Lysogenic bacteriophages encoding arsenic resistance determinants promote bacterial community adaptation to arsenic toxicity

Xiang Tang[1, 2], Linrui Zhong[1], Lin Tang[1], Changzheng Fan[1, *], Baowei Zhang[1], Mier Wang[1], Haoran Dong[1], Chengyun Zhou[1], Christopher Rensing[2], Shungui Zhou[2] and Guangming Zeng[1, *]

[1] College of Environmental Science and Engineering, Hunan University and Key Laboratory of Environmental Biology and Pollution Control (Hunan University), Ministry of Education, Changsha 410082, P.R. China

[2] Fujian Provincial Key Laboratory of Soil Environmental Health and Regulation, College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, P.R. China

[*] Corresponding author:
fancz@hnu.edu.cn (C.Z. Fan);
zgming@hnu.edu.cn (G.M. Zeng)

ABSTRACT

Emerging evidence from genomics gives us a glimpse into the potential contribution of lysogenic bacteriophages (phages) to the environmental adaptability of their hosts. However, it is challenging to quantify this kind of contribution due to the lack of appropriate genetic markers and the associated controllable environmental factors. Here, based on the unique transformable nature of arsenic (the controllable environmental factor), a series of flooding microcosms was established to investigate the contribution of *arsM*-bearing lysogenic phages to their hosts' adaptation to trivalent arsenic [As(III)] toxicity, where *arsM* is the marker gene associated with microbial As(III) detoxification. In the 15-day flooding period, the concentration of As(III) was significantly increased, and this elevated As(III) toxicity visibly inhibited the bacterial population, but the latter quickly adapted to As(III) toxicity. During the flooding period, some lysogenic phages re-infected new hosts after an early burst, while others persistently followed the productive cycle (i.e., lytic cycle). The unique phage-host interplay contributed to the rapid spread of *arsM* among soil microbiota, enabling the quick recovery of the bacterial community. Moreover, the higher abundance of *arsM* imparted a greater arsenic methylation capability to soil microbiota. Collectively, this study provides experimental evidence for lysogenic phages assisting their hosts in adapting to an extreme environment, which highlights the ecological perspectives on lysogenic phage-host mutualism.

INTRODUCTION

Phages (prokaryotic viruses) are capable of shaping the structure and function of microbial communities by influencing nutrient metabolism and mortality [1], which have profound impacts on microbial ecosystems [2]. Lysogenic phages are able to integrate their genome stably into the host genome to form prophages (i.e., latent lysogenic phages) rather than rapidly kill their hosts as lytic phages [3-5]. With respect to lysogenic phages, the establishment of lysogeny effectively decouples infection from host death, minimizing harmful environmental exposure [6, 7], and allows virions to be released in time for optimal reproduction and persistence [8]. In some context, the host in return benefits from the auxiliary metabolic genes carried by lysogenic phages [9-13]. The transformation of phage-bacteria interaction from parasitism to protective mutualism in extreme environments demonstrates the influential role of lysogenic phages in bacterial adaptation to environmental stress [14]. Despite previous studies highlighting the significance of lysogenic phages in assisting their host' environmental adaptability in different habitats, most of these are based on sequence alignment of viral genome sequences as a basis to further speculate on this process [13-15]. To date, the authentic contribution of lysogenic phages to the environmental adaptability of their hosts is still poorly understood.

Arsenic (As) is a ubiquitous metalloid displaying high toxicity [16], and inorganic arsenic (the most abundant species present in the environment) has been shown to mainly exist in one of two forms, arsenate [As(V)] and arsenite [As(III)], depending on redox potential (Eh) and pH in the environment [17, 18]. Typically, the toxicity and

bioavailability of As(III) is far greater than that of As(V). A lower oxidation-reduction potential value (E_h) can promote the conversion of As(V) to As(III), which inevitably render greater stress on microorganisms in anoxic soils such as wetlands and flooded soils [19]. This readily interchangeable nature inherent to arsenic is considered to pose a threat to microbial populations, and the subsequent elevated As(III) toxicity is a challenge to microbial fitness [20, 21]. In response to environmental arsenic stress, microorganisms have evolved sophisticated microbial arsenic resistance systems that incorporate precipitation, chelation, compartmentalization, extrusion or biochemical transformation [22-24]. Arsenic methylation, a function encoded by the As(III) methyltransferase gene (i.e., *arsM*) is a vital intracellular As(III) detoxification pathway [25]. Microorganisms expressing *arsM* were highly diverse in soil, both phylogenetically and ecologically [26-28]. Shimen realgar mine (SM), located in Hunan Province of China, is the largest realgar ore deposit in Asia, with approximately 1500 years of mining history. The long-term history of arsenic contamination recruited lysogenic phages to establish bacterial arsenic resistance system, including abundant *arsM*-bearing lysogenic phages [29, 30]. The quantifiable presence of *arsM* in lysogenic phages and the controllable As(III) toxicity (by reducing Eh) allows for the dynamic surveillance of phage-host interplay, giving us the opportunity to quantify the contribution of *arsM*-bearing lysogenic phages to the As(III) adaptability of their hosts.

In this study, the flooding microcosms inoculated with lytic phage-free SM soil were first established to investigate the development of the active bacterial community (by 16S rRNA gene sequencing) and the lysogenic phage population (including

prophage and free phage derived from initial lysogenic phage) under increasing arsenic toxicity. Then, virome sequencing was subsequently conducted to determine the profile of different phage population. Subsequently, the abundance and diversity of viral-encoded *arsM* was determined and the connections of their presence to changes in the active bacterial community were evaluated in order to corroborate the significance of phage-mediated horizontal gene transfer (HGT) to the adaptive evolution of the bacterial community. Finally, the copy number of *arsM* acquired by the soil microbiota after flooding and the resulting arsenic methylation capacity were determined to infer the authentic contribution of lysogenic phages on the arsenic detoxification capability of soil microbiota. This empirical study combines insights into phage-host interplay and speciation transformation of heavy metal(loid) and we anticipate our findings to be helpful to better understand the contribution of lysogenic phages to bacterial adaptability.

MATERIALS AND METHODS

Establishment of batch incubation microcosm

In this work, soil was collected from the core region of SM (29°39′30″N, 111°02′20″E), and the procedures of soil sampling and its main characterization were detailed in Supplementary Information. In order to focus on the interactions between lysogenic phages and their host, potassium citrate buffer (10.0 g/L $C_6H_5K_3O_7$, 1.92 g/L $Na_2HPO_4 \cdot 12H_2O$, and 0.24 g/L KH_2PO_4, pH = 7) was used to elute free phages before establishing microcosms to eliminate the interference from the lytic phages in free phages. In brief, 50 g air-dried soil was re-suspended with 300 mL potassium citrate buffer and the virus was desorbed with ice bath sonication for 3 min (47 kHz, with 30 s of manual shaking at every minute). The mixture was then centrifuged at 11,000× g for 10 min to recover soil and soil bacteria. The supernatant was discarded after centrifugation and the procedure was repeated one more time to ensure reliable removal of free phages. After that, the precipitate containing soil and bacteria was hold for establishing batch incubation microcosm. The pretreated soil was re-suspended in a 250 mL Erlenmeyer flask by adding 100 mL of a sodium acetate solution (10 mM in sterilized distilled water). Acetate was selected here in order to create low Eh rapidly. Meanwhile, the distilled water was through 0.02-μm PVDF filter but not nitrogen-purged. Vials were then capped with a butyl rubber septum to prevent moisture loss, and incubated stationary at room temperature in the dark. The reduction batch incubations were conducted for 15 days under anoxic conditions, and specimens [soil and soil solution (if available)] were sampled by sacrificing three bottles each at day 0,

1, 2, 5, 10 and 15. The procedures of the determination of different arsenic species, the quantification and sequencing of 16S rRNA gene (after RNA reverse transcription) were detailed in Supplementary Information.

Step extraction of free phages and prophages

In order to reveal the dynamics of two lysogenic phage subpopulations (i.e., the free phages derived from lysogenic phage and the prophages remaining in the host), we performed separate-extraction of different phages in the microcosm during flooding period (Figure 1).

Several steps were undergone to extract and enrich free phages from the microcosm following previous protocols [31]. Briefly, the mixture of the microcosm (i.e., soil slurry) was suspended in 300 mL of 4 °C sterilized potassium citrate buffer and incubated at 4 °C for 15 min. After bath sonication in a water-ice mixture for 3 min at 47 kHz (with 30 s of manual shaking at every minute), the suspension was centrifuged (11,000× g, 4 °C, 10 min) to obtain the virus-containing supernatant. Such supernatant was subsequently followed by centrifugation (11,000× g for another 15 min) and filtration (1.0, 0.45, and 0.22 μm in order) to remove impurities larger than 0.22 μm [32]. This filtration process inevitably excludes viruses larger than 0.22 μm, leading them not to be included in further analysis. The virus particles in the filtrate were initially concentrated by tangential flow filter cartridge (Sartorius Vivaflow 200R with 100,000 MWCO PES membrane). The residual DNA was digested with DNase I (Thermo Scientific, USA; final concentration: 1.5 U/mL) at 37 °C, and this reaction

was terminated by 20 mM of EDTA after 30 min. Thereafter, the digested virus mixture was precipitated by polyethylene glycol 8000 (in 1 M NaCl) at a final concentration of 10% (4 °C, 6 h). Subsequently, the phage particles were further concentrated using ultracentrifugation at 25,000 rpm (BECKMAN OptimaTM L-100XP, USA). Then, the virus pellets were centrifugated and resuspended in 1/20 (v/v) SM buffer [50 mL/L Tris-Cl (1 M, pH=7.5), 5 mL/L gelatin solution (2%), 5.8 g/L NaCl, and 2.0 g/L $MgSO_4 \cdot 7H_2O$]. Lastly, the obtained phage particles were stored at -80 °C in a 30% glycerol-medium solution before counting and nucleic acid extraction.

In order to provide carbon, energy and other nutrient resources to support viral production during prophage induction assays, sterile water-soluble organic carbon (WSOC) solutions were prepared. For the preparation of the WSOC solution, each soil sample was suspended in deionized water and blended at the maximum speed for 3 minutes. The soil extractions were then centrifuged at 5,000 × g for 20 minutes, and the supernatants were filtered through 0.02-μm PVDF filter and then autoclaved for use in induction assays. The induction of prophages was performed as previously described with modifications [33]. After being shaken for 20 min at maximum speed, the soil slurry was inoculated into sterile WSOC solution (1:2), and this mixture contained mitomycin-C at a final concentration of 1 μg/mL. Subsequently, this culture was incubated at room temperature, with shaking at 200 rpm for 24 h under in the dark. The culture was concentrated following the above procedures to harvest the induced free phages (Figure 1). Note that the subsequent obtained results of prophages were calculated based on the results obtained by treatment with mitomycin-C and then

subtracting the control treatment (where mitomycin-C was replaced with sterile water).

Enumeration of phage particle and extraction of viral DNA

The harvested phage particles were treated with 4 °C precooled electron-microscopy-grade glutaraldehyde (final concentration was 0.5%) as a fixative at 4 °C for 20 min prior to staining. After that, this viral suspension was vacuum filtered (less than 13 kPa pressure) through a 0.02-μm-pore-size Whatman Anodisc filter. The inverted fluorescence microscope (Olympus BX 61, Japan) was used to observe phage particles stained with SYBR Gold fluorescent dyes (phenylenediamine as antifade) as previously described [34]. Viral DNA was extracted via the TIANamp Virus DNA/RNA Kit (TIANGEN, China) following the manufacturer's instructions. It should be mentioned that the enrichment method of viral nucleic acid in this study is not perfect in acquisition of RNA viral information. In addition, transmission electron microscopy was used to determine phage acquisition (Figure S1).

