

Article

A Genome of Temperate *Enterococcus* Bacteriophage Placed in a Space of Pooled Viral Dark Matter Sequences

Ivan M. Pchelin ^{1,*} , Pavel V. Tkachev ¹ , Daniil V. Azarov ¹, Andrey N. Gorshkov ^{2,3}, Daria O. Drachko ^{4,5}, Vasily V. Zlatogursky ⁵, Alexander V. Dmitriev ¹ and Artemiy E. Goncharov ¹ 

¹ Scientific and Educational Center “Molecular Bases of Interaction of Microorganisms and Human” of the WCRC “Center for Personalized Medicine”, Institute of Experimental Medicine, 197022 Saint Petersburg, Russia

² Smorodintsev Research Institute of Influenza, Ministry of Health of the Russian Federation, 197376 Saint Petersburg, Russia

³ Laboratory of Pathomorphology, Almazov National Research Centre, 197341 Saint Petersburg, Russia

⁴ Laboratory of Cellular and Molecular Protistology, Zoological Institute of the Russian Academy of Sciences, 199034 Saint Petersburg, Russia

⁵ Department of Invertebrate Zoology, Faculty of Biology, St. Petersburg State University, 199034 Saint Petersburg, Russia

* Correspondence: arcella.oraia@gmail.com

Abstract: In the human gut, temperate bacteriophages interact with bacteria through predation and horizontal gene transfer. Relying on taxonomic data, metagenomic studies have associated shifts in phage abundance with a number of human diseases. The temperate bacteriophage VEsP-1 with siphovirus morphology was isolated from a sample of river water using *Enterococcus faecalis* as a host. Starting from the whole genome sequence of VEsP-1, we retrieved related phage genomes in blastp searches of the tail protein and large terminase sequences, and blastn searches of the whole genome sequences, with matches compiled from several different databases, and visualized a part of viral dark matter sequence space. The genome network and phylogenomic analyses resulted in the proposal of a novel genus “Vespunovirus”, consisting of temperate, mainly metagenomic phages infecting *Enterococcus* spp.

Keywords: bacteriophage; *Enterococcus*; siphovirus; phylogeny; sequence space



Citation: Pchelin, I.M.; Tkachev, P.V.; Azarov, D.V.; Gorshkov, A.N.; Drachko, D.O.; Zlatogursky, V.V.; Dmitriev, A.V.; Goncharov, A.E. A Genome of Temperate *Enterococcus* Bacteriophage Placed in a Space of Pooled Viral Dark Matter Sequences. *Viruses* **2023**, *15*, 216. <https://doi.org/10.3390/v15010216>

Academic Editor: Breck A. Duerkop

Received: 21 December 2022

Revised: 10 January 2023

Accepted: 11 January 2023

Published: 12 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Temperate bacteriophages are estimated to constitute at least 20–50% of free phages in the human gut [1] or even outnumber the lytic ones [2]. While integrated in the bacterial chromosome, they offer their hosts the encoded adhesion factors, CRISPR arrays, toxins, and stress tolerance determinants, contributing therefore to bacterial adaptation [3]. Through packaging host DNA into virions and subsequent recombination within a new host, temperate phages facilitate horizontal gene transfer and drive bacterial evolution [4].

The composition of phage populations varies from person to person, but in adult individuals, it is generally stable at the time scale of days [5] and years [2]. The relative contribution of bacteriophages to the viral community of the gut is associated with inflammatory bowel disease, where the changes in the virome composition reflect alterations in the bacterial composition [6]. The bacteriophage abundance increases in individuals with Crohn’s disease [7]. Statistical analysis, including not only clinical or bacterial data, but also data on the fecal viral diversity, allows for a more accurate identification of patients with severe nonalcoholic fatty liver disease and fibrosis [8]. In alcoholic hepatitis, specific bacteriophage taxa were associated with an increased disease severity and 90-day mortality [9]. The virome signatures associated with colorectal cancer have a potential to be used for developing screening tools and predicting disease outcomes [10]. The mentioned

studies relied on a phage classification at the levels of families [6,7], species [8,9], or species and genera [10] and illustrate why bacteriophage taxonomy is not only of a theoretical, but also of an applied significance.

In the studies on the human gut microbiome, the undescribed viruses may account for 75–99% of viral sequence reads [11]. The major part of this diversity is constituted by temperate bacteriophages. Due to their prevalence, a resistance to culturing techniques and sophisticated evolutionary patterns, these phages are often referred to as “the viral dark matter” [12]. The description and classification of bacteriophages lags behind the generation of metagenomic sequence data, hampering the interpretation of the experimental results [13,14].

Therefore, we aimed to describe a novel temperate *Enterococcus* phage VEsP-1. The taxonomic analysis of VEsP-1 and the related sequences from the GenBank substantiated a proposal for a new genus, “Vespunovirus”, represented mainly by metagenomic phage sequences. The genome network, calculated from sequences retrieved in nine sequential blast searches, allowed us to visualize the relationships between very dissimilar bacteriophages.

2. Materials and Methods

2.1. Isolation and Description of *Enterococcus* Phage VEsP-1

Bacteriophage isolation, host spectrum determination, transmission electron microscopy, and DNA sequencing were done with the use of protocols described in our previous publication [15]. Briefly, the bacteriophage VEsP-1 was isolated from a sample of river water enriched with the *E. faecalis* strain Serg. The same strain was used for bacteriophage propagation. The sample originated from Tô Lịch River (Hanoi, Vietnam). The host range was determined with the use of an in-house collection of 80 bacterial strains identified by MALDI-TOF mass spectrometry. The collection included 65 *Enterococcus* spp. strains and 15 isolates belonging to other bacterial genera, mainly *Staphylococcus* and *Streptococcus* (Table S1). The lytic activity was determined on a bacterial lawn by a spot test with eight serial 10-fold dilutions of the phage stock preparation. The efficiency of plating values were calculated by dividing the phage titer determined on a bacterial lawn by a titer of phage on a lawn of *E. faecalis* Serg.

The size of viral particles ($n = 8$) was measured from transmission electron microscopy images. The probable natural range of the mean capsid diameter and tail length values was predicted by calculating the standard errors. DNA isolation was performed by the phenol/chloroform method with PEG 8000/NaCl precipitated bacteriophage preparation. The phage genome sequence was obtained with the use of the Illumina MiSeq platform (Illumina, San Diego, CA, USA). The library preparation was carried out with a Nextera XT DNA Library Preparation Kit (Illumina), resulting in 300 bp paired-end reads. Quality control was performed using FastQC 0.11.9 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 1 May 2020). A whole genome sequence was assembled de novo by SPAdes 3.13.0 [16] and deposited at the NCBI Nucleotide database with the accession number MZ333456. Coding sequences were predicted by the RAST genome annotation service (<https://rast.nmpdr.org/rast.cgi>, accessed on 7 December 2022) [17–19]. The tool was set to follow the RASTtk annotation scheme with frameshift correction. The possible spontaneous induction of VEsP-1 was excluded by searching the genome of the *E. faecalis* strain Serg (BioProject ID PRJNA884994) with the phage sequence.

2.2. Genome Network Analysis

To determine the position of VEsP-1 in the space of the related phage genomes, the VEsP-1 genome-centered sequence sample was obtained by performing three BLAST searches against the sequences of Caudoviricetes deposited in GenBank (NCBI:txid2731619). Two blastp searches were performed with the amino acid sequences of the terminase large subunit QYI86497.1 (TLS) and tail protein QYI86480.1 (TP) of VEsP-1. The retrieved accessions were used to query the NCBI Identical Protein Groups database and prepare the accession lists of the corresponding phage genome sequences. The obtained accession

lists were pooled with the results of the blastn search with the VEsP-1 genome sequence MZ333456 (WGS) and were used to download sequence data from the NCBI Nucleotide database. To increase the sample size, the VEsP-1 genome-centered sample was pooled with sequences, obtained through similar searches with TLS, major capsid proteins and the WGS of Streptococcus phage P7954 (KY705280.1), and Brevibacillus phage Powder (KT151958.1). The searches were performed in July 2022. In all cases, every sequence producing a significant alignment was taken for a further analysis. The genomes shorter than 33,512 bp, 85% of the VEsP-1 sequence MZ333456.1, were considered to be incomplete and were thus rejected. For a taxonomic annotation, we downloaded the full NCBI Taxonomy entries associated with Nucleotide accessions in the XML file format with the use of Batch Entrez feature. To connect the sequence accessions to the taxonomic data, the associated taxonomy identifiers were retrieved using Entrez ESummary utility.

