

Enhanced Mutualistic Symbiosis Between Soil Phages and Bacteria With Elevated Chromium-induced Environmental Stress

Dan Huang

Institute of Soil Science Chinese Academy of Sciences

Pingfeng Yu (✉ pingfeng.yu@rice.edu)

Rice University <https://orcid.org/0000-0003-0402-773X>

Mao Ye

Institute of Soil Science Chinese Academy of Sciences

Cory Schwarz

Rice University

Xin Jiang

Institute of Soil Science Chinese Academy of Sciences

Pedro J.J. Alvarez

Rice University

Research

Keywords: Soil virome, phage-bacteria interactions, Lysogenic conversion, Auxiliary metabolic genes, Chromium Stress

Posted Date: February 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-207906/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Microbe-virus interactions have broad implications on the composition, function, and evolution of microbiomes. Elucidating effects of environmental stresses on these interactions is critical to identify the ecological function of viral communities and understand microbiome environmental adaptation. Heavy metal-contaminated soils represent a relevant ecosystem to study the interplay between microbes, viruses and environmental stressors.

Results: Metagenomic analysis revealed that Cr pollution adversely altered the abundance, diversity and composition of viral and bacterial communities. Host-phage linkage based on CRISPR indicated that, in soils with high Cr contamination, the abundance of phages associated with heavy metal tolerant hosts increased, as did the relative abundance of phages with broad host ranges (identified as host-phage linkages across genera), which would facilitate transfection and broader distribution of resistance genes in the bacterial community. Examining variations along the pollutant gradient, phage-bacteria interactions shifted to a more protective mutualistic relationship in the face of greater environmental stresses. Specifically, the fractions of lysogens in bacterial communities (identified by integrase genes within bacterial genomes and prophage induction assay by mitomycin-C) were positively correlated with Cr contamination levels. Furthermore, viral genomic analysis demonstrated that lysogenic phages under higher Cr-induced stresses carried more auxiliary metabolic genes regulating microbial heavy metal detoxification.

Conclusion: With the intensification of Cr-induced environmental stresses, the composition, replication strategy, and ecological function of the phage community all evolve alongside the bacterial community to adapt to extreme habitats. This results in a transformation of the phage-bacteria interaction from parasitism to protective mutualism in extreme environments, and underscores the influential role of phages in bacterial adaptation to pollution-related stress and in related biogeochemical processes.

Background

Phages, bacterial viruses that constitute the most abundant and genetically most diverse biological entities in the biosphere, are critical components of microbial communities [1, 2]. The rapid development of viral metagenomics has advanced our understanding of phage abundance and composition in various ecosystems [3-8], allowing us to identify their functional importance more comprehensively. The sheer abundance and diversity of phages, together with their multifaceted impacts on microbial hosts, strongly indicate that the phages can influence entire macroscopic ecosystems by affecting the native microbial communities [9-13]. However, previous virome studies mainly focused on phage species and genetic profiles from different ecosystems, while bacteria-phage interactions have been heretofore relatively understudied [6, 14]. To better understand the ecological and evolutionary roles of phages in ecosystems, it is important to elucidate bacteria-phage interactions and their response to environmental stresses.

Recent studies have shed light on manifold modes of phage-bacteria interactions in various ecosystems [4, 15-17]. Phages are obligate intracellular bacterial parasites who rely on the presence of appropriate hosts to persist and propagate [18, 19]. The classical view of phage-bacteria parasitism holds that phages propagate at sacrifice of bacterial hosts and redirects host resources to produce virions [20]. Beside the parasitic pattern often observed under favorable conditions, mutualistic phage-bacteria interactions also occur under unfavorable circumstances (e.g., solar radiation or nutrient limitation) [21, 22]. For example, lysogenic conversion associated with horizontal gene transfer and host phenotype alteration can expand the host metabolic profile and enhance microbial environmental adaptability [23-26]. In return, the phages that are integrated into the host genomes in dormant forms (i.e., prophages) can take advantage of microbial environmental resistance to avoid direct exposure to harsh environments [27]. Typically, the extent of phage-bacteria mutualism is positively correlated with the abundance of lysogenic phages in a virome, lysogenized bacteria in a microbiome, and auxiliary metabolic genes (AMGs) carried by lysogenic phages [23].

Heavy metal-contaminated soils generally exhibit a decreasing geochemical gradient from the contamination source [28, 29]. Such soils are recognized as ecosystems under complex stresses and are both relevant and convenient to study the interplay between phages, bacteria and environmental stressors. However, the viromes in heavy metal-contaminated soils remain poorly characterized and the response of phage-bacteria interactions to heavy metal-induced environmental stress is largely unknown. This critical knowledge gap should be addressed to inform microbiome adaptation to hostile environments and the role of phages in related biogeochemical processes.

The objective of this study was to explore the variability of soil viromes and phage-bacteria interactions along gradients of environmental stress. Chromium (Cr) contaminated sites were chosen to represent heavy metal soil contamination, due to the high frequency of occurrence and well-recognized ecotoxicity [30, 31]. Soil samples were taken along Cr gradients in one lightly polluted site (Luzhou, LZ for short) and one heavily polluted site (Zhangye, ZY for short). Metagenomic analyses for total environmental DNA and viral DNA were separately conducted to inform indigenous bacterial and viral profiles. Bacteria-phage linkages, commonality of lysogenicity within the virome, abundance of phage-carried AMGs and the lysogenized fraction of bacteria were investigated to demonstrate how variations in the virome and its functions were related to the indigenous bacteria under Cr-generated environmental stress. This metagenomic study reveals an enhanced mutualistic symbiosis between phages and bacteria with elevated Cr-induced environmental stress, and highlights the roles of phages in mediating microbiome adaptation to hostile environments.

Results

Bacterial Community Profiles of Cr-Contaminated Sites

There was considerable variability in physicochemical and nutritional properties of the contaminated soil, including gradients in Cr contamination (**Fig.S1 and Table S1**). Based on the Cr pollution level and

nutrient conditions, these seven samples can be roughly divided into three groups: a slightly-contaminated group (L1, L2 and Z1 with available Cr concentrations of 0.11, 0.27, and 0.91 mg/kg, respectively), a moderately-contaminated group (L3 and Z2 with Cr concentrations of 6.76 and 7.09 mg/kg, respectively) and a highly-contaminated group (Z3 and Z4 with Cr concentrations of 413.84 and 465.42 mg/kg, respectively). Accordingly, there was a logical decrease in microbial biomass with the increase of Cr contamination and decrease in nutrient availability. The soil bacterial abundance, quantified by 16S rRNA abundance, fell dramatically by several orders of magnitude along the increasing Cr concentration gradient, from 9.78 to 5.97 log (copies)/gdw (per gram dry weight) in LZ, and from 8.89 to 5.96 log (copies)/gdw in ZY (**Fig.1A**).

Metagenomic analysis of dominant bacterial species and beta diversity revealed a significant difference in the bacterial communities of soils with different levels of Cr pollution. Phylogenetic analysis and heatmap of the 565 dominant species (with a relative abundance of > 0.1%) in LZ and ZY sites showed that the bacterial species from the comparable contamination levels (slight, moderate, and severe) were more closely related, while those from different contamination levels differed observably (**Fig.1B**), although they were all mainly from *Proteobacteria* and *Actinobacteria*. Consistently, redundant analysis (RDA) of bacterial communities showed that bacterial profiles in L1, L2 and Z1 with low available Cr concentration were more similar, and those in L3 and Z2 with moderate pollution levels were clustered together. The bacterial communities in Z3 and Z4 were dramatically different from those in other sites, possibly due to the severe degree of pollution and associated poor nutritional status (**Fig.S2**).

Virome Patterns along Cr contamination gradients.

In total, 5099 viral clusters (VCs) were identified from the 7257 viral contigs (>5 kb) originating from site LZ and 3021 VCs were identified from the 4561 viral contigs (>5 kb) originated from site ZY (**Fig.S3A & B**). As with most characterized viromes in various habitats [14, 32], the annotated phages in Cr-contaminated soil predominately belonged to *Siphoviridae*, *Podoviridae* and *Myoviridae* families (**Fig.2A**). The relative abundance of *Siphoviridae*, *Podoviridae* and *Myoviridae* families among the seven different soil samples was 33.3-93.4%, 0.7-31.9% and 0.5-12.2%, respectively (**Fig.2A**). Similar to the trend of bacterial abundance, the abundance of viral-like particles (VLP, determined by fluorescence microscopy imaging) decreased from 9.58 to 7.18 log(VLP)/gdw in site LZ and from 9.13 to 7.86 log(VLP)/gdw in site ZY with the increase of Cr contamination (**Fig.2B**).

Consistent with the trends observed in the bacterial communities, the phage populations among these seven different locations were significantly affected by soil Cr contamination and soil nutrient status. Cluster analysis of the top 20 viruses (with a relative abundance of > 1.0 %) showed that viruses from samples with similar physical-chemical properties were more closed related: L1 and L2 clustered together, Z1, Z2 and L3 clustered together, and Z3 was most closely clustered to Z4 (**Fig.S4**). RDA corroborated the viral composition patterns with virome of soil with the same level of Cr contamination clustered together (**Fig.S5**).

Compared with the prokaryotic viruses in the viral database (RefSeq, version 75) [33], the viromes of Cr-contaminated sites were highly novel and variant in terms of genetic profiles (**Fig.3**). Specifically, approximately 92.3% of viral contig sequences are not related to RefSeq virus sequences. Only 7.7% of the viral contigs had high genetic similarity to viruses recorded in RefSeq (weighted pairwise similarity scores of ≥ 1). Whereas 24.6% of the matched viral contigs in the contaminated soil were linked with viruses found in soil, the values decreased from 38.0% to 20.9% with the increase of pollution concentration in each sample (**Fig.S6**). Consistent with the similarities in viral profiles, the soil viral contigs with similar levels of Cr contamination were assigned with relevant habitats (**Fig.S6**). Interestingly, 14.6% of the matched viral contigs were linked with aquatic viruses. Besides, there were less than 10.0% viral contigs linked with viruses from sediment, human body and other habitats.

