Linking Genes to Microbial Biogeochemical Cycling Lessons from Arsenic


Published in:
Environmental science & technology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright © 2017 American Chemical Society. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Linking genes to microbial biogeochemical cycling: lessons from arsenic

Yong-Guan Zhu, †,‡,#,* Xi-Mei Xue, †,# Andreas Kappler, † Barry P Rosen, § Andrew A Meharg ‴

† Key Laboratory of Urban Environment and Health, Institute of Urban Environment,
Chinese Academy of Sciences, Xiamen, China
‡ State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China
# Geomicrobiology, Center for Applied Geosciences, University of Tübingen,
Tübingen 72076, Germany
§ Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199, U.S.A.
″ Institute for Global Food Security, Queen’s University Belfast, Belfast BT9 5HN,
United Kingdom

# These authors contributed equally to the article.

* Address correspondence to Yong-Guan Zhu

Address: Institute of Urban Environment, Chinese Academy of Sciences, 1799 Jimei Road, Xiamen 361021, China, Phone number: +86(0)592 6190997, Fax number: +86(0)592 6190977, Email address: ygzhu@rcees.ac.cn
Abstract

The biotransformation of arsenic is highly relevant to the arsenic biogeochemical cycle. Identification of the molecular details of microbial pathways of arsenic biotransformation coupled with analyses of microbial communities by meta-omics can provide insights into detailed aspects of the complexities of this biocycle. Arsenic transformations couple to other biogeochemical cycles, and to the fate of both nutrients and other toxic environmental contaminants. Microbial redox metabolism of iron, carbon, sulfur and nitrogen affects the redox and bioavailability of arsenic species. In this critical review we illustrate the biogeochemical processes and genes involved in arsenic biotransformations. We discuss how current and future metagenomic-, metatranscriptomic-, metaproteomic-, and metabolomic-based methods will help to decipher individual microbial arsenic transformation processes, and their connections to other biogeochemical cycle. These insights will allow future use of microbial metabolic capabilities for new biotechnological solutions to environmental problems. To understand the complex nature of inorganic and organic arsenic species and the fate of environmental arsenic will require integrating systematic approaches with biogeochemical modeling. Finally, from the lessons learned from these studies of arsenic biogeochemistry, we will be able to predict how the environment changes arsenic, and, in response, how arsenic biotransformations change the environment.
1. Introduction

Biogeochemical cycles are interconnected through redox reactions and other biotransformations. Biogeochemical processes such as the cycling of a particular element are likely to be mediated by multiple microbes and are often linked to other biogeochemical processes. For example, redox changes of arsenic are mediated by diverse arsenate (As(V))-reducing and arsenite (As(III))-oxidizing microbes. Arsenic biogeochemical cycling is often coupled to the cycling of iron (Fe), carbon (C) and nitrogen (N), and to the dynamics of elements/ions associated with the arsenic redox cycle, such as sulfur (S). Coupling of biogeochemical cycles has recently received attention. The study of coupled biogeochemical cycles offers a scientific basis for major current environmental problems.

The proteins catalyzing physiological processes in living organisms are influenced by geological, physical and chemical forces and therefore continuously evolve and redistribute chemical species involved in biogeochemical cycles. Genetic analysis is the key to understand the arsenic biogeochemical cycle. Once the genes associated with the reactions and the environmental signals that affect gene expression are understood, we will be able to predict how microbial metabolism influences arsenic biogeochemical cycling. In this review, we focus on the known genes involved in arsenic biotransformations and the effect of other elements on arsenic biogeochemistry. We highlight the effects of other elements on arsenic metabolism and the current state of meta-omics research in microbial arsenic metabolism. Finally, we discuss how integration of meta-omics information into biogeochemical models can allow us to predict the possible biotransformation of other elements.
2. Arsenic metabolism: from genes to biogeochemical processes

Organisms have evolved various strategies to transform arsenic for detoxification or energy metabolism. An overview about the enzymatic pathways for arsenic biotransformations is presented in Figure 1, and related microbial genes are summarized in Table 1.

2.1 The arsenic redox cycle

The earliest microorganisms evolved in an anoxic environment, where the predominant arsenic species was most probably reduced As(III), with little oxidized As(V). The physiological activities of the earliest microorganisms were, therefore, largely driven by anaerobic metabolic processes, and we propose that As(III) bioavailability was a driving force for the evolution or acquisition of genes encoding anaerobic respiratory pathways. For example, the photosynthetic purple sulfur bacterium *Ectothiorhodospira* PHS-1 carries out anoxygenic photosynthesis using As(III) as an electron donor in the light and uses As(V) as an electron acceptor in the dark. The chemolithoautotrophic As(III)-oxidizer *Alkalilimnicola ehrlichii* MLHE-1 utilizes As(III) as an electron donor and nitrate as an electron acceptor in energy-generating respiratory chains. Microorganisms with similar metabolic versatility probably evolved quite early. These microbes could cope with extreme growth conditions, such as high concentrations of As(III) or low oxygen, similar to those that existed in an primordial anoxic biosphere.