Virome sequencing and analysis

To provide more information regarding the development of phage-host interaction in the flooding period, viral metagenomic sequencing was performed at three points when the active bacterial community was undergoing significant changes (i.e., day 0, 2 and 15). Only one subset (i.e., prophages on day 0) was included on day 0, whereas two subsets (i.e., prophages and free phages) were included on day 2 and 15. Before sequencing, whole-genome amplification [KAPA HiFi HotStart ReadyMix (Fisher

182 scientific, USA)] of the extracted viral DNA was subjected to meet the virome
183 sequencing requirements. Then, five amplification products were sent for genomic
184 libraries construction and sequencing on Hiseq 4000 system (2 × 150 bp, paired-end
185 reads) at MAGIGENE Biotechnology Co., Ltd. (Guangdong, China). Finally, 13.9 ±
186 1.3 GB raw bases were obtained for each of the samples. The quality control of raw
187 reads was conducted by Trimmomatic, and high-quality reads [(85.9 ± 1.1)% of raw
188 reads, the number of clean reads for each group see Table S1] were left to assemble
189 sequences with Megahit (v1.1.2) [35]. The length of assembled contigs ranged from
190 300 bp to 331503 bp. After that, CheckV (AII used for high/medium-confidence contigs
191 and HMM used for lower-bound contigs) and Virsorter2 (as the supplement) were used
192 to identify viral contigs from 308638 ± 290148 assembled raw contigs [36]. Finally, A
193 total of 131738 non-redundant viral contigs were included for further taxonomy
194 annotation (including abundance analysis; gene prediction and functional annotation)
195 and host prediction. In this work, taxonomy annotations of viral contigs were assigned
196 using VPF-Class [37] (against IMG/VR database) and CheckV as previously described
197 [38]. The abundance of each contig was normalized to Reads Per Kilobase per Million
198 mapped reads (RPKM) for comparison; the gene prediction of viral contigs was
199 conducted via Prokka (v1.13) [39]. The functional annotation of virial contigs was
200 performed by two methods: 1) BLASTP (v2.9.0+) against UniProtKB/Swiss-
201 Prot_ViralZone database (e-value threshold of 10^{-3} and 95% nucleotide identity); and
202 2) diamond against KEGG database (e-value threshold of 10^{-3} and best-hit). Given the
203 potential interference of some remaining free phages (less than 2E+4 VLPs/g soil)

204 associated with microbial cells and soil particles, we only focused on changes in
205 lysogenic phage population. For this end, we conducted the further quality control of
206 viral contigs based on lysogenic phage biomarkers (i.e., transposase, integrase,
207 excisionase, resolvase, and recombinase) [40], and a total of 1,075 viral contigs with an
208 average length of 8,344 bp were obtained. The composition of dominant putative
209 lysogenic phages in different times was quite different (Figure S2), indicating the rapid
210 developments in lysogenic phages. For host prediction, two methods were utilized: 1)
211 Viral contigs were aligned to the spacers in microbial CRISPR regions by
212 SpacePHARER to link viral contig with their potential bacterial hosts in the public
213 database [41]; 2) VPF-Class was also used to predict phage-bacteria linkages [37].
214
215 **Quantification of *arsM* and amplicon sequencing of viral *arsM***
216 As in previous reports, the primers arsMF1 [primer sequences (5'-3'):
217 "TCYCTCGGCTGCGGCAAYCCVAC"] and arsMR2 [primer sequences (5'-3'):
218 "CGWCCGCCWGGCTTWAGYACCCG"] were used to amplify bacterial and viral
219 *arsM* in the specimens [29, 30]. The standard curve preparation (Figure S3) and
220 execution conditions for quantitative PCR were available in Supplementary
221 Information. Furthermore, the copy number of *arsM* was normalized to the microbial
222 biomass (counted as 16S rRNA gene copy) and the total number of phage particles
223 (counted as VLPs) to minimize variances caused by different background bacterial/viral
224 abundance.
225 In the flooding period (day 0, 1, 2, 5, 10, and 15), the viral DNA extracted from

free phages and prophages was amplified by barcoded primer pair arsMF1 and arsMR2. These amplicons were pooled and sequenced on MiSeq system (Majiorbio, China). The raw sequence data of paired-end reads were denoised through DADA2 within the QIIME2 environment [42]. Specifically, the filtered sequences (26363 ± 6662 for each independent sample) were clustered into amplicon sequence variants (ASVs) against the NT database using an open-reference Bayes feature classifier using 0.7 as the minimum confidence threshold. However, the annotation coverage of viral *arsM* was extremely limited, in which over 99% ASVs were unannotated (data not shown). Therefore, phylogenetic analysis was used to identify the putative origin of viral *arsM*. In short, similar sequences of the top 10 most abundant ASVs in all samples (24 non-redundant ASVs), which met the e-value threshold of 10^{-5} and 80% nucleotide identity/coverage from NCBI database, were retrieved using BLASTP. MEGA 11 was used to construct a Neighbor-Joining tree (alignment with ClustalW), and further visualized in iTOL.

Arsenic methylation capacity assays

The ArsM determinant encoded by *arsM* is responsible for methylating intracellular As(III) for detoxification. In this gradual process, As(III) is eventually methylated into volatile nontoxic trimethylarsine [TMAs(III)], thereby alleviating As(III) toxicity. Here, the arsenic methylation ability of the flooded soil microbiota (from 15-days flooded microcosm as described in “Establishment of batch incubation microcosm”) was assayed in other microcosms. The experimental set-up was consistent

with previously reported [43]. In short, each Erlenmeyer flask (250 mL) contained 50 g of air-dried flooded soil with 15 mL of 10 mM sodium acetate solution to maximize arsenic methylation [44]. The control group was spiked with untreated SM soil to reveal the arsenic methylation in pristine soil, and the abiotic control group was spiked with gamma-ray treated flooded soil (50 kGy). Trapping tubes for volatile arsenic were prepared by filling the oven-dried (overnight at 70 °C) silica gel beads (0.5~1.0 mm) impregnated with 10% $AgNO_3$ (24 h) into glass tubes and connected to the flask. The headspace was refreshed by pumping filtered air with pumps at intervals (24 h) for 30 min each time. Each microcosm was replicated five times. All flasks were shaken in the dark at 150 rpm at room temperature for 7 days. All trapping tubes were taken off, and the captured TMAs(III) on silica gel beads was extracted by 5.0 mL of 1% HNO_3 (60 °C for 10 min, 80 °C for 10 min, 100 °C for 30 min). Finally, the produced TMAs(III) gas was identified by oxidizing it to TMAs(V)O with H_2O_2 [45] since both trimethylarsine oxide (TMAO) and As(III) showed similar retention times in an anion exchange chromatogram [46].

Statistical analyses

For virome sequencing, viral genomes extracted from three parallel microcosms were combined into one nucleic acid sample (e.g., day 0_pro). Otherwise, all experiments were performed independently in triplicate and the results were expressed as mean ± standard deviation. ANOVA and Student's t-test for multiple comparisons were used to determine statistical significance (SPSS 23 software).

RESULTS

Evolution of active bacterial community under increasing arsenic toxicity

The distribution profile of arsenic species in SM soil was profoundly changed as flooding progressed. Specifically, the PO_4-As(V) (PO_4- represents 0.1% phosphoric acid extractable) was the dominant arsenic specie on day 0, and PO_4-As(III) were also detected (Figure S4A), whereas dissolved-As(III) (dissolved- represents directly detectable) and dissolved-As(V) were minimal (Figure S4B). With the proceeding of flooding, the concentrations of dissolved-As(III) and PO_4-As(III) were consistently and rapidly elevated within the first 5 days. For example, after 24 h (i.e., day 1), the concentration of dissolved-As(III) increased from 0.2 ± 0.0 mg/L to 26.0 ± 0.9 mg/L and the corresponding concentration of PO_4-As(III) was also raised from 20.8 ± 4.7 mg/L to 100.2 ± 8.6 mg/L. Furthermore, the concentration of dissolved-As(III) further increased about 8-fold on day 5 compared with that on day 1, and the corresponding concentration of PO_4-As(III) also increased 3-fold. The total As(III) concentration increased on day 10, while the corresponding concentration of dissolved-As(III) decreased to 185.7 ± 18.9 mg/L (Figure S4C), which indicated the formation of precipitable As(III) [e.g., As(III) adsorbed on iron (oxyhydr)oxides] [47, 48]. When looking at the influence of Eh and pH, the former was the dominating driver for the transformation of arsenic speciation in arsenic-contaminated soil since no significant change in pH was observed during the flooding period (Figure S5), but the *Eh* at the soil-water interface decreased continuously (Figure S6).

The active bacterial community was significantly changed during the flooding period. Within the domain Bacteria, 12 distinct phyla and more than 39 genera were detected with a relative abundance >1% (at least in one group). The most dominant active bacterial phyla were quite different at different times. For instance, the top 3 dominant phyla on day 0 were Proteobacteria, Actinobacteriota, and Planctomycetota, yet the dominant evolving phyla were Firmicutes, Actinobacteriota, and Proteobacteria on day 1 and day 2 (Figure S7). The divergence of the active bacterial community was more prominent at the genus level (Figure 2A). In the comparison of the top 10 dominant genera, there was only one genus (i.e., *Marmoricola*) shared on day 0 and day 1. Furthermore, the percentage of sequences annotated as *Bacillus* increased from (0.1 ± 0.1) % to (29.8 ± 4.6) % within only 24 h (day 0 to day 1). Typically, *Bacillus* sp. displayed high As(III) resistance [49, 50], partly attributed to their spore-forming capacity [51]. The abundance increment of *Bacillus* suggested their hyper-adaptability to rapidly increasing As(III) toxicity. The microbial community between day 2 and day 5 also displayed a very different development. For example, the relative abundance of genus *Magnetospirillum* increased from (0.1 ± 0.0) % on day 2 to (37.4 ± 10.0) % on day 5 (Figure 2A), which suggested the potential transformation of metal compounds in the microcosm. Both *Pseudomonas* and *Anaeromyxobacter* have been frequently detected in arsenic-contaminated sites [20, 52], and some of them were able to respire As(V) under anoxic conditions [53]. They were both detected in SM soil (day 0), but *Pseudomonas* was not well adapted to As(III) toxicity since its abundance decreased rapidly (Figure 2A).

The network analysis visualized the differentiation of active microbial communities during the 15-day flooding period (Figure 2B). The node clusters at the genus level explained the differentiation of active microbial communities on day 0/1 and day 2/5, but could not explain the corresponding differentiation on day 1/2 and day 5/10/15 (nor could principal component analysis, Figure S8). Therefore, the development of the active bacterial community could be divided into stage-I (day 0), stage-II (day 1 to 2) and stage-III (day 5-15). Redundancy analysis further identified the environmental factors driving the evolution of microbial community. When examining different arsenic species, the concentration of dissolved-As(III) displayed the highest correlation with development of active microbial populations, followed by the content of PO_4-As(III) (Figure 2C). To discriminate between the deterministic and stochastic processes in the assembly of the active bacterial community as flooding proceeded, the βNTI of every sample was calculated. Variable selection assembly processes were dominant during the whole flooding period (βNTI > 2, see Table S2). These results suggested that the environmental stress derived from As(III) toxicity was the main driving force of bacterial community evolution. As shown in the dynamics of the copy number of the 16S rRNA gene (Figure 2D), As(III) toxicity impeded the health of the bacterial community within stage-I and stage-II, and this suppression was mitigated in stage-III. Moreover, the changes in richness, evenness and diversity of the active bacterial community followed the same trends as the copy number of 16S rRNA gene, which illustrated that the bacterial community in flooding microcosm was restored (Table S3).

335 **Dynamics of phage-host interplay during flooding**

336 The lysogenic phage population developed alongside soil microbiota. Here, (93.5
337 ± 2.9)% of the recovered viral contigs were identified as phage, suggesting that VLP
338 can serve as a reasonable indicator to characterize phage density (Figure 3A). Most of
339 the free phages had been eluted since the number of residual free phages on day 0 was
340 less than 2E+4 virus-like particles (VLPs)/g soil (without detectable DNA). Meanwhile,
341 in the no-flooding control microcosm, the prophage concentration, free phage
342 concentration, and *arsM* abundance in viral genome and bacterial genome in eluted SM
343 soil did not significantly change (determined every 8 h in a 3-day period, data not
344 shown), which confirmed that the subsequent experimental results were not affected by
345 potential unreleased lytic phages (in virocells). In the SM soil used in this work, the
346 number of prophages (i.e., day 0) was (6.6 ± 0.1)E+8 VLPs/g soil, and these prophages
347 predominantly belonged to the order *Caudovirales* (Figure S9). Moreover, the viromes
348 of SM soil were highly novel and variant in terms of genetic profiles, and approximately
349 (90.9 ± 6.2)% of the lysogenic phage sequences did not matched to recorded viral
350 sequences at the genus level (Figure 3B). After 24 h, a significant new virion release
351 was observed, where the number of free phages increased to (3.9 ± 0.1)E+9 VLPs/g
352 soil, which illustrated that many prophages were induced within the first day (Figure
353 3C). The results of host prediction indicated that the prophages separated from day 0
354 (i.e., day 0_pro) mainly resided in the phylum Proteobacteria (Figure 3D). Therefore,
355 the decrease in the abundance of Proteobacteria from day 0 to day 1 (Figure S7) could
356 be partially explained by prophage induction. Meanwhile, the corresponding proportion

of lysogenic phages out of total phages declined to (8.1 ± 0.2) % (Figure 3C). The abundance of the genus *Bacillus*, which is often reported to bear abundant prophages [also as one of the main putative hosts for phages in SM soil (Figure 3D)] [54-56], was significantly increased on day 1 (Figure 2A). Such result undoubtedly corroborated the adaptability of *Bacillus* genus to stronger As(III) toxicity [49, 50]. On day 2, the number of prophages recovered was (1.3 ± 0.1)E+9 VLPs per g soil (Figure 3C), which indicated that a fraction of the As(III)-induced lysogenic phages re-infected and resided in new hosts. Such rise could not be attributed to the clonal expansion of As(III)-resistant taxa, as the copy number of the 16S rRNA gene declined steadily on day 2 (Figure 2D) and *unclassified_f_Oxalobacteracese* (i.e., the genus with the most significant increment on day 2) was not identified as the putative host of any lysogenic phage (Figure 3D).