Variation in host-specific features

The host spectra of viruses were estimated via matching CRISPR spacers in the IMG/VR database to viral contigs (**Fig.4**). A total of 1044 hits were obtained between 646 phage contigs and 279 bacterial genera. Following the typical parasite-host pattern, the matched 279 host genera were mainly from eight bacterial phyla (**Fig.4A & Fig.S7A**), which were comprised of the most abundant phyla in the bacterial community (**Fig.S7B**). For example, the phage contigs linking to *Proteobacteria* and *Actinobacteria*, the most abundant phyla (36.4% and 33.0%) in the bacterial community, accounted for 42.4% and 34.3% of the total phyla, respectively. Whereas most soil viruses (731 contigs) seemed to be specific with contigs linked to one specific bacterial genus, a considerable fraction of viruses (363 contigs) may have broad host ranges with phage contigs linked to more than two bacterial genera. The relative abundance of the possibly polyvalent phages (having links to two or more distinct genera), which had higher fitness under high environmental stresses and low bacterial abundance [34, 35], was higher in soils with higher Cr levels (Kruskal-Wallis (W-K) test, $p < 0.05$). Specifically, the fractions of polyvalent phages in LZ site increased from 26.3% in L1 to 33.3% in L3, and their fractions in ZY site increased from 21.3% in Z1 to 55.2% in Z4 (**Fig.4B**).

In the low and medium contaminated sites, the relative abundance of *Pseudomonas*, *Cronobacter*, *Klebsiella* and *Halomona* genera within the bacterial community increased with the available Cr concentrations in soil (**Fig.S8**). These four bacterial genera are characterized by their biofilm formation capabilities, low membrane permeability and strong efflux pump systems conducive to bacterial survival under heavy metal stress [36, 37]. Accordingly, the relative abundance of phages linking to these four genera increased from 3.6% -13.9% in the low contamination sites (L1, L2 and Z1) to 27.3- 40.3% in the medium contamination sites (L3 and Z2) (**Fig.4 & Fig.S8**). However, the dominant bacterial genera linked by phage in Z3 and Z4 with high Cr levels (413.8 and 465.4 mg/kg) were significantly different from those in the previous slightly and moderately contaminated sites. The dominant genera of Z3 included *Micropruina*, *Brevibacterium*, *Acidithiobacillus*, and *Zoogloea*, and that of Z4 included *Bacillus*, *Stenotrophomonas*, *Comamonas* and *Lysinibacillus*. These bacteria are often detected in heavy metal-contaminated environments because of their strong resilience and adaptivity to high levels of salts and

heavy metals [38-40]. Overall, the host-phage linkage corroborated phages as bacterial parasites depend heavily on the appropriate host flora to persist in Cr contaminated soils.

Variation in the relative abundance of lysogenic phage indicators (LPis)

The annotation of viral genomes was performed by comparing the predicted proteins against the PFAM database (33.1) to explore the lysogenicity of virome. Viral contigs with integrase genes (**Table S2**) were counted as lysogenic phages [41, 42]. The relative abundance of lysogenic phages increased with Cr levels in LZ site: L1 (7.0%) < L2 (7.5%) < L3 (35.1 %) (**Fig.5A**). For ZY, the relative abundance of lysogenic phage formed a bell-shaped curve as the Cr level increased. The relative abundance of lysogenic phages increased from Z1 (8.3%) to Z2 (23.1%), while decreased in Z3 (14.7%) and Z4 (13.3%) (**Fig.5A**). There was a significant increase in the relative abundance of viral integrase genes in bacterial genomes in response to the increased available Cr levels (W-K test, $p < 0.01$): L1 (0.1%) \approx L2 (0.1%) < L3 (0.3%) in LZ site, and Z1 (0.1%) < Z2 (0.4%) < Z3 (0.6%) < Z4 (0.7%) in ZY site (**Fig.5B**). Furthermore, more phages could be chemically induced from bacteria in more contaminated soil, which corroborated the host genome contained more prophage as Cr levels increased (W-K test, $p < 0.05$). In L1, L2 and Z1, the lysogenic fractions were all lower than 2 VLP/cell, while the lysogenic fractions increased to 5.45 and 4.14 VLP/cell in L3 and Z2, respectively. The corresponding fractions further increased to 19.14 and 22.84 VLP/cell in Z3 and Z4, respectively (**Fig.5C**).

Community-wide comparative genetic profiles of viromes

The annotation of viral genomes was performed by comparing the predicted proteins against the Non-supervised Orthologous Groups (eggNOG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Nonmetric multidimensional scaling (NMDS) analysis of viral genomes revealed that the viral contigs from different locations were more similar in functional composition when the spatially distinct locations had similar Cr levels (**Fig.S9**). Specifically, the low Cr level samples (i.e., L1, L2, and Z1) clustered together, medium Cr level samples (L3 and Z2) clustered together, while high Cr level samples (i.e., Z3 and Z4) but distinct from the other samples.

Moreover, viruses carried multiple AMGs related to host metabolism, and their relative abundance increased significantly with available Cr concentrations in slight and medium contamination sites (**Fig.S10**). For example, the relative abundance of genes related to energy production and conversion (Group C) and lipid transport and metabolism (Group I) increased within the virome with available Cr concentrations increased from 0.11 to 6.97 mg/kg. However, at severe pollution locations (Z3 and Z4), the relative abundance of these genes in virome decreased significantly compared to Z2. Similar trends were observed for genes related to inorganic ion transport and metabolism (Group P), secondary metabolites biosynthesis, transport and catabolism (Group Q), intracellular trafficking, secretion, and vesicular transport (Group U) and defense mechanisms (Group V). These genes are well recognized for mediating host transport and ion secretion. Upregulation of these genes is generally beneficial to microbial energy conversion under environmental stress, and to enhance detoxification processes (e.g., efflux and chromate reduction) [43, 44].

Bacterial Cr detoxification mainly involves efflux of intracellular Cr (VI) ions through membrane transporters and enzymatic reduction of Cr (VI) into less toxic Cr (III) [45, 46]. KEGG databases were used to screen functional genes corresponding to microbial heavy metal detoxification (i.e., membrane transporter gene (**Table S3**) and reductase gene (**Table S4**) in viromes. As shown in **Fig.6A**, the relative abundance of reductase genes in viromes increased from 0.03% to 3.8% and membrane transporter increased from 0.1% to 16.5% as the concentrations of available Cr increased from 0.11 to 6.76 mg/kg, which demonstrated that phages can be important reservoirs of bacterial heavy metal resistance genes (MRGs) in contaminated sites. However, the relative abundance of bacterial MRGs in Z3 and Z4 decreased to 0.2% and < 0.1%, respectively. Notably, over 77.0% of the genes encoding membrane transporters and 61.7% of reductases were located on lysogenic phages (**Fig.6B**), by investigating whether these genes are carried by viral contigs with the integrase genes. This suggests that lysogenic phages may serve as carriers of MRGs in highly contaminated sites.

Case study of lysogenic phages with MRGs.

To corroborate that putative lysogenic phage-encoded AMGs (especially MRGs), four viral contigs (VC-2, VC-16, VC-20, VC-22) with integrase gene were randomly selected from the viromes as representatives and their genomic profiles were systematically analyzed. All VCs carried prokaryotic genes associated with heavy metal detoxification (**Fig.6C**). Notably, viral contig VC-2 encoded genes including Arsenate reductase, Gutathione S-transferase, Ferric iron ABC transporter, Fe-S oxidoreductase, Heavy metal RND efflux outer membrane protein, Cobalt/zinc/cadmium efflux RND transporter, and Molybdenum ABC transporter. Viral contig VC-20 encoded Dehydrogenase oxidoreductase protein and Permease of the drug/metabolite transporter (DMT) superfamily. For VC-22, there were genes encoding Ferredoxin-NADPH reductases, Sulfonate ABC transporter and Reductase.

Discussion

Phages, being the most common biological entity on earth, are a great untapped resource with possible applications in medical, industrial and agricultural fields [47]. To identify the ecological and evolutionary role of phages in the biosphere and reveal the tactics of microbial adaptivity to hostile environments [48, 49]. We systematically investigated variations of the viral community profile and phage-bacteria interactions along horizontal gradients of Cr-induced environmental stresses at two contaminated sites. The abundance and composition of viral communities varied greatly alongside their host communities across contamination and geochemical gradients. As Cr-induced environmental stresses increased, phage communities exhibited an enhanced beneficial relationship with the bacterial communities that is reflected by elevated relative abundance of lysogenic phages in the viromes and higher lysogenic fractions of bacteria. Furthermore, viral genomic analysis showed that phages under higher Cr-induced stresses carried more AMGs that can contribute to microbial heavy metal detoxification and survival in hostile environments. These findings inform variations in interactions between phages and microbes under adverse conditions and advance our understanding of the contribution of phages in biogeochemical cycles.

As is common for heavy metal contamination sites, increased levels of Cr were associated with decreased nutrient availability at both site LZ and ZY (**Table S1**). Accordingly, the toxicity of Cr and the accompanying change in trophic conditions significantly altered both soil microbial and viral communities (**Fig.1 & Fig.2**). Notably, viral and bacterial communities from soil samples with similar Cr contamination levels rather than spatial proximity had more similarities in composition and abundance of functional genes (**Fig.S2, Fig.S5, & Fig.S9**). Interestingly, the viromes in Cr-contaminated soils were mostly not linked to known viruses recorded in the RefSeq database (**Fig.3**), possibly due to the unique physical and chemical properties of Cr-contaminated sites compared to other characterized habitats (uncontaminated soil, surface water, human gut and sediment) [50]. This suggests that viromes of metal-contaminated soils (and possibly other extreme environments) might represent an untapped gene reservoir of biological agents with unknown potential for medical and environmental applications.