Since As(III) was probably the primary bioavailable arsenic species on the early anoxic Earth, it was the inevitable choice for organisms to utilize As(III) as an electron donor to produce energy. As(III) oxidation is catalyzed by the enzyme As(III) oxidase, which is composed of two different subunits, a large subunit (α) having molybdopterin and a [3Fe-4S] cluster (AioA) and a smaller subunit (β)
incorporating a Rieske-type [2Fe-2S] cluster (AioB). The cluster of aioA and aioB genes (aio operon) usually consists of aioS and aioR genes, encoding for a two-component signal transduction pair, AioS (sensor histidine kinase)/AioR (transcriptional regulator), which regulates expression of aio genes via recognizing As(III). The operon sometimes has an aioX gene that encodes an As(III)-binding protein involved in As(III)-based signaling and regulation of As(III) oxidation, a cytC gene encoding a cytochrome c that is required for efficient As(III) oxidation in Ochrobactrum tritici SCII24, or a moeA gene encoding MoeA protein that synthesizes the molybdenum cofactor of AioAB oxidase. Recently, a new type of As(III) oxidase, ArxA that exhibited both As(V) reductase and As(III) oxidase activities in vitro, was identified in A. ehrlichii MLHE-1. In Ectothiorhodospira sp. PHS-1 these genes code for As(III) oxidation coupled to photosynthesis. In addition to arxA, the MLHE-1 and PHS-1 arx operons, contain four other genes arxB2, arxB, arxC, and arxD, that encode two proteins with [4Fe-4S] centers, a membrane anchoring and quinol oxidoreductase subunit and a TorD-like molybdoenzyme chaperone respectively. An adjacent and divergent gene cluster, arxSR, encodes putative regulatory proteins, a periplasmic substrate-binding protein specific for phosphate (ArxX), a two-component histidine kinase sensor (ArxS), and a response regulator (ArxR). ArxA has higher sequence similarity to the ArrA subunit than to AioA, and fills the phylogenetic gap between As(III) oxidases and As(V) reductases.

Note that As(III) oxidation by anaerobes would have produced As(V) in the absence of an oxygen-containing atmosphere, which opened a niche for As(V)-respiring microbes prior to the Great Oxidation Event (GOE). Dissimilatory As(V)-respiring prokaryotes (DARPs) evolved pathways to take advantage of the
appearance of As(V) as a terminal electron acceptor. This new energy-generating respiratory chain utilized the respiratory As(V) reductase, ArrAB, that reduce the less toxic As(V) to the more toxic and potentially more mobile As(III). ArrAB is a heterodimer consisting of a large catalytic subunit (ArrA) and a small subunit (ArrB). 

The arr operon also includes arrC, arrD, arrS, and arrR. Their gene products are ArrC, a membrane-bound As(V) reductase subunit, ArrD, a As(V) reductase chaperon, ArrS, a sensor histidine kinase and ArrR, a transcriptional regulator respectively. A phylogenetic analysis was conducted to search for molybdenum-bis (pyranopterin guanine dinucleotide)-containing catalytic subunits of representative enzymes. This complex iron sulfur molybdoenzyme family includes Arr, Aio, Arx, polysulfide reductase, and nitrate reductase. The results indicate that Arr clusters most likely evolved from polysulfide reductases.

After the GOE, As(III) in oceans mostly oxidized to As(V), a new environmental toxin. As(V) enters the cells of most organisms adventitiously via phosphate uptake systems. As a consequence, early life had to evolve novel strategies for coping with new (potentially toxic) arsenic species. As described in more detail below, nearly every extant microbe has ArsB or Acr3 efflux permeases for As(III) detoxification, so it is reasonable to assume that organisms that arose before the GOE already had an As(III) efflux system. When As(V) became the predominant soluble species, all cells had to do was to reduce As(V) to As(III), the substrate of ArsB or Acr3, and they would become resistant to As(V). A number of independently-evolved As(V) reductases arose in a variety of organisms using a small molecular mass protein As(V) reductases (one of several types of ArsC or Acr2 reductases). The ArsC system conferred by the ars operon is the most well studied mechanism of arsenic detoxification and resistance (for details see the previous review). Most recently, a
glutathione S-transferase B (GstB) was found to mediate an alternate pathway which conferred As(V) resistance to *E. coli* mutant cells lacking *arsC* by directly reducing As(V) to As(III). These enzymes all use small molecule proteins such as glutaredoxin (Grx) or reduced glutathione (GSH) coupled to thioredoxin (Trx) as electron donor. The Acr2 reductases evolved from proteins that incorporated the phosphate binding loop of phosphoprotein tyrosine phosphatases related to the cell phosphatase CDC25. These phosphatase can be converted into As(V) reductases by just a few mutations, indicating a facile evolutionary path.

### 2.2 The arsenic methylation cycle

In addition to oxidation and reduction of inorganic arsenic species, pathways for biotransformation of arsenic, including methylation and demethylation, organoarsenical degradation, evolved in early organisms. Interest in arsenic biomethylation began in 1800’s with the observation that inorganic arsenic compounds used as wallpaper pigments were converted into Gosio gas (trimethylarsine) by fungi. More recent reports of methylated arsenical showed that arsenic methylation was widespread in the environment and detected in bacteria, cyanobacteria, algae, and protozoa. Arsenic methylation is a common stratagem to detoxify arsenic. The highly toxic trivalent products are rapidly oxidized nonenzymatically in air to the less toxic pentavalent methylated arsenic species. Also, gaseous end-products such as trimethylarsine will emit to air, thus removing the product. Methylation is catalyzed by the enzyme As(III) S-adenosylmethionine (SAM) methyltransferase (EC 2.1.1.137), designated as AS3MT in animals and as ArsM in microorganisms. Expression of typical prokaryotic and archaeal *arsM* genes are regulated by the As(III)-responsive transcriptional repressor ArsR, consistent with arsenic methylation being a detoxification pathway in the microbes. Expression of
arsM in some cyanobacteria appears to be constitutive, indicating that alternate
detoxification pathways are used by microorganisms in which the expression of arsM
is not regulated.

The degradation of environmental organoarsenicals has been documented for
some time, while few molecular mechanisms for these reactions have been
demonstrated. Recently, a two-step pathway of MSMA reduction and demethylation
was elucidated. Although no reductases of pentavalent organoarsenicals have been
identified as yet, the enzyme, ArsI, which catalyzes demethylation of trivalent
organoarsenicals, was identified and characterized from the environmental isolate
bacterium Bacillus sp. MD1 and from the cyanobacterium Nostoc sp. 7120. ArsI,
a non-heme iron-dependent dioxygenase with C-As lyase activity, cleaves the C-As
bond in MAs(III), trivalent roxarsone, and other trivalent aromatic arsenicals. Putative
ArsI orthologs were found only in bacterial species, suggesting that alternate
pathways of organoarsenical demethylation might exist in other organisms.

