UNIVERSITY OF CALIFORNIA
SANTA CRUZ

USING SYSTEMS BIOLOGY APPROACHES TO ELUCIDATE THE MECHANISMS OF ARSENIC REDUCTION IN *SHEWANELLA* SP. ANA-3

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELL, AND DEVELOPMENTAL BIOLOGY

by

Ruth Pamela Tilus Watson

June 2015

The Dissertation of Ruth Pamela Tilus Watson

is approved:

_____

Professor Chad W. Saltikov, Chair

_____

Distinguished Professor Susan Strome

_____

Assistant Professor Needhi Bhalla

_____

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

List of Supplementary Figures and Tables

Abstract of the dissertation

Using systems biology approaches to elucidate the mechanisms of arsenic reduction in *Shewanella* sp. ANA-3

By Ruth Pamela Tilus Watson

Doctor of Philosophy in Molecular, Cell and Developmental Biology

University of California, Santa Cruz, 2015

Professor Chad W. Saltikov, Chair

Arsenic is a naturally occurring ubiquitous metalloid that is usually associated with Iron, sulfur and other compounds in the earth's crust. In some places around the world the bio-geochemical conditions can cause the mineral bound form of arsenic (arsenate) to be reduced to a more water-soluble form (arsenite). In its reduced state, arsenic can seep from the soil down into ground water aquifers and contaminate drinking water supplies. The effects of drinking arsenic tainted water are devastating, however the current methods of detecting arsenic contamination are difficult to implement in most communities that rely on individual tube wells for drinking water.

Dissimilatory arsenic reducing microbes can release terminal electrons, from their electron transport chain, onto arsenic in the nearby environment. These electrons then cause arsenate to be reduced to arsenite leading to eventual arsenic pollution of drinking water. However, arsenic reduction only occurs when more favorable terminal electron acceptors, like oxygen or nitrate, are depleted. An understanding of

the genetic and biochemical triggers that activate dissimilatory arsenic reduction will help future detection and bioremediation efforts. This dissertation is a compilation of several systems level studies I performed in order to understand the molecular mechanisms regulating arsenic reduction in *Shewanella* sp. ANA-3. The techniques include transcriptomics, molecular biology and comparative genomics. Using systems biology techniques I identified important regulatory parameters in activating respiratory arsenic reduction in three forms:

- Transcriptomics of ANA-3 during arsenic reduction vs. growth on oxygen or fumarate
- Characterization of the role of cAMP as a global regulator of cellular respiration in ANA-3
- Isolation of a novel arsenic respiring microbe from an arsenic rich hot spring

Systems biological approaches in *Shewanella* represent a new frontier that considers the sum of biological processes that allow this genre of microbes to thrive in some of the most hostile environments on earth. My findings contribute to both a broad understanding of cellular respiration in *Shewanella* and the mechanistic criteria for arsenic reduction.

Dedications and Acknowledgements

This work is dedicated to my amazingly supportive husband / counselor Eddie Lamar Watson and our gorgeous children Jada Rose-Marie Watson and Elijah Jean-Reese Watson who have made countless sacrifices along with me to realize this dream.

I give special recognition to my Parents Jean and Rose Tilus and Siblings Soledad, Laury, David, Emmanuel, Mardochee, and Washingtonia for their faith in me. Thank you Laury for not accepting any excuses. Special thanks to my nieces and nephews Stanley, Raymond, Thierry, Emmanuela, Kensy, and Jenny-Jean for motivation. Special thanks to my In–laws Maurice and Deidre Childress for their encouragement and my church family at Calvary Chapel Capitola for years of steady love and support throughout this process. God answers prayers.

I acknowledge Professor Chad Saltikov for his leadership and support throughout this dissertation process in mentoring me through my research project with great consideration for my personal wellbeing and unique career goals. I also acknowledge Susan Strome as a strong example of a distinguished scientist with a vested interest in recruiting and cultivating me and other minority scientist. I acknowledge Needhi Bhalla for serving on my committee and challenging me to think more deeply about the implications of my research findings. I acknowledge the faculty, staff, and students in the MCD and METX departments for taking interest in my success. I acknowledge my past mentors Dr. Chetty, Dr. Johnson, and Dr. Purnell of Savannah State University for zealously encouraging me to pursue an advanced degree in biology. I acknowledge Dr. Zia Isola, Malika Bell, and Yualina Ortega for

their professional development and helping me navigate social aspects of the PhD. I acknowledge Prestina Smith for countless hours of scientific discourse sometimes into the wee hours of the morning and being a great friend. I acknowledge my colleagues Jaime Hernandez, Allison Boren, Jesica, Carolina Reyas, Oscar, Nate, Brendon, Dr. Ramatchandirane, Dr. Ersoy Omeroglu, and Dr. Sultana for maintaining an amiable and cultured lab environment which was always buzzing with interesting conversations about environmental microbes, kids and great food from around the world. I acknowledge the undergraduates who worked with me directly allowing me to practice my mentoring on them and their contributions to my progress: Janey Aiello, Marcelo Quiroz, April VillaReal, and Indra Lazcano.

I cannot thank you all enough!

Chapter 1. Introduction, Background, Chapter Descriptions

Introduction

This work investigates a fundamental question in environmental microbiology. How does a microbe "breathe" a toxic substance? Microbes account for most of the biomass on our planet yet their ability to drive the biogeochemical cycling of toxins in the environment has only been appreciated in the last few decades (Colwell, 2009). It is now well understood that microbes play a unique role in shaping earth's geochemistry by driving the conversion of pollutants such as chromium, uranium, nitrate, or chlorate into less toxic forms (Tiedje, 2002). Most microbes reduce or oxidize pollutants simply to avoid their toxic effects. There are also other more "exotic" microbes that can thrive in contaminated environments by using environmental pollutants as a substrate during cellular respiration. However, the mechanisms used to respire these toxins are by no means fully understood. Developing a clear understanding of the mechanisms required to respire environmental pollutants can help us engineer methods to exploit these mechanisms for environmental decontamination and even alternative energy development. One major roadblock to studying the mechanisms involved in the bio-transformation of toxins is finding a suitable model organism as most microbes with the capability are difficult to grow and manipulate in a laboratory setting. Our lab studies *Shewanella* because several *Shewanella* sp. can respire pollutants and *Shewanella* are easy to manipulate in lab. In fact *Shewanella* sp. ANA-3 can actually respire arsenic reducing it to a more toxic and water-soluble form (Figure 1.1) Arsenic respiration only occurs

when more favorable terminal electron acceptors, such as oxygen, are depleted. In near neutral ground water aquifers respiratory arsenic reduction increases the propensity for arsenic to seep into ground water supplies (Figure 1.2). However, arsenic contaminated ground water is difficult to detect (Smedley and Kinniburgh, 2002, Barringer and Reilly, 2013). Therefore, it is important to understand the microbial mechanisms that allow arsenic respiration to proceed in order to implement new detection methods. ANA-3 is currently the only genetically tractable model developed to study respiratory arsenic reduction. Understand the mechanism ANA-3 uses to respire arsenic could help inform future efforts to for both decontamination and preventing human exposure. Few studies of cellular respiration specifically in ANA-3 have been published in the last several years. However other Shewanella have been shown to have similar mechanisms to regulate respiration (Murphy and Saltikov, 2009 Murphy et al., 2009, Gralnick et al., 2005, Charania et al., 2009). We can use the collective body of knowledge of all *Shewanella* spp. to inform our investigations of respiratory arsenic reduction in ANA-3.

The field of environmental microbiology has made great strides in studying cellular respiration in the genus *Shewanella*. *Shewanella* have been isolated from stratified soil, marine and fresh water environments from around the world (Nealson and Scott, 2006, Hau and Gralnik, 2007). These stratified environments have redox conditions, such as concentrations of dissolved oxygen and useable carbon sources, that vary by depth and can change quickly based on geochemical and microbial activity. *Shewanella* sp. ANA-3 in particular has been developed into a model for

studying arsenic reduction (Saltikov et al., 2003). ANA-3, isolated from arsenic treated wood pier pilings, is able to respire arsenic under anaerobic conditions and is genetically tractable. ANA-3 uses arsenic as a terminal electron acceptor in its electron transport chain when oxygen becomes depleted, which leads to the reduction of pentavalent arsenate (oxidized) into its trivalent form arsenite (reduced). Arsenite is a more potent toxin and has a greater propensity to contaminate ground water than the oxidized form arsenate. In ANA-3, the ArrAB protein complex is responsible for arsenate respiratory reduction. Moreover, the *arrAB* genes encoding this arsenic reductase have been identified and thoroughly investigated (Saltikov et al., 2003, Malasaran et al., 2008, Murphy and Saltikov, 2007, Murphy and Saltikov, 2008). However, there are still some open-ended questions about the regulatory signals that trigger the production of ArrAB (Murphy et al., 2009, Charania et al., 2009). Identifying how *arrAB* is transcriptionally regulated will contribute to the knowledge of arsenic metabolism. This information will help predict when microbes will reduce arsenic in the environment and possibly translate to less amenable environmental microbes that respire other environmental pollutants (Charania et al., 2009). To address the question of how the arsenate reduction pathway is activated I used a systems biology approach.

Systems biology approaches further our understanding of arsenic respiration by looking holistically at all of the genetic, biochemical, and environmental cues that work in concert to promote arsenic respiration. By evaluating expression of genes within a species

during arsenic respiration and comparing genomes from different arsenic respiring microbes, and the literature on cellular respiration in *Shewanella* I am able to piece together the network that drives arsenic metabolism. We have learned a great deal about cellular respiration in the genus *Shewanella* from other recent systems level studies.

Overview of Recent findings in Shewanella's Metabolism Using Systems Biology

Recent advances in the study of cellular respiration in the genus *Shewanella* using systems biology

As a genus, *Shewanella* is most famous for its collective ability to use a diverse array of metabolic substrates (Nealson and Scott, 2006). Notably, this includes respiring several substrates that are pollutants, toxins, or are otherwise unsuitable for many other life forms (Myers and Nealson, 1988, Saltikov and Newman, 2003, Sheng and Fein, 2014). One hallmark of *Shewanella* is the huge array of *c*-type cytochrome genes that are involved in electron transfer to this diverse set of metabolic substrates. A *c*-type cytochrome is easily identified by the common hemeC binding motif, CXXCH, where X is any amino acid, and an addition distal H residue (Thöny-Meyer et al., 1995). Moreover, the heme cofactor is usually covalently bound to the apo-protein. It is believed that unique c-type cytochromes containing multiple heme cofactors (some with as many as ten hemes) permit respiration of uncommon substrates such as metal oxides of Mn(IV) and Fe(III). In support of this notion, researchers identified nearly 8 times more cytochrome domain containing genes in the genome of *S. oneidensis*, which respires multiple substrates, compared to *E. coli* which has a much more limited respiratory capacity (Meyer et al., 2004, Myers and

Myers, 2007, Murphy and Saltikov, 2007). Taken together, these and other studies show that each species of *Shewanella* has a unique arsenal of terminal reductases that allows it to thrive on the available substrates produced in its environmental niche. In addition to respiratory abilities, *Shewanella* can detoxify the cell toxic byproducts of these reactions from the cell. Because detoxification is also an important limiting factor in respiring environmental pollutants, mechanisms that regulate detoxification pathways have also been a subject of research in *Shewanella*.

Shewanella's ability to biologically transform and detoxify a diverse array of environmental substrates makes it a favorable model for future alternative energy and bioremediation applications. Sequencing of nearly thirty of the *Shewanella* spp. genomes in the last decade has lead to a recent rise in systems level studies of cellular respiration in *Shewanella*. These studies include, microarray, deep sequencing, proteomic and metabolomics experiments that have all enriched our understanding of cellular respiration in *Shewanella*. The pioneering systems level studies of the genus *Shewanella* were well reviewed in Nealson and Scott in 2006 and again by Hau and Gralnick 2007 and Fredrickson et al., 2008. These writings highlight the advances that underline our understanding of cellular respiration in *Shewanella* over the previous decade. However, several large-scale studies involving both respiration and toxin resistance have been published since then. Next, I will highlight some of the more recent findings that employed systems biological approaches.

Comparative genomic analysis of central metabolism

A versatile respiration gives *Shewanella* the advantage of being able to thrive in environments where the availability of oxygen and other respiratory substrates is constantly changing; *Shewanella* continue to grow when more favorable conditions are limited and less metabolically diverse organisms may die off. In the recent literature much attention has been given to identifying the systems that allow *Shewanella* to have such a diverse metabolism and how these genes and enzymes are regulated.

Each *Shewanella* species has some unique metabolic abilities from other *Shewanella*. When comparing their genomes one can identify many of the gene regulons that allow each strain to thrive in its unique niche. Several *in silico* studies map out the common features between strains as well as the unique traits of each species (Rodinov et al., 2010, Karpinets et al.,2010).

Other studies have employed comparative genomic techniques to deduce which genes may be important for specific cellular respiration in *Shewanella* (Deutshbauer et al., 2011, Rodrigues et al., 2011). One such study compares phenotypic arrays with the genomic differences between five *Shewanella* spp. isolated from diverse environments (Rodrigues et al., 2011). They found that very few carbon-containing compounds were used by all five *Shewanella* strains. They also analyzed each strain's ability to metabolize nitrogen, phosphorus and sulfur containing compounds. Using the available genome sequences the authors were able to identify genes that were responsible for several of these metabolic phenotypes. In summary, they found that evolutionary distance did produce greater phenotypic differences in nitrogen source

utilization but not carbon source utilization. Overall this work highlighted some of the phenotypic traits that allow *Shewanella* to thrive in each environment. Another recent study compared the genome of *S. putrefaciens* W3-18-1 to that of *S. oneidensis* to determine what makes their metabolic capabilities so different (Qiu et al., 2013). They found that W3-18-1 lacked several of the *c*-type cytochrome genes found in *S. oneidensis* and hypothesized that the lack of certain *c*-type cytochromes is linked to the W3-18-1's inability to respire several of the terminal electron acceptors that *S. oneidensis* can. The authors also found three terminal oxidases, CcoNOQP, COX, and CydAB, which were common between *S. oneidensis* and W3-18-1. We also found these three oxidases in our transcriptomic data set from ANA-3. These oxidases may be present in other *Shewanella* as well (see Chapter 2). Each of these terminal oxidases has a different affinity for oxygen and is thought to help the cell cope with changing redox conditions in the environment (Zhou et al., 2013). Our transcriptomic study suggests that these "aerobic" cytochromes may also be playing a role in metabolism during anaerobic conditions as they are highly expressed in anaerobic conditions as well.

Phenotypic arrays can also be used to compare mutant with wild type to identify regulons that are important for specific metabolic processes such as coping with the stress caused by changes in redox conditions. A phenotypic array of EnvZ and OmpR mutants in *S. onidensis* was used to determine how *Shewanella* copes with osmotic stress (Yuan et al., 2011). The authors tested mutants of the EnvZ/OmpR two component system, which is known to regulate expression of genes involved in

osmotic stress response. The authors also used a phenotypic array (PM9 and PM10) to determine if the *S. oneidensis* mutants showed similar phenotypes to *E. coli*. They found that unlike *E.coli* the *S. oneidensis* mutants did not have much reaction to changes in carbon source but were sensitive to alkaline stress. Previous studies involving sensory kinase/response regulator systems like EnvZ/OmpR usually measure phosphorylation as an indicator for activity. The authors rationalized studying the transcriptional profile of this two-component regulon because an increase in transcription may be an additional strategy for the cell to insure that enough signaling molecules will be available to induce the response regulator under stress. The authors found that in fact EnvZ/OmpR were up-regulated during osmotic stress but EnvZ/OmpR mutants had no phenotypic consequence indicating that this is either not the actual osmotic stress coping mechanism or there is another redundant mechanism used to overcome these mutations. Highly expressed genes in metabolic and stress response pathways with no phenotypic consequence may be a recurring theme in *Shewanella* as we also noticed several highly expressed genes with no phenotype during arsenic reduction in *Shewanella* sp. ANA-3 (see Chapter 2).

Deep-sequencing studies have become a new frontier in studying respiration and stress response in *Shewanella*. Three recent studies focus on identifying the genetic factors that are important for aerobic vs. anaerobic respiration (Brutinel and Gralnick., 2012, Clark et al., 2014, and Watson et al., in preparation). Studying the difference between aerobic and anaerobic respiration may be the key to understanding the regulation of many pathways that are important for dissimilatory respiration of

environmental pollutants, which is only active anaerobically (Beliaev et al., 2005, Gau et al., 2010). In a recent study by Brutinel and Gralnick (2012) the classical transposon mutagenesis technique was combined with high-throughput sequencing to identify mutations that render *Shewanella oneidensis* more or less fit for growth during aerobic or anaerobic conditions. The authors found several genes in the TCA cycle normally essential for aerobic respiration are not required during anaerobic respiration. The difference in the phenotype between transposon mutants grown aerobically and anaerobically, such as citrate synthase or fumarate reductase mutants, was attributed to an alternate anaerobic TCA cycle that bypasses the need for these steps during anaerobic respiration. Transposon Seq linked the differences between aerobic and anaerobic respiration directly to many of the genes that are involved in making *S. oneidensis* fit in different redox environments.

Aerobic versus anaerobic respiration using high-throughput sequencing

Clark et al. (2014) took several systems level approaches to study aerobic and anaerobic respiration in *S. algae* strain ACDC, a newly isolated chlorate reducer (Clark et al., 2014). Through transcriptomics they identified several components of central metabolism that were similar during low oxygen growth and growth on chlorate. They also found that much of the electron transport chain is not differentially expressed in aerobic conditions when compared to chlorate reducing conditions; this is not surprising since oxygen is a byproduct of chlorate reduction. This is not the case for other anaerobic terminal reductases such as nitrate reductase, which had several uniquely expressed genes in its electron transport chain. Of the

genes that were uniquely expressed in chlorate respiring conditions, most were involved in oxidative stress response or disulfide bond repair. In nitrate respiring conditions there was also a specific stress response to nitric oxide stress. The transcriptional regulators Fnr and NsrR also seem to play a role in the chlorate and nitrate regulons respectively.

In our transcriptomic study of *Shewanella* sp. ANA-3 we found hundreds of genes that were differentially expressed between aerobic and anaerobic conditions. However there was much less variation when comparing transcription between arsenic and fumarate, the two anaerobic conditions. Like the *S. algae* str. ACDC study, we found that stress response was a major part of the anaerobic transcriptional profile under arsenic respiring conditions with many genes in the sulfur assimilation pathway being highly transcribed during anaerobic respiration possibly due to the greater need for sulfur containing proteins.. The sulfur assimilation and scavenging pathways were even more highly expressed in arsenic respiring conditions relative to anaerobic respiration on fumarate. This may be another stress response mechanism as arsenic like chlorate, tends to interact with sulfur containing proteins (Shen et al., 2013). Moreover, cellular reductants like glutathione are often used by Gram negative bacteria for the detoxifying arsenate reductase ArsC. (More discussion on the transcriptional regulation of arsenic respiration in Chapter 3)

Microarray based studies of respiration

Microarray studies are the most widely used systems level tool to study cellular respiration in *Shewanella* and continue to highlight new findings. Binnenkade, Lassak

and Thormann (2011) used microarray analysis to study the global effects of a sensory kinase/response regulator system BarA/UvrY on cellular respiration in *Shewanella*. Studies in *E. coli* and *Pseudomonas* show that BarA/UvrY is a global regulator but it was not clear if this two-component system works as a global regulator in *Shewanella or* if this system played a direct role in central metabolism in *Shewanella. Shewanella* regulate cellular respiration much differently then other Gama-proteobacteria. Instead of using the global regulators FNR or ARC to sense aerobic/anaerobic conditions and regulate gene expression accordingly, *Shewanella* use the cyclic AMP receptor protein (CRP) to activate anaerobic respiration (Murphy et al., 2009, Charania et al.,2009). CRP does not sense anaerobic conditions and therefore may be working in concert with a yet unidentified respiratory sensor. The authors hypothesized that the BarA/UvrY system may sense and regulate cellular respiration. In this study, homologs to the *E. coli bar*/*uvrY* genes were identified in *S. oneidensis*. Similar homologs are found in all sequenced *Shewanella species*. The authors showed that an *uvrY* mutant in *S. oneidensis* does have a global transcriptomic effect and likely plays an important role in anaerobic respiration. Upon further analysis the authors found that several of the genes regulated by UvrY were involved in central metabolism. Interestingly, the mutants grew faster than wild type aerobically in defined media, with a N-acetylglucosamine (NAG) carbon source, however their growth was suppressed anaerobically with the same media. This study yet again highlights the vast differences between aerobic and anaerobic respiration in *Shewanella* and identified a sensory regulator that may help the cell sense aerobic vs.

anaerobic conditions.

Rosenbaum et al. (2012) also used microarrays to study transcriptional differences between aerobic and anaerobic respiration on iron or an electrode (Rosenbaum et al., 2012). Although their data sets were very different for the three experimental conditions they were still able to identify some meaningful transcriptional trends between each form of respiration. They found that several genes that are known to be important for anaerobic respiration in *Shewanella cymA, mtrA, mtrB, mtrC/omcB, and omcA* were up in anaerobic conditions relative to oxygen as in previous studies (Reyes et al 2010). **Stress response due to environmental conditions**

Shewanella sp. ANA-3's diverse metabolic capabilities pose a unique problem for the cell. For example, the reduced form of arsenic (arsenite) is more cytotoxic than the oxidized form (arsenate). Other metal-reducing *Shewanella* sp. also have to deal with this problem as well because reduction of Fe(III) or sulfur containing minerals can also cause the release of other toxic metals or organic pollutants that are associated with these substrates. Even nitrate reduction can cause nitrite and nitric oxide to accumulate. Chlorate and other substrates can cause oxidative stress in their reduced form as well. Several of the recent systems level studies identify mechanisms that can help the cell cope with the stress associated with these metabolisms. Using microarrays McLean et al. (2008), found that growth in oxygen rich environments increases the expression of genes that promote auto-aggregation in order to block free radicals from entering the cell along with several genes that encode efflux pumps, superoxide dismutase and DNA damage repair. Another study of *S. onidensis* found

that oxidative stress caused by H_2O_2 was handled by the transcriptional regulator OxyR, which activates the expression of genes encoding H_2O_2 scavenging and disulfide reduction enzymes (Jiang et al., 2014. Interestingly they found that OxyR did not bind DNA in the canonical fashion that it does in *E. coli*, *Pseudomonas aeruginosa*, *Erwinia carotovora*, and *Vibrio cholera.* OxyR could not bind many of the genes identified in its regulons directly. There is likely another general transcription factor that helps activate oxidative stress response in *Shewanella*. In *S. algae* str. ACDC, genes for sulfur assimilation or iron sulfur bond repair were specifically expressed in the anaerobic conditions (Clark et al., 2014). In our study we found several genes involved in the sulfur assimilation pathway to be expressed anaerobically and even more so in arsenate-respiring conditions. Our follow-up experiments showed that these highly expressed genes were also necessary for dealing with the toxicity of arsenite which is not surprising considering this form of arsenic is known to have a strong affinity for sulfur and ability to bind sulfur containing proteins. Perhaps expression of sulfur assimilation genes has dual roles during anaerobic respiration: ensure enough cytochromes are made to scavenge any toxic byproducts. Taken together it is evident that *Shewanella*'s diverse metabolic abilities rely not only on common central metabolic pathways but also on specific stress response mechanisms for each substrate.

Chapter descriptions

The second chapter describes the unique transcriptional profile of ANA-3 during dissimilatory arsenic reduction compared to respiration on fumarate or

oxygen. This work includes the analysis of RNAseq experiments and a discussion about their interpretation in the context of the current literature. The RNAseq study also provides several new hypotheses about the requirements for arsenic respiration and detoxification. Chapter 2 also includes the results of several reverse genetic experiments using targets identified by RNAseq. In addition, the RNAseq study validates our current knowledge on the expression of the arsenic island identified in Saltikov and Newman (2003), Saltikov et al. (2003), and Saltikov et al. (2005). Lastly, several new genes that were not previously known to be involved in arsenic respiration were identified in the RNAseq experiments.

In Chapter 3, I describe a study were I investigated the role of the global signaling molecule cAMP in regulating the switch between aerobic and anaerobic respiration. Also, Chapter 3 highlights the role of cAMP in activating arsenic reduction and identifies the minimum amount of cAMP required for anaerobic growth. Finally, I provide a measurement of baseline levels of cAMP during aerobic and anaerobic respiration.

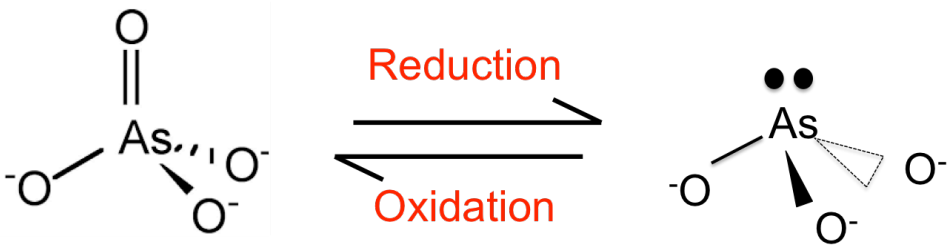
Lastly, in Chapter 4, I describe my work on the isolation of a "novel" arsenic respiring microbe from a naturally occurring arsenic rich environment. I describe the methodologies used to isolate and characterize the organism as well as a comparison of the arsenic respiratory reductase of this novel organism to that of ANA-3 and compared to other arsenic respiring microbes.

Taken together the methodologies, results and discussions presented in this dissertation take a holistic approach to identify several new avenues of research in the

regulation of respiratory arsenic reduction. Many past studies have focused on the regulation of genes within the arsenic island of ANA-3 directly however an unbiased genome-wide approach using RNAseq identified several genes that were not previously implicated in arsenic reduction opening several new avenues of research. Several past studies also implicate the cyclic AMP receptor protein and cyclic AMP in triggering anaerobic respiration; for the first time the role of cAMP on gene expression and the actual levels of cAMP during aerobic and anaerobic respiration in *Shewanella* are quantified in this study. Currently ANA-3, isolated from arsenic treated wood, is the only genetic model for studying arsenic respiration. My studies use microbial ecology paired with molecular biology techniques to isolate and characterize a novel organism that respires arsenic. The new organism may possibly respire arsenic through a unique regulatory mechanism, which would provide another model for studying arsenic respiration in microbes from a naturally arsenic rich ecosystem.

Chapter 1 figures

Figure 1.1 Properties of arsenate and arsenite.
Arsenate and arsenite are the most common naturally occurring forms of arsenic. Each has a different molecular structure, mechanisms of toxicity and solubility



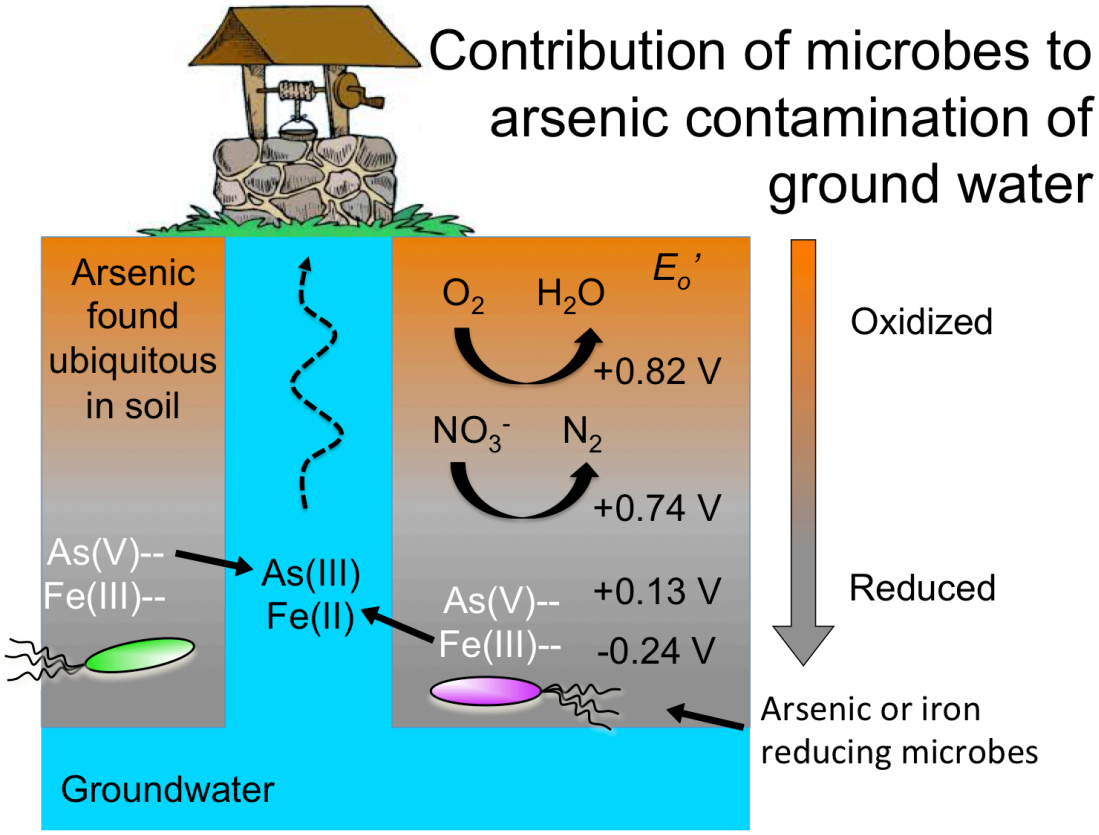
Arsenate As(V)

- Mimics PO_4^{3-}
- Bound to iron and other compounds at near neutral pH

Arsenite As(III)

- Strong electrophile
- More water soluble at near neutral pH

Figure 1.2 Cartoon of microbial mediated arsenic contamination of groundwater.
Respiratory arsenic (or iron) reduction by bacteria in oxygen depleted environments leads to arsenic reduction. More oxidized soil shown in orange and more reduced soil shown in grey. The standard reduction potential of each available terminal electron acceptor is shown on the right.



Chapter 1 References

Malasarn, D., Saltikov, C. W., Campbell, K. M., Santini, J. M., Hering, J. G., & Newman, D. K. (2004). arrA is a reliable marker for As(V) respiration. *Science (New York, N.Y.)*, *306*(5695), 455–455. http://doi.org/10.1126/science.1102374

Beliaev, A. S., Klingeman, D. M., Klappenbach, J. A., Wu, L., Romine, M. F., Tiedje, J. M., et al. (2005). Global transcriptome analysis of Shewanella oneidensis MR-1 exposed to different terminal electron acceptors. *Journal of Bacteriology*, *187*(20), 7138–7145. http://doi.org/10.1128/JB.187.20.7138-7145.2005

Binnenkade, L., Lassak, J., & Thormann, K. M. (2011). Analysis of the BarA/UvrY two-component system in Shewanella oneidensis MR-1. *PLoS ONE*, *6*(9), e23440–e23440. http://doi.org/10.1371/journal.pone.0023440

Brutinel, E. D., & Gralnick, J. A. (2012). Anomalies of the anaerobic tricarboxylic acid cycle in Shewanella oneidensisrevealed by Tn-seq. *Molecular Microbiology*, *86*(2), 273–283. http://doi.org/10.1111/j.1365-2958.2012.08196.x

Charania, M. A., Brockman, K. L., Zhang, Y., Banerjee, A., Pinchuk, G. E., Fredrickson, J. K., et al. (2009). Involvement of a membrane-bound class III adenylate cyclase in regulation of anaerobic respiration in Shewanella oneidensis MR-1. *Journal of Bacteriology*, *191*(13), 4298–4306. http://doi.org/10.1128/JB.01829-08

Clark, I. C., Melnyk, R. A., Iavarone, A. T., Novichkov, P. S., & Coates, J. D. (2014a). Chlorate reduction in Shewanella algae ACDC is a recently acquired metabolism characterized by gene loss, suboptimal regulation and oxidative stress. *Molecular Microbiology*, *94*(1), 107–125. http://doi.org/10.1111/mmi.12746

Cleiss-Arnold, J., Koechler, S., Proux, C., Fardeau, M.-L., Dillies, M.-A., Coppee, J.-Y., et al. (2010). Temporal transcriptomic response during arsenic stress in Herminiimonas arsenicoxydans. *BMC Genomics*, *11*(1), 709–709. http://doi.org/10.1186/1471-2164-11-709

Conservation of Transcription Start Sites within Genes across a Bacterial Genus. (2014). Conservation of Transcription Start Sites within Genes across a Bacterial Genus, 1–14. http://doi.org/10.1128/mBio.01398-14
Construction and characterization of regulated. (1999). Construction and characterization of regulated, 1–7.

