Diversity and Host Interactions among Virulent and Temperate Baltic Sea *Flavobacterium* Phages

Emelie Nilsson [1], Oliver W. Bayfield [2], Daniel Lundin [1], Alfred A. Antson [2] and Karin Holmfeldt [1,*]

[1]   Centre for Ecology and Evolution in Microbial Model Systems (EEMiS), Department of Biology and Environmental Science, Faculty of Health and Life Sciences, Linnaeus University, SE-39231 Kalmar, Sweden; emelie.nilsson@lnu.se (E.N.); daniel.lundin@lnu.se (D.L.)

[2]   York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, UK; oliver.bayfield@york.ac.uk (O.W.B.); fred.antson@york.ac.uk (A.A.A.)

*   Correspondence: karin.holmfeldt@lnu.se

check for updates

Abstract: Viruses in aquatic environments play a key role in microbial population dynamics and nutrient cycling. In particular, bacteria of the phylum *Bacteriodetes* are known to participate in recycling algal blooms. Studies of phage–host interactions involving this phylum are hence important to understand the processes shaping bacterial and viral communities in the ocean as well as nutrient cycling. In this study, we isolated and sequenced three strains of flavobacteria—LMO6, LMO9, LMO8—and 38 virulent phages infecting them. These phages represent 15 species, occupying three novel genera. Additionally, one temperate phage was induced from LMO6 and was found to be competent at infecting LMO9. Functions could be predicted for a limited number of phage genes, mainly representing roles in DNA replication and virus particle formation. No metabolic genes were detected. While the phages isolated on LMO8 could infect all three bacterial strains, the LMO6 and LMO9 phages could not infect LMO8. Of the phages isolated on LMO9, several showed a host-derived reduced efficiency of plating on LMO6, potentially due to differences in DNA methyltransferase genes. Overall, these phage–host systems contribute novel genetic information to our sequence databases and present valuable tools for the study of both virulent and temperate phages.

Keywords: genomics; aquatic; lytic; lysogeny; host range

1. Introduction

In the majority of aquatic environments, bacteriophages (phages) vastly outnumber their bacterial hosts [1] and have a significant impact on the ecosystem as a result [2]. The impact on the microbial community depends on a variety of factors, including the host range and efficiency of infection on different strains [3,4], the number and character of metabolic genes carried by the phage [5–7], the metabolic reprogramming of the host population [8–10], and the replication strategy of the phage [11,12]. Phages have been estimated to kill ~20% of the bacterial population in the ocean every day [2], thereby contributing to the available pool of nutrients for the microbial community and preventing their transport to higher trophic levels [13,14]. In addition to lysing their microbial hosts, temperate phages have the ability to integrate into the genome of their hosts and remain dormant until they are induced to undergo lytic replication. In this way, temperate phages stay protected during times of harsh conditions [15]. Temperate phages can also offer benefits to the host, such as homoimmunity that prevents further phage infections [11], toxin production [16], and changes to the host gene expression [17], all of which increases the fitness and competitiveness of the lysogenised bacteria. Whilst a large proportion of bacteria isolated to date appear to contain integrated prophages [18–20],

the patterns of lysogenic replication in aquatic environments do not show clear patterns. Large seasonal variations in the lytic-lysogenic switch have been observed, and it has been suggested that lysogenic replication is favoured under conditions of low bacterial abundance and productivity [21–23].

Bacteria of the phylum *Bacteroidetes* inhabit a wide range of environments, ranging from Antarctic soil, to Sahara sand dust [24], and the human gut microbiome [25]. In aquatic environments, *Bacteriodetes* constitutes a large proportion of the microbial community [26,27], where they play important roles in phytoplankton bloom recycling [28] and degradation of biopolymers [29]. So far, relatively few phages infecting *Bacteroidetes* have been characterised, for example, compared to phages infecting *Gammaproteobacteria*, *Actinobacteria*, and *Firmicutes* [30], and hence there is little known about the virus–host interactions influencing this major member of the ocean microbial community. Among isolated *Bacteroidetes* phages, the majority infect members of *Flavobacteriaceae*—the largest aquatic family within *Bacteroidetes* [31]. These phages originate from various aquatic environments [32–35]. They include a large collection of *Cellulophaga baltica* phages [36,37], which contains phages related to the abundant human gut crAssphage [38–40]. The *Cellulophaga* phages have proved to be an important model system for investigations of phage–host interactions, ranging from host range variations to transcriptomic analyses [4,8,10,41–43]. *Bacteroidetes* phages have also shown their potential as bio-control agents in aquatic environments [44,45], and a large diversity of both virulent and temperate phages infecting *Flavobacterium psychrophilum* have been isolated from fish farms [46–48]. As *F. psychrophilum* is a fish pathogen that causes large economic losses [49], the potential of using phage therapy to control this disease is of great importance. Overall, these comprehensive phage collections constitute an ideal model system for both the investigation of genetic diversity and phage–host interactions [48,50,51].

Given the importance of *Bacteroidetes* and their phage–host interactions in aquatic environments, we aim to isolate new *Flavobacterium* strains and phages. By isolating bacteria of different strains within the same genus, we wanted to explore the diversity within phages infecting them as well as the interactions between these phages and the different bacterial isolates.

2. Materials and Methods

2.1. Bacterial Isolation and DNA Extraction

Water for bacterial isolation was collected at 2 m depth at the Linnaeus Microbial Observatory (LMO; N 56° 55.8540′, E 17° 3.6420), situated 10 km east off the coast of Öland, Sweden, with a Ruttner sampler on 16 September 2015. Bacteria within the water were grown by spreading 100 μL Baltic Sea water onto Zobell agar plates (1 g yeast extract (Becton, Dickinson and Company (BD), Franklin Lakes, NJ, USA), 5 g bacto-peptone (BD), and 15 g of bacto-agar (BD) in 800 mL filtered (Whatman glass microfiber filer, GF/C, GE Healthcare, Chicago, IL, USA) Baltic Sea water and 200 mL Milli-Q water: this, and all other media used, were autoclaved at 121 °C for 20 min), from where colonies were picked and clean-streaked three times. DNA extraction of the genomic DNA of bacteria LMO6, LMO8, and LMO9 was conducted using the E.Z.N.A. Tissue DNA kit (Omega bio-tek, Norcross, GA, USA) according to the manufacturer's instructions. Briefly, cells were lysed, protease treated, and bound to a HiBind DNA Mini column. The bound DNA was washed with DNA wash buffer diluted with 100% ethanol before it was eluted in 10 mM Tris. The quality and quantity of the DNA were verified with Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA) and Qubit (high sensitivity DNA kit; Invitrogen, Life Technologies, Carlsbad, CA, USA), respectively, and sequenced at SciLife/NGI (see Section 2.7 "Sequencing of Viral and Bacterial Genomes").

2.2. Viral Isolation

Water for viral isolation was collected on 16 September 2015 at LMO. To remove bacteria and larger organisms, 10 L of water was prefiltered through 0.22 μm sterivex filters (Merck Millipore, Darmstad, Germany) before viral particles were concentrated using a 30 kDa tangential flow filter

(Merck Millipore). The resulting 100–150 mL was further concentrated using a 50 kDa Amicon Ultra centrifugal filter (Merck Millipore) to 13–30 mL.

