Chemical Engineering Journal 510 (2025) 161877

Contents lists available at ScienceDirect

Chemical Engineering Journal

journal homepage: www.elsevier.com/locate/cej

Plasticizers determine a deeper reshape of soil virome than microplastics

Mengyuan Ji [a], Laura Treu [a,*], Stefano Campanaro [a]

[a] Department of Biology (DIBIO), University of Padova, Via Ugo Bassi 58/B, 35131 Padova, Italy

ARTICLE INFO

Keywords:
Plastic pollution
Virome analysis
Auxiliary metabolic genes
3-oxoadipate enol lactonase
Assisted degradation

ABSTRACT

Viruses play a crucial role in shaping local and global biogeochemical cycles, supporting bacterial survival in diverse environments by encoding auxiliary metabolic genes involved in energy acquisition, stress tolerance, and the degradation of organics. However, how plastic pollution influences soil viromes remains largely unexplored, in particular when microplastics and plasticizers are involved in the process. In this study, we conducted an incubation experiment where soil samples from rice fields were exposed to microplastics—polyethylene and polyvinyl chloride and the plasticizer diethyl phthalate to assess their effects on viral communities. After controlled incubation, second- and third-generation sequencing, along with advanced bioinformatics, were used to determine whether viral taxa were impacted by these contaminants. Our results revealed that diethyl phthalate exposure led to a 3.15-fold increase in the proportion of viral sequences in the treated samples compared to control soils, significantly surpassing the modest increases observed for polyethylene (13.08%) and polyvinyl chloride (48.59%). These shifts were accompanied by changes in viral diversity, functional gene content, and virus-host interactions. Notably, we identified virus-encoded auxiliary metabolic genes, such as the 3-oxoadipate enol-lactonase (PcaD) gene, which are critical for phthalate degradation. This finding underscores the direct role of phages in facilitating microbial adaptation and pollutant degradation in contaminated soils, suggesting that viral auxiliary metabolic genes could be harnessed for targeted bioremediation strategies to mitigate the environmental impact of plastic pollutants.

1. Introduction

Agricultural soils, including those used for rice cultivation, are facing rising contamination from different plastic pollutants [1], such as microplastics and plastic additives. These pollutants, due to their persistence and potential for accumulation [2], can have profound impacts on soil health and microbial communities. While the effects of plastic pollution on soil bacteria and fungi have been relatively well-studied [3], the influence on soil virome—the viral community within the soil—remains largely unexplored.

Viruses, especially bacteriophages, play a pivotal role in soil ecosystems by regulating microbial populations, facilitating horizontal gene transfer, and influencing biogeochemical cycles [4]. Through processes such as lytic infection and lysogenic conversion, viruses can significantly impact the composition and functionality of microbial communities [5]. In agricultural soils, where microbial activity is essential for nutrient cycling and plant health, understanding how plastic pollutants affect the virome is crucial.

Recent studies have highlighted the potential for plastic pollutants to alter soil properties, disrupt microbial habitats, and influence the diversity, abundance and activity of microbial taxa. For example, microplastics in soil have been found to reduce soil porosity and alter water retention characteristics, impacting microbial activity and nutrient cycling [6]. The presence of polyethylene (PE), polyvinyl chloride (PVC) and polypropylene (PP) residues in agricultural soils has been shown to disrupt microbial community structures, leading to a variation in specific bacteria such as nitrogen-fixing species [7]. Phthalate contamination derived from plasticizers has been demonstrated to increase the abundance of plastic-degrading microbes while reducing overall microbial diversity [8]. Given the integral role of viruses in microbial ecology, it is hypothesized that the influence of plastic pollution on soil viromes can have a strong impact on diversity, composition and functional dynamics of all the other species present. Currently, research on this topic is very limited, with only a few studies reporting the presence of phages, such as those from the *Autographiviridae* and *Podoviridae* families, on the surface of plastic beads [9]. Additionally, polystyrene microplastics have been shown to act as carriers, extending the survival time and infectivity of adsorbed viruses in water environment [10].

* Corresponding author.
E-mail address: laura.treu@unipd.it (L. Treu).

https://doi.org/10.1016/j.cej.2025.161877
Received 8 January 2025; Received in revised form 16 March 2025; Accepted 22 March 2025
Available online 24 March 2025

However, there is a lack of studies clearly demonstrating the effects of plastic pollutants on soil viromes. Different types of plastic pollutants may exert distinct effects on the soil virome due to their varying chemical properties and degradation pathways. Therefore, it is crucial to study the comprehensive impact of different types of plastic-based pollutants on agricultural soil viromes.

Auxiliary metabolic genes (AMGs) carried by viruses play a pivotal role in modulating the metabolic functions and environmental adaptability of their microbial hosts [11]. These genes, beyond the core viral genes that encode structural proteins, enhance the metabolic capacity of host cells, thereby facilitating survival under various stress conditions. AMGs are known to be involved in critical biochemical processes, including photosynthesis, translation machinery, carbon metabolism, phosphate metabolism, and sulfur cycling [12]. While the role of AMGs has been investigated in detail in marine ecosystems, where they contribute significantly to nutrient cycling and energy flow [13,14], their diversity and functional significance in soil environments remain less understood. Recent research suggests that AMGs can provide a competitive advantage to microbial hosts in contaminated soil environments [15], such as soils polluted with organochlorine pesticides, by influencing degradation pathways and promoting bacterial growth even at low contaminant concentrations [16]. This mutualistic interaction, often described as the "Piggyback the Winner" effect [17], indicates that viruses are not merely predators but also essential contributors to the ecological function and resilience of microbial communities. Understanding the role of AMGs in soil ecosystems is crucial for deciphering the complex interactions between viruses, their hosts, and plastic pollutants, and for developing innovative phage-assisted microbial strategies to eliminate plastic pollutants from soil.

Based on these existing issues, in the present study, we combined metagenomics and a lab-scale incubation experiment to explore how exposure to different plastic-based pollutants affects viral diversity, functional gene content, and virus-host interactions in rice field soil. Additionally, we focused on the relationship between plastic pollutant exposure and the virus-encoded metabolic and organic degradation AMGs. Multiple bioinformatics methods indicated that a candidate viral AMG, 3-oxoadipate enol-lactonase (PcaD), assists its host in metabolizing benzoate, a downstream product of phthalates. Overall, our findings enhance the understanding of how soil viromes are influenced by different plastic pollutants and suggest that virus-encoded AMGs can help bacteria resist plasticizer contamination.

2. Materials and methods

2.1. Sample Preparation and incubation

Soil samples were collected from rice fields located in Grumolo delle Abbadesse, VI Province, Italy (45°30′36.8″N 11°39′29.3″E), which has a humid subtropical climate classification according to the Köppen climate classification system, at a depth of 2–20 cm. This region is characterized by its long-standing use as paddy fields, with no plastic pollution from surrounding areas. The collected samples were dried, ground, and sieved through a 40-mesh sieve for further use. Two types of microplastic particles and one plasticizer were selected as plastic pollutants. The PE and PVC materials were provided by Euronewpack S.r.l. (an Italian company specializing in high-quality polymer products). Prior to the experiment, plastics were crushed using a crusher and sieved (<5 mm) to obtain microplastic pollutants. Phthalate plasticizer was purchased from Sigma-Aldrich (C_6H_4-1,2-$(CO_2C_2H_5)_2 \geq 99$ %).

Microcosm incubation experiments were conducted in 120 mL serum bottles, each containing 50 g of soil and 80 mL of sterile ultrapure water. To each bottle, 1 % w/w of PE, PVC, and DP (2.25 mmol) were added separately, with three replicates for each treatment. The serum bottles were placed in a shaking incubator set to 25 °C and incubated while avoiding light exposure for 24 h to ensure thorough mixing of the pollutants with the soil matrix. Following this, the microcosms were wrapped in aluminum foil to prevent evaporation and continued to be incubated at 25 °C, avoiding light exposure. After the first week of incubation, the liquid phase was collected using a 2 mL syringe and centrifuged at 4000 rpm for 10 min using an Eppendorf 5804R centrifuge. This sampling process was repeated every two weeks. The upper liquid layer was preserved at −20 °C for subsequent total organic content analysis.

2.2. DNA Extraction, Sequencing, Co-assembly and Binning Strategy

After long-term incubation, homogeneous soil mixtures were collected from each microcosm bottle for microbial analysis. The extraction method was as mentioned in our previous studies, using DNeasy PowerSoil® (QIAGEN GmbH, Hilden, Germany) to extract DNA from soil slurry [14]. NanoDrop (ThermoFisher Scientific, Waltham, Massachusetts) and Qubit fluorometer (Life Technologies, Carlsbad, California) were used to assess the quantity and quality of nucleic acids.

Metagenomic sequencing combined second and third-generation methods to achieve high-quality soil microbial and viral genomes. Shotgun sequencing was performed on the Illumina NovaSeq platform (DiBio Sequencing Facility, Department of Biology, University of Padua). For total DNA pooled from all samples, long-read libraries were constructed using the Rapid Barcoding Kit (SQK-RBK004). These long-read libraries were then sequenced on a MinION device with FLO-MIN106 R9 flow cells (Oxford Nanopore Technologies, Oxford, UK). The hybrid sequencing approach, leveraging both Illumina and Oxford Nanopore technologies, ensured high accuracy and comprehensive coverage of the soil genomes. Sequence data have been submitted to the NCBI Sequence Read Archive (SRA) under Bioproject PRJNA892267.

The quality of the raw data was assessed using FastQC v0.11.9 [18]. Low-quality reads and adapter sequences were filtered out using Trimmomatic v0.39 [19], applying adapter trimming and an average quality threshold of 20. Nanopore reads were assembled using Flye v2.9–0 [20], while Illumina reads were assembled into contigs using Megahit v1.2.9 [21] and setting the sensitive model. The Illumina reads were aligned to the Nanopore assembly using Bowtie2 [22], and the long-read assembly was then cleaned and corrected using Pilon v1.24–0 [23]. Finally, the Illumina and Nanopore assembly results were merged using Quickmerge v0.3–1 [24] to produce the final assembly.

Contigs were clustered to recover bins using MaxBin2 v2.2.7 [25], CONCOCT v1.0.0 [26], VAMB v3.0.2 [27], and MetaBAT2 v2.0 [28]. Subsequently, DAS_Tool 1.1.7 [29] was employed to dereplicate and aggregate the prokaryotic bins. The completeness and contamination levels of the bins were assessed using CheckM v1.1.3 [30], and low-quality bins (completeness < 50 % and contamination > 10 %) were filtered out. For prokaryotes taxonomic annotation, GTDB-Tk v1.3.0 [31] was used, based on the Genome Taxonomy Database (GTDB) Release 207.