Phages as obligate intracellular parasites rely on the presence of appropriate hosts to persist and propagate [18, 19], which is corroborated by a high degree of consistency between the predicted phage host and the dominant bacteria in soils (**Fig.4A, Fig.S7 & Fig.S8**). Relative abundance of bacterial genera (e.g., *Micropruina*, *Brevibacterium*, and *Bacillus*) with strong heavy metal resistance and survival capability under adverse conditions [51] increased in heavily Cr polluted environments; CRISPR-based host linkage data revealed a tandem increase in the relative abundance of phages dependent on these resistant hosts. From an evolutionary perspective, bacteria with stronger resistance will have higher biomass and more stable intracellular environments under adverse conditions, providing better survival and reproduction conditions for their corresponding phages [16]. Furthermore, broad host range phages can use multiple hosts to propagate are enriched in nutrient poor environments or in conditions of low host density [35]; this research corroborates this trend for the first time within a Cr contaminated area showing that the polyvalent strategy is viable to enhance phage fitness under severe environmental stress (**Fig.4B**) .

Furthermore, the capacity and tendency of phages towards a lysogenic lifestyle were generally enriched when faced with elevated levels of heavy metal pollution (**Fig.5**). This corroborated that lysogeny is a highly refined relationship whose development may be closely related to growth in unfavorable or even hostile conditions [23, 32, 52]. Both computational and experimental analyses demonstrated that the closer to the source of contamination, the higher abundance of prophages within the genomes of bacterial members of the community (**Fig.5B & C**). When faced with hostile environments, phages capable of entering lysogeny have shown a higher inclination to integrate into their host genome as a prophage, establishing a more mutualistic relationship [20].

The relative abundance of lysogenic phages within the free virome in highly contaminated samples (Z2, Z3, Z4, and L3) increased significantly above that of low contamination samples (Z1, L1 and L2) (**Fig.5A**). Interestingly, when examining the ratio of lysogenic to lytic phages in the free virome, the extremely contaminated samples (Z3 and Z4), had lower relative abundances of lysogenic phages, decreasing to nearly one-half of that for moderately contaminated Z2 at the same sampling site. To explain this drop, we posit that with further deterioration of the surrounding environment, lysogenic

phages may transform into prophages and lurk in the host body [53], which was corroborated by the increase of prophage in bacterial genome (**Fig.5B & C**). Therefore, the number of free but lysogenic capable phages in the ultra-harsh environment would decrease. These findings inform the potential preference of lysogenic viruses in more extreme environments, suggesting that the lysogeny should be favorable for phages under heavy metal pollution stress.

Previous studies reported that phage-carried AMGs can contribute to the genetic and phenotypic evolution of their hosts and ultimately provide the hosts with diversified competitiveness and environmental adaptability [20, 53]. In our study, many MRGs encoding machinery for the detoxification of heavy metals from bacteria were detected in the virome (**Fig.6A & Fig.6C**). Our analysis also highlights a consistent trend between the relative abundance of MGRs and lysogenic phages within the virome (**Fig.5A & Fig.6A**). That is mainly because most MGRs are distributed on lysogenic phages (**Fig.6B**). Therefore, when lysogenic phages integrate into host genomes in prophage form, the relative abundance of MGRs in the free virus particles also decreases. Lysogenic phages carrying AMGs (e.g., MRGs) suggest the potential of lysogenic phages to transfer beneficial genes in the microbial community and thus improve the survival ability of hosts in extreme environments. Therefore, there is strong tendency toward mutualism relationship between phage and their hosts based on lysogenic integration with the further aggravation of Cr pollution in soil. In this relationship, lysogenic bacteria provide more stable genetic environment for phages [24], in return for which they may benefit from the expression of phage-carried AMGs [54, 55]. Furthermore, the discovery of MRGs in virome highlights the influence of viral communities on biogeochemical process in contaminated sites, which may inspire phage-based strategies (e.g., virome transplant) for microbiome engineering [56].

Conclusion

This study of virome and metagenome interactions in soils subjected to various levels of stress associated with Cr contamination revealed that phage-bacteria interactions change from antagonistic (parasitic) at low stress towards a more protective mutualistic relationship with increasing Cr concentrations. This mutualistic symbiosis is mainly based on the lysogenic integration of lysogenic phages. That is, the lysogenic phages carrying MGRs can improve their host's ability to resist toxicity of heavy metals by integrating themselves into the host genome, which in turn facilitates expression of resistance genes and reproduction of related phages. This demonstrates that viral populations and their functions have a profound influence on the adaptation of prokaryotic communities and their ability to withstand adverse geochemical conditions. In addition, efforts to remediate and restore heavy-metal contaminated sites could benefit from advanced understanding of interactions between bacterial and phages and their manipulations to enhance redox processes that reduce metal toxicity and mobility.

Methods

Site description and sampling

Seven soil samples were collected from two Cr residue-contaminated sites, located in Zhangye City (ZY, N 38° 25' 49" E 100° 48' 41") with relatively high chromium pollution and Luzhou City (LZ, N 28° 54' 24", E 105° 31' 26") with light chromium pollution, in September 2019. Soil samples are collected at intervals of 50 meters, from the Cr-slag dump center towards the edge of the site. According to the site area, 3 soil samples (L1, L2 and L3) were collected in LZ and 4 soil samples (Z1, Z2, Z3 and Z4) were collected in ZY. Each soil sample was a manually homogenized composite of five subsamples collected at the surface layer (5-10 cm) and sieved through a size 20 mesh. The details of sampling locations are shown in **Fig.S1**. The physical and chemical properties (i.e., pH, organic matter, total phosphorus, total sulfur, total potassium, available phosphorus, nitrogen, nitrate nitrogen) of soil samples were determined by standard methods [57]. The total Cr, available Cr (i.e., the total amount of water-soluble and exchangeable Cr), hexavalent Cr concentrations were measured by ICP-MS (Agilent 7700x) according to the USEPA Method 7196A [58]. The concentrations of available Cr in the seven samples varied from 0.11 to 465 mg/kg (**Table S1**).

Quantification of bacteria and virus in soil

Fluorogenic quantitative PCR reactions was conducted with 16S rRNA gene V4-V5 primer pairs to determine the biomass of bacteria in soil samples from ZY and LZ (Primer sequences (5'-3') were 16S-515F: "GTGCCAGCMGCCGCGG" and 16S-907R: "CCGTCAATTCMTTTRAGTTT"). To determine the biomass of viruses in soil samples from ZY, LZ, the inverted fluorescence microscope (Nikon, eclipse Ti-S) was used to observe soil virus particles stained with SYBR Gold fluorescent dyes as previously described [59], and **Fig.S11** is a typical image, in which viruses are the most numerous tiny dots, while bacteria are larger with definite shapes and edges.

Total DNA extraction, metagenomic sequencing and analysis

FastDNA Spin Kit for Soil (MP Bio, Irvine, CA) was used to extract total DNA from the soil samples according to the manufacturer's instructions. DNA quantification was performed with a Qubit 3.0 fluorometer using the dsDNA system (Invitrogen, Waltham, Massachusetts). According to the Illumina TruSeq DNA Sample Preparation Guide, By using NEBNext Ultra II DNA Prep Kit (New England Biolabs, Ipswich, MA) and NovaSeq 6000 S4 Reagent Kit (300 cycles) (Illumina, San Diego, CA), genomic libraries were constructed and sequenced on Illumina Hiseq 4000 platform (2 x 150 bp, paired-end reads) at Shanghai Personal Biotechnology Co., Ltd. (Shanghai, China). Finally, 16 GB sequence data was obtained for each of the samples. The quality control of raw reads was conducted by Cutadapt (v1.17) and FASTP (0.20.0) [60], and high-quality reads were used to assemble sequences with MEGAHIT v1.0. [61] Minimap2 [62], Samtools [63], and blast2lca [64] were used to complete the construction of non-redundant sequence sets. After predicting the open reading frame (ORF) in sequences using MetaGeneAnnotator [65], redundant sequences were removed based on 90% sequence similarity by using CD-HIT [66]. Soap.coverage (<http://soap.genomics.org.cn/>) was used to calculate contig abundance and gene TPM (transcripts per million) value using a script (https://github.com/EnvGen/metagenomics-workshop/blob/master/in-house/tpm_table.py). The functional annotation of predicted ORFs was

performed through assigning predicted ORF sequences to the PFAM 33.1 [67], KEGG database [68], and eggNOG v5.0.0 with a cutoff of E-value < 0.001 [69]. Taxonomic classification at the species level was annotated via the NCBI-NT database (v2016-6-19) using Kaiju [70].

Virus particle elution and verification

The extraction of virus particles followed previous protocols (**Fig.S12**) [71]. Briefly, fresh soil sample was suspended in 4 °C potassium citrate buffer (10g/L C₆H₅K₃O₇, 1.92g/L Na₂HPO₄·12H₂O, and 0.24g/L KH₂PO₄, pH=7). After bath sonication in an ice-water mixture for 3 minutes at 47 kHz, the suspension was centrifuged (7,000 rpm, 4 °C, 10 min) to obtain the virus-containing supernatant. This was subsequently followed by centrifugation and filtration (0.2 µm) to remove impurities larger than 0.2 µm [71]. It should be noted that this filtration process may exclude viruses larger than 0.2 µm, leading them not to be included in further analysis of the viral population. The virus particles in the filtrate were initially concentrated by tangential flow filter cartridge (TFF, Sartorius Vivaflow 50 30000MWCO PES) followed by precipitation by polyethylene glycol 8000. The obtained virus pellets were treated with DNase and RNase to remove free DNA and RNA before viral DNA extraction [72]. In addition, virus particles were verified by transmission electron microscope (Zeiss LSM710) with the phosphotungstic acid counterstaining method (**Fig.S13**) [73]. It should be mentioned that the enrichment method of virome in this study is not perfect in the acquisition of giant phages.