The viral enriched water was used in plaque assays with the three bacteria isolated in this study, LMO6, LMO8 and LMO9, as hosts to discover viral infection. Bacteria were grown overnight in liquid Zobell medium (1 g yeast extract (BD) and 5 g bacto-peptone (BD) in 800 mL filtered Baltic Sea water and 200 mL Milli-Q water) with gentle agitation. The bacteria were mixed with the viral enriched water and molten top-agar (marine sodium magnesium (MSM) buffer: 450 mM NaCl (Sigma, St Loius, MO, USA), 50 mM MgSO$_4$ × 7H$_2$O (Merck Millipore) and 50 mM Trizma base (Sigma), pH 8; with 0.5% low melting point agarose (Thermo Fisher Scientific)). The mixture was spread evenly onto Zobell agar plates. The plates were incubated at room temperature (RT) for two days and plaque formation was monitored on a daily basis. On two of the bacteria, LMO6 and LMO9, 30 plaques were originally isolated for each bacterium, while only the five plaques that were formed on LMO8 were isolated. Phage isolates were purified by picking individual plaques with a sterile 100 µL pipette tip. The end of the pipette tip (containing the phages) was thereafter submerged in MSM to disperse the phages. The phages were then re-plated as described above and the procedure was repeated three times. Pure isolates were harvested by adding 5 mL MSM to fully lysed plates, the top agar-layer was shredded with an inoculation loop, and the plate incubated on a shaking table (40 rpm) for at least 30 min. The phages suspended in MSM were collected into a falcon tube, centrifuged for 10 min at 10,000× *g*, filtered through 0.2 µm syringe filter (BD) and stored at 4 °C.

2.3. Host Range Investigation

In order to determine the host range of the isolated phages, we used the three bacteria, LMO6, LMO8 and LMO9, as well as four other bacteria that shared 16S rRNA gene similarity with the previous three; Flavobacteriaceae bacterium BAL304 (accession: KM586862), Flavobacteriaceae bacterium BAL314 (accession: KM586887), Flavobacteriaceae bacterium BAL330 (accession: KM586875) and Flavobacteriaceae bacterium BAL346 (accession: KM586918). In order to produce plaque assays, they were grown in MLB (0.5 g Casein hydrolysate (Merck Millipore), 0.5 g bacto-peptone (BD), 0.5 g yeast extract (BD), 3 mL glycerol (Sigma) in 800 mL filtered Baltic Sea water and 200 mL Milli-Q water) overnight. They were grown in MLB instead of Zobell medium since the bacteria aggregated in Zobell medium but stayed more dispersed in MLB. Overnight culture (300 µL) was mixed with top-agar (3.5 mL) and dispersed on an agar plate, then a 100 µL pipette tip was dipped in the viral isolate and a line was drawn across the prepared agar plate according to the molten streaking for singles method [52]. This was repeated three times, after which a new pipette tip was used to draw three new lines intersecting the previous lines in an angle; this was repeated so that four sets of lines were drawn. When this method provided clear plaques, standard plaque assays were performed in order to count the efficiency of plating. The efficiency of plating was also performed with all phages on their original host of isolation. Plates were incubated for 24–48 h. The results were visualised with Tidyverse (version 1.2.1) [53] in R (version 3.5.1) [54] using RStudio (version 1.1.383) [55].

2.4. Prophage Induction Experiments with LMO6

LMO6 was grown in MLB until exponential growth was established, as measured with an optical density at 600 nm with a CO8000 cell density meter (WPA, Cambridge, UK). Thereafter, 25 mL of LMO6 was dispersed into a sterile petri dish (covering the bottom to approximately 1 mm depth) and subjected to UV irradiation with wavelength of 254 nm for 10 s by a handheld UV-lamp (Model UVGL-58 mineral light lamp, 6W, Upland, CA, USA) at a distance of 15 cm without the lid of the petri dish. In addition, 5 mL of LMO6 was dispersed in a similar manner, but not UV-irradiated, to act as control. After the treatments, the bacteria were diluted 10x in fresh MLB in sterile falcon tubes and kept in the dark with agitation for five hours and OD was measured every hour. After five hours, 2 mL of bacteria was filtered through a 0.22 µm syringe filter (Merck Millipore) to collect the supernatant containing free viral particles. The flow-through was used to enumerate viral particles by plaque

assays, as previously described, with LMO6, LMO8 and LMO9 as hosts. The plates were kept up to 36 h in order to verify plaque formation. To obtain enriched stocks of only the temperate phage, flow-through from the non-UV treated LMO6 was used to perform plaque assays on LMO9. Here, the flow-through was diluted in a 10-times dilution series and fully lysed plates were harvested as described above. This was repeated until a high-titre phage stock was received.

2.5. Morphological Examination Using Transmission Electron Microscopy (TEM)

TEM was conducted on high-titre phage lysate. Carbon-formvar coated copper grids (Agar Scientific Ltd., Stansted, UK) were plasma cleaned in a PELCO easiGlow for 60 s at 0.38 mbar (air) and 20 mA. Lysate (5 μL) was applied to a grid for 60 s then removed by wicking. The grid was washed with deionised water (5 μL) and stained with 2% *w/v* uranyl acetate (5 μL (Agar Scientific Ltd.)). Grids were imaged using an FEI Tecnai 12 G2 BioTWIN microscope with tungsten filament operating at 120 kV. Phage particles were measured using ImageJ (version 1.52a).

2.6. Viral DNA Extraction

DNA extraction from the phages was performed with the Wizard PCR DNA Purification Resin and Minicolumns (Promega, Madison, WI, USA). First, newly harvested, high-titre phages were added to 1 mL resin and thoroughly mixed before being pushed through a minicolumn. Second, the column was washed twice with 1 mL 80% isopropanol and residues of isopropanol were removed by centrifuging the column for 2 min at 10,000× *g*. Last, to elude the DNA, pre-heated (80 °C) 10 mM Tris (Trizma base (Sigma), pH 8) was added, the column was vortexed and then centrifuged at 10,000× *g* for 30 s. The DNA was verified with regards to quality and quantity with Nanodrop (Thermo Fisher Scientific) and Qubit (high sensitivity DNA kit; Invitrogen, Life Technologies), respectively.

2.7. Sequencing of Viral and Bacterial Genomes

Library preparation for the extracted DNA from both bacteria and phages was done with Nextera XT (Illumina Inc., San Diego, CA, USA) and sequencing was performed by SciLifeLab/NGI (Solna, Sweden) on a HiSeq 2500 (Illumina Inc.), which resulted in paired-end 125 bp reads (average 1.2 million reads per sample).

2.8. Assembly, Annotation, and Pan-Genome Analysis

Raw reads were trimmed to remove adapters and poor quality reads (Trimmomatic version 0.30, settings: -PE –threads 2 -phred33 ILLUMINACLIP:nextera_linkers.txt:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:30) [56]; quality was evaluated with FastQC [57]. Trimmed reads were then assembled into contigs with Spades (version 3.6.0, settings: –careful –t 8 –pe1-1 –pe1-2 -o) [58] and Abyss (version 1.3.5, settings: -np 16 –k61 –q3 –coverage-hist –s –o) [59], and the assembler providing best quality was chosen (Supplementary Table S1). CheckM was used to assess bacterial genome completeness [60]. Bacterial genomes were annotated with NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) [61] while open reading frames (ORFs) in the phage genomes were called with Phanotate (version 1.0.1; default settings) [62] and manually inspected. Viral ORFs were then annotated with Diamond (version 0.9.10, settings: blastp -p 5 -q -k 100 -e 0.001 –sensitive –tmpdir -f 6 -o) [63] against the NCBI non-redundant database (3 May 2019) and hmmsearch (version 3.2.1, settings: –tblout -E 1e-3 –cpu 1) [64] against the Pfam database (24 February 2017) [65]. A significant alignment had an e-value of less than 0.001. Core- and pan-genomes were calculated with Roary (version 3.6.2, settings: -s -i 70) [66]. The comparison of phage genomes based on nucleotide identity was performed with Gegenees (version 3.0.0, settings: fragment size of 500 and step size of 500, blastn) [67], which fragments the genomes and compares identity across the entire genomes. The similarities were then plotted with Tidyverse (version 1.2.1) [53] with R (version 3.5.1) [54] using RStudio (version 1.1.383) [55]. Bacterial genomes were compared with Pyani (version 0.2.7, settings:

average_nucleotide_identity.py -m ANIb) [68], and run through PHASTER [69] to identify potential prophage regions.