The structure of the sequence space in the general sample of the genomes was determined by a vContact2 0.11.3 analysis with the default parameters [20,21]. The calculations were carried out with the protein sequences predicted by prokka 1.14.6 [22]. The obtained pairwise distances were based on the probability of sharing a certain number of protein clusters between the genomes [20]. The resulting network file was visualized with the use of Cytoscape 3.9.1 [23] with edge weights interpreted as negative Log values. The group of VEsP-1-related viruses was predicted by vContact2 as a cluster of sequences with a statistically significant number of shared genes [20].

2.3. Phylogenomic Analysis

The phylogenomic analysis was performed with the VICTOR web service (<https://victor.dsmz.de>, accessed on 5 September 2022) [24]. The analysis included pairwise comparisons of the nucleotide sequences using the Genome-BLAST Distance Phylogeny (GBDP) method [25] optimized for prokaryotic viruses [24]. The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME, including SPR postprocessing [26] for each of the formulas D0, D4, and D6, respectively. Branch support was inferred from 100 pseudo-bootstrap replicates each. Taxon boundaries were estimated with the OPTSIL program [27], with the clustering thresholds described by Meier-Kolthoff and Göker [24] and the fraction of links required for a cluster fusion of 0.5 [28]. Intergenomic similarities were calculated using the VIRIDIC tool with the default parameters (<http://rhea.icbm.uni-oldenburg.de/VIRIDIC/>, accessed on 9 January 2023) [29]. Additional methods are described in the Supplementary Materials [30–32].

2.4. Functional Analysis

The search for homologous protein-coding genes was performed with the CoreGenes 5.0 webserver (<http://coregenes.ngrok.io/>, accessed on 27 July 2022) using the Group Clustering: Bidirectional Best Hit Algorithm [33]. The prediction of capsid structure was done using the code described by Lee et al. on the basis of the sequence of the gene coding for a major capsid protein. The result was expressed as the T-number, which is a discrete index determining the possible capsid surfaces compatible with icosahedral symmetry and 1/60th of the estimated number of the MCP molecules needed to complete the head of phage particle [34]. The search for the known antibiotic resistance determinants and bacterial virulence factors was performed using the ABRicate 1.0.1 tool [35] and the databases ARG-ANNOT, CARD, MEGARes 2.00, NCBI AMRFinderPlus, Resfinder, and VFDB [36–41].

Bacteriophage protein immunogenicity prediction was done with the use of two web servers: VaxiJen 2.0 (http://www.ddg-pharmfac.net/vaxijen/scripts/VaxiJen_scripts/VaxiJen3.pl, accessed on 10 August 2022) [42,43] and VirVACPRED (<https://virvacpred.herokuapp.com/>, accessed on 10 August 2022) [44] with the thresholds recommended in the original publications. Both web servers are designed to perform alignment-free searches for protective immunogens with whole protein sequences as an input.

3. Results

3.1. Phenotype of Enterococcus Phage VEsP-1

The bacteriophage VEsP-1 was isolated from a sample of river water. It was able to produce plaques on four *Enterococcus faecalis* strains (11% of tested *E. faecalis* isolates). The results of a spot test assay with other bacterial isolates were negative (Table S1). Transmission electron microscopy revealed viral particles with a siphovirus morphology, possessing a long flexible tail and an icosahedral head (Figure 1). The average head size was 50.3 ± 0.6 nm and the average tail length was 148.1 ± 11.6 nm. The presence of the gene coding for integrase implied the temperate lifestyle.

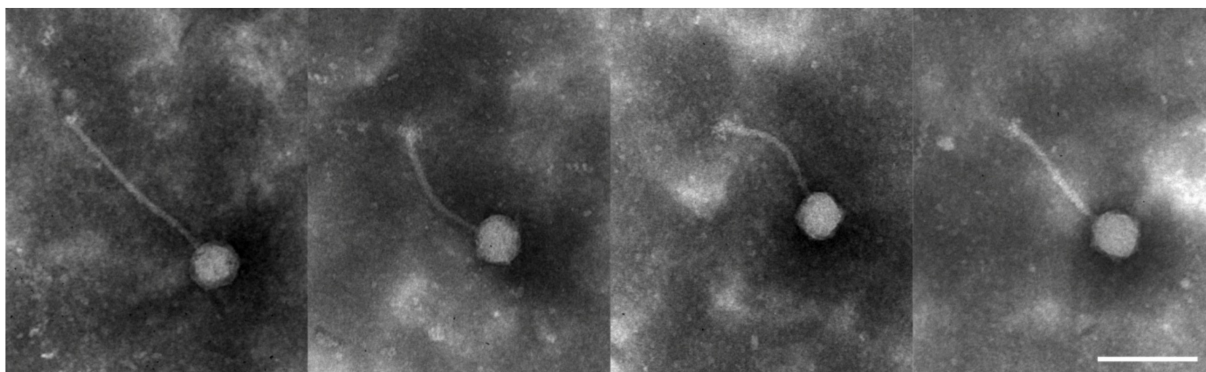


Figure 1. Transmission electron microscopy of Enterococcus phage VEsP-1. Scale bar, 100 nm.

As a result of DNA sequencing, 1,399,584 paired-end reads were obtained. The viral genomic sequence was assembled de novo into a single contig with $60\times$ coverage and was manually corrected. The length of the genome was 39,221 bp with the GC content at 35.1%. The genome contained 63 protein-coding genes and 1 tRNA gene.

3.2. Taxonomic Analysis

The closest match to the obtained whole genome sequence of VEsP-1 was the metagenomic sequence of bacteriophage cts681 (BK058556.1) with a 53% similarity to the VEsP-1 genome as determined by the multiplication of the percent identity by the percent coverage. The closest genome with known taxonomic position had a 14% similarity and belonged to the Enterococcus phage phiFL3B, a member of the genus *Phifelvirus* (GQ478087.1). Consequently, bacteriophage VEsP-1 could not be identified at the species or genus levels.

Aiming at elucidating the position of the VEsP-1 genome in the space of the related sequences, we performed a genome network analysis. The preliminary vContact2 analysis was done with the results of the BLAST searches with the VEsP-1 whole genome sequence, tail protein, and terminase large subunit sequences. The analysis placed the VEsP-1 genome in a cloud of the phage sequences. To increase the sample size, the BLAST searches were repeated with the sequences of the topologically close Streptococcus phage P7954 (*Brussowvirus*) and Brevibacillus phage Powder (*Jimmervirus*, not shown). The two genomes shared homologous sequences coding for a terminase small subunit and portal protein with VEsP-1. During the sequential BLAST searches, we obtained an increase in the sample size from 1214 to 6739 sequences (at a factor of 5.6) moving from VEsP-1 WGS to the general sample.

All nine sequences used in the BLAST searches to prepare the general sample contributed a noticeable number of unique matches (Figure 2).