2.3. Identification of viral contigs and their hosts.

Viral sequences were identified from assembled contigs using VIBRANT v1.2.1 [32] (default parameters), VirSorter2 v2.2.4 [33] (cutoff score ≥ 0.5), and CheckV v1.5 [34]. Only contigs identified as viral sequences by at least two of these three tools were retained. Contigs longer than 1 kb, with 95 % identity and 80 % sequence coverage, were clustered into representative viral operational taxonomic units (vOTUs) using Stampede-ClusterGenomes (https://bitbucket.org/MAVERICLab/stampede-clustergenomes/src/master/). Additionally, vOTUs with a length of 2 kb or less were filtered out to reduce false positives, except for those identified as complete genomes by CheckV v1.5 [34]. Abundance of vOTUs was estimated using CoverM v0.7.0 (https://github.com/wwood/CoverM). Then, the proteins encoded in vOTUs were clustered using vConTACT v2.0 [35] with the NCBI RefSeq v211 database and default parameters. vOTUs were further taxonomically classified using PhaGCN2.0 [36] based on its built-in database. Following the

clustering and subdivision process, the resulting viral clusters and their distribution patterns were visualized and analyzed using Cytoscape v3.7.18 [37]. For complete or high-quality vOTUs, viral lifestyles (lysogenic/lytic) were predicted using BACPHLIP [38] (a random forest-based tool) and PhaTYP [39] (a Transformer-based method). vOTUs predicted by BACPHLIP with high confidence (>90 %) and concurrently validated by PhaTYP were classified as either virulent or temperate viruses.

Three methods were used to identify viral hosts: homology matching, tRNA similarity, and CRISPR spacer similarity. In this study, the entire sequences, tRNA regions, and CRISPR spacer regions of vOTUs were searched against the recovered high-quality MAGs. Blastn v2.5.0+ [40] (coverage = 100, identity = 100), tRNAScan-SE v1.23 [41], and CRISPRCasFinder [42] were used to align sequences, identify tRNA, and search for CRISPR spacer regions, respectively. Furthermore, the aforementioned viral contigs identification process was independently conducted on the binned MAGs, deriving potentially integrated vOTUs at the MAGs level.

To determine potential induction events, we employed a log-transformed abundance model to assess the relative abundance of prophages compared to their corresponding host MAGs, as shown in the formula below:

$$\text{Coverage Ratio} = \log\left(\frac{\text{hosts relative abundance} + 1}{\text{phages RPKM} + 1}\right)$$

A threshold of 1 was applied to this log-transformed ratio to define induction events; for more details, see the Supplementary Information. This model enabled robust detection of potential prophage activation events, which were analyzed in the context of different environmental conditions, such as exposure to Diethyl Phthalate (DP). The script used for this analysis can be found on the Github: https://github.com/mengyuan09876/Induction-Events-Detection.

2.4. Viral gene annotation

Prodigal v2.6.1 [43] was used to predict the open reading frames (ORFs) of the final viral assemblies. Non-redundant proteins of the viral genomes were annotated using the KEGG (kobas3.0.3) [44] and eggNOG v2.1.1.2 [45] databases. To estimate functions related to organic carbon metabolism, the ORFs of the viral genomes were annotated using run_dbcan v4.1.4 with default parameters based on the Carbohydrate-Active enZYmes (CAZy) database [46]. Functional genes related to carbon, nitrogen, sulfur metabolism, and plastic pollutant degradation were identified according to the metabolic pathways mapped by KEGG orthologs.

AMGs were identified from assembly and each MAG using VIBRANT (v1.2.1, default parameters) [32] and DRAM-v (v1.2.0) [47]. The plastic degradation-associated auxiliary metabolic genes were further validated for viral origin using VIRSorter2 (v2.2.4, default parameters) [33] and CheckV (v1.5, default parameters) [34]. AMG identification with DRAM-v [47] was performed using default parameters. The resulting AMG markers were categorized as follows: 'V' for virus, 'M' for metabolic marker, 'K' for known AMG, 'E' for experimentally validated AMG, 'A' for virus host attachment and entry, 'P' for peptidase utilized by the virus, F for proximity to the contig end, and B for a cluster of consecutive genes (≥3) containing the metabolic marker 'M'. For VIBRANT, the obtained VIBRANT_AMGs.tsv lists all KEGG KOs designated as AMGs; while VIBRANT_categories.tsv lists the VOG, KEGG, and Pfam accession numbers and their respective 'v scores'. The AMG annotation results at the MAG level were summarized using a custom python script.

Genes promoter regions were identified using PhagePromoter [48]. Sequences with a palindromic score greater than 6 were used to predict Rho-independent terminators [49] while Rho-dependent terminators were identified separately. For the selected AMG phylogenetic analysis, protein sequences were retrieved from the NCBI RefSeq database using BLASTp [50], specifically targeting sequences from members of same host *Pseudomonas nitroreducens*. A total of nine protein sequences from the NCBI RefSeq database were included in the analysis. Sequences were then aligned using ClustalW [51], and a maximum likelihood tree was constructed using MEGA [52], with 500 bootstrap replicates, which was then visualized in iTOL [53]. Additionally, protein domain predictions were performed using the NCBI Conserved Domain Database (CDD) [54]. Protein 3D structure of preoteins involved in degradation were predicted using AlphaFold3 [55].

2.5. Statistical analyses

Data statistics and visualization in this study were conducted using Python (v3.6.2) (https://www.python.org/). Microbial rarefaction curves, alpha and beta diversity analyses (including alpha indices, UPGMA, and NMDS) were performed using scikit-bio and scikit-learn Python packages. The explanatory power (R-value) and significance (p-value) between samples were calculated based on Adonis analysis. Samples were clustered using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA), implemented with the SciPy and scikit-learn packages. Pearson correlations between viral contigs and MAGs were calculated using the pandas package, with a correlation threshold set at $r > 0.6$ and $p < 0.05$. The resulting network was generated using Gephi (v0.8.2) [56].

3. Results

3.1. Soil characteristics and organic carbon changes

Soil samples were collected from a depth of 2–20 cm in a paddy field located in Grumolo delle Abbadesse, Italy (45°30′36.8″N 11°39′29.3″E). The soil exhibits a homogenous texture, with no skeleton fraction larger than 2 mm, consisting of 316 g/kg of sand, 560 g/kg of silt, and 124 g/kg of clay, indicating a balanced loam texture.

The slightly alkaline pH of 8.1 is likely to positively influence nutrient availability and microbial activity, as nutrients like phosphorus and potassium become more soluble in such conditions, thereby supporting enhanced microbial processes. The soil is rich in organic carbon (16.62 g/kg) and organic matter (29 g/kg), with a total nitrogen content of 1.70 g/kg and a carbon/nitrogen ratio of 9.8, suggesting efficient organic matter decomposition and nutrient cycling. The cation exchange capacity of 13.1 meq/100 g, exchangeable potassium at 0.36 meq/100 g, and available phosphorus at 67 mg/kg indicate a fertile soil capable of retaining essential nutrients. Electrical conductivity (0.19 dS/m) and salinity (1.7 dS/m) are moderate, posing no significant hindrance to plant growth. Overall, this fertile and balanced soil provides an excellent medium for studying microbial activity (**Table S1**).

Fig. 1 presents the changes in Total Organic Carbon (TOC) concentrations of soil samples over a 49-day incubation period. The DP samples exhibit a pronounced increase in TOC, starting from approximately 754 mg/kg at 7 days and reaching about 4195 mg/kg by 49 days, indicating a significant contribution of DP to TOC levels. This sharp increase can be attributed to DP's high solubility and great microbial degradability, which facilitate its rapid breakdown and integration into the organic carbon pool of the soil. In contrast, PVC samples show a more modest increase from around 60 mg/kg at 7 days to roughly 100 mg/kg by day 35, before a slight decline. PE samples follow a similar trend, with TOC rising from about 61 mg/kg at 7 days to approximately 93 mg/kg by 49 days. These MPs consist of large polymeric structures resistant to degradation, leading to a slower and less significant release of organic carbon into the soil [57]. The decrease in TOC in the MPs groups compared to clean soil in later stages may be due to the adsorption of organic carbon onto the MPs [58].