Viral DNA extraction and metagenomic sequencing

Viral DNA was extracted using the AllPrep PowerViral DNA/RNA Kit (MP Bio) following the manufacturer's instructions. The extracted DNA was subjected to whole genome amplification (KAPA HiFi HotStart ReadyMix) to meet the metagenomic sequencing requirements. Libraries were constructed before sequencing on Illumina Hiseq4000 platform as described in "**Total DNA extraction, metagenomic sequencing and analysis**". The protocol of quality control and assembly of virome is also the same as above. It should be mentioned that the enrichment method of viral nucleic acid in this study is not perfect in the acquisition of RNA viral information.

Viral contig identification, taxonomic classification and functional gene annotation

Viral contigs > 5 Kb in length were recovered from the virome according to the "Virsorter+vHMM" method [18, 74]. Firstly, Virsorter was used, and the categories 1, 2, 4 and 5 of Virsorter predictions were considered as viral contigs. Secondly, the protein sequences of the remaining contigs were scanned with vHMM pipeline under 3 criteria (refer to

http://portal.nersc.gov/dna/microbial/prokpubs/EarthVirome_DP/): a, the contigs had at least 5 hits to viral protein families, and the total number of genes covered with KO terms on the contig was < 20%; the total number of genes covered with Pfams ≤ 40%; and the number of genes covered with viral protein families > 10%; b, the contigs were selected as viral contigs when the number of viral protein families on the contig were equal or higher than the number of Pfams (bacterial genes); c, the contigs for which the number of viral protein families was equal or higher than 60% of the total genes were classified as viral

contigs. Viral contigs must satisfy at least one of the above three criteria. Then, vcontact2 was used to annotate the taxa of these viral contigs [23]. The functional annotation of non-redundant proteins of viral contigs was performed through assignment of predicted proteins to the PFAM, KEGG and EggNOG databases as described above.

CRISPR-based phage-bacteria linkage

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) were used to predict phage-bacteria linkage based on the sequence similarity between the spacers in microbial CRISPR regions and the protospacers in viral genomes [23]. The sequences of all viral contigs were aligned to the spacer sequences deposited in the IMG/VR database (<https://img.jgi.doe.gov/vr/>). Specifically, all viral contigs and spacers sequences from the shared genomic regions were queried by BLASTn (e-value 10^{-10} and 100% nucleotide identity) to obtain the information about the potential host of the virus and finally link the virus to its bacterial host.

Prophage induction assay by mitomycin-C

Induction of prophages was performed as previously described [52]. Briefly, 1 gram of each of the original soil samples were suspended in 9ml sterilized deionized water, and shaken for 20 minutes at the maximum speed, 20,000 RPM. Then, these resuspensions were inoculated into sterile water-soluble organic carbon (WSOC) solution both with and without mitomycin-C (final concentration of 1 µg/ml). These cultures were incubated at room temperature, shaking at 200rpm for 24 hours under in the dark. To harvest the induced and free phages, the cultures were passed through a 0.2 µm sterile membrane. The filtrates were then treated with glutaraldehyde (final concentration was 2.5%) as a fixative and stored at 4°C for 24 hours prior to staining.

Staining and fluorescence counting were performed previously described as Anand, et al. [59]. After counting, the number of phages induced from a single bacterium was calculated using the following formula:

$$X = (V_i - V_{ck})/B$$

Here, V_i is the abundance of viral like particles in the treatment with mitomycin-C, V_c is the abundance of viral like particles in control treatments, and B is the bacterial abundance in the corresponding soil sample (by fluorogenic quantitative PCR reactions).

Data analysis

The data analysis and visualization process involved in this study were conducted by Excel 2013 and R x64 3.6.2. In order to clarify the correlation of viral contigs recovered in Cr-contaminated soil and other known virus, the co-occurrence network analysis was conducted between the acquired virus contigs and known bacterial and archaeal viruses (NCBI RefSeq, version 75) using Cytoscape3.7.1 (<http://cytoscape.org>) [33]. The phylogenetic tree and heat maps of dominant bacteria (relative abundance greater than 0.1%) were produced using MEGA7 and iTOL (<https://itol.embl.de/>). The RDAs of viruses and bacteria were performed using R packages ggbiplot to analyze the viral and bacterial composition and diversity of different samples. In addition, functional genes annotated in KEGG databases corresponding to microbial heavy metal detoxification (i.e., membrane transporter gene (**Table S3**) and reductase gene (**Table S4**)) in the metagenome and virome were isolated to determine the detoxification potential of the microbial community. The differences between genetic profiles of viral communities were assessed by NMDS analysis utilizing R package vegan [75]. Based on PFAM databases, phage integrase family genes (as listed in **Table S2**) in the metagenome and virome were screened out to investigate changes in the relative abundance of lysogenic phages [29]. The statistical significance of differences at different Cr-contamination levels (i.e., the lysogenicity fraction of bacterial community, the relative abundance of MRGs, lysogenic phages and polyvalent phages among soils with different pollution levels) were determined at the 95% confidence level ($p < 0.05$) using the non-parametric Kruskal-Wallis test (K-W test) [32].

Abbreviations

Cr: chromium; AMG: auxiliary metabolic gene; MRG: heavy metal resistance gene; VC: Viral cluster; NCBI: National Center for Biotechnology Information; eggNOG: Evolutionary genealogy of genes: Non-supervised Orthologous Groups; KEGG: Kyoto Encyclopedia of Genes and Genomes; RDA: redundant analysis; NMDS: nonmetric multidimensional scaling; PCR: Polymerase chain reaction; WSOC: water-soluble organic carbon.

Declarations

Acknowledgements

We thank Dr. Jinfeng Wang from China Academy of Sciences for experimental design.

Authors' contributions

All authors contributed intellectually to and agreed to this submission. HD, YPF and YM designed the experiments. HD conducted the experiments and collected the data. HD, YPF and YM analyzed the data. HD, YPF, YM and CS wrote the initial draft of the manuscript while XJ and PA provided substantial feedback.

Funding

This work was financially supported by National Key Research and Development Program of China (2018YFC1803100), the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2018350) and NSF ERC on Nanotechnology-Enabled Water Treatment (EEC-1449500).

Availability of data and materials

All raw sequence data generated in this research has been deposited in the Genome Sequence Archive (GSA) in BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number CRA003088 and is publicly accessible at <https://bigd.big.ac.cn/gsa>.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare no conflict of interest

References

1. Martha RJC, Andrew DM, Andrey VL, Shaun H. Phages in nature, *Bacteriophage*. 2011;1(1):31-45.
2. Wang YL, Jiang XT, Liu L, Li B, Zhang T. High-Resolution Temporal and Spatial Patterns of Virome in Wastewater Treatment Systems. *Environ Sci Technol*. 2018;52(18):10337-10346.
3. Petrovich ML, Zilberman A, Kaplan A, Eliraz GR, Wang YB, Langenfeld K, Duhaime M, Wigginton K, Poretsky R, Avisar D, Wells GF. Microbial and Viral Communities and Their Antibiotic Resistance Genes Throughout a Hospital Wastewater Treatment System. *Front Microbiol*. 2020;11:13.
4. Cao B, Xu H, Mao CB. Phage as a template to grow bone mineral nanocrystals. *Methods Mol Biol*. 2014;1108:123-35.
5. Kimura M, Jia ZJ, Nakayama N, Asakawa S. Ecology of viruses in soils: Past present and future perspectives. *J. Soil Sci. Plant Nut*. 2008;54(1):1-32.
6. Suttle Curtis A. Viruses in the sea. *Nature*. 2005;437(7057):356-361.
7. Motlagh AM, Bhattacharjee AS, Coutinho FH, Dutilh BE, Casjens SR, Goel RK. Insights of Phage-Host Interaction in Hypersaline Ecosystem through Metagenomics Analyses. *Front Microbiol*. 2017;8:352.
8. Manrique P, Dills M, Young MJ. The Human Gut Phage Community and Its Implications for Health and Diseases Viruses-Basel. 2017;9(6):141.
9. Pan D, Watson R, Wang D, Tan ZH, Snow DD, Weber K. Correlation between viral production and carbon mineralization under nitrate-reducing conditions in aquifer sediment. *Isme J*. 2014;8(8):1691-

1703.

