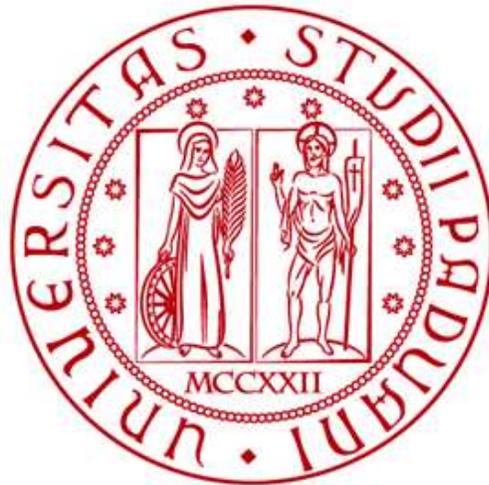# Università degli Studi di Padova

Dipartimento di Biologia

Corso di Laurea Magistrale in Molecular Biology

Tesi di Laurea

Exploring the Role of Phages in the Anaerobic Digestion Microbial Community through Metagenomics and Single-Cell Analysis

Relatore: Dott. Stefano Campanaro
Dipartimento di Biologia

Correlatore: Dott. Esteban Orellana
Dipartimento di Biologia

Laureanda: Retno Wahidah Supardi

ANNO ACCADEMICO 2023/2024

INDEX

Abstract

Phages play a crucial yet underexplored role in shaping microbial communities within anaerobic digestion systems. This study investigates the role of phages on a simplified microbial community under various stress conditions, including UV light exposure, antibiotic treatment, heavy metals, and hydrogen peroxide (H_2O_2). Among these stressors, oxidative stress induced by H_2O_2 had the most significant effect as it reduced microbial growth by 30%. Metagenomic analysis of DNA extracted from both pellets and supernatants enabled the reconstruction of 33 metagenome-assembled genomes (MAGs) and 129 viral MAGs. The most abundant organism was *Methanothermobacter thermautotrophicus*, a hydrogenotrophic methanogen capable of capturing carbon dioxide (CO_2) producing methane. The second most abundant organism was *Caldanaerobacter subterraneus*, a putative competitor of *M. thermautotrophicus* with metabolic capabilities of capturing CO_2 producing acetate through Wood-Ljungdahl (WL) pathway. A phage-host interaction analysis revealed that *C. subterraneus* was infected by *Caudoviricetes* sp. 124 which carried auxiliary metabolic genes involved in WL. Single-cell analysis enabled the confirmation of 2 phage-host interactions predicted from metagenomics data. Furthermore, single-cell analysis allowed the discovery of new phage-host interactions between 9 MAGs and 28 virMAGs. The integration of metagenomics and single-cell analysis offers a powerful approach for elucidating the dynamics of phage-host relationships and their influence on microbial community structure in carbon-dioxide methanation systems.

1.   INTRODUCTION

1.1. Biomethane Production and Microbial Community in Anaerobic Digestion

The global shift towards renewable energy sources has sparked interest in biomethane production as a sustainable alternative. Biomethane (referred from now as methane - CH_4) can be generated through anaerobic digestion (AD), a process that relies on the dynamics of diverse microbial communities (Kouzuma et al., 2015). Therefore, the intricate interplay of microbial communities is key to unlocking the full potential of biogas production.

Essentially, biogas production through anaerobic digestion consists of biological degradation and gasification processes. In detail, the anaerobic digestion process involves several stages, including hydrolysis, acidogenic fermentation, hydrogen-producing acetogenesis, and methanogenesis. During the initial stage, complex organic materials are broken down into soluble monomers and oligomers like sugars, fatty acids, and amino acids. In the second stage, known as acidogenesis, acidogenic fermentation bacteria convert soluble monomers into smaller organic molecules such as short-chain volatile fatty acids (VFAs), alcohols, acetate, hydrogen (H_2), and carbon dioxide (CO_2). Moreover, in the third stage which is acetogenesis, hydrogen-producing acetogenic bacteria metabolize and oxidize the VFA and alcohols produced during acidogenesis into acetate, CO_2, and H_2. In the final stage, methane (CH_4) is generated primarily from acetate via acetotrophic methanogenesis or from H_2 and CO_2 via hydrogenotrophic methanogenesis. Additionally, acetate can be synthesized from H_2 and CO_2 through homoacetogenesis, or it can be further degraded to H_2 and CO_2 by acetate-oxidizing bacteria (Li et al., 2019). This final step is crucial for biogas production, as methane is the major component of biogas.

As methanogenesis can occur from CO_2 or acetate, methane is produced in different pathways. The methanogenesis from CO_2 can be explained from a process called the Wolfe cycle. This process begins with CO_2 being reduced to formyl group (CHO), which attaches to the carrier molecule methanofuran (MFR). This formyl group is transferred to tetrtamethanopterin (H_4MPT), then undergoes dehydration and cyclization to form CH_3-H_4MPT. The methenyl group is reduced to methylene, and then to a methyl group. The methyl group is transferred to coenzyme M (HS-CoM) forming (CH_3-S-CoM), alongside the export of sodium ion (Na+) from the cell. The methyl group is finally reduced to methane (CH_4) through the oxidation of CH_3-S-CoM and coenzyme B (HS-CoB), forming heterodisulfide (CoM-S-S-CoB) which is later reduced to regenerate HS-CoM and HS-CoB by an enzyme called heterodisulfide reductase (Hdr). In the case of hydrogenotrophic methanogenesis, the term hydrogenotrophic refers to the process in which heterodisulfide (CoM-S-S-CoB)

is reduced by Heterodisulfide reductase (Hdr) with H_2 as electron donors. In the case of acetotrophic methanogenesis, the methane production from acetate begins from the binding of acetate with coenzyme A (CoA). Then, the acetyl group in acetyl-CoA is split into methyl group (CH_3) and carbonyl group (CO). The methyl group of acetate enters the Wolfe cycle in the level of CH_3-H_4MPT. The carbonyl group of acetate is then oxidized to form CO_2 which provides electrons for methyl reduction (**Figure 1)**.
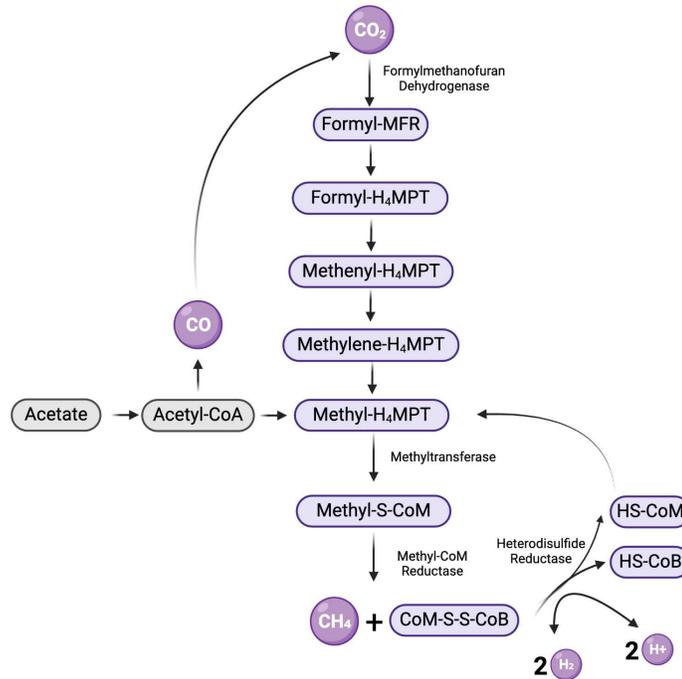
Figure 1. Graphical depiction of the methanogenesis pathway (Created with BioRender.com)

The acetate formed through homoacetogenesis does not necessarily represent a loss of H_2 and CO_2, as it can be further converted into methane through acetotrophic methanogenesis. In particular, homoacetogenesis plays a crucial role in biomethanation systems due to the utilization of substrates that are similar to those used in hydrogenotrophic methanogenesis (Kofoed et al., 2021). Homoacetogens are microorganisms that perform homoacetogenesis, a process where they convert CO_2 into acetate using the Wood-Ljungdahl pathway. This process is considered an indirect method for biogas upgrading because, instead of directly removing or converting CO_2 from biogas into methane, it produces acetate as an intermediary product. Acetate can subsequently serve as a substrate for other microorganisms, such as methanogens, which then convert it into methane, thereby increasing the methane content in the biogas. This two-step approach—first forming acetate and then methane—makes homoacetogenesis an indirect contribution to biogas upgrading, in

contrast to direct CO_2-to-CH_4 conversion methods like hydrogenotrophic methanogenesis.

Different processes of methanogenesis are related to the interaction of microorganisms within the microbial community. Methanogen, the major metabolic organism responsible for the methanogenesis of anaerobic digestion, is classified into three groups according to the substrate for their metabolism and their role in different methanogenesis processes. These three groups include hydrogenotrophic methanogens, acetotrophic methanogens, and homoacetogens (Wu et al., 2021).

Hydrogenotrophic methanogens are responsible for the production of methane (CH_4) from CO_2 and H_2. Specifically, the methane from hydrogenotrophic methanogenesis can relieve the inhibitory effect of H_2 on acetogenic microorganisms. Methane from this reaction is also crucial in the biogas production system as it can solve the problem derived from H_2 due to its explosion risk. The representative methanogenic species in this process includes *Methanobacterium bryantii, Methanobacterium thermoalcaliphium,* and *Methanosarcina barkeri.* Besided H_2, some hydrogenotrophic methanogens can also use compounds likemethanol to reduce CO_2 and produce methane.

Acetotrophic methanogens are microorganisms that convert acetic acid to CH_4 and CO_2. The performance of these microorganisms are crucial during anaerobic conversion of acetate. Among all the species of acetotrophic methanogens, *Methanosaeta* spp. And *Methanosarcina* spp. are the most abundant in the anaerobic digestion system. The abundance of these methanogens is impacted by H_2 availability due to their competition with homoacetogens. However, regardless of the capacity of homoacetogens to metabolize CO_2 and H_2, they cannot compete over methanogens under theomorphic condition. Therefore, the consumption of CO_2 by homoacetogens is ruled out in thermophilic anaerobic digestion system.

Regarding the biological biogas upgrading, in the context of anaerobic digestion, it can be obtained from two processes, one is mediated by hydrogenotropic methanogenic archaea and the other one is mediated by homoacetogenic bacteria. Microbial analysis conducted in biogas upgrading system revealed that the most common hydrogenotrophic methanogenic genera were *Methanobacterium, Methanomassiliicoccus,* and *Methanothermobacter* (Xu et al., 2020). In practical applications, using mixed adapted culture is more advantageous than pure cultures because mixed culture are more resilient, do not require sterile condition, and can produce higher methane yield.

Microorganisms in biogas upgrading systems are highly responsive to environmental changes. A previous study reported that the presence of various archaeal genera such

as *Methanoculleus* within the Methanomicrobiales order was dominant in the system, but its relative abundance dropped from 36% to 24% as the temperature increased from 37°C to 55°C (Bassani et al., 2015). Factors such as temperature, pH, gas composition, and nutrient availability can influence the shape of microbial community. Additionally, the interactions between microorganisms, such as competition may also have a major effect on the taxonomic and functional composition of microbial communities. For example, species from the archaeal family Methanosaetaceae have enzymes with a high acetate affinity, which allows them to potentially suppress other microorganisms that also utilize acetate (Vrieze et al., 2012). Aside from interactions between microorganisms, the presence of phages has a major influence on the shape of the microbial community in the anaerobic digestion process. This is due to the fact that the replication of phages, through the lysis of host microorganisms can cause the removal of essential microbial groups.

1.2. Role of Phages in Microbial Communities

Viruses that infect and replicate within bacteria and archaea, known as bacteriophages or phages, are the most abundant entities on Earth. They play a critical role in shaping microbial diversity and composition. A study conducted in four full-scale anaerobic digesters at wastewater treatment plants across China reported that phages accounted for 40.6% of the variation in the prokaryotic community, which is substantially higher than the influence of abiotic factors, such as temperature or pH, which explained only 14.5% of the variation (Zhang et al., 2017). These findings highlight that phages play a much larger role than previously thought in regulating the composition and activity of prokaryotic communities in anaerobic digesters. Phages are considered as a key biotic factor that influences both microbial dynamics and the overall performance of anaerobic digestion, which suggests their potential role in optimizing biogas production processes.

Phages affect microbial communities in three main ways depending on the life cycle. Phages can modulate the abundance of certain bacterial species and release dissolved organic materials into the environment (viral shunt) through lysis, as well as facilitate horizontal gene transfer (HGT) between microbial genomes through lysogenic conversion, specialized transduction, or generalized transduction (Touchon et al., 2017). Therefore, understanding the phage life cycle is essential in order to study their role in the microbial community.

1.2.1. Phage Life Cycle

Phages act as obligate intracellular parasites during their initial phase. Their life cycles include lytic and lysogenic cycles. In the lytic cycle, right after phages infect the host, phages immediately begin to replicate by hijacking the host's cell to synthesize new viral particles. These viral particles are further assembled and released as new phages by lysing the host cell. On the other side, in the lysogenic cycle, the phage genome is either incorporated into the host's genome or exists in a plasmid-like form. The incorporated phage genome is known as prophage. The prophage replicates alongside the host's DNA and can be passed down to the daughter cells. However, under stress conditions, prophages can transition from the lysogenic to the lytic cycle. This process is known as phage induction (**Figure 2**).
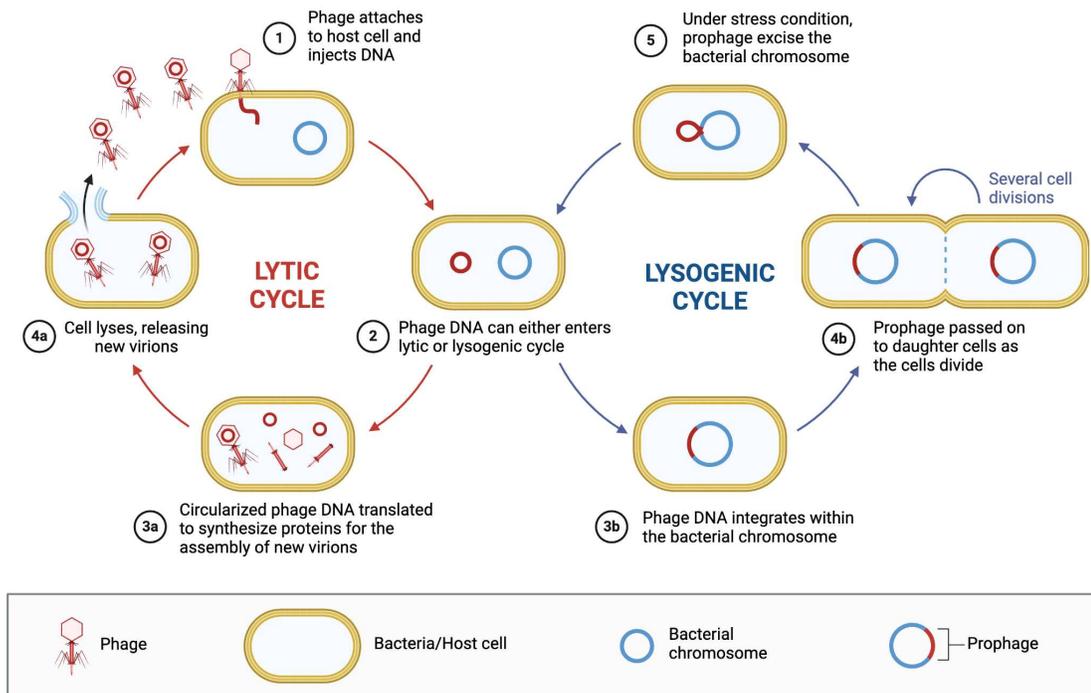
Figure 2. Scheme of the phage life cycle (Created with BioRender.com)

The transition from the lysogenic to a lytic cycle of the phages largely determines their interaction with the host. As an example, the lytic cycle is ecologically important in the marine food web because it affects the biogeochemical cycling of major elements, such as the flow of carbon into dissolved organic matter (DOM) and particulate organic matter (POM) (Perez Sepulveda et al., 2016). On the other side, the lysogenic cycle of the phages can impact the host's adaptability by either

facilitating horizontal gene transfer (HGT) or providing immunity against other homologous phage infections.

There are three major mechanisms for phage-mediated HGT, which include lysogenic conversion, specialized transduction, and generalized transduction (**Figure 3)**. In lysogenic conversion, temperate phage DNA typically integrates into the host's chromosome, though some prophages can exist as plasmids within the cell. When prophage genes are active, they can induce phenotypic change in the host's traits in terms of virulence, motility, and competition with other bacteria. Understanding how phages decide between lysis and lysogeny can reveal how lysogeny influences bacterial evolution. A recent study in phage metagenomics found that virus-to-microbe ratios decrease as microbial populations grow (Wigington et al., 2016). Additionally, the percentage of temperate phages rises with bacterial density which suggests that lysogeny is favored at high cell densities, as phages benefit more from replicating within fast-growing bacterial hosts.

Figure 3. Scheme of major phage-mediated HGT mechanisms (Created with BioRender.com)
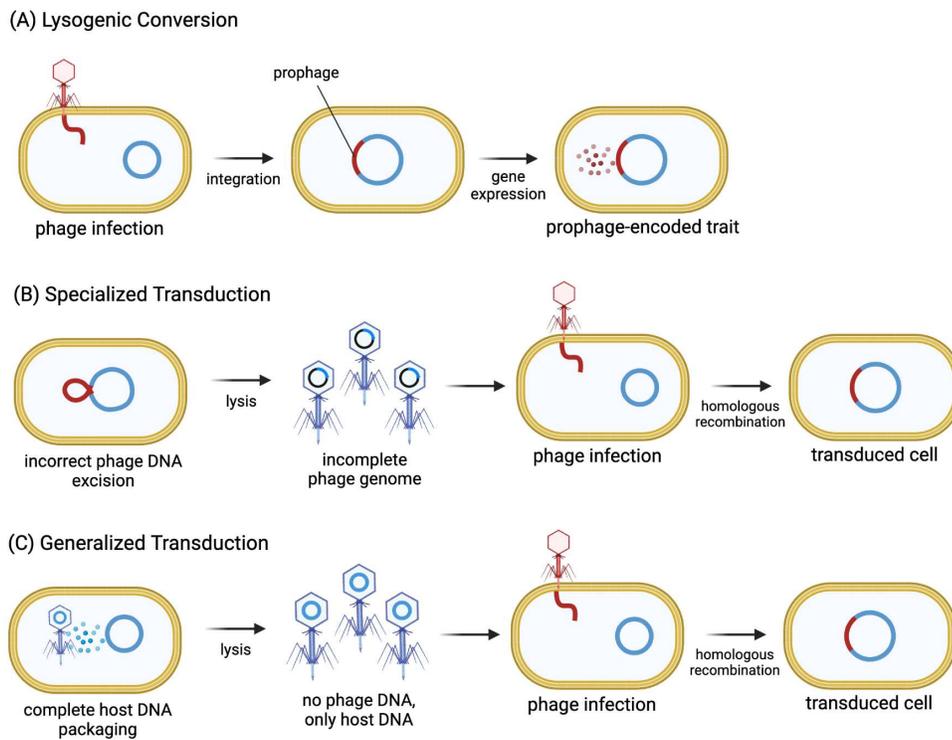
Specialized transduction occurs when a prophage inaccurately excises from the bacterial chromosome, often due to "illegitimate" recombination. This process causes a segment of both prophage and neighboring chromosomal DNA to be packaged into

a new phage particle. The probability of specialized transduction can vary greatly depending on the prophage's location on the chromosome due to the local density of repeated sequences. Interestingly, phage capsids can accommodate DNA slightly longer than the original phage genome, which can sometimes enable nearly complete phage genomes to be transferred. However, this may come at the cost of transferring minimal bacterial DNA.

