Exploring viral diversity in fermented vegetables through viral metagenomics

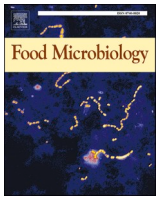
Julia Cantuti Gendre, Claire Le Marrec, Stéphane Chaillou, Lysiane Omhover-Fougy, Sophie Landaud, Eric Dugat-Bony

Food Microbiology 128 (2025) 104733

Contents lists available at ScienceDirect

Food Microbiology

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



Exploring viral diversity in fermented vegetables through viral metagenomics

Julia Cantuti Gendre [a], Claire Le Marrec [b], Stéphane Chaillou [c], Lysiane Omhover-Fougy [d], Sophie Landaud [a], Eric Dugat-Bony [a,*]

[a] UMR 0782 SAYFOOD, Université Paris-Saclay, INRAE, AgroParisTech, 22 place de l'agronomie, 91120, Palaiseau, France
[b] UMR 1366 OENOLOGIE, Université Bordeaux, INRAE, Bordeaux INP, Bordeaux Sciences Agro, Institut des Sciences de la Vigne et du Vin, 210 Chemin de Leysotte, 33140, Villenave-d'Ornon, France
[c] UMR 1319 MICALIS INSTITUTE, Université Paris-Saclay, INRAE, AgroParisTech, Domaine de Vilvert, 78350, Jouy-en-Josas, France
[d] Aerial, 250 rue Laurent Fries, 67300, Illkirch, France

ARTICLE INFO

Keywords:
Bacteriophages
Fermented vegetables
Food microbial ecology
Metavirome
Metagenomics
Epifluorescence microscopy

ABSTRACT

Fermented vegetables are traditionally produced using the endogenous microorganisms present in raw ingredients. While the diversity of bacteria and fungi in fermented vegetables has been relatively well studied, phage communities remain largely unexplored. In this study, we collected twelve samples of fermented cabbage, carrot, and turnip after fermentation and analyzed the microbial and viral communities using shotgun and viral metagenomic approaches. Assessment of the viral diversity also benefited from epifluorescence microscopy to estimate viral load. The viral metagenomics approach targeted dsDNA, ssDNA, and RNA viruses. The microbiome of fermented vegetables was dominated by lactic acid bacteria and varied according to the type of vegetable used as raw material. The analysis of metagenome-assembled-genomes allowed the detection of 22 prophages of which 8 were present as free particles and therefore detected in the metaviromes. The viral community, estimated to range from 5.28 to 7.57 log virus-like particles per gram of fermented vegetables depending on the sample, was mainly composed of dsDNA viruses, although ssDNA and non-bacterial RNA viruses, possibly originating from the phyllosphere, were also detected. The dsDNA viral community, primarily comprising bacteriophages, varied depending on the type of vegetable used for fermentation. The bacterial hosts predicted for these phages mainly belonged to *Lactobacillaceae* and *Enterobacteriaceae* families. These results highlighted the complex microbial and viral composition of fermented vegetables, which varied depending on the three types of vegetables used as raw material. Further research is needed to deepen our understanding of the impact of these viruses on the microbial ecology of fermented vegetables and on the quality of the final products.

1. Introduction

Fermented foods – *i.e.* foods made through desired microbial growth and enzymatic conversions of food components (Marco et al., 2017) – are largely consumed worldwide and represent 20% of the human diet (Tamang and Kailasapathy, 2010; Varzakas et al., 2017). It is estimated that more than 5000 varieties of fermented foods and beverages are currently produced and consumed worldwide (Tamang et al., 2016; Tamang and Kailasapathy, 2010). Fermentation allows a longer shelf life of raw food material, reducing the presence of anti-nutrients or even toxic compounds. It is also considered an efficient way to enhance food products' flavor, texture, and aroma (Marco et al., 2017; Molfetta et al., 2022). Furthermore, many potential health benefits have been attributed to fermented foods, although establishing a direct correlation with their consumption still requires rigorous clinical studies in many cases (Marco et al., 2017; Rul et al., 2022; Sanlier et al., 2019).

Vegetables are rich sources of beneficial compounds and are used to produce various fermented products. Fermentation of cabbage, turnip, radish, carrot, and others, dates back more than 2000 years. The renewed interest for fermented vegetables observed in the last decades is due to the demand for food products with healthy properties found in plant-based products (Gunawardena et al., 2024; Siddiqui et al., 2023). Their production mostly relies on spontaneous fermentation, wherein biochemical transformations are carried out by endogenous microbial

* Corresponding author.
E-mail address: eric.dugat-bony@inrae.fr (E. Dugat-Bony).

https://doi.org/10.1016/j.fm.2025.104733
Received 19 October 2024; Received in revised form 16 January 2025; Accepted 17 January 2025
Available online 17 January 2025

populations originating from the autochthonous microbiota of raw vegetables (Di Cagno et al., 2008). The bacterial community of raw vegetables is usually dominated by taxa belonging to the *Pseudomonadota* phylum (formerly *Proteobacteria*), including members of the *Enterobacteriaceae* family (Leff and Fierer, 2013; Mudoor Sooresh et al., 2023). However, while lactic acid bacteria (LAB) represent less than 1% of the initial microbial community in raw materials (Buckenhueskes, 2015), they are the main microorganisms involved in the fermentation of vegetables, a process commonly referred to as lactic acid fermentation (Thierry et al., 2023a). From an ecological perspective, the fermentation of vegetables is usually characterized by the successive development of hetero- and homo-fermentative LAB populations (Buckenhueskes, 2015). Microbiological and biochemical monitoring of sauerkraut, one of the most popular fermented vegetables in Europe and North America, have been reported earlier (Plengvidhya et al., 2007). This is also the case in Asia with kimchi, a spicy analog of sauerkraut made from Chinese cabbage or radish (Cheigh et al., 1994). In sauerkraut, rapid growth of heterolactic LAB species such as *Leuconoctoc mesenteroides* is responsible for the rapid acidification of the medium through the production of lactic and acetic acids. They also create anaerobic conditions through the production of CO_2. This first step provides favorable conditions for the growth of more acid-tolerant homolactic LAB species, primarily *Lactiplantibacillus plantarum,* which are responsible for the complete degradation of sugars. Similar microbial dynamics were reported during the fermentation of carrot juice with *Leuconostoc* spp. and *Lactobacillus* spp. being detected as dominant populations along fermentation (Wuyts et al., 2018). In a recent study, the microbiota associated to 75 homemade plant-based fermented foods from 23 different types (carrot, cabbage, onion, beet, celeriac and mix of different vegetables), were analyzed through a combination of cultural methods and targeted metagenomics (16S metabarcoding) (Thierry et al., 2023b). The authors observed no relationship between the structure of the bacterial community and the nature of the vegetable used as raw material, nor the common manufacturing practices, such as salt content or the number of ingredients, confirming previous work on Kimchi reaching similar conclusions (Lee et al., 2017).

Bacteriophages – *i.e.* viruses that infect bacteria - are ubiquitous in nature where they play a role in modulating the composition of microbial communities (Breitbart and Rohwer, 2005; White et al., 2022). These entities are therefore naturally present in fermented foods where their ecology and overall impact remain understudied (Ledormand et al., 2021; Paillet and Dugat-Bony, 2021). Bacteriophages are commonly perceived as potential risks to industrial fermentation, capable of inducing fermentation delays or failures. They also represent promising perspectives for the food industry as biocontrol agents to prevent the proliferation of spoilage and pathogenic bacteria (de Melo et al., 2023; Ranveer et al., 2024; Wu et al., 2023). In complex communities, such as undefined starter cultures, it was demonstrated that phages can affect the population dynamics at the strain level and help maintain a high degree of genetic diversity (Erkus et al., 2013; Spus et al., 2023). Previous studies reported the isolation of hundreds of virulent LAB phages from fermented vegetables, such as sauerkraut (Lu et al., 2003) and fermented cucumbers (Lu et al., 2012; Yoon et al., 2007), thereby offering first evidence of the existence of a diverse phage community within this type of ecosystem. The principal bacterial hosts of these isolated phages primarily encompassed genera such as *Leuconostoc*, *Levilactobacillus*, *Lactiplantibacillus*, *Weissella* and *Pediococcus*. Notably, recent work on fermented cucumber has extended this spectrum, with phages infecting various Gram-negative bacteria (*Pseudomonadota* phylum) (Lu et al., 2020). While perfectly adapted for isolating and precisely characterizing representative phages, the culture-based approach employed may have offered only a partial description of the true phage diversity within the studied fermented vegetables. Such an approach indeed heavily depends on the diversity of the collection of bacterial isolates utilized as potential hosts. Furthermore, it is widely acknowledged that certain phages are challenging, if not impossible, to cultivate (Breitbart et al., 2002).

With the advance of high throughput sequencing methods, viral metagenomics has proven to be particularly effective in reliably describing the composition of viral communities from various environmental samples, especially in the discovery of novel and previously unknown viruses (Breitbart et al., 2002; Rosario and Breitbart, 2011). Consequently, it offers a complementary approach to cultivation methods for detecting and characterizing a broader range of phages in fermented foods (Ledormand et al., 2022; Paillet et al., 2024; Park et al., 2011). To date, such an approach was successfully applied on a few fermented vegetables including sauerkraut and kimchi (Jung et al., 2018; Park et al., 2011). Park et al. study focused on the dsDNA viral composition in sauerkraut and kimchi, while Jung et al. studied both ds and ssDNA viruses in kimchi revealing the presence of *Circoviridae* (infecting birds and pigs), *Genomoviridae* (infecting plants and fungi), and *Microviridae* (infecting bacteria) along with dsDNA phages in this type of products. These studies led to the discovery of large numbers of previously unknown viral sequences (Jung et al., 2018; Park et al., 2011). Furthermore, a recent study compared the microbial community larger than 0.22 μm and the ultra-small microbiome smaller than 0.22 μm in kimchi using a combination of metataxonomic and metatranscriptomic approaches, revealing the presence and abundance of RNA viruses in the smallest fraction (H.-W. Lee et al., 2022). However, to our knowledge, no study has provided a comprehensive description of viral communities in fermented foods that includes both DNA and RNA viruses, as well as viruses present as free particles and those integrated into the genome of their hosts.

This study aimed to investigate the potential impact of the type of vegetables used as raw material on the composition of the viral community in the resulting fermented product. We conducted a comparative analysis of twelve commercially available fermented vegetables, comprising four fermented cabbages (sauerkraut), four fermented carrots, and four fermented turnips (sauerruben) sourced from six different producers. For each sample, we quantified the viral load by epifluorescence microscopy and conducted a comprehensive analysis of the viral and microbial composition. This analysis included viral metagenomics targeting dsDNA, ssDNA, and RNA viruses to capture encapsidated viruses, as well as shotgun metagenomics to explore the diversity of phages existing as prophages within bacterial genomes.

2. Material and methods

2.1. Sampling procedure and microbiological analyses

All the fermented vegetables studied were collected at the end of fermentation and obtained either directly from the producers or purchased in local grocery stores (Table 1; Samples metadata: https://doi.org/10.57745/DODROX). In total, the samples were sourced from six different producers. Each type of vegetable (cabbage, carrot, turnip) included in the study was obtained from at least two different producers, thus ensuring variability in fermentation conditions and the microbiological characteristics of the analyzed samples. Samples were immediately stored at 4 °C until processing. For each sample, 10 g of fermented vegetable was homogenized with 90 ml of sterile saline solution (9 g/l NaCl) in a stomacher for 2 min at the maximal speed (BagMixer® 400W, Interscience). Serial dilutions were performed in sterile saline solution and microorganisms were enumerated by surface plating in duplicate on specific agar base media as follows. Plate Count Agar (PCA, Biokar Diagnostics) was used for the determination of the total number of live, aerobic bacteria in samples. LAB were enumerated on Man-Rogosa-Sharpe agar (MRS, pH 6.5, Biokar Diagnostics). Both media were supplemented with 22.5 mg/l amphotericin B to inhibit fungi. Fungal populations were enumerated on yeast extract-glucose-chloramphenicol (YEGC, Biokar Diagnostics). All incubations were carried out for 3–5 days at 28 °C. Anaerobic jars with 1 bag of GENbox anaer (Biomerieux) were used to incubate LAB under anaerobic conditions. A 10 ml sample

Table 1
Microbial, viral load estimation and pH results for each sample. Total bacteria were enumerated on PCA medium under aerobic conditions. LAB were enumerated on MRS medium under anaerobic conditions. Yeasts were enumerated on YEGC medium under aerobic conditions. The concentration of virus-like particles was estimated by epifluorescence microscopy.

Sample	Producer	Total bacteria (log CFU/g)	LAB (log CFU/g)	Yeast (log CFU/g)	Viruses (log VLP/g)	pH
Cabbage_1	A	6.83	7.10	<2.00	7.33	3.39
Cabbage_2	A	5.94	6.12	3.97	7.01	3.36
Cabbage_3	A	6.67	6.75	4.04	6.65	3.53
Cabbage_4	E	<2.00	6.72	5.20	7.57	3.46
Carrot_1	B	5.51	5.66	3.87	6.79	3.34
Carrot_2	B	6.77	6.81	5.28	7.15	3.84
Carrot_3	D	<2.00	<2.00	<2.00	6.43	3.47
Carrot_4	D	<2.00	<2.00	<2.00	5.28	3.75
Turnip_1	C	7.45	7.57	4.03	7.10	3.67
Turnip_2	B	7.24	7.30	<2.00	6.84	3.55
Turnip_3	C	<2.00	<2.00	5.55	–	3.30
Turnip_4	F	6.50	7.51	6.52	7.19	3.55

of the ten-fold dilution (equivalent to 1 g of the original sample) was centrifuged at 10,000×*g* for 30 min. The microbial pellet was washed using sterile saline solution. It was collected upon centrifugation, and stored à −20 °C for metagenomic DNA extraction.

2.2. Shotgun metagenomic analysis

The microbial pellet was resuspended in 180 µl of pre-lysis buffer (20 mM Tris-HCl [pH 8], 2 mM sodium EDTA, 1.2% Triton X-100, 20 mg/ml lysozyme) and 10 U mutanolysin was add and incubated for 30 min at 37 °C. DNA extraction was performed using the DNeasy® PowerFood® kit (Qiagen) following the manufacturer's instructions except for the vortexing step which was replaced by homogenisation in a Precellys® Evolution homogenizer (Bertin Instruments) for two 45 s mixing steps at a speed of 6500 rpm. DNA was quantified using the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific) sent to Eurofins Genomics (Konstanz, Germany) for shotgun sequencing using a NovaSeq platform (Illumina, San Diego, CA, USA). A minimum of 5 million of 150 bp paired-end reads was produced per sample.

Taxonomic affiliation of the raw reads was carried out using Kaiju v.1.9.2 (Menzel et al., 2016). Raw reads were processed using the SnakeMAGs v.1.1.1 workflow with default parameters (Tadrent et al., 2022) which enabled the reconstruction of metagenome-assembled genomes (MAGs). Quality control of the bins was performed using CheckM v.1.1.3, and bins exhibiting >50% completion and <10% contamination were considered medium-quality MAGs (Bowers et al., 2017) and further studied. They were taxonomically classified using the Genome Taxonomy Database Toolkit v2.1.0 (GTDB-Tk; (Chaumeil et al., 2022). The presence of prophage sequences in the MAGs was predicted using geNomad v.1.7.4 4 (Camargo et al., 2024) using both the end-to-end and standard options. We retained sequences annotated as proviruses by geNomad, which corresponds to regions containing virus-specific markers flanked by host-specific genes. The predicted prophages were classified with PhaGCN tool from PhaBOX v.2.0 using standard options (Shang et al., 2021, 2023) and compared to the viral Core Nucleotide database of NCBI using BLASTn v.2.13.0 (Camacho et al., 2009).

