bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Meta-omics-aided isolation of an elusive anaerobic arsenic-methylating soil bacterium**

2

3 Karen Viacava[1,2], Jiangtao Qiao[1], Andrew Janowczyk[3], Suresh Poudel[4], Nicolas Jacquemin[5],

4 Karin Lederballe Meibom[1], Him K. Shrestha[4,6], Matthew C. Reid[7], Robert L. Hettich[4], Rizlan

5 Bernier-Latmani[1*].

6

7 [1] Environmental Microbiology Laboratory, School of Architecture, Civil and Environmental

8 Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

9 [2] Soil Science Group, Institute of Geography, University of Bern, Bern, Switzerland.

10 [3] Bioinformatics Core Facility, Swiss Institute of Bioinformatics, Lausanne, Switzerland.

11 [4] BioSciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States.

12 [5] Translational Bioinformatics and Statistics, Department of Oncology, Université de

13 Lausanne, Lausanne, Switzerland.

14 [6] Genome Science and Technology Graduate School, University of Tennessee, Knoxville, TN,

15 United States.

16 [7] School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, United States.

17 *Corresponding author. E-mail address rizlan.bernier-latmani@epfl.ch. Postal address EPFL

18 ENAC IIE EML CH A1 375 (Bâtiment CH) Station 6 CH-1015 Lausanne Switzerland. Tel. +41

19 21 69 35001.

20 Competing Interests Statement: The authors declare there are no competing financial

21 interests in relation to the work described. The work was funded by the Swiss National Science

22 Foundation (SNSF) grant 310030_176146-1 and the SNSF NCCR Microbiomes (grant #

23 51NF40_180575). The work at ORNL was conducted under the Plant-Microbe Interface

24 Science Focus Area, as supported by the U.S. Department of Energy, Office of Science, Office

25 of Biological and Environmental Research, Genomic Science Program.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

26 Key words: metagenomics, metaproteomics, metatranscriptomics, *Paraclostridium*, arsenic
27 metabolism.

28

29 **Abstract**

30 Soil microbiomes harbor unparalleled functional and phylogenetic diversity and are sources
31 of novel metabolisms. However, extracting isolates with a targeted function from complex
32 microbiomes is not straightforward, particularly if the associated phenotype does not lend itself
33 to high-throughput screening. Here, we tackle the methylation of arsenic (As) in anoxic soils.
34 By analogy to mercury, As methylation was proposed to be catalyzed by sulfate-reducing
35 bacteria. However, to date, there are no anaerobic isolates capable of As methylation, whether
36 sulfate-reducing or otherwise. The isolation of such a microorganism has been thwarted by
37 the fact that the anaerobic bacteria harboring a functional arsenite S-adenosylmethionine
38 methyltransferase (ArsM) tested to date did not methylate As in pure culture. Additionally,
39 fortuitous As methylation can result from the release of non-specific methyltransferases upon
40 lysis. Thus, we combined metagenomics, metatranscriptomics, and metaproteomics to
41 identify the microorganisms actively methylating As in anoxic soil-derived microbial cultures.
42 Based on the metagenome-assembled genomes of microorganisms expressing ArsM, we
43 isolated *Paraclostridium* sp. strain EML, which was confirmed to actively methylate As
44 anaerobically. This work is an example of the application of meta-omics to the isolation of
45 elusive microorganisms.

46

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

47 **Introduction**

48 Soil microbiomes represent a rich source of novel metabolisms and taxa. However, isolating
49 microorganisms to study specific functions from these microbiomes can be challenging, and
50 even more so in cases for which the phenotype is not identifiable with high-throughput
51 methods. An example of challenging microorganisms to identify are anaerobic As-methylating
52 strains. Arsenic is a toxic metalloid that is naturally-occurring, extensively distributed in the
53 environment, and the substrate of numerous microbial transformations [1]. One such
54 transformation is As methylation, catalyzed by arsenite (As(III)) S-adenosylmethionine
55 methyltransferase (ArsM) [2], which entails the binding of one to three methyl group(s) to the
56 As atom. Arsenic methylation occurs in flooded rice paddy soils, resulting in the accumulation
57 of methylated As in rice grains [3]. The bioaccumulation of methylated As in rice grains is
58 considerably more efficient than that of inorganic As species [4, 5].

59 The gene encoding ArsM (*arsM*) has been identified in phylogenetically diverse soil
60 microorganisms [6–9]. The production of toxic trivalent monomethylated As (MMAs(III)) by
61 anaerobic prokaryotes has been proposed as a microbial warfare strategy, to inhibit microbial
62 competitors with what amounts to an arsenic-containing antibiotic [10]. If that is the case, it is
63 conceivable that As methylation may not occur in pure cultures but only in microbial
64 communities, triggered by metabolites produced by the microbiome. Alternatively, *arsM*-
65 harboring microorganisms that express As(III) efflux pump(s), the major pathway of As
66 resistance within bacteria [11], may not methylate As due to the efficient removal of As(III)
67 from the cytoplasm, which is the location of ArsM [12, 13]. Either occurrence would render the
68 isolation of pure cultures of As-methylating anaerobes very challenging using standard
69 approaches. The latter hypothesis is supported by recent work showing the lack of As
70 methylation by anaerobic pure cultures harboring functional ArsM enzymes [13].

71 An additional complexity is the evidence for the fortuitous methylation of As upon cell lysis and
72 release of methyltransferases. This occurrence was suggested by considering the
73 methanogen *Methanosarcina mazei* for which As methylation was initiated only when cell
74 viability decreased [13]. Thus, presence of As methylation for cultures incubated beyond the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

75 exponential phase may only be an experimental artefact. Finally, the detection of methylated
76 As requires relatively complex analytical tools (high pressure liquid chromatography coupled
77 to inductively-coupled plasma mass spectrometry, HPLC-ICP-MS) that do not lend
78 themselves readily to high-throughput screening of a large number of colonies. As a result of
79 these challenges, there are no anaerobic microorganisms known to actively methylate As
80 despite many efforts to identify them. In one instance, researchers had identified a Gram-
81 positive sulfate-reducing bacterium (SRB) [14] that was reported to methylate As but this
82 isolate is no longer available, precluding further investigation.
83 Thus, this study aimed to conclusively identify an active anaerobic As methylator in soil-
84 derived microbial cultures using a multi-omics approach. The experimental strategy was to
85 build Metagenome-Assembled Genomes (MAGs) from metagenomic data of the microbiome
86 and to identify the subset of MAGs harboring the gene *arsM* that also expressed *arsM*
87 (metatranscriptomics) and/or ArsM (metaproteomics). Based on the genetic information from
88 the target MAG, an informed isolation strategy was devised that allowed the recovery of a
89 pure culture later confirmed to be a novel anoxic As-methylating strain.
90

91 **Materials and methods**

92 *Rice paddy soil microbiomes*

93 The soil-derived cultures consist of two anaerobic cultures derived from a Vietnamese rice
94 paddy soil and introduced in Reid *et al.* [15]. The first soil-derived microbiome was grown in
95 ¼ strength tryptic soy broth (TSB) medium (7.5 g l^{-1} TSB), used previously to enrich arsenic-
96 methylating microbes from a lake sediment [16], and will be referred to as the TSB culture.
97 The medium from the second soil-derived microbiome, in addition to ¼ strength TSB, included
98 electron acceptors and two additional carbon sources to simultaneously allow the growth of
99 nitrate-, iron-, and sulfate-reducers, as well as microbes with fermentative and methanogenic
100 metabolisms (EA medium: 5 mM $NaNO_3$, 5 mM Na_2SO_4, 5 mM ferric citrate, 0.2 g l^{-1} yeast
101 extract (Oxoid, Hampshire, UK) and 1 g l^{-1} cellobiose, pH 7). This enrichment will be referred
102 to as the EA culture. Both media were boiled, cooled down under 100% N_2 gas and 50 ml of

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

103 medium were dispensed into 100-ml serum bottles. The bottle headspace was flushed with
104 100% N_2 gas prior to autoclaving. All culture manipulations were carried out using thoroughly
105 N_2-flushed syringes and needles. Cultures were grown at 30°C. Growth was quantified using
106 optical density at 600 nm (OD_{600}).
107
108 *Arsenic methylation assays*
109 Pre-cultures from each culture were started from -80°C glycerol stocks. The EA culture started
110 from the glycerol stock was transferred only after a dark precipitate, presumably iron sulfide
111 (suggesting potentially active sulfate reduction), was formed. The first experimental set-up
112 consisted of the inoculation of bottles containing medium amended with As(III) as $NaAsO_2$
113 (+As condition) or unamended (no-As control). Cell pellets were sampled during the stationary
114 phase for DNA sequencing and proteome characterization and at the mid-exponential growth
115 phase for RNA sequencing (see Figures S1, S2 and S3 Supplementary Information (SI) for
116 precise times). In a second experimental set-up, cultures were grown in medium without As(III)
117 and, at the mid-exponential growth phase, As(III) was added. Cell pellets were sampled before
118 (no-As control) and 30 min after As spiking (+As condition) and were used for a second
119 transcriptomic analysis only. All experiments were performed in biological triplicates. Sampling
120 for soluble As species, determination of As speciation and total As concentration were
121 performed as described in [13] using an Agilent 8900 ICP-QQQ instrument coupled to an
122 HPLC 1260 Infinity II (Agilent Technologies, CA, USA). Instrument settings in Table S1. The
123 DNA and RNA sequencing, metaproteome characterization, metagenomic,
124 metatranscriptomic, and metaproteomic analyses are described in SI.
125
126 *Isolation of* As-methylating microorganism
127 The isolation of the anaerobic *arsM*-expressing microorganism was conducted by using serial
128 dilution agar plate method in an anaerobic chamber (Coy Laboratories, Grass Lake, MI, USA)
129 containing 90% N_2:10% H_2 with less than 5 ppm of O_2. Briefly, 1 ml of cell suspension of the
130 EA culture was serially diluted in a 10-fold series (10^{-1} to 10^{-5} dilutions) using sterile Reinforced

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

131 Clostridial Broth (RCB) (Oxoid Ltd., Basingstoke, UK). Consecutively, 100 μl of EA cell
132 suspension from each dilution was spread uniformly over the surface of Reinforced Clostridial
133 agar (RCA) (Oxoid Ltd., Basingstoke, UK). The inoculated RCA plates were incubated at 30°C
134 for 24 hours. The single colonies were transferred with sterile toothpicks to Tryptose-sulfite-
135 cycloserine (TSC) agar (Merck, Darmstadt), containing 5 g l^{-1} sucrose, 0.04 g l^{-1} bromocresol
136 purple (Sigma-Aldrich), and 0.4 g l^{-1} de-hydrated D-cycloserine (Sigma-Aldrich), and grown at
137 30°C for 24 hours.

