UNIVERSITY OF CALIFORNIA
SANTA CRUZ

Genomic and Physiologic Characterization of a Freshwater Photoarsenotroph, *Cereibacter azotoformans* str. ORIO, Isolated From Sediments Capable of Light-Dark Arsenic Redox Cycling

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

DOCTOR OF PHILOSOPHY

in

MICROBIOLOGY AND ENVIRONMENTAL TOXICOLOGY

by

Sanjin Mehic

March 2022

The Dissertation of Sanjin Mehić is approved:

_____
Professor Chad W. Saltikov

_____
Professor Karen M. Ottemann

_____
Professor Jonathan P. Zehr

_____
Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of contents

List of Figures

Abstract

Genomic and physiologic characterization of a freshwater photoarsenotroph, *Cereibacter azotoformans* str. ORIO, from sediments capable of light-dark arsenic redox cycling

Sanjin Mehić

Thesis statement:

Photoarsenotrophy is an anoxygenic photosynthesis-dependent arsenite oxidation pathway encoded by the Arx gene cluster and is linked to light-dark cycling of arsenic in freshwater environments.

This dissertation uses techniques from microbiology, molecular genetics, DNA sequencing, and analytical chemistry to characterize photoarsenotrophy in freshwater environments. The hypothesis is that photoarsenotrophy occurs in freshwater environments and is associated with light-dark cycling of arsenic, which may have a different effect on arsenic biogeochemical cycling when compared to Arx-type arsenotrophy. Light-dark arsenic redox cycling is defined herein as variations in arsenic species, arsenite and arsenate, that correlate to light or dark phases.

Since the discovery of photoarsenotrophy in 2008, less than a dozen Arx-dependent arsenotrophs have been isolated, and light-dependent arsenite oxidation has only been detected in three genera (*Ectothiorhodospira sp.* strains MLW-1, PHS-

1, BSL-9, *Ect. shaposhnikovii* strains DSM 243 and DSM 2001, *Halorhodospira halophila* SL-1, and now *Cereibacter azotoformans* str. ORIO). The other studied Arx-dependent arsenotrophs couple arsenite oxidation to anaerobic respiration (i.e., nitrate reduction) instead of anoxygenic photosynthesis (*Alkalilimnicola ehrlichii* str. MLHE-1, *Azoarcus sp.* CIB, *Sterolibacteraceae* strain M52, *Desulfotomaculum* strain TC-1, and *Halomonas* sp. ANAO440). In this thesis, I define the light requirement by referring to photosynthesis-dependent arsenite-oxidation as photoarsenotrophy, and photosynthesis-independent Arx-type arsenite-oxidation as Arx-type arsenotrophy. This difference in light requirement is important as it introduces the possibility that light-dark cycling of arsenic could occur in environments containing a photoarsenotroph and arsenate reducer, since Arx-type arsenotrophs are able to oxidize arsenite in the dark so long there is availability of a terminal electron acceptor.

The first question investigated was: Does photoarsenotrophy occur in freshwater environments? This question was studied through the isolation and genetic characterization of a novel photoarsenotroph from freshwater sediments in Owens River, CA, USA. Our results show the photoarsenotrophs are present in freshwater environments and harbor the *arx* genes required for photoarsenotrophy.

The second question investigated was: Can arsenic be light-dark cycled in freshwater environments? In other words, do concentrations of different arsenic forms correlate to the light phase or dark phase of a day? To answer this question, we performed anaerobic microcosm studies with sediment collected from Owens River.

Arsenic speciation was measured over light-dark cycles and was followed by metagenome sequencing and analysis. The results provide evidence of light-dark cycling in freshwater sediments and the potential genes involved in the cycle.

The third question investigated was: Can we develop a new model organism for studying photoarsenotrophy? We successfully determined that *Cereibacter azotoformans* str. ORIO is genetically malleable using traditional cloning techniques and can serve as a model for studying the biological mechanism underlying photoarsenotrophy. This was achieved by validating the role of the *arxA* gene in ORIO and characterizing the physiology surrounding arsenite oxidation.

Taken together, these studies show photoarsenotrophy occurs in freshwater environments, *arx* genes are more widespread than previously known, ORIO can serve as a model organism for studying photoarsenotrophy, and evidence of light-dark arsenic redox cycling in freshwater sediments.

Dedication

The thesis work is dedicated to my family, who have supported me and my scientific journey by being extremely patient and supportive through all my struggles. Hvala puno mojoj familiji, mojoj ženi, i familiju od moje žene. Bez vas nema nista od mene, i volim vas kao oci u glavi. I would also like to dedicate this work to the late Dr. Ronald Oremland, who helped pioneer the field of microbial arsenic transformations and inspire many of us to investigate these incredible and important processes.

Acknowledgments

First and foremost, I would like to thank my advisor, Dr. Chad W. Saltikov for allowing me to join his laboratory and providing me the guidance needed to succeed in studying arsenic-metabolizing bacteria. Chad has truly been a role model and friend during my time in graduate school. I will forever cherish our time together and the valuable lessons I learned. I would like to also acknowledge Jaime Hernandez-Maldonado for his support when I joined the lab. I am thankful for the Saltikov lab members Nicole Schrad, Juliana Nzongo, as well as former member Esther Munoz, for their friendship and support in the lab. I also thank former Salitkov Lab member Esra Mescioglu, for initial culturing efforts of strain ORIO.

I would like to thank the programs and fellowships that have supported me financially throughout my time as UCSC. This dissertation was made possible through the funding of the UC Regents, the Moore Foundation, UCSC STARS, and the UCSC Presidents Dissertational Year Fellowship. In addition, financial and academic support from the UCSC STEM Diversity Program helped prepare me for graduate school. This includes all the staff and student workers who help run and maintain these modes of support on campus.

I would like to thank and commemorate the late Dr. Ronald Oremland for his guidance and insight, as part of my thesis committee. In addition, I thank my entire thesis committee, Dr. Jonathan P. Zehr, Dr. Karen M. Ottemann, and Dr. Chad Saltikov. I am thankful to have to support of such outstanding scientists.

I give many thanks to the MAL facility and Brian Dreyer for expert guidance in our arsenic speciation analyses.

Finally, I give thanks to my Summer Internship Program 2020 high school students Isaac Chang, Madeline Day, and Pranav Kirti for their help processing the JGI dataset.

Chapter 1. Thesis overview

Section 1.01 Introduction

Microbiology and environmental microbiology are the main fields of interest in this dissertation. Specifically, this dissertation aims to advance the field of photosynthetic arsenic metabolism, termed "photoarsenotrophy," by investigating the occurrence of this process in freshwater environments, developing a model organism for studying the biological mechanisms underlying this process, and providing an update to the occurrence of *arx* genes around the world. I aim to also provide evidence and genetic basis for light-dark arsenic redox cycling in freshwater environments.

Section 1.02 Current knowledge

The first reports of microbes converting arsenite to arsenate (As(V)) in a sunlight-dependent manner came in 2008 (Budinoff and Hollibaugh, 2008; Kulp *et al.*, 2008). Termed "anoxygenic photosynthetic arsenite oxidation" or "photoarsenotrophy," this process has been shown to be mediated by bacteria containing the *arx* gene cluster, typically from extreme environments such as hot springs, soda lakes, or hypersaline lakes (Oremland *et al.*, 2002; Hoeft *et al.*, 2007; Budinoff and Hollibaugh, 2008; Kulp *et al.*, 2008; Hamamura *et al.*, 2014; Hernandez-Maldonado *et al.*, 2016; Hoeft-McCann *et al.*, 2017; Wu *et al.*, 2017). Of the eleven previously studied photoarsenotrophs only two were found in non-extreme environments, but do not require photosynthesis for the process and instead perform a light-independent version of photoarsenotrophy that couple arsenite oxidation to

nitrate reduction (Ospino *et al.,* 2019; Durante-Rodríguez *et al.,* 2019). This non-photosynthetic Arx-type arsenite oxidation is herein referred to as Arx-type arsenotrophy. The difference between photoarsenotrophy and Arx-type arsenotrophy give rise to unique ecological niches, since one depends on sunlight and is restricted to the day while the other relies on the availability of a terminal electron acceptor (e.g., nitrate).

Section 1.03 Knowledge gaps

Prior to this report, no photoarsenotrophs had been isolated from freshwater environments, and autotrophy was the main mode of metabolism associated with bacteria participating in this process. A major impediment for advancing the study of photoarsenotrophs has been that traditional genetic manipulation techniques are difficult due to the high pH, temperature, and salinity needed to mimic the extreme environments from which previous photoarsenotrophs were isolated. The possibility of freshwater environments harboring photoarsenotrophs provides the opportunity for isolating more robust genetic systems. Furthermore, the co-occurrence of photoarsenotrophs and arsenate-reducing bacteria could allow for a mutualistically shared arsenic pool. In other words, environments containing a photoarsenotroph and arsenate reducer could result in a light-dark arsenic redox cycle where arsenite is oxidized during the light phase, and arsenate reduction during the dark phase. No reports have documented the biological cycling of arsenic in freshwater ecosystems or used metagenomics to investigate light-dark variations of arsenic concentrations.

Section 1.04 Hypotheses

The purpose of this study was to develop a more robust genetic system for future studies aimed at characterizing the biological mechanisms underlying photoarsenotrophy, to investigate the occurrence of photoarsenotrophy in freshwater systems, and to determine if light-dark cycling of arsenic can occur in freshwater sediments. The main hypotheses are (i) photoarsenotrophs can be isolated from arsenic-rich freshwater sediments, (ii) light-dark arsenic redox cycling occurs in environments containing photoarsenotrophs and arsenate-reducers, and (iii) freshwater photoarsenotrophs can be genetically manipulated using traditional cloning techniques.

Section 1.05 Research design

To test our hypotheses and provide evidence of freshwater photoarsenotrophy, light-dark arsenic cycling, and a genetic system, we first chose a study site that could harbor such microbes. Owens River in the Long Valley Caldera (USA) was chosen, due to its active hydrothermal system that brings naturally occurring arsenic from the Earth's crust to the surface via hot springs. Samples collected from Owens River water and sediment taken in July 2021 were shown to contain 7 μM As(III), or about 900 ppb which is almost 100 times more than EPA drinking water standard. Preliminary analysis of an arsenite-oxidizing bacteria isolated from Owens River showed a conserved *arxA* and requirement of infrared (850 nm) light for anaerobic growth. In addition, 16S rRNA analysis suggested the isolated bacteria was closely related to the genus *Cereibacter* (formerly known as *Rhodobacter*) which has served as a model organism for studying anoxygenic photosynthesis for over 40 years. For

these reasons, we chose to develop this isolate, named *Cereibacter azotoformans* str. ORIO, as a model organism for the study of photoarsenotrophy.

To study the possibility of light-dark arsenic redox cycling, we again chose Owens River to prospect for arsenite-oxidizing and arsenate-reducing bacteria. Instead of performing isolations for individual oxidizers and reducers, we took sediment samples from Owens River and made arsenite-spiked microcosms that were light-dark cycled and measured for arsenic speciation. Arsenic speciation and metagenomic analysis of the microcosms provided the first evidence and genetic basis for a light-dark arsenic redox cycle in freshwater environments.

Section 1.06 Scientific merit of this research

This report demonstrates photoarsenotrophy and light-dark arsenic cycling can occur in freshwater sediments and provides a more robust genetic system through the genomic and physiological characterization of a novel freshwater photoarsenotroph named *Cereibacter azotoformans* str. ORIO (**O**wens **R**iver **I**solate **O**xidizer). Detection of freshwater photoarsenotrophy and light-dark arsenic cycling has implications for the arsenic biogeochemical cycle and future studies should investigate the co-occurrence of arsenite-oxidizers and arsenate-reducers in the environments, and how their interactions can give rise to previously unknown arsenic biogeochemical cycling. The ability to understand such processes in the environment is of utmost importance for public health because arsenic-contaminated groundwater and rice affects hundreds of millions of people worldwide and among the known

pollutants in drinking water it poses the highest risk factor for causing cancer in humans.

Section 1.07 Chapter overviews

Chapter 2 provides a review of photoarsenotrophy, and a brief overview of other relevant arsenite oxidation and arsenate reduction pathways. Chapter 3 investigates the occurrence of photoarsenotrophy and light-dark arsenic cycling in freshwater sediments. First, a novel freshwater photoarsenotroph from Owens River (CA, USA) was isolated and characterized, then freshwater sediment collected from the same river was investigated for light-dark arsenic cycling. Chapter 4 includes further analysis and characterization of photoarsenotrophy in ORIO by validating the role of the *arxA* gene using traditional cloning techniques. Together, this work establishes ORIO as a robust model organism for the study of photoarsenotrophy. Chapter 5 concludes what was learned through these investigations and provides future directions for photoarsenotrophy research.

Section 1.08 References

Budinoff, C. R., & Hollibaugh, J. T. (2008). Arsenite-dependent photoautotrophy by an Ectothiorhodospira-dominated consortium. ISME Journal, 2(3), 340–343.

Durante-Rodríguez, G., Fernández-Llamosas, H., Alonso-Fernandes, E., Fernández-Muñiz, M.N., Muñoz-Olivas, R., Díaz, E., & Carmona, M. (2019) ArxA From Azoarcus sp. CIB, an Anaerobic Arsenite Oxidase From an Obligate Heterotrophic and Mesophilic Bacterium. Front. Microbiol. 10: 1699.

Hamamura, N., Itai, T., Liu, Y., Reysenbach, A. L., Damdinsuren, N., and Inskeep, W. P. (2014) Identification of anaerobic arsenite-oxidizing and arsenate-reducing bacteria associated with an alkaline saline lake in Khovsgol, Mongolia. Environ. Microbiol. Rep. 6: 476-482.

Hernandez-Maldonado, J., Sanchez-Sedillo, B., Stoneburner, B., Boren, A., Miller, L., McCann, S., … Saltikov, C. (2016). The genetic basis of anoxygenic photosynthetic arsenite oxidation. Environmental Microbiology, 0, 1–29.

Hoeft, S. E., Blum, J. S., Stolz, J. F., Tabita, F. R., Witte, B., King, G. M., … Oremland, R. S. (2007). Alkalilimnicola ehrlichii sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. International Journal of Systematic and Evolutionary Microbiology, 57(3), 504–512.

Hoeft McCann, S., Boren, A., Hernandez-Maldonado, J., Stoneburner, B., Saltikov, C., Stolz, J., *et al.*, (2017) Arsenite as an electron donor for anoxygenic

photosynthesis: description of three strains of Ectothiorhodospira from Mono Lake, California and Big Soda Lake, Nevada. Life 7: 1–14.

Kulp, T. R., Hoeft, S. E., Asao, M., Madigan, M. T., Hollibaugh, J. T., Fisher, J. C., … Oremland, R. S. (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. Science, 321(5891), 967–970.

Oremland, R. S., Hoeft, S. E., Santini, J. M., Bano, N., Hollibaugh, R. a, & Hollibaugh, J. T. (2002). Anaerobic Oxidation of Arsenite in Mono Lake Water and by a Facultative, Arsenite-Oxidizing Chemoautotroph, Strain MLHE-1. Society, 68(10), 4795–4802.

Ospino MC, Kojima H and Fukui M (2019) Arsenite Oxidation by a Newly Isolated Betaproteobacterium Possessing arx Genes and Diversity of the arx Gene Cluster in Bacterial Genomes. Front. Microbiol. 10: 1210.

Wu, G., Huang, L., Jiang, H., Peng, Y., Guo, W., Chen, Z., *et al.*, (2017) Thioarsenate formation coupled with anaerobic arsenite oxidation by a sulfate-reducing bacterium isolated from a hot spring. Front. Microbiol. 8: 1336.

Chapter 2. Bacterially mediated arsenic metabolism: A review with emphasis on photoarsenotrophy

Section 2.01 Abstract

Of all the pollutants found in drinking water, for humans, arsenic poses the highest risk factor for developing cancer. Arsenic contaminated water is not a problem geographically restricted to areas such as Bangladesh or India. Hundreds of millions of people on Earth are likely consuming arsenic contaminated water or ingesting arsenic contaminated food. Studies in the United States have detected arsenic in baby food, juices, and drinking water across the country, and implicated low levels of arsenic-contaminated drinking water in various cancers. The need to understand how arsenic is transformed in the environment is more important than ever. One main driver of arsenic transformation in nature are prokaryotic microorganisms. Bacteria and archaea possess certain genes which allow them to transform arsenic between its various forms, some forms are more toxic to life and some are less toxic to life. This review aims to provide the current knowledge of genes and mechanisms of the major forms of prokaryotic arsenic metabolism and transformation, with emphasis on anoxygenic photosynthetic arsenite oxidation (photoarsenotrophy). Traditional and genetic classifications of arsenic-metabolizing bacteria are discussed. Arsenic metabolism regulatory mechanisms are briefly explained. Lastly, we provide a brief review of metal redox cycling in nature and hypothesize a connection to photoarsenotrophy.

Section 2.02 Introduction

Over 100 years have passed since the first publication of microbially-mediated arsenic transformation (Green 1918). Since then, researchers have uncovered several key pathways for arsenic metabolism in bacteria, began to understand the ecology of arsenic, and even detected the ability of gut microbiome to protect the host from arsenic toxicity (Oremland and Stolz, 2003; Amend *et al.*, 2014; Coryell et al., 2018; McDermott et al., 2020). This review will cover the genes, enzymes, and mechanisms involved in bacterial arsenic metabolism, with a focus on anoxygenic photosynthetic arsenite oxidation and the environmental implications. The regulation of these processes will also be reviewed. The need to understand how and when microbes can use arsenic in the environment is more important than ever, as millions of people in Asia, South America and North America are consuming arsenic contaminated water (Harvey *et al.,* 2005; Fendorf *et al.,* 2010; Burgess *et al.,* 2010; Mailloux and Trembath-Reichert, 2013). Further, rice is a staple food for over a billion people, many of which rely on arsenic contaminated ground water for agricultural applications (Li *et al.,* 2011; Zhao *et al.,* 2010). There is a growing concern for mitigating the absorption of arsenic by plants, while maintaining quality yields and low costs. Several recent studies have focused on the microbial effects on arsenic uptake in rice plants (Abedi and Mojiri, 2020; Moens *et al.,* 2020; Mawia *et al.*, 2021).

Arsenic in the environment is due to naturally occurring processes, and the forms that are most relevant to public health are: 1) arsenite (AsO_3^{3-}, As(III)), and 2) arsenate (AsO_4^{3-}, As(V)). Arsenite is the more toxic form of the two, because it can

bind sulfur groups in proteins and thus disrupt the structure and function of many different enzymes and proteins. Another contributing factor to the greater toxicity of arsenite is the fact that it is much less likely to form arsenic minerals and precipitate out of solution than arsenate, thus is more mobile in aquatic systems. However, arsenate can disrupt ATP synthesis by mimicking phosphate, which leads to reduced ATP production thereby poisoning the cell. Although microorganisms have been recognized for their capabilities to transform arsenic for over 100 years, further investigations aimed at understanding how and when bacteria transform arsenic in the environment will ultimately help identify new forms of bioremediation (Verma *et al.*, 2016; Satyapal *et al.*, 2016), agricultural practices (Das et al 2016; Das *et al.*, 2016; Lin *et al.*, 2017; Boye *et al.*, 2017), and even provide novel biosensing tools for detecting arsenic in the environment (Turner *et al.*, 2019).

Traditionally, the microbial biology of arsenic metabolism can be classified into four different groups (Oremland and Stolz, 2003): 1) arsenate-resistant microorganisms (ARMs), 2) dissimilatory arsenate-reducing prokaryotes (DARPs), 3) chemoautotrophic arsenite oxidizers (CAOs), 4) heterotrophic arsenite oxidizers (HAOs). However, recent investigations have revealed a new group of bacteria, photosynthetic arsenite oxidizer (PAOs). The key to these classifications is whether metabolism is coupled to detoxification or energy generation, but heterotrophy and lithotrophy also play a role in these assignments. These classifications alone are not enough to capture the complexity of bacterial arsenic metabolism. Recent investigations have demonstrated some bacteria fall into multiple categories, such as

heterotrophic arsenite oxidizers that use electrons from arsenite for growth, or photoheterotrophic arsenite oxidizers that detoxify arsenite through oxidation. Therefore, the gene clusters used during arsenic metabolism are another way to classify arsenic transforming bacteria. The *arr* genes are involved in anaerobic arsenate respiration, the *aio* genes in aerobic and anaerobic arsenite oxidation, and the *arx* genes in arsenite oxidation coupled to anoxygenic photosynthesis or anaerobic respiration (Figure 1).

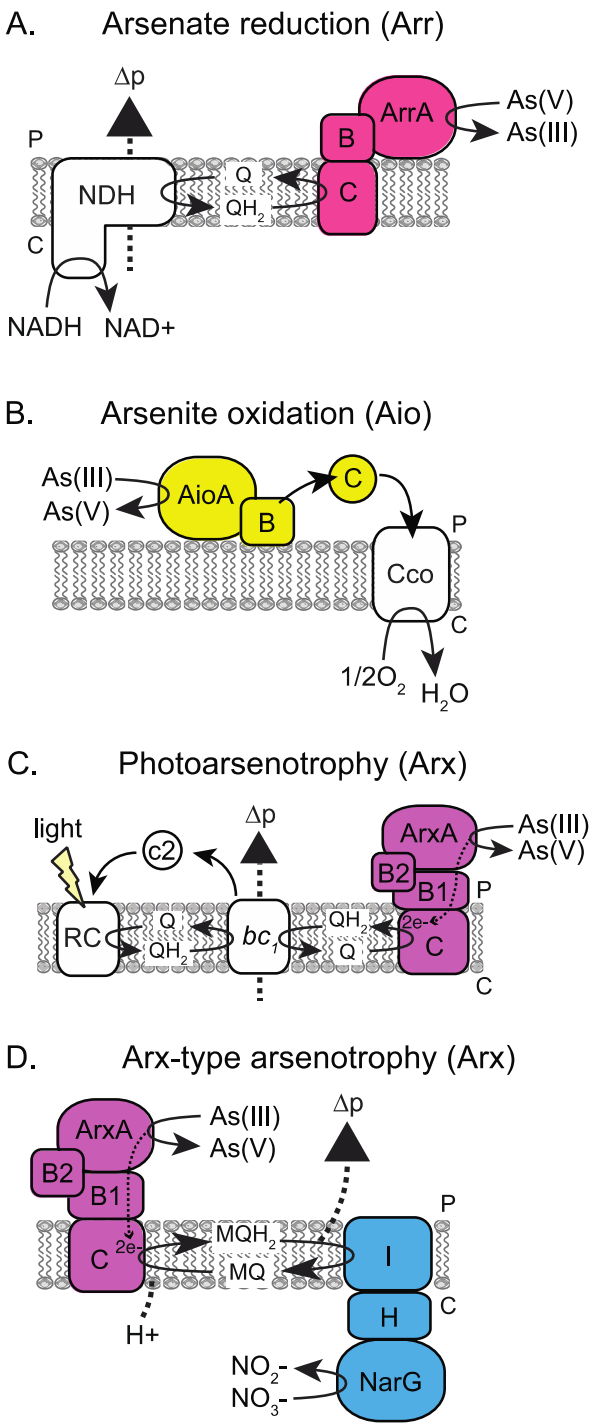


Figure 1. Diagrams of enzyme subunits involved in arsenotrophy. (A) The Arr pathway in strain *Bacillus selenitireducens* str. MLS-10. (B) The Aio pathway in *Rhizobium sp.* NT-26. (C) Photosynthetic Arx pathway in *Cereibacter azotoformans*

str. ORIO. (D) The Arx pathway in *Alkalilimnicola ehrlichii* str. MLHE-1. P = periplasm, C = cytoplasm, NDH = NADH dehydrogenase, Cco = cytochrome c oxidase, RC = reaction center of light harvesting complex, c2 = cytochrome c2. Bc_1 = bacteriochlorophyll a. Adapted from Oremland *et al.*, 2017.

Section 2.03 Arsenite oxidation

Chemoautotrophs

CAOs can gain energy from transforming arsenic and do so via arsenite oxidation. These microorganisms are strict chemolithoautotrophs, which typically use oxygen as the terminal electron acceptor, but studies have shown that nitrate, selenate, and chlorate can also be used for this purpose (Oremland *et al.,* 2002; Rhine *et al.,* 2006; Hoeft *et al.,* 2007; Zargar *et al.,* 2010; Rodriguez-Freire *et al.,* 2012; Amend *et al*., 2014). The energy generated from this process is coupled to fixing CO_2 into organic cellular matter via Arx or Aio pathways.

Under both aerobic (Santini *et al.,* 2000; Santini *et al.,* 2002; Rhine *et al.,* 2008; Garcia-Dominguez *et al.,* 2008) and anaerobic nitrate-reducing conditions (Oremland *et al.,* 2002; Rhine *et al.,* 2006), bacteria are able to use energy and reducing power from As(III) oxidation to fix CO_2 and generate ATP. One example, *Alkalilimnicola ehrlichii* str. MLHE-1 was isolated from an alkaline hypersaline soda lake, Mono Lake, in California, USA, and shown to grow anaerobically with inorganic electron donors (arsenite, sulfide, thiosulfate, hydrogen) while reducing nitrate to nitrite, and fixing CO_2 via ribulose 1,5-bisphosphate carboxylase (RuBisCO) (Hoeft *et al.,* 2007). This led to the discovery of a novel arsenite oxidase

gene, *arxA*, which encodes for a protein that is phylogenetically more closely related to ArrA than AioA (Zargar *et al.*, 2010; Zargar *et al.*, 2012). The *Alphaproteobacteria Rhizobium sp*. str. NT-26 is a chemolithoautotrophic arsenite oxidizer isolated from a gold mine in Australia. Although NT-26 is more closely related to other plant-associated microbes, it has lost the ability to act as a symbiont with legumes, and instead gained the ability to oxidize arsenite through acquisition of a mega-plasmid. Interestingly, flagellar motility genes are located on the chromosome, but the genes associated with arsenite oxidation on the mega-plasmid are in direct control of regulating motility (Andres *et al.*, 2013).