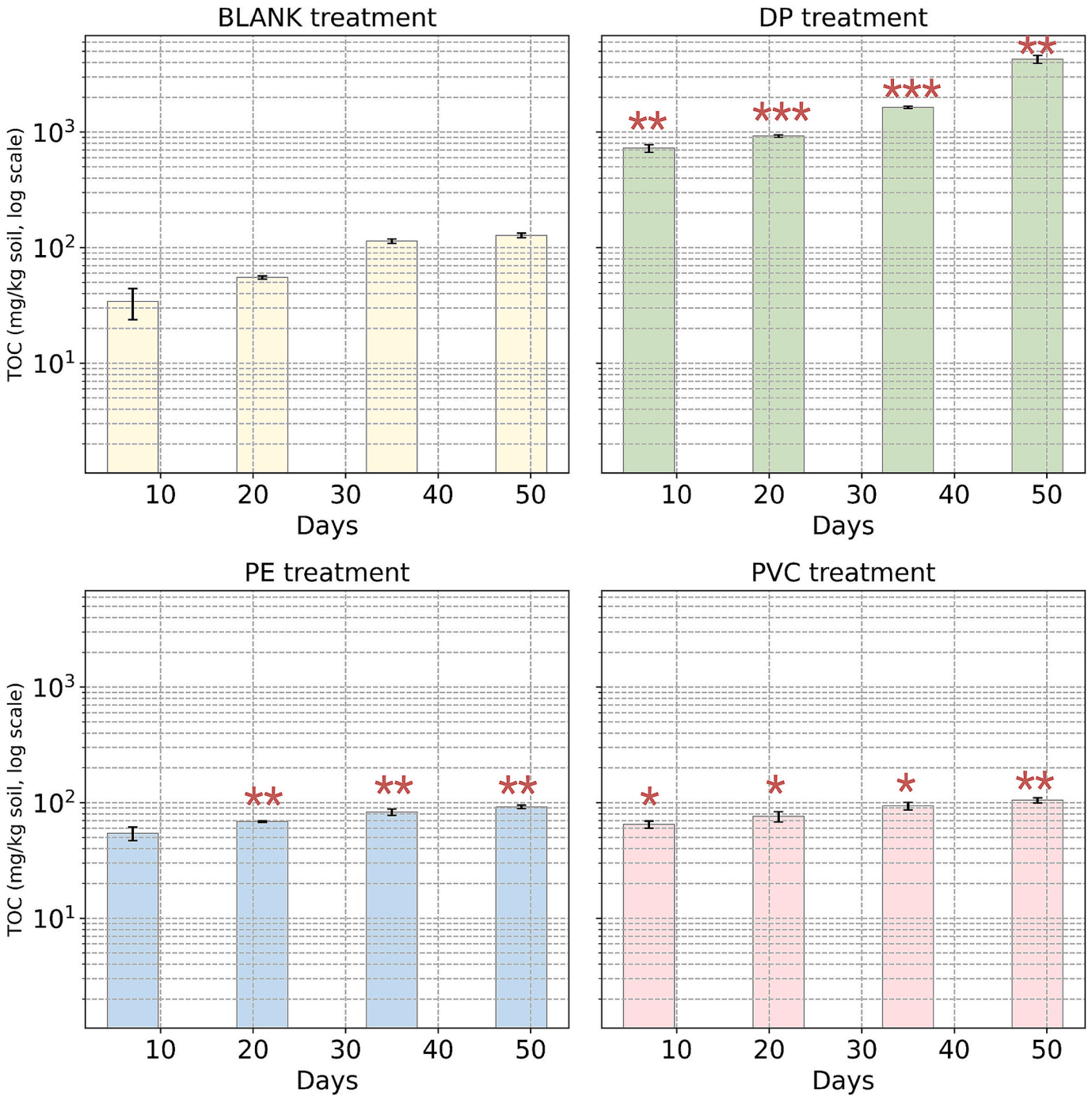

Fig. 1. Changes in total organic carbon (TOC) levels during extended incubation with a logarithmic y-axis for a wide TOC range (ANOVA: $F = 17.72$, $p < 0.0001$; pairwise T-tests comparing each treatment to BL indicated as * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

3.2. Viral diversity in different plastic-based polluted environments

vOTUs were extracted from the assembly to assess the diversity and have a functional overview of viral communities. A total of 2139 vOTUs (1992 larger than 1000 bp; 93 %) were identified from contaminated soil (Table S2). Compared to pristine soil, DP contamination led to a 3.15-fold increase in the proportion of viral sequences in the group, significantly higher than the more modest increases observed with polyethylene (13.08 %) and polyvinyl chloride (48.59 %) contamination. The representative, de-replicated vOTUs were clustered into 939 viral clusters (VCs) using the vConTACT2 method, which employs the ClusterONE algorithm to identify overlapping clusters based on shared protein gene clusters. These VCs were then compared with viral genomes in the NCBI RefSeq v211 database for taxonomic classification. (Fig. 2**a;** Table S3 – "viral_cluster_overview"). A total of 204 VCs were identified across the three treatments, with the DP treatment exhibiting a significantly higher number of unique VCs (44) compared to the other two treatments. Interestingly, all VCs from the PVC treatment had corresponding matches in the DP and PE treatments. The viral community network was reconstructed by mapping gene-sharing interactions between viral clusters (VCs). The network in the DP zone (network degree = 0.001) displayed a looser structure compared to the overall network (0.017), indicating fewer interactions or connections among the VCs in this treatment. The observation that only 7 known VCs (Table S3) were

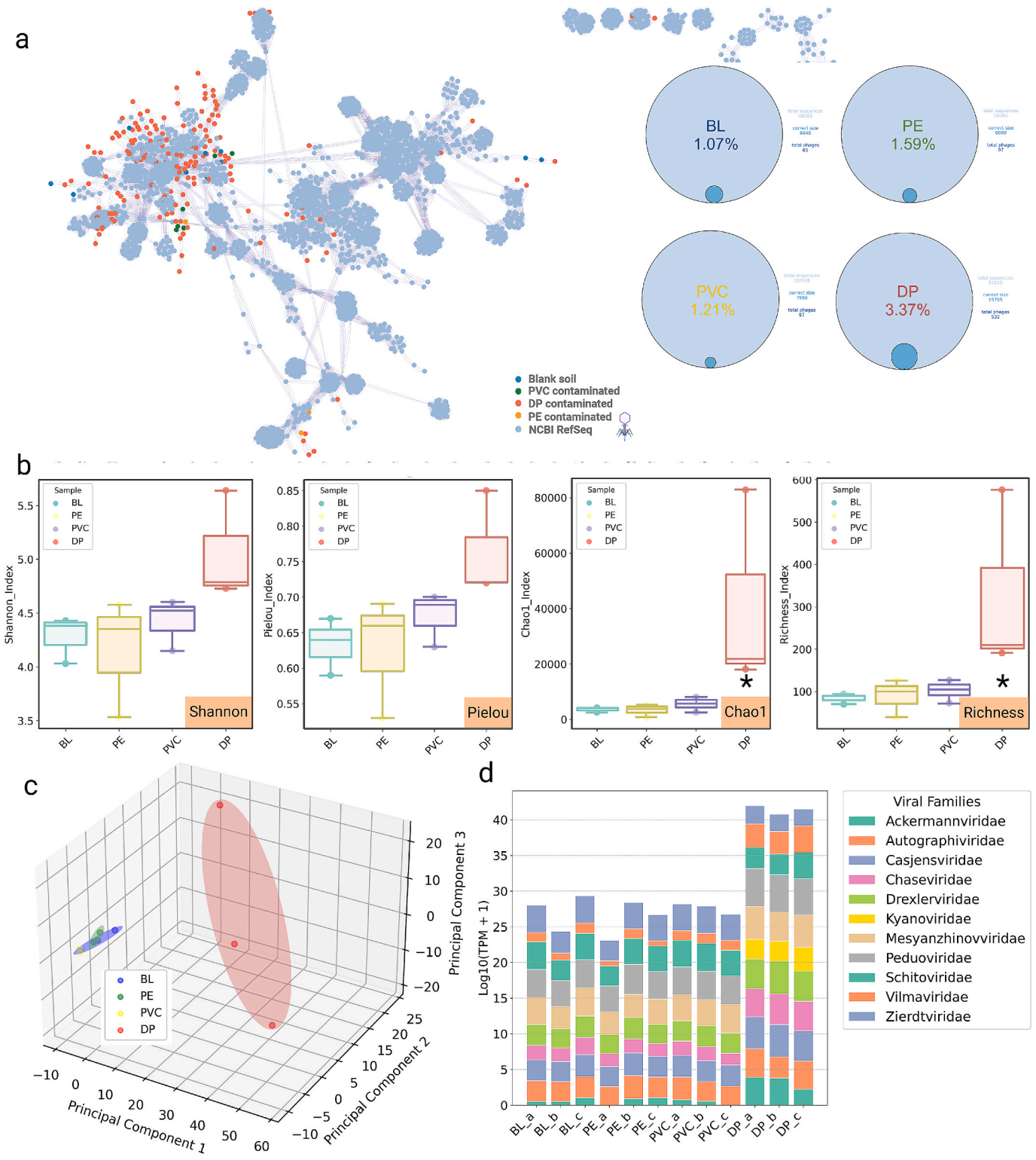

Fig. 2. A. a gene-sharing network for viral contigs (>10 kb) retrieved from clean (dark blue), PE (orange), PVC (green), and DP-contaminated (red) soils, as well as the NCBI RefSeq database (light blue). Nodes (circles) represent viral genome contigs, and edges indicate shared protein content. The nested circle diagram in the right shows the proportion of viral sequences in clean, MPs, and DP-contaminated soils. b. Differences in alpha diversity among clean, MPs, and DP-contaminated soils (One-way ANOVA: $F = 21.29$, $p = 0.0004$; pairwise T-tests comparing each treatment to BL indicated as * $P < 0.05$). c. PCoA analysis was performed to compare beta diversity among clean, MPs, and DP-contaminated soils ANOVA, followed by Tukey's multiple comparisons test, was used to assess differences in beta diversity values between groups. d. Bar plot of viral taxonomic composition at the family level in clean, MPs, and DP-contaminated soils based on $\log_{10}(TPM + 1)$ values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shared with the NCBI RefSeq database suggests that our samples contain a significant number of novel viruses, particularly in soils exposed to plastic-based pollutants. This could indeed be due to changes in the microbiome, as plasticizers and other pollutants may alter the microbial community, leading to the emergence of new bacteriophages. However, other factors could also contribute to this diversity, such as the selection pressure exerted by the pollutants themselves, which might drive the rise of novel viral entities. Additionally, the complexity and variability of the soil environment could inherently support a diverse viral community, which includes members that might not have been captured in existing databases [59]. Overall, the soil viral community under DP contamination is more diverse (i.e., higher Chao1 and Richness indices) and more even (i.e., higher Shannon and Pielou indices; Fig. 2**b**). Three-dimensional PCA analysis further confirmed that DP resulted in a more unique viral community structure among the three plastic-based pollutants (ANOSIM, $R^2 = 0.9788$, $P = 0.011$, Fig. 2**c**).

Considering that vcontact2 relies on protein assignments, the viral taxonomic predictions were further refined utilizing the PhaGCN-equipped GCN model. As depicted in Fig. 2**d**, *Mesyanzhinovviridae*, *Schitoviridae*, and *Zierdtviridae* emerge as phage families within clean soil samples. The introduction of PE and PVC MPs led to the emergence of the *Peduoviridae* family, accounting for 7.32 % in PE-contaminated soil and 13.12 % in PVC-contaminated soil. Conversely, the PVC treatment resulted in a significant increase in the relative abundance of *Schitoviridae* (an increase of 27.5 %). Notably, DP-contaminated soil exhibited a higher diversity in viral families, with the relative abundance of the *Peduoviridae* family increasing by 44.3 % compared to the PE treatment, and by 17.3 % compared to the PVC treatment. Furthermore, the exclusive presence of seven viral families (*Ackermannviridae*, *Autographiviridae*, *Casjensviridae*, *Chaseviridae*, *Drexlerviridae*, *Kyanoviridae*, and *Vilmaviridae*) solely in DP-contaminated soil underscores the profound impact of DP contamination on the viral community composition in soil.