138 Colony PCR was performed on black colonies picked from purple TSC agar using the
139 designed specific primers arsM-9F: 5'-TCTAATCTAAGTTGTTGTGGGGAAG-3' and arsM-
140 9R: 5'-TGATATAGATAACCTACCTCCGCC-3', generating a 500 bp amplicon of the *arsM*
141 gene from MAG 9 from the EA culture (Table 1-A) (protein id k119_30669_28, Table S2).
142 Before direct colony PCR, the black colonies were first picked, diluted into 15 μl lysis buffer
143 (0.1% triton X-100 + TE buffer) and boiled at 95°C for 10 min, to release DNA, and then
144 centrifuged (13,000 g, 10 min) to spin down cell debris. The supernatant of the lysate was
145 used as the DNA template for colony PCR.

146 Each colony PCR consisted of a 25-μl reaction containing: 12.5 μl 2x GoTaq Green Master
147 mix (Promega, UK), 0.5 μl of the reverse and forward primers (10 μM each), 1.5 μl of lysate
148 supernatant as the DNA template, 0.25 μl of 20 mg ml^{-1} bovine serum albumin (BSA) (Sigma-
149 Aldrich), and 10.75 μl sterile DNase RNase Free water. The thermocycling program consisted
150 of an initial denaturation at 95°C for 5 min, 30 cycles of denaturation at 95°C for 40s, annealing
151 at 53°C for 40 sec, and extension at 72°C for 40s, and a final extension at 72°C for 10 min.
152 After purification of the PCR product with Wizard® SV Gel and PCR Clean-Up System
153 (Promega, UK), the *arsM* amplicon was sequenced at Microsynth (Balgach, Switzerland). The
154 amplification of full-length 16S rRNA gene was performed using the primers 27F [17] and
155 1542R [18]. The As methylation assay of the isolate is described in SI.

156

157 **Results**

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

158 *Arsenic methylation by soil-derived microbiomes*

159 As described in materials and methods, two experimental set-ups were used to probe the As-
160 methylating cultures. The first set-up compared the soil-derived microbial communities grown
161 in the presence or absence of As(III). Samples for the metagenome, metaproteome and one
162 of the metatranscriptomes (labeled metatranscriptome G for 'growth in the presence of As')
163 were obtained from this set-up (Figures S1, S2 and S3). In the second set-up, the intention
164 was to capture the short-term response of the community to As(III). Thus, the cultures were
165 grown to mid-exponential phase in the absence of As, sampled, amended with As(III), and
166 sampled a second time after 30 minutes. Samples obtained from this set-up were used for the
167 second metatranscriptome (labeled metatranscriptome R for 'response to arsenic addition')
168 (Figures S1-A and S2-A).

169 Both cultures exhibited As methylation, reaching an efficiency of transformation of the initial
170 As(III) of 27.7% and 19.5% for the EA and TSB cultures, respectively (Figures S1 and S2 and
171 SI). The TSB and EA microbiomes compare favorably to previous studies of anaerobic
172 enrichments [16] and single-strain cultures [13, 14, 19], and represent the most efficient anoxic
173 As-methylating microbial cultures reported.

174

175 *Microbiome composition*

176 The taxonomic classification of small subunit (SSU) 16S rRNA sequences show that, although
177 eukaryotic DNA was also identified, the main fraction of the communities is bacterial (>89.0
178 ±0.8% for EA cultures and >98.5 ±0.3% for TSB cultures, relative abundance) and is
179 distributed amongst eight operational taxonomic units (OTUs) at the order level:
180 Acidaminococcales, Bacillales, Bacteroidales, Clostridiales, Desulfovibrionales,
181 Enterobacterales, Lactobacillales and Selenomonadales (Figure 1 and Tables S3-S6).
182 Statistically significant changes in the OTUs relative abundances, +As condition versus no-As
183 control, are described in SI.

184 The read-assembled contiguous sequences (contigs) from the four metagenomes, EA (+As,
185 no-As control) and TSB (+As, no-As control), were clustered separately into bins. High-quality

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

186 (≥95% completeness and ≤5% contamination) and good-quality (≥74% completeness and
187 ≤10% contamination) bins were designated as MAGs. For the +As condition, the parsing
188 process led to a total of 40 MAGs: 19 in EA culture (Table 1-A), and 21 in TSB culture (Table
189 1-B). Additionally, matching bins for the +As condition, based on the MAG groups, were found
190 in the bins from the no-As control cultures (Tables S7 and S8). Only three of the 40 MAGs in
191 the +As condition were left unpaired (EA MAGs 7 and 12 and TSB MAG 10).
192 For each MAG group, a lineage was assigned by CheckM, based on lineage-specific marker
193 genes [20]. The MAGs identified belonged mainly to the phylum Firmicutes (orders
194 Clostridiales, Selenomonadales and Lactobacillales, and the genus *Clostridium*).
195 Proteobacteria MAGs included the *Enterobacteriaceae* family and the Deltaproteobacteria
196 class. Finally, in each metagenome, one MAG from the order Bacteroidales was present.
197 Fifteen MAGs presented strain heterogeneity, an index of the phylogenetic relatedness of
198 binned contigs based on the amino acid identity of the encoded proteins. For ten MAGs, the
199 value is >50%, suggesting some phylogenetic relation of the contaminating strains. Five MAGs
200 had heterogeneity values <33.33%, suggesting contamination with microorganisms that are
201 not closely related. In the remaining 25 MAGs, the strain heterogeneity is 0%, i.e., no strain
202 heterogeneity or no contamination (Tables S7 and S8).
203 Ultimately, to evaluate the relatedness of the EA and TSB microbial communities, matching
204 bins between the EA and TSB +As condition MAGs were found by pairwise comparison of the
205 predicted genomes. Of the 40 MAGs identified across EA and TSB cultures, 22
206 (corresponding to 11 pairs) exhibited similarities >98.47% (Table S9), including a
207 Deltaproteobacteria, a *Clostridium*, three Clostridiales, a Bacteroidales, two Lactobacillales, a
208 Selenomonadales, and two Firmicutes. Thus, approximately half of the MAGs of the EA
209 culture are also present in the TSB culture and vice-versa.
210 Changes in the relative abundance of MAGs (no-As control vs. +As condition), as well as the
211 presence, transcription and translation of genes encoding key enzymes from major metabolic
212 pathways from each MAG in the +As condition cultures are included in SI.
213

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

214 *Arsenic resistance genes*

215 The metagenomic libraries from the +As condition of the EA and TSB cultures were mined for
216 arsenic resistance (*ars*) genes and their encoded proteins (pipeline described in SI). A gene
217 was considered to be present in the culture if DNA reads represented >5 TPM-DNA
218 ('transcripts per million' (TPM), referred to as TPM-DNA when used for gene abundance, see
219 SI), considered to be transcribed if >5 TPM-RNA (referred as TPM-RNA when employed for
220 transcript abundance) were detected, and as translated if protein abundance could be
221 calculated from the detected peptides in at least two of the three biological replicates.
222 Additionally, increased expression in the RNA and protein in the +As condition relative to the
223 no-As condition, was considered when the absolute $\log_2$ fold change was ≥1 (i.e., 0.5 ≥ fold
224 change ≥2) and the adjusted q-value ≤0.05 (refer to SI for pipeline) (Figure 2). A total of 309
225 and 282 genes were annotated as *ars* genes in the EA and TSB +As metagenomic libraries,
226 respectively (Tables S10 and S11). Of those, 255 and 226 were considered correctly
227 annotated as *ars* genes based on BLAST® and HMMER (refer to SI for pipeline), and 225 and
228 147 had above-threshold DNA abundances, respectively (Figure 2). Individual values of
229 transcript and protein abundance in the +As condition and the no-As control and their
230 abundance values in the +As condition relative to the no-As control for each MAG group,
231 biological replicate and *ars* proteins from the EA and TSB cultures are available in Tables S2
232 and S12, respectively.

233 The *ars* genes encode proteins involved in the detoxification of As oxyanions: *arsB* and *acr3*
234 encoding As(III)-efflux systems; *arsA* encoding the ATPase energizing the efflux of As(III) and
235 mostly found alongside the gene for a As(III) chaperone and weak *ars* operon repressor (*arsD*)
236 [21]; *arsC1* and *arsC2* encoding As(V) reductases coupling As reduction to the oxidation of
237 glutaredoxin or thioredoxin, respectively; and *arsR* genes encoding As(III)-regulated
238 repressors (ArsR1, ArsR2, and ArsR3) which are distinguishable based on the location of the
239 As(III)-binding cysteine residues [22–24].

