Unexpected genetic and microbial diversity for arsenic cycling in deep sea cold seep sediments

Chuwen Zhang[1], Xinyue Liu[1, 2], Ling-Dong Shi[3], Jiwei Li[4],

Xi Xiao[5], Zongze Shao[1, 6], Xiyang Dong[1, 6] *,

[1] Key Laboratory of Marine Genetic Resources, Third Institute of Oceanography, Ministry of Natural Resources, Xiamen, China

[2] School of Marine Sciences, Sun Yat-Sen University, Zhuhai, China

[3] College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, China

[4] Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences, Sanya, China

[5] Key Laboratory of Marine Mineral Resources, Ministry of Natural Resources, Guangzhou Marine Geological Survey, China Geological Survey, Guangzhou, China

[6] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China

* Correspondence can be addressed to Xiyang Dong (dongxiyang@tio.org.cn).

Abstract

Cold seeps, where cold hydrocarbon-rich fluid escapes from the seafloor, showed strong enrichment of toxic metalloid arsenic (As). The toxicity and mobility of As can be greatly altered by microbial processes that play an important role in global As biogeochemical cycling. However, a global overview of genes and microbes involved in As transformation at seeps remains to be fully unveiled. Using 87 sediment metagenomes and 33 metatranscriptomes derived from 13 globally distributed cold seeps, we show that As detoxification genes (*arsM*, *arsP*, *arsC1*/*arsC2*, *acr3*) were prevalent at seeps and more phylogenetically diverse than previously expected. Asgardarchaeota and a variety of unidentified bacterial phyla (e.g. 4484-113, AABM5-125-24 and RBG-13-66-14) may also function as the key players in As transformation. The abundances of As-cycling genes and the compositions of As-associated microbiome shifted across different sediment depths or types of cold seep. The energy-conserving arsenate reduction or arsenite oxidation could impact biogeochemical cycling of carbon and nitrogen, via supporting carbon fixation, hydrocarbon degradation and nitrogen fixation. Overall, this study provides a comprehensive overview of As-cycling genes and microbes at As-enriched cold seeps, laying a solid foundation for further studies of As cycling in deep sea microbiome at the enzymatic and processual levels.

Introduction

Cold seeps are characterized by the emission of subsurface fluids into the sea floor and occur widely at active and passive continental margins[1, 2]. The upward fluids are often rich in methane and other hydrocarbons which sustain sea bed oasis composed of various microorganisms and faunal assemblages[3, 4]. The primary process that fuel complex cold seep ecosystems is the anaerobic oxidation of methane (AOM), conjointly operated by a consortium of anaerobic methane-oxidizing archaea (ANME) and sulfate-reducing bacteria (SRB)[5, 6]. AOM removes approximately 80% of upward venting methane, acting as an efficient methane filter[7]. Additionally, deep-sea cold seep sediments also contain diverse and abundant diazotrophs that might contribute substantially to the global nitrogen balance[8]. Cold seeps are therefore biologically and geochemically significant on a global scale.

The venting fluids can significantly influence the sedimentary environment of seep sites, resulting in changes of chemical characteristics of sediments[9]. In particular, arsenic (As), one of the most abundant elements in the Earth's crust, are anomalously enriched in seep sediments[10-14]. The anomalous As enrichment could be attributed to the ascending fluids that could capture As and other metals when passing through thick shaly formations[10, 14]; or the so-called particulate iron shuttle effect[9, 11, 13]. As is also toxic metalloid in nature that, upon exposure, can cause negative effects for all living things[15]. Depending on the physicochemical conditions, As can be found in

different oxidation and methylation states, showing various levels of toxicity and bioavailability[16]. In marine environments, arsenate (As(V)), and arsenite (As(III)) are the dominant forms of inorganic As[17]. It is assumed that microbes have evolved a genetic repertoire related to As cycling, dated back to at least 2.72 billion years ago[18, 19]. As biotransformation processes include As detoxification to mitigate toxicity and As respiration to conserve energy. The As detoxification is mainly achieved by two steps: reduction of As(V) to As(III) by cytoplasmic As(V) reductases (*arsC* gene) with homology to either glutaredoxin (*arsC1* gene) or thioredoxin (*arsC2* gene) family and subsequent extrusion of As(III) via As(III) efflux permeases (*arsB* and *acr3* genes)[20, 21] (**Figure 1**). Another As detoxification mechanism involves the methylation of As(III) to methylarsenite (MAs(III)) by the As(III) S-adenosylmethionine (SAM) methyltransferase (*arsM* gene)[22] (**Figure 1**). Although MAs(III) intermediates are more toxic than As(III), they do not accumulate in cells and can be detoxified through several different pathways. MAs(III) can be further methylated by ArsM and volatilized, extruded from cells via the MAs(III) efflux permease (*arsP* gene)[23], oxidated to less toxic MAs(V) by the MAs(III)-specific oxidase (*arsH* gene)[24], or demethylated to less toxic As(III) by the C-As lyase (*arsI* gene)[25]. As respiration consists of the chemolithotrophic oxidation of As(III) by As(III) oxidase (*aioAB*/*arxAB* gene cluster) and dissimilatory As(V) reduction by respiratory As(V) reductase (*arrAB* gene cluster)[15, 26] (**Figure 1**). Taken together, microbes have a huge potential effect on the biogeochemical cycling and toxicity of As.

So far, As-transforming microbes and As-related genes have been widely investigated in various natural environments, including polluted and pristine soils[27, 28], terrestrial geothermal springs[29-31], wetlands[32, 33], pelagic oxygen-deficient zones[34], groundwater[35], etc. For example, metagenomic and metatranscriptomic analyses revealed that Aquificae are the key players for the *arsC*-based detoxification in Tengchong geothermal springs[29]. A global survey also described the phylogenetic diversity, genomic location, and biogeography of As-related genes in soil metagenomes[36]. Only recently, the behavior of As biotransformation has been reported in the deep-sea realms, i.e. hadal trench of the Challenger Deep[37]. Deep-sea ecosystems cover 67% of Earth surface and have extremely high densities of microbes (up to 1000× greater than surface waters) which play a critical role for long-term controls on global biogeochemical cycles[38, 39]. The environmental conditions in deep seafloor cold seeps differ greatly from those in the aforementioned ecosystems, such as low temperatures, high pressure, darkness and the presence of seepage activities[1]. Thus, the investigation of As-related genes and microbes at seeps will expand our current knowledge on As metabolisms and allow us to discover new lineages containing As-related genes.

The purpose of this study was to decipher the microbial transformation of As in cold seep sediments at a global scale. Here, we applied a relatively comprehensive data set of 87 sediment metagenomes and 33 metatranscriptomes derived from 13 geographically diverse cold seeps across global oceans (**Supplementary Data 1**), to

investigate As-associated genes and their host microbes. This study aims to address the following questions: (i) biogeography of As-cycling genes across global cold seeps; (ii) phylogenetic diversity and distribution of As-cycling genes across global cold seeps; (iii) interactions between As metabolisms and biogeochemical cycles of carbon and nitrogen.

Results and discussion

As cycling genes are widespread and active across global cold seeps

To gain a broad view on biogeography of As cycling genes, we determined their abundance from 87 sediment metagenomes collected from 13 globally distributed cold seeps. Considering that sulfate respiration is one of the most important microbial redox processes in cold seep sediments[40], the *dsrA* was used as the target gene to compare with As cycling genes. We found that genes related to As detoxification were prevalent in these cold seep samples (**Figure 2**) and their abundances were higher than those of *dsrA* genes (**Supplementary Figures 2-3; Supplementary Data 2**). The *arsM* and *arsP* genes that respectively produce volatilized methylated organoarsenicals and mediate its subsequent expulsion outside cell, were the most abundant ones. The *arsC1*/*arsC2* genes for cytoplasmic As(V) reduction and the *acr3* gene for As(III) extrusion also dominated most cold seep samples. Moreover, As detoxification genes, i.e. *arsM*, *arsP*, *arsC1*/*arsC2*, *acr3*, were actively expressed in the sediment metatranscriptomes from Haima and Jiaolong seeps along with gas hydrate deposit

zones of Qiongdongnan and Shenhu, revealing *in situ* microbial activities on As detoxification (**Supplementary Figure 4**). Seep microbes might utilize both methylation and cytoplasmic As(V) reduction strategies to overcome potential toxic effects of exceptional As accumulation at cold seeps. Alternatively, methylation is not strictly a detoxification pathway but also an antibiotic-producing process with MAs(III) being a primitive antibiotic[41], which could provide additional competitive advantages. However, the function of *arsM* in anoxic environments and its contribution to As cycling have yet to be verified. Our results contradict previous findings demonstrating that *arsM* are less common in soils[36] and hot springs[29] than *arsC*, but in line with those found in hadal sediments[37]. The discrepancy in As detoxification mechanisms between terrestrial and deep-sea ecosystems could be attributed to their huge variations in the habitats and geographical locations. When comparing the abundance of As(III) efflux pumps, we observed that *arsB* was in much lower abundance than *acr3* (**Figure 2a**). Previous studies also reported an abundance of *acr3* over *arsB* in forest soils and wetlands[32, 36, 42]. This is likely because Acr3 proteins are more ancient and have greater phylogenetic distribution as compared with ArsB[19]. Conversely, genes related to energetic As respiratory oxidation (*aioA*) and reduction (*arrA*/*arxA*) were less abundant in all cold seep samples as compared with As detoxification genes. Despite of this, respiratory genes were transcriptionally active, as evidenced by the detection of *arrA* transcripts in Jiaolong seep (up to 15.9 TPM, **Supplementary Figure 4**).