Deutschbauer, A., Price, M. N., Wetmore, K. M., Shao, W., Baumohl, J. K., Xu, Z., et al. (2011). Evidence-based annotation of gene function in Shewanella oneidensis

MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genetics*, *7*(11), e1002385–e1002385. http://doi.org/10.1371/journal.pgen.1002385

Ehlert, K., Mikutta, C., & Kretzschmar, R. (2014). Impact of Birnessite on Arsenic and Iron Speciation during Microbial Reduction of Arsenic-Bearing Ferrihydrite. *Environmental Science & Technology*, *48*(19), 11320–11329. http://doi.org/10.1021/es5031323

Fredrickson, J. K., Romine, M. F., Beliaev, A. S., Auchtung, J. M., Driscoll, M. E., Gardner, T. S., Nealson, K. H., Osterman, A. L., Pinchuk, G., Reed, J. L., Rodionov, D. A., Rodrigues, J. L. M., Saffarini, D. A., Serres, M. H., Spormann, A. M., Zhulin, I. B., & Tiedje, J. M. (2008a). Towards environmental systems biology of Shewanella. *Nature Reviews Microbiology*, *6*(8), 592–603. http://doi.org/10.1038/nrmicro1947

Gralnick, J. A., Brown, C. T., & Newman, D. K. (2005b). Anaerobic regulation by an atypical Arc system in Shewanella oneidensis. *Molecular Microbiology*, *56*(5), 1347–1357. http://doi.org/10.1111/j.1365-2958.2005.04628.x

Hau, H. H., & Gralnick, J. A. (2007). Ecology and biotechnology of the genus Shewanella. *Annual Review of Microbiology*, *61*, 237–258. http://doi.org/10.1146/annurev.micro.61.080706.093257

Jiang, Y., Dong, Y., Luo, Q., Li, N., Wu, G., & Gao, H. (2014). Protection from oxidative stress relies mainly on derepression of OxyR-dependent KatB and Dps in Shewanella oneidensis. *Journal of Bacteriology*, *196*(2), 445–458. http://doi.org/10.1128/JB.01077-13

Karpinets, T. V., Romine, M. F., Schmoyer, D. D., Kora, G. H., Syed, M. H., Leuze, M. R., et al. (2010). Shewanella knowledgebase: integration of the experimental data and computational predictions suggests a biological role for transcription of intergenic regions. *Database*, *2010*, baq012–baq012. http://doi.org/10.1093/database/baq012

Malasarn, D., Keeffe, J. R., & Newman, D. K. (2008). Characterization of the arsenate respiratory reductase from Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *190*(1), 135–142. http://doi.org/10.1128/JB.01110-07

McLean, J. S., Pinchuk, G. E., Geydebrekht, O. V., Bilskis, C. L., Zakrajsek, B. A., Hill, E. A., et al. (2008). Oxygen-dependent autoaggregation in Shewanella oneidensis MR-1. *Environmental Microbiology*, *10*(7), 1861–1876. http://doi.org/10.1111/j.1462-2920.2008.01608.x

Murphy, J. N., & Saltikov, C. W. (2007). The cymA gene, encoding a tetraheme c-type cytochrome, is required for arsenate respiration in Shewanella species. *Journal of Bacteriology*, *189*(6), 2283–2290. http://doi.org/10.1128/JB.01698-06

Murphy, J. N., & Saltikov, C. W. (2009). The ArsR repressor mediates arsenite-dependent regulation of arsenate respiration and detoxification operons of Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *191*(21), 6722–6731. http://doi.org/10.1128/JB.00801-09

Murphy, J. N., Durbin, K. J., & Saltikov, C. W. (2009). Functional roles of arcA, etrA, cyclic AMP (cAMP)-cAMP receptor protein, and cya in the arsenate respiration pathway in Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *191*(3), 1035–1043. http://doi.org/10.1128/JB.01293-08

Myers, C. R., & Myers, J. M. (1997). Cloning and sequence of cymA, a gene encoding a tetraheme cytochrome c required for reduction of iron(III), fumarate, and nitrate by Shewanella putrefaciens MR-1. *Journal of Bacteriology*, *179*(4), 1143–1152.

Myers, C. R., & Nealson, K. H. (1988). Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor. *Science (New York, N.Y.)*, *240*(4857), 1319–1321. http://doi.org/10.1126/science.240.4857.1319

Nealson, K. H., & Scott, J. (2006). Ecophysiology of the Genus Shewanella. In *The Prokaryotes* (pp. 1133–1151). Springer New York. http://doi.org/10.1007/0-387-30746-X_45

Qiu, D., Wei, H., Tu, Q., Yang, Y., Xie, M., Chen, J., et al. (2013). Combined genomics and experimental analyses of respiratory characteristics of Shewanella putrefaciens W3-18-1. *Applied and Environmental Microbiology*, *79*(17), 5250–5257. http://doi.org/10.1128/AEM.00619-13

Reyes, C., Murphy, J. N., & Saltikov, C. W. (2010). Mutational and gene expression analysis of mtrDEF, omcA and mtrCAB during arsenate and iron reduction in Shewanella sp. ANA-3. *Environmental Microbiology*, *12*(7), 1878–1888. http://doi.org/10.1111/j.1462-2920.2010.02192.x

Rodionov, D. A., Yang, C., Li, X., Rodionova, I. A., Wang, Y., Obraztsova, A. Y., et al. (2010). Genomic encyclopedia of sugar utilization pathways in the Shewanella genus. *BMC Genomics*, *11*(1), 494–494. http://doi.org/10.1186/1471-2164-11-494

Rodrigues, J. L. M., Serres, M. H., & Tiedje, J. M. (2011). Large-scale comparative phenotypic and genomic analyses reveal ecological preferences of shewanella species and identify metabolic pathways conserved at the genus level. *Applied*

and Environmental Microbiology, *77*(15), 5352–5360. http://doi.org/10.1128/AEM.00097-11

Saltikov, C. W., & Newman, D. K. (2003). Genetic identification of a respiratory arsenate reductase. *Pnas*, *100*(19), 10983–10988. http://doi.org/10.1073/pnas.1834303100

Saltikov, C. W., Cifuentes, A., Venkateswaran, K., & Newman, D. K. (2003). The ars detoxification system is advantageous but not required for As(V) respiration by the genetically tractable Shewanella species strain ANA-3. *Applied and Environmental Microbiology*, *69*(5), 2800–2809. http://doi.org/10.1128/AEM.69.5.2800-2809.2003

Saltikov, C. W., Wildman, R. A., & Newman, D. K. (2005). Expression dynamics of arsenic respiration and detoxification in Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *187*(21), 7390–7396. http://doi.org/10.1128/JB.187.21.7390-7396.2005

Sheng, L., & Fein, J. B. (2014). Uranium reduction by Shewanella oneidensis MR-1 as a function of NaHCO3 concentration: surface complexation control of reduction kinetics. *Environmental Science & Technology*, *48*(7), 3768–3775. http://doi.org/10.1021/es5003692
Sigma Factors. (2013). *Sigma Factors*.

Smedley, P. L., & Kinniburgh, D. G. (2002). A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry*, *17*(5), 517–568. http://doi.org/10.1016/S0883-2927(02)00018-5

Soria, S., de Anda, R., Flores, N., Romero-Garcia, S., Gosset, G., Bolívar, F., & Báez-Viveros, J. L. (2012). New insights on transcriptional responses of genes involved in carbon central metabolism, respiration and fermentation to low ATP levels in Escherichia coli. *Journal of Basic Microbiology*, *53*(4), 365–380. http://doi.org/10.1002/jobm.201100525

Thöny-Meyer, L., Fischer, F., Künzler, P., Ritz, D., & Hennecke, H. (1995). Escherichia coli genes required for cytochrome c maturation. *Journal of Bacteriology*, *177*(15), 4321–4326.

Tiedje, J. M. (2002). Shewanella--the environmentally versatile genome. *Nature Biotechnology*, *20*(11), 1093–1094. http://doi.org/10.1038/nbt1102-1093

Yuan, J., Wei, B., Shi, M., & Gao, H. (2011). Functional assessment of EnvZ/OmpR two-component system in Shewanella oneidensis. *PLoS ONE*, *6*(8), e23701–

e23701. http://doi.org/10.1371/journal.pone.0023701

Zhou, G., Yin, J., Chen, H., Hua, Y., Sun, L., & Gao, H. (2013). Combined effect of loss of the caa3 oxidase and Crp regulation drives Shewanella to thrive in redox-stratified environments. *ISME Journal*, *7*(9), 1752–1763. http://doi.org/10.1038/ismej.2013.62

Chapter 2. Systems level study of respiration in the arsenate reducing bacterium Shewanella sp. strain ANA-3

R. Pamela Watson[1], Indra Lazcano[1] Jaime Hernandez [2], David Bernick[3], Patricia Chen[4], and Chad W. Saltikov[2*]

1.Molecular, Cellular and Developmental Biology
University of California, Santa Cruz
Santa Cruz, CA 95064

2.Microbiology and Environmental Toxicology Department
University of California, Santa Cruz
Santa Cruz, CA 95064

3.Biomolecular Engineering Department
University of California, Santa Cruz
Santa Cruz, CA 95064

4.Maverix Biomics
1875 South Grant Street, Suite 700
San Mateo, CA 94402

***Corresponding author:**
(tel.) 831-459-5520, (fax) 831-459-3524
(email) saltikov@ucsc.edu
Running Head: Cellular respiration of *Shewanella*

Abstract

Dissimilatory metal reducing bacteria play a unique role in landscaping the earth's geochemistry by donating terminal electrons to the environment leading to metal transformation. Understanding how these microbes can alter the chemical state of a nearby metal has implications for both bioremediation and fuel cell studies. However, in order to exploit these microbes' full potential for biotechnology, it is necessary to have a grasp of the biological processes that lead to metal reduction. For this purpose, *Shewanella* was developed as a model for its genetic tractability and diverse metabolic capability, including the ability to reduce metals during anaerobic respiration. It is now well understood that *Shewanella* does not rely solely on the common anaerobic sensors *fnr* or *arcA* to regulate metal reduction. However, exactly how *Shewanella* regulates anaerobic respiration has yet to be elucidated. In this study we used a transcriptomic approach to identify genes that have a potential role in general anaerobic respiration and, more specifically, respiration of the toxic metal arsenic. We found more than 10% of the genome is differentially regulated between aerobic and anaerobic conditions and a subset of these genes are specifically regulated under arsenic respiring conditions. This study represents progress towards understanding the regulatory networks that govern respiration in this environmentally relevant genus.

Introduction

Members of the *Shewanella* genus are facultative anaerobes best know as "respiratory specialists" because of their diverse anaerobic metabolic capabilities (Nealson and Scott, 2006). The diverse nature of *Shewanella's* metabolic capacity has led to the isolation of numerous species from a variety of diverse environments such as anoxic sediments, various cold deep-sea environments, near hydrothermal vents in the iron-rich Lohi seamount (Hawaii), lakes, rivers, and estuaries (summarized in (Hau and Gralnick, 2007)). The ability to respire solid electron acceptors like oxides of Fe(III) and Mn(IV) in addition to their genetic tractability have propelled *Shewanella* as a favored model organism used in biotechnology, alternative energy, and bioremediation applications. Several *Shewanella* strains are also known for their ability to reduce radionuclides of uranium and technetium, toxic heavy metals and metalloids such as cobalt and chromate (Hau and Gralnick, 2007) and arsenate (Saltikov and Newman, 2003).

Arsenate respiration is limited to organisms that contain *arrAB* gene homolog (Malasarn, 2004). The genes encoded by *arrA* and *arrB* include, a Mo-containing arsenate reductase and an iron sulfur protein subunit, respectively. Dissimilatory arsenate reduction has been implicated in processes that can lead to high concentrations of arsenic in groundwater. *Shewanella* sp. ANA-3 along with a few other *Shewanella* sp. have the arrAB genes. *Shewanella* sp. ANA-3 in particular, has served as a model organism to investigate the environmental cycling of arsenic. Through laboratory based experiments, arsenic mobilization from solid phase Fe(III) minerals is impeded in *Shewanella* that lack the *arrA* gene (Campbell, et al., 2006;

Tufano, et al., 2008). Moreover, *arr* containing *Shewanella* strains can also mobilize arsenic from natural sediments collected from arsenic-containing Bangladesh aquifers sand (Dhar, et al., 2011). The arsenate reduction activities of *arrAB* containing *Shewanella* are expressed mainly under anoxic conditions and further induced in the presence of arsenic. Although *Shewanella* is not the likely microbe catalyzing arsenic mobilization within groundwater, microbial ecology studies have demonstrated the occurrence and/or expression of *arrA* genes within environments that have arsenic problems (Malasarn et al., 2004; Kulp et al., 2006; Diversity and abundance of arsenic biotransformation genes in paddy soils from southern china., 2015; Biogeochemical environments of streambed-sediment pore waters with and without arsenic enrichment in a sedimentary rock terrain, New Jersey Piedmont, USA., 2015; Microbial ecology of arsenic-mobilizing Cambodian sediments: lithological controls uncovered by stable-isotope probing., 2014).

A variety of genetic and systems level biology studies have been done in *Shewanella* to better understand the regulation of anaerobic metabolism and reduction of toxic substrates like chromate, uranium, and chlorate (Bencheikh-Latmani et al., 2005; Clark et al., 2014). Other systems-level studies have probed deeper into the mechanisms of general metabolism and cellular respiration in *Shewanella.* For example, proteomic and transcriptomic studies were used to identify the core genomes of several *Shewanella* (Konstantinidis et al., 2009). Subsequent bioinformatic analyses have identified *Shewanella*'s core regulon as well (Rodionov et al., 2011), and high throughput phenotype arrays have linked many of the core

metabolic activities to the responsible genes (Rodrigues et al., 2011). However the pathway for anaerobic respiration in *Shewanella* has not yet been completely resolved.

Numerous reverse genetic studies have been conducted within the *Shewanella* genus to investigate anaerobic respiration. Contrary to the *E. coli* model of anaerobic respiration, several studies have shown that neither of the more common anaerobic sensors *arcA* or *fnr* are sufficient to guide anaerobic respiration in *Shewanella* (Gralnick et al., 2005) (Murphy et al., 2009). Furthermore, biochemical studies of individual enzymes implicated cAMP and the Cyclic AMP Receptor Protein (CRP) are important for the regulation of anaerobic respiration (Saffarini et al., 2003; Murphy et al., 2009; Charania et al., 2009). However our unpublished observations indicate that exogenous cAMP is not sufficient to induce expression of arsenate reductase in aerobic conditions (R. P. Watson and C. W. Saltikov, unpublished results); thus the sensory trigger leading to anaerobic switch has not yet been identified.

In this study, we conducted a transcriptome investigation in the arsenate-respiring *Shewanella* sp. ANA-3 for the purposes of identifying gene expression patterns unique to anaerobic growth on arsenate. Although microarray approaches have been used in past transcriptome studies in *Shewanella,* RNAseq represents a more sensitive method of identifying the transcriptional profile of the cell without hybridization bias and allows the identification of the full range of transcriptional differences without saturation (Sîrbu et al., 2012). In this study RNAseq combined

with quantitative RT-PCR was used to identify transcriptional patterns involved in aerobic vs. anaerobic metabolism. Bioinformatic analyses confirmed differential expression of genes previously identified under aerobic/anaerobic conditions. In addition to other known arsenic reduction pathways (e.g. *arr* and *ars*), we were also able to identify several "new" genes that may play a role in aerobic/anaerobic metabolisms as well as metabolism of the toxic metalloid arsenate. In addition, our results complement past microarray studies in other *Shewanella* by providing a higher resolution of the *Shewanella* transcriptome for oxygen, fumarate, and arsenate.

Materials and Methods

Cell culture conditions. *Shewanella* sp. Strain ANA-3 was grown aerobically overnight in minimal salts medium (TME) (described in Murphy et al 2009). The culture was diluted in medium at a 0.6 OD_{600nm} (OD, optical density) with a Spectronic 20+ spectrophotometer. One ml of culture per replicate at 1:100 was then used to inoculate 1L triplicate anaerobic cultures containing either 10 mM arsenate or 10 mM fumarate, and triplicate aerobic cultures for a total of 3 separate experimental conditions equaling 9 total cultures. Once the individual replicates reached exponential growth phase (represented by an $OD_{600\ nm}$ of 0.1) the cells were spun down and total RNA extracted using Trizol reagent.

RNA isolation. After Trizol purification 8 µg of RNA from each experiment was treated with Promega RQ DNase1 then purified again using the Qiagen RNeasy RNA

cleanup protocol, according to the manufacture's instructions. RNA purity was verified by analyzing 1 µL of the sample on a 1% agarose gel stained with ethidium bromide

Ribosomal RNA was depleted immediately after purification using the Ribo-Zero™ kit for Gram-negative bacteria. One microgram of purified RNA was treated with the rRNA removal solution for 10 minutes at 68°C followed by 15 minute incubation at 22°C. Each sample was then incubated with the re-suspended Microsphere reagent and RNase inhibitor solution for 10 minutes at 22°C with a final incubation for 10 minutes at 50°C.

RNASeq Library Preparation. Libraries were generated using the Illumina® NEBNext® Ultra Directional RNA library kit. Briefly, each of the nine ribosoml RNA-depleted samples was fragmented, converted to cDNA, then purified with 1.8 X Agencourt® AMPure® XP beads. Illumina sequencing adaptors were ligated onto the ends of the transcripts. The ligated transcripts were size selected using 0.6 X Agencourt AMPure XP beads and enriched with 15 cycles of PCR. Finally the reactions were purified with 1X Agencourt AMPure XP beads and sent to University of California Davis Genome Center for sequencing HI seq TM 2000

High Throughput Sequencing. Prepared libraries were further purified with 0.8X Agencourt AMPure XP beads to minimize un-ligated adaptor contamination. Samples were analyzed on a bioanalyzer with the high sensitivity Agilent DNA chip to ensure

purity. Finally, UC Davis Genome Center normalized each sample to equal amounts, pooled each library, then loaded the samples onto one lane of the illumina HiSeq gene platform reading single end 50 base pair reads.

Data Analysis. Sequencing data was analyzed using RNA-seq Analysis Kit for Bacteria 1.0 by Maverix Biomics. Raw sequencing reads (50 bp single end) generated by Illumina HiSeq were quality checked for potential sequencing issues and contaminants using FastQC ((Anders and Huber, 2010)Andrews, 2012). The reads were trimmed to remove adaptor sequences, primers, Ns, and bases with quality score below 28 at the ends and as an overall average using fastq-mcf of ea-utils (Aronesty, 2013) and PRINSEQ (Schmieder and Edwards, 2011). Reads with a remaining length below 20 bases after trimming were discarded. These trimmed sequences were then mapped to *Shewanella* sp. ANA-3 genome (NCBI ReqSeq accession NC_008577) and megaplasmid (NCBI RefSeq accession NC_08573) using EDGE-pro (Magoc et al., 2013) with Bowtie2 (Langmead and Salzberg, 2012). Estimated read coverage of overlapping genes was done using default read mapping parameters with a window size of 50 bp. Read coverage on forward and reverse strands for genome browser (Chan et al., 2012) visualization was computed using SAMtools (Li et al., 2009) and BEDtools (Quinlan and Hall, 2010). The reads per kilo base per million (RPKM) (Mortazavi et al., 2008) generated by EDGE-pro (Magoc et al., 2013) was used to compare all 9 samples, grouped into three experiments (arsenic vs. fumarate vs. oxygen). A t-test was performed for each pair of conditions among the three

experiments under the 'homogeneous' assumption that all three groups have equal variances. Read counts per annotated gene in NCBI RefSeq computed by EDGE-pro (Magoc et al., 2013) were normalized among samples of each growth condition using DEseq (Anders and Huber, 2010) for differential gene expression analysis. Adjusted p-values and fold change in log2 scale were calculated for each gene. The expression values were filtered with a False Discovery Rate (adjusted p-value) of less than 0.05 with a fold change greater than 2 to identify significantly differentially expressed genes. The expression values were then Log10 transformed to produce scatter diagrams.

Verification of transcription patterns with Quantitate Reverse Transcriptase PCR (q-RT-PCR). Five genes that were up regulated in arsenic conditions, in addition to *recA* (Shewana3_1126), were selected as targets for q-RT-PCR Previously frozen RNA extracted from cells grown in each of the 9 conditions was then DNase treated using ambigon DNase1 then Reverse transcribed using the TaqMan based MultiScribe™ reverse transcriptase. Dual quenched fluorescent probes (IDT, San Jose CA) were designed to produce a 100 bp PCR amplicon within the coding sequence of test each gene. The 5' end of each probe was labeled with FAM; the 3' end was labeled with Ioa Black with an internal ZEN quencher. cDNAs was then diluted 1:4 for use in realtime (quantitative) PCR on an MJ Research Opticon2. Each q-RT-PCR reaction was performed in triplicate with duplicate cDNAs prepared from each biological replicate for a total of 6 PCRs per RNA preparation. qPCR reactions

were standardized using genomic DNA diluted to 0.1 ng/µl, 0.05 ng/µl, and 0.01 ng/µl. No PCR products were observed in RT-PCRs lacking reverse transcriptase verifying the lack of contaminating genomic DNA in the RNA preparations.

Generation of Gene Disruption Mutants. Gene disruption mutants were made as described in Zargar et al., 2010. Briefly, an internal portion of each gene was PCR amplified and inserted into an antibiotic resistance containing plasmid. Each plasmid was then cloned in an *E.coli* DH5α λ pir vector and plated on X-gal containing plates. Clones with the insert were selected using a blue white screen. Then the plasmid was purified and transformed into WM3064. Finally WM3064 was mated with wild type ANA-3. Single recombinant mutant ANA-3 with the gene disruption were picked from plates containing kanamycin and no diaminopelmic acid (DAP).

Growth Experiments of Gene Disruption Mutants. Aerobic cultures were grown overnight in TME medium. The cultures were then diluted in TME until the optical density of each culture was 0.6. The cells were then inoculated into anaerobic bulch tubes containing 10 mM of either arsenate or fumarate and control cultures with no terminal electron acceptor. Growth was measured using a Spectronic 20D at 600nm wavelength.

Arsenite Sensitivity of Gene Disruption Mutants. *Shewanella* sp. Strain ANA-3 disruption mutants were grown aerobically in TME media with no arsenite or

enriched with 5, 2.5, or 1.25 mM As(III). The cells were grown at 30°C on a VersaMax microplate reader for 12 hours while being aerated by shaking for 1 minute each 5 minute interval. Growth was assessed by reading the optical density at 600nm.

Results/Discussion

Experimental Design and Validation of RNAseq Library preparation

The goal of this work was to identify genes in *Shewanella* sp. ANA-3 that are uniquely expressed during arsenate respiration conditions relative to growth on fumarate and oxygen as terminal electron accepters. To achieve our goal we grew *Shewanella* sp. ANA-3 in triplicate cultures to a similar OD_{600nm} of 0.1, which represents mid-log phase of growth and had been used in past gene expression studies in ANA-3 (Murphy and Saltikov, 2009; Murphy et al., 2009; Saltikov et al., 2005). Several quality control measures were taken to verify the fidelity of each sequenced cDNA library. First, adaptors and ambiguous reads were trimmed out leaving between 84-92% of high quality sequencing reads with PHRED scores > 28. From these trimmed/filtered reads, > 3 million reads remained per library. Unique reads were then mapped to the genome (Supplementary Table 1). Mapped read counts were then normalized to avoid gene length bias using reads per kilobase per million (RPKM) as described in Mortazavi *et al.* (2008). Once we verified the fidelity of the RNAseq libraries using principal component analysis (Supplementary Figure 1), we tested to see each replicate library group together based on the experimental conditions from which the library was derived as in Yuan et. al, 2012. The six

anaerobic RNAseq libraries grouped closer to each other than to the three aerobic replicates further validating the biological significance of each data set. One of the arsenate-specific replicates, however, did not group as closely with the other two corresponding replicates (Supplementary Figure 1). This was likely caused by a less effective ribosomal RNA depletion step relative to the other replicates as shown in the reads mapping distribution profile (Supplementary Figure 2). Overall, each replicate grouped closer with other replicates from the same condition than to replicates from different conditions as expected.

Validation of RNAseq transcriptional patterns with quantitative reverse transcriptase PCR

To verify that transcriptional levels were not an artifact of the library preparation process we preformed qRT-PCR on several genes that were differentially expressed by at least two-fold in arsenic conditions relative (Table 1). We chose five candidates genes from this list in Table 1, which to our knowledge were not previously implicated in arsenic respiration. Q-RT-PCR expression was normalized to *recA*, which was similarly expressed in all three conditions. Of these transcripts we saw similar transcriptional trends as in the transcriptome data (Figure 1A and B). Again, when plotted against the aerobic transcription levels the anaerobic transcriptional patterns were more similar to each other than to aerobically grown cells. Because the RNAseq expression data showed consistent results with at least five of the genes we analyzed by more traditional qRT-PCR methods, we concluded that RNAseq would

be a promising approach to finding additional genes that are unique to the arsenate respiration pathway of *Shewanella* sp. ANA-3.

Transcriptome analysis using several bioinformatics tools

In this study we set out to identify changes in the global transcriptional patterns in *Shewanella* in response to differential respiratory conditions. We used several commercial (CLC Genomics Workbench version 7.5.1) (http://www.clcbio.com**)** and open source bioinformatics tools (DEseq)(Anders and Huber, 2010) to identify differentially expressed genes that are important for aerobic vs. anaerobic respiration and arsenate specific respiration (Figure 2). Using CLC Genomics Workbench we compared the results of two different statistical analyses, t-test vs. EDGE because each approach could produce different outcomes (Fernandes et al., 2013)(Anders and Huber, 2010) . In all cases we used the same criteria to consider a gene as differentially regulated: (i) a false discovery rate (FDR) p-value less than 0.05 and (ii) a greater than two-fold difference in the average normalized expression compared to the average normalized expression of the replicates in another condition. Using DESeq, we found 85 genes that were up-regulated under arsenate respiring conditions. Almost all of the DE (differentially expressed) genes identified by CLC Genomic Workbench were also found by DESeq (Figure 3A). For t-test and EDGE analyses, we found that 42 and 89 genes, respectively, were highly expressed in arsenate respiration conditions in ANA-3 (Figure 3B and 3C). For anaerobic conditions (arsenate and fumarate) using the CLC Genomics Workbench t-test and

EDGE we found 203 and 289 genes that were common to both conditions. For growth on oxygen, 303 and 302 genes were unique to that aerobic growth conditions (Figure 3B and 3C). For DEseq similar analyses revealed 278 genes expressed under fumarate and arsenate respiring conditions with 295 genes expressed uniquely to aerobic conditions. We also compared the two anaerobic conditions and found 31 (44 DESeq) genes were up regulated in fumarate grown cells whereas 42 were up regulated in arsenic conditions (Figure 3C). Our data analysis shows that DESeq and EDGE produced the most similar results compared to t-test analysis through CLC Genomics Workbench. Taken together, analyzing transcriptomic data with several statistical methods is useful and helps prevent loss of statistically significant data due to the assumptions of each method (Fernandes et al., 2013).

Overview of transcriptomic trends in *Shewanella* grown on various terminal electron acceptors

We found about 12% of the genome was differentially expressed when ANA-3 was grown aerobically vs. under anaerobic conditions. In a similar study done with *Staphylococcus aureus*, a switch from aerobic to anaerobic respiration conditions resulted in the differential expression of only ~6% of the genome (Fuchs et al., 2007). In contrast to the aerobic vs. anaerobic comparison, only 1% of the genome is expressed specifically in arsenic respiring conditions in ANA-3 (Figure 2), which is similar to the patterns observed for metal reduction in *Shewanella* oneidensis str. MR-1 (Bencheikh-Latmani et.al, 2005). In the arsenate-respiring hyperthermophilic

archaeon, *Pyrobaculum aerophilium*, 14% of its genes were differentially expressed when comparing growth on different respiratory conditions including aerobic vs. arsenate, nitrate, and Fe(III)-citrate (Transcriptional map of respiratory versatility in the hyperthermophilic crenarchaeon Pyrobaculum aerophilum., 2009). Transcriptomic studies in the arsenite-oxidizing microbe, *H. arsenicoxydans,* reported up to 12% of the genome differentially expressed under arsenate stressed conditions (Cleiss-Arnold et al., 2010).

I. Pairwise comparison of arsenate, fumarate, and oxygen conditions

We did a pairwise comparison of ANA-3's Transcriptome in the three respiratory conditions based on the TIGR gene ontology (Figure 4) (Kumar et al., 2004). We found differential expression based on functional groups; as expected, there were many more genes that were differentially expressed in pairwise comparisons between aerobic and either anaerobic condition than differentially expressed genes between the two anaerobic groups. Most notably there were several genes that were differentially expressed between all groups involved in electron and carbohydrate transport, DNA interactions, amino acid metabolism and central metabolism, which highlights the unique need for specific genes involved in transport, regulation and metabolism to respire each different terminal electron acceptor.

II. Genetic response to anaerobic conditions

Of the 569 genes that were differentially expressed between the aerobic and anaerobic conditions, based on a corrected t-test, many of the genes that were highly expressed anaerobically were directly involved in energy metabolism (Figure 4), including sixteen genes for electron transport, seven involved in fermentation, and three in the tricarboxylic acid cycle. Our results reflect the trends of a similar microarray based study of *S. aureus* (Fuchs et al., 2007). Moreover, genes that were annotated in TIGR as transport and binding genes were heavily represented in anaerobic transcripts including several NADH and sulfate/thiosulfate transporters.

III. Electron transport chain response to terminal electron acceptor

III. A. Oxygen specific response

One of the goals of this study was to identify the genes that are important for aerobic respiration. Using CLC's t-test, we first identified 293 genes that were specifically up-regulated when ANA-3 was grown in oxygen. Our expression data for these genes corroborate trends found in other recently published transcriptomic studies of aerobic respiration in *Shewanella* (Clark et al., 2014) (Brutinel and Gralnick, 2012). Several of the genes in the tricarboxylic cycle were highly expressed. For example, the *gltA* gene encoding for citrate synthase (Shewana3_1705; [SO_1926 and ACDC_00013790]) was more highly expressed under aerobic relative to arsenate (fold change (FC), 5.5) and fumarate (FC, 5.46) conditions. In the genome wide study of *S. oneidensis,* using high throughput sequencing of transposon mutants (TN-seq),

gltA was shown to be critical for aerobic growth but non-essential for growth under anaerobic conditions (Brutinel and Gralnick, 2012).