10. Nair RR, Vasse M, Wielgoss S, Sun L, Yu YTN, Velicer GJ. Bacterial predator-prey coevolution accelerates genome evolution and selects on virulence-associated prey defences. *Nat Commun.* 2019;10:4301.
11. Kotay SM, Datta T, Choi JD, Goel R. Biocontrol of biomass bulking caused by *Haliscomenobacter hydrossis* using a newly isolated lytic bacteriophage. *Water Res.* 2011;45(2):694-704.
12. Calero-Caceres W, Ye M, Balcazar JL. Bacteriophages as Environmental Reservoirs of Antibiotic Resistance. *Trends Microbiol.* 2019;27(7):570-577.
13. Bhattacharjee AS, Choi J, Motlagh AM, Mukherji ST, Goel R. Bacteriophage therapy for membrane biofouling in membrane bioreactors and antibiotic-resistant bacterial biofilms. *Biotechnol Bioeng.* 2015;112(8):1644-1654.
14. Shkorporov AN, Hill C. Bacteriophages of the Human Gut: The “Known Unknown” of the Microbiome. *Cell Host Microbe.* 2019;25(2):195-209.
15. Knowles B, Silveira CB, Bailey BA, Barott K, Cantu VA, Cobian-Guemes AG, Coutinho FH, Dinsdale EA, Felts B, Furby KA, George EE, Green KT, Gregoracci GB, Haas AF, Haggerty JM, Hester ER, Hisakawa N, Kelly LW, Lim YW, Little M, Luque A, McDole-Somera T, McNair K, de Oliveira LS, Quistad SD, Robinett NL, Sala E, Salamon P, Sanchez SE, Sandin S, Silva GGZ, Smith J, Sullivan C, Thompson C, Vermeij MJA, Youle M, Young C, Zgliczynski B, Brainard R, Edwards RA, Nulton J, Thompson F, Rohwer F. Lytic to temperate switching of viral communities. *Nature.* 2016;531(7595):466-470.
16. Shapiro OH, Kushmaro A, Brenner A. Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater. *Isme J.* 2010;4(3):327-336.
17. Knowles B, Rohwer F. Lysis, lysogeny and virus-microbe ratios Reply. *Nature.* 2017;549(7672):E3-E4.
18. Paez-Espino D, Eloë-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, Rubin E, Ivanova NN, Kyrpides NC. Uncovering Earth’s Virome. *Nature.* 2016;536(7617):425-430.
19. Rodriguez-Brito B, Li LL, Wegley L, Furlan M, Angly F, Breitbart M, Buchanan J, Desnues C, Dinsdale E, Edwards R, Felts B, Haynes M, Liu H, Lipson D, Mahaffy J, Martin-Cuadrado AB, Mira A, Nulton J, Pasic L, Rayhawk S, Rodriguez-Mueller J, Rodriguez-Valera F, Salamon P, Srinagesh S, Thingstad TF, Tran T, Thurber RV, Willner D, Youle M, Rohwer F. Viral and microbial community dynamics in four aquatic environments. *Isme J.* 2010;4(6):739-751.
20. Warwick-Dugdale J, Buchholz HH, Allen MJ, Temperton B. Host-hijacking and planktonic piracy: how phages command the microbial high seas. *Virol J.* 2019;16:13.
21. Scanlan PD, Hall AR, Blackshields G, Friman VP, Davis MR, Goldberg J, Joanna B, Buckling A. Coevolution with Bacteriophages Drives Genome-Wide Host Evolution and Constrains the Acquisition of Abiotic-Beneficial Mutations. *Mol Biol Evol.* 2015;32(6):1425-1435.
22. Roossinck MJ. The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol.* 2011;9(2):99-108.
23. Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB, Singleton CM, Soden LM, Naas AE, Boyd JA, Hodgkins SB, Wilson RM, Trubl G, Li CS, Frokings S, Pope PB, Wrighton KC, Crill PM,

- Chanton JP, Saleska SR, Tyson GW, Rich VI, Sullivan MB. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat Microbiol.* 2018;3(8):870-880.
24. Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. Lysogeny in nature: mechanisms impact and ecology of temperate phages. *Isme J.* 2017;11(7):1511-1520.
 25. Zhang JY, Gao Q, Zhang QT, Wang TX, Yue HW, Wu LW, Shi J, Qin ZY, Zhou JZ, Zuo JE, Yang YF. Bacteriophage-prokaryote dynamics and interaction within anaerobic digestion processes across time and space. *Microbiome.* 2017;5:10.
 26. Walker CB, Stolyar S, Chivian D, Pinel N, Gabster JA, Dehal PS, He ZL, Yang ZK, Yen HCB, Zhou JZ, Wall JD, Hazen TC, Arkin AP, Stahl DA. Contribution of mobile genetic elements to *Desulfovibrio vulgaris* genome plasticity. *Environ Microbiol.* 2009;11(9):2244-2252.
 27. Maslov S, Sneppen K. Well-temperate phage: optimal bet-hedging against local environmental collapses. *Sci Rep-UK.* 2015;5:10523.
 28. Wu G, Kang HB, Zhang XY, Shao HB, Chu LY, Ruan CJ. A critical review on the bio-removal of hazardous heavy metals from contaminated soils: Issues progress eco-environmental concerns and opportunities. *J Hazard Mater.* 2010;174(1-3):1-8.
 29. Liu XY, Bai ZK, Shi HD, Zhou W, Liu XC. Heavy metal pollution of soils from coal mines in China. *Nat Hazards.* 2019;99(2):1163-1177.
 30. Zou YD, Wang XX, Khan A, Wang PY, Liu YH, Alsaedi A, Hayat T, Wang XK. Environmental Remediation and Application of Nanoscale Zero-Valent Iron and Its Composites for the Removal of Heavy Metal Ions: A Review. *Environ Sci Technol.* 2016;50(14):7290-7304.
 31. Ma Y, Zhong H, He ZG. Cr (VI) reductase activity locates in the cytoplasm of *Aeribacillus pallidus* BK1 a novel Cr(VI)-reducing thermophile isolated from Tengchong geothermal region China. *Chem Eng J.* 2019;371:524-534.
 32. Gao SM, Schippers A, Chen N, Yuan Y, Zhang MM, Li Q, Liao B, Shu WS, Huang LN. Depth-related variability in viral communities in highly stratified sulfidic mine tailings. *Microbiome.* 2020;8(1):89.
 33. Trubl G, Jang HB, Roux S, Emerson JB, Solonenko N, Vik DR, Solden L, Ellenbogen J, Runyon AT, Bolduc B, Woodcroft BJ, Saleska SR, Tyson GW, Wrighton KC, Sullivan MB, Rich VI. Soil Viruses Are Underexplored Players in Ecosystem Carbon Processing. *Msystems.* 2018;3(5):21.
 34. Yu PF, Mathieu J, Yang Y, Alvarez PJJ. Suppression of Enteric Bacteria by Bacteriophages: Importance of Phage Polyvalence in the Presence of Soil Bacteria. *Environ Sci Technol.* 2017;51(9):5270-5278.
 35. Yu PF, Mathieu J, Lu GW, Gabiatti N, Alvarez PJ. Control of Antibiotic-Resistant Bacteria in Activated Sludge Using Polyvalent Phages in Conjunction with a Production Host. *Environ Sci Tech Let.* 2017;4(4):137-142.
 36. Yang JX, Wei W, Pi SS, Ma F, Li A, Wu D, Xing J. Competitive adsorption of heavy metals by extracellular polymeric substances extracted from *Klebsiella sp J1*. *Bioresource Technol.* 2015;196:533-539.

37. Dogan NM, Kantar C, Gulcan S, Dodge CJ, Yimaz BC, Mazmanci MA. Chromium (VI) Bioremoval by Pseudomonas Bacteria: Role of Microbial Exudates for Natural Attenuation and Biotreatment of Cr(VI) Contamination. *Environ Sci Technol*. 2011;45(6):2278-2285.
38. Ng TW, Cai QH, Wong CK, Chow AT, Wong PK. Simultaneous chromate reduction and azo dye decolourization by *Brevibacterium casei*: Azo dye as electron donor for chromate reduction. *J Hazard Mater*. 2010;182(1-3):792-800.
39. Anderson C, Jakobsson AM, Pedersen K. Influence of in situ biofilm coverage on the radionuclide adsorption capacity of subsurface granite. *Environ Sci Technol*. 2007;41(3):830-836.
40. Kumar R, Singh R, Kumar N, Bishnoi K, Bishnoi NR. Response surface methodology approach for optimization of biosorption process for removal of Cr (VI), Ni (II) and Zn (II) ions by immobilized bacterial biomass *sp Bacillus brevis*. *Chem Eng J*. 2009;146(3):401-407.
41. Lamont I, Richardson H, Carter DR, Egan JB. Genes for the establishment and maintenance of lysogeny by the temperate coliphage 186. *J Bacteriol*. 1993;175(16):5286-5288.
42. Villafane R, Black J. Identification of 4 genes involved in the lysogenic pathway of the *salmonella Newington* bacterial-virus epsilon(34). *Arch Virol*. 1994;135(1-2):179-183.
43. Yin H, He BY, Peng H, Ye JH, Yang F, Zhang N. Removal of Cr(VI) and Ni(II) from aqueous solution by fused yeast: Study of cations release and biosorption mechanism. *J Hazard Mater*. 2008;158(2-3):568-576.
44. Nies DH. Efflux-mediated heavy metal resistance in prokaryotes. *Fems Microbiol Rev*. 2003;27(2-3):313-339.
45. Miao Y, Liao RH, Zhang XX, Wang Y, Wang Z, Shi P, Liu B, Li AM. Metagenomic insights into Cr (VI) effect on microbial communities and functional genes of an expanded granular sludge bed reactor treating high-nitrate wastewater. *Water Res*. 2015;76:43-52.
46. Malaviya P, Singh A. Metagenomic insights into Cr (VI) effect on microbial communities and functional genes of an expanded granular sludge bed reactor treating high-nitrate wastewater. *Water Res*. 2015;76:43-52.
47. Martinez-Garcia M, Santos F, Moreno-Paz M, Parro V, Anton JL. Unveiling viral-host interactions within the 'microbial dark matter'. *Nat Commun*. 2014;5:8.
48. Forterre P. The virocell concept and environmental microbiology. *Isme J*. 2013;7(2):233-236.
49. Breitbart M, Bonnain C, Malki K, Sawaya NA. Phage puppet masters of the marine microbial realm. *Nat Microbiol*. 2018;3(7):754-766.
50. Dhal B, Thatoi HN, Das NN, Pandey BD. Chemical and microbial remediation of hexavalent chromium from contaminated soil and mining/metallurgical solid waste: A review. *J Hazard Mater*. 2013;250:272-291.
51. Liu YR, Delgado-Baquerizo M, Bi L, Zhu J, He JZ. Consistent responses of soil microbial taxonomic and functional attributes to mercury pollution across China. *Microbiome*. 2018;6:12.