240 The most common *ars* genes in EA and TSB culture metagenomes were *arsR*, *arsC,* and *arsP*
241 (Figure 2). The first two genes are part of the canonical *ars* operon *arsRBC* [25], whilst *arsP*,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

242 encoding a more recently discovered membrane transporter, has been found to be widely
243 distributed in bacterial genomes [11]. Most of the surveyed *arsP* genes, 57% in EA and 50%
244 in TSB, are encoded in putative *ars* operons, represented by *ars* genes contiguously encoded
245 in the same contig (Tables S2 and S12), supporting their As-related function and correct
246 annotation. Unsurprisingly, the genes responsible for As(III) efflux (*arsB*, *acr3,* and *arsA*),
247 typically found in organisms living in reducing environments, were amongst the most common
248 *ars* genes along with *arsC* [8, 26]. Finally, *arsM* and the two genes, *arsI* and *arsH*, encoding
249 oxygen-dependent MMAs(III)-resistance mechanisms, were the least recurrent genes in the
250 metagenomes. In soils, *arsM* has previously been observed to be less common in comparison
251 to other surveyed *ars* genes [6, 8, 27]. The results of gene and protein relative expression vs.
252 the no-As control of the *ars* genes involved in the metabolism of inorganic As in the MAGs can
253 be found in SI.
254 It is striking to note that while there are a large number of *ars* genes in the metagenome, a
255 small proportion is expressed (whether as mRNA transcripts or as proteins) in the presence
256 of As when compared to the no As control (Figure 2). This contrast is particularly evident for
257 the gene responsible for As methylation, *arsM*.
258
259 *Arsenic-methylating MAGs*
260 The aim of the present study was to identify the microorganisms catalyzing As methylation in
261 two anaerobic soil-derived microbiomes. The *arsM* gene can be expressed at similar, or
262 slightly different levels in the absence or presence of As(III) in some organisms [28, 29], but
263 expressed at significantly higher levels in the presence of As(III) in others [30–33]. Thus, we
264 sought to identify *arsM* genes transcribed and ArsM proteins showing increased expression
265 in the +As condition relative to the no-As control (Figure 3) but also irrespective of whether
266 their expression was reported as increased relative to the no-As control (Figure S4).
267 Sixteen phylogenetically distinct *arsM* genes were identified in the EA +As metagenome, but
268 increased transcriptome reads or peptides (relative to the no-As control) were only detected
269 for three genes (Figure 3) (see SI for details of calculation). The first is an *arsM* in Clostridiales

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

270 EA MAG 9 classified by GhostKOALA as belonging to *Paeniclostridium sordellii* (EA MAG 9,
271 *arsM*-1, psor type strain, in Table S2). The second was found in *Clostridium* EA MAG 10, also
272 detected in the metaproteome, and the GhostKOALA taxonomic classification of the
273 corresponding gene (EA MAG 10, *arsM-1* in Table S2) revealed that it was attributed to the
274 unclassified species *Ruminococcaceae bacterium* CPB6 (Figure 3, Table S2), also referred
275 as *Clostridium bacterium* CPB6 [34]. Finally, the third *arsM* was obtained from transcriptomic
276 data but not clustered in any EA MAG (EA unbinned, *arsM-5* in Table S2) and likewise
277 classified as pertaining to *Paeniclostridium sordellii*. Thus, in EA, all three *arsM* genes showing
278 evidence of involvement in active As methylation pertain to fermenting microorganisms from
279 the order Clostridiales.
280 Similarly, in the TSB +As metagenome, 9 distinct *arsM* genes were identified but none were
281 detected in the metatranscriptome and only one exhibited increased expression in the
282 metaproteome (Figure 3). It corresponds to an *arsM* gene from MAG 11 (TSB MAG11, *arsM-*
283 *2* in Table S12). The expressed ArsM protein was assigned by GhostKOALA to a Clostridiales
284 strain: *Clostridium botulinum* (*cby* type strain) (TSB MAG11, *arsM-2*) (Figure 3, Table S12).
285 Finally, there was one *arsM* expressed in the TSB +As metaproteome but with no increased
286 expression relative to the no-As control, it was classified as *Ruminococcaceae bacterium*
287 CPB6 (TSB MAG 11, *arsM-1*) (Figure S4), the same organism identified in the EA culture (EA
288 MAG 10, *arsM-1*). Thus, in TSB as in EA soil-derived microbiomes, As methylation appears
289 to be catalyzed by various fermenting bacteria pertaining to the order Clostridiales such as
290 members of the genera *Paeniclostridium* and *Clostridium*, and the family *Ruminococcaceae*.
291 In addition to evidence for active arsenic methylation, there was evidence for active
292 detoxification of methylated arsenic. Indeed, the metagenome included genes encoding
293 proteins involved in the metabolism of methylated arsenic such as *arsH*, *arsI, arsP,* and *arsR4*
294 (Figures 2 and 3). These genes encode proteins involved in the detoxification of methylated
295 arsenic such as MMAs(III) and roxarsone: the oxidase ArsH, responsible for the oxidation of
296 trivalent methylated As to the less toxic pentavalent form [35]; the demethylase ArsI that
297 removes methyl groups from the As atom [36]; and the transmembrane transporter ArsP,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

298 thought to efflux methylated As [37]. The *arsR4* gene encodes an atypical MMAs(III)-
299 responsive ArsR repressor, containing only two conserved cysteine residues [38]. The
300 *Enterobacteriaceae* TSB MAG 16 exhibited activity of the oxygen-dependent ArsH protein [35]
301 (Figure 3), a fact that is difficult to reconcile with the anoxic conditions. It is conceivable that
302 this protein is capable of additional functions under anoxic conditions. An *arsR4*, shown to
303 induce expression of *arsP* in the presence of MMAs(III) [38], had increased transcription along
304 with an *arsP* encoded in the same contig in the Selenomonadales TSB MAG 21 (Figure 3,
305 Table S12). Both genes transcripts were <5 TPM-RNA (Table S12) and thus, were not
306 considered as transcribed in Figure S4. Finally, an ArsI protein, taxonomically related to class
307 *Clostridia* ([*Eubacterium*] *rectale*), was expressed but it was encoded in an unbinned gene
308 from the EA culture (Figure 3, Table S12). The identification of MAGs exhibiting a
309 detoxification response to methylated As supports the hypothesis, raised above, of the role of
310 monomethylated As as an arsenic antibiotic.
311
312 *Isolation of an arsenic-methylating anaerobic microorganism*
313 Based on the analysis of the active metabolic activity from the EA MAG 9, expressing an As(III)
314 methyltransferase (Figure S5), an appropriate selective medium was identified for its isolation.
315 We utilized the fact that this MAG harbors and expresses the anaerobic assimilatory sulfite
316 reductase encoded in the *asrABC* operon which is responsible for the NADH-dependent
317 reduction of sulfite to sulfide [39–41] in sulfite-reducing Clostridia (SRC). It was the only
318 member of Clostridiales expressing this capability in the EA microbiome (Figure S5). Thus,
319 the isolation relied on growing the EA culture on agar medium selective for the SRC
320 phenotype. In TSC agar, designed for the enumeration of *Clostridium perfringens* in food [42],
321 the colonies from SRC are black, as the ammonium ferric citrate forms iron sulfide during
322 sulfite reduction. Additionally, D-cycloserine acts a selective agent for the isolation of Clostridia
323 strains [43] while inhibiting facultative anaerobes [42]. Finally, the bromocresol purple
324 contained in the agar allows further differentiation between negative and positive sucrose
325 fermenters, the latter changing the purple color of the agar to yellow. As none of the genes

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

326 involved in sucrose transport or hydrolysis were binned in EA Clostridiales MAG 9 (Figure S5),
327 only non-sucrose fermenting black colonies were considered. Those colonies were selected
328 and using a colony PCR screen specifically targeting the *arsM* gene of EA MAG 9, we isolated
329 a Clostridiales strain encoding the gene of the expressed ArsM in the EA MAG 9 (protein id
330 k119_30669_28, Table S2) (Figure S6).
331 The isolate consists of non-sucrose-fermenting, rod-shaped and spore-forming bacteria
332 forming convex and circular black colonies on TSC agar (Figures S7 and S8). The BLAST®
333 (NCBI) search of the 16S rRNA sequence gives >99% identity to *Paraclostridium* strains
334 (Table S13). On the basis of the 16S rRNA sequence, we assign the following name to the
335 bacterium: "*Paraclostridium* species str. EML". Strain EML was tested for As methylation
336 under anaerobic conditions with 25 μM As(III). The growth of strain EML was hindered by
337 As(III) (Figure 4-A) and starting from ~4 hours, the isolate transformed As(III) to
338 monomethylated soluble As representing 48.3±1.5% of the soluble arsenic in the culture after
339 83 h (panels B and C from Figure 4). A fraction (14.7±0.6 μM) of the arsenic was found
340 associated with biomass almost exclusively as inorganic arsenic (Figure 4-D).
341

342 **Discussion**

343 Our results demonstrate the successful translation of multi-omic information to a specific
344 strategy for targeted microbial isolation. While the metagenomes from the anaerobic soil-
345 derived microbiomes identified the potential for As methylation in microorganisms from diverse
346 taxa, the post-genomic approaches of community gene and protein expression clearly pointed
347 to the active role of Clostridiales microorganisms in both cultures. This information paved the
348 way for the identification of As-methylating microorganisms and the successful isolation of an
349 anaerobic As methylator.
350 The EA and TSB soil-derived cultures offered the opportunity to study active As methylation
351 from paddy-soil microbiomes in an environment that is less complex than soil but that remains
352 environmentally relevant. In contrast to soil slurries, the lack of soil minerals facilitated the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