Similar to the citrate synthase gene, *gltA*, we also observed increased aerobic expression patterns of the two genes annotated as aconitate hydratase: *acnB* (Shewana3_0433, SO_0432) and *acnD* (Shewana3_3827; SO_0343; ACDC_00027580). Aconitate hydratase functions within the TCA cycle to catalyze the reversible conversion of citrate to iso-citrate with aconitate as an intermediate. However, AcnD appears to be involved in proprionate utilization within the *prp* operon, *prpBCacnDprpF*. AcnD is known to catalyze a related reaction of 2-methylcitrate to 2-methylate-*cis*-aconitate followed by hydration via AcnA or AcnB to 2-methylisocitrate (2-MIC). The resulting 2-MIC can be converted into pyruvate and succinate, which are then fed back into the TCA cycle. In ANA-3, both *acnB* and *acnD* showed differential expression in aerobic conditions relative to growth in arsenate (FC 33.8 and 6.8 , respectively) and fumarate (FC 38.31 and 8, respectively). In the TN-seq study by Brutinel and Gralnick (2012) no aconitate hydratase mutants were detected in their TN libraries (which were prepared under aerobic conditions) indicating that this function is essential for aerobic growth in *S. oneidensis*. However, Clark et al. (2014) reported the aconitate hydratase as being repressed aerobically in lactate grown cells relative to acetate grown conditions. In *E. coli,* there are three aconitate hydratase encoding genes: *acnA*, *acnB*, and *acnC* (Transcriptional regulation of the aconitase genes (acnA and acnB) of Escherichia coli., 1997). The *E. coli acnB* gene appears to be the dominant aconitate hydratase

exhibiting anaerobic repression mediated by ArcA, Fis, and Fru but activated by CRP. In our studies with ANA-3, *acnB* expression was consistent with that of *E. coli*, however ANA-3 lacks *acnA* and instead has the paralog *acnD*. From our RPKM data the *acnB* and *acnD* showed interesting expression patterns. The *acnB* (Shewana3_0433) aconitate hydratase was expressed nearly 10-fold higher than *acnD* (Shewana3_3827) in both arsenate and fumarate conditions relative to oxygen. Grimek et al. (Grimek and Escalante-Semerena, 2003) demonstrated that the *Shewanella oneidensis* AcnD could catalyze the first half of the aconitase hydratase reaction, citrate to *cis*-aconitate. The final conversion to isocitrate required AcnA or AcnB protein thereby providing a metabolic connection to the TCA cycle via proprionate acid metabolism. Brutinel and Gralnick (2012) further demonstrated that *acnB* was likely the main aconitate hydratase in *Shewanella* and showed that genes in the *prp* operon were functioning also as an alternative citrate synthase. Both *acnB* and *acnD* expression could provide additional respiratory flexibility to *Shewanella* when grown on fermentation end products produced in anoxic environments.

In addition to the TCA cycle genes, a number of additional genes involved in oxygen respiration were also differentially expressed (Figure 5). Two distinct NADH dehydrogenase gene clusters were differential expressed, which appear to be substituted for the *ndh/nqo*-type typically associated with Complex I ((Melo et al., 2004)). One particular NADH-quinone reductase encoding complex, *nqrABCDEF* (Shewana3_0938-0943), found in *V. alginolyticus* and other marine bacteria (Unemoto and Hayashi, 1993) (Hayashi et al., 1995), was highly expressed in oxygen

respiring conditions. Conversely, the other sodium translocating NADH-quinone reductase was differentially expressed under anaerobic conditions (Shewana3_3374-3378). Unlike ANA-3 that lacks the typical H+ dependent NADH-quinone reductase, *V. alginolyticus* and other marine and halophilic bacteria, although they rely heavily on Na+ for respiration, they also have a H+ dependent NADH-quinone reductase complex (Melo et al., 2004). Other *Shewanella* spp. also have two gene clusters encoding putative type-III sodium translocating NADH dehydrogenases. However the canonical (type-I) rotenone sensitive NADH dehydrogenase (NUO/NDH) is not present in ANA-3. Similarly, another related marine microbe, *Vibrio cholerae*, also lacks the type-I NADH dehydrogenase. Moreover, of the 24 sequenced *Shewanella* genomes, only *Shewanella oneidensis* str. MR-1 contains the type-I NADH dehydrogenase (SO_1010-1021) in addition to one type-II (SO_3517) and two type-III NADH dehydrogenases (SO_1103-1108 and SO_0902-0907).

The succinate dehydrogenase complex (Complex II) is also an important component of aerobic respiration that is used during the conversion of succinate to fumarate and also recycling FADH to FAD^{+}. We found that the genes encoding Complex II (Shewana3_1706-1709) were expressed ~5 fold higher under aerobic conditions relative to growth on arsenate and fumarate. Little is known about the role of *sdh* in *Shewanella*. However, our expression profile for Complex II is similar to *E. coli* where the *sdh* gene cluster is induced aerobically and repressed under anaerobiosis (Park et al., 1995).

Complex III: Several Complex III genes were up regulated aerobically. The *petABC* homologs (Shewana3_0601-0603) were up regulated more than 2-fold in aerobic conditions compared to anaerobic conditions (Figure 5). The *petABC* complex encodes the cytochrome bc_1 respiratory complex that functions as a proton pump and reducing agent for cytochrome *c*. In MR-1 the *petABC* genes were also up regulated aerobic relative to anaerobic conditions. In MR-1, a *petC* mutant strain (Shewana3_0603 (SO-0610)) exhibited aerobic growth defects and was unable to use TMAO or MnO_2 as electron acceptors (Gao et al., 2010). Luo et al. (2013) found that similar to the *petC* mutant reported in Gao et al. (2010) *petA* mutants also had aerobic growth defects but were able to grow on fumarate and DMSO. Based on this latter result, we predicted that *petC* mutant in ANA-3 would not likely affect growth on arsenate because of commonalities in the electron transport to growth on fumarate (e.g. requirement for CymA) and similarities in their regulation.

Complex IV: In prokaryotes, there are two families of terminal oxidases for O_2 reduction, the more broad heme copper oxidases (HCO) that can receive electrons from quinone or other cytochromes and the more stringent quinol only cytochrome *bd*-type oxidases (Morris and Schmidt, 2013). Most *Shewanella* species have one *bd*-type cytochrome *c* oxidases two HCO's, cbb_3-type and A-type. ANA-3 in particular has genes encoding two putative *bd*-type quinol oxidases and one heme copper oxidase of the cbb_3-type (Figure 5). We found that genes for the ccb_3-type oxidase (Shewana3_2107-2110) *ccoNOQP* were highly expressed in cells grown aerobically

and as well as for anaerobic growth on fumarate or arsenate (Figure 5). However, Zhou et al. (2013) found that the expression of *cbb*$_3$-type cytochrome genes in *Shewanella* sp. MR-1 (SO_2364-2360) was highest in aerobic growth but ~10-fold less in microaerophilic (1% oxygen) conditions. Using biochemical and genetic approaches Le Laz et al. (2014) showed that in *Shewanella* sp. MR-1 the *cbb*$_3$-type cytochrome oxidase was the most dominant in aerobically grown cells (Zhou et al., 2013; Le Laz et al., 2014). The *bd*-type quinol oxidase encoding genes, *cydAB* (Shewana3_1241-1242, SO_3286-3285) were expressed 9 and 11-fold higher in arsenate and fumarate growth conditions, relative to aerobically grown cells, respectively. Both Zhou and Laz found the A-type HCO encoding genes, *coxABGC* (Shewana3_3991-3994, SO_4606-4609) were not expressed aerobically. This gene cluster was not expressed in the respiratory conditions we tested either. However, ANA-3 has a second gene cluster (Shewana3_4031-Shewan3_4034) encoding another putative cytochrome *c* oxidase *cyoBACD,* which is not found in MR-1. This alternative *cyoBACD* gene cluster was mostly expressed aerobically in ANA-3 with minimal to no detectable expression in fumarate (~21-fold change relative to fumarate) and arsenate (~87-fold change relative to arsenate) conditions relative to aerobically grown cultures respectively. In *E. coli* the cyo$_3$, is the preferred cytoplasmic oxidase in high oxygen conditions therefore it is not surprising that this oxidase is differentially expressed with many more transcripts during aerobic respiration (Yap et al., 2009) (Figure 5). The diversity in Shewanella's terminal

oxidase expression profiles indicates a fine tuned ability to respire a diverse set of substrates under changing redox conditions.

Transcriptomic response during arsenate respiration

One of the main goals of our study was to use differential gene expression profiles of ANA-3 to gain a better understanding of how microbes utilize the toxic metalloid arsenate as a terminal electron acceptor for anaerobic respiration. The terminal reductase that is required for arsenic respiration, ArrAB, has been studied extensively (Saltikov and Newman, 2003; Konstantinidis et al., 2009; Jiang et al., 2013). We know that the genes encode by *arrAB* are regulated in response to anaerobiosis, arsenite, and elements of central metabolism such cAMP its receptor protein, CRP (Malasarn et al., 2008; Murphy et al., 2009; Murphy and Saltikov, 2009). However, there are other possible regulatory elements involved in controlling arsenate respiration in ANA-3. We focused our efforts on identifying genes that were specifically expressed in arsenate-respiring conditions. Our data reflect many of the previously established observations in respiratory arsenate reduction for *Shewanella sp.* ANA-3 and other arsenic-metabolizing bacteria. For example, most of the genes in the "arsenic island" of ANA-3 (Shewana3_2327-Shewana3_2351) were uniquely expressed under arsenate respiring conditions. Many of these genes showed the highest differential expression relative to fumarate or oxygen growth conditions (Table 3 and Figure 4D) (Saltikov). Similarly, expression patterns previously determined for genes required for Fe(III) and Mn(IV) reduction (*mtr* genes) also

showed similar patterns by RNAseq (Reyes et al., 2010). We also observed up regulation of sulfur related transcripts (Figure 8), which were previously identified in a microarray-based transcriptomic study on arsenic induced stress response with *H. arsenicoxydans* (Cleiss-Arnold et al., 2010). The following is an overview of the trends in our RNAseq data relevant to arsenic metabolism.

The arsenic island: Consistent with earlier accounts, the majority of the top 20 highly expressed transcripts mapped to a cluster of arsenate-respiration and detoxification genes know collectively as the "arsenic island" (Figure 7A) (Murphy and Saltikov, 2009; 2007; Silver and Phung, 2005). Within the arsenic island there are four genes encoding putative permease/facilitator/transport like proteins. Shewana3_2344 shows the greatest homology to known ArsB, arsenite efflux pumps. This gene was shown to be required for arsenite resistance and enhanced anaerobic growth on arsenate (Saltikov et al., 2003). This gene most likely encodes an ArsB because the neighboring genes are also highly similar to other genes in previously characterized *ars* operons such as the *arsRDABC* operon on the R773 conjugative resistance plasmid of *E. coli* (Xu, 1998). In ANA-3, the *ars* operon is composed of *arsDABC* with several genes for *arsR* and *arsC* nearby. The first gene of the ANA-3 *ars* operon encodes a putative arsenite-chaperone, ArsD (Shewana3_2342), which was differentially expressed (~146-fold increase) in arsenate relative to fumarate and aerobically grown cultures. An *arsD* mutation strain of ANA-3 was previously shown to only affect growth in the presence of arsenite and not when arsenate was

provided as a sole electron acceptor (Murphy and Saltikov, 2009). The two neighboring genes encoding ArsC-like proteins, Shewana3_2334 and Shewana3_2331 (annotated as protein-tyrosine phosphatases), were also highly expressed in arsenate-respiring conditions. There are a number of other genes encoding ArsC-like proteins; many of these are highly expressed under arsenate growth conditions. Shewana3_2345 appears to encode the primary detoxifying arsenate reductase because deleting this gene in an *arrA* background strain of ANA-3 eliminated the ability to reduce arsenate (Saltikov et al., 2005). There may be other putative *arsC* encoding genes, which are annotated as either glyceraldehyde 3-phosphate dehydrogenase or low molecular weight protein tyrosine phosphatase (LMW PTPase). The former enzyme is known to catalyze arsenate reduction (Nemeti., 2006). For the latter, ArsCs are related to LMW PTPase and often have residual phosphatase activity (Zegers et al., 2001).

Transcriptional regulators in the ArsR family of arsenite dependent repressors, Shewana3_2342 and Shewana3_2335, were also highly expressed during anaerobic growth on arsenate. Our past studies showed that the gene product of Shewana3_2342, ArsR2, binds the intergenic region of *arr* and *ars* operons and is a likely repressor for their transcription (Murphy and Saltikov, 2009). An ANA-3 *arsR2* deletion strain exhibited enhanced anaerobic growth on high arsenate concentrations (100 mM) and increased resistance to arsenite (3 mM). In the other *arsRC* gene cluster (Shewana3_2327-2338), which occurs adjacent to *arsR2C2*, expression (only *arsR1*) showed lower expression relative to the other *ars* encoding

genes within the arsenic island. Moreover, *arsR1* was not differentially expressed relative to fumarate and oxygen growth conditions. Unlike *arsR2*, deletion in *arsR1* had little effect on the growth on arsenate and resistance to arsenite(Murphy and Saltikov, 2009). A putative ArcA-family regulator, encoded by Shewan3_2327, was also differentially expressed relative to fumarate and oxygen. The function of this putative protein remains to be determined.

Finally, the arsenic respiratory reductase gene cluster *arrAB* (Shewana3_2341 and Shewana3_2340) was among the top twenty genes identified by RNAseq expressed in arsenic respiring conditions. These two genes are absolutely essential for anaerobic growth with arsenate as a terminal electron acceptor (Saltikov and Newman, 2003), thus confirming the utility of RNAseq and transcriptomics for identifying essential genes in the arsenate respiration pathway.

Sulfur metabolism. The reduced form of arsenic, arsenite, is known to exert its toxicity via binding to exposed sulfhydryl groups on proteins and also by forming conjugates to -SH containing peptides like glutathione (Shen et al., 2013). A number of genes involved in sulfur assimilation were among our top most differentially expressed in arsenate growth conditions. For example, the chromosomal region containing assimilatory sulfate reductase genes (Shewana3_0855-Shewana3_0867) were highly expressed in arsenate relative to oxygen conditions. Most of these genes were expressed 15-40-fold and to a lesser degree 2-4 fold higher relative to fumarate growth conditions (Fig. 5). A well-known pathway for sulfur assimilation from

sulfate involves uptake through an ABC-type transporter. ANA-3 genome contains a gene cluster encoding for a putative sulfate ABC-type transporter (Fig. 5, *cysPUWA*, Shewana3_3189, Shewana3_3190, Shewana3_3191, Shewana3_3192). These genes were highly expressed anaerobically and even greater in arsenate relative to fumarate growth conditions. Once in the cell, sulfate is activated by the sulfate adenylyltransferase subunits 1 (Shewana3_0864, 3.7-fold change, As/Fum) and 2 (Shewana3_0863, 2.4-fold change, Fum/As) enzyme complex, which catalyzes the formation of adenylphosphosulfate (APS) from ATP and sulfate. The next step in sulfur assimilation involves APS phosphorylation to phosphoadenosine phosphosulfate (PAPS) by adenylylsulfate kinase (Shewana3_0866, 4.4-fold change, As/Fum). This pathway is common in a variety of bacteria and archaea and can provide a major sulfide for biosynthetic pathways requiring sulfur (Klassen and Boles, 2007). Because these genes were also differentially expressed in fumarate vs. oxygen conditions, arsenite toxicity may not be the only reason for increased anaerobic expression of sulfur assimilatory genes. However, the up-regulation of these genes in arsenate growth conditions relative to fumarate indicates that arsenic toxicity may impose a further impact on biosynthesis of sulfur containing biomolecules.

Our expression data also suggests that ANA-3 may have other pathways for arsenite detoxification. In addition to *arsB* of the *arsDABC* gene cluster, we observed differential expression of the two annotated glutathione (GSH) biosynthesis genes, Shewana3_1824 and Shewana3_2339. The former appears to be constitutively

expressed among all three conditions whereas the latter is differential expressed in arsenate relative to both fumarate (11-fold greater) and aerobically (24-fold greater) grown cells. In bacterial studies with ArsC arsenate reductase, reduced GSH is known to function as a reductant in Gram-negative bacteria (Mukhopadhyay et al., 2002). In eukaryotes, sulfur-containing biomolecules such as GSH have been shown to form intracellular complexes with arsenite as a means of arsenic detoxification (Scott et al., 1993; Hughes, et al., 2002). The ANA-3 genome contains seven genes predicted to encode for glutathione-S-transferase (GST) genes. However, only two seem to be differentially expressed but in opposite ways. Shewana3_0755 is differentially expressed ~3 higher in aerobic conditions relative to arsenate and fumarate. Shewana3_2865. In a transcriptome study with *Arabidopsis*, arsenic exposure was shown to induce the expression of 14 genes encoding for GSTs. Despite the lack of information about the formation of GSH-As (III) conjugates in bacteria, the increased expression of GSH biosynthesis and GST genes could be an additional factor affecting growth on high concentrations of arsenic.

In particular, the gene encoding a putative thiol-disulfide interchange protein (Shewana3_3226) was highly expressed in arsenate respiring conditions. This protein is important for correcting non-native disulfide bonds. This interaction can possibly serve as a detoxification mechanism to protect the cell, as arsenate is known to react chemically with sulfur containing amino acids (Brosnan and Brosnan, 2006). Alternatively the thiol-disulfide interchange protein may be directly interacting with the ArrAB arsenate reductase or other cysteine rich proteins involved in arsenic

respiration to ensure proper protein confirmation as overexpression of DsbC and DsbG has been shown to improve the function of soluble cysteine rich proteins by preventing autoaggregation in *E. coli* (Zhang et al., 2002).

We also surveyed other systems-level biology studies involved arsenic exposure in prokaryotes for trends in sulfur metabolism. However the analysis in this regard is often focused elsewhere making it difficult to draw a general conclusion about the effects of arsenic on sulfur metabolism in prokaryotes without further mining supplemental material. Although sulfur metabolism was not a topic that surfaced in their discussion, Cozen et al., 2009 characterized the transcriptome of the arsenate respiring hyperthermophilic archaeon, *Pyrobaculum aerophilium* with the focus on validating genes involved in oxygen, nitrate, arsenate, and Fe(III)-citrate respiration. Nearly two-thirds of the *Pyrobaculum* genome lacks reliable annotations and hence, the analysis using microarray approaches, identified numerous genes encoding for various terminal oxidases and reductases for nitrate and arsenate. They were unable to identify a specific "reductase" for iron (III). An arsenic-specific transcriptomic and proteomic study was conducted with chemolithoautotrophic arsenite-oxidizer, *Rhizobium* sp. NT-26 (Andres et al., 2013). Genes involved in sulfur oxidation (*soxGVWXYZ*) and the sulfate/thiosulfate transporter (*cysT*) were down regulated in NT-26 when grown under arsenite-oxidizing conditions. In the non-arsenic respiring cyanobacterium *Synechocystis* sp. PCC 6803, no changes in sulfur metabolism genes were observed in cells exposed to arsenate or arsenite (Sánchez-Riego et al., 2014). Transcriptome studies indicated that sulfur metabolism

might be impacted by arsenic exposure, however the physiological effects remain to be determined.

II.C-II Electron transport

It is well established that the tetraheme *c*-type cytochrome, CymA, is essential for arsenate respiration (Murphy and Saltikov, 2007). Our study showed that CymA (Shewana3_3977) and cytochrome biogenesis factor (Shewana3_2348). We also found a blue-copper domain containing protein (Shewana3_3478), which was previously unrecognized in arsenic metabolism, to be significantly expressed in arsenic conditions. Typically blue copper domain containing proteins are found in all kingdoms and are known to be important for electron transfer (Vasin et al., 2013). Increased expression of blue copper domain containing protein in the arsenic respiring context may be important for shuttling electrons towards arsenic during respiratory arsenic reduction.

Genes repressed in arsenate respiration conditions

We identified 24 differentially expressed genes that were lower in arsenate growth conditions relative to fumarate and aerobic conditions (Table 2S). Many of these genes are predicted to encode transport and solute binding proteins, regulators, proteins involved in amino acid biosynthesis, and proteins of unknown function. Despite the overall high expression of the heat shock protein HtpX, the heat shock sigma factor *rpo*E (Shewana3_3032) was significantly down regulated in arsenic compared to fumarate and oxygen grown conditions. Other transcripts that were

specifically down regulated in arsenic conditions include the Protoheme IX farnesyltransferase (Shewana3_4035). This enzyme is known to be important for producing hemO, which in turn causes heme cleavage for iron acquisition (Zhu et al., 2000). Activation of hemO could be detrimental in arsenic respiring cells since cytochromes play such an integral role in arsenic reduction. The gene cluster spanning Shewana3_0565 to Shewana3_0567 genes encodes a putative transporter. The putative ABC-3 protein transcriptional regulator, shown to release vacuolar iron into the cytoplasmic space when the cells are starving for iron in *S. pombe*, is also down regulated in arsenic respiring conditions (Pouliot et al., 2009). Both hemO and ABC-3 are iron scavenging mechanisms that make iron more bioavailable It is possible that the cell may also be using this repression to sequester iron away from arsenic since arsenic and iron are so reactive with each other (Shen, 2013). Many of these genes were hypothetical yet only a few have an obvious relationship to arsenic respiration (Supplementary table 2 (down in As(V)) .

Despite the fact that several phosphatases were up regulated in arsenic respiring conditions still one was down regulated in arsenic respiring conditions (Shewana3_0904). Modulating the expression of various phosphatases could represent a mechanism to elude misuse of arsenic for phosphate, an essential element which shares an uncanny resemblance to arsenic (Elias et al., 2013).

Phenotypic analysis of several genes expressed in arsenic respiring conditions

To study the phenotypic significance of differentially expressed genes several genetic mutants were generated. Gene disruption mutants were generated as described in Zargar et al 2010 with one modification (using pSMV10) (Zargar et al., 2010). The 12 mutants were selected because they had significantly higher expression in arsenic respiring conditions than the other respiratory conditions. We were interested in determining how these genes affected growth on arsenic relative to fumarate and oxygen respiring conditions. These mutants were grown anaerobically with 10 mM arsenate or fumarate. (Figure 9A and B). Despite the strong expression patterns, 8 of the 12 mutants showed no growth defects in either anaerobic condition. Paradoxically, 4 mutants exhibited a growth defect in both anaerobic conditions despite low expression levels in fumarate grown cells. This suggested that expression levels, although very informative for elucidating relevant genes, might not always point to a clear phenotype. Nonetheless, because arsenic is notoriously cytotoxic, we suspected that many of the genes that were up regulated in arsenic grown conditions during exponential growth were important for mitigating arsenite's toxic effects rather than involved in cellular respiration. To identify genes that are involved in toxin resistance we grew several of the disruption mutants aerobically in serial 2-fold dilutions of arsenite from 5mM to 1.25mM. Of the mutants tested for arsenite sensitivity, we found that mutants phosphoadenosine phosphosulfate reductase (Shewana3_0857), inner membrane transport (Shewana3_2062), TetR family transcriptional regulator (Shewana3_2864), were hyper-sensitive to arsenite compared to wild type ANA-3. These hyper-sensitive mutants also exhibited general

growth defects and severe anaerobic growth defects in both arsenate and fumarate, despite their low expression values in aerobic or fumarate grown cells. The non-iron sulfur phosphoadenosine phosphosulfate reductase (Kopriva et al., 2002) (Shewana3_0857) was highly expressed in arsenic, intermediately in fumarate and low in oxygen however this mutant produced striking growth phenotypes in all three conditions. The phosphosulfate reductase functions to catalyzes the reduction of 3'-phosphoadenylyl sulfate into sulfite although this gene seems to be important for all three forms of metabolism perhaps arsenic reduction requires more free sulfite to facilitate arsenic reduction. In eukaryotic cells glutathione is important for conferring resistance to arsenic's toxicity (Davison et al., 2003). Perhaps glutathione is also functioning in arsenic detox by providing a thiol group for arsenic scavenging.

The protein product of the Thiol: disulfide interchange, was shown to interact substantially with the detoxifying arsenate reductase ArsC in *Bacillus subtilis* to maintain a reduced state in ArsC (Li et al., 2007). In our study a Thiol: disulfide interchange mutant, Shewana3_3226, also produced some degree of arsenite sensitivity when grown aerobically in 5mM arsenite despite no significant growth defect in cells grown anaerobically in 10mM arsenate.

Conclusion

In this work we sought to identify transcriptional trends that are important for anaerobic respiration as well as the respiration of the toxic metal arsenic. Our study shows that a notable portion (~12%) of the genome is differentially expressed in aerobic vs. anaerobic respiration in ANA-3 and only a fraction of those were specific

for the metabolism of a particular anaerobic substrate. The prevalence of many transcripts that are associated with sulfur metabolism also suggests that transcription of genes involved in the sulfur metabolism pathway plays an important role in respiratory arsenic reduction. If translated, these thiol containing proteins may be involved in sequestering arsenic to avoid its toxic effects however this must be verified biochemically. In our study we also noticed that using more than one form of bioinformatic analysis (proprietary work-bench vs. open-source workflow) could produce slightly different datasets. This is difficult to tease out since the details for the experimental methodology are not readily available for a genomics workbench. In spite of this fact, both data sets verified similar trends found in the publication records although the open source DE-seq analysis provided twice as many (85 vs 42) genes that are specifically up regulated more than 2 fold in arsenic respiring conditions.

Chapter 2 Figures and Tables

Figure 2.1. Quality control and experimental validation.
A. Expression values based on DEseq analysis of five genes that were highly expressed in arsenic respiring conditions. Y-axis represents reads per kilobase per million (RPKM) in either arsenic or fumarate normalized to RPKM values under aerobic conditions. X-axis represents the suffix of the gene locus tag number prefixed by Shewana3_. Error bars represent the standard deviation of three replicate experiments after pairwise normalization to aerobic experiments. **B.** Q-RT-PCR verification of specific transcripts. Expression of five transcripts that were highly expressed only during arsenic reduction in transcriptomic experiments was measured using qRT-PCR. Expression levels of specific genes (X-axis) were plotted relative to the delta delta Ct in oxygen grown conditions. Red bars represent transcript levels in arsenate respiring cells and green bar represents transcript levels in fumarate respiting cells. Error bars represent standard deviation of triplicate samples.

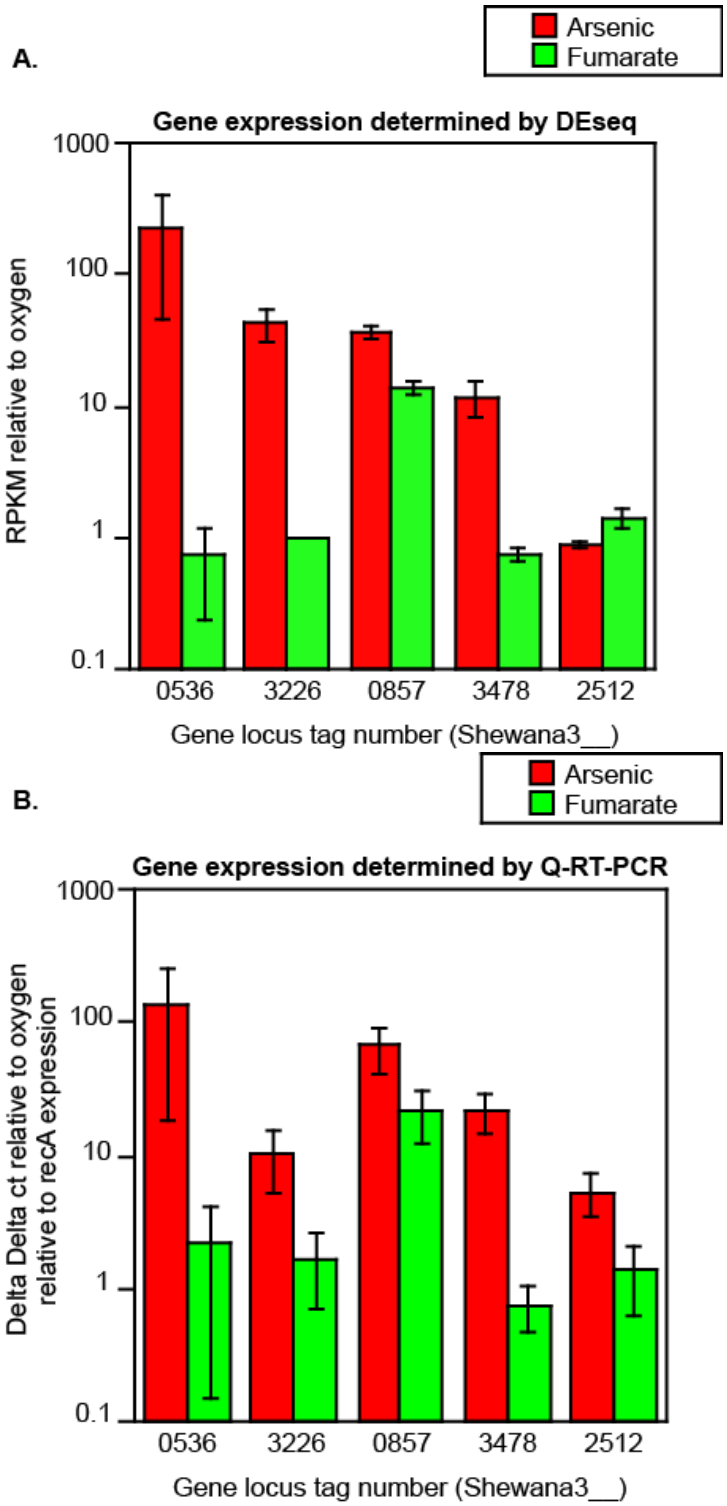

A.
Arsenic
Fumarate
Gene expression determined by DEseq
RPKM relative to oxygen
1000
100
10
1
0.1
0536
3226
0857
3478
2512
Gene locus tag number (Shewana3__)
B.
Arsenic
Fumarate
Gene expression determined by Q-RT-PCR
Delta Delta ct relative to oxygen relative to recA expression
1000
100
10
1
0.1
0536
3226
0857
3478
2512
Gene locus tag number (Shewana3__)

Figure. 2.2 Pairwise scattergraphs of genomewide RPKM values. comparisons of the RPKM values for each gene in each growth condition (arsenate, fumarate and oxygen). Each data point represents the average of three biological replicates for each gene. The colored data points (red, green and blue) represent genes that are differentially expressed more that 2 fold between the two conditions based on the DESeq analysis with corrected p-values < 0.05.