2.9. Phylogenetic Analysis

GTDBtk (version 0.3.2) [70] was used to place the bacterial genomes of LMO6, LMO8 and LMO9 in the Genome Taxonomy Database [71] tree, which was conducted through analysis of a set of marker genes, and calculate shared average nucleotide identity with the reference genomes. The tree was visualised with Dendroscope (version 3.5.9) [72].

To perform phylogenetic analysis of the phage genomes, a set of reference genomes were selected based on the criteria that they should share at least one gene with significant amino acid similarity (e-value less than 0.001) to any of the phages in this study. For the lytic phages, only other viral genomes were selected, while bacterial contigs with several genes with significant amino acid similarity to the temperate phage were also selected. The protein sequences of this reference set, containing 37 genomes, were then acquired through NCBI and phylogenetic analysis of the references and LMO-phages were conducted through VICTOR [73]. For bacterial genomes with potential prophage regions with similarity to the temperate phage, only the relevant, potential prophage region was used for the VICTOR analysis. VICTOR was performed using pairwise comparisons of the nucleotide sequences, with settings recommended for phages [73], through the Genome-BLAST Distance Phylogeny (GBDP) method [74]. The distances were used with FastME to infer a tree with branch support [75] for the D0 formula. Also, VICTOR estimated taxon boundaries (species, genus, family) with the OPTSIL program [76] with recommended clustering thresholds [73].

2.10. Mapping of Environmental Reads to the Phage Genomes and Host Bacteria

Reads from viral metagenomes from LMO, sampled between 2012 and 2015 (BioProject accession number: PRJNA474405) [77], were mapped against the type phages for each genus (mumin9-1, pippi8-1, tant8-1, laban6-1) with bowtie2 (v2.3.5.1, settings: –ignore-quals –mp = 1,1 –np = 1 –rdg = 0,1 –rfg = 0,1 –score-min = L,0,-0.1) [78]. For a phage to be counted as present in the dataset, the genome needed to be covered by at least 75% of its length by reads with at least 90% sequence identity [79], where the first part was calculated with SAMtools [80] and BEDtools [81] and the second was acquired through the bowtie2 settings. Viral metagenomes were also gathered from publicly available datasets (Global Ocean Sampling (GOS): CAM_P_0000914, CAM_P_0000915, CAM_P_0001109, and Tara Ocean Viromes: ERR594353-4, ERR594356, ERR594358-9, ERR594361-2, ERR594364-5, ERR594368-70, ERR594374, ERR594376, ERR594378, ERR594384, ERR594391-3, ERR594395-6, ERR594398, ERR594400-1, ERR594403-4, ERR594406, ERR594410, ERR594412) and recruited against the genomes in the same manner, except for data from the GOS dataset where unpaired data were mapped (setting -f).

Reads from bacterial metagenomes from LMO [82] were mapped against the three bacterial strains LMO6, LMO8, and LMO9 (settings as above, except: –score-min = L,0,-0.03 for species identity, –score-min = L,0,-0.2 for genus identity). To assign a read to a particular bacterial strain, the read had to share at least 97% identity to the reference genome, while 80% identity was required to assign the read to a bacterial genus, and 60% of the genome should be covered. Also, the 16S rRNA genes from LMO6, LMO8 and LMO9 were compared with 16S amplicon sequencing data from LMO [77] (blastn) and the counts for the 100% identical amplicon sequencing variant (ASV) were retrieved from the dataset.

2.11. Mapping of Reads to Confirm Prophage Induction

Mapping of reads was performed in order to confirm that the temperate phage was induced prior to DNA extraction and sequencing, and not only sequenced as an integrated prophage within the bacterial genome. During phage DNA extraction, there might occur contamination from bacterial DNA, which, after sequencing, could result in low abundances of bacterial host sequences. Therefore, reads from one sequencing sample (well number 103 that contained tooticki6-1 and laban6-1, the lytic and temperate phage identified after sequencing, respectively) were mapped with bowtie2 (as above)

against the lytic phage genome (tooticki6-1) and the bacterial contig of LMO6 that contained the prophage (contig name: NODE_12_length_95344_cov_81.199_ID_11796, specific accession number: WIBH01000012.1). Coverage and depth were calculated as described above. The mapping against the bacterial contig was plotted with tidyverse with R (version 3.5.1) [54] using RStudio (version 1.1.383) [55].

2.12. Accession Numbers

All sequences in this study are deposited in bioproject PRJNA566314 at DDBJ/ENA/GenBank.

Bacterial strains, assembled as high-quality draft genomes, were deposited with the specific accession numbers WIBH00000000 for *Flavobacterium* sp. LMO6, WIBI00000000 for *Flavobacterium* sp. LMO9 and WIBJ00000000 for *Flavobacterium* sp. LMO8.

Viral genomes were deposited under the accession numbers MN812203–MN812241.

Reads from well 103 were deposited under the accession number SAMN13638723.

3. Results

3.1. Characteristics of Flavobacterium spp. Strains

Three bacterial strains (LMO6, LMO8, and LMO9) were isolated from the Baltic Sea in September 2015. All three had similar colony morphology, with a yellow to orange colour. LMO8 had colonies that were easily retrieved from the agar plates, whereas both LMO6 and LMO9 colonies were sticky and difficult to pick from the agar plates. While the 16S rRNA genes for the three strains were 99.7–100% identical at the nucleotide level, LMO6 and LMO9 shared more than 99% average nucleotide identity (ANI) to each other across 98.4–98.9% of the genomes while LMO8 shared 85% ANI with the other two across 60–64% of the genomes. Genome comparison and placement in the reference tree showed that all strains were placed within *Flavobacterium* and were most similar to *Flavobacterium sp000169355* (Figure 1): LMO6 and LMO9 with 96.8% ANI and LMO8 with 86.1% ANI. The bacterial isolates are therefore named *Flavobacterium* sp. strain LMO6, *Flavobacterium* sp. strain LMO8 and *Flavobacterium* sp. strain LMO9. Of the three bacterial strains isolated in this study, no prophage region was detected in LMO8, an incomplete prophage region of 8.9 kb was detected in LMO9, while LMO6 had an incomplete prophage region of 37.3 kb.
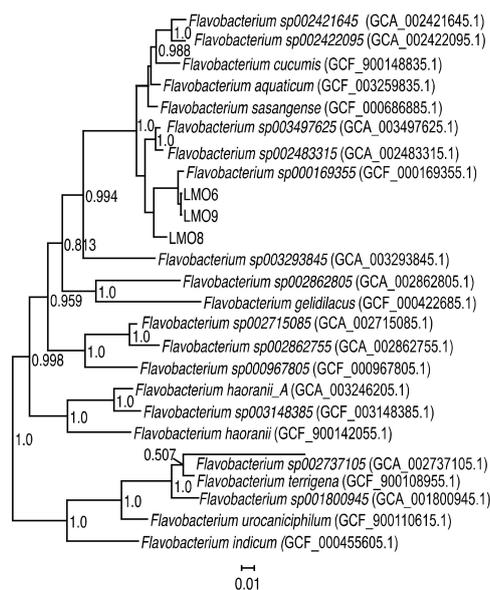

Figure 1. Phylogenomic placement of the three strains from this study—LMO6, LMO8, and LMO9—in the GTDB phylogeny [71]. The three strains were placed with GTDBtk [70] and here visualised together with their closest *Flavobacterium* spp. relatives. Numbers at nodes are bootstrap values (values below 0.5 are not shown).