353 detection of soluble methylarsenicals and the facile extraction of DNA, RNA and proteins. The
354 multi-omic approach made it possible to identify the putative microorganisms driving As
355 methylation and their metabolism. Targeting a specific *arsM* gene rather than the synthesis of
356 methylarsenicals greatly accelerated colony screening, as colony PCR could be employed
357 instead of analytical detection of methylated As by HPLC-ICP-MS.
358 Had only the metagenomic approach been implemented, the data would have pointed to SRB
359 MAGs as putative As methylators, as they harbored the most abundant *arsM* genes (Figure
360 6). Indeed, SRB have been proposed as drivers of As methylation in rice paddy soils based
361 on the correlation in the abundance of the *arsM* and *dsrB* genes, and a decrease in As
362 methylation by the addition chemical inhibitors of DSR [44, 45]. Yet, in the present findings,
363 the SRB Deltaproteobacteria MAGs, although actively reducing sulfate (Figures S5 and S9),
364 did not exhibit As-methylating activity as their *arsM* genes were neither transcribed nor
365 translated (Figures 5). A recent study has shown that the abundance of *ars* genes in high- vs.
366 low-As paddy soils was comparable whilst their transcriptomic activity was significantly
367 impacted [46], bolstering the emerging view that to identify active As methylators in natural
368 environments, the exclusive use of genomic data is insufficient.
369 Previous work had identified another As-methylating Clostridiales strain, *Clostridium* sp. BXM
370 [14], that performed fermentation and DSR. Amendment with organic matter [7, 47, 48], and
371 an increase in dissolved organic carbon [49] have been shown to increase the As methylation
372 efficiency, suggesting a correlation between the enrichment of fermenting communities, via
373 increased availability of organic substrates, and As methylation. These observations along
374 with the isolation of the present strain, point to a key role for fermenting Clostridiales
375 microorganisms in As methylation.
376 It was previously proposed that the As-methylating activity of anaerobic microorganisms may
377 be limited by efficient efflux of intracellular As(III) [13], or that it may function as a defensive
378 response against nutrient competition [10]. However, up until now, the lack of anaerobic
379 microbial isolates able to methylate As *in vitro* [13] precluded the investigation of this question.
380 This study, capitalizing on the promise of omics-aided cultivation, paves the way for the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

381 elucidation of the controls on anaerobic As methylation. Further work is needed to elucidate
382 why ArsM expression was restricted to members of Clostridiales fermenters and did not occur
383 in other organisms harboring *arsM* genes.
384

385 **Acknowledgements**
386 The work was funded by the Swiss National Science Foundation (SNSF) grant
387 310030_176146-1 and the SNSF NCCR Microbiomes (grant # 51NF40_180575). The work at
388 ORNL was conducted under the Plant-Microbe Interface Science Focus Area, as supported
389 by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental
390 Research, Genomic Science Program.
391

392 **Data availability**
393 Metagenomic and metatranscriptomic raw sequencing reads are available at the National
394 Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject
395 PRJNA714492 (post publication). Data from the meta-omic analyses and source data from
396 figures are available in Zenodo data repository (10.5281/zenodo.4605527, post publication)
397 or by request from the authors.
398

399 **Conflict of interest**
400 The authors declare no conflict of interest.
401

402 **Figures Legends**
403 **Figure 1. Operational taxonomic units (OTUs) at order and genus level (with > 1%
404 relative abundance at genus level) identified from SSU 16S rRNA sequences from soil-
405 derived cultures.** Abbreviations: EA no As: EA culture no-As control EA +As: EA culture +As
406 condition, TSB no As: TSB culture no-As control TSB +As: TSB culture +As condition. OTUs
407 at the order level are indicated in bold in the legend. Plotted values are the average relative
408 abundance and together with SD values and student t-test results are available in SI Tables

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

409 S14 and S15.

410

411 **Figure 2. Number of *ars* genes, encoded in the +As condition cultures and with**
412 **increased expression in metatranscriptomes/metaproteome relative to no-As controls.**
413 Number of *ars* genes encoded in metagenomes and with increased expression in
414 metatranscriptomes, R or G, or metaproteomes and the non-redundant overlap between
415 genes with increased expression in metatranscriptomes and/or metaproteomes from +As
416 condition EA (left panels) and TSB (right panels) cultures. Bar length and numbers on the right
417 side of the bars correspond to the number of genes per *ars* gene category. Bold numbers on
418 the lower left corner of each panel correspond to the sum of all *ars* genes per category.

419

420 **Figure 3. Distribution of *ars* genes involved in methylated arsenic metabolism encoded**
421 **in MAGs from the +As condition and differentially expressed in**
422 **metatranscriptomes/metaproteome relative to the no-As EA control.** Each numbered box
423 represents an *ars* gene. The number in each box corresponds to the “Numbering” column in
424 Tables S2 and S12 where individual gene abundance and fold change values can be found.
425 % Com: % Community as defined in caption from Table 1.

426

427 **Figure 4. Isolate “*Paraclostridium* sp. EML”:** (A) growth as OD_{600} with 25 μM As(III) and
428 without, (B) percentage of soluble arsenic species in filtered medium containing 25 μM As(III),
429 (C) concentration of arsenic species soluble in filtered medium containing 25 μM As(III) (solid
430 lines) and biomass-bound (dashed lines) and (D) percentage of biomass-bound arsenic
431 species. OD_{600} corresponds to optical density at 600 nm. Data points and bars represent the
432 mean value and error bars, plus and minus one standard deviation. Individual values for each
433 measurement and biological replicate are available in Tables S23 and S24.

434 **Figure 5. Gene abundance of *arsM* genes in MAGs from the +As condition cultures.**
435 Colored bars correspond to *arsM* genes with increased expression in the metaproteome (blue-
436 green) or in the metatranscriptome G (purple) from +As condition relative to the no-As control

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

437 in EA (left panel) and TSB (right panel) cultures. The taxonomic classification shown on the
438 right side of the error bars for selected *arsM* genes corresponds to the individual gene
439 classification assigned by GhostKOALA - "Genus" column in Tables S2 and S12. Columns
440 with matching symbols on the right side of the error bars, correspond to matching *arsM* genes
441 between the cultures. Individual gene abundance can be found in Tables S2 and S12.
442 Numbers inside bar and bar length represent mean and error bars one standard deviation.
443

444 **Table Legend**

445 **Table 1. Metagenome assembled genomes (MAGs) from EA (upper Table A) and TSB**
446 **(lower Table B) cultures in the +As condition.** Marker lineage: taxonomic rank set by
447 CheckM. Completeness and contamination %: estimated completeness and contamination of
448 genome as determined by CheckM from the presence/absence of single-copy marker genes
449 and the expected colocalization of these genes. Strain heterogeneity: index between 0 and
450 100 where a value of 0 means no strain heterogeneity, high values suggest the majority of
451 reported contamination is from closely related organisms (i.e., potentially the same species)
452 and low values suggest the majority of contamination is from phylogenetically diverse sources.
453 % of binned proteins assigned to MAG: number of protein-coding genes assigned to the MAG
454 divided by the total number of protein-coding genes binned. % Community: sum of the number
455 of reads mapped to the contigs in each MAG divided by the total number of reads mapped to
456 all contigs including the unbinned contigs, and normalized to MAG size, assuming an average
457 genome size for all unbinned populations. High-quality MAGs are denoted by bolded numbers,
458 good-quality MAGs by italicized numbers.
459

460 **SI Figures Legends**

461 **Figure S1. Arsenic methylation in the +As condition EA culture. A) Growth curves from**
462 **experiments as OD_{600}.** Red arrow indicates the samples used for metagenome and
463 metaproteome analyses. Blue arrow indicates the samples used for metatranscriptomes G
464 and R. B) Total soluble arsenic in medium. C) Percentage of arsenic species soluble in

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

465 medium for metagenome and metaproteome experiments. D) Percentage of arsenic species
466 soluble in medium for metatranscriptome G and metatranscriptome R experiments. OD_{600}
467 corresponds to optical density at 600 nm. Points and bar heights represent mean and error
468 bars plus, minus one standard deviation. Individual values for each measurement and
469 biological replicate are available in Tables S25 and S26.
470
471 **Figure S2. Arsenic methylation in the +As condition TSB culture. A) Growth curves from**
472 **experiments as OD_{600}.** Red arrow indicates the samples used for metagenome and
473 metaproteome analyses. Blue arrow indicates the samples used for metatranscriptomes G
474 and R. B) Total soluble arsenic in medium. C) Percentage of arsenic species soluble in
475 medium for metagenome and metaproteome experiments. D) Percentage of arsenic species
476 soluble in medium for metatranscriptome G and metatranscriptome R experiments. OD_{600}
477 corresponds to optical density at 600 nm. Points and bar heights represent mean and error
478 bars plus, minus one standard deviation. Individual values for each measurement and
479 biological replicate are available in Tables S25 and S26.
480
481 **Figure S3. Growth in no-As controls. Growth curves as OD_{600} from no-As controls in EA**
482 **(left panel) and TSB (right panel) cultures.** Red arrow indicates the samples used for
483 metagenome and metaproteome analyses. Blue arrow indicates the samples used for
484 metatranscriptome G. OD_{600} corresponds to optical density at 600 nm. Points represent mean
485 and error bars plus, minus one standard deviation. Individual values for biological replicate
486 are available in Table S25.
487
488 **Figure S4. Distribution of *ars* genes involved in methylated arsenic metabolism**
489 **encoded in MAGs from the +As condition and expressed in**
490 **metatranscriptomes/metaproteome.** Each numbered box represents an *ars* gene. The
491 number in each box corresponds to the “Numbering” column in Tables S2 and S12 where

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

492 individual gene abundance and fold change values can be found. % Com: % Community as
493 defined in caption from Table 1.
494
495 **Figure S5. Key enzymes from major metabolic pathways in MAGs from the +As**
496 **condition EA culture.** Vertical bold lines correspond to the grouping of the MAGs with same
497 lineage (Table 1-A). Pathway abbreviations: dissimilatory nitrate reduction to ammonia
498 (DNRA), organic carbon metabolism (Org. C), propionic acid fermentation (propionic acid f.),
499 and acetone-butanol-ethanol (ABE) fermentation. Refer to SI Table S16 for individual gene,
500 transcript and protein abundance values and further enzymes.
501
502 **Figure S6. Colony PCR agarose gel from *Paraclostridium* sp. EML.** Left lane: PCR product
503 from the amplification of *arsM* (protein id k119_30669_28, Table S2) in the PCR reaction using
504 a *Paraclostridium* sp. EML colony, right lane: ladder corresponding to (from top to bottom)
505 1000, 750, 500, 300, 150 and 50 bp.
506
507 **Figure S7.** Light microscopy of *Paraclostridium* sp. EML cells, 48-h culture.
508
509 **Figure S8.** Growth of *Paraclostridium* sp. EML isolate in TSC agar.
510
511 **Figure S9. Key enzymes from major metabolic pathways in MAGs from the +As**
512 **condition TSB culture.** Vertical bold lines correspond to the grouping of the MAGs with same
513 lineage (Table 1-B). Pathway abbreviations: dissimilatory nitrate reduction to ammonia
514 (DNRA), organic carbon metabolism (Org. C), propionic acid fermentation (propionic acid f.),
515 and acetone-butanol-ethanol (ABE) fermentation. Refer to SI Table S17 for individual gene,
516 transcript and protein abundance values and further enzymes.
517
518 **Figure S10. % Community of MAGs. % Community of MAGs in +As condition and no-**
519 **As control from EA (top panel) and TSB (low panel) cultures.** Statistical differences

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

520 between +As condition vs. no-As control were identified by unpaired Student t-test with p-
521 value <0.05. Symbols: NA: no matching MAG in no-As control was found, one or more
522 asterisks (*): significant difference and ns: no significant difference (p-value >0.05) (see Table
523 S14 for P value symbol summary) . Points represent individual values from three biological
524 replicates. Bar heights represent mean and horizontal lines plus, minus one standard
525 deviation.