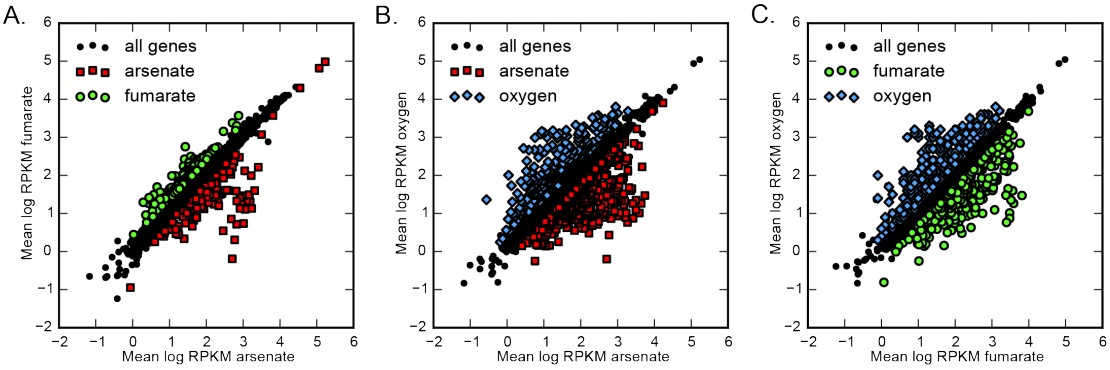


Figure 2.3. Physiological trends in transcriptomic expression profiles. Venn diagrams of up regulated genes found in each statistical analysis type. In each diagram transcriptional patterns are based on a fold change greater than 2. **A.** Venn diagram of up regulated genes identified by on DESeq analysis with p-value of less than 0.05 and fold change greater than 2. **B.** Up regulated genes identified by the EDGE-test analysis generated with CLC genomics workbench 3.2.6, a false discovery rate (FDR) p-value <0.01 **C**. Up regulated genes identified by t-test analysis generated with CLC genomics workbench 3.2.6, with a false discovery rate (FDR) p-value <0.05.

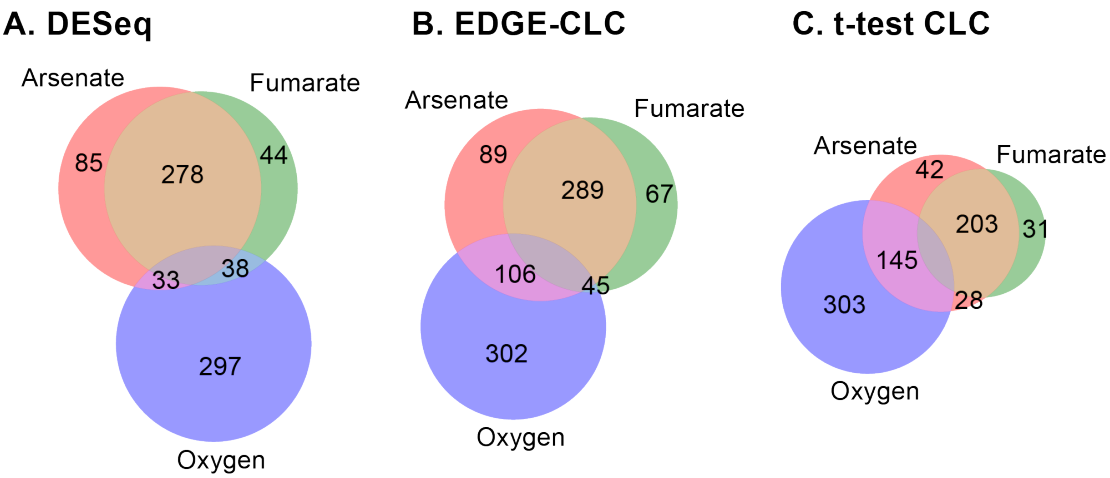
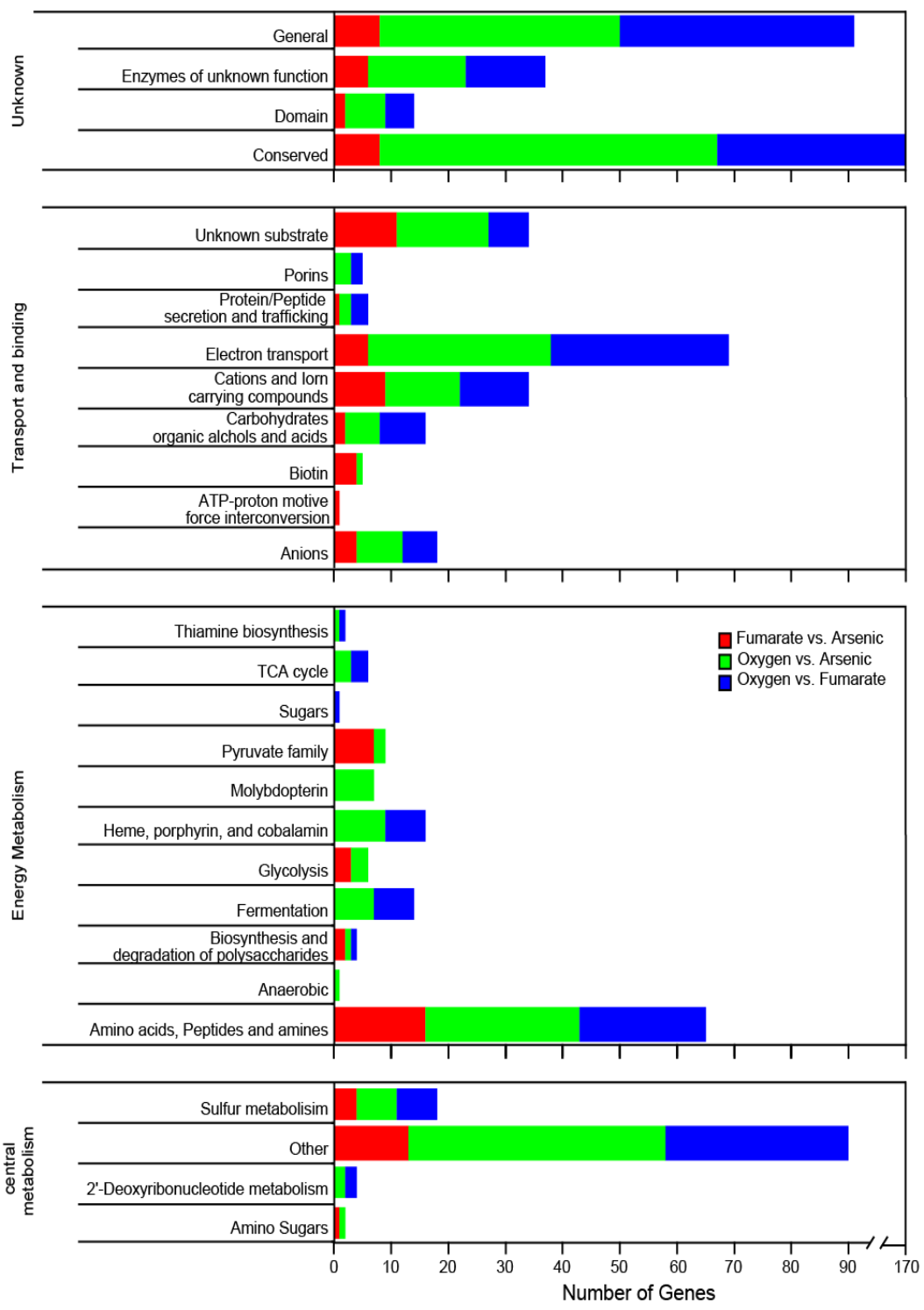


Figure 2.4. Gene ontology of differentially expressed genes. Functional gene ontology clustering of differentially expressed genes assigned using J. Craig Venter Institute (JCVI), The Institute of Genomic Research (TIGR) gene product annotation. Blue bars represent genes that are differentially expressed more than 2 fold in oxygen relative to fumarate respiring conditions. Red bars represent genes that are differentially expressed in fumarate relative to arsenic. And green bars represent genes that are differentially expressed more than two fold in oxygen relative to arsenic. A. Differentially expressed genes that are annotated as unknown functions or annotated with a main role in transport and binding, energy metabolism or central metabolism. B. Differentially expressed genes with a main role in cellular processes

A.


Unknown
General
Enzymes of unknown function
Domain
Conserved
Transport and binding
Unknown substrate
Porins
Protein/Peptide secretion and trafficking
Electron transport
Cations and Iorn carrying compounds
Carbohydrates organic alchols and acids
Biotin
ATP-proton motive force interconversion
Anions
Energy Metabolism
Thiamine biosynthesis
TCA cycle
Sugars
Pyruvate family
Molybdopterin
Heme, porphyrin, and cobalamin
Glycolysis
Fermentation
Biosynthesis and degradation of polysaccharides
Anaerobic
Amino acids, Peptides and amines
central metabolism
Sulfur metabolisim
Other
2'-Deoxyribonucleotide metabolism
Amino Sugars
Fumarate vs. Arsenic
Oxygen vs. Arsenic
Oxygen vs. Fumarate
0
10
20
30
40
50
60
70
80
90
170
Number of Genes

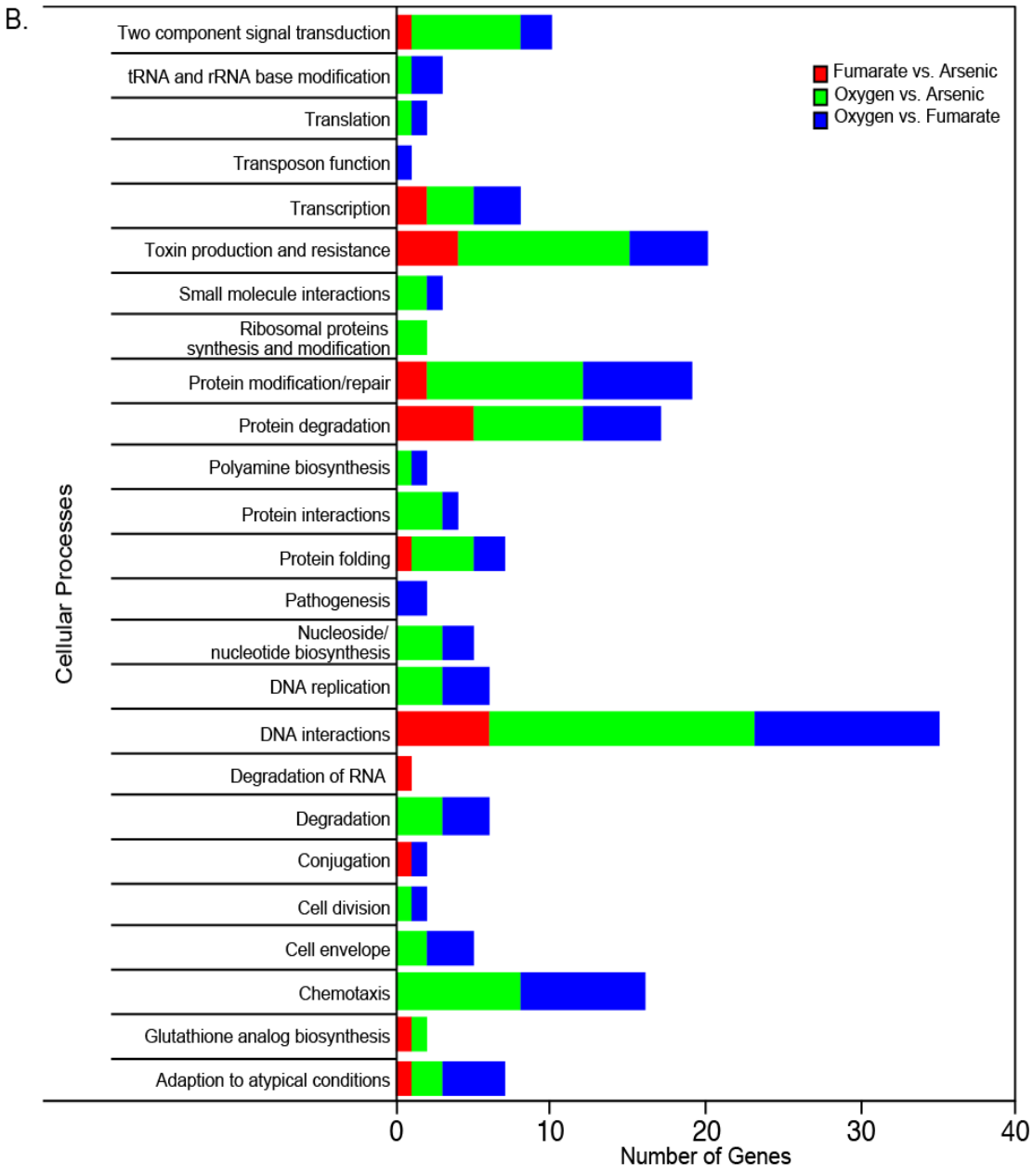

B.
Two component signal transduction
tRNA and rRNA base modification
Translation
Transposon function
Transcription
Toxin production and resistance
Small molecule interactions
Ribosomal proteins synthesis and modification
Protein modification/repair
Protein degradation
Polyamine biosynthesis
Protein interactions
Protein folding
Pathogenesis
Nucleoside/ nucleotide biosynthesis
DNA replication
DNA interactions
Degradation of RNA
Degradation
Conjugation
Cell division
Cell envelope
Chemotaxis
Glutathione analog biosynthesis
Adaption to atypical conditions
Cellular Processes
Fumarate vs. Arsenic
Oxygen vs. Arsenic
Oxygen vs. Fumarate
0
10
20
30
40
Number of Genes

Figure 2.5. Electron Transport Chain expression analysis. Distribution of average reads for each operon within the electron transport chain in ANA-3. Red represents percentage of total reads for each operon that mapped during arsenic respiration. Green represents percentage of total reads for each operon that mapped during fumarate respiration and blue represents percentage of total reads for each operon that mapped during aerobic respiration. Total reads mapped for each operon in all three conditions combined is showed in black bar graph. Operon locus tags: *nqrABCDEF* Shewana3_3373-Shewana3_3378, *nqrABCDEF2* Shewana3_0938- Shewana3_0943, *rnfABCDGE* Shewana3_2168- Shewana3_2173, *ndhII* Shewana3_3123, *srdABCD* Shewana3_1706- Shewana3_1709, *petABC* Shewana3_0601-Shewana3_0603, *cydAB* Shewana3_0927-Shewana3_0928, *cydABX* Shewana3_1241-Shewana3_1243, *cyoABCDE* Shewana3_4031-Shewana3_4035, *ccoPQON* Shewana3_2107-Shewana3_2110, *coxABGC* Shewana3_3991-Shewana3_3994, *frdABCD* Shewana3_0401- Shewana3_0404, *frd* Shewana3_3318, *arrAB* Shewana3_2340-Shewana3_2341. ND stands for none detected

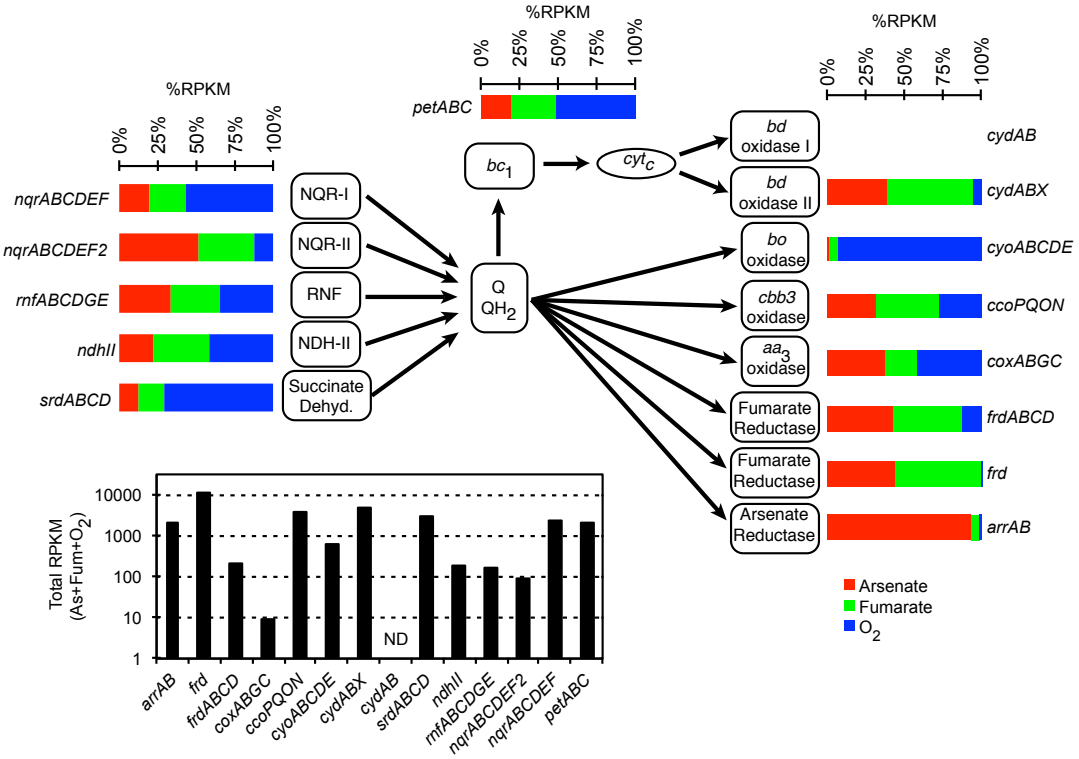


Figure 2.6. Heatmap analysis of genes that encode two MTR gene clusters. RPKM represents Reads per Kilobase per Million reads. In the left heat maps each column represents one of the nine replicates. Each row represents expression values for the specific gene listed on the right. Red color represents high expression (more reads mapped) and blue color represents low expression (less reads mapped). The right heat map represents average Log $_2$ fold change which compares the mean expression levels of the first condition over second condition. Blue color represents a fold change that is higher in the first condition. Red color represents a fold change that is higher in the second condition. Green represents genes that are not differentially expressed between the two conditions. The individual expression values can be obtained from the supplementary data. For the right heatmap, the data represents the $\log_2$ fold-change in expression based on pairwise comparisons of each growth condition, arsenate, fumarate, or aerobic.

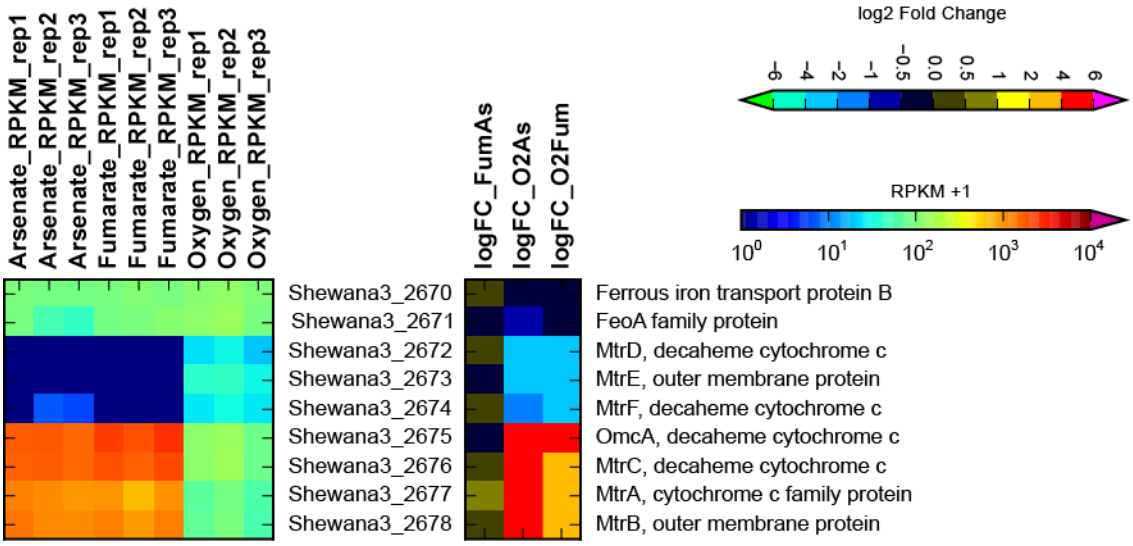


Figure 2.7. Heatmap analysis of differentially expressed genes in arsenate growth conditions. Heatmaps of "arsenic island" and other gene clusters that are expressed in arsenic respiring conditions. Layout and scales are the same as figure 6. **A.** Heatmaps of genomic region known collectively as the arsenic island containing the *arr* and *ars* regulons along with several regulatory genes. **B.** Two gene clusters that encode differentially expressed MTR genes. **C.** Putative cation transport gene cluster expressed in anaerobic conditions.

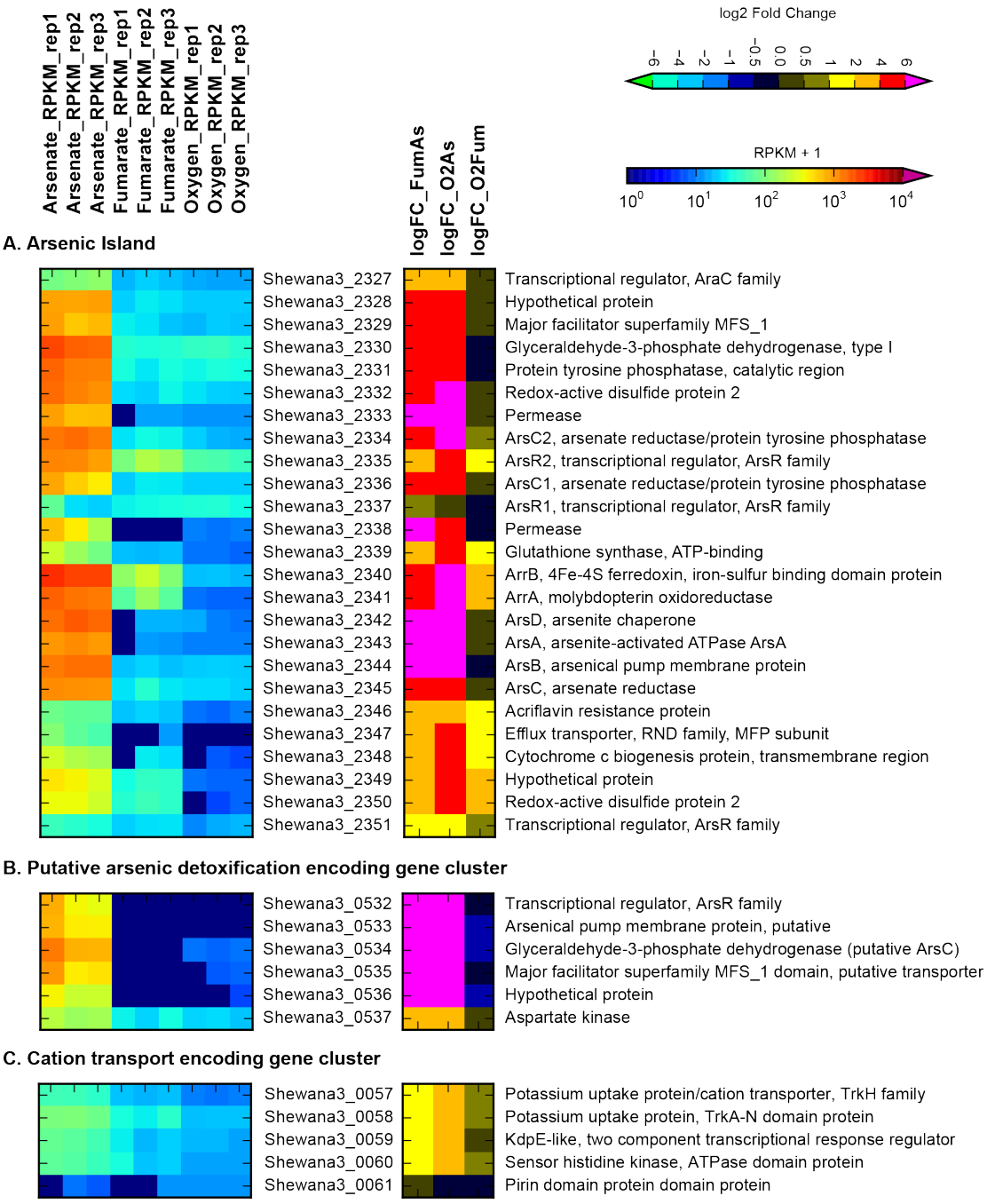

Arsenate_RPKM_rep1
Arsenate_RPKM_rep2
Arsenate_RPKM_rep3
Fumarate_RPKM_rep1
Fumarate_RPKM_rep2
Fumarate_RPKM_rep3
Oxygen_RPKM_rep1
Oxygen_RPKM_rep2
Oxygen_RPKM_rep3
log2 Fold Change
−6 −4 −2 −1 −0.5 0.0 0.5 1 2 4 6
RPKM + 1
10^0 10^1 10^2 10^3 10^4
logFC_FumAs
logFC_O2As
logFC_O2Fum
A. Arsenic Island
Shewana3_2327 Transcriptional regulator, AraC family
Shewana3_2328 Hypothetical protein
Shewana3_2329 Major facilitator superfamily MFS_1
Shewana3_2330 Glyceraldehyde-3-phosphate dehydrogenase, type I
Shewana3_2331 Protein tyrosine phosphatase, catalytic region
Shewana3_2332 Redox-active disulfide protein 2
Shewana3_2333 Permease
Shewana3_2334 ArsC2, arsenate reductase/protein tyrosine phosphatase
Shewana3_2335 ArsR2, transcriptional regulator, ArsR family
Shewana3_2336 ArsC1, arsenate reductase/protein tyrosine phosphatase
Shewana3_2337 ArsR1, transcriptional regulator, ArsR family
Shewana3_2338 Permease
Shewana3_2339 Glutathione synthase, ATP-binding
Shewana3_2340 ArrB, 4Fe-4S ferredoxin, iron-sulfur binding domain protein
Shewana3_2341 ArrA, molybdopterin oxidoreductase
Shewana3_2342 ArsD, arsenite chaperone
Shewana3_2343 ArsA, arsenite-activated ATPase ArsA
Shewana3_2344 ArsB, arsenical pump membrane protein
Shewana3_2345 ArsC, arsenate reductase
Shewana3_2346 Acriflavin resistance protein
Shewana3_2347 Efflux transporter, RND family, MFP subunit
Shewana3_2348 Cytochrome c biogenesis protein, transmembrane region
Shewana3_2349 Hypothetical protein
Shewana3_2350 Redox-active disulfide protein 2
Shewana3_2351 Transcriptional regulator, ArsR family
B. Putative arsenic detoxification encoding gene cluster
Shewana3_0532 Transcriptional regulator, ArsR family
Shewana3_0533 Arsenical pump membrane protein, putative
Shewana3_0534 Glyceraldehyde-3-phosphate dehydrogenase (putative ArsC)
Shewana3_0535 Major facilitator superfamily MFS_1 domain, putative transporter
Shewana3_0536 Hypothetical protein
Shewana3_0537 Aspartate kinase
C. Cation transport encoding gene cluster
Shewana3_0057 Potassium uptake protein/cation transporter, TrkH family
Shewana3_0058 Potassium uptake protein, TrkA-N domain protein
Shewana3_0059 KdpE-like, two component transcriptional response regulator
Shewana3_0060 Sensor histidine kinase, ATPase domain protein
Shewana3_0061 Pirin domain protein domain protein

Figure 2.8. Heatmap analysis of sulfur metabolism genes expressed in response to arsenic.

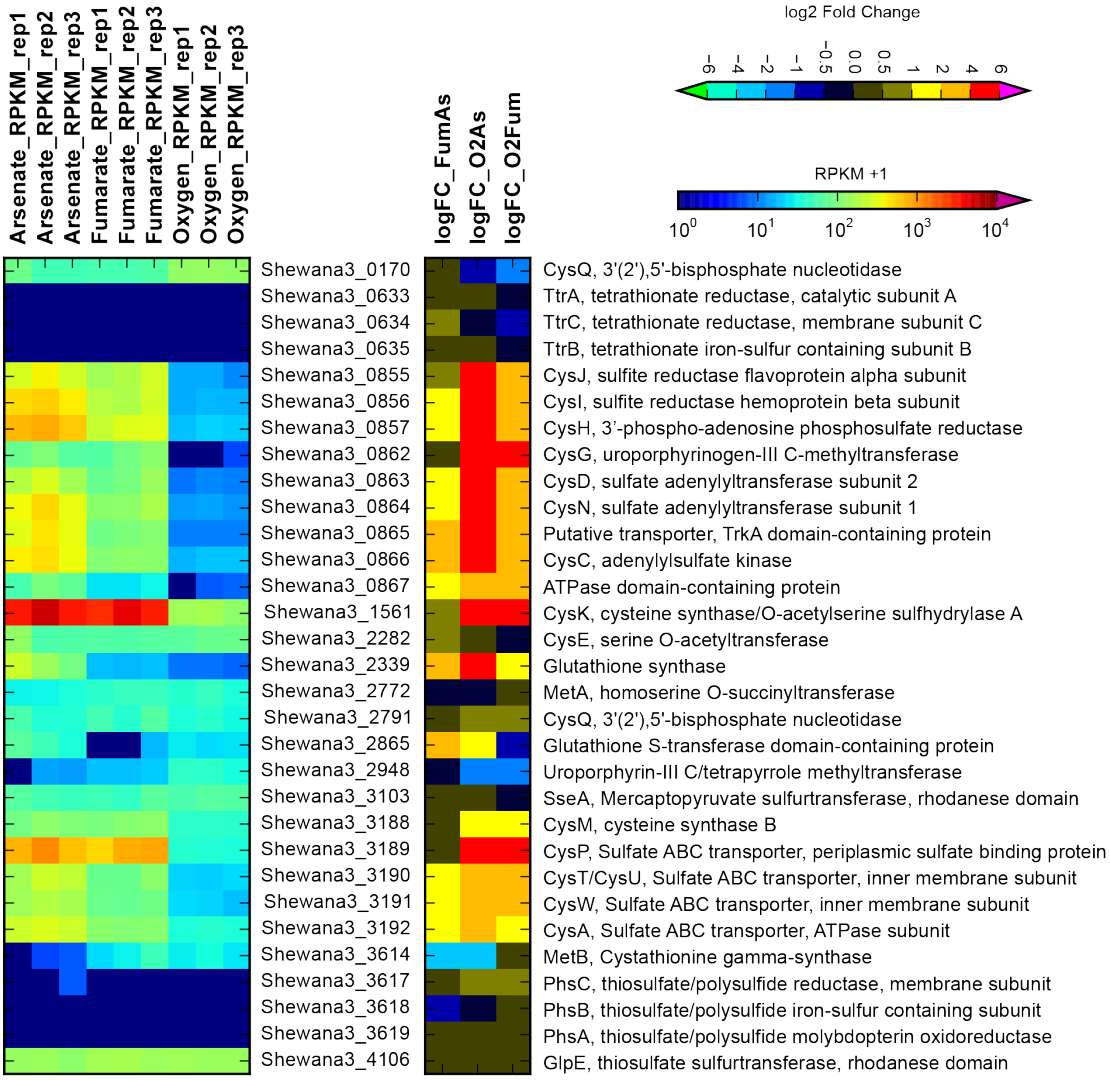


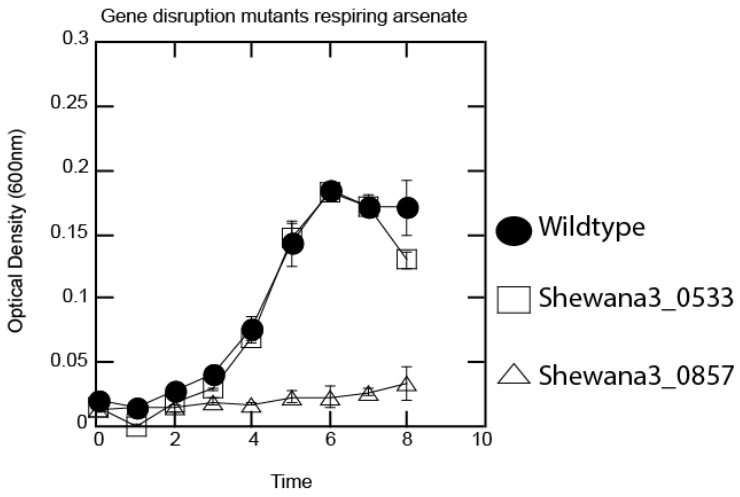
Figure 2.9. Phenotypic analysis of several gene disruption mutants.