3.2. Genomic Characteristics of Lytic Phages

Of the 65 phages that were originally picked for isolation on the three bacterial strains, 38 lytic phages could be purified and sequenced. All sequenced phages had double-stranded DNA genomes and identical k-mers at both ends of the genomes, indicating that they were circular. The lytic phages' genomes ranged between 34,627 and 40,542 nucleotides, and they contained between 53 and 76 predicted genes (Supplementary Table S1). The majority of the phages (34) were isolated from LMO6 or LMO9 and shared >50% average nucleotide identity (ANI) to each other. The remaining four phages were isolated with LMO8 as host and three were highly similar (>90% ANI) while the last phage shared no similarity to the others. With a genus cut-off of more than 50% ANI [83] across all genomes and a species cut-off of more than 95% ANI, these newly isolated phages are placed in three new genera and 15 new species based on all-versus-all genome comparisons, as well as VICTOR analysis (Figure 2, Supplementary Table S2). These newly isolated phages were not sufficiently similar to any reference genome to be able to establish further phylogenetic relationships (Figure 3). The three genera are proposed to be named Muminvirus, Pippivirus and Tantvirus based on the type species for each genus, Flavobacterium virus Mumin, Flavobacterium virus Pippi and Flavobacterium virus Tant. The isolates are named according to the ICTV's recommendations [84], e.g., Flavobacterium phage vB_FspS_mumin9-1, where 9 stands for the bacterial host of isolation (LMO9) and 1 is the phage isolate number.
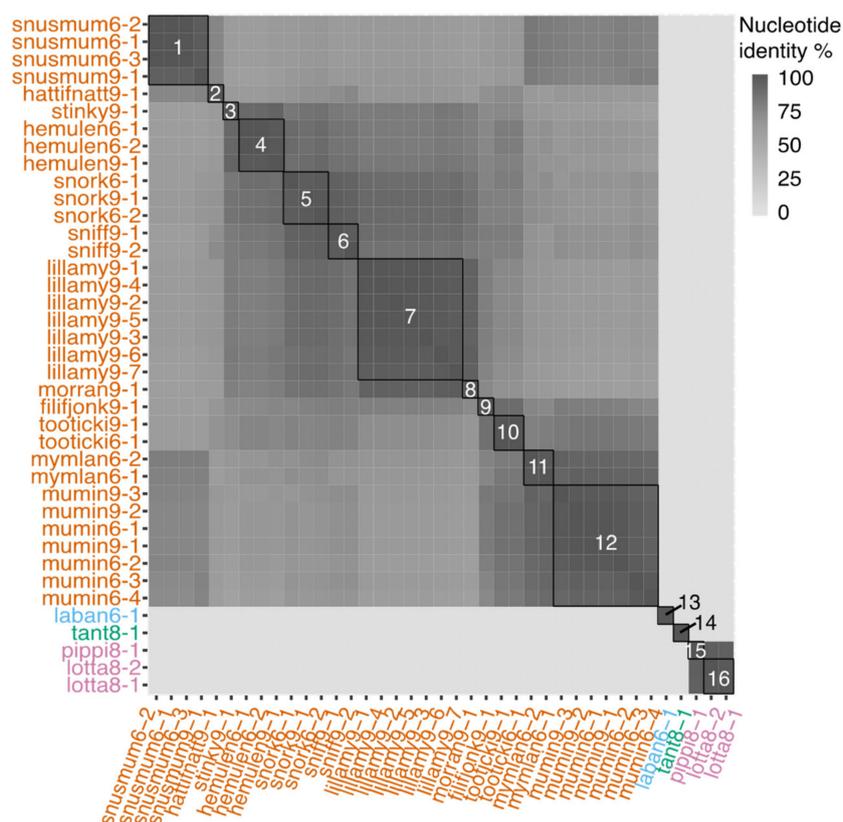
Figure 2. Similarity between the 38 lytic phages and the temperate phage based on Gegenees, which does a fragmented blastn between the entire genomes [67]; for details see Supplementary Table S3. Species are boxed with black while genera are coloured in the labels. Phages belonging to Muminvirus are orange, the isolate that constitutes Labanvirus is blue, Tantvirus is green and the isolates that belong to Pippivirus are pink. The species are named with the Flavobacterium virus and then 1 = Snusmum, 2 = Hattifnatt, 3 = Stinky, 4 = Hemulen, 5 = Snork, 6 = Sniff, 7 = Lillamy, 8 = Morran, 9 = Filifjonk, 10 = Tooticki, 11 = Mymlan, 12 = Mumin, 13 = Laban, 14 = Tant, 15 = Pippi, 16 = Lotta.

Figure 3. VICTOR tree produced with FastME based on nucleotide identity calculated with the D0-formula [73], with the newly isolated phages and reference genomes that share gene similarity to the novel phages presented in this study (accession number in parenthesis). The phages from this study are coloured based on the genera, as described in Figure 2. Bootstrap values above 50 are shown.

Genes belonging to the phage in Tantvirus did not share any similarities with either Mumin- or Pippivirus genes and Muminvirus and Pippivirus shared only four genes with low (27–29%) amino acid similarity (e-value less than 0.001; Supplementary TableS S4 and S5). Thus, the pan-genomes (based on more than 70% amino acid similarity) were calculated within each genus. The pan-genome

of Muminvirus consisted of 142 genes, including 26 genes in the core-genome, and the pan-genome of Pippivirus consisted of 56 genes, including 50 genes in the core-genome. Tantvirus only consisted of one phage isolate with 62 genes. Within the pan-genome of Muminvirus, 66 genes had significant hits (e-value less than 0.001) to the NCBI nr database but only 37 could be assigned a function (Supplementary Table S4). The pan-genome of Pippivirus had a slightly higher ratio of hits to the NCBI nr database (33 hits, 59%, e-value less than 0.001) and 13 genes could be assigned a function (Supplementary Table S5). For Tantvirus, 39 of the 62 predicted genes had hits to the NCBI nr database and 14 of those were assigned a function (Supplementary Table S6). Despite the dissimilarities between the genera, the annotations of the genes all belong to the same categories, generally even the same predicted functions. These were mainly functions involved in DNA processing, replication and recombination, as well as structural proteins and proteins involved in DNA packaging (Figure 4).
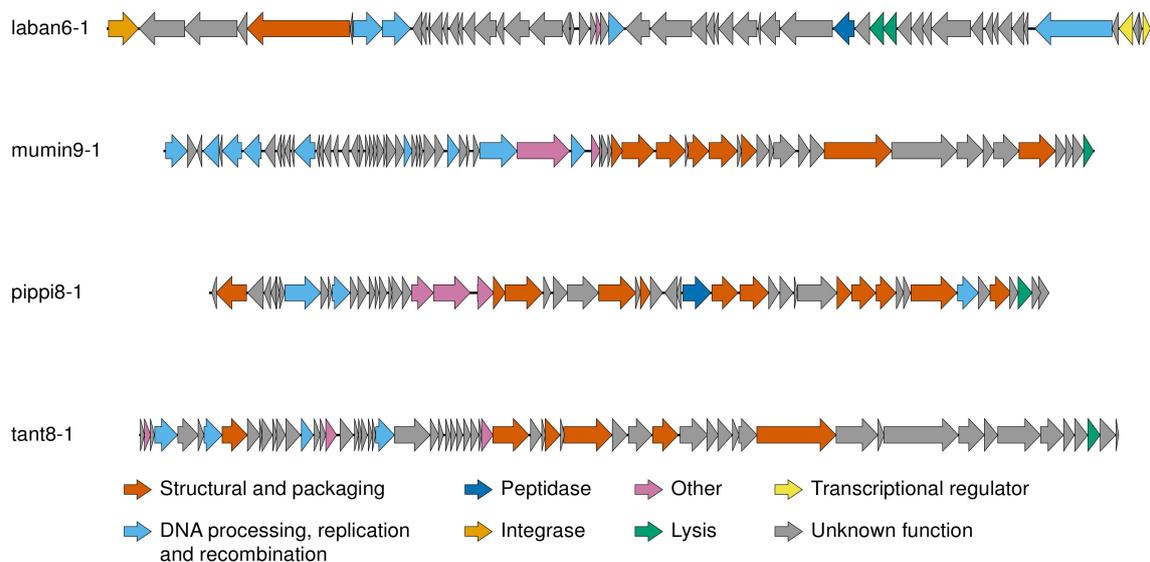
Figure 4. Genomic organisation of the type phages for each genus described in this study. Figure created with EasyFig [85]; for detailed information regarding the genomes and genes, see Supplementary Tables S1, S4–S7. The genomes are displayed linearly for visualisation.

