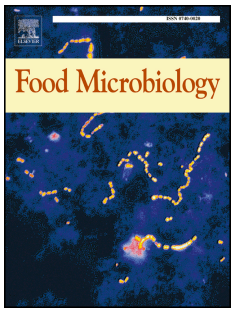
Viral metagenomic analysis of the cheese surface: A comparative study of rapid procedures for extracting viral particles

Eric Dugat-Bony, Julien Lossouarn, Marianne de Paepe, Anne-Sophie Sarthou, Yasmina Fedala, Marie-Agnès Petit, Stephane Chaillou

Accepted Manuscript



Viral metagenomic analysis of the cheese surface: A comparative study of rapid procedures for extracting viral particles

Eric Dugat-Bony, Julien Lossouarn, Marianne De Paepe, Anne-Sophie Sarthou, Yasmina Fedala, Marie-Agnès Petit, Stéphane Chaillou

PII: S0740-0020(19)30046-2

DOI: https://doi.org/10.1016/j.fm.2019.103278

Article Number: 103278

Reference: YFMIC 103278

To appear in: *Food Microbiology*

Received Date: 11 January 2019

Revised Date: 18 July 2019

Accepted Date: 23 July 2019

Please cite this article as: Dugat-Bony, E., Lossouarn, J., De Paepe, M., Sarthou, A.-S., Fedala, Y., Petit, Marie.-Agnè., Chaillou, Sté., Viral metagenomic analysis of the cheese surface: A comparative study of rapid procedures for extracting viral particles, *Food Microbiology* (2019), doi: https://doi.org/10.1016/j.fm.2019.103278.

Viral metagenomic analysis of the cheese surface: a comparative study of rapid procedures for extracting viral particles

Eric Dugat-Bony, [a,*] Julien Lossouarn,[b] Marianne De Paepe,[b] Anne-Sophie Sarthou,[a] Yasmina Fedala,[c] Marie-Agnès Petit,[b] and Stéphane Chaillou[b]

[a]UMR782 GMPA, INRA, AgroParisTech, Université Paris-Saclay, Avenue Lucien Brétignières, 78850, Thiverval-Grignon, France.

[b]UMR1319 MICALIS, INRA, AgroParisTech, Université Paris-Saclay, Domaine de Vilvert, 78350, Jouy-en-Josas, France.

[c]Institut Langevin, ESPCI Paris, CNRS, PSL University, 1 rue Jussieu, 75005, Paris, France

*Corresponding author: Eric Dugat-Bony, UMR782 GMPA, INRA, AgroParisTech, Université Paris-Saclay, Avenue Lucien Brétignières, 78850, Thiverval-Grignon, France, eric.dugat-bony@inra.fr

Keywords

Cheese rind; viral metagenomic; viral particles extraction procedure

Abstract

The structure and functioning of microbial communities from fermented foods, including cheese, have been extensively studied during the past decade. However, there is still a lack of information about both the occurrence and the role of viruses in modulating the function of this type of spatially structured and solid ecosystems. Viral metagenomics was recently applied to a wide variety of environmental samples and standardized procedures for recovering viral particles from different type of materials has emerged. In this study, we adapted a procedure originally developed to extract viruses from fecal samples, in order to enable efficient virome analysis of cheese surface. We tested and validated the positive impact of both addition of a filtration step prior to virus concentration and substitution of purification by density gradient ultracentrifugation by a simple chloroform treatment to eliminate membrane vesicles. Viral DNA extracted from the several procedures, as well as a vesicle sample, were sequenced using Illumina paired-end MiSeq technology and the subsequent clusters assembled from the virome were analyzed to assess those belonging to putative phages, plasmid-derived DNA, or even from bacterial chromosomal DNA. The best procedure was then chosen, and used to describe the first cheese surface virome, using Epoisses cheese as example. This study provides the basis of future investigations regarding the ecological importance of viruses in cheese microbial ecosystems.

1. Introduction

The cheese surface hosts dense and diverse microbial communities composed of bacteria, yeasts and filamentous fungi. This microbiota, whose structure and function is influenced by biotic interactions and abiotic factors, contributes to the quality of the products (Irlinger and Mounier, 2009). Composition of these communities has been studied for decades (see (Irlinger et al., 2015) and (Montel et al., 2014) for reviews). With the help of high throughput sequencing techniques, we now have detailed pictures of the communities present in a large panel of cheese varieties, and from all over the world (Dugat-Bony et al., 2016; Murugesan et al., 2018; Quigley et al., 2012; Wolfe et al., 2014). However, like many other microbial ecosystems, there is still a lack of knowledge on whether and how viral diversity controls the structure of cheese surface microbiota.

Viruses infect all forms of life (Hyman et al., 2012), from prokaryotes (Clokie et al., 2011) to eukaryotes (Koonin et al., 2015) and, in some particular cases, viruses themselves (La Scola et al., 2008). In microbial ecosystems, viral predation is known to greatly influence the structure and functioning of microbial communities (de Melo et al., 2018; Fernández et al., 2018; Sime-Ngando, 2014). Nevertheless, since virus genomes lack a single marker sequence for phylogenetic analysis, the structure and role of viral communities in nature are rarely evaluated. Recently, viral shotgun metagenomics helped to describe viral communities from environmental samples including ocean (Angly et al., 2006), freshwater (Roux et al., 2012), soil (Zablocki et al., 2014), mammalian gut (Minot et al., 2011) and, to a minor extent, fermented food products (Park et al., 2011). Regarding the dairy environment, a recent viral metagenomic study was conducted to describe the virome present in whey - the watery part of milk that remains after coagulation - from dairies using different starter cultures, revealing the dominance of *Lactococcus lactis* phages of the 936, P335 and c2 species (Muhammed et al.,

2017). Nevertheless, there is currently no literature on the viral diversity present on the surface of mature cheese.

Protocols for extraction of viruses have been optimized for different environments (Castro-Mejía et al., 2015; Conceição-Neto et al., 2015; Kleiner et al., 2015; Thurber et al., 2009). For food samples, most available procedures have been designed for the recovery of foodborne viruses potentially affecting human health such as noroviruses, rotaviruses (RoV) and hepatitis viruses (Stals et al., 2012), but not for microbial viruses such as bacteriophages. Furthermore, because the cheese surface has peculiar characteristics such as the presence of caseins at high concentration, high fat and salt content, it proves necessary to set up a dedicated procedure. Ideally, the method should be easy-to-use and rapid enough to be compatible with medium (dozens to hundred samples) to large-scale studies (hundreds to thousands samples) such as those performed to describe microbial communities (Wolfe et al., 2014).

Taking into account all these constraints, we compared four procedures for the isolation of viruses from cheese surfaces. The backbone of the protocol followed the PEG-based protocol already evaluated by Castro-Mejia et al. on fecal samples (Castro-Mejía et al., 2015), which was also used for dairy bacteriophages recovery from whey after some adaptation (Muhammed et al., 2017). The authors evaluated phage recovery at the different steps of the protocol by spiking with known phages, and concluded that virtually no spiked phages were lost until the ultracentrifugation step (density gradient). In our experimental design, presented in Fig. 1, protocol P1 was constructed from the first steps of this protocol except the replacement of the buffer used for sample suspension, and was thus considered as a control maximizing phage recovery. In parallel, we tested both the effect of adding a filtration step before PEG precipitation in order to deplete microbial cells (protocol P3) and/or the substitution of the expensive and time-consuming density gradients for viral fraction's

purification by a simple chloroform treatment (protocols P2 and P4) which is used for membrane vesicles removal (Biller et al., 2017; Forterre et al., 2013). We evaluated the usefulness of the four procedures for cheese virome analysis (limited to the viruses with DNA genomes, expectedly the most abundant) using three different types of cheese, namely Camembert, Epoisses and Saint-Nectaire, and following two criteria: particles recovery and particles purity. Finally, we selected Epoisses cheese for producing the first cheese surface virome and assessing the level of microbial DNA contamination in virome sequencing data.

2. Material and methods

2.1. Sampling procedure

Three types of French surface-ripened cheese were studied, namely Camembert (CAM, bloomy rind), Epoisses (EP, washed rind) and Saint-Nectaire (SN, natural rind). For each type of cheese, three cheeses produced at the same date and from a unique producer were purchased and analyzed as replicates. Rind was gently separated from the core using sterile knives, and mixed using a blender.

2.2. Microbiological analysis

One gram of the cheese surface was diluted 1:10 in sterile saline solution (9 g/l NaCl) and homogenized with an Ultra Turrax Homogenizer (Labortechnik, Staufen,Germany) at full speed for 1 min. Serial dilutions were performed in 9 g/l NaCl and microorganisms were enumerated by surface plating in duplicate on specific agar base medium. Cheese-surface bacteria were enumerated on brain heart infusion agar (Biokar Diagnostics) supplemented with 22.5 mg/l amphotericin B after 3 to 5 days of incubation at 25°C under aerobic conditions. Lactic acid bacteria were enumerated on de Man-Rogosa-Sharpeagar (pH 6.5, Biokar Diagnostics) supplemented with 22.5 mg/l amphotericin B after 3 days of incubation at 30°C under anaerobic conditions. Finally, fungal populations were enumerated on yeast

extract-glucose-chloramphenicol (Biokar Diagnostics) supplemented with 2,3,5-triphenyltetrazolium chloride (10 mg/l) after 3 to 5 days of incubation at 25°C under aerobic conditions.

2.3. Extraction of the viral fraction from cheese

Four protocols, named P1, P2, P3 and P4, were tested in parallel starting from the same material (Fig. 1). Six grams of cheese rind was diluted 1:10 with cold trisodium citrate (2% w/v) into a sterile bag and mixed for 1 min using a BagMixer (Interscience). Citrate is a complexing agent for calcium and allows casein solubilization. It is largely used in nucleic acid extraction protocols for recovering microbial cells from casein network in milk and cheese (Randazzo et al., 2002; Ulve et al., 2008). The solution was centrifuged at 300 × g for 10 min at 4°C in order to pellet big aggregates. The supernatant was centrifuged at 5,000 × g for 45 min at 4°C in order to pellet microbial cells. At this stage, the supernatant containing free viral particles was diluted 1:5 with cold SM buffer (200 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris pH 7.5) and split, half being used for filtration-free protocols (protocols P1 and P2) and half being successively filtrated using 0.45µm and 0.2µm polyethersulfone membranes and glass vacuum filter holders (Millipore) (protocols P3 and P4). Samples, filtrated or not, were then supplemented with 10% PEG 6,000 (Sigma) and kept at 4°C over-night after dissolution for viral particles precipitation. After centrifugation at 6,000 × g for 1h at 4°C, pellets were resuspended with 2 ml of cold SM buffer and split again, half being stocked at 4°C (protocols P1 and P3) and half being treated with chloroform in order to eliminate membrane vesicles (protocols P2 and P4). More specifically, 1 volume of fresh, non-oxidized chloroform was added to the sample and mixed thoroughly for 1 min using a vortex to create an emulsion. After centrifugation at 15,000 × g for 5 min at 4°C, the aqueous phase containing viral particles was recovered.