Heterotrophs

HAOs, which make up a significant portion of known arsenic transformers, also oxidize arsenite. However, since many do not oxidize arsenite for the purpose of energy production, the more general term Detoxifying Arsenite Oxidizers (DAOs) can be used in place. In this case, the Aio or Arx enzymes are used but do not generate energy. While *Alkalilimnicola ehrlichii* str. MLHE-1 was able to grow heterotrophically, it could not oxidize arsenite during this mode of growth (Hoeft *et al.*, 2007). Alternatively, *Agrobacterium tumefaciens* GW4 was shown to grow heterotrophically while oxidizing arsenite (or antimonite, Sb(III)), via AioA, to harness energy and reducing power to generate NADH, proton motive force, and ATP (Wang *et al.*, 2015; Wang *et al.*, 2017). *Hydrogenophaga sp*. str. NT-14 was found to couple arsenite oxidation to oxygen reduction for energy production and believed to use a c-type cytochrome for accepting electrons from the arsenite oxidase AioA

(Hoven *et al.,* 2004). A recent study reported the first heterotrophic arsenite oxidizing bacterium, *Azoarcus sp.* CIB, which uses the extra electrons from arsenite for growth (Durante-Rodriguez *et al.,* 2019). Arsenite oxidation in CIB is coupled to nitrate reduction, therefore it is an Arx-type arsenotroph and not a photoarsenotroph. More recently, the first suspected freshwater light-dependent arsenite oxidizing bacterium was isolated from Owens River in California, USA (this work). This *Alphaproteobacteria* belonging to the *Cereibacter* genus is a heterotrophic purple non-sulfur bacterium and a photoarsenotroph. *Cereibacter* are formerly the *Rhodobacter* genus, which have been studied as model organisms for photosynthesis for over 40 years. This new addition to the photoheterotrophic group of arsenite oxidizers will be extremely useful for future studies aimed at understanding the ecology and biological mechanisms of photoarsenotrophy.

Phototrophs

Phototrophs convert solar energy into chemical energy by absorbing sunlight with pigments in their light-harvesting (LH) antenna system. The absorbed energy is transferred to the reaction center (RC), which initiates the electron transfer process through redox reactions in the electron transport chain. Electron donors and acceptors help with proton translocation, ultimately creating proton motive force to drive ATP formation and other energy-consuming processes needed to grow. A newer, photosynthesis coupled, form of arsenite oxidation was discovered from biofilms containing *Ectothiorhodospiraceae* in the anoxic brine pools of Paoha Island at Mono Lake (Kulp *et al.,* 2008). These purple bacteria, named *Ectothiorhodospira* strains

MLW-1, BSL-9 and PHS-1, were shown to oxidize arsenite anaerobically in an anoxygenic and light dependent manner, using ArxA to generate proton motive force, ATP, and fix CO_2 (Hoeft *et al.*, 2010; Hernandez-Maldonado *et al.*, 2016; Heoft-McCann *et al.*, 2017). More interestingly, experiments showed light-dependent arsenite oxidation, which has new implications for light-dark cycling of metals in nature (Hernandez-Maldonado *et al.*, 2016). Simulating this light-dark metabolism with an oxidizer and reducer would be interesting to identify how arsenic might be cycled in the environment and to further understand bacteria-bacteria interactions in biogeochemical cycles. Several studies investigating how bacteria influence arsenic uptake in plants have been conducted recently, but it is rare to find any that specifically aim to cycle arsenic with an oxidizer and reducer to identify changes in arsenic uptake by plants (Das *et al.*, 2016; Lin *et al.*, 2016; Das *et al.*, 2016; Reid *et al.*, 2017).

Section 2.04 Arsenate reduction

DARPs use arsenate as a terminal electron acceptor for anaerobic respiration. They can reduce arsenate to conserve energy and are predominantly heterotrophic strict anaerobes. Typically, complex organic compounds (e.g., yeast extract, peptone, tryptone) serve as electron donors, but some organic acids (e.g., acetate, lactate, pyruvate) and inorganic compounds (e.g., hydrogen, reduced sulfur) can be used (Amend *et al*., 2014). Ultimately, electrons from an organic substrate will be oxidized and passed onto the electron transport chain, where they can then be used by ArrA to reduce arsenate to arsenite (Figure 1A).

Arsenic Detoxification

ARMs can reduce arsenate solely for detoxification purposes. Microorganisms containing an *arsC* gene, which codes for the ArsC protein, are usually assigned to this group. Some exceptions occur, for example *Bacillus sp.* str. 3.9 has an *arsC* gene but no arsenate reduction capability. Conversely, some microorganisms have been found to reduce arsenate while no *arsC* gene is identified (e.g., *Arthrobacter*) (Amend *et al.,* 2014). Therefore, since ARMs are typically determined by mere presence of an *arsC* gene or ArsC protein homolog, we can use the term Heterotrophic Arsenate Reducers (HARs) to classify all microbes which are verified to reduce arsenate, but do not gain energy from this process.

Section 2.05 Genes of arsenic metabolism and resistance

ArrAB

The *arrAB* genes encode for the arsenate reductase ArrA and ArrB subunits. Belonging to the prokaryotic molybdopterin-containing oxidoreductase family, ArrAB contain predicted 4Fe-4S cluster and molybdenum ion binding sites. In *Shewanella* str. ANA-3, a model organism for arsenate respiration, a TAT signal is found in *arrA*, and a Fnr binding site upstream of the gene. This suggests periplasmic activity for ArrAB, and transcriptional control via anaerobic conditions. Genetic studies have shown that *ars* detoxification genes aid but are not required for arsenate respiration in this bacterium (Saltikov *et al.,* 2003a), while further studies identified *arrAB* as the critical genes involved in this process (Saltikov *et al.,* 2003b). In addition, investigations have recognized that during arsenate respiration, ArrAB

facilitate electron transfer from cytochromes in the electron transport chain, while limiting the activity of other anaerobic pathways (Saltikov *et al.,* 2003b; Croal *et al.,* 2004; Murphy and Saltikov, 2007). A crystal structure of ArrAB was recently described by Glasser *et al.,* (2018) and provided clues about the enzymatic directionality of ArrA and ArxA. Specifically, Glasser *et al.,* suggest the active site motifs of ArrA (R/K-GRY) and ArxA (XGRGWG) give rise to arsenate reduction and arsenite oxidation, respectively. Until this structural distinction was made, many believed ArrA and ArxA to be the same class of bidirectional redox enzymes.

Arr regulation

In *Shewanella* str. ANA-3, *arr* genes have been found to be regulated in two ways. First, an ArsR family protein, ArsR2, binds the promoter region of the *arr* gene cluster and represses transcription under anaerobic arsenite absent conditions (Murphy and Saltikov, 2009). Similarities can be drawn between this regulation and that of many bacterial arsenic detoxification operons (*ars*). Several conserved cysteine residues have been shown to be critical for the binding of arsenite in both detoxifying and arsenate respiring bacteria (Shi *et al.,* 1994; Busenlehner *et al.,* 2003; Murphy and Saltikov, 2009). The second mechanism for regulating the *arr* operon is via cyclic AMP (cAMP)-cyclic AMP receptor (CRP). A *Shewanella sp.* ANA-3 strain containing a CRP mutation is unable to transcribe *arrA* when shifting arsenate containing cultures from aerobic to anaerobic, and further computational analysis found putative CRP binding motifs within the *arr* promoter region which was subsequently verified with EMSA (Murphy *et al.,* 2009). Furthermore, putative

adenylate cyclase (*cya*) genes were found in the genome, and when mutations were created in these genes, ANA-3 lost the ability to respire arsenate (Murphy and Saltikov, 2007). This suggests that the catabolite repression system has an essential role in regulating arsenate respiration.

AioAB

Part of the molybdenum-containing dimethyl sulfoxide (DMSO) reductase family of oxidoreductases, AioA was the first bacterial arsenite oxidase reported which coupled arsenite oxidation to energy conservation and growth (Rhine *et al.*, 2006; Hoven *et al.*, 2004; Santini *et al.*, 2007). Along with a molybdenum cofactor, AioA contains iron-sulfur clusters which are thought to aid in the transfer of electrons from arsenite to the electron transport chain, likely via an oxidized c-type cytochrome (Figure 1B). AioA is a periplasmic protein, which is common among the other two types of arsenic metabolizing enzymes (ArxA, ArrA). Aio-encoding genes typically appear in genomic islands but have also been reported in plasmids, and thus may be attained through horizontal gene transfer (Andres *et al.*, 2013; Bertin *et al.*, 2012; Emmenegger *et al.*, 2001). Although *aioAB* are known to aid in detoxification, it is now clear that some organisms utilize these genes for energy production (Shi *et al.*, 2020).

It was suggested that a separate periplasmic oxidoreductase, AioE, acts as a novel electron transporter between the molybdenum cofactor of AioA and the electron transport chain in *Agrobacterium tumefaciens* GW4 (Wang *et al.*, 2017). Deletion of the *aioE* gene resulted in a decrease in NADH levels when GW4 was

grown with As(III). AioE played a direct role in As(III) resistance and detoxification, but it was determined that AioA is still the primary enzyme catalyzing the oxidation activity (Wang *et al.,* 2015). In the few organisms that AioE has been detected, it was found that this gene typically clustered in the *ars* operon and not the *aio* operon, suggesting a role in both detoxification and energy gain (Wang *et al.,* 2017).

AioXSR

To date, the most studied arsenite oxidase regulatory operon is *aioXSR*. Interestingly, the three component regulator genes *aioXSR* are only found in *Proteobacteria* and have opposite transcriptional orientation between *Alpha-* and *Betaproteobacteria* (Li *et al.,* 2013). In one of the first genetic studies into arsenite oxidation in *Agrobacterium tumefaciens* GW4, a transposon mutagenesis screen found a response regulator, AioR, critical to responding to arsenite, in which quorum sensing was also a regulator of arsenite oxidation (Kashyap *et al.,* 2006). Further studies have revealed that RpoN alternative sigma factor (σ^{54}) plays a role in *aio* regulation (Andres *et al.,* 2013; Li *et al.,* 2013). Several bacteria were shown to have these binding motifs directly upstream of the *aio* operon. Phosphate transport genes, oxidative stress genes, and chemotaxis genes have been reported to be upregulated in response to arsenite exposure (Andres *et al.,* 2013; Muller *et al.,* 2005; Weiss *et al.,* 2009). In *Agrobacterium,* the chemotactic response to arsenite was abolished in a Δ*aioR* mutant but recovered when complemented with this gene (Shi *et al.,* 2017). This discovery revealed how a motility protein, methyl-accepting chemotaxis protein (mcp), interacts with AioR to provide a chemotactic response to arsenite.

Furthermore, this chemotaxis defect also showed that arsenite oxidation is reduced compared to wild-type, suggesting a pivotal role in generating energy (Shi *et al.,* 2017). Lastly, the arsenite sensing protein AioX has been shown to be essential for upregulation of *aioBA* in *Agrobacterium tumefaciens* (Lui *et al.,* 2012). AioX is a membrane-associated periplasmic protein which uses conserved cysteine residues for interacting with arsenite and facilitating interaction with AioS to generate a signal transduction (Lui *et al.,* 2012). It remains unclear however, if all *aio* and *arx* containing bacteria require the XSR cluster to respond to arsenite.

$ArxB_2AB_1CD$

The most recently discovered arsenite oxidase, ArxA, has been found in chemoautotrophs, photoautotrophs, heterotrophs, and photoheterotrophs (Zargar *et al.,* 2010; Zargar *et al.,* 2012; Hernandez-Maldonado *et al.,* 2016; Durante-Rodriguez *et al.,* 2018). Interestingly, phylogenetic analyses show ArxA protein sequences cluster more closely to ArrA arsenate reductases rather than AioA arsenite oxidases (Zargar *et al.,* 2010; Oremland *et al.,* 2017). While also in the DMSO family of oxidoreductases, this molybdenum-containing subunit also has a predicted 4Fe-4S cluster binding domain. The *arx* operon features two ferrodoxin-like proteins adjacent to the oxidase ($ArxB_2$, $ArxB_1$) both also containing predicted 4Fe-4S cluster binding sites. It is believed this complex is anchored to the cytoplasmic membrane with a NrfD-like membrane subunit (ArxC). The last gene in the operon encodes for a TorD-like molybdoenzyme chaperone protein (ArxD) and is thought to be required for protein folding. It is believed that the arx oxidase is positioned in the cytoplasmic

membrane and interacting with the electron transport chain through the Q pool, however it is not known whether this enzyme complex clusters to distinct regions or structures in the membrane. Further, no crystals structures have been published, nor is the exact mechanism or electron transfer known. We distinguish photoarsenotrophy from Arx-type arsenotrophy because these two varying forms of arsenite oxidation have different implications on arsenic biogeochemical cycling since one relies on light and carbon while the other relies on nitrogen availability (Figure 1CD). Future investigations on how photosynthesis, and light in general, regulate the ArxA oxidase would help advance our understanding of the mechanism of this process.

ArxXSR

The *arx* operon typically features an adjacent regulatory operon coding for three proteins in the opposite direction. This regulatory operon has a putative ABC transporter substrate-binding protein (ArxX), which is structurally similar to the *E. coli* periplasmic binding protein (PhnD) and is believed to be involved in periplasmic shuttling of arsenite to the sensor histidine kinase (ArxS). It is not known if ArxX or AioX hold any phosphonate uptake capabilities like PhnD, but it is worth exploring due to the need for developing new phosphonate biosensors (Alicea *et al.*, 2011). One recent report characterized the binding properties of purified AioX, ArrA, and ArxX proteins using isothermal titration calorimetry and detected patterns in ligand specificity (Badilla *et al*., 2018). The experiments demonstrated AioX and ArxX could only bind arsenite, while ArrX could only bind arsenate and phosphate. Furthermore, ArxS has not been tested for its dependence on ArxX for sensing

arsenite, but one study showed an arsenic oxidation defect in a Δ*aioX* background (Lui *et al.,* 2012). The location of ArxX (membrane-associated versus periplasmic) is also of interest because it would shed light on its dependence by ArxS to respond to arsenite. Research on AioX and ArxX sensing of other arseno-compounds is lacking.

Last in the operon is a putative response regulator (ArxR), annotated as a sigma-54-dependent Fis family transcriptional regulator. ArxSR make up a putative two-component system, where ArxS is thought to activate ArxR via phosphate transfer after being activated by arsenite. While studies on arsenite oxidase (*aio*) regulation have been published, currently none have investigated the regulation of the Arx oxidase. In the photoarsenotroph *Ectothiorhodospira sp.* str. BSL-9, a disruption of the *aioR* gene abolished arsenite oxidation, suggesting that this gene is involved in regulation of photoarsenotrophy (Appendix A Figure 2). This is consistent with previous reports on regulation of arsenite oxidation by AioR (Kashyup *et al.,* 2006).

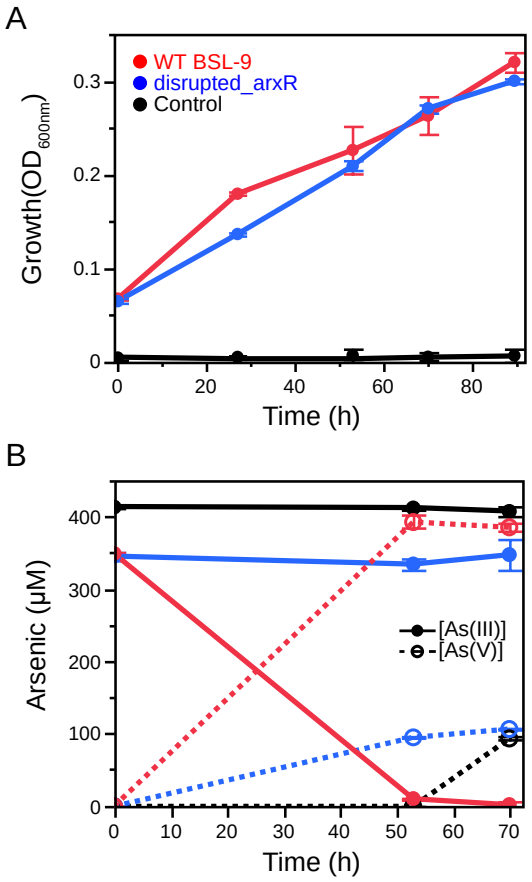


Figure 2. Disruption of the *arxR* gene in *Ectothiorhodospira sp.* str. BSL-9 abolishes arsenite oxidation. (A) Growth of wild-type BSL-9 (red), disrupted_*arxR* mutant (blue), and negative control (black), in BSM broth containing 10 mM acetate, is shown as optical density (600_{nm}) versus time (hours). (B) Arsenic speciation is shown for wild-type BSL-9 (red), disrupted_*arxR* mutant (blue), and negative control (black) as micromolar concentrations of arsenite (solid line) and arsenate (dotted line) versus time (hours). Cultures were grown in BSM broth containing 0.5 mM arsenite as the electron donor. Error bars are shown as one standard error from the mean (n=3). See Appendix A Figure 1 for PCR verification of *arxR* disruption.

Ars

The *ars* operon, typically three (*arsRBC*) or five (*arsRDABC*) genes that operate as a single transcriptional unit, provide arsenic tolerance, and are found in many different bacteria (Rosen 1999; Ben Fekih et al., 2018). This detoxification mechanism can be broken down into two key steps: reduction and transport. The detoxifying arsenate reductase ArsC, is encoded by the *arsC* gene, in the *ars* operon is a key enzyme for this mechanism. First, the cytoplasmic ArsC protein reduces arsenate to the more mobile form, arsenite. Next, a transport protein ArsB can remove the arsenite from the cell, or other enzymes can transform the arsenite into other arsenic compounds for transport. In some cases, energy is required for the transport step and so the ATPase ArsA is utilized (Rosen *et al.,* 1999; Zhou *et al.,* 2000; Meng *et al.,* 2004). The *ars* operon from *Escherichia coli* plasmid R773 features *arsD*, which encodes for the ArsD protein. ArsD is annotated as a metallochaperone and can use two cysteine pairs to bind three arsenite molecules, to facilitate transporting arsenic out of the cell via ArsAB. (Yang *et al.,* 2010; Shilpa *et al.,* 2017).

Some bacteria have evolved other forms of arsenic resistance by adding to their *ars* operon. The gene *arsM* and its protein product ArsM (As(III) S-adenosylmethyltransferase) provides the ability to methylate inorganic arsenic to methylated arsenicals (e.g., MAs(III)), which can then be pumped out of the cell (Qin *et al.,* 2006; Qin *et al.,* 2009; Chen *et al.,* 2014; Ajees and Rosen 2015). The membrane transporter ArsJ, has also recently been described as being involved in a novel form of arsenate resistance. A glycolytic gene *gapdh* and *arsJ* appear together

in many *ars* operons, and expression in bacteria sensitive to arsenate provides resistance (Chen *et al.*, 2016). It is thought that this glycolysis enzyme, glyceraldehyde-3-phosphate dehydrogenase, forms 1-arseno-3-phosphoglycerate, which can then be pumped out of the cell via ArsJ (Chen *et al.*, 2016). One model has proposed that 1-arseno-3-phosphoglycerate rapidly degrades to arsenic and 3-phosphoglycerate in the extracellular space. However, it remains unclear what that fate of this arsenophosphoglycerate compound is once outside of the cell. These various arsenic resistance mechanisms are not limited to bacteria; many eukaryotic cells contain orthologues of *ars* genes. Some bacteria have been found to contain an *arsH* gene, which encodes for an organoarsenical oxidase protein, ArsH. Cells expressing ArsH are resistant to trivalent organoarsenicals (monosodium methylarsenate, roxarsone, and phenylarsenite), while purified ArsH can oxidize these compounds in vitro (Chen *et al.*, 2015; Yang and Rosen, 2016). Another recently described arsenic resistance gene is the C-As lyase encoding gene *arsI*, which cleaves MAs(III) to produce As(III) and a methyl group (Yang and Rosen, 2016).

Ars Regulation

Regulation of the *ars* operon is typically controlled by the ArsR protein, encoded by the *arsR* gene, which acts as a repressor in the presence of arsenite (Wu and Rosen, 1991). ArsR features both an arsenic sensing domain, as well as a regulatory domain. As mentioned earlier, in *Shewanella sp*. str. ANA-3, both *arr* and *ars* are regulated by the ArsR family protein, ArsR2 (Murphy and Saltikov, 2009).

Other bacteria like *Acidithiobacillus ferrooxidans* may use an atypical ArsR-like regulator, for regulating *ars* genes that have diverged from the normal operon orientation (Butcher and Rawlings, 2002). In this case, *A. ferrooxidans* features *arsCR* and *arsBH* genes facing in different directions.

Section 2.06 Light-dark metal redox cycling

Light-dark metal cycling occurs when dissolved metal concentrations correlate to the light-dark phases of a day. Few studies have investigated light-dark metal cycling, and little is known about the biological component of these processes. Fuller and Davis (1989) were the first to report what they called "diel variation in arsenic concentrations" by measuring arsenate concentrations over the course of two days at Whitewood Creek (SD, USA). Their study demonstrated arsenate concentrations increase from 6:00 A.M. until about 6:00 P.M. and decreases from 6:00 P.M. until 6:00 A.M. So-called diel variations in total Mn, Zn, Ni, and As concentrations, were detected in freshwater streams Prickly Pear Creek and High Ore Creek (MT, USA; Nimick *et al.,* 2003). Similar results were reported in several other studies, but never attributed to bacterial processes such as arsenite oxidation or arsenate reduction (Nimick *et al.,* 2005; Sarmiento *et al.,* 2007; Dicataldo *et al.,* 2011). Another study reported phototrophic biofilm activity as a driver of pH and O_2 dynamics, and subsequently diurnal Cd cycling (Beck *et al.,* 2009). These early studies concluded that photosynthesis-induced changes in pH were the driver for arsenic variation, and never suspected bacteria or photosynthesis-coupled arsenic transformations. Today, we know bacteria harboring certain genes can directly

oxidize or reduce arsenic and can be the primary drivers of the biological component of arsenic cycling. In microcosms prepared with water from Big Soda Lake (NV, USA), researchers detected light-dark cycling of arsenic that was likely biologically driven (Hernandez-Maldonado, 2017). The experiments using Big Soda Lake microcosms found arsenite oxidation occurred during a 15-day light phase, then arsenate reduction occurred during a 12-day dark phase, and finally arsenite oxidation occurred again during a 7-day light phase. The light-dark arsenic cycling was not observed in microcosms that were heat-inactivated prior to the start of the experiment. This study was the first to link the biological component of light-dark arsenic cycling to phototrophic bacteria. It remains to be determined whether this process occurs in non-extreme environments like freshwater sediments, an analog to the rice paddy environment.

Section 2.07 Conclusion and future directions

While the last century of investigating arsenic transformations in nature has identified hundreds of different participating microbes and the four key gene clusters (*aio, arx, arr,* and *ars*) responsible for these transformations, the mechanism and ecology of photoarsenotrophy (*arx*) in non-extreme environments has just begun. Early studies detected light-dark variations in arsenic concentrations, but no research since has identified bacteria or genes possibly driving this activity. By researching the mechanism and ecology of photoarsenotrophy, we can begin to unravel the complexity of the arsenic biogeochemical cycle. Research on the mechanism of Arx-type arsenotrophy has been stalled due to slow and laborious genetic manipulations of

extremophilic bacteria, and a more robust genetic system would help move it forward. Moreover, investigations on these enzymes and regulatory pathways may uncover novel methods for attenuating and manipulating agricultural practices to favor reduced arsenic exposure and ingestion by humans. Since arsenic-contaminated groundwater is often used for agricultural purposes, staple crops like rice can contain high levels of arsenic.

Section 2.08 References

Abedi, T., and Mojiri, A., (2020). Arsenic Uptake and Accumulation Mechanisms in Rice Species. Plants 9, no. 2: 129.

Ajees, A. A., & Rosen, B. P. (2016). As(III) S-adenosylmethionine methyltransferases and other arsenic binding proteins. HHS Public Access, 32(7), 570–576.

Alicea, I., Marvin, J. S., Miklos, A. E., Ellington, A. D., Looger, L. L., & Schreiter, E. R. (2011). Structure of the Escherichia coli Phosphonate Binding Protein PhnD and Rationally Optimized Phosphonate Biosensors. Journal of Molecular Biology, 414(3), 356–369.

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

Andres, J. *et al.*, (2013) Life in an arsenic-containing gold mine: genome and physiology of the autotrophic arsenite-oxidizing bacterium Rhizobium sp. NT-26. Genome. Biol. Evol. 5, 934–953.

Andres, J., & Bertin, P. N. (2016). The microbial genomics of arsenic. *FEMS Microbiology Reviews*, *40*(2), 299–322.

Badilla, C., Osborne, T. H., Cole, A., Watson, C., Djordjevic, S., & Santini, J. M. (2018). A new family of periplasmic-binding proteins that sense arsenic oxyanions. Scientific reports, 8(1), 1-12.

Beck, A.J., Janssen, F., Polerecky, L., Herlory, O., and De Beer, D., (2009). Phototrophic biofilm activity and the dynamics of diurnal Cd cycling a freshwater system. Environ. Sci. Technol. 43, 7245–7251.

Ben Fekih, I., Zhang, C., Li, Y. P., Zhao, Y., Alwathnani, H. A., Saquib, Q., ... & Cervantes, C. (2018). Distribution of arsenic resistance genes in prokaryotes. Front. in Microbio, 2473.

Bertin PN, *et al.,* (2012). Microbial arsenic response and metabolism in the genomics era. In: Santini JM, Ward SA, editors. The metabolism of arsenite. London: CRC Press. 99–114.

Boye, K. (2017). Relevance of Reactive Fe:S Ratios for Sulfur Impacts on Arsenic Uptake by Rice. Soils. 1(1).

Burgess, W. G., Hoque, M. A., Michael, H. A., Voss, C. I., Breit, G. N., & Ahmed, K. M. (2010). Aquifer system to contamination by arsenic. Nature Geoscience, 3(2), 83–87.

Busenlehner, L. S., M. A. Pennella, and D. P. Giedroc. (2003). The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. FEMS Microbiol. Rev. 27:131–143

Butcher, B. G., & Rawlings, D. E. (2002). The divergent chromosomal ars operon of Acidithiobacillus ferrooxidans is regulated by an atypical ArsR protein. 3983–3992.

Chen, J., Bhattacharjee, H., & Rosen, B. P. (2015). ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium

methylarsenate and the poultry growth promoter roxarsone, 96(April), 1042–1052.

Chen, J., Yoshinaga, M., Garbinski, L. D., & Rosen, B. P. (2017). Synergistic interaction of glyceraldehydes-3-phosphate dehydrogenase and ArsJ, a novel organoarsenical efflux permease, confers arsenate resistance. HHS Public Access, 100(6), 945–953.

Cleiss-Arnold, J. *et al.,* Temporal transcriptomic response during arsenic stress in Herminiimonas arsenicoxydans. BMC. Genomics. 11, 709 (2010).

Coryell, M., McAlpine, M., Pinkham, N. V., McDermott, T. R., & Walk, S. T. (2018). The gut microbiome is required for full protection against acute arsenic toxicity in mouse models. Nature communications, 9(1), 1-9.

Croal, L.R., Gralnick, J.A., Malasarn, D., Newman, D.K. (2004) The Genetics of Geochemistry. Annu. Rev. Genet., 38 (1): 175-202.