52. Liang XL, Zhang YY, Wommack KE, Wilhelm SW, DeBruyn JM, Sherfy AC, Zhuang J, Radosevich M. Lysogenic reproductive strategies of viral communities vary with soil depth and are correlated with bacterial diversity. *Soil Biol Biochem.* 2020;144:10.
53. Touchon M, de Sousa JAM, Rocha EPC. Rocha Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr Opin Microbiol.* 2017;38:66-73.
54. Wang XX, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, Wood TK. Cryptic prophages help bacteria cope with adverse environments. *Nat Commun.* 2010;1:9.
55. Moon K, Jeon JH, Kang IN, Park KS, Lee KY, Cha CJ, Lee SH, Cho JC. Freshwater viral metagenome reveals novel and functional phage-borne antibiotic resistance genes. *Microbiome.* 2020;8(1): 5.
56. Braga LPP, Spor A, Kot W, Breuil MC, Hansen LH, Setubal JC, Philippot L. Freshwater viral metagenome reveals novel and functional phage-borne antibiotic resistance genes. *Microbiome.* 2020;8(1):15.
57. Rowell DL. *Soil Science: Methods & Applications*; Longman Scientific & Technical, Harlow, 1994.
58. S. EPA. *Alkaline Digestion for Hexavalent Chromium (EPA Method 3060a)*. U.S. EPA.1996
59. 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(2):69-276.
60. Chen SF, Zhou YQ, Chen YR, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):884-890.
61. Li DH, Luo RB, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. MEGAHIT v1, 0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods.* 2016;102:3-11.
62. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics.* 2018;34(18):3094-3100.
63. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools, *Bioinformatics.* 2009;25(16):2078-2079.
64. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *Genome Res.* 2007;17(3):377-386.
65. Noguchi H, Taniguchi T, Itoh T. Meta Gene Annotator: Detecting Species-Specific Patterns of Ribosomal Binding Site for Precise Gene Prediction in Anonymous Prokaryotic and Phage Genomes. *DNA Res.* 2008;15(6):387-396.
66. Li WZ, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006;22(13):1658-1659.
67. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD. The Pfam protein families database. *Nucleic Acids Res.* 2012;40(D1):D290-D301.

68. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J Mol Biol.* 2016; 428(4):726-731.
69. Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. eggNOG 5.0: a hierarchical functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 2019;47(D1): D309-D314.
70. Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun.* 2016;7:9.
71. Adriaenssens EM, Kramer R, Van Goethem MW, Makhalanyaane TP, Hogg I, Cowan DA. Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome.* 2017;5:14.
72. Yu DT, He JZ, Zhang LM, Han LL. Viral metagenomics analysis and eight novel viral genomes identified from the Dushanzi mud volcanic soil in Xinjiang China. *J Soil Sediment.* 2019;19(1):81-90.
73. Fuhrman JA. Marine viruses and their biogeochemical and ecological effects. 1999;399(6736):541-548.
74. Roux S, Hallam SJ, Woyke T, Sullivan MB. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *Elife.* 2015;4:20.
75. De Sanctis M, Fanelli G, Gjeta E, Mullaj A, Attorre F. The forest communities of Shebenik-Jabllanice National Park (Central Albania). *Phytocoenologia.* 2018;48(1):51-76.

Figures

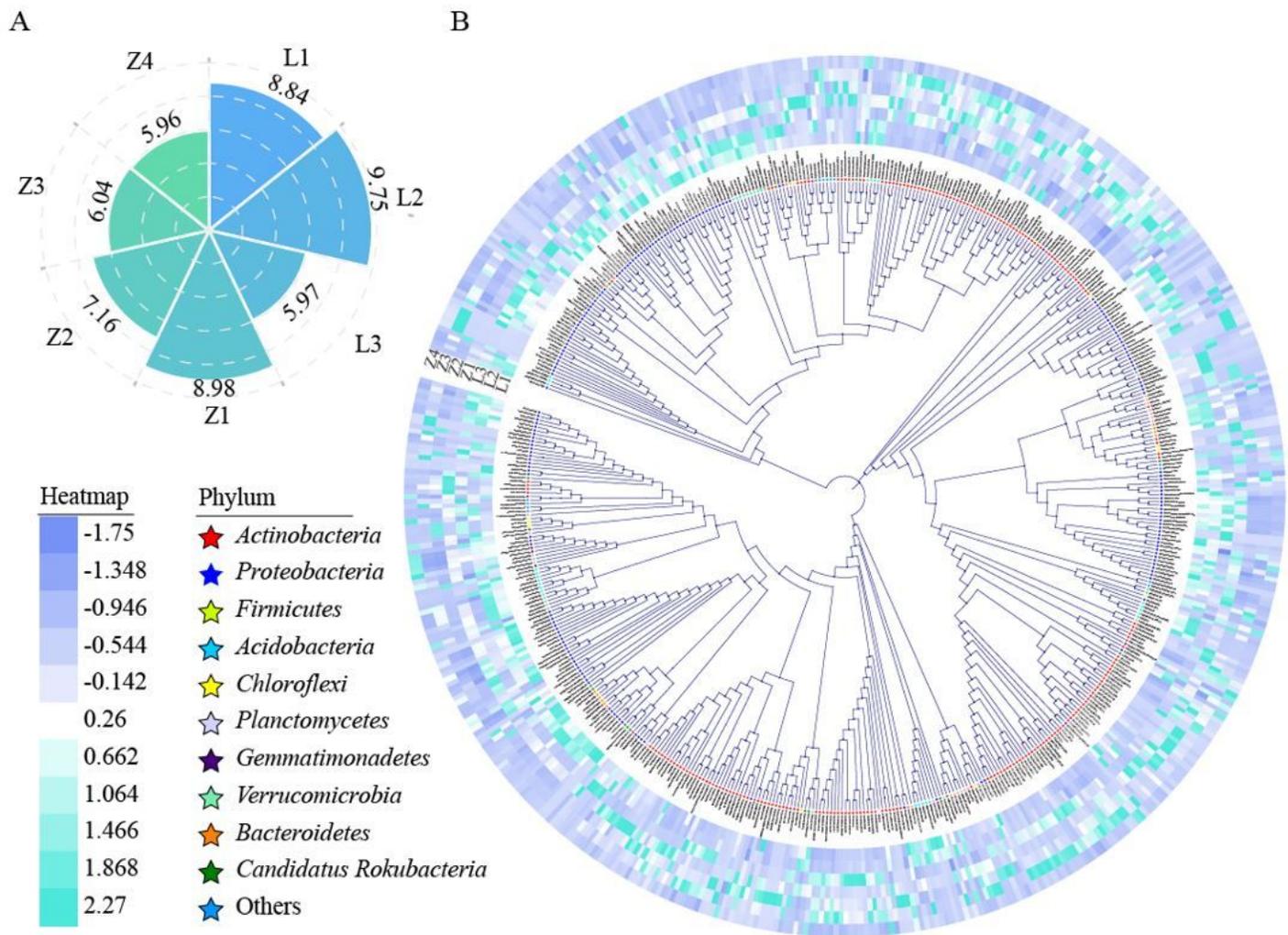


Figure 1

Profile of bacterial communities in Cr-contaminated soils. (A) The biomass of bacteria in Cr-contaminated soil. The petal lengths of the rose represent the value of log (copies/g soil), which generally decreased with the increase of Cr contamination in both sites. (B) The phylogenetic tree of dominant bacterial species with relative abundance greater than 0.1% and heatmap of their relative abundance. Metagenomic analysis of dominant bacterial species and beta diversity revealed a significant difference in the bacterial communities of soils with different levels of Cr pollution.

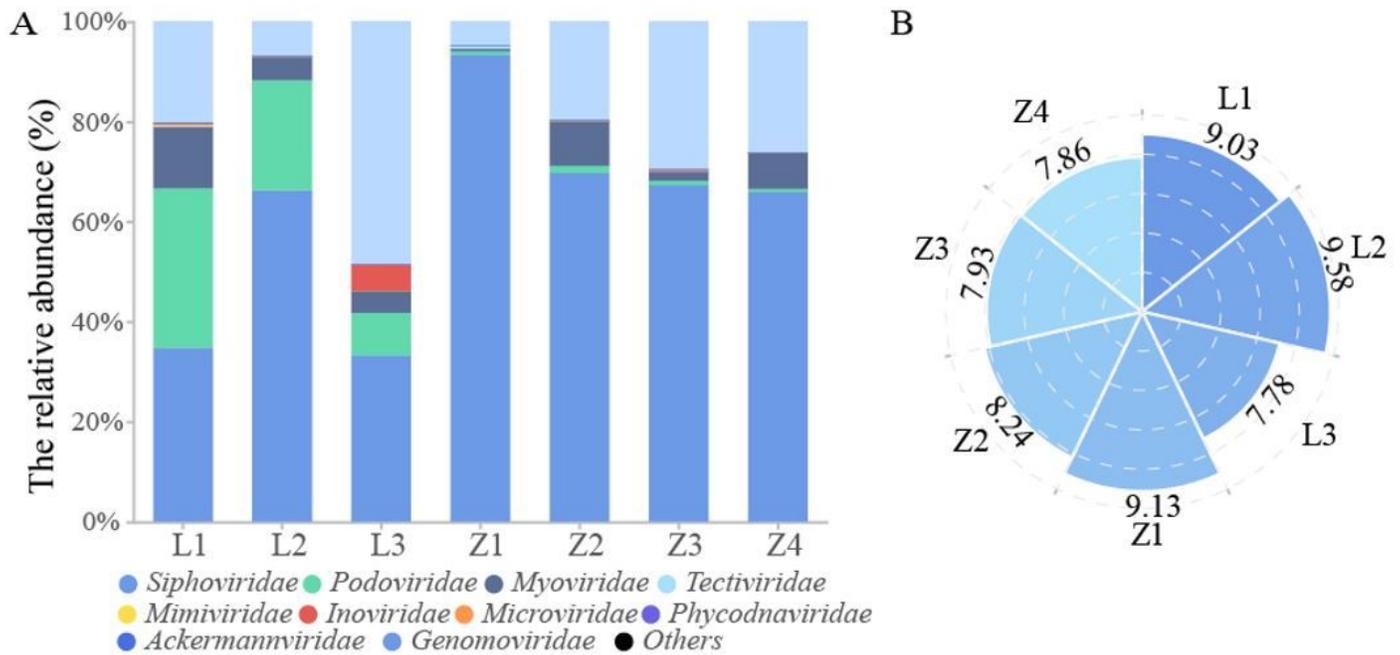


Figure 2

Profile of viral communities in Cr-contaminated soils. (A) The composition of the viral community at the family level from seven different locations at two Cr-contaminated sites shows that viral populations between these 7 different locations exhibited distinct profiles. (B) The abundance of viral particles in Cr-contaminated soil counted by fluorescence microscope. The petal lengths of the rose represent the abundance of viral particles per gram of soil (log (VLP/g soil)), which decreased with the increase of Cr contamination in both sites.