As for generalized transduction, it occurs when the phage mistakenly packages bacterial DNA instead of its own, allowing any part of the bacterial chromosome, including rDNA, to be transferred. This DNA can then integrate into a new host's chromosome, typically through homologous recombination. Generalized transduction is common among phages that use the headful (pac) mechanism, which allows slightly more DNA to be packaged than the phage genome itself. Transduction rates tend to decrease in environments with high phage densities, as new infections by viable phages often kill the host before gene transfer can occur. This high transduction rate has raised concerns in phage therapy, as it can potentially spread antibiotic resistance or virulence genes between bacteria.

1.2.2. Factors Affecting Phage Induction

There are factors that can affect the transition of the phages from lysogenic to lytic cycle, including temperature, nutrient availability, salinity, aeration, ultraviolet radiation, heavy metals, environmental pollutants, superinfection, and host density. Regarding the effect of temperature, it has been reported that the phage life cycle showed changes according to seasonal patterns, which involve changes in temperatures, salinity, and nutrient availability. In the case of aeration, it is known that temperate phages are dominant in the anaerobic natural environment, such as the gut, deep sea, and deep soil (Knowles et al., 2016). However, it is hard to link the phage life cycle with the seasonal patterns and other conditions in a natural environment as it involves many complex factors.

Among all the stress conditions for phage induction, the use of UV light and antibiotics such as Mitomycin C are the most common and widely applied in the laboratory experiments. The exposure of UV is a DNA-damaging stress that can activate the host's SOS response and induce the switch from the lysogenic to lytic cycle. This is due to the fact that when the host experiences irreparable DNA damage, most prophages detect the stress signals and shift from dormant state to active replication for their own survival by escaping the host cell that is likely to die (Brady et al., 2021). Another abiotic compound that also induces bacterial SOS response is antibiotics like Mitomycin C (Jancheva & Böttcher, 2021). Additionally, several studies indicate heavy metals such as copper ions in freshwater environments

significantly trigger the induction of phages. It has been reported that copper ions from the addition of copper sulfate is a significant inducer for lysogenic cyanobacterial cells as it increases the phage release (Lee et al., 2006).

1.3. Strategies for Studying Phages in Microbial Community

1.3.1. Metagenomic Analysis

Metagenomics is a powerful approach for studying phages in complex environments. This approach allows researchers to directly analyze the collective genetic material from environmental samples. This is particularly advantageous for studying phages because many of them are difficult or impossible to culture in the laboratory. By using metagenomics, researchers can uncover the diversity, abundance, and functional roles of phages in ecosystems, such as oceans, soils, and the human microbiome. This method involves sequencing the DNA from environmental samples, which includes both phage and microbial DNA, followed by bioinformatic analysis to identify viral sequences, reconstruct viral genomes, and investigate phage-host interactions.

Additionally, metagenomic analysis has been instrumental in understanding phage dynamics within the biogas microbiome. It has been reported that *Firmicutes* and *Actinobacteria* are particularly vulnerable to phage infection in biogas production systems (Willenbücher et al., 2022). Therefore, metagenomic analysis helped uncover the interaction between phages and their hosts. Metagenomics can be a powerful approach to predict phage-host interactions by analyzing the genetic material within microbial communities. One primary method is CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) spacer analysis, which leverages the CRISPR-Cas system found in bacterial and archaeal genomes. These systems capture and store short sequences because when a bacterium encounters a phage, it may incorporate small fragments (spacers) of the phage's DNA into its own CRISPR array, and by matching spacers to viral sequences in the metagenomic data, it can be used to identify specific phage-host relationships based on past interactions. Therefore, spacers can be the representation of biological records of past phage-bacteria interactions, linking specific phages to their hosts (Dion et al., 2021). On the other side, metagenomics can provide insights into the functional roles of phages. For example, it can reveal genes involved in host metabolism, virulence, or stress response, which are useful to understand how phages influence the efficiency and stability of the biogas process.

1.3.2. Single-Cell Analysis

Single-cell analysis is an emerging and powerful technique for studying phages, offering detailed insights into the interactions between individual phages and their bacterial hosts. Metagenomic analysis for studying phages relies on bulk measurements, which average the behavior of large populations of cells and phages. In contrast, single-cell analysis enables researchers to examine the dynamics of phage-host interactions at the individual cell level, uncovering the variability and heterogeneity within microbial communities.

Single-cell techniques, such as microfluidics and single-cell genomics, allow researchers to isolate and analyze individual host cells and their associated phages. This approach provides critical information about how phages infect, replicate, and influence their bacterial hosts under various conditions. It also sheds light on the diversity of phage strategies, such as lytic and lysogenic cycles, within single cells that may be masked in population-level studies.

1.4. Aims of the Work

This study represents a first attempt of an innovative approach that combines metagenomics and single-cell analysis to investigate the diversity and potential roles of phages within the anaerobic digestion microbiota, particularly in biogas production systems. By characterizing phage populations and their interactions with microbial hosts, this study seeks to enhance the understanding of how phages influence microbial community. Specifically, the purpose of this study is to identify the various phages present in anaerobic digestion system, examining both virulent and temperate types, as well as to determine phage-host relationships by linking specific phages to their bacterial or archaeal hosts. Stressful conditions, including UV exposure, Mitomycin C, heavy metal ($CuCl_2$), and hydrogen peroxide (H_2O_2) were imposed in order to induce temperate phages within the microbial community. These treatments can induce the phages in the community, facilitating their identification and allowing for a more comprehensive assessment of viral diversity, and abundance, as well as linking specific phages with their bacterial or archaeal hosts.

2. MATERIALS AND METHODS

2.1. Experimental Design

2.1.1. Inoculum

The microbial community used in this experiment was derived through a simplification process of an anaerobic microbiota from pilot-scale reactors to a fed-batch culture. This batch consists of basal medium (BA), vitamin solution, yeast extract, and sodium sulfide (Na_2S), with feeding of hydrogen gas (H_2) and carbon dioxide (CO_2) in 4:1 proportion. The culture was incubated at 55°C with continuous agitation at 130 rpm. In this experiment, the experimental batches were reinoculated by using 15% of the final volume.

2.1.2. Preliminary Selection of Phage Induction

The choice of treatments for phage induction was based on a thorough review of existing scientific literature, especially from the previous research regarding metagenomic analysis of the virome composition in anaerobic digestion (Rossi et al., 2022). Four different treatments were selected including UV exposure, Mitomycin C, $CuCl_2$, and H_2O_2. These DNA-damaging stresses were applied to induce temperate viruses to enter the lytic cycle. The duration of UV exposure, as well as the concentration of Mitomycin C, $CuCl_2$, and H_2O_2 were selected based on the result of pre-trials considering the growth rate and methane production.

2.1.3. Experimental Set Up and Phage Induction

The experimental set-up was designed with four different treatments conducted in 3 biological replicates in 1 L bottles with an initial working volume of 250 mL. Before inoculation, the basal medium was flushed with N_2 for 10 minutes in order to ensure an anaerobic environment. Each bottle was inoculated with 37.5 mL of inoculum and 2.5 mL of yeast extract, vitamin solution, and Na_2S. For each condition, a bottle without inoculum was set as blank in order to check the reliability of the growth rate. Fed-batch reactors were incubated at 55°C with a constant agitation of 130 rpm in an orbital shaker (NBIOTEK NB-T205L, South Korea). The treatments refer to four different stress conditions for phage induction, including ultraviolet light exposure (1 hour), antibiotic (2 µg/ml of Mitomycin C), heavy metal (0.1 mg/L of $CuCl_2$), and hydrogen peroxide (0.1 mM of H_2O_2). The cultures were fed daily with 150 mL of carbon dioxide (CO_2) and 600 mL of hydrogen gas (H_2) (**Figure 4).**
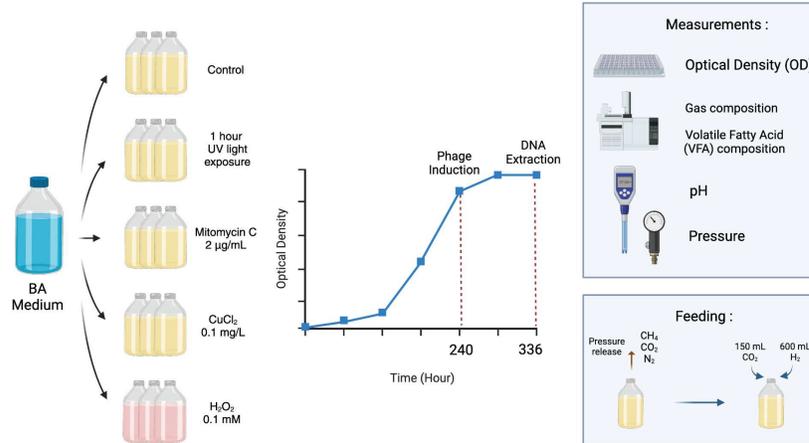
Figure 4. SCheme of the experimental design with applied treatments and timing (Created with BioRender.com)

2.2. Analytic Measurements

The gas composition in the batches was assessed daily, which includes the measurement of gas pressure and gas composition. The gas pressure was measured pre and post feeding. As for the measurement of gas composition, 10 mL of gas from fed-batch bottles headspace were collected before feeding. The measurement was done by using a gas chromatograph (8860 GC, Agilent Technologies, CA, USA) equipped with a thermal conductivity detector (TCD). The gas separation was performed using three micropacked columns, namely Hayesep Q (1.5 m, 1/16" OD, 1.0 mm ID), Hayesep N (0.5 m, 1/16" OD, 1 mm ID), and MolSieve 5Å (1.5 m, 1/16" OD, 1.0 mm ID), all using helium as carrier gas.

The growth rate assessment of the cultures was done by measuring optical density (OD) daily. The OD measurement was conducted using the Tecan Spark Multimode Microplate Reader (Tecan Life Sciences, Switzerland). The pH of the liquid sample was also measured before the OD measurement to maintain the optimal pH of the culture. Additionally, 1 mL of liquid sample was also collected every 3 days before the induction and collected daily after the induction for VFA assessment. As for the VFA analysis, 40 µL of orthophosphoric acid was mixed with the sample and centrifuged at 13.000 rcf at 4°C for 12 minutes. Next, 1 mL of sample was transferred to a glass vial with the addition of 100 µL isocaproic acid as internal standard. VFA concentrations were determined using a gas chromatograph (8860 GC, Agilent Technologies, CA, USA) equipped with a flame ionization detector (FID) and hydrogen as the carrier gas.

15

2.3. DNA Isolation and Sequencing

Prior to the metagenomics analysis, DNA extraction was performed 3 days after the phage induction, which is day 13 after inoculation. The sample was centrifuged at 13,000 g to separate the cells from the liquid medium resulting pellets at the bottom of the tube and the supernatant. The pellets contains prokaryotic cells and viruses meanwhile the supernatant contains the culture medium, free-living viruses, and other soluble components such as salt. Therefore, the supernatant was collected for freeze-drying prior to DNA extraction in order to separate the sample from the culture medium. The freeze-dried supernatant was resuspended by 200 μL of phenol (Sigma-Aldrich, St. Louis, MO, USA).

DNA from the pellet and supernatant were extracted by using the DNeasy PowerSoil® Pro Kit (QIAGEN GmbH, Hilden, Germany). Qualitative and quantitative assessment of each extracted DNA was done by using Nanodrop2000 (ThermoFisher Scientific, MA, USA) and Qubit Fluorometers (Invitrogen, CA, USA). As for the library preparation, it was performed using the Nextera DNA Flex Library Prep Kit (Illumina Inc, CA, USA) at the sequencing facility of the Department of Biology (University of Padova, Italy) and the samples were sequenced using the Illumina NovaSeq platform.

2.4. Metagenomic Analysis

DNA sequences of the supernatant were assembled into viral MAGs (VirMAGs). On the other hand, DNA sequences of the pellet were assembled into both MAGs and VirMAGs because they also contain DNA derived from prokaryotic cells. The raw Illumina reads from pellet and supernatant were initially processed using Trimmomatic (v0.39-1) and bbmap (38.90) to eliminate low-quality sequences. The filtered reads from the pellet were assembled into scaffolds employing SPAdes v3.15.50, utilizing both the forward and reverse fastq files of the filtered reads. Quast (v5.2) was used to inspect the quality of assembly. The obtained scaffolds were classified into MAGs through five distinct binning tools, including Metabat v1, Metabat v2, CONCOCT (v1.1.0), MaxBin (2.2.7), and Vamb (v3.0.9), employing a sorted bam file. A dereplication step was performed in order to identify and remove redundant MAGs. The quality of the remaining MAGs was assessed using CheckM (v1.1.3-1) which calculated the completeness and contamination levels. Only high-quality (completeness > 90% and contamination < 5%) and medium-quality (completeness > 50% and contamination < 1%) MAGs according to the MIMAG. were retained for further analysis (Bowers et al., 2017). To measure the coverage of each MAG, the reads from the samples were aligned onto the MAGs using Bowtie2

and SAMtools (v1.12-1) with default parameters. CoverM was then used to calculate the coverage of each MAG based on the alignment results. The taxonomic classification was performed using GTDB-Tk (v1.7.015). Gene sequences in each MAG were identified with Prodigal (2.6.319) and their functionality was evaluated with eggNOG-mapper (2.0.120). Functional characterization of stress response genes was conducted by performing a BLAST search against a dedicated stress response gene database (Gorecki et al., 2021). Viral metagenome-assembled genomes (virMAGs) were constructed using both pellet and supernatant samples following an identical protocol. Raw reads were assembled using SPAdes with metaviral parameters, and contigs were processed through several tools for prophage identification and viral genome prediction. These included VIBRANT, VirSorter2, DeepVirFinder, CenoteTaker3, PPR-Meta, CheckV, CoCo-net, Jaeger, PHASTEST (via API), Phold, and Genomad. Contigs were filtered based on length, quality, and scoring thresholds specific to each tool. For dereplication, an all-versus-all BLASTn approach was used with an ANIm-based clustering threshold of ≥95% identity. Coverage was assessed using Bowtie2 and samtools, while taxonomy assignment utilized PhaGCN2.0. The identification of the CRISPR-Cas system was carried out by detecting and analyzing the cas operons with the help of CRISPRCasFinder v4.3.2 (https://github.com/dcouvin/CRISPRCasFinder) and CRISPRDetect v3.0 (https://github.com/ambarishbiswas/CRISPRDetect_3.0). CRISPR spacer matches were determined through BLASTn, and functional annotations were refined using rpsblast+ against conserved domain databases. Results from the pellet and supernatant workflows were merged during dereplication to produce high-quality, non-redundant virMAGs for downstream analysis. Additional phage-host prediction was done based on CHERRY database by using PhaBOX (https://github.com/KennthShang/PhaBOX)

2.5. Single-Cell Analysis

Single-cell analysis involves several essential steps to study the molecular profile of individual cells. These steps include pre-treatment such as sonication and filtering. Sonication is a commonly used method to disaggregate cell samples without damaging the membranes and iwas performed with 3 sonication cycles and 10-second pause between each cycle. Filtering using a 5 μm filter was done to remove unwanted cell aggregates and result in a more homogeneous suspension of individual cells. Then, in order to assess both the quality and quantity of cells, a fluorescence microscopy assay with cell counting by using a Burker chamber was carried out. Afterward, the sample was subjected to encapsulation in semi-permeable capsules (SPC) by using the Onyx Droplet Generator (Atrandi Bioscience Inc., Vilnius, Lithuania) to isolate individual cells and minimize cross-contamination, as well as

allowing for the parallel processing of thousands of cells. Next, DNA extraction, whole genome amplification, and library preparation were performed using the Single-Microbe DNA Barcoding kit (Atrandi Bioscience Inc., Vilnius, Lithuania).

The initial step in the computational analysis involves demultiplexing single-cell sequencing data. This process separates raw sequencing reads based on barcodes located on the reverse strand of each paired-end read. Each cell is ideally associated with a unique barcode, enabling the identification and segregation of sequences from individual cells into separate FASTQ files. Demultiplexing was conducted using Ultraplex (v1.2.9) for each sample, leveraging a comprehensive library of 331,776 barcodes. A minimum threshold of 2,500 reads per barcode was applied to filter sequences for further analysis. This threshold was chosen to provide approximately 0.5X genome coverage for a genome size of 1 million base pairs, striking a balance between sensitivity and reducing noise for subsequent steps. The filtered reads undergo preprocessing with Trimmomatic (v0.39), which removed the first 10 bases from each read and trimmed low-quality bases at both ends until a quality score of at least 20 was achieved (parameters: HEADCROP:10, LEADING:20, TRAILING:20). Additionally, a 4-bases sliding window approach was used to cut reads when the average quality dropped below 20 (SLIDINGWINDOW:4:20), and reads shorter than 70 bases were discarded (MINLEN:70). Quality assessment of the raw, demultiplexed, and trimmed reads was performed using FastQC (v0.12.1), and a summary of the results was compiled into a report with MultiQC (v1.17).

3.   RESULTS AND DISCUSSION

3.1. Biochemical Results

In this study, the biochemical parameters of the microbial community were measured to understand various aspects of microbial activity, composition, and the overall stability of the batches. Biochemical measurements such as methane (CH_4) production and volatile fatty acids (VFAs) are essential to understand the metabolic functions and efficiency of the microbial community.

3.1.1. Growth Rate and Cumulative CH_4 Production

The growth rate of the cultures were assessed on a daily basis by the measurement of optical density (OD) at - 600nm. The primary objective was to evaluate the growth of the culture and how different treatments affected the culture's growth, as well as to compare these effects with the controls (**Figure 5).** Before induction, all cultures were maintained under identical conditions to ensure consistency, with the induction process carried out after 240 hours. Following induction, the experiment was divided into five treatment groups, each with three replicates: Control (untreated), UV, Mitomycin C, $CuCl_2$, and H_2O_2.

Prior to induction, all cultures exhibited similar growth rates, confirming uniform initial conditions. However, immediately after induction at hour 240, cultures treated with UV, Mitomycin C, and $CuCl_2$ showed a temporary decline in growth (**Figure 5**). Despite this, growth steadily increased during the remaining observation period, reaching levels comparable to the untreated control by the end of the 300-hour experiment. This recovery suggests that the microbial communities were able to withstand and adapt to the phage induction treatments. Moreover, this result indicate that these induction factors did not significantly inhibit microbial growth in the long term, suggesting their effects on growth were either neutral or only mildly disruptive.

