Novel and 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 resistance genes (*arsM*, *arsP*, *arsC1*/*arsC*2, *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 resistance (detoxification) to mitigate toxicity and As respiration to conserve energy. The arsenic resistance is mainly achieved by two steps: cytoplasmic reduction of As(V) to As(III) by the ArsC protein

and subsequent extrusion of As(III) via ArsB or Acr3 pump[20, 21] (**Figure 1**). Another As detoxification mechanism involves the methylation of As(III) to methylarsenite (MAs(III)) by the ArsM enzyme[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 ArsP pump[23], oxidated to less toxic MAs(V) by ArsH[24], or demethylated to less toxic As(III) by the ArsI lyase[25]. As respiration consists of the chemolithotrophic oxidation of As(III) by AioAB/ArxAB and dissimilatory As(V) reduction by ArrAB[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 (**Table S1**), 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. For comparison, we also calculated the abundance of *dsrA* genes whose products catalyzing the well-studied and predominated metabolic processes in cold seep ecosystems, i.e. sulfur oxidation and reduction[40]. We found that genes related to As resistance are prevalent in these cold seep samples (**Figure 2**) and their abundances are higher than those of *dsrA* genes (**Figure S2-S3; Table S2**). 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 resistance 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 (**Figure S4**). 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 survival 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 arsenic respiratory oxidation (*aioA*) and reduction (*arrA*/*arxA*) were less abundant in all cold seep samples as compared with arsenic 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, **Figure S4**).

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 (**Figure S2**). 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 as compared with those in seeps and volcanos (**Figure S3**). 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 (**Table S3**). 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 (**Table S4**). Metagenomic read recruitments revealed that the recovered 1083 arsenic-related MAGs accounted for 1.8-62.8% cold seep communities (**Figure 3 and Table S4**). 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, arsenic-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 arsenic-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].

Expanded diversity of microbial lineages containing As resistance genes

Among these As resistance genes, *acr3*, *arsC1*/*arsC2*, *arsM* and *arsP* were widely distributed in bacteria and archaea, while other As resistance 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, 45]. 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 resistance 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, 45]. 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*). In general, these data advance our understanding on the phylogenetical diversity of As resistance genes and highlight the potentially important role played by archaea in As cycling, Asgardarchaeota particularly.

Novel 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*)[46, 47]. 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 resistance 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, 48]. By assigning the taxonomy, *aioA*/*arxA* genes recovered here belong 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, 48]. 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[49, 50]. 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. Phylogenetic analyses revealed that identified AssA sequences were phylogenetically close to archaea-type and Group V AssA[51] (**Figure S5**). 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[52]. 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[53, 54]. The data present here further complement that diazotrophs could also fix N_2 using energy obtained from dissimilatory As(V) reduction.

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 resistance genes to a broader microbial community including Asgardarchaeota and a great number of candidate bacterial phyla. In addition, novel 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. The identification of arsenotrophic microbes will also enable proxies for arsenic-based metabolisms in early anoxic oceans.

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 (**Figure S1**). Among them, 65 metagenomes and 10 metatranscriptomes were compiled from our previous publications[8, 55], 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 **Table S1**.

Bioinformatic analyses

DNA reads pre-processing, metagenomic assembly and binning were performed with the function modules of metaWRAP (v1.3.2)[56]. First, the metaWRAP Read_qc module was used to trim raw sequencing DNA reads. Then the filtered DNA reads were

303 individually assembled with the metaWRAP Assembly module using Megahit[57] or
304 metaSPAdes[58] with default settings (detailed assembly statistics are summarized in
305 Table S1). In addition, metagenomic reads from the same sampling station (n=10) were
306 also co-assembled using Megahit with the default settings. Thereafter, MAGs were
307 recovered from contigs with the length longer than 1kb using the metaWRAP Binning
308 module (parameters: -maxbin2 -concoct -metabat2) or the VAMB tool[59] (v3.0.1;
309 default parameters; detailed binning statistics are summarized in **Table S1**). Further
310 refinement of MAGs was performed by the Bin_refinement module of metaWRAP
311 (parameters: -c 50 -x 10), and CheckM (v1.0.12)[60] was used to estimate the
312 completeness and contamination of these MAGs. All MAGs were dereplicated at 95%
313 average nucleotide identity (ANI) using dRep (v3.4.0; parameters: -comp 50 -con 10)[61]
314 to obtain representative species MAGs. This analysis provided a non-redundant
315 genome set consisting of 1,741 species-level MAGs.

316 Raw metatranscriptomes were quality filtered with the Read_qc module of metaWRAP
317 (v1.3.2)[56] as described above. The removal of ribosomal RNAs was conducted with
318 sortmeRNA (v2.1)[62] in the quality-controlled metatranscriptomic reads.

319 **Non-redundant gene catalog construction**

320 Genes were predicted on contigs (≥1kb) from the assemblies using the METABOLIC
321 pipeline (v4.0)[63], which resulted in 33,799,667 protein-coding genes. Clustering of the
322 predicted proteins was performed with MMseqs2 (v13.45111)[64] using the cascaded
323 clustering algorithm at 95% sequence similarity and 90% sequence coverage
324 (parameters: -c 0.95 -min-seq-id 0.95 -cov-mode 1 -cluster-mode 2) following the ref.[65].
325 This process yielded a total of 17,217,131 non-redundant gene clusters.