526

527 **Figure S11. Volcano plots of metatranscriptomes.** Dots represent individual genes
528 transcribed in metatranscriptomes R (upper panels) or G (lower panels) from +As condition
529 EA (left panels) and TSB (right panels) cultures. Genes considered statistically differentially
530 transcribed in the +As condition vs. no-As controls, based on the adjusted p value (q value),
531 are represented as magenta (decreased transcription), green (increased transcription) and
532 yellow (-1< $\log_2$ fold changes < 1) dots. Grey dots are genes with non-statistically significant
533 changes in transcription. In bold numbers, the percentage of genes in magenta or green.
534 Individual fold-change values are available in Tables S19 and S20.

535

536 **Figure S12. Volcano plots of metaproteomes.** Dots represent individual genes expressed
537 in metaproteomes from +As condition EA (left panel) and TSB (right panel) cultures. Genes
538 considered statistically differentially expressed in the +As condition vs. respective no-As
539 controls, based on the adjusted p value (q value), are represented as magenta (decreased
540 expression), green (increased expression) and yellow (-1< $\log_2$ fold changes < 1) dots. Grey
541 dots are genes with non-statistically significant changes in expression. In bold numbers, the
542 percentage of genes in magenta or green. Individual fold-change values are available in
543 Tables S19 and S20.

544

545 **Figure S13. Distribution of *ars* genes encoded in MAGs from the +As condition EA**
546 **culture and differentially expressed in metatranscriptomes/metaproteome relative to**
547 **the no-As EA control.** Each numbered box represents an *ars* gene. The number in each box

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

548 corresponds to the “Numbering” column in Table S2 where individual gene abundance and
549 fold change values can be found.
550
551 **Figure S14. Distribution of *ars* genes encoded in MAGs from the +As condition TSB**
552 **culture and differentially expressed in metatranscriptomes/metaproteome relative to**
553 **the no-As TSB control.** Each numbered box represents an *ars* gene. The number in each
554 box corresponds to the “Numbering” column in Table S12 where individual gene abundance
555 and fold change values can be found.
556
557 **Figure S15. Edwards-Venn diagrams of *ars* genes with increased expression in +As**
558 **condition EA and TSB relative to no-As control cultures.** Number of *ars* genes encoded
559 in metagenomes, with increased expression in metatranscriptomes R and G or/and
560 metaproteomes from +As condition EA culture (left panel) and +As condition TSB (right panel)
561 cultures.
562

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

563 **References**

564 1. Zhu Y-G, Yoshinaga M, Zhao F-J, Rosen BP. Earth Abides Arsenic
565 Biotransformations. *Annu Rev Earth Planet Sci* 2014; **42**: 443–467.
566 2. Lomax C, Liu WJ, Wu L, Xue K, Xiong J, Zhou J, et al. Methylated arsenic species in
567 plants originate from soil microorganisms. *New Phytol* 2012; **193**: 665–672.
568 3. Zhao FJ, Zhu YG, Meharg AA. Methylated arsenic species in rice: Geographical
569 variation, origin, and uptake mechanisms. *Environ Sci Technol* 2013; **47**: 3957–3966.
570 4. Zheng MZ, Li G, Sun GX, Shim H, Cai C. Differential toxicity and accumulation of
571 inorganic and methylated arsenic in rice. *Plant Soil* 2013; **365**: 227–238.
572 5. Abedin MJ, Feldmann J, Meharg AA. Uptake kinetics of arsenic species in rice plants.
573 *Plant Physiol* 2002; **128**: 1120–1128.
574 6. Dunivin TK, Yeh SY, Shade A. A global survey of arsenic-related genes in soil
575 microbiomes. *BMC Biol* 2019; **17**.
576 7. Jia Y, Huang H, Zhong M, Wang F-H, Zhang L-M, Zhu Y-G. Microbial Arsenic
577 Methylation in Soil and Rice Rhizosphere. *Environ Sci Technol* 2013; **47**: 3141–3148.
578 8. Xiao KQ, Li LG, Ma LP, Zhang SY, Bao P, Zhang T, et al. Metagenomic analysis
579 revealed highly diverse microbial arsenic metabolism genes in paddy soils with low-
580 arsenic contents. *Environ Pollut* 2016; **211**: 1–8.
581 9. Zhang SY, Su JQ, Sun GX, Yang Y, Zhao Y, Ding J, et al. Land scale biogeography
582 of arsenic biotransformation genes in estuarine wetland. *Environ Microbiol* 2017; **19**:
583 2468–2482.
584 10. Chen J, Rosen BP. The Arsenic Methylation Cycle: How Microbial Communities
585 Adapted Methylarsenicals for Use as Weapons in the Continuing War for Dominance.
586 *Front Environ Sci* . 2020. Frontiers Media S.A. , **8**: 43
587 11. Yang Y, Wu S, Lilley RM, Zhang R. The diversity of membrane transporters encoded
588 in bacterial arsenic-resistance operons. *PeerJ* 2015; **3**: e943.
589 12. Yang P, Ke C, Zhao C, Kuang Q, Liu B, Xue X, et al. ArsM-mediated arsenite
590 volatilization is limited by efflux catalyzed by As efflux transporters. *Chemosphere*

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

591 2020; **239**: 124822.

592 13. Viacava K, Meibom KL, Ortega D, Dyer S, Gelb A, Falquet L, et al. Variability in
593 Arsenic Methylation Efficiency across Aerobic and Anaerobic Microorganisms.
594 *Environ Sci Technol* 2020; **54**: 14343–14351.

595 14. Wang PP, Bao P, Sun GX. Identification And Catalytic Residues Of The Arsenite
596 Methyltransferase From a Sulfate-Reducing Bacterium, Clostridium sp. BXM. *FEMS*
597 *Microbiol Lett* 2015; **362**: 1–8.

598 15. Reid MC, Maillard J, Bagnoud A, Falquet L, Le Vo P, Bernier-Latmani R. Arsenic
599 Methylation Dynamics in a Rice Paddy Soil Anaerobic Enrichment Culture. *Environ*
600 *Sci Technol* 2017; **51**: 10546–10554.

601 16. Bright DA, Brock S, Reimer KJ, Cullen WR, Hewitt GM, Jafaar J. Methylation of
602 arsenic by anaerobic microbial consortia isolated from lake sediment. *Appl*
603 *Organomet Chem* 1994; **8**: 415–422.

604 17. DeLong EF. Archaea in coastal marine environments. *Proc Natl Acad Sci U S A* 1992;
605 **89**: 5685–5689.

606 18. Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. Isolation and direct complete
607 nucleotide determination of entire genes. Characterization of a gene coding for 16S
608 ribosomal RNA. *Nucleic Acids Res* 1989; **17**: 7843–7853.

609 19. Wang PP, Sun GX, Zhu YG. Identification And Characterization Of Arsenite
610 Methyltransferase From An Archaeon, Methanosarcina Acetivorans C2A. *Environ Sci*
611 *Technol* 2014; **48**: 12706–12713.

612 20. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: Assessing
613 the quality of microbial genomes recovered from isolates, single cells, and
614 metagenomes. *Genome Res* 2015; **25**: 1043–1055.

615 21. Lin Y-F, Walmsley AR, Rosen BP. An arsenic metallochaperone for an arsenic
616 detoxification pump. *Proc Natl Acad Sci U S A* 2006; **103**: 15617–22.