3.3. DP pollution triggers a new pattern of virus-host interaction

Using tRNA matches and CRISPR spacer similarity searches (see Methods), we predicted hosts for 825 vOTUs, linking them to 80 MAGs across 9 bacterial and 1 archaeal phylum (see **Table S4**). Only hosts with a prediction score above 3, indicating strong alignment and confidence in host-virus associations, were selected for further analysis [60]. Notably, this analysis facilitated the linkage of 18 host taxa within the soil to their respective viral counterparts (Fig. 3**a**). Interestingly, a distinct differentiation in host abundance across soils with varying levels of contamination was observed, with some species more influenced by MPs contamination, and others affected by DP pollution. Specifically, taxa such as *Caulobacter* sp. *upd137*, *Hyphomicrobiaceae* sp. *upd130*, and *Rhodocyclaceae* sp. *upd106* were identified in pristine soils or those characterized by MPs contamination. Conversely, in soil contaminated with DP, there is a higher diversity of hosts identified at high-confidence, with *Comamonas terrigena* sp. *upd013*, *Pseudomonas* sp. *upd084*, *Pigmentiphaga* sp. *upd102*, and *Achromobacter* sp. *upd101* showing significantly higher relative abundances. These hosts exhibited multiple associations with viral contigs from various VCs. Additionally, the host species *Pigmentiphaga* sp. upd102 has been previously characterized for its potential activity in DP degradation [8], suggesting a

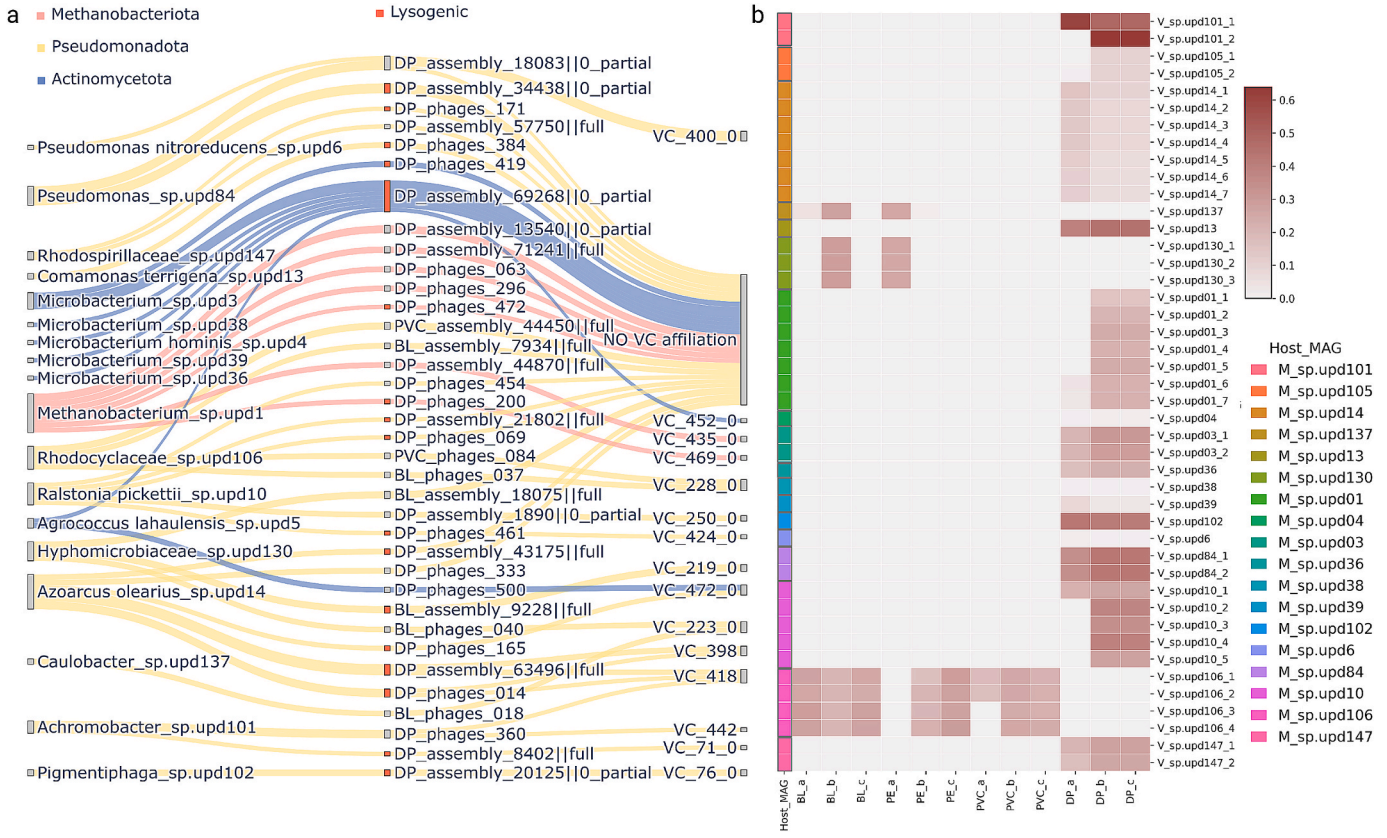

Fig. 3. A.Viral contigs taxonomically assigned at the family level and associated with different host species, with line colors indicating host phyla. The "No VC affiliation" group represents viruses without any affiliation to viral clusters. The line thickness connecting viruses and prokaryotic species corresponds to the predicted final score based on tRNA matching and CRISPR spacer analysis. The red boxes denote lysogenic phages. b. The ratio coverage of the predicted host MAGs alongside the RPKM values for the integrative prophage components. The heatmap values are represented on a logarithmic scale. MAGs are identified by their sequence numbers, with 'M' denoting the prokaryotic portion and 'V' indicating the integrated viral component. The colored annotations on the left distinguish the different MAGs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

plausible correlation with specific phages harboring AMGs that facilitate degradation. The heatmap (Fig. 3**b**) visualizes the coverage of host MAGs (labeled with an 'M' prefix) and the RPKM (Reads Per Kilobase of transcript, per Million mapped reads) values of their integrated prophages (labeled with a 'V' prefix) across various plastic treatments. For each host MAG (first column), one or more prophages (next columns) are shown. A threshold of 1 in the virus/MAG coverage ratio is used to identify prophage induction (**Table S5**), indicating that prophage abundance has surpassed that of the corresponding host. This suggests that DP acts as a potent environmental stressor, triggering the activation of latent prophages. In contrast, fewer induction events are observed under BL, PE, and PVC treatments, indicating a more stable host-phage dynamic. These findings highlight the potential role of induced phages in response to contamination, possibly influencing microbial functions such as pollutant degradation in affected environments. **Fig. S1** also shows the significantly enriched taxa from phylum to genus under different treatments. For more detailed descriptions, see the Supplementary Information.

A correlation network based on the abundance of viruses and hosts is depicted in **Fig. S2**. Notably, in the virus-host correlation network of the samples exposed to DP pollution a strong positive correlation between the abundances of phages and their hosts was observed. Additionally, some vOTU exhibited positive correlations with hosts across different phyla, this suggests that DP pollution induces more complex interactions between viruses and prokaryotes, emphasizing multiple partnerships. This finding highlights the ecological impact of DP pollution, as it appears to promote the formation of complex viral-host networks, potentially influencing microbial community dynamics and contributing to

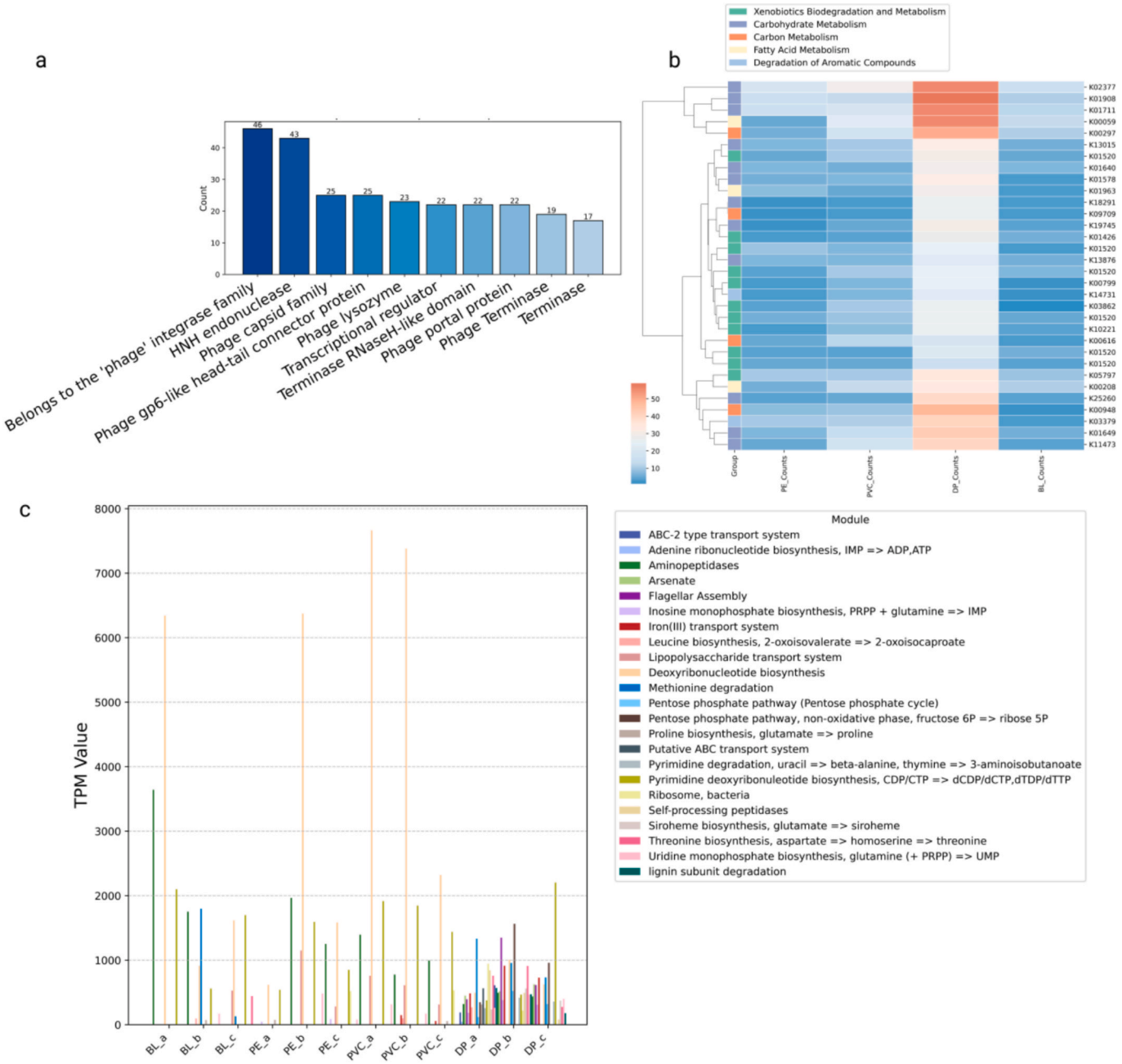

Fig. 4. A. top ten most abundant functional gene classes identified in viral sequences. b. variations in mapped gene counts for plastic-associated pollutant degradationpathways across different samples. c. Distribution of metabolic pathway modules encoded by assembly-level predicted auxiliary metabolic genes, as determined by VIBRANT and DRAM-v, across different samples (displayed using gene TPM values for all three replicates).

shifts in host populations in response to environmental stressors. In addition, the results in Fig. S3 indicate generally weak and non-significant relationships between phage abundance and MAGs relative abundance across the BL, PE, and PVC groups, while only the DP group shows a modest but statistically significant positive correlation, further supporting the idea that DP pollution would trigger a distinct pattern of virus-host interactions.

