_RS18085	3.046	4.87E-06	H+-ATPase subunit H	NA
18 m	WP_043739885.1	TVNIR_RS17025	2.755	2.04E-03	hypothetical protein	ArsR family transcriptional regulator
18 m	WP_015260081.1	TVNIR_RS15870	2.754	3.59E-06	TRAP ABC transporter permease	NA
18 m	WP_015258817.1	TVNIR_RS09565	2.46	1.61E-05	Universal stress protein family	NA
18 m	WP_015259787.1	TVNIR_RS14390	2.396	7.93E-03	ribosomal protein S1	NA
18 m	WP_043740268.1	TVNIR_RS05620	2.347	1.55E-07	LysR family transcriptional regulator	NA
18 m	WP_015257410.1	TVNIR_RS02550	2.31	4.72E-03	hypothetical protein	Peptidase propeptide and YPEB domain (CDD)
18 m	WP_043739837.1	TVNIR_RS15865	2.282	1.55E-07	TRAP ABC transporter substrate-binding protein	NA
18 m	WP_015257685.1	TVNIR_RS03970	-2.983	8.48E-05	glutamate synthase	NA
18 m	WP_015259323.1	TVNIR_RS12110	-3.022	1.85E-04	50S ribosomal protein L22	NA
18 m	WP_015257690.1	TVNIR_RS03995	-3.078	1.40E-03	tRNA 5-methylaminomethyl-2-thiouridine synthase TusC	NA
18 m	WP_015257686.1	TVNIR_RS03975	-3.102	3.01E-05	Sulfite reduction-associated complex DsrMKJOP protein DsrK (HmeD)	DsrK
18 m	WP_015258557.1	TVNIR_RS08315	-3.106	1.09E-04	Cytochrome c-552 precursor	Ethylbenzene dehydrogenase (Pfam)
18 m	WP_015259322.1	TVNIR_RS12105	-3.113	2.37E-04	30S ribosomal protein S3	NA
18 m	WP_006746310.1	TVNIR_RS17160	-3.131	2.12E-03	MULTISPECIES: hypothetical protein	NA
18 m	WP_015257688.1	TVNIR_RS03985	-3.265	8.72E-05	sulfite reductase subunit gamma	DsrC
18 m	WP_015259417.1	TVNIR_RS12615	-3.35	4.42E-04	dihydrolipoyllysine succinyltransferase	NA
18 m	WP_015260292.1	TVNIR_RS16955	-3.649	5.66E-04	ribonucleotide reductase of class II (coenzyme B12-dependent), subunit alpha	NA
25 m	WP_015256938.1	TVNIR_RS00285	4.992	5.27E-38	DUF302 domain-containing protein	NA
25 m	WP_043740268.1	TVNIR_RS05620	4.038	2.50E-26	LysR family transcriptional regulator	NA
25 m	WP_043739884.1	TVNIR_RS17020	3.472	8.85E-04	hypothetical protein	NA
25 m	WP_015256896.1	TVNIR_RS00080	3.294	2.49E-14	DUF86 domain-containing protein	NA
25 m	WP_043739885.1	TVNIR_RS17025	3.015	2.07E-04	hypothetical protein	ArsR family transcriptional regulator