2.4. Particles counting using Interferometry

An interferometric light microscope (ILM) (Boccara et al., 2016) was used to count nanoparticles (i.e. both vesicles and viruses) present in our viral fractions as previously described (Roose-Amsaleg et al., 2017). Briefly, 5 µl of each sample were used to collect a stack of 200 images (CMOS camera). We developed a simple ImageJ (Schindelin et al., 2012) script that allows first background subtraction and image quality enhancement, then particles localization on each image. An average number of nanoparticles per frame was obtained. To calibrate the concentration estimates, crude lysates of phages P1, T4, T5, T7, λCI857 and φX174 were titrated on *Escherichia coli* MG1655, and counted with the ILM device. These counts were converted into concentrations, assuming that all particles present in the observation volume (10^{-8} ml) were detected with the ILM device. Table S1 and Fig. S1 showed a good match between the two measurements, usually within a +/- 3 fold range for phages P1, T4, T5, T7 and λCI857, whereas φX174 was underestimated by 20 fold. We next compared the ILM estimates with epifluorescence microscopy on two viral samples obtained from Epoisses cheese (Fig. S2). As shown in the result section, virome samples of Epoisses rind also contain vesicles, which are detected by ILM, but not by epifluorescence, since they mostly do not contain DNA. In accordance with expectations, 1.5 to 1.9-fold more particles were counted by ILM than by epifluorescence. We concluded that the ILM measurements give reasonable estimates of nanoparticle concentrations (both vesicles and viruses) of viromes. Statistical comparison between ILM measurements (log transformed data) from two extraction procedures was performed using a one-tailed t-test for two dependent means with a significant threshold of 0.05 after testing the data normality using a Shapiro-Wilk test (significant threshold of 0.05). For comparison between two cheese types, a one-tailed t-test for two independent means was used with a significant threshold of 0.05.

2.5. Particles purity evaluation using iodixanol gradients

Two-layer iodixanol discontinuous gradients of 45% and 20% (w/v) in SM Buffer (50 mM Tris pH7.5, 100 mM NaCl, 10 mM MgCl2) were prepared by adding first 6 ml of the 20% fraction in a 12 ml ultracentifugation tube, and then under layering 4.5 ml of the 45% fraction in the bottom of the tube, using a glass transfer pipette and a pipette pump. Each viral fraction (1 ml diluted 1:2 with SM Buffer) was finally added on top of the gradient and the tubes were centrifuged at 200,000 × g for 5 h at 10°C, in an SW41 rotor, using a Beckman XL-90 ultracentrifuge.

2.6. Transmission electron microscopy

Ten microliters of the viral fraction obtained from an Epoisses cheese with protocol P4 was directly spotted onto a Formwar carbon coated copper grid. Particles were allowed to adsorb to the carbon layer for 5 min and excess of liquid was removed. Ten microliters of a staining uranyl acetate solution (1%) was then spotted to the grid for 10 seconds and excess of liquid was removed again. The grid was imaged at 80 kV in a Hitachi HT7700 transmission electron microscope.

2.7. Viral DNA extraction

The twelve Epoisses samples (three cheeses treated by protocols P1, P2, P3 and P4) were used for viral DNA extraction and shotgun sequencing. For one Epoisses sample (Epb) treated with P3, the vesicle band was also recovered from iodixanol gradient for subsequent DNA extraction and sequencing. 500 µl of the viral fraction (or vesicles) were treated for 30 min at 37°C with 1 U of TURBO DNase (Invitrogen) in order 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. The entire content of the tube was transferred to a 2 ml tube containing a gel that improves separation between the aqueous and organic phases (Phase Lock Gel Heavy; Eppendorf AG, Hamburg, Germany). One volume of phenol-chloroform (1:1; saturated with 10 mM Tris pH 8.0, and 1 mM EDTA) was added to the sample and, after gentle mixing for 1

min by inversion, the tube was centrifuged at 10,000 × g for 10 min at 4°C.The aqueous phase was then transferred to another Phase Lock Gel tube. After adding one volume of phenol-chloroform and gentle mixing for 1 min by inversion, centrifugation was performed at 10,000 × g for 10 min at 4°C. The aqueous phase was then transferred to another Phase Lock Gel tube. After adding one volume of chloroform and gentle mixing for 1 min by inversion, centrifugation was performed at 10,000 × g for 10 min at 4°C. The aqueous phase (approximately 300 µl) was recovered in a 1.5 ml tube and DNA was precipitated by adding two volumes of absolute ethanol, 50 µl of sodium acetate (3 M, pH 5.2) and 3 µl of glycogen (5 mg/ml; Invitrogen) as carrier. After incubation on ice for 10 min, the DNA was recovered by centrifugation at 13,000 × g for 30 min at 4°C. The DNA pellet was subsequently washed with 800 µl of 70% (vol/vol) ethanol. After centrifugation at 13,000 × g for 10 min at 4°C, the DNA pellet was dried at room temperature for 30 min and dissolved in 20 µl of 10 mM Tris pH 7.5. Purified DNA was quantified using the Qubit DNA Broad Range assay (ThermoFisher Scientific).

2.8. Viral DNA sequencing and bioinformatic analysis

Library preparation and sequencing were performed at the GeT-PlaGe platform (Toulouse, France). Briefly, libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit (New England Biolabs) following manufacturer's instructions and using up to 200 ng dsDNA (QuBit quantified, Thermofisher) for each sample, when possible. DNA was fragmented using covaris M220 (Covaris) leading to libraries with an average insert size of 575 bp. After adapter ligation, 9 PCR cycles were applied to add Illumina adaptators and indexes and to amplify the library. Libraries were then individually quantified by qPCR using the KAPA Library Quantification Kits for NGS (KAPA Biosystems), equimolary pooled and sequenced on the Illumina MiSeq system (MiSeq reagent kit v3, Illumina). Raw reads were quality filtered using Trimmomatic v0.36 (Bolger et al., 2014) in order to remove paired reads

with bad qualities (with length < 125 bp, with remnant of Illumina adapters and with average quality values below Q20 on a sliding window of 4 nt). The overall filtering process yielded an average of 2.72 million reads per sample (35.5 million reads in total). The strategy for assembly and mapping is presented in Fig. S3, as well as the number of reads kept at each step. For assembly, only the quality-filtered and paired reads (2 × 1.13 million reads per sample in average, 83% of the total) were kept while those being unpaired at the end of the filtering process (0.46 million reads per sample in average, 17% of the total) were kept aside and used only for mapping. A global assembly was performed using the quality-filtered and paired reads from the pooled thirteen sequenced samples. Therefore, reads from identical phages and present in several samples were binned together during the assembly process. Assembly was performed using SPAdes v3.13.0 with the *meta* option and increasing *kmer* values -k 21,33,55,77,99,127 (Bankevich et al., 2012). Contigs with length below 2,500 bp, unlikely to encode complete viruses, were discarded. The resulting 910 filtered contigs were first analyzed with VirSorter v1.0.3 (Roux et al., 2015) using the Viromes database (all bacterial and archaeal virus genomes in Refseq (as of January 2014), together with non-redundant predicted genes from viral metagenomes), which returned 92 putative viral contigs (phages or prophages). A list of “most abundant and pertinent” contigs was constructed using three criteria: 1) detected as viral with VirSorter, after filtering out those placed in categories 3 or 6 (not so sure) with coverage below 10 (remaining 78 contigs). Coverage was used as a proxy for abundance instead of percentage of reads because this criterion is independent of the contig's size. Contigs with a coverage below 10 corresponded approximately to contigs present in the first quartile when ordered by coverage, so the 25% less abundant contigs. 2) Detected as circular contig, even if not detected as viral (adds 34 contigs). 3) Coverage above 100, even if not detected as viral or circular (final list of 124 contigs). Contigs with a coverage above 100 corresponded approximately to contigs present in the third quartile when

ordered by coverage, so the 25% more abundant contigs. The 124 final contigs were further analyzed using PHASTER (Arndt et al., 2016) and compared to both the Viruses section of nucleotide database at NCBI and to the complete database using Blast with default parameters (Altschul et al., 1990) in order to provide a first characterization of putative phage-encoding contigs. Then, quality filtered reads (both paired and unpaired) were mapped individually for each sample against the 124 potential viral contigs using Bowtie2 aligner v2.2.6 with default parameters values (Langmead and Salzberg, 2012) in order to produce an abundance table. Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarity, composition analysis and heatmap were processed with the R package Phyloseq v1.26.1 (McMurdie and Holmes, 2013). In order to measure the effect of the extraction procedure on the composition of the viral metagenome, Spearman correlations were calculated for each pair of protocols (P1 vs P2, P1 vs P3, P1 vs P4, P2 vs P3, P2 vs P4 and P3 vs P4) on the mean log10 abundance values and visualized with a scatter plot using the ggpubr package v0.2 (Kassambara, 2018). The level of bacterial DNA contamination in the cheese viromes was estimated by detection of ribosomal DNAs among the reads using SortMeRNa v2.0 (Kopylova et al., 2012) and the SILVA v129 database (Quast et al., 2013) with default parameters (search in both strands, e-value threshold = 1, similarity threshold = 0.97, query coverage = 0.97) . The three most prevalent contigs present in the final virome, namely Epvir4, 949_Epvir1 and Epvir8, were further annotated using RAST (Aziz et al., 2008) using genetic code 11 and "virus" options. Homologs of the predicted proteins in the contigs were additionally compared to the NCBI *nr* database with BlastP (E-value cutoff of 10^{-8}) (Altschul et al., 1990), and to the PDB database using HHpred (Zimmermann et al., 2018) with a probability cutoff of 99%.

2.9. Accession numbers

Raw sequence data were deposited at the Sequence Read Archive of the National Center for Biotechnology Information under the accession numbers SRR8080803 to SRR8080815 (bioproject PRJNA497596).

