

ARTICLE OPEN



Benzo[a]pyrene stress impacts adaptive strategies and ecological functions of earthworm intestinal viromes

Rong Xia^{1,6}, Mingming Sun^{1,6}, José Luis Balcázar^{2,3}, Pingfeng Yu⁴, Feng Hu¹ and Pedro J. J. Alvarez⁵

© The Author(s) 2023

The earthworm gut virome influences the structure and function of the gut microbiome, which in turn influences worm health and ecological functions. However, despite its ecological and soil quality implications, it remains elusive how earthworm intestinal phages respond to different environmental stress, such as soil pollution. Here we used metagenomics and metatranscriptomics to investigate interactions between the worm intestinal phages and their bacteria under different benzo[a]pyrene (BaP) concentrations. Low-level BaP (0.1 mg kg^{-1}) stress stimulated microbial metabolism (1.74-fold to control), and enhanced the antiphage defense system ($n = 75$) against infection (8 phage-host pairs). Low-level BaP exposure resulted in the highest proportion of lysogenic phages (88%), and prophages expressed auxiliary metabolic genes (AMGs) associated with nutrient transformation (e.g., amino acid metabolism). In contrast, high-level BaP exposure (200 mg kg^{-1}) disrupted microbial metabolism and suppressed the antiphage systems ($n = 29$), leading to the increase in phage-bacterium association (37 phage-host pairs) and conversion of lysogenic to lytic phages (lysogenic ratio declined to 43%). Despite fluctuating phage-bacterium interactions, phage-encoded AMGs related to microbial antioxidant and pollutant degradation were enriched, apparently to alleviate pollution stress. Overall, these findings expand our knowledge of complex phage-bacterium interactions in pollution-stressed worm guts, and deepen our understanding of the ecological and evolutionary roles of phages.

The ISME Journal (2023) 17:1004–1014; <https://doi.org/10.1038/s41396-023-01408-x>

INTRODUCTION

Earthworms are vital to soil ecosystems due to their widespread distribution and major contributions to the maintenance of the physical structure, drainage, nutrient availability, and microbial activity in soils [1]. Moreover, earthworms may improve bioremediation of contaminated soils by enhancing soil aeration through burrowing, and through enzymatic biotransformation of various contaminants, such as organic pollutants [2] and microplastics [3]. Recent advances in metagenomics analysis have shed light on the importance of the earthworm intestinal microbiome in the worm's growth, behavior, and ecological roles [4–6]. The earthworm intestine is an anoxic microenvironment with abundant organic carbon and nitrogen, and is prone to accumulate soil contaminants [7, 8]. Therefore, to harness the full ecological and biotechnological benefits of earthworms, there is a critical need to advance our understanding of factors that shape their intestinal microbiome structure and functional diversity.

The virome contains the most abundant and genetically diverse biological entities in the biosphere [9]. There is an increasing recognition that the virome can substantially affect intestinal microbiome structure and functions through various phage-bacterium interactions [10–13]. Previous studies suggest that intestinal viromes are important reservoirs of bacterial metabolic genes and may modulate important microbial processes such as

metabolic activities, quorum sensing, and chemotaxis pathways [14]. Other studies suggest that environmental stress such as drought [15] and pollutants [16] could affect the virome profile and modify the phage-bacterium symbiotic or antagonistic relationships. However, it remains largely unexplored how earthworm intestinal virome and phage-bacterium relationship responds to different environmental stress conditions. This represents a critical knowledge gap in exploring the roles of viromes in intestinal microbial homeostasis and in the adaptive mechanisms of earthworms facing various environmental stresses.

We used a multi-omics approach, including metagenomics, viromics, metatranscriptomics, and bioinformatics, to investigate the earthworm intestinal virome and its interaction with bacteriome under environmental pollution stress. Benzo[a]pyrene (BaP), a carcinogenic polycyclic aromatic hydrocarbon (PAH) was selected as a test soil pollutant because of its widespread presence and recalcitrance in soils [17] as well as its tendency to accumulate in earthworm gut [18, 19]. Multiple phage identification assays were adopted to characterize the earthworm intestinal phage community profile and its response to BaP-induced stress. Prognostic computational approaches were used for phage host prediction and to infer the phage-bacterium relationship. The phage-encoded AMG profile was also characterized to assess the potential of intestinal virome for microbiome resistance to

¹Soil Ecology Lab, Key Laboratory of Plant Immunity, Jiangsu Collaborative Innovation Center for Solid Organic Waste Resource Utilization and Jiangsu Key Laboratory for Solid Organic Waste Utilization, Nanjing 210095, China. ²Catalan Institute for Water Research (ICRA), 17003 Girona, Spain. ³University of Girona, 17004 Girona, Spain. ⁴College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310085, China. ⁵Civil and Environmental Engineering Department, Rice University, Houston, TX 77005, USA. ⁶These authors contributed equally: Rong Xia, Mingming Sun. ✉email: sunmingming@njau.edu.cn; yupf@zju.edu.cn

Received: 15 January 2023 Revised: 24 March 2023 Accepted: 30 March 2023

Published online: 17 April 2023

environmental stress, which was further verified by metatranscriptomics. Overall, this study reveals adaptive strategies of intestinal viromes under BaP stress, and highlights the potential of phages to assist intestinal prokaryotic metabolism and resistance.

MATERIALS AND METHODS

Experimental design and BaP exposure

Surface farmland soil (depth 0–20 cm) without BaP contamination was collected from Nanjing, China (118°52'55" E, 32°2'21" N). The United States Environmental Protection Agency (USEPA) screening levels of BAP is between 0.1 and 1.8 mg kg⁻¹ for resident soil [20]. Accordingly, in the exposure experiments, BaP (J&K Scientific; Guangdong, China) was amended to the collected clean soil at final concentrations of 0.1 (B1), 2.0 (B2), 20.0 (B3), and 200 (B4) mg BaP per kg soil, while soil without BaP contamination (CK) was used as the control. Based on BaP concentrations in soils, B1 was referred to as “low-level”, B2 and B3 as “medium-level”, and B4 as “high-level” BaP treatments, respectively. Five microcosms were prepared at each concentration following the standard process [21] and ten domesticated earthworms (*Metaphire guillelmi* from Yilong Earthworm Farm in Jiangsu, China) with clear clitellum and uniform body weight (~4 g each) were placed in each microcosm for BaP exposure. The microcosms with punctured lids were incubated at 20 °C in darkness for 28 days with a constant humidity of 15–20% [21]. The details of the microcosm preparation and exposure experiments are provided in the Supporting Information as Text S1 and Fig. S1.

After 28 days of incubation, earthworms were collected and the intestinal contents were obtained as previously described [22]. Briefly, the earthworms were anesthetized with anhydrous ethanol and rinsed with sterile water. The complete intestinal contents were collected by dissecting below the clitellum region using sterile scissors and scalpels. Intestinal contents were mixed and divided into several portions, which were stored at different temperature conditions for further analysis. Characterization of lethality and bioconcentration of BaP in the culture soil and earthworm gut after 28 days are shown in Table S1. The reactive oxygen species (ROS) content and the activity of antioxidant enzymes, including superoxide dismutase (SOD), catalase (CAT), and peroxidase (POD) in intestinal sample were determined to assess the response of earthworms to different levels of BaP. The details of the measurement are described in the Supporting Information as Text S2.

Bacterial and phage DNA extraction, metagenome sequencing and assembly

Earthworm intestinal contents (0.25 g) from each treatment were collected for the extraction of total bacterial DNA using DNeasy PowerSoil Pro Kit (Qiagen, Germany). The qualified DNA was used for determining bacterial biomass in earthworm gut by quantitative PCR using universal primers targeting the V4 region of the 16S rRNA gene. Primer sequences were 515F (5'-GTG CCA GCM GCC GCG GTA A-3') and 806R (5'-GGA CTA CHV GGG TWT CTA AT-3'). For phage DNA extraction, virus-like particles (VLP) in the earthworm intestinal contents were first purified and concentrated by filtration and centrifugation following the previously described protocol [23, 24]. Briefly, earthworm intestinal contents were suspended with 1% potassium citrate buffer (10 g l⁻¹ C₆H₅K₃O₇, 1.92 g l⁻¹ Na₂HPO₄·12H₂O, 0.24 g l⁻¹ KH₂PO₄, pH = 7), and then centrifuged at 7000 rpm 4 °C for 10 min to obtain the supernatant. The supernatant was filtered through 0.22 μm filter (Anpel hydrophilic PTFE syringe filter, China) and then concentrated by tangential flow filtration (TFF, Sartorius Vivaflow50 30000 MWCO PES, USA) [24]. Phages were precipitated using polyethylene glycol (PEG 8000, Biofroxx, Germany) and incubated in ice at 4 °C overnight, followed by centrifugation (8000 rpm for 30 min). Then VLP pellets were resuspended in TE buffer (Biosharp, China) and treated with DNase (TaKaRa Recombinant DNase I (RNase-free) 2270A) and RNase (Takara Ribonuclease A 2158). To determine the biomass of phages in the earthworm gut, fluorescence microscopy (Zeiss Axio Imager A1) was used to observe and quantify VLP stained with SYBR Gold fluorescent dye as previously described [16]. The ratio of VLP counts to bacterial 16S rRNA copy numbers was used to estimate virus-to-bacteria ratio (VBR). Thereafter, MagPure Viral DNA/RNA Mini LQ Kits (Angen Biotech Co., China) were used to extract phage DNA and the whole phage genome was amplified by REPLI-g Cell WGA & WTA Kit (Qiagen, Germany). The quality of phage DNA amplified products was verified by 1% agarose gel, Qubit 3.0, and Nanodrop One.