A. Anaerobic growth curve of representative gene disruption mutants grown with 10 mM As(V). Wild type is represented by filled circles. The arsenic respiring gene disruption mutant ΔShewana3_0533 (showing similar growth pattern to ΔShewana3_0057, ΔShewana3_0058, ΔShewana3_0060, ΔShewana3_0536, ΔShewana3_1204, ΔShewana3_3226, ΔShewana3_3478) is represented by squares. The growth defective mutant ΔShewana3_0857 (showing similar growth pattern to ΔShewana3_0537, ΔShewana3_2062, ΔShewana3_2864) is represented by triangles.

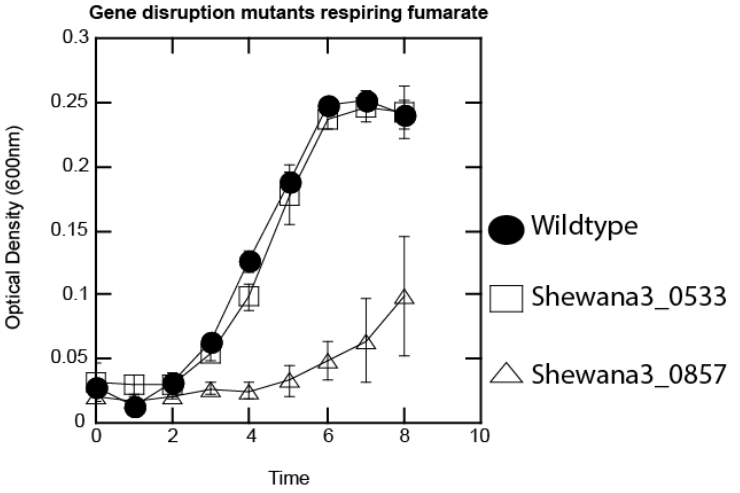
B. Anaerobic growth curve of representative gene disruption mutants grown with 10 mM Fumarate. Wild type is represented by filled circles. The arsenic respiring gene disruption mutant ΔShewana3_0533 (showing similar growth pattern to ΔShewana3_0057, ΔShewana3_0058, ΔShewana3_0060, ΔShewana3_0536, ΔShewana3_1204, ΔShewana3_3226, ΔShewana3_3478) is represented by squares. The growth defective mutant ΔShewana3_0857 (showing similar growth pattern to ΔShewana3_0537, ΔShewana3_2062, ΔShewana3_2864) is represented by triangles.

C. Aerobic growth curve of representative gene disruption mutants grown with (grey) or without (black) 5mM arsenite. Wild type is represented by filled circles. The arsenic respiring gene disruption mutant ΔShewana3_0533 (showing similar growth pattern to ΔShewana3_0057, ΔShewana3_0058, ΔShewana3_0060, ΔShewana3_0536, ΔShewana3_1204, ΔShewana3_3478) is represented by squares. The growth defective mutant ΔShewana3_0857 (showing similar growth pattern to, ΔShewana3_2062, ΔShewana3_2864, ΔShewana3_3226) is represented by triangles. Mutant ΔShewana3_0537 was not tested for arsenite sensitivity. Mutant ΔShewana3_3226 grew well anaerobically but exhibits some degree of arsenite sensitivity.

A.


Gene disruption mutants respiring arsenate
Optical Density (600nm)
0.3
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
10
Time
Wildtype
Shewana3_0533
Shewana3_0857

B.


Gene disruption mutants respiring fumarate
Optical Density (600nm)
0.3
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
10
Time
Wildtype
Shewana3_0533
Shewana3_0857

C.

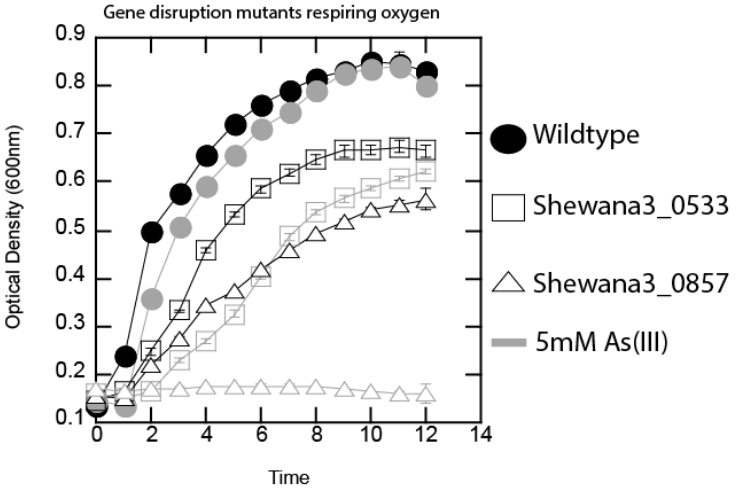

Gene disruption mutants respiring oxygen
Optical Density (600nm)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
4
6
8
10
12
14
Time
Wildtype
Shewana3_0533
Shewana3_0857
5mM As(III)

Table 1. List of arsenic specific transcripts used in q-RT-PCR validation

Gene ID	Description	Arsenic Expression (CLC)	Fumarate Expression (CLC)	Oxygen Expression (CLC)
Shewana3_0536	Hypothetical DUF: near major facilator subfamily	286.91	3.5	5.17
Shewana3_3226	thiol: disulfide interchange	45.729	6.6	34.14
Shewana3_0857	phosphoadenosine phosphosulfate reductase	792.75	298.87	21.12
Shewana3_3478	blue(type 1) copper domain	650.26	38.92	56.45
Shewana3_2512	HtpX heat shock protein	465.27	106.74	97.05

Table 2. Phenotypic analysis of gene disruption mutants

		Growth rate Hours-1			
Locus Tag	Tigr based gene ontology	Arsenic	Fumarate	Oxygen	Oxygen +As(III)
Shewana3_0057	TrkH family potassium uptake protein	1.5 ± 0.39	1.62±0.18	2.32 ±0.04	3.05 ± 0.13
Shewana3_0058	TrkA domain-containing protein	1.6 ± 0.1	1.83 ± 0.09	2.09 ± 0.03	3.36 ±0.07
Shewana3_0060	ATPase domain-containing protein Transport and binding proteins	0.96 ± 0.02	1.0 ±0.1	2.25 ± 0.21	3.38 ± 0.08
Shewana3_0533	bile acid:sodium symporter (arsenical-resistance protein arsB)	0.86 ± 0.03	1.18 ± 0.06	2.59 ± 0.04	3.16 ± 0.20
Shewana3_0536	conserved hypothetical protein	1.0 ± 0.17	1.1 ±0.14	2.19 ± 0.09	2.80 ± 0.07
Shewana3_0537	aspartate kinase	NG	NG	NG	NG
Shewana3_0857	phosphoadenosine phosphosulfate reductase	NG	NG	2.71±0.002	NG
Shewana3_1204	hypothetical protein(Cell envelope (inner membrane protein YicG))	1.1 ± 0.1	1.17 ±0.08	2.63 ± 0.11	3.39 ± 0.11
Shewana3_2062	inner membrane transport protein YdhC	NG	NG	4.00 ± 0.008	NG
Shewana3_2864	TetR family transcriptional regulator	NG	NG	3.33 ± 0.06	NG
Shewana3_3226	thiol:disulfide interchange protein, putative	1.2 ± 0.08	1.21 ± 0.006	2.47 ± 0.01	NG
Shewana3_3478	blue (type1) copper domain-containing protein	1.2 ± 0.35	1.2 ±0.12	2.72± 0.09	3.18 ± 0.07
	ANA-3 Wild Type	1.1 ±0.18	1.3 ±0.04	0.52 ± 0.009	0.65 ± 0.02

Chapter 2 Supplementary figures

Figure 2S1. Principal component analysis of RNAseq Data: Each dot represents expression data for all gene expression data for each replicate in Principal component space.

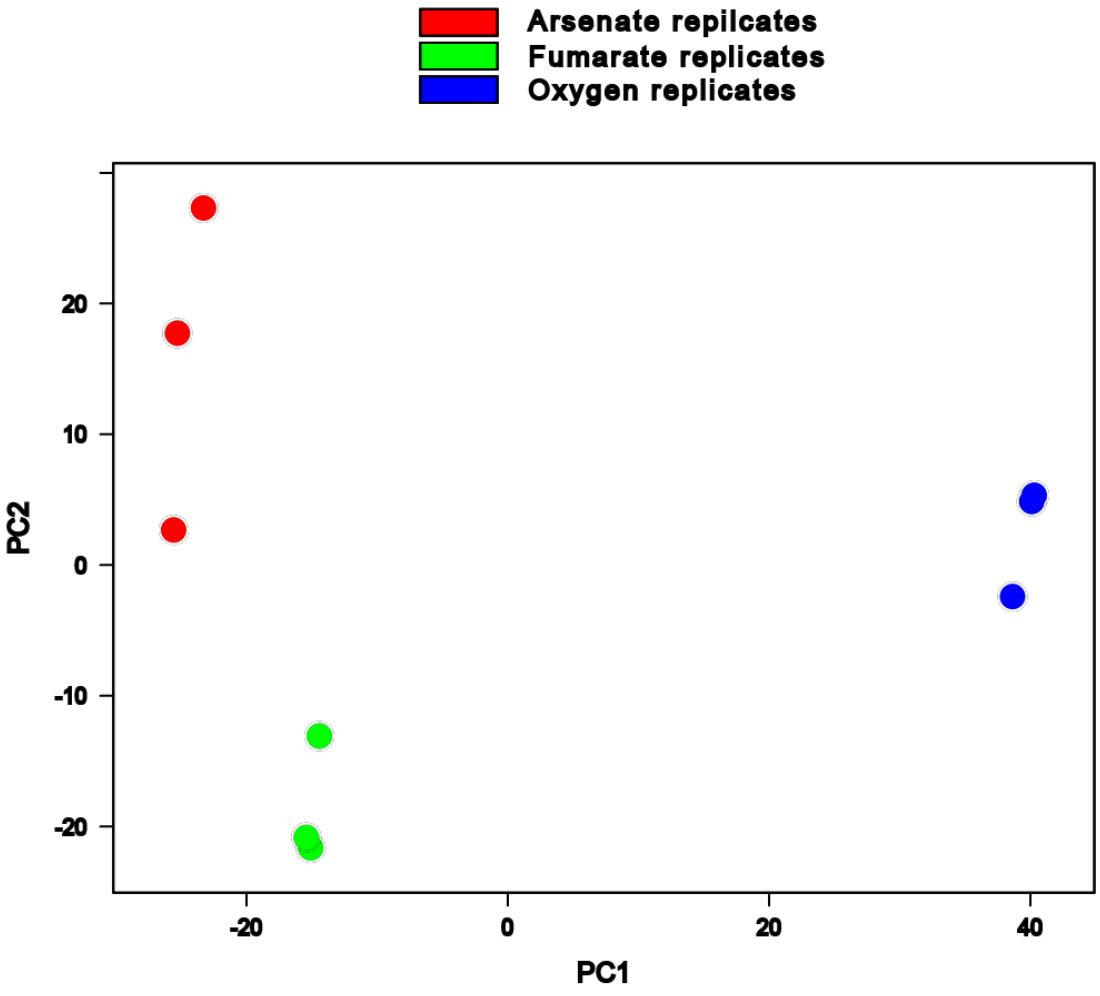
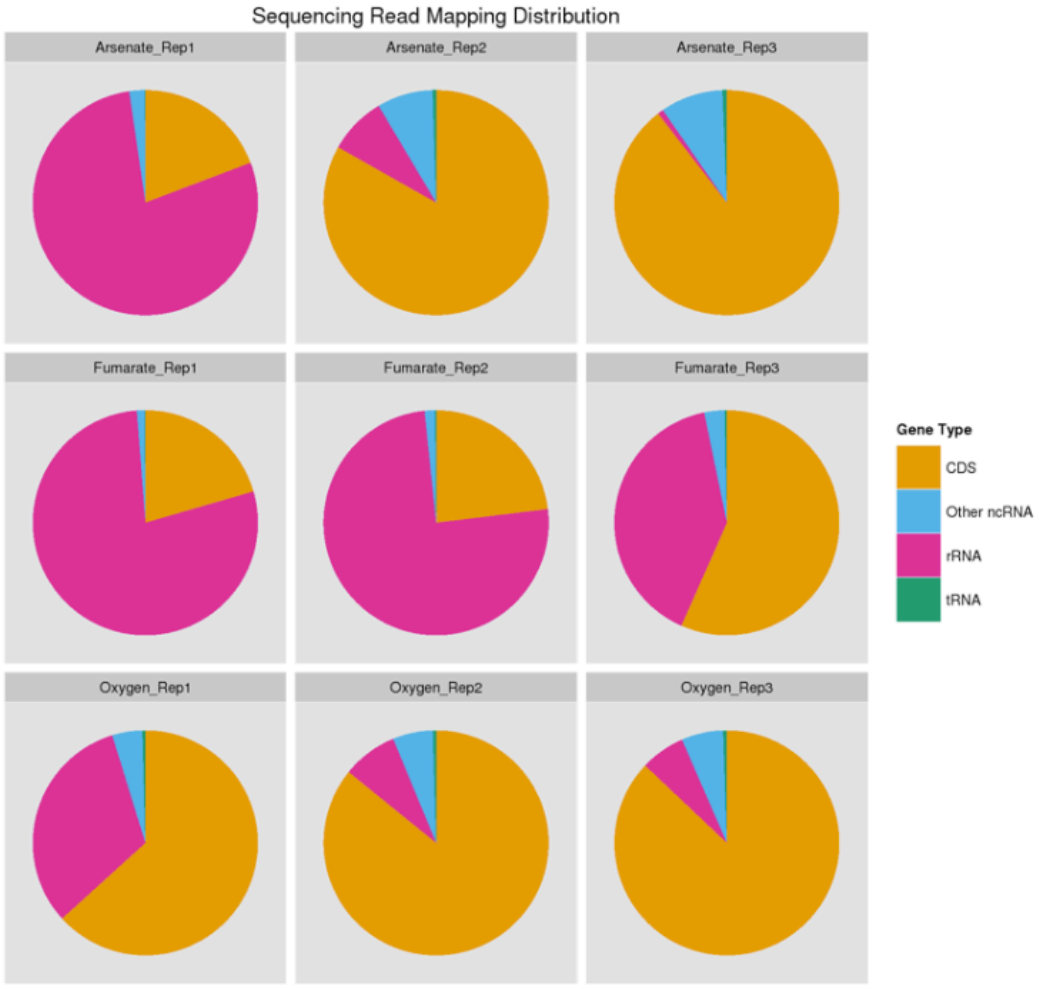


Figure 2S2. Distribution of reads obtained from the illumina sequencer: Orange (CDS) is coding sequence. Blue is other non-coding sequences. Pink (rRNA) is ribosomal RNA, and green (tRNA) is transfer RNA



Supplementary tables

Table 2.S1. Raw reads obtained from the illumina sequencer and number of reads mapped after trimming off adaptors and low quality sequences.

Library (Barcode)	1 Arsenic (CAGATC)	2 Arsenic (ACTTGA)	3 Arsenic (GATCAG)	4 Fumarate (TAGCTT)	5 Fumarate (GGCTAC)	6 Fumarate (CTTGTA	7 Oxygen (TGACCA)	8 Oxygen (ACAGTG)	9 Oxygen (GCCAAT)
Mapped Reads	3,204,554	9,340,131	10,596,348	3,215,149	4,161,453	5,370,704	6,723,937	9,372,423	10,113,043
Un-mapped Reads	16,134,761	5,753,506	5,593,951	17,380,218	18,116,665	7,394,343	7,160,995	4,782,277	5,197,717
Total Reads	19,339,315	15,093,637	16,190,299	20,595,367	22,278,118	12,765,047	13,884,932	14,154,700	15,310,760

Table 2S2. Genes down regulated in arsenate growth conditions relative to fumarate and oxygen. Log_2FC, log_2-fold change. Negative value inducates genes is down regulated in relative to the condition on the right of the colon.

Gene	Product	Log2FC Fum:As	Log2FC O2:As	Log2FC O2:Fum
Shewana3_0442	zinc-responsive transcriptional regulator	-1.2	-1.2	0.0
Shewana3_0565	hypothetical protein	-1.4	-1.8	-0.4
Shewana3_0566	periplasmic solute binding protein	-1.3	-1.9	-0.6
Shewana3_0567	ABC-3 protein	-1.5	-1.7	-0.1
Shewana3_0651	hypothetical protein	-1.3	-1.2	0.1
Shewana3_0918	Na+/H+ antiporter NhaC	-1.7	-1.0	0.6
Shewana3_1018	hypothetical protein	-1.8	-1.1	0.7
Shewana3_1658	RND family efflux transporter MFP subunit	-1.1	-2.0	-0.9
Shewana3_1931	alpha-glucosidase	-1.2	-1.7	-0.5
Shewana3_1959	two component transcriptional regulator	-1.4	-2.1	-0.7
Shewana3_2070	dihydroxy-acid dehydratase	-1.0	-1.6	-0.6
Shewana3_2404	hypothetical protein	-1.4	-1.0	0.4
Shewana3_3031	anti sigma-E protein, RseA	-1.3	-2.1	-0.7
Shewana3_3032	RNA polymerase sigma factor RpoE	-1.4	-1.8	-0.4
Shewana3_3089	amino acid/peptide transporter	-1.4	-1.7	-0.3
Shewana3_3470	LysR family transcriptional regulator	-2.3	-1.5	0.8
Shewana3_3472	TonB-dependent siderophore receptor	-1.3	-1.2	0.1
Shewana3_3613	bifunctional aspartate kinase II/homoserine dehydrogenase II	-1.8	-1.9	0.0
Shewana3_3614	cystathionine gamma-synthase	-2.4	-2.2	0.2
Shewana3_3675	hypothetical protein	-2.7	-2.5	0.2
Shewana3_4179	hypothetical protein	-1.1	-1.2	-0.1
Shewana3_4182	hypothetical protein	-1.8	-1.2	0.7
Shewana3_4219	hypothetical protein	-1.9	-1.6	0.3
Shewana3_4376	hypothetical protein	-1.1	-1.1	0.0

Chapter 2 refrences

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology* 11, R106–R106. doi:10.1186/gb-2010-11-10-r106.

Andres, J., Arsene-Ploetze, F., Barbe, V., Brochier-Armanet, C., Cleiss-Arnold, J., Coppee, J. Y., Dillies, M. A., Geist, L., Joublin, A., Koechler, S., et al. (2013). Life in an Arsenic-Containing Gold Mine: Genome and Physiology of the Autotrophic Arsenite-Oxidizing Bacterium Rhizobium sp. NT-26. *Genome Biology and Evolution* 5, 934–953. doi:10.1093/gbe/evt061.

Andrews S. (2012) FastQC: A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Aronesty E. (2013) Comparison of sequencing utility programs. The Open Bioinformatics Journal 7:1-8.

Bencheikh-Latmani, R., Williams, S. M., Haucke, L., Criddle, C. S., Wu, L., Zhou, J., and Tebo, B. M. (2005). Global Transcriptional Profiling of Shewanella oneidensis MR-1 during Cr(VI) and U(VI) Reduction. *Applied and Environmental Microbiology* 71, 7453–7460. doi:10.1128/AEM.71.11.7453-7460.2005.

Mumford, A.C., Barringer, J.L., Reilly, P.A., Eberi, D.D., Blum, A.E., Young L.Y., Biogeochemical environments of streambed-sediment pore waters with and without arsenic enrichment in a sedimentary rock terrain, New Jersey Piedmont, USA. (2015). Biogeochemical environments of streambed-sediment pore waters with and without arsenic enrichment in a sedimentary rock terrain, New Jersey Piedmont, USA. 505, 1350–1360. doi:10.1016/j.scitotenv.2014.07.104.

Brosnan, J. T., and Brosnan, M. E. (2006). The sulfur-containing amino acids: an overview. *J. Nutr.* 136, 1636S–1640S.

Brutinel, E. D., and Gralnick, J. A. (2012). Anomalies of the anaerobic tricarboxylic acid cycle in Shewanella oneidensisrevealed by Tn-seq. *Molecular Microbiology* 86, 273–283. doi:10.1111/j.1365-2958.2012.08196.x.

Chan PP, Holmes AD, Smith AM, Tran D, and Lowe TM. (2012) The UCSC Archaeal Genome Browser: 2012 update. Nucleic Acid Res. 40:D646-D652.

Charania, M. A., Brockman, K. L., Zhang, Y., Banerjee, A., Pinchuk, G. E., Fredrickson, J. K., Beliaev, A. S., and Saffarini, D. A. (2009). Involvement of a membrane-bound class III adenylate cyclase in regulation of anaerobic respiration in Shewanella oneidensis MR-1. *Journal of Bacteriology* 191, 4298–4306. doi:10.1128/JB.01829-08.

Clark, I. C., Melnyk, R. A., Iavarone, A. T., Novichkov, P. S., and Coates, J. D. (2014). Chlorate reduction in Shewanella algae ACDC is a recently acquired metabolism characterized by gene loss, suboptimal regulation and oxidative stress. *Molecular Microbiology* 94, 107–125. doi:10.1111/mmi.12746.

Cleiss-Arnold, J., Koechler, S., Proux, C., Fardeau, M.-L., Dillies, M.-A., Coppee, J.-Y., Arsène-Ploetze, F., and Bertin, P. N. (2010). Temporal transcriptomic response during arsenic stress in Herminiimonas arsenicoxydans. *BMC Genomics* 11, 709–709. doi:10.1186/1471-2164-11-709.

Cozen, A. E., Weiraeuch, M.T., Pollard, K. S., Bernick, D.L., Stuart, J. M., Lowe, T. M.,Transcriptional map of respiratory versatility in the hyperthermophilic crenarchaeon Pyrobaculum aerophilum (2009). 191, 782–794. doi:10.1128/JB.00965-08.

Cunningham, L., Gruer, M.J., Guest, J.R.,Transcriptional regulation of the aconitase genes (acnA and acnB) of Escherichia coli. (1997). 143 (Pt 12), 3795–3805. doi:10.1099/00221287-143-12-3795.

Dhar, H.K., Zheng, Y., Saltikov, C.W., Radloff, K.A., Mailloux, B. J., Ahmed, K.M., Van Geen, A. Bangladesh. (2011). Microbes enhance mobility of arsenic in pleistocene aquifer sand from Bangladesh. 45, 2648–2654. doi:10.1021/es1022015.

Elias, M., Wellner, A., Goldin-Azulay, K., Chabriere, E., Vorholt, J. A., Erb, T. J., and Tawfik, D. S. (2013). The molecular basis of phosphate discrimination in arsenate-rich environments. *Nature* 490, 134–137. doi:10.1038/nature11517.

Fernandes, A. D., Macklaim, J. M., Linn, T. G., Reid, G., and Gloor, G. B. (2013). ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PLoS ONE* 8, e67019–e67019. doi:10.1371/journal.pone.0067019.

Fuchs, S., Pané-Farré, J., Kohler, C., Hecker, M., and Engelmann, S. (2007). Anaerobic gene expression in Staphylococcus aureus. *Journal of Bacteriology* 189, 4275–4289. doi:10.1128/JB.00081-07.

Gao, H., Barua, S., Liang, Y., Wu, L., Dong, Y., Reed, S., Chen, J., Culley, D., Kennedy, D., Yang, Y., et al. (2010). Impacts of Shewanella oneidensis c-type cytochromes on aerobic and anaerobic respiration. *Microb Biotechnol* 3, 455–466. doi:10.1111/j.1751-7915.2010.00181.x.

Gralnick, J. A., Brown, C. T., and Newman, D. K. (2005). Anaerobic regulation by an atypical Arc system in Shewanella oneidensis. *Molecular Microbiology* 56, 1347–1357. doi:10.1111/j.1365-2958.2005.04628.x.

Grimek, T. L., and Escalante-Semerena, J. C. (2003). The acnD Genes of Shewenella oneidensis and Vibrio cholerae Encode a New Fe/S-Dependent 2-Methylcitrate Dehydratase Enzyme That Requires prpF Function In Vivo. *Journal of Bacteriology* 186, 454–462. doi:10.1128/JB.186.2.454-462.2004.

Hau, H. H., and Gralnick, J. A. (2007). Ecology and biotechnology of the genus Shewanella. *Annu. Rev. Microbiol.* 61, 237–258. doi:10.1146/annurev.micro.61.080706.093257.

Hayashi, M., Hirai, K., and Unemoto, T. (1995). Sequencing and the alignment of structural genes in thenqroperon encoding the Na+-translocating NADH-quinone reductase fromVibrio alginolyticus. *FEBS Lett* 363, 75–77. doi:10.1016/0014-5793(95)00283-F.

Hery, M., Rizoulis, A., Sanguin, H., Cooke, D. A., Pancost, R.D., Lloyd J. R., Microbial ecology of arsenic-mobilizing Cambodian sediments: lithological controls uncovered by stable-isotope probing. (2014). Microbial ecology of arsenic-mobilizing Cambodian sediments: lithological controls uncovered by stable-isotope probing. n/a–n/a. doi:10.1111/1462-2920.12412.

Hughes, M. F. (2002). Arsenic toxicity and potential mechanisms of action. *Toxicol. Lett.* 133, 1–16. doi:10.1016/S0378-4274(02)00084-X.

Jiang, S., Lee, J.-H., Kim, D., Kanaly, R. A., Kim, M.-G., and Hur, H.-G. (2013). Differential arsenic mobilization from As-bearing ferrihydrite by iron-respiring Shewanella strains with different arsenic-reducing activities. *Environ Sci Technol* 47, 8616–8623. doi:10.1021/es400534z.

Konstantinidis, K. T. K., Serres, M. H. M., Romine, M. F. M., Rodrigues, J. L. M. J., Auchtung, J. J., McCue, L.-A. L., Lipton, M. S. M., Obraztsova, A. A., Giometti, C. S. C., Nealson, K. H. K., et al. (2009). Comparative systems biology across an evolutionary gradient within the Shewanella genus. *Audio, Transactions of the IRE Professional Group on* 106, 15909–15914. doi:10.1073/pnas.0902000106.

Kulp, T. R., Hoeft, S. E., Miller, L. G., Saltikov, C., Murphy, J. N., Han, S., Lanoil, B., and Oremland, R. S. (2006). Dissimilatory arsenate and sulfate reduction in sediments of two hypersaline, arsenic-rich soda lakes: Mono and Searles Lakes, California. *Applied and Environmental Microbiology* 72, 6514–6526. doi:10.1128/AEM.01066-06.

Kumar, A., Smith, B., and Borgelt, C. (2004). Dependence relationships between Gene Ontology terms based on TIGR gene product annotations. ... *Workshop on Computational Terminology ...*.

Langmead B and Salzberg SL. (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357-359.

Le Laz, S., Kpebe, A., Bauzan, M., Lignon, S., Rousset, M., and Brugna, M. (2014). A biochemical approach to study the role of the terminal oxidases in aerobic respiration in Shewanella oneidensis MR-1. *PLoS ONE* 9, e86343. doi:10.1371/journal.pone.0086343.

Li H, Handsaker B, Wysoker A, et al. (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25:2078-9.

Li, Y., Hu, Y., Zhang, X., Xu, H., Lescop, E., Bin Xia, and Jin, C. (2007). Conformational fluctuations coupled to the thiol-disulfide transfer between thioredoxin and arsenate reductase in Bacillus subtilis. *Journal of Biological Chemistry* 282, 11078–11083. doi:10.1074/jbc.M700970200.

Luo, Q., Dong, Y., Chen, H., and Gao, H. (2013). Mislocalization of Rieske Protein PetA Predominantly Accounts for the Aerobic Growth Defect of tat Mutants in Shewanella oneidensis. *PLoS ONE* 8, e62064–e62064. doi:10.1371/journal.pone.0062064.

Magoc T, Wood D, and Salzberg SL. (2013) EDGE-pro: Estimated degree of gene expression in prokaryotic genomes. Evolutionary Bioinformatics 9:27-136.

Malasarn, D., Keeffe, J. R., and Newman, D. K. (2008). Characterization of the arsenate respiratory reductase from Shewanella sp. strain ANA-3. *Journal of Bacteriology* 190, 135–142. doi:10.1128/JB.01110-07.

Malasarn, D., Saltikov, C. W., Campbell, K. M., Santini, J. M., Hering, J. G., and Newman, D. K. (2004). arrA is a reliable marker for As(V) respiration. *Science* 306, 455–455. doi:10.1126/science.1102374.

Melo, A. M. P., Bandeiras, T. M., and Teixeira, M. (2004). New insights into type II NAD(P)H:quinone oxidoreductases. *Microbiology and Molecular Biology Reviews* 68, 603–616. doi:10.1128/MMBR.68.4.603-616.2004.

Metalloid resistance mechanisms in prokaryotes (1998). Metalloid resistance mechanisms in prokaryotes. 123, 16–23.

Mukhopadhyay, R., Rosen B. P., Phung, L.T., Silver, S., Microbial arsenic: from geocycles to genes and enzymes (2002). Microbial arsenic: from geocycles to genes and enzymes. 26, 311–325. Available at: http://apps.isiknowledge.com/InboundService.do?product=WOS&action=retrieve&SrcApp=Papers&UT=000177621600006&SID=3EKDGmeoLid45f1b4J5&SrcAuth=mekentosj&mode=FullRecord&customersID=mekentosj&DestFail=http%

3A%2F%2Faccess.isiproducts.com%2Fcustom_images%2Fwok_failed_auth.html.

Morris, R. L., and Schmidt, T. M. (2013). Shallow breathing: bacterial life at low O(2). *Nature Reviews Microbiology* 11, 205–212. doi:10.1038/nrmicro2970.

Murphy, J. N., and Saltikov, C. W. (2009). The ArsR repressor mediates arsenite-dependent regulation of arsenate respiration and detoxification operons of Shewanella sp. strain ANA-3. *Journal of Bacteriology* 191, 6722–6731. doi:10.1128/JB.00801-09.

Murphy, J. N., and Saltikov, C. W. (2007). The cymA gene, encoding a tetraheme c-type cytochrome, is required for arsenate respiration in Shewanella species. *Journal of Bacteriology* 189, 2283–2290. doi:10.1128/JB.01698-06.

Murphy, J. N., Durbin, K. J., and Saltikov, C. W. (2009). Functional roles of arcA, etrA, cyclic AMP (cAMP)-cAMP receptor protein, and cya in the arsenate respiration pathway in Shewanella sp. strain ANA-3. *Journal of Bacteriology* 191, 1035–1043. doi:10.1128/JB.01293-08.

Nealson, K. H., and Scott, J. (2006). "Ecophysiology of the Genus Shewanella," in *The Prokaryotes* (Springer New York), 1133–1151. doi:10.1007/0-387-30746-X_45.