To determine the distribution characteristics of As cycling genes, each metagenome was categorized in terms of its sediment depth (i.e. surface: <1 mbsf; shallow: 1-10 mbsf; deep: >10 mbsf). Metagenomes were also grouped based on the type of cold seep, including gas hydrate, seep (i.e. oil and gas/methane seep), and volcano (mud/asphalt volcano)[1], respectively. The partial least squares discrimination analysis (PLS-DA) revealed dissimilarity in As cycling genes among different sediment layers (**Figure 2b**; F = 4.3504, p = 0.001, R^2 = 0.10267, 999-permutations PERMANOVA test). The distribution traits of As cycling genes in surface sediments were separated from deep sediments and more similar to those in shallow ones (**Figure 2b**). The abundance of prevalent As cycling genes such as *acr3*, *arsC2* and *arsM* in deep sediments were significantly higher as compared with those in shallow and surface sediments (**Supplementary Figure 2**). As cycling genes in different types of cold seep were also different from each other (**Figure 2b**; F = 3.5246, p = 0.004, R^2 = 0.07742, 999-permutations PERMANOVA test). Dominant As cycling genes in gas hydrates displayed higher abundances relative to those in seeps and volcanos (**Supplementary Figure 3**). Hence, the distributions of As-associated genes were influenced by a combination of sediment depths and types of cold seep. The higher As-cycling gene abundance observed in our deep or gas hydrate-associated samples could be correlated with a high level of environmental As, as what described in As-rich altiplanic wetland[32]. In the Nankai Trough, As with unknown sources was demonstrated to actively release into sediments layers where methane hydrates occur (As concentration of 14 ppm in gas

hydrate-bearing sediments vs av. 6.4 ppm for the whole sediment core)[17].

Microbes involved in As cycling varied between seep habitats

To profile taxonomic diversity of As-related microbes, a total of 1741 species-level metagenome-assembled genomes (MAGs, 95% average nucleotide identity) were reconstructed from these 87 cold seep metagenomes (**Supplementary Data 3**). Of these, 1083 MAGs spanning 9 archaeal and 63 bacterial phyla as well as one unclassified bacterial phylum were potentially involved in As cycling at cold seeps (**Supplementary Data 4**). Metagenomic read recruitments revealed that the recovered 1083 As-related MAGs accounted for 1.8-62.8% cold seep communities (**Figure 3 and Supplementary Data 4**). Taxonomic compositions of As-related microbiomes across different types of cold seep displayed pronounced variations (**Figure 3**). In the sediments derived from oil and gas/methane seep, As-related microbes contained mostly Methanogasteraceae (i.e. ANME-2c) and Methanocomedenaceae (i.e. ANME-2a) within Halobacteriota phylum, ETH-SRB1 within Desulfobacterota phylum, JS1 within Atribacterota phylum as well as Anaerolineae and Dehalococcoidia within Chloroflexota phylum. The As-related microbes in gas hydrate sediments were dominated by bacterial lineages, highlighted by Atribacterota (JS1) and Chloroflexota (Anaerolineae and Dehalococcoidia). Nevertheless, in asphalt/mud volcano sediments, the compositions of As-related microbes were diverse in different samples. The clear distinctions in As-related microbiomes between different seep habitats suggested an

important role driven by environment selection. Multiple parameters, including sediment temperature, sediment depth, water depth, methane concentration, and geographic distance have been demonstrated to cause these variations[43, 44]. Additionally, our results show that Chloroflexota outnumbered Atribacterota in sediment samples with lower Fe(II) concentration (av. 16.51 μmol/L), while Atribacterota dominated over Chloroflexota in sediment samples with higher Fe(II) concentration (av. 81.54 μmol/L). It is possible that iron oxyhydroxides control the mobilization of As[45] and thus affect As-related microbial communities (**Figure 3**).

Expanded diversity of microbial lineages containing As detoxification genes

Among these As detoxification genes, *acr3*, *arsC1*/*arsC2*, *arsM* and *arsP* were widely distributed in bacteria and archaea, while other As detoxification genes (*arsB*, *arsI* and *arsH*) were sparsely distributed (**Figure 4**). The *acr3* gene is typically affiliated with Proteobacterial, Firmicutes, Actinobacterial and other bacterial sequences[36, 42, 46]. Our study observed an unexpectedly wider phylogenetical diversity of *acr3* than previously reported. Notably, Asgardarchaeota including Lokiarchaeia, Thorarchaeia, Sifarchaeia, LC30, along with Heimdallarchaeia and Wukongarchaeia described as the most likely sister group of eukaryotes, are firstly documented to have genetically capability for As(III) extrusion. The greater diversity of As detoxification genes found in Asgardarchaeota phylum further point to their ancient origin[19]. Furthermore, a considerable number of candidate bacterial phyla without cultured representatives (e.g.

4484-113, AABM5-125-24 and RBG-13-66-14) were also equipped with such an ability. Though their functional redundancy as As(III) efflux pumps, *arsB* was more phylogenetically conserved as compared with *acr3* and simply restricted to Alphaproteobacteria, Gammaproteobacteria and Campylobacterota (**Figure 4**). This observation is in agreement with previous reports comparing the diversity of *arsB* to *acr3*[36, 42, 46]. The *arsM* gene was relatively uncommon in terrestrial soil microorganisms[29, 36]. In contrast, this study showed that the *arsM* genes in seep microbes have a great taxonomic diversity similar to *acr3* gene, including Chloroflexota, Proteobacteria, Atribacterota, Asgardarchaeota, Hydrothermarchaeota, Thermoplasmatota, Thermoproteota as well as other currently unidentified bacterial phyla (e.g. 4484-113, AABM5-125-24 and RBG-13-66-14) (**Figure 4**). Among these, Atribacterota, Asgardarchaeota and the candidate bacterial phyla stated above have not previously been implicated in As methylation[19, 36]. For cytoplasmic As(V) reduction, Asgard archaeal lineages all lacked corresponding genes (*arsC1* and *arsC2*). The underlying causes to their absence in Asgardarchaeota are unclear. It's likely that Asgardarchaeota lost cytoplasmic As(V) reduction genes during evolution or possess different enzyme systems. In general, these data advance our understanding on the phylogenetical diversity of As detoxification genes and highlight the potentially important role played by archaea in As cycling, Asgardarchaeota particularly.

Diverse seep lineages are identified to perform As respiration

In addition to mitigating toxicity, some microorganisms can respire the redox-sensitive element of As to reap energetic gains (i.e. arsenotrophy), either via chemoautotrophic As(III) oxidation (*aioAB*/*arxAB*) or anaerobic As(V) respiration (*arrAB*)[47, 48]. The alpha subunits of these arsenotrophic enzymes form distinct clades with the dimethylsulfoxide (DMSO) reductase superfamily[34]. This superfamily also includes other enzymes critical in respiratory redox transformations, e.g. Nap and Nar. Here, we identified two AioA, three ArxA and 17 ArrA protein sequences, respectively. A phylogenetic analysis of recovered arsenotrophic protein sequences showed that they all clustered together with known AioA/ArxA and ArrA proteins (**Figure 5a**). Functional As bioenergetic *aioA*/*arxA* and *arrA* genes are generally found together with other necessary accessory genes. The *aioA* of As(III) oxidizing microorganisms always forms an operon with *aioB* and other genes involved in As detoxification and metabolisms (e.g. *aioD*, *aioXSR*, *arsR*)[15, 26]. Arx is demonstrated to be a variant of Arr and these two enzymes have a similar genetic arrangement. The *arrA*/*arxA* gene is always found together with the *arrB*/*arxB* and often with the *arrC*/*arxC* and *arrD*/*arxD*[15, 26]. The genomic organization analysis showed that identified two *aioA*, three *arxA* and 11 of 17 *arrA* genes all have corresponding accessory genes (**Figure 5b**), further confirming their potential identities as arsenotrophic enzymes.

The *aioA*/*arxA* genes are uncommon in soil microbiomes and mostly found in

Proteobacteria[15, 36, 49]. By assigning the taxonomy, *aioA*/*arxA* genes recovered here belonged to Gammaproteobacteria (*n* = 3) and Alphaproteobacteria (*n* = 2), consistent with previous findings (**Figure 3 and Figure 5c**). Nevertheless, 17 *arrA* genes were phylogenetically affiliated with seven distinct bacterial lineages: Bacteroidota (*n* = 1), Chloroflexota (*n* = 4), Deferribacterota (*n* = 1), Desulfobacterota (*n* = 8), Desulfobacterota_I (*n* = 1), Gammaproteobacteria (*n* = 1), and Nitrospirota (*n* = 1) (**Figure 3 and Figure 5c**). Despite that several other bacterial lineages (i.e. Deferribacterota, Firmicutes and Chrysiogenetes) are reported to contain *arrA* genes, most known As(V)-respiring microorganisms are assigned to proteobacterial clades[15, 36, 49]. Our findings, the *arrA*-containing Bacteroidota, Chloroflexota and Nitrospirota, expand the database of putative dissimilatory As(V) reducers.