2.3 The organoarsenical cycle

The arsenic concentration in seawater is around 1 to 2 µg per liter, mainly inorganic
arsenic that is usually transformed into complex organoarsenical compounds by
marine organisms. Arsenosugars, first identified in 1981, are commonly detected
water-soluble arsenic species present in marine algae; arsenobetaine is the most
abundant arsenic species in the majority of marine animals. More complex
organoarsenicals have been identified with the improvement of analytical techniques
in recent years. Since the structure of an arsenosugar phospholipid (AsPL) from a
brown alga Undaria pinnatifida was first identified, AsPL has been found in algae
and cyanobacteria. Arsenic-containing fatty acids (AsFA) that were first
identified in cod liver oil have now been found in algae and various fish species.
AsHC that were first reported in capelin have been detected in fish and algae. A new class of arsenolipids, trimethylarsenio fatty alcohols (TMAsFOH), was reported in Capelin oil. Two new groups of arsenolipids, arsenic-containing phosphatidylcholines (AsPC) and arsenic-containing phosphatidylethanolamine (AsPE) from herring caviar, were characterized. In total, more than 20 arsenosugars and 70 arsenolipids have been identified in marine organisms that live in low phosphate and high salt environments. These organoarsenicals are not toxic for the marine plants and animals, but their function in these marine organisms is not known. Phytoplankton can substitute sulfur- and nitrogen-containing membrane lipids for membrane phospholipids, arsenolipids might be used in membranes in place of phospholipids due to the more similarity of As(V), than sulfate and nitrate, to inorganic phosphate. Thus As(V) could be used as a phosphate-sparing substitute in phosphate-limiting environments. A recent study on Ectocarpus siliculosus that was found to produce more arsenosugar phospholipids under low-phosphate than under normal phosphate conditions supports this hypothesis.

These organoarsenicals may be toxic to organisms that cannot biosynthesize them. In vitro toxicological characterization of three arsenic-containing hydrocarbons showed that cytotoxicity of the arsenic-containing hydrocarbons was comparable to that of As(III) for cultured human bladder and liver cells, and arsenolipids were metabolized by humans to dimethylated arsenical species (DMAs) and other small molecular arsenic compounds, then excreted in the urine.

Although several potential pathways have been proposed for the synthesis of complex organoarsenicals, few genes involved in these biotransformation have been identified. Even less is known about the degradation of these organoarsenicals, and more studies are needed on this front.
2.4 The arsenic thiolation cycle

Thioarsenates (H₃AsSnO₄₋ₙ) are the dominant arsenic species in alkaline, arsenic-rich, sulfidic environments. These play a significant role in the arsenic biochemical cycle in sulfidic geothermal environments. ⁷⁴, ⁷⁵, ⁷⁶ Thioarsenates are transformed to As(V) and/or As(III) via exposing to oxidizing agents or increased pH, ⁷⁷ by biological conversion by sulfur-oxidizing bacteria, ⁷⁵, ⁷⁸ or by abiotic decomposition (desulfidation) with subsequent biological oxidation. ⁷⁷, ⁷⁹ Thioarsenates contain reduced S²⁻ and oxidized As(V). They can serve both as electron donors and electron acceptors. For example, monothioarsenate can be used as an electron donor by *Thermocrinis ruber* OC 14/7/2, ⁸⁰ and for anoxygenic photosynthesis by phototrophic purple sulfur bacteria growing in an alkaline environment. ⁸¹ Recently, the haloalkaliphilic bacterium MLMS-1 can grow chemolithotrophically by oxidizing the S²⁻ of monothioarsenate to S⁰ or SO₄²⁻, while concurrently reducing As(V) to As(III). ⁷⁸ In summary, various microbes have evolved to utilize thioarsenates that are widespread in sulfidic environments.

2.5 Arsenic efflux pathways

The best way to deal with toxic arsenicals in cells is acquisition of an efficient efflux system. As(III) efflux systems have been intensively studied in both microbes and higher organisms. ⁸², ⁸³ As(III) efflux in most bacteria is mediated by ArsB in an energy-dependent process, driven in *Staphylococcus aureus* by the membrane potential ⁸⁴ and in *E. coli* by ATP hydrolysis that ArsA binds to ArsB to an ATP-driven arsenic-specific pump. ²⁶ In the legume symbiont *S. meliloti*, an aquaglyceroporin (AqpS), instead of ArsB, has been identified to extrude As(III) from cells. ²⁷ Acr3 has been shown to be an As(III)-efflux transporter in both bacteria and yeast, and provides a pathway for As(III) extrusion from cells. ¹⁹ In fact, genes for
Acr3 are more widespread in bacteria and archaea than are \textit{arsB} genes. The cytosolic As(III)/glutathione complex sequestered into vacuoles by an ABC-type transporter, Ycf1p (yeast cadmium factor protein), is the second pathway for As(III) detoxification in yeast \textit{S. cerevisiae}.\textsuperscript{28}

Moreover, a novel mechanism for As(V) resistance was identified in a variety of microbes including \textit{Pseudomonas aeruginosa}.\textsuperscript{29} In these bacteria there are two genes that always go together, one encoding a typical glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and the second one, called \textit{arsJ}, that encodes an organoarsenical efflux permease (ArsJ). GAPDH uses As(V) and glyceraldehyde 3-phosphate to form the extremely unstable organoarsenical 1-arsenso-3-phosphoglycerate, which is extruded from cells by ArsJ and immediately breaks down into As(V) and 3-phosphoglycerate. The net reaction is effectively As(V) extrusion, and the coupled reaction confers As(V) resistance to these microbes, the only known efflux pathway for As(V). Meanwhile the bacterial permease, ArsP, from \textit{Campylobacter jejuni}, was demonstrated to be an efflux system specific for trivalent organoarsenicals.\textsuperscript{30} It is more selective for the ancient organoarsenical MAs(III) than for the recently anthropogenically-developed antimicrobial aromatic arsenical growth promoters such as trivalent roxarsone. More and more arsenic reductases and trivalent arsenic-specific transporters identified show that arsenic reduction and efflux play an important role in arsenic biogeochemical cycling.