3. Results

3.1. Quantification and purity of viral particles recovered from the cheese surface.

Among the four tested procedures, only those containing chloroform treatments (P2 and P4) resulted in viral fractions sufficiently pure to allow nanoparticle counting using the interferometric light microscope (Table 1). We observed 1×10^9 to 4×10^{10} nanoparticles per gram of cheese surface. The counts observed in samples treated with P4 were significantly lower than those for samples treated with P2 (one-tailed t-test for two dependent means, $p < 0.05$) indicating a loss of particles due to the filtration step (Table 1). Counts were significantly higher in Camembert than in Saint-Nectaire cheese (one-tailed t-test for two independent means, $p < 0.05$). For Epoisses cheese, counts were slightly higher than in Saint-Nectaire and slightly lower than in Camembert but those differences were not statistically significant (one-tailed t-test for two independent means, $p > 0.05$). Camembert had the highest nanoparticles per microbial cell ratios. This type of cheese had the lowest bacterial counts (Table S2), whether such a community leads to higher cell lysis mediated by viruses or higher levels of membrane vesicles remains to be investigated. Viral fractions prepared using procedures P1 and P3, without chloroform treatment, were very dense, milk-white and led to hard noise in interferometer's films with a myriad of spots larger than those typically observed for viruses. Microbial cells clearly contaminated the viral fractions when using procedures P1 and P2, without filtration (Table 1, column 3). It has to be mentioned we also observed bacterial cells in one filtrated sample, *i.e.* Camembert cheese with P3, indicating probable post-filtration contamination.

Two-layer iodixanol gradients were used as quality controls in order to separate microbial cells, debris and membrane vesicles from viruses, enabling to visually observe the effect of the four procedures on the quantity and purity of viral fractions (Fig. 2). For the three types of cheese tested, similar profiles were observed after ultracentrifugation. A strong band located at the top of the lightest density layer was present in samples prepared without filtration and chloroform treatment (P1) suggesting high contamination with microbial cells, debris and membrane vesicles (Fig. 2, red arrow). This band was still present – albeit slightly less intense – in filtered samples (P3), suggesting that cheese rind is rich in membrane vesicles. Finally, this band was completely absent from samples prepared with procedures including chloroform treatment (P2 and P4) confirming the efficiency of such treatment for removing membrane vesicles from viral fractions (Biller et al., 2017). A visible band containing viruses formed at the proximity of the 45% iodixanol cushion (Fig. 2, blue arrow), at a density corresponding to 40% iodixanol, as estimated with a refractometer. The virus band was barely visible, however, in samples treated with chloroform (P2 and P4) when compared to untreated samples (P1 and P3). Nanoparticles quantification using interferometry in the viral band of Epoisses samples after dialysis indicated that samples treated with cholorofom (P2 and P4) contained approximately ten times less nanoparticles than untreated samples (P1 and P3). For Camembert cheese and procedure P3, we observed two distinct bands near the 45% iodixanol cushion. Transmission electronic microscopy revealed that viruses were more abundant in the lowest one.

Altogether, these results suggest that filtration is an important step for the extraction of viruses from cheese in order to deplete microbial cells, which represent the major source of DNA contamination in virome studies. Furthermore, chloroform treatment is very beneficial for viral fractions' cleaning and removal of membrane vesicles, which might contain pieces of microbial DNA (see below), but at the cost of reducing the nanoparticle recovery.

3.2. Effect of the extraction procedure on the composition of the Epoisses virome

We selected Epoisses cheese for further virome analysis because it had both high counts of nanoparticles and bacteria (Table 1 and Table S2), bacteria being expected as the main targets of microbial viruses (*e.g.* bacteriophages or phages) in the cheese ecosystem. DNA was successfully extracted from viral fractions produced from Epoisses cheese using the four procedures and prior the gradient step. DNA yields are available in Table S3. DNA samples were then sequenced, in order to assess the impact of both filtration and chloroform treatment on the final sequence dataset (see Material and Methods section). In addition, for sample P3, the vesicle band was recovered from iodixanol gradient, DNA was extracted using the same protocol as viruses, and sequenced. The virome sequence data from each of the 12 samples (4 protocols x 3 biological replicates) and that of the vesicles sample were binned together for assembly of the virome (see Material and Methods section and Fig. S3 for more details). Quality filtered reads were pooled and assembled into 910 contigs of length greater than 2,500 bp, a threshold used for keeping putative small virus genomes (Varsani and Krupovic, 2018). Of these, 124 contigs were selected for further analysis based on VirSorter results and coverage criteria (see Material and Methods section). These 124 contigs were ranging in size from 2.5 kb to 122 kb with a N50 of 24.4 kb and were designated as the Epoisses cheese virome. Mapping of the quality filtered reads against this virome revealed that microbial cell contamination was higher with protocol P1 and in the vesicles sample than in any of the other three protocols (Fig. 3 and Fig S3).

The contigs were further analyzed and separated into four classes based on PHASTER and Blast results (*i.e.* probable phage contigs, plasmid-derived contigs, putative plasmid-derived contigs and unclassified contigs) (Table S4). Principal coordinate analysis performed on the Bray-Curtis dissimilarity index calculated from the abundance table of the 124 contigs (Fig. 4A) showed an ordination of the samples according to the extraction procedure, with samples

obtained with P1 and P3 being separated from samples obtained with P2 and P4. This indicated an impact of the chloroform treatment on the quality of the resulting virome. Datasets produced with protocols P2, P3 and P4 were strikingly enriched for reads mapping to probable phage contigs (>99% in average) when compared to viral fraction obtained with P1 (94% in average) (Fig. 4B, Fig. S3), which was largely explained by the higher proportion of unmapped reads present in P1 samples (Fig. 4F). The relative proportion of reads mapping to the 19 plasmid-derived contigs was greatly reduced in samples treated with P3 (filtration) and even more so in samples treated with P2 (chloroform) and P4 (chloroform and filtration), when compared to P1 (Fig. 4C), reflecting the highly positive impact of chloroform treatment and, to a lesser extent of filtration, on the quality of the final virome. The relative proportions of reads matching putative plasmid-derived contigs (Fig. 4D) and unclassified contigs (Fig. 4E) remained similar regardless of the extraction procedure.

Heatmap representation was used to visualize the detailed composition of each virome (Fig.5). Spearman correlations confirmed that the abundance profile of phage contigs obtained with the different extraction procedures was almost identical (Rho values >0.99) suggesting that neither the filtration step nor the chloroform treatment biased the sample composition (Fig. S4A). When considering only the profile of non-phage contigs, Rho values dropped (e.g. 0.59, 0.58, 0.65 for the comparisons P1 vs P2, P1 vs P3 and P1 vs P4, respectively) (Fig. S4B), indicating that filtration and chloroform treatment had more impact on the detection of these contigs. In fact, the major differences were observed for the 39 contigs belonging to the categories plasmid and putative plasmid (Fig. 5). Most of them were undetected or sporadically detected in samples treated with chloroform (P2, P4). They were on the contrary much higher with protocols P1 (and to a lesser extend P3) and highly abundant in the vesicles sample (39% of the reads). Interestingly, NODE-134, the most abundant contig detected in the vesicles sample (24 % relative abundance) and assigned to the

plasmid-derived contig category, was only marginally detected in samples treated with protocols P1, P2, P3 and P4 (<0.005% relative abundance).

Much less DNA was present in membrane vesicles, compared to total viromes (Table S3). This has to be kept in mind when comparing the vesicle line to the virome lines in Fig. 5, since colors reflect only relative abundances. Almost 51% of the reads from the vesicles sample mapped to contigs assigned to potential phages, which indicated that phage DNA was present in the vesicles. Interestingly, the reads corresponding to plasmid-derived contigs were specifically enriched in membrane vesicles (10% of total), relative to the viromes (0.03%). It seems therefore that the plasmid content of samples P1 and P3, observed in Fig. 5, originates from membrane vesicles, which are still present since no chloroform treatment was applied in these protocols. The total DNA content obtained from the vesicle fraction was low, relative to viromes, with only 3.8 ng (Table S3), and corresponded to an estimated amount of 1.3×10^{11} nanoparticles according to interferometry measures. Supposing all DNA content is plasmidic, with an average size of 5 kb for the plasmids (and a corresponding weight of 0.55×10^{-17} g), 3.8 ng of DNA would amount to 7×10^{8} plasmidic DNA molecules, suggesting that only 0.5% of the vesicles effectively contain a DNA molecule. The same calculation for phage DNA (~80 kb, corresponding to the average size of the two major contigs Epvir4 and 949_Epvir1 present in the Epoisses virome) would result in 16-fold less membrane vesicles containing viral DNA, compared to plasmidic DNA.

3.3. Composition of the Epoisses cheese virome

Three repeats of the extraction-sequencing on Epoisses cheese bought simultaneously from the same manufacturer gave highly reproducible results. Two contigs, NODE-4 and NODE-1, now named Epvir4 and 949_Epvir1, were prevalent, representing relative mean abundances of 63% and 28%, respectively. Epvir4 genome is 42 kb long. It was assigned by VirSorter as a category 2 phage (« quite sure ») but did not share similarity with already known phages. The

annotation by RAST, HHpred and BLAST revealed an organization into functional modules (Fig. S5A). Amongst the ORFs with predicted homologues in the NCBI *nr* database, most best BlastP matches (63%) were found in the genus *Glutamicibacter* (formerly *Arthrobacter*) suggesting that the host of this abundant phage might be a *Glutamicibacter* species (Table S5). 949_Epvir1 genome (Fig. S5B) is 122 kb long and shares 95% nucleotide identity and 95% coverage with *Lactococcus lactis* phage 949, a virulent phage (despite two predicted integrase genes) isolated from cheese whey in New Zealand (Samson and Moineau, 2010). The third most abundant contig (1.1% of the reads), Epvir8, has a genome 41 kb in length, and shares a small 1.1 kb region with 88% nucleotide identity with *Pseudoalteromonas* phage SL25 (Liu et al., 2018). Amongst the ORFs with predicted homologues, more than half (56%) were found in *Pseudoalteromonas* phages while another quarter (24%) were found in *Vibrio* phages (Table S5 and Fig. S5C). All other contigs exhibited a relative abundance lower than 1%, most of them being difficult to assign to a putative host. Nevertheless, few contigs were closely related to *Lactococcus lactis* phages (*e.g.* NODE-15, NODE-59, NODE-142), *Pseudoalteromonas* phages (*e.g.* NODE-48, NODE-276), *Vibrio* phages (*e.g.* NODE-12, NODE-116), *Leuconostoc* phages (*e.g.* NODE-18, NODE-658) and *Halomonas* phage (NODE-636). All of these bacterial genera are frequently detected as part of cheese surface (Dugat-Bony et al., 2016). The best host assignments carried out for all contigs are available in supplementary Table S4.