326 **Searching for arsenic cycling genes**

327 In this study, 11 well-characterized marker genes[66, 67] were selected to assess their

potential influence to the arsenic biogeochemical cycle. These genes include eight arsenic resistance genes (*acr3*, *arsB*, *arsC1*, *arsC2*, *arsP*, *arsH*, *arsI*, and *arsM*) and three arsenic respiratory genes (*aioA*, *arrA*, and *arxA*). A hidden Markov model (HMM)-based search was performed to identify arsenic-related genes in non-redundant gene catalogue by using hmmsearch function in HMMER package (v3.1b2)[68]. The HMM profile searches and score cutoffs for 11 arsenic-related genes were taken from Lavy et al. (2020)[67].

Taxonomic and functional profiling of MAGs

Arsenic-related MAGs were taxonomically annotated using the classify_wf function of the GTDB-Tk toolkit (v2.1.1)[69] 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)[63]. Functional annotation of genomes was also carried out by searching against KEGG, Pfam, MEROPS and dbCAN databases using DRAM (v1.3.5)[70]. The identification of arsenic-related genes in MAGs was performed by searching against As-related HMM profiles from Lavy et al. (2020)[67] 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[51].

Abundance calculations

For contig level, the relative abundance of genes related to arsenic cycling across 87 metagenomes were calculated from non-redundant gene catalog using the program Salmon (v1.9.0)[71] 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.[70]. 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 abundance of As-related genes, we also mapped clean reads from the 33 metatranscriptomes to non-redundant gene catalog. The transcript abundance of each gene was calculated as the metric-TPM (transcripts per million).

Phylogenetic analyses of functional genes

For phylogeny inference, protein sequences of functional genes were aligned with MAFFT (v7.490, -auto option)[72], and gap sequences were trimmed using trimAl (v. 1.2.59, -automated1 option)[73]. Maximum likelihood phylogenetic trees were constructed for each genes using IQ-TREE (v2.12)[74] 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)[51]. All the tree files were uploaded to Interactive tree of life (iTOL; v6)[75] 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 arsenic-cycling genes with R package 'mixOmics'. The permutational multivariate

analysis of variance (PERMANOVA) was employed to test whether arsenic-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

Genes and MAGs with As cycling genes, and files for the phylogenetic trees were deposited in Figshare (https://doi.org/10.6084/m9.figshare.21550431).

References

1. Joye, S. B., The Geology and Biogeochemistry of Hydrocarbon Seeps. *Annual Review of Earth and Planetary Sciences* **2020,** 48(1), 205-231.
2. Feng, D.; Qiu, J.-W.; Hu, Y.; Peckmann, J.; Guan, H.; Tong, H.; Chen, C.; Chen, J.; Gong, S.; Li, N.; Chen, D., Cold seep systems in the South China Sea: An overview. *Journal of Asian Earth Sciences* **2018,** 168(3-16.
3. Dubilier, N.; Bergin, C.; Lott, C., Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nature Reviews Microbiology* **2008,** 6(10), 725-740.
4. Levin, L. A., ECOLOGY OF COLD SEEP SEDIMENTS: INTERACTIONS OF FAUNA WITH FLOW, CHEMISTRY AND MICROBES. *Oceanography and Marine Biology* **2005,** 43(11-56.
5. Dong, X.; Rattray, J. E.; Campbell, D. C.; Webb, J.; Chakraborty, A.; Adebayo, O.; Matthews, S.; Li, C.; Fowler, M.; Morrison, N. M.; MacDonald, A.; Groves, R. A.; Lewis, I. A.; Wang, S. H.; Mayumi, D.; Greening, C.; Hubert, C. R. J., Thermogenic hydrocarbon biodegradation by diverse depth-stratified microbial populations at a Scotian Basin cold seep. *Nature communications* **2020,** 11(1), 5825-5825.
6. Kleindienst, S.; Herbst, F.-A.; Stagars, M.; von Netzer, F.; von Bergen, M.; Seifert, J.; Peplies, J.; Amann, R.; Musat, F.; Lueders, T.; Knittel, K., Diverse sulfate-reducing bacteria of the *Desulfosarcina*/*Desulfococcus* clade are the key alkane degraders at marine seeps. *The ISME Journal* **2014,** 8(10), 2029-2044.
7. Reeburgh, W. S., Oceanic methane biogeochemistry. *Chemical reviews* **2007,** 107(2), 486-513.
8. Dong, X.; Zhang, C.; Peng, Y.; Zhang, H.-X.; Shi, L.-D.; Wei, G.; Hubert, C. R. J.; Wang, Y.; Greening, C., Phylogenetically and catabolically diverse diazotrophs reside in deep-sea cold seep sediments. *Nature Communications* **2022,** 13(1), 4885.
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. *Marine and Petroleum Geology* **2019,** 100(263-269.
10. Tribovillard, N.; du Châtelet, E. A.; Gay, A.; Barbecot, F.; Sansjofre, P.; Potdevin, J.-L., Geochemistry of cold seepage-impacted sediments: Per-ascensum or per-descensum trace metal enrichment? *Chemical Geology* **2013,** 340(1-12.