25 m	WP_043739084.1	TVNIR_RS02010	3.01	3.53E-05	nitrogen regulatory protein P-II 1	NA
25 m	WP_043739767.1	TVNIR_RS13665	2.901	9.53E-07	RND transporter	NA
25 m	WP_015257641.1	TVNIR_RS03755	2.854	7.05E-10	peroxiredoxin	NA
25 m	WP_015259146.1	TVNIR_RS11240	2.817	1.10E-09	hypothetical protein	NA
25 m	WP_015257742.1	TVNIR_RS04245	2.778	2.49E-14	XRE family transcriptional regulator	NA
25 m	WP_015260357.1	TVNIR_RS17275	-3.601	7.49E-03	MFS transporter	NA
25 m	WP_015257682.1	TVNIR_RS03955	-3.645	6.58E-05	Sulfite reduction-associated complex DsrMKJOP protein DsrP (HmeB)	DsrP
25 m	WP_015259736.1	TVNIR_RS14145	-3.669	1.23E-04	Ribulose bisphosphate carboxylase small chain	NA
25 m	WP_015258739.1	TVNIR_RS09175	-3.75	1.17E-05	NADH-ubiquinone oxidoreductase chain G	NA
25 m	WP_015258736.1	TVNIR_RS09165	-3.834	1.80E-03	NADH-ubiquinone oxidoreductase chain I	NA
25 m	WP_015260435.1	TVNIR_RS17665	-4.024	2.57E-09	ATP synthase subunit gamma	NA
25 m	WP_015259857.1	TVNIR_RS14725	-4.146	3.42E-07	NADH-ubiquinone oxidoreductase chain N	NA
25 m	WP_015257683.1	TVNIR_RS03960	-4.312	6.89E-04	4Fe-4S ferredoxin iron-sulfur-binding domain-containing protein	DsrO
25 m	WP_015257688.1	TVNIR_RS03985	-4.475	3.40E-08	sulfite reductase subunit gamma	DsrC
25 m	WP_015258557.1	TVNIR_RS08315	-4.671	1.63E-08	Cytochrome c-552 precursor	Ethylbenzene dehydrogenase (Pfam)
31 m	WP_015256938.1	TVNIR_RS00285	5.105	2.38E-39	DUF302 domain-containing protein	NA
31 m	WP_015256896.1	TVNIR_RS00080	4.06	1.07E-21	DUF86 domain-containing protein	NA
31 m	WP_043740268.1	TVNIR_RS05620	4.019	1.68E-25	LysR family transcriptional regulator	NA
31 m	WP_043739884.1	TVNIR_RS17020	3.769	2.27E-04	hypothetical protein	NA
31 m	WP_015256895.1	TVNIR_RS00075	3.756	3.21E-13	nucleotidyltransferase	NA
31 m	WP_015257641.1	TVNIR_RS03755	3.68	1.56E-16	peroxiredoxin	NA
31 m	WP_043739885.1	TVNIR_RS17025	3.501	1.02E-05	hypothetical protein	ArsR family transcriptional regulator
31 m	WP_015259146.1	TVNIR_RS11240	3.292	4.65E-13	hypothetical protein	NA
31 m	WP_043739767.1	TVNIR_RS13665	3.251	3.93E-08	RND transporter	NA
31 m	WP_015258019.1	TVNIR_RS05640	3.191	8.06E-05	AsnC family transcriptional regulator	NA
31 m	WP_015259736.1	TVNIR_RS14145	-3.949	0.000387	Ribulose bisphosphate carboxylase small chain	NA
31 m	WP_015257688.1	TVNIR_RS03985	-4.022	4.82E-06	sulfite reductase subunit gamma	DsrC

31 m	WP_015257886.1	TVNIR_RS04925	-4.05	0.000406	2-octaprenyl-6-methoxyphenol hydroxylase	NA
31 m	WP_015258557.1	TVNIR_RS08315	-4.078	4.51E-06	Cytochrome c-552 precursor	Ethylbenzene dehydrogenase (Pfam)
31 m	WP_015258739.1	TVNIR_RS09175	-4.169	2.68E-05	NADH-ubiquinone oxidoreductase chain G	NA
31 m	WP_015257691.1	TVNIR_RS04000	-4.363	5.06E-05	tRNA 5-methylaminomethyl-2-thiouridine synthase TusD	NA
31 m	WP_015260284.1	TVNIR_RS16920	-4.389	0.000143	ATP/GTP-binding protein; DUF299	NA
31 m	WP_015259849.1	TVNIR_RS14680	-4.407	3.90E-08	phosphatidylethanolamine N-methyltransferase	NA
31 m	WP_015259378.1	TVNIR_RS12420	-4.655	1.75E-05	histidinol dehydrogenase	NA
31 m	WP_015260435.1	TVNIR_RS17665	-4.734	1.68E-07	ATP synthase subunit gamma	NA

Table 3.S4. Differential transcription analysis of *Thioalkalivibrio* sulfur- and arsenic-related transcripts. Sulfur genes are highlighted in yellow and arsenic genes are highlighted in blue. Locus tags are sorted in order of appearance in genome, and highlighted arbitrarily to show adjacent genes. For arsenite oxidase (arxA) operon, the *Alkalilimnicola ehrlichii* MLHE-1 homolog locus tag is listed in the first column. Locus number corresponds to NCBI locust tag (TVNIR_RS#####). Log2± is log2 fold change. Values > ±2 are in bold. FDR is false discovery rate (adjusted *p*-value). Significant (<0.05) values are in bold.