617 22. Shi W, Wu J, Rosen BP. Identification of a putative metal binding site in a new family
618 of metalloregulatory proteins. *J Biol Chem* 1994; **269**: 19826–19829.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

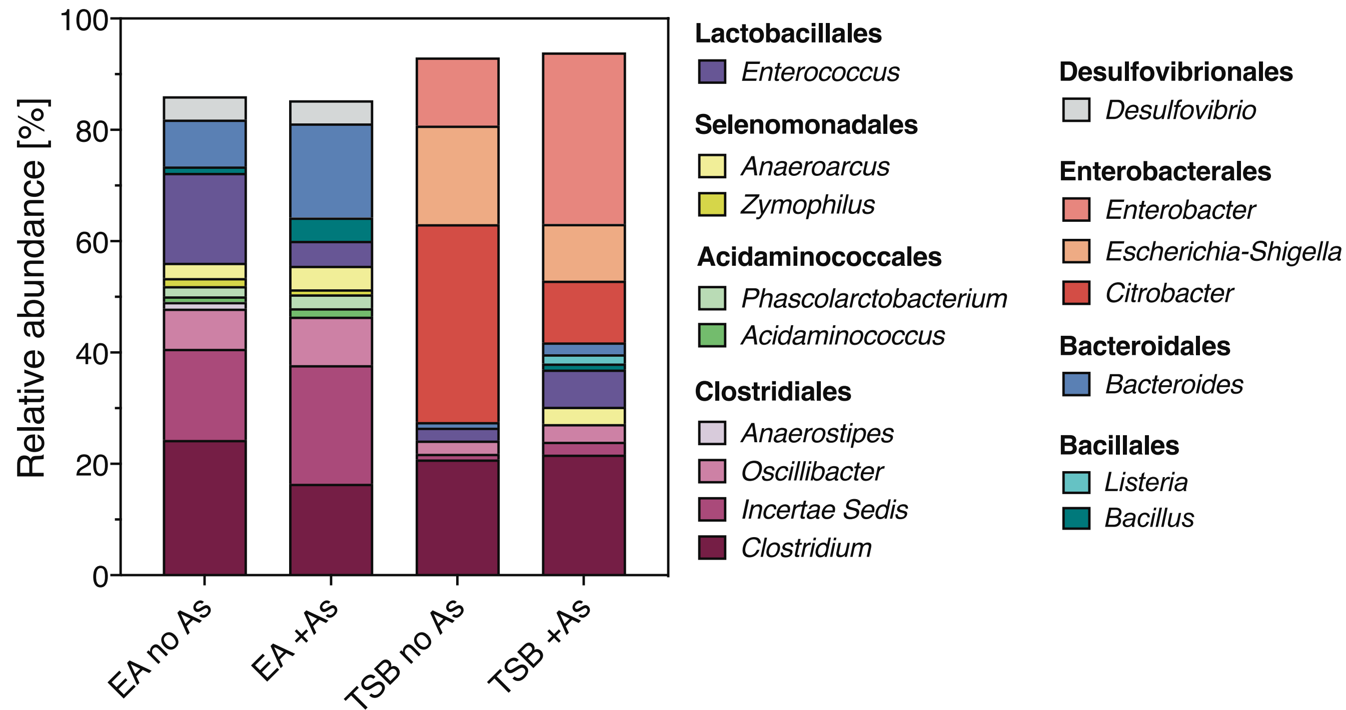
619 23. Murphy JN, Saltikov CW. The ArsR repressor mediates arsenite-dependent regulation
620 of arsenate respiration and detoxification operons of Shewanella sp. strain ANA-3. *J*
621 *Bacteriol* 2009; **191**: 6722–6731.
622 24. Santha S, Pandaranayaka EPJ, Rosen BP, Thiyagarajan S. Purification,
623 crystallization and preliminary X-ray diffraction studies of the arsenic repressor ArsR
624 from Corynebacterium glutamicum. *Acta Crystallogr Sect F Struct Biol Cryst Commun*
625 2011; **67**: 1616–1618.
626 25. Fekih I Ben, Zhang C, Li YP, Zhao Y, Alwathnani HA, Saquib Q, et al. Distribution of
627 arsenic resistance genes in prokaryotes. *Front Microbiol* 2018; **9**: 2473.
628 26. Cai L, Yu K, Yang Y, Chen BW, Li XD, Zhang T. Metagenomic exploration reveals
629 high levels of microbial arsenic metabolism genes in activated sludge and coastal
630 sediments. *Appl Microbiol Biotechnol* 2013; **97**: 9579–9588.
631 27. Zhang SY, Zhao FJ, Sun GX, Su JQ, Yang XR, Li H, et al. Diversity and abundance of
632 arsenic biotransformation genes in paddy soils from Southern China. *Environ Sci*
633 *Technol* 2015; **49**: 4138–4146.
634 28. Huang K, Xu Y, Packianathan C, Gao F, Chen C, Zhang J, et al. Arsenic methylation
635 by a novel ArsM As(III) S-adenosylmethionine methyltransferase that requires only
636 two conserved cysteine residues. *Mol Microbiol* 2018; **107**: 265–276.
637 29. Zhang J, Cao T, Tang Z, Shen Q, Rosen BP, Zhao FJ. Arsenic methylation and
638 volatilization by arsenite s-adenosylmethionine methyltransferase in Pseudomonas
639 alcaligenes NBRC14159. *Appl Environ Microbiol* 2015; **81**: 2852–2860.
640 30. Huang K, Chen C, Zhang J, Tang Z, Shen Q, Rosen BP, et al. Efficient Arsenic
641 Methylation and Volatilization Mediated by a Novel Bacterium from an Arsenic-
642 Contaminated Paddy Soil. *Environ Sci Technol* 2016; **50**: 6389–6396.
643 31. Yin XX, Chen J, Qin J, Sun GX, Rosen BP, Zhu YG. Biotransformation and
644 volatilization of arsenic by three photosynthetic cyanobacteria. *Plant Physiol* 2011;
645 **156**: 1631–1638.
646 32. Zhao C, Zhang Y, Chan Z, Chen S, Yang S. Insights into arsenic multi-operons

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

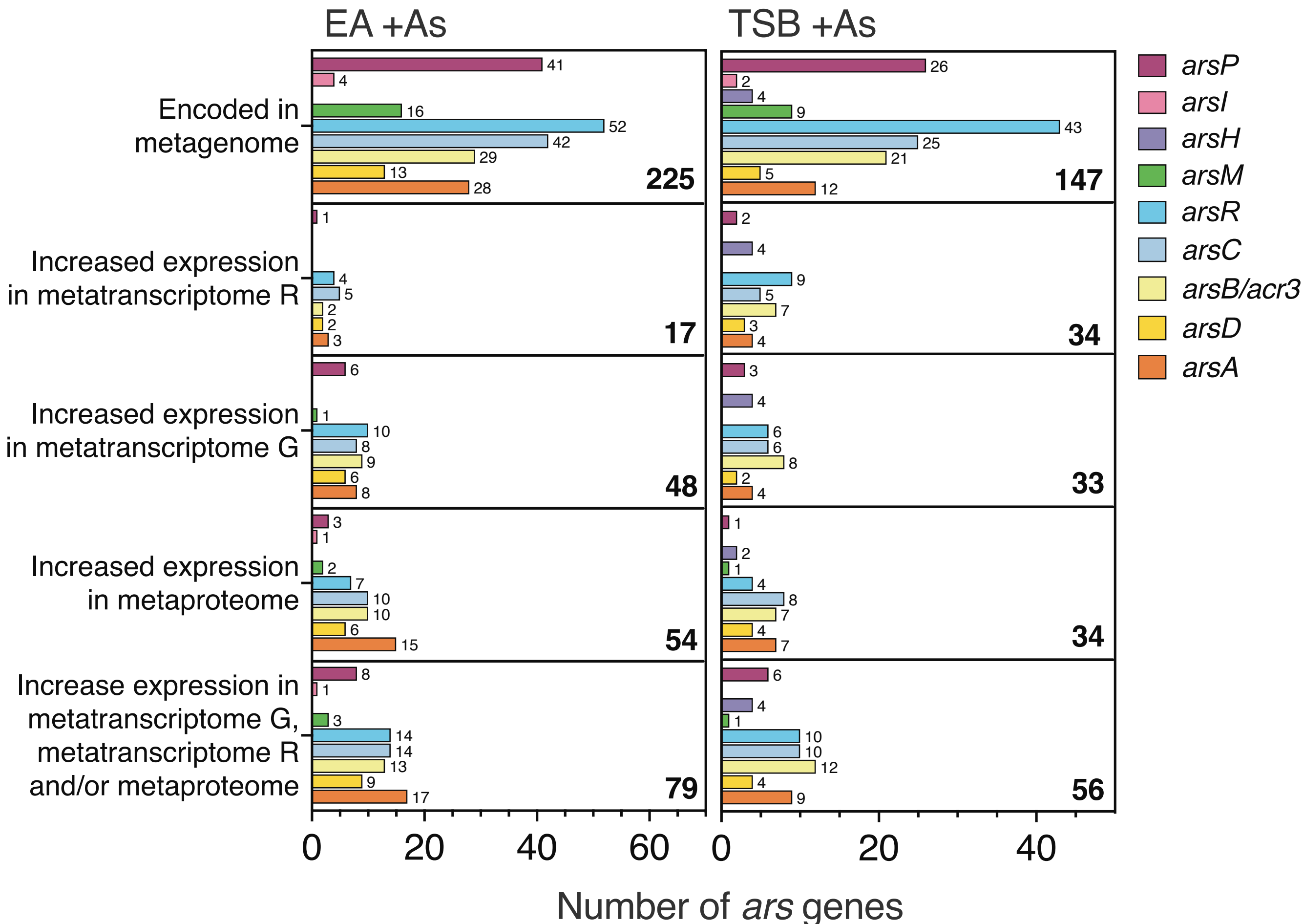
647 expression and resistance mechanisms in Rhodopseudomonas palustris CGA009.
648 *Front Microbiol* 2015; **6**: 986.
649 33. Wang G, Kennedy SP, Fasiludeen S, Rensing C, DasSarma S. Arsenic Resistance in
650 Halobacterium sp. Strain NRC-1 Examined by Using an Improved Gene Knockout
651 System. *J Bacteriol* 2004; **186**: 3187–3194.
652 34. Subhraveti P, Ong Q, Keseler I, Kothari A, Caspi R, Karp PD. Summary of Clostridium
653 sp. CPB-6 CPB6, version 23.5. https://biocyc.org/organism-
654 summary?object=GCF_002119605. Accessed 30 Mar 2020.
655 35. Chen J, Bhattacharjee H, Rosen BP. ArsH is an organoarsenical oxidase that confers
656 resistance to trivalent forms of the herbicide monosodium methylarsenate and the
657 poultry growth promoter roxarsone. *Mol Microbiol* 2015; **96**: 1042–1052.
658 36. Yoshinaga M, Rosen BP. A C-As lyase for degradation of environmental
659 organoarsenical herbicides and animal husbandry growth promoters. *Proc Natl Acad*
660 *Sci U S A* 2014; **111**: 7701–6.
661 37. Chen J, Madegowda M, Bhattacharjee H, Rosen BP. ArsP: A methylarsenite efflux
662 permease. *Mol Microbiol* 2015; **98**: 625–635.
663 38. Chen J, Nadar VS, Rosen BP. A novel MAs(III)-selective ArsR transcriptional
664 repressor. *Mol Microbiol* 2017; **106**: 469–478.
665 39. Doyle CJ, O'Toole PW, Cotter PD. Genomic characterization of sulphite reducing
666 bacteria isolated from the dairy production chain. *Front Microbiol* 2018; **9**: 1507.
667 40. Czyzewski BK, Wang DN. Identification and characterization of a bacterial
668 hydrosulphide ion channel. *Nature* 2012; **483**: 494–497.
669 41. Huang CJ, Barrett EL. Sequence analysis and expression of the Salmonella
670 typhimurium asr operon encoding production of hydrogen sulfide from sulfite. *J*
671 *Bacteriol* 1991; **173**: 1544–1553.
672 42. Harmon SM, Kautter DA, Peeler JT. Improved medium for enumeration of Clostridium
673 perfringens. *Appl Microbiol* 1971; **22**: 688–692.
674 43. George WL, Sutter VL, Finegold SM. Toxigenicity and antimicrobial susceptibility of