3.4. Bacteriophage morphologies

Transmission Electronic Microscopy (TEM) analysis was carried out on a sample prepared using the extraction procedure P4 and representative pictures are presented in Fig. 6. Frequencies of the different morphotypes were calculated based on the analysis of 25 pictures totaling 43 phage particles (Table S6). The most frequent morphotype (79%) exhibited an icosahedral capsid 60.7 (±4.4) nm in diameter, and a noncontractile tail 161.3 (±25.7) nm in

length, as illustrated in Fig. 6C, D and E. A larger bacteriophage presenting characteristics similar to *Lactococcus lactis* phage 949, *i.e.* an icosahedral capsid of 89 (±6) nm diameter and a noncontractile tail of length 482 (±36) nm, was observed with a frequency of 9% (Fig. 6F). Two other morphotypes were also detected, one with a capsid of 71 (±2) nm diameter and a noncontractile tail of length 173 (±6) nm (Fig. 6B) and the other with a capsid of 51 (±1) nm diameter and a tail of length 140 (±8) nm, at frequencies of 7 and 5%, respectively.

4. Discussion

With the increasing attraction for virome sequencing, protocols for extracting viruses from diverse environmental samples have been developed during the past decade. Prior to nucleic acid extraction and sequencing, these protocols usually included sample pretreatment to make viral particles accessible for extraction, virus concentration and finally virus purification (Thurber et al., 2009). Each step has to be adapted, according to the type of samples studied and the type of virus targeted. For cheese, we chose to blend the samples after a dilution step in trisodium citrate, in order to maximize the chance for recovering viral particles from the matrix. Indeed, citrate is a complexing agent for calcium and allows casein solubilization. It is largely used in nucleic acid extraction protocols for recovering microbial cells from casein network in milk and cheese (Randazzo et al., 2002; Ulve et al., 2008). In order to complete our control condition (protocol P1), we then followed the first steps of the PEG-based protocol proposed by Castro-Mejia et al. (2015), namely centrifugation and PEG precipitation, considering that it were already validated and provided very high recovery rates for all the tested spiked-phages.

We decided to evaluate the potential benefit of adding a filtration step (0.22 µm) on the quality of viral fractions and of the resulting virome. Filtration is preferably avoided in virus extraction protocols from environmental sources because some viruses can be very large, even

larger than microbial cells (Aherfi et al., 2016). Nevertheless, our choice was motivated by the fact that (i) such large viruses are generally hosted by organisms which are not part of the cheese ecosystem such as amoebas, protists and microalgae (ii) it ensures complete removal of microbial cells which generally account for the major source of contaminating DNA sequences in virome studies (Roux et al., 2013). Our results demonstrated that adding filtration to the extraction procedure reduced microbial contamination of the cheese viral fractions by microbial cells, without major modification neither in the particle counts nor in the final virome profile.

The best way to purify viruses includes a density gradient ultracentrifugation step (Thurber et al., 2009), in which viruses are separated from other components of the extract based on their physical properties. This technique is, however, very time-consuming, expensive and requires specific technical skills and lab equipments. Chloroform treatment represents a possible alternative for rapid and efficient virus purification (Biller et al., 2017; Guo et al., 2012) and has already been used in the viral metagenomic context (Willner et al., 2011). This solvent permeabilizes the membranes of bacterial cells and membrane vesicles, disrupting their structural integrity and making nucleic acids available for digestion by nucleases. Our results indicate that such treatment, in combination with filtration, is mandatory for successful quantification of viruses extracted from cheese samples by interferometry. Even if viral capsids are said to be resistant to chloroform (Biller et al., 2017; Jurczak-Kurek et al., 2016), it is important to mention that some viruses, in particular enveloped viruses, might be lost after such treatment (Biller et al., 2017; Forterre et al., 2013; Thurber et al., 2009; Weynberg et al., 2014). Comparing the Epoisses cheese viral fractions obtained with procedures including chloroform treatment or not, indicated a loss of particles upon chloroform treatment (as illustrated in Fig. 2). Analyzing the nanoparticle counts in the viral band of one Epoisses cheese after density gradient ultracentrifugation and dialysis indicated that we recovered

approximately ten times less nanoparticles in chloroform-treated samples versus non-treated samples. However, this was accompanied by a limited reduction of the final DNA yields (<3 times less DNA in average for treated versus untreated samples, Table S3) and more importantly, we did not observe qualitative differences in the abundances of viral contigs as a function of chloroform treatment (Fig. 5 and Fig. S4). Furthermore, adding chloroform treatment to the extraction procedure resulted in a significant enrichment of the virome in actual viral sequences while drastically reducing the proportion of plasmids and sequences from microbial origin (unmapped reads).

For one Epoisses, we sequenced the DNA content of a membrane vesicle fraction, separated from phage particles on an analytical density gradient. Only 0.5% of the total virome DNA yield (3.8 ng out of 548 ng) was recovered in this fraction (Table S3). Compared to total vesicles counts, we estimated that at most a few percent of the membrane vesicles contained DNA. Indeed, the transport of nucleic acids only constitutes one of many potential roles for membrane vesicles (Gill et al., 2018). Half of vesicle reads mapped to phage contigs, in line with the proposal that membrane vesicles serve as decoys and trap phages, as recently reported in *Vibrio cholerae* (Reyes-Robles et al., 2018). Thus, compared to total viral DNA, only 0.35% got trapped into such “decoys”. Another 10% of vesicles reads corresponded to predicted plasmids, a property already reported for *Escherichia coli* membrane vesicles (Yaron et al., 2000). The third category of DNA content (also around 10%) in this vesicle fraction was composed of circular contigs that might be plasmidic. Interestingly, their relative proportion was constant in the global viromes across the various protocols, irrespective of the chloroform treatment, suggesting they might not belong to vesicles.

The rapid procedure presented in this study enabled to obtain sufficient quantity of viral DNA from Epoisses cheese (equivalent of 1.5 gram as starting material for each protocol) for direct library preparation prior to sequencing, avoiding the use of multiple displacement

amplification (MDA) (DNA quantities obtained for all samples presented in this study are available in Table S3). MDA is widely used in viral metagenomic studies due to the insufficient DNA material extracted from the samples (Thurber et al., 2009). However, amplification bias has been documented for this technique (Pinard et al., 2006) which might provoke distortion in the viral community profiles and making preferable direct DNA sequencing when possible.

The Epoisses virome analyzed in this study was highly dominated by few contigs (only 3 with relative abundance >1%), contrasting with the viromes produced from whey mixture which generally contain a wide diversity of abundant and equally distributed contigs (Muhammed et al., 2017). Cheese is a peculiar environment where the high nutrient availability and strong abiotic factors (*e.g.* pH, salt, ripening temperature) select from a quite large reservoir of microbial diversity present in milk and the dairy environment, only few dominant populations (Irlinger et al., 2015). The resulting microbial community in cheese rind after the maturation step is thus highly dominated by few taxa (Dugat-Bony et al., 2016; Wolfe et al., 2014). The same rule may apply to viruses but additional experiments, such as time-series virome analysis during an entire cheese production cycle (from milk to mature cheese) would be necessary to confirm this hypothesis.

The Epoisses virome was composed of many contigs sharing high sequence identity with known *Lactococcus lactis* phages including the widely known 936, P335 and C2-like groups. Interestingly, the second most abundant one (28% of reads), named 949_Epvir1, was very similar to the genome of *L. lactis* phage 949, a virulent phage isolated from cheese whey in New Zealand, which is phylogenetically distant from those groups of commonly isolated *L. lactis* phages (Samson and Moineau, 2010). We also detected a phage morphotype whose characteristics, namely a head of 89 nm and a long noncontractile tail of 482 nm, are compatible with *L. lactis* phage 949. *Lactococcus lactis* is used as starter in the manufacture

of Epoisses cheese. In a previous work describing the microbial diversity in twelve french cheese varieties, it was detected as the dominant bacterial taxa in the core of Epoisses cheese and was also highly detected in the rind (Dugat-Bony et al., 2016). In the same study, other dominant bacterial genera of the Epoisses cheese rind included not-deliberately inoculated taxa such as, by order of importance, *Psychrobacter*, *Marinomonas*, *Vibrio*, *Pseudoalteromonas*, *Glutamicibacter*, *Mesonia*, *Enterococcus*, *Lactobacillus* and *Halomonas*. Interestingly, phages potentially infecting some of such non-starter bacterial taxa, including *Glutamicibacter* (Epvir4, most abundant contig in the Epoisses virome, 63% of the reads), *Pseudoalteromonas* (Epvir8, third most abundant contig, 1.1% of the reads), *Vibrio*, *Leuconostoc* and *Halomonas*, were also detected in the Epoisses cheese virome.

The cheese surface microbiota is in constant evolution during the ripening process, and is characterized by the successive development of different microbial groups. Microbial interactions between different species have been observed in cheese (Irlinger and Mounier, 2009; Kastman et al., 2016; Mounier et al., 2008; Zhang et al., 2018), providing the first elements of comprehension regarding biotic forces sustaining microbial assemblage in this peculiar environment. Our results indicate that many microbial species living on the cheese surface are also subjected to viral predation, and shed light on the need of careful evaluation of the impact of viruses on the dynamic of the cheese microbial ecosystem.

5. Conclusions

In this study, we adapted a viral extraction protocol originally developed for fecal samples to the cheese matrix and evaluated its usefulness for subsequent virome analysis. We demonstrated its efficiency by extracting viruses from three different type of cheese and produced the first cheese surface virome using Epoisses cheese as a model. Our results emphasize the positive impact of chloroform treatment and, to a lesser extent filtration, on the

final virome quality in the cheese context. In the future, we anticipate virome analysis will expand our knowledge on cheese microbial ecosystems and possibly give the opportunity to better understand the rules of microbial assembly that occur in this fermented food.

Funding

This work was conducted in the frame of the Virome Access project supported by the French National Institute for Agricultural Research (INRA) and the Meta-omics and Microbial Ecosystems (MEM) programme. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Acknowledgements

We are grateful to Martine and Claude Boccara for their invaluable help with the interferometer device. We thank the INRA GeT-PlaGE platform (https://get.genotoul.fr/la-plateforme/get-plage) for sequencing the Epoisses virome and the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources and support.