As respiration are potentially critical to central metabolisms in cold seeps

Microbially mediated As respiration has been verified to influence biogeochemical cycles of carbon and nitrogen, e.g. chemoautotrophic As(III) oxidation coupled with denitrification[50, 51]. Here, functional annotation identified near-complete calvin and reductive TCA carbon fixation pathways in *aioA*/*arxA*-carrying Alphaproteobacteria (*n* = 2) and Gammaproteobacteria MAGs (*n* = 3) (**Figure 5c**). Terminal reductase systems were also recognized in *aioA*/*arxA*-carrying MAGs, i.e. nitrate reductase (*narGHI*). The cooccurrence of these genes suggests that the As(III) oxidation may help support autotrophic carbon fixation and nitrate reduction.

In addition, five *arrA*-carrying MAGs possessed genes for AssA (**Figure 5c**), which mediate the first step of anaerobic activation of alkanes via fumarate addition[52]. Phylogenetic analyses revealed that identified AssA sequences were phylogenetically close to archaea-type and Group V AssA[53] (**Supplementary Figure 5**). These potential hydrocarbon degraders were classified as Chloroflexota (*n* = 2), Deferribacterota (*n* = 1), Desulfobacterota (*n* = 1), and Bacteroidota (*n* = 1). Methane, the simplest hydrocarbon, has been demonstrated to stimulate As(V) respiration during the process of anaerobic oxidation of methane[54]. Similarly, the occurrence of both AssA and ArrA indicated that heterotrophic MAGs stated above may also employ As(V) as electron acceptor for anaerobic degradation of multi-carbon alkanes. Genes encoding carbohydrate-active enzymes (CAZymes) targeting various complex carbohydrates were also present in these arsenotrophic MAGs, including chitin, pectin, starch and plyphenolics (**Figure 5c**).

Notably, arsenotrophic MAGs may function as potential nitrogen fixers introducing new N to local environment. Genes encoding for the catalytic component of nitrogenase (i.e. *nifHDK*) were detected in one *arxA*-carrying (Gammaproteobacteria, *n* = 1) and six *arrA*-carrying (Gammaproteobacteria, *n* = 1; Desulfobacterota, *n* = 4; Desulfobacterota_I, *n* = 1) MAGs (**Figure 5c**). It has been previously reported that As(III) oxidation can fuel biological nitrogen fixation in tailing and metal(loid)-contaminated soils[55, 56]. The data present here further complement that diazotrophs could also fix N_2 using energy obtained from dissimilatory As(V)

reduction.

The metatranscriptomic reads were mapped against arsenotrophic MAGs to depict the gene expression profile at the genome level. Genes for dissimilatory As(V) reduction were transcriptionally active in the Jiaolong seep, as evidenced by the detection of *arrA* transcripts (21.7−340.1 TPM, **Supplementary Data 7**). In addition, transcripts of *assA* (9.2−6306.5 TPM) and *nifH* (155.1−28016.9 TPM) were identified in the Jiaolong seep, Shenhu area and Qiongdongnan Basin, implying anaerobic degradation of hydrocarbons and nitrogen fixation were actively expressed for these arsenotrophic microbes. No transcriptomic sequences related to *arrA*, *assA* and *nifH* genes were detected in the Haima seep. Nevertheless, it does not mean that genes of interest are not transcribed *in situ* because it is difficult to recover enough RNA from deep-sea samples and RNA can get lost during the process of deep-sea sampling[57].

Our findings point towards a previously unrecognized arsenotrophs at seeps, impacting both carbon and nitrogen cycling. However, we acknowledge that cultivation experiments with As-respiring isolates are ultimately needed both to elucidate their lifestyle and confirm functionality for As-dependent carbon fixation, hydrocarbon and carbohydrate degradation as well as nitrogen fixation.

Conclusions

Microbial transformation of As has been well documented and characterized in environments in such as ocean water, groundwater and geothermal springs, but the

knowledge on gene- and genome-level As cycling in deep sea (e.g. cold seep) is limited. Our study demonstrated that As methylation and cytoplasmic As(V) reduction were the predominant detoxification mechanisms employed by cold seep microbiomes. These results substantially expanded the diversity of As detoxification genes to a broader microbial community including Asgardarchaeota and a great number of candidate bacterial phyla. In addition, diverse arsenotrophic lineages are also identified, including Bacteroidota, Chloroflexota, Nitrospirota, etc, which also potentially participate in carbon and nitrogen biogeochemical cycling. This study provides a detailed understanding of As biotransformation in a complex microbiome in deep-sea realms, which could have significant implications for addressing environmental issues. Our result will also provide insights for microbial evolution in the early ocean with harmful metal(loids), e.g. As, as a driving force[58].

Methods

Metagenomic and metatranscriptomic data sets

The 87 metagenomes and 33 metatranscriptomes analyzed in this study are derived from 13 globally distributed cold seep sites (**Supplementary Figure 1**). Among them, 65 metagenomes and 10 metatranscriptomes were compiled from our previous publications[8, 59], and other 22 metagenomes were downloaded from NCBI Sequencing Read Archive (SRA). A detailed description of sampling locations and sequencing information for metagenomic and metatranscriptomic data is given in **Supplementary**

Data 1.

Bioinformatic analyses

DNA reads pre-processing, metagenomic assembly and binning were performed with the function modules of metaWRAP (v1.3.2)[60]. First, the metaWRAP Read_qc module was used to trim raw sequencing DNA reads. Then the filtered DNA reads were individually assembled with the metaWRAP Assembly module using Megahit[61] or metaSPAdes[62] with default settings (detailed assembly statistics are summarized in **Supplementary Data 1**). In addition, metagenomic reads from the same sampling station (n = 10) were also co-assembled using Megahit with the default settings. Thereafter, MAGs were recovered from contigs with the length longer than 1kb using the metaWRAP Binning module (parameters: -maxbin2 -concoct -metabat2) or the VAMB tool[63] (v3.0.1; default parameters; detailed binning statistics are summarized in **Supplementary Data 1**). Further refinement of MAGs was performed by the Bin_refinement module of metaWRAP (parameters: -c 50 -x 10), and CheckM (v1.0.12)[64] was used to estimate the completeness and contamination of these MAGs. All MAGs were dereplicated at 95% average nucleotide identity (ANI) using dRep (v3.4.0; parameters: -comp 50 -con 10)[65] to obtain representative species MAGs. This analysis provided a non-redundant genome set consisting of 1,741 species-level MAGs.

Raw metatranscriptomes were quality filtered with the Read_qc module of metaWRAP (v1.3.2)[60] as described above. The removal of ribosomal RNAs was conducted with

sortmeRNA (v2.1)[66] in the quality-controlled metatranscriptomic reads.

Non-redundant gene catalog construction

Genes were predicted on contigs (≥1kb) from the assemblies using the METABOLIC pipeline (v4.0)[67], which resulted in 33,799,667 protein-coding genes. Clustering of the predicted proteins was performed with MMseqs2 (v13.45111)[68] using the cascaded clustering algorithm at 95% sequence similarity and 90% sequence coverage (parameters: -c 0.95 -min-seq-id 0.95 -cov-mode 1 -cluster-mode 2) following the ref.[69]. This process yielded a total of 17,217,131 non-redundant gene clusters.

Searching for As cycling genes

In this study, 11 well-characterized marker genes[70, 71] were selected to assess their potential influence to the As biogeochemical cycle. These genes include eight As detoxification genes (*acr3*, *arsB*, *arsC1*, *arsC2*, *arsP*, *arsH*, *arsI*, and *arsM*) and three As respiratory genes (*aioA*, *arrA*, and *arxA*). A hidden Markov model (HMM)-based search was performed to identify As-related genes in non-redundant gene catalogue by using hmmsearch function in HMMER package (v3.1b2)[72]. The HMM profile searches and score cutoffs for 11 As-related genes were taken from Lavy et al. (2020)[71].

Taxonomic and functional profiling of MAGs

As-related MAGs were taxonomically annotated using the classify_wf function of the GTDB-Tk toolkit (v2.1.1)[73] with default parameters against the GTDB r207 release. For all MAGs, gene calling and metabolic pathway prediction were conducted with the METABOLIC pipeline (v4.0)[67]. Functional annotation of genomes was also carried out by searching against KEGG, Pfam, MEROPS and dbCAN databases using DRAM (v1.3.5)[74]. The identification of As-related genes in MAGs was performed by searching against As-related HMM profiles from Lavy et al. (2020)[71] as reported above. Genes involved in anaerobic hydrocarbon degradation were screened using BLASTp (identity >30%, coverage >90%, $e < 1 \times 10^{-20}$) against local protein databases[53].

Abundance calculations

For contig level, the relative abundance of genes related to As cycling across 87 metagenomes were calculated from non-redundant gene catalog using the program Salmon (v1.9.0)[75] in the mapping-based mode (parameters: -validateMappings -meta). GPM (genes per million) values were used as a proxy for gene abundance as describe in ref.[74]. For genome level, the relative abundance of each MAG was profiled by mapping quality-trimmed reads from the 87 metagenomes against the MAGs using CoverM in genome mode (https://github.com/wwood/CoverM) (v0.6.1; parameters: -min-read-percent-identity 0.95 -min-read-aligned-percent 0.75 -trim-min 0.10 -trim-max 0.90 -m relative_abundance).