Das, S., Chou, M., Jean, J., Liu, C., & Yang, H. (2016). Science of the Total Environment Water management impacts on arsenic behavior and rhizosphere bacterial communities and activities in a rice agro-ecosystem. Science of the Total Environment, The, 542, 642–652.

Das, S., Jean, J., Chou, M., Rathod, J., & Liu, C. (2016). Arsenite-oxidizing bacteria exhibiting plant growth promoting traits isolated from the rhizosphere of Oryza sativa: Implications for mitigation of arsenic contamination in paddies. J Haz Mat, 302, 10–18.

Dicataldo, G. et al., (2011) Diel variation of selenium and arsenic in a wetland of the Great Salt Lake, Utah. Appl. Geochemistry 26, 28–36.

Emmenegger, L., Sigg, L., & Sulzberger, B. (2001). Light-induced redox cycling of iron in circumneutral lakes, 46(1), 49–61.

Fendorf, S., Michael, H. A., & Geen, A. Van. (2010). Spatial and Temporal Variations of Groundwater Arsenic in South and Southeast Asia, (May), 1123–1128.

Fuller, C. C., & Davis, J. A. (1989). and Photosynthetic Cycles in Natural Waters. Nature, 340(JULY), 1987–1989.

Garcia-Dominguez, E., Mumford, A., Rhine, E. D., Paschal, A., & Young, L. Y. (2008). Novel autotrophic arsenite-oxidizing bacteria isolated from soil and sediments. FEMS Microbiology Ecology, 66(2), 401–410.

Glasser, N. R., Oyala, P. H., Osborne, T. H., Santini, J. M., & Newman, D. K. (2018) Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations. Proc Natl Acad Sci USA 115(37): E8614–E8623.

Green, H. H., Rep. Dir. Vet. Res. S. Afr., 5–6, 593 (1918)

Harvey, C. F., Swartz, C. H., Bohran, A., Badruzzaman, M., Niedan, V., Brabander, D., … Ashfaque, K. N. (2005). Groundwater arsenic contamination on the Ganges Delta: biogeochemistry, hydrology, human perturbations, and human suffering on a large scale, 337, 285–296.

Hernandez-Maldonado, J., Sanchez-Sedillo, B., Stoneburner, B., Boren, A., Miller, L., McCann, S., … Saltikov, C. (2016). The genetic basis of anoxygenic photosynthetic arsenite oxidation. Environmental Microbiology, 0, 1–29.

Hernandez-Maldonado, J. (2017). Ecology and molecular genetics of anoxygenic photosynthetic arsenite oxidation by *arxA*. Doctoral thesis. UC Santa Cruz. ProQuest ID: HernandezMaldonado_ucsc_0036E_11312. Merritt ID: ark:/13030/m5wq503x.

Hoeft SE, Kulp TR, Han S, *et al.*, Coupled arsenotrophy in a photosynthetic hot spring biofilm from Mono Lake, California. Appl Environ Microbiol 2010; 76: 4633 – 4639

Hoeft, S. E., Blum, J. S., Stolz, J. F., Tabita, F. R., Witte, B., King, G. M., … Oremland, R. S. (2007). Alkalilimnicola ehrlichii sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. International Journal of Systematic and Evolutionary Microbiology, 57(3), 504–512.

Hoven, R. N., & Santini, J. M. (2004). Arsenite oxidation by the heterotroph Hydrogenophaga sp. str. NT-14: the arsenite oxidase and its physiological electron acceptor. Biochim Biophys Acta, 1656(3), 148–155.

Kashyap, D. R., Botero, L. M., Franck, W. L., Hassett, D. J. & McDermott, T. R. Complex regulation of arsenite oxidation in Agrobacterium tumefaciens. J. Bacteriol. 188, 1081–1088 (2006).

Kulp TR , Hoeft SE, Asao M, *et al.,* Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. Science 2008; 321: 967 – 970.

Li, G., Sun, G., Williams, P. N., Nunes, L., & Zhu, Y. (2011). Inorganic arsenic in Chinese food and its cancer risk. Environment International, 37(7), 1219–1225.

Li, H., Li, M., Huang, Y., Rensing, C. & Wang, G. In silico analysis of bacterial arsenic islands reveals remarkable synteny and functional relatedness between arsenate and phosphate. Front. Microbiol. 4, 347 (2013).

Lin, T., Wei, C., Huang, C., Chang, C., Hsu, F., & Liao, V. H. (2016). Both Phosphorus Fertilizers and Indigenous Bacteria Enhance Arsenic Release into Groundwater in Arsenic-Contaminated Aquifers.

Liu, G. *et al.,* A periplasmic arsenite-binding protein involved in regulating arsenite oxidation. Environ. Microbiol. 14, 1624–1634 (2012).

Mailloux, B. J., Trembath-Reichert, E., Cheung, J., Watson, M., Stute, M., Freyer, G. A., … van Geen, A. (2013). Advection of surface-derived organic carbon fuels microbial reduction in Bangladesh groundwater. Proc Natl Acad Sci, 110(14), 5331–5335.

Mawia, A. M., Hui, S., Zhou, L., Li, H., Tabassum, J., Lai, C., … Hu, P. (2021). Inorganic arsenic toxicity and alleviation strategies in rice. J Haz Mat, 408, 124751.

McDermott, T. R., Stolz, J. F., & Oremland, R. S. (2020). Arsenic and the gastrointestinal tract microbiome. Environmental Microbiology Reports, 12(2), 136-159.

Meng YL, Liu Z, Rosen BP. (2004) As(III) and Sb(III) uptake by GlpF and efflux by ArsB in Escherichia coli. J Biol Chem. 279:18334–18341.

Moens, M., Branco, R. & Morais, P.V. (2020) Arsenic accumulation by a rhizosphere bacterial strain *Ochrobactrum tritici* reduces rice plant arsenic levels. World J Microbiol Biotechnol 36, 23.

Muller, D. *et al.,* (2005) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. PloS. Genet. 3, e53.

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.

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.

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.

Nimick, D. A. et al., (2003) Diel cycles in dissolved metal concentrations in streams: Occurrence and possible causes. Water Resour. Res. 39.

Nimick, D. A., Cleasby, T. E. & McCleskey, R. B. (2005) Seasonality of diel cycles of dissolved trace-metal concentrations in a Rocky Mountain stream. Environ. Geol. 47, 603–614.

Oremland RS, Hoeft SE, Bano N, *et al.,* Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, strain MLHE-1. Appl Environ Microbiol 2002; 68: 4795 – 4802.

Oremland, R. S., & Stolz, J. F. (2003). The Ecology of Arsenic. Science, 300(9), 939–944.

Oremland, R. S., Saltikov, C. W., Stolz, J. F., & Hollibaugh, J. T. (2017). Autotrophic microbial arsenotrophy in arsenic-rich soda lakes. FEMS Microbiology Letters, 364(15), 1–11.

Qin, J., Lehr, C. R., Yuan, C., Le, X. C., Mcdermott, T. R., & Rosen, B. P. (2009). Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga, 106(13), 5213–5217.

Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S., & Rensing, C. (2006). Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase, 103(7).

Reid, M. C., Maillard, J., Bagnoud, A., Falquet, L., Le Vo, P., & Bernier-Latmani, R. (2017). Arsenic methylation dynamics in a rice paddy soil anaerobic enrichment culture. Environmental science & technology, 51(18), 10546-10554.

Rhine ED, Phelps CD, Young LY. Anaerobic arsenite oxidation by novel denitrifying isolates. Environ Microbiol 2006; 8: 899 – 908.

Rhine, E. D., Onesios, K. M., Serfes, M. E., Reinfelder, J. R., & Young, L. Y. (2008). Arsenic transformation and mobilization from minerals by the arsenite oxidizing strain WAO. Environmental Science and Technology, 42(5), 1423–1429.

Rodríguez-Freire, L., Sun, W., Sierra-Alvarez, R. et al. (2012) Flexible bacterial strains that oxidize arsenite in anoxic or aerobic conditions and utilize hydrogen or acetate as alternative electron donors. Biodegradation 23, 133–143.

Rosen, B. P. (1999). Families of arsenic transporters. *Trends in microbiology*, *7*(5), 207-212.

Yang, H. C., & Rosen, B. P. (2016). New mechanisms of bacterial arsenic resistance. Biomedical journal, 39(1), 5-13.

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, 69(5), 2800–2809.

Saltikov, C. W., & Newman, D. K. (2003). Genetic identification of a respiratory arsenate reductase. Proc Natl Acad Sci USA 100(19).

Santini, J. M., Sly, L. I., Schnagl, R. D., & Macy, J. M. (2000). A New Chemolithoautotrophic Arsenite-Oxidizing Bacterium Isolated from a Gold

Mine: Phylogenetic, Physiological , and Preliminary Biochemical Studies, 66(1), 92–97.

Santini JM, *et al.,* (2002) New arsenite-oxidizing bacteria isolated from Australian gold-mining environments - phylogenetic relationships. Geomicrobiol J 19: 67-76.

Santini, J. M., Kappler, U., Ward, S. A., Honeychurch, M. J., vanden Hoven, R. N., & Bernhardt, P. V. (2007). The NT-26 cytochrome c552 and its role in arsenite oxidation. *Biochimica et biophysica acta*, *1767*(2), 189–196.

Sarmiento, A. M., Oliveira, V., Gómez-Ariza, J. L., Nieto, J. M. & Sánchez-Rodas, D. (2007) Diel cycles of arsenic speciation due to photooxidation in acid mine drainage from the Iberian Pyrite Belt (Sw Spain). Chemosphere 66, 677–683.

Satyapal, G.K., Kumar, N. (2016). Potential Role of Arsenic Resistant Bacteria in Bioremediation : Current Status and Future Prospects Microbial & Biochemical Technology Potential Role of Arsenic Resistant Bacteria in Bioremediation: Current, (January), 1–4.

Shi, W., J. Wu, and B. P. Rosen. (1994). Identification of a putative metal binding site in a new family of metalloregulatory proteins. J. Biol. Chem. 269:19826–19829.

Shi, K., Fan, X., Qiao, Z., Han, Y., Mcdermott, T. R., Wang, Q., & Wang, G. (2017). Arsenite oxidation regulator AioR regulates bacterial chemotaxis towards arsenite in Agrobacterium tumefaciens GW4. Nature Publishing Group, (January), 1–11.

Shi, K., Wang, Q., & Wang, G. (2020). Microbial oxidation of arsenite: regulation, chemotaxis, phosphate metabolism and energy generation. Frontiers in Microbiology, 11, 2235.

Shilpa, T., Rosen, B., Abdul Ajees. A. (2017). Structural studies of the ArsD arsenic metallochaperone using molecular dynamics. Journal of Computational Methods in Sciences and Engineering, vol. 17, no. 2, pp. 227-233, 2017.

Turner, K., Joel, S., Feliciano, J., Feltus, A., Pasini, P., Wynn, D., Dau, P., Dikici, E., Deo, S. K., & Daunert, S. (2017). Transcriptional regulatory proteins as biosensing tools. Chemical communications (Cambridge, England), 53(51), 6820–6823.

Verma, S., Verma, P. K., Meher, A. K., Dwivedi, S., Bansiwal, A. K., Pande, V., … Chakrabarty, D. (2016). A novel arsenic methyltransferase gene of Westerdykella aurantiaca isolated from arsenic. Metallomics, 8, 344–353.

Wang, Q., Warelow, T. P., Kang, Y. S., Romano, C., Osborne, T. H., Lehr, C. R., Bothner, B., McDermott, T. R., Santini, J. M., & Wang, G. (2015). Arsenite oxidase also functions as an antimonite oxidase. Applied and environmental microbiology, 81(6), 1959–1965.

Wang, Q., Han, Y., Shi, K., Fan, X., Wang, L., Li, M., & Wang, G. (2017). An Oxidoreductase AioE is Responsible for Bacterial Arsenite Oxidation and Resistance. Nature Publishing Group, (December 2016), 1–10.

Weiss, S. *et al.*,(2009) Enhanced structural and functional genome elucidation of the arsenite-oxidizing strain Herminiimonas arsenicoxydans by proteomics data. Biochimie. 91, 192–203.

Wu, J., and Rosen B.P., (1991). The ArsR protein is a trans-acting regulatory protein. Molecular microbiology. 5(6), 1331–1336.

Yang, J., Rawat, S., Stemmler, T.L., Rosen, B. (2010). Arsenic Binding and Transfer by the ArsD As(III) Metallochaperone. Biochemistry. 2010 May 4; 49(17): 3658–3666.

Zargar K, Hoeft S, Oremland R, *et al.,* (2010) Genetic identification of a novel arsenite oxidase, arxA, in the haloalkaliphilic arsenite oxidizing bacterium Alkalilimnicola ehrlichii strain MLHE-1. J Bacteriol 192: 3755 – 3762.

Zargar K, Conrad A, Bernick DL, *et al.*, (2012) ArxA, a new clade of arsenite oxidase within the DMSO reductase family of 3 molybdenum oxidoreductases. Environ Microbiol; 14: 1635 – 1645.

Zhao, F., Mcgrath, S. P., & Meharg, A. A. (2010). Arsenic as a Food Chain Contaminant : Mechanisms of Plant Uptake and Metabolism and Mitigation Strategies.

Zhou, T., Radaev, S., Rosen, B. P., & Gatti, D. L. (2000). Structure of the ArsA ATPase : the catalytic subunit of a heavy metal resistance pump. 19(17).

Chapter 3. Genomic characterization of a freshwater photoarsenotroph, *Cereibacter azotoformans* str. ORIO, from sediments capable of light-dark arsenic redox cycling

Section 3.01 Abstract

Photoarsenotrophic microbes have the potential for remediating arsenic contamination in aquatic environments by their ability to convert arsenite to arsenate, a form with greater affinity for soil and sedimentary minerals. Here we present the complete genome sequence of *Cereibacter azotoformans* strain ORIO, a freshwater photosynthetic arsenite-oxidizing bacterium isolated from Owens River, CA, USA. The genome has a length of 4.8 Mb with 2 chromosomes, 6 plasmids, and a 68 % G+C content. ORIO is closely related to other well-studied *Cereibacter* strains (formerly *Rhodobacter*). The ORIO genome contains *arxB2AB1CD* (encoding an arsenite oxidase*), arxXSR (arx* regulators*),* and several *ars* resistance genes all co-localized on a 136 kb plasmid, named pORIO3. Taxonomic analysis of pORIO3 determined it is related to previously sequenced *Cereibacter sphaeroides* plasmid sequences. Phylogenetic analysis of *arx* genes demonstrate photoarsenotrophy is likely to occur within the *Alphaproteobacteria* clade. ORIO is a mixotroph, oxidizes arsenite (As(III)) to arsenate (As(V)) photoheterotrophically, and expresses *arxA* in cultures grown with As(III). Ecophysiology studies with Owens River sediment demonstrated the occurrence of light-dark cycling of arsenite to arsenate. *arxA* and *arrA* (arsenate respiratory reductase) genes were detected in the light-dark cycled

sediment metagenomes suggesting syntrophic interactions among arsenotrophs. This work establishes *Cereibacter azotoformans* str. ORIO as a new model organism for studying the effects of photoarsenotrophy on arsenic biogeochemical cycling.

Section 3.02 Introduction

Over 230 million people worldwide are likely consuming arsenic polluted water that exceeds the recommended limit of 10 μg/L (Shaji *et al.,* 2021). Arsenic in water supplies is usually caused by naturally occurring processes, with arsenite (As(III)) and arsenate (As(V)) as the predominant inorganic species found in the environment. Microbes are the major drivers of cycling arsenic between the As(III) and As(V) species. Consequently, microbes that reduce and oxidize arsenic have strong influences on the fate of arsenic (Oremland and Stolz, 2003). Bacteria exposed to arsenic have evolved distinct forms of arsenic metabolism and resistance, which contribute to arsenic biogeochemical cycling. However, contributions from a recently described form of arsenic metabolism, anoxygenic photosynthetic As(III)-oxidation (Arx), termed photoarsenotrophy, are largely uncharacterized. Our knowledge of how bacteria transform arsenic in nature can better inform our ability to mitigate arsenic contamination.

Photoarsenotrophy is a process by which bacteria oxidize As(III) to As(V) while performing anoxygenic photosynthesis (Kulp *et al.,* 2008). This process has been shown to rely on the Arx pathway, and to date has been studied in eleven cultivated strains (Oremland *et al.,* 2002; Hoeft *et al.,* 2007; Hamamura *et al.,* 2014; Hernandez-

Maldonado *et al.,* 2016; Hoeft-McCann *et al.,* 2017; Wu *et al.,* 2017; Ospino *et al.,* 2019; Durante-Rodríguez *et al.,* 2019). The distribution of Arx genes is concentrated in the Proteobacteria clade and is dominated by *Gammaproteobacteria*. Of the eleven studied Arx-based arsenotrophs, eight are *Gammaproteobacteria* (*Ectothiorhodospira sp.* strains MLW-1, PHS-1, BSL-9, *Ect. shaposhnikovii* strains DSM 243 and DSM 2001, *Halorhodospira halophila* SL-1, *Alkalilimnicola ehrlichii* str. MLHE-1, and *Halomonas sp.* ANAO440) from extreme environments (high pH, salinity, temperature, and arsenic), two are *Betaproteobacteria* (*Azoarcus sp.* CIB, and *Sterolibacteriaceae* strain M52) from freshwater sources, and one is a thermophilic *Firmicutes* (*Desulfotomaculum* strain TC-1) from a sulfidic hot spring. The study of arsenite-oxidation coupled to photosynthesis has only been investigated in two extremophilic genera, while the other are chemosynthetic Arx-based arsenotrophs (not coupled to photosynthesis). This difference in requiring photosynthesis versus a terminal electron acceptor (e.g., nitrate for *Alkalilimnicola ehrlichii* str. MLHE-1 (Oremland *et al.,* 2002; Zargar *et al.,* 2010) means that Arx-based arsenotrophy can either be controlled by light-dark cycles or availability of a terminal electron acceptor. These two variations of the Arx pathway likely have different effects on the arsenic biogeochemical cycle because they have different requirements for arsenite oxidation. One is dependent on light and carbon (either CO_2 or dissolved organic carbon (DOC)), while the other is dependent on nitrate and carbon (DOC).

A major impediment for advancing the study of photoarsenotrophs has been that traditional genetic manipulation techniques are difficult due to the high pH, temperature, and salinity needed to mimic the extreme environments from which previous photoarsenotrophs were isolated. In this study, *Cereibacter azotoformans* str. ORIO, a photosynthetic As(III)-oxidizing *Alphaproteobacterium*, was isolated from arsenic-rich freshwater river sediments of Owens River (CA, USA). We provide the complete genome sequence and physiological attributes of ORIO. Whole genome sequencing revealed ORIO contains the *arx* gene cluster, 2 chromosomes, 6 plasmids, and has a G+C content of 68% consistent with *Cereibacter/Rhodobacter* species. Genome and phylogenetic analyses suggest ORIO is a member of the *Cereibacter* genus and contains an *arx* gene cluster with similarity to *Ectothiorhodospira* sp. PHS-1. ORIO is resistant to up to 2 mM arsenite and 4 mM arsenate, and transcription of *arxA* is induced when grown with arsenite. We used the ORIO ArxA protein sequence to search for homologs in metagenomes in the JGI-IMG database. Metagenome mining confirms the presence of *arx* genes in metagenomes of samples from all seven continents and common types of environments. Further ecophysiology studies with Owens River sediments were done to determine impacts of light-dark cycling on arsenic geochemistry. The results revealed the potential for an arsenic redox cycle among a syntrophic microbial community composed of arsenite-oxidizing and arsenate-reducing arsenotrophs.

Section 3.03 Results

Sample Site and Environment

The goal of this work was to determine if photoarsenotrophs could be cultured from arsenic-rich freshwater environments and isolate a representative mesophilic phototroph for future work on the biology of photoarsenotrophy. Owens River sediments were selected because arsenic levels exceed USEPA drinking water standards and are of public health concerns to the people of Southern California. This waters system is a major source of Los Angeles drinking water via the LA Aqueduct (Wikie and Hering, 1998; Kneebone and Hering, 2000). Owens River sediment was collected at the edge of the river (37.667000 N 118.750167 W) and used to establish anaerobic enrichment cultures with arsenite added as an electron donor. Following multiple rounds of enrichment culturing with arsenite-spiked freshwater medium at 30°C illuminated with infrared light-emitting diodes (850 nm IR-LEDs), a photosynthetic arsenite-oxidizing bacterium was isolated. Various pigmented (dark green and purple) colonies were streaked purified multiple times on agar-solidified medium and grown in an anaerobic chamber (Coy Labs, 95% N_2:5% H_2). A purple-colored strain was selected and named: ORIO, **O**wens **R**iver **I**solate **O**xidizer. This strain was selected for a variety of reasons described below.

Genome Sequencing

While we isolated several strains with green and purple colonies, one was further chosen based on initial PCR detection of *arxA* and a preliminary 16S rRNA gene identification as a *Rhodobacter* species. A draft genome was completed using Illumina MiSeq and further confirmed the presence of an *arx* gene cluster. We used the Oxford Nanopore Minion sequencer to generate longer sequencing reads (7.96G,

1,709,248 reads, mean quality score 13.26). This approach of combining long reads from the Minion and high accuracy short reads (MiSeq PE150: 6,881,173 reads, Q20 of 95.76%) was successful in assembling a complete genome of ORIO (sequencing quality information can be found in ORIO_Genome_SuppInfo.html). Sequencing yielded over 900x coverage from long reads and 428x coverage from short reads. The genome was determined to be 4.8 Mb in size, containing 2 chromosomes, 6 plasmids, with an average GC content of 68%. Genome annotation determined there are 4,711 coding sequences, 56 tRNA, 12 rRNA, and 3 ncRNA sequences.

Genome and Plasmid Analysis

Distance analysis with the ORIO 16S *rRNA* gene shows a close relation to *Cereibacter sphaeroides* and *C. azotoformans* (Figure 1A). The *arx* gene cluster (*arxXSR*, *arx*B_1*AB*$_2$*CD*) is located on a single 136 kb plasmid, pORIO3 and appears within an "island" of other arsenic detoxification encoding genes. Phylogenetic distance analysis of the ORIO ArxA protein sequence with other previously studied Arx-type arsenotrophs shows closest relation to *Ectothiorhodospira sp.* str. PHS-1 (Figure 1B). This arsenic island is flanked by transposase encoding genes. The top five pORIO3 nucleotide BLAST search results were aligned to pORIO3 (Figure 1C). These five closely related plasmid sequences are assigned to the plasmid taxonomic group PTU-Rho15, and have a host range of III meaning they are broad host range plasmids (Redondo-Salvo *et al.*, 2020). The *arx* gene cluster is flanked by several arsenic resistance gene clusters (*arsR*, *arsA*, *arsC*, *acr3*, and *arsP*), and phosphate stress pathway genes (*pstSCABC* and *phoUB*). We detected several sigma-54 (RpoN)

binding motifs in a 190 base pair region between *arxR* and *arxB2*. The genome does not contain any genes encoding for the Aio-type arsenite oxidases or Arr-type arsenate reductases. The genome analysis suggests ORIO is well adapted to living in an arsenic rich environment. The proximity to transposases and analysis of pORIO3 further suggests arsenic metabolic genetic traits were likely acquired by horizontal gene transfer.

Physiology

We tested a variety of carbon substrates and conditions to better understand the physiology and metabolism of *C. azotoformans* str. ORIO. We observed that ORIO is mixotrophic and mesophilic, with the most optimal growth occurring at 30°C. ORIO grew aerobically in the dark using succinate, acetate, glycerol, glucose, lactate, and pyruvate as electron donors. ORIO appeared to ferment sucrose without forming gas. ORIO also grew anaerobically with infrared light (850 nm) using organic carbon sources such as succinate (preferred), acetate, citrate, glycerol, glucose, or yeast extract. The ORIO genome contains genes for carbon fixation (RuBisCO), which suggests ORIO could also be autotrophic, yet under our conditions tested required an organic carbon source. ORIO also has the genetic capacity for denitrification (*napA*, *nirK*, *norBC*, *nosZ*), nitrogen fixation (*nif*), and sulfur metabolism (*sox*). Antibiotic sensitivity has been demonstrated with 10 μg/mL kanamycin and 1 μg/mL chloramphenicol. Antibiotic resistance genes for erythromycin (*macA*) and daptomycin (*pgsA* and *gpdP*) were detected in the genome. Under aerobic conditions ORIO demonstrated resistance to high levels of arsenic (>1 mM arsenite and

arsenate), but during anaerobic growth, inhibition occurred at concentrations greater than 0.5 mM arsenite (Figure 3). In other Arx-type arsenotrophs, inhibition of growth by arsenite can occur at concentrations ranging from 0.5 mM (Durante-Rodriguez *et al.,* 2019) to 5 mM (Heoft-McCann *et al.,* 2017). Arsenite oxidation and transcription of *arxA* was detected in cultures grown phototrophically with 1 mM As(III) for 2 days (Figure 4).

Light-dark arsenic cycling in Owens River sediment

We next addressed the environmental context for photoarsentrophy and its potential role in the diurnal biogeochemical cycling of arsenite to arsenate. We hypothesized that photoarsenotrophy depends on the availability of arsenite as an electron donor, which is produced by dissimilatory arsenate reducing bacteria. Microcosms with Owens River sediments were established with Owens River water and arsenite spiked in at environmentally relevant concentrations. Arsenic speciation was measured over the course of 6 days and found to cycle based on light-dark phases. In the microcosms that were not heat-killed, arsenite was oxidized to arsenate during the light phases, while arsenate was reduced to arsenite during the dark phases (Figure 6A). Each cycle exhibited the same pattern of light-dependent arsenite oxidation followed by arsenate reduction in the dark. The arsenic speciation in heat-killed controls remained relatively constant over the course of the experiment indicating that light-dark cycling of arsenic was biologically mediated (Figure 6B).