2.3. Purification of the viral fraction

Purification of the viral fraction from fermented vegetables was achieved using a protocol previously developed and optimized for cheese samples (Dugat-Bony et al., 2020), after minor changes. Briefly, 10 g of sample were diluted in 90 ml of Tris-HCl 10 mM, NaCl 200 mM buffer into a sterile bag and mixed for 2 min using a stomacher (BagMixer® 400 W by Interscience). The mixture was transferred in a glass bottle and mixed with 900 ml of Tris-HCl 10 mM, NaCl 200 mM buffer. The sample was kept at 4 °C overnight under agitation in order to recover the maximum number of viruses from the solid matrix. The suspension was next centrifuged at 7000×*g* for 15 min at 4 °C to pellet big aggregates and microbial cells. The supernatant containing free viral particles was filtered using 0.45 µm polyethersulfone membranes (Metricel®, Pall Life Sciences) and glass vacuum filter holders (Millipore). Viral particles were precipitated overnight at 4 °C after the addition of 10% (w/v) PEG 8000 (Sigma-Aldrich). After centrifugation at 5000×*g* for 30 min at 4 °C, pellets were resuspended in 1 ml of SM buffer (Tris-HCl 50 mM, NaCl 200 mM, $MgSO_4$ 10 mM) and the viral fraction was stored at 4 °C. A buffer control (Tris-HCl 10 mM, NaCl 200 mM) which received no sample was run in parallel to evaluate the background noise of our procedure in the viral metagenomes.

In order to validate the extraction method, three phage isolates with different morphologies were added to three sauerkraut samples at the beginning of the protocol: *Psychrobacter* phage d'Alembert (Myophage morphology) infecting *Psychrobacter aquimaris*, *Glutamicibacter* phage Voltaire (Podophage morphology) and *Glutamicibacter* phage Montesquieu (Siphophage morphology) both infecting *Glutamicibacter arilaitensis* (Paillet et al., 2022). One hundred µl of a lysate of each phage with a titer of 10^9 PFU/ml was added to the sample directly in the stomacher bag and the extraction of viral particles was processed according to the procedure described previously. Samples were collected at different steps of the protocol and a dedicated plaque assay was used to enumerate each phage. Briefly, serial dilutions of the viral extract were performed into SM buffer (Tris-HCl 50 mM, NaCl 200 mM, $MgSO_4$ 10 mM). Then eight microliters of each dilution were spotted on Brain Heart Infusion (Biokar Diagnostics) soft agarose 0.3% (Basic agarose Premier, MP Biomedicals) previously inoculated with the sensitive bacterial strain and supplemented with $MgSO_4$ (10 mM) and $CaCl_2$ (1 mM). After incubation for 24 h at 23 °C, the phage titer was determined by counting lysis plaques at the lowest possible dilution. At each step of the extraction protocol, phage titers were compared using a Kruskal-Wallis test, which revealed a significant effect of the protocol step on phage recovery (Supplementary Fig. S2). To identify the step with the highest impact on phage recovery, pairwise comparisons were conducted using Dunn's test, which showed a significant reduction in phage titers following PEG precipitation. However, no differences were observed at any step based on phage morphology (Kruskal-Wallis test), suggesting that, according to this criterion, the procedure did not selectively bias the representativeness of the viral extract from fermented vegetables.

2.4. Estimation of virus-like particles' concentration by epifluorescence microscopy

A volume of 100 µl of the viral fraction was tenfold diluted in SM buffer (Tris-HCl 50 mM, NaCl 200 mM, $MgSO_4$ 10 mM) before filtration on 0.02 µm Anodisc filters (Anodisc 25 mm diameter, Whatman). Each filter was then incubated on a 75 µl drop of 1X SYBR Gold (Invitrogen) in the dark for 15 min. After removal, filters were dried in the dark before being mounted on a glass slide with Fluoromount-G (Invitrogen) and a coverslip. Slides were stored at −20 °C until observation. Microscopic observations were carried out using an Axiostar Plus microscope (Zeiss) with a 100 × oil objective CP-ACHROMAT (Zeiss). Ten images were captured per slide in the bright field and GFP fluorescence channels (with an excitation filter wavelength of 472/30 nm and emission filter wavelength of 520/35 nm). Emission was collected using interference filters and the images were captured using Dino-Eye AM4023CT C-mount Camera (Dino-Lite) and the associated software. The number of virus-like particles was counted using ImageJ v1.54d (Schroeder et al., 2021) and then related to the initial weight of the fermented vegetable sample, taking into account the ratio between the image area and the filter area, as well as the sample dilution factor.

2.5. Viral metagenomic analysis

Eight hundred µl of viral fraction or molecular-biology grade water (water control) was transferred to a 2 ml Phase Lock Gel tube (5PRIME) and mixed thoroughly with 800 µl of fresh, non-oxidized chloroform in order to eliminate membrane vesicles. After centrifugation at 13,000×*g* for 10 min at 4 °C, the aqueous phase containing viral particles was recovered and treated for 30 min at 37 °C with 1 U of TURBO DNase (Invitrogen) to digest free DNA. DNase was inactivated by the addition of 5.5 µl of 100 mM EDTA and the sample was placed on ice for 5 min. Nucleic acids were released from the capsids by adding 7 µl of SDS 20% (05030, Sigma) and 21 µl of proteinase K (20 µg/µl) and incubating the sample for 1 h at 56 °C. The entire content of the tube was transferred to a 2 ml Phase Lock Gel tube and further processed using the procedure described in our previous work (Dugat-Bony et al., 2020), which includes phenol-chloroform extraction and ethanol precipitation. Purified nucleic acids were finally resuspended in 20 µl of nuclease-free water and quantified using both the Qubit dsDNA HS Assay and Qubit RNA BR Assay kits (ThermoFisher Scientific).

Viral DNA and RNA were amplified separately using the REPLI-g® Cell WGA & WTA Kit (Qiagen) following the manufacturer's instructions. We used this kit to analyze both DNA and RNA viruses and due to the low quantity of nucleic acids obtained from the samples, which was insufficient for direct sequencing. After amplification, all samples were purified using AMPure® XP beads (Beckman Coulter). DNA and RNA-amplified samples were finally sent to Eurofins Genomics (Konstanz, Germany) for shotgun sequencing using a NovaSeq platform (Illumina, San Diego, CA, USA). A minimum of 5 million of 150 bp paired-end reads was produced per sample. Raw reads were decontaminated to eliminate reads matching with cabbage, carrot and turnip genomes (accession numbers: GCA_000695525.1, GCA_001625215.1, GCA_018901965.1) according to the procedure detailed in Yeoh et al. (Yeoh, 2021). Quality filtering, assembly, clustering of contigs into species-level vOTUs and selection of viral contigs were performed as described in our previous work (Paillet et al., 2024). Briefly, raw reads were quality-filtered using Trimmomatic v0.39 (Bolger et al., 2014). Decontaminated and trimmed reads were subsampled to 100k, 400k, and 700k using seqtk version 1.3 (https://github.com/lh3/seqtk). Decontaminated and trimmed reads and all subsampled read sets were independently assembled into contigs using SPAdes 3.15.3 (Bankevich et al., 2012). Contigs with length below 2000 bp were discarded. This threshold corresponds to almost half the size of the smallest phage genome already described, and a little bit lower than the minimal size observed for a metagenome-derived complete phage genome (Olo Ndela et al., 2023). Selected contigs were clustered following the approach previously described (Shah et al., 2023). Succinctly, a pairwise alignment was first performed for all contigs using BLAT (Kent, 2002). Then, contigs with a self-alignment >110% of contig length, corresponding to chimeras, were removed. The remaining contigs were clustered at the species level (90% identity × coverage) and the longest contig within each cluster was selected as the representative sequence. To predict viral contigs, contigs were submitted to VIBRANT version 1.2.1 (Kieft et al., 2020), VirSorter2 version 2.2.4 (Guo et al., 2021) and CheckV version 0.8.1 (Nayfach et al., 2021). Contigs were selected and declared vOTUs if they meet at least one of the following criteria: declared "complete", "high" or "medium" quality by either VIBRANT or CheckV, declared "full" by VirSorter2.To evaluate the relative abundance of each vOTU in each sample, trimmed reads were mapped against the vOTUs using bwa-mem2 v2.2.1 (Vasimuddin et al., 2019). Alignments were filtered using msamtools v1.1.3 (https://github.com/arumugamlab/msamtools) in order to ultimately retain counts only for contigs which were detected with high confidence. First, msamtools filter was used with the option -l 80, -p 95, -q 80, –besthit. Then, msamtools coverage was used with the –sumary option to estimate sequence coverage and only alignments with more than 50% of coverage were kept to construct the final abundance table. Then, msamtools coverage estimates the coverage of each sequence, and the –summary option reports the fraction of sequence covered in percentage. Finally, only alignments with more than 50% of coverage were kept to construct the final abundance tables.

To better characterize the vOTUs, we predicted their bacterial hosts using iPHoP-1.2.0 (Roux et al., 2023), classified them with PhaGCN tool from PhaBOX v.12.0 using standard options (Shang et al., 2021, 2023), and searched for sequence similarity with known bacteriophages using BLASTn (Camacho et al., 2009) against the Core Nucleotide BLAST and International Committee on Taxonomy of Viruses (ICTV) databases. For unknown phages, we performed a rapid annotation using Pharokka (Bouras et al., 2023). Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarity, permutational analysis of variance (PERMANOVA), composition analysis, and heatmap visualization were processed within R using the following packages: Phyloseq v1.44.0 (McMurdie and Holmes, 2013), vegan_2.6–8 (Oksanen et al., 2024), pheatmap_1.0.12 (Raivo Kolde, 2010), ggplot2_3.5.1 (Wickham, 2016), tidyverse_2.0.0 (Wickham et al., 2019).

3. Results

3.1. Fermented vegetables characterization

To evaluate the microbial composition of our fermented vegetables in terms of bacteria, yeasts, and viruses, we used a combination of plate counts and epifluorescence microscopy (Table 1). First, total aerobic bacteria were enumerated on PCA medium, a non-selective rich medium supplemented with amphotericin B to inhibit fungi. Counts were above the detection limit of 100 CFU/g for eight of the twelve samples, ranging from 5.83 to 7.45 log CFU/g. LAB, enumerated on MRS medium supplemented with amphotericin B after incubation in anaerobic jars, were detected in nine samples and ranged from 5.66 to 7.57 log CFU/g. Yeasts were detected using YEGC medium in eight samples, at concentrations ranging from 3.87 to 6.52 log CFU/g. Virus-like particles (VLP), counted using epifluorescence microscopy, were present in all but one sample (Turnip 3) and reached concentrations similar to those of LAB in the analyzed samples, ranging from 5.28 to 7.57 log VLP/g of vegetable. The pH value, a classical indicator used for assessing the lactic acid fermentation process, ranged from 3.30 to 3.84. Whatever the variable (microbial count, VLP count, and pH), values did not significantly vary depending on the three types of vegetables used for fermentation (Kruskal-Wallis test, $p > 0.05$) (Supplementary Fig. S3).

3.2. Composition of the fermented vegetable microbiome

Shotgun metagenomic was applied to the same samples to examine the microbial composition of our fermented vegetables. Two samples of carrots (3 and 4) were amplified prior sequencing due to the very low concentration of DNA after extraction. The average sequencing depth was 10,019,400 reads per sample (interquartile range: 9,134,234–11,001,433). Taxonomic affiliation of the reads using Kaiju revealed that between 29.10% and 54.84% of the reads corresponded to Bacteria, and 0.56%–32.48% corresponded to Eukaryota. Only a small proportion of the reads, between 0.15% and 4.84%, were affiliated with viruses (Fig. 1A). Reads affiliated with Archaea were detected in seven samples but in low abundance, with a maximum of 0.009% in carrot 3. A large number of reads were unclassified, possibly corresponding to vegetable reads not identified by Kaiju or reads that could not be classified due to lack of matching reference sequences. The bacterial community was predominantly composed of lactic acid bacteria (*Bacillota* phylum), including *Leuconostoc, Secundilactobacillus, Lactiplantibacillus, Latilactobacillus, Levilactobacillus, Lactobacillus, Lactococcus, Lentilactobacillus* and *Pediococcus* genera. In addition, some *Pseudomonadota* including *Rahnella, Bradyrhizobium, Sphingomonas, Variovorax* and *Enterobacter* were detected in all samples (Fig. 1B). The bacterial community structure, after aggregation of the data at the genus level, varied

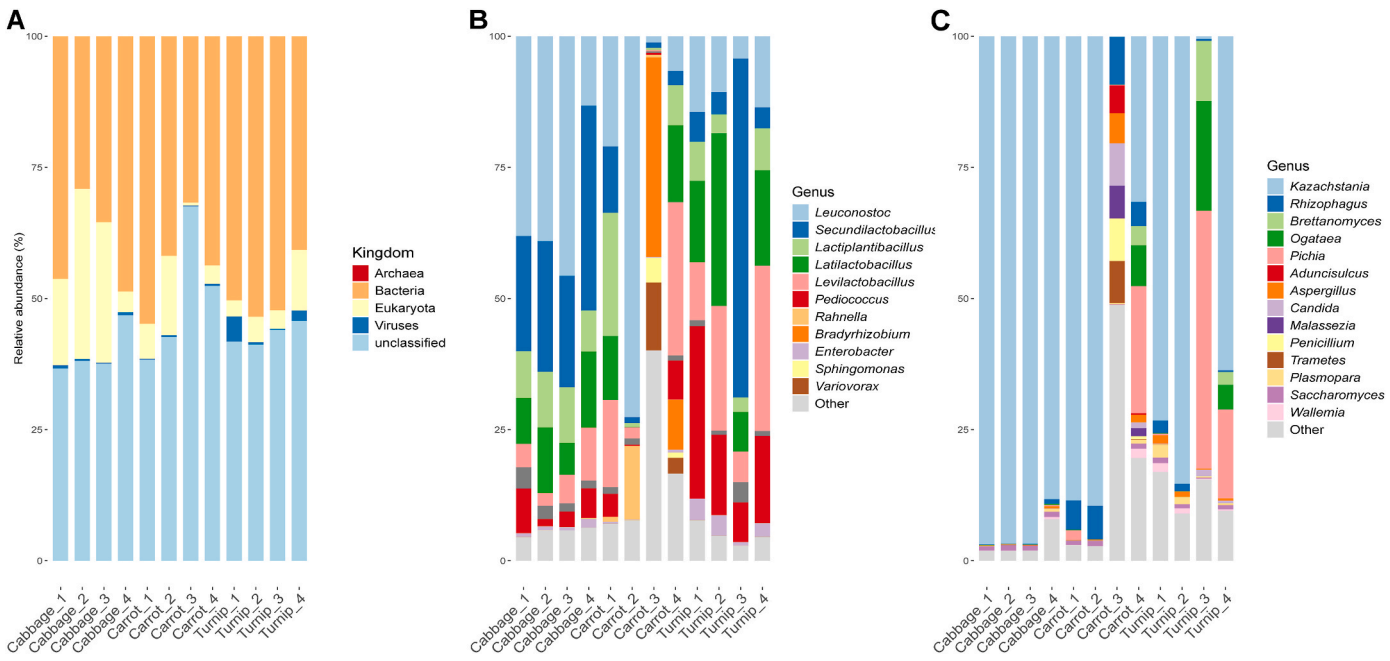


Fig. 1. Bar plot composition of the microbial community of fermented vegetables. A) Relative abundance of reads per sample (%) at the domain level. B) Relative abundance of bacterial reads per sample (%) at the genus level. C) Relative abundance of eukaryotic reads per sample (%) at the genus level.

according to the type of vegetable used as raw material for the fermentation ($p = 0.018$*, $p < 0.05$, PERMANOVA based on Bray-Curtis dissimilarity matrix) (Supplementary Fig. S1). Regarding yeasts, *Kazachstania* was the main genus detected in our samples (Fig. 1C), except for carrot 3 and turnip 3. The profile of the fungal community of carrot 3 is different from that of the other samples, with most reads affiliated with *Rhizophagus, Aduncisulcus, Aspergillus, Candida, Malassezia, Penicillium, Trametes* and *Trichosporon.* However, as mentioned earlier, DNA quantity was very low for this sample and DNA amplification was necessary prior to sequencing, which may explain this difference. Regarding turnip 3, *Pichia, Ogataea, and Brettanomyces* were the main detected fungal genera.