Surprisingly, the number of free phages also showed the highest concentration on day 2 throughout the flooding period. The simultaneous growth of prophages and free phages suggested that a fraction of phages followed a productive cycle, and further multiplied the number of free phages in the microcosm by releasing new virions. The top 20 abundant viral contigs in different times were compared, and the result showed that seven contigs were shared in day 0_pro and day 2_free (Figure 4A), which suggested that the bursting phages originated from the latent prophages on day 0. Given that most of the viral contigs cannot be annotated, principal coordinate analysis was performed to visualize the differences among different phage populations. Here, day 0_pro and day 2_free were very similar (Figure 4B), which as we can expected from

Figure 4A. This compositional homology supported an influential role for prophage induction in the expansion of lysogenic phages. Interestingly, day 15_pro and day 15_free also presented a similar composition (Figure 4B). Besides, the number of prophages was significantly correlated to the concentrations of dissolved-As(III) (r = 0.91, $p < 0.01$), which suggested the positive selection of lysogenic phage re-infection by soil microbiota under elevated As(III) toxicity (Figure 4C). Comparatively, the number of free phages was poorly correlated with the concentrations of dissolved-As(III) (Figure 4D) (r = -0.04, p = 0.86), implying that there are other drivers of the development influencing the free phage population except As(III) toxicity.

The *arsM*-bearing lysogenic phages facilitated the restoration of bacterial community

The transduction of *arsM* by lysogenic phages aided their hosts' adaptation to As(III) toxicity. On day 0, the copy number of *arsM* in prophages was (2.2 ± 0.1)E+7 copies/g soil (Figure 5A), which accounted for (4.9 ± 0.2) % of the *arsM* copy number on bacterial genomes. These results suggested that lysogenic phages participated in the establishment of the arsenic resistance system of the microbial community and, alternatively, indicated that methylation was an indispensable arsenic detoxification pathway under high arsenic stress [29]. On day 1, the abundance of *arsM* in prophages was decreased to (6.8 ± 0.3)E+5 copies/g soil. Meanwhile, a large number of prophages were also induced at this time (see "Dynamics of phage-host interplay during flooding"). To avoid bias caused by changes in phage numbers, the *arsM* copy numbers in phage

genomes were normalized to the number of *arsM* copies carried by each VLPs. The relative abundance of *arsM* in prophages declined from (3.3 ± 0.2)E-2 copies/VLP on day 0 to (2.0 ± 0.1)E-3 copies/VLP on day 1 (Figure 5B). These results implied that most *arsM*-bearing prophages were induced within 24 h. After a transitory decline, the abundance of *arsM* in prophages increased constantly in subsequent times. For example, on day 15, the absolute and relative abundance of *arsM* in prophages increased to (1.2 ± 0.0)E+9 copies/g soil and (3.1 ± 0.1)E-1 copies/VLP, respectively. The growth of relative abundance of *arsM* in prophages indicated that the induced *arsM*-bearing lysogenic phages have a strong infectious capacity. The relative abundance of *arsM* in free phages also increased continuously. For example, the relative abundance of *arsM* in free phages was (2.9 ± 0.1)E-1 copies/VLP on day 15, while the corresponding abundance on day 1 was only (2.8 ± 0.2)E-2 copies/VLP (Figure 5B). Figure 5C showed a significant positive correlation between the relative abundance of *arsM* in prophages and that in free phages, which suggested that the free phages were released from bacteria that newly infected with lysogenic phages [57]. This phenomenon was consistent with the synergistic development of the community structure between prophages and free phages (Figure 4B).

In order to provide direct evidence that prophages were repeatedly lysogenized to spread *arsM*, lysogenic phage-like viral contigs that persisted in the prophage subpopulation and whose abundance increased by more than 10 (RPKM) were searched for the presence of *arsM*. Although the *arsM* fragment cannot be obtained by searching from these viral contigs because it contained many variable regions [26] (that is *arsM*

423 fragments were separated in overlapping process). However, primer arsMF1/arsMR2
424 that targeted on conserved region of *arsM* fragments were able to dig potential *arsM*
425 fragments from persisted viral contigs (match at least 18 bases at 3' of primers or 21
426 bases at 5 ends of primers, allowing single base mismatches). Many viral contigs carried
427 *arsM* fragments and their abundance increased during flooding (Table S4). For example,
428 the concentration of *pro.15|contig_527121* in day 0_pro was 0.599 (RPKM), but
429 increased to 3939.846 in day 15_pro; the concentration of *pro.15|contig_5203864* in
430 day 0_pro was 0.002, but increased to 221.883 in day 15_pro. These results strongly
431 support that prophages were repeatedly lysogenized to spread *arsM*.

432 To reveal the potential impact of *arsM*-bearing phages on their hosts from a
433 genomic perspective, the changes in the genetic diversity of viral *arsM* were identified.
434 As exhibited in Figure 6A, viral *arsM* had undergone dramatic alterations during
435 flooding. For example, on day 0, the most abundant *arsM* taxa were ASV13 and ASV7,
436 while they rapidly evolved to ASV1 and ASV2 on day 2. Similar to the changes in the
437 active bacterial community, the changes in the viral *arsM* taxa became gentle in the
438 Stage-III, where ASV11 and ASV8 were the dominant *arsM* taxa shared on day 5, 10
439 and 15. Overall, the genetic diversity of viral *arsM* showed a clear downward trend, as
440 evidenced by the α-diversity of the *arsM* taxa (Table S5). Canonical correlation analysis
441 further identified the environmental factors that had the greatest impact on *arsM*
442 evolution. The As(III) concentration of [dissolved-As(III) and PO_4-As(III)] had
443 powerful influence on the development of viral *arsM* (Figure 6B), which was consistent
444 with the main influencing factors on the evolution of the active bacterial community

(Figure 2C). According to phylogenetic tree analysis, the main putative sources of viral *arsM* were Proteobacteria and Actinobacteria phyla (Figure 6C). Although several viral *arsM* with increased abundance (e.g., ASV153 and ASV10) were from Actinobacteria [whose abundance decreased in the stage-III (Figure S7)], the predominant viral *arsM* (e.g., ASV1 and ASV115) originated from Proteobacteria being the most abundant phylum in stage-III. An intriguing question was whether those viral *arsM* contributors also had higher abundance or not, thus the correlation between the relative abundance between-group variation of ASVs (with putative hosts) and the relative abundance between-group variation of corresponding hosts was analyzed. It could be shown in Figure 6D, they were significantly correlated ($r = 0.13$, $p = 0.03$), which implied that the viral *arsM* contributors also had higher activity. These observations demonstrated that *arsM*-bearing phages assist their hosts in adapting to a significant As(III) threat. To further confirm this hypothesis, correlation analyses were performed to reveal the connections between the relative abundance of *arsM* in prophages (per VLP) and the biomass (i.e., the copy number of 16S rRNA gene) and the diversity (i.e., ace index) of active bacterial communities. The copy number of 16S rRNA gene in soil microbiota was positively correlated to the copy number of *arsM* in prophages (Figure 6E, $r = 0.99$, $p < 0.01$) after day 1, which indicated *arsM*-bearing prophages favored host survival. Moreover, the α-diversity index of the active bacterial community was positively correlated to the copy number of *arsM* in prophage from day 5 to day 15, i.e., Shannon index for diversity (Figure S10), Chao index for richness (Figure S11) and Shannon

evenness index for evenness (Figure S12), which indicated *arsM*-bearing prophages favored the restoration of active bacterial community.

Enhanced arsenic methylation capacity of flooded soil microbiota

The *arsM* acquired by soil microbiota enabled them to methylate As(III) more rapidly. During the flooding period, the change of copy number of *arsM* in soil microbiota undergone two stages: firstly, it decreased (day 0 to day 1), and subsequently increased (day 2 to day 15). This profile was in accord with the change of *arsM* abundance in prophages (Figure 7A). It was important to note that such data was measured after the removal of free phages, thus the influence of *arsM* carried by free phages was excluded. Furthermore, the copy number of *arsM* in prophages increased by a net (1.2 ± 0.1)E+9 after 15 days, which accounted for 70.7 ± 1.2 % of the net increase in *arsM* in the soil microbial genome (Figure 7B). These results demonstrated the influential role of phage-mediated HGT in *arsM* acquisition by the soil microbiota. In addition, the contribution of transduction in *arsM* propagation may be underestimated given the negative bias introduced by prophage induction methods (some prophages cannot be induced by mitomycin-C) and enrichment procedure (resulting from exclusion of large viruses and the inability to see small RNA or single-stranded DNA viruses). The proliferation of chromosomal *arsM* enabled increased effective arsenic methylation. After 7 days of incubation, the concentration of captured TMAs(III) in the microcosm of that spiked flooded soil was 216.4 ± 17.0 μg/kg soil, while the corresponding concentration in the microcosm spiked pristine soil was 28.1

488 ± 3.5 μg/kg soil (Figure 7C). Meanwhile, the abiotic control group further demonstrated

489 the contribution of microbial populations to arsenic methylation (Figure 7C).

DISCUSSION

This work aimed to characterize the interaction between lysogenic phages and their hosts during elevating arsenic toxicity. This study was motivated by the observed phage-host collaboration in arsenic contaminated SM soil from our previous studies [29, 30] and the unique transformable nature of arsenic. Here, we were able to rule out the possible impact of lytic phages by eliminating free phages from SM soil beforehand, while we have to pay attention about the impacts of the remaining free viral particles on the soil fraction on the dataset. By analyzing the temporal development of the active bacterial community and the of lysogenic phage population, we found that the elevated As(III) toxicity drove the transduction of lysogenic phages containing plentiful beneficial genes (i.e., *arsM*) into their bacterial hosts, which then enabled survival of the soil microbiota under high levels of As(III). These results firstly provided the empirical evidence of lysogenic phages to the environmental adaptability of their hosts in a complex soil environment.

The proliferative ability of lysogenic phages under limited host density is based on their switchable life cycle [58]. Upon detection of host cell damage, prophages can excise themselves from the host chromosome to commence a potential lytic cycle [59]. In this scenario, the released formerly lysogenic turned lytic phage makes a last-ditch opportunity for propagation. Typically, As(III) has a strong affinity for protein sulfhydryl groups. Indeed, the redox status of cysteine residues can affect both the structure and the activity of numerous enzymes, receptors and transcription factors. Besides, As(III) toxicity has also been linked to the capacity of As(III) to oxidize

512 reduced glutathione, which is the major cellular antioxidant. This oxidation leads to an
513 increase in reactive oxygen species that have been shown to damage macromolecules
514 such as proteins, lipids and DNA [60, 61]. Such damage can arrest the cell cycle and
515 induce DNA repair and SOS response, further leading to the inactivation of a prophage
516 repressor. Therefore, prophages can be induced by As(III), thereby contributing to the
517 release of free phages from stage-I to stage-II. Intuitively, it may be deemed beneficial
518 for lysogenic phages to take refuge in a new host after being induced to be lytic because
519 the toxicity of As(III) hinders further reproduction by repressing potential hosts [the
520 dissolved As(III) reached 91.0 ± 3.1 mg/L on day 2]. The existing studies also
521 confirmed that lysogenic lifestyle of phages was more favored by in unfavorable
522 environments [15, 62]. In this work, some released phages continued to follow the
523 productive cycle (i.e., lytic cycle), not only because of the synchronous rise of both
524 prophages and free phages on day 2 compared to day 1, but also because of the increase
525 in the relative abundance of *arsM* (per VLP) in free phages. The functional gene carried
526 by viral contigs provided additional evidence of this critical proliferative capacity
527 (Figure S13). The viral contigs carrying polymerase-coding gene accounted for 9.2 %
528 of all viral contig considered as lysogenic phage in day 2_free but this was only 2.8%
529 in day 0_pro, which indicated that the induced free phages have stronger proliferative
530 capacity. On day 15, same proportion of viral contigs carrying polymerase-coding gene
531 in day 15_pro and day 0_pro supported that a new equilibrium between lysogenic
532 phages and their hosts was emerged, which consistent with the increase of the
533 proportion of prophages within total phages [recovered to (88.0 ± 0.2) % on day 15].