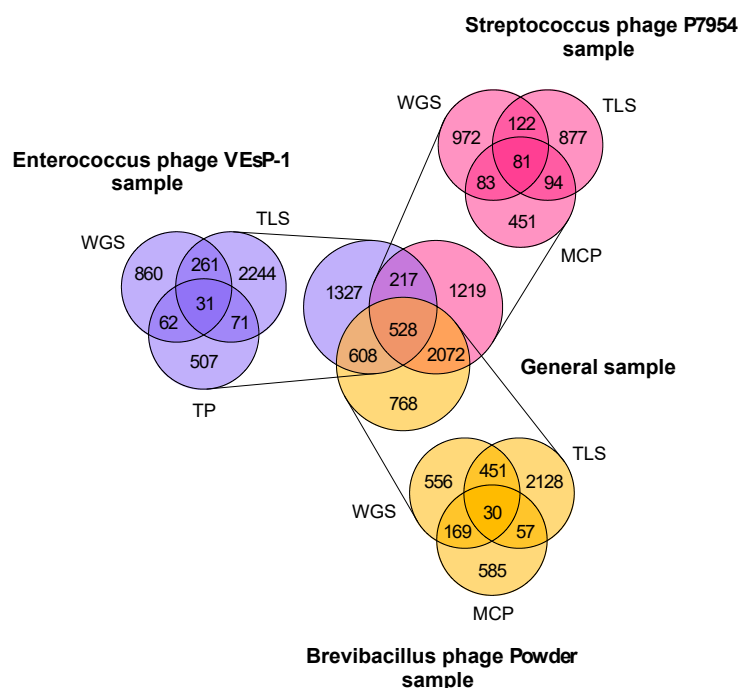


Figure 2. Sequence sample preparation. The numbers of whole genome sequences retrieved in nine BLAST searches with bacteriophage whole genome sequences (WGS), amino acid sequences of major capsid protein (MCP), terminase large subunit (TLS), and tail protein (TP) are shown. The general sample was obtained by pooling the results of the searches. The diagrams were visualized in VennDiagram 1.7.3 package [45] for R 4.1.2 [46].

The vContact2 network calculated from our general sample of genomes ($n = 6739$) contained 6711 nodes; the main interconnected part counted 6359 nodes or 94% genomes of the original sample, including representatives of 112 bacteriophage genera. The only revealed group of genomes with ICTV-accepted family names belonged to *Herelleviridae* (Figure 3). The occasional sequences with family annotations in the main part of the network belonged to the genera *Eclunavirus*, *Webervirus* (*Drexelvirus*), one sequence for each genus), *Cellubavirus* (*Assiduviridae*, one sequence), and *Baltivirus* (*Pachyviridae*, two sequences). In the smaller fragmented parts of the network, there were two *Pleakleyvirus* genomes (*Zierdtviridae*) and one *Nohivirus* genome (*Autographiviridae*). In the results of the vContact2 cluster analysis, the VEsP-1-containing group comprised nine members (Data S1). The genus confidence score for the respective subcluster was calculated at 0.8889 (Data S2). To find whether it is possible to interpret this group from a taxonomic point of view, we performed a phylogenomic analysis with VICTOR.

To prepare the VICTOR dataset, we selected the genomes (1) belonging to bacteriophages from the VEsP-1 cluster obtained during vContact2 analysis; (2) coding for the related tail protein sequences, forming a branch with a VEsP-1 cluster on a preliminary TP tree, together with the closest outgroup (Figure S1); and (3) belonging to genera, topologically close to VEsP-1 on the vContact2 phylogenetic network and at the same time containing at least one homologous sequence with VEsP-1 cluster members. The following bacteriophages were initially selected as topologically close but did not share homologous genes with VEsP-1 cluster members: *Colneyvirus*, *Fernvirus*, *Lacnuvirus*, *Lilyvirus*, *Lubbockvirus*, *Moineauvirus*, *Phifelvirus*, *Spizizenvirus*, and *Sukhumvitvirus*.

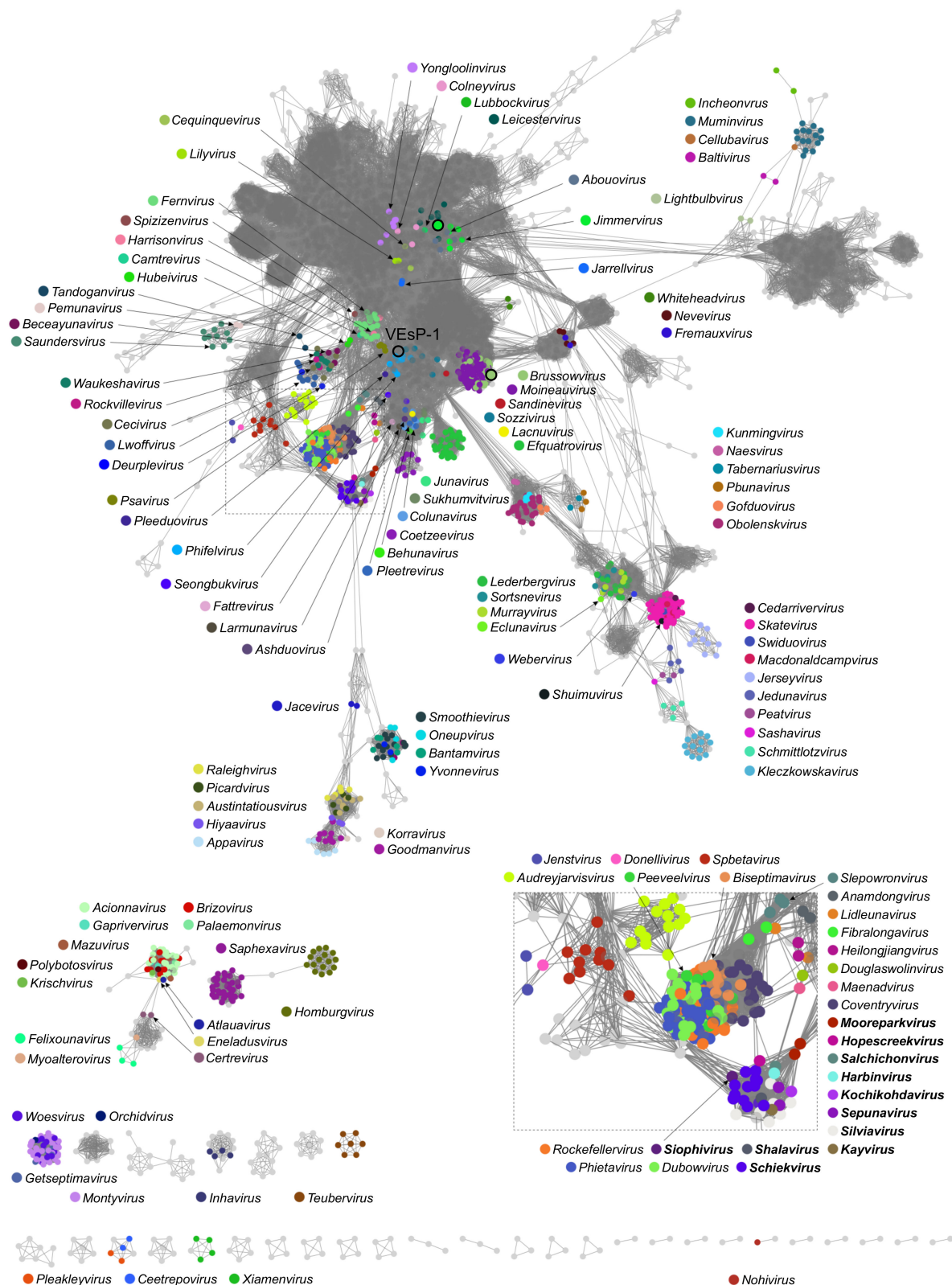


Figure 3. Network analysis of bacteriophage genomes. The inset shows a part of the network containing *Herelleviridae* members (given in boldface). The circles with outer boundaries show starting points for BLAST searches: Enterococcus phage VEsP-1, Streptococcus phage P7954 (*Brussowvirus*), and Brevibacillus phage Powder (*Jimmervirus*). The network was calculated in vContact2 0.11.3 and visualized with the use of Cytoscape 3.9.1 program.