To calculate the transcript abundances of As-related genes, we also mapped clean reads from the 33 metatranscriptomes to non-redundant gene catalog or arsenotrophic MAGs. The transcript abundance of each gene was calculated as the metric-TPM (transcripts per million). GPM or TPM values were normalized based on the gene length and sequencing depth.

Phylogenetic analyses of functional genes

For phylogeny inference, protein sequences of functional genes were aligned with MAFFT (v7.490, -auto option)[76], and gap sequences were trimmed using trimAl (v. 1.2.59, -automated1 option)[77]. Maximum likelihood phylogenetic trees were constructed for each genes using IQ-TREE (v2.12)[78] with the following options: -m TEST -bb 1000 -alrt 1000. Branch support was estimated using 1000 replicates of both ultrafast bootstrap approximation (UFBoot)) and Shimodaira-Hasegawa (SH)-like approximation likelihood ratio (aLRT). Reference protein sequences for As-based respiratory cycle were obtained from Saunders et al. (2019)[34]. Reference protein sequences for fumarate addition were derived from Zhang et al. (2021)[53]. All the tree files were uploaded to Interactive tree of life (iTOL; v6)[79] for visualization and annotation.

Statistical analyses

Statistical analyses were done in R (v4.0.4-v4.1.0) with the following descriptions.

Normality and homoscedasticity of data were evaluated using Shapiro-Wilk test and Levene's test, respectively. One-way analysis of variance (ANOVA) and least significant difference (LSD) test were conducted to evaluate the variations of each gene across different sediment depths and types of cold seeps. The partial least squares discrimination analysis (PLS-DA) was performed based on the GPM values of As-cycling genes with R package 'mixOmics'. The permutational multivariate analysis of variance (PERMANOVA) was employed to test whether As-cycling genes shifted among different sediment depths and types of cold seeps using 'adnois' function in vegan package. All PERMANOVA tests were performed with 9999 permutations based on Bray–Curtis dissimilarity.

DATA AVAILABILITY

Non-redundant gene catalog, assemblies, MAGs containing As cycling genes, and raw tree files have been uploaded to Figshare (https://figshare.com/s/833c3dc27319617e76ed). Arsenotrophic MAGs have also been deposited in NCBI under accession numbers SAMN33581604-33581625 (BioProject ID PRJNA831433).

CODE AVAILABILITY

The present study did not generate codes, and mentioned tools used for the data analysis were applied with default parameters unless specified otherwise.

References

1. Joye, S.B. The geology and biogeochemistry of hydrocarbon seeps. *Annu. Rev. Earth Planet. Sci.* **48**, 205-231 (2020).
2. Feng, D. et al. Cold seep systems in the South China Sea: An overview. *J. Asian Earth SCI.* **168**, 3-16 (2018).
3. Dubilier, N., Bergin, C. & Lott, C. Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat. Rev. Microbiol.* **6**, 725-740 (2008).
4. Levin, L.A. Ecology of cold seep sediments: Interactions of fauna with flow, chemistry and microbes. *Oceanogr. Mar. Biol.* **43**, 11-56 (2005).
5. Dong, X. et al. Thermogenic hydrocarbon biodegradation by diverse depth-stratified microbial populations at a Scotian Basin cold seep. *Nat. Commun.* **11**, 5825-5825 (2020).
6. Kleindienst, S. et al. Diverse sulfate-reducing bacteria of the *Desulfosarcina*/*Desulfococcus* clade are the key alkane degraders at marine seeps. *ISME J.* **8**, 2029-2044 (2014).
7. Reeburgh, W.S. Oceanic methane biogeochemistry. *Chem. Rev.* **107**, 486-513 (2007).
8. Dong, X. et al. Phylogenetically and catabolically diverse diazotrophs reside in deep-sea cold seep sediments. *Nat. Commun*. **13**, 4885 (2022).
9. Wang, Q., Chen, D. & Peckmann, J. Iron shuttle controls on molybdenum, arsenic, and antimony enrichment in Pliocene methane-seep carbonates from the southern Western Foothills, Southwestern Taiwan. *Mar. Pet. Geol.* **100**, 263-269 (2019).
10. Tribovillard, N. et al. Geochemistry of cold seepage-impacted sediments: Per-ascensum or per-descensum trace metal enrichment? *Chem. Geol.* **340**, 1-12 (2013).
11. Hu, Y., Feng, D., Peckmann, J., Roberts, H.H. & Chen, D. New insights into cerium anomalies and mechanisms of trace metal enrichment in authigenic carbonate from hydrocarbon seeps. *Chem. Geol.* **381**, 55-66 (2014).
12. Carvalho, L. et al. Vertical distribution of major, minor and trace elements in sediments from mud volcanoes of the Gulf of Cadiz: evidence of Cd, As and Ba fronts in upper layers. *Deep Sea Res. Pt I. Oceanographic Res. Pap.* **131**, 133-143 (2018).
13. Tribovillard, N. Arsenic in marine sediments: how robust a redox proxy? *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **550**, 109745 (2020).
14. Cangemi, M. et al. Geochemistry and mineralogy of sediments and authigenic carbonates from the Malta Plateau, Strait of Sicily (Central Mediterranean): relationships with mud/fluid release from a mud volcano system. *Chem. Geol.* **276**, 294-308 (2010).
15. Andres, J. & Bertin, P.N. The microbial genomics of arsenic. *FEMS Microbiol. Rev.* **40**, 299-322 (2016).
16. Rahman, M.A. & Hassler, C. Is arsenic biotransformation a detoxification mechanism for microorganisms? *Aquat. Toxicol.* **146**, 212-219 (2014).
17. Masuda, H., Yoshinishi, H., Fuchida, S., Toki, T. & Even, E. Vertical profiles of arsenic and arsenic species transformations in deep-sea sediment, Nankai Trough, offshore Japan. *Prog. Earth Planet. Sc.* **6**, 28 (2019).
18. Sforna, M.C. et al. Evidence for arsenic metabolism and cycling by microorganisms 2.7 billion years ago. *Nat. Geosci.* **7**, 811-815 (2014).

19. Chen, S.-C. et al. The Great Oxidation Event expanded the genetic repertoire of arsenic metabolism and cycling. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10414 (2020).
20. Rosen, B.P. Biochemistry of arsenic detoxification. *FEBS Lett.* **529**, 86-92 (2002).
21. Mukhopadhyay, R. & Rosen, B.P. Arsenate reductases in prokaryotes and eukaryotes. *Environ. Health Perspect.* **110**, 745-748 (2002).
22. Qin, J. et al. Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2075-2080 (2006).
23. Chen, J., Madegowda, M., Bhattacharjee, H. & Rosen, B.P. ArsP: a methylarsenite efflux permease. *Mol. Microbiol.* **98**, 625-635 (2015).
24. Chen, J., Bhattacharjee, H. & Rosen, B.P. ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone. *Mol. Microbiol.* **96**, 1042-1052 (2015).
25. Yoshinaga, M. & Rosen, B.P. A C·As lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7701-7706 (2014).
26. Rascovan, N., Maldonado, J., Vazquez, M.P. & Eugenia Farías, M. Metagenomic study of red biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. *ISME J.* **10**, 299-309 (2016).
27. Zhang, S.Y. et al. High arsenic levels increase activity rather than diversity or abundance of arsenic metabolism genes in paddy soils. *Appl. Environ. Microbiol.* **87**, e0138321 (2021).
28. Shi, L.-D. et al. Coupled aerobic methane oxidation and arsenate reduction contributes to soil-arsenic mobilization in agricultural fields. *Environ. Sci. Technol.* (2022).
29. Yin, Z., Ye, L. & Jing, C. Genome-resolved metagenomics and metatranscriptomics reveal that Aquificae dominates arsenate reduction in Tengchong geothermal springs. *Environ. Sci. Technol.* (2022).
30. Hug, K. et al. Microbial contributions to coupled arsenic and sulfur cycling in the acid-sulfide hot spring Champagne Pool, New Zealand. *Front. Microbiol.* **5** (2014).
31. Zhang, S.-Y. et al. Diversity and abundance of arsenic biotransformation genes in paddy soils from southern China. *Environ. Sci. Technol.* **49 7**, 4138-4146 (2015).
32. Castro-Severyn, J. et al. Living to the high extreme: unraveling the composition, structure, and functional insights of bacterial communities thriving in the arsenic-rich Salar de Huasco Altiplanic ecosystem. *Microbiol. Spectr.* **9**, e00444-00421 (2021).
33. Zhang, S.-Y. et al. Land scale biogeography of arsenic biotransformation genes in estuarine wetland. *Environ. Microbiol.* **19**, 2468-2482 (2017).
34. Saunders, J.K., Fuchsman, C.A., McKay, C. & Rocap, G. Complete arsenic-based respiratory cycle in the marine microbial communities of pelagic oxygen-deficient zones. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9925 (2019).
35. Wang, L., Yin, Z. & Jing, C. Metagenomic insights into microbial arsenic metabolism in shallow groundwater of Datong basin, China. *Chemosphere* **245**, 125603 (2020).
36. Dunivin, T.K., Yeh, S.Y. & Shade, A. A global survey of arsenic-related genes in soil microbiomes. *BMC Biol.* **17**, 45 (2019).