Such a “rampant” infection strategy enabled lysogenic phages to infect more potential hosts, that is, some lysogenic phages that follow the production life cycle acted as a devotee of protecting the interests of the lysogenic phage population as a whole. A similar devotee has been documented in bacterial populations, e.g., Snoussi *et al*. observed a fraction of *E. coli* cells rapidly absorbed and retained a large number of antimicrobial peptides upon the inhibition of their growth, which increased population survivability [63]. Figure 8A presented the inferred life strategy of lysogenic phages during flooding, and such strategy is predicted to ensure the overall success of lysogenic phages and their hosts through the aggressive expansion of the phage subpopulation. Noteworthy, the selection of lysogenic phages following productive cycles was random. Otherwise, on the one hand, there would be a dominating taxon among free phages that differed greatly from that of prophages; One the other hand, there should be a decrease in corresponding bacterial taxon due to continued predation.

Lysogens (infected by *arsM*-bearing phages) were favored by selection during the flooding period since viral *arsM* was more likely to originate from those bacteria containing a higher abundance of *arsM* in the flooding period. There are two pathways for the spread of the *arsM* in prophages, namely lysogenic phage re-infection and host replication. The host replication may be the main contributor to the observed increase in *arsM*-containing prophages in the early stage of flooding since the abundance of 16S rRNA gene was decreased within the first 2 days. In contrast, the lysogenic phage infection may have been important from day 2 to day 5 because the relative abundance of *arsM* in prophages (copies/VLP) was elevated. In comparison to day 0, the number

of prophages on day 15 increased by 5.9 times while the copy number of *arsM* in prophages increased by 55.3 times. Such non-synchronized development showed that *arsM*-bearing lysogenic phages possess a stronger proliferative potential, which was clearly correlated to the benefits they confer to the host. A recent work showed that arsenic resistant determinants encoded by a prophage harbored in *Citrobacter portucalensis* strain Sb-2, were upregulated under arsenic exposure [64]. This observation supported that viral *arsM*, which resided new hosts with lysogenic phages, was able to expressed to facilitate their hosts survive, which was also confirmed by the higher methylation rate of soil microbiota after flooding. The rapid adaption of soil microbiota to toxic levels of As(III) is surprising because it intuitively takes longer for the soil microbiota to adapt to such a significant As(III) toxicity.

Conjugation, transduction and transformation are the three main routes of HGT [65, 66], and there has been an increasing attention on phage-mediated transduction considering that phages are the most abundant replicating entities in the biosphere [15]. Our findings suggest that lysogenic phages make an important contribution to bacterial acquisition of arsenic resistance gene under As(III) exposure. The path analysis revealed a strong positive correlation between the copy number of *arsM* in soil microbiota and these in prophages (Figure 8B). The priority given to phage-mediated transduction in this work is not only based on transduction being less energy intensive than conjugation [67] but also because the resting aqueous environment is propitious to small-molecule communication of phage populations [68, 69]. Moreover, the overall correlation between As(III) formation and increments of *arsM* in soil microbiota was

explained by an indirect linkage rather than by a direct effect—that is, it was mediated through phage-mediated *arsM* transduction. In other words, the phage-mediated HGT of *arsM* enhanced the restoration of fitness of soil microbiota (Figure 8B). All in all, our findings highlighted the significance of lysogenic phages to the adaptability of their host in changing environments in community level. In extensive habitats, the prevalence of phage-host collaboration has consistently been shown [15, 29]. It appears coherent that a phage would be able to provide the genetic information needed to allow this rapid adaptation to a variety of environments because viruses exhibit extreme levels of diversity and are able to evolve rapidly to encode new functions [70]. In this study, *arsM*-bearing lysogenic phages can transduce a large number of *arsM* in a short period of time, so that soil microbiota obtained an enhanced As(III) methylation capability, which implies a potential opportunity for a reforming method of environmental microbial community based on phage-host collaboration. A better understanding regarding the evolution of phage-host collaboration will enable future attempts to modify microbial populations by forming beneficial endosymbionts via phage-mediated transduction of specific functional genes.

Lysogenic phages achieving steps in adaptive evolution of their host by variation in gene content [71, 72], but the environment, in turn, will affect the phage-host interaction. Accordingly, the phage-host collaboration in changing extreme environments merits particular attention. For example, permafrost environment imposes multiple stresses on its microbial inhabitants, including low temperature, water availability, and low thermal energy. Lysogenic phages have been shown to be deeply

involved in the transduction of functional genes in psychrophiles [73, 74]. Climate change might be driving the evolution of phage-host collaboration, especially considering a scenario where temperature has been implicated in prophage activation [75, 76]. The potential collapse or the development of phage-host collaboration has been shown to impact the emission of microbially generated greenhouse gases and thereby exacerbate climate change [77, 78]. Therefore, the influences of temperature on phage-host interactions ought to be assessed in similar environments. To date, although the *arsM*-bearing lysogenic phage and the metagenomic sequencing of different subsets of the phage population allowed us to monitor the temporal dynamics of lysogenic phages, the description of active lysogenic phages remains scarce (neither in this work). A more efficient coupled analysis method will be necessary in the future to give more specific information on phage-host collaboration. For instance, stable isotope probing (SIP) analysis relies on the incorporation of a substrate that is highly enriched in a stable isotope (e.g., ^{13}C), and the subsequent identification of active microbial populations by selective recovery and analysis of isotope-enriched cellular components [79]. Previously, SIP was used to target the phage-related genes in soil [80]. Combining SIP with viral metagenomics could potentially enable us to evaluate the active phage populations and help us to uncover this huge biological resource bank of phages. On this basis, further combined with single-cell Raman technique may yield important insights in the phage-host interactions [81, 82].

ACKNOWLEDGMENTS

We acknowledge financial support from the project of National Natural Science Foundation of China (NSFC): 52070075, and Natural Science Foundation of Hunan Province: 2020JJ4187. The authors want to dedicate the paper to the celebration of the forthcoming wedding of Dr. Xiang Tang and Miss Man Zhou.

DATA AVILABILITY

The active bacterial raw sequence data generated in this study are archived at the NCBI database under BioProject number: PRJNA823829. The raw sequence data of viral *arsM* genetic diversity in this study are archived at the NCBI database under BioProject number: PRJNA886312. The raw sequence data of viral metagenome in this study are archived at the NCBI database under BioProject number: PRJN PRJNA896864.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. Sun CL, Thomas BC, Barrangou R, Banfield JF. Metagenomic reconstructions of bacterial CRISPR loci constrain population histories. ISME J. 2016;10:858-70.

2. Weinbauer MG. Ecology of prokaryotic viruses. FEMS Microbiol Rev. 2004;28:127-81.

3. Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR. One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 2018;115:4987.

4. Nobrega FL, Costa AR, Kluskens LD, Azeredo J. Revisiting phage therapy: new applications for old resources. Trends Microbiol. 2015;23:185-91.

5. Mathieu J, Yu P, Zuo P, Da Silva MLB, Alvarez PJJ. Going viral: emerging opportunities for phage-based bacterial control in water treatment and reuse. Acc Chem Res. 2019;52:849-57.

6. Parvathi A, Jasna V, Aparna S, Pradeep Ram AS, Aswathy VK, Balachandran KK, et al. High incidence of lysogeny in the oxygen minimum zones of the arabian sea (southwest coast of India). Viruses. 2018;10:588.

7. Wommack KE, Colwell RR. Virioplankton: viruses in aquatic ecosystems. Microbiol Mol Biol Rev. 2000;64:69-114.

8. Scott MF, Otto SP. Why wait? Three mechanisms selecting for environment-dependent developmental delays. J Evol Biol. 2014;27:2219-32.

9. Breitbart M, Bonnain C, Malki K, Sawaya NA. Phage puppet masters of the marine microbial realm. Nat Microbiol. 2018;3:754-66.

10. Paul JH. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? ISME J. 2008;2:579-89.

11. Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science. 1996;272:1910-4.

12. Boyd EF, Davis BM, Hochhut B. Bacteriophage-bacteriophage interactions in the evolution of pathogenic bacteria. Trends Microbiol. 2001;9:137-44.

13. He T, Li H, Zhang X. Deep-sea hydrothermal vent viruses compensate for microbial metabolism in virus-host interactions. mBio. 2017;8:e00893-17.

14. Mann N, Cook A, Millard A, Bailey S, Clokie M. Bacterial photosynthesis genes in a virus. Nature. 2003;424(6950):741.

15. Zheng X, Jahn MT, Sun M, Friman V-P, Balcazar JL, Wang J, et al. Organochlorine contamination enriches virus-encoded metabolism and pesticide degradation associated auxiliary genes in soil microbiomes. ISME J. 2022;16:1397-1408.

16. Podgorski J, Berg M. Global threat of arsenic in groundwater. Science. 2020;368:845-50

673 17. Bissen M, Frimmel FH. Arsenic — a review. Part I: occurrence, toxicity, speciation,
674 mobility. Acta hydrochim hydrobiol. 2003;31:9-18

675 18. Sharma VK, Sohn M. Aquatic arsenic: toxicity, speciation, transformations, and
676 remediation. Environ Int. 2009;35:743-59

677 19. Tamaki S, Frankenberger WT. Environmental biochemistry of arsenic. In: Ware GW
678 (ed). Springer New York. 1992;1:79-110.

679 20. Qiao J-T, Li X-M, Hu M, Li F-B, Young LY, Sun W-M, et al. Transcriptional activity
680 of arsenic-reducing bacteria and genes regulated by lactate and biochar during
681 arsenic transformation in flooded paddy soil. Environ Sci Technol. 2018;52:61-70

682 21. Wang J, Halder D, Wegner L, Brüggenwirth L, Schaller J, Martin M, et al. Redox
683 dependence of thioarsenate occurrence in paddy soils and the rice rhizosphere.
684 Environ Sci Technol. 2020;54:3940-50

685 22. Páez-Espino D, Tamames J, de Lorenzo V, Cánovas D. Microbial responses to
686 environmental arsenic. BioMetals. 2009;22:117-30

687 23. Chen S-C, Sun G-X, Yan Y, Konstantinidis KT, Zhang S-Y, Deng Y, et al. The Great
688 Oxidation Event expanded the genetic repertoire of arsenic metabolism and
689 cycling. Proc Natl Acad Sci U S A. 2020;117:10414-21

690 24. Garbinski LD, Rosen BP, Chen J. Pathways of arsenic uptake and efflux. Environ
691 Int. 2019;126:585-97

692 25. Slyemi D, Bonnefoy V. How prokaryotes deal with arsenic†. Environ Microbiol
693 Rep. 2012;4:571-86

694 26. Jia Y, Huang H, Zhong M, Wang F-H, Zhang L-M, Zhu Y-G. Microbial arsenic
695 methylation in soil and rice rhizosphere. Environ Sci Technol. 2013;47:3141-8

696 27. Zhao F-J, Harris E, Yan J, Ma J, Wu L, Liu W, et al. Arsenic methylation in soils
697 and its relationship with microbial *arsM* abundance and diversity, and as
698 speciation in rice. Environ Sci Technol. 2013;47:7147-54

699 28. Dunivin TK, Yeh SY, Shade A. A global survey of arsenic-related genes in soil
700 microbiomes. BMC Biol. 2019;17:45

701 29. Tang X, Yu P, Tang L, Zhou M, Fan C, Lu Y, et al. Bacteriophages from arsenic-
702 resistant bacteria transduced resistance genes, which changed arsenic speciation
703 and increased soil toxicity. Environ. Sci. Technol. Lett. 2019;6:675-80

704 30. Tang X, Zhou M, Fan C, Zeng G, Lu Y, Dong H, et al. The arsenic chemical species
705 proportion and viral arsenic biotransformation genes composition affects
706 lysogenic phage treatment under arsenic stress. Sci Total Environ.
707 2021;780:146628

708 31. Adriaenssens EM, Kramer R, Van Goethem MW, Makhalanyane TP, Hogg I, Cowan
709 DA. Environmental drivers of viral community composition in Antarctic soils
710 identified by viromics. Microbiome. 2017;5:83

32. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun. 2016;13:11257

33. Liang X, Zhang Y, Wommack KE, Wilhelm SW, DeBruyn JM, Sherfy AC, et al. Lysogenic reproductive strategies of viral communities vary with soil depth and are correlated with bacterial diversity. Soil Biol Biochem. 2020;144:107767

34. Patel A, Noble RT, Steele JA, Schwalbach MS, Hewson I, Fuhrman JA. Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. Nat Protoc. 2007;2:269-76

35. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, et al. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods. 2016;102:3-11

36. Huang D, Yu P, Ye M, Schwarz C, Jiang X, Alvarez PJJ. Enhanced mutualistic symbiosis between soil phages and bacteria with elevated chromium-induced environmental stress. Microbiome. 2021;9:150.

37. Pons JC, Paez-Espino D, Riera G, Ivanova N, Kyrpides NC, Llabrés M. VPF-Class: taxonomic assignment and host prediction of uncultivated viruses based on viral protein families. Bioinformatics. 2021;37:1805-13.

38. Zhao JL, Jing HM, Wang ZM, Wang L, Jian HH, Zhang R, et al. Novel viral communities potentially assisting in carbon, nitrogen, and sulfur metabolism in the upper slope sediments of mariana trench. mSystems. 2022;7: e0135821.

39. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068-9.

40. Muscatt G, Hilton S, Raguideau S, Teakle G, Lidbury IDEA, Wellington EMH, et al. Crop management shapes the diversity and activity of DNA and RNA viruses in the rhizosphere. Microbiome. 2022;10:181.

41. Zhang R, Mirdita M, Levy Karin E, Norroy C, Galiez C, Söding J. SpacePHARER: sensitive identification of phages from CRISPR spacers in prokaryotic hosts. Bioinformatics. 2021;37:3364-6.

42. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581-3.

43. Huang H, Jia Y, Sun G-X, Zhu Y-G. Arsenic speciation and volatilization from flooded paddy soils amended with different organic matters. Environ Sci Technol. 2012;46:2163-8.

44. Bentley R, Chasteen Thomas G. Microbial methylation of metalloids: arsenic, antimony, and bismuth. Microbiol Mol Biol Rev. 2002;66:250-71.

45. Chen J, Qin J, Zhu Y-G, de Lorenzo Vc, Rosen Barry P. Engineering the soil bacterium *Pseudomonas putida* for arsenic methylation. Appl Environ Microbiol. 2013;79:4493-5.

46. Mestrot A, Uroic MK, Plantevin T, Islam MR, Krupp EM, Feldmann J, et al. Quantitative and qualitative trapping of arsines deployed to assess loss of volatile arsenic from paddy soil. Environ Sci Technol. 2009;43:8270-5.

47. Wang S, Xu L, Zhao Z, Wang S, Jia Y, Wang H, et al. Arsenic retention and remobilization in muddy sediments with high iron and sulfur contents from a heavily contaminated estuary in China. Chem Geol. 2012;314-317:57-65.

48. Masscheleyn PH, Delaune RD, Patrick WH. Effect of redox potential and pH on arsenic speciation and solubility in a contaminated soil. Environ Sci Technol. 1991;25:1414-9.

49. Mujawar SY, Vaigankar DC, Dubey SK. Biological characterization of *Bacillus flexus* strain SSAI1 transforming highly toxic arsenite to less toxic arsenate mediated by periplasmic arsenite oxidase enzyme encoded by *aioAB* genes. BioMetals. 2021;34:895-907.

50. Huang K, Chen C, Shen QR, Rosen BP, Zhao FJ. Genetically engineering *Bacillus subtilis* with a heat-resistant arsenite methyltransferase for bioremediation of arsenic-contaminated organic waste. Appl Environ Microbiol. 2015;81:6718-24.

51. Yang T, Chen ML, Liu LH, Wang JH, Dasgupta PK. Iron(III) modification of *Bacillus subtilis* membranes provides record sorption capacity for arsenic and endows unusual selectivity for As(V). Environ Sci Technol. 2012;46:2251-6.

52. Chen Z, Wang Y, Jiang X, Fu D, Xia D, Wang H, et al. Dual roles of AQDS as electron shuttles for microbes and dissolved organic matter involved in arsenic and iron mobilization in the arsenic-rich sediment. Sci Total Environ. 2017;574:1684-94.

53. Kudo K, Yamaguchi N, Makino T, Ohtsuka T, Kimura K, Dong Dian T, et al. Release of arsenic from soil by a novel dissimilatory arsenate-reducing bacterium, *Anaeromyxobacter sp.* strain PSR-1. Appl Environ Microbiol. 2013;79:4635-42.

54. Dragoš A, Priyadarshini B, Hasan Z, Strube ML, Kempen PJ, Maróti G, et al. Pervasive prophage recombination occurs during evolution of spore-forming *Bacilli*. ISME J. 2021;15:1344-58.

55. Kohm K, Floccari VA, Lutz VT, Nordmann B, Mittelstädt C, Poehlein A, et al. The *Bacillus* phage SPβ and its relatives: a temperate phage model system reveals new strains, species, prophage integration loci, conserved proteins and lysogeny management components. Environ Microbiol. 2022;24:2098-118.

56. Abe K, Kawano Y, Iwamoto K, Arai K, Maruyama Y, Eichenberger P, et al. Developmentally-regulated excision of the SPβ prophage reconstitutes a gene required for spore envelope maturation in *Bacillus subtilis*. PLoS Genet. 2014;10:e1004636.

57. Gómez P, Buckling A. Bacteria-phage antagonistic coevolution in soil. Science. 2011;332:106-9.

58. Tang X, Fan C, Zeng G, Zhong L, Li C, Ren X, et al. Phage-host interactions: The

neglected part of biological wastewater treatment. Water Res. 2022;226:119183.

59. Tzipilevich E, Habusha M, Ben-Yehuda S. Acquisition of phage sensitivity by bacteria through exchange of phage receptors. Cell. 2017;168:186-99.

60. Liu SX, Athar M, Lippai I, Waldren C, Hei TK. Induction of oxyradicals by arsenic: implication for mechanism of genotoxicity. Proc Natl Acad Sci U S A. 2001;98:1643-8

61. Dopp E, von Recklinghausen U, Diaz-Bone R, Hirner AV, Rettenmeier AW. Cellular uptake, subcellular distribution and toxicity of arsenic compounds in methylating and non-methylating cells. Environ Res. 2010;110:435-42

62. Gao S-M, Schippers A, Chen N, Yuan Y, Zhang M-M, Li Q, et al. Depth-related variability in viral communities in highly stratified sulfidic mine tailings. Microbiome. 2020;8:89

63. Snoussi M, Talledo JP, Del Rosario N-A, Mohammadi S, Ha B-Y, Košmrlj A, et al. Heterogeneous absorption of antimicrobial peptide LL37 in *Escherichia coli* cells enhances population survivability. Elife. 2018;7:e38174

64. Yu Y, Xie Z, Yang J, Yang R, Li Y, Zhu Y, et al. *Citrobacter portucalensis* Sb-2 contains a metalloid resistance determinant transmitted by *Citrobacter* phage Chris1. J Hazard Mater. 2023;443:130184

65. Cyriaque V, Madsen JS, Fievez L, Leroy B, Hansen LH, Bureau F, et al. Lead drives complex dynamics of a conjugative plasmid in a bacterial community. Front Microbiol. 2021;12:655903

66. Johnston C, Martin B, Fichant G, Polard P, Claverys J-P. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol. 2014;12:181-96

67. Sun R, Yu P, Zuo P, Alvarez PJJ. Bacterial concentrations and water turbulence influence the importance of conjugation versus phage-mediated antibiotic resistance gene transfer in suspended growth systems. ACS Environ Au. 2021;2:152-65

68. Aframian N, Omer Bendori S, Kabel S, Guler P, Stokar-Avihail A, Manor E, et al. Dormant phages communicate via arbitrium to control exit from lysogeny. Nat Microbiol. 2022;7:145-53

69. Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, et al. Communication between viruses guides lysis–lysogeny decisions. Nature. 2017;541:488-93

70. Roossinck MJ. The good viruses: viral mutualistic symbioses. Nat Rev Microbiol. 2011;9:99-108

71. Reno ML, Held NL, Fields CJ, Burke PV, Whitaker RJ. Biogeography of the *Sulfolobus islandicus* pan-genome. Proc Natl Acad Sci U S A. 2009;106:8605-10

72. Aertsen A, Michiels CW. Diversify or die: generation of diversity in response to

stress. Crit Rev Microbiol. 2005;31:69-78

73. Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB, et al. Host-linked soil viral ecology along a permafrost thaw gradient. Nat Microbiol. 2018;3:870-80

74. Zhong Z-P, Rapp JZ, Wainaina JM, Solonenko NE, Maughan H, Carpenter SD, et al. Viral ecogenomics of arctic cryopeg brine and sea ice. mSystems. 2020;5:e00246-20

75. Brum JR, Hurwitz BL, Schofield O, Ducklow HW, Sullivan MB. Seasonal time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics. ISME J. 2016;10:437-49

76. Lunde M, Aastveit AH, Blatny JM, Nes IF. Effects of Diverse Environmental conditions on φLC3 prophage stability in *Lactococcus lactis*. Appl Environ Microbiol. 2005;71:721-27

77. Schuur EAG, McGuire AD, Schädel C, Grosse G, Harden JW, Hayes DJ, et al. Climate change and the permafrost carbon feedback. Nature. 2015;520:171-79

78. Ji M, Fan X, Cornell CR, Zhang Y, Yuan MM, Tian Z, et al. Tundra soil viruses mediate responses of microbial communities to climate warming. mBio. 2023;0:e03009-22

79. Dam HT, Sun W, McGuinness L, Kerkhof LJ, Häggblom MM. Identification of a chlorodibenzo-*p*-dioxin dechlorinating *Dehalococcoides mccartyi* by stable isotope probing. Environ Sci Technol. 2019;53:14409-19

80. Li Y, Watanabe T, Murase J, Asakawa S, Kimura M. Identification of the major capsid gene (g23) of T4-type bacteriophages that assimilate substrates from root cap cells under aerobic and anaerobic soil conditions using a DNA–SIP approach. Soil Biol Biochem. 2013;63:97-105

81. Li H-Z, Yang K, Liao H, Lassen SB, Su J-Q, Zhang X, et al. Active antibiotic resistome in soils unraveled by single-cell isotope probing and targeted metagenomics. Proc Natl Acad Sci U S A. 2022;119:e2201473119

82. Yang K, Xu F, Zhu L, Li H, Sun Q, Yan A, et al. An isotope-labeled single-cell raman spectroscopy approach for tracking the physiological evolution trajectory of bacteria toward antibiotic resistance. Angew Chem Int Ed Engl. 2023;62:e20221

Figure legends

Figure 1. The workflow of separate-extraction of different phage subsets in the microcosm, and we were able to obtain the free phages derived from lysogenic phage and the prophages remaining in the host via this workflow.

Figure 2. (A) The composition of the active bacterial community at the genus level, the abundance is presented as the average percentage of three replicates; (B) Network analysis of the active bacterial community based on the genus level; (C) The arsenic species affecting differences in the composition of the active microbial community as revealed by redundancy analysis (RDA); (D) Dynamics of the copy number of the 16S rRNA gene in soil microbiota during 15-day flooding period. Error bars represent standard deviations of triplicate tests.

Figure 3. (A) The proportion of phages to all assembled contigs identified as viruses, and this result indicated that VLP can be used to characterize the number of phage; (B) The composition of the lysogenic phages at the genus level from five samples in three sampling times, the viral contigs containing transposase, integrase, excisionase, resolvase or recombinase were considered as lysogenic phages; (C) Dynamics of the numbers of prophages and free phages in microcosm during the 15-day flooding period, where the numerical digit up the column indicates the percentage of prophages in the total phages (i.e., prophages and free phages); (D) Predicted virus-host linkages in the

flooding period, where worth noting that the main viral contigs that cannot be annotated is not displayed. Error bars represent standard deviations of triplicate tests.

Figure 4. (A) The shared viral contigs in the top 20 contigs between the prophages on day 0 (day 0_pro) and the free phages on day 2 (day 2_free); (B) The differences in the composition of the viral community revealed by principal co-ordinates analysis (PCoA) based on Bray–Curtis distances; The correlation between the number of (C) prophages and (d) free phages and the dissolved-As(III) concentration (N = 18).