Metagenomic analysis of Owens River microcosms

Since our microcosm assay suggested a biological component to the light-dark arsenic cycling, we hypothesized we would detect relevant arsenic metabolizing genes such as *arxA* and *arrA.* We conducted a metagenomic analysis of the light-dark cycled microcosms and searched for taxa and genes relevant to arsenic biogeochemical cycling. Nanopore sequencing generated 5.96G of data, 1,533,746 reads, an N50 of 9,314 bp, and a mean quality score of 12.50. Metagenome assembled contigs were binned by taxonomy (genus level) into one of 12 bins. Moreover, we detected over 3000 taxonomic hits among all the assembled contigs. Bacteria dominated the metagenome, with 88% of the sequences belonging to this clade. The *Proteobacteria* clade made up 66% of the bacterial hits, and 58% of the entire metagenome. In the *Proteobacteria* clade, *Gammaproteobacteria* (35 %) were the most abundant, followed by *Betaproteobacteria* (29 %), and 15% for *Alphaproteobacteria* and *Delta*/*Epsilonproteobacteria*, respectively. We detected one *arxA*, two *arrA*, and two *aioA* homologs. The *arxA* gene was detected in a *Gammaproteobacteria* belonging to the *Allochromatium* genus, while the *arrA* gene was detected in two *Deltaproteobacteria* belonging to the *Desulfatitalea* and *Desulfococcus* genera (Figure 7C). We also detected *arrA* fragments in a *Betaproteobacteria* belonging to the *Rubrivivax* genus, however we could not determine a complete *arrA* sequence through polishing attempts and excluded it from phylogenetic distance analysis. Lastly, we detected *aioA* homologs in the *Alphaproteobacteria Rhodopseudomonas* and the *Deltaproteobacteria Desulfococcus*. Interestingly, *Desulfococcus* had an *aio* gene cluster near the *arr* gene

cluster. These results are the first to provide a genetic link to the observed variations in arsenic concentrations in freshwater conditions. Previous studies detected so-called "diurnal arsenic cycling", where arsenate concentrations increased during the day, and decreased during the night, but never have implicated arsenic-metabolizing bacteria as the drivers of the process (Fuller and Davis, 1989).

Metagenome Mining and *arxA* Phylogenetic Analysis

No recent studies have investigated the occurrence of *arx* genes in metagenomes in the JGI-IMG database. At the time of analysis, JGI-IMG had over 35,000 metagenome assemblies in the database. We felt this was a large enough dataset to capture the occurrence of *arx* in potentially novel environments. The ORIO ArxA sequence was used to BLAST search 35,335 public metagenomes on JGI-IMG and we grouped our results by our broad and narrow criteria. The broad search hits contained the full-length *arxA* gene and the predicted active site motif XGRGWG (Figure 5A). The narrow search hits contained the full length *arxA*, predicted active site motif, and the predicted *arx* gene clusters *arxXSRB2AB1CD* (Figure 5B). A total of 283 hits were detected, with hits on all seven continents and common environment types (Figure 5AB). When looking at the sample environments of total hits, 43.8% were aquatic, 25.09% were sediment, 16.25% were biofilm, 14.84% were other (i.e., soil, activated sludge). Freshwater environments accounted for 29.67% of the total hits. Hot spring environments accounted for 28.62% of the total hits and contained the most hits in both broad and narrow BLAST searches. Of the total hits, 10.25% were marine environments, 8.12% were wastewater environments, while 6.71% were

soil environments. Hot spring environments appeared to be the dominant source of *arx*-containing microbial communities, which could be a result of generally higher arsenic concentrations of hot springs. Although the *arx* pathway is not commonly associated with freshwater, groundwater, or wastewater in published literature, we found 37.44% of the total hits came from these environments. These findings greatly expand on the previously knowledge of the distribution of *arx* genes in nature. Photoarsenotrophy and Arx-type arsenotrophy has largely been associated with extreme environments like hypersaline and alkaline lakes, and sulfidic hot springs. Further, there is a clear bias for an active site containing a Y in the X position of the active site. It remains to be validated how the sequence of ArxA affects function.

Section 3.04 Discussion

The overall goal of this work was to investigate the occurrence of arsenite-linked photosynthesis in a freshwater environment. A potentially useful freshwater environment to investigate photoarsenotrophy is Owens River in the Long Valley Caldera (USA), where an active hydrothermal system brings naturally occurring arsenic from the Earth's crust to the surface via hot springs. Owens River receives arsenic-rich water from Hot Creek (2.7 μM average, 350 ppb) and Little Hot Creek (9 μM average, 1,169 ppb; Salmassi *et al.,* 2002). Samples collected from Owens River water and sediment taken in July 2021 were shown to contain 7 μM As(III) (909 ppb), for context that is about 100 times more than the US EPA allows for drinking water (10 ppb). Owens River flows to Lake Crowley, a reservoir used as drinking water for Los Angeles County. Microbial prospecting within the Long Valley Caldera

region has yielded the isolation of numerous arsenotrophic and arsenic resistant microbes. For example, the aerobic arsenite-oxidizing bacteria, *Agrobacterium albertimagni* str. AOL15 (Salmassi *et al.*, 2002) was isolated from aquatic plants in the arsenic rich waters of Hot Creek. Various *Gammaproteobacteria*, that respire arsenate or oxidize arsenite coupled to nitrate reduction, were isolated from Mono Lake (Hoeft *et al.*, 2007; Kulp *et al.*, 2008; Zargar *et al.*, 2010; Heoft-McCann *et al.*, 2017). Surprisingly, photosynthetic bacteria and *Alphaproteobacteria* are yet to be cultured. Since the Long Valley Caldera is a region of the United States with the some of the highest annual solar irradiation (>5.2 kWh/m^2/Day), we predicted phototrophic As(III)-oxidizing bacteria would be present (Sengupta *et al.*, 2018). Therefore, we incubated enrichment cultures with constant infrared (IR) light and varying arsenite concentrations to select for anoxygenic phototrophic bacteria that tolerate arsenite. The isolation of ORIO is the first freshwater photoarsenotroph from this region to be sequenced and characterized, introducing the potential for further exploration of how light impacts the arsenic biogeochemical cycle.

ORIO Genome Sequencing

We sequenced the ORIO genome to characterize the underlying genetics of arsenite-oxidation. Our draft ORIO genome assembled into 8 contigs, using MinION nanopore sequencing data with 972x coverage. We further polished the nanopore assembly using merged and quality trimmed MiSeq sequencing data resulting in a total of 428x coverage. The final assembly contains 2 chromosomes and 6 plasmids. Many strains of *Cereibacter* and *Rhodobacter* are known to have multiple

chromosomes and plasmids (GenBank sequence accession numbers: CP090021-CP090026, CP030271-CP030276, CP000577-CP000579, CP031750-CP031755). The size and G+C composition of the ORIO genome are also typical among the *Cereibacter* genus. The genus name *Cereibacter* is relatively new, but studies on members in this group date back to 1907 when Hans Molisch described 16 strains of phototrophic bacteria isolated from mud collected in California and Cuba (Molisch, 1907). In 1944, C.B. van Niel described purple non-sulfur and brown bacteria, which were then reclassified twice, once in 1984 to the *Rhodobacter* name, and again in 2020 to the current *Cereibacter* name (van Niel, 1944; Imhoff *et al.,* 1984; Hördt *et al.,* 2020).

Genome and Plasmid Analysis

Comparative genome analysis suggests the ORIO genome is closely related to *Cereibacter azotoformans* strain YLK20 (Figure 1A). The first *C. azotoformans* was discovered in 1996, was isolated from photosynthetic wastewater sludge, and performed denitrification (Hiraishi *et al.,* 1996). The other related genomes belong to *Cereibacter sphaeroides* strains, and 16S rRNA phylogenetic analysis confirms a close relation (Figure 1B). The *Cereibacter* genus (formerly *Rhodobacter*) has been an exceptional model organism for studying photosynthesis for over 60 years. *Cereibacter* are commonly used in photosynthesis, bioremediation, and biotechnology studies, meaning ORIO could be a potential model organism for studying photoarsenotrophy (Orsi *et al.,* 2020).

The *arx* gene cluster encodes the ability to perform photoarsenotrophy (Hernandez-Maldonado *et al.*, 2016), and was found in the ORIO genome on a single 136 kb plasmid, pORIO3. The arsenic island on this plasmid was likely acquired through a horizontal gene transfer (HGT) event. The sigma-54 factor (RpoN) is thought to play a role in regulation of Aio-based arsenite-oxidation (Kang *et al.*, 2012). Previous studies on Arr pathway have found cAMP-CRP and *arsR* to be involved, while studies on Aio pathway have found *aioXSR* and *pho* genes to be involved (Murphy *et al.*, 2009; Murphy and Saltikov, 2009; Liu *et al.*, 2012; Wang *et al.*, 2018; Corsini *et al.*, 2018). While recent biochemical studies with ArxX have demonstrated it specifically binds arsenite, future studies on *arxXSR* are required to fully elucidate their role in regulating the Arx pathway (Badilla *et al.*, 2018). ORIO can serve as a model organism for characterizing to biological mechanisms underlying photoarsenotrophy.

Physiology

The physiologic versatility of *C. azotoformans* str. ORIO makes ORIO advantageous for investigating the biological mechanism of photoarsenotrophy. ORIO is mesophilic and can grow in a range of 0-10% salinity. Growth can occur aerobically in the dark using succinate, acetate, glycerol, glucose, lactate, and pyruvate as electron donors. ORIO ferments sucrose without forming gas. ORIO grows anaerobically with infrared light and an organic carbon source like succinate, acetate, citrate, glycerol, glucose, or yeast extract. We have not been able to grow ORIO anaerobically without organic carbon or light. The ORIO genome contains genes for denitrification (*napA*,

nirK, *norBC*, *nosZ*) but has not been shown to denitrify, suggesting culture conditions did not support that mode of growth. Antibiotic sensitivity has been demonstrated with 5 μg/mL kanamycin and 1 μg/mL chloramphenicol, which allows for genetic manipulations using traditional techniques (Saltikov and Newman, 2003; Jaschke *et al.*, 2011).

Under aerobic conditions ORIO can resist levels of arsenic up to 2 mM, but during anaerobic growth is inhibited by concentrations greater than 0.5 mM arsenite (Figure 3). Compared to other arsenite-oxidizing and arsenate-reducing bacteria, ORIO is slightly more sensitive to arsenite (Saltikov and Newman, 2003; Zargar *et al.*, 2010; Salmassi *et al.*, 2010; Heoft-McCann *et al.*, 2017; Durante-Rodriguez *et al.*, 2019) Future RNA-seq studies can highlight systems that are differentially affected by As(III) during anaerobic conditions and provide the genetic context to variations in arsenic sensitivity. As we predicted, transcription of *arxA* was detected in cultures grown with As(III) (Figure 4). Future studies on Arx in ORIO should aim to determine how light-dark cycling affects arsenic transformations in co-culture experiments. Nature is complex and likely contains both arsenite-oxidizers and arsenate-reducers. Since photoarsenotrophy is restricted to light conditions, environments which have a light-dark cycle could have increased arsenite-oxidation during the light phase and decreased arsenite-oxidation during the dark phase. Thus, a net flux toward arsenate during the light phase and toward arsenite during the dark phase.

Light-dark Cycling in Owens River Sediment

Previous studies using microcosms prepared with water from Big Soda Lake (NV, USA), an arsenic-rich soda lake, detected a light-dark arsenic cycle over the course of weeks-long light-dark phases (Hernandez-Maldonado, 2017). It remained to be determined if light-dark cycling could occur in a non-extreme environment like Owens River. Based on the isolation of ORIO, we predicted that arsenic metabolism would be detectable in Owens River sediment. Surprisingly, we detected arsenic speciation in Owens River sediment followed a light-dark cycle in alive microcosms (Figure 6A). Arsenite was oxidized during the light phase, while arsenate was reduced during the dark phase. Heat-inactivated microcosms were autoclaved for 20 minutes prior to the start of the experiment and did not show the same pattern of arsenic cycling (Figure 6B). The sediment spiked with 50 μM arsenite contained everything needed to perform light-dark cycling of arsenic. The reasoning for using filtered Little Hot Creek water in combination with Owens River sediment is because arsenic-rich hot springs like LHC are the source of water for Owens River. Eventually, arsenic-rich water and microbes adapted to hot springs will travel downstream and can adapt to new environmental conditions or participate in horizontal gene transfer. Either situation would yield a similar result of delivering arsenic, and the genes needed to metabolize it, downstream. We predict that the agricultural fields in Owens Valley may have the ability to light-dark cycle arsenic under the right environmental conditions. This needs further exploration and validation within the natural environment.

Metagenomic analysis of Owens River microcosms

Since light-dependent arsenite oxidation could be a mechanism of light-dark cycling of arsenic, we wanted to determine if *arx* genes could be detected in the light-dark cycled microcosms. The metagenome from microcosms prepared with Owens River sediment microcosm experiments contained around 3000 taxonomic hits, and contigs were binned by taxonomy (genus level) into 1 of 12 bins. Arsenic genes were detected in several bins (Figure 7AB). The *Gammaproteobacteria, Allochromatium warmingii*, contained a homolog to the ORIO *arxA* gene and was found in a cluster that closely resembles reference *arx* clusters previously described (Figure 7C). The nature of third-generation sequencing approaches, like MinION Nanopore, are that high error rates causes indels in the assemblies. Polishing helps to remove these low-quality regions with errors. Currently, the only way to significantly reduce the errors is with hybrid assemblies that also use short reads from other next-generation sequencing platforms (e.g., Illumina). Nonetheless, we were able to detect a fragmented *arxA* homolog in the *Gammaproteobacteria Allochromatium warmingii*, and fragmented *arrA* homologs in the *Deltaproteobacteria Desulfococcus oleovarans* and *Desulfatitalea tepidiphila.* By manually polishing the fragmented *arxA* and *arrA* regions, we were able to recover complete sequences. Although we detected *arrA* fragments in a *Betaproteobacteria, Rubrivivax gelatinosus,* we were unable to recover a complete sequence through polishing attempts and therefore left it out of phylogenetic analysis. This is the first report to demonstrate light-dark cycling of arsenic in freshwater sediment coupled with metagenomics. This approach provides

critical insight into the potential arsenic functional genes driving light-dark arsenic redox cycling. Light-dark cycling of arsenic has been detected in Mono Lake's Paoha Island hot springs red mat material (Kulp *et al.,* 2008). The microbial members of this community were dominated by *Ectothiorhodospira* and several halophilic Archaea. *Ectothiorhodospira* are known to use ArxA for arsenite oxidation. However, there are no known confirmed studies of arsenate-respiring haloarchaea. Previous reports have also highlighted the mutualism between Fe-oxidizers and Fe-reducers co-cultured together, and we suspect a sustained common pool of arsenic is also shared amongst As-oxidizers and As-reducers (Cooper *et al.,* 2020).

Metagenome Mining and *arxA* Phylogenetic Analysis

Previous studies have provided accounts for the distribution of *arx* genes in nature but have sometimes included sequences which are likely not ArxA. The crystal structure of the *Shewanella sp.* ANA-3 arsenate reductase ArrA shed light on what likely gives ArxA directionality toward arsenite oxidation, a XGRGWG motif (Glasser *et al.,* 2018). Using this motif information, we performed a broad and narrow BLAST search of publicly available JGI-IMG metagenomes. The broad search hits contained only the full-length *arxA* gene and the predicted active site motif XGRGWG. We detected 283 *arxA* hits in the broad search (Figure 4A). Further analysis of the active site residues for all 283 hits provides insight into the frequency of the amino acid in the X position (figure 4B). The narrow search hits contained the full length *arxA*, predicted active site motif, and the putative *arx* gene clusters *arxXSRB2AB1CD.* We included *arxXSR* as a constraint because we have found *arxR* to be critical for arsenite

oxidation in *Ectothiorhodospira sp.* str. BSL-9, and therefore likely to be involved in regulating Arx-based arsenotrophy (Appendix A Figure 2). The narrow search contained 63 hits to metagenomes from all seven continents (Figure 4C). Our results indicate that *arxA* can be found in microbial communities from many different biomes (i.e., aquatic, sediment, freshwater, salt water, wastewater bioreactor sludge, acid-mine drainage/tailings, and soil; Figure 4C). We used the ArxA sequences from our narrow JGI-IMG BLAST search to build a phylogenetic distance tree with other known *arxA*-containing arsenotrophs (Figure 5). Phylogenetic analysis of the ArxA protein sequences show clustering of ArxA is mainly by taxonomy. The results also show some metagenome sample environment types cluster together while others do not. For example, hits from saline sample environments tended to cluster with known halophiles, while hits from freshwater environments clustered with bacteria isolated from neutral pH and low-salt environments. The *arxA* detected in the Owens River metagenome clustered with other *Gammaproteobacteria*, while the ArrA that were detected clustered outside of the ArxA tree and closer to the *Shewanella* ANA-3 ArrA sequence. This analysis expands our knowledge of the distribution of *arx* genes in nature and the environments they are detected in.

Section 3.05 Conclusion

Isolation of *Cereibacter azotoformans* str. ORIO indicates that photoarsenotrophy is occurring in freshwater sediment environments, an analog to flooded rice paddies. Genome sequencing revealed multiple chromosomes and plasmids, consistent with previously sequenced *Cereibacter* species. The Arx gene cluster was likely acquired

through horizontal gene transfer. Metagenomic mining detected the *arx* gene clusters in samples from soil, sediment, and wastewater environments, which are typically associated with the Arx pathway. Our light-dark cycling studies suggest that environments containing a photoarsenotroph and arsenate reducer can share a pool of arsenic in a mutualistic manner. Since Owens Valley was once a booming agricultural area that relied on Owens River, it is highly likely that photosynthetic arsenite-oxidizing bacteria have been delivered to the soil throughout the area. Previous studies on rice paddy environments have characterized the role of arsenite-oxidizing bacteria that use the Aio pathway (Dittmar *et al.*, 2007, Roberts *et al.*, 2007, Zhang *et al.*, 2015). Future studies should aim to characterize the role of photoarsenotrophy in agricultural settings, such as rice paddies.

Section 3.06 Experimental Procedures

Enrichment and Isolation

Strain ORIO was isolated from sediment samples collected from Owens River, CA. Samples were stored on ice during transport to UCSC. Enrichment cultures were performed anaerobically in freshwater medium (g/L): (0.33) KH_2PO_4, (0.33) $MgSO_4$ x $7H_2O$, (0.33) NaCl, (0.5) NH_4Cl, (0.05) $CaCl_2$ x $2H_2O$, (5) yeast extract and supplemented with 0.5 ml (0.02%) $FeSO_4$ x $7H_2O$, 10 ml trace elements (SL10), 10 ml vitamin solution (Wolfe), and 1 mM As(III) (Table01_SuppInfo.xlsx). Following incubation, the cultures turned green. An aliquot was transferred to fresh enrichment medium and repeated several times. The final cultures were streaked onto agar plates containing similar medium above with 1 mM arsenite. Purified isolates were stored in

25% glycerol at -80˚C. In July 2021, water and sediment were collected from Owens River with a sterile 50 mL conical tube and prepared for arsenic analysis by vortexing followed by filtering 10 mL with a 0.22 μm STERIVEX filter (Millipore).

Media Composition and Growth Conditions

Media composition can be found in Table01_SuppInfo.xlsx. ORIO was routinely grown in Luria-Bertani (LB) medium, purple non-sulfur (PNS) medium, or modified-PNS, using aseptic techniques. Agar (15 g/L) was added to solidify media when needed. Electron donors were added as needed. ORIO cultures were grown at 30°C, either aerobically in the dark or anaerobically illuminated with infrared (850 nm IR) LED light (described in Hernandez-Maldonado *et al.*, 2016). Anaerobic cultures were prepared in an anaerobic chamber (COY; 95% nitrogen, 5% hydrogen) using Balch tubes with blue butyl rubber stoppers crimped with aluminum seals. Liquid cultures were shaken at 200 rpm on an orbital shaker to simulate aerobic conditions. Anaerobic tubes were not shaken while incubating.

Genome Sequencing, Assembly, and Annotation

Short reads: A 5 mL culture was grown for 2 days in Purple Non-Sulfur media (PNS) containing 10 mM succinate. Following collection of cell pellets via centrifugation, cells were washed twice with PNS media lacking an electron donor. DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit protocol. Eluted DNA was shipped directly to Novogene for Illumina Miseq 150 bp paired-end sequencing. The sequence data was checked for quality with FastQC v0.11.9. The sequencing run produced 2.07 Gb of data and 6,891,792 raw reads.

Long reads: Starting from a glycerol stock, wild-type ORIO was grown phototrophically on mPNS agar with 5 mM succinate for two days, then overnight in a 3 mL mPNS broth culture. 50 mL of mPNS with 10 mM succinate was inoculated with 1 mL of the overnight culture, and subsequently grown phototrophically for 2 days. Using sterilized Sorvall centrifuge tubes, the 50 mL culture was centrifuged at room temperature for 10 minutes at 7500 rpm. The supernatant was removed, and DNA was extracted from the pellets using a technique previously described by Wright *et al.,* 2017 with the following modifications: we performed four parallel extractions and combined 100 µL of the aqueous phase from each extract before the final chloroform wash step. The DNA was spooled onto the tip of a glass Pasteur pipette, then soaked in a graded ethanol wash, before being allowed to air dry for 10 minutes at room temperature. The dried spool of DNA was then resuspended in 50 µL TE buffer and allowed to dissolve for several weeks at room temperature. QA/QC was done with nanodrop and gel electrophoresis. 1 µg of genomic DNA was end-prepped and ligated following the protocol for the ONT Ligation Kit (SQK-LSK-109). The concentration of the sample was measured with a QuBit prior to loading into the MinION flow cell (Mk1b, R9), following the manufacturers protocol. The MinION voltage was -170 mV and sequenced for 18 hours. Basecalling was performed by Guppy v5.0.17 software on super high accuracy. The sequence data was checked for quality with ToulligQC v2.1.1 and stats can be found in ORIO_Genome_SuppInfo.html (https://github.com/GenomiqueENS/toulligQC). The

sequencing run produced 7.36 Gb of data, 1,709,248 raw reads, N50 of 5,726, and a mean PHRED quality score of 13.26 for passed reads.

Assembly and annotation: The basecalled reads were concatenated into a single file and co-assembled with Illumina MiSeq 150 bp reads, using Unicycler v.0.4.8 via the PatricBRC web service with the following parameters: min. contig length=1000 bp, min. contig coverage=10, racon polish iterations=2 (Davis *et al.,* 2020). The assembled contigs were annotated using NCBI Prokaryotic Genome Annotation Pipeline (Tatusova *et al.,* 2016; Haft *et al.,* 2018; Li *et al.,* 2021) and deposited in NCBI under the accession number PRJNA747035 and BioSample SAMN20253214.

pORIO3 Plasmid Analysis

The iPro54-PseKNC web tool was used to analyze the pORIO3 DNA sequence for predicted sigma-54 binding sites (http://lin-group.cn/server/iPro54-PseKNC; Lin *et al.*, 2014).

Phylogenetic Analyses and Metagenome Mining

Sequences were obtained from JGI-IMG metagenome samples or NCBI via protein BLAST search of the ORIO ArxA protein sequence, 16S rRNA sequence, and pORIO3 DNA sequence (Nordberg *et al.*, 2014; Chen *et al*., 2019). Only top hits containing the XGRGWG active site motif, Mo-coordinating cysteine-259, and a tyrosine-209 residue, were used for phylogenetic analysis of ArxA. Alignments of ArxA and 16S rRNA sequences were made using EMBL-EBI Clustal Omega (Madeira *et al.*, 2019). Phylogenetic inference using maximum likelihood was

accomplished with RaxmlGUI v2.0 and the default analysis pipeline (Edler *et al.*, 2020). Phylogenetic trees were constructed using R version 4.1.2 and the packages: treeio, tidytree, ggtree, ggtreeExtra, ggstar, ggnewscale, ggpubr, and tidyverse.

RT-PCR Transcriptional Expression Analysis

RNA was harvested using the manufacturer's protocol (Qiagen RNeasy). Briefly, 500 ng of RNA was first treated with DNase (Promega RQ1) and then reverse transcribed using the TaqMan kit (Applied Biosystems). Complementary DNA (cDNA) was purified and used for PCR. Herculase Fusion II polymerase (Agilent) was used per the manufacturer's protocol for cDNA with high G+C and contained 8% DMSO per PCR reaction. Primers internal_*arxA*_F (5'-AACAGCCTGCGCGACAA-3') and internal_*arxA*_R (5'- AACCGTAGAGCTTGCCGAAG-3'), targeting a 109 bp region of *arxA*, were used to detect transcription of *arxA*. The primers annealed at 62°C.

Analytical Techniques

All samples were filtered (0.2 μm nylon membrane) and diluted 1/10 in sterile water. Arsenic speciation was determined and quantified using tandem High-Performance Liquid Chromatography Inductively Coupled Plasma Mass Spectroscopy (HPLC-ICP-MS). Elemental arsenic was detected using a Thermo X-Series2 ICP-MS instrument. The isocratic mobile phase contained 30 mM phosphoric acid, which was also used in the autosampler flush. The flow rate was 1 min/mL with a 100 μL sample injection loop. An anion exchange column (Hamilton PRP-X-100, particle size

10 μm, size 4.1 by 50 mm) was used with a 2 min/sample run-time at an ambient temperature. This method has been reported in Hernandez-Maldonado *et al.*, 2016.

Owens River Microcosm Light-Dark Arsenic Cycling Assay

Sediment from Owens River, CA, USA, was collected in March 2021, from the edge of the river (37°41'57.0"N 118°45"53.0"W), using sterile 50 mL conical tubes. Water was collected from Little Hot Creek, CA, USA, using a clean 1-gallon plastic bottle (37° 41'25.8" N 118° 50' 40.2" W). Microcosms were prepared in triplicate inside an anaerobic chamber (COY; 97% nitrogen 3% hydrogen), with 5 mL Owens River sediment slurry, 5 mL of 0.2 μm nylon Sterivex-filtered (Millipore) Little Hot Creek water, and a final concentration of 50 μM As(III). Heat-inactivated controls were prepared by autoclaving the microcosms for 20 minutes prior to the start of the experiment. Microcosms were sealed with a rubber stopper and aluminum crimp cap, then incubated at 30°C with varying infrared light-dark cycles. Samples for arsenic speciation were filtered with non-sterile 0.2 μm nylon syringe filters (Costar) prior to arsenic speciation analysis.

Metagenomic Analysis of Owens River Microcosms

DNA was extracted at the end of the microcosm assay using the Qiagen PowerMax Soil Kit. Briefly, the microcosms were vortexed and DNA was extracted from 4 grams of slurry, using the manufacturer's standard protocol. Extracted DNA was verified with gel electrophoresis for quality. We used the MinION SQK-109 Ligation Kit in combination with the PCR-free Native Barcoding Kit (NBD-104), to prepare barcoded and ligated DNA for metagenomic sequencing on a MinION Mk3 9.4.1

flow cell (Oxford Nanopore). Guppy v5.1.14 with the SUP accuracy configuration was used for basecalling the reads. Metagenome sequencing quality can be found in Metagenome-QC_SuppInfo.html. Prior to assembly, FiltLong v0.2.1 was used to filter out reads smaller than 250 bp and the lowest quality 5% of reads (github.com/rrwick/Filtlong). Due to low coverage, we merged all barcodes together and assembled the reads using Flye v2.9 with the metagenome setting (Kolmogorov *et al.,* 2020). The assembly was polished twice with Racon, and once with BlockPolish (Vaser *et al.,* 2017; Huang *et al.,* 2021). The metagenome binning service on PatricBRC was used to bin contigs by taxonomic group at the genus level (Brettin *et al.,* 2015; Davis *et al.,* 2020). This produced twelve bins of taxonomically related contigs. Kraken2 was used for taxonomic analysis of the polished assembly (Wood *et al.,* 2019). For phylogenetic analysis, genes encoding for arsenotrophic functions (e.g., *arx*, *arr*, *aio*) from metagenome bins were further polished manually by aligning the metagenomic sequence to a reference genome sequence identified by BLAST searches (Figure07_SuppInfo.xlsx). The unassembled, unfiltered/unpolished barcoded sequencing reads were submitted to the NCBI BioSample database under the accessions: SAMN26547118, SAMN26547119, and SAMN26547120. The twelve assembled and binned contigs were submitted to the NCBI database under the accession number PRJNA808117.