Both bacterial and phage sequencing libraries were generated using NEB Next Ultra DNA Library Prep Kit for Illumina (New England Biolabs, MA, USA) following the manufacturer's recommendations. The libraries were then sequenced using the NovaSeq6000 system (Illumina, San Diego, CA) to yield 150 bp paired-end reads. This approach yielded 78,071,036–108,731,722 reads (average 99,860,922 reads/sample) for metagenome and 36,296,118–38,203,828 reads (average 36,739,983 reads/sample) for virome, respectively. Trimmomatic (v0.36) was used for raw data processing [25] and BWA-MEN (v0.7.17) was used to remove possible eukaryotic genome sequence with parameter $-k = 30$ [26]. Afterwards, the contigs from the bacterial and phage fractions were assembled separately using Megahit (v1.1.2) with parameters as: $k\text{-min } 35$, $k\text{-max } 95$, $k\text{-step } 20$ for bacteria and $--\text{presets meta-large, } -\text{min-contig-len} = 300$ for phage [27].

Bacterial taxonomic assignment and functional annotation

MetaGeneMark (v3.38) [28] was used to predict open reading frames (ORFs) with length information shorter than 90 bp filtered out using default parameters. The CD-HIT (v4.7) [29] was adopted to remove redundancy and obtain the unigenes (i.e., the nucleotide sequences coded by unique and continuous genes) catalog with default parameter: $-c 0.95 -a5 0.8$. The longest sequences in each catalog were chosen to be the representative sequences. Then, the gene catalogs were mapped to clean data using BBMap (v38.90, <https://github.com/BioInfoTools/BBMap>) as default parameter in order to determine the abundance of genes in each sample. MetaPhlan2 (v2.7.6, <https://github.com/biobakery/MetaPhlan2>) [30] was used for the subsequent analysis of bacterial classification and abundance at both phylum and genus levels [31]. Functional genes were annotated by matching to functional gene database KEGG (Release 101.0) [32] and eggNOG (v4.5.1) using DIAMOND with $e\text{-value} \leq 0.001$ [33].

Phage contig identification for taxonomic assignment and functional annotation

CD-HIT (v4.7) was used to cluster the contigs longer than 5 kb of all samples to obtain unique phage contigs with parameter $-c$ at 0.95 and $-a5$ at 0.8. Two independent methods, VirSorter2 (v2.1.0) [34] based on gene enrichment and DeepVirFinder (v1.0) [35] based on designing convolutional neural networks to automatically learn phage genomic signature, were used to identify phage contigs comprehensively. The contig included in output file “final-viral-combined.fa” of virsorter2 and satisfied $p \leq 0.05$ in DeepVirFinder were considered as phages. The phage contigs detected by one or both methods were classified by querying protein sequence against the HMM models with specific phage taxon from the ViPhOG database using “hmmscan” from HMMER (v3.1b2) [36]. The BLASTp mode incorporated in vConTACT2 (v0.9.17) [37] was used to complement phage classification and explore the overlap of earthworm gut phages with phage genomes from human gut [38], soil [39], honey bee gut [14], and RefSeq [40] database utilizing genome gene-sharing profiles. The identified phage contigs were then further analyzed by BWA-MEN (v0.7.17) to compare clean reads and the reads per kilobase of exon model per million mapped reads (RPKM) value of each contig was calculated to obtain the phage abundance. Prodigal (v2.6.3) was used for gene prediction of phage contigs using the meta option [41] and DRAMV (v1.2.0) combined with VIBRANT (v1.2.0) was used for phage-encoded AMG identification with default parameters [42]. The possible function of AMGs was further supported by protein structure prediction. Phyre2 (v2.0) was used for secondary and tertiary structural homology searches for each AMG with alignment coverage above 70% and SWISS-MODEL server (<https://swissmodel.expasy.org/>, accessed on 11 December 2021) was then used to predict the quaternary structure of each protein with a Global Model Quality Estimation (GMQE) score above 0.5. Then BPROM server (<http://www.softberry.com/>, accessed on 2 January 2022) and FindTerm server (<http://www.softberry.com/>, accessed on 2 January 2022) were used to predict the promoter and terminator of AMGs, respectively [43].

Phage-bacterium interaction analysis

Three prognostic computational approaches (CRISPR-match, tRNA match, and genome homology match) were adopted to improve the accuracy of phage host prediction [44]. Specifically, the bacterial scaffolds were used for searching CRISPR spacers by CRISPR recognition tool (CRT, v2.1) [45]. The identified CRISPR spacers were compared with phage contigs using BLASTn (v2.9.0+), and the thresholds of satisfying $\geq 95\%$ identity and ≤ 2 single nucleotide polymorphisms (SNPs) were selected to serve as

putative phage hosts. In parallel, tRNA genes in phage contigs were predicted by tRNAscan-SE (v1.23) with default settings and blasted against the bacterial sequences using the BLASTn, keeping only best hits with at least 95% sequence identity [46]. Moreover, phage genomic signatures in microbial genomes, lysogens, were identified via a search against the bacterial scaffolds through BLASTn with following parameters: bitscore ≥ 50 , e-value $\geq 10^{-3}$, identity $\geq 70\%$, and matching length ≥ 2500 bp [47]. Phage hosts obtained by each method were combined to analyze the host range comprehensively given the high recognition of the three host prediction methods [44]. Phages linked to multiple bacterial genera were regarded as polyvalent phages (i.e., the phage able to infect diverse hosts) [48]. To better understand the variation in phage-bacterium association, DefenseFinder server (<https://defense-finder.mdmparis-lab.com/>, accessed on 13 August 2022) was used to identify antiphage defense systems in bacteria [49], which classified the defense mechanisms as previously suggested [50].

To assess the symbiosis of bacteria and phages in different treatments, the proportion of lysogenic phages in the gut virome and lysogenized bacteria in the gut bacteriome were separately identified. Specifically, lysogenic phages were identified by DeepPhage (v1.0) using 0.6 as the cutoff. The lysogenic marker proteins in phage sequence (transposase, integrase, excisionase, resolvase, and recombinase proteins downloaded from Pfam (v35.0); Table S10) were identified using "hmmScan" from HMMER (v3.1b2) with e-value threshold of 10^{-5} when the DeepPhage probability score was not deemed confident [51]. Additionally, prophages were identified in bacterial genomes according to known phage signatures with CheckV (v0.8.1) and Prophage hunter server (<https://pro-hunter.genomics.cn/index.php/Home>, accessed on 4 May 2022).

Transcriptomic analysis of bacterial and phage functional genes

The transcriptome of earthworm gut microbiome was obtained to investigate the expression of bacterial genes and phage-encoded AMGs at different BaP levels. In brief, the earthworm gut contents were frozen in liquid nitrogen immediately after dissection. Then total RNA was extracted using RNeasy Mini Kit (Qiagen, Germany) from 30 mg sample, which quality and quantity were determined by Agilent 4200 system (Agilent Technologies, Waldbronn, Germany) and NanoDrop2000 (Thermo Fisher, MA, USA), respectively. Whole mRNAseq libraries were generated using NEB Next Ultra Nondirectional RNA Library Prep Kit for Illumina (New England Biolabs) following manufacturer's recommendations. The rRNA transcripts were eliminated by ALFA-SEQ rRNA depletion Kits (for Bacteria). The mRNA was converted into cDNA, followed by purification and amplification to obtain the sequencing library. Paired-end sequencing was conducted on a NovaSeq6000 system to generate 150 bp paired-end reads with low quality and short filtered with Trimmomatic (v0.36). Clean reads were mapped to NCBI Rfam (v14.9) databases to remove non-protein-coding rRNA sequences from eukaryotes and prokaryotes by Bowtie2 (v2.33). Afterwards, the remaining clean reads were assembled into a de novo transcriptome using Trinity (v2.4.0) with the parameters: `--pre_correction --mink 20 --maxk 60 --step 10`. Prodigal (v2.6.3) was used to predict ORFs for assembled Scaffolds longer than 200 bp and the unigenes were clustered using methods consistent with metagenomics [29]. BLASTx (v2.9.0+) was used to compare the unigenes with eggNOG (v4.5.1), CAZy (v2016.7.15), and KEGG (Release 101.0) database for functional annotation. Bowtie2 (v2.33) was used to align the clean reads back to the unigenes with default parameters for quantification of gene expression levels. The expression profiles of phage-encoded AMGs were also determined by mapping the clean metatranscriptomic reads to AMG sequences using BMap (v38.70) with strict criterion (100% identity and 100% coverage) to avoid false alignment to bacterial transcripts [15, 52, 53]. The relative abundance of genes was determined by RPKM for normalization [53].

Statistical analysis

The alpha diversity indexes of the microbes in each sample were determined using the "diversity" function in vegan package in R (v4.0.3). ANOSIM analysis was used to determine the significance of the differences in intestinal microbial communities. Differential expression (DE) of two treatments was estimated using "DESeq" function in DESeq package in R (v4.0.3). KEGG enrichment analysis of significantly augmented genes were implemented using Fisher's Exact Test (<https://www.omicshare.com/tools/>). The significance of differences in parameters were compared using *t*-test in SPSS 26.0. Data visualization was performed in the website: <https://www.chiplotonline/> and R (v4.0.3).