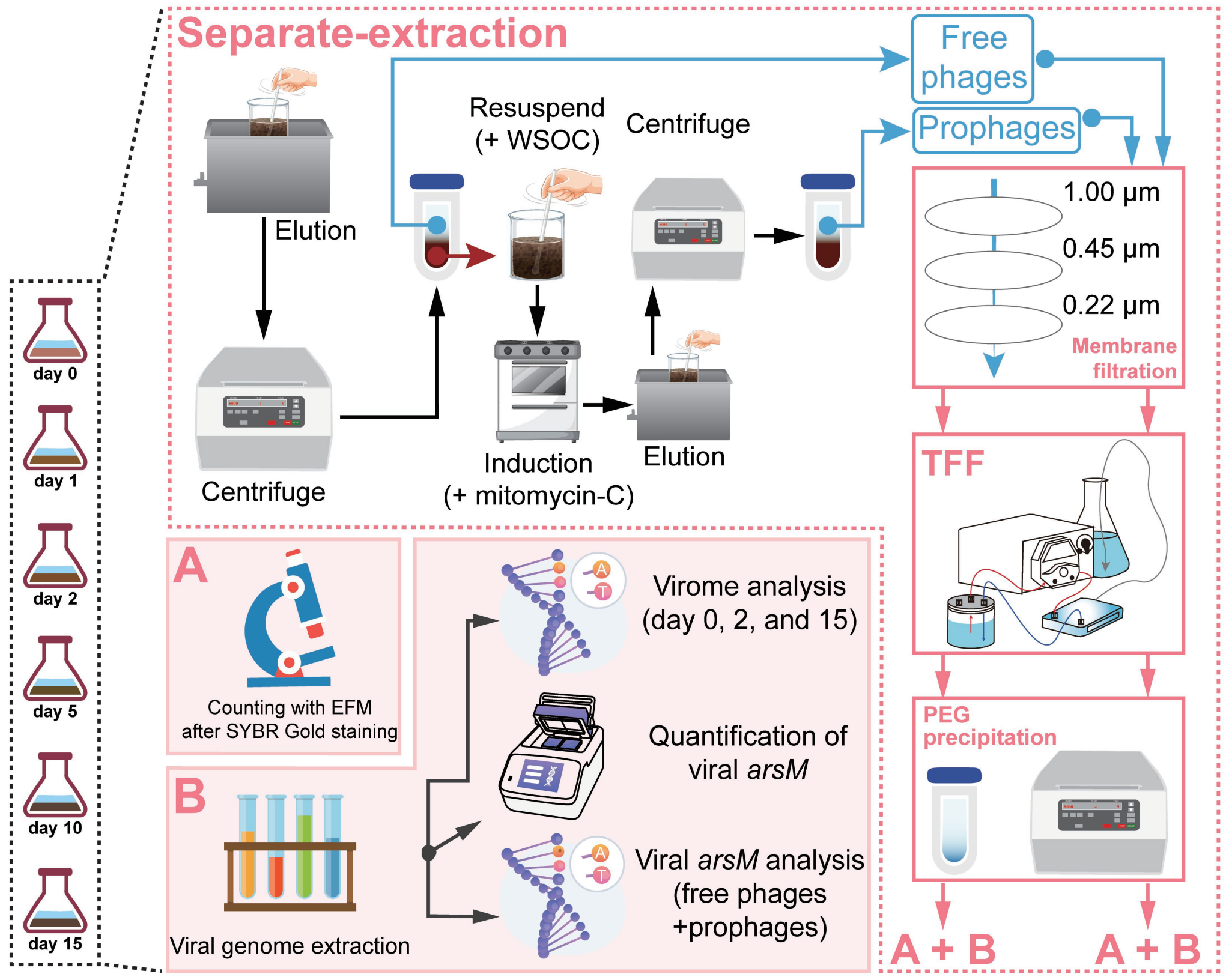
Figure 5. (A) Absolute (gene copies) and (B) relative abundance (per VLP) of *arsM* in prophages; (C) The correlation between the relative abundance of arsM in prophages and the relative abundance of *arsM* in free phages (N = 18).

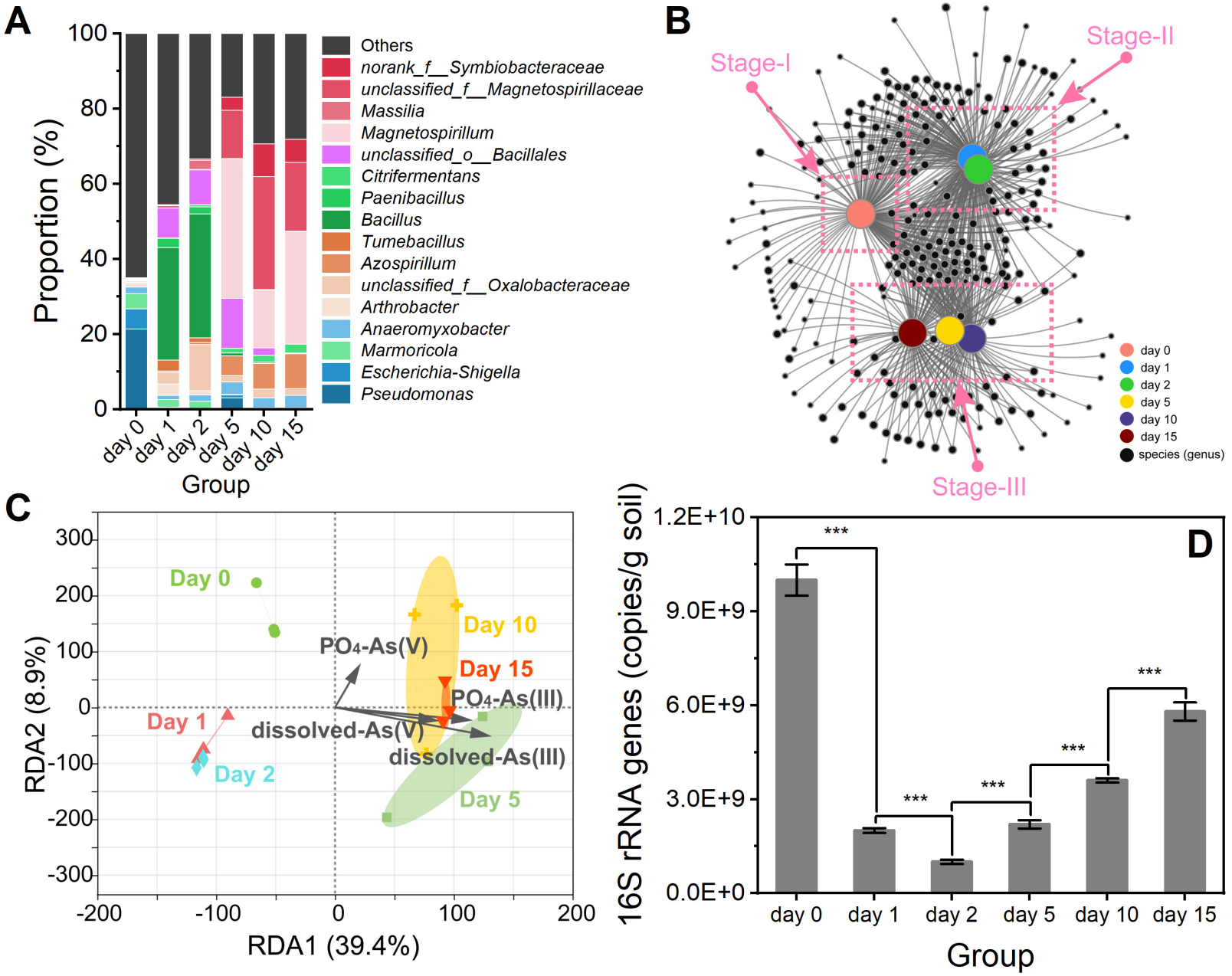
Figure 6. (A) The composition of viral *arsM* taxa at the ASV level, the abundance is presented as the average percentage of three replicates; (B) The main environmental factors affecting differences in the composition of viral *arsM* taxa revealed by canonical correlation analysis; (C) The viral *arsM* phylogenetic tree based on *arsM* amplicon sequencing where differently colored spots denote different changing trends; (D) The correlation between the between-group variations of top 10 ASV abundance and the between-group variations of their putative host, where the between-group variation is the difference between one sample and another sample at the previous point-in-time [e.g., $V_{0\text{-}1}$ (between-group variation between day 1 and day 0)= C_1 (relative abundance

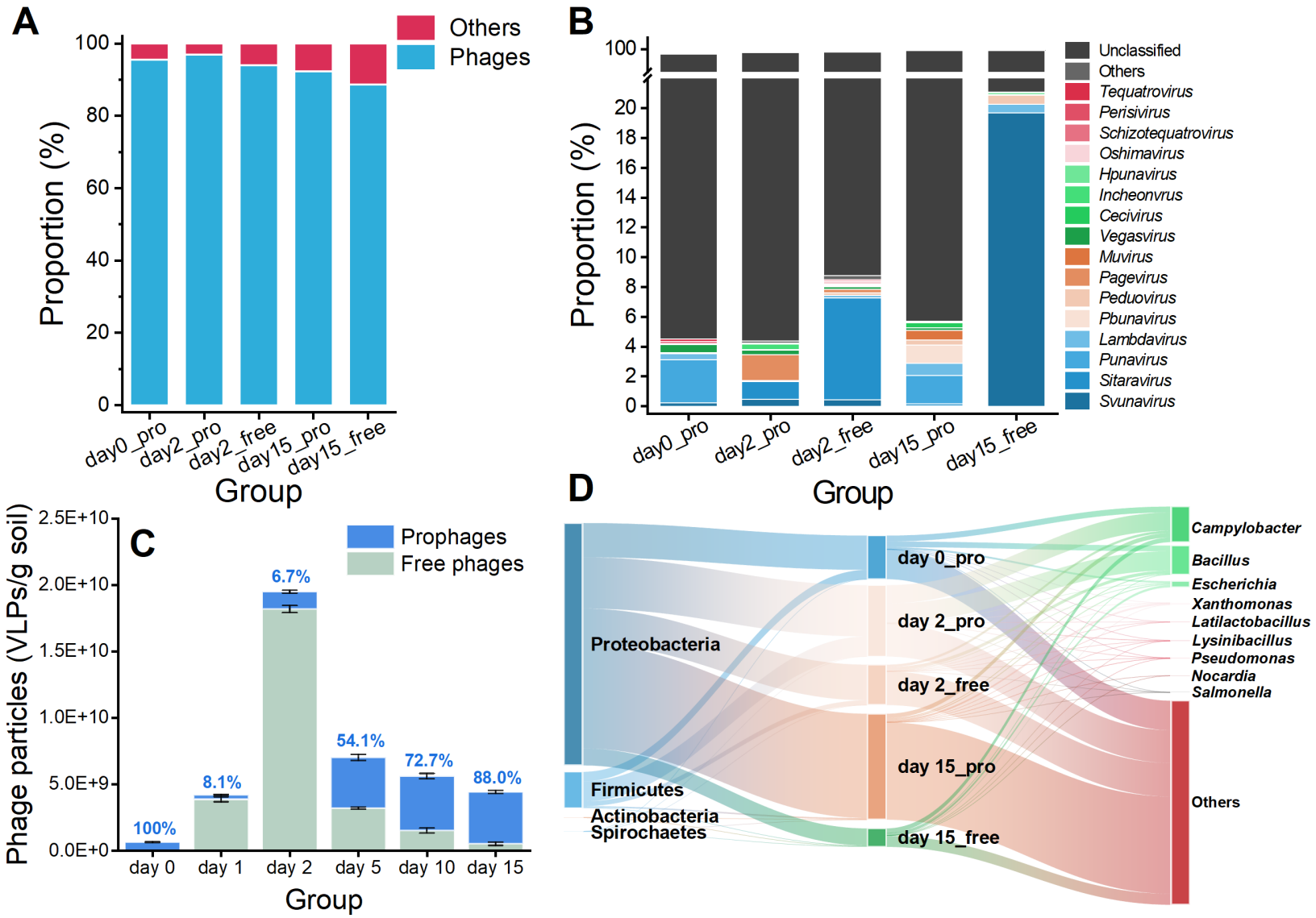
of ASV or their putative host on day 1) – C_0 (relative abundance on day 1)]; (E) The correlation between the copy number of 16S rRNA gene and the copy number of *arsM* in prophages.

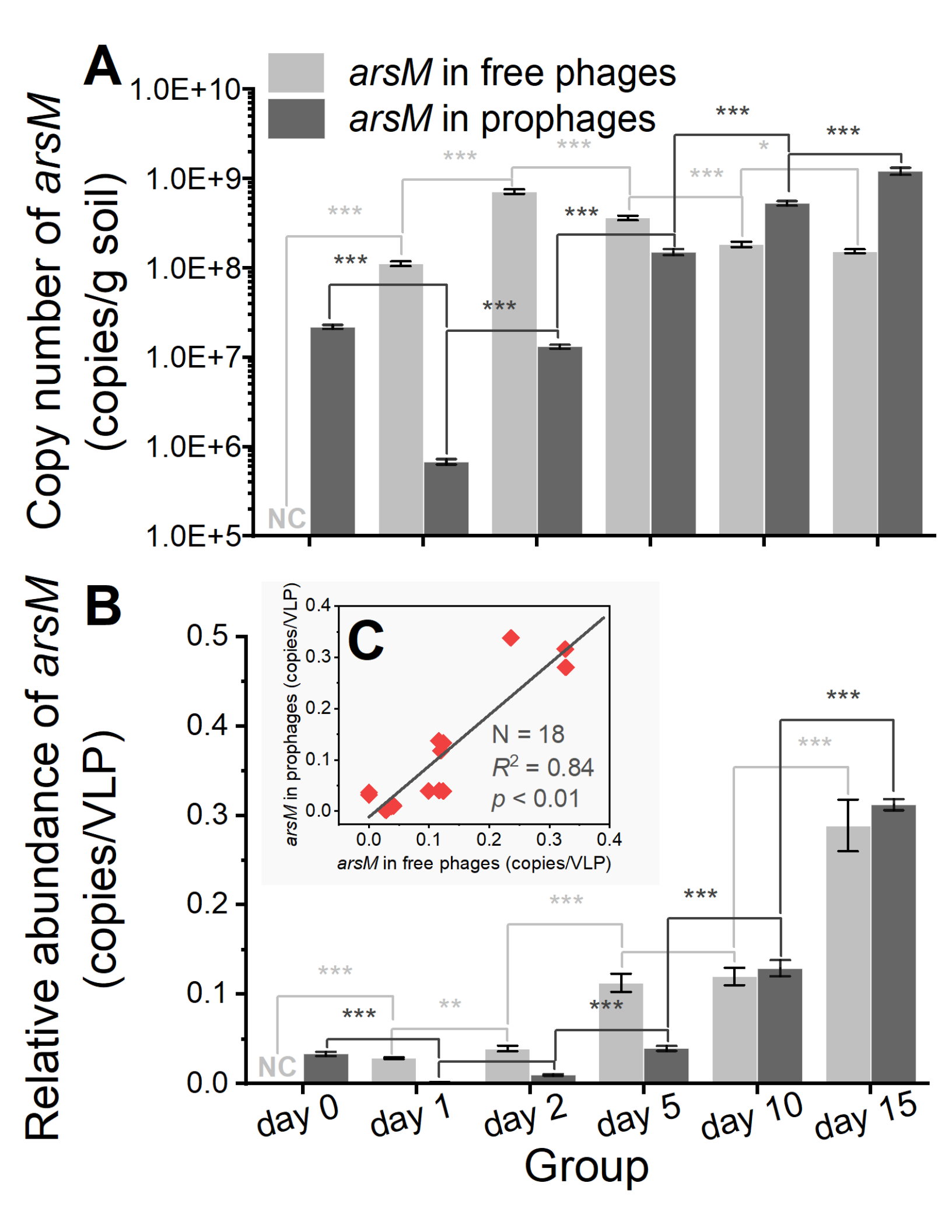
Figure 7. (A) Dynamics of the copy numbers of *arsM* in soil microbiota; (B) The contribution of *arsM*-bearing prophages to the increment of *arsM* in soil microbiota; (C) The yield of TMAs(III) from different soil after 15-day incubation. Error bars represent standard deviations of triplicate tests.