In contrast, the H_2O_2-treated group shows a markedly different pattern (**Figure 5**). While it initially follows the same growth trend as the other treatments, its OD declines sharply after the phage induction, indicating a reduction in cell concentration. This suggests that H_2O_2 exerts a strong inhibitory or toxic effect on the culture over time, likely due to oxidative stress harming the microbial cells. This result is expected given the anaerobic nature of these microorganisms. By the end of the experiment, the H_2O_2-treated group has the lowest OD among all groups, highlighting its detrimental impact. Overall, the graph demonstrates that while UV,

Mitomycin C, and CuCl$_2$ have minimal effects on microbial growth, H$_2$O$_2$ significantly disrupts it.
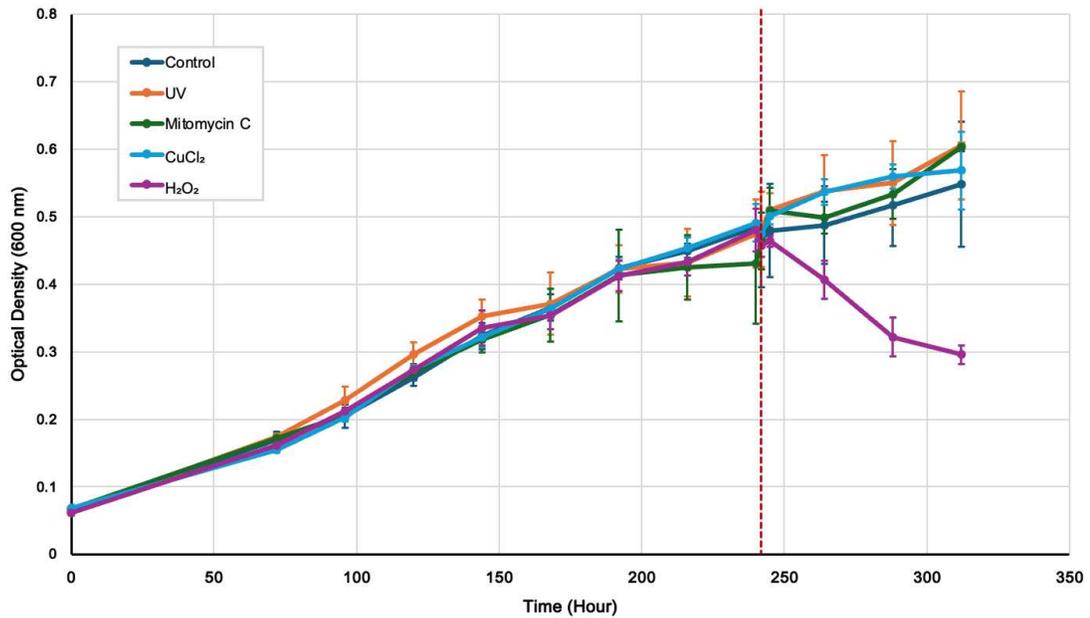
Figure 5. **Growth rate measured by optical density (OD) at 600 nm**. Points correspond to the average of experimental data for each treatment and the bars are the standard deviation. The red dash corresponds to the starting time of induction (240 hours after inoculation).

The CH$_4$ accumulation graph shows the production of methane (CH$_4$) over time across different treatment groups (**Figure 6**). In the early stages, before the induction, all treatments including the Control, show similar rates of methane production. However, after the induction, differences between treatments become more apparent. As for the Control, UV, Mitomycin C, and CuCl$_2$ treatments, methane production continues to increase steadily after the induction. CuCl$_2$-treated cultures achieve the highest methane production by the end of the observation period, while the other groups (Control, UV, Mitomycin C) exhibit slightly lower but comparable levels of methane accumulation. The similar levels of methane production observed across the Control, UV, Mitomycin C, and CuCl$_2$ treatments, despite slight variations, suggest that the phage-inducing treatments did not significantly disrupt the core methanogenic activity of the microbial community.
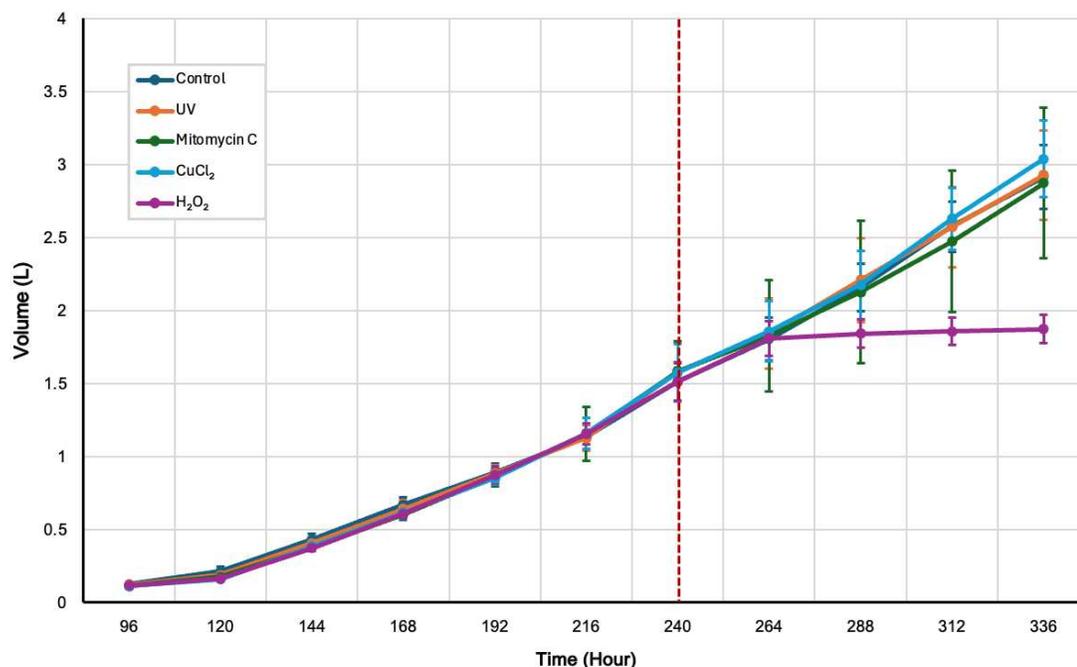
Figure 6. **Accumulated CH$_4$ production.** Points correspond to the average of experimental data for each treatment and the bars are the standard deviation. The red dash corresponds to the starting time of induction.

In contrast, the H$_2$O$_2$ treatment shows a notable different pattern. After the induction, methane production in the H$_2$O$_2$-treated group leveled off and eventually stops increasing. This plateau suggests that H$_2$O$_2$ has a strong effect on the culture in terms of methane production, potentially due to oxidative stress impacting the microbial community responsible for methane production. The strong effect of H$_2$O$_2$ might also occur because it can oxidize the organic compounds that serve as substrates for microbial metabolism, which leads to substrate depletion. This dual impact, oxidative stress and substrate depletion, can result in the disruption of biomethane production (Achouri et al., 2021).

3.1.2. Volatile Fatty Acid Profile

The production and consumption of volatile fatty acid (VFA) were monitored throughout the experiment every three days before the phage induction and every day after the induction. VFAs are critical intermediates in anaerobic digestion process, representing the fermentation of organic matter before methane production. The x-axis represents the time elapsed in hours, ranging from 0 to 312 hours, while the y-axis shows the amount of each VFA.
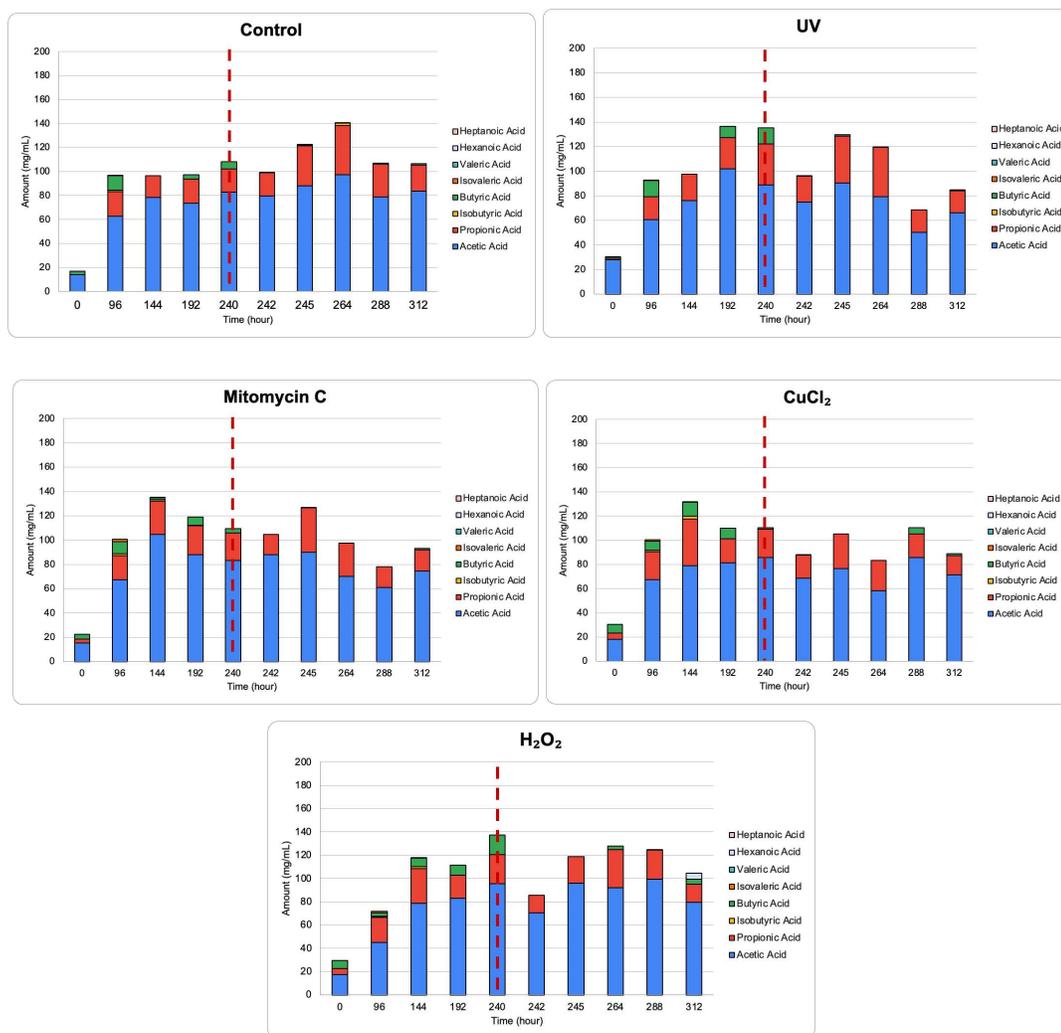
Figure 7. **Volatile Fatty Acids (VFAs) Profile.** Stacked bar charts showing the concentrations (average of replicates) of each VFAs (mg/mL) measured at different time points for each treatment condition. The red dash corresponds to the starting time of induction.

The result shows that acetate was the dominant VFA across all the treatments. In the Control setup, the VFAs demonstrate a relatively stable trend over time. Acetic acid dominates the VFA profile. Propionic acid and butyric acid are present in smaller but consistent quantities. Under UV treatment, a similar VFA profile is observed, with acetic acid remaining dominant. However, fluctuations in VFA concentrations occur after the phage induction which was performed at 240 hours. The UV-induced stress may have triggered shifts in the microbial community that led to changes in the amounts of VFA being produced. However, the microbial community appears to recover, as evidenced by the increase of VFAs at later stages. In the Mitomycin C-treated condition, a noticeable reduction in total VFAs is observed after phage

induction, particularly in acetic acid concentrations. Mitomycin C, known for its ability to induce prophages and damage the bacterial DNA, which can lead to the decrease of VFA production. A similar trend was also shown in $CuCl_2$-treated condition, after the phage induction, the VFA concentrations decreased, suggesting that prolonged exposure to heavy metal stress has negatively impacted the microbial community. As for the H_2O_2-treated condition, a sharp decline was shown in total VFA concentration after the phage induction though it increased again in less than 24 hours. Compared to other conditions, the H_2O_2-treated samples are have higher concentration of acetate and total VFA. As previously found, the methane yield stopped increasing after the addition of H_2O_2. The low methane yield and the accumulation of VFA are coherent with the idea that the disruption of acetoclastic archaea and reflected that there is a kinetic imbalance between acid producers and consumers in the culture (Boe et al., 2010).

3.2. Microbiome Dynamic and Activity

A total of 33 metagenome-assembled genomes (MAGs) were identified across three experimental conditions, including UV, Mitomycin C, and $CuCl_2$. These MAGs were assembled from the DNA of the pellets which contain prokaryotic cells and phages. Initially, a treatment with H_2O_2 (hydrogen peroxide) was included to induce phage activity. However, in this case, the DNA concentration extracted from the H_2O_2-treated samples was extremely low. As a result, there was insufficient DNA for library preparation and sequencing, meaning that the samples could not be analyzed further to assess phage induction or microbial community composition in this treatment group. The graph representing the taxonomic composition and relative abundance of 33 MAGs was reconstructed **(Figure 9)**. This visualization provides an overview of the microbial community, allowing for the identification of dominant organisms and their potential roles in the community.

The most abundant MAG is *Methanothermobacter thermautotrophicus,* a thermophilic, hydrogenotrophic archaeon. *M. thermautotrophicus* was the dominant methanogen in H_2/CO_2 enriched mixed cultures ensuring the methane production through hydrogenotrophic methanogenesis (Bu et al., 2018). It is a crucial organism for industrial production of biomethane in the thermophilic anaerobic digestion process and biogenic methane in the natural gas field. This archeon is essential for efficient CO_2 conversion in biogas upgrading systems and its abundance in this community highlights its crucial role in increasing the methane yield. The second most abundant MAG is identified as *Caldanaerobacter subterraneus.* This bacterium is thermophilic and fermentative, producing primarily H_2, CO_2 and acetate. A previous study suggested that horizontal gene transfer (HGT) during the evolution of

Caldanaerobacter subterraneus has boosted its inherent adaptability, enabling it to thrive in naturally occurring unfavorable conditions (Sant'Anna et al., 2015). This increased adaptability, conferred by acquired genetic material, has likely provided *C. subterraneus* with specialized functions and metabolic pathways that allow it to tolerate environmental stressors. This HGT may have involved interactions with phages, as they are well-known vectors of genetic material between bacteria and can introduce new genes that boost host resilience.

Figure 8. Phylogenetic Analysis and **Relative Abundance of MAGs.** The outer ring highlights the average relative abundance of MAGs reported as a color gradient. The inner ring shows the taxonomic assignment of the MAGs at the phylum level.

Z-score analysis of MAG abundances across different conditions provides insights into how each MAG responds to different phage induction treatments. The purpose of this comparison is to assess how the abundance of specific MAGs changes under different treatments relative to the baseline (Control). A Z-score below 0 indicates that the abundance in a given sample is lower than in the Control condition, providing

a quantitative measure of a decline microbial population. Positive z-scores indicate higher-than-average abundance relative to the control condition, while negative z-scores suggest lower-than-average abundance compared to the control. In the UV-treated culture, the most abundant MAG was identified as *Bacillota_G* sp. 28, as indicated by their relatively high positive z-score (**Figure 9**). The high z-score suggests that *Bacillota_G* sp. 28 not only survived but potentially thrived under UV-treated conditions compared to other microbial species in the community. The resilience of *Bacillota_G* sp. 28 to UV treatment could point to specific adaptations, such as efficient DNA repair mechanisms or protective cellular structures, that mitigate UV-induced damage. This is confirmed by the functional gene annotation which revealed that *Bacillota_G* sp. 28 have 7 genes related to the radiation protection (**Figure S5).**
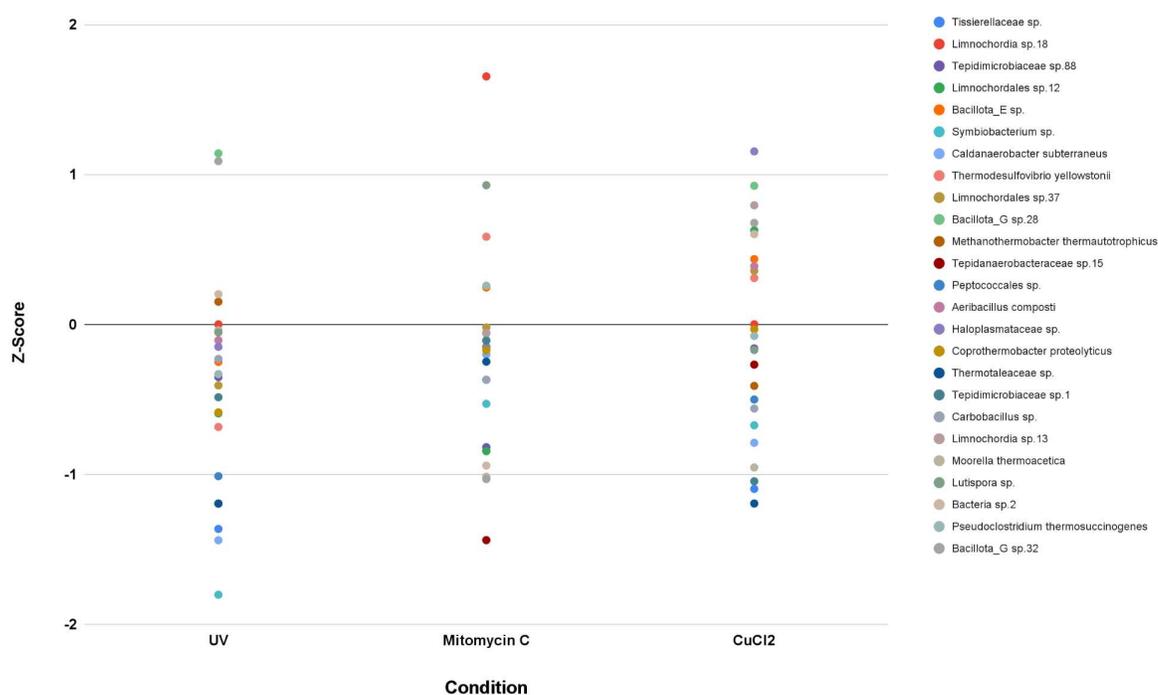
Figure 9. **Z-Scores of MAGs Across Experimental Conditions.** Each dot represents a different MAG, with colors and symbols distinguishing between MAGs.

As for Mitomycin C, *Limnochordia* sp. 18 was notably abundant, with one of the highest positive z-scores observed in the dataset (**Figure 9**), indicating that this microbe continued to grow despite the addition of antibiotics. The class Limnochordia harbors a single cultivated member, the mesophilic *Limnochorda pilosa* and it was previously described as potential homoacetogenic bacteria (Campanaro et al., 2020). The high abundance of *Limnochordia* sp. 18 after

Mitomycin C treatment can be attributed to their efficient DNA repair systems. Limnochordia has genomic region that contains ExbBD/TonB machinery essential for nutrient acquisition and adaptation to specific harsh environmental conditions and cohesin gene that involved in structural and functional roles such as DNA repair and gene regulation (Taib et al., 2020) Finally, in $CuCl_2$ condition, the most abundant MAG is *Haloplasmataceae sp.* (**Figure 9**). In details, Haloplasmataceae is a family within the order Haloplasmatales and currently includes one single genus and species, which is *Haloplasma contractile*. This organism is known to be capable of thriving in extreme environments, particularly under hypersaline or alkaline conditions (Antunes et al., 2011). Its metabolic versatility and ability to adapt to harsh environmental stressors may explain its abundance after $CuCl_2$. Furthermore, as $CuCl_2$ can induce phage activity (Lee et al., 2006), this organism might possess phage resistance mechanisms or other adaptive strategies that allow it to persist and thrive despite potential phage-induced disruptions.