RESULTS AND DISCUSSION

BaP induced oxidative stress and altered the intestinal bacterial community

Although low-level BaP did not noticeably change ROS production, high-level BaP significantly induced ROS production in the earthworm gut. When earthworms were exposed to 2.0, 20.0, and 200 mg kg⁻¹ of BaP, intestinal ROS concentrations increased by 1.42-, 1.22-, and 2.22-fold, respectively (Fig. 1A). In response to this oxidative stress, the intestinal intracellular antioxidant systems were stimulated ($p < 0.05$) as BaP concentrations increased (Figs. 1B and S2). SOD activity was 1.7- and 3.0-fold higher than that of control group when earthworms were exposed to 2.0 and 20.0 mg kg⁻¹, respectively (Fig. 1B). However, excessive oxidative stress could deactivate antioxidant enzymes [54, 55] and result in a significant decrease in SOD activity, as observed under high-level (200 mg kg⁻¹) BaP exposure ($p < 0.05$; Fig. 1B). Accordingly, the functionally active intestinal microbial profile altered significantly under BaP-induced stress ($p < 0.05$, Fig. S3). Particularly, bacterial genera with high environmental adaptivity (*Bacillus* and *Pseudomonas*) [56, 57] were enriched as the BaP levels increased (Fig. 1C). BaP exposure decreased the intestinal microbial diversity including community evenness (i.e., Shannon, Simpson, and Pielou indices) and community richness (i.e., Chao1 and ACE indices) (Fig. 1D and Table S2). In addition, we observed the lowest community evenness at low-level BaP concentration. However, these indices increased with the pollutant levels, which was consistent with the previous microbiome study [58].

Transcriptomic analysis corroborated the toxic effects of BaP on microbial functions inferred by metagenome sequencing. Low-level BaP exposure enriched the expression of genes coding for bacterial energy production pathways (e.g., carbohydrate, lipid, and amino acid metabolism) [59], but the relative abundance of these pathways decreased dramatically under high-level BaP exposure (Fig. 2A). Metabolic pathways closely related to bacterial survival (nucleotide metabolism and genome replication and repair) [60–62] and those associated with bacterial sensing (signal transduction) [63] were significantly enriched in the high-level group (Fig. 2A). Transcriptome analysis also corroborated that the genes associated with energy production were significantly enriched by low-level BaP exposure (1.74-fold) compared to the control ($p < 0.05$), but were significantly depleted when BaP increased further to 200 mg kg⁻¹ ($p < 0.05$; Fig. 2B). However, the relative abundance of genes associated with microbial adaptation (e.g., *sbcb* involved in DNA repair and recombination and *XPF* encoding proteins for DNA excision repair) increased with BaP concentration and was 1.65-fold higher than that of control at 200 mg kg⁻¹ BaP ($p < 0.05$; Fig. 2C). The functional genes that contributed to BaP degradation such as dioxygenase *hcaD*, dehydrogenase *XDH*, and hydrolase *mhpC* also enriched in low-level BaP exposure (1.72-fold higher than that of control, Fig. 2D).

KEGG enrichment analysis was conducted for the genes that were significantly enriched in each BaP concentration to explore the major metabolic pathways where these genes are located (Figs. 2E, F and S4A, B). Low-level BaP exposure significantly enriched genes in glycolysis/gluconeogenesis, carbon metabolism, starch and sucrose metabolism, and citrate cycle (TCA cycle), which are the pathways closely related to microbial energy production [64] (Fig. 2E). In contrast, enriched genes in the high-level BaP group were distributed in spliceosome, DNA damage repair, RNA degradation, proteasome, and other metabolic pathways related to microorganisms adaptation and survival [65] (Fig. 2F). Overall, metagenomic and metatranscriptomic analyses show that BaP-induced stress significantly affected the composition, diversity, and functional profile of the earthworm intestinal bacterial community, which is consistent with previous studies showing that earthworm gut microbes are sensitive to pollutant exposure [66]. As a microbial adaptive strategy to environmental stress, the enrichment in the functional genes that contributed to

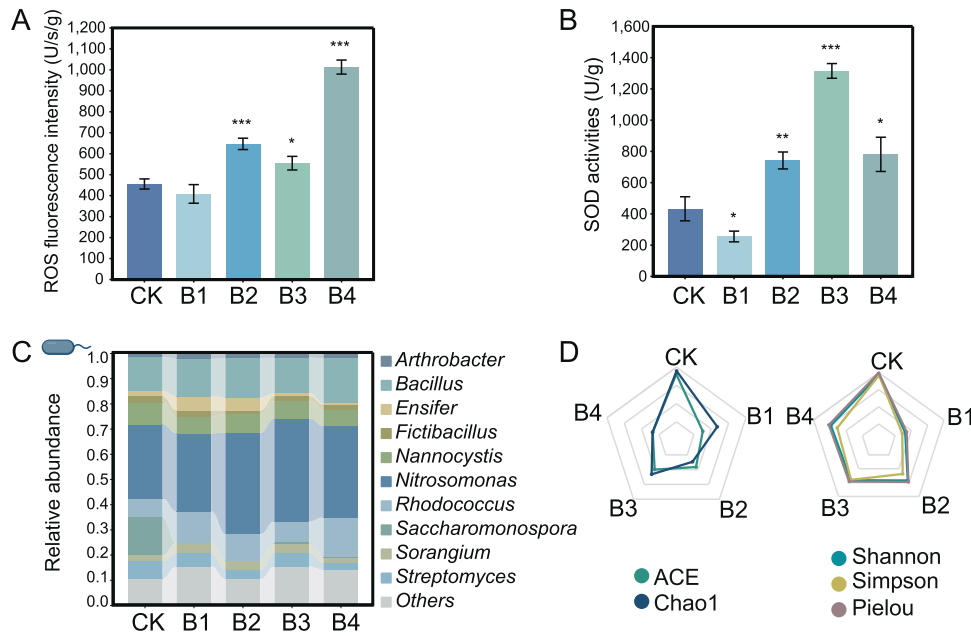


Fig. 1 Toxicity of Benzo[a]pyrene (BaP) to earthworm intestines and gut bacterial community. Reactive oxygen species (ROS) (A) and superoxide dismutase (SOD) activity (B) of earthworm gut exposed to BaP at 0 (CK), 0.1 (B1), 2.0 (B2), 20 (B3), and 200 (B4) mg per kg soil. C Relative abundance of the top 10 abundant intestinal bacteria genus. D Alpha diversity (ACE and Chao1 indexes for richness, and Shannon, Simpson, Pielou indexes for evenness) of the bacterial community in the earthworm intestines. The radar map is drawn according to the data in Table S2 after normalization. The error bars in this study mean standard deviation of three replicates. * $p < 0.05$ indicates significant differences between the BaP exposure treatment and control, based on the t -test.

pollutant degradation and oxidative stress alleviation corroborated that microbes in earthworm gut could facilitate host earthworms to survive hostile environments (Fig. 2D).

BaP exposure affected the composition and lifestyle of phage community in the earthworm gut

A total of 2629 non-redundant phage contigs larger than 5 kb were identified from all the five groups of earthworms intestinal viromes (Table S4). These intestinal phages formed 325 viral clusters (VC) using the existing phage datasets and presented high similarity to human gut and soil phages (Fig. 3A–C). Among them, a total of 41% phage contigs could obtain species annotation information. Taxonomic classification suggested that the dominant taxa in earthworm gut virome include *Siphoviridae*, *Microviridae*, *Myoviridae*, and *Podoviridae* (Fig. 3D), which was corroborated by morphology analysis using TEM imaging (Fig. 3E). When earthworms were exposed to 0, 0.1, 2.0, 20.0, and 200 mg kg⁻¹ BaP, the relative abundance of *Siphoviridae* were 33.1%, 5.8%, 36.6%, 23.8%, and 43.7%, *Microviridae* were 3.0%, 1.2%, 2.2%, 2.7%, and 5.7%, and *Myoviridae* were 3.0%, 0.2%, 1.3%, 0.9%, and 2.0%, respectively (Fig. 3D). The species diversity of phage community was the lowest under low level BaP exposure, and then increased along with the elevated BaP concentrations (Table S5). This could be caused by the activation of prophages and the release of virions under high-level BaP exposure.

BaP exposure remarkably affected the lifecycles of intestinal virome. Lysogenic phages accounted for 56%, 88%, and 43% of total phages in control, low-level BaP, and high-level BaP exposure, respectively (Fig. 3F). Similarly, the relative abundance (average log₁₀ RPKM) of prophages also exhibited bell-shaped stimulatory effect, which was the highest in low level BaP (3.10, $p < 0.05$) (Fig. 3G), and was consistent with the relative abundance of lysogenic bacteria (Fig. S5). Moreover, evidence from the metatranscriptome sequencing also showed that lysogenic marker genes (transposase, integrase, excisionase, resolvase, and recombinase) were significantly enriched under low-level BaP exposure, indicating that phages were prone to integrate into bacterial

genome and became lysogenic phages in worm gut at low BaP level. But the relative abundance of these lysogenic marker genes decreased substantially under high-level BaP exposure (Fig. S6), suggesting the decrease in lysogenic phages. These results suggest that phages adopted a lysogenic adaptive strategy at low BaP levels and converted toward a lytic lifestyle at high BaP levels. High-level BaP, similar to other environmental stressors (e.g., high temperatures, nutrient depletion, and anomalous pH) [67, 68], can induce prophages conversion to a lytic lifecycle. The significant enrichment of SOS response genes *dinF* and *ssb* [69] and depletion of repressor gene *lexA* [70] indicate that high-level BaP exposure could activate the SOS system and trigger a phage lifestyle switch from lysogenic to lytic (Fig. S7).