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. *Chemical Geology* **2014,** 381(55-66.

12. Carvalho, L.; Monteiro, R.; Figueira, P.; Mieiro, C.; Almeida, J.; Pereira, E.; Magalhães, V.; Pinheiro, L.; Vale, C., 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 Research Part I: Oceanographic Research Papers* **2018,** 131(133-143.

13. Tribovillard, N., Arsenic in marine sediments: how robust a redox proxy? *Palaeogeography, Palaeoclimatology, Palaeoecology* **2020,** 550(109745.

14. Cangemi, M.; Di Leonardo, R.; Bellanca, A.; Cundy, A.; Neri, R.; Angelone, M., 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. *Chemical Geology* **2010,** 276(3), 294-308.

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

16. Rahman, M. A.; Hassler, C., Is arsenic biotransformation a detoxification mechanism for microorganisms? *Aquatic Toxicology* **2014,** 146(212-219.

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. *Progress in Earth and Planetary Science* **2019,** 6(1), 28.

18. Sforna, M. C.; Philippot, P.; Somogyi, A.; van Zuilen, M. A.; Medjoubi, K.; Schoepp-Cothenet, B.; Nitschke, W.; Visscher, P. T., Evidence for arsenic metabolism and cycling by microorganisms 2.7 billion years ago. *Nature Geoscience* **2014,** 7(11), 811-815.

19. Chen, S.-C.; Sun, G.-X.; Yan, Y.; Konstantinidis, K. T.; Zhang, S.-Y.; Deng, Y.; Li, X.-M.; Cui, H.-L.; Musat, F.; Popp, D.; Rosen, B. P.; Zhu, Y.-G., The Great Oxidation Event expanded the genetic repertoire of arsenic metabolism and cycling. *Proceedings of the National Academy of Sciences* **2020,** 117(19), 10414.

20. Rosen, B. P., Biochemistry of arsenic detoxification. *FEBS Letters* **2002,** 529(1), 86-92.

21. Mukhopadhyay, R.; Rosen, B. P., Arsenate reductases in prokaryotes and eukaryotes. *Environmental health perspectives* **2002,** 110 Suppl 5(Suppl 5), 745-8.

22. Qin, J.; Rosen, B. P.; Zhang, Y.; Wang, G.; Franke, S.; Rensing, C., Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proceedings of the National Academy of Sciences* **2006,** 103(7), 2075-2080.

23. Chen, J.; Madegowda, M.; Bhattacharjee, H.; Rosen, B. P., ArsP: a methylarsenite efflux permease. *Molecular Microbiology* **2015,** 98(4), 625-635.

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. *Molecular Microbiology* **2015,** 96(5), 1042-1052.

25. Yoshinaga, M.; Rosen, B. P., A C As lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters. *Proceedings of the National Academy of Sciences* **2014,** 111(21), 7701-7706.