In addition to highlighting the MAGs that are most abundant in each condition, it is equally important to identify those with the lowest abundance values in each stress condition compared to the Control. In the UV-treated samples, *Symbiobacterium* sp. exhibited the lowest abundance value among all MAGs (**Figure 9**). This decline in abundance may be attributed to UV exposure triggering prophage activation in lysogenic bacteria, leading to phage-induced lysis and a reduction in the bacterial population. Additionally, this pattern was not unique to UV treatment, *Symbiobacterium* sp. also ranked among the lowest abundance MAGs in samples treated with Mitomycin C and $CuCl_2$. This consistent population decline across multiple treatments suggests that *Symbiobacterium* sp. may be particularly susceptible to external stressors that induce prophage activation, or it may lack robust mechanisms to recover from such disturbances compared to other MAGs.

Interestingly, despite being the second most abundant MAGs, *Caldanaerobacter subterraneus* was among the lowest abundance MAGs in stress condition compared to the Control (**Figure 9**). This could indicate that each phage-induction treatment may have triggered the lytic cycle of phages infecting this organism and resulted in a decline of its population. Alternatively, it could also imply that *Caldanaerobacter subterraneus* lacks effective mechanisms to recover from phage-related stressors since this organism is lack of genes related to stress resistance (**Appendix Figure S5)**, making it more vulnerable to phage-induced lysis compared to other MAGs in the community.

3.3. Viral MAGs in the Microbial Community

A total of 129 viral MAGs (virMAGs) were reconstructed. Based on their life cycle category, 98 virMAGs are temperate, and 30 virMAGs are virulent (**Appendix Table S4**). The high number of temperate phages suggests that the phages integrate their genome into the host's DNA, as a prophage and suggests the possibility of lytic activation under stress conditions. According to a model explaining transition between lytic and lysogenic life cycles, virus switches from lytic to lysogenic when its population grows faster as prophage than as virions while stress conditions can also trigger lytic activation, leading to virus-mediated horizontal gene transfer (HGT) through transduction because the phage carries not just its own viral genes, but also fragments of the host's genetic material (Roughgarden, 2024). Although the stress conditions applied to the microbial community did not significantly inhibit microbial growth, it still created fluctuations in the microbial community and unstable environment. The stress conditions induce the lytic activation which enable phage-mediated transduction leading to the addition of new genetic material of the host, including genes that confer stress resistance or enhance metabolic versatility (Maslov & Sneppen, 2015). In contrast, the smaller number of virulent virMAGs, which follow a lytic cycle by rapidly replicating and lysing host cells, reflects a less favorable life cycle under the given conditions. The relative scarcity of virulent phages may indicate that while some level of active infection and lysis occurs, it is not the dominant mechanism shaping microbial populations. Instead, the result suggests that temperate phages play a more significant role in maintaining viral presence within this community, offering insight into the ecological balance between phages and their bacterial hosts in these microbial communities.
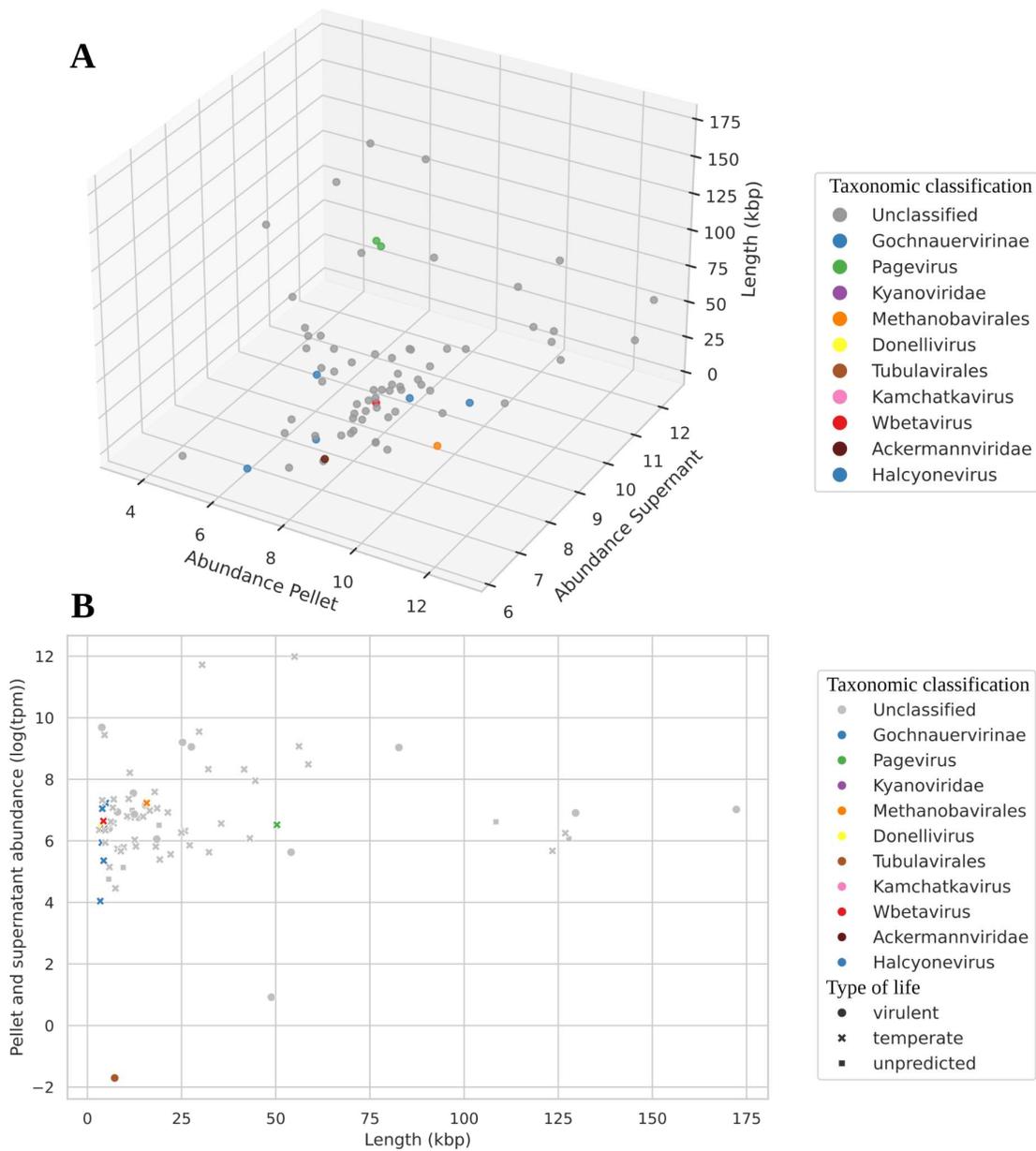
Figure 10. **Abundance of Viral MAGs Relative to Genomic Length. (A)** 3D Scatter plot illustrating the relationship between the abundance of viral MAGs (y-axis) and their genomic lengths (x-axis). Each dot represents a different Viral MAG, with colors distinguishing between Viral MAGs. The abundance measured is the log of the median between supernatant and pellet in TPM units. **(B)** Distribution of virMAGs by the genome length.

The scatter plot consisting of virMAGs abundance reveals that the viral community is predominantly composed of unclassified phages (**Figure 10**). This abundance of unclassified phages suggests that this aspect of the microbiota is understudied. This

also highlights the importance of genomic and functional studies of phages to uncover the roles, host specificity, and potential impacts of these phages in shaping microbial communities, especially in biogas production systems. The most abundant genera identified virMAGs are *Pagevirus*. This genera of viruses are known to have a complete genome size of around 45-50 kbp with 40.5% GC content (Ladzekpo et al., 2015). Gene annotation results from this experiment reveal that one of the two *Pagevirus* virMAGs carries an integrase gene within its genome. The discovery of integrase in this context may reflect its role in horizontal gene transfer (HGT) and indicate its lysogenic capabilities. Moreover, in the context of the genome length, these virMAGs genome length matches the references which can indicate that the assembly and binning processes successfully captured most, if not all, of the viral genome.
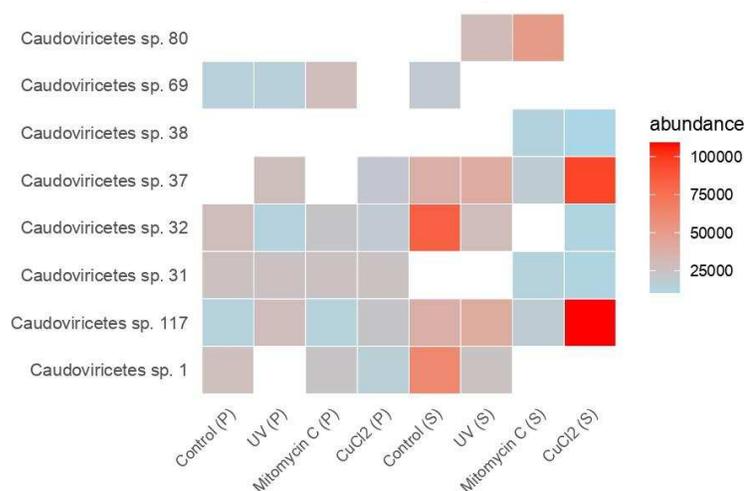
Figure 11. **Top 5 Most Abundant Viral MAGs in Each Different Condition**. Viral MAGs assembled from pellets (P) and Viral MAGs assembled from supernatants (S)

There is a limit to how much DNA can be packaged in a capsid, and therefore phage genomes are generally short, typically in the range 20–100 kbp (McNair et al., 2019). In order to reduce the number of mispredicted and partially assembled viral genomes, virMAGs with a length higher than 3 kbp were selected (**Appendix Figure S2**). Throughout all the conditions, virMAGs that are related to Caudoviricetes were the most abundant (**Figure 11**). Caudoviricetes is a class of viruses with head-tail morphology, which includes several families such as Straboviridae, Drexlerviridae, Demerecviridae, and Ackermannviridae (Zhu et al., 2022). Interestingly, while the 3D scatter plot reveals that one of the most abundant virMAGs is identified as *Pagevirus*, this particular virMAG does not rank among the top five most abundant virMAGs in any specific condition. This discrepancy may be explained by the classification process, where the unclassified virMAGs were ultimately assigned a name based on

the closest identifiable taxon, in this case, *Caudoviricetes* and apparently these unclassified virMAGs are very abundant which made them present among the highest abundance virMAGs in all different conditions.

3.4. Functional Categories of Proteins Encoded in Viral MAGs

Phage gene annotation was performed to clarify the genes encoding proteins that are involved in infection, replication, and host manipulation. This step is important to explore the genes that are potentially involved in horizontal gene transfer (HGT) and contribute to enhancing the metabolic process of the host. Functional analysis was performed on protein-coding genes (**Appendix Figure S4**). The genes were categorized according to their function. The category includes DNA, RNA and nucleotide metabolism, head and packaging, auxiliary, integration and excision, transcription regulation, connector, tail, and lysis. Genes encoding proteins that did not fall into the predefined categories were collectively classified under the group labeled "Other" and genes with unidentified functions were categorized under "Unknown Function."

In addition to the analysis of functional genes annotation, the host of the phages were predicted through CRISPR-based and AAI-based. The prediction was done by CRISPR spacer matching. However, CRISPR-based tools for predicting phage-host relationships often struggle with low recall, meaning they fail to identify many true phage-host interactions. This limitation arises primarily from two factors. First, not all bacteria possess CRISPR systems and some entire groups of bacteria lack CRISPR arrays entirely, making it impossible to use this method to predict their phage hosts. Second, even among bacteria with CRISPR systems, the stored spacer sequences, which is the record of past phage infections, tend to be highly variable. This variability exists even between closely related bacterial strains, reducing the effectiveness of these tools. Consequently, the absence of comprehensive and consistent CRISPR data across bacterial genomes significantly limits the ability of CRISPR-based methods to identify phage-host interactions (Roux et al., 2023). Therefore, in order to cover the limitation from CRISPR-based prediction, AAI-based prediction was also carried out. AAI (Amino Acid Identity) is a measure of the genetic similarity between the protein sequences of two organisms, and it is commonly used to assess relatedness between bacterial genomes. In the context of phage-host interaction studies, AAI-based methods allows the prediction of potential bacterial hosts for viruses. These methods involve comparing the protein sequences encoded by phage genomes to those found in bacterial genomes. This approach is particularly relevant for temperate phages, which often integrate into the host genome and share conserved protein sequences with their bacterial hosts.

The top three phages carrying the highest number of genes from each category were selected. The distribution of the gene annotations of these phages are visualized in a stacked bar graph (**Figure 12**). Phages with the highest number of genes encoding lysis-associated proteins are *Caudoviricetes* sp. 132, *Caudoviricetes* sp. 80, and *Pagevirus* sp. 46. The high number of lysis-associated proteins suggested that the phages have high genetic capability for lysis. CRISPR-based host prediction revealed that *Caudoviricetes* sp. 132, *Caudoviricetes* sp. 80, *Pagevirus* sp. 46 have infected *Moorella* sp. and *Thermohydrogenium kirishiense,* and *Priestia megaterium* respectively (**Appendix Table S4)**. However, both *Thermohydrogenium kirishiense* and *Priestia megaterium* were not identified in this experiment suggesting that the host prediction might not have been fully accurate and the prediction model was probably based on limited data. On the other side, a species in *Moorella* genera was identified in this experiment which is *Moorella thermoacetica.* This MAG was one of the lowest abundance after phage induction with Mitomycin C and $CuCl_2$. *M. thermoacetica* is notable for utilizing the Wood–Ljungdahl pathway, a metabolic process that enables the fixation of CO_2 or CO into acetate, making it a critical player in carbon cycling and bioenergy production (Redl et al., 2017). Despite its functional significance, *M. thermoacetica* was observed to have one of the lowest abundances following phage induction treatments with Mitomycin C and $CuCl_2$. The fact that *M. thermoacetica* was infected by *Caudoviricetes* sp. 132 which carries eight lysis-related genes, provides significant insight into the potential reasons for the low abundance of *M. thermoacetica* after phage induction treatments. The identification of eight lysis-related genes in *Caudoviricetes* sp. 132 strongly suggests that this phage possesses a robust lytic machinery (**Appendix Figure S4**). These genes encode proteins such as endolysins and holins, which coordinate the degradation of the bacterial cell layer and facilitate the release of newly formed phage particles. This indicates that *M. thermoacetica* was likely subjected to phage-mediated lysis, which can explain its decline in population under these conditions.
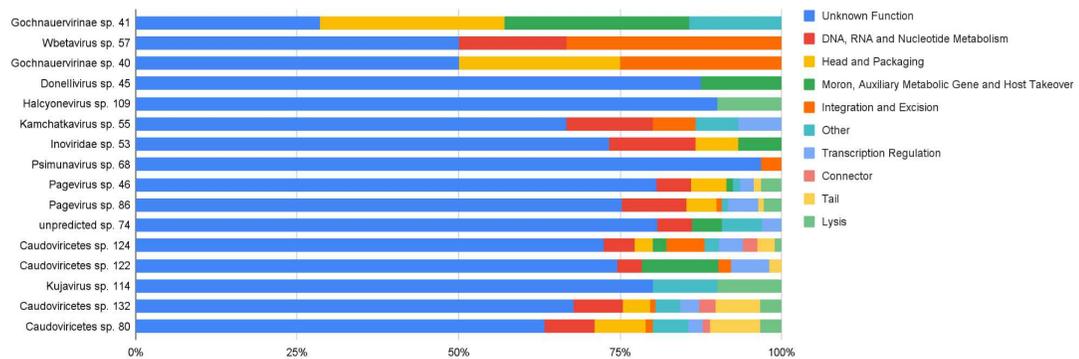
Figure 12. **Distribution of Gene Annotations by Functional Category.** The proportion of genes with established roles. Selected top three phages carrying the highest number of genes from each category. Those with unidentified roles grouped in 'Unknown Function', and those grouped under 'Other' due to their lack of association with predefined categories.

Additionally, phages with the highest number of genes encoding integration and excision-related proteins are *Caudoviricetes* sp. 124, *Wbetavirus* sp. 57, and *Gochnanauervirinae* sp. 40. According to CRISPR-based host prediction (**Appendix Table S3)**, *Caudoviricetes* sp. 124 infected *Caldanaerobacter subterraneus* and *Tepidanaerobacteraceae*. Both of these organisms were detected in this experiment, with *Caldanaerobacter subterraneus* standing out as the second most abundant MAG identified. The phage's substantial number of integration-related genes suggests a strong capacity for lysogeny, allowing it to integrate its genome into the host's DNA. Additionally, *Caudoviricetes* sp. 124 also carried 4 auxiliary metabolic genes (AMG) which are genes that phages can transfer to their host bacteria, potentially altering the host's metabolism. AMGs are known to confer various benefits, such as enhancing host survival, modulating the host's metabolic pathways, or increasing the phage's own replication efficiency. This further suggests that *Caudoviricetes* sp. 124 may play a role in shaping the metabolic activity of its host, providing a possible mechanism for symbiotic or beneficial interactions within the microbial community. Moreover, AAI-based host prediction (**Appendix Table S4**) suggested that *Wbetavirus* sp. 57 infected *Bacillus* genus and CRISPR-based host prediction revealed that *Gochnauervirinae* sp. 40 infected *Bacteroides ilei* which is not identified in this microbial community. These observations suggest that both AAI and CRISPR-based prediction methods might be biased or limited by the available reference data. Phage-host interactions are highly specific and when the reference dataset is incomplete or skewed, the predictions can misidentify hosts or assign hosts that are not relevant to the current microbial community.