Phage-encoded AMGs were associated with PAH degradation and antioxidation

To understand the role of intestinal phages in alleviating BaP stress, we investigated the diversity and functions of phage-encoded AMGs in earthworm gut. A total of 203 phage AMGs encoding 44 types of metabolic functions were identified through VIBRANT and DRAMV analysis (Fig. 4A and Table S6). The most common AMGs in lysogenic phages were associated with amino acid metabolism, while AMGs associated with nucleotide metabolism were relatively more abundant in lytic phages (Fig. 4B). These AMGs involved in carbohydrate, glycan, amino acid, and sulfur transformation, which have been widely recognized in marine, terrestrial, and other environments [39, 43, 71], were conducive to microbial metabolism and energy production. Several AMGs, such as *dcd*, *dut*, and *nrdD* that are related to pyrimidine metabolism are known to promote bacterial genome replication and reparation [72] (Fig. 4A). Three AMGs associated with cell membrane integrity and robustness were detected for the first time in the viromes, including *lpxD*, *vanY*, and *NEU1* that were related to lipopolysaccharide biosynthesis, peptidoglycan biosynthesis, and sphingolipid metabolism, respectively [73, 74] (Fig. 4A, C). Additionally, 19 novel AMGs were associated with biosynthesis of folate (such as *dfrB*, *queC*, *queD*, and *queE* etc.),

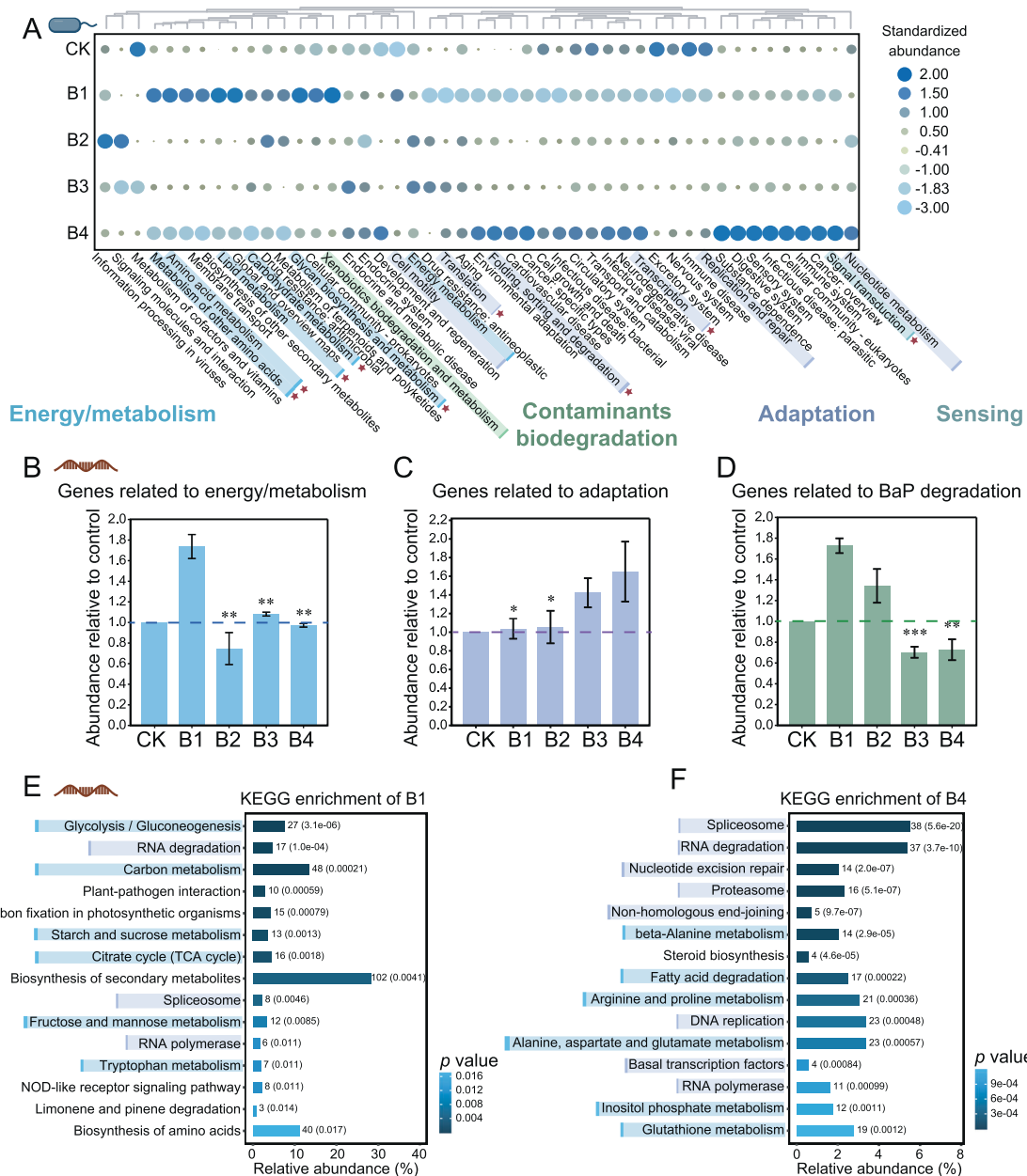


Fig. 2 Functional profile and metatranscriptome validation of prokaryotic community in earthworm intestines. **A** Heatmap shows the relative abundance of KEGG pathway (Level 2) of earthworm intestinal prokaryotic community. Abundance is normalized among different treatments based on Z-score. Red stars represent pathways with significant differences in expression levels in metatranscriptome. The regulation of genes related to energy and metabolism (**B**), stress adaptation (**C**) and BaP degradation (**D**) in metatranscriptomes relative to the control group. The histogram shows the top 15 enriched KEGG orthology (KO) pathways of enriched genes under low- (**E**) and high- (**F**) level BaP exposure. The number of enriched genes in each pathway is labeled in front of the histogram, and the bars are ranked by *p* values. “Relative abundance” represents the proportion of significantly enriched genes to all genes in the KEGG pathway. The dark red single stranded nucleic acid shown in the diagram represents data obtained from the metatranscriptome.

which has potent antioxidant properties that alleviate ROS stress [75]. Among them, *queC*, *queD*, and *queE* were located in the same genomic neighborhood and shared the same promoter and terminator (Fig. 4C), indicating that the functions of these three phage-encoded AMGs were complementary [76]. These AMGs that are associated with microbial adaptation were enriched as BaP stress elevated (Fig. 4A). Notably, AMGs for PAH biodegradation (*ubiE*) [77] were also identified and their relative abundance increased by 55.4-fold in the high-level BaP group (Fig. 4A, C). The structural model prediction of phage-encoded *ubiE* at Phyre2 showed 99.8% confidence, suggesting the AMG sequence could encode a structurally intact protein and perform its function

(Fig. 4C and Table S7). This is the first detection of *ubiE* as a phage-encoded AMG, which encoded benzoquinol methylase to catalyze the anaerobic degradation of phenanthrene [61].

Transcriptomic analysis further confirmed that phage-carried AMGs could facilitate microbial resistance to BaP-induced stress. Among them, 103, 113, 116, and 115 AMGs were expressed when BaP exposure levels were 0.1, 2.0, 20.0, and 200 mg kg⁻¹, respectively (Fig. S8A). The corresponding sum expression abundance (RPKM value) of these AMGs were 1.16E + 05, 5.87E + 04, 8.55E + 04, and 9.19E + 04, respectively (Fig. S8B). Compared to the control group, *ubiE* was enriched by 1.35- and 5.17-folds at 0.1 and 2.0 mg kg⁻¹ BaP exposure, respectively (Fig. S8A). Furthermore, *lpxD*