Gene Name or Homolog	Arsenic or Sulfur Gene	Locus	RefSeq Accession	Log2± (10 m)	FDR	Log2± (15 m)	FDR	Log2± (25 m)	FDR	Log2± (31 m)	FDR
SoxZ	thiosulfate oxidation carrier complex protein SoxZ	00345	WP_015256949.1	-0.759	1.64E-01	-1.381	**6.65E-03**	**-2.469**	**1.95E-08**	**-2.224**	**9.29E-07**
SoxY	thiosulfate oxidation carrier protein SoxY	00350	WP_015256950.1	-1.065	**3.80E-02**	0.474	4.46E-01	-0.871	1.02E-01	-0.294	6.21E-01
FccB	Sulfide dehydrogenase	01205	WP_015257122.1	-0.881	5.55E-01	-1.340	NA	-1.395	2.99E-01	-1.748	1.44E-01
FccA	cytochrome c subunit of flavocytochrome c sulfide dehydrogenase	01210	WP_015257123.1	-1.296	NA	-1.317	NA	-0.952	NA	-1.443	NA
CysQ	3\'(2\'),5\'-bisphosphate nucleotidase	01340	WP_015257154.1	1.795	1.15E-01	-0.305	NA	-1.614	2.64E-01	-0.568	7.01E-01
FccB	Sulfide dehydrogenase	02535	WP_015257406.1	-0.137	8.02E-01	-1.028	**1.12E-02**	-1.726	**6.91E-07**	**-2.147**	**1.24E-09**
FccA	cytochrome c, class I	02540	WP_015257407.1	0.807	1.96E-01	0.002	9.99E-01	-1.513	**7.32E-03**	-0.963	1.09E-01
CysM	hypothetical protein	03880	WP_015257666.1	0.413	6.73E-01	1.460	**2.69E-02**	1.095	1.14E-01	1.577	**7.42E-03**
DsrB	sulfite reductase subunit beta	04005	WP_015257692.1	**-2.724**	**1.14E-03**	-1.969	**1.08E-02**	**-3.516**	**3.42E-07**	**-3.927**	**4.42E-07**
DsrA	sulfite reductase subunit alpha	04010	WP_015257694.1	**-2.071**	**1.02E-04**	-1.026	**4.03E-02**	**-3.343**	**2.07E-14**	**-3.289**	**2.16E-11**
CysE	serine acetyltransferase	04055	WP_015257703.1	0.965	2.17E-01	-0.012	9.97E-01	0.126	9.24E-01	0.086	9.33E-01
mlg_0213	arxD	04415	WP_052316670.1	**-4.382**	**6.44E-07**	-0.828	2.96E-01	-1.906	**2.03E-03**	-1.449	**2.17E-02**