Phages which carried the highest number of auxiliary metabolic genes (AMG) and moron genes are *Caudoviricetes* sp. 122 and *Gochanauervirinae* sp. 41. *Caudoviricetes* sp. 122, CRISPR-based host prediction reported that this virus infected *Hungateiclostridium thermocellum* which is also not identified in this experiment. As for *Gochanauervirinae* sp. 41, this phage infected *Thermoanaerobacteraceae* which is a family of bacteria consisting several genera such as *Caldanaerobacter*, *Thermoanaerobacter*, and *Desulfothermobacter*.
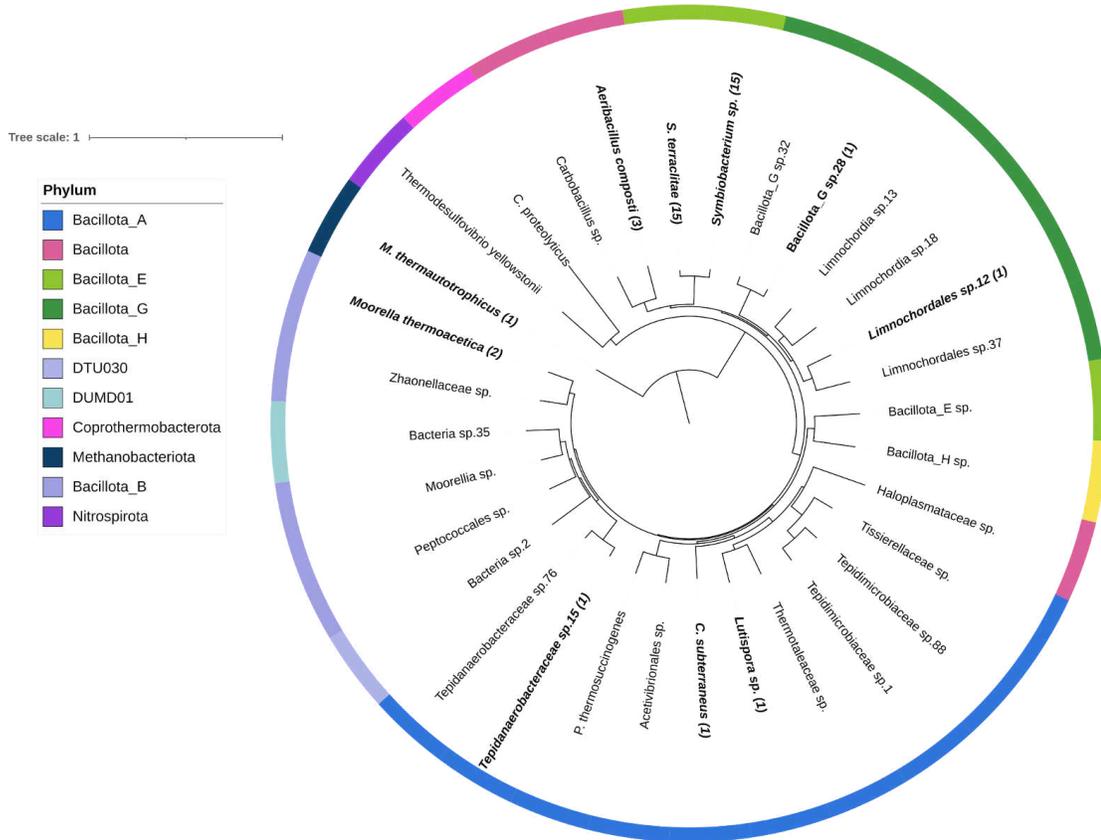
Figure 13. **Host-Phage Interactions in the Microbial Community.** The ring shows the microbial diversity according to their phylum. The number associated with each MAG represents the predicted number of phages that infect it.

There are 26 phages predicted to have a host. Out of all 33 identified MAGs, 10 of them were infected by phages (**Figure 13**). The phage-host prediction was done by CRISPR-based prediction and database from CHERRY using PhaBOX. *Symbiobacterium* spp. were the MAGs with the highest number of viruses (n=15) (**Appendix Table S4**). A previous study reported that *Symbiobacterium* correlate positively with biogas production and CH_4 in biogas, meaning that this genera is likely to contribute to enhancing the microbial processes involved in biogas production, particularly in the process of methanogenesis (Damtie et al., 2021). The

fact that *Symbiobacterium* is infected by 15 phages, of which 4 are virulent and 11 are temperate, along with the observation that *Symbiobacterium* were among the lowest abundance MAGs after phage induction treatments (UV, Mitomycin C, and $CuCl_2$), suggests that phage infection can highly impacts this bacterial population. Virulent phages, which follow the lytic cycle, would cause immediate bacterial cell lysis, leading to a reduction in *Symbiobacterium* abundance after induction. On the other side, the 11 temperate phages can integrate into the bacterial genome and enter the lysogenic cycle, where they remain dormant until induced. When phage induction occurs, these temperate phages likely enter the lytic cycle, causing the release of new phage particles and the lysis of infected *Symbiobacterium* cells, further decreasing their abundance.

One of the notable finding is that *Caldanaerobacter subterraneus*, the second most abundant MAGs in this experiment, were infected by a single temperate phage which is *Caudoviricetes* sp. 124 (**Appendix Table S4**). This phage is reported to have the highest number of integration-related genes (11 genes) and carried 4 AMGs, which is consistent with the fact that this phage is temperate. *Caldanaerobacter subterraneus* also shows the low abundance value after phage induction treatments, especially in the UV treatment of which *Caldanaerobacter subterraneus* has the lowest abundance out of all MAGs. Since temperate phages integrate into the bacterial genome and remain dormant in the lysogenic cycle, phage induction particularly UV exposure likely triggers these phages to enter the lytic cycle, causing the rapid replication of phages and subsequent lysis of the host cells. The fact that *Caldanaerobacter subterraneus* shows the lowest abundance after UV induction indicates that this process effectively activates the temperate phage, leading to a high decline in *Caldanaerobacter subterraneus* population.

The only identified archeon in this microbial community, *Methanothermobacter thermautotrophicus,* were predicted to be infected by *Psimunavirus* sp. 68 (**Appendix Table S4**). based on AAI prediction. *Psimunavirus* sp. 68 carried one integration-related gene which is consistent to the fact that this phage is temperate. It has been previously reported that this virus infected *Methanothermobacter*. It is a non-enveloped virus with head-tail structure. The head is about 55 nm in diameter. Tail is 210 nm in length and 10 nm in width, with a terminal knob and the genome size of this virus is 31 kbp with GC content of 45.5% (Luo et al., 2001).

3.5. Single-Cell Analysis Result

Single-cell analysis provides a powerful tool for uncovering detailed insights into phage-host interactions. Unlike metagenomics, which analyzes community-wide genomic data, single-cell analysis isolates and examines individual cells, enabling the direct observation of interactions between specific phages and their bacterial hosts. The results obtained from single-cell analysis can serve to validate and corroborate findings derived from metagenomic studies, enhancing the robustness of conclusions regarding phage-host interaction.

In the initial step of single-cell, semi-permeable capsule (SPC) was generated and it serves as an independent reaction chamber that is both easy to manipulate and monitor, enabling highly efficient amplification processes. Once created, SPCs remain stable under a wide range of reaction conditions and are well-suited for downstream analyses. These analyses can be conducted on millions of SPCs in parallel using standard laboratory equipment, allowing for bulk-scale processing. The experimental design aimed to sequence approximately 2,500 cells per run, with a target average coverage of 6X. Assuming an average microbial genome size of 2 Mbp and a read length of 220 bp per paired-end read, the sequencing goal was set at 130 million (M) paired reads. Sequencing of the control and mitomycin-treated samples was performed across two combined runs, yielding 232 M and 243 M raw paired reads, respectively (**Table 1**). The reason for the higher sequencing effort is that only 66.3% and 52.8% of the reads were successfully assigned to the samples during demultiplexing, likely due to barcode ligation inefficiencies that affected overall read recovery. After quality trimming, the paired reads from all samples were distributed across 4,718 to 11,737 barcodes, each containing more than 2,500 reads **(Table 1)**. This resulted in an estimated average coverage ranging from 0.73X to 1.8X, which is below the intended 6X target. The shortfall was likely caused by inaccuracies in cell concentration estimates obtained through standard microscopic methods. Meeting the target coverage will require additional sequencing efforts. This preliminary experiment highlights the challenges of optimizing single-cell sequencing workflows and underscores the need to fine-tune experimental conditions for improved results. The study was intentionally designed to identify such obstacles and refine sequencing parameters, paving the way for more reliable and effective methodologies in future experiments.

Table 1. Number of paired reads and estimated average coverage of each sample based on single-cell sequencing.

	Sequenced raw pair-reads	Reads recovered during demultiplexing	Number of barcodes with >2,500 reads	Number of barcodes with alignment rate on metagenomic assembly above threshold	Estimated average coverage (X)
Control	232 M	66.3% - 154 M	11,737	9,501 (th: 80%)	1.4
MMC	155 M	87.6% - 137 M	8,447	7,922 (th: 72%)	1.8
UV	37 M	83.7% - 32 M	4,718	1909 (th: 68%)	0.73
CuCl$_2$	243 M	52.8% - 129 M	9,826	5758 (th: 84%)	1.4

In order to match the single-cell analysis results with the metagenomic findings, single-cell reads were aligned on MAGs and virMAGs. This alignment process ensures that the microbial and viral entities identified through metagenomic analysis match with the corresponding individual cells. The result from the alignment was then compared with metagenomic result and visualized in a heatmap showing the phage-host interactions revealed within these two methods (**Figure 14)**. Phage-host interaction within the heatmap was based on the CRISPR-Cas spacer matching from the metagenomics data, BLAST, and single-cell findings.
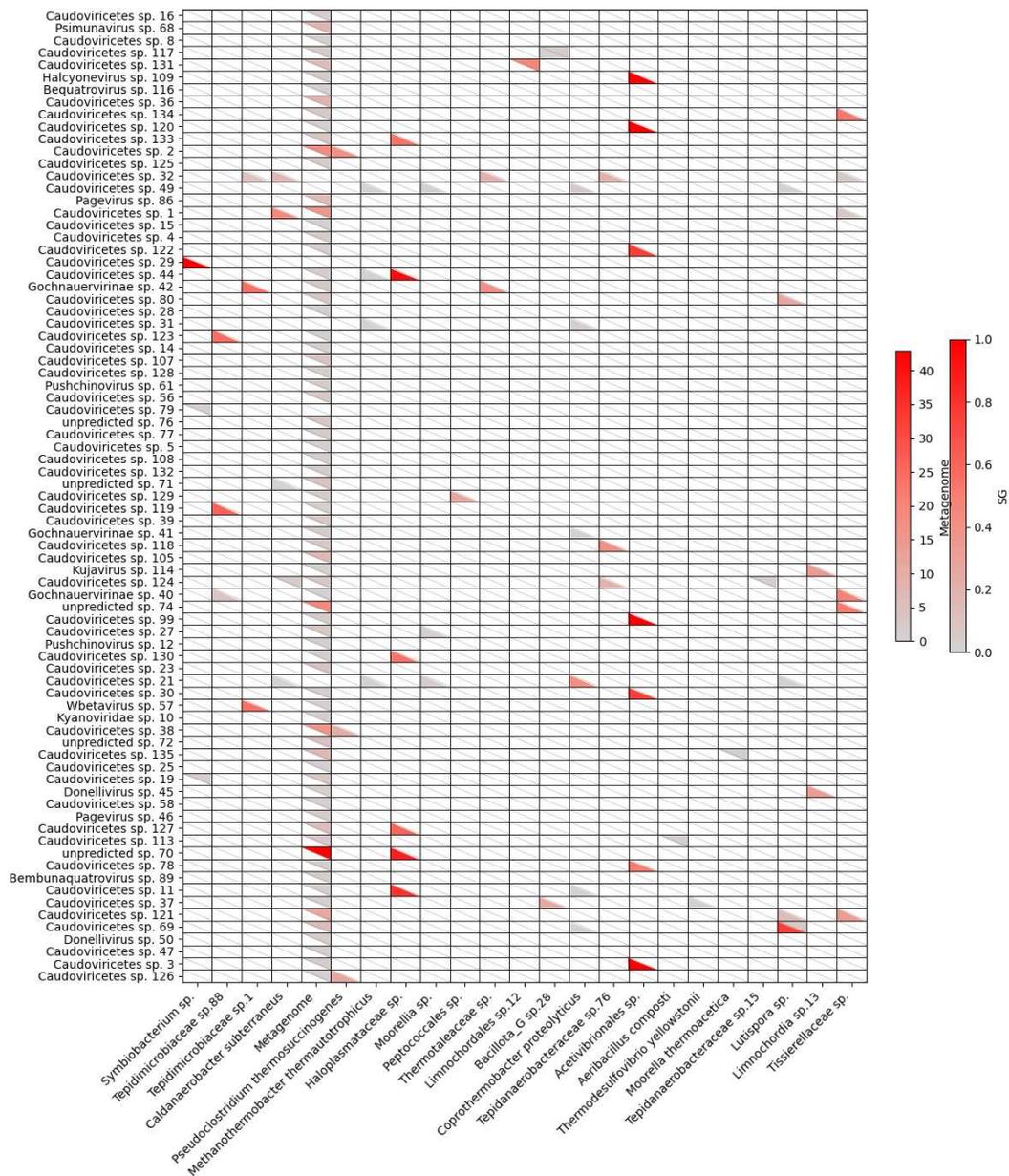
Figure 14. **Comparison of Phage-Host Interactions Between Metagenomic and Single-Cell Analysis.** Phage-host interactions from metagenomic (upper triangle) and single-cell (lower triangle) analyses. Color intensity represents the interaction, with red indicating higher numbers. Metagenomic interactions are predicted by CRISPR-spacer or AAI matching, while single-cell interactions are based on the number of host cells with ≥1% phage genome coverage, normalized by the total cell counts per species.

Lutispora sp. is the host for *Caudoviricetes* sp. 69 according to metagenomic data. The identification of this relationship was further supported by the results from

single-cell analysis as the sequences from *Lutispora* sp., *Caudoviricetes* sp. 69 were found in the same cells. In addition to that, the single cell result also reported more interaction between *Lutispora* sp. with other phages such as *Caudoviricetes* sp. 80 and *Caudoviricetes* sp. 121 (**Figure 14**). These phages carry both integration-related genes and lysis-associated genes, which suggests their ability to switch between lysogenic and lytic phases, depending on the environmental conditions. Interestingly, *Caudoviricetes* sp. 69 and *Caudoviricetes* sp. 80 were among the top five most abundant virMAGs identified in this experiment. In terms of abundance across the experimental conditions, *Caudoviricetes* sp. 69 was found to be particularly abundant in the Control (both pellet and supernatant), UV-treated samples (pellets), and Mitomycin C samples (pellets), indicating that it thrives under these conditions. On the other hand, *Caudoviricetes* sp. 80 was most abundant in the UV-treated and Mitomycin C-treated samples, but specifically in the supernatant, suggesting a possible release of phage particles into the surrounding medium after the phage induction. This result provide a possible explanation of why *Lutispora* sp. showed low abundance in all of phage-inducting conditions compared to the Control.

Furthermore, the metagenomic findings indicated that *Bacillota_G* sp. 28 was infected by *Caudoviricetes* sp. 117, which was also corroborated by single-cell analysis as both of their sequences are found in one cell (**Figure 14**). *Caudoviricetes* sp. 117 is a virulent phage and carries a lysis-associated gene, which is endolysin. This phage was found among the top five most abundant virMAGs in all phage-inducting condition, particularly in the supernatant which suggests that this phage is a free living virus. Despite this, *Bacillota_G* sp. 28 was reported to remain abundant even after phage induction treatments, which raises questions about the dynamics of their interaction. One possible explanation for this apparent result is that while *Caudoviricetes* sp. 117 is actively lysing its host, the population of *Bacillota_G* sp. 28 may be resilient due to high growth rates or efficient population recovery.

In addition to confirm the findings from metagenomics, single cell analysis also provides more information regarding phage-host interaction that were not reported from metagenomic analysis. Single cell result revealed new interactions between 9 MAGs with 28 virMAGs (**Figure 14**). Among this result, the interaction between *Haloplasmataceae* sp. and *Caudoviricetes* sp. 44 stood out with the highest counts particularly in CuCl$_2$-treated samples. This indicates that the phage-host pair was frequently detected in the analyzed single cell data, suggesting a prominent interaction within the microbial community. Aside from its interaction with *Caudoviricete*s sp. 44, *Haloplasmataceae* sp. were found to be infected by several other phages such as *Caudoviricetes* sp. 11, *Caudoviricetes* sp. 127, *Caudoviricetes* sp. 130, *Caudoviricetes* sp. 133, and *Caudoviricetes* sp. 134 (**Figure 14**). Notably,

Caudoviricetes sp. 44 was also found to infect *Methanothermobacter thermautotrophicus,* the most abundant MAG and the only identified archeon within the microbial community. However, this result is not particularly reliable because the count was extremely low. This finding may point to potential limitations of the method because the low counts could represent noise or a spurious alignment rather than a biologically meaningful interaction. Similarly, the alignment revealed that *Coprothermobacter proteolyticus* was infected by three different phages including *Caudoviricetes* sp. 11, *Caudoviricetes* sp. 69, and *Gochnauervirinae* sp. 41 (**Figure 14)** but also with extremely low counts.
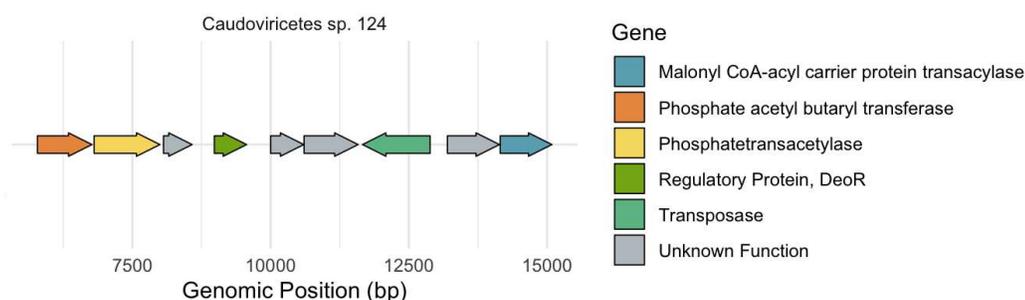
Figure 15. **Genome Map of Significant Phage.** Annotated genes are depicted in colors, genes with unknown function are depicted in grey.

Another significant finding from the single-cell analysis is the identification of additional phage-host interactions involving *Caldanaerobacter subterraneus*, which is an important species in methane production within this microbial community. The analysis revealed that *Caldanaerobacter subterraneus* interacts with several phages, including *Caudoviricetes* sp. 1, *Caudoviricetes* sp. 21, *Caudoviricetes* sp. 32, and *unpredicted* sp. 71. In particular, the interaction between *Caldanaerobacter subterraneus* with *Caudoviricetes* sp. 1 was present in all conditions with high counts compared to interactions with other phages. This result strongly indicates phage-host relationship between *Caldanaerobacter subterraneus* with *Caudoviricetes* sp. 1. These newly discovered interactions from single cell analysis broaden the scope of viral relationships involving this species. However, despite these additional interactions, the single-cell analysis did not corroborate the metagenomic result, which had identified *Caldanaerobacter subterraneus* as being infected by *Caudoviricetes* sp. 124. This discrepancy highlights the complexity and potential variability of phage-host interactions. In details, the interaction between *Caldanaerobacter subterraneus* and *Caudoviricetes* sp. 124 is particularly intriguing because the gene annotation revealed that this phage carried auxiliary metabolic genes (AMG). Among these AMGs, it includes genes encoding Phosphate acetyl butaryl transferase and Phosphotransacetylase (**Figure 15**). In the context of

biomethane production, these enzymes are essential. Phosphate acetyl/butaryl transferase, plays a crucial role in converting acetyl-CoA and butyryl-CoA into their respective phosphate esters, acetyl-phosphate and butyryl-phosphate. This enzyme indirectly supporting the overall metabolic balance in microbial communities involved in biomethane production, particularly in the regulation of acetyl-CoA and butyryl-CoA. On the other side, Phosphotransacetylase plays a direct role in the production of acetate, a key precursor for methanogenesis as the acetyl-phosphate produced by phosphotransacetylase can be converted into acetate, which is then utilized by homoacetogenic bacteria for methane generation.