The resulting dataset for VICTOR included phage sequences sharing at least one homologous sequence with the VEsP-1 cluster. The hosts of the analyzed bacteriophages belonged to 11 Firmicutes genera, including two genera from the order Eubacteriales (*Clostridioides* and *Clostridium*), four genera from the order Bacillales (*Bacillus*, *Brevibacillus*, *Listeria*, and *Paenibacillus*), and five genera from the order Lactobacillales (*Enterococcus*, *Lactobacillus*, *Leuconostoc*, *Oenococcus*, and *Streptococcus*).

The genome-BLAST distance phylogenomic trees inferred using the formulas D0, D4, and D6 had an average support of 65%, 8%, and 70%, respectively. At the genus level, the OPTSIL clustering yielded eighteen (D0), fourteen (D4), and eighteen (D6) taxonomic hypotheses. Since the D4 tree had a low support and the supposed division of the dataset into genera with D0 and D6 formulas was identical, we considered only the D0 results (Figure 4). Ten of the 15 genus hypotheses matched the current ICTV-accepted genera. *Anamdongvirus*, *Jarrellvirus*, *Pleeduovirus*, and *Seongbukvirus* were represented by one sequence each. For the remaining genera reproduced in the OPTSIL analysis, we calculated intergenomic similarity values: *Harrisonvirus*, 50–98%; *Leicestervirus*, 45–100%; *Yongloolinivirus*, 49–87%; *Sozzivirus*, 79–82%; *Cequinguevirus*, 78–88%; and *Brussowvirus*, 54–88%. Here, four ICTV-accepted genera united viral genomes with a sequence similarity below the canonical 70% threshold. The *Lactococcus* phage BK5-T (*Sandinevirus*) and *Lactococcus* phage bIL285 (not classified) belonged to the same hypothetical genus. The remaining four OPTSIL genus hypotheses united 2–4 accepted genera. In all cases, the genera were monophyletic and in every accepted genus, the members belonged to one individual genus hypothesis. The VEsP-1 cluster genomes formed a highly supported monophyletic clade and belonged to an independent genus hypothesis, providing further evidence for a description of a taxonomic group (Figure 4). The nucleotide sequence similarity between the nine representative genomes was 17–77%.

3.3. The Proposed Genus “Vespunovirus”

The nine “Vespunovirus” genomes belonged mainly to uncharacterized bacteriophages found in metagenomic studies [47]. Since all VEsP-1-related viruses originated from fecal samples, VEsP-1 itself probably has a fecal origin. The genomes ranged in size from 35,370 to 42,320 bp and coded for 51–66 proteins (Table 1). The capsid structure was uniform, with T-numbers predicted at 7.0. The GC content varied from 33.5 to 37.9%. The group was monophyletic on a phylogenetic tree calculated from major capsid protein sequences (Figure S2). The CoreGenes analysis revealed 16 conservative coding sequences, including the portal protein gene in the DNA packaging module, the gene coding for integrase, and 14 genes in the structural module (Figure 5).

In “Vespunovirus”, several different terminase families were represented (Table 1). Six representatives possessed the XtmB super family terminase, sharing a 74–82% similarity with the sequence YP_004301277.1 of the *Brochothrix* phage NF5, suggestive of the P22-like headful packaging mechanism [48]. The metagenomic phages ct40X5 and ctCNj1 had COG5362 super family terminases homologous to the sequence NP_112694.1 of the *Lactococcus* phage TP901-1. In these two cases, a pac site-dependent packaging mechanism may be expected [49]. The TLS of bacteriophage ctyrd11 was 53% similar to the entry YP_009195762.1 from the genome of the *Clostridium* phage phiCD505 and belonged to the DEAD-like helicases superfamily. For the *Clostridium* phage phiCD505, a T4-like packaging mechanism was predicted [50].

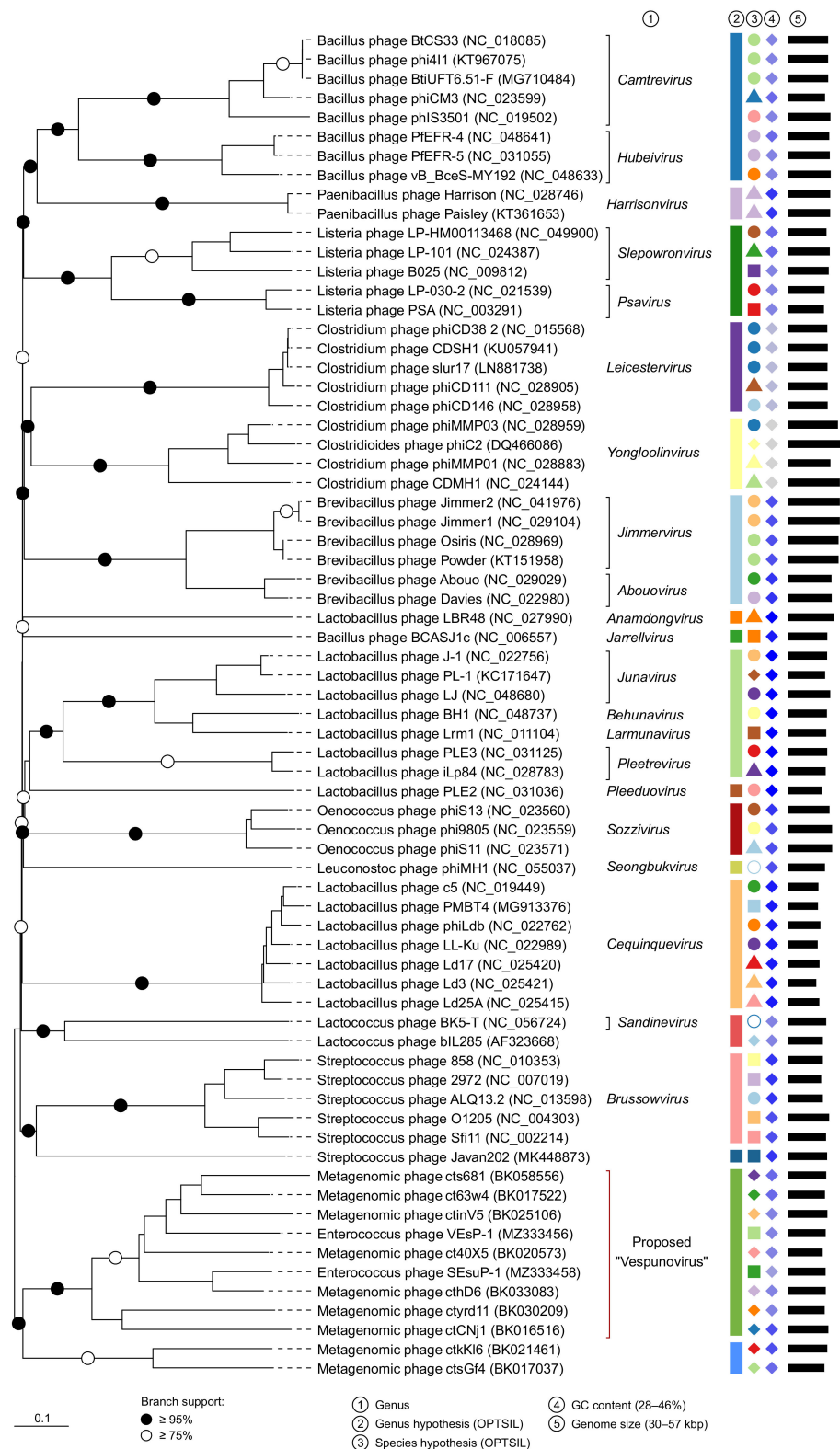


Figure 4. VICTOR phylogenomic tree. The group of VEsP-1-related bacteriophages predicted in vContact2 analysis belong to the same “Vespunovirus” genus hypothesis.

Table 1. The genomes belonging to the proposed genus “Vespunovirus”.