37. Zhou, Y.-L., Mara, P., Cui, G.-J., Edgcomb, V.P. & Wang, Y. Microbiomes in the Challenger Deep slope and bottom-axis sediments. *Nat. Commun.* **13**, 1515 (2022).

38. Jørgensen, B.B. & Boetius, A. Feast and famine — microbial life in the deep-sea bed. *Nat. Rev. Microbiol.* **5**, 770-781 (2007).

39. Orcutt, B.N., Sylvan, J.B., Knab, N.J. & Edwards, K.J. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiol. Mol. Biol. Rev.* **75**, 361-422 (2011).

40. Müller, A.L., Kjeldsen, K.U., Rattei, T., Pester, M. & Loy, A. Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J* **9**, 1152-1165 (2015).

41. Chen, J., Yoshinaga, M. & Rosen, B.P. The antibiotic action of methylarsenite is an emergent property of microbial communities. *Mol. Microbiol.* **111**, 487-494 (2019).

42. Achour, A.R., Bauda, P. & Billard, P. Diversity of arsenite transporter genes from arsenic-resistant soil bacteria. *Res. Microbiol.* **158**, 128-137 (2007).

43. Ruff, S.E. et al. Global dispersion and local diversification of the methane seep microbiome. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4015 (2015).

44. Inagaki, F. et al. Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2815 (2006).

45. Yamaguchi, N., Ohkura, T., Takahashi, Y., Maejima, Y. & Arao, T. Arsenic distribution and speciation near rice roots influenced by iron plaques and redox conditions of the soil matrix. *Environ. Sci. Technol.* **48**, 1549-1556 (2014).

46. Cai, L., Liu, G., Rensing, C. & Wang, G. Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. *BMC Microbiol.* **9**, 4 (2009).

47. Silver, S. & Phung Le, T. Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Appl. Environ. Microbiol.* **71**, 599-608 (2005).

48. Oremland, R.S., Saltikov, C.W., Stolz, J.F. & Hollibaugh, J.T. Autotrophic microbial arsenotrophy in arsenic-rich soda lakes. *FEMS Microbiol. Lett.* **364**, fnx146 (2017).

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

50. Zhang, J. et al. Anaerobic arsenite oxidation by an autotrophic arsenite-oxidizing bacterium from an arsenic-contaminated paddy soil. *Environ. Sci. Technol.* **49**, 5956-5964 (2015).

51. Rhine, E.D., Phelps, C.D. & Young, L.Y. Anaerobic arsenite oxidation by novel denitrifying isolates. *Environ. Microbiol.* **8**, 899-908 (2006).

52. Callaghan, A.V., Wawrik, B., Ní Chadhain, S.M., Young, L.Y. & Zylstra, G.J. Anaerobic alkane-degrading strain AK-01 contains two alkylsuccinate synthase genes. *Biochem. Biophys. Res. Commun.* **366**, 142-148 (2008).

53. Zhang, C. et al. Marine sediments harbor diverse archaea and bacteria with the potential for anaerobic hydrocarbon degradation via fumarate addition. *FEMS Microbiol. Ecol.* **97**, fiab045 (2021).

54. Shi, L.-D. et al. Coupled anaerobic methane oxidation and reductive arsenic mobilization in wetland soils. *Nat. Geosci.* **13**, 799-805 (2020).

55. Li, Y. et al. *Serratia* spp. are responsible for nitrogen fixation fueled by As(III) oxidation, a

novel biogeochemical process identified in mine tailings. *Environ. Sci. Technol.* **56**, 2033-2043 (2022).

56. Li, Y. et al. *Thiobacillus* spp. and *Anaeromyxobacter* spp. mediate arsenite oxidation-dependent biological nitrogen fixation in two contrasting types of arsenic-contaminated soils. *J. Hazard. Mater.* **443**, 130220 (2022).

57. Orsi, W.D., Edgcomb, V.P., Christman, G.D. & Biddle, J.F. Gene expression in the deep biosphere. *Nature* **499**, 205-208 (2013).

58. Li, Y.P. et al. Antimicrobial activity of metals and metalloids. *Annu. Rev. Microbiol.* **75**, 175-197 (2021).

59. Zhang, C. et al. The majority of microorganisms in gas hydrate-bearing subseafloor sediments ferment macromolecules. *Microbiome* **11**, 37 (2023).

60. Uritskiy, G.V., DiRuggiero, J. & Taylor, J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158 (2018).

61. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674-1676 (2015).

62. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P.A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824-834 (2017).

63. Nissen, J.N. et al. Improved metagenome binning and assembly using deep variational autoencoders. *Nat. Biotechnol.* **39**, 555-560 (2021).

64. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P. & Tyson, G.W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043-1055 (2015).

65. Olm, M.R., Brown, C.T., Brooks, B. & Banfield, J.F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864 (2017).

66. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211-3217 (2012).

67. Zhou, Z., Tran, P., Liu, Y., Kieft, K. & Anantharaman, K. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome* **10**, 33 (2022).

68. Steinegger, M. & Söding, J. Clustering huge protein sequence sets in linear time. *Nat. Commun.* **9**, 2542 (2018).

69. Delgado, L.F. & Andersson, A.F. Evaluating metagenomic assembly approaches for biome-specific gene catalogues. *Microbiome* **10**, 72 (2022).

70. Keren, R. et al. Global genomic analysis of microbial biotransformation of arsenic highlights the importance of arsenic methylation in environmental and human microbiomes. *Comput. Struct. Biotechnol. J.* **20**, 559-572 (2022).

71. Lavy, A. et al. Taxonomically and metabolically distinct microbial communities with depth and across a hillslope to riparian zone transect. Preprint at https://doi.org/10.1101/768572 (2020).

72. Johnson, L.S., Eddy, S.R. & Portugaly, E. Hidden Markov model speed heuristic and iterative

HMM search procedure. *BMC Bioinformatics* **11**, 431 (2010).

73. Chaumeil, P.-A., Mussig, A., Philip, H. & Parks, D. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925-1927 (2019).

74. Shaffer, M. et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* **48**, 8883-8900 (2020).

75. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417-419 (2017).

76. Katoh, K. & Standley, D.M. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* **32**, 1933-1942 (2016).

77. Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).

78. Nguyen, L.T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol Evol.* **32**, 268-274 (2015).

79. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242-W245 (2016).

ACKNOWLEDGMENTS

This study was supported by the Scientific Research Foundation of Third Institute of Oceanography, MNR (No. 2022025), State Key Laboratory of Marine Geology, Tongji University (No. MGK202303), China Postdoctoral Science Foundation (2022M723709), the Science and Technology Projects in Guangzhou (No. 202102020970), the Guangdong Basic and Applied Basic Research Foundation (No. 2019B030302004, 20201910240000691) and the Marine Geological Survey Program of China Geological Survey (DD20221706).

COMPETING INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

XD designed this study with input from JL. CZ and XL analyzed omic data. XD, CZ, XL, LDS and ZS interpreted data. JL and XX contributed to data collection. XD and CZ drafted the paper, with input from all other authors.

Figure Legends

Figure 1. Diagram of the microbial transformations of As. As(III), arsenite; As(V), arsenate; MMAs(III), trivalent methylarsenite; MMAs(V), pentavalentmethylarsenate. As(III) efflux permease: ArsB/Acr3; cytoplasmic As(V) reductase: ArsC; respiratory As(V) reductase: ArrA; As(III) oxidase: AioA/ArxA; As(III) S-adenosylmethionine (SAM) methyltransferase: ArsM; C-As lyase: ArsI; MAs(III) efflux permease: ArsP; MAs(III)-specific oxidase: ArsH.