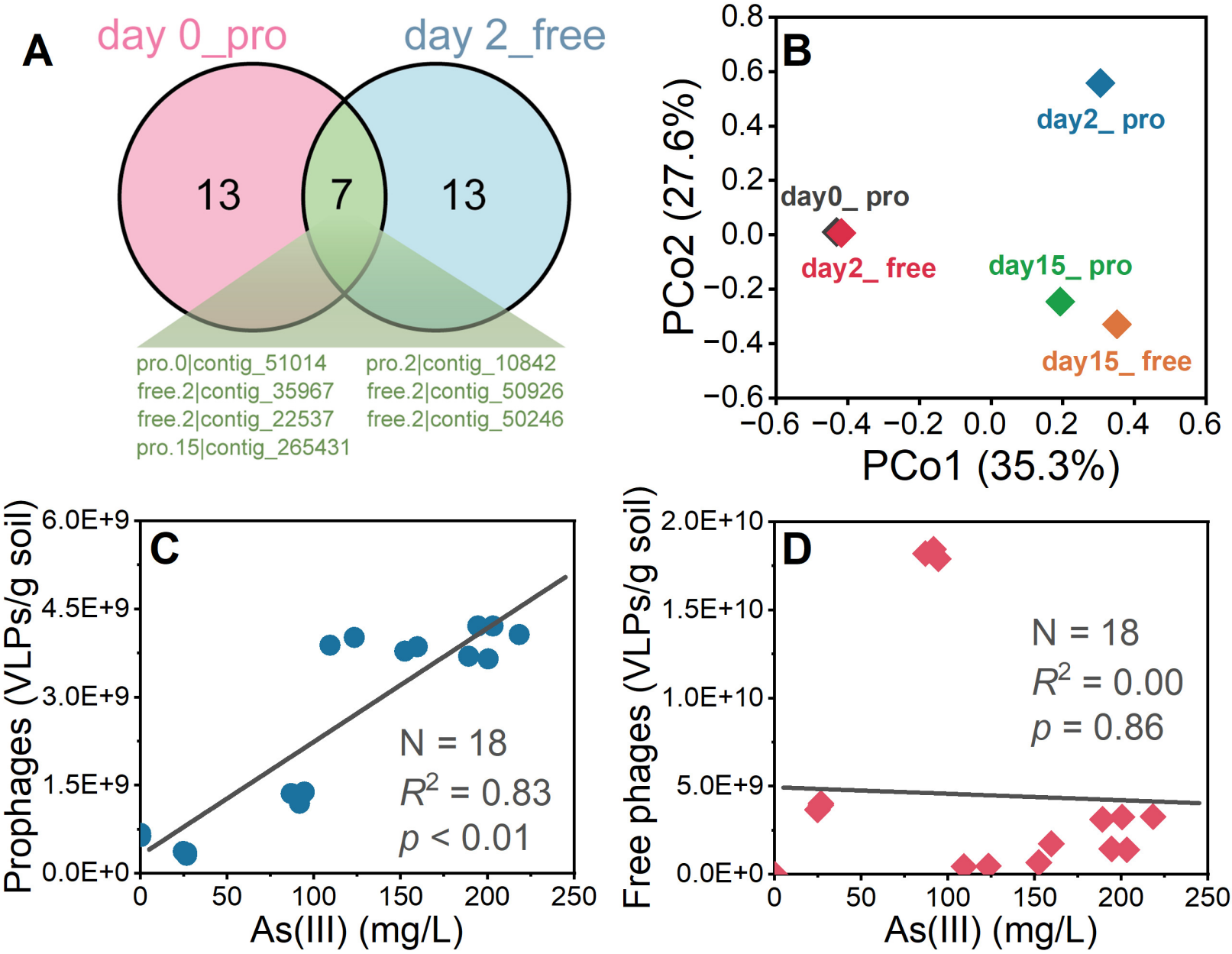
Figure 8. (A) The presumed survival strategy of lysogenic phages in the microcosm during flooding, in which some of As(III)-induced lysogenic phages maintained the productive life cycle rather than enter a new lysogenic life cycle, and this strategy facilitated the spread of *arsM*. Error bars represent standard deviations of triplicate tests. (B) Path analysis showing the direct and indirect effects of the As(III) toxicity on the abundance of *arsM* in soil microbiota. Indirect effects of As(III) toxicity are mediated through *arsM*-bearing prophages. Numbers above paths represent standardized coefficients in flooding period. Thickness and color of lines correspond to coefficient magnitude and direction, respectively.

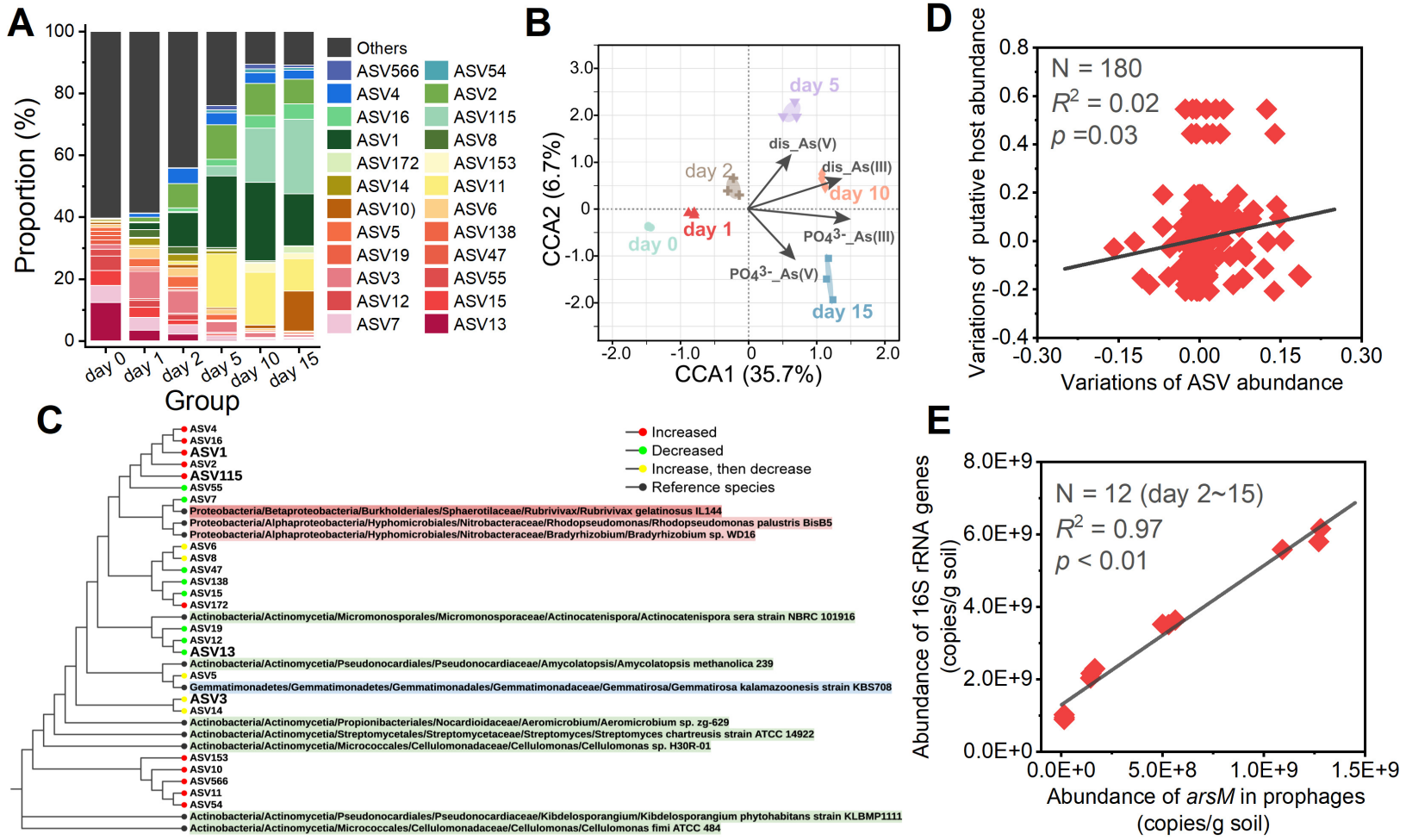

Separate-extraction
Free phages
Prophages
Resuspend (+ WSOC)
Centrifuge
Elution
Centrifuge
Induction (+ mitomycin-C)
Elution
1.00 μm
0.45 μm
0.22 μm
Membrane filtration
TFF
PEG precipitation
A + B
A + B
day 0
day 1
day 2
day 5
day 10
day 15
A
Counting with EFM after SYBR Gold staining
B
Viral genome extraction
Virome analysis (day 0, 2, and 15)
Quantification of viral arsM
Viral arsM analysis (free phages +prophages)