26. Rascovan, N.; Maldonado, J.; Vazquez, M. P.; Eugenia Farías, M., Metagenomic study of red

456 biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. *The ISME Journal*
457 **2016,** 10(2), 299-309.
458 27. Zhang, S. Y.; Xiao, X.; Chen, S. C.; Zhu, Y. G.; Sun, G. X.; Konstantinidis, K. T., High Arsenic
459 Levels Increase Activity Rather than Diversity or Abundance of Arsenic Metabolism Genes in Paddy
460 Soils. *Appl Environ Microbiol* **2021,** 87(20), e0138321.
461 28. Shi, L.-D.; Zhou, Y.-J.; Tang, X.-J.; Kappler, A.; Chistoserdova, L.; Zhu, L.-Z.; Zhao, H.-P.,
462 Coupled Aerobic Methane Oxidation and Arsenate Reduction Contributes to Soil-Arsenic Mobilization
463 in Agricultural Fields. *Environmental Science & Technology* **2022**.
464 29. Yin, Z.; Ye, L.; Jing, C., Genome-Resolved Metagenomics and Metatranscriptomics Reveal that
465 Aquificae Dominates Arsenate Reduction in Tengchong Geothermal Springs. *Environmental Science &*
466 *Technology* **2022**.
467 30. Hug, K.; Maher, W. A.; Stott, M. B.; Krikowa, F.; Foster, S.; Moreau, J. W., Microbial
468 contributions to coupled arsenic and sulfur cycling in the acid-sulfide hot spring Champagne Pool, New
469 Zealand. *Frontiers in Microbiology* **2014,** 5(
470 31. Zhang, S.-Y.; Zhao, F.-j.; Sun, G.; Su, J.; Yang, X.-R.; Li, H.; Zhu, Y.-g., Diversity and abundance
471 of arsenic biotransformation genes in paddy soils from southern China. *Environmental science &*
472 *technology* **2015,** 49 7(4138-46.
473 32. Castro-Severyn, J.; Pardo-Esté, C.; Mendez Katterinne, N.; Fortt, J.; Marquez, S.; Molina, F.;
474 Castro-Nallar, E.; Remonsellez, F.; Saavedra Claudia, P., Living to the High Extreme: Unraveling the
475 Composition, Structure, and Functional Insights of Bacterial Communities Thriving in the
476 Arsenic-Rich Salar de Huasco Altiplanic Ecosystem. *Microbiology Spectrum* **2021,** 9(1), e00444-21.
477 33. Zhang, S.-Y.; Su, J.-Q.; Sun, G.-X.; Yang, Y.; Zhao, Y.; Ding, J.; Chen, Y.-S.; Shen, Y.; Zhu, G.;
478 Rensing, C.; Zhu, Y.-G., Land scale biogeography of arsenic biotransformation genes in estuarine
479 wetland. *Environmental Microbiology* **2017,** 19(6), 2468-2482.
480 34. Saunders, J. K.; Fuchsman, C. A.; McKay, C.; Rocap, G., Complete arsenic-based respiratory
481 cycle in the marine microbial communities of pelagic oxygen-deficient zones. *Proceedings of the*
482 *National Academy of Sciences* **2019,** 116(20), 9925.
483 35. Wang, L.; Yin, Z.; Jing, C., Metagenomic insights into microbial arsenic metabolism in shallow
484 groundwater of Datong basin, China. *Chemosphere* **2020,** 245(125603.
485 36. Dunivin, T. K.; Yeh, S. Y.; Shade, A., A global survey of arsenic-related genes in soil microbiomes.
486 *BMC Biology* **2019,** 17(1), 45.
487 37. Zhou, Y.-L.; Mara, P.; Cui, G.-J.; Edgcomb, V. P.; Wang, Y., Microbiomes in the Challenger Deep
488 slope and bottom-axis sediments. *Nature Communications* **2022,** 13(1), 1515.
489 38. Jørgensen, B. B.; Boetius, A., Feast and famine — microbial life in the deep-sea bed. *Nature*
490 *Reviews Microbiology* **2007,** 5(10), 770-781.
491 39. Orcutt, B. N.; Sylvan, J. B.; Knab, N. J.; Edwards, K. J., Microbial ecology of the dark ocean
492 above, at, and below the seafloor. *Microbiol Mol Biol Rev* **2011,** 75(2), 361-422.
493 40. Müller, A. L.; Kjeldsen, K. U.; Rattei, T.; Pester, M.; Loy, A., Phylogenetic and environmental
494 diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *Isme j* **2015,** 9(5), 1152-65.
495 41. Chen, J.; Yoshinaga, M.; Rosen, B. P., The antibiotic action of methylarsenite is an emergent
496 property of microbial communities. *Molecular microbiology* **2019,** 111(2), 487-494.
497 42. Achour, A. R.; Bauda, P.; Billard, P., Diversity of arsenite transporter genes from arsenic-resistant

soil bacteria. *Research in microbiology* **2007,** 158(2), 128-37.

43. Ruff, S. E.; Biddle, J. F.; Teske, A. P.; Knittel, K.; Boetius, A.; Ramette, A., Global dispersion and local diversification of the methane seep microbiome. *Proceedings of the National Academy of Sciences* **2015,** 112(13), 4015.

44. Inagaki, F.; Nunoura, T.; Nakagawa, S.; Teske, A.; Lever, M.; Lauer, A.; Suzuki, M.; Takai, K.; Delwiche, M.; Colwell, F. S.; Nealson, K. H.; Horikoshi, K.; D'Hondt, S.; Jørgensen, B. B., Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103(8), 2815.

45. Cai, L.; Liu, G.; Rensing, C.; Wang, G., Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. *BMC Microbiology* **2009,** 9(1), 4.

46. Silver, S.; Phung Le, T., Genes and Enzymes Involved in Bacterial Oxidation and Reduction of Inorganic Arsenic. *Applied and Environmental Microbiology* **2005,** 71(2), 599-608.

47. Oremland, R. S.; Saltikov, C. W.; Stolz, J. F.; Hollibaugh, J. T., Autotrophic microbial arsenotrophy in arsenic-rich soda lakes. *FEMS Microbiology Letters* **2017,** 364(15), fnx146.

48. van Lis, R.; Nitschke, W.; Duval, S.; Schoepp-Cothenet, B., Arsenics as bioenergetic substrates. *Biochimica et biophysica acta* **2013,** 1827(2), 176-88.

49. Zhang, J.; Zhou, W.; Liu, B.; He, J.; Shen, Q.; Zhao, F. J., Anaerobic arsenite oxidation by an autotrophic arsenite-oxidizing bacterium from an arsenic-contaminated paddy soil. *Environ Sci Technol* **2015,** 49(10), 5956-64.

50. Rhine, E. D.; Phelps, C. D.; Young, L. Y., Anaerobic arsenite oxidation by novel denitrifying isolates. *Environ Microbiol* **2006,** 8(5), 899-908.

51. Zhang, C.; Meckenstock, R. U.; Weng, S.; Wei, G.; Hubert, C. R. J.; Wang, J.-H.; Dong, X., Marine sediments harbor diverse archaea and bacteria with the potential for anaerobic hydrocarbon degradation via fumarate addition. *FEMS Microbiology Ecology* **2021,** 97(5), fiab045.