Niemeti, B., Csanaky, I., Gregus, Z., Effect of an inactivator of glyceraldehyde-3-phosphate dehydrogenase, a fortuitous arsenate reductase, on disposition of arsenate in rats. (2006). Effect of an inactivator of glyceraldehyde-3-phosphate dehydrogenase, a fortuitous arsenate reductase, on disposition of arsenate in rats. 90, 49–60. doi:10.1093/toxsci/kfj058.

Park, S. J., Tseng, C. P., and Gunsalus, R. P. (1995). Regulation of succinate dehydrogenase (sdhCDAB) operon expression in Escherichia coli in response to carbon supply and anaerobiosis: role of ArcA and Fnr. *Molecular Microbiology* 15, 473–482. doi:10.1111/j.1365-2958.1995.tb02261.x.

Quinlan AR and Hall IM. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841-2.

Scott, N., Hatlelid, K.M., MacKenzie, N.E., Carter, D.E., Reactions of arsenic(III) and arsenic(V) species with glutathione. (1993). Reactions of arsenic(III) and arsenic(V) species with glutathione. 6, 102–106. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=8448339&retmode=ref&cmd=prlinks.

Tufano, J.K., Reyes, C.R., Saltikov, C.W., Fendorf, S. Reductive processes

controlling arsenic retention: revealing the relative importance of iron and arsenic reduction. (2008). Reductive processes controlling arsenic retention: revealing the relative importance of iron and arsenic reduction. 42, 8283–8289. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19068807&retmode=ref&cmd=prlinks.

Reyes, C., Murphy, J. N., and Saltikov, C. W. (2010). Mutational and gene expression analysis of mtrDEF, omcA and mtrCAB during arsenate and iron reduction in Shewanella sp. ANA-3. *Environmental Microbiology* 12, 1878–1888. doi:10.1111/j.1462-2920.2010.02192.x.

Rodionov, D. A., Novichkov, P. S., Stavrovskaya, E. D., Rodionova, I. A., Li, X., Kazanov, M. D., Ravcheev, D. A., Gerasimova, A. V., Kazakov, A. E., Kovaleva, G. Y., et al. (2011). Comparative genomic reconstruction of transcriptional networks controlling central metabolism in the Shewanella genus. *BMC Genomics* 12 Suppl 1, S3–S3. doi:10.1186/1471-2164-12-S1-S3.

Rodrigues, J. L. M., Serres, M. H., and Tiedje, J. M. (2011). Large-scale comparative phenotypic and genomic analyses reveal ecological preferences of shewanella species and identify metabolic pathways conserved at the genus level. *Applied and Environmental Microbiology* 77, 5352–5360. doi:10.1128/AEM.00097-11.

ROSENTEL, J. K., HEALY, F., MAUPINFURLOW, J. A., LEE, J. H., and SHANMUGAM, K. T. (1995). Molybdate and Regulation of Mod (Molybdate Transport), Fdhf, and Hyc (Formate Hydrogenlyase) Operons in Escherichia-Coli. *Journal of Bacteriology* 177, 4857–4864.

Saffarini, D. A., Schultz, R., and Beliaev, A. (2003). Involvement of cyclic AMP (cAMP) and cAMP receptor protein in anaerobic respiration of Shewanella oneidensis. *Journal of Bacteriology*. doi:10.1128/JB.185.12.3668-3671.2003.

Saltikov, C. REGULATION OF ARSENIC METABOLIC PATHWAYS IN PROKARYOTES. *researchgate.net*

Saltikov, C. W., and Newman, D. K. (2003). Genetic identification of a respiratory arsenate reductase. *Proc Natl Acad Sci U S A* 100, 10983–10988. doi:10.1073/pnas.1834303100.

Saltikov, C. W., Cifuentes, A., Venkateswaran, K., and Newman, D. K. (2003). The ars detoxification system is advantageous but not required for As(V) respiration by the genetically tractable Shewanella species strain ANA-3. *Applied and Environmental Microbiology* 69, 2800–2809. doi:10.1128/AEM.69.5.2800-2809.2003.

Saltikov, C. W., Wildman, R. A., and Newman, D. K. (2005). Expression dynamics

of arsenic respiration and detoxification in Shewanella sp. strain ANA-3. *Journal of Bacteriology* 187, 7390–7396. doi:10.1128/JB.187.21.7390-7396.2005.

Sánchez-Riego, A. M., López-Maury, L., and Florencio, F. J. (2014). Genomic responses to arsenic in the cyanobacterium Synechocystis sp. PCC 6803. *PLoS ONE* 9, e96826–e96826. doi:10.1371/journal.pone.0096826.

Schmieder R and Edwards R. (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863-4.

Shen, S., Li, X.-F., Cullen, W. R., Weinfeld, M., and Le, X. C. (2013). Arsenic Binding to Proteins. *Chem. Rev.* 113, 7769–7792. doi:10.1021/cr300015c.

Silver, S., and Phung, L. T. (2005). Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Applied and Environmental Microbiology* 71, 599–608. doi:10.1128/AEM.71.2.599-608.2005.

Campbell, K.M., Malasaran, D., Saltikov, C.W., Newman D. K., Hering., J.G., Simultaneous microbial reduction of iron(III) and arsenic(V) in suspensions of hydrous ferric oxide. (2006). Simultaneous microbial reduction of iron(III) and arsenic(V) in suspensions of hydrous ferric oxide. 40, 5950–5955. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17051784&retmode=ref&cmd=prlinks.

Sîrbu, A., Kerr, G., Crane, M., and Ruskin, H. J. (2012). RNA-Seq vs Dual- and Single-Channel Microarray Data: Sensitivity Analysis for Differential Expression and Clustering. *PLoS ONE* 7, e50986–e50986. doi:10.1371/journal.pone.0050986.

Unemoto, T., and Hayashi, M. (1993). Na(+)-translocating NADH-quinone reductase of marine and halophilic bacteria. *J Bioenerg Biomembr* 25, 385–391.

Vasin, A., Klotchenko, S., and Puchkova, L. (2013). Phylogenetic Analysis of Six-Domain Multi-Copper Blue Proteins. *PLoS Curr*. doi:10.1371/currents.tol.574bcb0f133fe52835911abc4e296141.

Zargar, K., Hoeft, S., Oremland, R., and Saltikov, C. W. (2010). Identification of a novel arsenite oxidase gene, arxA, in the haloalkaliphilic, arsenite-oxidizing bacterium Alkalilimnicola ehrlichii strain MLHE-1. *Journal of Bacteriology* 192, 3755–3762. doi:10.1128/JB.00244-10.

Zhang, S. Y., Zhao, F.J., Sun, J.Q., Yang, X. R., Li, H., Zhu, Y. G., Diversity and abundance of arsenic biotransformation genes in paddy soils from southern china. (2015). Diversity and abundance of arsenic biotransformation genes in paddy soils from southern china. 49, 4138–4146. doi:10.1021/acs.est.5b00028.

Zhou, G., Yin, J., Chen, H., Hua, Y., Sun, L., and Gao, H. (2013). Combined effect of loss of the caa3 oxidase and Crp regulation drives Shewanella to thrive in redox-stratified environments. *ISME J* 7, 1752–1763. doi:10.1038/ismej.2013.62.

Zhu, W. M., Wilks, A., and Stojiljkovic, I. (2000). Degradation of heme in gram-negative bacteria: the product of the hemO gene of Neisseriae is a heme oxygenase. *Journal of Bacteriology* 182, 6783–6790.

Zegers,I., Martins, J. C., Willem, Wyns, LA Messenren., (2001). Arsenate reductase from S. aureus plasmid pI258 is a phosphatase drafted for redox duty. 8, 843–847. doi:10.1038/nsb1001-843.

Chapter 3. The role of cAMP in anaerobic respiration in *Shewnella* sp. ANA-3

Abstract

Arsenic is a ubiquitous toxin found in low levels worldwide. However, respiratory arsenic reducing microbes have been shown to contribute to the release of harmful levels of this metalloid into drinking water aquifers. In Shewanella sp. ANA-3, the model for respiratory arsenic reduction, this reductase is not always active and its expression depends on the activity of the cyclic AMP receptor protein (CRP). CRP is a global transcriptional regulator that is activated in the presence of sufficient cAMP. In this study the role of cAMP in signaling the expression of the respiratory arsenic reductase is assessed.

Introduction

Bacterial respiratory arsenic reduction causes harmful levels of arsenic to leach into drinking water (Barringer and Reilly 2013). However, arsenic contamination of ground water is difficult to detect leaving millions of people worldwide vulnerable to arsenic poisoning. It is important to understand the conditions that favor microbial arsenic reduction, which will enable better detection methods and will inform predictions of when microbes are most likely to reduce the toxins.

Respiratory arsenic reduction only happens under a strict set of environmental conditions. Arsenic respiration only takes place in environments where there is a useable carbon source and a lack of a more suitable redox substrate (Kulp, Hoeft and Oremland., 2004). For example, when there is oxygen available, a more favorable terminal electron acceptor, arsenic respiration does not take place and the expression of the respiratory arsenic reductase *arrAB* is regulated accordingly. During aerobic respiration, the *arrAB* genes are not transcribed and therefore the reductase is never made.

In *Shewanella,* several studies have pursued the mechanisms used to sense aerobic vs. anaerobic conditions and regulate gene expression of the respective terminal reductases. Studies of both ANA-3 and *S. oneidensis* found that the two known redox sensing anaerobic transcriptional regulators ArcA and EtrA are not essential for anaerobic respiration (Gralnick, Brown and Newman., 2005, Murphy et al., 2009).

Conversely, in both ANA-3 and *Shewanella oneidensis,* a closely related metal reducer, the cyclic AMP receptor protein is essential for expressing anaerobic terminal reductases (Saffarini, Schultz and Beliaev, 2003). In *S. oneidensis* exogenous cAMP is sufficient to activate the terminal reductases aerobically Indicating the importance of cAMP in regulating these reductases (Charania et al., 2009). How cAMP and CRP work together to regulate anaerobic respiration is not clear since this system is not know to sense redox conditions directly. In both biochemical and genetic studies the cyclic AMP receptor protein (CRP) was shown to be vital for anaerobic respiration in *Shewanella* (Murphy et al., 2009, Charania et al., 2009). However, CRP lacks the sensory domains that other anaerobic regulators ArcA and FNR would use to identify aerobic/ anaerobic conditions. Therefore, some other factor may be acting on CRP to provide the sensory signal to activate anaerobic respiration. The activity of CRP is dependent upon cellular cAMP levels; cAMP's ability to serve as the cellular signal of change in the availability of oxygen or anaerobic terminal electron acceptors has not previously been tested. In this study cAMP is assessed as the possible trigger for the redox specific activation of the arsenic reductase and other anaerobic reductases by measuring how the arsenic reductase responds to cAMP and characterize the changes in cAMP concentrations in the cell during aerobic and anaerobic respiration.

Methods

Growth conditions.

Shewanella sp. ANA-3 (ANA-3) was grown overnight in Luria-Bertani medium then diluted down to an Optical Density of 0.6 at 600 nM wavelength to inoculate growth experiments at a ratio of 1:100. Cultures were then grown at 30° C in the minimal salts media described in Murphy and Saltikov, 2007. In aerobic conditions, cultures were maintained in a shaking incubator at 250 RPM and either 10mM arsenic or 10mM cAMP was added. For anaerobic experiments, medium was prepared then dispensed into Balch tubes. The media was then made anaerobic by bubbling under a stream of N_2 for ten minutes. Next the tubes were amended with either 10mM of an anaerobic terminal electron acceptor from an anaerobic stock using N_2 filled syringe. Finally, the tubes were sealed with butyl rubber stoppers and sterilized in the autoclave. Cyclic AMP (cAMP) was added from a freshly prepared 100mM stock immediately before growth experiments.

Gene expression analysis.

Wild-type *Shewanella* sp. ANA-3 or an adenylate cyclase mutant was grown up either aerobically or anaerobically until their optical density read at 600nM wavelength read exactly 0.1. One mL of culture was then spun down and stored at -80° C until RNA extraction could be performed simultaneously. RNA was extracted using the standard protocol outlined in the Qiagen RNeasy mini kit. Next 16 uL of each RNA extract was DNase treated using the Promega RQ1 DNase with the standard protocol suggested by Promega. Then cDNA was synthesized using the

Applied Biosystems TaqMan Reverse Transcriptase kit. DNase treatment was verified with standard PCR using the arrA-F4 primers with non Reverse Transcriptase treated cDNA. Finally all cDNA was diluted down to 1 nanogram and gene expression was quantified by measuring cycle threshold of the SyberGreen *Taq* polymerase reaction on the MJ Research Opticon2.

PCR conditions: 95°C 10 minutes (95°C 30 Seconds,60°C 60 Seconds) X40 cycles

Primers ArrA-F4 Forward AATGGTCAGATACCTCACCGCAG

ArrA-F4 Reverse GCTATTCCACACCCCTTTTTGC

GyrB-F1 Forward ACGAGCGTGACAATTAAGAATGA

GyrB-R1 Reverse ACGTCTTTGTTTACTGGCGTTT

Serial 10-fold dilutions of 1- 0.001 nanograms of genomic DNA was used to produce a standard curve of the Cycle Threshold. Expression was calculated by Cycle Threshold normalized to DNA Gyrase expression (measured with the GyrB primer set)

Growth Experiments.

The adenylate cyclase mutant described in Murphy Durbin and Saltikov, 2009 (AN-CYA-ALL) was grown anaerobically in TME with either arsenic or fumarate supplemented with 10, 1, 0.1, 0.01, 0.001 or 0 mM cyclic AMP and 20mM lactate as the carbon source. Growth was monitored over a 12-hour time course. For cAMP measurements in wild-type cells Shewanella sp. ANA-3 was grown either aerobically by shaking at 250 RPM or anaerobically in the same TME media supplemented with

10mM of an alternative terminal electron acceptor. The carbon source was 20mM Sucrose or lactate in the cAMP measurement experiments.

Cyclic AMP measurements.

In each condition cells were grown up to an optical density of 0.1 at 600nM then immediately flash frozen with dry ice until all other cultures also reached 0.1. for the experiments measuring cAMP with growth on several different terminal electron acceptors all cultures were then thawed and serially diluted 1X10^-6. cAMP concentrations were measured for 180ul samples using the standard non-acetylation protocol for total cellular cAMP with the GE Healthcare Life Science "cAMP Direct Biotrak™ EIA" system in duplicate for triplicate biological replicates of each condition making a total of 6 replicates per sample. For the experiments measuring difference in cAMP based on carbon source cells were not diluted before adding 180 ul of cells to the assay. Final cAMP concentration was calculating by determining the slope of the line of best fit for each set of standards and using this formula to calculate the concentration of each replicate. After concentration was determined for each replicate they were averaged. The standard deviation represents the difference between replicates for each experimental condition.

Results/ discussion

Transcriptional regulation of the arsenic reductase by cyclic AMP

Previously the cyclic AMP receptor protein was shown to be necessary for anaerobic growth on arsenic in *Shewanella* sp. ANA-3 (Murphy et al., 2009). The CRP is also necessary for induction of other anaerobic reductases in *Shewanella oneidensis* MR-1 (Charania et al., 2009). In vitro analysis of the intragenic region upstream of the arsenic reduct determined that there may be a CRP binding site upstream of the anaerobic arsenic reductase. Therefore we wanted to test if the arsenic reductase is regulated directly by CRP or is the lack of anaerobic growth of arsenate caused by some pleotropic effect of mutating this global transcriptional regulator. To test the ability of CRP to directly regulate the activity of the arsenic reductase exogenous cAMP was used to trigger the activity of CRP in an adenylate cyclase mutant. By measuring the gene expression of *arrA,* the first subunit of the arsenic reductase, under various growth conditions we got a clearer idea of the role of cAMP and CRP in directly regulating this reductase (Figure 1). We found that adding 10mM cAMP was sufficient to activate the arsenic reductase in arsenic respiring conditions to wild-type levels. However, adding 10mM cAMP to wild-type cells was not sufficient to overexpress the arrA reductase. These finding indicate that CRP binding is directly involved in regulating the arsenic reductase but there may be some other levels of regulation preventing the promiscuous expression of the arrAB operon.

Aerobic regulation of the arsenic reductase

Since the arsenic reductase is not made during aerobic respiration we wanted to also test if this lack of expression is due to lower cAMP levels. To test for cAMP's ability to express the arsenic reductase aerobically exogenous cAMP was added to the

adenylate cyclase mutant then the expression of *arrA* was measured aerobically in the presence of cAMP (Figure 1). Adding cAMP was not sufficient to induce the expression of the arsenic reductase aerobically. The ArsR repressor is also known to bind the intergenic region near the promoter of *arrA* when there is no arsenic present (Murphy and Saltikov, 2009B). Therefore we also measured the expression of the arsenic reductase aerobically in the adenylate cyclase mutant with both 10mM cAMP and 10mM arsenic added. Still, cAMP and arsenic together were not sufficient to express the arsenic reductase on oxygen respiring conditions above wild-type levels and not be overexpressed even in wild-type with both substrates added in excess. This indicates that there may also be additional factors preventing the expression of the arsenic reductase during aerobic respiration as well such as the degradation of cAMP.

Minimum cAMP requirements for anaerobic growth

An adenylate cyclase mutant is capable of growing aerobically with no additional cAMP however the adenylate cyclase mutant does not grow anaerobically on arsenic. In the case of carbon catabolite repression, altering the concentration of cAMP is sufficient to activate different sets of catabolic genes. In order to determine if a similar mechanism of altering cAMP concentrations is used to activate anaerobic respiration the minimum amount of cAMP needed for growth on arsenic was defined (Figure 2). When an adenylate cyclase mutant was grown anaerobically on arsenic the cells required at least 1mM of cAMP added to the media to grow although at least 10mM cAMP was required for the cells to grow to a similar cell density to wild-type (Figure 2A). The adenylate cyclase mutant was grown anaerobically on fumarate test

if the cAMP dependent activation of a terminal reductase is unique to arsenic respiring conditions or if this was a global mechanism of regulating anaerobic respiration. When the adenylate cyclase mutant was grown anaerobically with fumarate at least 1mM of cAMP was also required for growth on fumarate whereas 10mM cAMP allowed the fumarate respiring cells to grow robustly. Media amended with 0.1mM cAMP or less did not grow in either arsenic or fumarate respiring condition. This suggests that cells grown on either arsenic or fumarate reductases have similar cAMP requirements. This is not suprizing since arsenic and fumarate have similar redox potentials.

cAMP production during respiration

It is possible that the cell increases the amount of available cAMP in anaerobic conditions in order to activate the anaerobic reductases based on their redox potential much like carbon catabolite repression modulates cAMP levels to use the most favorable carbon sources first. To determine if different amounts of cAMP are produced based on redox potential of the terminal electron acceptor total cAMP was quantified for cells grown with oxygen or anaerobically using nitrate, arsenic, fumarate, or iron citrate (figure 3). Surprisingly there was very little difference in total cAMP levels between cells grown on lactate in either aerobic or anaerobic conditions. Since there was not a clear relationship between the amount of cAMP measured in each experimental condition and the ability to respire each substrate we wanted to know if modulating cAMP levels had any relationship to respiring terminal electron acceptors. We grew ANA-3 in conditions known to decrease cAMP levels in

the cell. When ANA-3 was cultured anaerobically on either succinate or sucrose it did not respire either arsenic or fumarate. Figure 3B shows that sucrose does decrease cAMP levels in ANA-3. Taken together this study does show that there is a minimum requirement for cAMP during anaerobic respiration to activate terminal reductases. Interestingly this work also shows that there are also other regulatory requirements to activate anaerobic reductases since cAMP is not sufficient to activate these reductaces alone.

Future directions

In the RNAseq analysis several other genes were differentially expressed during aerobic versus anaerobic respiration. It will be interesting to investigate the cAMP phospodiesterase that was specifically down regulated

Chapter 3 Figures

Figure 3.1. q-RTPCR expression of arrA in adenylate cyclase mutant. Gene expression of the *arrA* subunit of the respiratory arsenic reductase from the adenylate cyclase mutant grown aerobically or anaerobically with or without the addition of 10mM cAMP. Gene expression is measured in the adenylate cyclase mutant in all aerobic conditions. In the anaerobic conditions either adenylate cyclase mutant or wild type expression was measured. All expression data was normalized to DNA gyrase. Error bars represent three technical replicates for each experiment grown in triplicate.

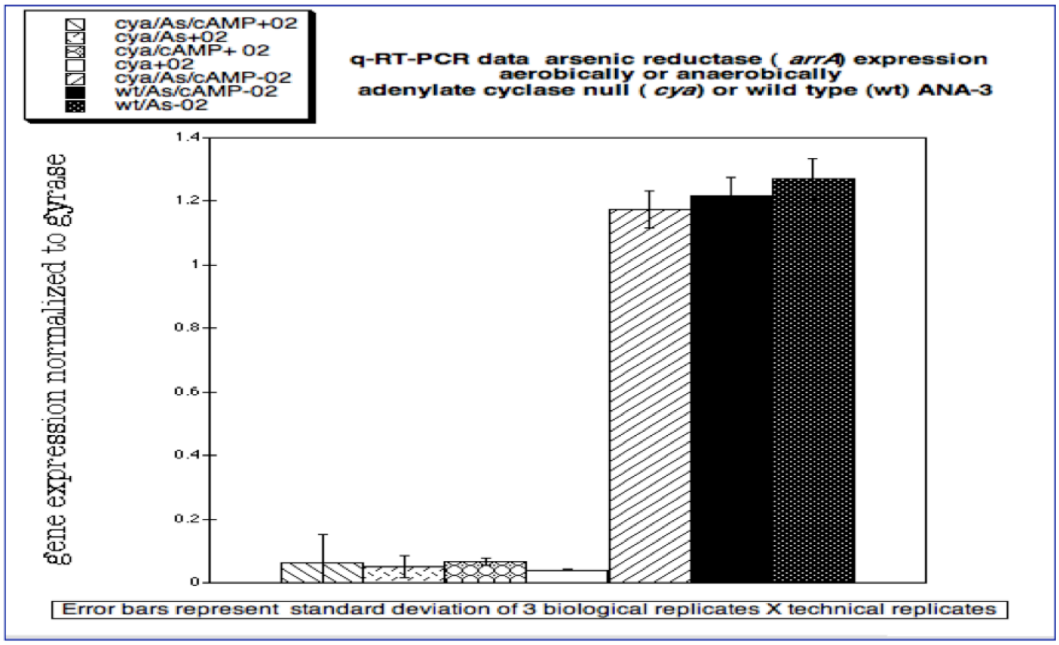


Figure 3.2. Minimum cAMP required for growth on arsenate or fumarate. Anaerobic growth of the adenylate cyclase mutant respiring arsenic or fumarate. Cells were grown with the addition of 10mM, 1mM, or 0.1mM cAMP. Each line represents triplicate growth experiments. The more intense color representing higher concentration of cAMP. Growth is measured by optical density at 600nM over 12 hours.

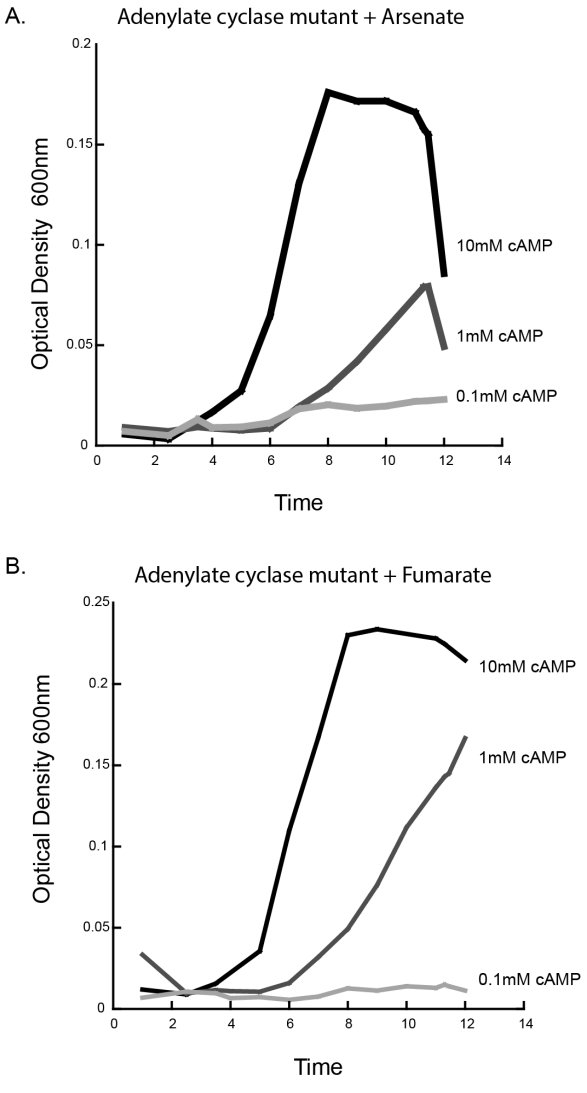
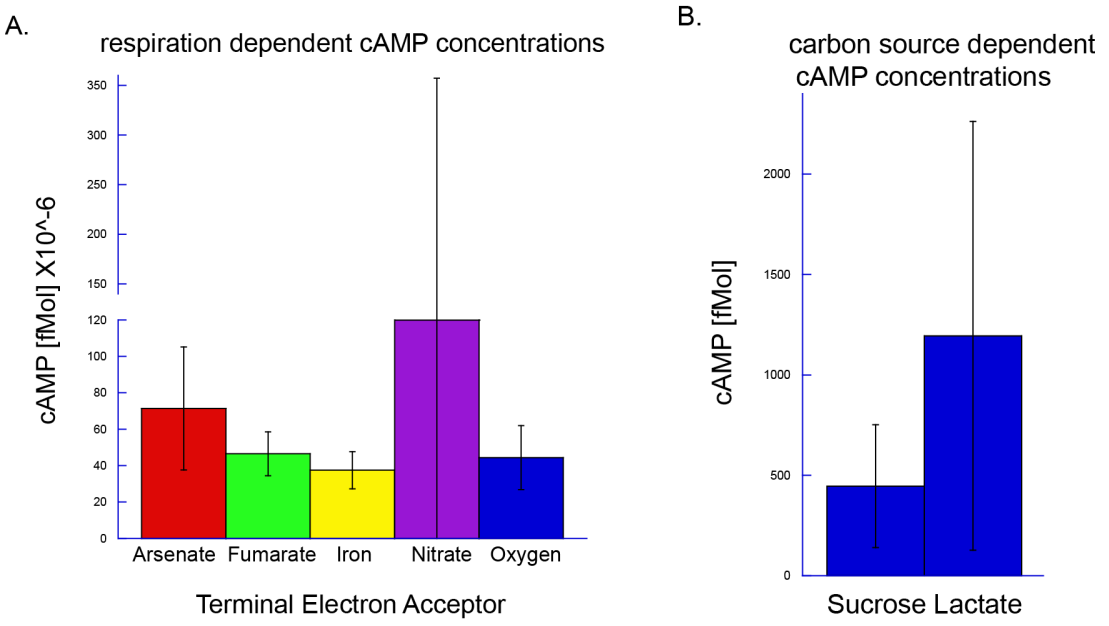


Figure 3.3. Cellular cAMP concentrations based on terminal electron acceptor or carbon source. Bar graphs of cyclic AMP levels from cells grown in the redox conditions listed below. A) using lactate as a carbon source cells respired arsenic, fumarate, iron citrate, nitrate, or oxygen. The cAMP concentratin was determined using a standard curve of optical density from know cAMP concentration. Cells were diluted down 1X10^-6. Error bars represent experiments were performed in three biological replicates measured based on an EIA immunoassay in two technical relicates. Error bars represent the standard deviation of these measurements. B) Cells were grown aerobically either with lactate or sucrose and cAMP measurements are displayed in a similar fashion to 3A only these cells were not previously diluted.



Chapter 3 references

Barringer, J. L., & Reilly, P. A. (2013). Arsenic in groundwater: a summary of sources and the biogeochemical and hydrogeologic factors affecting arsenic occurrence and mobility. http://doi.org/10.5772/5535

Charania, M. A., Brockman, K. L., Zhang, Y., Banerjee, A., Pinchuk, G. E., Fredrickson, J. K., et al. (2009). Involvement of a membrane-bound class III adenylate cyclase in regulation of anaerobic respiration in Shewanella oneidensis MR-1. *Journal of Bacteriology*, *191*(13), 4298–4306. http://doi.org/10.1128/JB.01829-08

Gralnick, J. A., Brown, C. T., & Newman, D. K. (2005). Anaerobic regulation by an atypical Arc system in Shewanella oneidensis. *Molecular Microbiology*, *56*(5), 1347–1357. http://doi.org/10.1111/j.1365-2958.2005.04628.x

Kulp, T. R., Hoeft, S. E., & Oremland, R. S. (2004). Redox transformations of arsenic oxyanions in periphyton communities. *Applied and Environmental Microbiology*, *70*(11), 6428–6434. http://doi.org/10.1128/AEM.70.11.6428-6434.2004
Measurements of Rates of Adenosine 3“:5-”Cyclic Monophosphate Synthesis in Intact Escherichia coli B. (1999). Measurements of Rates of Adenosine 3“:5-”Cyclic Monophosphate Synthesis in Intact Escherichia coli B, 1–4.

Murphy, J. N., Durbin, K. J., & Saltikov, C. W. (2009). Functional roles of arcA, etrA, cyclic AMP (cAMP)-cAMP receptor protein, and cya in the arsenate respiration pathway in Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *191*(3), 1035–1043. http://doi.org/10.1128/JB.01293-08
Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in Streptomyces coelicolor. (1998). Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in Streptomyces coelicolor, 1–14.

Saffarini, D. A., Schultz, R., & Beliaev, A. (2003). Involvement of cyclic AMP (cAMP) and cAMP receptor protein in anaerobic respiration of Shewanella oneidensis. *Journal of Bacteriology*. http://doi.org/10.1128/JB.185.12.3668-3671.2003

Schwab, C., Walter, J., Tannock, G. W., Vogel, R. F., & Gänzle, M. G. (2007). Sucrose utilization and impact of sucrose on glycosyltransferase expression in Lactobacillus reuteri. *Systematic and Applied Microbiology*, *30*(6), 433–443. http://doi.org/10.1016/j.syapm.2007.03.007
Soria, S., de Anda, R., Flores, N., Romero-Garcia, S., Gosset, G., Bolívar, F., & Báez-Viveros, J. L. (2012). New insights on transcriptional responses of genes involved in carbon central metabolism, respiration and fermentation to low ATP

levels in Escherichia coli. *Journal of Basic Microbiology*, *53*(4), 365–380. http://doi.org/10.1002/jobm.201100525
The relationship between external glucose concentration and cAMP levels inside. (2008). The relationship between external glucose concentration and cAMP levels inside, 1–10.

Chapter 4. Isolation of an arsenic respiring microbe from Mono Lake, CA

Abstract

Over the last 20 years *Shewanella* sp. ANA-3 has been developed as a model to study the genetic mechanisms of respiratory arsenic reduction. However the use of ANA-3 as a model of environmental arsenic release may become problematic since ANA-3 was originally isolated from a marine environment yet microbes that are associated with environmental release of arsenic are typically found in freshwater aquifers and soda lake environments (Lloyd and Oremland, 2006). In this study I sought to identify a novel arsenic respiring microbe directly from an arsenic rich soda lake in Mono county California that is easily grown in lab and possibly genetically tractable as well.