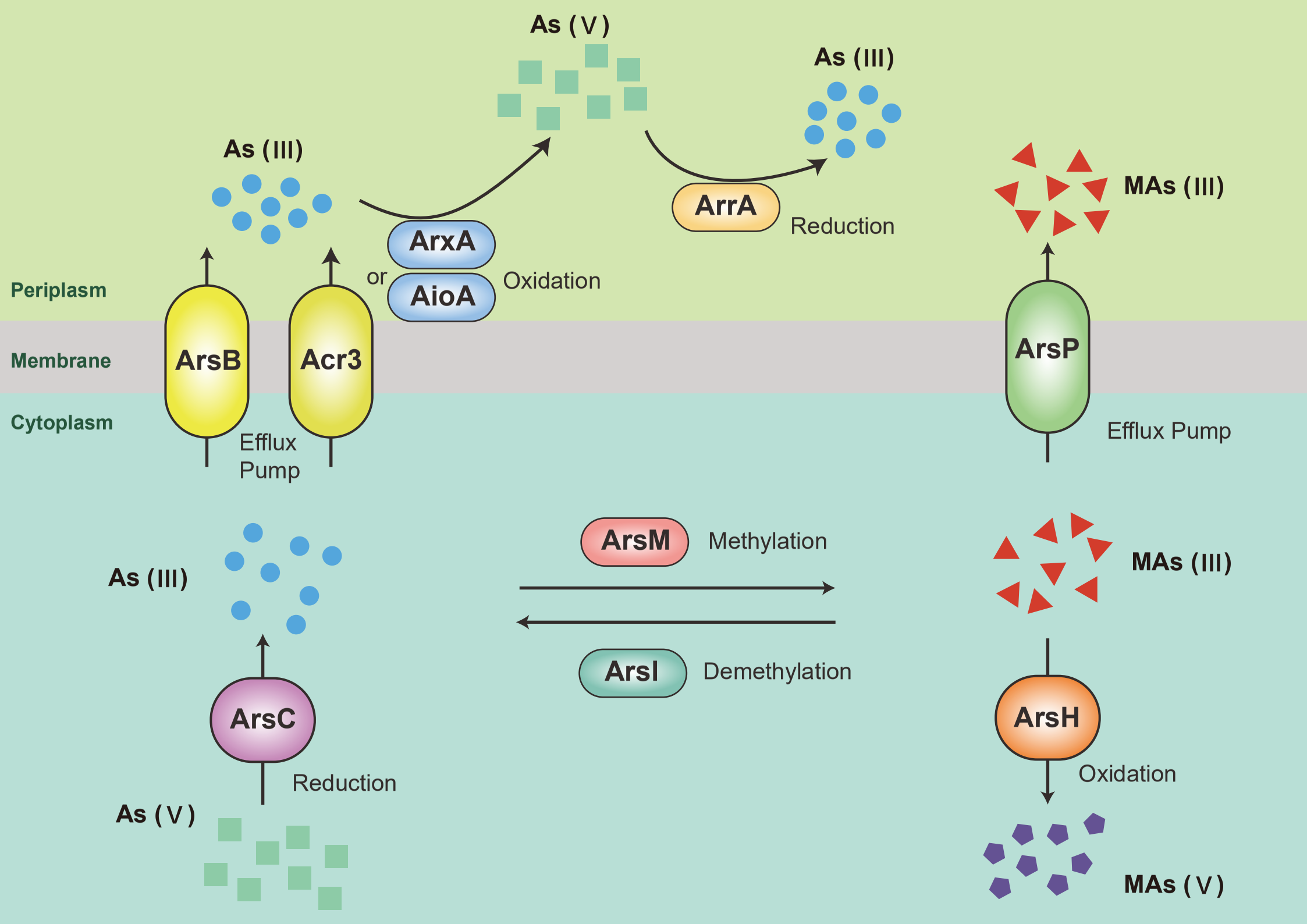
Figure 2. The global distribution of potential genes involved in As cycling at cold seeps. (a) The abundances of As cycling genes across the 87 cold seep metagenomes. The abundance of each gene was normalized by the gene length and sequencing depth and represented as GPM (genes per million) value. (b) The partial least squares discrimination analysis (PLS-DA) plots based on the abundances of As-cycling genes (n = 87). Similarity values among the samples of different sediment depths and types of cold seep were examined using a 999-permutation PERMANOVA test. Source data is available in **Supplementary Data 2**.

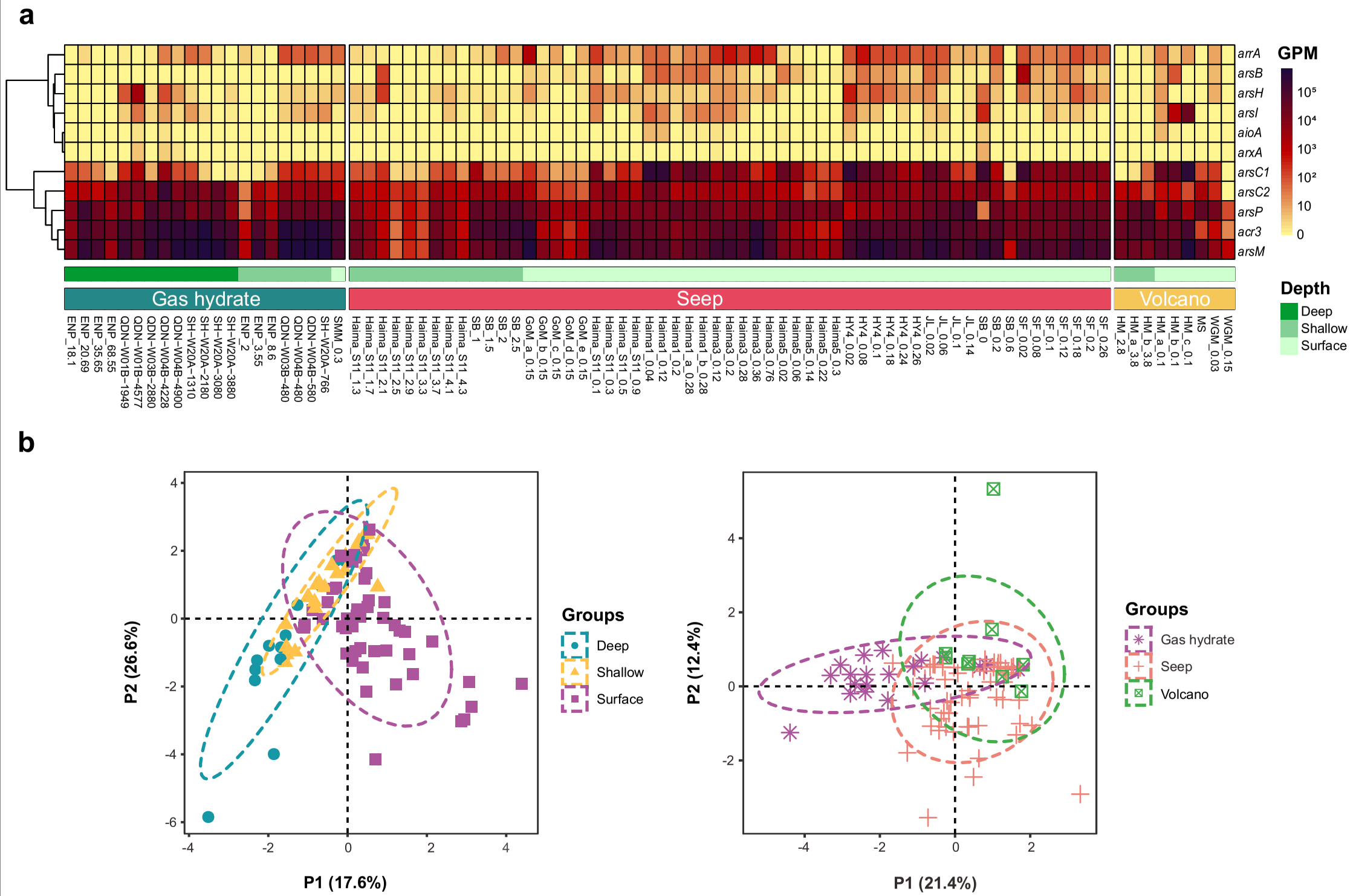
Figure 3. The community structures of microbiome involved in As cycling at cold seeps. The relative abundance of each MAG was estimated using CoverM. The compositions of microbiome involved in As cycling across different types of cold seep were clustered based the Bray–Curtis distance. Orange and red asterisks denote

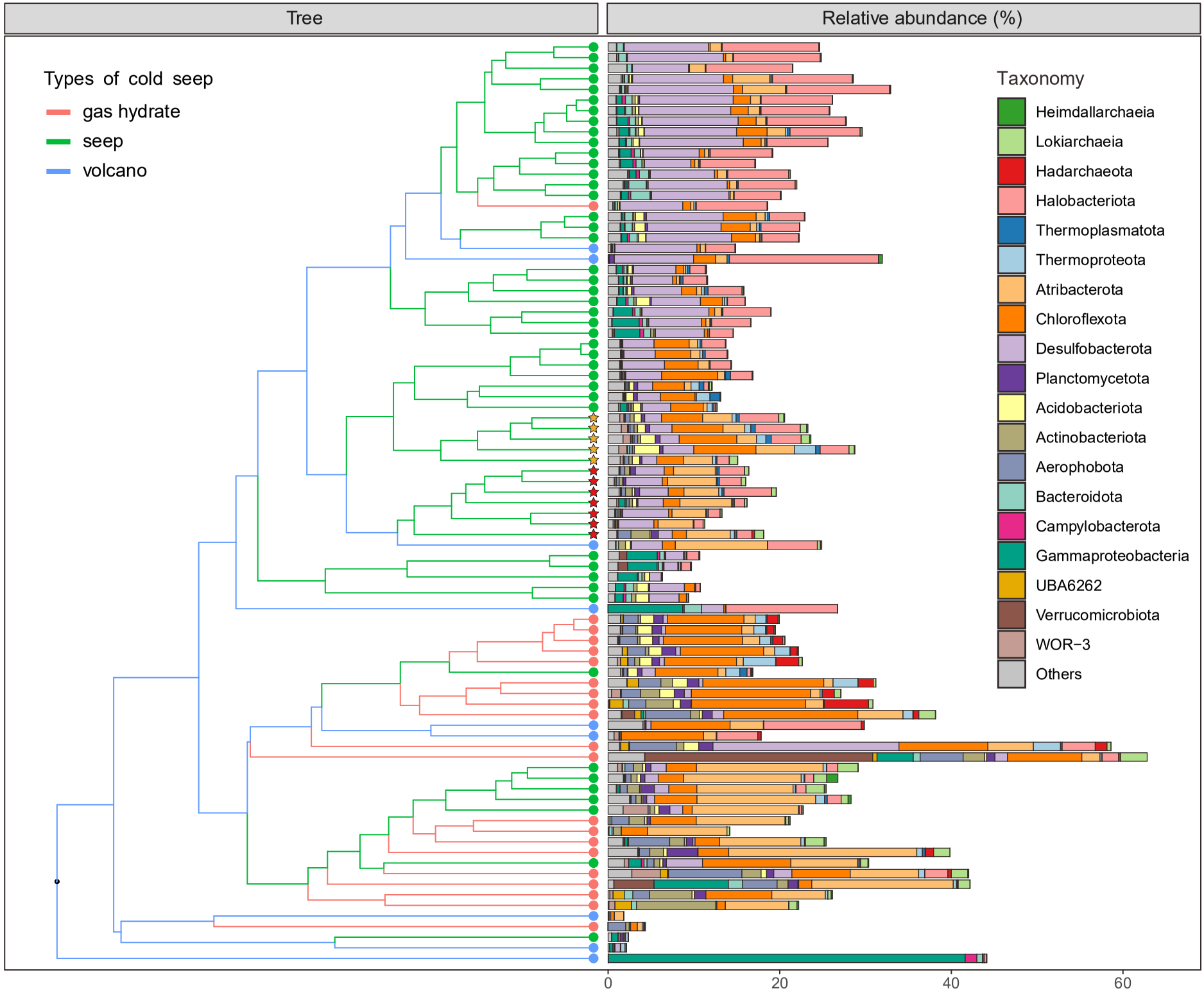
samples with lower and higher concentrations of Fe(II), respectively. Detailed statistics for As-related microbiome are provided in **Supplementary Data 4**.