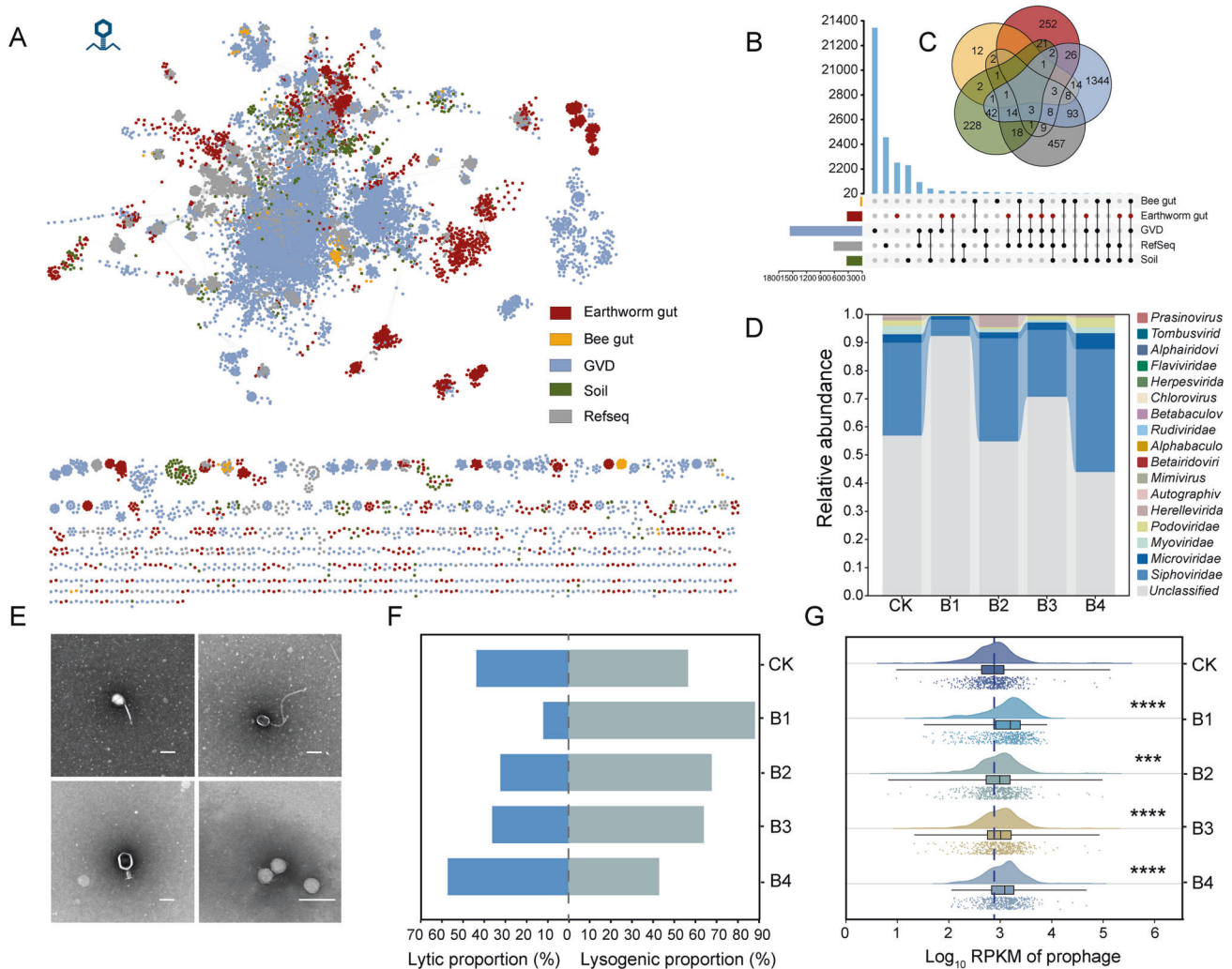


Fig. 3 Composition and lifestyle of earthworm gut virome. **A** Gene-sharing network associates viral contig in earthworm gut (red nodes) with datasets of viral genomes that includes viral sequence from bee gut (yellow nodes), human gut (blue nodes), soil (green nodes), and RefSeq (gray nodes). The edges indicate similarity based on shared protein clusters. **B** The matrix layout shows the number of VCs that are exclusive (one circle) or shared (multiple circles) between the five different datasets used for clustering. **C** Venn diagram of shared VC among the five different datasets. **D** Community compositions (family level) of phages in earthworm gut. **E** Transmission electron microscopy (TEM) image of typical phage in earthworm intestines. The scale bars shown in the figures are 100 nm. **F** Proportions of lytic phage and lysogenic phage in earthworm gut. **G** Abundance (Log₁₀ RPKM) of prophage.

related to microbial resistance to BaP was enriched by 14.1 folds when earthworms were exposed to 2.0 mg kg⁻¹ of BaP ($p < 0.05$, Fig. 4D). In addition, genes associated with microbial resistance to oxidative stress (e.g., *dfxB*, *NAMPT*, and *nadM* [78, 79]) were significantly enriched under BaP-induced stress ($p < 0.05$, Fig. 4D, E).

With the expression of approximately half of the annotated AMGs in viromes, the transcriptomic results verified the important roles of these AMGs in mediating bacterial metabolism in worm gut, especially bacterial resistance to BaP stress. Those AMGs that were not detected in the transcriptome could still serve as functional repositories and play an ecological role in the microbial community through horizontal gene transfer (HGT) [14]. To capture the temporal characteristic of AMG expression, it is important to apply multi-omics analysis to comprehensively reveal the influence of different stress on phages activity, contribution of phages to host bacterial metabolism, and the interactions between phages and bacteria [15, 52, 80]. Although omics analysis greatly advances our understanding of unculturable microbes and their interactions, microbial experiments with representative culturable species can often provide more insights into the microbial processes and underlying mechanisms [80].

BaP contamination affected phage-bacterium interactions and phage adaptive strategies

To elucidate the response of phage-bacterium interactions to BaP stress, we investigated putative phage-host linkages and anti-phage defense systems carried by bacteria and VBR across all the intestinal samples. Based on the CRISPR-match, tRNA match, and genome homology match approaches, we could link 66 phage contigs distributed in *Mimivirus*, *Myoviridae*, *Podoviridae*, and *Siphoviridae* to 4 host phyla and 37 genera (Table S9). Putative host bacteria were *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, and *Firmicutes*, which were the dominant members in the bacterial community (Figs. 5A and S2C). A total of 93 phage-host pairs were matched in the five treatments of these 66 contigs, with more pairings were predicted as BaP stress elevated ($n = 8$ and 37 at low and high BaP exposure, respectively) (Fig. 5A). These results suggest that high-level BaP exposure resulted in more phage-host associations relative to low-level BaP exposure. More phages-host pairings indicated the increased infectivity of phage. Given the increased number of free phages, we can infer that the bacteria were subjected to greater phage predation pressure at high BaP concentrations. The abundance and diversity of anti-phage

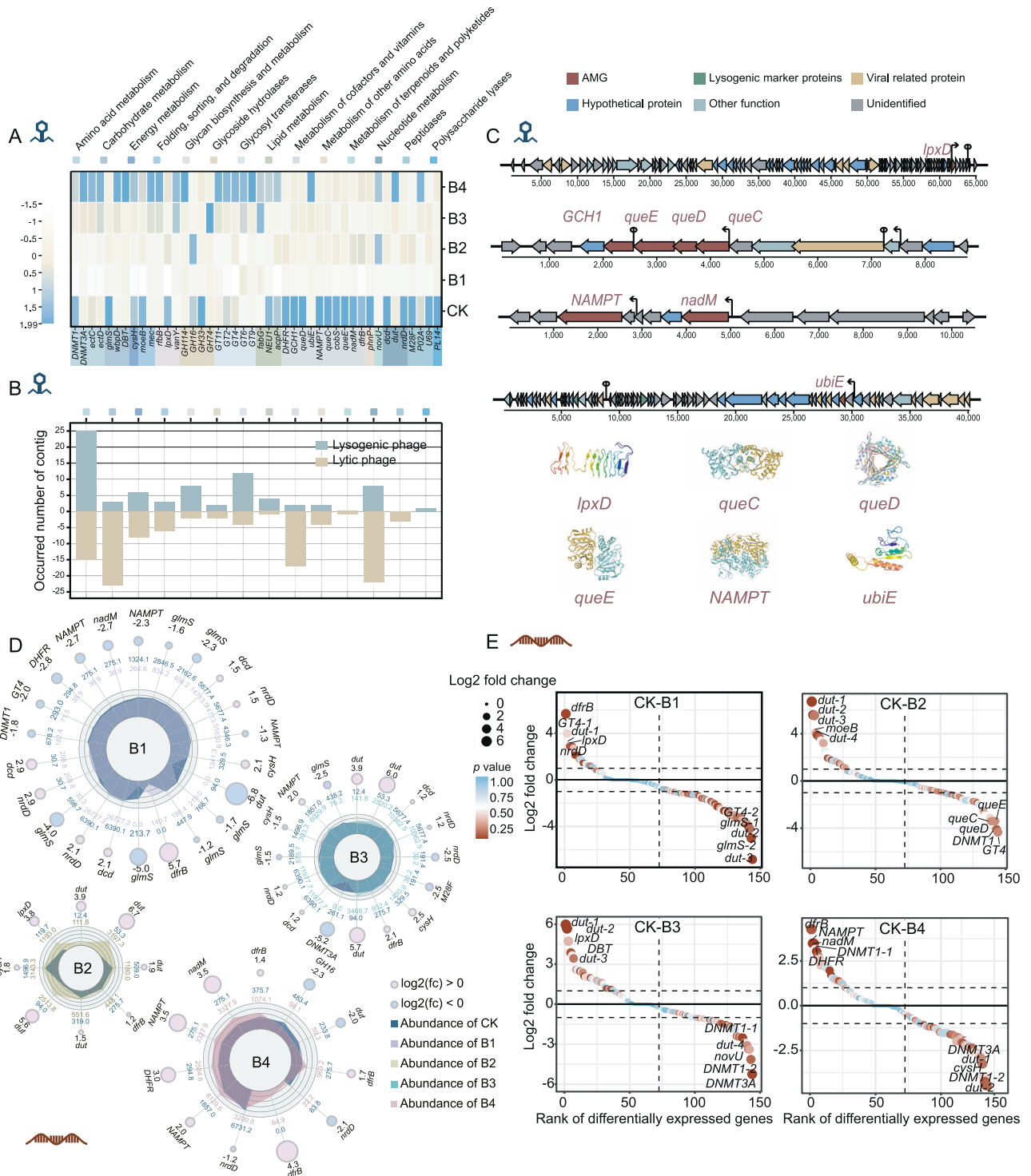


Fig. 4 Phage-encoded auxiliary metabolic genes (AMGs) and their transcriptome profiles under BaP stress. **A** Heatmap shows the classification of AMG in KEGG level 2 pathways and their relative abundance. **B** Number of AMGs in lysogenic and lytic viral contig respectively. “Occurred number of contig” means the number of contigs carrying a certain class of AMGs in different lifestyles. **C** Genomic context and protein structure of six phage encoded AMGs associated with BaP resistance and degradation. **D** Significantly up-regulated or down-regulated gene abundance in different samples and its log₂ fold change (fc) compared with the control. When log₂ (fc) >0 represents that the gene in the sample is significantly up-regulated compared with control, and vice versa. **E** Rank the differentially expressed genes according to the log₂ fold change. Each circle represents a gene, and the color of the circle represents the significance of the difference in gene expression. The genes annotated to the same function were distinguished by different symbols with detailed information listed in Table S8.