Phage Name	Sample Location	Sample Type	GenBank Accession	Length, bp	ORFs	tRNAs	GC Content, %	Terminase Large Subunit	Top Bacterial BLAST Hit; Query Cover/Identity, %
cts681	ND *	Human feces	BK058556	40,006	63	2	35.4	XtmB	<i>Enterococcus faecalis</i> ; 71/99
ct63w4	Estonia	— “ —	BK017522	38,654	59	2	35.4	XtmB	<i>E. faecalis</i> ; 72/96
ctinV5	Denmark	— “ —	BK025106	41,201	66	3	35.0	XtmB	<i>E. faecalis</i> ; 69/95
VEsP-1	Vietnam	River water	MZ333456	39,221	63	1	35.1	XtmB	<i>E. faecalis</i> ; 48/95
SEsP-1	Russia	Human feces	MZ333458	39,319	51	2	33.5	XtmB	<i>E. faecalis</i> ; 96/100
cthD61	USA	— “ —	BK033083	39,498	54	1	34.1	XtmB	<i>E. faecalis</i> ; 99/100
ct40X5	Russia	— “ —	BK020573	35,370	52	2	33.9	COG5362	<i>E. faecalis</i> ; 100/100
ctyrd11	Ethiopia	— “ —	BK030209	38,401	52	0	33.8	DEXDc	<i>E. hirae</i> ; 83/98
ctCNj1	USA	— “ —	BK016516	42,320	66	1	37.9	COG5362 **	<i>E. avium</i> ; 23/100

* ND, no data; ** in the genome, there is one terminase-coding gene.

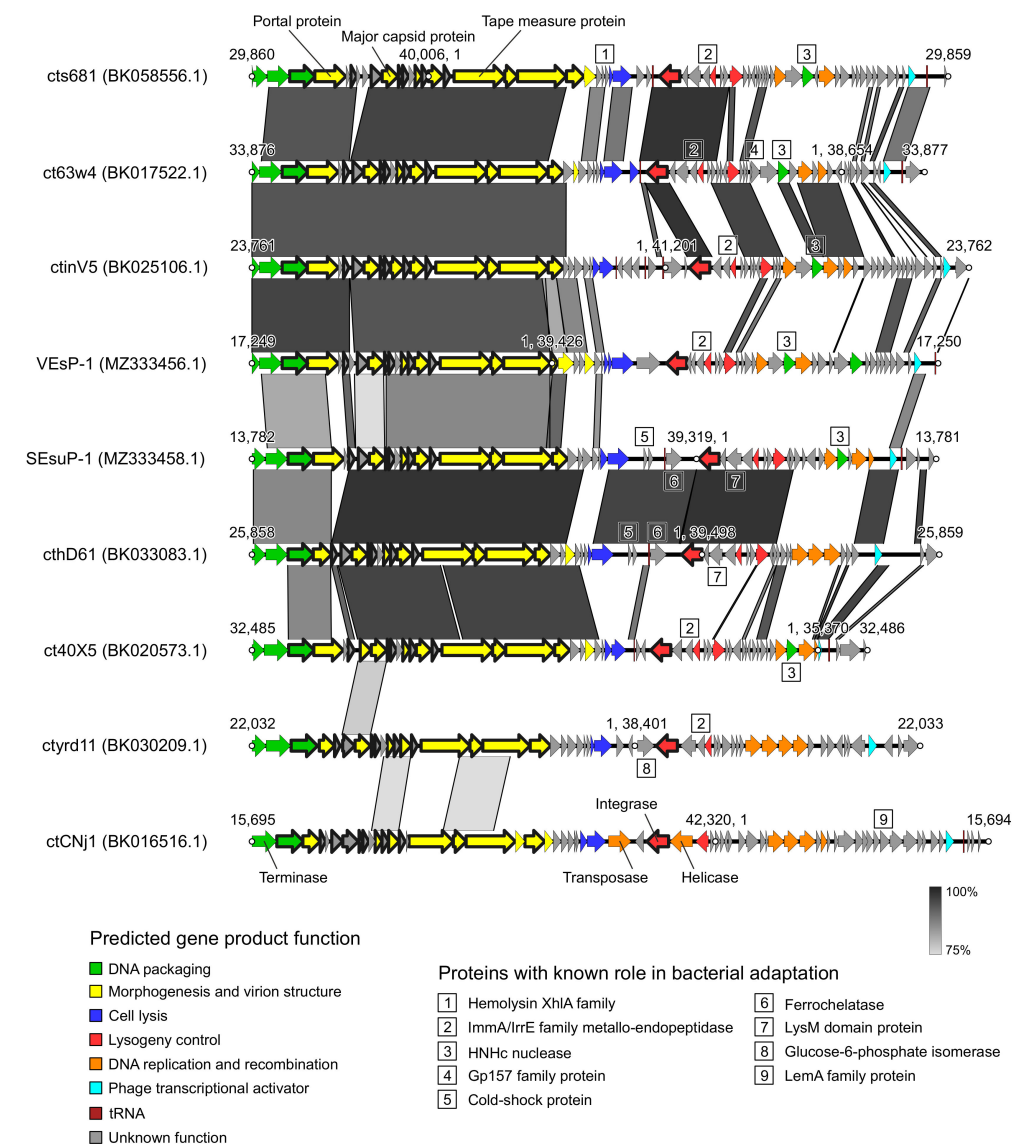


Figure 5. Genome structures in the proposed “Vespunovirus” genus. The conserved protein coding sequences are marked with a thick outline. The shades of grey show nucleotide sequence identity. The image was prepared using Easyfig 2.2.5 [51] on the basis of Prokka 1.14.6 annotations.

The “Vespunovirus” genomes lacked RNA polymerase. There was some degree of variability in the genome structures. For example, the ctCNj1 genome coded for one

terminase subunit unlike other “Vespunovirus” genomes with two TER subunits. In the ctCNj1 genome, the putative DNA replication and recombination module genes alternated with lysogeny control genes. In the cthD61 genome, the DNA replication and recombination module was not interrupted by HNHc endonuclease, unlike the closely related SEsuP-1 and ct40X5 genomes (Figure 5).

Analyzing the relationships of what will later become closely related but separate genera *Slepovronvirus* and *Psavirus*, Denes et al. compared the lengths of tail tape measure proteins [52]. In our cluster of VEsP-1-related viruses, the genome of the bacteriophage ctCNj1 encoded the tape measure protein of 849 amino acids. In the metagenomic phage ctyrd11, the protein consisted of 938 residues and in other “Vespunovirus” members, the length of the protein was 965–974 amino acids. The shortest tape measure protein counted 87% amino acids of the longest one. In *Slepovronvirus*, we found the same ratio of 0.87 between the lengths of the tail tape measure protein of *Listeria* phage LP-HM00113468 (1420 amino acids) and the protein of the *Listeria* phage B025 (1640 amino acids). In *Psavirus*, the *Listeria* phage LP-030-2 and *Listeria* phage PSA had tail tape measure proteins of the same length of 1026 amino acids.

Using the ABRicate tool, we did not find antibiotic resistance determinants and bacterial virulence factors in the “Vespunovirus” genomes. However, they contained a number of genes potentially contributing to a bacterial adaptation. The most common genes of this kind coded for the ImmA/IrrE family metallo-endopeptidases and putative HNHc nucleases, each found in six genomes. In *Bacillus*, ImmA-dependent proteolysis of ImmR repressors is considered to be a conserved mechanism for regulating a horizontal gene transfer [53]. In *Geobacillus*, the ImmA/IrrE family protein is a part of the toxin-antitoxin ToxN/AbiQ system [54]. HNHc nuclease domain-containing proteins may contribute to a bacterial virulence by diminishing the reactive oxygen species, and also by participating in the DNA repair [55]. Each other potential bacterial adaptation factor occurred in one or two genomes. The XhlA protein is a bacterial cell surface-associated hemolysin. It is capable of lysing rabbit and horse erythrocytes [56]. The Gp157 family proteins may contribute to a bacterial resistance to the bacteriophages [57]. In *Escherichia coli*, the cold shock proteins of the CspA family are involved in a cold-shock acclimation and are probably important for growth under optimal conditions [58]. Ferrochelatase catalyzes the insertion of the ferrous form of iron into protoporphyrin IX in the heme synthesis pathway. The LysM domain containing the proteins of *Enterococcus faecium* have unknown cellular functions. However, a transcriptomic analysis of *E. faecium* in an animal infection model revealed a significant increase in the number of corresponding transcripts, suggesting a role in the infection process [59]. Glucose-6-phosphate dehydrogenase is a part of the SoxR oxidative stress regulon of *E. coli* [60]. The LemA protein family was characterized in the genus *Pseudomonas*, where its members were identified as transmembrane histidine protein kinase sensor-regulators, involved in the formation of lesions in the host [61].