4.   CONCLUSION

Phages play a pivotal role in anaerobic digestion system by influencing the composition and the shape of microbial communities that drive the methanogenesis process. In the same way, the environmental condition can affect the viral community and the rate of induction of temperate viruses. This study investigates the role of phages on a simplified microbial community under various stress conditions, including UV light exposure, antibiotic treatment (Mitomycin C), heavy metals ($CuCl_2$), and hydrogen peroxide (H_2O_2) through metagenomics and single-cell analysis. Genome-centric metagenomics enabled the reconstruction of 33 MAGs and 129 virMAGs. Among the reconstructed virMAGs, 98 were predicted as temperate, and 30 as virulent. The majority of virMAGs could only be taxonomically identified up to the class level (Caudoviricetes) highlighting the limitations of current classification frameworks and the need for more comprehensive reference databases to achieve finer taxonomic resolution. Phage-host prediction based on metagenomics data revealed that 10 out of 33 MAGs were infected by 26 phages. Meanwhile, single-cell findings enabled the confirmation of 2 phage-host interactions and allowed the discovery of new phage-host interactions between 9 MAGs with 28 virMAGs. *Methanothermobacter thermautotrophicus*, a hydrogenotrophic methanogen, was the dominant species in the microbial community. According to metagenomic findings, this archeon was infected by *Psimunavirus* sp. 68. However, the single-cell findings did not confirm this interaction indicating the complexity of accurately characterizing phage-host dynamics. *Caldanaerobacter subterraneus,* an important homoacetogens in methane production was among the most negatively impacted species after phage-inducting condition. Metagenomics and single-cell findings reported that *Caldanaerobacter subterraneus* was infected by *Caudoviricetes* sp. 124, *Caudoviricetes* sp. 1, *Caudoviricetes* sp. 21 and *Caudoviricetes* sp. 32. Functional gene annotation revealed that *Caudoviricetes* sp. 124 carried auxiliary metabolic genes encoding Phosphate acetyl butaryl transferase and Phosphotransacetylase, enzymes that catalyze the conversion of acetyl-CoA into acetyl phosphate and subsequently into acetate through Wood-Ljungdahl pathway. This underscores the role of phages in promoting horizontal gene transfer, thereby potentially enhancing methane production, by facilitating the transfer of essential genes for metabolic function. Additionally, it presents promising opportunities for developing phage-based tools to enhance CO_2 conversion and methane yield in industrial settings. Overall, the combined approach of metagenomics and single-cell analysis provides a more comprehensive and robust method to uncover phage-host interactions and explore the role of phage in the microbial community.

5. ACKNOWLEDGEMENT

6. APPENDIX

Table S1. List of different treatments for each sample groups.

Sample	Condition
Control	-
UV	1 hour exposure
Mitomycin C	2 ug/mL
CuCl2	0.1 mg/L
H2O2	0.1 mM

Table S2. Average VFAs and standard deviation in each sample groups.

Condition	Sample	VFA	0	96	144	192	240	242	245	264	288	312
CONTROL	MEAN	Acetic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Propionic Acid	32.69	64.11	46.44	34.02	61.88	37.13	57.36	103.55	54.37	26.26
		Isobutyric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.23
		Butyric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Isovaleric Acid	9.15	32.17	50.65	48.05	56.70	41.45	57.04	50.16	55.28	74.17
		Valeric Acid	2.52	15.13	19.42	11.30	16.45	9.01	14.02	23.97	14.54	9.24
		Isocaproic Acid	0.00	1.27	1.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Hexanoic Acid	3.57	3.99	5.56	5.01	18.13	0.00	0.00	2.58	0.10	4.50
		Heptanoic Acid	0.00	33.06	48.18	64.00	80.00	80.67	81.67	88.00	96.00	104.00
	SD	Acetic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Propionic Acid	26.33	41.68	15.74	25.16	35.48	26.78	39.20	52.21	33.58	12.76
		Isobutyric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.42
		Butyric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Isovaleric Acid	11.93	18.01	39.96	49.37	54.74	41.12	55.17	59.17	62.49	7.35
		Valeric Acid	3.42	9.29	14.56	11.94	12.29	8.35	12.13	12.82	14.80	9.01
		Isocaproic Acid	0.00	0.48	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Hexanoic Acid	4.67	1.51	2.65	5.09	2.07	0.00	0.00	0.25	0.04	0.18
		Heptanoic Acid	0.00	54.51	82.98	110.85	138.56	139.72	141.45	152.42	166.28	180.13
UV	MEAN	Acetic Acid	28.06	60.45	76.24	101.87	88.85	74.96	90.55	79.13	50.16	66.16
		Propionic Acid	1.01	18.68	21.15	25.62	33.13	20.85	37.81	40.10	18.34	17.75
		Isobutyric Acid	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Butyric Acid	1.42	12.77	0.00	8.96	13.34	0.13	1.48	0.62	0.00	0.80
		Isovaleric Acid	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Valeric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Isocaproic Acid	15.15	49.41	93.22	55.12	70.04	60.06	128.54	127.25	49.49	61.19
		Hexanoic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Heptanoic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	SD	Acetic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Propionic Acid	16.02	23.56	15.41	14.78	19.34	15.58	22.42	32.18	18.96	6.76
		Isobutyric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.55
		Butyric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Isovaleric Acid	1.97	10.01	7.56	0.93	1.39	0.24	1.32	6.37	5.10	47.25
		Valeric Acid	0.64	4.13	3.44	0.45	2.94	0.47	1.34	7.88	0.18	0.16
		Isocaproic Acid	0.00	0.56	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Hexanoic Acid	0.78	1.75	2.06	0.06	11.36	0.00	0.00	1.65	0.04	3.05
		Heptanoic Acid	0.00	32.00	48.50	64.76	80.95	81.63	82.64	89.05	97.14	105.24
MMC	MEAN	Acetic Acid	15.12	67.16	104.74	87.80	83.10	88.04	89.83	70.19	60.97	74.29
		Propionic Acid	3.48	19.95	27.22	23.90	22.65	16.81	36.62	27.51	17.28	17.48
		Isobutyric Acid	0.00	1.64	1.38	0.35	0.00	0.00	0.00	0.00	0.00	0.00
		Butyric Acid	3.81	9.98	1.40	6.92	3.88	0.00	0.07	0.00	0.00	1.43
		Isovaleric Acid	0.00	2.08	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Valeric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Isocaproic Acid	125.25	59.16	72.67	54.27	76.99	90.43	129.58	84.10	49.19	64.68
		Hexanoic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

		Heptanoic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	SD	Acetic Acid	16.20	34.90	44.02	58.81	51.30	43.28	52.28	45.68	28.96	38.20
		Propionic Acid	12.73	12.12	3.22	6.13	8.72	5.60	9.32	10.09	8.63	5.50
		Isobutyric Acid	0.00	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.71
		Butyric Acid	0.82	7.37	0.00	5.17	7.70	0.08	0.86	0.36	0.00	0.46
		Isovaleric Acid	6.40	8.82	21.23	28.24	31.21	23.67	31.48	32.48	34.70	25.42
		Valeric Acid	1.82	4.66	7.61	6.77	6.42	4.69	6.65	6.47	8.49	5.15
		Isocaproic Acid	8.75	28.23	53.47	31.82	40.44	34.67	74.21	73.47	28.57	35.33
		Hexanoic Acid	2.50	0.95	1.39	2.92	6.05	0.00	0.00	0.89	0.02	1.71
		Heptanoic Acid	0.00	27.39	41.69	55.69	69.61	70.19	71.06	76.57	83.53	90.49
CuCl2	MEAN	Acetic Acid	10.44	34.02	49.59	48.87	44.80	43.77	47.37	38.63	29.98	37.50
		Propionic Acid	10.74	18.54	15.28	14.94	16.90	12.67	22.78	23.26	14.96	9.91
		Isobutyric Acid	0.00	0.61	0.46	0.12	0.00	0.00	0.00	0.00	0.00	1.75
		Butyric Acid	1.54	5.78	0.47	4.03	3.86	0.03	0.31	0.12	0.00	0.63
		Isovaleric Acid	2.79	6.97	9.73	9.72	10.87	7.97	10.93	12.95	13.27	24.22
		Valeric Acid	0.82	2.93	3.68	2.41	3.12	1.72	2.66	4.78	2.89	1.77
		Isocaproic Acid	44.66	29.31	42.34	28.70	39.14	41.70	67.93	52.52	25.92	33.34
		Hexanoic Acid	1.09	0.90	1.15	0.99	5.80	0.00	0.00	0.85	0.02	1.59
		Heptanoic Acid	0.00	19.80	30.06	40.15	50.19	50.61	51.23	55.21	60.22	65.24
	SD	Acetic Acid	9.06	33.59	52.59	44.74	41.93	44.02	45.11	35.63	30.50	37.15
		Propionic Acid	6.50	5.85	12.00	8.89	7.28	6.15	13.65	11.65	5.55	6.58
		Isobutyric Acid	0.00	0.90	0.80	0.20	0.00	0.00	0.00	0.00	0.00	1.52
		Butyric Acid	2.01	5.18	0.81	3.60	3.85	0.04	0.47	0.21	0.00	0.73
		Isovaleric Acid	3.28	4.28	10.59	16.04	17.63	13.60	17.80	17.21	18.74	23.65
		Valeric Acid	0.92	2.55	3.81	3.78	3.21	2.58	3.52	4.20	4.85	2.93
		Isocaproic Acid	69.92	29.31	37.16	27.27	38.51	45.63	65.02	45.80	24.70	32.39
		Hexanoic Acid	1.28	0.88	1.05	1.67	5.68	0.00	0.00	0.82	0.02	1.53
		Heptanoic Acid	0.00	17.30	26.26	35.07	43.83	44.20	44.74	48.21	52.60	56.98
H2O2	MEAN	Acetic Acid	17.59	44.91	78.91	82.96	95.41	70.52	96.05	92.01	99.47	79.37
		Propionic Acid	4.94	21.70	29.72	19.74	25.14	14.91	22.59	33.03	25.00	15.61
		Isobutyric Acid	0.00	0.93	1.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Butyric Acid	6.88	2.92	7.43	8.61	16.67	0.00	0.00	2.76	0.07	4.37
		Isovaleric Acid	0.00	1.16	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Valeric Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Isocaproic Acid	46.79	79.11	59.06	45.46	66.80	48.71	86.43	123.73	80.53	40.93
		Hexanoic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.35
		Heptanoic Acid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	SD	Acetic Acid	3.79	0.67	4.35	7.24	4.80	0.38	3.66	5.16	0.78	0.53
		Propionic Acid	3.18	6.35	6.23	4.50	5.19	3.93	6.87	7.20	4.80	2.30
		Isobutyric Acid	0.00	0.36	0.40	0.10	0.00	0.00	0.00	0.00	0.00	0.63
		Butyric Acid	0.60	1.13	0.41	0.81	2.22	0.03	0.28	0.12	0.00	0.13
		Isovaleric Acid	1.96	2.28	6.41	9.41	10.36	7.96	10.46	10.27	11.14	0.91
		Valeric Acid	0.55	1.12	2.23	2.23	1.88	1.53	2.10	1.18	2.84	1.72
		Isocaproic Acid	30.74	0.63	8.33	2.33	0.98	5.55	4.70	14.43	1.98	1.50
		Hexanoic Acid	0.77	0.04	0.18	0.98	0.19	0.00	0.00	0.03	0.00	0.09
		Heptanoic Acid	0.00	5.26	8.04	10.74	13.43	13.54	13.71	14.77	16.12	17.46

Table S3. Details of DNA concentration of each sample from pellets and supernatants.

Sample	Nanodrop (ng/µL)	Qubit (ng/µL)	Source
Control 1	1186.7	12.3	Pellet
Control 2	1216	10	Pellet
Control 3	1247.7	11.4	Pellet
UV 1	1476.3	10.3	Pellet
UV 2	1360	11.8	Pellet
UV 3	1233.1	11.9	Pellet
Mitomycin C 1	1436.8	13.8	Pellet
Mitomycin C 2	1148.6	10.4	Pellet
Mitomycin C 3	1203.7	12.4	Pellet
$CuCl_2$ 1	1321.4	12.4	Pellet
$CuCl_2$ 2	90.2	8.3	Pellet
$CuCl_2$ 3	1299	13.7	Pellet
H_2O_2 1	108	8.98	Pellet
H_2O_2 2	150.5	11.4	Pellet
H_2O_2 3	137.7	10.2	Pellet
Control Mix	213.7	8.69	Supernatant
Control Pre	43.8	0.61	Supernatant
UV Mix	139.2	10.8	Supernatant
UV Pre	11.9	1.04	Supernatant
Mitomycin C Mix	140.1	7.95	Supernatant
Mitomycin C Pre	2.8	0.94	Supernatant
$CuCL_2$ Mix	35.4	9.78	Supernatant

Table S4. Details of Viral MAGs host prediction with Phabox.

VirMAGs	Length	Type	Host Prediction	Prediction Method
Abouovirus sp. 136	1940	temperate	Brevibacillus_laterosporus	AAI-based
Abouovirus sp. 84	2552	temperate	Brevibacillus_laterosporus	AAI-based
Appavirus sp. 92	1413	virulent	Microbacterium_foliorum	AAI-based
Arawnvirus sp. 102	1128	temperate	Butyrivibrio_fibrisolvens	AAI-based
Bastillevirinae sp. 85	2524	temperate	Symbiobacterium_thermophilum	CRISPR-based
Bembunaquatrovirus sp. 89	2071	virulent	Bacillus_thuringiensis	AAI-based
Bequatrovirus sp. 116	1762	virulent	Bacillus	AAI-based
Caudoviricetes sp. 1	29656	temperate	Cycloclasticus_sp._DSM_27168	CRISPR-based
Caudoviricetes sp. 100	1170	temperate	Vagococcus_entomophilus	CRISPR-based
Caudoviricetes sp. 101	1134	temperate	Thermaerobacter_subterraneus	CRISPR-based
Caudoviricetes sp. 103	1108	virulent	Cytophaga_sp._DMA-K-7a	CRISPR-based
Caudoviricetes sp. 105	14801	temperate	Streptomyces_sp._CNR698	CRISPR-based
Caudoviricetes sp. 106	6991	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 107	6950	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 108	4779	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 11	4768	temperate	Methanocaldococcus_vulcanius	CRISPR-based
Caudoviricetes sp. 110	2055	temperate	Emticicia_sp._17J42-9	CRISPR-based
Caudoviricetes sp. 111	1461	virulent	Chloroflexus_islandicus	CRISPR-based
Caudoviricetes sp. 112	1400	temperate	Lactobacillus_sp._ESL0246	CRISPR-based
Caudoviricetes sp. 113	48824	virulent	Aeribacillus_pallidus	CRISPR-based
Caudoviricetes sp. 117	82690	virulent	Corallococcus_sp._CA040B	CRISPR-based
Caudoviricetes sp. 118	12557	temperate	Sporomusa_acidovorans	CRISPR-based
Caudoviricetes sp. 119	12455	virulent	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 120	54068	virulent	Thermoanaerobacterium_thermosaccharolyticum	CRISPR-based
Caudoviricetes sp. 121	44580	temperate	Streptococcus_sanguinis	CRISPR-based
Caudoviricetes sp. 122	32286	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 123	18518	temperate	Thermoanaerobacteraceae_bacterium_A05MB	CRISPR-based
Caudoviricetes sp. 124	123454	temperate	Thermovenabulum_gondwanense	CRISPR-based
Caudoviricetes sp. 125	129600	virulent	Tepidanaerobacter_syntrophicus	CRISPR-based
Caudoviricetes sp. 126	32091	temperate	Clostridium_autoethanogenum	CRISPR-based
Caudoviricetes sp. 127	24922	temperate	Thermoanaerobacterium_thermosaccharolyticum	CRISPR-based
Caudoviricetes sp. 128	6213	temperate	Tepidiphilus_sp._J10	CRISPR-based
Caudoviricetes sp. 129	16556	temperate	Thermoanaerobacterium_thermosaccharolyticum	CRISPR-based
Caudoviricetes sp. 13	1478	temperate	Sulfodiicoccus_acidiphilus	CRISPR-based
Caudoviricetes sp. 130	12876	temperate	Acinetobacter_sp._RF14B	CRISPR-based
Caudoviricetes sp. 131	172268	virulent	Streptomyces_sp._142MFCol3.1	CRISPR-based
Caudoviricetes sp. 132	126871	temperate	Moorella_sp._Hama-1	CRISPR-based
Caudoviricetes sp. 133	21360	temperate	Clostridium_tepidum	CRISPR-based
Caudoviricetes sp. 134	35519	temperate	Tepidibacillus_fermentans	CRISPR-based
Caudoviricetes sp. 135	43127	temperate	Desulfotomaculum_ferrireducens	CRISPR-based
Caudoviricetes sp. 14	2501	virulent	Corynebacterium_amycolatum	CRISPR-based
Caudoviricetes sp. 15	2350	virulent	Mycobacteroides_chelonae	CRISPR-based
Caudoviricetes sp. 16	2188	temperate		-
Caudoviricetes sp. 19	7982	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 2	41620	temperate	Cycloclasticus_sp._DSM_27168	CRISPR-based
Caudoviricetes sp. 21	54988	temperate		-
Caudoviricetes sp. 22	2050	temperate	Planctomycetes_bacterium_Pla175	CRISPR-based
Caudoviricetes sp. 23	9639	temperate	Tessaracoccus_sp._OH4464_COT-324	CRISPR-based
Caudoviricetes sp. 25	17870	temperate	Thermaerobacter_sp._FW80	CRISPR-based
Caudoviricetes sp. 27	11268	temperate	Vagococcus_entomophilus	CRISPR-based
Caudoviricetes sp. 28	12198	virulent	Bacillus_coagulans	CRISPR-based
Caudoviricetes sp. 29	10947	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 3	22100	temperate	Hungateiclostridium_thermocellum	CRISPR-based