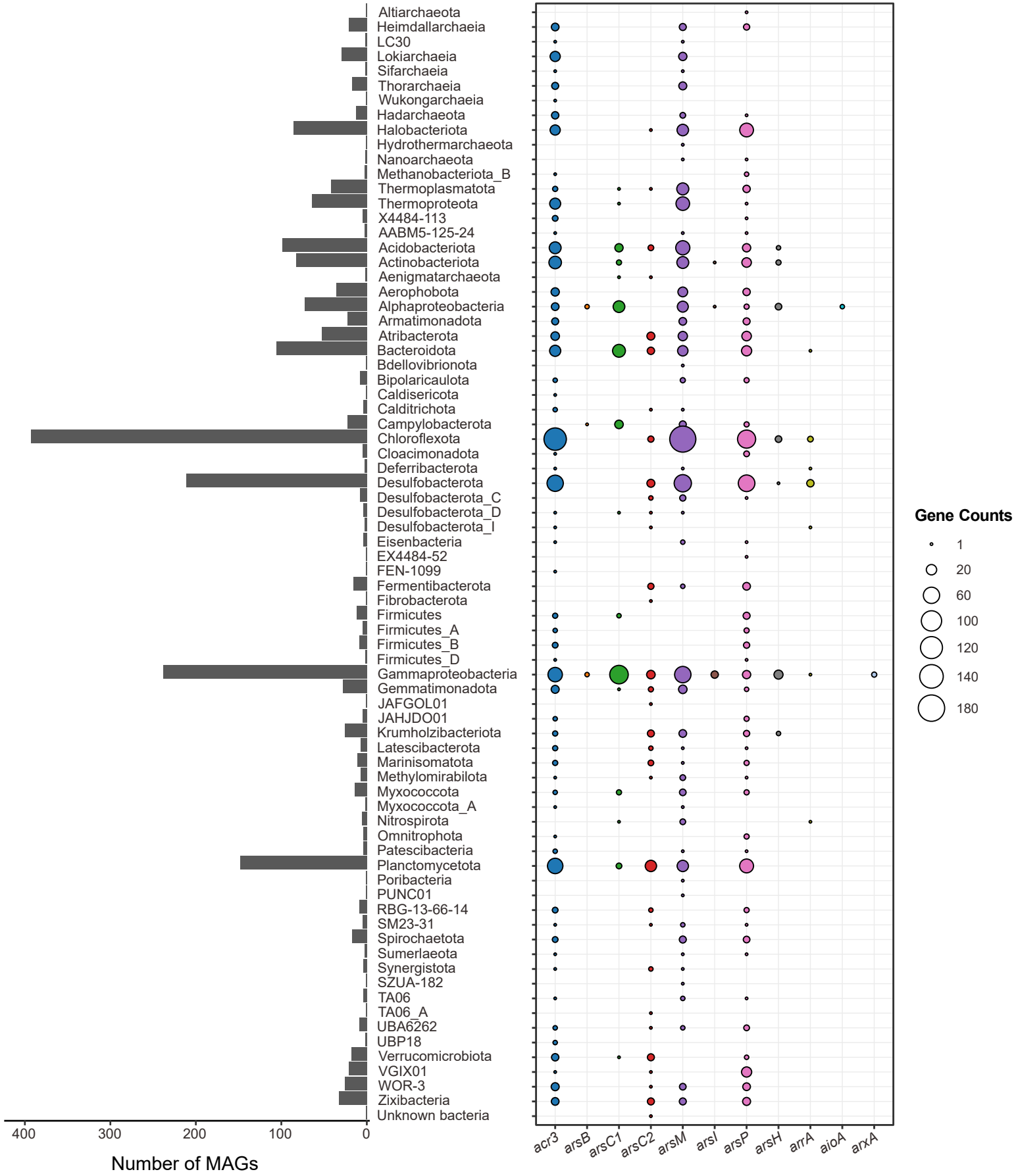
Figure 4. Phylogenetic distribution of As-cycling genes. Left bar plot showing the total number of genomes encoded in each phylogenetic cluster assigned by GTDB-Tk based on GTDB r207 release. Right bubble plot showing the number of As-cycling genes encoded within each phylogenetic cluster. Detailed information on phylogenetic diversity of As-cycling genes is provided in **Supplementary Data 5**.

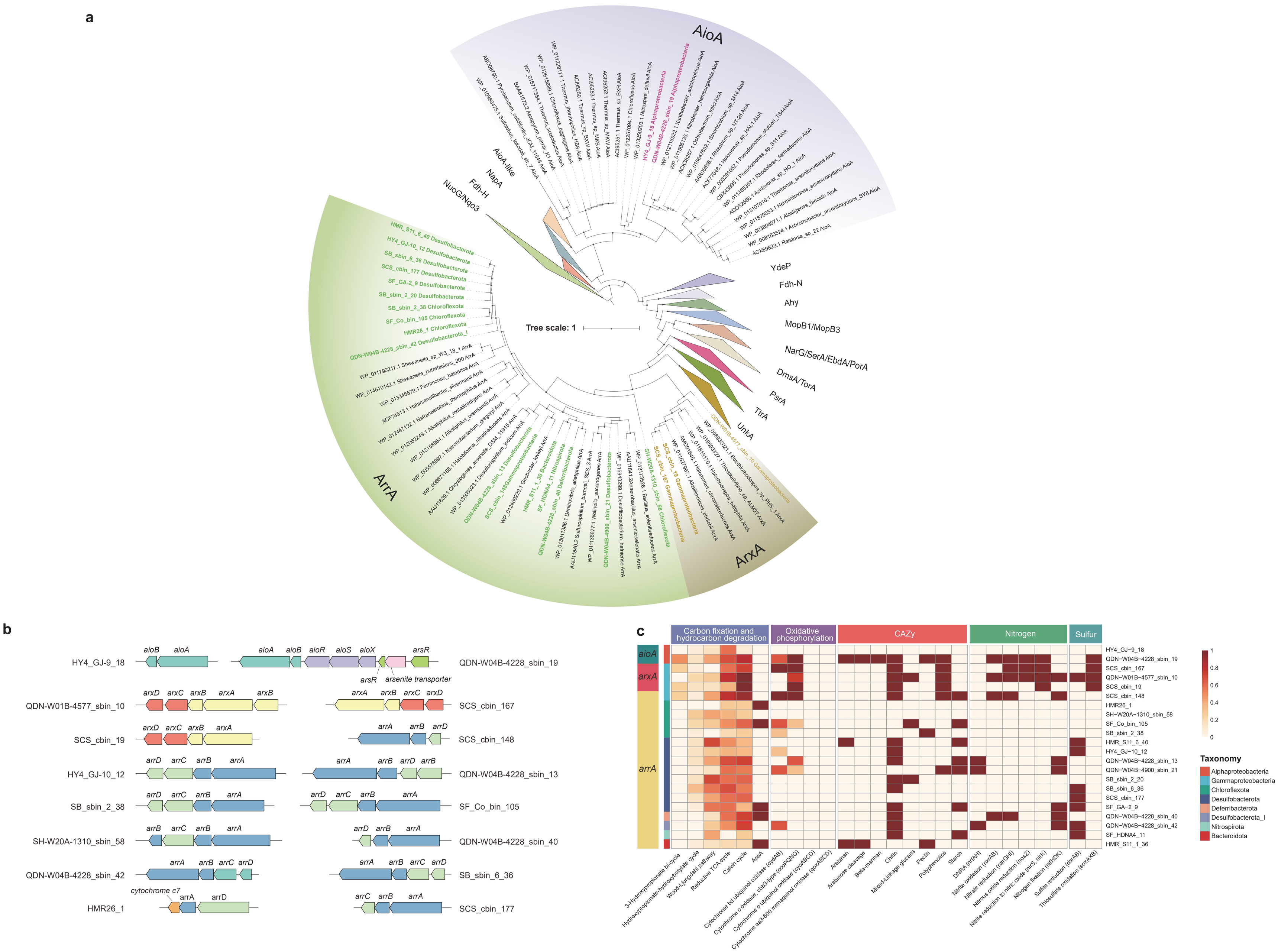
Figure 5. The metabolic potential of the arsenotrophic gene-carrying MAGs. (a) A maximum-likelihood tree of the DMSO reductase family, with protein sequences identified as associated with arsenotrophic enzymes in this study. Bootstrap values are generated from 1000 replicates. Bootstrap values≥70 are shown. Scale bar indicates amino acid substitutions per site (b) The genomic context of the *aioA*, *arxA* and *arrA* clusters in MAGs containing arsenotrophic genes. (c) Heatmap showing the predicted metabolism in potential As-respiring microbes. Detailed annotation is presented in **Supplementary Data 6**. The completeness of each pathway was calculated using the DRAM Distill function.