mlg_0214	arxC	04420	WP_015257775.1	**-4.292**	**5.79E-15**	-1.036	1.15E-01	**-2.425**	**3.84E-06**	-1.902	**4.06E-04**
mlg_0215	arxB	04425	WP_043740181.1	**-4.878**	**2.12E-11**	-0.863	2.41E-01	**-2.328**	**3.53E-05**	-1.595	**6.17E-03**
mlg_0216	arxA	04430	WP_015257777.1	**-4.615**	**8.05E-18**	-1.280	**1.80E-02**	**-2.144**	**3.60E-06**	-1.768	**1.84E-04**
mlg_0217	arxB2	04435	WP_015257778.1	**-3.845**	**8.72E-09**	-1.640	**4.04E-04**	**-2.453**	**5.44E-09**	**-2.433**	**7.97E-08**
mlg_0218	arxX	04445	WP_015257780.1	**-2.030**	**1.95E-02**	0.137	9.18E-01	-0.911	2.66E-01	-0.678	3.76E-01
mlg_0219	arxS	04450	WP_015257781.1	**-2.680**	**1.76E-02**	-0.292	8.17E-01	-0.915	3.61E-01	-1.529	7.63E-02
mlg_0220	arxR	04455	WP_015257782.1	-1.627	**4.51E-02**	-0.400	6.84E-01	-1.381	**4.95E-02**	**-2.126**	**2.05E-03**
TssA/SseA	Thiosulfate sulfurtransferase rhodanese	04460	WP_015257783.1	-0.431	4.57E-01	-0.151	8.34E-01	-1.331	**1.69E-03**	-1.127	**1.01E-02**
ND	arsenic efflux pump protein	04480	WP_015257790.1	-1.704	2.56E-01	-0.041	NA	0.163	9.44E-01	0.885	4.80E-01
ND	Rhodanese-related sulfurtransferase (part of arrA operon)	04525	WP_015257802.1	**-2.467**	**1.00E-02**	0.026	9.86E-01	-0.197	8.67E-01	0.600	3.55E-01
ND	arrC	04530	WP_015257803.1	**-2.676**	**8.25E-05**	-1.182	**4.71E-02**	**-2.130**	**2.40E-05**	-1.941	**2.72E-04**
ND	arrB	04535	WP_015257804.1	**-2.970**	**3.75E-03**	-0.742	4.61E-01	-1.641	**4.43E-02**	-0.905	2.83E-01
ND	arrA	04540	WP_015257805.1	**-2.713**	**4.27E-05**	-1.109	9.21E-02	**-2.191**	**5.27E-05**	-1.524	**6.85E-03**
Sqr	pyridine nucleotide-disulfide oxidoreductase	05510	WP_043739419.1	**-2.139**	**1.63E-02**	1.039	8.58E-02	0.388	6.64E-01	0.712	2.24E-01
FccB	Sulfide dehydrogenase	07250	WP_015258343.1	1.741	**4.32E-02**	-0.184	9.18E-01	-1.826	7.73E-02	-1.528	1.49E-01
MetB	cystathionine gamma-synthase - like protein	07995	WP_015258491.1	-1.162	4.26E-01	0.398	8.00E-01	0.589	7.17E-01	0.935	3.86E-01
Sqr	pyridine nucleotide-disulfide oxidoreductase	08015	WP_015258495.1	-0.284	NA	-0.475	NA	-0.534	NA	-0.428	NA

ND	ArsR family transcriptional regulator	08035	WP_015258499.1	-1.987	**2.87E-02**	0.434	5.82E-01	0.526	5.00E-01	0.892	1.16E-01
DsrA	sulfite reductase subunit alpha	08305	WP_015258555.1	-1.195	NA	-1.368	NA	-1.415	NA	-1.332	NA
ND	arsenate reductase-like protein	08685	WP_015258635.1	0.053	9.73E-01	-0.780	NA	-1.270	2.95E-01	-1.577	1.50E-01
Sor	sulfur oxygenase reductase	12880	WP_015259470.1	-0.340	NA	0.227	NA	-0.660	NA	0.566	NA
FccA	sulfide dehydrogenase cytochrome subunit	13010	WP_015259498.1	0.000	NA	0.000	NA	0.000	NA	0.858	NA
FccB	Sulfide dehydrogenase	13015	WP_015259499.1	-0.628	NA	**-2.116**	NA	**-2.911**	NA	**-2.571**	NA
MetZ	O-succinylhomoserine sulfhydrylase	15095	WP_043739807.1	0.895	2.61E-01	0.527	5.77E-01	-0.199	8.89E-01	-0.058	9.55E-01
SoxB	thiosulfohydrolase SoxB	15415	WP_015259997.1	-0.706	2.31E-01	-0.271	7.14E-01	**-2.863**	**3.62E-10**	**-2.388**	**9.29E-07**
Sqr	Sulfide-quinone reductase	15910	WP_015260090.1	0.000	NA	1.013	NA	0.000	NA	0.000	NA
ND	arsenate reductase	16090	WP_015260128.1	1.028	4.45E-01	0.766	NA	0.514	7.87E-01	1.154	3.02E-01
SoxX	sulfur oxidation c-type cytochrome SoxX	16210	WP_015260153.1	1.454	**3.73E-03**	-1.715	**1.04E-03**	**-3.510**	**2.49E-14**	**-3.110**	**8.11E-11**
SoxA	sulfur oxidation c-type cytochrome SoxA	16215	WP_015260154.1	-0.174	7.52E-01	-0.537	2.62E-01	**-3.171**	**4.01E-19**	**-3.071**	**1.50E-15**
ND	Arsenical-resistance protein ACR3	17030	WP_015260308.1	1.736	1.68E-01	-0.327	NA	-0.112	9.68E-01	-0.985	5.13E-01
AprA	adenylylsulfate reductase alpha-subunit	17165	WP_015260335.1	**-2.528**	**1.95E-04**	-1.099	1.51E-01	-0.895	2.77E-01	**-2.391**	**1.33E-04**
AprB	adenylylsulfate reductase	17170	WP_043739894.1	**-2.994**	**6.10E-04**	0.206	8.47E-01	0.334	7.70E-01	-0.139	8.73E-01
Sat	sulfate adenylyltransferase	17180	WP_015260338.1	-0.734	2.84E-01	-0.801	2.06E-01	-0.355	6.85E-01	-0.511	4.05E-01
MetA	homoserine O-succinyltransferase	17270	WP_015260356.1	1.245	1.23E-01	-1.710	6.68E-02	-1.618	5.78E-02	**-2.270**	**1.06E-02**
ND	ArsR family transcriptional regulator	17730	WP_015260448.1	0.564	3.51E-01	**2.039**	**1.64E-06**	1.999	**4.24E-07**	**2.581**	**1.59E-11**
FccA	cytochrome c subunit of flavocytochrome c sulfide dehydrogenase	18010	WP_015260507.1	-0.675	NA	-0.142	NA	-0.593	NA	-0.209	NA