Caudoviricetes sp. 30	8872	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 31	4554	temperate	Runella_zeae	CRISPR-based
Caudoviricetes sp. 32	3830	virulent	Streptomyces_sp._AZ1-7	CRISPR-based
Caudoviricetes sp. 34	49415	virulent	Marinobacterium_lutimaris	CRISPR-based
Caudoviricetes sp. 35	2560	temperate	Limnoraphis_robusta	CRISPR-based
Caudoviricetes sp. 36	2422	temperate	Cycloclasticus_sp._DSM_27168	CRISPR-based
Caudoviricetes sp. 37	27608	virulent	Thermovenabulum_gondwanense	CRISPR-based
Caudoviricetes sp. 38	25259	virulent	Mycoplasma_dispar	CRISPR-based
Caudoviricetes sp. 39	25903	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 4	11492	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 43	3292	virulent	Acinetobacter_sp._CFCC_10889	CRISPR-based
Caudoviricetes sp. 44	3946	temperate	Pseudoclostridium_thermosuccinogenes	CRISPR-based
Caudoviricetes sp. 47	3167	temperate	Ruminococcus_sp._1xD21-23	CRISPR-based
Caudoviricetes sp. 48	2550	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 49	30463	temperate		-
Caudoviricetes sp. 5	8539	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 52	7452	temperate	Streptomyces_sp._Root63	CRISPR-based
Caudoviricetes sp. 54	6746	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 56	5748	virulent	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 58	2869	temperate		-
Caudoviricetes sp. 59	2558	temperate		-
Caudoviricetes sp. 6	5853	temperate	Lactobacillus_delbrueckii	CRISPR-based
Caudoviricetes sp. 60	2490	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 63	1838	virulent	Mycobacterium_simiae	CRISPR-based
Caudoviricetes sp. 64	1818	temperate	Aminipila_sp._JN-39	CRISPR-based
Caudoviricetes sp. 67	1547	virulent	Hungateiclostridium_saccincola	CRISPR-based
Caudoviricetes sp. 69	56172	temperate	Thermovenabulum_gondwanense	CRISPR-based
Caudoviricetes sp. 73	18417	virulent	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 77	18169	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 78	19208	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 79	15427	virulent	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 8	10561	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 80	58644	temperate	Thermohydrogenium_kirishiense	CRISPR-based
Caudoviricetes sp. 81	2770	temperate	Clostridium_tyrobutyricum	CRISPR-based
Caudoviricetes sp. 82	2690	temperate	Salinicola_salarius	CRISPR-based
Caudoviricetes sp. 83	2653	unpredicted	Pyrobaculum_aerophilum	CRISPR-based
Caudoviricetes sp. 87	2424	temperate	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 88	2217	temperate	Frankia_symbiont_of_Datisca_glomerata	CRISPR-based
Caudoviricetes sp. 9	7998	virulent	Symbiobacterium_thermophilum	CRISPR-based
Caudoviricetes sp. 91	1471	temperate	Staphylococcus_delphini	CRISPR-based
Caudoviricetes sp. 93	1364	temperate	Nocardia_farcinica	CRISPR-based
Caudoviricetes sp. 95	1205	temperate	Hungateiclostridium_thermocellum	CRISPR-based
Caudoviricetes sp. 96	1199	temperate		-
Caudoviricetes sp. 97	1198	virulent	Thermoclostridium_stercorarium	CRISPR-based
Caudoviricetes sp. 99	27156	temperate	Thermoanaerosceptrum_fracticalcis	CRISPR-based
Donellivirus sp. 45	3712	temperate	Priestia_megaterium	AAI-based
Donellivirus sp. 50	1989	temperate	Priestia_megaterium	AAI-based
Fernvirus sp. 138	1654	temperate	Paenibacillus_larvae	AAI-based
Gochnauervirinae sp. 33	3411	temperate	Thermicanus_aegyptius	CRISPR-based
Gochnauervirinae sp. 40	3953	temperate	Bacteroides_ilei	CRISPR-based
Gochnauervirinae sp. 41	3868	temperate	Thermoanaerobacteraceae_bacterium_A05MB	CRISPR-based
Gochnauervirinae sp. 42	4878	temperate	Thermoanaerobacteraceae_bacterium_A05MB	CRISPR-based
Halcyonevirus sp. 109	4316	temperate	Paenibacillus_larvae	AAI-based
Halcyonevirus sp. 115	1919	temperate	Paenibacillus_larvae	AAI-based
Inoviridae sp. 53	7231	virulent	Delftia_tsuruhatensis	CRISPR-based
Kamchatkavirus sp. 20	2878	temperate	Aeribacillus_pallidus	AAI-based

48

|				
Kamchatkavirus sp. 55	6209	virulent	Aeribacillus_pallidus	AAI-based
Kujavirus sp. 114	4607	temperate	Vibrio_cholerae	AAI-based
Kyanoviridae sp. 10	7770	virulent	Anoxybacillus_flavithermus	CRISPR-based
Mushuvirus sp. 98	1192	temperate	Faecalibacterium_prausnitzii	AAI-based
Pagevirus sp. 46	50327	temperate	Priestia_megaterium	AAI-based
Pagevirus sp. 86	50187	temperate	Priestia_megaterium	AAI-based
Peduoviridae sp. 66	1548	temperate	Halomonas_sp._WRN001	CRISPR-based
Peduovirus sp. 51	1852	temperate	Salmonella	AAI-based
Phitrevirus sp. 17	2022	temperate	Pseudomonas_aeruginosa	AAI-based
Phitrevirus sp. 18	1814	temperate	Pseudomonas_aeruginosa	AAI-based
Phitrevirus sp. 62	1965	temperate	Pseudomonas_aeruginosa	AAI-based
Psimunavirus sp. 68	15737	temperate	Methanothermobacter_thermautotrophicus	AAI-based
Pushchinovirus sp. 12	2499	temperate	Bacillus_thuringiensis	AAI-based
Pushchinovirus sp. 61	2458	temperate	Bacillus_thuringiensis	AAI-based
Sherbrookevirus sp. 137	1883	temperate	Clostridioides_difficile	AAI-based
Spizizenvirus sp. 65	1688	temperate	Bacillus_subtilis	AAI-based
Triavirus sp. 90	1493	temperate	Staphylococcus_aureus	AAI-based
Triavirus sp. 94	1249	temperate	Staphylococcus_aureus	AAI-based
Wbetavirus sp. 57	4257	temperate	Bacillus	AAI-based
Webervirus sp. 104	1073	virulent	Klebsiella	AAI-based
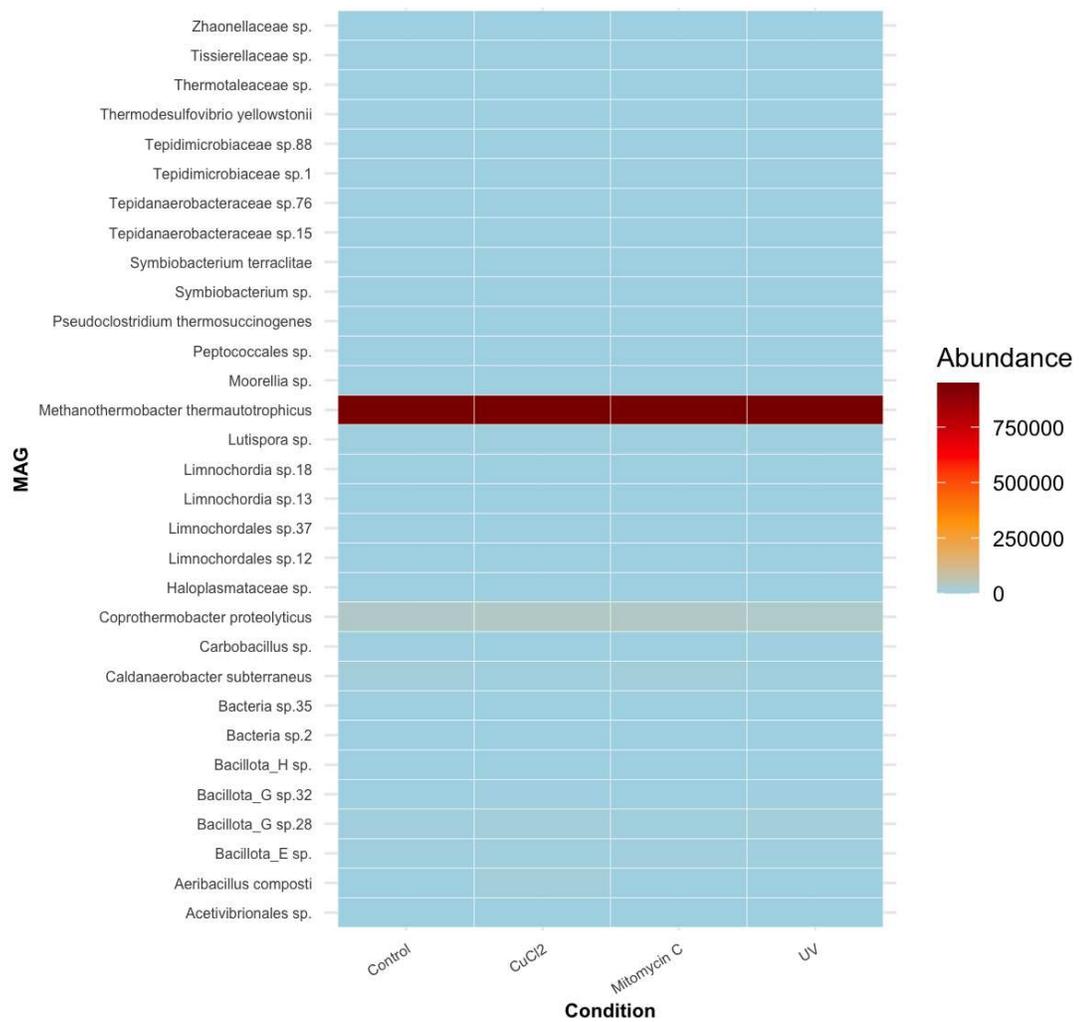
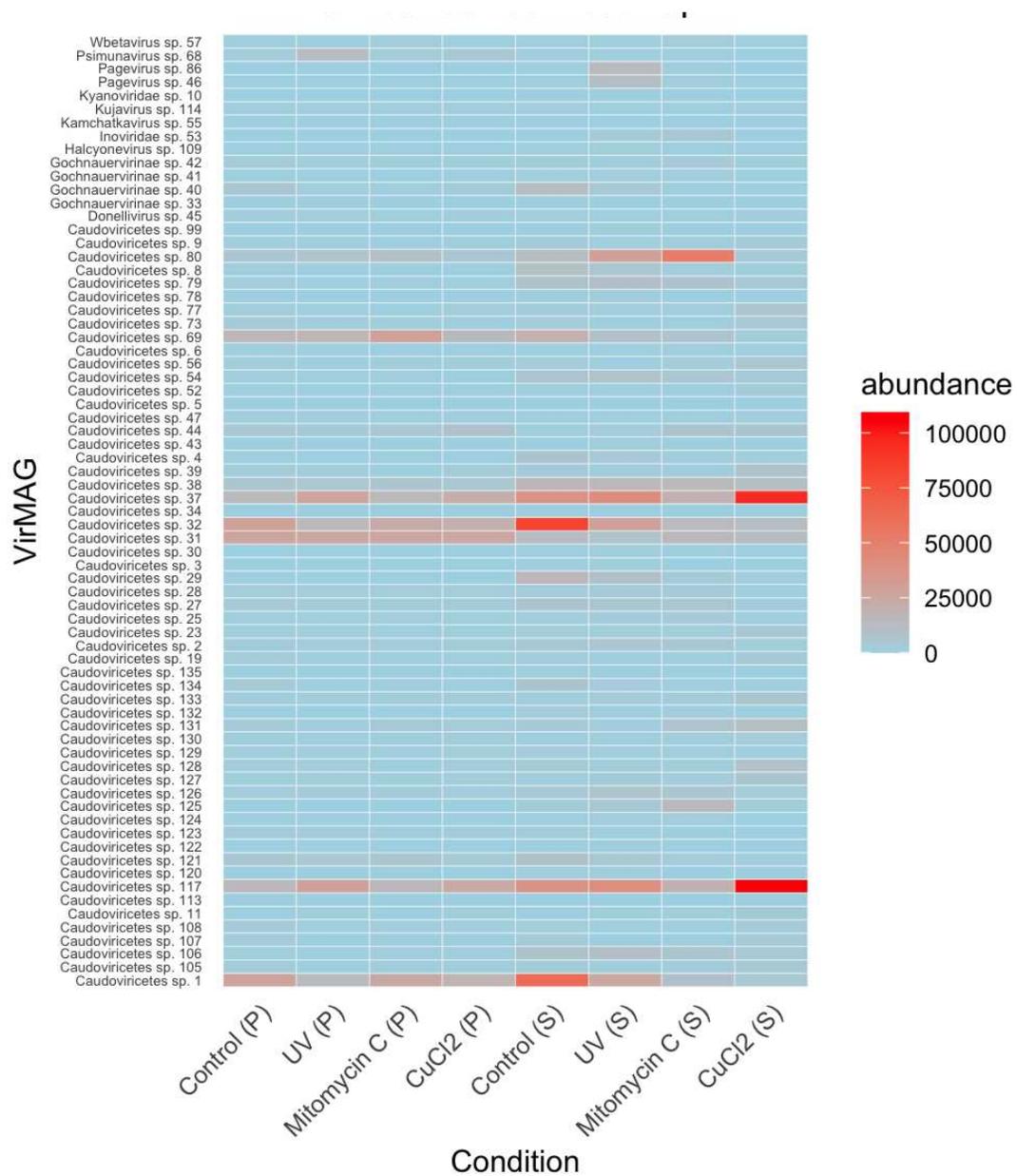**Figure S1.** Heatmap of MAGs abundance across conditions

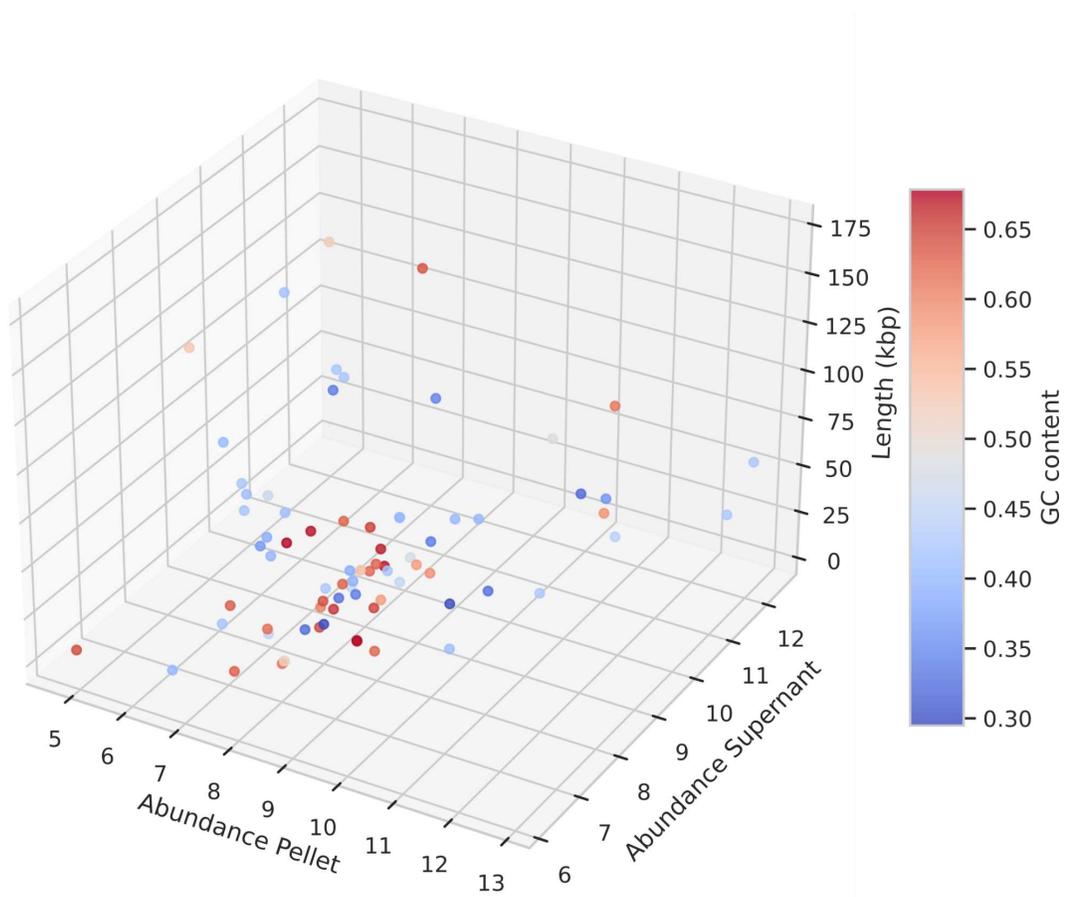**Figure S2.** Heatmap of Viral MAGs abundance across conditions

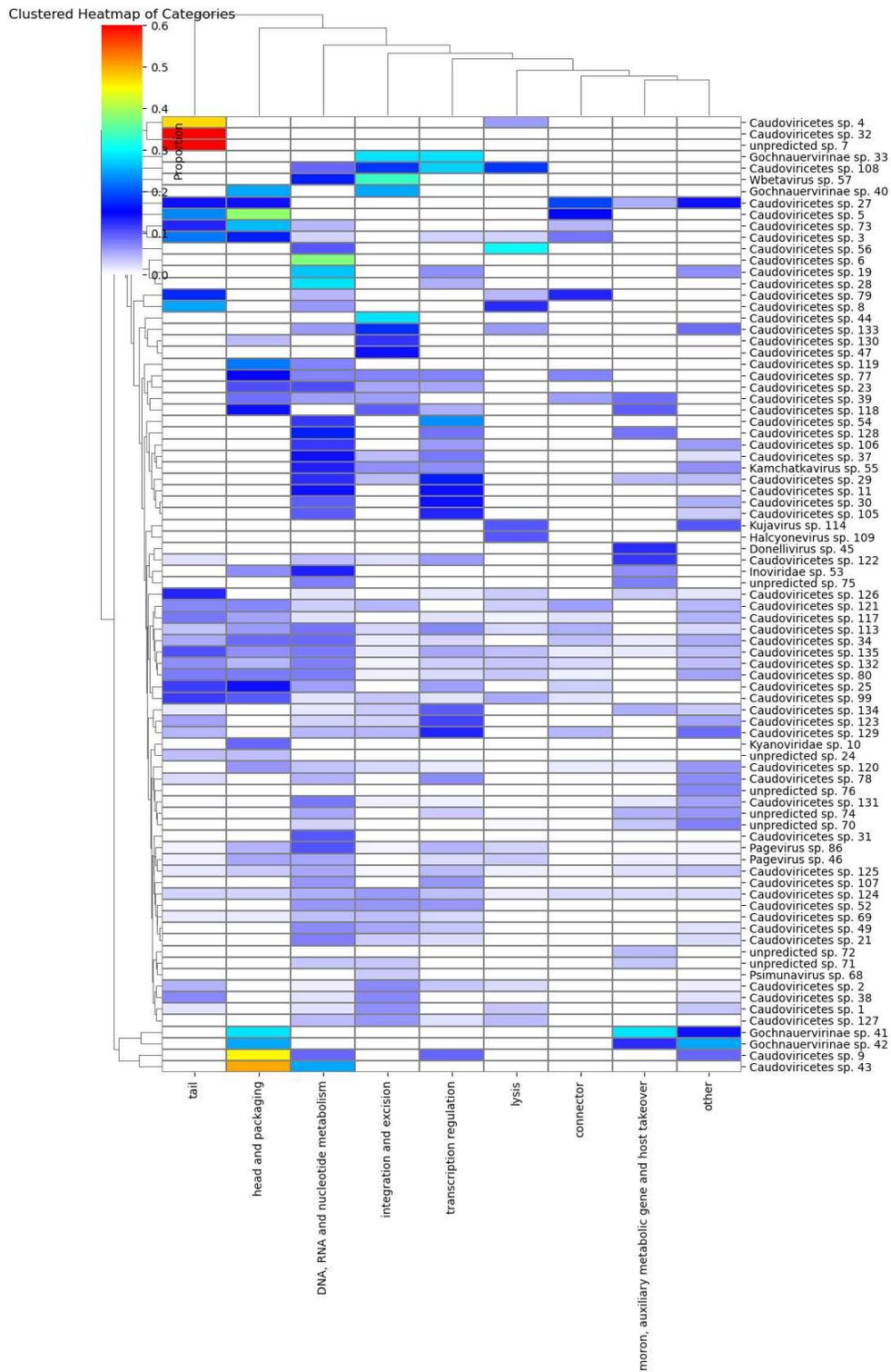**Figure S3.** Abundance of the phages with respect to GC content