References

Aherfi, S., Colson, P., La Scola, B., Raoult, D., 2016. Giant viruses of amoebas: an update. Front Microbiol 7. https://doi.org/10.3389/fmicb.2016.00349

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Angly, F.E., Felts, B., Breitbart, M., Salamon, P., Edwards, R.A., Carlson, C., Chan, A.M., Haynes, M., Kelley, S., Liu, H., Mahaffy, J.M., Mueller, J.E., Nulton, J., Olson, R., Parsons, R., Rayhawk, S., Suttle, C.A., Rohwer, F., 2006. The marine viromes of four oceanic regions. PLoS Biol 4. https://doi.org/10.1371/journal.pbio.0040368

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44, W16–W21. https://doi.org/10.1093/nar/gkw387

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O.,

Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9, 75. https://doi.org/10.1186/1471-2164-9-75

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

Biller, S.J., McDaniel, L.D., Breitbart, M., Rogers, E., Paul, J.H., Chisholm, S.W., 2017. Membrane vesicles in sea water: heterogeneous DNA content and implications for viral abundance estimates. The ISME Journal 11, 394–404. https://doi.org/10.1038/ismej.2016.134

Boccara, M., Fedala, Y., Bryan, C.V., Bailly-Bechet, M., Bowler, C., Boccara, A.C., 2016. Full-field interferometry for counting and differentiating aquatic biotic nanoparticles: from laboratory to Tara Oceans. Biomed Opt Express 7, 3736–3746. https://doi.org/10.1364/BOE.7.003736

Castro-Mejía, J.L., Muhammed, M.K., Kot, W., Neve, H., Franz, C.M.A.P., Hansen, L.H., Vogensen, F.K., Nielsen, D.S., 2015. Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of phage communities in the human gut. Microbiome 3, 64. https://doi.org/10.1186/s40168-015-0131-4

Clokie, M.R.J., Millard, A.D., Letarov, A.V., Heaphy, S., 2011. Phages in nature. Bacteriophage 1, 31–45. https://doi.org/10.4161/bact.1.1.14942

Conceição-Neto, N., Zeller, M., Lefrère, H., Bruyn, P.D., Beller, L., Deboutte, W., Yinda, C.K., Lavigne, R., Maes, P., Ranst, M.V., Heylen, E., Matthijnssens, J., 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. Scientific Reports 5, 16532. https://doi.org/10.1038/srep16532

de Melo, A.G., Levesque, S., Moineau, S., 2018. Phages as friends and enemies in food processing. Current Opinion in Biotechnology 49, 185–190. https://doi.org/10.1016/j.copbio.2017.09.004

Dugat-Bony, E., Garnier, L., Denonfoux, J., Ferreira, S., Sarthou, A.-S., Bonnarme, P., Irlinger, F., 2016. Highlighting the microbial diversity of 12 French cheese varieties. International Journal of Food Microbiology 238, 265–273. https://doi.org/10.1016/j.ijfoodmicro.2016.09.026

Fernández, L., Rodríguez, A., García, P., 2018. Phage or foe: an insight into the impact of viral predation on microbial communities. The ISME Journal 12, 1171. https://doi.org/10.1038/s41396-018-0049-5

Forterre, P., Soler, N., Krupovic, M., Marguet, E., Ackermann, H.-W., 2013. Fake virus particles generated by fluorescence microscopy. Trends in Microbiology 21, 1–5. https://doi.org/10.1016/j.tim.2012.10.005

Gill, S., Catchpole, R., Forterre, P., 2018. Extracellular membrane vesicles in the three domains of life and beyond. FEMS Microbiol Rev 43, 273–303. https://doi.org/10.1093/femsre/fuy042

Guo, P., El-Gohary, Y., Prasadan, K., Shiota, C., Xiao, X., Wiersch, J., Paredes, J., Tulachan, S., Gittes, G.K., 2012. Rapid and simplified purification of recombinant adeno-associated virus. J Virol Methods 183, 139–146. https://doi.org/10.1016/j.jviromet.2012.04.004

Hyman, P., Abedon, S.T., Hyman, P., Abedon, S.T., 2012. Smaller fleas: viruses of microorganisms. Scientifica 2012, e734023. https://doi.org/10.6064/2012/734023

Irlinger, F., Layec, S., Hélinck, S., Dugat-Bony, E., 2015. Cheese rind microbial communities: diversity, composition and origin. FEMS Microbiol. Lett. 362, 1–11. https://doi.org/10.1093/femsle/fnu015

Irlinger, F., Mounier, J., 2009. Microbial interactions in cheese: implications for cheese quality and safety. Current Opinion in Biotechnology 20, 142–148. https://doi.org/10.1016/j.copbio.2009.02.016

Jurczak-Kurek, A., Gąsior, T., Nejman-Faleńczyk, B., Bloch, S., Dydecka, A., Topka, G., Necel, A., Jakubowska-Deredas, M., Narajczyk, M., Richert, M., Mieszkowska, A., Wróbel, B., Węgrzyn, G., Węgrzyn, A., 2016. Biodiversity of bacteriophages: morphological and biological properties of a large group of phages isolated from urban sewage. Scientific Reports 6, 34338. https://doi.org/10.1038/srep34338

Kassambara, A., 2018. ggpubr: "ggplot2" Based Publication Ready Plots. R package version 0.2. https://CRAN.R-project.org/package=ggpubr.

Kastman, E.K., Kamelamela, N., Norville, J.W., Cosetta, C.M., Dutton, R.J., Wolfe, B.E., 2016. Biotic interactions shape the ecological distributions of Staphylococcus species. mBio 7, e01157-16. https://doi.org/10.1128/mBio.01157-16

Kleiner, M., Hooper, L.V., Duerkop, B.A., 2015. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. BMC Genomics 16. https://doi.org/10.1186/s12864-014-1207-4

Koonin, E.V., Krupovic, M., Yutin, N., 2015. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. Ann. N. Y. Acad. Sci. 1341, 10–24. https://doi.org/10.1111/nyas.12728

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211–3217. https://doi.org/10.1093/bioinformatics/bts611

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., Merchat, M., Suzan-Monti, M., Forterre, P., Koonin, E., Raoult, D., 2008. The virophage as a unique parasite of the giant mimivirus. Nature 455, 100–104. https://doi.org/10.1038/nature07218

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Liu, Z., Li, H., Wang, M., Jiang, Y., Yang, Q., Zhou, X., Gong, Z., Liu, Q., Shao, H., 2018. Isolation, characterization and genome sequencing of the novel phage SL25 from the Yellow Sea, China. Marine Genomics 37, 31–34. https://doi.org/10.1016/j.margen.2017.09.008

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

Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S.A., Wu, G.D., Lewis, J.D., Bushman, F.D., 2011. The human gut virome: Inter-individual variation and dynamic response to diet. Genome Res 21, 1616–1625. https://doi.org/10.1101/gr.122705.111

Montel, M.-C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D.A., Desmasures, N., Berthier, F., 2014. Traditional cheeses: Rich and diverse microbiota with associated benefits. International Journal of Food Microbiology 177, 136–154. https://doi.org/10.1016/j.ijfoodmicro.2014.02.019

Mounier, J., Monnet, C., Vallaeys, T., Arditi, R., Sarthou, A.-S., Helias, A., Irlinger, F., 2008. Microbial interactions within a cheese microbial community. Applied and Environmental Microbiology 74, 172–181. https://doi.org/10.1128/AEM.01338-07

Muhammed, M.K., Kot, W., Neve, H., Mahony, J., Castro-Mejía, J.L., Krych, L., Hansen, L.H., Nielsen, D.S., Sørensen, S.J., Heller, K.J., Sinderen, D. van, Vogensen, F.K., 2017. Metagenomic analysis of dairy bacteriophages: extraction method and pilot study on whey samples derived from using undefined and defined mesophilic starter cultures. Appl. Environ. Microbiol. 83, e00888-17. https://doi.org/10.1128/AEM.00888-17

Murugesan, S., Reyes-Mata, M.P., Nirmalkar, K., Chavez-Carbajal, A., Juárez-Hernández, J.I., Torres-Gómez, R.E., Piña-Escobedo, A., Maya, O., Hoyo-Vadillo, C., Ramos-Ramírez, E.G., Salazar-Montoya, J.A., García-Mena, J., 2018. Profiling of bacterial and fungal communities of Mexican cheeses by high throughput DNA sequencing. Food Research International 113, 371–381. https://doi.org/10.1016/j.foodres.2018.07.023

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

Pinard, R., de Winter, A., Sarkis, G.J., Gerstein, M.B., Tartaro, K.R., Plant, R.N., Egholm, M., Rothberg, J.M., Leamon, J.H., 2006. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics 7, 216. https://doi.org/10.1186/1471-2164-7-216

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41, D590–D596. https://doi.org/10.1093/nar/gks1219

Quigley, L., O'Sullivan, O., Beresford, T.P., Ross, R.P., Fitzgerald, G.F., Cotter, P.D., 2012. High-throughput sequencing for detection of subpopulations of bacteria not previously associated with artisanal cheeses. Applied and Environmental Microbiology 78, 5717–5723. https://doi.org/10.1128/AEM.00918-12

Randazzo, C.L., Torriani, S., Akkermans, A.D.L., de Vos, W.M., Vaughan, E.E., 2002. Diversity, dynamics, and activity of bacterial communities during production of an artisanal Sicilian cheese as evaluated by 16S rRNA analysis. Appl Environ Microbiol 68, 1882–1892. https://doi.org/10.1128/AEM.68.4.1882-1892.2002

Reyes-Robles, T., Dillard, R.S., Cairns, L.S., Silva-Valenzuela, C.A., Housman, M., Ali, A., Wright, E.R., Camilli, A., 2018. *Vibrio cholerae* outer membrane vesicles inhibit bacteriophage infection. Journal of Bacteriology 200. https://doi.org/10.1128/JB.00792-17

Roose-Amsaleg, C., Fedala, Y., Vénien-Bryan, C., Garnier, J., Boccara, A.-C., Boccara, M., 2017. Utilization of interferometric light microscopy for the rapid analysis of virus abundance in a river. Research in Microbiology 168, 413–418. https://doi.org/10.1016/j.resmic.2017.02.004

Roux, S., Enault, F., Hurwitz, B.L., Sullivan, M.B., 2015. VirSorter: mining viral signal from microbial genomic data. PeerJ 3, e985. https://doi.org/10.7717/peerj.985

Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime-Ngando, T., Debroas, D., 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. PLOS ONE 7, e33641. https://doi.org/10.1371/journal.pone.0033641

Roux, S., Krupovic, M., Debroas, D., Forterre, P., Enault, F., 2013. Assessment of viral community functional potential from viral metagenomes may be hampered by contamination with cellular sequences. Open Biology 3, 130160. https://doi.org/10.1098/rsob.130160

Samson, J.E., Moineau, S., 2010. Characterization of Lactococcus lactis phage 949 and comparison with other lactococcal phages. Appl. Environ. Microbiol. 76, 6843–6852. https://doi.org/10.1128/AEM.00796-10

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nature Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Sime-Ngando, T., 2014. Environmental bacteriophages: viruses of microbes in aquatic ecosystems. Front. Microbiol. 5, 355. https://doi.org/10.3389/fmicb.2014.00355

Stals, A., Baert, L., Van Coillie, E., Uyttendaele, M., 2012. Extraction of food-borne viruses from food samples: A review. International Journal of Food Microbiology 153, 1–9. https://doi.org/10.1016/j.ijfoodmicro.2011.10.014