FccB	Sulfide dehydrogenase	18015	WP_015260508.1	0.615	NA	-0.274	NA	-1.040	NA	-1.282	NA
ND	ArsR family transcriptional regulator	18280	WP_015260564.1	-1.431	NA	-0.774	NA	0.409	NA	0.356	NA

APPENDIX C

SUPPORTING INFORMATION FOR CHAPTER 4

Table 4.S1. Sources and growth conditions for pure cultures used in this study

Culture	Medium	Source	Reference
***Halorhodospira halophila* SL1 (DSM-244)**	DSMZ # 253 (omitted: yeast extract, sodium succinate, sodium sulfide)	DSMZ[a]	DSMZ
***Halorhodospira abdelmalekii* (DSM-2110)**	DSMZ # 430 (omitted: sodium acetate, sodium sulfide)	DSMZ	DSMZ
***Ectothiorhodospira shaposhnikovii* BN9512 (DSM-2111)**	DSMZ # 431 (omitted: sodium malate, sodium sulfide, sodium thiosulfate)	DSMZ	DSMZ
***Thioalkalivibrio jannaschii* (DSM-14478)**	DSMZ # 1272	DSMZ	DSMZ
***Alkalilimnicola ehrlichii* MLHE-1**	Artificial Mono Lake Water	Ron S. Oremland, USGS, Menlo Park, CA	(Hoeft *et al.* 2007)
***Ectothiorhodospira* sp. PHS-1**	Artificial Mono Lake Water	Ron S. Oremland, USGS, Menlo Park, CA	(Kulp *et al.* 2008)
***Ectothiorhodospira* sp. Bogoria Red**	Artificial Mono Lake Water	Michael T. Madigan, Southern Illinois University, Carbondale, IL	N/A
Allochromatium vinosum	'0' Medium	Christiane Dahl, University of Bonn, Germany	(Hensen *et al.* 2006)

[a]Deutsche Sammlung von Mikroorganismen und Zellkulturen

Table 4.S2. Top blastn hits of enrichment cultures clone libraries.