Sequencing reads were individually assembled and binned by sample, resulting in the reconstruction of 54 metagenome-assembled genomes (MAGs), with those meeting quality thresholds of >50% completeness and <10% contamination being classified as medium-quality MAGs according to international standards (Bowers et al., 2017) (Supplementary Table S1, MAG metadata: https://doi.org/10.57745/LHXYD9). Medium-quality MAGs were recovered from all samples except carrot 3, one of the two samples subjected to amplification prior to sequencing. The MAGs exhibited an average length of 1.6 Mbp (from 0.5 to 3.6 Mbp) and corresponded to multiple species, including *Latilactobacillus curvatus* (8 MAGs), *Pediococcus parvulus* (8 MAGs), *Secundilactobacillus malefermentans* (7 MAGs), *Latilactobacillus sakei* (5 MAGs), *Secundilactobacillus silagei* (5 MAGs), *Leuconostoc mesenteroides* (3 MAGs), *Leuconostoc pseudomesenteroides* (2 MAGs), *Levilactobacillus parabrevis* (2 MAGs), *Leuconostoc fallax* (1 MAGs). Those species covered the main bacterial taxa detected through direct taxonomic affiliation of raw reads (Fig. 1).

To characterize the prophage composition in our samples, we predicted prophage sequences in the previously described medium-quality MAGs (Table 2). In total, twenty-two prophages were predicted in thirteen MAGs (24% of the medium-quality MAGs), which belonged to the genera *Lactococcus* (1 MAG with 3 prophages), *Secundilactobacillus* (5 MAGs with 1–3 prophages), *Leuconostoc* (4 MAGs with 1 prophages), *Levilactobacillus* (1 MAG with 1 prophages), and *Pediococcus* (2 MAGs with 1 and 2 prophages). The sizes of these prophages ranged from 7336 to 36,981 bp. Notably, five of these prophages from *Secundilactobacillus silagei* were completely identical with each other (100% identity, 100% coverage), while assembled from different samples (Cabbage_1_22_2, Cabbage_2_17_2, Cabbage_3_23_2, Cabbage_4_2_2, Carrot_1_22_1), and showed only weak sequence homology with known phages. While encoding an integrase and a transcriptional regulator of phage origin, their short sizes, ranging from 7.8 kb to 10.5 kb and the absence of predicted genes encoding phage structural proteins, leave us uncertain whether they constitute complete prophages or other mobile genetic elements. The smallest predicted prophage from our dataset, Turnip_3_7_1 (7.3 kb), located at the extremity of a contig from MAG_N3_M_7, showed no sequence homology with known phages at the nucleotidic level. However, at the proteic level, seven predicted proteins encoded on this sequence shared homology with phage tail module proteins but the sequence lacked predicted proteins for capsid structure, indicating that this sequence may represent an incomplete prophage. The largest predicted prophage, Cabbage_2_17_3 (40 kb), showed weak homology with known phages. geNomad predicted 43 genes on this sequence, among which only two were annotated as encoding phage proteins. One belonged to the cluster of orthologous genes (COG) 5283 encoding a phage-related tail protein, and one encoded a protein possessing a prophage endopeptidase tail domain (pfam: PF06605). In addition, a BLASTp comparison of other predicted protein sequences against nr database also revealed distant homology with other known phage proteins (phage tail, distal tail protein, fiber tail, spike protein). This prophage sequence might therefore represent an uncharacterized prophage of *Secundilactobacillus silagei*. The second-largest predicted prophage, Turnip_1_5_1 (36 kb), showed no sequence homology with known phages at the nucleotidic level. However, geNomad predicted 54 genes including genes encoding structural phage proteins such as major capsid proteins, portal proteins and terminase. This prophage might also represent an uncharacterized prophage of *Leuconostoc*. All cabbage samples contained MAGs with prophages, while two turnip samples (turnip 1 and 3) and one carrot sample (carrot 1) had prophages in their MAGs. Comparing prophage sequences with known viral sequences revealed that prophage Cabbage_1_15_1 from *Lactococcus lactis* matched the complete genome of *Lactococcus* phage r1t (e-value: 0, 94.58% identity, 79% coverage), the first *Lactococcus* phage of the P335-species (Van Sinderen et al., 1996) (Table 2). On the contrary, three predicted

Table 2
List of prophages detected on MAGs from fermented vegetables. Bacterial host corresponds to the taxonomic identification of the MAGs containing the prophage sequence. Each prophage was compared with the viral Core Nucleotide database to find any match with existing sequences. The best hit is indicated with the percentage of identity and coverage.

Id_prophage	Length (bp)	Bacterial host (MAG_id)	Best hit (BLASTn)	% identity	% coverage	Taxonomy (PhaBOX)	Score taxonomy (PhaBOX)
Cabbage_1_15_1	20,947	*Lactococcus lactis* (C1_M_15)	NC_004302.1 *Lactococcus* phage r1t, complete genome	94.58	79	*Caudoviricetes*	1.00
Cabbage_1_15_2	18,340	*Lactococcus lactis* (C1_M_15)	KX160218.1 *Lactococcus* phage 98204, complete genome	97.06	20	*Caudoviricetes*	1.00
Cabbage_1_15_3	8489	*Lactococcus lactis* (C1_M_15)	LT993227.1 unidentified phage genome assembly, complete genome: monopartite	93.95	40	*Caudoviricetes*	0.87
Cabbage_1_22_1	11,991	*Secundilactobacillus silagei* (C1_M_22)	BK020715.1 TPA: *Caudoviricetes* sp. isolate ctkN83, partial genome	85.68	16	*Caudoviricetes*	1.00
Cabbage_1_22_2	10,517	*Secundilactobacillus silagei* (C1_M_22)	NC_070781.1 *Streptomyces* phage Tomas, complete genome	100.00	0	*Caudoviricetes*	1.00
Cabbage_2_17_1	11,322	*Secundilactobacillus silagei* (C2_M_17)	BK045548.1 TPA: *Caudoviricetes* sp. isolate ctQoP2, partial genome	86.85	19	*Caudoviricetes*	1.00
Cabbage_2_17_2	10,517	*Secundilactobacillus silagei* (C2_M_17)	NC_070781.1 *Streptomyces* phage Tomas, complete genome	100.00	0	*Caudoviricetes*	1.00
Cabbage_2_17_3	40,158	*Secundilactobacillus silagei* (C2_M_17)	MK504445.1 *Lactobacillus* phage ATCCB, complete genome	85.52	5	*Caudoviricetes*	0.96
Cabbage_3_17_1	30,882	*Leuconostoc mesenteroides* (C3_M_17)	BK021696.1 TPA: Caudoviricetes sp. isolate ct8S34, partial genome	80.92	15	*Caudoviricetes*	0.63
Cabbage_3_23_1	12,598	*Secundilactobacillus silagei* (C3_M_23)	PP453647.1 *Listeria* phage vB_Lmo_2389_typeII, complete genome	71.26	3	*Caudoviricetes*	1.00
Cabbage_3_23_2	7773	*Secundilactobacillus silagei* (C3_M_23)	No hit	–	–	*Caudoviricetes*	1.00
Cabbage_3_23_3	11,535	*Secundilactobacillus silagei* (C3_M_23)	BK020715.1 TPA: *Caudoviricetes* sp. isolate ctkN83, partial genome	85.68	17	*Caudoviricetes*	1.00
Carrot_1_22_1	10,517	*Secundilactobacillus silagei* (Cat1_M_22)	NC_070781.1 *Streptomyces* phage Tomas, complete genome	100.00	0	*Caudoviricetes*	1.00
Carrot_1_6_1	14,460	*Leuconostoc pseudomesenteroides* (Cat1_M_6)	No hit	–	–	*Sandinevirus*	0.58
Cabbage_4_2_1	10,749	*Secundilactobacillus silagei* (Cat4_M_2)	BK020715.1 TPA: *Caudoviricetes* sp. isolate ctkN83, partial genome	85.68	18	*Caudoviricetes*	1.00
Cabbage_4_2_2	10,517	*Secundilactobacillus silagei* (Cat4_M_2)	NC_070781.1 *Streptomyces* phage Tomas, complete genome	100.00	0	*Caudoviricetes*	1.00
Cabbage_4_6_1	23,307	*Leuconostoc mesenteroides* (Cat4_M_6)	MN855893.1 *Inoviridae* sp. isolate 382, complete genome	90.42	25	*Fadolivirus*	0.57
Turnip_1_2_1	21,995	*Levilactobacillus parabrevis* (N1_M_2)	BK045548.1 TPA: *Caudoviricetes* sp. isolate ctQoP2, partial genome	74.70	5	*Caudoviricetes*	0.88
Turnip_1_5_1	36,981	*Leuconostoc* (N1_M_5)	BK046201.1 TPA: *Caudoviricetes* sp. isolate ctbhY2, partial genome	76.62	0	*Caudoviricetes*	0.64
Turnip_3_7_1	7336	*Pediococcus parvulus* (N3_M_7)	No hit	–	–	*Caudoviricetes*	1.00
Turnip_3_9_1	31,508	*Pediococcus parvulus* (N3_M_9)	BK045548.1 TPA: *Caudoviricetes* sp. isolate ctQoP2, partial genome	74.70	3	*Caudoviricetes*	0.77
Turnip_3_9_2	21,044	*Pediococcus parvulus* (N3_M_9)	TPA: *Caudoviricetes* sp. isolate ctyPQ2, partial genome	87.09	23	*Caudoviricetes*	1.00

prophages identified in *Secondilactibacillus silagei* (Cabbage_3_23_2), *Leuconostoc pseudomesenteroides* (Carrot_1_6_1) and *Pediococcus parvulus* (Turnip_3_7_1) MAGs did not match any known viruses and are therefore totally new. The remaining prophages partially matched with known viral sequences but always with a low sequence coverage (<40%). Only two prophage sequences could be assigned to a viral genus with a low confidence score using PhaBOX, indicating the lack of reference genomes with a known taxonomy for temperate phages infecting LAB from fermented vegetables.

These results showed that these fermented vegetables are mainly composed of lactic acid bacteria and yeasts. Among MAGs corresponding to lactic acid bacteria, several complete prophage sequences were detected, which is an indication that lysogeny is a common phenomenon in fermented cabbage, carrot, and turnip.

3.3. Overview of the fermented vegetables' virome

Having observed that many bacterial genomes present in fermented vegetables host prophage sequences, we decided to explore the diversity of viruses present in the same samples as free viral particles using a comprehensive viral metagenomic approach. Nucleic acids were purified from the viral fraction extracted from the products, DNA (ssDNA and dsDNA), and RNA were amplified separately, and shotgun sequenced. The average sequencing depth was 3,256,769 reads per sample (interquartile range: 1,551,885–4,258,484). After assembly, contigs larger than 2 kb were selected and clustered based on sequence homology (remove chimeras self-alignment >110% of contig length, clustering 90% identity × coverage, and keep the longest contig within each cluster). Among the resulting 4458 clusters, 1214 were identified as viral. Finally, 1143 species-level viral operational taxonomic units (vOTUs) remained in the dataset after filtering contaminants (*i.e.* clusters highly abundant in negative controls) (vOTUs metadata: https://doi.org/10.57745/DODROX).

The vOTU sizes ranged from 2 to 145 kb (Fig. 2A), with 185 contigs (16%) exceeding 10 kb. Five vOTUs were qualified as complete genomes according to CheckV (Nayfach et al., 2021), 5 as high-quality and 27 as medium-quality genomes. In total, 29 vOTUs (1.6%) were identified as temperate phages by VIBRANT (Kieft et al., 2020). According to VirSorter2 predictions (Guo et al., 2021), the majority of vOTUs (72%) belonged to dsDNA phages (72%) (Fig. 2B). Yet, we were also able to detect ssDNA viruses (179 vOTUs, 16% of the total) and RNA viruses (15 vOTUs, 1.3% of the total). We also detected double-stranded dsDNA

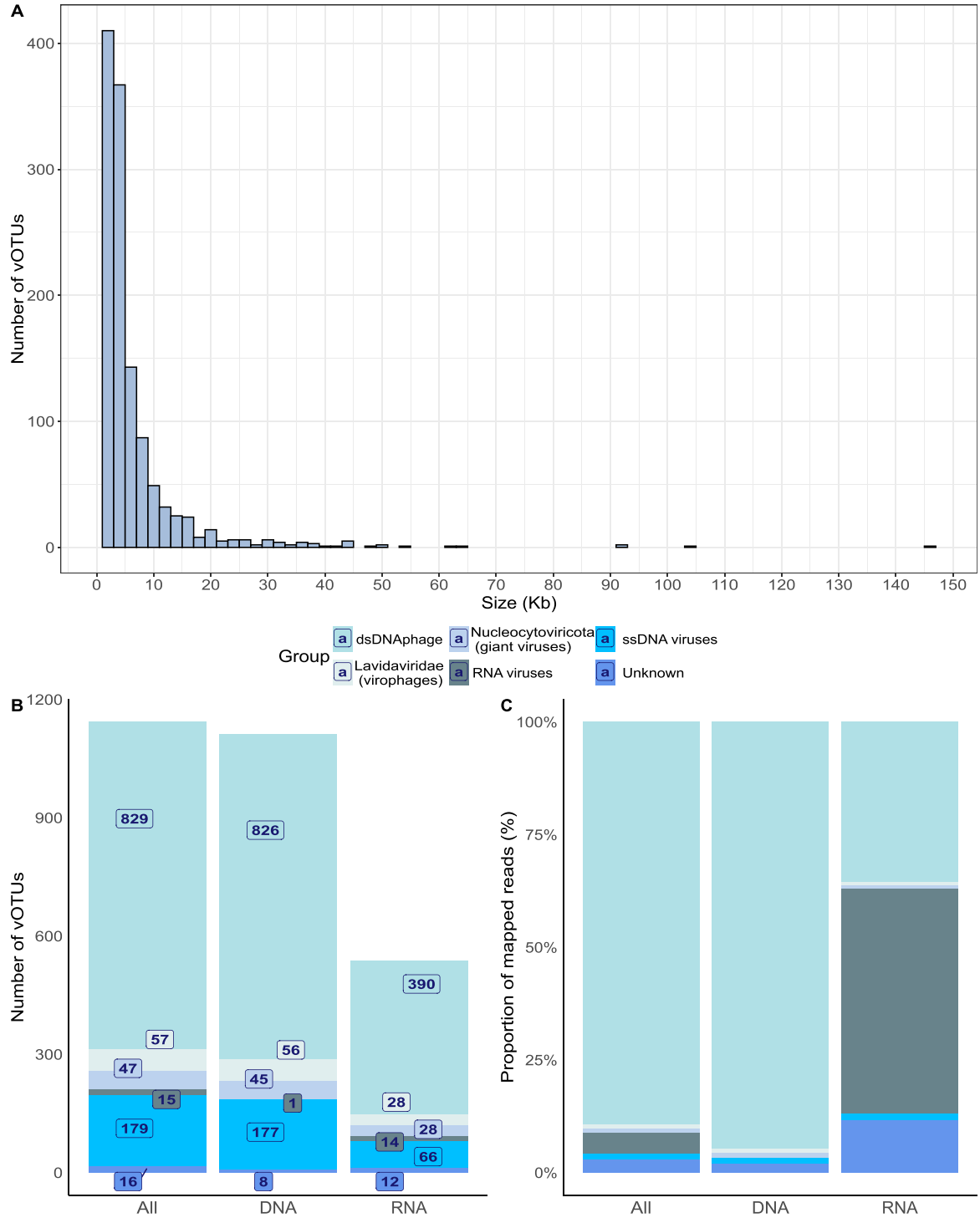


Fig. 2. Overview of the fermented vegetables virome. A) Size distribution of the vOTUs. B) Distribution of the vOTUs per type of virus (VirSorter2 predictions) in the complete dataset (All), DNA or RNA samples. C) Proportion of reads from the different types of viruses in the complete dataset (All), DNA or RNA samples.

viruses belonging to the family *Lavidaviridae* (57 vOTUs, 5% of the total), commonly known as virophages, as well as nucleocytoplasmic large DNA viruses (Nucleocytoviricota) (47 vOTUs, 4% of the total), a diverse group often referred to as giant viruses. To identify the proportion of known vOTUs we compared these 1143 vOTUs against the genomes of viruses present in the ICTV database. Only 20% had sequence homology with known viruses (>50% identity and >50% coverage), while 43% had no hit at all and could therefore correspond to new viruses. Mapping back the sequencing reads to the vOTUs sequences revealed that the relative abundance of the different types of viruses varied according to the nucleic acid used as target for amplification (DNA or RNA) (Fig. 2C). Indeed, as expected, dsDNA viruses were highly abundant in DNA-amplified samples whereas RNA viruses were enriched in RNA-amplified samples.