\section*{3. Coupling of arsenic biogeochemical cycling to other elements}

Any biogeochemical process, such as the cycling of a particular element, is likely to be mediated by more than one organism, and often linked to other fundamental biogeochemical processes. Arsenic biotransformations are often coupled to the
cycling of C, Fe, S and N. The effect of turnover of these elements on microbes involved in arsenic biogeochemical cycling is summarized in Figure 2 which contains data based on previous studies and as described below.

3.1 The effects of iron on arsenic biogeochemical cycling

The chemical speciation of arsenic and arsenic mobility in natural environments are strongly dependent on redox potential and pH. Under oxic conditions, As(V) is the predominated arsenic species, present mainly as H₂AsO₄⁻ at acid pH or HAso₄²⁻ at alkaline pH. In anoxic environments, arsenic occurs primarily as reduced As(III) (As(OH)₃ at neutral pH or H₂AsO₃⁻ at alkaline pH), and more mobile than As(V).

Moreover, pH will impact arsenic interactions with Fe, the sorption of As(V) onto amorphous iron oxide and goethite is higher than that of As(III) below pH 5-6, As(V) and As(III) sorption onto iron oxide are both relatively high at neutral pH, As(III) is more easily adsorbed to iron oxide than As(V) above pH 7-8.

Transformation of arsenic-bearing Fe mineral phases strongly affects the bioavailability of arsenic within soils and aquifers due to direct and indirect interactions between the arsenic and Fe cycles including mineral formation, transformation, dissolution and redox reactions. Previous studies from our laboratory showed that Fe plaque consisting of Fe(III) (oxyhydr)oxides, which was induced artificially through adding ferrous iron in solution to paddy soils, has high affinity for As(V), and reduced arsenic uptake by rice.

Fe(III)-reducing bacteria modulate arsenic mobility in the rhizosphere. Dissimilatory reduction of Fe(III) (oxyhydr)oxides to Fe(II) by dissimilatory iron-reducing bacteria (DIRB) can result either in the release of As(V) from poorly crystalline or more crystalline ferric minerals as well as from sorption sites within sediments, or in the binding of arsenic to the formed Fe(II) minerals.
DIRB are commonly present in rice paddy soil, and mediate dissimilatory reduction of Fe(III) on the rice root-plaque. A study on the role of DIRB in arsenic release under a range of biogeochemical regimes indicated that Fe(III) reduction was stimulated by addition of acetate as a potential electron donor that resulted in a marked increase in the number of DIRB, reduction of As(V) to As(III), and arsenic release after Fe(III) reduction. If DARPs were used as Fe(III)-reducers, *Shewanella* sp. ANA-3 or *Sulfurospirillum barnesii* could release both As(III) and Fe(II) from ferrihydrite containing As(V) by reducing solid-phase As(V) and Fe(III). Eventually, most of the ferrihydrite matrix was liberated as Fe(II) and As(III) if sufficient organic electron donor was present. There is more aluminum in the crust than iron. However, *S. barnesii* does not reductively dissolve the As(V)-aluminum hydroxide precipitate, so we did not include a detailed description of the effect of aluminum on arsenic biogeochemical cycling.

In addition to Fe(III) reduction that has the potential to mobilize or immobilize arsenic depending on geochemical conditions that lead to the formation of either dissolved Fe$^{2+}$ or Fe(II) minerals, the formation of Fe(III) minerals under Fe(II)-oxidizing conditions has the potential to significantly immobilize arsenic and thus to lower its bioavailability. In particular for nitrate-reducing Fe(II)-oxidizing bacteria it has been shown that they form poorly soluble Fe(III) minerals and efficiently co-precipitate arsenic (Figure 2). Besides nitrate-dependent Fe(II) oxidation, also microaerophilic Fe(II) oxidation has the potential to influence arsenic mobility and could even be used in biotechnical applications for arsenic removal, e.g. in drinking water filters although it has been shown that in commercial drinking water filters the formation of iron biominerals by Fe(II)-oxidizing bacteria lowers arsenic removal from the water. Moreover, a
thermo-acidophilic iron-oxidizing archaeon *Acidianus brierleyi*, has been used to immobilize As(III) in the copper refinery process by producing thermodynamically stable crystalline scorodite ($\text{FeAsO}_4\cdot2\text{H}_2\text{O}$).

### 3.2 The effects of sulfur on arsenic biogeochemical cycling

More than 200 different arsenic-containing minerals have been found in the Earth’s crust, and 20% are sulfides and sulfosalts. The behavior of arsenic is affected by abiotic or biological redox of sulfur, which can either release or immobilize arsenic. Abiotic sulfide, a strong reductant under sulfate-reduced conditions, plays a critical role in arsenic solubility by forming pyrite ($\text{FeS}_2$), realgar ($\text{AsS}$), orpiment ($\text{As}_2\text{S}_3$), arsenopyrite ($\text{FeAsS}$), or by reducing As(V).

In addition to abiotic factors, sulfate-reducing bacteria (SRB) also cause dramatic changes in Fe, sulfide and arsenic species by generating hydrogen sulfide or elemental sulfur from sulfate, or by localized reprecipitation of released arsenic as $\text{As}_2\text{S}_3$ or $\text{FeAsS}$, which have low solubility. Sulfur-oxidizing bacteria have the potential to immobilize arsenic by using free or arsenic-bound sulfur as an electron donor to directly or indirectly transform As(III) and thioarsenates to As(V), or reduce As(V). In brief, transformations involving sulfur have the potential to significantly impact the fate of environmental arsenic.