Enrichment Culture Clone Library	Number of Clones (% of total library)	Top RefSeq hit (Accession)	% Identity	Top nr hit (Accession)	% Identity
EC "ML Ecto"	9 (100%)	*Ectothiorhodospira variabilis* WN22 (NR_042700)	99	*Ectothiorhodospira* sp. 'Bogoria Red' (AF384207)	99
Mono Lake EC8 (late transfer)	18 (100%)	*Ectothiorhodospira* variabilis WN22 (NR_042700)	99	*Ectothiorhodospira* sp. 'ML Ecto' (EU341299)	99
Mono Lake EC12	7 (87.5%)	*Ectothiorhodospira* variabilis WN22 (NR_042700)	99	*Ectothiorhodospira* sp. 'Bogoria Red' (AF384207)	99
	1 (12.5%)	*Tindallia texcoconensis* IMP300 (NR_043664)	97	Same as cultured hit.	N/A
Big Soda Lake EC13	7 (87.5%)	*Ectothiorhodospira shaposhnikovii* BN 9512 (NR_104733)	99	*Ectothiorhodospira shaposhnikovii* 17G (EU700082)	99
	1 (12.5%)	*Bacillus cytotoxicus* strain NVH 391-98 (NR_074914)	82	Uncultured low G+C Gram-positive bacterium clone ML602J-25 (AF507895)	94
Mono Lake EC8 Library 1	16 (53.3%)	*Ectothiorhodospira variabilis* WN22 (NR_042700)	99	*Ectothiorhodospira* sp. 'ML Ecto' (EU341299)	99
	5 (16.7%)	*Alkaliflexus imshenetskii* Z-7010 (AJ784993)	84	Uncultured Bacteroidetes bacterium clone ML635J-56 (AF507862)	98
	4 (13.3%)	*Natronincola histidinovorans* Z-7940 (NR_026455)	91	Uncultured bacterium clone SINH899 (HM128341)	98
	1 (3.3%)	*Alkaliphilus transvaalensis* SAGM1 (NR_024748)	92	Uncultured bacterium clone SINH934 (HM128363)	96
	1 (3.3%)	*Gracilimonas tropica* CL-CB462 (NR_044361)	90	Uncultured Bacteroidetes bacterium clone CSS3 (JX240587)	96
	1 (3.3%)	*Gracilimonas rosea* CL-KR2 (NR_109751)	87	Uncultured Bacteroidetes bacterium clone CSS3 (JX240587)	94

	1 (3.3%)	*Halanaerobium hydrogeniformans* (NR_074850)	97	Uncultured Firmicutes bacterium clone CSS73 (JX240655)	98
	1 (3.3%)	*Spirochaeta americana* ASpG1 (NR_028820)	97	Uncultured bacterium clone SA_123 (JQ739016)	99
Mono Lake EC8 Library 2	19 (73.1%)	*Ectothiorhodospira variabilis* WN22 (NR_042700)	99	*Ectothiorhodospira* sp. 'ML Ecto' (EU341299)	99
	7 (26.9%)	*Natronincola histidinovorans* Z-7940 (NR_026455)	91	Uncultured bacterium clone SINH899 (HM128341)	97
Mono Lake EC8 Library 3	11 (40.7%)	*Ectothiorhodospira variabilis* WN22 (NR_042700)	99	*Ectothiorhodospira* sp. 'ML Ecto' (EU341299)	99
	9 (33.3%)	*Tindallia texcoconensis* IMP-300 (NR_043664)	98	Same as cultured hit.	N/A
	2 (7.4%)	*Natronincola histidinovorans* Z-7940 (NR_026455)	92	Uncultured bacterium clone SINH934 (HM128363)	98
	2 (7.4%)	*Alkaliphilus transvaalensis* SAGM1 (NR_024748)	92-93	Uncultured bacterium clone SINH934 (HM128363)	96
	1 (3.7%)	*Natronincola histidinovorans* Z-7940 (NR_026455)	93	Uncultured bacterium clone Bac_SB_118 (JQ739060)	98
	1 (3.7%)	*Natronoflexus pectinivorans* AP1 (NR_108635)	87	Uncultured Bacteroidetes bacterium clone ML635J-56 (AF507862)	99
	1 (3.7%)	*Tindallia magadiensis* Z-7934 (NR_026446)	91	Uncultured bacterium clone SINH934 (HM128363)	95

NCBI GenBank accession numbers are in parentheses.

Supporting Table References

Hensen D, Sperling D, Truper HG, Brune DC, Dahl C (2006). Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochromatium vinosum*. *Mol Microbiol* **62:** 794-810.

Hoeft SE, Blum JS, Stolz JF, Tabita FR, Witte B, King GM *et al.* (2007). *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. *Int J Syst Evol Microbiol* **57:** 504-512.