These results provided a comprehensive overview of the viral diversity present in these fermented vegetables and confirmed that using a combination of DNA and RNA amplification enabled accessing different types of viruses (RNA, ssDNA and dsDNA).

3.3.1. DNA viruses

To analyze with more detail the composition of DNA viruses in these three fermented vegetables, we first focused on DNA samples. A high variation of the percentage of mapped reads was observed between samples, ranging from 0.19% to 51.51%. We decided to keep only samples having more than 5000 mapped reads for further comparisons and therefore excluded two samples (carrot 1 and turnip 3). The potential impact of the type of vegetable used as raw material for

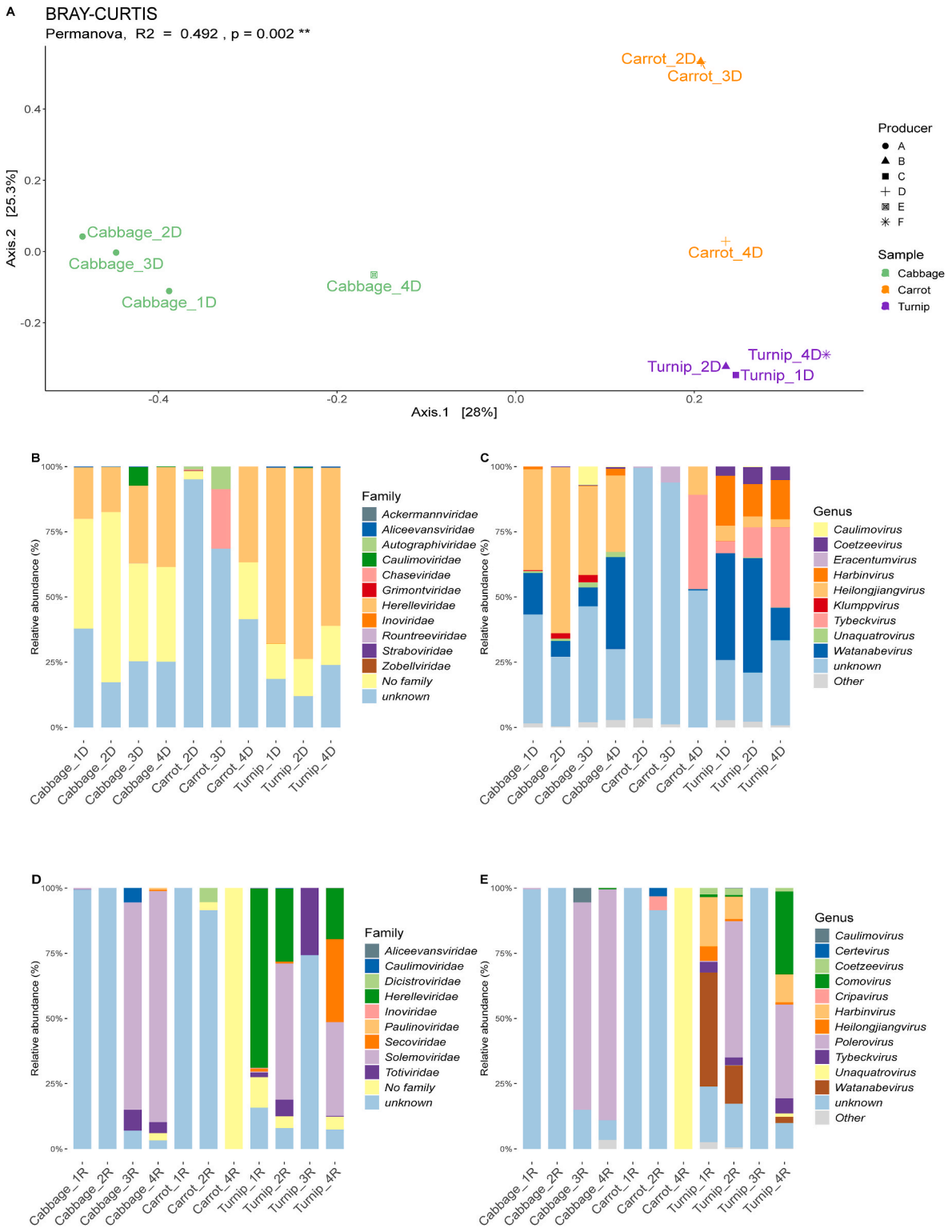


Fig. 3. Structural variations of the viral communities of fermented vegetables. A) Principal coordinate analysis (PCoA) was applied on the Bray-Curtis dissimilarity matrix derived from DNA viral community profiles. Samples are colored according to the vegetable matrix, and shaped according to their producer. B and C) Composition of the DNA viral communities at the family and genus level, respectively (relative abundance of vOTUs). D and E) Composition of the RNA viral communities at the family and genus levels, respectively (relative abundance of vOTUs). "No family" refers to vOTUs associated with subfamilies or genera that are not classified at the family level (ICTV classification).

fermentation on the composition of the DNA-viral community was evaluated by computing Bray-Curtis dissimilarity and principal coordinate analysis (Fig. 3A). A clear separation of the samples was observed according to the three types of vegetables, regardless of the producer. This effect was confirmed by a PERMANOVA test ($p = 0.002$, $r2 = 0.4925$). Interestingly, the principal axis helped discriminate the leaf (cabbage samples) from the root vegetables (carrot and turnip samples). The taxonomic affiliation of vOTUs based on ICTV classification also revealed differences between samples depending on the type of vegetable. A significant proportion of viral communities remained unclassified at both the family and genus levels in all samples (Fig. 3B and C). At the family level, most samples, except for carrot 2D and 3D were dominated by the *Herelleviridae* family or vOTUs affiliated with subfamily or genera lacking a defined family classification (Fig. 3B). In carrot 3D, vOTUs affiliated with the *Chaseviridae* family were present but lacked genus-level classification. Differences between samples were more pronounced at the genus level. *Heilonggjiangvirus* dominated viral communities of cabbage samples, while turnip samples were primarily dominated by *Tybeckvirus*, *Watanabevirus* or *Harbinvirus* genera (Fig. 3C). Most vOTUs in carrots 2D and 3D lacked taxonomic affiliation at this level, whereas *Heilonggjiangvirus* and *Tybeckvirus* were dominant in carrot 4D (Fig. 3C). To further explore the DNA virome composition we represented the distribution of the most abundant vOTUs across our samples through a heatmap (Fig. 4). This visualization method revealed that samples clustered according to the three types of vegetables, confirming the previous result observed using Bray-Curtis dissimilarity

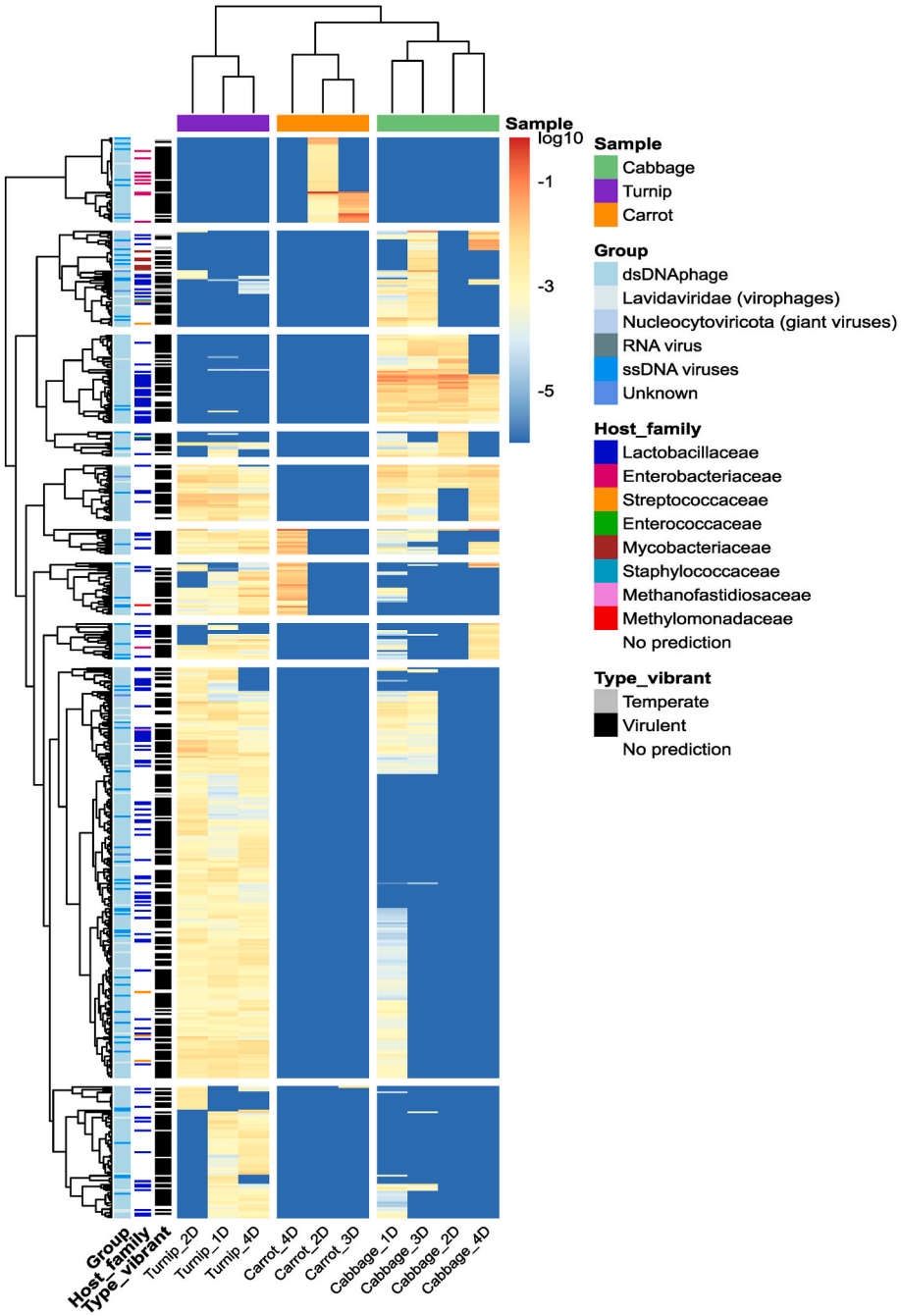


Fig. 4. Distribution of the most abundant vOTUs across fermented vegetables. The colors on the heatmap represent the log-transformed relative abundance and range from blue (low relative abundance) to red (high relative abundance). Host family, viral group, and lifestyle of the different vOTUs are indicated when available. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(Fig. 3). Additionally, distinct patterns among abundant vOTUs were observed, with 10 clusters determined using the gap statistics method (Tibshirani et al., 2001). There were no clear relationships between the vOTU clustering and the viral group, the type of phage (temperate versus virulent) or the bacterial host taxonomy except for cluster 1, 3 and 10. Cluster 1, which was exclusively detected in carrots 2 and 3, predominantly consisted of dsDNA phages with a predicted bacterial host belonging to the *Enterobacteriaceae* family. The most abundant vOTU in carrot 2 (68% relative abundance), which is 41 kb in size, was predicted as a temperate dsDNA phage infecting *Rahnella*. This vOTU was also abundant in carrot 3 (14% relative abundance). The most abundant vOTU detected in carrot 3 (25% relative abundance) was 2 kb in size and did not have a predicted host, nor a viral group or lifestyle prediction. Sequence comparison with known viral genomes revealed a high similarity with viruses affiliated to the *Genomoviridae* (best hit: *Sewage-associated Gemycircularvirus*, 99% identity, 100% coverage), a viral family composed of eukaryote-infecting ssDNA viruses possessing a small genome (usually 2.2–2.4 kb) (Krupovic et al., 2016). Cluster 3 predominantly consisted of dsDNA phages with predicted bacterial hosts from the *Lactobacillaceae* family. The vOTUs belonging to this cluster were mainly detected in cabbage samples. The five most abundant vOTUs in cabbage samples 1 and 2 (5%–12% relative abundance) were dsDNA phages, ranging from 7 to 32 kb in size, all predicted to infect members of the *Lactobacillaceae* family, such as *Lentilactobacillus*, *Lacticaseibacillus*, and *Secundilactobacillus*. This result is consistent with the composition of the bacterial community in our cabbage samples, where the most abundant bacterial genera also belonged to the *Lactobacillaceae* family (Fig. 1). Cluster 10 contained vOTUs primarily detected in turnip samples, consisting mainly of dsDNA phages. Those having host prediction were mainly associated with phages infecting the *Lactobacillaceae* family.

Altogether, these results indicate that the diversity and the relative abundance of DNA viruses vary according to the three types of vegetables used for fermentation. Furthermore, as summarized in Fig. 5, phage sequences, whether originating from free particles (vOTUs) or prophages, were predicted to infect almost all bacterial genera present in the tested fermented vegetables.