The revealed conservation in the structural module of the “Vespunovirus” genomes led to a question about the immunological properties of its proteins. To answer this question and provide a basis for a discussion on the significance of bacteriophage taxonomy at the genus level, we performed a computational prediction of immunogenicity in three related genera of bacteriophages sharing a conserved integrase sequence (Figure 6). In all three genera, there were confidently predicted immunogenic structural proteins. There was no clear difference in the overall distribution of the predicted immunogens according to the genus. However, unlike the major capsid proteins of *Sozzivirus* and *Psavirus* isolates, the MCPs of the “Vespunovirus” members apparently were not immunogenic.

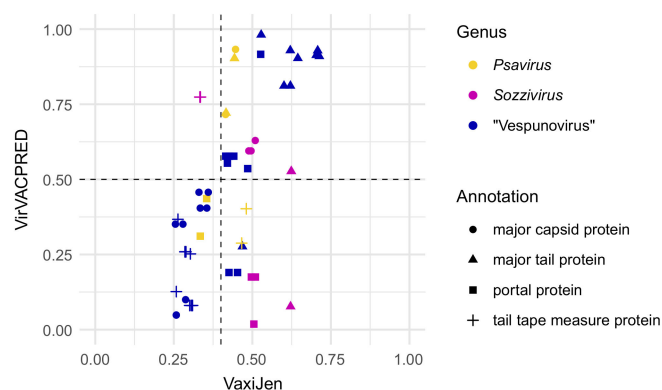


Figure 6. Probability of immunogenic potential of bacteriophage structural proteins in “Vespunovirus”, *Psavirus*, and *Sozzivirus* members. The thresholds for probability values are shown as dashed lines. The proteins were homologous within the genera, but not between the genera. The plot was visualized in ggplot2 3.3.5 package [62] for R 4.1.2 [46].

4. Discussion

Until recently, the demarcation of species and genus-level groups in bacteriophage taxonomy relied on nucleotide sequence similarity thresholds and molecular phylogenetic analysis. The bacteriophages sharing at least a 95% identity between the whole genome sequences were classified as the same species. The threshold for the genera was set at 70% [63]. Making the use of strict thresholds problematic, the genomes of most temperate bacteriophages have regions of a high similarity within the pairs of otherwise very dissimilar genomes, resulting in reticulate patterns of evolution [64–66]. Consequently, the current ICTV guidelines suggest the use of group-specific demarcation criteria rather than rigid similarity thresholds (<https://ictv.global/taxonomy/about>, accessed on 13 September 2022) [67].

Since the applicability of species rank to mosaic bacteriophages has been put under question [20,68], we refrain from proposing a species for the described bacteriophage VEsP-1. The revealed group of VEsP-1-related viruses shows a remarkable mixture of traits usually seen at different levels of phage taxonomy. On the one hand, clustering in vContact2 and VICTOR analyses is definitely evidence of a genus group. The nine viruses are highly likely to share hosts belonging to the same genus *Enterococcus*. The degree of differences in the lengths of the tape measure proteins support the genus status of the cluster. Additionally, the group was monophyletic in a signature gene phylogenetic analysis. On the other hand, three different packaging mechanisms as well as a slightly varying genomic organization favor a description of a family-level taxon. The number of 16 shared genes may also be interpreted as a family-level trait when compared to 31 shared genes in *Wbetavirus* [69]. Due to the limited number of representatives and the apparently minor significance of the proposed taxon, we suggest a grouping of the genus level. Though the intergenomic similarities between “Vespunovirus” members are low, there are at least four related ICTV-accepted genera with minimum nucleotide sequence similarity values of the genomes below the 70% threshold. Nearly all conserved in the “Vespunovirus” genes were structural, as it was earlier found in *Brussowvirus* and *Moineauvirus* members [70,71]. Therefore, we speculate the mentioned viral genera may have a significance for immunological studies.

Our two-step pooling approach to preparing a nucleotide sequence sample allowed us to increase the number of analyzed sequences by a factor of 5.6. The vContact2 analysis presented here predicted statistically significant similarities with at least one another genome in 94% of the genomes of our general sample. Therefore, with the mosaic viruses in question, the pooling results of several BLAST searches may seem reasonable.

Through the sequential steps of a significant similarity between the neighboring genomes, our vContact2 network connected very dissimilar phages with the hosts from

at least three different orders of Firmicutes. The obtained network depicts otherwise elusive relationships between bacteriophages sharing no homologous genes and possessing different genome organizations and can be thought of as a kind of substitute to higher taxonomic ranks in the studied part of viral dark matter. Probably, the meaning behind this would be in organizing the inventory of existing taxonomic names. It remains to be seen whether functional traits can be ascribed to phage groups of this magnitude.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v15010216/s1>, Figure S1: Maximum likelihood phylogenetic tree inferred from tail protein sequences; Figure S2: Maximum likelihood phylogenetic tree inferred from major capsid protein sequences; Table S1: Host range of Enterococcus phage VEsP-1; Data S1: Results of vContact2 cluster analysis; Data S2: vContact2 Genome_by_genome_overview file.