52. Shi, L.-D.; Guo, T.; Lv, P.-L.; Niu, Z.-F.; Zhou, Y.-J.; Tang, X.-J.; Zheng, P.; Zhu, L.-Z.; Zhu, Y.-G.; Kappler, A.; Zhao, H.-P., Coupled anaerobic methane oxidation and reductive arsenic mobilization in wetland soils. *Nature Geoscience* **2020,** 13(12), 799-805.

53. Li, Y.; Guo, L.; Häggblom, M. M.; Yang, R.; Li, M.; Sun, X.; Chen, Z.; Li, F.; Su, X.; Yan, G.; Xiao, E.; Zhang, H.; Sun, W., Serratia spp. Are Responsible for Nitrogen Fixation Fueled by As(III) Oxidation, a Novel Biogeochemical Process Identified in Mine Tailings. *Environ Sci Technol* **2022,** 56(3), 2033-2043.

54. Li, Y.; Guo, L.; Yang, R.; Yang, Z.; Zhang, H.; Li, Q.; Cao, Z.; Zhang, X.; Gao, P.; Gao, W.; Yan, G.; Huang, D.; Sun, W., Thiobacillus spp. and Anaeromyxobacter spp. mediate arsenite oxidation-dependent biological nitrogen fixation in two contrasting types of arsenic-contaminated soils. *Journal of hazardous materials* **2022,** 443(Pt A), 130220.

55. Zhang, C.; Fang, Y.-X.; Yin, X.; Lai, H.; Kuang, Z.; Zhang, T.; Xu, X.-P.; Wegener, G.; Wang, J.-H.; Dong, X., The majority of microorganisms in gas hydrate-bearing subseafloor sediments ferment macromolecules. *bioRxiv* **2022**, 2022.05.19.492412.

56. Uritskiy, G. V.; DiRuggiero, J.; Taylor, J., MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **2018,** 6(1), 158.

57. 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* **2015,** 31(10), 1674-1676.

58. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P. A., metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **2017,** 27(5), 824-834.

59. Nissen, J. N.; Johansen, J.; Allesøe, R. L.; Sønderby, C. K.; Armenteros, J. J. A.; Grønbech, C. H.; Jensen, L. J.; Nielsen, H. B.; Petersen, T. N.; Winther, O.; Rasmussen, S., Improved metagenome binning and assembly using deep variational autoencoders. *Nature Biotechnology* **2021,** 39(5), 555-560.

60. 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* **2015,** 25(7), 1043-1055.

61. 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. *The Isme Journal* **2017,** 11(2864.

62. Kopylova, E.; Noé, L.; Touzet, H., SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **2012,** 28(24), 3211-3217.

63. Zhou, Z.; Tran, P.; Liu, Y.; Kieft, K.; Anantharaman, K., METABOLIC: a scalable high-throughput metabolic and biogeochemical functional trait profiler based on microbial genomes. *bioRxiv* **2019**, 761643.

64. Steinegger, M.; Söding, J., Clustering huge protein sequence sets in linear time. *Nature Communications* **2018,** 9(1), 2542.

65. Delgado, L. F.; Andersson, A. F., Evaluating metagenomic assembly approaches for biome-specific gene catalogues. *Microbiome* **2022,** 10(1), 72.

66. Keren, R.; Méheust, R.; Santini, J. M.; Thomas, A.; West-Roberts, J.; Banfield, J. F.; Alvarez-Cohen, L., Global genomic analysis of microbial biotransformation of arsenic highlights the importance of arsenic methylation in environmental and human microbiomes. *Comput Struct Biotechnol J* **2022,** 20(559-572.

67. Lavy, A.; Matheus Carnevali, P. B.; Keren, R.; Bill, M.; Wan, J.; Tokunaga, T. K.; Williams, K. H.; Hubbard, S. S.; Banfield, J. F., Taxonomically and metabolically distinct microbial communities with depth and across a hillslope to riparian zone transect. *bioRxiv* **2020**, 768572.

68. Johnson, L. S.; Eddy, S. R.; Portugaly, E., Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC bioinformatics* **2010,** 11(431.

69. Chaumeil, P.-A.; Mussig, A.; Philip, H.; Parks, D., GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **2019,** 36(6), 1925-1927.

70. Shaffer, M.; Borton, M. A.; McGivern, B. B.; Zayed, A. A.; La Rosa, Sabina L.; Solden, L. M.; Liu, P.; Narrowe, A. B.; Rodríguez-Ramos, J.; Bolduc, B.; Gazitúa, M. C.; Daly, R. A.; Smith, G. J.; Vik, D. R.; Pope, P. B.; Sullivan, M. B.; Roux, S.; Wrighton, Kelly C., DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* **2020,** 48(16), 8883-8900.

71. Patro, R.; Duggal, G.; Love, M. I.; Irizarry, R. A.; Kingsford, C., Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **2017,** 14(4), 417-419.