3.3.2. DNA viruses corresponding to temperate phages

To identify whether temperate phages are induced from lysogenic bacteria and packaged through fermentation, we compared vOTU sequences from the virome to prophage sequences from the MAGs (Table 3). We considered a given vOTU to correspond to a prophage when both identity and coverage were greater than 90%. Using these criteria, 8 out of the 22 prophages identified from MAGs were also detected through the viral metagenomic approach. The group of five

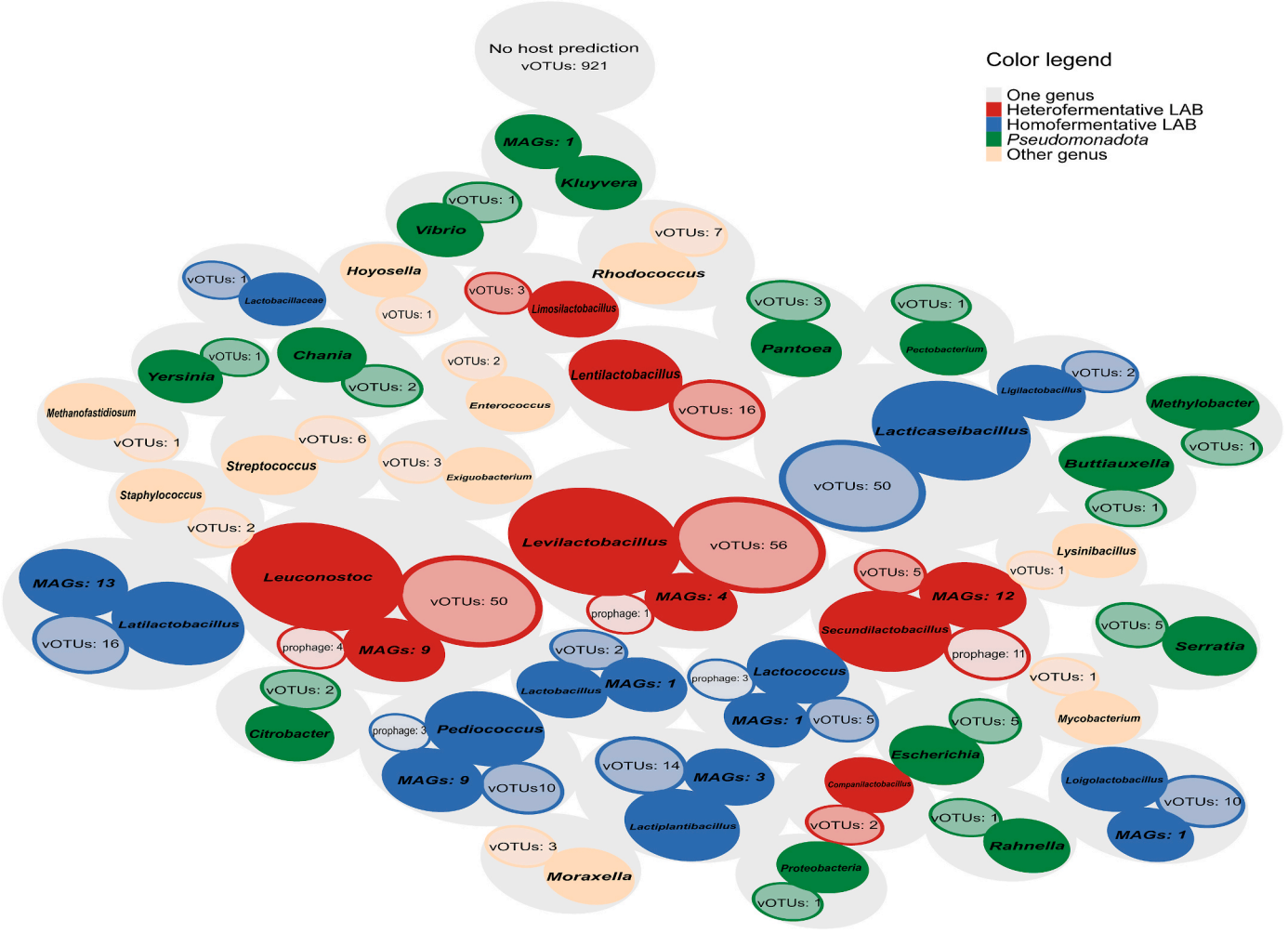


Fig. 5. Summary of phage diversity associated to their bacterial host genera found in fermented vegetables according to shotgun and viral metagenomics data. Each grey bubble represents a bacterial genus detected in the metagenome and contains the number of reconstructed MAGs (full color), the number of prophage sequences detected on MAGs (light color), and the number of associated vOTUs detected in the metavirome (light color and solid line). Bubble size represents the importance of the corresponding bacterial genus in terms of phage signal in the dataset. vOTUs without host prediction are represented in a separate dedicated bubble. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Comparison of vOTUs (metavirome) and prophage sequences (metagenome) from fermented vegetables. Sequence comparison between vOTUs and prophage sequences was performed using BLASTn.

vOTU ID	vOTU host prediction (iPhoP)	vOTU life style prediction (Vibrant)	vOTU's original sample (% relative abundance)	Prophage ID	MAGs species	% identity	% coverage
C4D_NODE_138	–	Temperate	Cabbage_3 (0.04%), Cabbage_4 (2.96%), Carrot_1 (3.77%), Turnip_3 (4.18%)	Cabbage_1_22_2, Cabbage_2_17_2, Cabbage_3_23_2, Cabbage_4_2_2, Carrot_1_22_1	*Secundilactobacillus silagei*	100	98
N4_VD_NODE_338	–	–	Turnip_1 (0.01%), Turnip_4 (0.03%)	Turnip_3_7_1	*Pediococcus parvulus*	91.71	96
C1_1D_NODE_255	*Levilactobacillus*	–	Cabbage_1 (0.05%), Turnip_1 (0.05%)	Cabbage_2_17_3	*Secundilactobacillus silagei*	95.612	100
C1_1D_NODE_459	*Secundilactobacillus*	Temperate	Cabbage_1 (0.01%)	Cabbage_3_23_1	*Secundilactobacillus silagei*	99.733	94

identical prophages detected in MAGs of *Secundilactobacillus silagei* perfectly matched with C4D_NODE_138. This vOTU, was correctly predicted as temperate by Vibrant, but lacked a predicted bacterial host. However, based on its detection in several MAGs identified as *Secundilactobacillus silagei*, we can infer this species as the bacterial host. Interestingly, this phage was detected in high abundance (>1%) in the viromes from three different fermented vegetables (cabbage 4, carrot 1 and turnip 3), and in low abundance in cabbage 3 (0.04%), demonstrating its wide distribution in fermented vegetables. Surprisingly, we did not detect this vOTU in the virome from cabbages 1 and 2, despite the fact that it had been assembled in MAGs from these samples. This suggests that there is probably no induction of the corresponding prophage in these samples but that it is well encapsidated in some others.

Prophage Turnip_3_7_1, identified in a MAG of *Pediococcus parvulus* assembled from turnip 3, showed high sequence homology with vOTU N4_VD_NODE_338 (91.71% identity, 96% coverage) for which no prediction was available regarding both its lifestyle and bacterial host. Interestingly, it was found in low abundance in turnip samples 1 and 4, but was not detected in turnip 3, where the MAG containing the corresponding prophage was assembled. Finally, two vOTUs, C1_1D_NODE_255 and C1_1D_NODE_459, matched with two different prophages detected in *Secundilactobacillus silagei* MAGs assembled from cabbages 2 and 3. C1_1D_NODE_255 was detected in low abundance in cabbage 1 and turnip 1, but not in cabbage 2 from which the MAG containing the corresponding prophage was assembled. Similar observation was made for C1_1D_NODE_459 which was only detected in cabbage 1.

Together, these results suggest that some temperate phages were present as virions in the samples, potentially reflecting basal prophage induction or active induction triggered by specific environmental conditions in fermented vegetables.

3.3.3. RNA viruses

We next focused our attention on RNA samples. In total, 538 vOTUs were detected in those samples, and the total number of mapped reads was low, ranging from 0 (carrot 3) to 353,914 (turnip 4), with an average of 87,897. A high number of vOTUs were detected in turnip samples 1, 2 and 4 (266, 283 and 352 vOTUs, respectively). On the contrary, the number of vOTUs in other samples ranged from 0 (carrot 3) to 19 (cabbage 4). The composition of the RNA viral communities varied between samples (Fig. 3D and E). All vOTUs from cabbage 1R, 2R, and carrot 1R samples remained unclassified by PhaBOX. The viral communities of turnip were diverse, except turnip 3R which was composed of either *Totiviradae* or unclassified vOTUs. In contrast, the viral communities in other turnip samples were dominated by the *Herelleviridae* family, with *Solemoviridae* also contributing in turnip 2R and 4R (Fig. 3D). Finally, carrot 4R was dominated by *Unaquatrovirus* (Fig. 3E). Although we selectively amplified RNA, some vOTUs were detected in both RNA and DNA samples. For example, the most abundant vOTU detected in the RNA sample from cabbage 1 (99% relative abundance), was also present in the corresponding DNA sample at a low relative abundance (0.06%). This vOTU of 53 kb was predicted to infect *Exiguobacterium*, but had no life cycle prediction, no taxonomic affiliation, and did not match with any known viral genome available. For turnip 1, the most abundant contig (10% relative abundance) was 63 kb in length and was also present in several DNA samples (cabbages 1, 3, 4 and in turnips 1, 2 and 4). This vOTU corresponded to a dsDNA phage closely related to *Lactobacillus* phage 3–521 (92% identity, 87% coverage) and belongs to *Watanabevirus* genus. For carrots, most abundant vOTUs in RNA samples were also detected in DNA samples. In total, only 29 out of the 538 vOTUs from the RNA virome were not detected in the DNA samples and were therefore likely to originate from RNA viruses. Among them were the most abundant vOTUs in RNA samples from cabbages 2, 3, and 4, and turnips 2, 3, and 4. Their sequences were next compared to the NCBI nt database and annotated with Pharokka.

Among these vOTUs, thirteen showed high sequence homology with known non-bacterial viruses (Table 4). Eight vOTUs corresponded to plant RNA viruses, such as turnip yellows virus and beet western virus. These vOTUs were present at different levels of abundance in cabbages 1, 3 and 4 and turnips 1, 2 and 4 and were not detected in carrot samples. Two vOTUs were homologous to insect and invertebrate viruses. N4_VR_700k_NODE_19, which was detected in both turnips 3 and 4, was closely related to *Drosophila* A virus (Webster et al., 2015). Ct2R_700k_NODE_2, present in carrot 2, was similar to a MAG assembled from lake sediment and was identified as *Cripavirus*, a genus of RNA viruses infecting invertebrates. Interestingly, a vOTU (N3R_100k_NODE_1) of 3 kb in size, detected in high relative abundance in turnip 3 (24.5%) and in lower abundance in cabbage 4 (0.54%), shared some degree of homology with a yeast virus (*Pichia membranifaciens* virus L-A, 73.83% sequence identity and 28% coverage). This virus belongs to the *Totiviridae* family and represents the first double-stranded RNA (dsRNA) virus infecting *Pichia membranifaciens* (M. D. Lee et al., 2022). It is noteworthy to mention that *Pichia* was the most dominant fungal genus detected in turnip 3 (Fig. 1C) and was also slightly detected in cabbage 4. Finally, two vOTUs (N4_VR_NODE_116 and N1R_NODE_20) matched only with uncultivated viral genomes (UViGs): one annotated as *Riboviria*, a realm of viruses that includes all viruses that use a homologous RNA-dependent polymerase for replication (Roux et al., 2019; Walker et al., 2019), and the second as *Genomoviridae*, a family of single-stranded DNA viruses that primarily infect fungi (Krupovic et al., 2016).

In addition to non-bacterial viruses, four vOTUs specific to cabbage RNA samples had positive hits with sequences originating from bacteria but were not identified as RNA phages (Table 5). The two first (C2_1R_NODE_9 and C2_1R_NODE_21), detected in cabbage 2 and identified as dsDNA phages by VirSorter2 and belong to *Caudoviricetes* class by PhaBOX, partially matched with putative prophage regions from *Moraxellaceae bacterium* detected in marine samples using single-cell genomics. The two other vOTUs (C4R_700k_NODE_40 and C4R_700k_NODE_49), both detected in cabbage 4 in low relative

Table 4
Description of vOTUs detected only in RNA samples and similar to known viruses.

vOTU ID	vOTU's original sample (% relative abundance)			Virus type	Best hit (NCBI nt)	% identity	% coverage	ICTV classification at the genus level (PhaBOX)
	Cabbage	Carrot	Turnip					
N4_VR_400k_NODE_2	Cabbage_3 (29%), Cabbage_4 (34%),	–	Turnip_1 (0.36%), Turnip_2 (21%), Turnip_4 (29%),	Plant RNA virus	OP797722.1 Turnip yellows virus	98.81	99	*Polerovirus*
N4_VR_700k_NODE_7	Cabbage_1 (0.46%) Cabbage_3 (31%), Cabbage_4 (46%),	–	Turnip_1 (0.07%), Turnip_2 (33%), Turnip_4 (11%)	Plant RNA virus	MN497804.1 Turnip yellows virus	99.63	99	*Polerovirus*
N4_VR_700k_NODE_3	–	–	Turnip_1 (1.39%), Turnip_2 (1.31%), Turnip_4 (18%)	Plant RNA virus	NC_078907.1 Turnip ringspot virus	92.50	99	*Comovirus*
N4_VR_700k_NODE_13	Cabbage_1 (0.13%), Cabbage_3 (5.52%), Cabbage_4 (3.46%),	–	Turnip_2 (33%), Turnip_4 (3.46%)	Plant RNA virus	ON603912.1 Beet western yellows virus	98.87	98	–
N4_VR_400k_NODE_8	Cabbage_1 (0.01%), Cabbage_3 (8.12%), Cabbage_4 (2.39%),	–	Turnip_2 (1.50%), Turnip_4 (1.74%)	Plant RNA virus	MN497828.1 Turnip yellows virus	99.10	98	–
N4_VR_NODE_16	Cabbage_4 (0.18%),	–	Turnip_4 (2.38%)	Plant RNA virus	NC_078907.1 Turnip ringspot virus	80.83	98	*Comovirus*
C3_1R_400k_NODE_4	Cabbage_1 (0.001%), Cabbage_3 (0.44%), Cabbage_4 (2.05%),	–	Turnip_2 (2.4%), Turnip_4 (0.31%)	Plant RNA virus	NC_078907.1 Turnip ringspot virus	99.61	99	–
C4R_NODE_124	Cabbage_3 (0.76%), Cabbage_4 (0.77%),	–	Turnip_2 (0.63%), Turnip_4 (0.28%)	Plant RNA virus	MN497827.1 Turnip yellows virus	99.61	99	–
N4_VR_700k_NODE_19	–	–	Turnip_3 (0.55%) Turnip_4 (0.01%)	Insect RNA virus	KP969923.1 Drosophila A virus	98.46	99	–
Ct2R_700k_NODE_2	–	Carrot_2 (1.6%)	–	Invertebrate RNA virus	MZ679063.1 MAG: *Cripavirus* sp. isolate 201-k141_128585	91.04	99	*Cripavirus*
N3R_100k_NODE_1	Cabbage_4 (3.76%)	–	Turnip_3 (24.5%)	Yeast RNA virus	OL687555.1 *Pichia membranifaciens* virus L-A	73.83	28	*Totiviridae* (family)
N4_VR_NODE_116	–	–	Turnip_4 (0.54%)	Unknown *Riboviria* virus	OQ423855.1 MAG: *Riboviria* sp. isolate crane70_contig_1245	76.36	47	–
N1R_NODE_20	–	–	Turnip_1 (0.09%)	Unknown *Genomoviridae* virus	MT309887.1 MAG: *Genomoviridae* sp. isolate 6433_227	92.37	100	–

abundance, were annotated as *Nucleocytoviricota* and ssDNA virus by VirSorter2, respectively and were not classified by PhaBOX (ICTV classification). However, sequence comparison with the nt database from NCBI revealed only partial homology with bacterial DNA and automatic annotation with Pharokka did not identify viral genes, thus probably reflecting background contamination.