### 3.3 The effects of N on arsenic biogeochemical cycling

Nitrate is an ecologically significant oxidant that can support microbial oxidation of As(III) in the absence of oxygen. The first evidence that microbes are capable of linking anoxic As(III) oxidation to denitrification came from a field study in anoxic lake water columns. The absence or presence of nitrate affected the redox state of arsenic. As(III) was present where nitrate was depleted, but As(V) was the dominant species during anoxic nitrate-rich periods. Subsequently, a nitrate-dependent As(III)
oxidation bacterium *A. ehrlichii* strain MLHE-1 was found to be capable of coupling As(III) oxidation with partial denitrification of nitrate to nitrite. \(^{38, 128}\) Two other anoxic chemolithoautotrophic strains, *Azoarcus* strain DAO1 and *Sinorhizobium* strain DAO10, were able to oxidize As(III) and fix CO\(_2\) via complete denitrification of nitrate to dinitrogen gas. \(^{129}\) Biological nitrate-dependent As(III) oxidation is widespread in the environment, and potentially plays a significant role in arsenic biogeochemical cycling. \(^{130}\)

As mentioned above, nitrate also influences the bioavailability and mobility of arsenic indirectly by linking nitrate reduction to Fe(II) oxidation. Previous studies showed tight coupling between N, Fe, and arsenic in paddy soil. \(^{85}\) Addition of nitrate reduced arsenic uptake by rice probably because i) the nitrate inhibited/reduced Fe(III) reduction leading to less arsenic mobilization and ii) nitrate-dependent Fe(II)-oxidizing bacteria stimulated Fe(II) oxidation, which led to arsenic co-precipitation with Fe(III) minerals in soil. Nitrate strongly affects arsenic cycling under anoxic conditions in nitrate-rich Upper Mystic Lake by microbially catalyzing As(III) to more particle-reactive As(V) and oxidizing Fe(II) to arsenic-sorbing particulate ferric oxides. \(^{4}\) Microbial nitrate-dependent Fe(II) oxidation in groundwater, \(^{131}\) freshwater sediments \(^{132}\) and marine sediments \(^{133}\) has the potential to contribute to the reduction of arsenic mobility in various ecosystems.

### 3.4 The effects of organic matter on arsenic biogeochemical cycling

Natural organic matter (NOM) is widely distributed in the environment. NOM consists of heterogeneous mixtures of organic compounds with various structural and functional properties, \(^{134}\) that influence the fate of arsenic by competitive adsorption and redox reactions, \(^{135, 136, 137}\) and by formation of arsenic-bearing organic-metal-complexes and mineral colloids. \(^{138, 139}\) NOM molecules have
combinations of carboxylic, amino, sulfhydryl, hydroxyl, esteric, phenolic, nitroso, and other functional groups. They are considered to be an efficient geochemical trap for arsenic both under oxic and reducing conditions. Whereas As(V) is immobilized by binding to protonated amino groups of NOM or a nucleophile substitution reaction between As(V) and phenolic OH groups of NOM, As(III) is associated with NOM via phenolic OH or carboxyl groups of NOM by H-bonding, hydrophobic As(III)-NOM interactions, or via ternary As(III)-Fe(III)-NOM complexes that form bridges between Fe(III), arsenic oxyanions and the functional entities of NOM. In contrast, under sulfate-reducing conditions, the formation of a trigonal-pyramidal complex between As(III) and sulfhydryl groups of NOM could potentially be a sequestration mechanism for arsenic.

Organic matter could also change the fate of arsenic by influencing microbial communities or activities. NOM is thought to drive the reductive dissolution of Fe(III) (oxyhydr)oxides, thereby causing redox transformations of As(V) to As(III) and facilitate arsenic release. In addition to the reductive dissolution of arsenic-bearing Fe(III) (oxyhydr)oxide phases, studies have revealed that addition of organic matter to paddy soil significantly increased arsenic methylation and volatilization. Application of rice straw into soil increased arsenic accumulation in rice by influencing microbial processes involved in arsenic redox.

Humic acids, which are forms of NOM, can reduce As(V). Small organic molecules, including lactate, pyruvate, fumarate, malate, succinate, butyrate, citrate, acetate, glycerol, ethanol, and formate, can be used an electron donor by DARPs for As(V) reduction. Desulfosporosinus sp. Y5 has been found to couple even complex aromatic substrates such as phenol, syringic acid, benzoate, ferulic acid, and toluene, to As(V) reduction. In fact, some DARPs can also respire
sulfate, selenite, Fe(III), thiosulfate, nitrate, or nitrite. This diversity of electron donors and acceptors may be of benefit to microbes grown in environments where Fe, N, S, or C coexist with arsenic.

Apart from iron oxides and organic materials, silicon significantly decreased As(III) concentration, but increased the concentration of DMAs in both the vegetative and reproductive tissues of rice. In brief, in addition to biological transformations, other inorganic elements, compounds/minerals and physico-chemical properties interact with arsenic in the environment. Coupling of arsenic with other elements makes it necessary to consider genes involved in cycling of other elements, such as Fe, N, S, and C when studying the arsenic biogeochemical cycling.

Known genes involved in arsenic biotransformation are readily determined in pure cultures. However, in the field, these genes cannot easily be quantified in bacterial communities with a limited number of primers, even though As(V)-reducing and As(III)-oxidizing bacteria are widely distributed in the environment. In addition, the speciation, fate and biogeochemical transformation processes of arsenic in the environment are much more complex than under laboratory conditions. It is therefore necessary to apply more systematic and more comprehensive approaches such as metagenomics, metatranscriptomics, metaproteomics, and metabolomics to understand interactions between environmental microbes. These approaches will take into account local geochemical surroundings and neighboring organisms by analyzing DNA, RNA, proteins, and metabolites extracted directly from environmental samples.