Author Contributions: Conceptualization, I.M.P.; formal analysis, I.M.P.; investigation, P.V.T., D.O.D. and A.N.G.; writing—original draft preparation, I.M.P.; writing—review and editing, I.M.P., P.V.T., D.V.A., A.E.G. and V.V.Z.; project administration, A.E.G., D.V.A. and A.V.D.; funding acquisition, A.E.G. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the Ministry of Science and Higher Education of the Russian Federation, no. 075-15-2022-302, Federal State Budgetary Scientific Institution “Institute of Experimental Medicine” (FSBSI “IEM”) Scientific and Educational Center “Molecular bases of interaction of microorganisms and human” of the World-Class Research Center “Center for Personalized Medicine” at FSBSI “IEM”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The whole genome sequence of the Enterococcus phage VEsP-1 was deposited in the NCBI Nucleotide database under the GenBank/ENA/DDBJ accession number MZ333456.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Sausset, R.; Petit, M.A.; Gaboriau-Routhiau, V.; De Paepe, M. New insights into intestinal phages. *Mucosal Immunol.* **2020**, *13*, 205–215. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Minot, S.; Bryson, A.; Chehoud, C.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. Rapid evolution of the human gut virome. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12450–12455. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Davies, E.V.; Winstanley, C.; Fothergill, J.L.; James, C.E. The role of temperate bacteriophages in bacterial infection. *FEMS Microbiol. Lett.* **2016**, *363*, fnw015. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Johnson, C.N.; Sheriff, E.K.; Duerkop, B.A.; Chatterjee, A. Let me upgrade you: Impact of mobile genetic elements on enterococcal adaptation and evolution. *J. Bacteriol.* **2021**, *203*, e0017721. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S.A.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. The human gut virome: Inter-individual variation and dynamic response to diet. *Genome Res.* **2011**, *21*, 1616–1625. [\[CrossRef\]](#)
6. Liang, G.; Conrad, M.A.; Kelsen, J.R.; Kessler, L.R.; Breton, J.; Albenberg, L.G.; Marakos, S.; Galgano, A.; Devas, N.; Erlichman, J.; et al. Dynamics of the stool virome in very early-onset inflammatory bowel disease. *J. Crohns Colitis* **2020**, *14*, 1600–1610. [\[CrossRef\]](#)
7. Clooney, A.G.; Sutton, T.D.S.; Shkoporov, A.N.; Holohan, R.K.; Daly, K.M.; O’Regan, O.; Ryan, F.J.; Draper, L.A.; Plevy, S.E.; Ross, R.P.; et al. Whole-virome analysis sheds light on viral dark matter in inflammatory bowel disease. *Cell Host Microbe* **2019**, *26*, 764–778.e5. [\[CrossRef\]](#)
8. Lang, S.; Demir, M.; Martin, A.; Jiang, L.; Zhang, X.; Duan, Y.; Gao, B.; Wisplinghoff, H.; Kasper, P.; Roderburg, C.; et al. Intestinal virome signature associated with severity of nonalcoholic fatty liver disease. *Gastroenterology* **2020**, *159*, 1839–1852. [\[CrossRef\]](#)
9. Jiang, L.; Lang, S.; Duan, Y.; Zhang, X.; Gao, B.; Chopyk, J.; Schwanemann, L.K.; Ventura-Cots, M.; Bataller, R.; Bosques-Padilla, F.; et al. Intestinal virome in patients with alcoholic hepatitis. *Hepatology* **2020**, *72*, 2182–2196. [\[CrossRef\]](#)
10. Nakatsu, G.; Zhou, H.; Wu, W.K.K.; Wong, S.H.; Coker, O.O.; Dai, Z.; Li, X.; Szeto, C.-H.; Sugimura, N.; Lam, T.Y.-T.; et al. Alterations in enteric virome are associated with colorectal cancer and survival outcomes. *Gastroenterology* **2018**, *155*, 529–541.e5. [\[CrossRef\]](#)
11. Aggarwala, V.; Liang, G.; Bushman, F.D. Viral communities of the human gut: Metagenomic analysis of composition and dynamics. *Mob. DNA* **2017**, *8*, 12. [\[CrossRef\]](#)

12. Fitzgerald, C.B.; Shkoporov, A.N.; Upadrastra, A.; Khokhlova, E.V.; Ross, R.P.; Hill, C. Probing the “dark matter” of the human gut phageome: Culture assisted metagenomics enables rapid discovery and host-linking for novel bacteriophages. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 616918. [CrossRef]
13. Shkoporov, A.N.; Hill, C. Bacteriophages of the human gut: The “known unknown” of the microbiome. *Cell Host Microbe* **2019**, *25*, 195–209. [CrossRef]
14. Deboutte, W.; Beller, L.; Yinda, C.K.; Maes, P.; de Graaf, D.C.; Matthijnsens, J. Honey-bee-associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 10511–10519. [CrossRef]
15. Tkachev, P.V.; Pchelin, I.M.; Azarov, D.V.; Gorshkov, A.N.; Shamova, O.V.; Dmitriev, A.V.; Goncharov, A.E. Two novel lytic bacteriophages infecting *Enterococcus* spp. are promising candidates for targeted antibacterial therapy. *Viruses* **2022**, *14*, 831. [CrossRef]
16. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef]
17. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.* **2008**, *9*, 75. [CrossRef]
18. Overbeek, R.; Olson, R.; Pusch, G.D.; Olsen, G.J.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Parrello, B.; Shukla, M.; et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206–D214. [CrossRef]
19. Brettin, T.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Olsen, G.J.; Olson, R.; Overbeek, R.; Parrello, B.; Pusch, G.D.; et al. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **2015**, *5*, 8365. [CrossRef]
20. Bolduc, B.; Jang, H.B.; Doulcier, G.; You, Z.Q.; Roux, S.; Sullivan, M.B. vConTACT: An iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. *PeerJ.* **2017**, *5*, e3243. [CrossRef]
21. Buchfink, B.; Reuter, K.; Drost, H.G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **2021**, *18*, 366–368. [CrossRef] [PubMed]
22. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]
23. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef] [PubMed]
24. Meier-Kolthoff, J.P.; Göker, M. VICTOR: Genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* **2017**, *33*, 3396–3404. [CrossRef] [PubMed]
25. Meier-Kolthoff, J.P.; Auch, A.F.; Klenk, H.-P.; Göker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **2013**, *14*, 60. [CrossRef]
26. Lefort, V.; Desper, R.; Gascuel, O. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* **2015**, *32*, 2798–2800. [CrossRef]
27. Göker, M.; García-Blázquez, G.; Voglmayr, H.; Tellería, M.T.; Martín, M.P. Molecular taxonomy of phytopathogenic fungi: A case study in *Peronospora*. *PLoS ONE* **2009**, *4*, e6319. [CrossRef]
28. Meier-Kolthoff, J.P.; Hahnke, R.L.; Petersen, J.; Scheuner, C.; Michael, V.; Fiebig, A.; Rohde, C.; Rohde, M.; Fartmann, B.; Goodwin, L.A.; et al. Complete genome sequence of DSM 30083T, the type strain (U5/41T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand. Genomic Sci.* **2014**, *9*, 2. [CrossRef]
29. Moraru, C.; Varsani, A.; Kropinski, A.M. VIRIDIC-A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses* **2020**, *12*, 1268. [CrossRef]
30. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate Maximum-Likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef]
31. Gouy, M.; Guindon, S.; Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **2010**, *27*, 221–224. [CrossRef]
32. Lemoine, F.; Domelevo Entfellner, J.-B.; Wilkinson, E.; Correia, D.; Dávila Felipe, M.; De Oliveira, T.; Gascuel, O. Renewing Felsenstein’s phylogenetic bootstrap in the era of big data. *Nature* **2018**, *556*, 452–456. [CrossRef]
33. Contreras-Moreira, B.; Vinuesa, P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.* **2013**, *79*, 7696–7701. [CrossRef]
34. Lee, D.Y.; Bartels, C.; McNair, K.; Edwards, R.A.; Swairjo, M.A.; Luque, A. Predicting the capsid architecture of phages from metagenomic data. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 721–732. [CrossRef]
35. Seemann, T. ABRicate. Available online: <https://github.com/tseemann/abricate> (accessed on 22 August 2022).
36. Gupta, S.K.; Padmanabhan, B.R.; Diene, S.M.; Lopez-Rojas, R.; Kempf, M.; Landraud, L.; Rolain, J.M. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* **2014**, *58*, 212–220. [CrossRef]