Regarding the 12 other RNA-specific vOTUs, sequence comparison with the nt database from NCBI at the nucleotidic level did not reveal any positive hit. We therefore performed an automatic annotation of the vOTU with Pharokka using the PHROGs viral protein database (Terzian et al., 2021). Among all the results, one of the 17 predicted proteins in vOTU C4R_NODE_16 (7 kb in length), was annotated as Zot-like toxin moron, while all others were hypothetical proteins. All PHROGs (protein orthologous groups) associated with the proteins of this vOTU were composed of *Inoviridae* proteins. Inoviruses are filamentous bacteriophages containing a circular, positive sense, single stranded DNA genome. They are distributed across all biomes and display an extremely broad host range spanning both prokaryotic domains of life, bacteria and archaea (Knezevic and Adriaenssens, 2021; Roux et al., 2019). This vOTU was found in cabbage 4 at 0.3% relative abundance. C4R_700k_NODE_2, which was 10 kb in size, was present only in Cabbage 4 (3% relative abundance). It was predicted to belong to *Lavidaviridae* (virophages), a double-stranded (ds) DNA virus family. Pharokka predicted 22 genes on this vOTU, four of which having known

Table 5
Description of vOTUs found only in RNA samples and similar to bacterial sequences.

vOTU ID	vOTU's original sample (% relative abundance)			Protein annotation (Pharokka)	Virus type (VirSorter2)	Best hit (NCBI nt)	% identity	% coverage
	Cabbage	Carrot	Turnip					
C2_1R_NODE_9	Cabbage_2 (18.8%)	–	–	29 (replication initiation O-like, NinB/Orf homologous recombination mediator, HNH endonuclease, terminase small subunit, terminase large subunit, portal protein, unknown function)	dsDNAphage	LC662981.1 Moraxellaceae bacterium, putative prophage region	95.60	47
C2_1R_NODE_21	Cabbage_2 (7.15%)	–	–	18 (exonuclease, Erf-like ssDNA annealing protein, transcriptional repressor, unknown function)	dsDNAphage	LC662967.1 Moraxellaceae bacterium, putative prophage region	91.40	50
C4R_700k_NODE_40	Cabbage_4 (0.32%)	–	–	5 (unknown function)	Nucleocytoviricota (giant viruses)	CP085993.1 Psychrobacter sanguinis strain, plasmid unnamed	86.15	28
C4R_700k_NODE_49	Cabbage_4 (0.25%)	–	–	6 (transposase, unknown function)	ssDNA viruses	CP069437.1 Moraxella osloensis strain, complete genome	86.13	75

functions but being non-essential for bacteriophages.

To summarize, most of the confirmed RNA viruses detected in our fermented vegetables correspond to non-bacterial viruses and probably originate from the phyllosphere.

4. Discussion

Both culture-based and metagenomic approaches were used in this study to investigate the composition of microbial communities in 12 fermented vegetables, showing that *Bacillota* was the main phylum in most samples. LAB from the *Lactobacillaceae* family dominated bacterial populations, with concentration measured by plate counting around 6.8 log CFU/g of vegetables. LAB also represented 83% relative abundance of the bacterial metagenomics reads on average. These findings align with microbial composition data reported in previous studies conducted on fermented vegetables (Bati and Boyko, 2017; Plengvidhya et al., 2007; Thierry et al., 2023b; Wuyts et al., 2018; Zabat et al., 2018). Among LAB, metagenomic results showed the presence of *Secundilactobacillus silagei* in all cabbage samples, a species first isolated from silage (Tohno et al., 2013; Zheng et al., 2020). The *Secundilactobacillus genus* was previously detected in fermented vegetables at relative abundance >1% using 16S metabarcoding and one isolate of *S. malefermentans* was obtained through cultivation-based approach (Thierry et al., 2023b). The second most abundant phylum was *Pseudomonadota,* among which *Bradyrhizobium* was abundant in two carrot samples. *Bradyrhizobium,* a bacterial genus commonly found in soils (Delgado-Baquerizo et al., 2018), was not previously associated with fermented vegetables. We hypothesize an environmental origin for these bacteria in our samples. *Enterobacteriaceae* were also present in our samples (2% on average). These bacteria are commonly found in fermented vegetables, especially in the early stages of the fermentation process (Junker et al., 2024; Leff and Fierer, 2013). Overall, the bacterial community structure in our samples varied depending on the three vegetables used as raw material, which contradicts recent findings (Thierry et al., 2023b). However, this discrepancy may be explained by the smaller number of samples and vegetables tested in our study. Finally, yeasts were also detected in all the studied samples through metagenomics and in over half the samples using the cultural approach, with a median count of 4.6 log CFU/g. This result was consistent with previous studies (Lu et al., 2003; Thierry et al., 2023b). *Kazachstania* was the most abundant fungal genus (67% of reads assigned to Eukaryota), followed by *Pichia* (7.7%). Both genera were already detected in Chinese sauerkraut (Liu et al., 2021). While some *Pichia* are sometimes considered as spoilage microorganisms, they were also detected in successful fermentations, highlighting the need for further research on the role of yeasts in vegetable fermentation (Ballester et al., 2022).

Although the presence of viruses in fermented vegetables has been demonstrated for more than two decades (Lu et al., 2003), there is currently a lack of information on their concentration in this type of food. For the first time, we estimated the abundance of VLPs in fermented vegetables using a non-selective method based on epifluorescence microscopy. The counts ranged from 5.28 to 7.57 log VLPs/g of vegetables, which was in the same range as the number of bacteria. By comparison, these values are lower than those observed for cheese rind (10 log VLPs/g) (Dugat-Bony et al., 2020), human feces (9–10 log VLPs/g) (Shkoporov and Hill, 2019), agricultural soils (8–9 log VLPs/g) (Roy et al., 2020; Williamson et al., 2017) and oceans (5–11 log VLPs/ml) (Culley and Welschmeyer, 2002; Tsai et al., 2021), but higher than those for Saudi Arabia desert sands (3–4 log VLPs/g) (Gonzalez-Martin et al., 2013) and rainwaters (4–5 log VLPs/ml) (Rahlff et al., 2023). Importantly, our data suggest that there is no significant variation in the VLP counts in fermented vegetables based on the three types of vegetables used as raw material. However, more samples from a broader range of vegetables should be analyzed to confirm it.

Thanks to the methods applied in this study, we provided a comprehensive overview of the viral composition in fermented vegetables, including prophages, ssDNA, and RNA viruses, in addition to the dsDNA viruses that are commonly described in metavirome studies. Regarding DNA viruses, it was not possible to predict a host for most of the detected vOTUs. However, as summarized in Fig. 5, the others mostly corresponded to bacteriophages infecting homofermentative and heterofermentative LAB, and to a lesser extent bacteria from the *Pseudomonadota* phylum. Our results also indicated that lysogeny is common in LAB from fermented vegetables, as many prophage sequences were detected in MAGs reconstructed from the metagenome. Few of them were also detected in the metavirome indicating they were also encapsidated and present as free particles in fermented vegetables. Future investigation would now be necessary to confirm whether these prophages are active or remnant sequences, through induction tests on representative bacterial isolates. It would also be interesting to follow the evolution of the prophages to confirm whether those present in virion form are induced or whether the decrease in diversity favors the lysogenic cycle due to the absence of predator/prey dynamics (Ledormand et al., 2022). Previous studies already reported the isolation of several dsDNA phages from fermented vegetables through cultural approach (Barrangou et al., 2002; Lu et al., 2003, 2010, 2012; Lu and Breidt, 2015; Yoon et al., 2001, 2007). For example, 171 virulent bacteriophages infecting LAB were isolated from sauerkraut sampled during the fermentation process (Lu et al., 2003). It revealed the succession of two phage populations infecting different LAB species, with a shift

occurring 3–7 days after the start of fermentation. A previous viral metagenomic study on kimchi also showed that the main viral sequences were predicted to infect LAB (Jung et al., 2018). The authors also demonstrated that the composition of the viral community varied according to the geographical origin of the samples. Among ssDNA viruses, we detected members of the *Genomoviridae* family as previously described in Korean and Chinese kimchi (Jung et al., 2018). This family of viruses mainly infect fungi (Krupovic et al., 2016). We also detected plant ssRNA viruses in both cabbage and turnip samples, which is consistent with previous metatranscriptomic analysis of kimchi (Kim et al., 2014; H.-W. Lee et al., 2022). Surprisingly, we did not detect such viruses in carrot samples. This might be due to some difficulties in obtaining sufficient amounts of nucleic acids from this type of sample. In addition, our methodological approach enabled the detection of insect viruses in fermented vegetables. Although not directly comparable, previous work on wine reported the isolation of virulent *Oenococcus oeni* bacteriophages that are more closely related to insect gut phages than any other LAB phages previously sequenced (Philippe et al., 2020). This suggests insects as a possible source for a wide diversity of viruses contaminating plant material used in fermentation processes including fruits and vegetables. Finally, we also observed sequences closely related to a dsRNA mycovirus, namely *Pichia membranifaciens* L-A virus. This virus was recently identified and represents a new species within the *Totiviridae* family related to yeast totiviruses (M. D. Lee et al., 2022). This type of virus was discovered in toxin-secreting strains of *Saccharomyces cerevisiae*, known as 'Killer yeasts', which secrete protein toxins that are lethal to sensitive strains (Ghabrial et al., 2015). In a killer yeast, production of such toxin is associated with the presence of a satellite virus, which requires a L-A helper virus to be stably maintained and replicated within the infected cell (Schmitt and Breinig, 2002). Although such types of mycoviruses appear to be symptomless for the infected yeast cells, they might be responsible for competition among yeast strains during food fermentation (Maqueda et al., 2012; Maske et al., 2022). Currently, there is an increasing interest in *Pichia* for food fermentation applications (Vicente et al., 2021; Zhang et al., 2021), so it could be relevant to isolate such viruses and investigate their effect on yeast growth and physiology, and ultimately on fermentation as well.

Viral metagenomics has gained increasing attention in recent years for understanding the microbial ecology in fermented foods (Ledormand et al., 2021; Paillet and Dugat-Bony, 2021; Zhang et al., 2024). However, one of the main caveats of such an approach is that it produces a large number of unclassified viral sequences preventing the accurate description of phage communities in fermented foods. In this study, 43% of the assembled vOTUs did not share any sequence homology to well classified viral sequences present in ICTV database, a similar proportion to what was observed in viromes of fermented shrimp, kimchi, and sauerkraut produced thirteen years ago (Park et al., 2011). Several methods for classifying phage sequences (both full genomes and partial sequences derived from metaviromes) already exist (Bin Jang et al., 2019; Jiang et al., 2023; Pons et al., 2021), but all depend on the sequence databases used for comparison. Although research on bacteriophages in dairy products and wine is currently very active (Chaïb et al., 2022; Paillet et al., 2022; Philippe et al., 2020), it would be beneficial for future research to focus on expanding viral sequence databases with new genomes from phages isolated from other fermented foods (Zhang et al., 2024). A previous study of our group on French smear-ripened cheeses revealed the dynamics of the viral community during cheese ripening (Paillet et al., 2024). For vegetable fermentation, the dynamics of microbial communities have been well described in previous studies using culture-based approaches and metabarcoding (16S rRNA gene sequencing) (Lu et al., 2003, 2012; Wuyts et al., 2018), while the dynamics of viral communities have been described by culture-based approaches (Lu et al., 2003, 2012) but remain largely unexplored. Gaining new insights into phage ecology is now essential for a comprehensive understanding of microbial succession during vegetable fermentation.

External environmental factors, such as acidity, temperature, and ions are known to play a role in phage-bacteria interactions (Ranveer et al., 2024), and could therefore impact the composition of viral communities in fermented products. Indeed, environmental changes can modify the metabolism of the host, including DNA replication, synthesis, and the activity of virion assembly machinery, ultimately affecting infection kinetics (Zaburlin et al., 2017). For example, a study on dairy phages infecting lactic acid bacteria revealed that phage infection was impacted by the availability of nitrogen sources, temperature, and pH (Zaburlin et al., 2017). However, the magnitude and direction of these effects varied across different phage-host pairs. Interestingly, some bacteriophages have the ability to switch from a lytic to a temperate lifestyle according to temperature (Shan et al., 2014). Phage inactivation by osmotic stress has also been reported, for example in the marine environment (Jonczyk et al., 2011). Fermented vegetables have distinct environmental characteristics in terms of temperature (ranging from 15 to 20 °C), pH (which decreases from approx. 6 in fresh vegetables to lower than 4 at the end of fermentation) and salinity (with typical salt levels comprised between 2 and 5%). Phages are present in this type of environment, as demonstrated in Sauerkraut (Lu et al., 2003) and as revealed in our viral metagenomic survey. However, future work should aim at studying the impact of such environmental factors on phages isolated from this type of environment and on the structure of phage communities during fermentation.

5. Conclusion

In conclusion, this study underscores that fermented vegetables harbor complex viral communities that vary depending on the vegetable used as raw material. These communities are predominantly composed of bacteriophages, with a notable presence of temperate phages infecting dominant lactic acid bacteria. In addition, the presence of other viruses such as plant, insect, invertebrate and yeast viruses was also demonstrated, providing, for the first time, a comprehensive understanding of the viral diversity in fermented vegetables as a whole. Future research should now investigate how these viruses, especially bacteriophages and yeast viruses, influence microbial succession during vegetable fermentation and their potential impact on the quality of the final product. Such knowledge would be of utmost importance for professionals in the field to mitigate the risks associated with the proliferation of phages in industrial fermented vegetable production settings.

CRediT authorship contribution statement

Julia Cantuti Gendre: Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Claire Le Marrec:** Writing – review & editing, Supervision, Methodology. **Stéphane Chaillou:** Writing – review & editing, Project administration, Funding acquisition. **Lysiane Omhover-Fougy:** Writing – review & editing, Validation, Resources. **Sophie Landaud:** Writing – review & editing, Validation, Supervision. **Eric Dugat-Bony:** Writing – review & editing, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Stephane Chaillou reports financial support was provided by French National Research Agency. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the French National Research Agency

(ANR) as part of the MetaSimFood project (ANR-21-CE21-0003). JG received a PhD grant from the French National Research Agency (ANR) as part of the same project. We are grateful to the INRAE MIGALE bioinformatics facility (https://doi.org/10.15454/1.5572390655343293E12) for providing computing and storage resources. We thank Margareth Renault (INRAE SAYFOOD) for technical assistance with microbial counts and epifluorescence microscopy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fm.2025.104733.

References

Ballester, E., Ribes, S., Barat, J.M., Fuentes, A., 2022. Spoilage yeasts in fermented vegetables: conventional and novel control strategies. Eur. Food Res. Technol. 248, 315–328. https://doi.org/10.1007/s00217-021-03888-7.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Barrangou, R., Yoon, S.-S., Breidt, F., Fleming, H.P., Klaenhammer, T.R., 2002. Characterization of six *Leuconostoc fallax* bacteriophages isolated from an industrial sauerkraut fermentation. Appl. Environ. Microbiol. 68, 5452–5458. https://doi.org/10.1128/AEM.68.11.5452-5458.2002.

Bati, B.B., Boyko, H.B., 2017. Microbiological analysis of sauerkraut in the process of its fermentation according to the traditional and modernized technologies. Microb. Biotechnol. 0, 90–100. https://doi.org/10.18524/2307-4663.2017.2(38).105108.

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister, J. R., Kropinski, A.M., Krupovic, M., Lavigne, R., Turner, D., Sullivan, M.B., 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat. Biotechnol. 37, 632–639. https://doi.org/10.1038/s41587-019-0100-8.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bouras, G., Nepal, R., Houtak, G., Psaltis, A.J., Wormald, P.-J., Vreugde, S., 2023. Pharokka: a fast scalable bacteriophage annotation tool. Bioinformatics 39, btac776. https://doi.org/10.1093/bioinformatics/btac776.