3.3. Genomic Characteristics of the Temperate Phage

During the sequencing of the lytic phages, two complete genomes were assembled from pure phage cultures isolated using LMO6 as host. Comparison to the sequenced phages from LMO9 and the sequenced genome of LMO6 concluded that one of the phage genomes was from a phage with lytic replication (described above), tooticki6-1. The extra phage genome was detected in the sequence of the bacterium LMO6 and overlapped the incomplete 37.3 kb prophage identified by PHASTER. Therefore, this proved to be a temperate phage integrated into the genome of LMO6. The complete temperate phage genome could be assembled from nine of the sequenced viral isolates from LMO6 and these sequences were identical at the nucleotide level. This temperate phage was named Flavobacterium phage vB_FspS_laban6-1. Mapping of reads from the sequenced sample, including both tooticki6-1 and laban6-1, resulted in high coverage to the lytic phage (1680 read coverage), while the bacterial contig containing laban6-1 only was covered to a larger extent (50-260 read depth) over the region where the temperate phage was integrated (Figure 5).
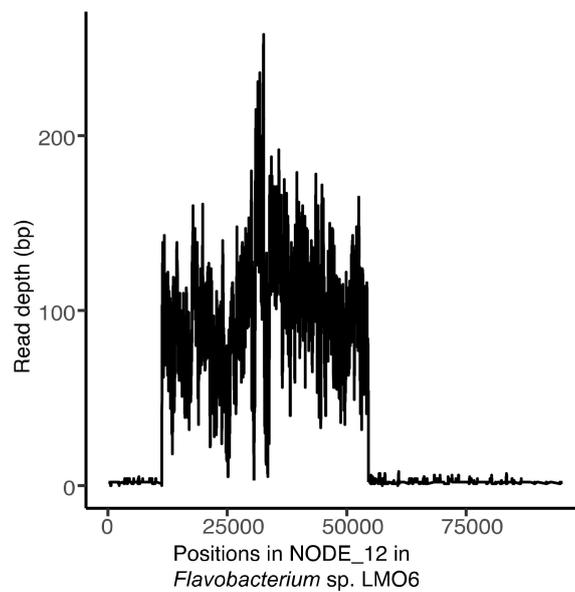
Figure 5. Coverage of viral reads (from well 103) against the bacterial contig that contained the prophage, laban6-1 (contig "NODE_12").

The temperate phage genome was 43,236 nucleotides long, and 57 potential genes could be predicted. The genome was double-stranded DNA and determined to be circular as both ends of the sequence had identical k-mers. Three of the predicted genes had sequence similarity to genes within Muminvirus, but showed only 25–37% amino acid identity (e-value less than 0.001; Supplementary Tables S4 and S7). Of the 57 genes, 44 showed amino acid similarity to proteins within databases but only 12 could be assigned to a function (Figure 4, Supplementary Table S7). As for the lytic phages, functions involved in DNA processing, replication and recombination as well as structural proteins and proteins involved in DNA packaging were identified, as well as a putative tyrosine site-specific recombinase involved in integration (Figure 4). VICTOR analysis using nucleotide similarities placed the temperate phage as a distinct species within the same genus of both previously isolated phages (e.g., *Flavobacterium* phage 6H [47]) and potentially integrated prophages (e.g., *Flavobacterium psychrophilum JIP02/86* [47]; Figure 3). However, Gegenees analysis based on nucleotide identity across the entire genomes (blastn analysis) did not report any similarity between the temperate phage and the references. Therefore, we propose that our new temperate phage should be placed in its own genus. The proposed name for this phage genus is Labanvirus, based on the type species Flavobacterium virus Laban, which is named after the isolated phage.

3.4. Morphology

In electron micrographs, it was observed that the type phage for Tantvirus, tant8-1, exhibited a siphovirus morphology, with a capsid diameter of 68 nm (standard deviation, ±1 nm, n = 13) and a tail length of 160 nm (±7 nm, n = 13; Figure 6a). The type phage for Pippivirus, pippi8-1, exhibited a myovirus morphology, with a capsid diameter of 63 nm (±2 nm, n = 9) and a tail length of 90 nm (±2 nm, n = 9; Figure 6b). mumin9-1, the type phage for Muminvirus, exhibited a siphovirus morphology, with only a slight curvature of the tail (Figure 6c), with a capsid diameter of 65 nm (±2 nm, n = 30) and a tail length of 153 nm (±4 nm, n = 30). Phage laban6-1 exhibited a siphovirus morphology, with a capsid of 68 nm (±2 nm, n = 12) and a tail length of 215 nm (±7 nm, n = 12; Figure 6d). In electron micrographs of the tooticki6-1 lysate samples, which should contain both the lytic phage tooticki6-1 and the temperate phage of LMO6, laban6-1, two different phage particles types with siphovirus morphology were indeed observed. Phage tooticki6-1 (Figure 6e(i)) had a head diameter of 64 nm (±2 nm, n = 15) and a tail length of 127 nm (±4 nm, n = 15), which was similar in appearance and lengths to mumin9-1 (compare Figure 6c to Figure 6e(i)). Present alongside were relatively few phage

particles possessing heads of diameter 65 nm (n = 1), and longer, flexible tails of 223 nm length (n = 1), characteristic of isolates of laban6-1 (compare Figure 6d to Figure 6e(ii)).
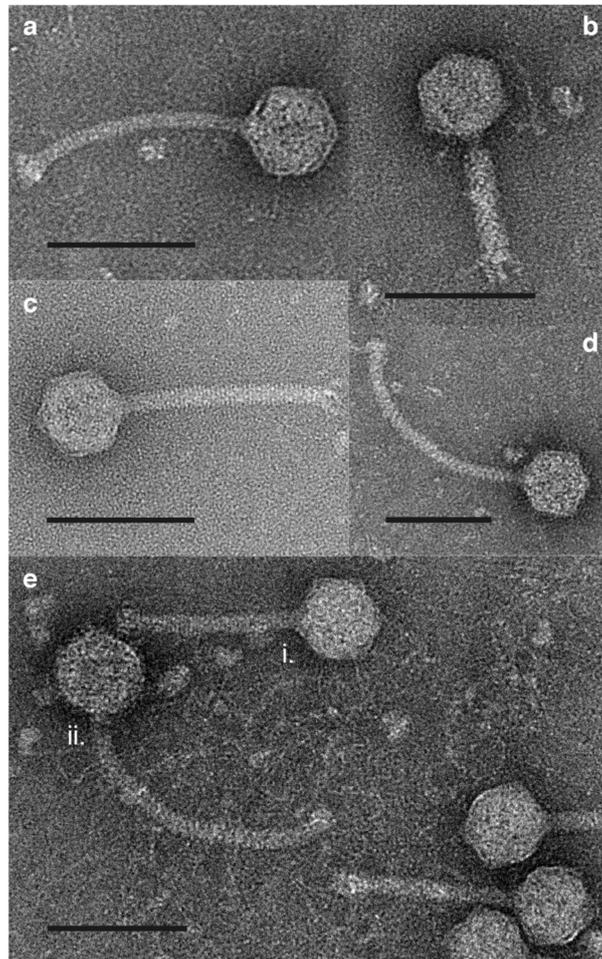

Figure 6. Negatively stained transmission electron micrographs of phage particles. (**a**) tant8-1, (**b**) pippi8-1, (**c**) mumin9-1, (**d**) laban6-1, (**e**) (i) tooticki6-1 and (**e**) (ii) laban6-1. Scale bars are 100 nm.