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

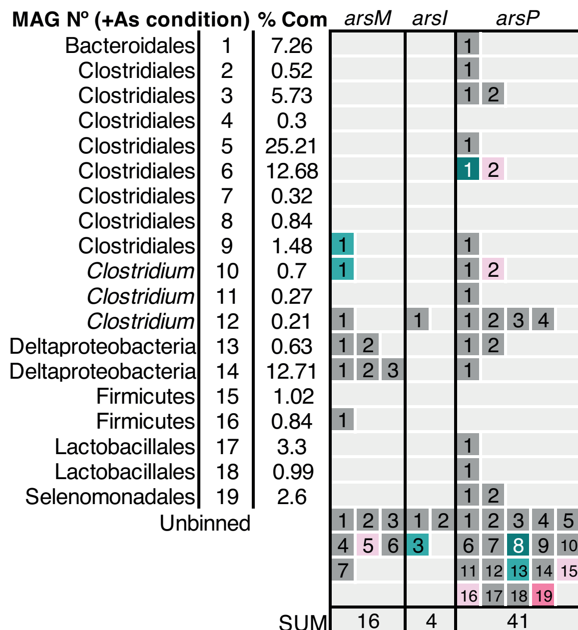
675 Clostridium difficile, a cause of antimicrobial agent-associated colitis. *Curr Microbiol*
676 1978; **1**: 55–58.
677 44. Chen C, Li L, Huang K, Zhang J, Xie WY, Lu Y, et al. Sulfate-reducing bacteria and
678 methanogens are involved in arsenic methylation and demethylation in paddy soils.
679 *ISME J* 2019; **13**: 2523–2535.
680 45. Wang M, Tang Z, Chen XP, Wang X, Zhou WX, Tang Z, et al. Water management
681 impacts the soil microbial communities and total arsenic and methylated arsenicals in
682 rice grains. *Environ Pollut* 2019; **247**: 736–744.
683 46. Zhang S-Y, Xiao X, Chen S-C, Zhu Y-G, Sun G-X, Konstantinidis KT. High arsenic
684 levels increase activity rather than diversity or abundance of arsenic metabolism
685 genes in paddy soils. *Appl Environ Microbiol* 2021.
686 47. Mestrot A, Feldmann J, Krupp EM, Hossain MS, Roman-Ross G, Meharg AA. Field
687 fluxes and speciation of arsines emanating from soils. *Environ Sci Technol* 2011; **45**:
688 1798–1804.
689 48. Huang H, Jia Y, Sun G-X, Zhu Y-G. Arsenic speciation and volatilization from flooded
690 paddy soils amended with different organic matters. *Environ Sci Technol* 2012; **46**:
691 2163–8.
692 49. Zhao F-J, Harris E, Yan J, Ma J, Wu L, Liu W, et al. Arsenic Methylation in Soils and
693 Its Relationship with Microbial arsM Abundance and Diversity, and As Speciation in
694 Rice. *Environ Sci Technol* 2013; 7147–7154.
695



Relative abundance [%]
0
20
40
60
80
100
EA no As
EA +As
TSB no As
TSB +As
Lactobacillales
Enterococcus
Selenomonadales
Anaeroarcus
Zymophilus
Acidaminococcales
Phascolarctobacterium
Acidaminococcus
Clostridiales
Anaerostipes
Oscillibacter
Incertae Sedis
Clostridium
Desulfovibrionales
Desulfovibrio
Enterobacterales
Enterobacter
Escherichia-Shigella
Citrobacter
Bacteroidales
Bacteroides
Bacillales
Listeria
Bacillus



EA +As
TSB +As
Encoded in metagenome
Increased expression in metatranscriptome R
Increased expression in metatranscriptome G
Increased expression in metaproteome
Increase expression in metatranscriptome G, metatranscriptome R and/or metaproteome
225
17
48
54
79
147
34
33
34
56
arsP
arsI
arsH
arsM
arsR
arsC
arsB/acr3
arsD
arsA
Number of ars genes



TSB +As

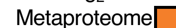
MAG Nº (+As condition)		% Com	*arsM*	*arsH*	*arsI*	*arsP*	*arsR4*
Bacteroidales	1	8.47				1	
Clostridiales	2	0.41				1 2	
Clostridiales	3	1.26			1	1 2	
Clostridiales	4	2.36				1	
Clostridiales	5	4.77				1	
Clostridiales	6	0.35				1	
Clostridiales	7	0.39	1			1	
Clostridiales	8	0.25	1			1	
Clostridiales	9	0.15					
Clostridiales	10	0.17					
Clostridium	11	1.42	1 2			1	
Clostridium	12	2.1	1		1	1 2 3 4	1
Deltaproteobacteria	13	0.21					
Deltaproteobacteria	14	0.81	1 2			1 2	
Enterobacteriaceae	15	0.39					
Enterobacteriaceae	16	6.77		1		1	
Firmicutes	17	0.37					1
Firmicutes	18	1.97	1			1	1
Lactobacillales	19	2.64				1	
Lactobacillales	20	1.94				1	
Selenomonadales	21	0.81				1	1
Unbinned			1	1 2 3		1 2 3 4	
SUM			9	4	2	26	4

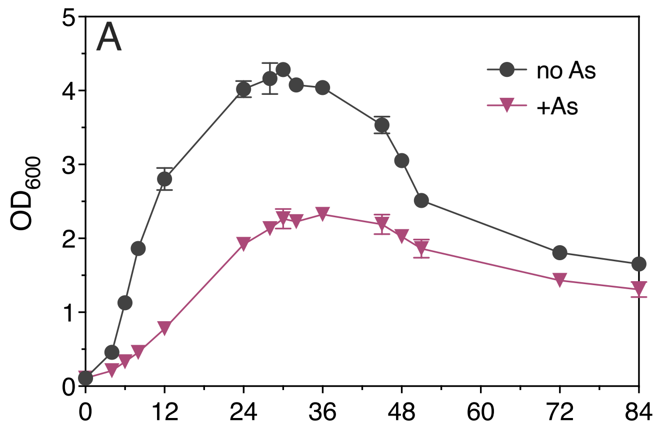
Present in (+As condition):

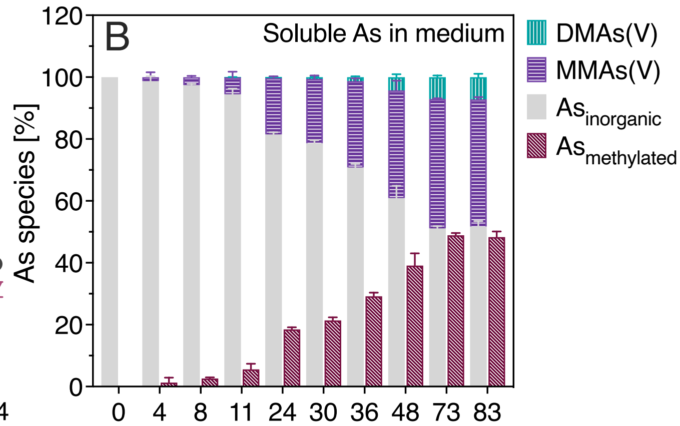
Expression fold change (FC) vs. no-As control:

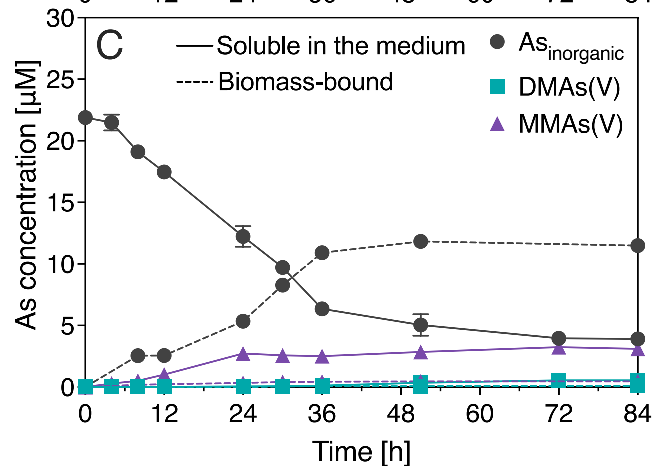
log2 FC ≥1

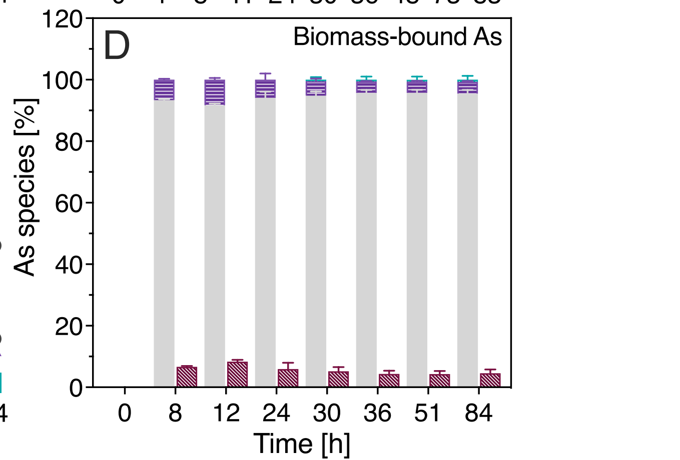
log_2 FC ≤-1



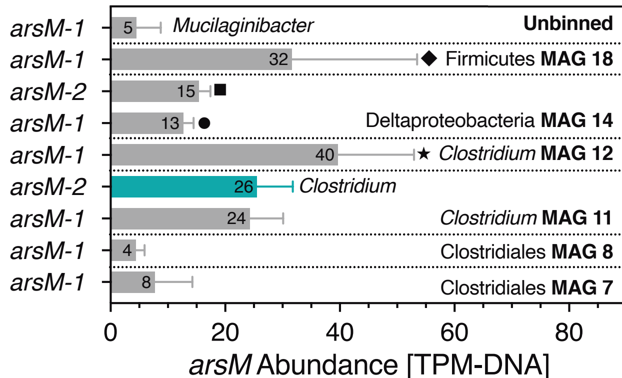

A
OD600
no As
+As
0
1
2
3
4
5
0
12
24
36
48
60
72
84


B
Soluble As in medium
DMAs(V)
MMAs(V)
As inorganic
As methylated
As species [%]
0
20
40
60
80
100
120
0
4
8
11
24
30
36
48
73
83


C
Soluble in the medium
Biomass-bound
As inorganic
DMAs(V)
MMAs(V)
As concentration [µM]
0
5
10
15
20
25
30
0
12
24
36
48
60
72
84
Time [h]


D
Biomass-bound As
As species [%]
0
20
40
60
80
100
120
0
8
12
24
30
36
51
84
Time [h]

EA +As
arsM-7
40 Desulfotomaculum
arsM-6
21 Desulfotomaculum
arsM-5
15 Paeniclostridium
arsM-4
11 Carboxydocella
arsM-3
32 Desulfitobacterium
arsM-2
10 Desulfovibrio
arsM-1
6 Carboxydothermus
Unbinned
arsM-1
26
Firmicutes MAG 16
arsM-3
433
arsM 2
334
Deltaproteobacteria
arsM-1
334
MAG 14
arsM-2
21
arsM-1
17
Deltaproteobacteria MAG 13
arsM-1
7
Clostridium MAG 12
arsM-1
24 Ruminococcaceae
Clostridium MAG 10
arsM-1
55 Paeniclostridium
Clostridiales MAG 9
0
20
40
60
400
600
arsM Abundance [TPM-DNA]



TSB +As
arsM-1
5 Mucilaginibacter
Unbinned
arsM-1
32
Firmicutes MAG 18
arsM-2
15
arsM-1
13
Deltaproteobacteria MAG 14
arsM-1
40
Clostridium MAG 12
arsM-2
26 Clostridium
arsM-1
24
Clostridium MAG 11
arsM-1
4
Clostridiales MAG 8
arsM-1
8
Clostridiales MAG 7
0
20
40
60
80
arsM Abundance [TPM-DNA]

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.25.477449; this version posted March 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Metagenome assembled genomes (MAGs) from EA (upper Table A) and TSB (lower Table B) cultures in the +As condition. Marker lineage: taxonomic rank set by CheckM. Completeness and contamination %: estimated completeness and contamination of genome as determined by CheckM from the presence/absence of single-copy marker genes and the expected colocalization of these genes. Strain heterogeneity: index between 0 and 100 where a value of 0 means no strain heterogeneity, high values suggest the majority of reported contamination is from closely related organisms (i.e., potentially the same species) and low values suggest the majority of contamination is from phylogenetically diverse sources. % of binned proteins assigned to MAG: number of protein-coding genes assigned to the MAG divided by the total number of protein-coding genes binned. % Community: sum of the number of reads mapped to the contigs in each MAG divided by the total number of reads mapped to all contigs including the unbinned contigs, and normalized to MAG size, assuming an average genome size for all unbinned populations. High-quality MAGs are denoted by bolded numbers, good-quality MAGs by italicized numbers.

A

MAG	Bin	Marker lineage*	Completeness (%)	Contamination (%)	Strain heterogeneity (%)	Genome size (Mbp)	% Community	GC content	N. of ArsM-encoding genes	Binner §
1	36	Bacteroidales (o)	98.5	0.4	0.0	3.8	7.26 ±0.32	39.1	0	A
2	15	Clostridiales (o)	98.7	0.0	0.0	2.2	0.52 ±0.06	58.4	0	B
3	21	Clostridiales (o)	95.2	0.0	0.0	4.3	5.73 ±0.24	28.5	0	C
4	4	Clostridiales (o)	90.7	0.0	0.0	2.1	0.30 ±0.03	57.5	0	A
5	24	Clostridiales (o)	97.8	0.3	0.0	2.0	25.21 ±0.23	43.2	0	A
6	9	Clostridiales (o)	100.0	1.3	50.0	3.2	12.68 ±0.68	54.9	0	A
7	51	Clostridiales (o)	85.6	2.2	80.0	4.3	0.32 ±0.12	44.4	0	BC
8	35	Clostridiales (o)	98.0	3.3	55.6	5.3	0.84 ±0.21	44.0	0	A
9	31	Clostridiales (o)	97.9	3.5	0.0	3.8	1.48 ±0.12	28.2	1	A
10	20	*Clostridium* (g)	97.2	2.2	0.0	3.4	0.70 ±0.06	30.1	1	C
11	18	*Clostridium* (g)	96.5	2.9	16.7	4.0	0.27 ±0.05	30.0	0	A
12	42	*Clostridium* (g)	93.4	6.4	0.0	4.6	0.21 ±0.05	32.3	1	A
13	11	Deltaproteobacteria (o)	99.2	0.7	100.0	3.4	0.63 ±0.05	57.4	2	BC
14	33	Deltaproteobacteria (o)	100.0	1.2	0.0	3.3	12.71 ±0.49	57.8	3	BC
15	28	Firmicutes (p)	99.9	0.0	0.0	2.5	1.02 ±0.10	47.2	0	B
16	27	Firmicutes (p)	91.9	3.3	92.3	3.1	0.84 ±0.07	49.2	1	BC
17	8	Lactobacillales (o)	99.6	0.0	0.0	2.7	3.30 ±0.61	36.8	0	A
18	1	Lactobacillales (o)	99.3	4.6	0.0	4.1	0.99 ±0.05	39.1	0	BC
19	16	Selenomonadales (o)	100.0	1.5	0.0	2.2	2.60 ±0.24	41.3	0	C

* (p) phylum, (o) order, or (g) genus.

§ A,B and C refer to MetaBAT 2, MaxBin 2.0 and CONCOCT respectively.

B

MAG	Bin	Marker lineage*	Completeness (%)	Contamination (%)	Strain heterogeneity (%)	Genome size (Mbp)	% Community	GC content	N. of ArsM-encoding genes	Binner §
1	11	Bacteroidales (o)	89.3	1.1	100.0	2.9	8.47 ±1.89	39.2	0	A
2	12	Clostridiales (o)	100.0	0.0	0.0	3.1	0.41 ±0.03	54.8	0	B
3	9	Clostridiales (o)	98.9	0.0	0.0	4.7	1.26 ±0.24	28.4	0	C
4	39	Clostridiales (o)	98.0	0.3	0.0	2.0	2.36 ±0.51	43.2	0	A
5	4	Clostridiales (o)	99.3	0.7	100.0	2.7	4.77 ±3.14	56.1	0	A
6	16	Clostridiales (o)	98.7	0.9	0.0	2.8	0.35 ±0.21	35.7	0	A
7	19	Clostridiales (o)	99.2	1.1	0.0	3.5	0.39 ±0.28	31.2	1	BC
8	1	Clostridiales (o)	98.7	1.3	50.0	2.6	0.25 ±0.07	56.1	1	A
9	2	Clostridiales (o)	74.6	1.3	50.0	2.1	0.15 ±0.02	61	0	A
10	15	Clostridiales (o)	97.3	2.5	16.7	2.7	0.17 ±0.02	60.5	0	C
11	28	*Clostridium* (g)	99.3	5.5	23.1	5.6	1.42 ±0.30	30.1	2	A
12	27	*Clostridium* (g)	98.6	6.9	0.0	4.6	2.10 ±0.60	32.3	1	A
13	32	Deltaproteobacteria (o)	94.8	0.0	0.0	3.1	0.21 ±0.06	59.3	0	BC
14	38	Deltaproteobacteria (o)	98.3	1.8	50.0	3.4	0.81 ±0.08	57.6	2	BC
15	10	*Enterobacteriaceae* (f)	96.6	0.7	33.3	4.3	0.39 ±0.08	52.8	0	B
16	42	*Enterobacteriaceae* (f)	95.7	2.1	12.5	5.1	6.77 ±0.35	56.3	0	BC
17	31	Firmicutes (p)	99.9	0.0	0.0	2.4	0.37 ±0.06	47.6	0	A
18	33	Firmicutes (p)	100.0	0.6	0.0	3.2	1.97 ±1.09	49.1	1	BC
19	7	Lactobacillales (o)	99.6	0.0	0.0	2.9	2.64 ±0.52	36.5	0	C
20	5	Lactobacillales (o)	98.9	4.2	0.0	4.1	1.94 ±0.81	39.1	0	AB
21	36	Selenomonadales (o)	100.0	1.5	0.0	2.3	0.81 ±0.11	41.1	0	A

* (p) phylum, (o) order, or (g) genus.

§ A,B and C refer to MetaBAT 2, MaxBin 2.0 and CONCOCT respectively.