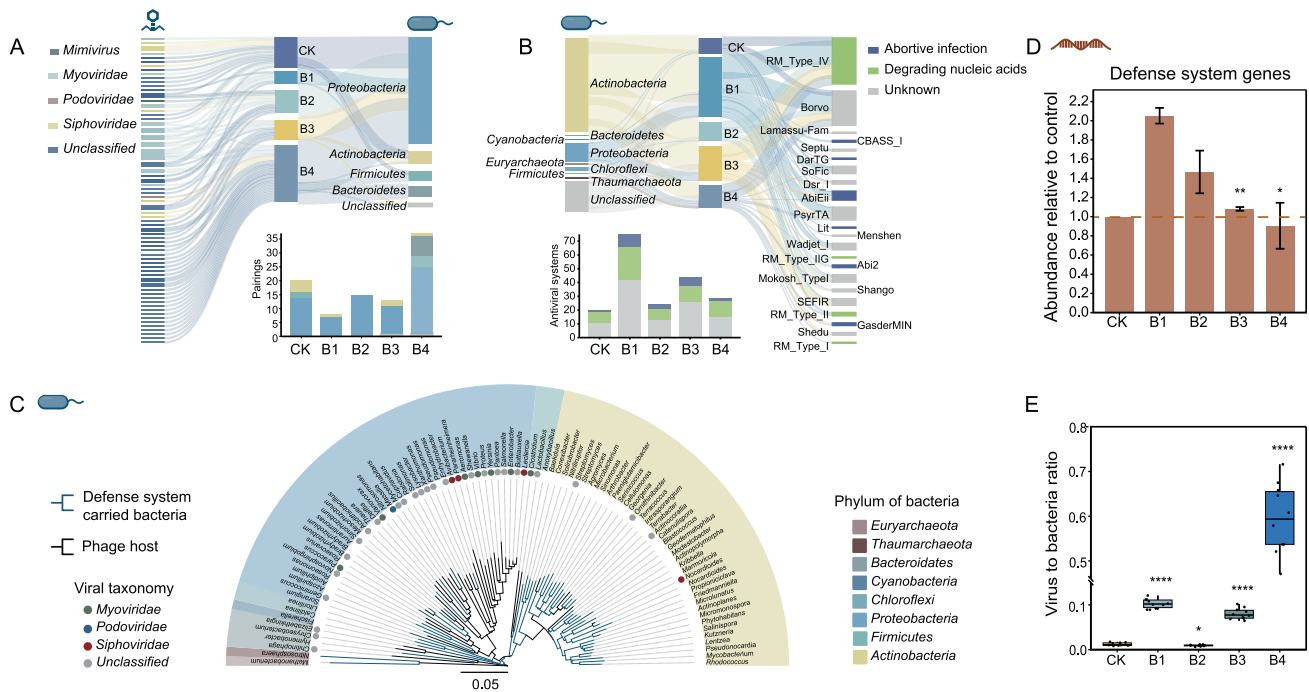


Fig. 5 Analysis of phage-bacterium interactions. **A** Predicted phage-host associations. Left: the viral contigs that can match to host. Different colors present different taxonomic classification at the family level. Middle: the number of phage-host pairings in different samples. Right: the distribution of bacterial hosts at the phylum level. The bar chart below counts the number of pairings in each sample, and the classification of the different hosts in the bar is consistent with the information of the hosts on the right. **B** Identification of antiviral defense systems in bacteria. Left: taxonomic classification of defense systems carrying bacteria at the phylum level. Middle: the number of defense systems in different samples. Right: the subtype and the defense mechanisms of the defense system. The bar charts below show the number of antiviral defence systems. **C** The phylogenetic tree of the phage host and the bacteria that carry the antiviral defense system. **D** The expression abundance of defense system genes in metatranscriptomes are presented by the abundance relative to control. **E** Phage-to-bacteria ratio in earthworm intestines at different concentrations of BaP.

defense systems carried by bacteria (and thus phage-bacterium interactions) changed as BaP stress increased. Specifically, 192 defense systems were observed in 55 bacterial genera in all treatments, corresponding to 22 subtypes and 2 known defense mechanisms “phage nucleic acid degradation” and “abortive infection” (Fig. 5B). The restriction-modification system (RM system), which protects against phage infection via nucleic acid degradation [81], was the most abundant defense system in all samples. Moreover, the most abundant and diverse defense systems were observed at low-level BaP exposure, possibly due to relatively high metabolic ability of bacteria. The largest number of defense systems were observed under low level BaP exposure ($n=75$), which was consistent with the least phage-bacterium association in all samples. As expected, the bacteria carrying the dominant antiphage defense systems were mainly distributed in the phylum *Actinobacteria*, which was different from the predicted dominant prokaryotic hosts (Fig. 5C). The relative abundance of expressed antiphage defense system genes was also found with the highest abundance at low-level BaP exposure while lowest abundance occurred with high-level BaP exposure (Fig. 5D). This was consistent with the increased proportion of prophages at low-level BaP exposure, as non-essential transcription regions within the prophages genome can encode defense systems to prevent host bacteria from subsequent infection [82, 83]. In addition, the strong metabolic capacity of bacteria under low-level BaP exposure also enhanced the expression of prophages and antiphage defense systems simultaneously (Figs. 2B, D and S6) [84]. When the concentration of BaP increased from low to high level, the bacterial defense capability against phage infection and lysis became weaker. The active antiphage systems under low-level BaP exposure could prevent phages from successfully infecting

bacteria, which explained the lower phage-bacteria linkages comparing with high-level BaP exposure.

VBR was also determined to explore phage-bacterium interactions. The VBR was 0.01, 0.1, and 0.6 in clean, low, and high BaP contamination, respectively (Figs. 5E and S9 and S10), indicating that high level of BaP could result in the release of virions from bacteria and the phages adopted an adaptive strategy hostile to the host bacteria. At low BaP concentration, the bacterial density was relatively high and the phage was inclined to lysogenic lifestyle, while the phage lifestyle changed to lytic at high BaP concentration and lower bacterial quantity (Figs. 3F and S10). This switching indicated that the phages in the earthworm gut exposure to BaP were more inclined to adopt the Piggyback-the-Winner (PtW) life strategy, where phages tended to integrate into the host genome instead of lysing the hosts for replication despite of high host abundance. This resulted in a reduced release of virions and ultimately a low VBR [85, 86]. Under pollution stress, this life strategy has profound significance for the adaptation of bacteria and phages. When adopted a lysogenic lifestyle, the phages replicate by taking advantage of the high metabolic capacity of the bacteria at low-level BaP, which is favorable for progeny reproduction [87]. Prophages would protect their host from reinfection through superinfection exclusion mechanism, which may explain the decreased phage-host associations observed at low-level BaP [86, 88]. To alleviate metabolic burden associated with foreign genome integration, the prophages carried functional genes that could enhance host bacterial fitness (Fig. 6A) [14, 56, 58, 59]. However, when exposed to high-level BaP, the intracellular environment of the bacteria was no longer suitable for sheltering lysogenic phages; and prophages were induced and released into the environment via lysing the host

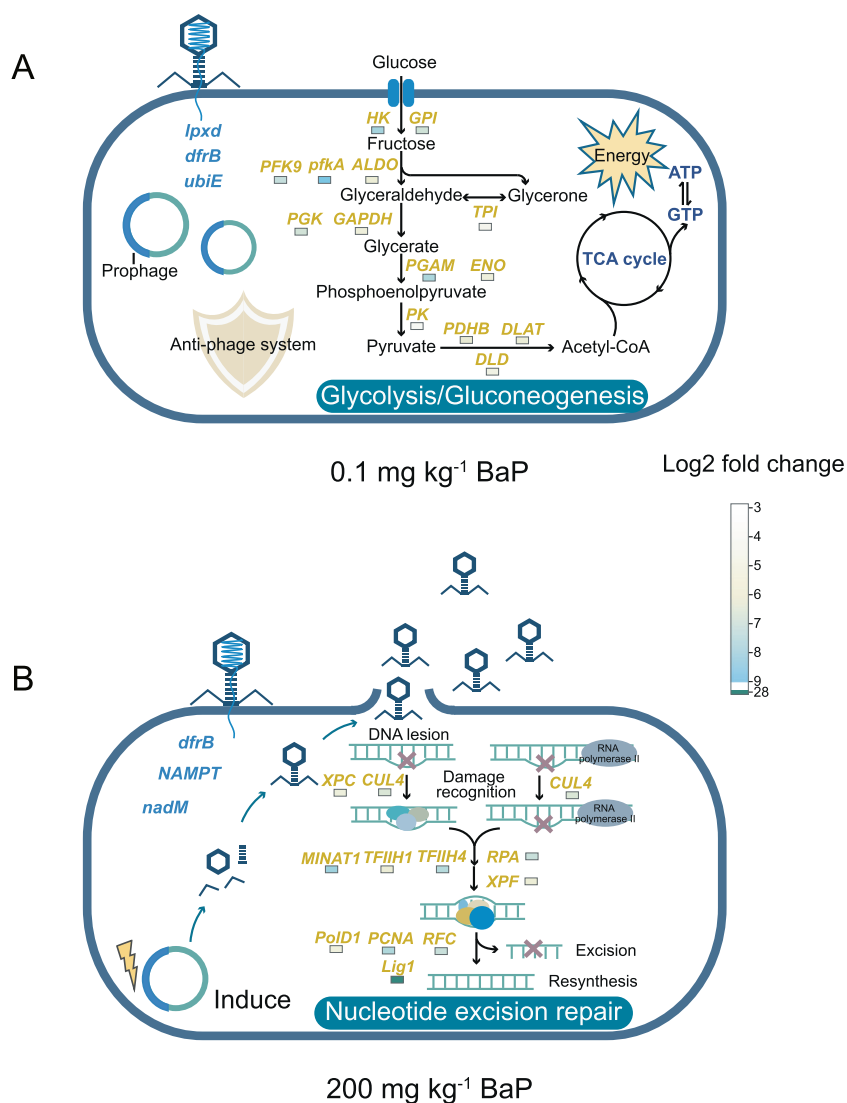


Fig. 6 Conceptual depiction of metabolic functions of bacteria and phages under different BaP stress. The pathways “glycolysis/gluconeogenesis” and “nucleotide excision repair” were respectively taken as examples to emphasize the differences in physiological metabolic states of bacteria under low (A) and high (B) BaP concentration conditions. The yellow fonts represent genes that are significantly enriched in the transcriptome, and the color squares below the genes represent the log₂ fold change compared to the control group. The blue fonts represent phage-encoded AMGs that are prominently expressed.

bacterial cells, leading to the phage-bacterium interaction converting from mutualism to antagonism (Fig. 6B) [89]. Similar conversion was observed in the human gut when inflammation occurred and intestinal oxidative stress increased [90].