3.5. Host Range

The host ranges of the different phage species varied between the different genera (Figure 7 and Supplementary Table S8). The type phages of the species pippi8-1, lotta8-1 and tant8-1, which were isolated on LMO8, could infect both LMO6 and LMO9 but with much lower efficiency of plating compared to their efficiency on their original host. None of the phages within Muminvirus were able to form plaques on LMO8, but they could infect both LMO6 and LMO9 independent of which of the two bacterial strains that were used for isolation. However, while all phages isolated with LMO6 as their original host had an equally high efficiency of plating on both LMO6 and LMO9, some of the phages isolated with LMO9 had a three to four orders of magnitude lower efficiency of plating on LMO6 (Figure 7 and Supplementary Table S8). To investigate if the last infected host could influence this behaviour, morran9-1 and lillamy9-1 were grown on LMO6 before new plaque assays were performed on LMO6 and LMO9. This resulted in similar efficiency of plating of morran9-1 and lillamy9-1 on both hosts: morran9-1 produced 2.5×10^8 plaque-forming units (PFU) ml^{-1} on LMO6 and 3.9×10^8 PFU ml^{-1} on LMO9, and lillamy9-1 produced 1.0×10^8 PFU ml^{-1} on LMO6 and 2.2×10^8 PFU ml^{-1} on LMO9.
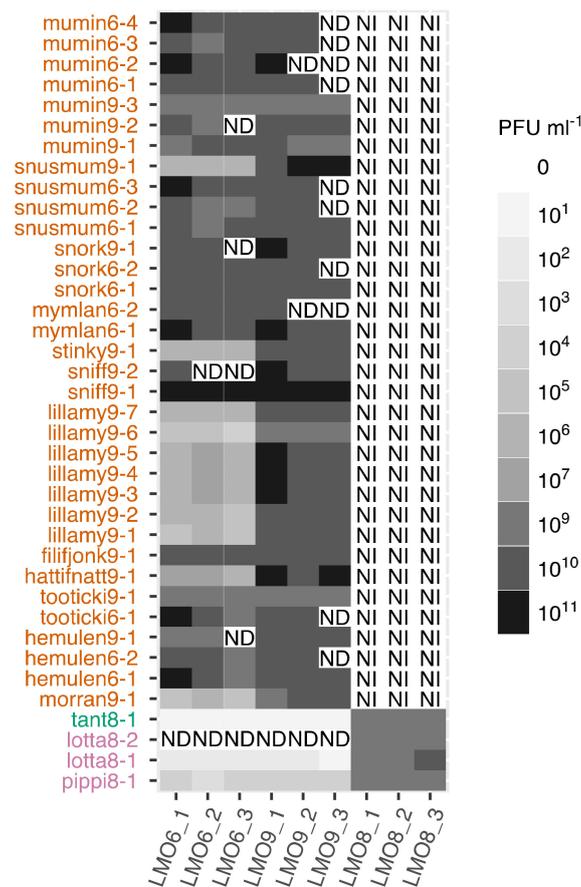
Figure 7. Host range tests for the type phages within each species with the three bacterial strains used within this study. ND = not determined, i.e., when the plaque assay has not been made; NI = no infection, i.e., when a plaque assay was performed but no plaques were detected. Phages are coloured according to genus, as in Figure 2. The number after, e.g., "mumin" indicates which bacterial strain the phage was originally isolated on, namely, 6 for LMO6, 8 for LMO8 and 9 for LMO9. The names on the *x*-axis are which bacterial strain the efficiency of plating was performed on, as well as if it was the first, second or third replicate (_1, _2, or _3).

3.6. Induction of laban6-1 From LMO6

During the induction experiment, the optical density of LMO6 and the UV-irradiated LMO6 continued to increase throughout the incubation (Supplementary Figure S1). Free, induced laban6-1 phages could be detected through plaque-formation on the lawn of LMO9, while no plaques were formed on LMO6 or LMO8. Free phages originating from bacteria-free filtrates from both the UV-radiated ($1.3 \times 10^8 \pm 1.1 \times 10^8$ PFU ml^{-1}, n = 3) and control LMO6 cultures ($2.0 \times 10^7 \pm 6.0 \times 10^6$ PFU ml^{-1}, n = 3) could be detected and the UV-radiated phages resulted in 3 to 6 times more plaques compared to the control.

3.7. Temporal and Spatial Variation

Mapping of viral metagenomic reads against the type phage of each genus resulted in the detection (>75% of the genome covered with reads [79]) of mumin9-1 in three LMO samples, August and September in 2013 and December in 2014 (Supplementary Table S9). None of the other phages showed great enough genome coverage to be regarded as present in any of the LMO metagenomes and none of the phages were detected in any of the GOS or Tara ocean viral metagenomes (Supplementary Table S10). Read-mapping of bacterial metagenomic reads from LMO showed that the bacterial isolates were not detected at species (97% read identity to the reference) or genus level (80% read identity to

the reference) (Supplementary Table S11). Even though the three bacterial strains did not have 100% identical 16S rRNA genes, they all match a 448-nucleotide long 16S V3V4 amplicon sequence variant (ASV) in the LMO amplicon sequencing data with 100% nucleotide identity. This ASV was detected in summer samples (May–July) at LMO in 2011–2013 and 2015 (Supplementary Table S12).

4. Discussion

The phages described in this study belong to four different genera whose members infect three strains of *Flavobacterium.* The proposed names for these phage genera are Muminvirus, Pippivirus, Tantvirus and Labanvirus (Supplementary Table S1). Despite their genetic (Figure 2; Figure 3) and morphological differences (Figure 6), they were quite similar in genome size, ranging between 34 and 43 kb (Supplementary Table S1). This limited range of genome sizes is more uniform than the range of genome sizes seen in e.g., the *Cellulophaga* phages, which range in genome size from 6 and 145 kb [37], or the mycobacteriophages, which range between 40 and 160 kb [86]. Considering the large number of isolated phages for the LMO6 and LMO9 host strains, it could be suggested that the natural diversity of phages, with lytic replication that infect these strains, was low at the time of sampling, as only members of one genus could be isolated. However, the intra-genus diversity was relatively large with twelve different species isolated from the same timepoint. A single bacterial species maintaining several phage species of the same genus has been seen before, e.g., for the *Cellulophaga* phages, where two genera were isolated from the same seawater sample [36] and consisted of either four or three different species (phi13:2, phi18:3, phi 19:3, and phi 46:3 in the first genus, and phi12:1, phi17:1, and phi18:1 in the second; Figure 3, Supplementary Table S2, and [37]). In contrast, while only managing to isolate five, and sequence four, different phages for LMO8, these were divided into both different genera and had different morphologies (Figure 6a,b; Supplementary Table S2). These results likely reflect the diversity of *Flavobacterium* spp. phages at LMO at the time of sampling, and the number of plaques originally produced on the different hosts would suggest a higher abundance of muminphages, originally isolated on LMO6 or LMO9, compared to pippi- or tantphages, originally isolated on LMO8. While there is no sequenced viral metagenome for the time when the phages were isolated, mumin9-1 could be detected in viral metagenomes from previous years (Supplementary Table S9). This is also an indication that muminphages are more commonly abundant than the pippi- or tantphages. Since LMO8 supports phages that belong to different genera and have different morphologies, the competition of these phages for their host might prevent any of them reaching an abundance that can be detected in the viral metagenomes. The presence of muminphages did not coincide with the presence of the detected ASV that was similar to the host strains (Supplementary Table S12). Coinciding phage–host patterns have previously been seen at LMO for the barbaphages, a recently reported phage genus with myovirus morphology infecting a *Gammaproteobacteria* host, isolated from the same sample as the phages reported in this study [77]. The lack of coinciding temporal variations for the *Flavobacterium* host-strains and their phages reported here implies that there might be other, undescribed phages for the LMO6, LMO8, and LMO9 hosts that are prevalent in the Baltic Sea. Thus, further, high-resolution time series investigations of phage–host systems are called for.