4 Understanding biogeochemical arsenic cycling by application of “omics” methods and integrated modeling

4.1 Metagenomics
Metagenomics provides an inestimable window into the microbial world by characterizing microorganisms involved in difficult-to-elucidate but important biochemical pathways, as the overwhelming majority of microbes in the environment cannot be cultured in the laboratory. From a metagenomic library two arsenic-resistant bacteria and one novel As(V) resistance gene (*arsN*), which encodes a protein similar to acetyltransferases, were identified. 22 Xiao et al applied metagenomic techniques to analyze genes associated with arsenic transformation. They analyzed five low-arsenic paddy soils using high-throughput sequencing and constructed a protein database of arsenic metabolizing genes. Their analysis shows that arsenic metabolism genes are ubiquitous and abundant, even in low-arsenic environments. 164 Metagenomics was also used to unravel the correlations between the microbes and arsenic transformation in different niches. 165, 166, 167, 168, 169 Although metagenomics provides taxonomic and functional profiles of a microbial community, it does not demonstrate the levels of expression of the genes nor their physiological activity. 170 Therefore, metatranscriptomics is needed to delineate the active functional genes and communities.

### 4.2 Metatranscriptomics

Metatranscriptomics offers novel insights into the expression of functional genes and microbial activities of complex microbial communities at a specific moment or under specific environmental conditions by sequencing the total mRNAs extracted from natural microbial communities. Functional metatranscriptomics has potential for isolation and characterization of novel genes involved in heavy metal transformation. 171 Metatranscriptomics enhances our understanding of microbial responses to their environment 172 and the functional profile of a microbial community. 173 Recently, a transcriptomics meta-analysis was used to unravel the effect of As(III) on the...
symbiotic interaction between the model legume *Medicago truncatula* and its symbiont *Ensifer* (syn. *Sinorhizobium*) *medicae* MA11. 174 This study identified the adaptive responses of the bacterial symbiont to arsenic exposure. This metatranscriptomic approach will be useful to study how microbes regulate their genes to adapt to the changes in environmental conditions, particularly arsenic concentrations. 175 The correlation between mRNA and protein inventories in environmental microbial communities is low when environmental conditions change rapidly. mRNA inventories respond rapidly and sensitively to the shift, while changes in protein inventories are slow. 176 As a consequence, microbial metaproteomics and metabolomics had to be used to identify the repertoire of proteins and small molecular metabolites that microbes use to adapt to complex and dynamic environments. In this way, the metabolic activities of a microbial community in specific environments at the moment of sampling could be elucidated. 170

**4.3 Metaproteomics and metabolomics**

Metaproteomics and metabolomics are the comprehensive methods by which proteins produced by microbes and metabolites released by microorganisms into the environment are characterized and quantified using a combination of liquid or gas chromatography-based separations and mass spectrometry-based identification techniques. 177, 178 The study of microbial proteome and metabolome can provide valuable information about the function of microbial communities and the interactions of the microbial communities with the environment. 179, 180 When the diversity of arsenic-adapted prokaryotic communities in mildly arsenic-contaminated sediments was analyzed using meta-proteomic and 16S rRNA amplification, the results indicated that the data sizes provided by metaproteomics are less than those provided by metagenomics and metatranscriptomics. 181 So far, metabolomics has been used
mainly to analyze low molecular mass metabolites within a tissue, biofluid, a cell or cell compartment of an organism including plants, animals, bacteria, and humans exposed to an environmental stressor. 180 High-throughput metabolomics has been applied to the analysis of metabolites in the liver of mice when co-exposed to high fat and cholesterol diets and arsenic-contaminated drinking water. 182 However, application of metaproteomic and metabolomic techniques to real environment is limited due to difficulties with amplification and the low quantities of extractable proteins and metabolites because of the interferences with many components present in complex environmental systems, such as soil. 180,183

4.4 Integrating meta-omic techniques

From the above, meta-omics are key techniques in elucidating the dynamic and complex interactions between microbial communities and the environment. Integrating multiple meta-omic datasets will provide a complete exhibition from genes to biogeochemical cycles. Metatranscriptomic and metagenomic techniques were combined to detect large numbers of novel genes from complex marine microbial communities. 173 Datasets of meta-genomics were integrated with metabolomics to reveal how a microbial community interacts with the environment and responds to environmental parameters. 184 Taken together, meta-omic technologies offer an unprecedented opportunity to elucidate the functions of microbes that are not readily cultured under normal laboratory conditions in biogeochemical cycles.

Recently, "Arsenomics" was termed as an approach to focusing on the analysis of alterations in transcriptome, proteome and metabolome occurring in microbes exposed to arsenic. 185,186 With the application of meta-omics to environmental
science, we believe that Arsenomics will evolve to include the analysis of metagenomic, metatranscriptomic, metaproteomic, and metabolomic changes in microbial communities from the real environment where they are exposed to arsenic.

4.5 Integrating environmental meta-omics into biogeochemical models for arsenic

Microbes are ubiquitous in diverse environmental niches including soil, oceans, sediments, freshwater environments, and inside the body of animals or plants, and exert great influence on biogeochemical cycles in these habitats. For example, microbes involved in arsenic transformation are ubiquitously distributed in paddy soils, resulting in various concentrations and percentage of inorganic and methylated arsenic species among different rice plants. Since higher plants appear not to methylate arsenic, microbial methylation is probably the primary source of methylated arsenic in plants, which occurs in soil prior to plant uptake.