Introduction

Arsenic reducing microbes are studied for their contributions to the release of arsenic into the environment. Dissimilatory arsenic respiration is one microbial process that reduces arsenic converting the generally mineral bound form (arsenate) to a more soluble form (arsenite)(Oremland and Stolz, 2003). This more water soluble arsenic is more likely to seep into and contaminate ground water aquifers depending on pH and adsorption to other geochemical features (Barringer and Reilly, 2013, Smedley and Kinniburgh, 2002). Therefore it is important to study the mechanisms of arsenic reduction by microbes to avoid human exposure to arsenic.

Currently Shewanella sp. ANA-3 is the only genetically tractable model for studying the genetic and biochemical mechanisms of dissimilatory arsenic respiration. ANA-3 was isolated from an artificially arsenic rich environment not a complex naturally arsenic rich ecosystem (Saltikov et al., 2003). Nonetheless, ANA-3 contains the same respiratory arsenic reductase enzymes found in soil samples associated with contaminated ground water aquifers (Malasarn et al., 2004); unlike the uncultured microbes found in these environments ANA-3 is a quickly growing facultative anaerobe that is readily amenable to genetic experiments that require mating with *E.coli*. Whereas many environmental microbes require strict anaerobic conditions, non-neutral pH, or high temperatures making it difficult to complete genetic experiments that require vector mediated mutations. Ideally we would find a microbe that can grow quickly under similar conditions to *E.coli* making it easier to develop a

genetic system. Once a microbe that is genetically amenable like ANA-3 but actually from naturally arsenic rich environments it can be used to test the longstanding hypothesis that ANA-3 uses a similar mechanism to environmentally relevant arsenic reducers can be tested.

Lee Vining, California is the site of Mono Lake, an arsenic rich soda lake amidst many hot springs (Bischoff, 2006). Mono Lake is teeming with microbial life and is a model for the biogeochemical cycling of arsenic due to microbial processes (Lloyd and Oremland, 2006, Kulp et al., 2006). Many of the microbes isolated from here require high temperatures, pH or salinity making them challenging to grow in lab as well (Blum et al., 1998). On the other hand, Navy Beach hot spring is located just yards from the coast of Mono Lake (Bischoff, 2006). Navy beach however has pockets of water of lower temperature and pH relative to the lake itself, which increases the likelihood of isolating an arsenic respiring microbe that is more amenable to laboratory conditions.

Wild Willy (Crowley) hot spring, a long valley hot spring, is located about 10 miles away from the coast of Mono Lake and also has similar geochemical features to navy beach and Mono Lake (Bischoff, 2006) and presumably is also arsenic rich (Oremland, Stolz and Hollibaugh, 2006). In this study I enriched for arsenic respiring microbe from 10 environmental samples. Eight were from Navy beach hot spring and two from Wild Willy in hopes to identify a more environmentally relevant arsenic reducer to compare to ANA-3.

Methods

Initial sample description.

A 50 Ml conical tube was filled to the brim with a soil core found in a neutral pH pond measuring about 20 inches wide on the coast of Navy Beach measuring 30° C. A second conical tube was also filled to the brim with water obtained from this pond. Samples were also taken from either water or soil collected from Wild Willy hot spring that met the same criteria. All samples were transported in a cooler then stored at 4 ° c until enrichment experiments could begin.

Enrichment Conditions.

One gram of soil or 1 ml of water from each sample was added to a separate Hungate tube containing anaerobic TME media without yeast extract amended with 20mM lactate as an electron donor 10 mM arsenate as an electron acceptor and 1mM L-cysteine to visualize arsenic reduction. These tubes were sealed in a coy glovebox in a nitrogen atmosphere then stored at 30° C.

Culture maintenance and isolation.

The soil settled over time, and after 24 hours the media appeared turbid and a yellow color was observed within the sediment on the bottom of the tubes. Without agitating the cultures 1 ml of supernatant was transferred into a new Hungate tube with the same media using an anaerobic sterile syringe and stored at 30°C in the dark. All samples were passaged again three days later. After eight more days they were passaged again in the same manner to ensure no soil particles were being transferred.

After one day a sterile loop was inserted into each culture and used to streak TME agar plates with 20mM lactate and 10mM arsenate. These plates were stored aerobically or anaerobically and as a control one culture was struck on an anaerobic TME lactate plate with no arsenate. Frozen stocks of all cultures were made then colonies were re-struck four more times on anaerobic plates containing TME lactate with 5mM arsenic to ensure a pure culture was obtained. After the fifth passage just one pink isolate and one white isolate remained from the Wild Willy culture.

Growth experiments. Pure cultures of the remaining Pink or White isolates from Wild Willy's hot spring were grown in anaerobic Hungate tubes containing TME amended with 20 mM lactate and 5 or 10mM arsenate or 10mM nitrate or 10mM fumarate.

16S sequencing. Pink isolate was grown aerobically up to an OD_{600} of 0.830. Then genomic DNA was purified from 1 ml of culture using the Qiagen DNeasy blood and tissue kit protocol for gram negative bacteria. Finally 16S sequence was amplified using 2mM of the 8F primer and 2mM of the1492 primer.

Arsenic reductase sequencing. The respiratory arsenic reductase of the Pink isolate was sequenced after amplifying from genomic DNA with the degenerate arrA primers first described in Malasarn et al 2004. This PCR produced a single band and the crude PCR product was sent for sequencing by sequetech company using the capillary DNA sequencer each strand was independently sequenced to verify the sequence accuracy and each strand produced the same results.

16s alignment. A NCBI blast search determined that the Pink isolate was closely related to *Shewanella*. The Isolate was also classified using RDP's RDP Naive Bayesian rRNA Classifier Version 2.10.

The Pink isolate was aligned to other *Shewanella* identified in the blast search including all *Shewanella* known to have the arrA reductase using the phylogeny.fr. platform below is a list of the experimental workflow generated directly from the program:

" 1) Sequences were aligned with MUSCLE (v3.8.31) configured for highest accuracy (MUSCLE with default settings).
2) After alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) using the following parameters:
-minimum length of a block after gap cleaning: 10
-no gap positions were allowed in the final alignment
-all segments with contiguous nonconserved positions bigger than 8 were rejected
-minimum number of sequences for a flank position: 85%
3) The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT). The HKY85 substitution model was selected assuming an estimated proportion of invariant sites (of 0.260) and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma=0.373). Reliability for internal branch was assessed using the aLRT test (SH-Like).
4) Graphical representation and edition of the phylogenetic tree were performed with TreeDyn (v198.3)."

As an alternative to bootstrapping the MABL program quantified homlology using a confidence index.

ArrA alignment. A partial sequence from the respiratory arsenic reductase was cloned using the primers described in Malasarn et al., 2004. This sequence was then identified using the crude PCR product as a template Next the sequence was aligned after a blast search determined its closest homologs. Putative amino acid sequence of

the Pink Isolate's arrA was determined by loading the partial arrA DNA sequence into NCBI's BlastX. From blast X the putative amino acid sequence was determined Finally the sequence was aligned and a tree was generated in a similar fashion to the first phylogenetic tree except it used the PhyML program (v3.1/3.0 aLRT). Method as described by phylogeny.fr:

"The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT). The WAG substitution model was selected assuming an estimated proportion of invariant sites (of 0.165) and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma=126.759). Reliability for internal branch was assessed using the aLRT test (SH-Like)".

Results

The goal of this study was to isolate an environmental microbe from the arsenic rich hot springs of Northern California that would grow easily and consistently in a laboratory environment both anaerobically respiring arsenic and aerobically. Once soli and water from either Navy Beach or Wild Willy was obtained enrichment cultures from these samples were maintained anaerobically in minimal salts media to ensure that isolates from these cultures could respire arsenic. Bacterial enrichments were cultivated in ten separate Isolation experiments to increase chances of isolating a facultative arsenic respiring microbe. Ten percent of each of the enrichments was passaged to new anaerobic media until particles of debris were no longer seen in the media. After the fourth passage a sterile loop was used to streak bacteria from each enrichment culture onto anaerobic plates with arsenic to obtain individual colonies. Two days after streaking aerobic and anaerobic cultures all plates had colonies of various morphologies. The no terminal electron acceptor control plate also had

colonies but these were needle-point colonies of a similar color to the arsenic plate streaked from the same culture. Colonies growing on the control plate with no terminal electron acceptor may be due to some carryover of arsenic or another terminal electron acceptor during streaking. To be certain the isolates could actually grow on arsenic they were re-struck several times.

All of the Navy beach plates produced cream, yellow or orange colonies. Several of the larger colonies are yellow in the center yet more cream towards the outer rim of the colony. Individual colonies have a circular raised morphology but colonies that are touching have a similar color but have a more irregular undulate morphology. After three subsequent streaking experiments none of the navy beach isolates remained viable despite their robust growth in liquid culture after several passages.

Most colonies from the original Wild Willies hot spring isolations were smaller that Navy Beach isolates but had similar orange/ cream coloring. Strikingly, one colony grown from a streak of the Wild Willy cultures was grapefruit pink. During successive streaks many of the isolates no longer grew until just two isolates remained both from the Wild Willy hot spring, one grapefruit pink isolate and one of the orange isolates that began to look more pale colored over time. Nonetheless, pure cultures were stored at -80° C for both the pink and white isolates. To ensure that the pure cultures from Wild Willy were facultative anaerobes that can respire arsenic the frozen stocks were re-struck again then were grown in liquid cultures either aerobically or anaerobically in Hungate tubes with arsenate nitrate or fumarate as the terminal electron acceptor over a 24 hour period (Figure 1). The cultures were grown

aerobically with 10mM arsenate to determine if the presence of this concentration of arsenic had any effects on growth. The aerobic culture with arsenic 10mM arsenic and actually grew even better with a maximum optical density of 0.29 (not shown). Finally the cultures were grown again anaerobically in the presence of five or 10 mM arsenic to ensure that the cultures were able to grow consistency and to determine the ideal concentration of arsenic for cellular respiration. At 5mM the cell did not grow at all and by this experiment the white isolate is no longer growing so the cell grows better with higher concentrations of As(V) (10mM) although the upper limit was not tested.

Only the grapefruit pink isolate grew consistently well anaerobically on arsenic and aerobically in the presence of arsenate over time; this isolate was identified using 16S sequencing. Sequencing 16S revealed this isolate is a previously unidentified *Shewanella* with 99% homology to *Shewanella* putrefaciens strain M-8m-1 using both BLAST and RDP. To determine how closely related this novel microbe is to other Shewanella species its 16S sequence was aligned to other Shewanella as well. (Figure 2).

The protein sequence for ArrA were also aligned once the amino acid sequence for JDA-W-1 was determined using blastx. JDA-W-1 was then aligned to other *Shewanella* with ArrA homologs. To test if the nonfunctional ArrAs were due to phylogenetic differences *Shewanella* that contained an *arrA* gene but lack the ability to respire arsenic were also aligned. An arsenic respiring strict anaerobe was also added to the phylogenetic tree to determine if this reductase was more or less similar

to the arsenic reductases found in the facultative *Shewanella* arsenic reductase (Figure 3).

Discussion

The genus *Shewanella* is well known for its diverse metabolic capabilities and is studied for its ability to respire diverse substrates including metals, pollutants, electrodes, and even radionuclides (Tiedje, 2002). Until now *Shewnaella* have been isolated from a diverse array of fresh water and marine environments including hydrothermal vents, but not arsenic rich hot springs (Rodionov et al.,2011). There are seven published sequences of *Shewanella* sp. that contain a homologous sequence to the *arrAB* arsenic reductase. ANA-3, WP3, W3-18-1 (Zargar et al., 2011), CN-32, 200, Har-4 and *waksmanii* two of which are not actually able to respire arsenic (WP3, W3-18-1) and one unknown (*waksmanii)* however in this study the first hot spring *Shewanella* sp. was Isolated, shown to contain an arrA and shown to respire arsenic. The advantage of the isolation method is that it selects for microbes that are facultative anaerobes that grow readily in lab conditions (neutral pH, 30 degrees). Isolating a *Shewanella* from the hot spring environment does have the advantage of being easily comparable to other arsenic respiring *Shewanella*. Alternatively, to avoid isolating another Shewanella the cultures could have been Isolated using acetate or succinate or any carbon source that *Shewanella* are not know to grow in.

It was surprising that a *Shewanella* was isolated from this desert hot spring environment when *Shewanella* are know for their psychrotolerance (Hau and Gralnick, 2007,Wang et al., 2008). The partial sequence from the respiratory arsenic

reductase along with growth experiments with arsenic as the sole terminal electron acceptor provide evidence that this pink isolate is a novel hot spring isolated arsenic reducing *Shewanella*. It would be interesting to determine if *Shewanella* from this naturally arsenic rich hot spring also has the same regulatory dependence on cAMP and CRP as other *Shewanella* during anaerobiosis (Murphy, Durbin, and Saltikov 2009, Charania et al., 2009) which will represent reduction mechanism from the first naturally occurring hot spring isolated *Shewanella*.

Future directions

The fact that this microbe grows well in lab and clearly can respire arsenic shows promise for future projects. The next interesting follow up experiment for this project would be to test if JDA-W-1 is actually genetically amenable. This can be done by attempting to insert plasmids that would allow JDA-W-1 to gain a function such as antibiotic resistance (once we determine which antibiotics JDA-W-1 is sensitive to). One could also test to see if JDA-W-1 has a similar cAMP dependent mechanism of anaerobic respiration to other *Shewanella* possibly by growing JDA-W-1 on non-PTS sugars or by knocking out the adenylate cyclases in JDA-W-1 once a system is developed. Another interesting direction would be to do comparative genomic studies with JDA-W-1 and other arsenic reducing *Shewanella* to identify which similaraties they have to each other and how these arsenic reducers are different from *Shewanella* who have arrA homologs yet can't respire arsenic. This third study could potentially identify any additional genetic prerequisites to arsenic respiration.

Chapter 4 figures

Figure 4.1. Growth of JDA-W-1 on different terminal electron acceptors. Optical density$_{600NM}$ reveals similar growth using various anaerobic terminal electron acceptors and higher cell densities in aerobic conditions. Each bar graph represents the average optical density of triplicate cultures for each terminal electron acceptor 24 hours after inoculating all cultures from the same starter culture.

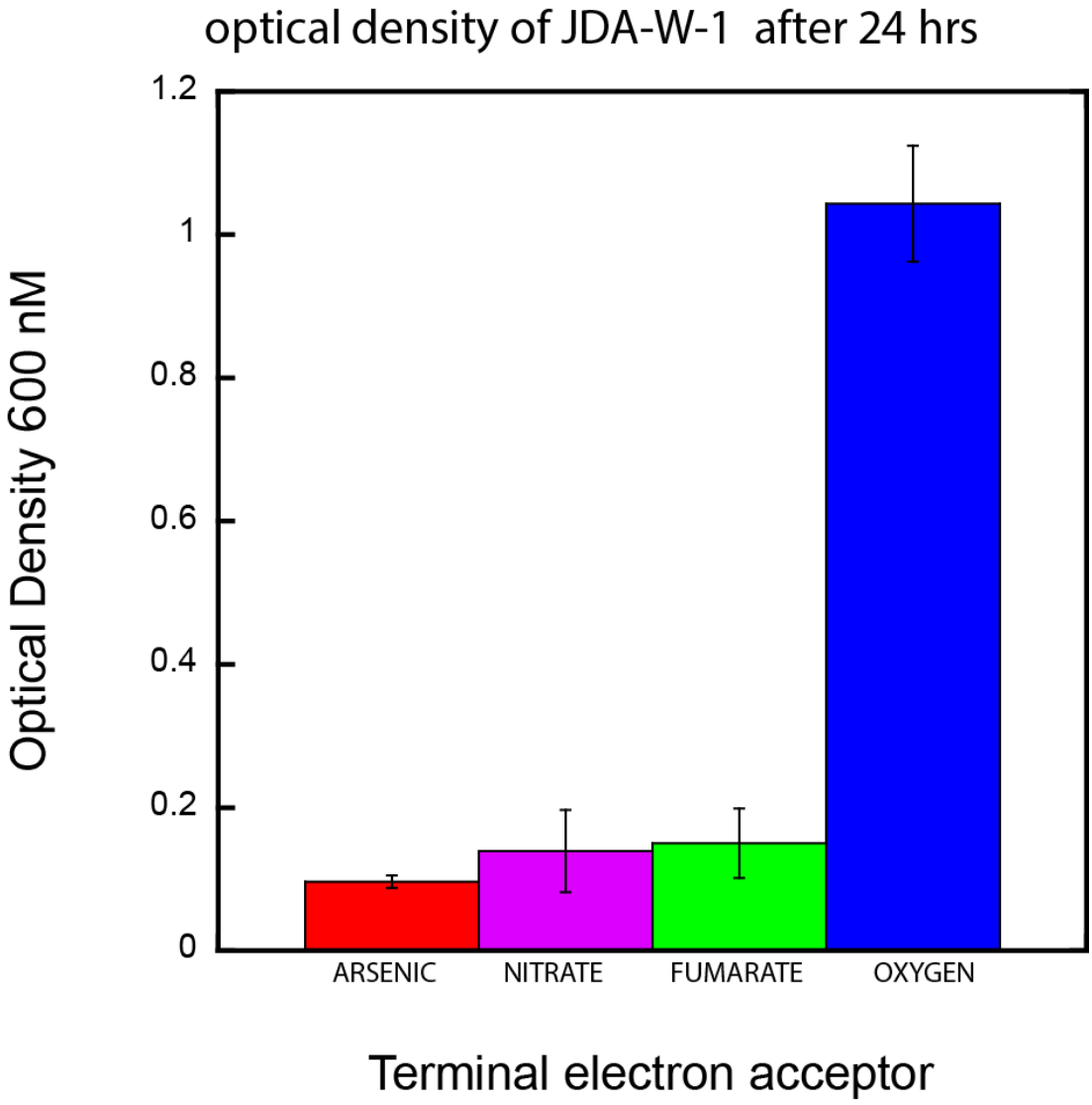


Figure 4.2. Ribosomal RNA 16S alignment of JDA-W-1 to Shewanella. 16S Cladogram alignment of pink isolate (*Shewanella* sp. JDA-W-1) 16s ribosomal sequence along with other *Shewanella* sp. actual branch lengths are not displayed. For all organisms species and strain name is displayed. JDA-W-1 is highlighted in red. Numbers displayed in red indicate percentage branched support value.

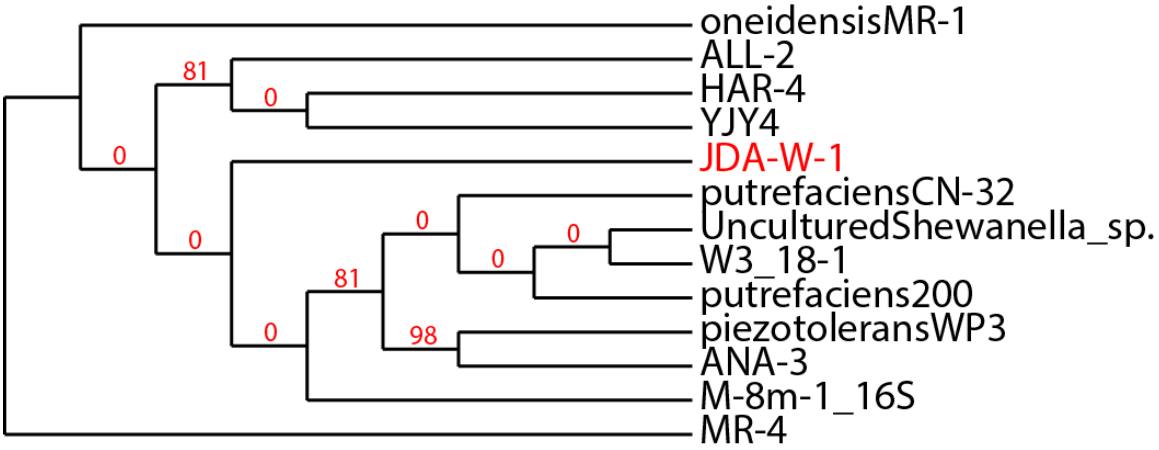
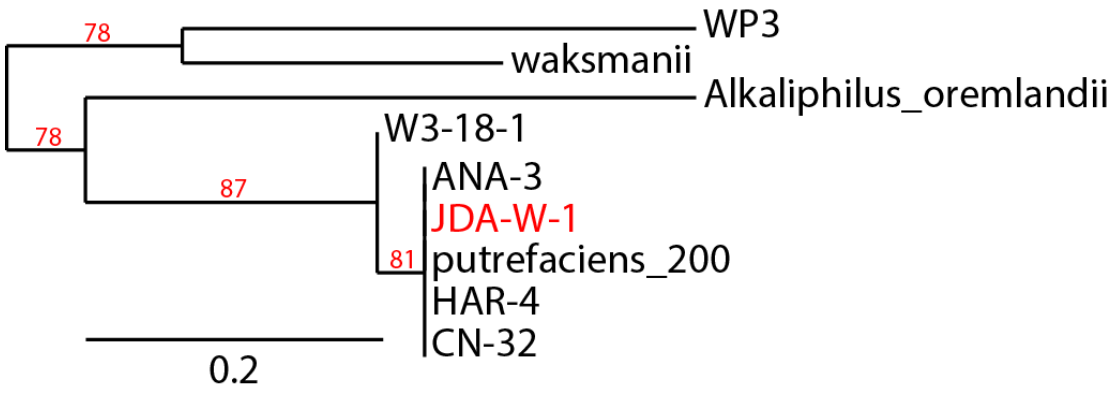


Figure 4.3: Protein alignment of arrA from JDA-W-1 to other arsenate/ non arsenate respiring microbes. Phylogram Alignment of protein sequence for the arrAB reductase subunit-A of the pink isolate (*Shewanella* sp. JDA-W-1) along with other *Shewanella* spp.. The strict anaerobe capable of respiring arsenate, *Alkaliphilus oremlandii,* is used as an out group For all *Shewanella* strains only species and strain name is displayed. JDA-W-1 is highlighted in red. Numbers displayed in red indicate percentage branched support value. Regions that have no number displayed have a branch support value of 100% less than 10%. Scale bar represents branch length



Malasarn Primers

ArrAfwd (5’-AAGGTG- TATGGAATAAAGCGTTTgtbgghgaytt-3’)

ArrArev (5’-CCTGTGATTTCAGGTGCC- caytyvggngt-3’)

JDA-W-1 16s sequence

NNGNGGCNGCNACACATGCAAGTCGAGCGGCAGCACAAGGGAGTTTACT
CCTGAGGTGGCGAGCGGCGGACGGGTGAGTAATGCCTAGGGATCTGCCC
AGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCCCT
ACGGGGGAAAGGAGGGGACCTTCGGGCCTTCCGCGATTGGATGAACCTA
GGTGGGATTAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCC
TAGCTGTTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCA
GACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACC
CTGATGCAGCCATGCCGCGTGTGTGAAGAAGGCCTTCGGGTTGTAAAGCA
CTTTCAGTAGGGAGGAAAGGGTAAGTCTTAATACGGCTTATCTGTGACGT
TACCTACAGAAGAAGGACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAT
ACGGAGGGTCCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCA
GGCGGTTTGTTAAGCGAGATGTGAAAGCCCTGGGCTCAACCTAGGAATAG
CATTTCGAACTGGCGAACTAGAGTCTTGTAGAGGGGGGTAGAATTCCAGG
TGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGC
GGCCCCCTGGACAAAGACTGACGCTCATGCACGAAAGCGTGGGGAGCAA
ACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTCTACTCGGA
GTTTGGTGTCTTGAACACTGGGCTCTCAAGCTAACGCATTAAGTAGAC
CGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGG
GCCCGCACAAGCGGNGGNNCATGGTGGTTTAATTCGATGCAACGCGAAN
NACCTTANCTACTNNNTGANNNTCCCACANANACTGCNNANNATNNGGT
NNNNTGCCCTTC

JDA-W-1 *arrA* partial DNA sequence
NNNNNNGTAAAANCTGTTTAAGCGGGTAAAACTGTCAGTGTCGAGAGCTT
TAAAGAAACCCATACCTACGGTTTAGTAGAATGGTGGAACCAAGCCCTTA
AAGACTATACCCCCGAGTGGGCACCTGAAATCACAGGA

JDA-W-1 ArrA inferred protein sequence
AGKTVSVESFKETHTYGLVEWWNQALKDYTPEWAPEITG

Strain name

The pink isolate from wild willies hot spring shall be called *Shewanella* sp. strain JDA-W-1 in honor of my daughter Jada Watson who's first birthday was spent on a sampling trip to an arid arsenic rich environment much like the one this isolate originates from.

Chapter 4 References

Anisimova M., Gascuel O. *Approximate likelihood ratio test for branchs: A fast, accurate and powerful alternative.* Syst Biol. 2006, Aug;55(4):539-52.

Arsenic in the environment : Effects on human health and possible prevention. (2009). Arsenic in the environment : Effects on human health and possible prevention, 1–7.

Barringer, J. L., & Reilly, P. A. (2013). Arsenic in groundwater: a summary of sources and the biogeochemical and hydrogeologic factors affecting arsenic occurrence and mobility. http://doi.org/10.5772/55354

Bruno W.J. , Socci N.D., and HalpernA.L. (2000). Weighted Neighbor Joining: A Likelihood-Based Approach to Distance-Based Phylogeny Reconstruction, *Mol. Biol. Evol.* 17 (1): 189-197.

Charania, M. A., Brockman, K. L., Zhang, Y., Banerjee, A., Pinchuk, G. E., Fredrickson, J. K., et al. (2009). Involvement of a membrane-bound class III adenylate cyclase in regulation of anaerobic respiration in Shewanella oneidensis MR-1. *Journal of Bacteriology*, *191*(13), 4298–4306. http://doi.org/10.1128/JB.01829-08

Castresana J. *Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.* Mol Biol Evol. 2000, Apr;17(4):540-52.

Chevenet F., Brun C., Banuls AL., Jacq B., Chisten R. *TreeDyn: towards dynamic graphics and annotations for analyses of trees.* BMC Bioinformatics. 2006, Oct 10;7:439.

Dereeper A., Audic S., Claverie J.M., Blanc G. *BLAST-EXPLORER helps you building datasets for phylogenetic analysis.* BMC Evol Biol. 2010 Jan 12;10:8.

Dereeper A.*, Guignon V.*, Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.F., Guindon S., Lefort V., Lescot M., Claverie J.M., Gascuel O. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist.* Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W465-9. Epub 2008 Apr 19. *: joint first authors

Edgar RC. *MUSCLE: multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Res. 2004, Mar 19;32(5):1792-7

Guindon S., Gascuel O. *A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.* Syst Biol. 2003, Oct;52(5):696-704

Hau, H. H., & Gralnick, J. A. (2007). Ecology and biotechnology of the genus Shewanella. *Annual Review of Microbiology*, *61*, 237–258.

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

Kulp, T. R., Hoeft, S. E., Miller, L. G., Saltikov, C., Murphy, J. N., Han, S., et al. (2006). Dissimilatory arsenate and sulfate reduction in sediments of two hypersaline, arsenic-rich soda lakes: Mono and Searles Lakes, California. *Applied and Environmental Microbiology*, *72*(10), 6514–6526. http://doi.org/10.1128/AEM.01066-06

Kulp, T. R., Hoeft, S. E., & Oremland, R. S. (2004). Redox transformations of arsenic oxyanions in periphyton communities. *Applied and Environmental Microbiology*, *70*(11), 6428–6434.

Malasarn, D., Keeffe, J. R., & Newman, D. K. (2008). Characterization of the arsenate respiratory reductase from Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *190*(1), 135–142.

Malasarn, D., Saltikov, C. W., Campbell, K. M., Santini, J. M., Hering, J. G., & Newman, D. K. (2004). arrA is a reliable marker for As(V) respiration. *Science (New York, N.Y.)*, *306*(5695), 455–455.

Murphy, J. N., Durbin, K. J., & Saltikov, C. W. (2009). Functional roles of arcA, etrA, cyclic AMP (cAMP)-cAMP receptor protein, and cya in the arsenate respiration pathway in Shewanella sp. strain ANA-3. *Journal of Bacteriology*, *191*(3), 1035–1043.

Oremland, R. S., & Stolz, J. F. (2003). The ecology of arsenic. *Science (New York, N.Y.)*, *300*(5621), 939–944. http://doi.org/10.1126/science.1081903

Oremland, R. S., Stolz, J. F., & Hollibaugh, J. T. (2004). The microbial arsenic cycle in Mono Lake, California. *FEMS Microbiology Ecology*, *48*(1), 15–27. http://doi.org/10.1016/j.femsec.2003.12.016

Rodionov, D. A., Novichkov, P. S., Stavrovskaya, E. D., Rodionova, I. A., Li, X., Kazanov, M. D., et al. (2011). Comparative genomic reconstruction of transcriptional networks controlling central metabolism in the Shewanella genus. *BMC Genomics*, *12 Suppl 1*(Suppl 1), S3–S3. http://doi.org/10.1186/1471-2164-

12-S1-S3

Saltikov, C. W., & Newman, D. K. (2003). Genetic identification of a respiratory arsenate reductase. *Pnas*, *100*(19), 10983–10988.

Saltikov, C. W., Cifuentes, A., Venkateswaran, K., & Newman, D. K. (2003). The ars detoxification system is advantageous but not required for As(V) respiration by the genetically tractable Shewanella species strain ANA-3. *Applied and Environmental Microbiology*, *69*(5), 2800–2809. http://doi.org/10.1128/AEM.69.5.2800-2809.2003

Smedley, P. L., & Kinniburgh, D. G. (2002). A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry*, *17*(5), 517–568. http://doi.org/10.1016/S0883-2927(02)00018-5
Tiedje, J. M. (2002). Shewanella--the environmentally versatile genome. *Nature Biotechnology*, *20*(11), 1093–1094.

Wiley E. O., Brooks D. R., Siegel-Causey D., Funk V. A. (1991). *The Compleat Cladist: A Primer of Phylogenetic Procedures*. Freely available at http://taxonomy.zoology.gla.ac.uk/teaching/CompleatCladist.pdf

Wang, F., Wang, J., Jian, H., Zhang, B., Li, S., Wang, F., et al. (2008a). Environmental adaptation: genomic analysis of the piezotolerant and psychrotolerant deep-sea iron reducing bacterium Shewanella piezotolerans WP3. *PLoS ONE*, *3*(4), e1937–e1937.

Zargar, K., Conrad, A., Bernick, D. L., Lowe, T. M., Stolc, V., Hoeft, S., et al. (2012). ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. *Environmental Microbiology*, *14*(7), 1635–1645.

Appendix 1 Curriculum Vitae

Ruth Pamela Tilus Watson

rtilus@ucsc.edu

University of California, Santa Cruz
METX mailstop
1156 High Street Santa Cruz, CA 95064
Office: 831-459-1885 Cell: 831-428-4135

Education

Period	Institution	Examinations
2010-present	University of California Santa Cruz	Ph.D. Molecular Cell and Developmental Biology Expected June 2015
2010	Savannah State University	B.S. Biology *summa cum laude*

Appointments

Period	Job Title	Duties

2010-present	Graduate Studer Researcher	Maintain a research project and advise undergraduate mentees while enrolled with 15 or more course credits.
Spring 2014	Teaching Assist	Genetics Lecture: Prepare lecture, reviews and quizzes to reinforce the course content, grade assignments and individual tutoring during office hours.
Summer 2013	Teaching Fellow	Inquiry institute: Teach Inquiry based bioinformatics short course, prepare prompts that highlight one biological phenomenon, facilitate hypothesis driven research.
Fall 2012	Teaching Assist	Microbiology Laboratory: Reinforce techniques and & concepts in microbiology, grade assignments.
Fall 2012	Guest Lecturer	Science Technology Engineering and Math journal club: lead discussion on a multidisciplinary journal article to students with various academic backgrounds.