About half, or more, of the genes for Mumin- (46%)-, Pippi (59%)-, and Tantvirus (62%) had similarities to previously known sequences, and the majority of best matches were to either various phages or to bacteria that belong to the same family as the bacterial host strains in this study, *Flavobacteriaceae*. Still, there was low similarity to any specific phage, which was confirmed by VICTOR analysis (Figure 3). The closest relationship to a previously known phage was a distant connection between tant8-1 and *Flavobacterium* phage Fpv5 (Figure 3), a similar-sized (42.6 kb) siphovirus isolated from fish farms in Denmark [46,48]. This points towards a shared ancestor of this viral type among phages infecting various *Flavobacterium* hosts in the Baltic Sea region. Besides this, no other close relationships were seen to previously isolated phages, even though phages infecting members of *Flavobacteriaceae* isolated from the Baltic Sea are relatively well-represented in culture collections [87],

including the diverse *Cellulophaga* phage–host collection [37] and sea ice *Flavobacterium* phage 1/32 [88] (Figure 3).

Only 21–38% of the predicted genes in the different genera could be assigned a function (Supplementary Tables S4–S7). Even though the genes between the different genera had low or non-existing similarities, the functions of the genes in each genus are classified in the same categories. As often seen among phages, these functions are involved in (*i*) DNA processing, replication and recombination, and (*ii*) structural and packaging (Figure 4). Among the structural genes, we were able to identify putative terminase genes in tant-, mumin-, and pippiphages, and putative major capsid proteins in tant- and muminphages. These are among the genes that are considered as viral hallmark genes [89,90], and even though hallmark genes are expected to exist in all phages, they are not always correctly annotated since their identification relies on having similarities to previously sequenced and properly annotated genes. The lack of identified hallmark genes has previously been noted for other aquatic phages infecting *Flavobacteriaceae* [37]. While aquatic phages in particular are known to have genes involved in various metabolic processes, e.g., DNA metabolism [9,91], no metabolic genes were detected in any of the phages. Further, no genes involved in lysogenic replication were detected, indicating that mumin-, pippi-, and tantphages are only capable of lytic replication. This is also true for *Flavobacterium* phage Fpv5, which is related to tant8-1, where no genes that would indicate a possible lysogenic lifestyle were detected [48].

Phage laban6-1 is a prophage in LMO6 and was sequenced as a result of induction, which was confirmed by read mapping (Figure 5). If the phage had been assembled by sequences of contaminating bacterial DNA, the coverage would have been even across the entire bacterial contig. That laban6-1 is indeed a temperate phage capable of lytic replication and not only a cryptic prophage—a prophage that has lost the ability to produce virions due to e.g., genome rearrangement [12,92]—was further confirmed by plaque assays on the closely related host, LMO9. Moreover, laban6-1 was induced from LMO6 grown in liquid medium without any additional stressors, producing high numbers of laban6-1 particles, around tens of millions. However, the induction did not affect the host's generation time (91 min, ±9), which was similar to that of LMO9 (92 min, ±21). Induction of temperate phages without any experimentally added stress is called spontaneous prophage induction (SPI) [93] and has been demonstrated previously [94]. SPI might give a bacterial host fitness advantages, for example, by promoting biofilm formation by the release of extracellular DNA or releasing toxins that increase the infectivity of non-induced bacteria (see [95] and references therein). Although none of these benefits from SPI were noted, LMO6 did not seem to be adversely affected by the induction, at least not considering its generation time. While none of the predicted genes in laban6-1 indicated any obvious benefit for the host (Supplementary Table S7), the failure of laban6-1 to produce plaques on LMO6 indicates that the presence of laban6-1 as a prophage in LMO6 resulted in resistance to new infections of laban6-1. This resistance could be caused by superinfection exclusion, i.e., prevention of similar phages entering the infected cell [96], or homoimmunity, i.e., resistance towards similar phage DNA within the cell [11]. Besides SPI, the increased phage titre in the UV-light exposure experiments accompanied by reduced bacterial OD (Supplementary Figure S1) indicates that UV can act as an inducing agent for laban6-1. However, it should be noted that the UV-light that LMO6 was exposed to in the experiment had a wavelength of 254 nm, i.e., UVC, which does not pass through the ozone layer and that solar radiation usually does not cause virus induction in all lysogenised bacteria [97]. Thus, it is unclear if natural UV can be a cause for laban6-1 induction, which therefore calls for future investigations. While SPI is clearly occurring in LMO6 cultures at room temperature (22 °C), laban6-1 could not be detected infecting LMO9 at the time of sampling, which indicates that there were no free laban6-1 in the water used for phage isolation. However, temperature increases are a known inducing agent for prophages [98] and given the lower temperature (16.2 °C) at the time of sampling compared to room temperature (22 °C), this temperature increase might lead to the induction levels we detect. While not investigated further in this study, this opens up for additional questions of the effect of

global temperature increases of the viral dynamics in aquatic environments and what effects these switches from lysogenic to lytic replication will have on microbial communities and nutrient cycling.

Of the predicted genes within laban6-1, 77% (44/57 genes) had hits to the queried databases, but only 21% (12 genes) could be assigned to a function. The large amount of hypothetical hits to bacterial genomes is likely a result of laban6-1 being a temperate phage and therefore sharing a lot of similarity with un-annotated prophages. The relatedness to sequenced prophages was confirmed by the VICTOR analysis, where laban6-1 is placed within the same genus as several potential prophage regions in *Flavobacterium* strains as well as potentially temperate *Flavobacterium* phages (Figure 3, Supplementary Table S2). These phages and bacteria have been isolated from various aquatic environments, including fish cultures (the phages and *Flavobacterium psychrophilum JIP02/86*), the freshwater Han River in Korea (*Flavobacterium* sp. CJ74), and mining wastewaters (*Flavobacteriales bacterium 32-34-25*). Thus, despite that laban6-1 was not detected in LMO or other viral metagenomes, this group of related phages is clearly prevalent in a large range of aquatic habitats and might have consequences for, e.g., economic interests through their infection of fish pathogens [47].

The host range of the *Flavobacterium* phages in this study was determined using efficiency of plating as a measure of successful propagation. The phages belonging to the Pippi- and Tantviruses, originally isolated on LMO8, could, besides infecting LMO8, also replicate on both LMO6 and LMO9; however, with a reduced efficiency of plating (Figure 7). This could be an indication that LMO6 and LMO9 are sub-optimal hosts for the Pippi- and Tantviruses, potentially by reducing the infection through defense mechanisms. These phages, which are able to infect two seemingly different species of bacteria, could be assumed to be more prevalent in the environment since they are able to infect several hosts [99]. However, this is contradicted by the mapping of viral metagenomes against tant8-1 and pippi8-1, which did not result in enough coverage to be able to count them as present at any time point. In contrast, the muminphages that only produced plaques on LMO6 and LMO9 (Figure 7) were detected in the viral metagenomes (Supplementary Table S9), potentially due to a specialisation on their host. Phages specialised on a specific host have been seen to result in larger burst sizes compared to more generalist phages infecting multiple hosts [43], which could explain why muminphages were more abundant in the viral metagenomes than the phages isolated on LMO8. While previous studies have reported that myoviruses have broader host ranges than siphoviruses [100], such patterns were not observed here. Instead, the siphovirus (tant8-1) and the myoviruses (pippiphages) isolated on LMO8 could infect the same host strains (Figure 7).