*In situ* measurements or prediction of arsenic transformations contribute to analysis of the dynamics of arsenic and prediction of arsenic bioavailability. *In situ* measurements of As(V) reduction in Mono Lake, California (dissolved inorganic arsenic ~ 200 μM), made with radiotracers (\(^{75}\)As and \(^{35}\)S) of mass balance considerations, revealed that As(V) reduction occurred in the monimolimnion waters with the highest rates between 18 and 19 meters (rate, ~ 5.9 μM/day) and sulfate reduction rates increased with depth at depths of 21 meters and below with the highest rates at 28 meters (rate, ~ 2.3 μM/day). The radioisotope method was further employed to examine the As(V) and sulfate reduction processes in sediments of two arsenic-rich soda lakes, Mono Lake (moderately salt, ~ 90 g/L) and Searles Lake (saturated salt, ~ 340 g/L). The rate constant \([k]\) of As(V) reduction was 0.103 to 0.04 h\(^{-1}\) in Mono Lake and 0.012 to 0.002 h\(^{-1}\) in Searles Lake, and sulfate reduction
was only detected in Mono Lake ($k=7.6 \times 10^4$ to $3.2 \times 10^{-6}$ h$^{-1}$). Denatured gradient gel electrophoresis (DGGE) of 16S rRNA genes amplified from Mono Lake and Searles Lake sediment DNA indicated that microbial communities from two sediments were distinct from each other. More *arrA* gene signal was found in Mono Lake than in Searles Lake, where higher As(V) reduction activity was observed, due to PCR biases, the presence of novel *arrA* genes, or higher expression of low-abundance *arrA* genes.  

Characterization of genes involved in arsenic biotransformation and application of multiple meta-omics in environment analysis will lead to insights into the microbial world, as limited information on functional genes cannot predict the status of arsenic in the environment. The information of field/*in situ* characterization of functional genes/functional microbial communities and biogeochemical fluxes should be integrated into biogeochemical models to complete the transition from lab to the field, from biochemistry to biogeochemistry, and from genes-genomics to microbial communities. This integration will help to predict the dynamics of arsenic in the environment, and to improve the effectiveness of mitigation technologies. Many strategies have been developed to model low complexity environments. For example, recent work integrating environmental genomics and qPCR in biogeochemical models explored the nexus between microbial community and geochemistry in the Arabian Sea oxygen minimum zone. Metabolic processes coupling C, N, and S transformations in the Saanich Inlet oxygen-starved zone were integrated using a biogeochemical model that integrates multi-omic information and geochemistry. These studies indicated that such integrated modeling approaches can provide a novel insight into microbial metabolic networks in water bodies, and pave a road for prediction of elemental cycling.
5 Perspectives

In summary, as arsenic biotransformations are catalyzed by a suite of enzymes from diverse environmental organisms, and these are coupled to biogeochemical cycles of other elements such as Fe, S, and N. As more and more genomes are being sequenced, more genes directly or indirectly involved in arsenic metabolism will be discovered and characterized. With the development of new technologies, we anticipate rapid advances in analytical chemistry, microbiology and genomics that will improve our understanding of how microbial metabolic pathways contribute to and govern complex environmental processes. In the future, integrating meta-omic datasets into biogeochemical models will improve the ability of prediction and offer a deeper insight into arsenic biogeochemical processes in diverse niches.

In future studies, it will be necessary to analyze the interaction between organisms and the environment using additional meta-omics approaches at different spatio-temporal scales. Geochemical analyses in combination with genetic analyses will provide insights into the specific roles of the complex biochemical pathways in the global arsenic biogeochemical cycle. More importantly, integrating modeling approaches linking arsenic biogeochemical cycle with meta-omics data should be developed to predict the dynamic of arsenic species in water, sediments and soils and provide our society and authorities with the tools necessary for limiting arsenic pollution, improving remediation and providing safe drinking water and food.

Acknowledgments
Our research was supported by National Natural Science Foundation of China (41571130063), Program of Chinese Academy of Sciences (XDB15020402), National Natural Science Foundation of China (21507125), and International Science & technology Cooperation Program of China (2011DFB91710). BPR was supported by NIH grants GM55425 and ES023779.

Reference


(88) Bertin, P. N.; Heinrich-Salmeron, A.; Pelletier, E.; Goulien-Chollet, F.; Arsène-Ploetz, F.;


(127) Hollibaugh, J. T.; Budinoff, C.; Hollibaugh, R. A.; Ransom, B.; Bano, N. Sulfide Oxidation


(153) Niggemyer, A.; Spring, S.; Stackebrandt, E. Isolation and Characterization of a Novel As (V)
Reducing Bacterium: Implications for Arsenic Mobilization and the Genus Desulfitobacterium.


Escudero, L. V; Casamayor, E. O.; Chong, G.; Pedro, C.; Demergasso, C. Distribution of Microbial Arsenic Reduction, Oxidation and Extrusion Genes along a Wide Range of Environmental Arsenic Concentrations. PLoS One 2013, 8 (10), e78890 DOI: 10.1371/journal.pone.0078890.


**Legends**

Figure 1. Proposed pathways for arsenic redox reactions and synthesis of novel organoarsenicals.