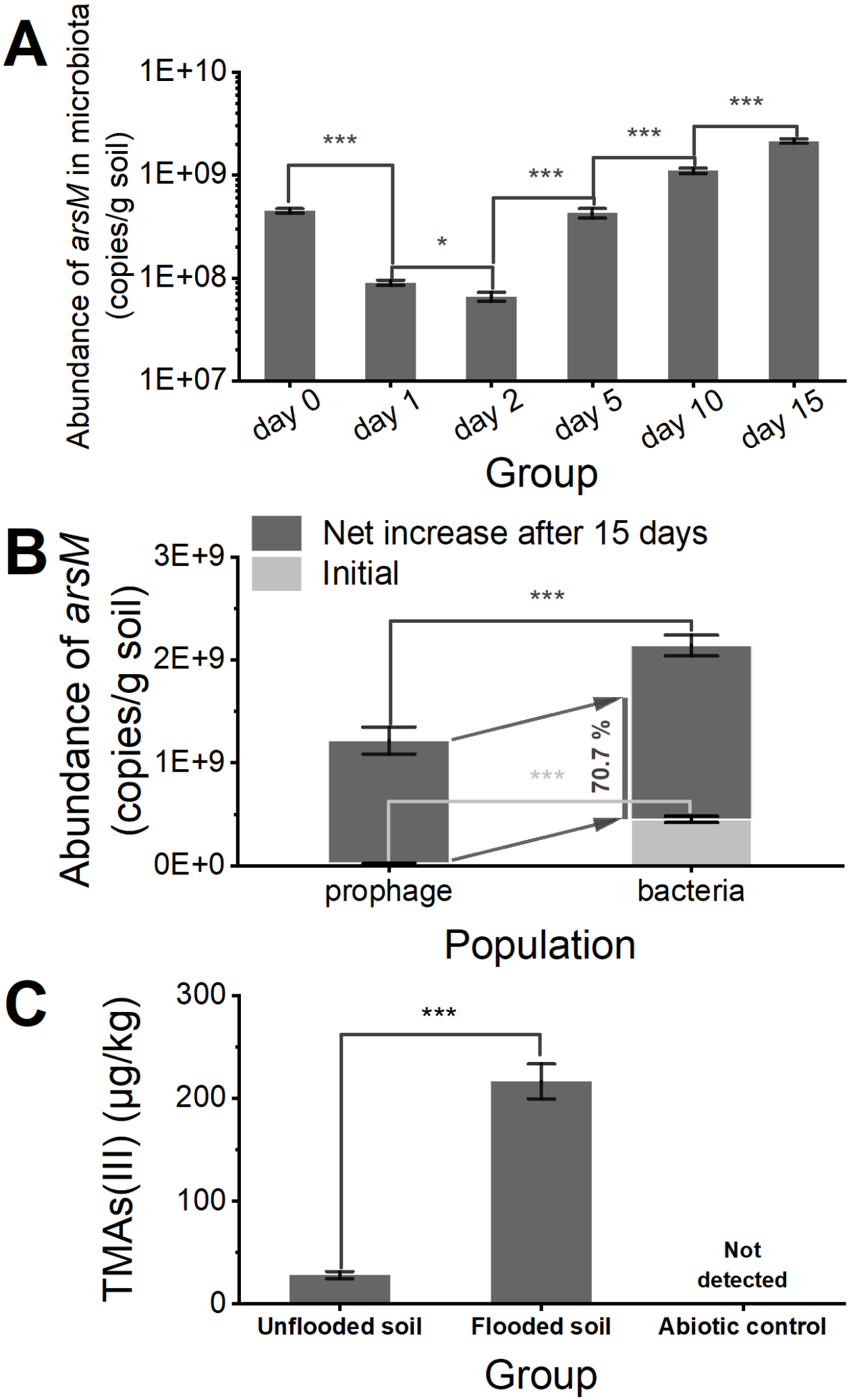

A
Proportion (%)
Group
day 0
day 1
day 2
day 5
day 10
day 15
Others
norank_f__Symbiobacteraceae
unclassified_f__Magnetospirillaceae
Massilia
Magnetospirillum
unclassified_o__Bacillales
Citrifermentans
Paenibacillus
Bacillus
Tumebacillus
Azospirillum
unclassified_f__Oxalobacteraceae
Arthrobacter
Anaeromyxobacter
Marmoricola
Escherichia-Shigella
Pseudomonas
B
Stage-I
Stage-II
Stage-III
day 0
day 1
day 2
day 5
day 10
day 15
species (genus)
C
RDA1 (39.4%)
RDA2 (8.9%)
Day 0
Day 1
Day 2
Day 5
Day 10
Day 15
PO4-As(V)
PO4-As(III)
dissolved-As(V)
dissolved-As(III)
D
16S rRNA genes (copies/g soil)
Group
***

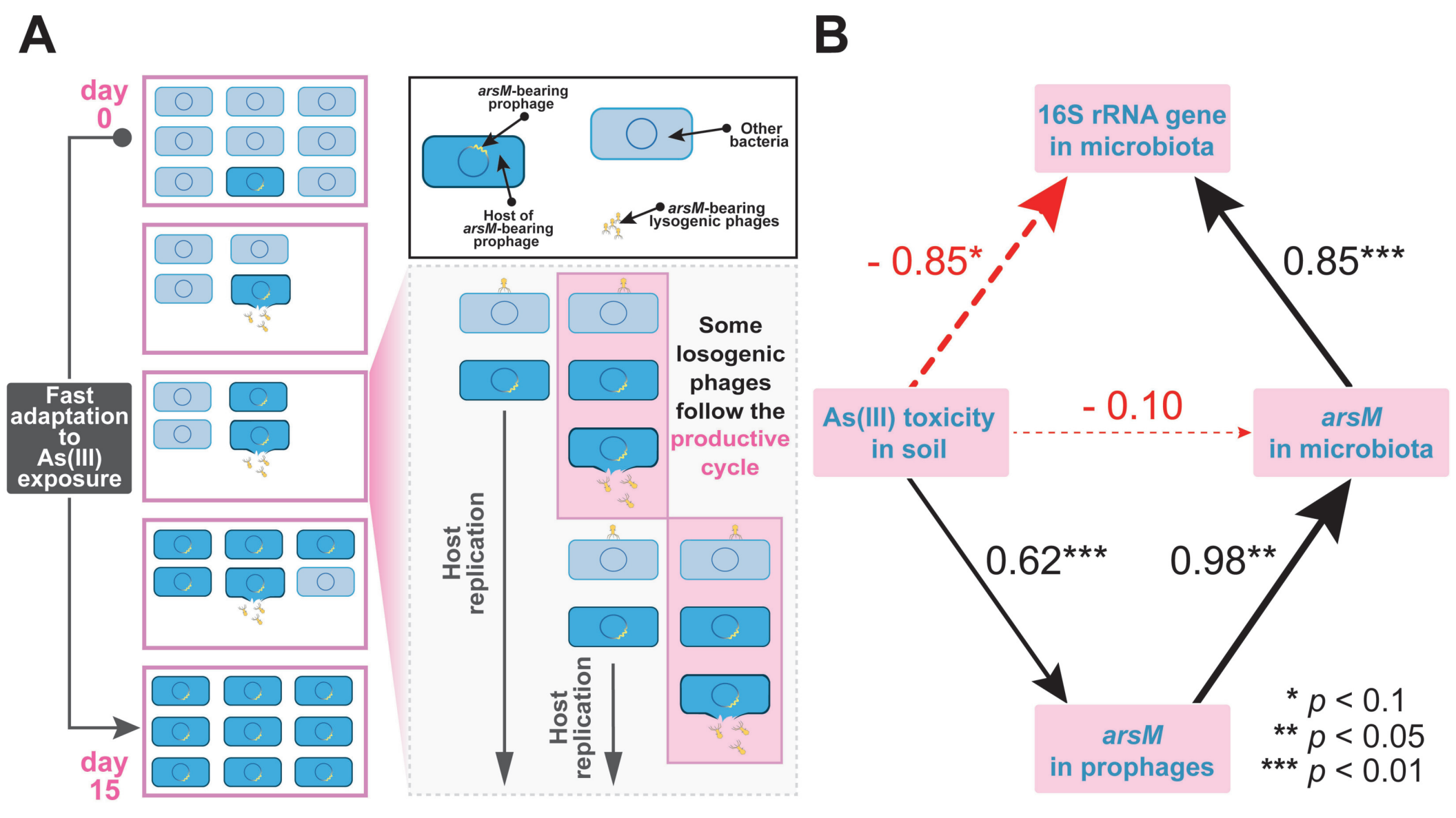

A
Others
Phages
Proportion (%)
100
80
60
40
20
0
day0_pro
day2_pro
day2_free
day15_pro
day15_free
Group
B
Unclassified
Others
Tequatrovirus
Perisivirus
Schizotequatrovirus
Oshimavirus
Hpunavirus
Incheonvrus
Cecivirus
Vegasvirus
Muvirus
Pagevirus
Peduovirus
Pbunavirus
Lambdavirus
Punavirus
Sitaravirus
Svunavirus
Proportion (%)
100
20
18
16
14
12
10
8
6
4
2
0
day0_pro
day2_pro
day2_free
day15_pro
day15_free
Group
C
Prophages
Free phages
Phage particles (VLPs/g soil)
2.5E+10
2.0E+10
1.5E+10
1.0E+10
5.0E+9
0.0E+0
100%
8.1%
6.7%
54.1%
72.7%
88.0%
day 0
day 1
day 2
day 5
day 10
day 15
Group
D
Proteobacteria
Firmicutes
Actinobacteria
Spirochaetes
day 0_pro
day 2_pro
day 2_free
day 15_pro
day 15_free
Campylobacter
Bacillus
Escherichia
Xanthomonas
Latilactobacillus
Lysinibacillus
Pseudomonas
Nocardia
Salmonella
Others


A
arsM in free phages
arsM in prophages
Copy number of arsM (copies/g soil)
1.0E+10
1.0E+9
1.0E+8
1.0E+7
1.0E+6
1.0E+5
NC
***
*
B
Relative abundance of arsM (copies/VLP)
0.5
0.4
0.3
0.2
0.1
0.0
**
C
arsM in prophages (copies/VLP)
arsM in free phages (copies/VLP)
N = 18
R2 = 0.84
p < 0.01
day 0
day 1
day 2
day 5
day 10
day 15
Group


A
day 0_pro
day 2_free
13
7
13
pro.0|contig_51014
free.2|contig_35967
free.2|contig_22537
pro.15|contig_265431
pro.2|contig_10842
free.2|contig_50926
free.2|contig_50246
B
day2_ pro
day0_ pro
day2_ free
day15_ pro
day15_ free
PCo2 (27.6%)
PCo1 (35.3%)
C
Prophages (VLPs/g soil)
As(III) (mg/L)
N = 18
R^2 = 0.83
p < 0.01
D
Free phages (VLPs/g soil)
As(III) (mg/L)
N = 18
R^2 = 0.00
p = 0.86


A
Proportion (%)
Group
day 0
day 1
day 2
day 5
day 10
day 15
Others
ASV566
ASV4
ASV16
ASV1
ASV172
ASV14
ASV10)
ASV5
ASV19
ASV3
ASV12
ASV7
ASV54
ASV2
ASV115
ASV8
ASV153
ASV11
ASV6
ASV138
ASV47
ASV55
ASV15
ASV13
B
CCA2 (6.7%)
CCA1 (35.7%)
day 5
dis_As(V)
dis_As(III)
day 2
day 10
day 1
day 0
PO43-_As(III)
PO43-_As(V)
day 15
C
Increased
Decreased
Increase, then decrease
Reference species
ASV4
ASV16
ASV1
ASV2
ASV115
ASV55
ASV7
Proteobacteria/Betaproteobacteria/Burkholderiales/Sphaerotilaceae/Rubrivivax/Rubrivivax gelatinosus IL144
Proteobacteria/Alphaproteobacteria/Hyphomicrobiales/Nitrobacteraceae/Rhodopseudomonas/Rhodopseudomonas palustris BisB5
Proteobacteria/Alphaproteobacteria/Hyphomicrobiales/Nitrobacteraceae/Bradyrhizobium/Bradyrhizobium sp. WD16
ASV6
ASV8
ASV47
ASV138
ASV15
ASV172
Actinobacteria/Actinomycetia/Micromonosporales/Micromonosporaceae/Actinocatenispora/Actinocatenispora sera strain NBRC 101916
ASV19
ASV12
ASV13
Actinobacteria/Actinomycetia/Pseudonocardiales/Pseudonocardiaceae/Amycolatopsis/Amycolatopsis methanolica 239
ASV5
Gemmatimonadetes/Gemmatimonadetes/Gemmatimonadales/Gemmatimonadaceae/Gemmatirosa/Gemmatirosa kalamazoonesis strain KBS708
ASV3
ASV14
Actinobacteria/Actinomycetia/Propionibacteriales/Nocardioidaceae/Aeromicrobium/Aeromicrobium sp. zg-629
Actinobacteria/Actinomycetia/Streptomycetales/Streptomycetaceae/Streptomyces/Streptomyces chartreusis strain ATCC 14922
Actinobacteria/Actinomycetia/Micrococcales/Cellulomonadaceae/Cellulomonas/Cellulomonas sp. H30R-01
ASV153
ASV10
ASV566
ASV11
ASV54
Actinobacteria/Actinomycetia/Pseudonocardiales/Pseudonocardiaceae/Kibdelosporangium/Kibdelosporangium phytohabitans strain KLBMP1111
Actinobacteria/Actinomycetia/Micrococcales/Cellulomonadaceae/Cellulomonas/Cellulomonas fimi ATCC 484
D
N = 180
R^2 = 0.02
p =0.03
Variations of putative host abundance
Variations of ASV abundance
E
N = 12 (day 2~15)
R^2 = 0.97
p < 0.01
Abundance of 16S rRNA genes (copies/g soil)
Abundance of *arsM* in prophages (copies/g soil)


A
Abundance of arsM in microbiota (copies/g soil)
1E+10
1E+09
1E+08
1E+07
***
*
***
***
***
day 0
day 1
day 2
day 5
day 10
day 15
Group
B
Abundance of arsM (copies/g soil)
Net increase after 15 days
Initial
3E+9
2E+9
1E+9
0E+0
***
***
70.7 %
prophage
bacteria
Population
C
TMAs(III) (μg/kg)
300
200
100
0
***
Not detected
Unflooded soil
Flooded soil
Abiotic control
Group


A
day 0
Fast adaptation to As(III) exposure
day 15
arsM-bearing prophage
Other bacteria
Host of arsM-bearing prophage
arsM-bearing lysogenic phages
Some losogenic phages follow the productive cycle
Host replication
Host replication
B
16S rRNA gene in microbiota
- 0.85*
0.85***
As(III) toxicity in soil
- 0.10
arsM in microbiota
0.62***
0.98**
arsM in prophages
* p < 0.1
** p < 0.05
*** p < 0.01