Thurber, R.V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory procedures to generate viral metagenomes. Nature Protocols 4, 470–483. https://doi.org/10.1038/nprot.2009.10

Ulve, V. m., Monnet, C., Valence, F., Fauquant, J., Falentin, H., Lortal, S., 2008. RNA extraction from cheese for analysis of in situ gene expression of Lactococcus lactis. Journal of Applied Microbiology 105, 1327–1333. https://doi.org/10.1111/j.1365-2672.2008.03869.x

Varsani, A., Krupovic, M., 2018. Smacoviridae: a new family of animal-associated single-stranded DNA viruses. Arch Virol 163, 2005–2015. https://doi.org/10.1007/s00705-018-3820-z

Weynberg, K.D., Wood-Charlson, E.M., Suttle, C.A., Oppen, V., H, M.J., 2014. Generating viral metagenomes from the coral holobiont. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00206

Willner, D., Furlan, M., Schmieder, R., Grasis, J.A., Pride, D.T., Relman, D.A., Angly, F.E., McDole, T., Mariella, R.P., Rohwer, F., Haynes, M., 2011. Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. PNAS 108, 4547–4553. https://doi.org/10.1073/pnas.1000089107

Wolfe, B.E., Button, J.E., Santarelli, M., Dutton, R.J., 2014. Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. Cell 158, 422–433. https://doi.org/10.1016/j.cell.2014.05.041

Yaron, S., Kolling, G.L., Simon, L., Matthews, K.R., 2000. Vesicle-mediated transfer of virulence genes from Escherichia coli O157:H7 to other enteric bacteria. Applied and Environmental Microbiology 66, 4414–4420. https://doi.org/10.1128/AEM.66.10.4414-4420.2000

Zablocki, O., Zyl, L. van, Adriaenssens, E.M., Rubagotti, E., Tuffin, M., Cary, S.C., Cowan, D., 2014. High-level diversity of tailed phages, eukaryote-associated viruses, and virophage-like elements in the metaviromes of antarctic soils. Appl. Environ. Microbiol. 80, 6888–6897. https://doi.org/10.1128/AEM.01525-14

Zhang, Y., Kastman, E.K., Guasto, J.S., Wolfe, B.E., 2018. Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. Nature Communications 9, 336. https://doi.org/10.1038/s41467-017-02522-z

Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., Alva, V., 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. Journal of Molecular Biology 430, 2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007

Figure legends

Fig. 1. Schematic representation of the experimental procedure for the extraction of viruses from cheese. The surface of three types of cheese was processed according to four routes named P1 to P4. Steps 1 (homogenization), 2 (centrifugation), 3 (dilution) and 5 (concentration) are common to all protocols, contrary to steps 4 (filtration) and 6 (chloroform treatment) which are optional. For each cheese type, three biological replicates (independent cheeses purchased at the same date and from a unique producer) were performed leading thus to 12 samples per cheese-type.

Fig. 2. Photographs of iodixanol density gradients after ultracentrifugation of cheese viral fractions. Cellular debris and membrane vesicles are concentrated at the top of the 20% iodixanol layer (red arrow) and viruses at the top of the 45% iodixanol layer (blue arrow). A schematic representation of the gradient is shown on the right. S: sample.

Fig. 3. Comparative plot showing the estimated level of bacterial DNA contamination in the Epoisses cheese virome. The expected percentage (x axis) is based on the % of rDNA reads detected by SortMeRNA software. This value was used to calculate the estimated % of bacterial chromosomal DNA using a multiplying factor of (x100) based on estimation that several copies of rRNA operons (*e.g.* 5 copies of ~5 kb in average) would represent about 1% of a bacterial chromosome of an average size of 2.5 Mb. The observed % of possible bacterial contamination was obtained by looking at the % of reads not matching the 124 contigs of the cheese virome. Yellow (P1), red (P2), green (P3), blue (P4) and orange (vesicles sample).

Fig. 4. Differences in the Epoisses cheese virome composition according to the extraction procedure. Principal coordinate analysis based on the Bray-Curtis dissimilarity index (A). Relative abundance of the contigs annotated as probable phages (B, 72 contigs), plasmids (C, 19 contigs), putative plasmids (D, 20 contigs), unclassified (E, 14 contigs) and unmapped reads (F).

Fig. 5. Heatmap of the 124 contigs composing the Epoisses virome. Samples were grouped according to the extraction procedure. Contigs were separated into four categories (phage, plasmid, putative plasmid and unclassified) according to the annotation results and sorted by abundance inside each group. Yellow to red: lowest to highest relative abundance. Grey: not detected. Ves: vesicles sample.

Fig. 6. Electron micrographs of representative phages from Epoisses cheese. The viral fraction was obtained with protocol P4.

Tables

Table 1. Quantification of viral particles using interferometry and microbial cells using plate numbering.

Type of Cheese and of protocol	Total nanoparticles (log per g) (SD)	Presence of cells	Presence of noise	Total biomass (log CFU per g) (SD)	Ratio nanoparticles / cell
CAM P1	NA	+++	+++	8.71 (0.11)	NA
EP P1	NA	+++	+++	10.10 (0.31)	NA
SN P1	NA	+++	+++	10.42 (0.07)	NA
CAM P2	10.38 (0.01)	++	-	8.71 (0.11)	47.29 (11.91)
EP P2	10.18 (0.23)	+	-	10.10 (0.31)	1.88 (2.07)
SN P2	10.04 (0.24)	+	-	10.42 (0.07)	0.66 (0.14)
CAM P3	NA	++	+++	8.71 (0.11)	NA
EP P3	NA	-	+++	10.10 (0.31)	NA
SN P3	NA	-	+++	10.42 (0.07)	NA
CAM P4	10.00 (0.13)	-	-	8.71 (0.11)	21.77 (12.15)
EP P4	10.13 (0.37)	-	-	10.10 (0.31)	2.40 (3.35)
SN P4	9.59 (0.40)	-	-	10.42 (0.07)	0.17 (0.10)

NA: not available because of noise in interferometric measures

SD: standard deviation

Supplementary data

Table S1. Comparison between phage titer and interferometry measure.

Table S2. Microbiological counts obtained for the nine cheese samples.

Table S3. DNA quantities obtained for the different samples studied.

Table S4. Sequence comparison results of the 124 contigs composing the Epoisses virome.

Table S5. Predicted function and best BlastP hit for ORFs of Epvir4 and Epvir8.

Table S6. Characteristics of the different phage morphotypes detected at the surface of Epoisses cheese.

Fig. S1. Comparison between phage concentration as measured by titration and using the interferometric light microscope (ILM). Six reference phages were used for the comparison, namely phages P1, T4, T5, T7, λCI857 and φX174.

Fig. S2. Epifluorescence image of a viral fraction obtained from Epoisses cheese. The viral fraction was diluted 10-fold in 1 ml of SM buffer with gluteraldehyde (2.5% final), incubated 10 minutes on ice and then flash frozen in liquid nitrogen. 100 µl of the nanoparticle solution was diluted into 5 ml of filtered water; and then filtered through a 25 mm anodisc membrane with pore sizes of 0.02 µm and stained with 100 µL of SYBR Gold (50×) for 15 min in the dark. Filters were placed on a microscopy slide with 100 µL of anti-fading solution Fluoromount, and observed immediately after preparation on a microscope Leica DMRA2 microscope equipped with a × 100 magnification oil□immersion objective and a COOLSNAP HQ camera (Roper Scientific, USA), using YFP filters. Images were captured and processed with METAMORPH V6.3r5.

Fig. S3. Schematic representation of the assembly and mapping strategy. The number of reads kept at each step of the process is indicated.

Fig. S4. Comparison of contigs abundance in the Epoisses virome profiles produced using four extraction protocols. Scatter plots produced from the log10 abundance values of the

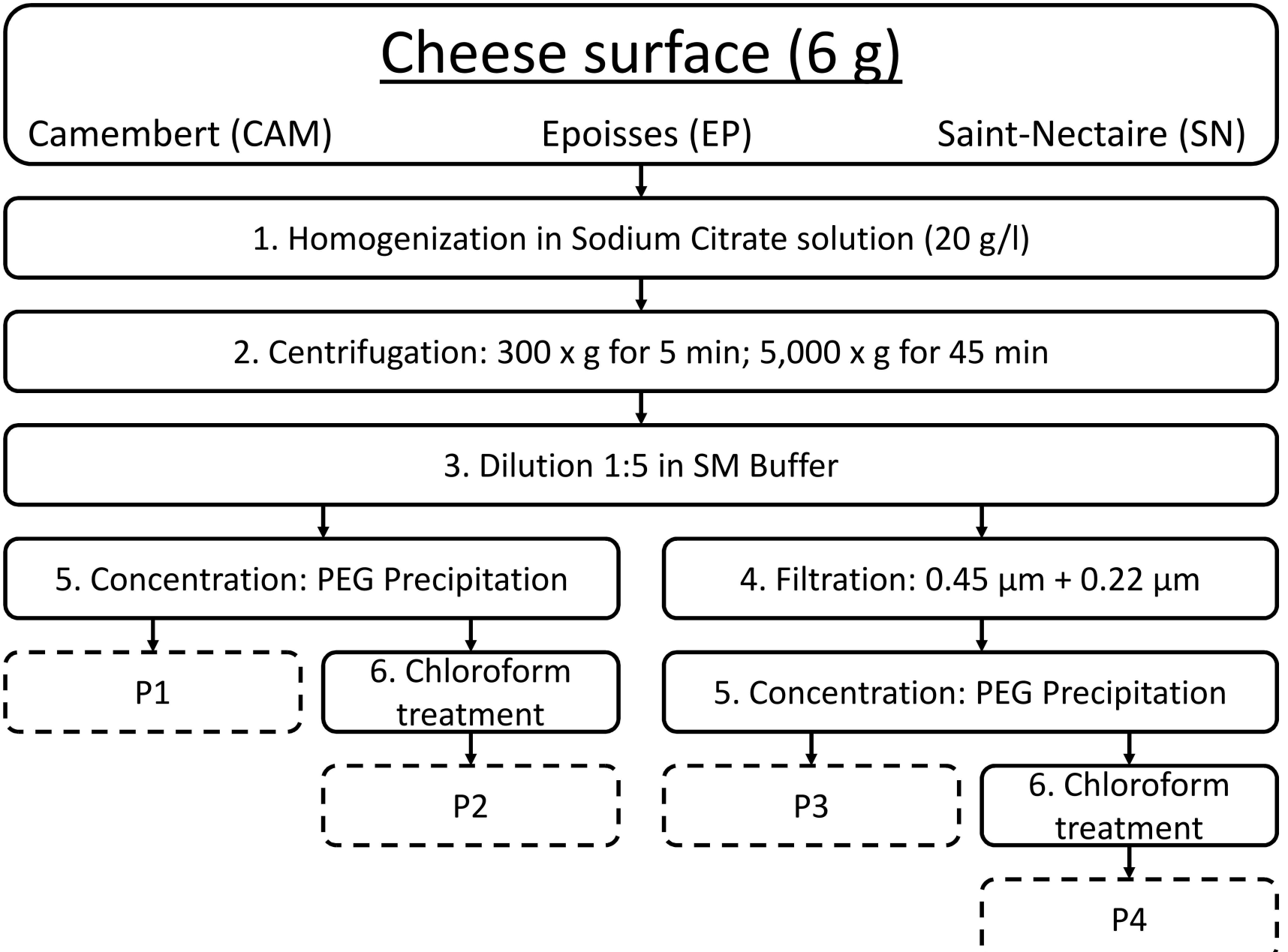
contigs identified as probable phages (A) or from the other contigs (B). R = Spearman rank correlation coefficients (Rho).