Figure 2. The model of effect of Fe, N, S and natural organic matter (NOM) on microbes involved in arsenic biogeochemical cycling highlights proteins associated with elemental metabolisms. Green ovals denote arsenic transporters, yellow ovals denote transmembrane enzymes. Red words are enzymes, blue words are related arsenic compounds. The full name of enzymes that were not mentioned in the text was provided in the follow, NarG, transmembrane nitrate reductase that drives the nitrate
reduction to nitrite; Nas, cytoplasmic-assimilatory nitrate reductase that drives the nitrate reduction to nitrite; NapA, periplasmic-dissimilatory nitrate reductase; Nir /Nrf (associated with NapA), nitrite reductase that drives the nitrite reduction to nitric oxide; NorB/C, nitric oxide reductase that drives the nitric oxide reduction to nitrous oxide; NosZ, nitrous oxide reductase that drives the nitrous oxide reduction to nitrogen; Nif, nitrogenase that catalyzes the nitrogen fixation to ammonia; Hs, hydrazine synthase that catalyzes the production of nitrogen from nitrous oxide and ammonia; cyt, cytochrome.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Protein abbreviation</th>
<th>function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>As(III) oxidation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>aioA</em> and <em>aioB</em></td>
<td>As(III) oxidase</td>
<td>AioAB</td>
<td>Oxidate As(III)</td>
<td>9</td>
</tr>
<tr>
<td><em>arxA</em></td>
<td>As(III) oxidase</td>
<td>ArxA</td>
<td>Oxidate As(III)</td>
<td>10</td>
</tr>
<tr>
<td><em>aioX/arxX</em></td>
<td>As(III)-binding protein</td>
<td>AioX/ArxX</td>
<td>Involved in As(III)-based signaling and regulation of As(III) oxidation</td>
<td>12, 11</td>
</tr>
<tr>
<td><em>aioS/arxS</em></td>
<td>sensor histidine kinase</td>
<td>AioS/ArxS</td>
<td>One part of two-component signal transduction system</td>
<td>9, 11</td>
</tr>
<tr>
<td><em>aioR/arxR</em></td>
<td>transcriptional regulator</td>
<td>AioR/ArxR</td>
<td>Regulate the expression of <em>aio/arx</em> operon</td>
<td>9, 11</td>
</tr>
<tr>
<td><em>moeA</em></td>
<td>molybdenum cofactor biosynthesis protein</td>
<td>MoeA</td>
<td>Synthesize the molybdenum cofactor of AioAB oxidase</td>
<td>9</td>
</tr>
<tr>
<td><em>arxC</em></td>
<td>membrane anchoring and quinol oxidoreductase subunit</td>
<td>ArxC</td>
<td>Involved in As(III) oxidation</td>
<td>11</td>
</tr>
<tr>
<td><em>arxD</em></td>
<td>TorD-like molybdoenzyme chaperone</td>
<td>ArxD</td>
<td>Involved in As(III) oxidation</td>
<td>11</td>
</tr>
<tr>
<td><em>arsH</em></td>
<td>organoarsenical oxidase</td>
<td>ArsH</td>
<td>Oxidate trivalent methylated and aromatic arsenicals, reduce chromium and iron</td>
<td>13, 14</td>
</tr>
<tr>
<td><strong>As(V) reduction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>arrA</em> and <em>arrB</em></td>
<td>As(V) respiratory reductases</td>
<td>ArrAB</td>
<td>Reduce As(V)</td>
<td>15, 16</td>
</tr>
<tr>
<td><em>arrC</em></td>
<td>As(V) reductase membranous subunit</td>
<td>ArrC</td>
<td>Involved in As(V) reduction</td>
<td>17</td>
</tr>
<tr>
<td><em>arrD</em></td>
<td>As(V) reductase chaperon</td>
<td>ArrD</td>
<td>Involved in As(V) reduction</td>
<td>17</td>
</tr>
<tr>
<td><em>arrS</em></td>
<td>sensor histidine kinase</td>
<td>ArrS</td>
<td>Regulate the expression of <em>arr</em> operon</td>
<td>17</td>
</tr>
<tr>
<td><em>arsC</em></td>
<td>As(V) reductase</td>
<td>ArsC</td>
<td>Reduce As(V)</td>
<td>18</td>
</tr>
<tr>
<td><em>ACR2</em></td>
<td>As(V) reductase</td>
<td>ACR2</td>
<td>Reduce As(V)</td>
<td>19</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Function</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>ACR1</td>
<td>transcriptional regulatory protein</td>
<td>Regulate the expression of ACR genes</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>arrR/arsR</td>
<td>arsenic-responsive repressor</td>
<td>Regulate the expression of arr/ars operon</td>
<td>17, 20</td>
<td></td>
</tr>
<tr>
<td>GstB</td>
<td>glutathione S-transferase B</td>
<td>Reduce As(V) to As(III) with reduced GSH</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>arsN</td>
<td>acetyltransferase</td>
<td>Putative As(V) reductase</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Arsenic methylation and demethylation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arsM</td>
<td>As(III) S-adenosylmethionine (SAM) methyltransferases</td>
<td>Methylate arsenic</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>arsI</td>
<td>A C · As lyase</td>
<td>Catalyze demethylation of trivalent organoarsenicals</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Arsenic transport</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arsB</td>
<td>As(III)-pump protein</td>
<td>Extrude As(III) from the cell</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>arsA</td>
<td>As(III)-pump ATPase</td>
<td>The catalytic subunit of an oxyanion-translocating ATPase</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>arsD</td>
<td>arsenical metallochaperone</td>
<td>Transfer trivalent metalloids to ArsA</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>apqS</td>
<td>aquaglyceroporin</td>
<td>Extrude As(III) from the cell</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>ACR3</td>
<td>As(III) permease</td>
<td>Extrude As(III) from the cell</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Ycf1p</td>
<td>yeast cadmium factor protein cytosolic</td>
<td>Sequester cytosolic As(III)/glutathione complex into vacuoles</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>arsJ</td>
<td>organoarsenical efflux permease</td>
<td>Extrude organoarsenicals from the cell</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>arsP</td>
<td>efflux system specific for trivalent organoarsenicals</td>
<td>Extrude trivalent organoarsenicals from the cell</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>pgpA</td>
<td>P-glycoprotein-related protein</td>
<td>Recognize and transport thiol-metal conjugates</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Unknown functions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arsO</td>
<td>putative flavin-binding monooxygenase</td>
<td>Unknown function</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>arsT</td>
<td>putative thioredoxin reductase</td>
<td>Unknown function</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>