Kulp TR, Hoeft SE, Asao M, Madigan MT, Hollibaugh JT, Fisher JC *et al.* (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science* **321:** 967-970.

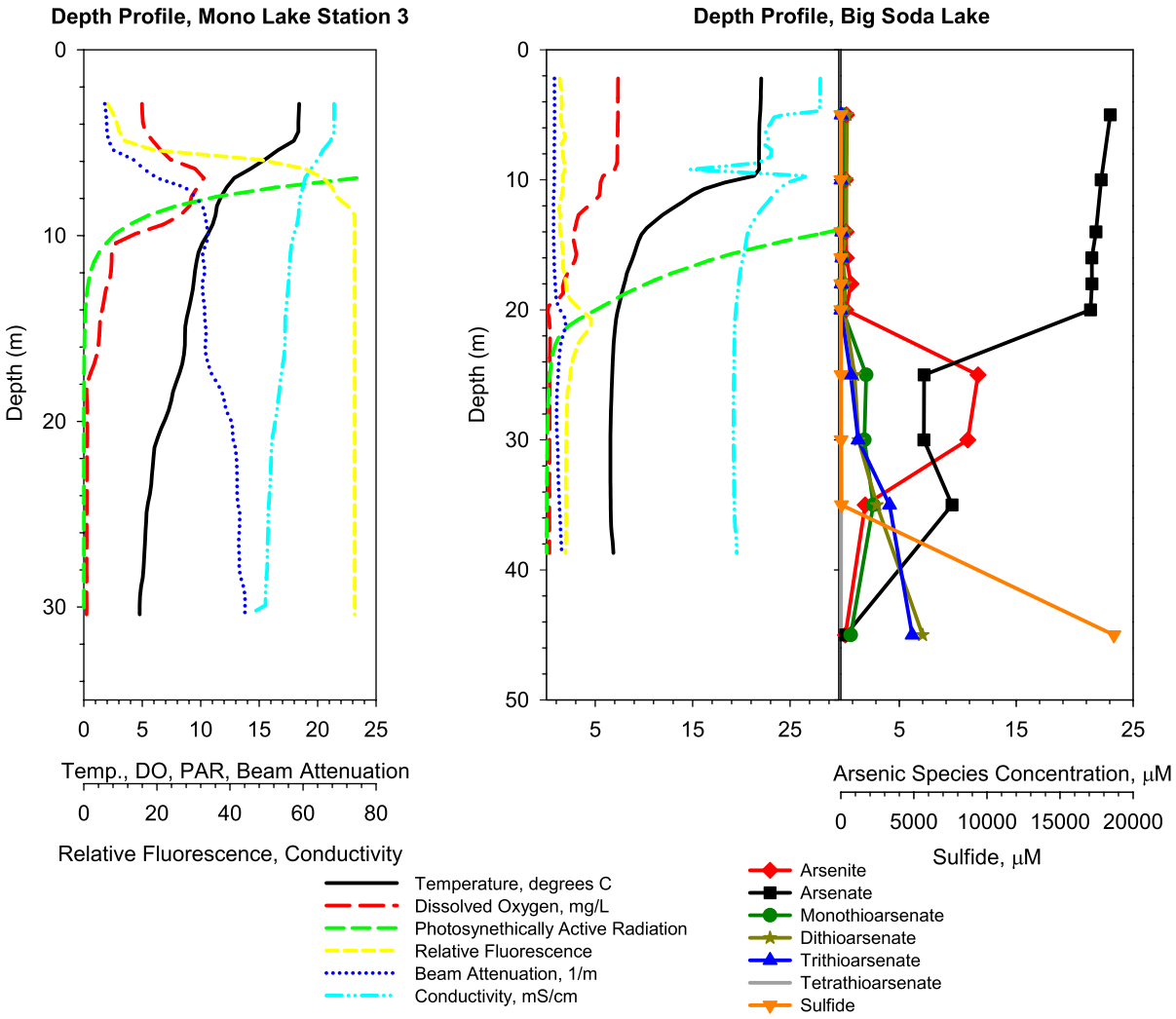


Figure 4.S1. CTD and chemical profiles of Mono Lake Station 3 (June 29, 2011), and Big Soda Lake (September 20, 2011).