Figure S4. Heatmap of functional gene annotations of Viral MAGs

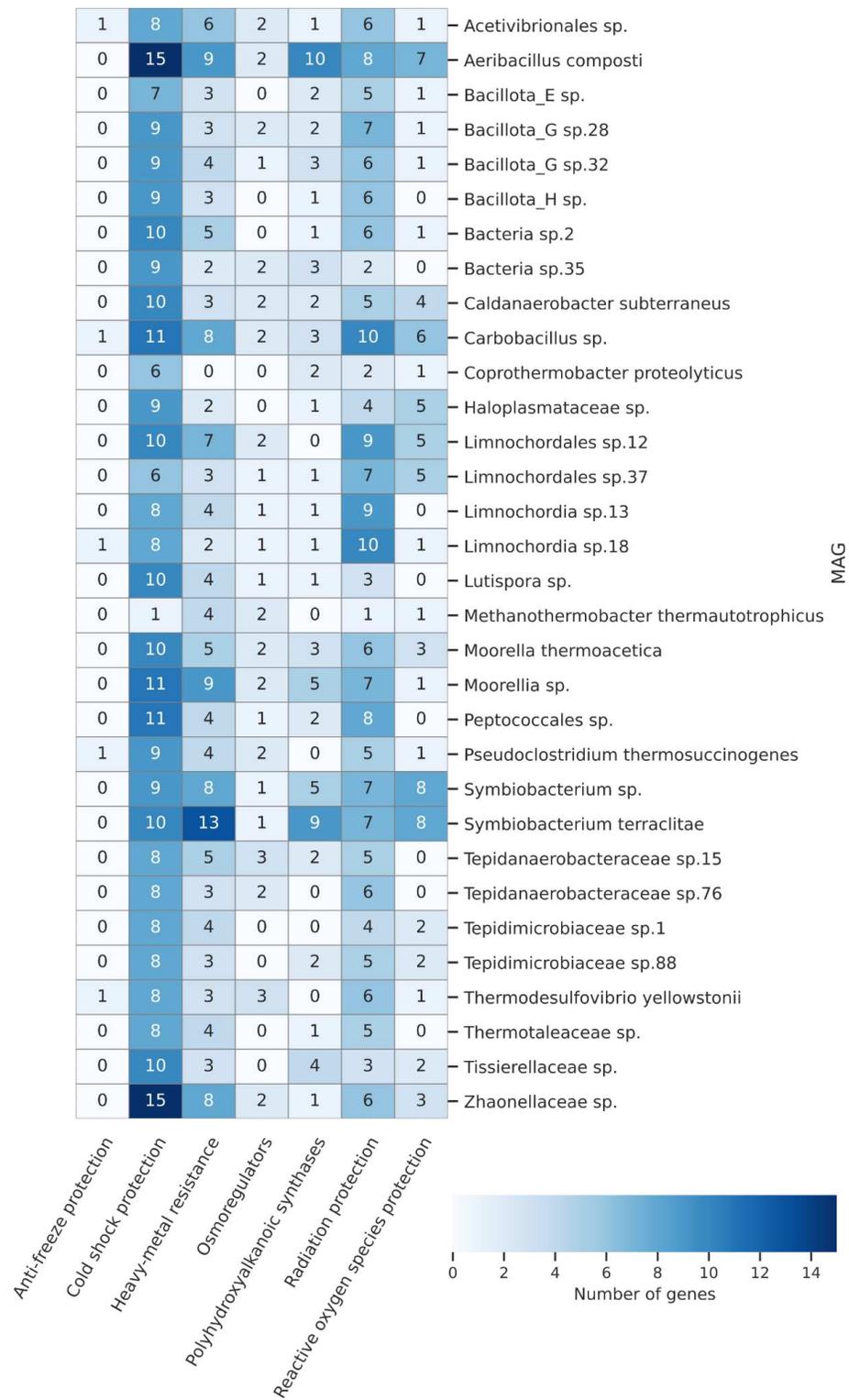| MAG	Anti-freeze protection	Cold shock protection	Heavy-metal resistance	Osmoregulators	Polyhydroxyalkanoic synthases	Radiation protection	Reactive oxygen species protection
Acetivibrionales sp.	1	8	6	2	1	6	1
Aeribacillus composti	0	15	9	2	10	8	7
Bacillota_E sp.	0	7	3	0	2	5	1
Bacillota_G sp.28	0	9	3	2	2	7	1
Bacillota_G sp.32	0	9	4	1	3	6	1
Bacillota_H sp.	0	9	3	0	1	6	0
Bacteria sp.2	0	10	5	0	1	6	1
Bacteria sp.35	0	9	2	2	3	2	0
Caldanaerobacter subterraneus	0	10	3	2	2	5	4
Carbobacillus sp.	1	11	8	2	3	10	6
Coprothermobacter proteolyticus	0	6	0	0	2	2	1
Haloplasmataceae sp.	0	9	2	0	1	4	5
Limnochordales sp.12	0	10	7	2	0	9	5
Limnochordales sp.37	0	6	3	1	1	7	5
Limnochordia sp.13	0	8	4	1	1	9	0
Limnochordia sp.18	1	8	2	1	1	10	1
Lutispora sp.	0	10	4	1	1	3	0
Methanothermobacter thermautotrophicus	0	1	4	2	0	1	1
Moorella thermoacetica	0	10	5	2	3	6	3
Moorellia sp.	0	11	9	2	5	7	1
Peptococcales sp.	0	11	4	1	2	8	0
Pseudoclostridium thermosuccinogenes	1	9	4	2	0	5	1
Symbiobacterium sp.	0	9	8	1	5	7	8
Symbiobacterium terraclitae	0	10	13	1	9	7	8
Tepidanaerobacteraceae sp.15	0	8	5	3	2	5	0
Tepidanaerobacteraceae sp.76	0	8	3	2	0	6	0
Tepidimicrobiaceae sp.1	0	8	4	0	0	4	2
Tepidimicrobiaceae sp.88	0	8	3	0	2	5	2
Thermodesulfovibrio yellowstonii	1	8	3	3	0	6	1
Thermotaleaceae sp.	0	8	4	0	1	5	0
Tissierellaceae sp.	0	10	3	0	4	3	2
Zhaonellaceae sp.	0	15	8	2	1	6	3

Figure S5. Heatmap of functional gene annotations related to stress resistance of MAGs

7. REFERENCES

1. Achouri, O., Panico, A., Bencheikh-Lehocine, M., Derbal, K., Arias, D., Iasimone, F., Padulano, R., Bouteraa, M., Rebahi, A., & Pirozzi, F. (2021). Role of H2O2 dosage on methane production from tannery wastewater: Experimental and kinetic study. *Journal of Water Process Engineering*, *43*, 102313. https://doi.org/10.1016/j.jwpe.2021.102313

2. Antunes, A., Alam, I., Dorry, H. E., Siam, R., Robertson, A., Bajic, V. B., & Stingl, U. (2011). Genome Sequence of Haloplasma contractile, an Unusual Contractile Bacterium from a Deep-Sea Anoxic Brine Lake. *Journal of Bacteriology*, *193*(17), 4551–4552. https://doi.org/10.1128/jb.05461-11

3. Bassani, I., Kougias, P. G., Treu, L., & Angelidaki, I. (2015). Biogas Upgrading via Hydrogenotrophic Methanogenesis in Two-Stage Continuous Stirred Tank Reactors at Mesophilic and Thermophilic Conditions. *Environmental Science & Technology*, *49*(20), 12585–12593. https://doi.org/10.1021/acs.est.5b03451

4. Boe, K., Batstone, D. J., Steyer, J.-P., & Angelidaki, I. (2010). State indicators for monitoring the anaerobic digestion process. *Water Research*, *44*(20), 5973–5980. https://doi.org/10.1016/j.watres.2010.07.043

5. Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., Schulz, F., Jarett, J., Rivers, A. R., Eloe-Fadrosh, E. A., Tringe, S. G., Ivanova, N. N., Copeland, A., Clum, A., Becraft, E. D., Malmstrom, R. R., Birren, B., Podar, M., Bork, P., … The Genome Standards Consortium. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology*, *35*(8), 725–731. https://doi.org/10.1038/nbt.3893

6. Brady, A., Felipe-Ruiz, A., Gallego del Sol, F., Marina, A., Quiles-Puchalt, N., & Penadés, J. R. (2021). Molecular Basis of Lysis–Lysogeny Decisions in Gram-Positive Phages. In *Annual Review of Microbiology* (Vol. 75, Issue Volume 75, 2021, pp. 563–581). Annual Reviews. https://doi.org/10.1146/annurev-micro-033121-020757

7. Bu, F., Dong, N., Khanal, S. K., Xie, L., & Zhou, Q. (2018). Effects of CO on hydrogenotrophic methanogenesis under thermophilic and extreme-thermophilic conditions: Microbial community and biomethanation pathways. *Bioresource Technology*, *266*, 364–373. https://doi.org/10.1016/j.biortech.2018.03.092

8. Campanaro, S., Treu, L., Rodriguez-R, L. M., Kovalovszki, A., Ziels, R. M., Maus, I., Zhu, X., Kougias, P. G., Basile, A., Luo, G., Schlüter, A., Konstantinidis, K. T., & Angelidaki, I. (2020). New insights from the biogas microbiome by comprehensive genome-resolved metagenomics of nearly 1600 species originating from multiple anaerobic digesters. *Biotechnology for Biofuels*, *13*(1), 25.

https://doi.org/10.1186/s13068-020-01679-y

9. Damtie, M. M., Shin, J., Jang, H. M., Cho, H. U., Wang, J., & Kim, Y. M. (2021). Effects of biological pretreatments of microalgae on hydrolysis, biomethane potential and microbial community. *Bioresource Technology*, *329*, 124905. https://doi.org/10.1016/j.biortech.2021.124905

10. Gorecki, A., Holm, S., Dziurzynski, M., Winkel, M., Yang, S., Liebner, S., Wagner, D., Dziewit, L., & Horn, F. (2021). Metaplasmidome-encoded functions of Siberian low-centered polygonal tundra soils. *The ISME Journal*, *15*(11), 3258–3270. https://doi.org/10.1038/s41396-021-01003-y

11. Jancheva, M., & Böttcher, T. (2021). A Metabolite of Pseudomonas Triggers Prophage-Selective Lysogenic to Lytic Conversion in Staphylococcus aureus. *Journal of the American Chemical Society*, *143*(22), 8344–8351. https://doi.org/10.1021/jacs.1c01275

12. Knowles, B., Silveira, C. B., Bailey, B. A., Barott, K., Cantu, V. A., Cobián-Güemes, A. G., Coutinho, F. H., Dinsdale, E. A., Felts, B., Furby, K. A., George, E. E., Green, K. T., Gregoracci, G. B., Haas, A. F., Haggerty, J. M., Hester, E. R., Hisakawa, N., Kelly, L. W., Lim, Y. W., … Rohwer, F. (2016). Lytic to temperate switching of viral communities. *Nature*, *531*(7595), 466–470. https://doi.org/10.1038/nature17193

13. Kofoed, M. V. W., Jensen, M. B., & Ottosen, L. D. M. (2021). Chapter 12—Biological upgrading of biogas through CO2 conversion to CH4. In N. Aryal, L. D. M. Ottosen, M. V. W. Kofoed, & D. Pant (Eds.), *Emerging Technologies and Biological Systems for Biogas Upgrading* (pp. 321–362). Academic Press. https://doi.org/10.1016/B978-0-12-822808-1.00012-X

14. Kouzuma, A., Kato, S., & Watanabe, K. (2015). Microbial interspecies interactions: Recent findings in syntrophic consortia. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00477

15. Ladzekpo, T. N., DeCrescenzo, A. J., Hernandez, A. C., & Everett, G. F. K. (2015). Complete Genome of Bacillus megaterium Podophage Pookie. *Genome Announcements*, *3*(1), e01432-14. https://doi.org/10.1128/genomeA.01432-14

16. Lee, L. H., Lui, D., Platner, P. J., Hsu, S.-F., Chu, T.-C., Gaynor, J. J., Vega, Q. C., & Lustigman, B. K. (2006). Induction of temperate cyanophage AS-1 by heavy metal – copper. *BMC Microbiology*, *6*(1), 17. https://doi.org/10.1186/1471-2180-6-17

17. Li, Y., Chen, Y., & Wu, J. (2019). Enhancement of methane production in anaerobic digestion process: A review. *Applied Energy*, *240*, 120–137. https://doi.org/10.1016/j.apenergy.2019.01.243

18. Luo, Y., Pfister, P., Leisinger, T., & Wasserfallen, A. (2001). The Genome of Archaeal Prophage ΨM100 Encodes the Lytic Enzyme Responsible for Autolysis of *Methanothermobacter wolfeii*. *Journal of Bacteriology*, *183*(19),

5788–5792. https://doi.org/10.1128/jb.183.19.5788-5792.2001

19. Maslov, S., & Sneppen, K. (2015). Well-temperate phage: Optimal bet-hedging against local environmental collapses. *Scientific Reports*, *5*(1), 10523. https://doi.org/10.1038/srep10523

20. McNair, K., Zhou, C., Dinsdale, E. A., Souza, B., & Edwards, R. A. (2019). PHANOTATE: a novel approach to gene identification in phage genomes. *Bioinformatics*, *35*(22), 4537–4542. https://doi.org/10.1093/bioinformatics/btz265

21. Perez Sepulveda, B., Redgwell, T., Rihtman, B., Pitt, F., Scanlan, D. J., & Millard, A. (2016). Marine phage genomics: The tip of the iceberg. *FEMS Microbiology Letters*, *363*(15), fnw158. https://doi.org/10.1093/femsle/fnw158

22. Redl, S., Sukumara, S., Ploeger, T., Wu, L., Ølshøj Jensen, T., Nielsen, A. T., & Noorman, H. (2017). Thermodynamics and economic feasibility of acetone production from syngas using the thermophilic production host Moorella thermoacetica. *Biotechnology for Biofuels*, *10*(1), 150. https://doi.org/10.1186/s13068-017-0827-8

23. Rossi, A., Morlino, M. S., Gaspari, M., Basile, A., Kougias, P., Treu, L., & Campanaro, S. (2022). Analysis of the anaerobic digestion metagenome under environmental stresses stimulating prophage induction. *Microbiome*, *10*(1), 125. https://doi.org/10.1186/s40168-022-01316-w

24. Roughgarden, J. (2024). Lytic/Lysogenic Transition as a Life-History Switch. *Virus Evolution*, *10*(1), veae028. https://doi.org/10.1093/ve/veae028

25. Roux, S., Camargo, A. P., Coutinho, F. H., Dabdoub, S. M., Dutilh, B. E., Nayfach, S., & Tritt, A. (2023). iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLOS Biology*, *21*(4), 1–26. https://doi.org/10.1371/journal.pbio.3002083

26. Sant'Anna, F., Lebedinsky, A., Sokolova, T., Robb, F., & Gonzalez, J. (2015). Analysis of three genomes within the thermophilic bacterial species Caldanaerobacter subterraneus with a focus on carbon monoxide dehydrogenase evolution and hydrolase diversity. *BMC Genomics*, *16*(1), 757. https://doi.org/10.1186/s12864-015-1955-9

27. Taib, N., Megrian, D., Witwinowski, J., Adam, P., Poppleton, D., Borrel, G., Beloin, C., & Gribaldo, S. (2020). Genome-wide analysis of the Firmicutes illuminates the diderm/monoderm transition. *Nature Ecology & Evolution*, *4*(12), 1661–1672. https://doi.org/10.1038/s41559-020-01299-7

28. Touchon, M., Moura de Sousa, J. A., & Rocha, E. P. (2017). Embracing the enemy: The diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Current Opinion in Microbiology*, *38*, 66–73. https://doi.org/10.1016/j.mib.2017.04.010

29. Vrieze, J. D., Hennebel, T., Boon, N., & Verstraete, W. (2012). Methanosarcina: The rediscovered methanogen for heavy duty biomethanation. *Bioresource

Technology, *112*, 1–9. https://doi.org/10.1016/j.biortech.2012.02.079

30. Wigington, C. H., Sonderegger, D., Brussaard, C. P. D., Buchan, A., Finke, J. F., Fuhrman, J. A., Lennon, J. T., Middelboe, M., Suttle, C. A., Stock, C., Wilson, W. H., Wommack, K. E., Wilhelm, S. W., & Weitz, J. S. (2016). Re-examination of the relationship between marine virus and microbial cell abundances. *Nature Microbiology*, *1*(3), 15024. https://doi.org/10.1038/nmicrobiol.2015.24

31. Willenbücher, K., Wibberg, D., Huang, L., Conrady, M., Ramm, P., Gätcke, J., Busche, T., Brandt, C., Szewzyk, U., Schlüter, A., Barrero Canosa, J., & Maus, I. (2022). Phage Genome Diversity in a Biogas-Producing Microbiome Analyzed by Illumina and Nanopore GridION Sequencing. *Microorganisms*, *10*(2). https://doi.org/10.3390/microorganisms10020368

32. Wu, L., Wei, W., Song, L., Woźniak-Karczewska, M., Chrzanowski, Ł., & Ni, B.-J. (2021). Upgrading biogas produced in anaerobic digestion: Biological removal and bioconversion of CO2 in biogas. *Renewable and Sustainable Energy Reviews*, *150*, 111448. https://doi.org/10.1016/j.rser.2021.111448

33. Xu, J., Bu, F., Zhu, W., Luo, G., & Xie, L. (2020). Microbial Consortiums of Hydrogenotrophic Methanogenic Mixed Cultures in Lab-Scale Ex-Situ Biogas Upgrading Systems under Different Conditions of Temperature, pH and CO. *Microorganisms*, *8*(5). https://doi.org/10.3390/microorganisms8050772

34. Zhang, J., Gao, Q., Zhang, Q., Wang, T., Yue, H., Wu, L., Shi, J., Qin, Z., Zhou, J., Zuo, J., & Yang, Y. (2017). Bacteriophage–prokaryote dynamics and interaction within anaerobic digestion processes across time and space. *Microbiome*, *5*(1), 57. https://doi.org/10.1186/s40168-017-0272-8

35. Zhu, Y., Shang, J., Peng, C., & Sun, Y. (2022). Phage family classification under Caudoviricetes: A review of current tools using the latest ICTV classification framework. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.1032186