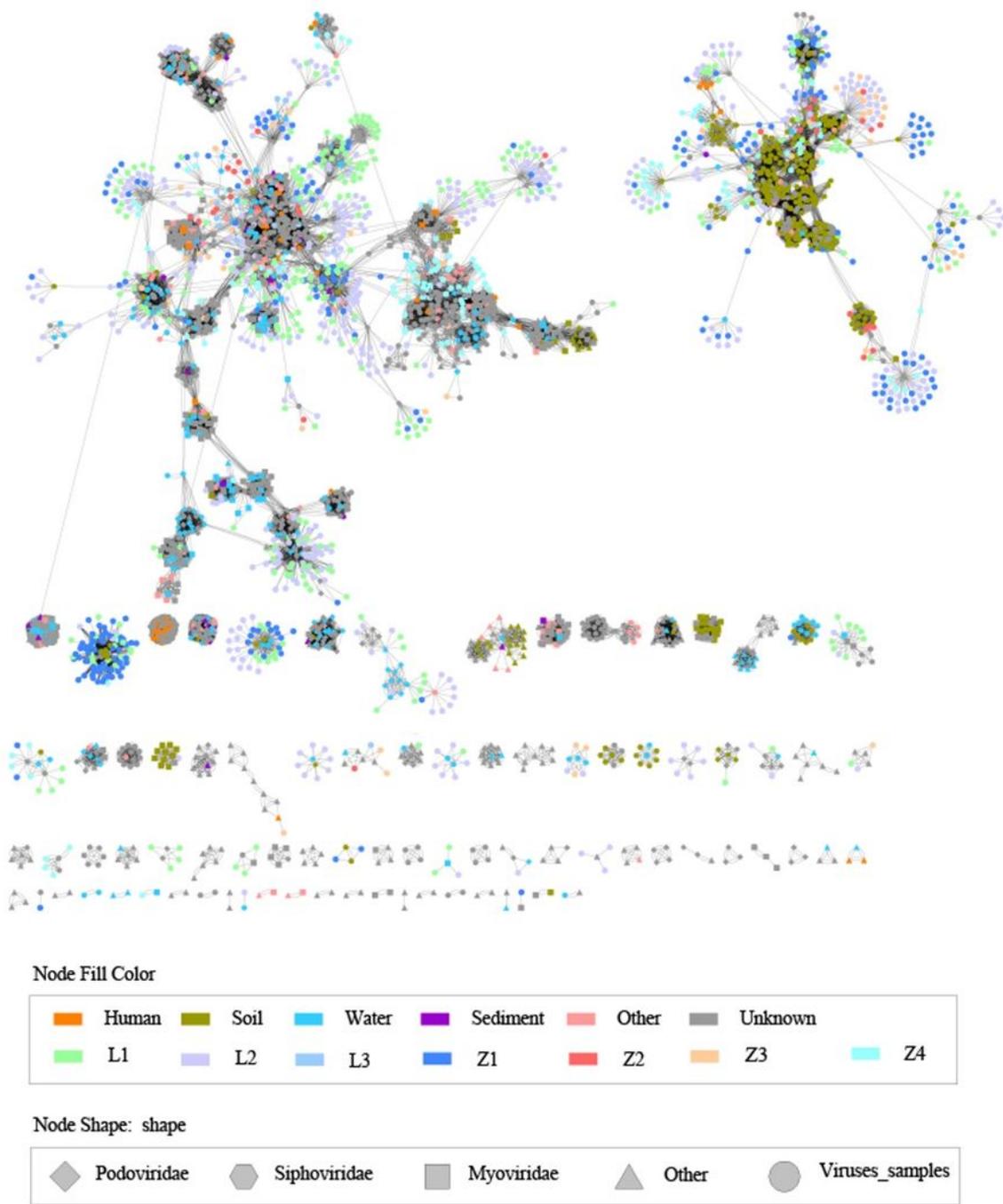


Figure 3

Relating viral contigs from Cr-contaminated soils to known viral sequence space. Viral contigs recovered from Cr-contaminated soils were clustered with all RefSeq (v 75) viral genomes or genome fragments with genetic connectivity to these data. Each node colors indicate sample and reference viral sequences' habitat of origin. The edges (lines) between the nodes represent statistically weighted pairwise similarity scores of ≥ 1 . The results show that viromes at different sites along a Cr contaminated soil gradient have

distinct specificity. And only a few of recovered viral contigs were associated with viruses recorded in NCBI databases (RefSeq, v 75) from soil, water, human and sediment habitats.

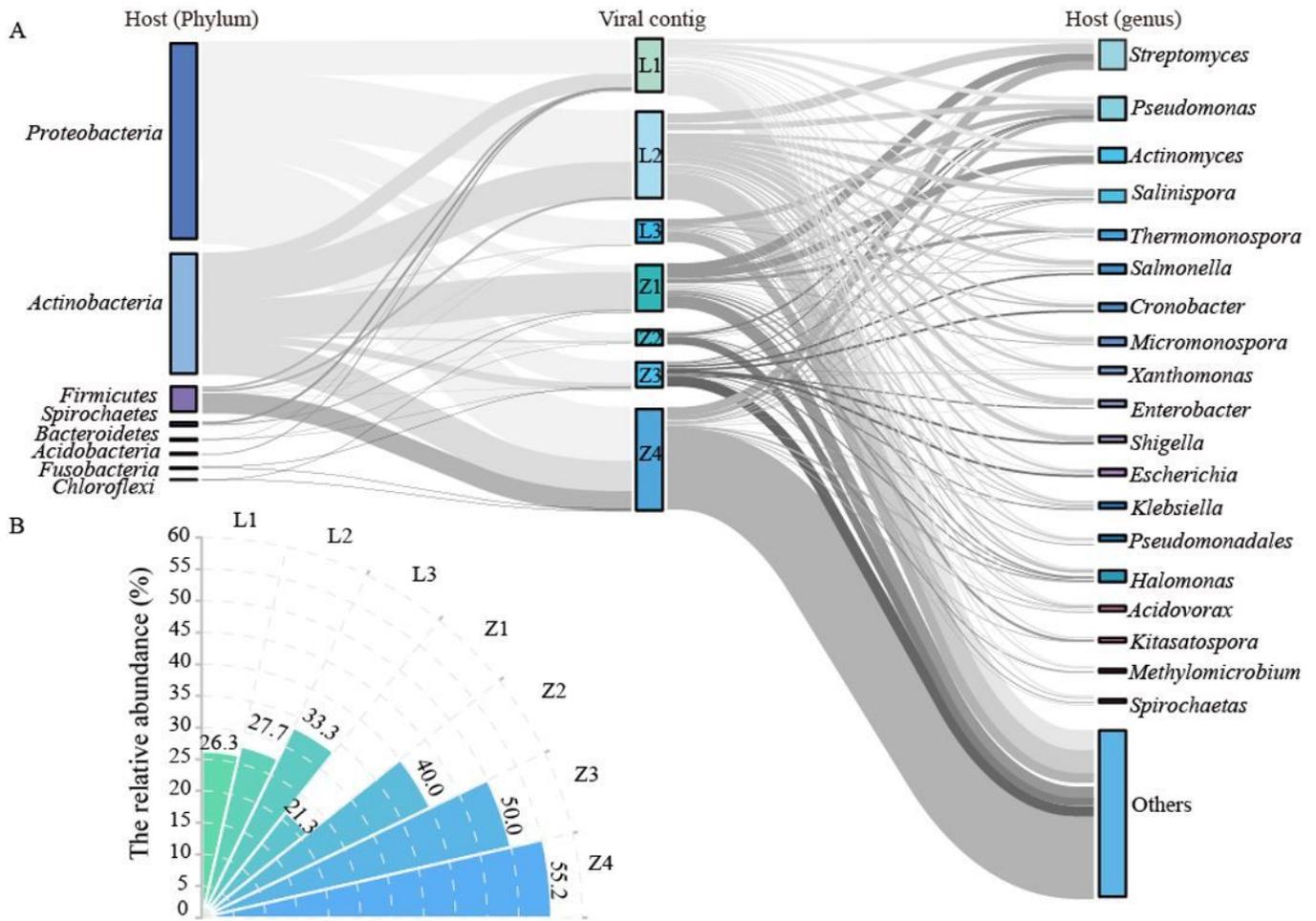


Figure 4

Predicted virus-host linkages in Cr-contaminated soils. (A) A total of 1044 hits were obtained between 646 phage contigs and 279 bacterial genera, and these host genera were mainly from 8 bacterial phyla. From left to right, and the first columns show the distribution of bacterial hosts at the phylum level; the second column depicts viral contigs linking to hosts in the seven soil samples; the third column shows the distribution of the bacterial host assigned with CRISPR approaches at the genus level. The height of the column section represents the abundance of viruses and hosts. The results revealed that these viral host species predicted in the soil polluted by high Cr concentration mostly belong to the bacterial genera which have been reported to have strong adaptivity to better survive in heavy metal-polluted habitats. (B) The relative abundance of polyvalent phages within the viral communities of Cr-contaminated soils. Polyvalence was found to be consistently higher in soil samples with higher available Cr concentrations.