3.4. Functional groups of the viral-encoded proteins

To explore the functional contribution of viruses to the ecology of bacterial communities, we compared the functional annotation of viral sequences in pristine soil and in soil contaminated with plastic-based pollutants. A thorough annotation of proteins functions was obtained by combining results from KEGG, CAZy and EGGNOG databases. Fig. S4 illustrates the counts of KEGG IDs among different treatments. Notably, the DP group exhibits distinct characteristics compared to other groups, underscoring significant increases in KO abundance levels. Non-metric multidimensional scaling (NMDS) analysis of viral genomes revealed greater similarity in the functional composition of viral contigs from PE and clean soil, whereas those from PVC and DP exhibited a distinct and broader range of genes spread across different functional categories (Fig. S4).

The proteins encoded by most viral ORFs (about 35 %) could only be annotated as “unclassified” or “poorly characterized” by EGGNOG. As expected, the highest proportion of annotated genes were related to viral replication, repair, or transcription, such as the ’phage’ integrase family (1.4 %) and HNH endonuclease (1.3 %) (Fig. 4**a**, Table S6). Fig. S5 further reveals the distribution of gene hit counts across various dbCAN subclass groups: GH (glycoside hydrolases), GT (glycosyltransferases), CE (carbohydrate esterases), CBM (carbohydrate-binding modules), AA (auxiliary activities), and PL (polysaccharide lyases). The GH and GT classes exhibited the highest gene hit counts, especially in the DP group. In contrast, the abundances of AA and PL were comparatively low, in agreement with the observation that GT, GH, CE, and CBM together account for about 95 % of the CAZyme genes present in various environmental contexts [61].

The relevant pathways involved in organic matter metabolism (xenobiotics, carbohydrate, fatty acid, and aromatic compounds metabolism) were further selected, and their gene copy numbers under different treatments were compared. Notably, “carbohydrate metabolism” and “xenobiotic metabolism” were significantly higher in the viral microbiome compared to other metabolic pathways (Fig. 4**b**). Specifically, genes involved in the degradation of aromatic compounds and xenobiotics, such as aldehyde dehydrogenase (NAD^+), cyclohexanone monooxygenase chnB, muconate cycloisomerase catB, and amidase amiE, showed a significantly higher abundance only in the viral groups exposed to DP contamination. Furthermore, several genes related to methane metabolism (three genes in DP-contaminated soil and one gene in PVC-contaminated soil) and sulfur metabolism (four genes in DP-contaminated soil) were identified in the contaminated soils. KEGG annotation results indicated that the viral genes encompassed various metabolic activities and pathways for organic matter degradation. Alterations in viral AMGs were compared across different treatments (refer to Methods and Table S7 for details). The AMGs encompass a total of 23 functional pathways. Compared to other treatments, DP-contaminated soil exhibited significantly more diverse viral AMGs. Notably, under DP treatment, TPM values for AMGs involved in flagellar assembly, the pentose phosphate pathway, and methionine degradation were significantly increased, whereas the abundance of genes associated with deoxyribonucleotide biosynthesis was markedly decreased, suggesting a strategic reallocation of metabolic resources by viruses under DP stress (Fig. 4**c**).

3.5. Structural and functional characteristics of virus-encoded genes in plastic-associated pollutant degradation

Previous research indicates that AMG identification at the individual bin level yields a higher number of detected AMGs compared to whole metagenome assemblies. This approach provides greater resolution, facilitating the discovery of additional metabolic genes that may be overlooked in broader metagenomic analyses [62]. Therefore, to conduct a more detailed examination of AMGs, we analyzed the proteins encoded by phages sequences integrated within individual MAGs, as shown in Table S8. Interestingly, this more detailed analysis confirmed the presence of the gene encoding 3-oxoadipate enol lactonase (PcaD, EC 3.1.1.24) in the integrated prophage of *Pseudomonas nitroreducens* sp. upd006. The completeness and classification of the viral genome (completeness: 20 %, contamination: 26 %) containing this gene were assessed using CheckV and Phabox (Table S10).

Multiple analyses were conducted to confirm the viral origin of the PcaD encoding gene (682123_fragment_3_48) located at positions 57603–58436 within the contig 682123_fragment_3 (See Fig. S6 for the viral genome map). This contig was assigned to the *Casjensviridae* family and categorized as a dsDNA phage by VirSorter2. Gene 682123_fragment_3_48 is also flanked by regions containing viral hallmark genes, such as VOG00741 (Protein P5) and other hypothetical proteins identified by VOGs (Table S9). The significant viral score (38.900/100) and low cellular score (11.100/100) predicted by VirSorter2 for sequence 682123_fragment_3 strongly support its viral origin. DRAM-v classified the sequence as Category 1 [63], indicating an enrichment of viral-like genes and possibly non-Caudovirales genes, reinforcing confidence in its viral classification beyond the more well-characterized Caudovirales (Table S10). However, DRAM-v did not recognize PcaD as an AMG, likely because its degradation activity is not part of the usual metabolic processes. This might be due to the limited number of studies focused on PcaD in prior research, resulting in its absence from the DRAM-v database (Table S10 – “DRAM-v”).

Promoter and terminator prediction analysis also revealed important regulatory elements associated with the 682123_fragment_3_48 gene, suggesting its role in phage-mediated processes. A phage promoter, spanning positions 57429–60068 with a score of 0.826, was located upstream of the PcaD gene. Downstream, a Rho-independent terminator at positions 58241–58319, with a score of 9.0, featured several palindromic sequences, contributing to its stability and efficiency in terminating transcription. Additionally, two potential Rho-dependent terminators were identified within the PcaD gene, which could influence the transcription of this gene or other viral genes located in this region. The spanin gene, essential for bacterial cell lysis during phage infection, was also predicted close to the PcaD gene region, strongly supporting its viral origin (Table S10). In summary, these results confirmed that PcaD gene is of viral origin and may be positively selected by phthalate plasticizers, as they were only found in DP-contaminated soils.

The enzyme 3-oxoadipate enol-lactonase (EC 3.1.1.24) catalyzes the hydrolysis of 3-oxoadipate enol-lactone to 3-oxoadipate. As a member of the hydrolase family, it acts on carboxylic ester bonds and is crucial in degrading phthalates and benzoates via hydroxylation, aiding the breakdown of these aromatic [64] (Fig. 5**a**). Since the viral gene encoding EC 3.1.1.24 was identified in *Pseudomonas nitroreducens* sp. upd006, we performed a phylogenetic analysis using PcaD protein sequences from the *Pseudomonas* species group, which includes *Pseudomonas nitroreducens*, available in the UniProtKB database (Fig. 5**b**). Most sequences cluster closely together, indicating high similarity among the host-encoded PcaD genes. However, the gene encoded by 682123_fragment_3_48 forms a unique evolutionary branch, distinct from most reference sequences within *Pseudomonas nitroreducens* despite sharing the same function. This distinct clustering further suggests that the PcaD gene encoded by the prophage within *Pseudomonas nitroreducens* has diverged significantly from those encoded by the host bacteria, potentially due to different evolutionary pressures. In addition,