72. Katoh, K.; Standley, D. M., A simple method to control over-alignment in the MAFFT multiple

582 sequence alignment program. *Bioinformatics* **2016,** 32(13), 1933-42.
583 73. Capella-Gutiérrez, S.; Silla-Martínez, J. M.; Gabaldón, T., trimAl: a tool for automated alignment
584 trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009,** 25(15), 1972-1973.
585 74. Nguyen, L. T.; Schmidt, H. A.; von Haeseler, A.; Minh, B. Q., IQ-TREE: a fast and effective
586 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **2015,** 32(1),
587 268-74.
588 75. Letunic, I.; Bork, P., Interactive tree of life (iTOL) v3: an online tool for the display and
589 annotation of phylogenetic and other trees. *Nucleic Acids Res* **2016,** 44(W1), W242-W245.
590

ACKNOWLEDGMENTS

This study was supported by the Scientific Research Foundation of Third Institute of Oceanography, MNR (No. 2022025), the National Science Foundation of China (No. 41906076), 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).

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.

COMPETING INTERESTS

The authors declare no competing interests.

Figures Legends

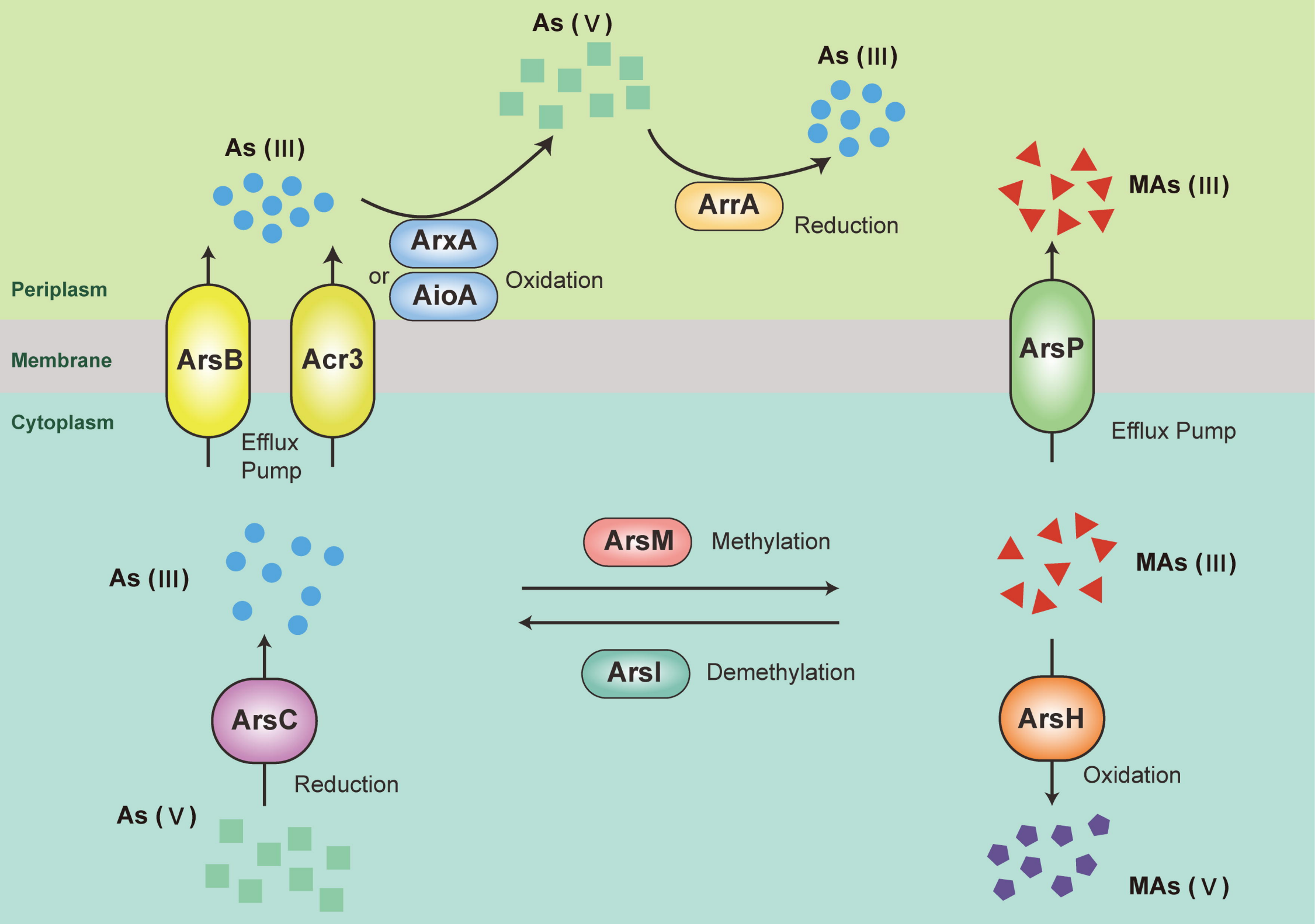
Figure 1. Diagram of the microbial transformations of As. As(III), arsenite; As(V), arsenate; MMAs(III), trivalent methylarsenite; MMAs(V), pentavalentmethylarsenate. As(III) efflux permease: ArsB/Acr3; cytoplasmic As(V) reductase: ArsC; respiratory As(V) reductase: ArrA; As(III) oxidase: AioA/ArxA; As(III) S-adenosylmethionine (SAM) methyltransferase: ArsM; non-heme iron-dependent dioxygenase: ArsI; MAs(III) efflux permease: ArsP; MAs(III) 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. 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 **Table S2**.