As (V)
As (III)
As (III)
ArrA
Reduction
MAs (III)
ArxA
or
AioA
Oxidation
Periplasm
Membrane
Cytoplasm
ArsB
Acr3
Efflux
Pump
ArsP
Efflux Pump
As (III)
ArsM
Methylation
MAs (III)
ArsI
Demethylation
ArsC
Reduction
ArsH
Oxidation
As (V)
MAs (V)


a
arrA
arsB
arsH
arsI
aioA
arxA
arsC1
arsC2
arsP
acr3
arsM
GPM
10⁵
10⁴
10³
10²
10
0
Gas hydrate
Seep
Volcano
Depth
Deep
Shallow
Surface
ENP_18.1
ENP_20.69
ENP_35.65
ENP_68.55
ENP_1949
QDN-W01B-4577
QDN-W03B-2880
QDN-W04B-4228
QDN-W04B-4900
SH-W20A-1310
SH-W20A-2180
SH-W20A-3080
SH-W20A-3880
ENP_2
ENP_3.55
ENP_8.6
QDN-W03B-480
QDN-W04B-480
QDN-W04B-580
SH-W20A-766
SMM_0.3
Haima_S11_1.3
Haima_S11_1.7
Haima_S11_2.1
Haima_S11_2.5
Haima_S11_2.9
Haima_S11_3.3
Haima_S11_3.7
Haima_S11_4.1
Haima_S11_4.3
SB_1
SB_1.5
SB_2
SB_2.5
GoM_a_0.15
GoM_b_0.15
GoM_c_0.15
GoM_d_0.15
GoM_e_0.15
Haima_S11_0.1
Haima_S11_0.3
Haima_S11_0.5
Haima_S11_0.9
Haima1_0.04
Haima1_0.12
Haima1_0.2
Haima1_a_0.28
Haima1_b_0.28
Haima3_0.12
Haima3_0.2
Haima3_0.28
Haima3_0.36
Haima3_0.76
Haima5_0.02
Haima5_0.06
Haima5_0.14
Haima5_0.22
Haima5_0.3
HY4_0.02
HY4_0.08
HY4_0.1
HY4_0.18
HY4_0.24
HY4_0.26
JL_0.02
JL_0.06
JL_0.1
JL_0.14
SB_0
SB_0.2
SB_0.6
SF_0.02
SF_0.08
SF_0.1
SF_0.12
SF_0.18
SF_0.2
SF_0.26
HM_2.8
HM_a_3.8
HM_b_3.8
HM_a_0.1
HM_b_0.1
HM_c_0.1
MS
WGM_0.03
WGM_0.15
b
P2 (26.6%)
P1 (17.6%)
Groups
Deep
Shallow
Surface
P2 (12.4%)
P1 (21.4%)
Groups
Gas hydrate
Seep
Volcano


Tree
Relative abundance (%)
Types of cold seep
gas hydrate
seep
volcano
Taxonomy
Heimdallarchaeia
Lokiarchaeia
Hadarchaeota
Halobacteriota
Thermoplasmatota
Thermoproteota
Atribacterota
Chloroflexota
Desulfobacterota
Planctomycetota
Acidobacteriota
Actinobacteriota
Aerophobota
Bacteroidota
Campylobacterota
Gammaproteobacteria
UBA6262
Verrucomicrobiota
WOR−3
Others
0
20
40
60


Altiarchaeota
Heimdallarchaeia
LC30
Lokiarchaeia
Sifarchaeia
Thorarchaeia
Wukongarchaeia
Hadarchaeota
Halobacteriota
Hydrothermarchaeota
Nanoarchaeota
Methanobacteriota_B
Thermoplasmatota
Thermoproteota
X4484-113
AABM5-125-24
Acidobacteriota
Actinobacteriota
Aenigmatarchaeota
Aerophobota
Alphaproteobacteria
Armatimonadota
Atribacterota
Bacteroidota
Bdellovibrionota
Bipolaricaulota
Caldisericota
Calditrichota
Campylobacterota
Chloroflexota
Cloacimonadota
Deferribacterota
Desulfobacterota
Desulfobacterota_C
Desulfobacterota_D
Desulfobacterota_I
Eisenbacteria
EX4484-52
FEN-1099
Fermentibacterota
Fibrobacterota
Firmicutes
Firmicutes_A
Firmicutes_B
Firmicutes_D
Gammaproteobacteria
Gemmatimonadota
JAFGOL01
JAHJDO01
Krumholzibacteriota
Latescibacterota
Marinisomatota
Methylomirabilota
Myxococcota
Myxococcota_A
Nitrospirota
Omnitrophota
Patescibacteria
Planctomycetota
Poribacteria
PUNC01
RBG-13-66-14
SM23-31
Spirochaetota
Sumerlaeota
Synergistota
SZUA-182
TA06
TA06_A
UBA6262
UBP18
Verrucomicrobiota
VGIX01
WOR-3
Zixibacteria
Unknown bacteria
400
300
200
100
0
Number of MAGs
acr3
arsB
arsC1
arsC2
arsM
arsI
arsP
arsH
arrA
aioA
arxA
Gene Counts
1
20
60
100
120
140
180


a
AioA
ArrA
ArxA
Tree scale: 1
AioA-like
NapA
Fdh-H
NuoG/Nqo3
YdeP
Fdh-N
Ahy
MopB1/MopB3
NarG/SerA/EbdA/PorA
DmsA/TorA
PsrA
TtrA
UnkA
b
HY4_GJ-9_18
QDN-W04B-4228_sbin_19
aioB
aioA
aioR
aioS
aioX
arsR
arsenite transporter
QDN-W01B-4577_sbin_10
SCS_cbin_167
arxD
arxC
arxB
arxA
SCS_cbin_19
SCS_cbin_148
arrD
arrC
arrB
arrA
HY4_GJ-10_12
QDN-W04B-4228_sbin_13
SB_sbin_2_38
SF_Co_bin_105
SH-W20A-1310_sbin_58
QDN-W04B-4228_sbin_40
QDN-W04B-4228_sbin_42
SB_sbin_6_36
HMR26_1
cytochrome c7
SCS_cbin_177
c
Carbon fixation and hydrocarbon degradation
Oxidative phosphorylation
CAZy
Nitrogen
Sulfur
aioA
arxA
arrA
3-Hydroxypropionate bi-cycle
Hydroxypropionate-hydroxybutylate cycle
Wood-Ljungdahl pathway
Reductive TCA cycle
Calvin cycle
AssA
Cytochrome bd ubiquinol oxidase (cydAB)
Cytochrome c oxidase, cbb3-type (ccoPQNO)
Cytochrome o ubiquinol oxidase (cyoABCD)
Cytochrome aa3-600 menaquinol oxidase (qoxABCD)
Arabinan
Arabinose cleavage
Beta-mannan
Chitin
Mixed-Linkage glucans
Pectin
Polyphenolics
Starch
DNRA (nrfAH)
Nitrite oxidation (nxrAB)
Nitrate reduction (narGHI)
Nitrous oxide reduction (nosZ)
Nitrite reduction to nitric oxide (nirS, nirK)
Nitrogen fixation (nifHDK)
Sulfite reduction (dsrAB)
Thiosulfate oxidation (soxAXB)
HY4_GJ-9_18
QDN-W04B-4228_sbin_19
SCS_cbin_167
QDN-W01B-4577_sbin_10
SCS_cbin_19
SCS_cbin_148
HMR26_1
SH-W20A-1310_sbin_58
SF_Co_bin_105
SB_sbin_2_38
HMR_S11_6_40
HY4_GJ-10_12
QDN-W04B-4228_sbin_13
QDN-W04B-4900_sbin_21
SB_sbin_2_20
SB_sbin_6_36
SCS_cbin_177
SF_GA-2_9
QDN-W04B-4228_sbin_40
QDN-W04B-4228_sbin_42
SF_HDNA4_11
HMR_S11_1_36
1
0.8
0.6
0.4
0.2
0
Taxonomy
Alphaproteobacteria
Gammaproteobacteria
Chloroflexota
Desulfobacterota
Deferribacterota
Desulfobacterota_I
Nitrospirota
Bacteroidota