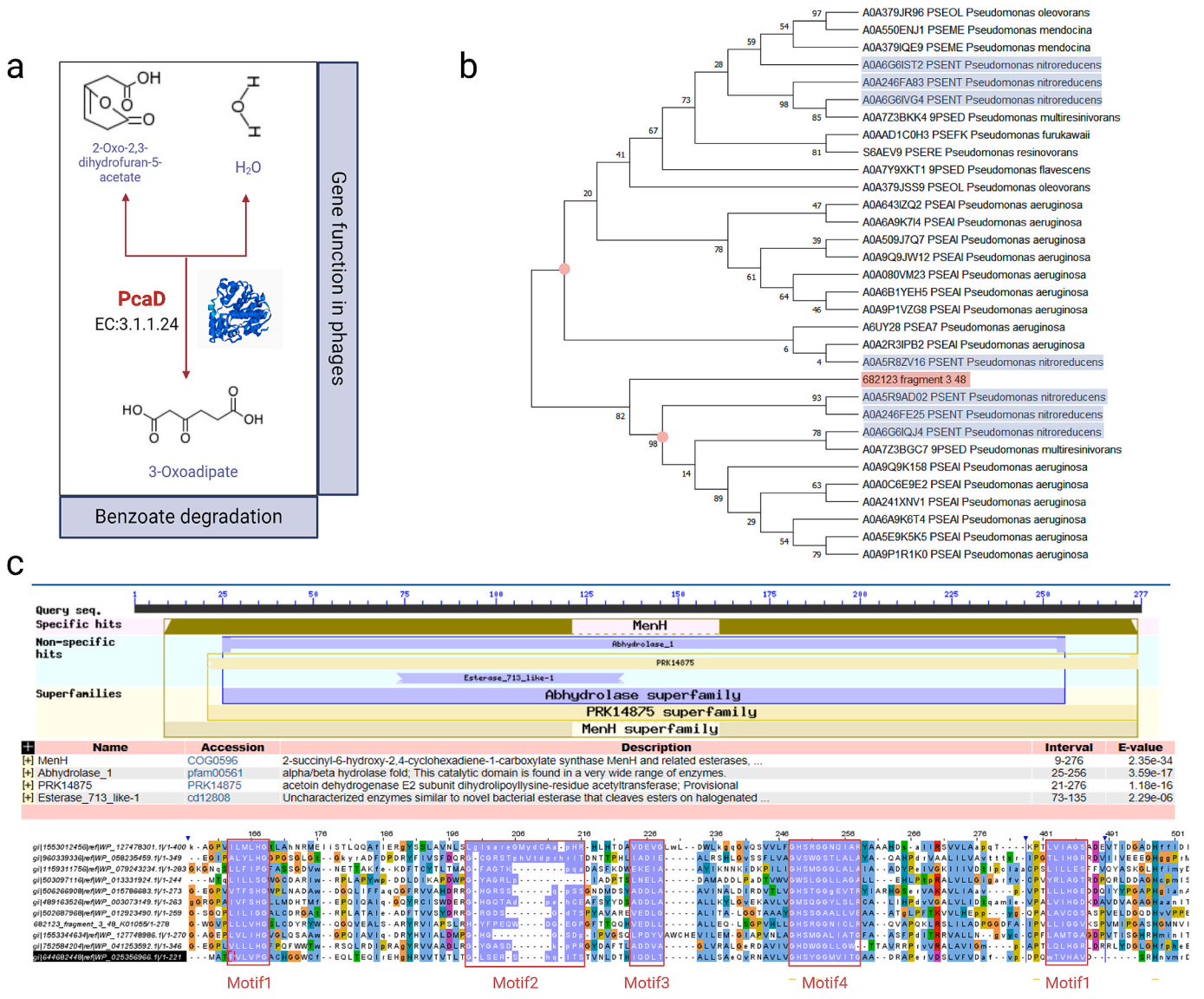

Fig. 5. A. overview of plasticizer degradation genes identified in the soil virome; the encoded proteins are linked to the KEGG pathway. Protein structures of pesticide degradation genes were modeled using AlphaFold3. b. Phylogenetic protein tree based on the PcaD protein identified in the phage 682123_fragment_3 (highlighted in red), along with similar protein sequences retrieved from UniProtKB database. c. Query and alignment results of protein conserved domains in the virus-encoded PcaD. The CD-Search standard results list the highest-scoring domain models from each source database. Signature motif sites are indicated by red frames. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the alpha/beta hydrolase family protein encoded by 682123_fragment_3_48 contains the nucleophile-His-acid catalytic triad characteristic of the alpha/beta fold hydrolase domain architecture (domain architecture ID 11426811) (Fig. 5**c**). The protein sequence closely matches the conserved domain MenH (COG0596), spanning residues 9–276, with a significant E-value of 2.41e-34. COG0596 is classified as a model that may encompass more than one domain and is the sole member of the superfamily cl43201. The catalytic core residues are highly conserved within the alpha/beta hydrolase family. Detailed analysis of the protein sequences from this superfamily identified four key motifs, each with distinct functional roles. Motif 1, characterized by glycine and serine residues, enhances flexibility, potentially aiding in substrate binding. Motif 2, a leucine-rich region, is likely involved in stabilizing hydrophobic interactions crucial for maintaining the structural integrity of the enzyme. Motif 3, enriched with aspartate residues, plays a central role in catalytic activity by facilitating substrate binding and reaction mechanisms. Motif 4, similar in composition to Motif 1, further contributes to the enzyme's flexibility and activity. We also performed motif identification on the PcaD protein from the common species *Pseudomonas nitroreducens*. Interestingly, the motifs present in this species are largely absent in 682123_fragment_3_48 (**Fig. S7**). This also suggests that the PcaD in this study is of viral origin rather than from the host *Pseudomonas nitroreducen*.

4. Discussion

Although viral metagenomics have been relatively well-explored in aquatic environments [65] and in the human gut [66], contaminated soil ecosystems have not received the same level of attention. The present study employed second- and third-generation hybrid metagenomic sequencing to investigate soil microbiomes, revealing a link between plastic pollutants and the diversity, composition, and function of viral communities. Notably, DP-contaminated soils exhibited a more diverse viral community, including unique viral taxa and increased predicted associations with host bacterial taxa, suggesting strong viral-bacterial interdependence. Additionally, viral metagenomes from DP-

contaminated soils were enriched in genes related to metabolism and organic matter degradation, further supporting this interdependence. It is hypothesized that DP's higher solubility facilitates the rapid release of dissolved organic matter, thereby enhancing microbial metabolism and triggering stress responses that may promote prophage induction or increased phage release; although direct evidence for this mechanism is currently lacking, future studies employing a gradient of DP concentrations and multi-dimensional analyses, including structural equation modeling, would address this gap. In addition, non-biological factors, such as the physical adsorption of DP altering soil permeability, may influence virus–host dynamics; however, these effects are likely minor compared to the substantial organic carbon changes [67]induced by DP, a limitation that future studies should address.

Under the influence of PE, PVC MPs, and the DP, a diverse array of host MAGs exhibited significant phage-induction events, reflecting their sensitivity to these pollutants. While both PE and PVC MPs triggered viral activation in several host MAGs, particularly those within the *Proteobacteria* and *Actinobacteria* phyla, the most pronounced effects were observed under DP exposure. DP, as a potent chemical stressor, possibly induced widespread prophage activation across diverse host lineages, including *Azoarcus olearius* sp. upd14, *Pseudomonas nitroreducens* sp. upd6, and multiple members of the *Microbacterium* genus. The induction of these phages may be linked to their metabolic versatility and potential ability to interact with or degrade phthalate compounds. For instance, *Microbacterium* spp., commonly present in contaminated environments, may engage in metabolic pathways that interact with phthalates. Such interactions could generate cellular stress, subsequently triggering prophage induction as part of a stress response mechanism. Based on the analysis of lysogenic and lytic phage ratios and TPM values, DP significantly increases TPM for both phage types and results in a higher proportion of lysogenic phages (Fig. S8). This effect is analogous to antibiotic-induced phage [68], suggesting that DP may trigger prophage induction by generating stress-inducing metabolites or by directly interfering with key cellular processes. Data further indicate that certain phages are associated with DP degradation pathways, which may favor lytic cycles under DP exposure. This pronounced response to DP underscores its significant impact on microbial metabolism and community dynamics, potentially disrupting ecological functions more profoundly than microplastics alone. This strong response to DP suggests its significant impact on microbial metabolism and community dynamics, potentially disrupting ecological functions more profoundly than microplastics alone. These findings highlight the need for further investigation into the mechanisms of prophage induction by phthalate and its broader ecological implications compared to those of microplastics.

According to Pfam, KEGG, and CAZy databases, a variety of genes having a metabolic role were identified in the viral metagenomes, suggesting that the degradation of plastic pollutants may drive the enrichment of virus-encoded AMGs in soil microbiomes. These viral AMGs might benefit bacteria by mitigating the toxic effects of pollutants or aiding in energy acquisition through plastic degradation. For instance, the substantial increase in glycoside hydrolase gene counts in the DP group highlights a shift in microbial enzymatic potential [69], indicating that microbial communities might be adapting to utilize complex available carbon sources more efficiently as a stress response to DP. Additionally, viruses play a role in regulating carbon cycling, through the presence of AMGs related to carbohydrate and xenobiotic metabolism in contaminated soil viral microbiomes, demonstrating their adaptability to plastic pollutants. The increased presence of these genes also shows that viruses can influence pollutant degradation through microbial co-metabolism. This highlights the potential for using viral capabilities in bioremediation, as viruses could be engineered to transfer essential genes to bacterial hosts [70], enhancing the breakdown of persistent pollutants [16].

To further investigate the potential role of virus-encoded AMGs in plastic pollutant degradation, we bioinformatically identified a virus-encoded 3-oxoadipate enol lactonase (PcaD) gene associated with a dsDNA phage from the *Casjensviridae* family. This gene was classified by DRAM-v as a Category 1 region, which could play a crucial role in environmental adaptation. The PcaD enzyme is essential for hydrolyzing 3-oxoadipate enol-lactone to 3-oxoadipate [71], a key step in benzoate degradation. This suggests that viral genes play a significant role in providing novel functional properties to bacteria which can become efficient in breaking down complex phthalates. According to these findings, phages can potentially support host bacteria in pollutant degradation. Phylogenetic analysis revealed that the virus-encoded PcaD gene is evolutionarily distinct from its ortholog counterpart in members of the *Pseudomonas aeruginosa* group, indicating unique viral adaptability. The alpha/beta hydrolase family protein encoded by this viral gene contains conserved catalytic motifs essential for its enzymatic activity, crucial for flexibility, binding, and catalysis, possibly an enol lactonase with peculiar functional properties.

The discovery of virus-encoded AMGs related to pollutant degradation opens new avenues for bioremediation strategies. While our current results provide insights into the potential role of the viral PcaD gene, our study is limited by the absence of direct transcriptional quantification and functional validation under controlled conditions. Future studies comparing the degradation efficiency between strains with and without phage integration, along with direct transcriptional quantification of the viral PcaD gene during phage induction events, will be highly meaningful in elucidating its independent contribution to pollutant degradation. Additionally, comparative studies of soils contaminated with different pollutants and clean soils can provide deeper insights into the evolutionary origins and adaptation mechanisms of these viral genes.

5. Conclusion

This study underscores the pivotal role of phages in influencing soil ecosystem responses to plastic pollution. Through advanced metagenomic analyses, we discovered that plasticizers, particularly phthalates, induce more significant changes in the soil virome compared to microplastics. These changes lead to greater alterations in viral diversity, functional gene content, and virus-host interactions. The identification of a phage-encoded 3-oxoadipate enol-lactonase gene further highlights the direct contribution of phages to the degradation of complex pollutants. These findings suggest that phages play a vital role in enhancing microbial adaptability to plastic contaminants, opening new possibilities for targeted bioremediation strategies that utilize viral capabilities to mitigate the environmental impact of plastic pollutants.

6. Environmental implications

This study highlights the critical role of phages in soil ecosystem responses to plastic pollution. Metagenomic analyses reveal that plasticizers, especially phthalates, cause more significant changes in the soil virome than microplastics, altering viral diversity, functional genes, and virus-host interactions. The discovery of a phage-encoded 3-oxoadipate enol-lactonase underscores phages' direct role in degrading complex pollutants. These findings indicate that phages may contribute to microbial adaptability to plastic contaminants, presenting potential avenues for exploring bioremediation strategies to mitigate the environmental impacts of plastic pollution.