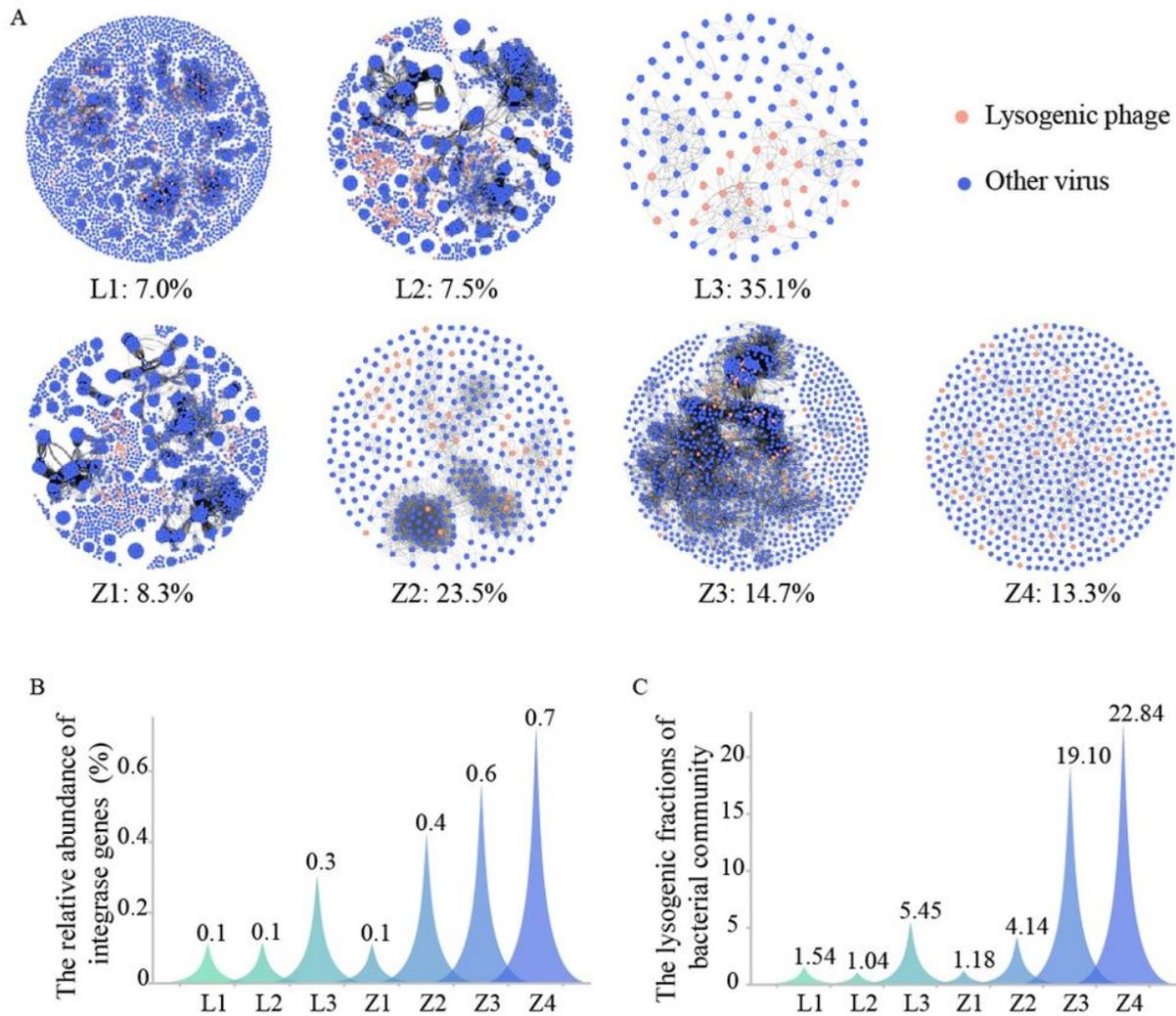


Figure 5

Lysogen occurrence within the soil virome along Cr concentration gradients. (A) Co-occurrence networks of viral contigs in seven soil samples. Nodes (circles) represent viral contigs. Orange represent the contigs containing lysogenic phage indicators; dark blue nodes are all other viral contigs. Shared edges (lines) between nodes indicate a correlation, and the more edges in a node indicate a closer genetic relationship between contigs. The values to the right of each sample name are the relative abundances of lysogenic phages within the corresponding sample. As the distance between sampling location and source of contamination grew, the relative abundance of lysogenic phages decreased in the lightly polluted LZ site. In the heavily polluted ZY site, an increasing trend from sample Z1 to Z2 was observed, which then showed a decreasing trend between Z2 and Z4. The relative abundance of phage integrase genes found

in bacteria (B) and inducible lysogenic fractions (C) at the sampling locations consistently increased in tandem with Cr-concentrations.

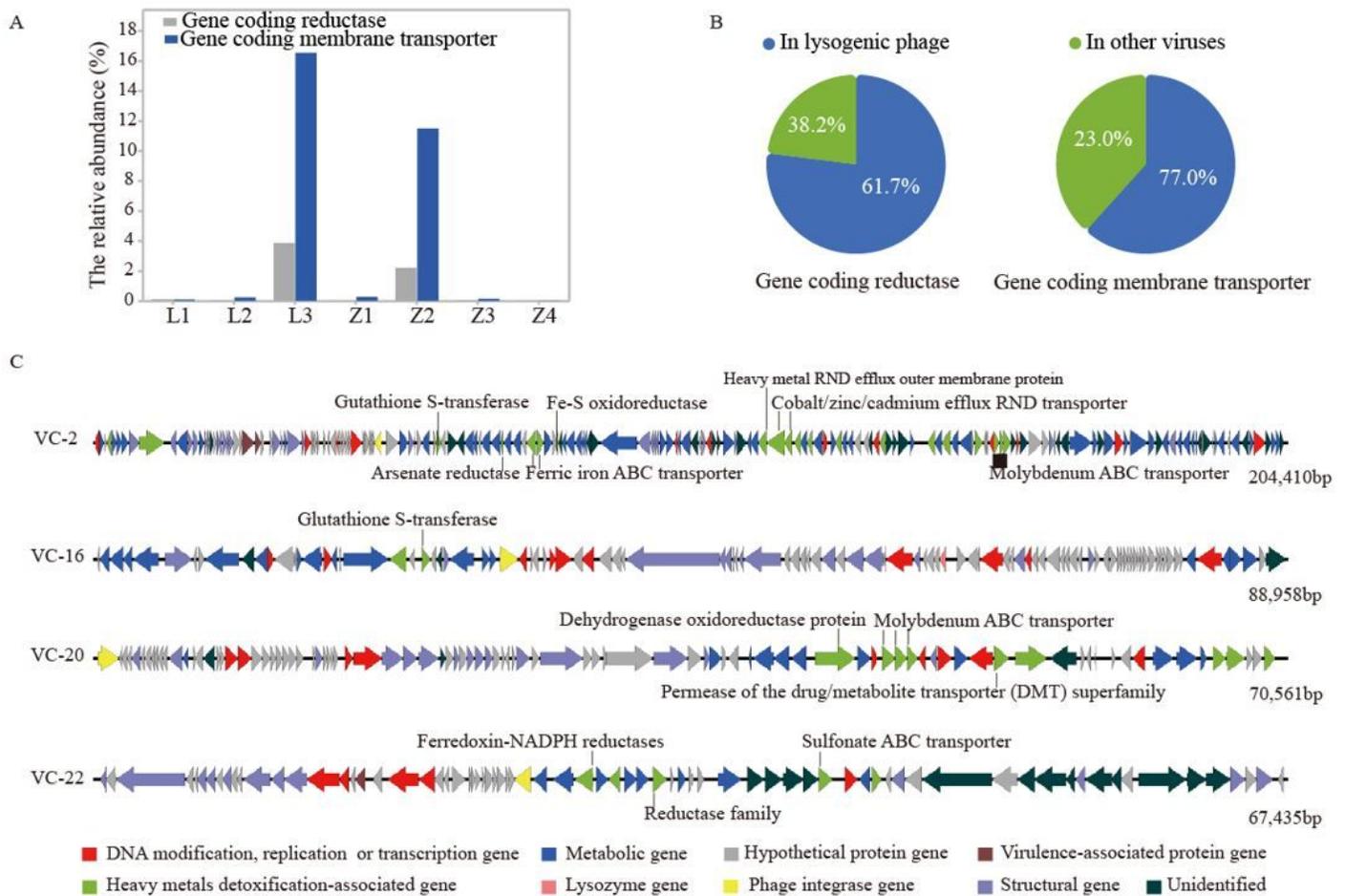


Figure 6

Genomic analysis of viral genomes in Cr-contaminated soils. (A) The relative abundance of viral genes coding membrane transporter (dark blue) and reductase (gray) were annotated by KEGG datasets (Table S3 and Table S4). There was a significant increase in the relative abundance of these genes associated with reduction and transport of heavy metals in L3 and Z2. (B) The distribution ratio of genes coding reductase (left) and membrane transporter (right) on lysogenic phages (blue) and other viruses (green), which were are calculated by determining whether these genes are carried by viral contigs with the integrase genes. It shows that over 60% of them were on lysogenic phages, so lysogenic phages are an important reservoir for MRGs. (C) Genome map of four viral contigs containing both the phage integrase gene and the metal detoxification gene. Key of gene families and their colors are found in the figure, and arrow length within the gene maps correspond to the size of the open reading frame (ORF). Detailed function descriptions of the seven viral contigs are listed in Table S5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)
- [Additionalfile2.docx](#)