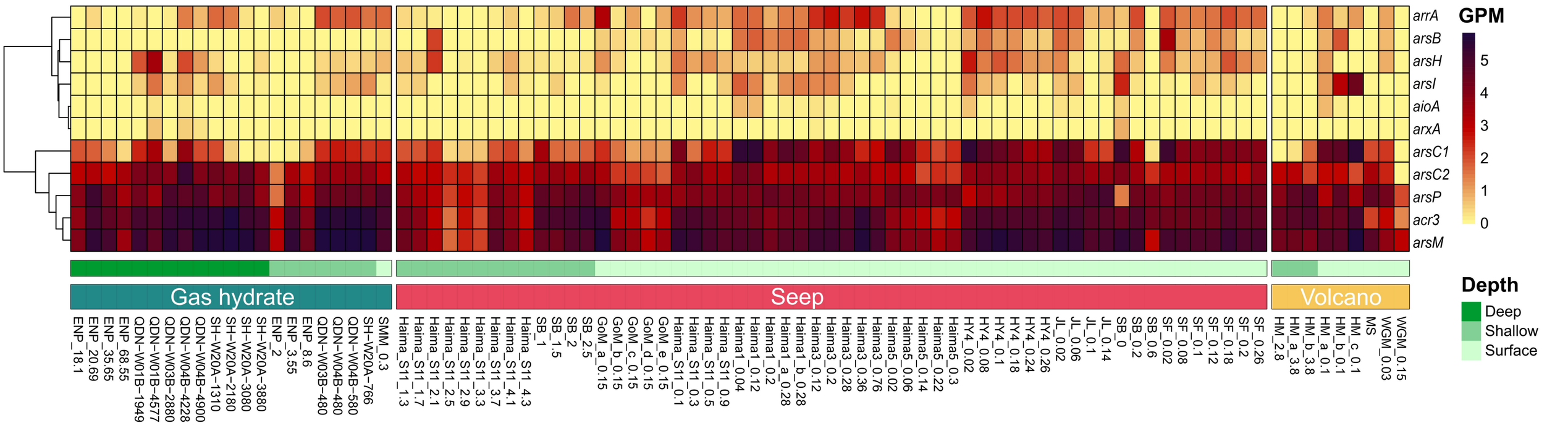
Figure 3. The community structures of microbiome involved in As cycling at cold seeps. The relative abundance of each MAG was estimated using CoverM. The compositions of microbiome involved in As cycling across different types of cold seep were clustered based the Bray-Curtis distance. Detailed statistics for As-related microbiome are provided in **Table S4**.

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 **Table S5**.

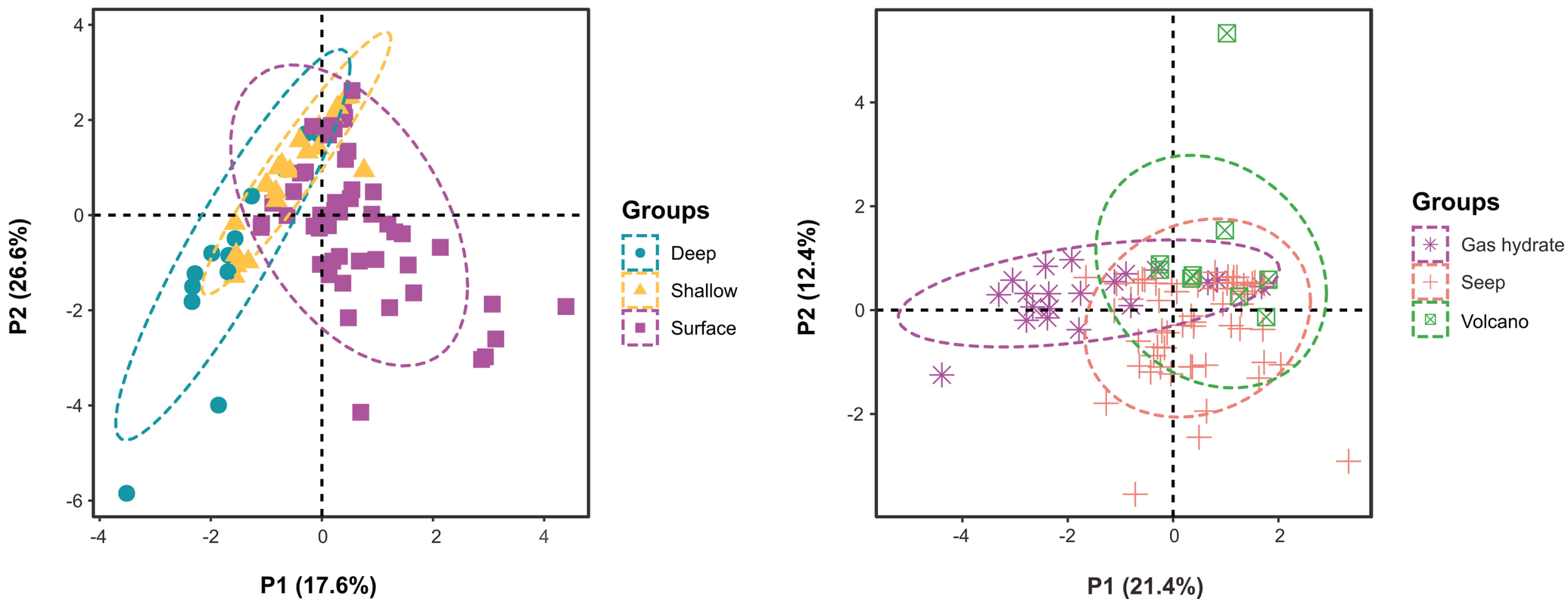
Figure 5. The metabolic potential of the arsenotrophic gene-carrying MAGs. (a) A maximum-likelihood tree of the DMSO reductase family, with protein sequences identified as associated with arsenotrophic enzymes in this study. Bootstrap values are generated from 1000 replicates. Bootstrap values≥70 are 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 **Table S6**. The completeness of each pathway was calculated using the DRAM Distill function.