CRediT authorship contribution statement

Mengyuan Ji: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Laura Treu:** Writing – review & editing, Resources, Project administration. **Stefano Campanaro:** Writing – review & editing, Supervision, Resources, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cej.2025.161877.

Data availability

Sequencing data has been submitted to SRA

References

[1] MacLeod, M.; Arp, H. P. H.; Tekman, M. B.; Jahnke, A. The global threat from plastic pollution. *Science* **2021**, *373* (6550), 61-65. Sa'adu, I.; Farsang, A. Plastic contamination in agricultural soils: a review. *Environmental Sciences Europe* **2023**, *35* (1), 13.

[2] J.J. Alava, A. Jahnke, M. Bergmann, G.V. Aguirre-Martínez, L. Bendell, P. Calle, G. A. Domínguez, E.M. Faustman, J. Falman, T.N. Kazmiruk, A call to include plastics in the global environment in the class of persistent, bioaccumulative, and toxic (PBT) pollutants, Environ. Sci. Tech. 57 (22) (2023) 8185–8188.

[3] Zhang, X.; Li, Y.; Ouyang, D.; Lei, J.; Tan, Q.; Xie, L.; Li, Z.; Liu, T.; Xiao, Y.; Farooq, T. H. Systematical review of interactions between microplastics and microorganisms in the soil environment. *Journal of hazardous materials* **2021**, *418*, 126288. Rillig, M. C.; Kim, S. W.; Zhu, Y.-G. The soil plastisphere. *Nature Reviews Microbiology* **2024**, *22* (2), 64-74. Zhu, D.; Li, G.; Wang, H.-T.; Duan, G.-L. Effects of nano-or microplastic exposure combined with arsenic on soil bacterial, fungal, and protistan communities. *Chemosphere* **2021**, *281*, 130998.

[4] Brum, J. R.; Hurwitz, B. L.; Schofield, O.; Ducklow, H. W.; Sullivan, M. B. Seasonal time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics. *The ISME journal* **2016**, *10* (2), 437-449. Irwin, N. A.; Pittis, A. A.; Richards, T. A.; Keeling, P. J. Systematic evaluation of horizontal gene transfer between eukaryotes and viruses. *Nature microbiology* **2022**, *7* (2), 327-336.

[5] Fernández, L.; Rodríguez, A.; García, P. Phage or foe: an insight into the impact of viral predation on microbial communities. *The ISME journal* **2018**, *12* (5), 1171-1179. Silveira, C. B.; Luque, A.; Rohwer, F. The landscape of lysogeny across microbial community density, diversity and energetics. *Environmental microbiology* **2021**, *23* (8), 4098-4111.

[6] Wang, F.; Wang, Q.; Adams, C. A.; Sun, Y.; Zhang, S. Effects of microplastics on soil properties: current knowledge and future perspectives. *Journal of Hazardous Materials* **2022**, *424*, 127531. Kumar, A.; Mishra, S.; Pandey, R.; Yu, Z. G.; Kumar, M.; Khoo, K. S.; Thakur, T. K.; Show, P. L. Microplastics in terrestrial ecosystems: Un-ignorable impacts on soil characterises, nutrient storage and its cycling. *TrAC Trends in Analytical Chemistry* **2023**, *158*, 116869.

[7] Huang, S.; Guo, T.; Feng, Z.; Li, B.; Cai, Y.; Ouyang, D.; Gustave, W.; Ying, C.; Zhang, H. Polyethylene and polyvinyl chloride microplastics promote soil nitrification and alter the composition of key nitrogen functional bacterial groups. *Journal of Hazardous Materials* **2023**, *453*, 131391. Wang, Q.; Feng, X.; Liu, Y.; Li, W.; Cui, W.; Sun, Y.; Zhang, S.; Wang, F.; Xing, B. Response of peanut plant and soil N-fixing bacterial communities to conventional and biodegradable microplastics. *Journal of Hazardous Materials* **2023**, *459*, 132142.

[8] M. Ji, G. Giangeri, M. Usman, C. Liu, M. Bosaro, F. Sessa, P. Canu, L. Treu, S. Campanaro, An integrated Metagenomic-Pangenomic strategy revealed native microbes and magnetic biochar cooperation in plasticizer degradation, Chem. Eng. J. 468 (2023) 143589.

[9] L. Niu, S. Zhao, Y. Chen, Y. Li, G. Zou, Y. Tao, W. Zhang, L. Wang, H. Zhang, Diversity and potential functional characteristics of phage communities colonizing microplastic biofilms, Environ. Res. 219 (2023) 115103.

[10] J. Lu, Z. Yu, L. Ngiam, J. Guo, Microplastics as potential carriers of viruses could prolong virus survival and infectivity, Water Res. 225 (2022) 119115.

[11] Sun, M.; Yuan, S.; Xia, R.; Ye, M.; Balcázar, J. L. Underexplored viral auxiliary metabolic genes in soil: diversity and eco-evolutionary significance. *Environmental Microbiology* **2023**, *25* (4), 800-810. Luo, X.-Q.; Wang, P.; Li, J.-L.; Ahmad, M.; Duan, L.; Yin, L.-Z.; Deng, Q.-Q.; Fang, B.-Z.; Li, S.-H.; Li, W.-J. Viral community-wide auxiliary metabolic genes differ by lifestyles, habitats, and hosts. *Microbiome* **2022**, *10* (1), 190.

[12] Anantharaman, K.; Duhaime, M. B.; Breier, J. A.; Wendt, K. A.; Toner, B. M.; Dick, G. J. Sulfur oxidation genes in diverse deep-sea viruses. *Science* **2014**, *344* (6185), 757-760. Hurwitz, B. L.; Hallam, S. J.; Sullivan, M. B. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome biology* **2013**, *14*, 1-14. Goldsmith, D. B.; Parsons, R. J.; Beyene, D.; Salamon, P.; Breitbart, M. Deep sequencing of the viral phoH gene reveals temporal variation, depth-specific composition, and persistent dominance of the same viral phoH genes in the Sargasso Sea. *PeerJ* **2015**, *3*, e997.

[13] Gazitúa, M. C.; Vik, D. R.; Roux, S.; Gregory, A. C.; Bolduc, B.; Widner, B.; Mulholland, M. R.; Hallam, S. J.; Ulloa, O.; Sullivan, M. B. Potential virus-mediated nitrogen cycling in oxygen-depleted oceanic waters. *The ISME Journal* **2021**, *15* (4), 981-998. Braga, L. P.; Orland, C.; Emilson, E. J.; Fitch, A. A.; Osterholz, H.; Dittmar, T.; Basiliko, N.; Mykytczuk, N. C.; Tanentzap, A. J. Viruses direct carbon cycling in lake sediments under global change. *Proceedings of the National Academy of Sciences* **2022**, *119* (41), e2202261119.

[14] J. Lopez-Simon, M. Vila-Nistal, A. Rosenova, D. De Corte, F. Baltar, M. Martinez-Garcia, Viruses under the Antarctic Ice Shelf are active and potentially involved in global nutrient cycles, Nat. Commun. 14 (1) (2023) 8295.

[15] Q. Xu, H. Zhang, P. Vandenkoornhuyse, S. Guo, Y. Kuzyakov, Q. Shen, N. Ling, Carbon starvation raises capacities in bacterial antibiotic resistance and viral auxiliary carbon metabolism in soils, Proc. Natl. Acad. Sci. 121 (16) (2024) e2318160121.

[16] X. Zheng, M.T. Jahn, M. Sun, V.-P. Friman, J.L. Balcazar, J. Wang, Y. Shi, X. Gong, F. Hu, Y.-G. Zhu, Organochlorine contamination enriches virus-encoded metabolism and pesticide degradation associated auxiliary genes in soil microbiomes, ISME J. 16 (5) (2022) 1397–1408.

[17] A.N. Shkoporov, C.J. Turkington, C. Hill, Mutualistic interplay between bacteriophages and bacteria in the human gut, Nat. Rev. Microbiol. 20 (12) (2022) 737–749.

[18] J. Brown, M. Pirrung, L.A. McCue, FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool, Bioinformatics 33 (19) (2017) 3137–3139.

[19] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (15) (2014) 2114–2120.

[20] C.D. Faulk, novo sequencing, diploid assembly, and annotation of the black carpenter ant, Camponotus pennsylvanicus, and its symbionts by one person for $1000, using nanopore sequencing, Nucleic Acids Res. 51 (1) (2023) 17–28.

[21] D. Li, C.-M. Liu, R. Luo, K. Sadakane, T.-W. Lam, MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, Bioinformatics 31 (10) (2015) 1674–1676.

[22] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods 9 (4) (2012) 357–359.

[23] Z. Chen, D.L. Erickson, J. Meng, Polishing the Oxford Nanopore long-read assemblies of bacterial pathogens with Illumina short reads to improve genomic analyses, Genomics 113 (3) (2021) 1366–1377.

[24] M. Chakraborty, J.G. Baldwin-Brown, A.D. Long, J. Emerson, Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage, Nucleic Acids Res. 44 (19) (2016) e147–e.

[25] Y.-W. Wu, B.A. Simmons, S.W. Singer, MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets, Bioinformatics 32 (4) (2016) 605–607.

[26] Alneberg, J.; Bjarnason, B. S.; de Bruijn, I.; Schirmer, M.; Quick, J.; Ijaz, U. Z.; Loman, N. J.; Andersson, A. F.; Quince, C. CONCOCT: clustering contigs on coverage and composition. *arXiv preprint arXiv:1312.4038* **2013**.

[27] J.N. Nissen, J. Johansen, R.L. Allesøe, C.K. Sønderby, J.J.A. Armenteros, C. H. Grønbech, L.J. Jensen, H.B. Nielsen, T.N. Petersen, O. Winther, Improved metagenome binning and assembly using deep variational autoencoders, Nat. Biotechnol. 39 (5) (2021) 555–560.

[28] D.D. Kang, F. Li, E. Kirton, A. Thomas, R. Egan, H. An, Z. Wang, MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies, PeerJ 7 (2019) e7359.

[29] C.M. Sieber, A.J. Probst, A. Sharrar, B.C. Thomas, M. Hess, S.G. Tringe, J. F. Banfield, Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy, Nat. Microbiol. 3 (7) (2018) 836–843.

[30] D.H. Parks, M. Imelfort, C.T. Skennerton, P. Hugenholtz, G.W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, Genome Res. 25 (7) (2015) 1043–1055.