Section 3.07 Figures

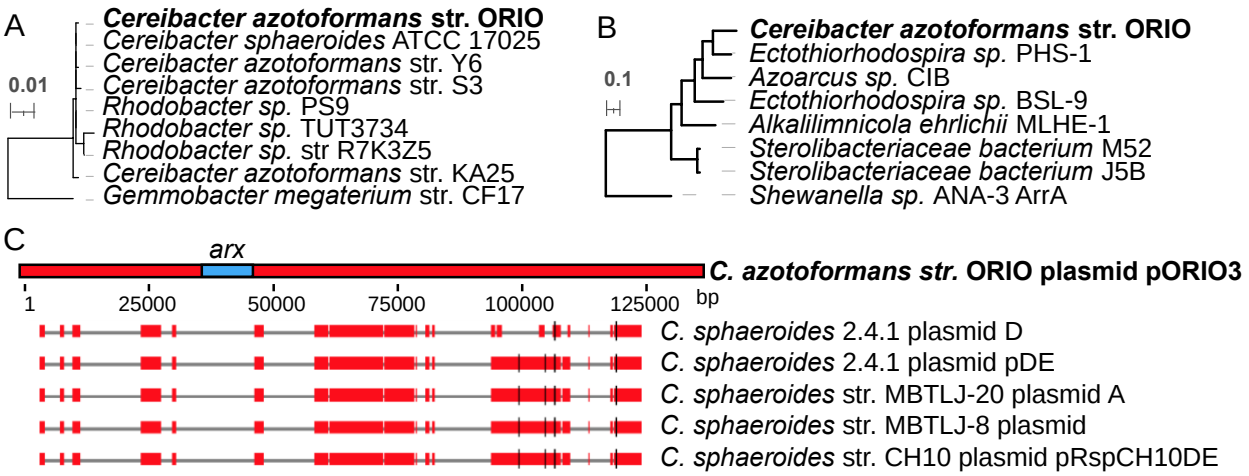


Figure 1. Taxonomic and genomic analysis of strain ORIO, a new *Cereibacter azotoformans* strain. Phylogenetic dendrogram of closely related (A) 16S *rRNA* genes, and (B) ArxA sequences from verified Arx-type arsenotrophs. Sequences were aligned with Clustal Omega and phylogenetic inferences were made using maximum likelihood method. *Gemmobacter megaterium* str CF17 was used as an outgroup to root the 16S *rRNA* phylogeny, while the *Shewanella* ANA-3 ArrA protein sequence was used to root the ArxA phylogeny. (C) An alignment of plasmid pORIO3 (red bar) with the top 5 blastn hits from NCBI (vertically from top to bottom: CP000147.2, CP030276.1, CP051473.1, CP015289.1, CP012962.1), showing regions of homology (red) and the location of the *arx* genes (blue) on pORIO3. Accession information can be found in Figure01_SuppInfo.xlsx

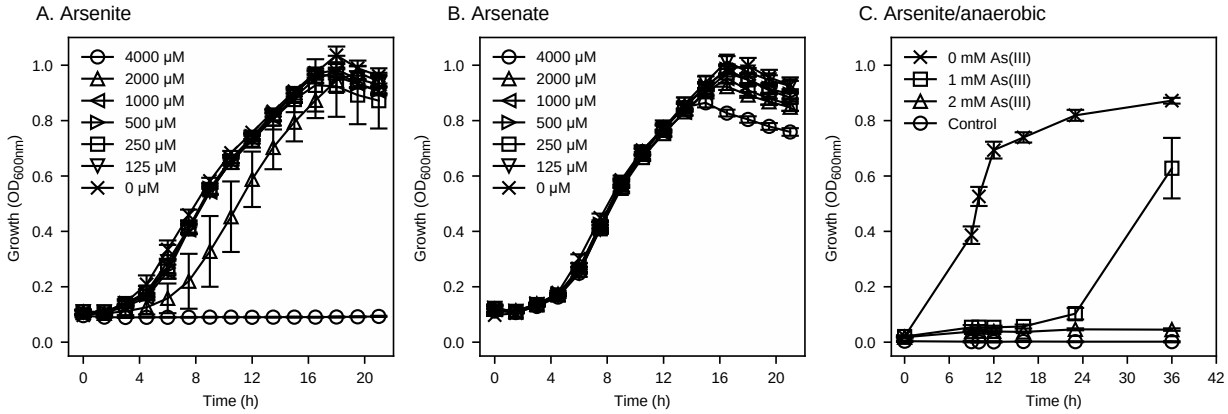


Figure 2. Effect of increasing arsenic concentration on growth. Resistance assays are shown as growth (OD_{600nm}) versus time (h). (A) Resistance to arsenite (As(III)) and (B) arsenate (As(V)), under aerobic conditions. (C) Resistance to arsenite (As(III)) under anaerobic conditions. Data points and error bars represent the mean and standard deviation of four biological replicates.

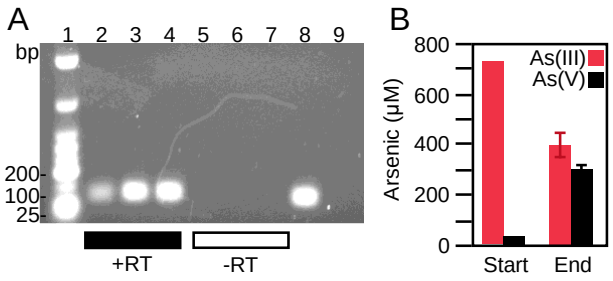


Figure 3. Detection of *arxA* transcription and As(III) oxidation in *C. azotoformans* str. ORIO cultures grown with As(III). (A) Agarose gel electrophoresis of RT-PCR products are shown. The *arxA* gene was targeted with internal primers. The lanes correspond to: New England Biolabs Low Molecular Weight DNA ladder (1), cDNA of ORIO RNA extracts with (2–4) and without (5–7) reverse transcriptase (respectively), ORIO genomic DNA (8), water negative control (9). (B) Arsenic speciation is shown for the start and end of the experiment. The start was measured from a single sample of arsenite-containing media used for the experiment. The end was measured from triplicate ORIO cultures incubated for 50 hours. Error bars are shown as 1 standard error from the mean.

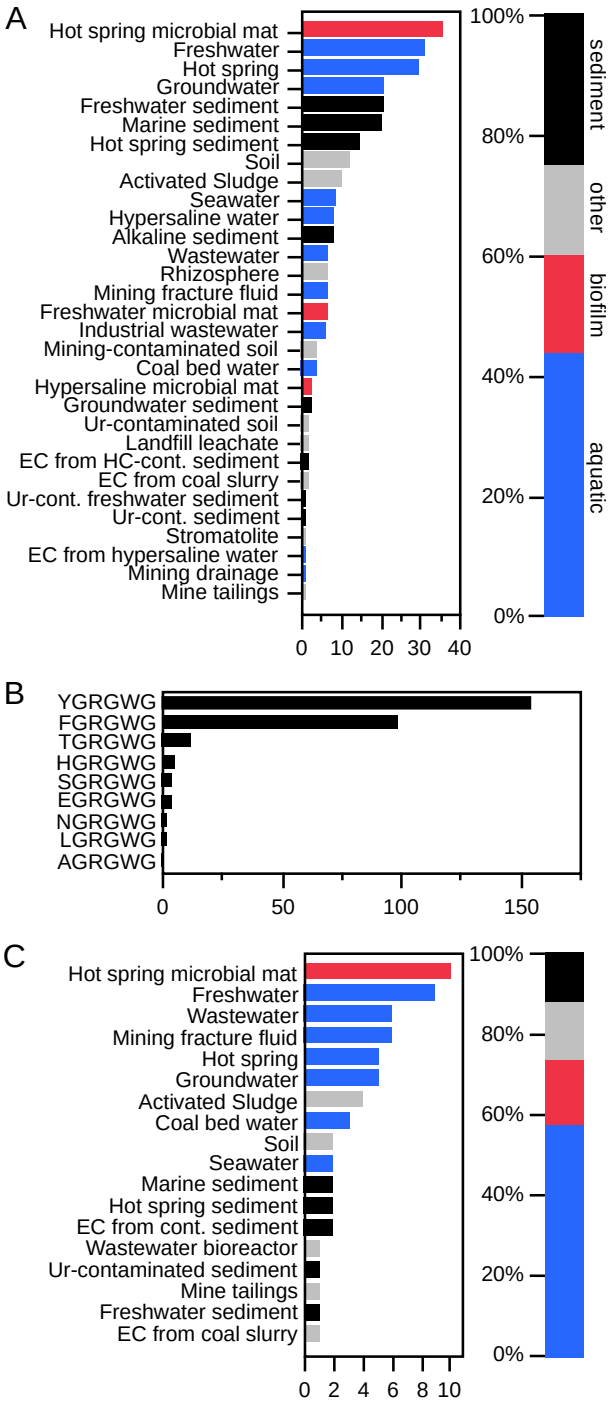


Figure 4. Detection of *arx* genes in metagenomes publicly available on JGI-IMG.

(A) The number of microbial communities containing the *arxA* gene with the predicted ArxA active site motif ("XGRGWG"), colored by sample environment

types. The stacked bar shows the percentage of 283 total hits colored by sample environment types. (B) The distribution of active site motifs among hits in panel A. (C) The number of microbial communities containing both a complete *arx* gene cluster and having an ArxA sequence with the predicted active site motif. Bars are colored by sample environment types. The stacked bar shows the percentage of 63 total hits colored by sample environment types. Abbreviations: HC = hydrocarbon, EC = enriched cells, Ur = Uranium, cont. = contaminated. See Figure04_SuppInfo.xlsx for accession information.

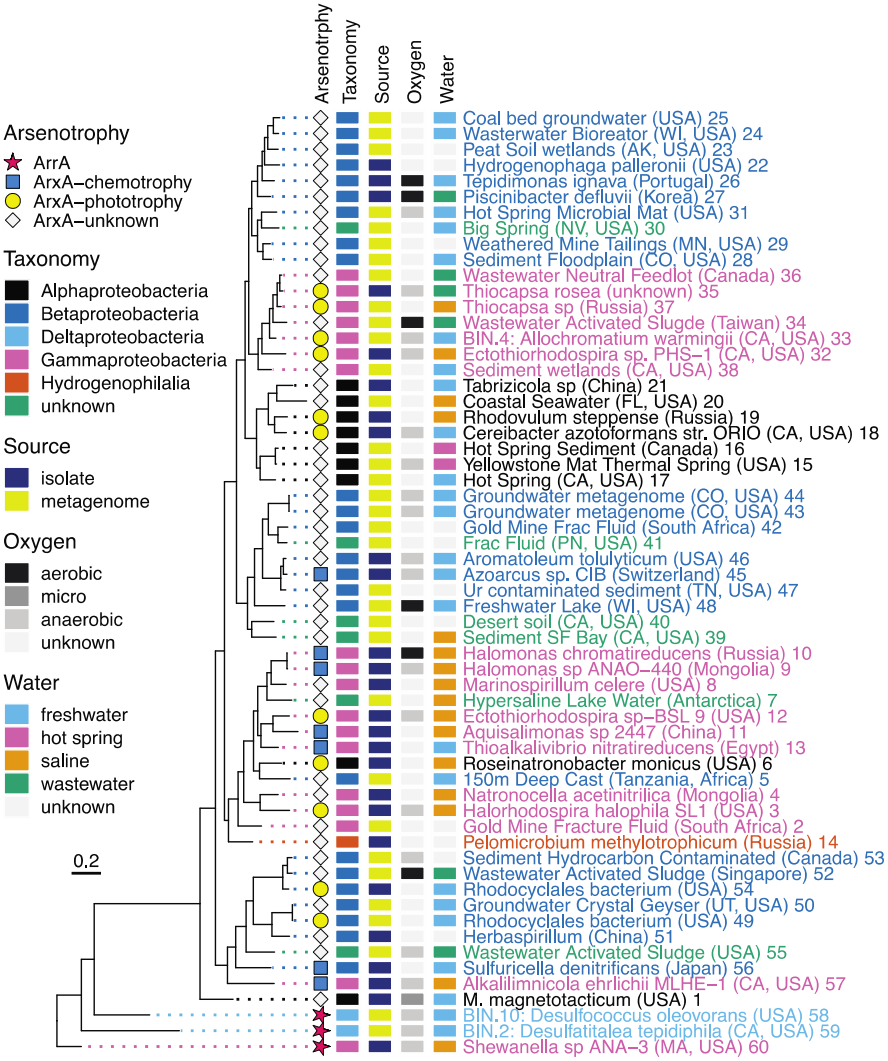


Figure 5. Phylogenetic dendrogram of strain ORIO, and other closely related As(III)-oxidizers and metagenomic hits, based on comparisons of ArxA protein sequence. Sequences were aligned with Clustal Omega and phylogenetic inferences were made using maximum likelihood method (RaxML). The *Shewanella* ANA-3 ArrA protein sequence was used as an outgroup to root the tree. Groupings were established based on if the sequence being from an isolate or metagenomic sequencing project, and whether the strain is a known chemotroph or phototroph. The

oxygen category is predicted from the sample source or type mode of metabolism. Accession numbers can be found in the Figure_ 5_SuppInfo.xlsx.

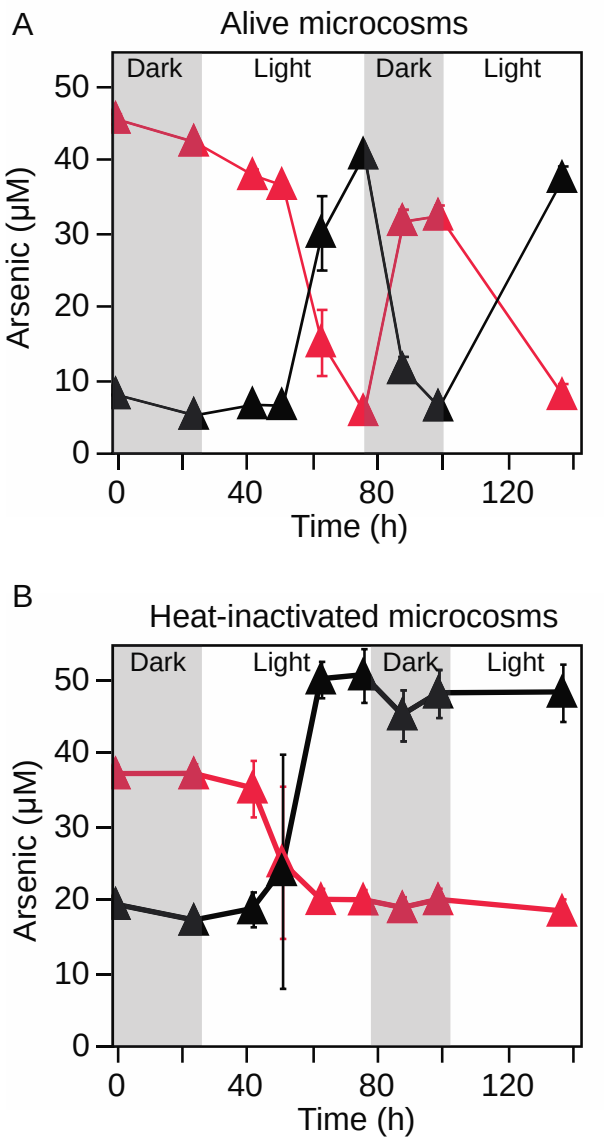


Figure 6. Effect of light-dark cycling on arsenic speciation in Owens River sediment microcosms. Owens River sediment microcosm assay is shown as arsenic speciation (μM) versus time (h) for (A) alive, and (B) heat-inactivated, microcosms. Arsenite concentrations are shown with red triangles, and arsenate with black. Shaded areas indicate dark phases and unshaded areas indicate infrared light phases. Microcosms were prepared in triplicate and contained 5 mL of Owens River sediment

slurry, 5 mL of 0.2 μm filtered Little Hot Creek water, and 50 μM As(III). Each error bar is constructed using 1 standard error from the mean.

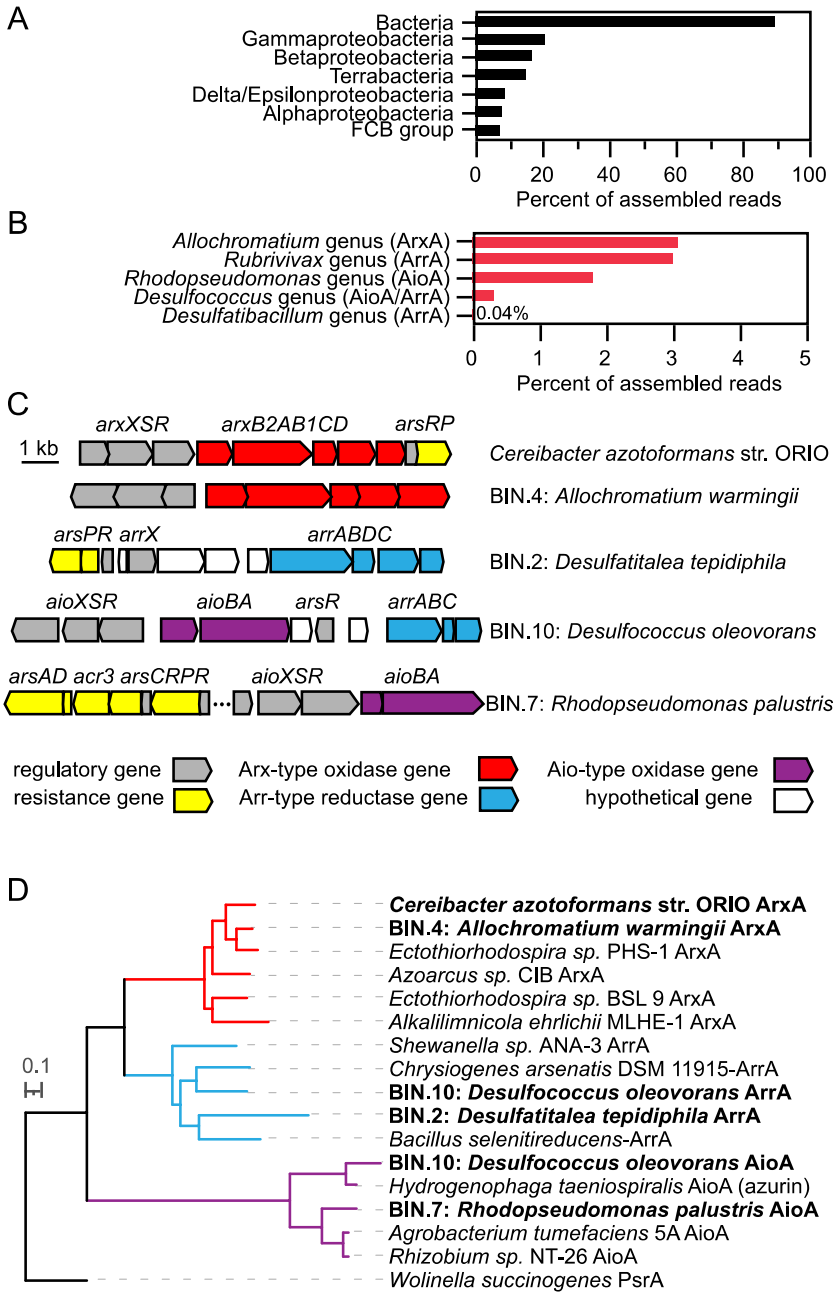


Figure 7. Detection of *arx*, *arr*, and *aio* gene clusters in the metagenome of light-dark cycled Owens River sediment microcosms. The distribution of (A) taxa, and (B) genera associated with arsenic genes detected in metagenome bins, are shown as percent abundance of total reads within the assembly. (C) Arsenotrophy genes

detected in representative bins, (D) the phylogenetic distance of arsenic genes recovered from assembled contigs binned by taxonomic group, and model organisms, for each respective mode of metabolism. Bold text indicates sequences generated in this study.

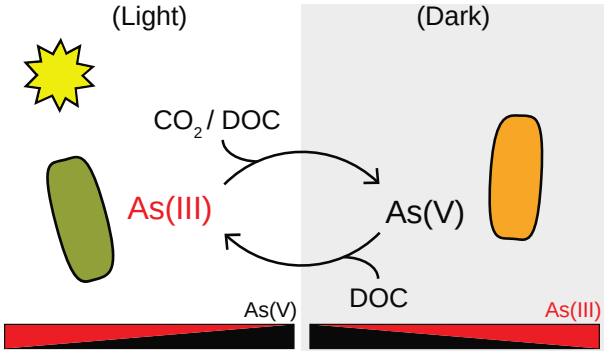


Figure 8. Model representation of light-dark arsenic cycling in freshwater sediments. Our proposed model of light-dark cycling of arsenic in freshwater sediment environments shows the cooperation of arsenite-oxidizing photoarsenotrophs (left), and an arsenate-reducing bacteria (right) in interconverting arsenite (red shading) to arsenate (black shading). Their relative concentrations are predicted to change during a respective light-dark phase, as demonstrated in the sediment microcosm experiments (Fig 6). For example, photoarsenotrophs are likely most active during the day and as the light phase ends, much of the arsenite is oxidized to arsenate. Conversely, during the dark phase, dissimilatory arsenate reducing bacteria convert arsenate back into arsenite, thereby replenishing the electron acceptor for the photoarsenotrophs. Participants of this cycle are likely using dissolved organic carbon (DOC) or fixing CO_2.

Section 3.08 References

Badilla, C., Osborne, T.H., Cole, A. *et al.,* (2018) A new family of periplasmic-binding proteins that sense arsenic oxyanions. Sci Rep 8: 6282.

Beck, A.J., Janssen, F., Polerecky, L., Herlory, O., and De Beer, D., (2009). Phototrophic biofilm activity and the dynamics of diurnal Cd cycling a freshwater system. *Environ. Sci. Technol.* **43**, 7245–7251.

Brettin, T., *et al.,* (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep. 5: 8365.

Chen, I. M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., ... & Kyrpides, N. C. (2019). IMG/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic acids research, 47(D1), D666-D677.

Cooper, R.E., Wegner, CE., Kügler, S. *et al.,* (2020) Iron is not everything: unexpected complex metabolic responses between iron-cycling microorganisms. ISME J 14: 2675–2690.

Corsini, P.M., Walker, K. T. & Santini, J. M. (2018) Expression of the arsenite oxidation regulatory operon in Rhizobium sp. str. NT-26 is under the control of two promoters that respond to different environmental cues. Microbiology open 7: 1–8.

Davis, J.J., *et al.,* (2020) The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. Nucleic Acids Res. 48(D1):D606-D612.

Dicataldo, G. *et al.,* (2011) Diel variation of selenium and arsenic in a wetland of the Great Salt Lake, Utah. *Appl. Geochemistry* **26**, 28–36.

Durante-Rodríguez, G., Fernández-Llamosas, H., Alonso-Fernandes, E., Fernández-Muñiz, M.N., Muñoz-Olivas, R., Díaz, E., & Carmona, M. (2019) ArxA From Azoarcus sp. CIB, an Anaerobic Arsenite Oxidase From an Obligate Heterotrophic and Mesophilic Bacterium. Front. Microbiol. 10: 1699.

Edler, D., Klein, J., Antonelli, A., and Silvestro, D. (2020) raxmlGUI 2.0: A graphical interface and toolkit for phylogenetic analyses using RAxML. Methods in Ecology and Evolution.

Fuller, C.C. & Davis, J.A. (1989) Photosynthetic Cycles in Natural Waters. *Nature.* 340: 1987–1989.

Glasser, N. R., Oyala, P. H., Osborne, T. H., Santini, J. M., & Newman, D. K. (2018) Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations. Proc Natl Acad Sci USA 115(37): E8614–E8623.

Haft DH, *et al.,* (2018) RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res. 46(D1):D851-D860.

Hamamura, N., Itai, T., Liu, Y., Reysenbach, A. L., Damdinsuren, N., and Inskeep, W. P. (2014) Identification of anaerobic arsenite-oxidizing and arsenate-reducing bacteria associated with an alkaline saline lake in Khovsgol, Mongolia. Environ. Microbiol. Rep. 6: 476-482.

Hao Lin, En-Ze Deng, Hui Ding, Wei Chen, Kuo-Chen Chou. (2014) iPro54-PseKNC: a sequence-based predictor for identifying sigma-54 promoters in prokaryote with pseudo k-tuple nucleotide composition. Nucleic Acids Research 42(21): 12961-12972

Hernandez-Maldonado, J., Stoneburner, B., Boren, A., Miller, L., Rosen, M., Oremland, R. S., *et al.*, (2016) Genome sequence of the photoarsenotrophic bacterium Ectothiorhodospira sp. strain BSL-9, isolated from a hypersaline alkaline arsenic-rich extreme environment. Genome Announc. 4: 9–10.

Hernández-Maldonado J., Sanchez-Sedillo B., Stoneburner B., Boren A., Miller L., McCann S., *et al.*, (2016) The genetic basis of anoxygenic photosynthetic arsenite oxidation. Environ. Microbiol. 19: 130–141.

Hernandez-Maldonado, J. (2017). Ecology and molecular genetics of anoxygenic photosynthetic arsenite oxidation by arxA. Doctoral thesis. UC Santa Cruz. ProQuest ID: HernandezMaldonado_ucsc_0036E_11312. Merritt ID: ark:/13030/m5wq503x.

Hiraishi, A., Muramatsu, K., & Ueda, Y. (1996) Molecular genetic analyses of Rhodobacter azotoformans sp. nov. and related species of phototrophic bacteria. Systematic and Applied Microbiology (19)2: 168-177.

Hoeft McCann, S., Boren, A., Hernandez-Maldonado, J., Stoneburner, B., Saltikov, C., Stolz, J., *et al.*, (2017) Arsenite as an electron donor for anoxygenic photosynthesis: description of three strains of Ectothiorhodospira from Mono Lake, California and Big Soda Lake, Nevada. Life 7: 1–14.

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

Hördt, A., *et al.*, (2020) Analysis of 1,000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of Alphaproteobacteria. Front. Microbiol. 11: 468.