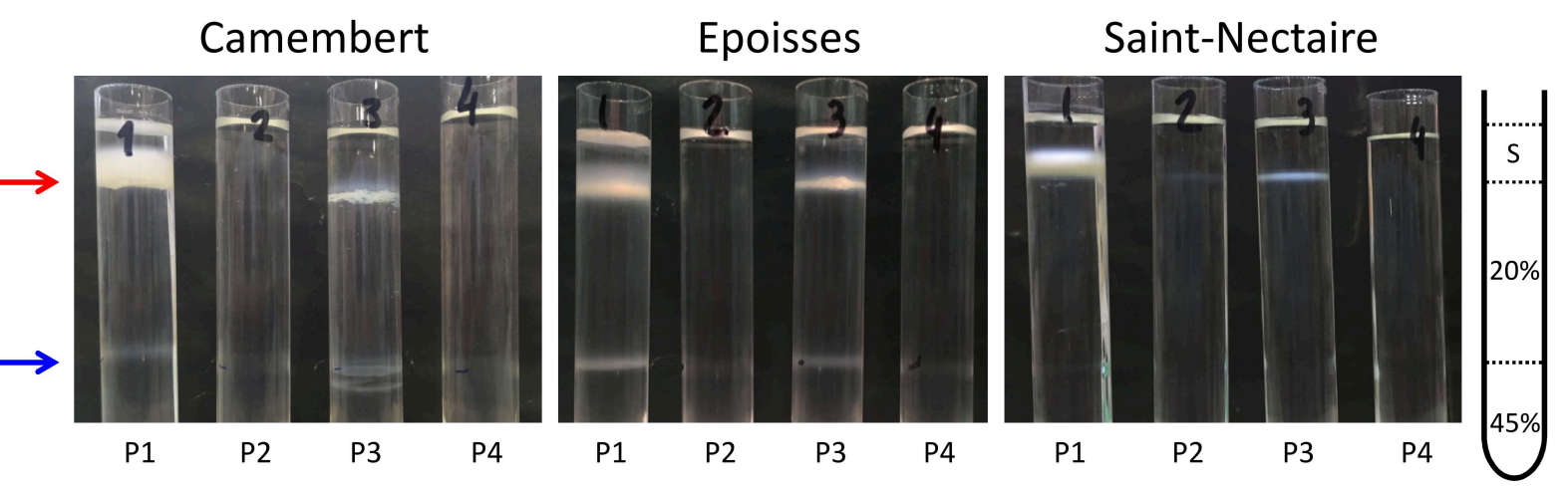
Fig. S5. Functional annotation of the three most abundant viral contigs detected in the Epoisses cheese virome. Predicted functions of the ORFs identified in Epvir4 (A), 949_Epvir1 (B) and Epvir8 (C) are color-coded as follow: red for integrase, yellow for transcriptional regulation, orange for replication and recombination, green for DNA packaging and head, light blue for connector, dark blue for tail, light pink for homing endonuclease, dark pink for lysis, grey for hypothetical proteins, violet for additional functions. Some of the predicted functions of the ORFs are detailed for each contig.

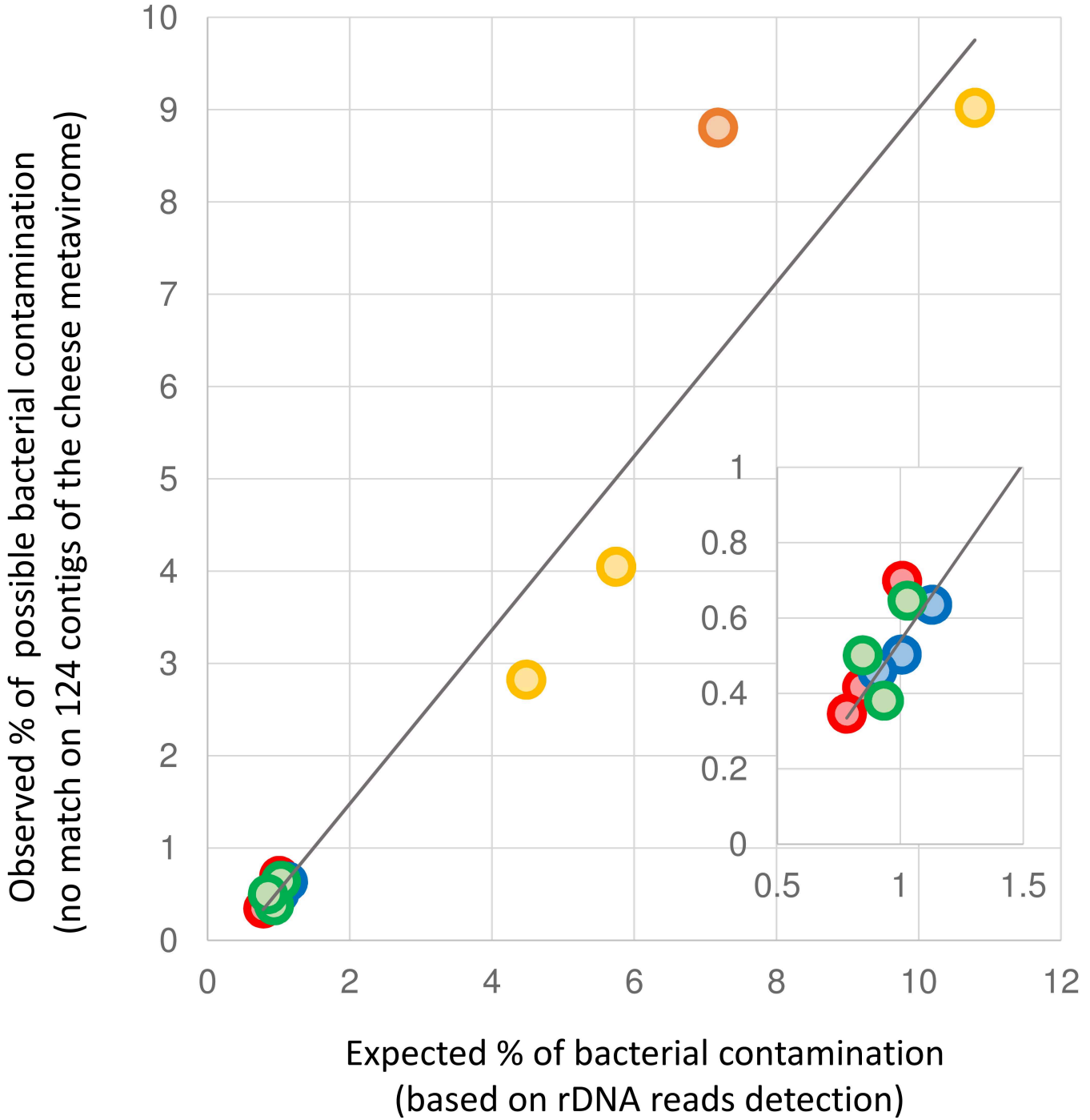
Highlights

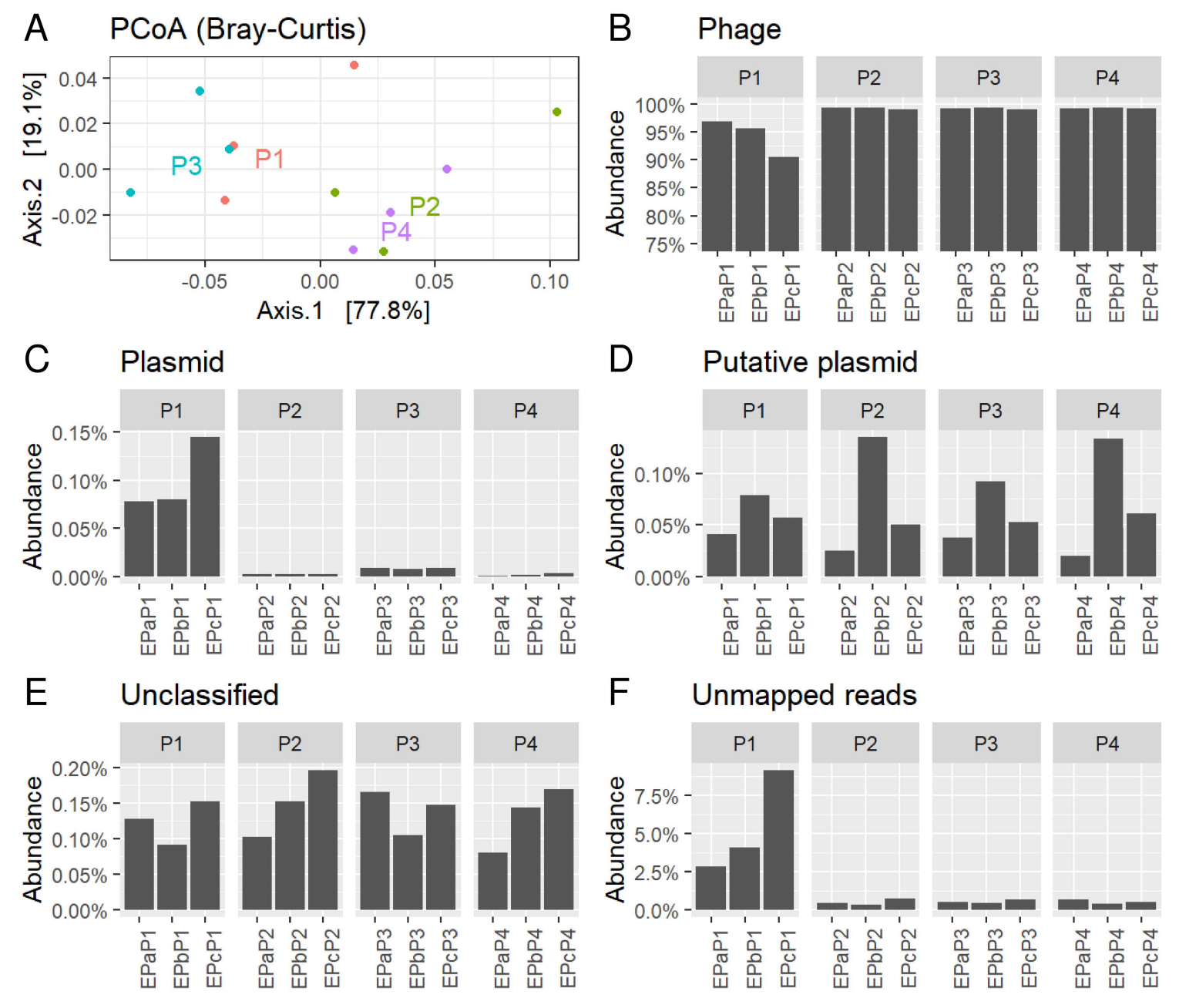
- set up of an efficient procedure for extracting viral particles from cheese surface
- extremely low contamination of the virome by microbial DNA
- first viral metagenomic dataset from the cheese surface
- genomic characterization of the dominant phages from Epoisses cheese