[31] P.-A. Chaumeil, A.J. Mussig, P. Hugenholtz, D.H. Parks, GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database, Oxford University Press, 2020.

[32] K. Kieft, Z. Zhou, K. Anantharaman, VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences, Microbiome 8 (2020) 1–23.

[33] J. Guo, B. Bolduc, A.A. Zayed, A. Varsani, G. Dominguez-Huerta, T.O. Delmont, A. A. Pratama, M.C. Gazitúa, D. Vik, M.B. Sullivan, VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses, Microbiome 9 (2021) 1–13.

[34] S. Nayfach, A.P. Camargo, F. Schulz, E. Eloe-Fadrosh, S. Roux, N.C. Kyrpides, CheckV assesses the quality and completeness of metagenome-assembled viral genomes, Nat. Biotechnol. 39 (5) (2021) 578–585.

[35] H. Bin Jang, B. Bolduc, O. Zablocki, J.H. Kuhn, S. Roux, E.M. Adriaenssens, J. R. Brister, A.M. Kropinski, M. Krupovic, R. Lavigne, Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks, Nat. Biotechnol. 37 (6) (2019) 632–639.

[36] J.-Z. Jiang, W.-G. Yuan, J. Shang, Y.-H. Shi, L.-L. Yang, M. Liu, P. Zhu, T. Jin, Y. Sun, L.-H. Yuan, Virus classification for viral genomic fragments using PhaGCN2, Brief. Bioinform. 24 (1) (2023) bbac505.

[37] M.S. Cline, M. Smoot, E. Cerami, A. Kuchinsky, N. Landys, C. Workman, R. Christmas, I. Avila-Campilo, M. Creech, B. Gross, Integration of biological networks and gene expression data using Cytoscape, Nat. Protoc. 2 (10) (2007) 2366–2382.

[38] A.J. Hockenberry, C.O. Wilke, BACPHLIP: predicting bacteriophage lifestyle from conserved protein domains, PeerJ 9 (2021) e11396.

[39] J. Shang, X. Tang, Y. Sun, PhaTYP: predicting the lifestyle for bacteriophages using BERT, Brief. Bioinform. 24 (1) (2023) bbac487.

[40] Y. Chen, W. Ye, Y. Zhang, Y. Xu, High speed BLASTN: an accelerated MegaBLAST search tool, Nucleic Acids Res. 43 (16) (2015) 7762–7768.

[41] P.P. Chan, B.Y. Lin, A.J. Mak, T.M. Lowe, tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes, Nucleic Acids Res. 49 (16) (2021) 9077–9096.
[42] D. Couvin, A. Bernheim, C. Toffano-Nioche, M. Touchon, J. Michalik, B. Néron, E. P. Rocha, G. Vergnaud, D. Gautheret, C. Pourcel, CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins, Nucleic Acids Res. 46 (W1) (2018) W246–W251.
[43] D. Hyatt, G.-L. Chen, P.F. LoCascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification, BMC Bioinf. 11 (2010) 1–11.
[44] Z. Jin, Y. Sato, M. Kawashima, M. Kanehisa, KEGG tools for classification and analysis of viral proteins, Protein Sci. 32 (12) (2023) e4820.
[45] J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S.K. Forslund, H. Cook, D.R. Mende, I. Letunic, T. Rattei, L.J. Jensen, eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses, Nucleic Acids Res. 47 (D1) (2019) D309–D314.
[46] J. Zheng, Q. Ge, Y. Yan, X. Zhang, L. Huang, Y. Yin, dbCAN3: automated carbohydrate-active enzyme and substrate annotation, Nucleic Acids Res. 51 (W1) (2023) W115–W121.
[47] M. Shaffer, M.A. Borton, B.B. McGivern, A.A. Zayed, S.L. La Rosa, L.M. Solden, P. Liu, A.B. Narrowe, J. Rodríguez-Ramos, B. Bolduc, DRAM for distilling microbial metabolism to automate the curation of microbiome function, Nucleic Acids Res. 48 (16) (2020) 8883–8900.
[48] M. Sampaio, M. Rocha, H. Oliveira, O. Dias, Predicting promoters in phage genomes using PhagePromoter, Bioinformatics 35 (24) (2019) 5301–5302.
[49] E.A. Lesnik, R. Sampath, H.B. Levene, T.J. Henderson, J.A. McNeil, D.J. Ecker, Prediction of rho-independent transcriptional terminators in Escherichia coli, Nucleic Acids Res. 29 (17) (2001) 3583–3594.
[50] R. Lavigne, D. Seto, P. Mahadevan, H.-W. Ackermann, A.M. Kropinski, Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools, Res. Microbiol. 159 (5) (2008) 406–414.
[51] J.D. Thompson, T.J. Gibson, D.G. Higgins, Multiple sequence alignment using ClustalW and ClustalX, Curr. Protoc. Bioinform. 1 (2003), 2.3. 1-2.3. 22.
[52] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (7) (2016) 1870–1874.
[53] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation, Nucleic Acids Res. 49 (W1) (2021) W293–W296.
[54] A. Marchler-Bauer, M.K. Derbyshire, N.R. Gonzales, S. Lu, F. Chitsaz, L.Y. Geer, R. C. Geer, J. He, M. Gwadz, D.I. Hurwitz, CDD: NCBI's conserved domain database, Nucleic Acids Res. 43 (D1) (2015) D222–D226.
[55] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A.J. Ballard, J. Bambrick, Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature (2024) 1–3.
[56] M. Bastian, S. Heymann, M. Jacomy, Gephi: an open source software for exploring and manipulating networks, in: Proceedings of the International AAAI Conference on Web and Social Media, 2009, pp. 361–362.
[57] T. Jin, J. Tang, H. Lyu, L. Wang, A.B. Gillmore, S.M. Schaeffer, Activities of microplastics (MPs) in agricultural soil: a review of MPs pollution from the perspective of agricultural ecosystems, J. Agric. Food Chem. 70 (14) (2022) 4182–4201.
[58] M. Munoz, D. Ortiz, J. Nieto-Sandoval, Z.M. de Pedro, J.A. Casas, Adsorption of micropollutants onto realistic microplastics: Role of microplastic nature, size, age, and NOM fouling, Chemosphere 283 (2021) 131085.
[59] J.R. Brister, D. Ako-Adjei, Y. Bao, O. Blinkova, NCBI viral genomes resource, Nucleic Acids Res. 43 (D1) (2015) D571–D577.
[60] A.C. Gregory, O. Zablocki, A.A. Zayed, A. Howell, B. Bolduc, M.B. Sullivan, The gut virome database reveals age-dependent patterns of virome diversity in the human gut, Cell Host Microbe 28 (5) (2020) 724–740, e728.
[61] Paixão, D. A. A.; Tomazetto, G.; Sodre, V. R.; Goncalves, T. A.; Uchima, C. A.; Büchli, F.; Alvarez, T. M.; Persinoti, G. F.; da Silva, M. J.; Bragatto, J. Microbial enrichment and meta-omics analysis identify CAZymes from mangrove sediments with unique properties. *Enzyme and Microbial Technology* **2021**, *148*, 109820. Santos-Pereira, C.; Sousa, J.; Costa, Â. M.; Santos, A. O.; Rito, T.; Soares, P.; Franco-Duarte, R.; Silvério, S. C.; Rodrigues, L. R. Functional and sequence-based metagenomics to uncover carbohydrate-degrading enzymes from composting samples. *Applied Microbiology and Biotechnology* **2023**, *107* (17), 5379-5401.
[62] D. Wang, L. Liu, X. Xu, C. Wang, Y. Wang, Y. Deng, T. Zhang, Distributions, interactions, and dynamics of prokaryotes and phages in a hybrid biological wastewater treatment system, Microbiome 12 (1) (2024) 134.
[63] S. Roux, F. Enault, B.L. Hurwitz, M.B. Sullivan, VirSorter: mining viral signal from microbial genomic data, PeerJ 3 (2015) e985.
[64] J. Bains, L. Kaufman, B. Farnell, M.J. Boulanger, A product analog bound form of 3-oxoadipate-enol-lactonase (PcaD) reveals a multifunctional role for the divergent cap domain, J. Mol. Biol. 406 (5) (2011) 649–658.
[65] A. Rastrojo, A. Alcamí, Aquatic viral metagenomics: Lights and shadows, Virus Res. 239 (2017) 87–96.
[66] Kavagutti, V. S.; Andrei, A.-Ş.; Mehrshad, M.; Salcher, M. M.; Ghai, R. Phage-centric ecological interactions in aquatic ecosystems revealed through ultra-deep metagenomics. *Microbiome* **2019**, *7*, 1-15. Nayfach, S.; Páez-Espino, D.; Call, L.; Low, S. J.; Sberro, H.; Ivanova, N. N.; Proal, A. D.; Fischbach, M. A.; Bhatt, A. S.; Hugenholtz, P. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nature microbiology* **2021**, *6* (7), 960-970.
[67] Y. Wang, Y. Liu, Y. Wu, N. Wu, W. Liu, X. Wang, Heterogeneity of soil bacterial and bacteriophage communities in three rice agroecosystems and potential impacts of bacteriophage on nutrient cycling, Environ. Microbiome 17 (1) (2022) 17.
[68] J.-L. Bru, B. Rawson, C. Trinh, K. Whiteson, N.M. Høyland-Kroghsbo, A. Siryaporn, PQS produced by the Pseudomonas aeruginosa stress response repels swarms away from bacteriophage and antibiotics, J. Bacteriol. 201 (23) (2019) 00383-00319, https://doi.org/10.1128/jb.
[69] O.Y. Costa, M. De Hollander, A. Pijl, B. Liu, E.E. Kuramae, Cultivation-independent and cultivation-dependent metagenomes reveal genetic and enzymatic potential of microbial community involved in the degradation of a complex microbial polymer, Microbiome 8 (2020) 1–19.
[70] E.P. Rocha, D. Bikard, Microbial defenses against mobile genetic elements and viruses: Who defends whom from what? PLoS Biol. 20 (1) (2022) e3001514.
[71] Y. He, Z. Wang, T. Li, X. Peng, Y. Tang, X. Jia, Biodegradation of phenol by Candida tropicalis sp.: Kinetics, identification of putative genes and reconstruction of catabolic pathways by genomic and transcriptomic characteristics, Chemosphere 308 (2022) 136443.