37. Jia, B.; Raphenya, A.R.; Alcock, B.; Wagglechner, N.; Guo, P.; Tsang, K.K.; Lago, B.A.; Dave, B.M.; Pereira, S.; Sharma, A.N.; et al. CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **2017**, *45*, D566–D573. [\[CrossRef\]](#)
38. Doster, E.; Lakin, S.M.; Dean, C.J.; Wolfe, C.; Young, J.G.; Boucher, C.; Belk, K.E.; Noyes, N.R.; Morley, P.S. MEGARes 2.0: A database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. *Nucleic Acids Res.* **2020**, *48*, D561–D569. [\[CrossRef\]](#)
39. Feldgarden, M.; Brover, V.; Haft, D.H.; Prasad, A.B.; Slotta, D.J.; Tolstoy, I.; Tyson, G.H.; Zhao, S.; Hsu, C.H.; McDermott, P.F. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob. Agents Chemother.* **2019**, *63*, e00483–19. [\[CrossRef\]](#)
40. Zankari, E.; Hasman, H.; Cosentino, S.; Vestergaard, M.; Rasmussen, S.; Lund, O.; Aarestrup, F.M.; Larsen, M.V. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **2012**, *67*, 2640–2644. [\[CrossRef\]](#)
41. Chen, L.; Zheng, D.; Liu, B.; Yang, J.; Jin, Q. VFDB 2016: Hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* **2016**, *44*, D694–D697. [\[CrossRef\]](#)
42. Doytchinova, I.A.; Flower, D.R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **2007**, *8*, 4. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Zaharieva, N.; Dimitrov, I.; Flower, D.; Doytchinova, I. Immunogenicity prediction by VaxiJen: A ten year overview. *J. Proteom. Bioinform.* **2017**, *10*, 11. [\[CrossRef\]](#)
44. Herrera-Bravo, J.; Fariás, J.G.; Contreras, F.P.; Herrera-Belén, L.; Norambuena, J.A.; Beltrán, J.F. VirVACPRED: A web server for prediction of protective viral antigens. *Int. J. Pept. Res. Ther.* **2022**, *28*, 35. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Chen, H. VennDiagram: Generate High-Resolution Venn and Euler Plots. R Package Version 1.7.3. 2022. Available online: <https://CRAN.R-project.org/package=VennDiagram> (accessed on 1 July 2022).
46. R Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: <https://www.R-project.org/> (accessed on 8 November 2021).
47. Tisza, M.J.; Buck, C.B. A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2023202118. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Kilcher, S.; Loessner, M.J.; Klumpp, J. *Brochothrix thermosphacta* bacteriophages feature heterogeneous and highly mosaic genomes and utilize unique prophage insertion sites. *J. Bacteriol.* **2010**, *192*, 5441–5453. [\[CrossRef\]](#)
49. Brøndsted, L.; Ostergaard, S.; Pedersen, M.; Hammer, K.; Vogensen, F.K. Analysis of the complete DNA sequence of the temperate bacteriophage TP901-1: Evolution, structure, and genome organization of lactococcal bacteriophages. *Virology* **2001**, *283*, 93–109. [\[CrossRef\]](#)
50. Garneau, J.R.; Depardieu, F.; Fortier, L.C.; Bikard, D.; Monot, M. PhageTerm: A tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Sci. Rep.* **2017**, *7*, 8292. [\[CrossRef\]](#)
51. Sullivan, M.J.; Petty, N.K.; Beatson, S.A. Easyfig: A genome comparison visualizer. *Bioinformatics* **2011**, *27*, 1009–1010. [\[CrossRef\]](#)
52. Denes, T.; Vongkamjan, K.; Ackermann, H.W.; Moreno Switt, A.I.; Wiedmann, M.; den Bakker, H.C. Comparative genomic and morphological analyses of *Listeria* phages isolated from farm environments. *Appl. Environ. Microbiol.* **2014**, *80*, 4616–4625. [\[CrossRef\]](#)
53. Bose, B.; Auchtung, J.M.; Lee, C.A.; Grossman, A.D. A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol. Microbiol.* **2008**, *70*, 570–582. [\[CrossRef\]](#)
54. Alkhalili, R.N.; Wallenius, J.; Canbäck, B. Towards exploring toxin-antitoxin systems in *Geobacillus*: A screen for type II toxin-antitoxin system families in a thermophilic genus. *Int. J. Mol. Sci.* **2019**, *20*, 5869. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Ortega, A.P.; Villagra, N.A.; Urrutia, I.M.; Valenzuela, L.M.; Talamilla-Espinoza, A.; Hidalgo, A.A.; Rodas, P.I.; Gil, F.; Calderón, I.L.; Paredes-Sabja, D.; et al. Lose to win: *marT* pseudogenization in *Salmonella enterica* serovar Typhi contributed to the *surV*-dependent survival to H₂O₂, and inside human macrophage-like cells. *Infect. Genet. Evol.* **2016**, *45*, 111–121. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Cowles, K.N.; Goodrich-Blair, H. Expression and activity of a *Xenorhabdus nematophila* haemolysin required for full virulence towards *Manduca sexta* insects. *Cell Microbiol.* **2005**, *7*, 209–219. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Da Silva Duarte, V.; Giaretta, S.; Campanaro, S.; Treu, L.; Armani, A.; Tarrah, A.; Oliveira de Paula, S.; Giacomini, A.; Corich, V. A cryptic non-inducible prophage confers phage-immunity on the *Streptococcus thermophilus* M17PTZA496. *Viruses* **2018**, *11*, 7. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Phadtare, S.; Alsina, J.; Inouye, M. Cold-shock response and cold-shock proteins. *Curr. Opin. Microbiol.* **1999**, *2*, 175–180. [\[CrossRef\]](#)
59. Cacaci, M.; Giraud, C.; Leger, L.; Torelli, R.; Martini, C.; Posteraro, B.; Palmieri, V.; Sanguinetti, M.; Bugli, F.; Hartke, A. Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in *Enterococcus faecium*. *Sci. Rep.* **2018**, *8*, 12412. [\[CrossRef\]](#)
60. Nunoshiba, T.; Demple, B. A cluster of constitutive mutations affecting the C-terminus of the redox-sensitive SoxR transcriptional activator. *Nucleic Acids Res.* **1994**, *22*, 2958–2962. [\[CrossRef\]](#)
61. Rich, J.J.; Kinscherf, T.G.; Kitten, T.; Willis, D.K. Genetic evidence that the *gacA* gene encodes the cognate response regulator for the *lemA* sensor in *Pseudomonas syringae*. *J. Bacteriol.* **1994**, *176*, 7468–7475. [\[CrossRef\]](#)
62. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; ISBN 978-3-319-24277-4.

63. Turner, D.; Kropinski, A.M.; Adriaenssens, E.M. A roadmap for genome-based phage taxonomy. *Viruses* **2021**, *13*, 506. [\[CrossRef\]](#)
64. Pope, W.H.; Bowman, C.A.; Russell, D.A.; Jacobs-Sera, D.; Asai, D.J.; Cresawn, S.G.; Jacobs, W.R.; Hendrix, R.W.; Lawrence, J.G.; Hatfull, G.F.; et al. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife* **2015**, *4*, e06416. [\[CrossRef\]](#)
65. Dion, M.B.; Oechslin, F.; Moineau, S. Phage diversity, genomics and phylogeny. *Nat. Rev. Microbiol.* **2020**, *18*, 125–138. [\[CrossRef\]](#)
66. Moura de Sousa, J.A.; Pfeifer, E.; Touchon, M.; Rocha, E.P.C. Causes and consequences of bacteriophage diversification via genetic exchanges across lifestyles and bacterial taxa. *Mol. Biol. Evol.* **2021**, *38*, 2497–2512. [\[CrossRef\]](#)
67. Lefkowitz, E.J.; Dempsey, D.M.; Hendrickson, R.C.; Orton, R.J.; Siddell, S.G.; Smith, D.B. Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.* **2018**, *46*, D708–D717. [\[CrossRef\]](#)
68. Morgan, G.J.; Pitts, W.B. Evolution without species: The case of mosaic bacteriophages. *Brit. J. Phil. Sci.* **2008**, *59*, 745–765. [\[CrossRef\]](#)
69. Nakonieczna, A.; Rutyna, P.; Fedorowicz, M.; Kwiatek, M.; Mizak, L.; Łobocka, M. Three novel bacteriophages, J5a, F16Ba, and z1a, specific for *Bacillus anthracis*, define a new clade of historical Wbeta phage relatives. *Viruses* **2022**, *14*, 213. [\[CrossRef\]](#)
70. Lucchini, S.; Desiere, F.; Brüssow, H. The structural gene module in *Streptococcus thermophilus* bacteriophage phi Sfi11 shows a hierarchy of relatedness to *Siphoviridae* from a wide range of bacterial hosts. *Virology* **1998**, *246*, 63–73. [\[CrossRef\]](#)
71. McDonnell, B.; Mahony, J.; Hanemaaijer, L.; Neve, H.; Noben, J.P.; Lugli, G.A.; Ventura, M.; Kouwen, T.R.; van Sinderen, D. Global survey and genome exploration of bacteriophages infecting the lactic acid bacterium *Streptococcus thermophilus*. *Front. Microbiol.* **2017**, *8*, 1754. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.