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., Weinstock, G.M., Garrity, G.M., Dodsworth, J.A., Yooseph, S., Sutton, G., Glöckner, F.O., Gilbert, J.A., Nelson, W.C., Hallam, S.J., Jungbluth, S.P., Ettema, T.J. G., Tighe, S., Konstantinidis, K.T., Liu, W.-T., Baker, B.J., Rattei, T., Eisen, J.A., Hedlund, B., McMahon, K.D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G.W., Rinke, C., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D.H., Murat Eren, A., Schriml, L., Banfield, J.F., Hugenholtz, P., Woyke, T., 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35, 725–731. https://doi.org/10.1038/nbt.3893.

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends Microbiol. 13, 278–284. https://doi.org/10.1016/j.tim.2005.04.003.

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., Rohwer, F., 2002. Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. USA 99, 14250–14255. https://doi.org/10.1073/pnas.202488399.

Buckenhueskes, H.J., 2015. Quality improvement and fermentation control in vegetables. In: Advances in Fermented Foods and Beverages. Elsevier, pp. 515–539. https://doi.org/10.1016/B978-1-78242-015-6.00022-0.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 421. https://doi.org/10.1186/1471-2105-10-421.

Camargo, A.P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., Chain, P.S.G., Nayfach, S., Kyrpides, N.C., 2024. Identification of mobile genetic elements with geNomad. Nat. Biotechnol. 42, 1303–1312. https://doi.org/10.1038/s41587-023-01953-y.

Chaïb, A., Philippe, C., Jaomanjaka, F., Barchi, Y., Oviedo-Hernandez, F., Claisse, O., Le Marrec, C., 2022. Phage-host interactions as a driver of population dynamics during wine fermentation: betting on underdogs. Int. J. Food Microbiol. 383, 109936. https://doi.org/10.1016/j.ijfoodmicro.2022.109936.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., Parks, D.H., 2022. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics 38, 5315–5316. https://doi.org/10.1093/bioinformatics/btac672.

Cheigh, H., Park, K., Lee, C.Y., 1994. Biochemical, microbiological, and nutritional aspects of kimchi (Korean fermented vegetable products). Crit. Rev. Food Sci. Nutr. 34, 175–203. https://doi.org/10.1080/10408399409527656.

Culley, A.I., Welschmeyer, N.A., 2002. The abundance, distribution, and correlation of viruses, phytoplankton, and prokaryotes along a Pacific Ocean transect. Limnol. Oceanogr. 47, 1508–1513. https://doi.org/10.4319/lo.2002.47.5.1508.

de Melo, A.G., Lemay, M.-L., Moineau, S., 2023. The impact of *Brevibacterium aurantiacum* virulent phages on the production of smear surface-ripened cheeses. Int. J. Food Microbiol. 400, 110252. https://doi.org/10.1016/j.ijfoodmicro.2023.110252.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D. J., Bardgett, R.D., Maestre, F.T., Singh, B.K., Fierer, N., 2018. A global atlas of the dominant bacteria found in soil. Science 359, 320–325. https://doi.org/10.1126/science.aap9516.

Di Cagno, R., Surico, R.F., Siragusa, S., De Angelis, M., Paradiso, A., Minervini, F., De Gara, L., Gobbetti, M., 2008. Selection and use of autochthonous mixed starter for lactic acid fermentation of carrots, French beans or marrows. Int. J. Food Microbiol. 127, 220–228. https://doi.org/10.1016/j.ijfoodmicro.2008.07.010.

Dugat-Bony, E., Lossouarn, J., De Paepe, M., Sarthou, A.-S., Fedala, Y., Petit, M.-A., Chaillou, S., 2020. Viral metagenomic analysis of the cheese surface: a comparative study of rapid procedures for extracting viral particles. Food Microbiol. 85, 103278. https://doi.org/10.1016/j.fm.2019.103278.

Erkus, O., de Jager, V.C., Spus, M., van Alen-Boerrigter, I.J., van Rijswijck, I.M., Hazelwood, L., Janssen, P.W., van Hijum, S.A., Kleerebezem, M., Smid, E.J., 2013. Multifactorial diversity sustains microbial community stability. ISME J. 7, 2126–2136. https://doi.org/10.1038/ismej.2013.108.

Ghabrial, S.A., Castón, J.R., Jiang, D., Nibert, M.L., Suzuki, N., 2015. 50-plus years of fungal viruses. Virology, 60th Anniversary Issue 479–480, 356–368. https://doi.org/10.1016/j.virol.2015.02.034.

Gonzalez-Martin, C., Teigell-Perez, N., Lyles, M., Valladares, B., Griffin, D.W., 2013. Epifluorescent direct counts of bacteria and viruses from topsoil of various desert dust storm regions. Res. Microbiol. 164, 17–21. https://doi.org/10.1016/j.resmic.2012.08.009.

Gunawardena, S., Nadeeshani, H., Amarasinghe, V., Liyanage, R., 2024. Bioactive properties and therapeutic aspects of fermented vegetables: a review. Food Production, Processing and Nutrition 6, 31. https://doi.org/10.1186/s43014-023-00176-7.

Guo, J., Bolduc, B., Zayed, A.A., Varsani, A., Dominguez-Huerta, G., Delmont, T.O., Pratama, A.A., Gazitúa, M.C., Vik, D., Sullivan, M.B., Roux, S., 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome 9, 37. https://doi.org/10.1186/s40168-020-00990-y.

Jiang, J.-Z., Yuan, W.-G., Shang, J., Shi, Y.-H., Yang, L.-L., Liu, M., Zhu, P., Jin, T., Sun, Y., Yuan, L.-H., 2023. Virus classification for viral genomic fragments using PhaGCN2. Briefings Bioinf. 24, bbac505. https://doi.org/10.1093/bib/bbac505.

Jonczyk, E., Kłak, M., Międzybrodzki, R., Górski, A., 2011. The influence of external factors on bacteriophages—review. Folia Microbiol. 56, 191–200. https://doi.org/10.1007/s12223-011-0039-8.

Jung, M.-J., Kim, M.-S., Yun, J.-H., Lee, J.-Y., Kim, P.S., Lee, H.-W., Ha, J.-H., Roh, S.W., Bae, J.-W., 2018. Viral community predicts the geographical origin of fermented vegetable foods more precisely than bacterial community. Food Microbiol. 76, 319–327. https://doi.org/10.1016/j.fm.2018.06.010.

Junker, R., Valence, F., Mistou, M.-Y., Chaillou, S., Chiapello, H., 2024. Integration of metataxonomic data sets into microbial association networks highlights shared bacterial community dynamics in fermented vegetables. Microbiol. Spectr. 12, e00312. https://doi.org/10.1128/spectrum.00312-24, 24.

Kent, W.J., 2002. BLAT—the BLAST-like alignment tool. Genome Res. 12, 656–664. https://doi.org/10.1101/gr.229202.

Kieft, K., Zhou, Z., Anantharaman, K., 2020. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome 8. https://doi.org/10.1186/s40168-020-00867-0.

Kim, D.S., Jung, J.Y., Wang, Y., Oh, H.J., Choi, D., Jeon, C.O., Hahn, Y., 2014. Plant RNA virus sequences identified in kimchi by microbial metatranscriptome analysis. J. Microbiol. Biotechnol. 24, 979–986. https://doi.org/10.4014/jmb.1404.04017.

Knezevic, P., Adriaenssens, E.M., 2021. ICTV virus taxonomy profile: Inoviridae. J. Gen. Virol. 102, 001614. https://doi.org/10.1099/jgv.0.001614.

Krupovic, M., Ghabrial, S.A., Jiang, D., Varsani, A., 2016. Genomoviridae: a new family of widespread single-stranded DNA viruses. Arch. Virol. 161, 2633–2643. https://doi.org/10.1007/s00705-016-2943-3.

Ledormand, P., Desmasures, N., Dalmasso, M., 2021. Phage community involvement in fermented beverages: an open door to technological advances? Crit. Rev. Food Sci. Nutr. 61, 2911–2920. https://doi.org/10.1080/10408398.2020.1790497.

Ledormand, P., Desmasures, N., Midoux, C., Rué, O., Dalmasso, M., 2022. Investigation of the phageome and prophages in French cider, a fermented beverage. Microorganisms 10, 1203. https://doi.org/10.3390/microorganisms10061203.

Lee, H.-W., Yoon, S.-R., Dang, Y.-M., Yun, J.-H., Jeong, H., Kim, K.-N., Bae, J.-W., Ha, J.-H., 2022a. Metatranscriptomic and metataxonomic insights into the ultra-small microbiome of the Korean fermented vegetable, kimchi. Front. Microbiol. 13. https://doi.org/10.3389/fmicb.2022.1026513.

Lee, M., Song, J.H., Jung, M.Y., Lee, S.H., Chang, J.Y., 2017. Large-scale targeted metagenomics analysis of bacterial ecological changes in 88 kimchi samples during fermentation. Food Microbiol. 66, 173–183. https://doi.org/10.1016/j.fm.2017.05.002.

Lee, M.D., Creagh, J.W., Fredericks, L.R., Crabtree, A.M., Patel, J.S., Rowley, P.A., 2022b. The characterization of a novel virus discovered in the yeast Pichia membranifaciens. Viruses 14, 594. https://doi.org/10.3390/v14030594.

Leff, J.W., Fierer, N., 2013. Bacterial communities associated with the surfaces of fresh fruits and vegetables. PLoS One 8, e59310. https://doi.org/10.1371/journal.pone.0059310.

Liu, Z., Li, J., Zhou, X., Wei, B., Xie, S., Du, T., Zhao, X., Jiang, L., Xiong, T., 2021. The lactic acid bacteria and yeast community of home-made sauerkraut from three provinces in Southwest China. Arch. Microbiol. 203, 3171–3182. https://doi.org/10.1007/s00203-021-02222-9.

Lu, Z., Altermann, E., Breidt, F., Kozyavkin, S., 2010. Sequence analysis of Leuconostoc mesenteroides bacteriophage φ1-A4 isolated from an industrial vegetable fermentation. Appl. Environ. Microbiol. 76, 1955–1966. https://doi.org/10.1128/AEM.02126-09.

Lu, Z., Breidt, F., 2015. Escherichia coli O157:H7 bacteriophage Φ241 isolated from an industrial cucumber fermentation at high acidity and salinity. Front. Microbiol. 6, 67. https://doi.org/10.3389/fmicb.2015.00067.

Lu, Z., Breidt, F., Plengvidhya, V., Fleming, H.P., 2003. Bacteriophage ecology in commercial sauerkraut fermentations. Appl. Environ. Microbiol. 69, 3192–3202. https://doi.org/10.1128/AEM.69.6.3192-3202.2003.

Lu, Z., Pérez-Díaz, I.M., Hayes, J.S., Breidt, F., 2020. Bacteriophages infecting gram-negative bacteria in a commercial cucumber fermentation. Front. Microbiol. 11. https://doi.org/10.3389/fmicb.2020.01306.

Lu, Z., Pérez-Díaz, I.M., Hayes, J.S., Breidt, F., 2012. Bacteriophage ecology in a commercial cucumber fermentation. Appl. Environ. Microbiol. 78, 8571–8578. https://doi.org/10.1128/AEM.01914-12.

Maqueda, M., Zamora, E., Álvarez, M.L., Ramírez, M., 2012. Characterization, ecological distribution, and population dynamics of Saccharomyces sensu stricto killer yeasts in the spontaneous grape must fermentations of southwestern Spain. Appl. Environ. Microbiol. 78, 735–743. https://doi.org/10.1128/AEM.06518-11.

Marco, M.L., Heeney, D., Binda, S., Cifelli, C.J., Cotter, P.D., Foligné, B., Gänzle, M., Kort, R., Pasin, G., Pihlanto, A., Smid, E.J., Hutkins, R., 2017. Health benefits of fermented foods: microbiota and beyond. Curr. Opin. Biotechnol. 44, 94–102. https://doi.org/10.1016/j.copbio.2016.11.010.

Maske, B.L., De Carvalho Neto, D.P., da Silva, G.B., De Dea Lindner, J., Soccol, C.R., de Melo Pereira, G.V., 2022. Yeast viruses and their implications in fermented foods and beverages. Curr. Opin. Food Sci. 47, 100879. https://doi.org/10.1016/j.cofs.2022.100879.

McMurdie, P.J., Holmes, S., 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8, e61217. https://doi.org/10.1371/journal.pone.0061217.

Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat. Commun. 7, 11257. https://doi.org/10.1038/ncomms11257.

Molfetta, M., Morais, E.G., Barreira, L., Bruno, G.L., Porcelli, F., Dugat-Bony, E., Bonnarme, P., Minervini, F., 2022. Protein sources alternative to meat: state of the art and involvement of fermentation. Foods 11, 2065. https://doi.org/10.3390/foods11142065.

Mudoor Sooresh, M., Willing, B.P., Bourrie, B.C.T., 2023. Opportunities and challenges of understanding community assembly in spontaneous food fermentation. Foods 12, 673. https://doi.org/10.3390/foods12030673.

Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S., Kyrpides, N.C., 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat. Biotechnol. 39, 578–585. https://doi.org/10.1038/s41587-020-00774-7.

Oksanen, Simpson, Kindt, Legendre, Minchin, O., Hara, Solymos, Stevens, Szoecs, Wagner, Barbour, Bedward, Bolker, Borcard, Carvalho, Chirico, De Caceres, Durand, Evangelista, FitzJohn, Friendly, Furneaux, Hannigan, Hill, Lahti, McGlinn, Ouellette, Ribeiro, Cunha, Smith, Stier, Ter Braak, Weedon, 2024. vegan: an R package for community ecologists • vegan [WWW Document]. URL. https://vegandevs.github.io/vegan/, 9.18.24.

Olo Ndela, E., Roux, S., Henke, C., Sczyrba, A., Sime Ngando, T., Varsani, A., Enault, F., 2023. Reekeekee- and roodoodooviruses, two different Microviridae clades constituted by the smallest DNA phages. Virus Evolution 9, veac123. https://doi.org/10.1093/ve/veac123.

Paillet, T., Dugat-Bony, E., 2021. Bacteriophage ecology of fermented foods: anything new under the sun? Curr. Opin. Food Sci. 40, 102–111. https://doi.org/10.1016/j.cofs.2021.03.007.

Paillet, T., Lamy-Besnier, Q., Figueroa, C., Petit, M.-A., Dugat-Bony, E., 2024. Dynamics of the viral community on the surface of a French smear-ripened cheese during maturation and persistence across production years. mSystems 0, e00201. https://doi.org/10.1128/msystems.00201-24, 24.

Paillet, T., Lossouarn, J., Figueroa, C., Midoux, C., Rué, O., Petit, M.-A., Dugat-Bony, E., 2022. Virulent phages isolated from a smear-ripened cheese are also detected in reservoirs of the cheese factory. Viruses 14, 1620. https://doi.org/10.3390/v14081620.

Park, E.-J., Kim, K.-H., Abell, G.C.J., Kim, M.-S., Roh, S.W., Bae, J.-W., 2011. Metagenomic analysis of the viral communities in fermented foods. Appl. Environ. Microbiol. 77, 1284–1291. https://doi.org/10.1128/AEM.01859-10.

Philippe, C., Chaïb, A., Jaomanjaka, F., Claisse, O., Lucas, P.M., Samot, J., Cambillau, C., Le Marrec, C., 2020. Characterization of the first virulent phage infecting Oenococcus oeni, the queen of the cellars. Front. Microbiol. 11, 596541. https://doi.org/10.3389/fmicb.2020.596541.