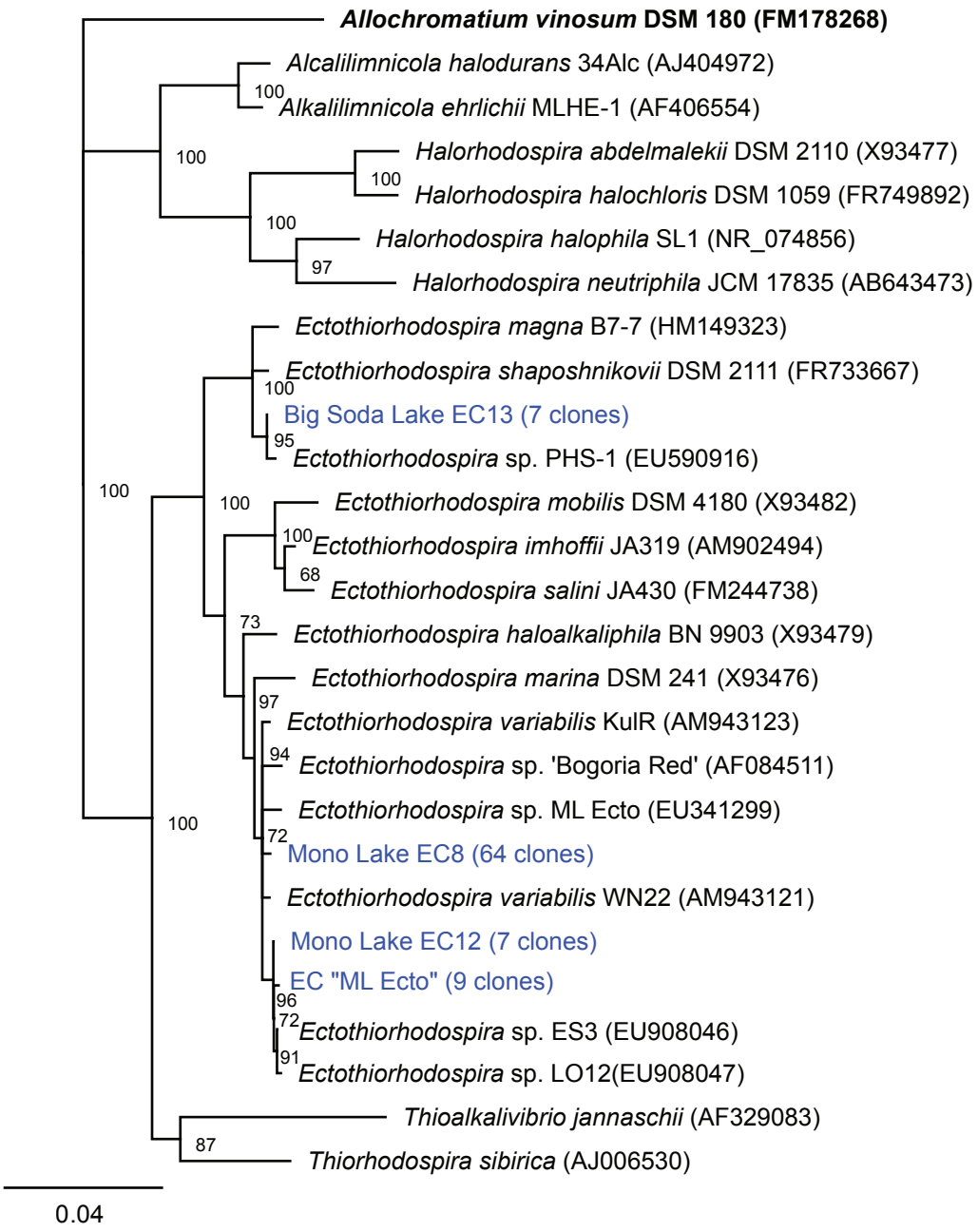


Figure 4.S2. Neighbor-Joining consensus tree based on 16S rRNA gene sequences. Node labels are consensus support values based on 100 bootstraps. Consensus sequences from enrichment cultures are in blue. The outgroup is in bold. Genbank accession numbers for reference sequences are given in parentheses.