Huang, N., Nie, F., Ni, P., Gao, X., Luo, F., Wang, J., (2021) BlockPolish: accurate polishing of long-read assembly via block divide-and-conquer. Briefings in Bioinformatics, bbab405.

Imhoff, J.F., Trüper, H.G., & Pfennig, N. (1984) Rearrangement of the species and genera of the phototrophic purple nonsulfur bacteria. Int. J. Syst. Bacteriol. 34: 340-343.

Jaschke, P.R., Saer, R.G., Noll, S., and Beatty, J.T. (2011) Modification of the Genome of Rhodobacter sphaeroides and Construction of Synthetic Operons. Methods in Enzymology 497: 519-538.

Kang, Y., Bothner, B., Rensing, C. & Mcdermott, T. R. (2012) Involvement of RpoN in Regulating Bacterial Arsenite Oxidation. Appl. Environ. Microbiol. 78: 5638–5645.

Kneebone, P.E., Hering, J.G. (2000) Behavior of arsenic and other redox sensitive elements in Crowley Lake, CA: A reservoir in the Los Angeles Aqueduct system. Environ Sci Technol 34: 4307–4312.

Kolmogorov, M., Bickhart, D.M., Behsaz, B. *et al.,* (2020) metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17: 1103–1110.

Kulp, T.R., Hoeft, S.E., Asao, M., Madigan, M.T., Hollibaugh, J.T., Fisher, J.C., *et al.,* (2008) Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. Science 321: 967–970.

Li, W., *et al.,* (2021) RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. Nucleic Acids Res. 49(D1):D1020-D1028.

Lin, H., Deng, E.Z., Ding, H., Chen, W., Chou , K.C., (2014) iPro54-PseKNC: a sequence-based predictor for identifying sigma-54 promoters in prokaryote with pseudo k-tuple nucleotide composition. Nucleic Acids Research 42(21): 12961-12972.

Liu, G. *et al.,* (2012) A periplasmic arsenite-binding protein involved in regulating arsenite oxidation. Environ. Microbiol. 14: 1624–1634.

Madeira F., Park Y.M., Lee J., *et al.,* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research (W1):W636-W641.

Mirdita, M., Steinegger, M., Söding, J., (2019) MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics 35(16):2856–2858

Molisch, H. (1907) Die Purpurbakterien nach neuen Untersuchungen. Gustav Fischer, Jena.

Murphy, J.N. & Saltikov, C.W. (2009) The ArsR Repressor Mediates Arsenite-Dependent Regulation of Arsenate Respiration and Detoxification Operons of Shewanella ANA-3. J. Bacteriol. 191: 6722–6731.

Murphy, J.N., Durbin, K.J. & Saltikov, C.W. (2009) Functional Roles of arcA, etrA, Cycli cAMP (cAMP)-cAMP Receptor Protein, and cya in the Arsenate Respiration Pathway in Shewanella ANA-3. J. Bacteriol. 191: 1035–1043.

Nimick, D. A. *et al.*, (2003) Diel cycles in dissolved metal concentrations in streams: Occurrence and possible causes. *Water Resour*. *Re*s. 39.

Nimick, D. A., Cleasby, T. E. & McCleskey, R. B. (2005) Seasonality of diel cycles of dissolved trace-metal concentrations in a Rocky Mountain stream. *Environ. Geol.* **47**, 603–614.

Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev IV, Dubchak I. (2014). The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. 42(1):D26-31.

Oremland, R.S. & Stolz, J.F. (2003) The Ecology of Arsenic. Science (80)300: 939–944.

Orsi, E., Beekwilder, J., Eggink, G., Kengen, S.W.M., Weusthuis, R.A. (2021) The transition of Rhodobacter sphaeroides into a microbial cell factory. Biotechnology and Bioengineering 118: 531– 541.

Ospino MC, Kojima H and Fukui M (2019) Arsenite Oxidation by a Newly Isolated Betaproteobacterium Possessing arx Genes and Diversity of the arx Gene Cluster in Bacterial Genomes. Front. Microbiol. 10: 1210.

Redondo-Salvo, S., Bartomeus, R., Vielva, L., Tagg, K.A., Webb, H.E., Fernandez-Lopez, R., de la Cruz, F. (2020) "COPLA, a taxonomic classifier of plasmids". bioRxiv

Roy, A., Yang, J., & Zhang, Y. (2012) COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research 40:W471-W477.

Salmassi, T.M., Venkateswaren, K., Satomi, M., Newman, D.K., & Hering, J.G. (2002) Oxidation of Arsenite by Agrobacterium albertimagni, AOL15, sp. nov., Isolated from Hot Creek, California. Geomicrobiology Journal 19(1): 53-66.

Saltikov, C. W., & Newman, D. K. (2003) Genetic identification of a respiratory arsenate reductase. Proc Natl Acad Sci USA 100(19): 10983–10988.

Sarmiento, A. M., Oliveira, V., Gómez-Ariza, J. L., Nieto, J. M. & Sánchez-Rodas, D. (2007) Diel cycles of arsenic speciation due to photooxidation in acid mine drainage from the Iberian Pyrite Belt (Sw Spain). *Chemosphere* **66**, 677–683.

Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., and Shelby, J. (2018) The National Solar Radiation Data Base (NSRDB). Renewable and Sustainable Energy Reviews 89(June): 51-60.

Shaji, E., Santosh, M., Sarath, K. V., Prakash, P., Deepchand, V., & Divya, B. V. (2020) Arsenic contamination of groundwater: A global synopsis with focus on the Indian Peninsula. Geoscience Frontiers 12(3).

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., Ostell, J. (2016) NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44(14):6614-24.

van Niel, C.B. (1944) The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. Bacteriological reviews, 8(1): 1–118.

Vaser, R., Sović, I., Nagarajan, N., & Šikić, M. (2017) Fast and accurate de novo genome assembly from long uncorrected reads. Genome research 27(5): 737–746.

Wang, Q. *et al.,* (2018) Phosphate starvation response controls genes required to synthesize the phosphate analog arsenate. Environ. Microbiol. 20: 1782–1793.

Wilkie J. and Hering J. (1998) Rapid oxidation of geothermal arsenic (III) in streamwaters of the eastern Sierra Nevada. Environ Sci Technol 32: 657–662.

Wood, D.E., Lu, J. & Langmead, B. (2019) Improved metagenomic analysis with Kraken 2. Genome Biol 20: 257.

Wright, M. H., Adelskov, J., & Greene, A. C. (2017) Bacterial DNA Extraction Using Individual Enzymes and Phenol/Chloroform Separation. *Journal of microbiology & biology education*, 18(2), 18.2.48.

Wu, G., Huang, L., Jiang, H., Peng, Y., Guo, W., Chen, Z., *et al.*, (2017) Thioarsenate formation coupled with anaerobic arsenite oxidation by a sulfate-reducing bacterium isolated from a hot spring. Front. Microbiol. 8: 1336.

Yang, J., Roy, A., Zhang, Y., (2013) BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. Nucleic Acids Research 41: D1096-D1103

Yang, J., Roy, A., Zhang, Y., (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29: 2588-2595.

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

Zargar, K., Conrad, A., Bernick, D.L., Lowe, T.M., Stolc, V., Hoeft, S., Oremland, R.S., Stolz, J., Saltikov, C.W., (2012) ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. Environ Microbiol. 14(7):1635-45.

Zhang, C., Freddolino, P.L., and Zhang, Y. (2017) COFACTOR: Improved protein function prediction by combining structure, sequence, and protein-protein interaction information. Nucleic Acids Research 45: W291-299.

Section 3.09 Acknowledgements

We thank Esra Mescioglu for initial culturing efforts of strain ORIO. We thank Dr. David Bernick for insightful discussion and guidance on the genome assembly of ORIO. A very special thanks to Brian Dryer and the UCSC Marine Analytical Lab for supporting the arsenic speciation analyses. We'd also like to acknowledge our Summer Internship Program 2020 interns (Pranav Kirti, Madeline Day, and Isaac Chang) for their assistance in processing the metagenomic dataset from JGI-IMG. We thank Christina Egami, Christie Kwok, Meghan Lamba, and Andrew Uyesugi for collecting samples for the Owens River microcosm/metagenome experiments in 2020. We also thank Tarik Stupac for help with setting up Google Colaboratory Notebooks for nanopore sequencing analyses. The work presented here was funded in part by the National Sciences Foundation (EAR-1349366) award to CWS.

Chapter 4. Characterization of a novel photoarsenotroph, *Cereibacter azotoformans* str. ORIO, a new model organism for photosynthetic arsenite oxidation

Section 4.01 Abstract

Arsenic is a poisonous element which naturally finds its way into water and food supplies. Consumption of arsenic-contaminated drinking water or food is linked to cancer and cardiovascular disease in humans. In the environment, bacteria are primary drivers of arsenic speciation. Some bacteria have the ability to reduce arsenate to arsenite, while other bacteria are able to oxidize arsenite to arsenate. This is important because arsenite is considered significantly more toxic to life than arsenate. Our knowledge of how bacteria transform arsenic in nature can better inform our ability to predict and mitigate arsenic contamination. In this study, the As(III) oxidation and physiology of *Cereibacter azotoformans* str. ORIO, a photosynthetic arsenite oxidizing *Alphaproteobacteria*, was characterized. Protein structure and ligand binding predictions provide insight into the organization of Arx subunits, electron flow, and the active site of the arsenite oxidase ArxA. Transcription of *arxA* is specific to anaerobic condition and the presence of arsenite. Growth assays, arsenic speciation measurements, and mutagenesis experiments determined ORIO requires light, organic carbon, and the *arxA* gene for photoarsenotrophy. In our studies, ORIO did not couple nitrate reduction to arsenite oxidation in the dark or perform carbon fixation. Because ORIO is a photoheterotroph, strict growth dependency on arsenite as a sole electron donor could not be established. Thus, due to

no observed growth advantage to cells grown with arsenite, this is the first report to describe photoarsenotrophy as a mechanism for arsenite resistance.

Section 4.02 Introduction

Arsenic is a naturally occurring element found in the Earth's crust and has provided microorganisms a metabolic niche for millions of years (Oremland and Stolz, 2003). The transformation of arsenic in the environment, or arsenic biogeochemical cycling, has broad implications for environmental and human health. Over 230 million people worldwide are likely consuming arsenic polluted water that exceeds the recommended limit of 10 μg/L (Shaji *et al.*, 2021). Arsenic in water supplies is usually caused by naturally occurring processes, with arsenite (As(III)) and arsenate (As(V)) as the predominant inorganic species. Arsenic is primarily cycled by prokaryotic organisms, and since the inorganic species are more toxic to humans, research tends to focus on the bacteria which are involved in redox reactions with inorganic arsenic. To date, bacterial arsenite oxidation is exclusive to the two pathways termed Aio and Arx (Ellis *et al.*, 2001; Inskeep *et al.*, 2007; Zargar *et al.*, 2012; Hernandez-Maldonado *et al.*, 2016). Arx differs from Aio in that it is a strictly anaerobic process, which requires light energy or nitrate reduction coupling to arsenite oxidation (i.e., *Ectothiorhodospira sp.* str BSL-9 or *Alkalilimnicola ehrlichii* MLHE-1, respectively; Zargar *et al.*, 2012; Hernandez-Maldonado *et al.*, 2016). This work focuses on the anoxygenic photosynthetic-coupled Arx pathway termed 'photoarsenotrophy', and not the nitrate-coupled pathway herein termed 'Arx-based

arsenotrophy'. To date, photoarsenotrophy has not been investigated in freshwater environments and a robust genetic model is lacking.

Photoarsenotrophy is an anaerobic process by which bacteria use arsenite as a source of electrons, converting it to the less toxic arsenate, while harnessing the ATP production of photosynthesis (Zargar *et al.,* 2012). This process utilizes the *arx* genes, specifically the *arxA* arsenite-oxidase gene (Hernandez-Maldonado *et al.,* 2016). ArxA is more similar to ArrA (arsenate reductase) than it is to AioA, and features a XGRGWG motif in its active site which is thought to provide directionality to the reaction, distinguishing it from a "reversible reductase capable of oxidation" (Glasser *et al.,* 2018). ArxA contains a molybdopterin co-factor, which is thought to be coordinated by a conserved Cysteine-192 in *Shewanella sp.* ANA-3 (Glasser *et al.,* 2018). It is thought that ArxA interacts with two iron-sulfur cluster containing proteins, ArxB1 and ArxB2, to shuttle electrons from As(III) to the quinone pool. ArxC contains transmembrane domains and is analogous to the nitrate reductase membrane protein NrfD. The protein folding chaperone ArxD is likely to be involved in assembly of this enzyme. To date, photoarsenotrophs have only been isolated from extreme environments, which has resulted in a major impediment for developing robust genetic systems (Hoeft-McCann *et al.,* 2017). Genetic manipulations with bacteria from extreme environments are extremely time consuming using traditional techniques. One recent study characterized the first heterotrophic bacteria that can use the Arx pathway coupled to nitrate, but this bacterium is not suitable for use as a model organism for studying photosynthetic arsenite-oxidation (Durante-Rodriguez et

al 2019). Here, we characterized a recently described photoheterotrophic freshwater photoarsenotroph from Owens River, California, *Cereibacter azotoformans* str. ORIO (formerly known as *Rhodobacter azotoformans*), and shown it is a suitable model organism for studying photoarsenotrophy. *Cereibacter* were recently reclassified from *Rhodobacter*, but nonetheless, over 40 years of research exists on the members of the *Cereibacter* genus, and ORIO adds the ability of *Cereibacter* to serve as model organisms. We report the characterization of physiology and photoarsenotrophy in ORIO.

Section 4.03 Results and Discussion

Arx enzyme structure prediction provides new insight for photoarsenotrophy

Since no Arx proteins have been crystallized to date, we wanted to perform a 3D structure prediction to gain insight into the mechanism of the Arx-type arsenite oxidase. Using a web-based version of AlphaFold v2.0 called ColabFold (Mirdita *et al.,* 2021), we predicted the 3D protein structures of Arx proteins thought to be involved in arsenite oxidation in ORIO, BSL-9 and MLHE-1 (Figure 1). AlphaFold first aligns the query protein to publicly available protein sequences, and then uses alignments to guide predictions. Per-residue estimates of confidence are called pLDDT, where lower scores are associated with lower confidence in the prediction. The putative iron-sulfur protein, ArxB2, had lower sequence coverage than the other subunits, which resulted in lower pLDDT (Figure 1A-C). Since ArxA, ArxB1, and ArxC all had high sequence coverage and pLDDT, we suspect ArxB2 was incorporated into this gene cluster after ArxB1. When looking at the five models

generated for ORIO, BSL-9, and MLHE-1, respectively, ArxB1 always forms an interface between ArxA and ArxC, whereas ArxB2 does not form a similar interface, and instead is positioned to the side of ArxAB1C. It is unclear if ArxB2 is needed for arsenite oxidation or if it plays an accessory role like stabilizing the complex, and future studies should aim to identify the role of ArxB1 and ArxB2. ArxC is predicted to encode a membrane protein with 10 transmembrane domains, believed to anchor the complex to the cytoplasmic membrane. All these subunits are detected as periplasmic proteins by Phobius, but a signal peptide was only detected in ArxA (Käll *et al.,* 2007). Structural analysis suggests ORIO ArxA is most closely related to the *Shewanella sp.* ANA-3 ArrA protein (Figure 2D). This finding is in line with the conclusion made by Glasser *et al.,* (2018) and confirmed by ArrA typically clustering close to ArxA sequences in phylogenetic tree analyses. The aligned 3D structures of ArxA and ArrA have a Tm-score of 0.93, where 1 is a perfect match (Figure 2D). ArxA and ArrA share some key structural similarities. For example, both proteins have a positively charged binding pocket which is shown for ORIO in the protein surface map (Figure 2C). The crystal structure of ArrA bound to arsenite (pdb: 6cz9) was used to predict the key residues involved in coordinating arsenite in ArxA by aligning the active site residues. In ArrA, a molybdenum ion is coordinated by a conserved cysteine-193, and a conserved tyrosine-210 coordinates arsenite with the help of histidine-189 (Glasser *et al.,* 2018). In ArxA, we predict arsenite is coordinated by residues X-160 (where X is tyrosine, phenylalanine, serine, or threonine), arginine-162, histidine-188, and tyrosine-209. The typical residue 160 is a

tyrosine, such as in ORIO, but other arsenotrophs have been found to have phenylalanine (*Azoarcus sp.* CIB) or threonine (*Sterolibacteriaceae bacterium* M52). An alignment of active sites using the predicted ORIO ArxA protein and the ANA-3 ArrA crystal structure is shown in Appendix A Figure 3. Future site-directed mutagenesis studies can identify the residues critical for arsenite-oxidation, and how directionality is maintained. Because this process is dependent on light in ORIO, we hypothesize electrons from arsenite are shuttled to the Q pool and participate in cyclic electron flow to replenish the RC of the P870 light-harvesting complex.

Arsenite and anerobic conditions are requires for *arxA* transcription in ORIO

To date, the transcription of *arxA* has only been detected in anaerobic cultures and environments which contained arsenite. Since ORIO can grow aerobically or anaerobically, we wanted to see if it had the ability to couple Arx-type arsenite oxidation to aerobic respiration. When we grew ORIO with 100 μM arsenite or 100 μM acetate under aerobic and anaerobic conditions, we detected a 10-fold induction of *arxA* transcription in the anaerobic arsenite treatment compared to aerobic and acetate conditions (Figure 3A). Unlike some Aio-type arsenite oxidase, the Arx-type continues to have a specific dependence on anaerobic conditions and the presence of arsenite. Next, we compared the level of *arxA* induction in the previously isolated photoarsenotroph *Ectothiorhodospira sp.* str. BSL-9 to better understand the difference in *arxA* transcription of an extremophile and a freshwater non-extremophile (Figure 3B). Cultures of ORIO and BSL-9 had a 10-fold increase in *arxA* transcription, compared to baseline, when grown anaerobically with 100 μM

arsenite. Transcription of *arxA* in BSL-9 might be induced by concentrations of arsenite between 10 μM and 100 μM, which for comparison is more sensitive than the induction of the detoxifying arsenate reductase *arsC* in *Shewanella* ANA-3 (100 μM) (Saltikov, 2011). Future studies should aim to determine the transcriptional sensitivity of *arxA* to arsenite in ORIO, and other freshwater photoarsenotrophs, to understand when important genes like *arxA* might be activated in freshwater environments.

ORIO requires light and organic carbon to grow with and oxidize arsenite.

To better comprehend the biological mechanisms underlying photoarsenotrophy in ORIO, we performed growth assays with wild-type ORIO, and *arxA* gene deletion mutant (Δ*arxA*), and various growth conditions. First, we found ORIO requires infrared light and organic carbon to grow with arsenite (Figure 2A). Organic carbon in the form of yeast extract was chosen for all experiments, but casamino acids, peptone, and tryptone could also be utilized for growth by ORIO. This finding contrasts with nearly all previously studied photoarsenotrophs since they fix inorganic carbon during photoarsenotrophy (*Ectothiorhodospira* sp. strains PHS-1, MLW-1, and BSL-9, and *Halorhodospira halophila* SL-1). Since ORIO did not use succinate or yeast extract for anaerobic growth in the dark, we wanted to know if other organic carbon sources would support dark anaerobic growth. We incubated ORIO anaerobically in the dark with 0.5 mM arsenite or 0 mM arsenite, and additions of glycerol or lactate, for 84 hours. After 84 hours, we turned on the IR light and observed immediate growth (Figure 2B). We saw opposite results with respect to

growth and arsenite oxidation. Growth with glycerol provided the highest specific growth rate, but about 100 μM arsenite remained after 120 hours, whereas lactate provided a lower specific growth rate but 0 μM arsenite remained (Figure 2C-D). One reason for this difference in growth rate and oxidation could be due to glycerol and arsenite both using aquaglyceroporins for entry into the cell, and therefore reducing the rate of arsenite uptake compared to lactate supplemented cultures (Meng *et al.*, 2004). A reduced rate of arsenite uptake corresponds with reduced rate of arsenite toxicity.

ORIO does not couple arsenite oxidation to nitrate reduction in the dark

Since some previous Arx-based arsenotrophs such as MLHE-1, CIB, and M52 coupled arsenite oxidation to nitrate reduction, and because periplasmic nitrate reductase genes (*nap*) are detected in the ORIO genome, we hypothesized ORIO may be able to grow with arsenite and nitrate under dark anaerobic conditions. wild-type ORIO was grown in light-dark cycles with varying treatments of arsenite and nitrate. During the dark phases, growth and arsenite oxidation ceased, while during the light phases growth and oxidation occurred (Figure 3). This differs from other previously studied Arx-based arsenotrophs such as MLHE-1, CIB, M52, and many other Aio-based arsenotrophs. These findings provide further evidence that ORIO requires light for photoarsenotrophy. This is an important observation because it provides a link to our previous report, where we detected light-dark arsenic redox cycling in microcosms prepared with Owens River sediment (Mehic and Saltikov, 2022; Chapter 3 of this thesis). Taken together our findings suggest there could be light-

dark variations in arsenite and arsenate concentrations in environments where light-dependent photoarsenotrophy dominates the arsenite oxidation activity.

Is photoarsenotrophy in ORIO a resistance mechanism?

When wild-type ORIO is grown photoheterotrophically, increasing arsenite additions do not correlate to growth increases, suggesting this form of photoarsenotrophy may not be coupled to growth directly (Figure 5A). This contrasts with some other Arx-based arsenotrophs, which do use arsenite as their sole electron source (BSL-9, PHS-1, MLHE-1), coupling light-dependent growth to arsenite oxidation. *Azoarcus sp.* CIB is an example of an obligate heterotroph which uses the extra electrons from arsenite for a slight growth increase. However, the mechanism appears to be largely for resistance like that of ORIO. If microbes like CIB and ORIO are using the *arx* gene cluster to resist the toxic effects of arsenite in the environment, they may have a competitive advantage since arsenite is toxic to cells lacking resistance mechanisms. Similar detoxifying resistance mechanisms have been described in a photoheterotroph capable of iron-oxidation (Jiao and Newman, 2007; Poulain and Newman, 2007). It remains to be determined, through future mutagenesis studies, where the electrons from arsenite go and why they do not provide a growth advantage in ORIO.

ORIO is a genetically malleable organism

Previous studies on the Arx pathway used gene disruptions to generate gene mutations (Zargar *et al.,* 2012; Hernandez-Maldonado *et al.,* 2016; Durante-Rodriguez *et al.,* 2018). This method not only limits our ability to make multiple mutations, but often has polar effects causing disruptions in downstream genes. Using

traditional cloning and mutagenesis techniques, we show that null gene deletions are possible in ORIO (Figure 4; Link *et al.,* 1997; Saltikov and Newman, 2003). Null mutations were obtained by first cloning PCR-amplified upstream and downstream regions of the gene of interest into a *sacB*-based mutagenesis vector containing a kanamycin resistance cassette. Next, the mutagenesis construct is conjugated into ORIO via an *E. coli* WM3064 donor strain. By plating the conjugation reaction on agar media containing kanamycin and lacking DAP, we select for single homologous recombination mutants, called exconjugants. Exconjugants are then grown on agar media containing 10% sucrose, and the *sacB* gene encoding levansucrase causes cell death, allowing for growth of cells which underwent a double homologous recombination event. Double homologous recombination efficiency was low, with around 1% of colonies grown on 10% sucrose plates being a mutant. Given the newest advancements in genetic engineering of *Cereibacter*/*Rhodobacter* genus with CRISPR, we suggest using newer techniques to obtain optimal genetic engineering efficiency (Mougiakos *et al.,* 2019). Complementation of *arxA* using the pBBR1MCS2 vector and 10 μg/mL kanamycin was highly efficient; lawns of mutants were easily attained. Our *arxA* complement strain was not able to oxidize arsenite, further RT-PCR analysis is needed to determine if *arxA* was heterologously expressed. Even though we did not successfully complement a functional *arxA*, this genetic system is more robust than those previously used to study the Arx pathway because null gene mutations do not cause polar effects and allows for multiple genes to be mutated in one strain. Another advantage of ORIO is its robust growth. It takes

four to six days to grow single colonies of BSL-9 on agar media, and an additional two days for liquid cultures to reach high cell density. ORIO on the other hand, can form colonies on agar media in two days, and reach high cell density in less than two days. The metabolic versatility of ORIO allows us to use a wider range of media and conditions. For example, ORIO grows in Luria-Bertani media, does not need sodium chloride, and can grow aerobically in the dark. Together this establishes ORIO as a model organism for studying photoarsenotrophy.

arxA is required for photoarsenotrophy in ORIO

We hypothesized the *arx* gene cluster was responsible for photoarsenotrophy and show through mutagenesis experiments that ORIO uses the gene *arxA* to oxidize arsenite. Deletion of *arxA* increased sensitivity to arsenite (Figure 5A). Previous studies did not test the sensitivity to arsenite in *arxA* disrupted MLHE-1 and BSL-9 strains, nor arsenate sensitivity in *arrA* null mutants of *Shewanella sp.* ANA-3. Therefore, this is the first report to demonstrate the role of photoarsenotrophy in resistance to arsenite. When growth phototrophically with arsenite, growth of the Δ*arxA* strain was reduced compared to wild-type (Figure 5B). In addition, arsenite oxidation was also abolished in the Δ*arxA* strain (Figure 5C). High-density cell suspension assays were used to test for growth-independent arsenite oxidation in ORIO. The Δ*arxA* mutant was unable to oxidize arsenite in the growth-independent manner whereas wild-type ORIO was able to oxidize all the arsenite (Figure 5D). During anaerobic and aerobic growth, ORIO is more resistant to arsenite than the freshwater Arx-type arsenotroph *Azoarcus sp.* CIB (Durante-Rodriguez *et al.,* 2019).

Arsenite oxidation rates are comparable to previously described photoarsenotrophs MLHE-1, BSL-9, and PHS-1. High cell density cultures can oxidize Compared to other photoarsenotrophs and Arx-type arsenotrophs, ORIO grows considerable faster and more robustly.