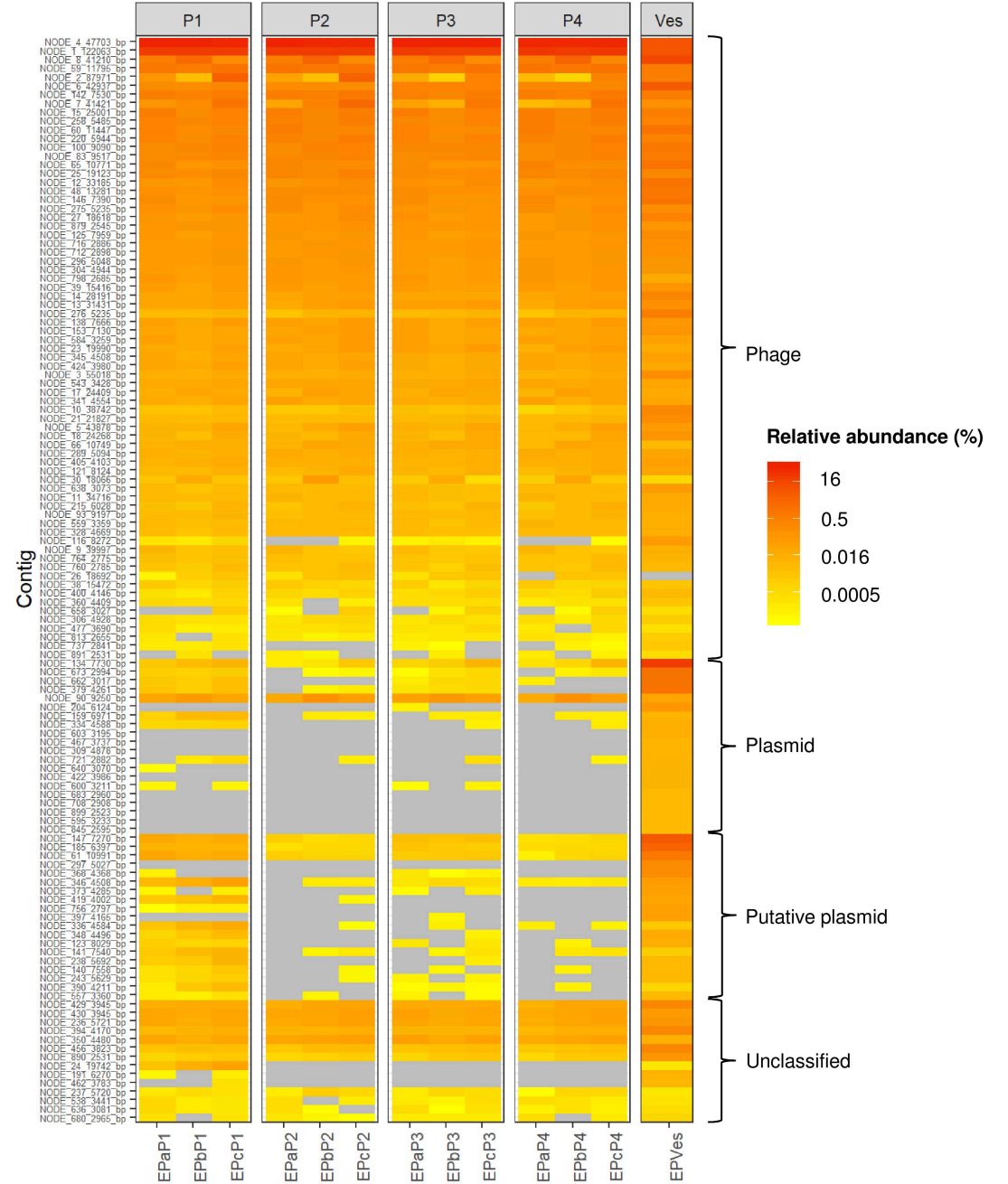
The authors declare no conflict of interest.

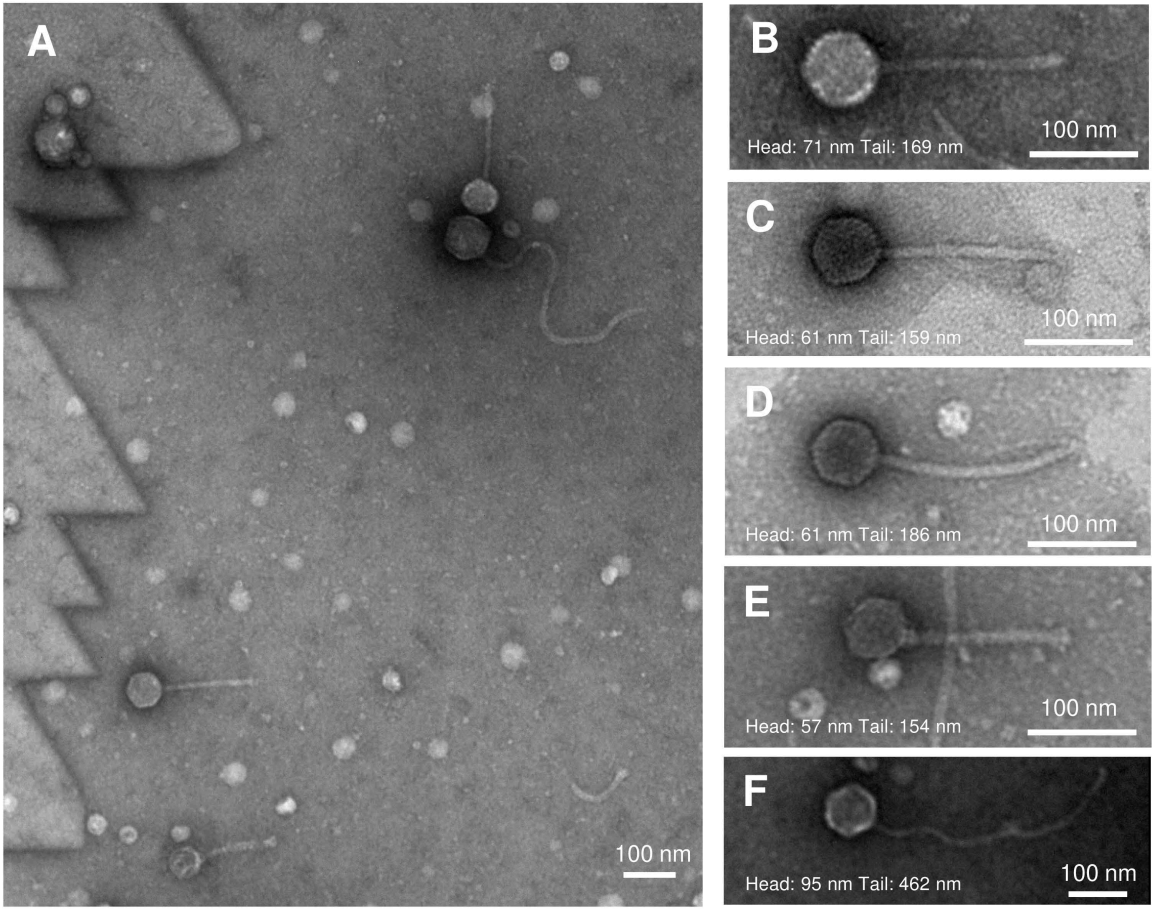

Cheese surface (6 g)
Camembert (CAM)
Epoisses (EP)
Saint-Nectaire (SN)
1. Homogenization in Sodium Citrate solution (20 g/l)
2. Centrifugation: 300 x g for 5 min; 5,000 x g for 45 min
3. Dilution 1:5 in SM Buffer
5. Concentration: PEG Precipitation
4. Filtration: 0.45 µm + 0.22 µm
P1
6. Chloroform treatment
5. Concentration: PEG Precipitation
P2
P3
6. Chloroform treatment
P4


Camembert
Epoisses
Saint-Nectaire
1
2
3
4
P1
P2
P3
P4
S
20%
45%


Observed % of possible bacterial contamination
(no match on 124 contigs of the cheese metavirome)
Expected % of bacterial contamination
(based on rDNA reads detection)
0
1
2
3
4
5
6
7
8
9
10
0
2
4
6
8
10
12
0
0.2
0.4
0.6
0.8
1
0.5
1
1.5


A
PCoA (Bray-Curtis)
Axis.2 [19.1%]
Axis.1 [77.8%]
P1
P2
P3
P4
B
Phage
Abundance
C
Plasmid
D
Putative plasmid
E
Unclassified
F
Unmapped reads
EPaP1
EPbP1
EPcP1
EPaP2
EPbP2
EPcP2
EPaP3
EPbP3
EPcP3
EPaP4
EPbP4
EPcP4


P1
P2
P3
P4
Ves
Contig
Phage
Plasmid
Putative plasmid
Unclassified
Relative abundance (%)
16
0.5
0.016
0.0005
EPaP1
EPbP1
EPcP1
EPaP2
EPbP2
EPcP2
EPaP3
EPbP3
EPcP3
EPaP4
EPbP4
EPcP4
EPVes


A
100 nm
B
Head: 71 nm Tail: 169 nm
100 nm
C
Head: 61 nm Tail: 159 nm
100 nm
D
Head: 61 nm Tail: 186 nm
100 nm
E
Head: 57 nm Tail: 154 nm
100 nm
F
Head: 95 nm Tail: 462 nm
100 nm