Research Experience

Period	**Mentor**	**Project Outcome**

2011- present	Chad Saltikov UC Santa Cruz	"Investigating arsenic based anaerobic metabolism in *Shewanella Sp. ANA-3* using molecular biology, genetics, and systems biology" Skills: RNAseq Library preparation, transcriptomic data processing and normalization, transposon mutagenesis, site-directed mutagenesis, multiplex q-RT-PCR, competitive ELISA, enrichment and isolation of chemoorganotrophs & photoautotrophs form arsenic rich field sites.
Spring 2011	David Feldheim UC Santa Cruz	"Producing a stable cell line for studying cadherin and other cell surface molecules" Skills: maintaining cell culture, human cell line transfection, immunocytochemistry, western blotting.
Winter 2011	Chad Saltikov UC Santa Cruz	"The effects of cyclic AMP on gene expression of the arsenic reduction genes" Skills: Sterile and anaerobic bacterial culture maintenance, physiological testing of bacterial cells, q-RT-PCR.
Fall 2010	Susan Strome UC Santa Cruz	"Study of the effects of temperature shift on fertility in the *C. elegans* model" Skills: Cultivating live animals, immunohistochemistry.
Summer 2010	Yu Zhang NYU	"Detecting Glass Infiltration in Graded Zirconia Materials for Dental Restorations." Skills: scanning electron microscopy, transmission electron microscopy.

Summer 2009	John Baatz Medical University Of South Carolina	"Functional Characterization of Dolphin (*Tursiops truncates*) Surfactant Protein B." Skills: PCR, Recombinant DNA manipulation and cloning.

Selected Presentations and Publications

Watson, R.P., Lazcano, I., Hernandez, J.M., Bernick, D., Chen,P., Saltikov, C.W. 2015. "Systems level study of respiration in the arsenate reducing bacterium *Shewanella* sp. Strain ANA-3." Publication in preparation

Watson, R.P., Lazcano, I., Saltikov, C.W. 2015. "New Insights into the role of cAMP modulation in the regulation of cellular respiration." Publication in preparation

- Watson, R.P., Saltikov, C.W. 2014. "Transcriptome analysis by RNAseq of *Shewanella* sp. ANA-3 grown on arsenate and other terminal electron acceptors." Society for Advancement of Hispanics/Chicanos and Native Americans in Science (SACNAS) **National Conference**, Los Angeles, CA, October 16-18. Oral Presentation
- Watson, R.P., Lazcano, I., Saltikov, C.W. 2014. "RNAseq roadmap to respiration in *Shewanella*." University of California Santa Cruz Program in Biomedical Science and Engineering (PBSci) **Annual Research Conference**, Aptos, CA, September 17. Poster Presentation
- Watson, R.P., Lazcano, I., Saltikov, C.W. 2014. "RNAseq roadmap to respiration in *Shewanella*." American Society for Microbiology (ASM) **114th General Meeting**, Boston Ma, May 17-20. Poster Presentation
- Watson, R.P., Lazcano, I., Saltikov, C.W. 2014. "RNAseq: Roadmap to the

arsenic respiration pathway in *Shewanella* sp. ANA-3." University of California Santa Cruz **10th Annual Graduate Research Symposium**, Santa Cruz, May 9. Poster Presentation

- Watson, R.P., Saltikov, C.W. 2013. "Molecular biology of arsenic metabolism in metal reducing bacteria." University of California Santa Cruz Molecular Cell and Developmental Biology (MCD) **Departmental Seminar**, Santa Cruz, CA, April 30. Oral Presentation
- Watson, R.P., Saltikov, C.W. 2012. "Regulation of respiratory arsenic reduction in *Shewanella* sp. ANA-3." Annual Biomedical Research Conference for Minority Students (ABRCMS) **Annual Research Conference**, San Jose, CA, November 7-10. Poster Presentation
- Watson, R.P., Saltikov, C.W. 2012. "Regulation of respiratory arsenic reduction in *Shewanella* sp. ANA-3." American Society for Microbiology (ASM) **112th General Meeting**, San Francisco, CA, June 16-19. Poster Presentation
- Watson, R.P., Saltikov, C.W. 2012. "Regulation of respiratory arsenic reduction in *Shewanella* sp. ANA-3." University of California Santa Cruz **8th Annual Graduate Research Symposium**, Santa Cruz, CA, May 11. Poster Presentation

Professional Development

Credential	**Outcome**
Cert. 2013	Institute for Science and Engineering Educators: Intensive training program in inquiry based teaching methods including curriculum development. Certificate granted after successfully designing and implementing a one-week short course.
Cert. 2012	Research Mentoring Institute Mentor Training Program: One-month course with rigorous training in personal development, conflict resolution, and professional skills. Certificate granted after completing all assignments and group discussions.

Mentoring Experience

Period	Outcome
2010-present	Mentoring Indra Lazcano, an upper level UCSC undergraduate, through an ongoing research project as well as her senior thesis. Indra has presented her work at both local and national conferences and has received travel awards based on her abstracts. She is currently working on completing her senior thesis.
2014	Academic bridge program for community college students pursuing a career on biomedical research (ACCESS). Mentored April Villareal, a community college student, through a 10 week summer research project which commenced in a poster symposium and oral presentation.
2012-2013	Initiative for Maximizing Student Diversity Program IMSD. Mentored Marcelo Quiroz, an upper level UCSC undergraduate. Marcelo presented his work at two national conferences in addition to several oral presentations for lab meetings and peer groups. Marcelo went on to become staff researcher at UCSF.
2011	Mentored Janie Alleio, a community college student through a 10 week summer research project which commenced in a poster symposium and oral presentation.

Synergistic Activities

Period	Activity
2014	Recruitment volunteer for the UCSC's physical and biological sciences program at the Society for Advancement of Hispanics/Chicanos and Native Americans in Science (SACNAS).
2014	Guest lecturer Hartnell Community College STEM speaker series. "Navigating big universities, big data and big dreams

#NONStopToTheTop."

2013 Poster judge for the California Alliance for Minority Participation in the Sciences (CAMP) statewide undergraduate research conference.

2012 Recruitment volunteer for UCSC's physical and biological sciences program at the ASM managed Annual Biomedical Research Conference for Minority Students (ABRCMS).

2012 K-12 Santa Cruz County Science Fair Judge provided feedback on project design and data analysis.

Fellowships and Awards

Period	**Award**
2010-201	Eugene Cota Robles Fellow: State funded merit-based fellowship for students who have overcome significant social or educational obstacles to achieve a college education.
2014	Institute for Science and Engineering Educators: Scholarship to attend teaching institute along with room and board based on essay articulating my desire to become a better educator and need for funding to attend the institute.
2010-201	Bridge To Doctorate Fellow: National Science Foundation sponsored fellowship. One of 12 fellows recruited for the inaugural cohort at UCSC. Bridges To Doctorate scholars were required to participate in leadership training and research training seminars and submit quarterly progress reports to the program manager in addition to regular course work.
2009-201	Louis Stokes Alliance for Minority Participation in the Sciences: National Science Foundation sponsored fellowship. One of 30 awardees tuition reduction stipend to participate in research and funding to present at both statewide and national conferences.

Graduate and Thesis Committee

Graduate advisor: Dr. Chad W. Saltikov,

Committee members: Dr. Susan Strome, Dr. Needhi Bhallah

Research/ Teaching Interests

microbiology, genetics, systems biology, transcriptional regulation, pathogenesis,

Other Languages/ Skills

Python Programing

Conversational Créole

Appendix 2 Teaching Statement

I knew I wanted to be a scientist before I started kindergarten. As a preschooler, when my older siblings came home from school I always had questions for them about how things worked. As I got older, I soaked up all the information I could in sciences classes throughout grade school. However, by high school, I got bored with the idea of becoming a scientist when I realized all the experiments we did in class were prefabricated with predictable outcomes. This all changed when I got to seniors physics. My physics teacher used many gadgets and gizmos to prompt her students into asking questions and taught us to integrate what we understood about math to help us answer those questions. Suddenly science was filled with discovery again. However, as an undergraduate I questioned whether there was a place for me, a Black woman, in science. There were volumes of textbooks already written, and most of the people who wrote those books did not look very much like me. Fortunately, I realized this was not true after becoming a part in the Louis Stokes Alliance for Minority Participation in science (LSAMP). As a LSAMP scholar, I saw many minorities engaging in research, which helped me to finally see that I could become the scientist I set out to be more than 15 years prior.

Because of my own experience in LSAMP, I see teaching as an opportunity to challenge and empower students to see themselves as scientists. In the classroom, I want to use both my academic and bench experience to implement a rigorous program that would promote inquiry-based learning, which sparks a desire to understand and learn more. As a minority woman, I am also heavily invested in promoting more underrepresented minorities in science by serving as a role model of

minorities doing high caliber research. I know that the FIRST program will afford me the experience of teaching this unique population of undergraduates and help cultivate the skills I need to meet their unique needs.

At UC Santa Cruz, I have gained hands-on teaching experience through work as a Teaching Assistant (TA), and as a participant in the Professional Development Program of the UCSC Institute for Science & Engineer Educators (ISEE PDP). I was trained in a variety of pedagogical strategies for effective teaching in the lab and in the classroom. As a TA I shadowed a professor during lecture and reinforced the course content in a weekly a problem solving seminar. I designed weekly quizzes that assessed each learner's retention of the information of both lab and lecture. Each quiz was designed to challenge the students to apply the information from class to solve a problem. From my TA experience I learned how to properly gauge learning through formative assessment. My teaching skills were further refined through the intensive training I received in the PDP. The PDP is a five-month institute that gives an introduction to inquiry based pedagogy through seminar series followed by hands-on teaching a short course for undergraduate students. Through the PDP I, along with two collaborators, designed a short course so that students would learn through generating and testing hypotheses. To achieve this, first I used prompts that students followed and encouraged them to generate their own hypotheses. Next, I helped each student test his or her hypothesis using bioinformatic tools. At the end of the course the students presented their findings to their peers. It was exciting to see all the students engage in the scientific method while they learn the course material. Overall,

I learned that achieving success in the classroom involved engaging the students and allowing students to teach themselves.

Although I have a strong passion for teaching, I'm also enamored by microbial genetics. My most extensive research experience stems from my thesis work studying the genetic cues that allow the bacterium *Shewanella* sp. ANA-3 to trigger arsenic respiration in oxygen depleted environments. In this field there was an open question about how Shewanella, an important model for toxic metal remediation, turns on arsenic reduction genes in the absence of oxygen. In the Saltikov lab, I executed and analyzed high-throughput transcriptomic sequencing experiments to identify genes that are important for anaerobic respiration. Results from these experiments identified several gene classes that were not previously known to be necessary for anaerobic respiration in *Shewanella* (manuscript in submission). This work added new insight into a recent dispute in the field and validated many of the previously published reverse genetic experiments. I also studied the role of the second messenger cyclic AMP, as a signal for transcriptional regulation of anaerobic respiration in *Shewanella*. By using a reverse-genetic approach, creating mutations and observing their biochemical consequence, I was able to target several genes in the cyclic AMP biosynthesis and degradation pathway and determine the exact amount of cAMP necessary for respiring arsenic and other terminal electron acceptors.

Before joining my thesis lab, I received a breadth of training in various subjects including Neuroscience, Developmental Biology, and Materials Science through

several 10 week rotation projects during my first year as a graduate student and REU programs in my junior and senior year of undergraduate school. These additional research opportunities developed my skills in cell culture, immunocytochemistry and microscopy, and give me a broad foundation for teaching. Taken together, my research experiences will allow me to take a creative and interdisciplinary approach to a future Post-doctorate experience in the lab while addressing how an organism can regulate gene expression to survive in hostile environments.

In order to incorporate my passions in both research and teaching, I plan to pursue a career as a professor. After fellowship training, I plan to run a lab at a minority serving institution and teach. In lab I would be teaching students the art of discovery through training and producing research that contributes to our field. In the classroom, I plan to design curriculum, which incorporates current technology relying heavily on primary literature to engage both future scientists and health professionals to understand biology as a dynamic science where their input is needed. The FIRST program will give me well balanced training in teaching and research. This training will develop the skills I need to engage a unique population of undergraduates while preparing me for a career at a research-intensive liberal arts institution thus allowing me to reach my goals and encourage the next budding researchers.

Appendix 3 Sample inquiry based student activity and lesson plan:

Summer 2013 I co-taught a one week summer elective as a part of a summer enrichment program for upper division STEM students with no previous research experience that are underrepresented in the STEM field. The three co-instructors, Myself, Dr. Mia Grifford and Dr Michael Halbisen, each came up a with a unique inquiry based lesson that pertained to our research while highlighting bioinformatics tools commonly used in our research that the students were not previously exposed to. First the student's background was assessed through a questioner that asked about their experience with bioinformatics tools and

Below are the worksheet I used to prompt the students to practice using the bioinformatics tools and the "starter exercise" I used to demonstrate how the tool works. Each student was then asked to come up with a hypothesis that they could test over the course of three days using any of the bioinformatics tools that they learned from the starters.

We sorted out and refined the questions they came up with as a group to questions that were actually feasible for a three day project. Next the students were allowed to pick from the refined pool of questions and work in groups of 2-3 to address their hypothesis while we guided their experiments to ensure they stayed on the right track and had a reasonable quality and amount of data to present at the final poster symposium. The poster symposium was attended by graduate students and faculty and they also had the option of presenting their research at conferences. The course was designed to provide the students with tools that they can use in future

research while giving them practice with several tools they would need to be successful in research such as hypothesis formation, collaboration, forming conclusions, and presenting findings. Finally the course ended with us giving them new questions to address in a larger group (4-5). And we had a lecture at the end to reveal the "right answer" by pointing out how each hypothesis lead to the biological phenomenon that similar genetic sequences indicate similar biological function and shared biological function is often encoded in the homologous regions of a sequence.

Worksheet/ demo

Characterizing novel enzymes

There are two very special classes of organisms that convert the metal arsenic into a more or less toxic form. Several years of genetic and biochemical manipulation by Chad Saltikov here at UCSC and other environmental microbiologist lead to the discovery of the essential enzymes for these processes. ArrA is the required enzyme for respiratory arsenic reduction which converts arsenic into a more toxic form and AoxB is the required enzyme for converting arsenic into a less toxic form when coupled to energy production during cellular respiration as well.

Using the 2007 review by Chad Saltikov (figure 2) I obtain the accession numbers for several of these enzymes in some common arsenic respiring organisms. Below I have included the accession numbers of a couple DMSO reductase class enzymes that are not classified as a reductase or oxidase.

Enzyme group: arsenate reductase

Organism	Accession Number
Shewanella piezotolerans WP3	YP_002311519
Shewanella sp. ANA-3	AAQ01672
Chrysiogenes	AAU11839
Geobacter lovleyi	ZP_ 01593421
Geobacter uraniireducans	ZP_01140714
Bacillus selenitireducens str. MLS10	AAQ19491
Bacillus arseniciselenatis str. E1H	AAU11841
Sulfurospirillum barnesii str. SES-3	AAU11840
Wolinella succinogenes	NP_906980
Desulfosporosinus	ABB02056
Desulfitobacterium	ZP_01372404

Enzyme group: arsenite oxidase

Organism	Accession Number
Agrobacterium tumefaciens	ABB51928
Ochrobactrum tritici	ACK38267
Xanthobacter autotrophicus	ZP_01198801
Nitrobacter hamburgensis	YP_571843
Alcaligenes	AAQ19838
Ralstonia sp. 22	ACX69823
Herminiimonas arsenicoxydans	AAN05581
Rhodoferax ferrireducens	YP_524325
Pseudomonas sp. TS44	ACB05943

Enzyme group: unclassified

Organism	Accession Number
Alkalilimnicola ehrlichii MLHE-1	YP_741061
Alcaligenes faecalis	AAQ19838

Enzyme Group: other types of reductases

Organism	Accession Number
Escherichia coli	WP_001431343
Thauera selenatis	CAB53372

Follow the link on our page to find the National Center for Biotechnology information.
Go to the main page and store the hyperlink information in favorites.

Search each enzyme individually using the unique identifier provided (accession number).

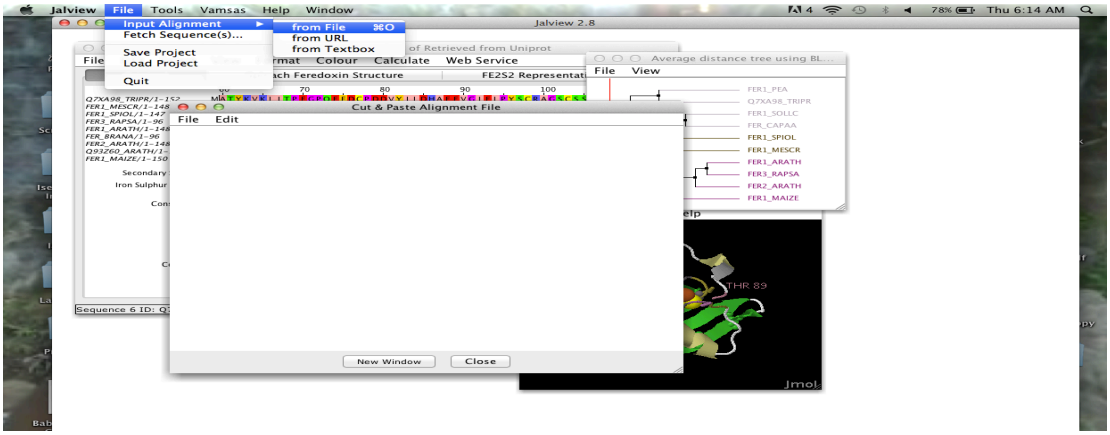
You can then get the protein sequence for each protein after selecting the protein returned from the NCBI search. Copy this sequence and paste into a text file. Save the file to the desktop. You may then use any of the computational biology tools to study these enzymes I will be demonstrating the use of Jalview.

Using Jalview to compare sequences

Save an additional copy of the text file, and change the file extension to save this copy as (YourFileName.fa) as opposed to (YourFileName.txt) save on desktop.

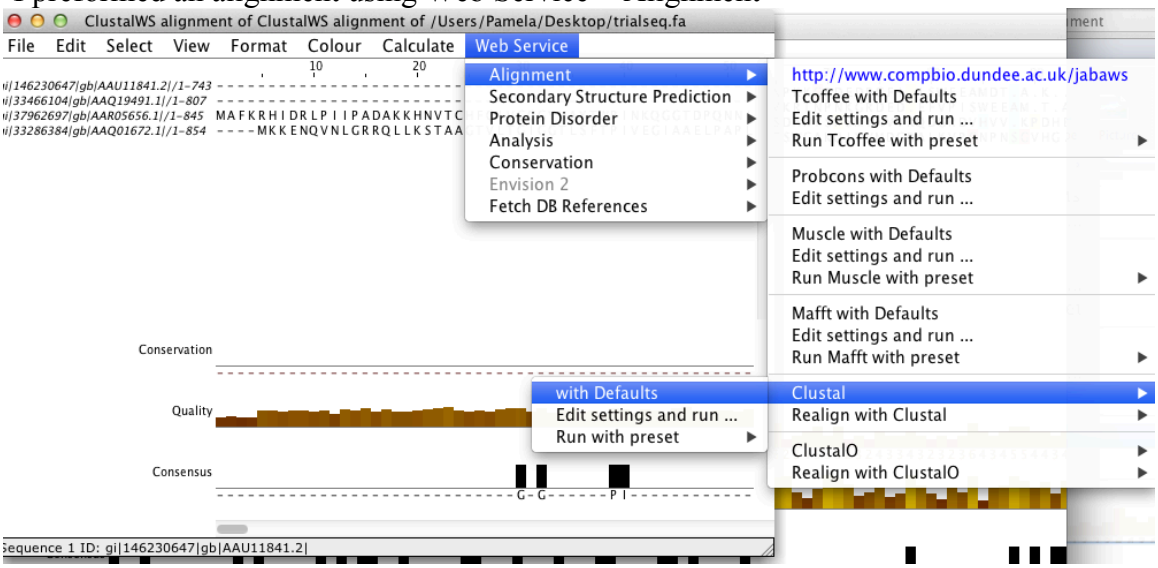
Run Jalview

Once loaded open Input this alignment in Jalview



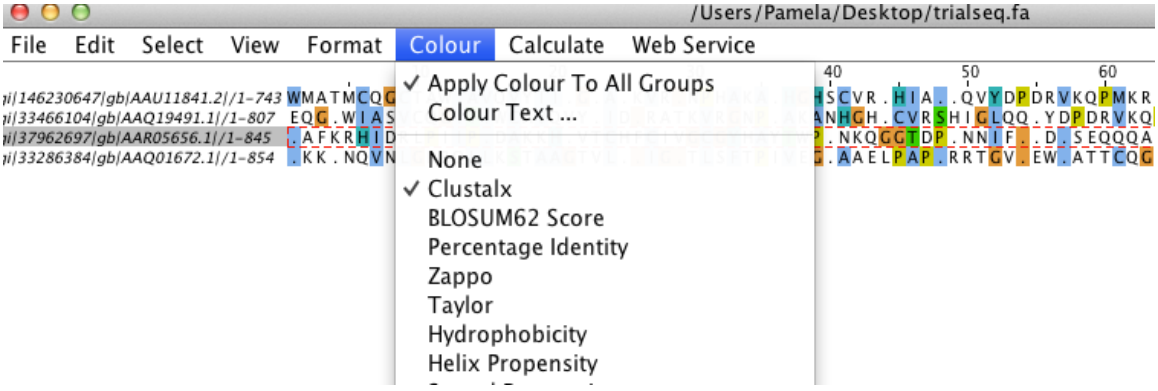
From the menu within the new window that appeared
You can analyze this data to your liking.

-I preformed an alignment using Web Service < Alignment



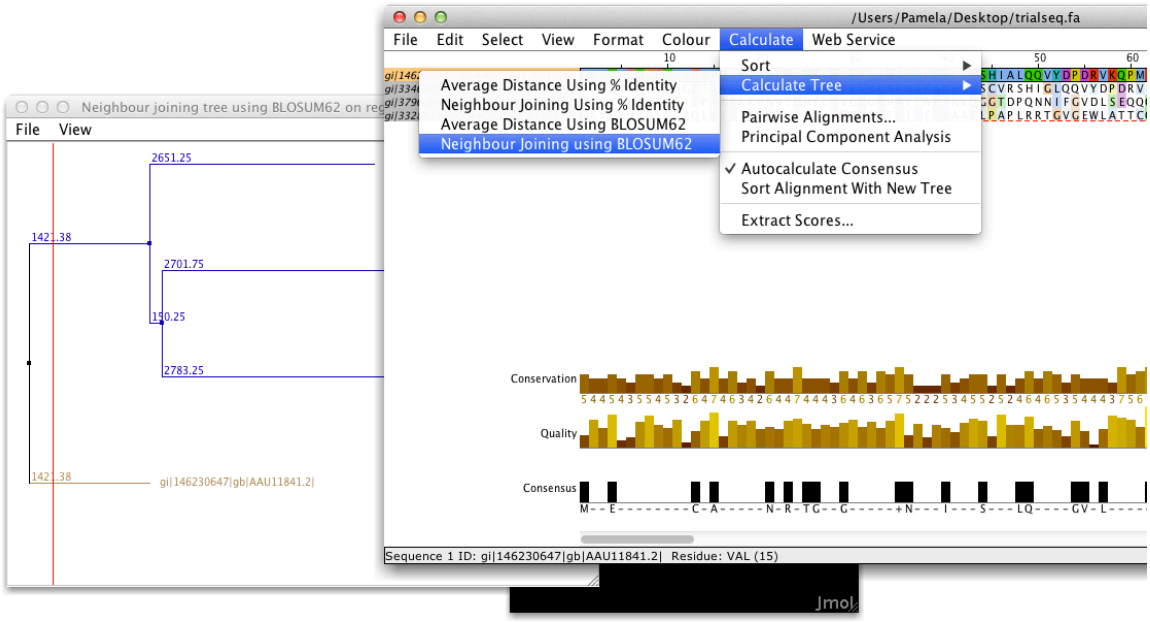
You can also change the colors to makes some features more apparent

-I used Colour < Clustalx



You can also explore many other options with this tool to help you analyze your data.

-I built a phylogenetic tree using Calculate < Calculate Tree



Be sure to **Document your Methods**

For your reference

Color scheme meaning: Institute of Biotechnology

http://ekhidna.biocenter.helsinki.fi/pfam2/clustal_colours

The column residue is given first in the round brackets; more than one may be specified, in which case the rules apply to each of these residues. Next, the rule or rules are given in curly braces; only one rule has to be met for the colour to be applied. The minimum percentage is given first, followed by the residue or residues which must meet or exceed this percentage within the column. If a group of residues is concatenated together, such as 'rstv', then any combination of these residues in total must meet or exceed the given percentage for the colour to be applied. For residues or residue groups separated by commas, at least one of these must by itself exceed the percentage.

ClustalX	Jalview	Pfam[2]	Rules:
blue	blue	blue	(W,L,V,I,M,F): {50%, p}{60%, wlvimafcyhp} (A): {50%, p}{60%, wlvimafcyhp}{85%, t,s,g} (C): {50%, p}{60%, wlvimafcyhp}{85%, s}
red	red	red	(K,R): {60%, kr}{85%, q}
green	green	green	(T): {50%, ts}{60%, wlvimafcyhp} (S): {50%, ts}{80%, wlvimafcyhp} (N): {50%, n}{85%, d} (Q): {50%, qe}{60%, kr}
pink	pink	pink	(C): {85%, c}
magenta	magenta	magenta	(D): {50%, de,n} (E): {50%, de,qe}
orange	orange	orange	(G): {always}
cyan	cyan	cyan	(H,Y): {50%, p}{60%, wlvimafcyhp}
yellow	yellow	yellow	(P): {always}

Revised Individual Lesson Plan

Pamela Watson

July 1, 2013 11:00 -12:30

- What do you want students to learn during the starter(s)?
    - Students will work on the STEM practice of **identification of the problem** (practice goal)
    - **Several biological phenomena can be observed using computational bio tools**
- What do you want students to learn or accomplish during the starter?
    - Students will be divided into groups of 2 or 3 (6 groups in all)
    - Students will share a terminal and follow along with me as I show a computational experiment using two related groups of enzymes described in a review article. (NCBI, Jalview- alignment)
    - Groups will make observations about each alignment and why these enzymes group together phylogenetically (they will record record questions that arise pertaining to biological phenomina in their google doc).
    - Students will observe my demonstration of using a computational tool to conduct an experiment
    - Learners will notice similarities and differences between related enzymes
- What is your rationale for designing the starter(s) this way?

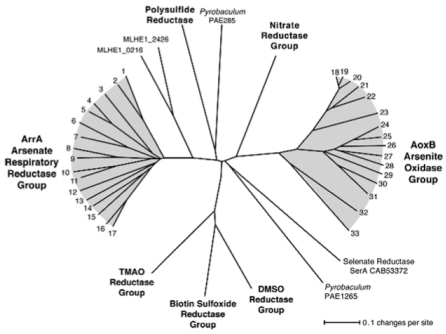
    - We're engaging students in STEM practices by having them:
    - find information from published literature
    - choose relevant enzymes and search web databases
    - form hypothesis driven questions
    - We want students to have ownership over their observations so I will allow them to further investigate with any tool they choose under the constraint of only related enzymes to ensure they can understand the content goal
- Describe the key points of contexting for the starters
    - These enzymes have distinct features despite their structural similarity.
- Describe how the starter(s) raise “how” or “why” questions that lead to investigations, and can be addressed by engaging in STEM practices
    - As students look at each 'sequence alignment', they should be interpreting the data presented,
    - They will then begin asking questions about the biology behind each data set however they will have to take time to compare each data point to understand the biology
    - Learners will investigate the characteristics that allow similar enzymes to have different functions.

Individual Instructional Plan (Pam)

- What facilitator will be saying, demonstrating, pointing out

- I will start off by introducing the arsenite oxidases and arsenate reductases and their functions using powerpoint slides
    - Next Show that these proteins look extremely similar (cartoon diagram from literature found in powerpoint presentation)
    - Finally Show that they do not group together phylogenetically (sequence tree from published literature) Tiffany mentioned that I should not explain in great detail. The students should interpret the phylogenetic tree themselves
- I'll ask the students to follow along using their worksheet as I look at the sequence of some individual enzymes accession numbers found in published review article)

- The list provided in the worksheet contains characterized Arsenic Oxidases, Arsenic Reductases, some uncharacterized enzymes that are either an oxidase or reductase, and some other respiratory emzymes that can be used as out-groups
    -
    - I will use a few assession numbers from each group and align in jalview(hand out has details on how to use jalview
- Students will reference Saltikov 2007 and identify 4 oxidases and 4 reeducates from figure 2.

- 

-
- They will use NCBI to get the protein sequence for these enzymes
- They will use clustal omega to align these sequences or another alignment software of choice
- What do you want to make sure learners notice?
    - I want the students to notice that although the protein structures are really similar some key residues are consistently different between the two enzyme classes.
    - I should not say this explicitly the students should notice this during inquiry
- What will students be doing?
    - I'll ask students to choose some other arsenite oxidases and arsenate reductases from the literature review and align those sequences and explore their questions from there

    - I'll ask the students to each come up with questions based on what they have seen so far
- What are the most important questions or observations you would like learners to raise?
    - What causes the two enzymes to group differently?
    - Do all enzymes in a group share characteristics?
    - Students should be able to pick out homologous and non-homologus regions of the enzymes

Jigsaw Revision notes:

The Jigsaw consisted of 4 different prompts. The students were assigned to groups containing up to 4 students that did not work together on the inquiry. They were then given a prompt that was different from the inquiry any of the students in the group pursued for the first two days. Very little facilitation will be given to encourage the students to assimilate all of their experiences to solve the jigsaw. At the of the jigsaw groups will be given 8 minutes to present there findings to the class orally and get questions from the students and facilators.

- What facilitator will be saying, demonstrating, pointing out
- The learners were told at the beginning of the Jigsaw that each group member must contribute in order to address the prompt in a timely manner.
- The learners were also told that we would like to hear from each group member during the oral presentation.

- Because the computers were limited it is important to facilitate all group members.
- Facilitators should ask questions of all 4 group members and encourage rotations of the learner who is working on the computer
-
- Originally some students came in late and others did not show up at all so the predetermined groups moved at very different speeds based on the number of group members or when they showed up. To correct for this I gave a few extra minutes for investigation. In the future the groups should be reassigned if some students do not show up instead of giving the disadvantaged groups more time.

Revision:

The smaller group initially struggled through jigsaw (because their pre-assigned group members were missing they started off with less ideas) while the complete groups finished quickly and became disruptive. I provided more facilitation to the smaller group (only 2 students) in order to point them in the right direction and help them catch up.

One alternative could have been taking someone from one of the other groups so there would be two groups of 3. In the end the smaller group was still able to address the prompt thoroughly with several lines of evidence. so several smaller groups (of

3) for the jigsaw probably would have worked in the same amount of time with less distraction.