Section 4.04 Conclusion

Our Arx multimer prediction is the first aimed at gaining structural insight to the function of ArxA. When comparing the Arx predictions for previously studied Arx-based arsenotrophs *Ectothiorhodospira sp.* str. BSL-9 and *Alkalilimnicola ehrlichii* str. MLHE-1 with ORIO, we see in each case that ArxB2 does not interface with ArxC the way ArxB1 does. ArxB1 always forms an interface with ArxA and ArxC, while ArxB2 interfaces with the side of the ArxAB1C complex. Furthermore, the large distance from ORIO's ArxB2 subunit to the electron transfer chain suggests the role of ArxB2 could be different from ArxB1, and future studies should aim to determine if ArxB2 is required for photoarsenotrophy. While previous studies have shown heterotrophic Arx-based arsenotrophy coupled to nitrate in freshwater conditions, our experiments are the first to demonstrate the occurrence of heterotrophic photoarsenotrophy in freshwater sediment. The difference in requirement of light between Arx-based arsenotrophy and photoarsenotrophy creates a unique ecological niche, where a photoarsenotroph and an arsenate reducer can mutualistically cycle arsenic based on the light-dark phases of the day. Our previous work has shown light-dark cycling occurs in freshwater sediments, and that *arxA* and *arrA* are detected in the metagenome (Mehic and Saltikov, 2022; Chapter 3 of this

thesis). Here, our characterization of photoarsenotrophy in ORIO helps to further inform this process by providing evidence of a freshwater photoarsenotroph that does not oxidize arsenite during the dark phase. In addition, this is the first report to suggest a light-dependent resistance pathway for arsenic. Our mutagenesis experiments show *arxA* confers resistance to arsenite, allows for growth in media with high levels of arsenite, and is required for arsenite oxidation. In addition, our successful editing of the genome establishes ORIO as a model organism for studying photoarsenotrophy in freshwater environments.

Section 4.04 Experimental Procedures

Bacterial strains, growth conditions, and mutagenesis

Strains can be found in supplementary file Strains_SuppInfo.xlsx. Growth media composition can be found in Table01_SuppInfo.xlsx. *E. coli* DH5α-λ*pir* and WM3064 were used for cloning and conjugation, respectively. Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, [pH 7.5]) was used to grow all strains during cloning and conjugation experiments. *E. coli* cultures were grown aerobically at 37°C and shaken at 200 rpm, and when needed chloramphenicol was used at 25 μg/mL and kanamycin was used at 50 μg/mL.

ORIO was grown in Purple Non-Sulfur (PNS) medium or modified PNS (mPNS) as described in Mehic and Saltikov, 2022 (Chapter 3 of this thesis). Filter sterilized electron donors (succinate, acetate, citrate, glycerol, lactate) were added as needed. ORIO cultures were grown at 30°C aerobically in the dark or anaerobically

illuminated with infrared (IR, 850 nm) light. The IR light was turned off during dark phases of phototrophic experiments. For phototrophic growth, cultures were prepared in an anaerobic chamber (95% nitrogen, 5% hydrogen; COY) using freshly autoclaved mPNS medium and Balch tubes with blue butyl rubber stoppers crimped with aluminum seals. When needed, kanamycin was used at 5 μg/mL. During aerobic growth, liquid cultures were shaken at 200 rpm on an orbital shaker to simulate aerobic conditions. During phototrophic (anaerobic) growth, liquid cultures were static.

For qPCR experiments, triplicate cultures of ORIO or BSL-9 were grown phototrophically in the presence of 10 mM acetate and varying arsenite concentrations for 12 hours, before RNA was extracted. BSL-9 was incubated at 35°C, using previously described methods (Hernandez-Maldonado *et al.*, 2016).

For mutagenesis and sucrose counter-selection, media was prepared with 10% ultrapure sucrose (MP Brand), and the entire process with done aerobically in the dark. Mutagenesis constructs were generated by splicing the vector pSMV20 (Zargar, *et al.,* 2010) with overlap extension PCR fragments containing 1000bp upstream and downstream fragments of the gene of interest (primers used can be found in the supplemental file Primers_SuppInfo.xlsx). The pSMV20 vector contains an origin of replication which is inoperative outside of *E. coli.* These gene deletion constructs were transformed into *E. coli* WM3064 (a strain for transferring mutations plasmids

to a recipient). Mating reactions with the *E. coli* donor strain and wild-type ORIO were mixed in a 1:1 ratio, centrifuged, spotted onto LB containing 10 mM succinate and 0.3 mM DAP, and then incubated at 30 °C aerobically in the dark overnight. The mating reactions were resuspended in 500 μL PNS and grown aerobically in the dark on PNS agar plates containing 10 mM succinate and 5 μg/mL kanamycin. The resulting colonies were spread onto peptone-yeast-succinate (PYS) agar media which contained: 1 g/L peptone, 1 g/L yeast, 2 g/L succinate, 10 % sucrose, 0.00001% $FeSO_4 \bullet 7H_2O$, 1 mL SL-10 solution, 1 mL Wolf's Vitamin mix, and pH 7.0. Colonies from the PYS-10% sucrose plates were screened for gene deletion via PCR using primers targeting a 109 bp region spanning the active site of *arxA*.

Plasmids and primer sets

Strains and plasmids used for this study are listed in supplementary file Strains_SuppInfo.xlsx. Primer composition and annealing temperatures are listed in supplementary file Primers_SuppInfo.xlsx. For PCR of ORIO DNA, Herculase II Fusion polymerase (Agilent) was routinely used with 8% DMSO per the manufacturer's protocol for DNA with high G+C.

Arx Protein Structure Prediction and Analysis

Arx protein sequences were used as inputs to predict the 3D protein structure via a Google Colaboratory notebook version of AlphaFold called ColabFold, using the multimer option (Jumper *et al.*, 2021; Evans *et al.*, 2021; Mirdita *et al.*, 2021). The predicted protein structure was analyzed for ligand binding with the COFACTOR2 and COACH web tools (Ambrish *et al.*, 2012; Yang *et al.*, 2013; Yang *et al.*, 2013;

Zhang *et al.*, 2017). The ArxA structure was aligned to ArrA using TM-align (Zhang *et al*. 2005). Structure predictions and ligands were overlayed, and measurements between ligands were made, using ChimeraX v1.3 (Pettersen *et al.*, 2021).

Expression analysis by quantitative RT-PCR

Total RNA was harvested using the manufacturer's protocol (Qiagen RNeasy Kit). 500 ng of RNA was first checked for quality on a 1.2% agarose gel, then treated with DNase I (Promega RQ1), and finally reverse transcribed using the TaqMan kit (Applied Biosystems) following the manufacturers protocols, respectively. For ORIO, complementary DNA (cDNA) was used for qPCR with the GoTaq 2X Master Mix (Promega; SYBR Green I dye) following the manufacture's protocol. ORIO primers targeting a 109 bp region spanning the active site of *arxA* were used in combination with 515-F and 926-R primers targeting the 16S rRNA gene. For BSL-9, the PrimeTime Mini qPCR Assay (Integrated DNA Technologies; FAM dye) probes targeting a 100 bp region of *arxA* and 16S rRNA genes were used in combination with a GoTaq 2X master mix (Promega). Primer information can be found in Primers_SuppInfo.xls. The expression was first normalized using the ΔCq method, where the target gene (*arxA*) is normalized to a reference gene (16S), and then the expression is calculated using the equation: ΔCq Expression = $2^{-\Delta Cq}$.

Cell suspension arsenite oxidation assay

Single colonies of ORIO wild-type and ORIO *arxA* mutant were picked and grown phototrophically overnight in 3 mL cultures in PNS media containing 10 mM succinate. 25 mL of fresh PNS containing 100 μM acetate was inoculated 1/10 with

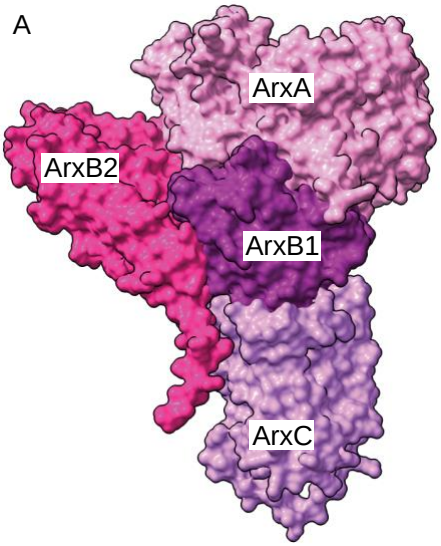
overnight cultures and allowed to grow phototrophically for 15 hours. After 15 hours, cultures were aseptically spiked with 100 μM arsenite and incubated for 6 hours. Samples for arsenic speciation were taken immediately before the spike, then 1 and 6 hours after the spike.

Arsenic Speciation Analysis – HPLC-ICP-MS

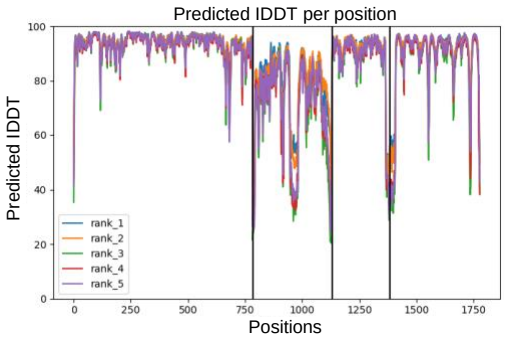
All samples were filtered (0.2 μm nylon membrane) and diluted 1/10 in sterile water. Arsenic speciation was determined and quantified using tandem High-Performance Liquid Chromatography Inductively Coupled Plasma Mass Spectroscopy (HPLC-ICP-MS) (described in Hernandez-Maldonado *et al.,* 2016). An anion exchange column (Hamilton PRP-X-100, particle size 10 μm, size 4.1 by 50 mm) was used with a 2 min/sample run-time at an ambient temperature. Elemental arsenic was detected using the Thermo X-Series2 ICP-MS. The isocratic mobile phase contained 30 mM phosphoric acid and MilliQ water was used in the autosampler flush. The flow rate was 1 min/mL with a 100 μL sample injection loop. Standard curves with different arsenite and arsenate standards were generated and used to calculate the arsenite and arsenate concentrations of various samples.

Section 4.05 Figures

A

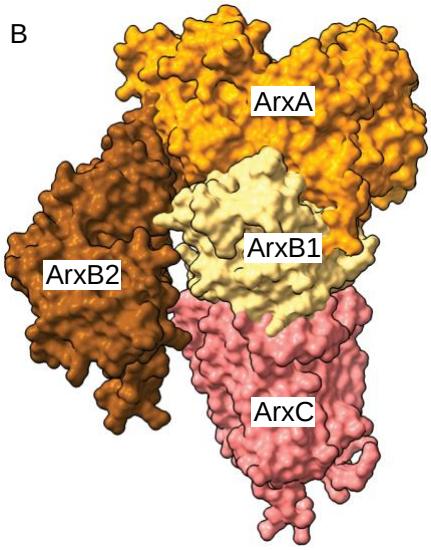


Cereibacter azotoformans str. ORIO

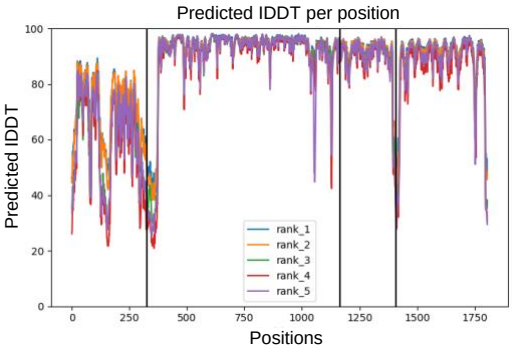


ArxA ArxB2 ArxB1 ArxC

B

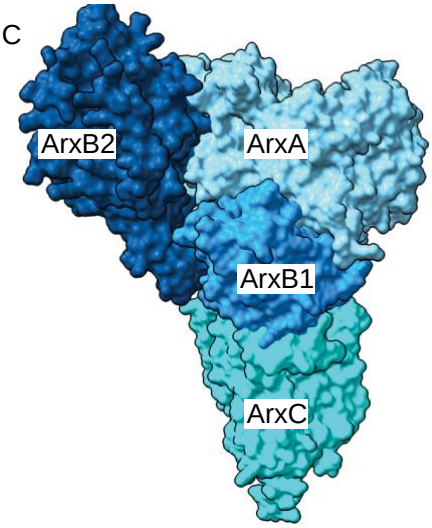


Ectothiorhodospira sp. str. BSL-9

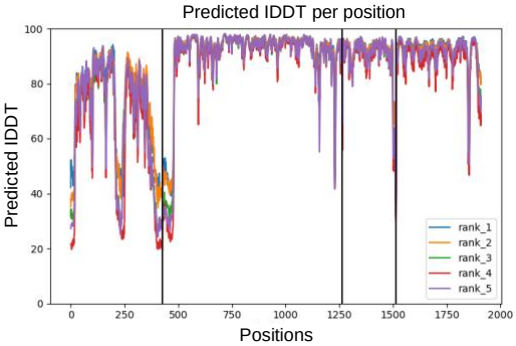


ArxB2 ArxA ArxB1 ArxC

C



Alkalilimnicola ehrlichii str. MLHE-1



ArxB2 ArxA ArxB1 ArxC

Figure 1. Prediction of the *Cereibacter azotoformans* str. ORIO Arx protein complex, and comparisons to previously isolated *Ectothiorhodospiraceae* members. The predicted ArxB2AB1C structure is shown with the predicted IDDT per position chart for (A) *Cereibacter azotoformans* str. ORIO, (B) *Ectothiorhodospira sp.* str. BSL-9, and (C) *Alkalilimnicola ehrlichii* str. MLHE-1. The predicted IDDT charts show the confidence scores for each amino acid position of each subunit, for the five models AlphaFold generates for each query sequence.

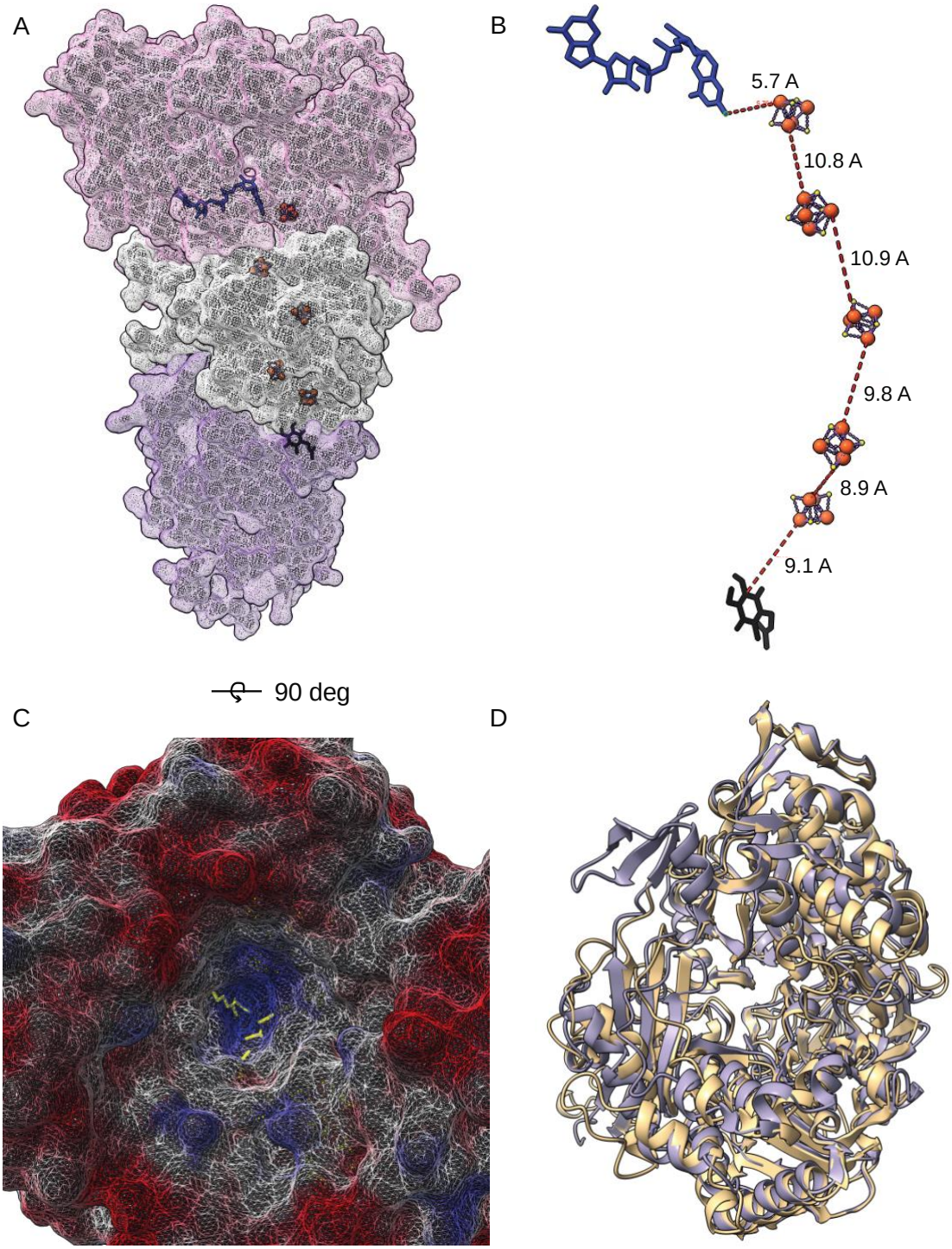


Figure 2. Structural analysis of the predicted ORIO ArxAB1C 3D structure. (A) The predicted ArxAB1C structure is shown with bound ligands molybdenum guanine

dinucleotide cofactor (blue), 4Fe-4S clusters (yellow/orange), and ubiquinone (black). (B) The electron path in ArxAB1C is shown with distances (in angstroms, A) between each predicted ligand. (C) A 90-degree rotated and magnified view from the top of ArxA, showing the electrostatic residues of the binding pocket and cofactors (yellow) in the active site. (D) Alignment of ORIO ArxA (grey) and *Shewanella* ANA-3 ArrA (tan) proteins, showing a view of the active site (as in panel C).

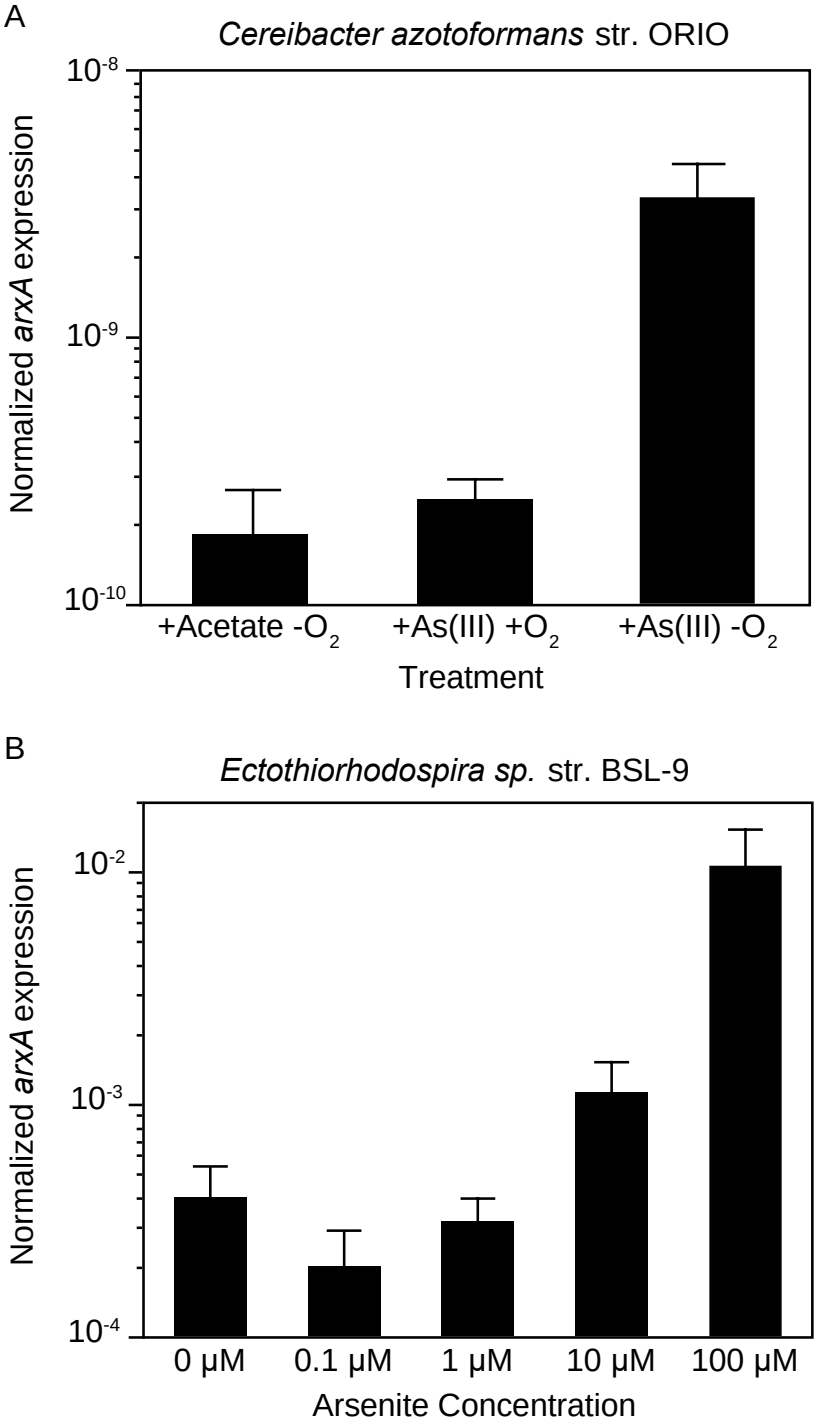


Figure 3. Quantitative RT-PCR analysis of *arxA* transcription in ORIO and BSL-9 during various growth conditions. (A) Quantitative RT-PCR analysis of

arxA in ORIO is shown as ΔCq expression versus growth conditions. ORIO was grown with 100 μM As(III) anaerobically, 100 μM As(III) aerobically, or 100 μM acetate anaerobically. Prior to reverse transcription, DNase I-treated RNA was tested for DNA contamination by PCR and gel electrophoresis. (B) Quantitative RT-PCR analysis of *arxA* in BSL-9 is shown as ΔCq expression versus growth conditions Expression of *arxA* was normalized to the 16S rRNA gene. Error bars show one standard error from the mean. Statistical significance was determined using Student's t-test.

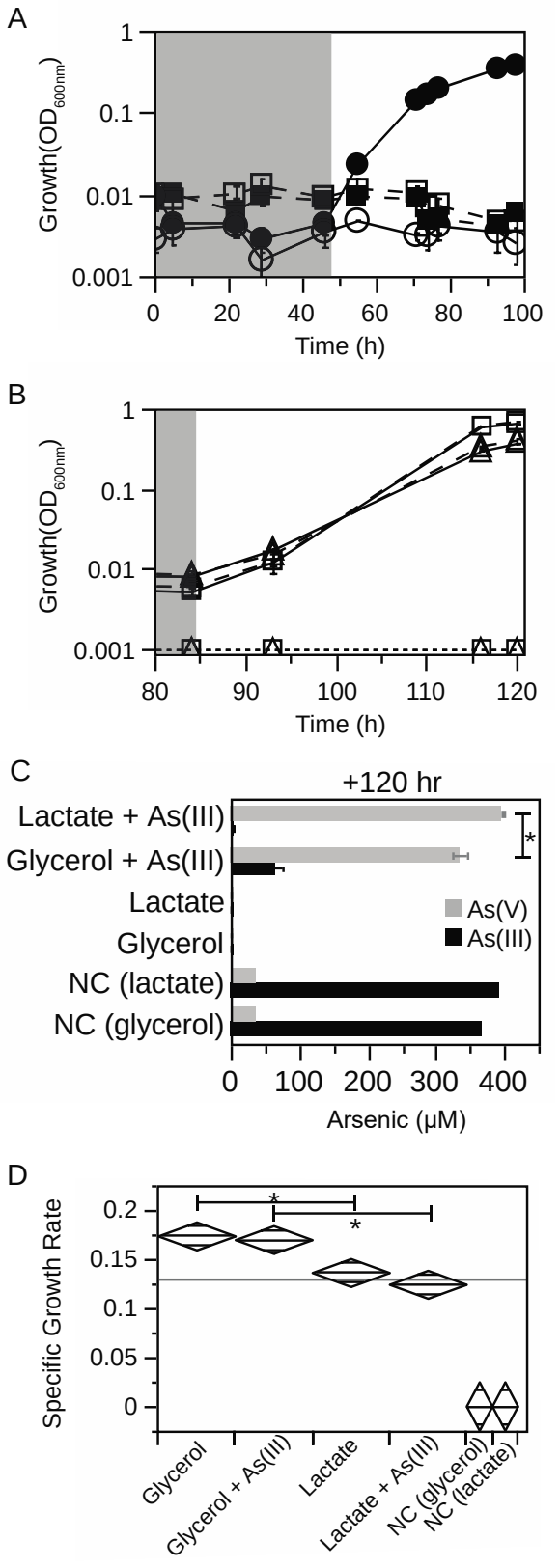

A
Growth(OD600nm)
1
0.1
0.01
0.001
0
20
40
60
80
100
Time (h)
B
Growth(OD600nm)
1
0.1
0.01
0.001
80
90
100
110
120
Time (h)
C
+120 hr
Lactate + As(III)
Glycerol + As(III)
Lactate
Glycerol
NC (lactate)
NC (glycerol)
As(V)
As(III)
*
0
100
200
300
400
Arsenic (μM)
D
Specific Growth Rate
0.2
0.15
0.1
0.05
0
*
*
Glycerol
Glycerol + As(III)
Lactate
Lactate + As(III)
NC (glycerol)
NC (lactate)

Figure 4. The effect of light and carbon source on photoarsenotrophy in ORIO. Anaerobically grown wild-type ORIO is shown as growth (OD_{600nm}) versus time (h), with grey areas indicating the infrared light source was off. Error bars are shown as one standard error from the mean. (A) wild-type ORIO growth versus time with 1 mM As(III) (filled symbols) or 0 mM As(III) (empty symbols), and organic (circles and solid lines) or inorganic (squares and dashed lines) carbon sources. (B) wild-type ORIO growth versus time with glycerol (square) or lactate (triangle), and 0.5 mM As(III) (solid lines) or 0 mM As(III) (dashed lines), controls were lacking cells (dotted lines). (C) Arsenic speciation for panel B is shown as treatment versus arsenite (black) and arsenate (grey) concentrations (μM), at the end of the experiment (hour 120). (D) The specific growth rate (SPG) for treatments in panels B and C were calculated using Absorbance (OD_{600nm}) during the exponential phase, using the equation $SPG = Ln(Absorbance_2 – Absorbance_1)/T_2 – T_1$. Statistical analysis was performed using the ANOVA and Student t-test methods. Negative controls (NC) lacked cells.