CONCLUSIONS

We propose that phage-bacteria interactions and phage-encoded AMGs play important roles in facilitating microbial adaption to BaP stress in earthworm intestines. Low-level BaP stress can stimulate microbial metabolism through exerting a bell-shaped stimulatory effect in earthworm intestinal bacterial community, which enhanced the antiphage defense system and cooperative interaction between phages and bacteria. In contrast, high-level BaP exposure damaged bacterial metabolism and the antiphage systems, resulting in the shift from cooperative to antagonistic relationship between phages and bacteria. Despite the fluctuating interactions between phages and bacteria, the active expression of phage-encoded AMGs related to pollutant resistance and degradation could assist the host bacteria in alleviating

pollution stress. Collectively, these results reveal the PtW life history strategy adopted by the earthworm intestinal phages under BaP exposure, and demonstrate the important ecological functions of intestinal phages in assisting symbiotic bacteria to survive pollution stress.

DATA AVAILABILITY

All raw sequence data generated in this research have been deposited in the NCBI's Sequence Read Archive (SRA) database and can be accessed under project accession no. PRJNA922612. All data are publicly accessible and can be downloaded from <https://www.ncbi.nlm.nih.gov/sra/PRJNA922612>.

REFERENCES

- Blume-Werry G, Krab EJ, Olofsson J, Sundqvist MK, Väisänen M, Klaminder J. Invasive earthworms unlock arctic plant nitrogen limitation. *Nat Commun.* 2020;11:1766.
- Xing Y, Meng X, Wang L, Zhang J, Wu Z, Gong X, et al. Effects of benzotriazole on copper accumulation and toxicity in earthworm (*Eisenia fetida*). *J Hazard Mater.* 2018;351:330–6.

3. Jiang X, Chang Y, Zhang T, Qiao Y, Klobučar G, Li M. Toxicological effects of polystyrene microplastics on earthworm (*Eisenia fetida*). *Environ Pollut*. 2020;259:113896.
4. Drake HL, Horn MA. As the worm turns: the earthworm gut as a transient habitat for soil microbial biomes. *Annu Rev Microbiol*. 2007;61:169–89.
5. Sun M, Chao H, Zheng X, Deng S, Ye M, Hu F. Ecological role of earthworm intestinal bacteria in terrestrial environments: a review. *Sci Total Environ*. 2020;740:140008.
6. Thakuria D, Schmidt O, Finan D, Egan D, Doohan FM. Gut wall bacteria of earthworms: a natural selection process. *ISME J*. 2010;4:357–66.
7. Jin B-J, Liu X-P, Roux XL, Bi Q-F, Li K-J, Wu C-Y, et al. Biochar addition regulates soil and earthworm gut microbiome and multifunctionality. *Soil Biol Biochem*. 2022;173:108810.
8. Zidar P, Kos M, Vogel-Mikuš K, van Elteren JT, Debeljak M, Žižek S. Impact of ionophore monensin on performance and Cu uptake in earthworm *Eisenia andrei* exposed to copper-contaminated soil. *Chemosphere*. 2016;161:119–26.
9. Liang G, Bushman FD. The human virome: assembly, composition and host interactions. *Nat Rev Microbiol*. 2021;19:514–27.
10. Mirzaei MK, Maurice CF. Ménage à trois in the human gut: interactions between host, bacteria and phages. *Nat Rev Microbiol*. 2017;15:397–408.
11. Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, et al. Uncovering earth's virome. *Nature*. 2016;536:425–30.
12. Pratama AA, van Elsas JD. The 'neglected' soil virome—potential role and impact. *Trends Microbiol*. 2018;26:649–62.
13. Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A, et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature*. 2016;537:689–93.
14. Deboutte W, Beller L, Yinda CK, Maes P, de Graaf DC, Matthijssens J. Honey-bee-associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential. *Proc Natl Acad Sci USA*. 2020;117:10511–9.
15. Wu R, Davison MR, Gao Y, Nicora CD, McDermott JE, Burnum-Johnson KE, et al. Moisture modulates soil reservoirs of active DNA and RNA viruses. *Commun Biol*. 2021;4:992.
16. 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.
17. Yu J, Luo H, Yang B, Wang M, Gong Y, Wang P, et al. Risk control values and remediation goals for benzo[a]pyrene in contaminated sites: sectoral characteristics, temporal trends, and empirical implications. *Environ Sci Technol*. 2023;57:2064–74.
18. Keum H, Kim J, Joo YH, Kang G, Chung N. Hemoglobin peroxidase reaction of hemoglobin efficiently catalyzes oxidation of benzo[a]pyrene. *Chemosphere*. 2021;268:128795.
19. Yu H, Li T, Liu Y, Ma L. Spatial distribution of polycyclic aromatic hydrocarbon contamination in urban soil of China. *Chemosphere*. 2019;230:498–509.
20. United States Environmental Protection Agency. Regional screening levels (RSLs). <https://semspub.epa.gov/work/HQ/403634.pdf>.
21. Wang H-T, Zhu D, Li G, Zheng F, Ding J, O'Connor PJ, et al. Effects of arsenic on gut microbiota and its biotransformation genes in earthworm *Metaphire sieboldi*. *Environ Sci Technol*. 2019;53:3841–9.
22. Chao H, Sun M, Ye M, Zheng X, Hu F. World within world: intestinal bacteria combining physiological parameters to investigate the response of *Metaphire guillelmi* to tetracycline stress. *Environ Pollut*. 2020;261:114174.
23. Mangalea MR, Paez-Espino D, Kieft K, Chatterjee A, Chriswell ME, Seifert JA, et al. Individuals at risk for rheumatoid arthritis harbor differential intestinal bacteriophage communities with distinct metabolic potential. *Cell Host Microbe*. 2021;29:726–39.e5.
24. 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;16:1397–408.
25. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
26. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. 2012;28:1420–8.
27. Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, 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.
28. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res*. 2010;38:e132.
29. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*. 2012;28:3150–2.
30. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods*. 2015;12:902–3.
31. Khan Mirzaei M, Khan MAA, Ghosh P, Taranu ZE, Taguer M, Ru J, et al. Bacteriophages isolated from stunted children can regulate gut bacterial communities in an age-specific manner. *Cell Host Microbe*. 2020;27:199–212.e5.
32. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 1999;27:29–34.
33. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. *BMC Bioinform*. 2003;4:41.
34. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*. 2021;9:37.
35. Ren J, Song K, Deng C, Ahlgren NA, Fuhrman JA, Li Y, et al. Identifying viruses from metagenomic data using deep learning. *Quant Biol*. 2020;8:64–77.
36. Camarillo-Guerrero LF, Almeida A, Rangel-Pineros G, Finn RD, Lawley TD. Massive expansion of human gut bacteriophage diversity. *Cell*. 2021;184:1098–109.e9.
37. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat Biotechnol*. 2019;37:632–9.
38. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. The gut virome database reveals age-dependent patterns of virome diversity in the human gut. *Cell Host Microbe*. 2020;28:724–40.e8.
39. 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.
40. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, et al. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res*. 2018;46:D851–60.
41. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
42. Li Z, Pan D, Wei G, Pi W, Zhang C, Wang JH, et al. Deep sea sediments associated with cold seeps are a subsurface reservoir of viral diversity. *ISME J*. 2021;15:2366–78.
43. Gazitúa MC, Vik DR, Roux S, Gregory AC, Bolduc B, Widner B, et al. Potential virus-mediated nitrogen cycling in oxygen-depleted oceanic waters. *ISME J*. 2021;15:981–98.
44. Jahn MT, Lachnit T, Markert SM, Stigloher C, Pita L, Ribes M, et al. Lifestyle of sponge symbiont phages by host prediction and correlative microscopy. *ISME J*. 2021;15:2001–11.
45. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinform*. 2007;8:209.
46. Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 2016;44:W54–7.
47. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. Computational approaches to predict bacteriophage–host relationships. *FEMS Microbiol Rev*. 2016;40:258–72.
48. de Jonge PA, Nobrega FL, Brouns SJJ, Dutilh BE. Molecular and evolutionary determinants of bacteriophage host range. *Trends Microbiol*. 2019;27:51–63.
49. Piel D, Bruto M, Labreuche Y, Blanquart F, Goudeñe D, Barcia-Cruz R, et al. Phage-host coevolution in natural populations. *Nat Microbiol*. 2022;7:1075–86.
50. Tesson F, Hervé A, Mordret E, Touchon M, d'Humières C, Cury J, et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat Commun*. 2022;13:2561.
51. Cook R, Hooton S, Trivedi U, King L, Dodd CER, Hobman JL, et al. Hybrid assembly of an agricultural slurry virome reveals a diverse and stable community with the potential to alter the metabolism and virulence of veterinary pathogens. *Microbiome*. 2021;9:65.
52. Luo XQ, Wang P, Li JL, Ahmad M, Duan L, Yin LZ, et al. Viral community-wide auxiliary metabolic genes differ by lifestyles, habitats, and hosts. *Microbiome*. 2022;10:190.
53. Yoshida T, Nishimura Y, Watai H, Haruki N, Morimoto D, Kaneko H, et al. Locality and diel cycling of viral production revealed by a 24 h time course cross-omics analysis in a coastal region of Japan. *ISME J*. 2018;12:1287–95.
54. Qiao Z, Zhang F, Yao X, Yu H, Sun S, Li X, et al. Growth, DNA damage and biochemical toxicity of cyantraniliprole in earthworms (*Eisenia fetida*). *Chemosphere*. 2019;236:124328.
55. Wang H, Zhang X, Wang L, Zhu B, Guo W, Liu W, et al. Biochemical responses and DNA damage induced by herbicide QYR301 in earthworm (*Eisenia fetida*). *Chemosphere*. 2020;244:125512.
56. Losick RM. *Bacillus subtilis*: a bacterium for all seasons. *Curr Biol*. 2020;30:R1146–50.
57. Wu X, Monchy S, Taghavi S, Zhu W, Ramos J, van der Lelie D. Comparative genomics and functional analysis of niche-specific adaptation in *Pseudomonas putida*. *FEMS Microbiol Rev*. 2011;35:299–323.
58. Liu Y-R, Delgado-Baquerizo M, Bi L, Zhu J, He J-Z. Consistent responses of soil microbial taxonomic and functional attributes to mercury pollution across China. *Microbiome*. 2018;6:183.