Plengvidhya, V., Breidt, F., Lu, Z., Fleming, H.P., 2007. DNA fingerprinting of lactic acid bacteria in sauerkraut fermentations. Appl. Environ. Microbiol. 73, 7697–7702. https://doi.org/10.1128/AEM.01342-07.

Pons, J.C., Paez-Espino, D., Riera, G., Ivanova, N., Kyrpides, N.C., Llabrés, M., 2021. VPF-Class: taxonomic assignment and host prediction of uncultivated viruses based on viral protein families. Bioinformatics 37, 1805–1813. https://doi.org/10.1093/bioinformatics/btab026.

Rahlff, J., Esser, S.P., Plewka, J., Heinrichs, M.E., Soares, A., Scarchilli, C., Grigioni, P., Wex, H., Giebel, H.-A., Probst, A.J., 2023. Marine viruses disperse bidirectionally along the natural water cycle. Nat. Commun. 14, 6354. https://doi.org/10.1038/s41467-023-42125-5.

Raivo Kolde, 2010. Pheatmap: pretty heatmaps. https://doi.org/10.32614/CRAN.package.pheatmap.

Ranveer, S.A., Dasriya, V., Ahmad, M.F., Dhillon, H.S., Samtiya, M., Shama, E., Anand, T., Dhewa, T., Chaudhary, V., Chaudhary, P., Behare, P., Ram, C., Puniya, D. V., Khedkar, G.D., Raposo, A., Han, H., Puniya, A.K., 2024. Positive and negative aspects of bacteriophages and their immense role in the food chain. npj Sci Food 8, 1. https://doi.org/10.1038/s41538-023-00245-8.

Rosario, K., Breitbart, M., 2011. Exploring the viral world through metagenomics. Current Opinion in Virology 1, 289–297. https://doi.org/10.1016/j.coviro.2011.06.004.

Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E.V., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Lavigne, R., Brister, J.R., Varsani, A., Amid, C., Aziz, R.K., Bordenstein, S.R., Bork, P., Breitbart, M., Cochrane, G.R., Daly, R.A., Desnues, C., Duhaime, M.B., Emerson, J.B., Enault, F., Fuhrman, J.A., Hingamp, P., Hugenholtz, P., Hurwitz, B.L., Ivanova, N.N., Labonté, J.M., Lee, K.-B., Malmstrom, R.R., Martinez-Garcia, M., Mizrachi, I.K., Ogata, H., Páez-Espino, D., Petit, M.-A., Putonti, C., Rattei, T., Reyes, A., Rodriguez-Valera, F., Rosario, K., Schriml, L., Schulz, F., Steward, G.F., Sullivan, M.B., Sunagawa, S., Suttle, C.A., Temperton, B., Tringe, S.G., Thurber, R.V., Webster, N.S., Whiteson, K.L., Wilhelm, S.W., Wommack, K.E., Woyke, T., Wrighton, K.C., Yilmaz, P., Yoshida, T., Young, M.J., Yutin, N., Allen, L.Z., Kyrpides, N.C., Eloe-Fadrosh, E.A., 2019. Minimum information about an uncultivated virus genome (MIUViG). Nat. Biotechnol. 37, 29–37. https://doi.org/10.1038/nbt.4306.

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 Biol. 21, e3002083. https://doi.org/10.1371/journal.pbio.3002083.

Roy, K., Ghosh, D., DeBruyn, J.M., Dasgupta, T., Wommack, K.E., Liang, X., Wagner, R. E., Radosevich, M., 2020. Temporal dynamics of soil virus and bacterial populations in agricultural and early plant successional soils. Front. Microbiol. 11. https://doi.org/10.3389/fmicb.2020.01494.

Rul, F., Béra-Maillet, C., Champomier-Vergès, M.C., El-Mecherfi, K.E., Foligné, B., Michalski, M.C., Milenkovic, D., Savary-Auzeloux, I., 2022. Underlying evidence for the health benefits of fermented foods in humans. Food Funct. 13, 4804–4824. https://doi.org/10.1039/D1FO03989J.

Sanlier, N., Gökcen, B.B., Sezgin, A.C., 2019. Health benefits of fermented foods. Crit. Rev. Food Sci. Nutr. 59, 506–527. https://doi.org/10.1080/10408398.2017.1383355.

Schmitt, M.J., Breinig, F., 2002. The viral killer system in yeast: from molecular biology to application. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Rev. 26, 257–276. https://doi.org/10.1111/j.1574-6976.2002.tb00614.x.

Schroeder, A.B., Dobson, E.T.A., Rueden, C.T., Tomancak, P., Jug, F., Eliceiri, K.W., 2021. The ImageJ ecosystem: open-source software for image visualization, processing, and analysis. Protein Sci. 30, 234–249. https://doi.org/10.1002/pro.3993.

Shah, S.A., Deng, L., Thorsen, J., Pedersen, A.G., Dion, M.B., Castro-Mejía, J.L., Silins, R., Romme, F.O., Sausset, R., Jessen, L.E., Ndela, E.O., Hjelmsø, M., Rasmussen, M.A., Redgwell, T.A., Leal Rodríguez, C., Vestergaard, G., Zhang, Y., Chawes, B., Bønnelykke, K., Sørensen, S.J., Bisgaard, H., Enault, F., Stokholm, J., Moineau, S., Petit, M.-A., Nielsen, D.S., 2023. Expanding known viral diversity in the healthy infant gut. Nat Microbiol 8, 986–998. https://doi.org/10.1038/s41564-023-01345-7.

Shan, J., Korbsrisate, S., Withatanung, P., Adler, N.L., Clokie, M.R.J., Galyov, E.E., 2014. Temperature dependent bacteriophages of a tropical bacterial pathogen. Front. Microbiol. 5, 599. https://doi.org/10.3389/fmicb.2014.00599.

Shang, J., Jiang, J., Sun, Y., 2021. Bacteriophage classification for assembled contigs using graph convolutional network. Bioinformatics 37, i25–i33. https://doi.org/10.1093/bioinformatics/btab293.

Shang, J., Peng, C., Liao, H., Tang, X., Sun, Y., 2023. PhaBOX: a web server for identifying and characterizing phage contigs in metagenomic data. Bioinformatics Advances 3, vbad101. https://doi.org/10.1093/bioadv/vbad101.

Shkoporov, A.N., Hill, C., 2019. Bacteriophages of the human gut: the "known unknown" of the microbiome. Cell Host Microbe 25, 195–209. https://doi.org/10.1016/j.chom.2019.01.017.

Siddiqui, S.A., Erol, Z., Rugji, J., Taşçı, F., Kahraman, H.A., Toppi, V., Musa, L., Di Giacinto, G., Bahmid, N.A., Mehdizadeh, M., Castro-Muñoz, R., 2023. An overview of fermentation in the food industry - looking back from a new perspective. Bioresour. Bioprocess. 10, 85. https://doi.org/10.1186/s40643-023-00702-y.

Spus, M., Wardhana, Y.R., Wolkers-Rooijackers, J.C.M., Abee, T., Smid, E.J., 2023. Lytic bacteriophages affect the population dynamics of multi-strain microbial communities. Microbiome Res Rep 2. https://doi.org/10.20517/mrr.2023.20.

Tadrent, N., Dedeine, F., Hervé, V., 2022. SnakeMAGs: a simple, efficient, flexible and scalable workflow to reconstruct prokaryotic genomes from metagenomes. F1000Res 11, 1522. https://doi.org/10.12688/f1000research.128091.2.

Tamang, J.P., Kailasapathy, K. (Eds.), 2010. Fermented Foods and Beverages of the World. CRC Press/Taylor & Francis, Boca Raton. https://doi.org/10.1111/j.1471-0307.2011.00717.x.

Tamang, J.P., Shin, D.-H., Jung, S.-J., Chae, S.-W., 2016. Functional properties of microorganisms in fermented foods. Front. Microbiol. 7. https://doi.org/10.3389/fmicb.2016.00578.

Terzian, P., Olo Ndela, E., Galiez, C., Lossouarn, J., Pérez Bucio, R.E., Mom, R., Toussaint, A., Petit, M.-A., Enault, F., 2021. PHROG: families of prokaryotic virus proteins clustered using remote homology. NAR Genom Bioinform 3, lqab067. https://doi.org/10.1093/nargab/lqab067.

Thierry, A., Baty, C., Marché, L., Chuat, V., Picard, O., Lortal, S., Valence, F., 2023a. Lactofermentation of vegetables: an ancient method of preservation matching new trends. Trends Food Sci. Technol. 139, 104112. https://doi.org/10.1016/j.tifs.2023.07.009.

Thierry, A., Madec, M.-N., Chuat, V., Bage, A.-S., Picard, O., Grondin, C., Rué, O., Mariadassou, M., Marché, L., Valence, F., 2023b. Microbial communities of a variety of 75 homemade fermented vegetables. Front. Microbiol. 14. https://doi.org/10.3389/fmicb.2023.1323424.

Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. J. Roy. Stat. Soc. B Stat. Methodol. 63, 411–423. https://doi.org/10.1111/1467-9868.00293.

Tohno, M., Kitahara, M., Irisawa, T., Masuda, T., Uegaki, R., Ohkuma, M., Tajima, K., 2013. Lactobacillus silagei sp. nov., isolated from orchardgrass silage. Int. J. Syst. Evol. Microbiol. 63, 4613–4618. https://doi.org/10.1099/ijs.0.053124-0.

Tsai, A.-Y., Lin, Y.-T., Gong, G.-C., 2021. Effect of the presence of virus-like particles on bacterial growth in sunlit surface and dark deep ocean environments in the southern east China sea. Water 13, 2934. https://doi.org/10.3390/w13202934.

Van Sinderen, D., Karsens, H., Kok, J., Terpstra, P., Ruiters, M.H.J., Venema, G., Nauta, A., 1996. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Mol. Microbiol. 19, 1343–1355. https://doi.org/10.1111/j.1365-2958.1996.tb02478.x.

Varzakas, T., Zakynthinos, G., Proestos, C., Radwanska, M., 2017. Fermented vegetables. In: Yildiz, F., Wiley, R.C. (Eds.), Minimally Processed Refrigerated Fruits and Vegetables, Food Engineering Series. Springer, US, Boston, MA, pp. 537–584. https://doi.org/10.1007/978-1-4939-7018-6_15.

Vasimuddin, Md, Misra, S., Li, H., Aluru, S., 2019. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In: 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). Presented at the 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 314–324. https://doi.org/10.1109/IPDPS.2019.00041.

Vicente, J., Calderón, F., Santos, A., Marquina, D., Benito, S., 2021. High potential of Pichia kluyveri and other Pichia species in wine technology. Int. J. Mol. Sci. 22, 1196. https://doi.org/10.3390/ijms22031196.

Walker, P.J., Siddell, S.G., Lefkowitz, E.J., Mushegian, A.R., Dempsey, D.M., Dutilh, B.E., Harrach, B., Harrison, R.L., Hendrickson, R.C., Junglen, S., Knowles, N.J., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Nibert, M., Rubino, L., Sabanadzovic, S., Simmonds, P., Varsani, A., Zerbini, F.M., Davison, A.J., 2019. Changes to virus taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2019). Arch. Virol. 164, 2417–2429. https://doi.org/10.1007/s00705-019-04306-w.

Webster, C.L., Waldron, F.M., Robertson, S., Crowson, D., Ferrari, G., Quintana, J.F., Brouqui, J.-M., Bayne, E.H., Longdon, B., Buck, A.H., Lazzaro, B.P., Akorli, J., Haddrill, P.R., Obbard, D.J., 2015. The discovery, distribution, and evolution of viruses associated with Drosophila melanogaster. PLoS Biol. 13, e1002210. https://doi.org/10.1371/journal.pbio.1002210.

White, K., Yu, J.-H., Eraclio, G., Bello, F., Nauta, A., Mahony, J., Van Sinderen, D., 2022. Bacteriophage-host interactions as a platform to establish the role of phages in modulating the microbial composition of fermented foods. Microbiome Research Reports 1. https://doi.org/10.20517/mrr.2021.04.

Wickham, H., 2016. ggplot2. Use R! Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-24277-4.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4, 1686. https://doi.org/10.21105/joss.01686.

Williamson, K.E., Fuhrmann, J.J., Wommack, K.E., Radosevich, M., 2017. Viruses in soil ecosystems: an unknown quantity within an unexplored territory. Annual Review of Virology 4, 201–219. https://doi.org/10.1146/annurev-virology-101416-041639.

Wu, Q., Li, L., Xiang, P., Zhang, T., Peng, L., Zou, L., Li, Q., 2023. Phages in fermented foods: interactions and applications. Fermentation 9, 201. https://doi.org/10.3390/fermentation9030201.

Wuyts, S., Van Beeck, W., Oerlemans, E.F.M., Wittouck, S., Claes, I.J.J., De Boeck, I., Weckx, S., Lievens, B., De Vuyst, L., Lebeer, S., 2018. Carrot juice fermentations as man-made microbial ecosystems dominated by lactic acid bacteria. Appl. Environ. Microbiol. 84, e00134. https://doi.org/10.1128/AEM.00134-18, 18.

Yeoh, Y.K., 2021. Removing host-derived DNA sequences from microbial metagenomes via mapping to reference genomes. In: Carvalhais, L.C., Dennis, P.G. (Eds.), The Plant Microbiome: Methods and Protocols, Methods in Molecular Biology. Springer US, New York, NY, pp. 147–153. https://doi.org/10.1007/978-1-0716-1040-4_13.

Yoon, S.-S., Barrangou-Poueys, R., Breidt, F., Fleming, H.P., 2007. Detection and characterization of a lytic Pediococcus bacteriophage from the fermenting cucumber brine. J. Microbiol. Biotechnol. 17, 262–270.

Yoon, S.S., Barrangou-Poueys, R., Breidt, F., Klaenhammer, T.R., Fleming, H.P., 2001. Isolation and characterization of bacteriophages from fermenting sauerkraut. Appl. Environ. Microbiol. 68, 973–976. https://doi.org/10.1128/AEM.68.2.973-976.2002.

Zabat, M., Sano, W., Wurster, J., Cabral, D., Belenky, P., 2018. Microbial community analysis of sauerkraut fermentation reveals a stable and rapidly established community. Foods 7, 77. https://doi.org/10.3390/foods7050077.

Zaburlin, D., Quiberoni, A., Mercanti, D., 2017. Changes in environmental conditions modify infection kinetics of dairy phages. Food Environ Virol 9, 270–276. https://doi.org/10.1007/s12560-017-9296-2.

Zhang, H., Wang, L., Tan, Y., Wang, H., Yang, F., Chen, L., Hao, F., Lv, X., Du, H., Xu, Y., 2021. Effect of *Pichia* on shaping the fermentation microbial community of sauce-flavor *Baijiu*. Int. J. Food Microbiol. 336, 108898. https://doi.org/10.1016/j.ijfoodmicro.2020.108898.

Zhang, Huadong, Zhang, Hongxia, Du, H., Yu, X., Xu, Y., 2024. The insights into the phage communities of fermented foods in the age of viral metagenomics. Crit. Rev. Food Sci. Nutr. 1–13. https://doi.org/10.1080/10408398.2023.2299323, 0.

Zheng, J., Wittouck, S., Salvetti, E., Franz, C.M.A.P., Harris, H.M.B., Mattarelli, P., O'Toole, P.W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M.G., Lebeer, S., 2020. A taxonomic note on the genus Lactobacillus: description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. Int. J. Syst. Evol. Microbiol. 70, 2782–2858. https://doi.org/10.1099/ijsem.0.004107.