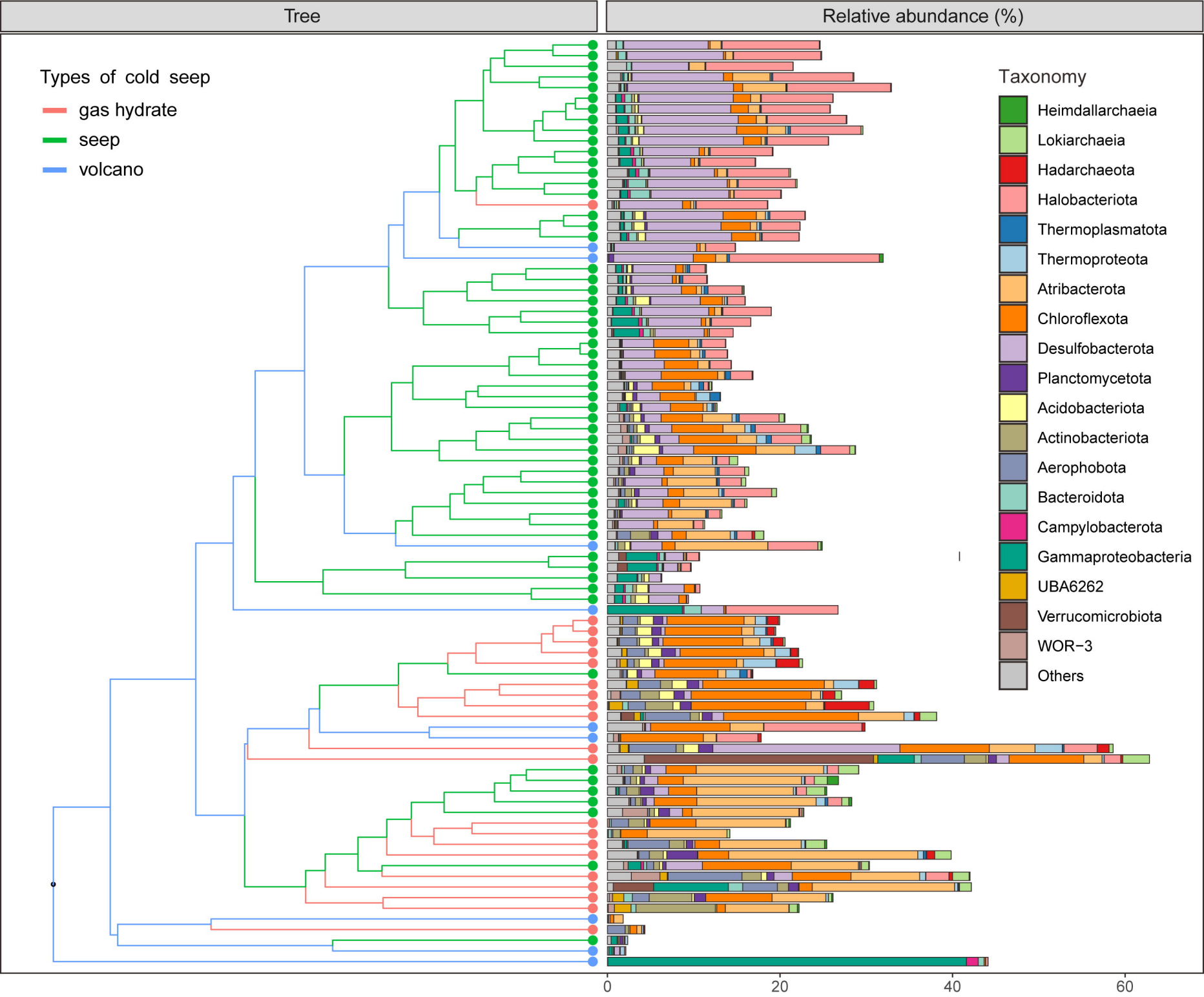

As (V)
As (III)
As (III)
ArrA
Reduction
ArxA
or
AioA
Oxidation
Periplasm
Membrane
Cytoplasm
ArsB
Acr3
Efflux
Pump
MAs (III)
ArsP
Efflux Pump
ArsM
Methylation
ArsI
Demethylation
As (III)
ArsC
Reduction
As (V)
MAs (III)
ArsH
Oxidation
MAs (V)

a



b




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