59. Busi SB, Bourquin M, Fodelianakis S, Michoud G, Kohler TJ, Peter H, et al. Genomic and metabolic adaptations of biofilms to ecological windows of opportunity in glacier-fed streams. *Nat Commun.* 2022;13:2168.
60. Fitzsimmons Liam F, Liu L, Kim J-S, Jones-Carson J, Vázquez-Torres A. *Salmonella* reprograms nucleotide metabolism in its adaptation to nitrosative stress. *mBio.* 2018;9:e00211–18.
61. Shafiei R, Lepince P, Sombolostani AS, Thonart P, Delvigne F. Effect of sequential acclimation to various carbon sources on the proteome of *Acetobacter senegalensis* LMG 23690T and its tolerance to downstream process stresses. *Front Microbiol.* 2019;10:608.
62. Cortez D, Neira G, González C, Vergara E, Holmes DS. A large-scale genome-based survey of acidophilic bacteria suggests that genome streamlining is an adaptation for life at low pH. *Front Microbiol.* 2022;13:803241.
63. Kinch LN, Cong Q, Jaishankar J, Orth K. Co-component signal transduction systems: fast-evolving virulence regulation cassettes discovered in enteric bacteria. *Proc Natl Acad Sci USA.* 2022;119:e2203176119.
64. Zhao Y, Han Y, Sun Y, Wei Z, Chen J, Niu X, et al. Comprehensive succinylome profiling reveals the pivotal role of lysine succinylation in energy metabolism and quorum sensing of *Staphylococcus epidermidis*. *Front Microbiol.* 2021;11:632367.
65. Selmoni O, Lecellier G, Magalon H, Vigliola L, Oury N, Benzoni F, et al. Seascape genomics reveals candidate molecular targets of heat stress adaptation in three coral species. *Mol Ecol.* 2021;30:1892–906.
66. Liu P, Yang Y, Li M. Responses of soil and earthworm gut bacterial communities to heavy metal contamination. *Environ Pollut.* 2020;265:114921.
67. Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* 2017;11:1511–20.
68. Correa AMS, Howard-Varona C, Coy SR, Buchan A, Sullivan MB, Weitz JS. Revisiting the rules of life for viruses of microorganisms. *Nat Rev Microbiol.* 2021;19:501–13.
69. Dubiel K, Henry C, Spenkelink LM, Kozlov AG, Wood EA, Jergic S, et al. Development of a single-stranded DNA-binding protein fluorescent fusion toolbox. *Nucleic Acids Res.* 2020;48:6053–67.
70. Hu J, Ye H, Wang S, Wang J, Han D. Prophage activation in the intestine: insights into functions and possible applications. *Front Microbiol.* 2021;12:785634.
71. Kieft K, Zhou Z, Anderson RE, Buchan A, Campbell BJ, Hallam SJ, et al. Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages. *Nat Commun.* 2021;12:3503.
72. He T, Li H, Zhang X. Deep-sea hydrothermal vent viruses compensate for microbial metabolism in virus-host interactions. *mBio.* 2017;8:e00893–17.
73. Buetow L, Smith TK, Dawson A, Fyffe S, Hunter WN. Structure and reactivity of *LpxD*, the N-acyltransferase of lipid A biosynthesis. *Proc Natl Acad Sci USA.* 2007;104:4321–6.
74. Meziane-Cherif D, Stogios PJ, Evdokimova E, Savchenko A, Courvalin P. Structural basis for the evolution of vancomycin resistance D,D-peptidases. *Proc Natl Acad Sci USA.* 2014;111:5872–7.
75. Tu H-C, Lin M-Y, Lin C-Y, Hsiao T-H, Wen Z-H, Chen B-H, et al. Supplementation with 5-formyltetrahydrofolate alleviates ultraviolet B-inflicted oxidative damage in folate-deficient zebrafish. *Ecotoxicol Environ Saf.* 2019;182:109380.
76. Howard-Varona C, Hargreaves KR, Solonenko NE, Markillie LM, White RA, Brewer HM, et al. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *ISME J.* 2018;12:1605–18.
77. Zhang Z, Sun J, Guo H, Wang C, Fang T, Rogers MJ, et al. Anaerobic biodegradation of phenanthrene by a newly isolated nitrate-dependent *Achromobacter denitrificans* strain PheN1 and exploration of the biotransformation processes by metabolite and genome analyses. *Environ Microbiol.* 2021;23:908–23.
78. Cameron AM, Castoldi A, Sanin DE, Flachsmann LJ, Field CS, Puleston DJ, et al. Inflammatory macrophage dependence on NAD⁺ salvage is a consequence of reactive oxygen species-mediated DNA damage. *Nat Immunol.* 2019;20:420–32.
79. Garten A, Schuster S, Penke M, Gorski T, de Giorgis T, Kiess W. Physiological and pathophysiological roles of NAMPT and NAD metabolism. *Nat Rev Endocrinol.* 2015;11:535–46.
80. Pratama AA, Bolduc B, Zayed AA, Zhong ZP, Guo J, Vik DR, et al. Expanding standards in viromics: in silico evaluation of dsDNA viral genome identification, classification, and auxiliary metabolic gene curation. *PeerJ.* 2021;9:e11447.
81. Cury J, Bernheim A. CRISPR-Cas and restriction-modification team up to achieve long-term immunity. *Trends Microbiol.* 2022;30:513–4.
82. Dedrick RM, Jacobs-Sera D, Bustamante CA, Garlena RA, Mavrich TN, Pope WH, et al. Prophage-mediated defence against viral attack and viral counter-defence. *Nat Microbiol.* 2017;2:16251.
83. Hampton HG, Watson BNJ, Fineran PC. The arms race between bacteria and their phage foes. *Nature.* 2020;577:327–36.
84. Liu X, Lu Q, Du M, Xu Q, Wang D. Hormesis-like effects of tetrabromobisphenol A on anaerobic digestion: responses of metabolic activity and microbial community. *Environ Sci Technol.* 2022;56:11277–87.
85. Knowles B, Silveira CB, Bailey BA, Barott K, Cantu VA, Cobián-Güemes AG, et al. Lytic to temperate switching of viral communities. *Nature.* 2016;531:466–70.
86. Kirsch JM, Brzozowski RS, Faith D, Round JL, Secor PR, Duerkop BA. Bacteriophage-bacteria interactions in the gut: from invertebrates to mammals. *Annu Rev Virol.* 2021;8:95–113.
87. Goldhill DH, Turner PE. The evolution of life history trade-offs in viruses. *Curr Opin Virol.* 2014;8:79–84.
88. Bondy-Denomy J, Qian J, Westra ER, Buckling A, Guttman DS, Davidson AR, et al. Prophages mediate defense against phage infection through diverse mechanisms. *ISME J.* 2016;10:2854–66.
89. Feiner R, Argov T, Rabinovich L, Sigal N, Borovok I, Herskovits AA. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nat Rev Microbiol.* 2015;13:641–50.
90. Shkoporov AN, Turkington CJ, Hill C. Mutualistic interplay between bacteriophages and bacteria in the human gut. *Nat Rev Microbiol.* 2022;20:737–49.

ACKNOWLEDGEMENTS

We thank Dr. Simon Roux, Dr. Mao Ye, Dr. Ruonan Sun, and Dan Huang for the valuable advice on the phage genome analyses and data interpretation. This work was supported by the National Natural Science Foundation of China (42077106, 42277115, 42177113, and 42277418), the Key R&D Project of Jiangsu Province (Modern Agriculture; BE2022322), and NSF ERC on Nanotechnology-Enabled Water Treatment (EEC-1449500).

AUTHOR CONTRIBUTIONS

RX constructed viral genome data and made the figures. RX, MS and PY performed bacteria and virus metagenome analyses and virus-host linkage analysis. JLB and PJJ assisted bacterial metagenome analyses. PY, JLB and PJJ analyzed the viral AMG annotation. MS, FH and PY supervised the study. RX, MS and PY wrote the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41396-023-01408-x>.

Correspondence and requests for materials should be addressed to Mingming Sun or Pingfeng Yu.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023