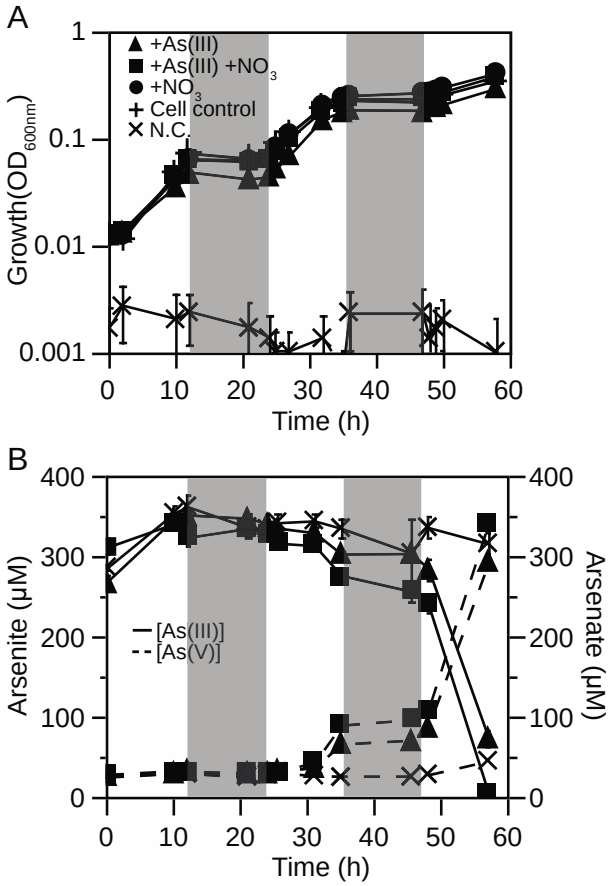


Figure 5. The effect of light and nitrate on photoarsenotrophy in ORIO. wild-type ORIO was grown anaerobically in mPNS with varying treatments. The infrared light source was on a light-dark cycle of 12 h:12 h (grey areas indicate infrared was off). Error bars show one standard error from the mean. (A) Growth of wild-type ORIO is shown as optical density (600_{nm}) versus time (h). Cell controls contained wild-type ORIO but no arsenite or nitrate. N.C. are negative controls lacking cells. (B) Arsenic speciation accompanying the data shown in panel A is shown as arsenic concentration (μM) versus time (h), with solid lines for arsenite and dashed lines for arsenate. Cultures lacking arsenite are not shown.

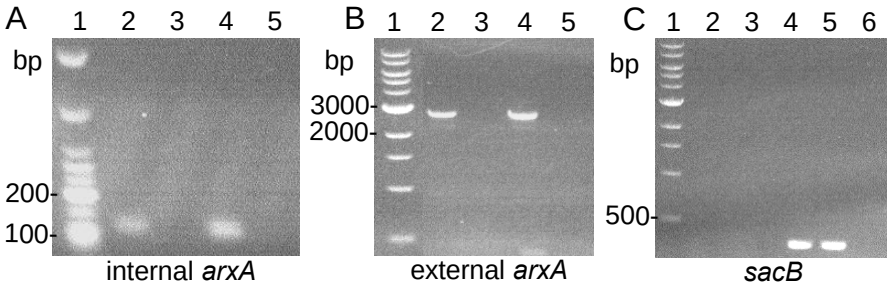


Figure 6. Mutagenesis of *arxA* in *Cereibacter azotoformans* str. ORIO. (A) 1% agarose gel of PCR reactions targeting the active site of *arxA.* Lanes correspond to 1=low molecular weight ladder, 2=wild-type ORIO DNA, 3=△*arxA* ORIO DNA, 4=positive control, 5=negative control. (B) 1% agarose gel of PCR reactions targeting the entire *arxA* gene. Lanes correspond to 1=1 kb ladder, 2=wild-type ORIO DNA, 3=△*arxA* ORIO DNA, 4=positive control, 5=negative control. (C) 1% agarose gel image of PCR reactions targeting the *sacB* gene of the mutagenesis plasmid. Lanes correspond to 1=low molecular weight ladder, 2=wild-type ORIO DNA, 3=△*arxA* ORIO DNA, 4=positive control (empty plasmid), 5=positive control (mutagenesis plasmid), 6=negative control.

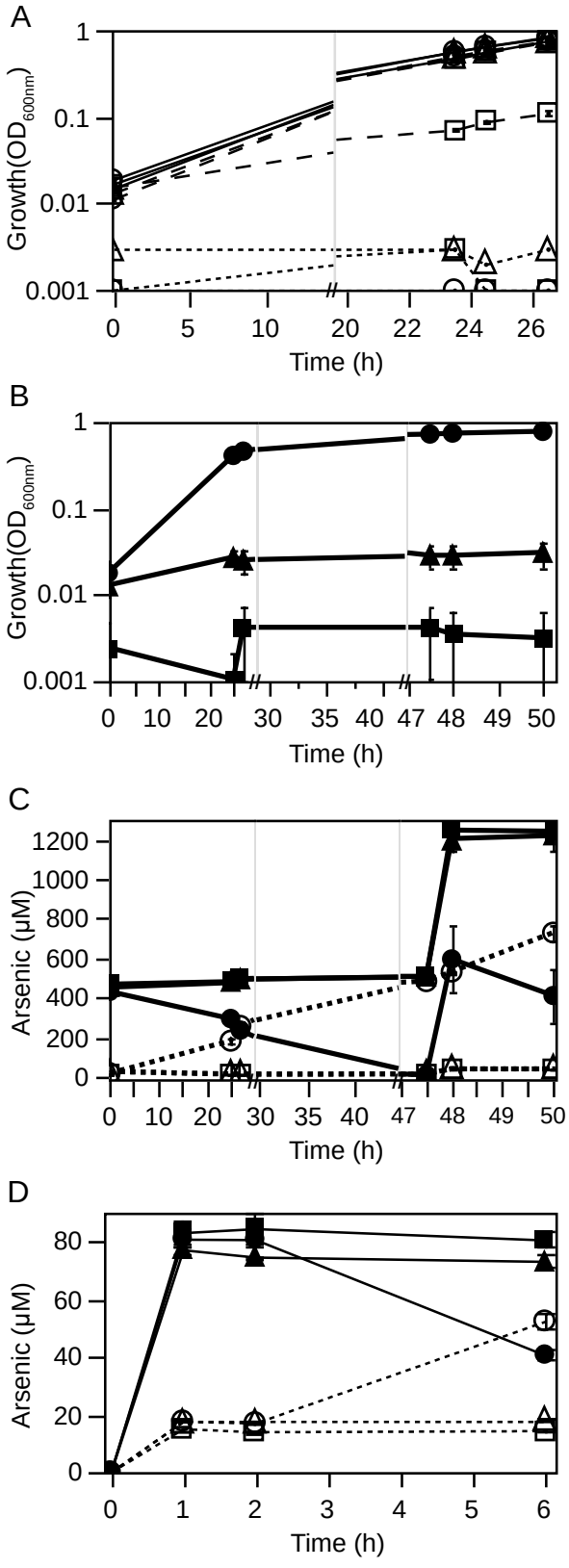

A
Growth(OD600nm)
1
0.1
0.01
0.001
0
5
10
20
22
24
26
Time (h)
B
Growth(OD600nm)
1
0.1
0.01
0.001
0
10
20
30
35
40
47
48
49
50
Time (h)
C
Arsenic (μM)
1200
1000
800
600
400
200
0
0
10
20
30
35
40
47
48
49
50
Time (h)
D
Arsenic (μM)
80
60
40
20
0
0
1
2
3
4
5
6
Time (h)

Figure 7. The role of *arxA* in photoarsenotrophy in ORIO. (A) The role of *arxA* in arsenite resistance is shown as growth (OD_{600nm}) versus time (h). wild-type ORIO (solid lines) and Δ*arxA* (dashed lines) were grown phototrophically in PNS with 10 mM succinate and either (○) 0 mM, (△) 0.1 mM, or (□) 0.5 mM arsenite. Cultures lacking cells (dotted lines) were used as the negative controls. (B) The role of *arxA* in growth on arsenite is shown as cell density (OD_{600nm}) over time (h). wild-type ORIO (●) and Δ*arxA* (▲) were grown phototrophically in PNS with 0.5 mM arsenite, and negative controls (■). (C) Arsenite (closed symbols) and arsenate (open symbols) concentrations are shown as arsenite (μM) versus time (h) for wild-type ORIO (●), Δ*arxA* (▲), and negative control (■). Samples were spiked with an additional 0.5 mM arsenite at hour 47.5. (D). High density cell suspensions require *arxA* for arsenite oxidation. wild-type and △*arxA* cultures were grown overnight phototrophically in PNS containing 100 nM acetate. After 15 hours, a 100 μM spike of As(III) was added. Breaks in the x-axis are shown with //.

Section 4.06 References

Chen, S. C. *et al.,* (2020) The Great Oxidation Event expanded the genetic repertoire of arsenic metabolism and cycling. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10414–10421.

Durante-Rodríguez, G., Fernández-Llamosas, H., Alonso-Fernandes, E., Fernández-Muñiz, M.N., Muñoz-Olivas, R., Díaz, E., & Carmona, M. (2019) ArxA From Azoarcus sp. CIB, an Anaerobic Arsenite Oxidase From an Obligate Heterotrophic and Mesophilic Bacterium. Front. Microbiol. 10: 1699.

Ellis, P.J., Conrads, T., Hille, R., and Kuhn, P. (2001) Crystal structure of the 100 kDa arsenite oxidase from Alcaligenes faecalis in two crystal forms at 1.64 angstrom and 2.03 angstrom. Structure 9: 125–132.

Evans et al. (2021). Protein complex prediction with AlphaFold-Multimer. Biorxiv. doi: 10.1101/2021.10.04.463034v1

Glasser, N. R., Oyala, P. H., Osborne, T. H., Santini, J. M., & Newman, D. K. (2018) Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations. Proc Natl Acad Sci USA 115(37): E8614–E8623.

Hamamura, N., Itai, T., Liu, Y., Reysenbach, A. L., Damdinsuren, N., and Inskeep, W. P. (2014) Identification of anaerobic arsenite-oxidizing and arsenate-reducing bacteria associated with an alkaline saline lake in Khovsgol, Mongolia. Environ. Microbiol. Rep. 6: 476-482.

Hernandez-Maldonado, J., Stoneburner, B., Boren, A., Miller, L., Rosen, M., Oremland, R. S., *et al.*, (2016) Genome sequence of the photoarsenotrophic bacterium Ectothiorhodospira sp. strain BSL-9, isolated from a hypersaline alkaline arsenic-rich extreme environment. Genome Announc. 4: 9–10.

Hernández-Maldonado J., Sanchez-Sedillo B., Stoneburner B., Boren A., Miller L., McCann S., *et al.*, (2016) The genetic basis of anoxygenic photosynthetic arsenite oxidation. Environ. Microbiol. 19: 130–141.

Hoeft McCann, S., Boren, A., Hernandez-Maldonado, J., Stoneburner, B., Saltikov, C., Stolz, J., *et al.*, (2017) Arsenite as an electron donor for anoxygenic photosynthesis: description of three strains of Ectothiorhodospira from Mono Lake, California and Big Soda Lake, Nevada. Life 7: 1–14.

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

Inskeep, W. P., Macur, R. E., Hamamura, N., Warelow, T. P., Ward, S. A., & Santini, J. M. (2007). Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. Environmental microbiology, 9(4), 934-943.

Jaschke, P.R., Saer, R.G., Noll, S., and Beatty, J.T. (2011) Modification of the Genome of Rhodobacter sphaeroides and Construction of Synthetic Operons. Methods in Enzymology 497: 519-538.

Jiao, Y., & Newman, D. K. (2007). The pio operon is essential for phototrophic Fe (II) oxidation in Rhodopseudomonas palustris TIE-1. Journal of bacteriology, 189(5), 1765-1773.

Jumper, J., Evans, R., Pritzel, A. *et al.* (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596,** 583–589.

Kulp, T.R., Hoeft, S.E., Asao, M., Madigan, M.T., Hollibaugh, J.T., Fisher, J.C., *et al.*, (2008) Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. Science 321: 967–970.

Link AJ, Phillips D, Church GM. Methods for generating precise deletions and insertions in the genome of wild-type Escherichia coli: application to open reading frame characterization. J Bacteriol. 1997;179(20):6228-6237. doi:10.1128/jb.179.20.6228-6237.1997

Käll, L., Krogh, A., and Sonnhammer, E.L.L. (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server Nucleic Acids Res., 35:W429-32

Madeira F., Park Y.M., Lee J., *et al.*, (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research (W1):W636-W641.

Meng YL, Liu Z, Rosen BP. (2004) As(III) and Sb(III) uptake by GlpF and efflux by ArsB in Escherichia coli. J Biol Chem. 279:18334–18341.

Mirdita, M., Steinegger, M., Söding, J., (2019) MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics 35(16):2856–2858

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S and Steinegger M. (2021) ColabFold - Making protein folding accessible to all. bioRxiv. doi: 10.1101/2021.08.15.456425

Mougiakos, I., Orsi, E., Ghiffary, M. R., Post, W., de Maria, A., Adiego-Perez, B., ... & van der Oost, J. (2019). Efficient Cas9-based genome editing of Rhodobacter sphaeroides for metabolic engineering. Microbial cell factories, 18(1), 1-13.

Oremland, R. S., & Stolz, J. F. (2003). The Ecology of Arsenic. Science, 300(9), 939–944.

Ospino MC, Kojima H and Fukui M (2019) Arsenite Oxidation by a Newly Isolated Betaproteobacterium Possessing arx Genes and Diversity of the arx Gene Cluster in Bacterial Genomes. Front. Microbiol. 10: 1210.

Poulain, A. J., & Newman, D. K. (2009). Rhodobacter capsulatus catalyzes light-dependent Fe (II) oxidation under anaerobic conditions as a potential detoxification mechanism. Applied and environmental microbiology, 75(21), 6639-6646.

Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. 2021 Jan;30(1):70-82.

Roy, A., Yang, J., & Zhang, Y. (2012) COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research 40:W471-W477.

Saltikov, C.W., & Newman, D. K. (2003) Genetic identification of a respiratory arsenate reductase. Proc Natl Acad Sci USA 100(19): 10983–10988.

Saltikov, C.W. (2011). “Regulation of arsenic metabolic pathways in prokaryotes” in Microbial metal and metalloid metabolism. eds. J. Stolz and R. Oremland (Washington DC: ASM Press), 195–210.

Shaji, E., Santosh, M., Sarath, K. V., Prakash, P., Deepchand, V., & Divya, B. V. (2020) Arsenic contamination of groundwater: A global synopsis with focus on the Indian Peninsula. Geoscience Frontiers 12(3).

Yang, J., Roy, A., Zhang, Y., (2013) BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. Nucleic Acids Research 41: D1096-D1103

Yang, J., Roy, A., Zhang, Y., (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29: 2588-2595.

Zargar, K., Conrad, A., Bernick, D.L., Lowe, T.M., Stolc, V., Hoeft, S., Oremland, R.S., Stolz, J., Saltikov, C.W., (2012) ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. Environ Microbiol. 14(7):1635-45.

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

Zhang, Y., Skolnick, J., (2005). TM-align: A protein structure alignment algorithm based on TM-score, Nucleic Acids Research, 33: 2302-2309.

Zhang, C., Freddolino, P.L., and Zhang, Y. (2017) COFACTOR: Improved protein function prediction by combining structure, sequence, and protein-protein interaction information. Nucleic Acids Research 45: W291-299.

Section 4.07 Acknowledgments

We thank Brian Dryer and the UCSC Marine Analytical Lab for supporting the arsenic speciation analyses.

Chapter 5. Conclusion and future directions

Section 5.01 Overarching goals

The research I have presented here demonstrates the occurrence, potential mechanisms, and ecology of photoarsenotrophy in freshwater sediments. Through a combination of experiments using a laboratory cultivated strain, microcosm studies with environmental samples, and *in silico* analyses, we have established the possibility of photoarsenotrophy and arsenic redox cycling in freshwater sediment environments.

Section 5.02 Isolation of a freshwater photoarsenotroph and metagenomics

Through the experiments presented in Chapter 3, we have begun to characterize a unique branch of the arsenic biogeochemical cycle where arsenic concentrations are influenced by light-dark cycling. This was accomplished by first isolating a novel freshwater heterotrophic photoarsenotroph, *Cereibacter azotoformans* str. ORIO. We successfully sequenced the genome, using a hybrid approach, found the *arx* gene cluster, and detected arsenite oxidation along with *arxA* transcription in cultures grown with arsenite. Then, our work showed ORIO was isolated from sediment that is capable of light-dark arsenic redox cycling. Previous reports in freshwater environments have detected cyclic changes in arsenic concentrations but did not investigate the biological component enough to implicate bacteria or their genes are drivers of this activity. Metagenomic analysis of light-dark cycled sediment microcosms provides insight into the likely bacterial taxa and genes associated with arsenite oxidation and arsenate reduction. Lastly, *in silico* analysis of over 35,000

metagenomes on JGI/IMG demonstrated the *arx* gene cluster can be detected in metagenomes from samples on all seven continents and major environment types (aquatic, sediment, wastewater, soil, freshwater, and salt water). Further expanding our knowledge of the occurrence of *arx* genes in nature.

Future studies on light-dark arsenic redox cycling in the environment are required to better understand the arsenic biogeochemical cycle. Although culture dependent experiments are essential to research, future studies should aim to utilize culture independent methods of DNA and RNA sequencing in the field. Metagenomic and RNA sequencing experiments are needed to better understand the key components in light-dark arsenic redox cycling. It is unclear which bacteria are participating in light-dark arsenic cycling, and transcriptional analysis of arsenic genes in the environment would help determine this. Furthermore, it is important to determine if light-dark arsenic cycling affects uptake of arsenic into plants, especially rice. We hypothesize that an arsenic pool being utilized by arsenite-oxidizers and arsenate-reducers may restrict availability to other organisms, such as plants.

Section 5.03 Characterization of photoarsenotrophy in a novel freshwater bacterium

To build on our work isolating freshwater photoarsenotrophs and detecting light-dark cycling of arsenic in Owens River sediment, in Chapter 4 we characterized photoarsenotrophy in ORIO and showed this process is dependent on light and organic carbon. We found that when ORIO is grown in a 12:12 hour light-dark cycle, arsenite oxidation occurs only during the light phase. This is consistent with the first

study on so-called “diurnal arsenic cycling” where Fuller and Davis (1989) detected arsenate concentrations increased during the day and decreased at night. Our results demonstrate ORIO cannot couple nitrate reduction to arsenite-oxidation in the dark, which implicates photoarsenotrophy as a mechanism for previously observed light-dark arsenic cycles. In addition, the presence of nitrate did not repress arsenite oxidation in ORIO. We then showed that *arxA* is responsible for arsenite oxidation in ORIO by genetically engineering a strain with an in-frame null deletion of *arxA*. As expected, our Δ*arxA* mutant lost the ability to oxidize arsenite. To our surprise however, *arxA* deletion resulted in significantly increased sensitivity to arsenite under anaerobic conditions. This finding suggested a light-dependent resistance mechanism to arsenite, which to our knowledge has not been reported for arsenic. Previously, it has been considered typical for heterotrophic bacteria to not gain an energy advantage from arsenic (Silver and Phung, 2005). In a recent report however, a heterotrophic bacterium that uses *arx* genes to both resist arsenite and generate a slight increase in energy generation was described (Durante-Rodriguez *et al.*, 2019). Our findings add to the abilities of heterotrophic arsenite-oxidizing bacteria.

Since no crystal structures for Arx proteins exist, we investigated the Arx protein structures using the latest advancement in protein prediction and provide a new view of the Arx enzyme complex. Comparison of predicted Arx protein complexes in ORIO, BSL-9, and MLHE-1 surprised us. In each case, the ArxB2 subunit did not interface with ArxAB1C in such a way that appeared would facilitate electron transfer. Based on measurements in the ORIO model, ArxB2 was too far from the

electron transfer chain to participate. Therefore, we hypothesize ArxB2 in ORIO does not participate in arsenite-oxidation and future studies should aim to characterize the role of all *arx* genes. Further studies on the electron flow through Arx proteins and the electron transfer chain will help identify crucial information regarding why this process is light-dependent in ORIO and not in other Arx-type arsenotrophs. Our structural analysis also provides insight into the similarities and differences between the predicted ArxA structure and the crystal structure of ArrA. Future site-directed mutagenesis studies should aim to determine which amino acid residues give rise to the enzymatic directionality of ArxA and ArrA, respectively.

Section 5.04 Future directions

Looking forward at the next 100 years of research on microbially mediated arsenic transformations, we can expect major progress toward our understanding of the arsenic biogeochemical cycle. The advancement of technologies such as nanopore sequencing will enable research to be carried out quicker, on-demand, and even in remote parts of the world. To be able to detect and understand which organisms and genes are present in a system is an essential part of monitoring efforts, which can enable better decision making when it comes to prevention and mitigation of arsenic contamination. Other advancements in technologies such as CRISPR will allow for faster genetic manipulations of microbes of interest, thereby allowing faster investigations of genes of interest. Genetically optimized strains can bring forth a new era of biological tools for agriculture that reduce arsenic contamination, increase

yield, and protect against crop diseases. Onward with the next 100 years of microbially-mediated arsenic transformation!

Section 5.05 References

Durante-Rodríguez, G., Fernández-Llamosas, H., Alonso-Fernandes, E., Fernández-Muñiz, M.N., Muñoz-Olivas, R., Díaz, E., & Carmona, M. (2019) ArxA From Azoarcus sp. CIB, an Anaerobic Arsenite Oxidase From an Obligate Heterotrophic and Mesophilic Bacterium. Front. Microbiol. 10: 1699.

Fuller, C. C., & Davis, J. A. (1989). and Photosynthetic Cycles in Natural Waters. Nature, 340(JULY), 1987–1989.

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

Appendix A

The experiments shown in Figure 1 and Figure 2 are preliminary work on the regulation of photoarsenotrophy in *Ectothiorhodospira sp*. str. BSL-9. To our knowledge, it is the first time *arx* regulatory genes have been shown to be involved in arsenite oxidation. Figure 3 is a continuation of the Arx protein analysis from Chapter 4 (Figures 1 and 2).

Figure 1. PCR verification of *arxR* gene disruption in *Ectothiorhodospira sp.* str. BSL-9. (A) Gel electrophoresis with a 1 % agarose gel was used to detect DNA amplification in wild-type BSL-9 and the *arxR*-disrupted mutant, using primers that target the entire *arxR* gene (1.3 kb product; See Strains_SuppInfo.xlsx and Primers_SuppInfo.xlsx). Lanes correspond to (1) 1 kb ladder, (2-5) disruption mutants, (6) water (negative control), (7) wild-type BSL-9 DNA (positive control). Disruption of *arxR* was achieved using the technique previously described by Hernandez-Maldonado *et al.*, 2016.

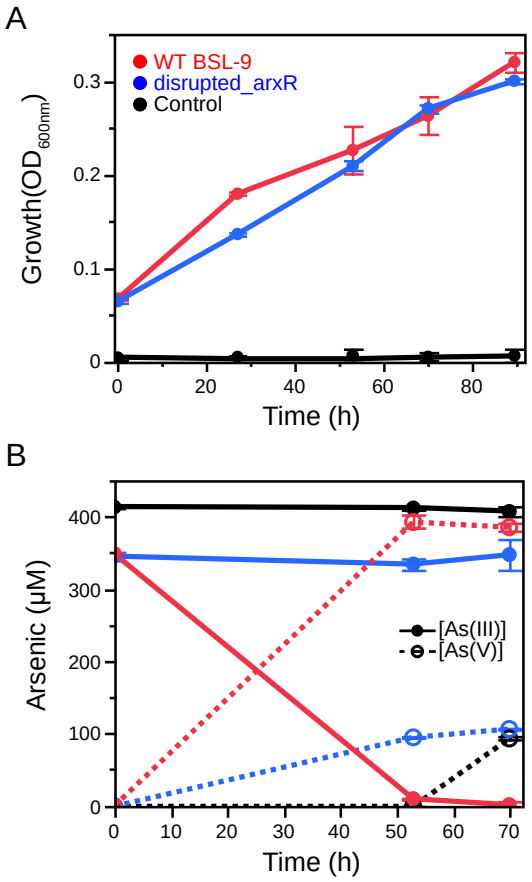


Figure 2. BSL-9 *arxR* disruption alters arsenite oxidation and not fitness. (A) Fitness of the disruption mutant was tested by measuring optical density (OD_{600nm}) over time (hours) of wild-type BSL-9 (red) and an *arxR* disruption mutant (blue) grown in BSM broth containing 5 mM acetate. Negative control is shown as "Control" (black) and contained media but no cells. (B) A separate arsenite oxidation assay is shown as arsenic concentrations (μM) versus time (hours) for wild-type BSL-9 (red), *arxA* disruption mutant (blue), and control (black). BSM containing 500 μM arsenite was spiked with washed cells and arsenic speciation was measured using HPLC-ICP-MS. Error bars show one standard error from the mean.

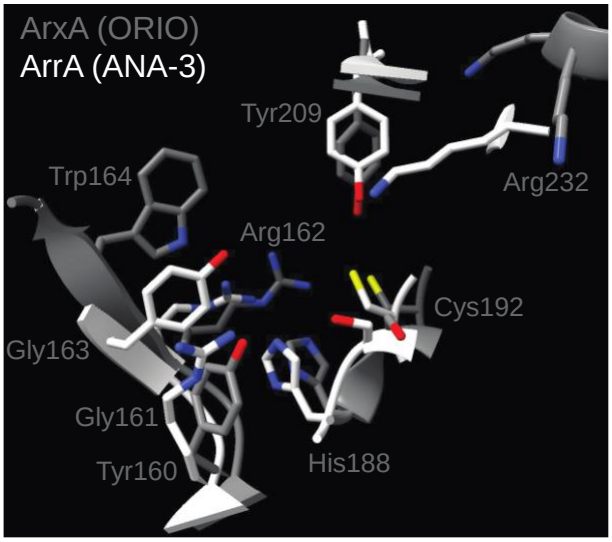


Figure 3. ArxA and ArrA active site alignment. The active sites of the AlphaFold predicted ArxA from *Cereibacter azotoformans* str. ORIO and ArrA crystal structure bound to arsenite from *Shewanella sp*. ANA-3 (PDB: 6CZ9) were aligned using TM-align. The alignment file was processed using ChimeraX v1.3 to only show the residues hypothesized to be relevant to the binding of arsenite and coordinating the molybdenum ion. ArxA is shown in dark grey color, and ArrA is shown in light grey color.

Appendix A references

Hernandez-Maldonado, J., Sanchez-Sedillo, B., Stoneburner, B., Boren, A., Miller, L., McCann, S., … Saltikov, C. (2016). The genetic basis of anoxygenic photosynthetic arsenite oxidation. Environmental Microbiology, 0, 1–29.

Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. (1):70-82.

Zhang, Y., Skolnick, J., (2005). TM-align: A protein structure alignment algorithm based on TM-score, Nucleic Acids Research, 33: 2302-2309.