While all phages in Muminvirus could infect both LMO6 and LMO9, they did so with varying efficiency of plating. Phages originally isolated on LMO6 had equal efficiency of plating on both hosts, while some of the phages originally isolated on LMO9 had lower efficiency of plating on LMO6 (Figure 7). Such results have previously been seen among other aquatic phages, e.g., *Lightbulbvirus*, and the variation in the efficiency of infection has been shown to depend on the last host that was infected [4]. This also appears to be the case for the muminphages, as the LMO9 phages with reduced efficiency of plating on LMO6 increased their efficiency after replication on LMO6. This shift in efficiency of plating depending on the latest infected host is commonly associated with the restriction-modification system and methylation or glycosylation of the phage genome [96,101,102]. Phages are also able to provide their own methylation to evade host restriction enzymes. While the methylation patterns of the LMO6 and LMO9 phages have not been investigated, all LMO9 phages which originally had reduced efficiency of plating on LMO6 lacked a specific methyltransferase that the LMO9 phages with high efficiency of infection on LMO6 possessed (Supplementary Table S13). Thus, it is likely that some of the Muminvirus species have the ability to methylate their genomes in a way that provides full protection against the LMO6 and LMO9 bacteria restriction enzymes, while others require additional host-derived methylation to be able to reach a high efficiency of plating on LMO6 [96,103].

The LMO *Flavobacterium* phage–host systems described here contribute to an ever-growing resource of phage–host model systems for the study of aquatic virology. The absence of metabolic genes in the phage genomes analysed, for example, those involved in nucleotide metabolism, was unexpected,

given their recent detection in many aquatic phage genomes. Bacterium LMO6 was found to be infected by virulent phages as well as harbouring a UVC-inducible temperate phage. Such phage–host systems appear to be quite rare among aquatic environmental isolates. Hence, the characterised LMO6–laban6-1 system represents a unique opportunity for future investigations focused on the effects of different environmental conditions on lytic and lysogenic infection cycles.

References

1. Bergh, O.; Borsheim, K.Y.; Bratbak, G.; Heldal, M. High abundance of viruses found in aquatic environments. *Nature* **1989**, *340*, 467–468. [CrossRef]

2. Suttle, C.A. Marine viruses-major players in the global ecosystem. *Nat. Rev. Microbiol.* **2007**, *5*, 801–812. [CrossRef]

3. Weitz, J.S.; Poisot, T.; Meyer, J.R.; Flores, C.O.; Valverde, S.; Sullivan, M.B.; Hochberg, M.E. Phage-bacteria infection networks. *Trends Microbiol.* **2013**, *21*, 82–91. [CrossRef] [PubMed]

4. Holmfeldt, K.; Solonenko, N.; Howard-Varona, C.; Moreno, M.; Malmstrom, R.R.; Blow, M.J.; Sullivan, M.B. Large-scale maps of variable infection efficiencies in aquatic *Bacteroidetes* phage-host model systems. *Environ. Microbiol.* **2016**, *18*, 3949–3961. [CrossRef] [PubMed]

5. Mann, N.H.; Cook, A.; Millard, A.; Bailey, S.; Clokie, M. Marine ecosystems: Bacterial photosynthesis genes in a virus. *Nature* **2003**, *424*, 741. [CrossRef]

6. Lindell, D.; Jaffe, J.D.; Coleman, M.L.; Futschik, M.E.; Axmann, I.M.; Rector, T.; Kettler, G.; Sullivan, M.B.; Steen, R.; Hess, W.R.; et al. Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* **2007**, *449*, 83–86. [CrossRef]

7. Breitbart, M.; Bonnain, C.; Malki, K.; Sawaya, N.A. Phage puppet masters of the marine microbial realm. *Nat. Microbiol.* **2018**, *3*, 754–766. [CrossRef] [PubMed]

8. Middelboe, M.; Holmfeldt, K.; Riemann, L.; Nybroe, O.; Haaber, J. Bacteriophages drive strain diversification in a marine Flavobacterium: Implications for phage resistance and physiological properties. *Environ. Microbiol.* **2009**, *11*, 1971–1982. [CrossRef] [PubMed]

9. Hurwitz, B.L.; Hallam, S.J.; Sullivan, M.B. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* **2013**, *14*, R123. [CrossRef]

10. Howard-Varona, C.; Hargreaves, K.R.; Solonenko, N.E.; Markillie, L.M.; White, R.A.; Brewer, H.M.; Ansong, C.; Orr, G.; Adkins, J.N.; Sullivan, M.B. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *ISME J.* **2018**, *12*, 1605–1618. [CrossRef]

11. Levin, B.; Lenski, R. Coevolution in bacteria and their viruses and plasmids. In *Coevolution*; Futuyama, D.a.S.M., Ed.; Sinauer Associates: Sunderland, MA, USA, 1983; pp. 99–127.

12. Howard-Varona, C.; Hargreaves, K.R.; Abedon, S.T.; Sullivan, M.B. Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. *ISME J.* **2017**, *11*, 1511–1520. [CrossRef] [PubMed]

13. Bratbak, G.; Thingstad, F.; Heldal, M. Viruses and the microbial loop. *Microb. Ecol.* **1994**, *28*, 209–221. [CrossRef] [PubMed]

14. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541–548. [CrossRef] [PubMed]

15. Warwick-Dugdale, J.; Buchholz, H.H.; Allen, M.J.; Temperton, B. Host-hijacking and planktonic piracy: How phages command the microbial high seas. *Virol. J.* **2019**, *16*. [CrossRef] [PubMed]

16. Waldor, M.K.; Mekalanos, J.J. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **1996**, *272*, 1910–1914. [CrossRef]

17. Chen, Y.; Golding, I.; Sawai, S.; Guo, L.; Cox, E.C. Population fitness and the regulation of *Escherichia coli* genes by bacterial viruses. *PLoS Biol.* **2005**, *3*, e229. [CrossRef]

18. Paul, J.H. Prophages in marine bacteria: Dangerous molecular time bombs or the key to survival in the seas? *ISME J.* **2008**, *2*, 579–589. [CrossRef]

19. Jiang, S.C.; Paul, J.H. Seasonal and diel abundance of viruses and occurrence of lysogeny/bacteriocinogeny in the marine environment. *Mar. Ecol. Prog. Ser.* **1994**, *104*, 163–172. [CrossRef]

20. Roux, S.; Hallam, S.J.; Woyke, T.; Sullivan, M.B. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *eLife* **2015**, *4*, e08490. [CrossRef]

21. Cochran, P.K.; Paul, J.H. Seasonal abundance of lysogenic bacteria in a subtropical estuary. *Appl. Environ. Microb.* **1998**, *64*, 2308–2312. [CrossRef]

22. Williamson, S.J.; Houchin, L.A.; McDaniel, L.; Paul, J.H. Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Appl. Environ. Microb.* **2002**, *68*, 4307–4314. [CrossRef] [PubMed]

23. Brum, J.R.; Hurwitz, B.L.; Schofield, O.; Ducklow, H.W.; Sullivan, M.B. Seasonal time bombs: Dominant temperate viruses affect Southern Ocean microbial dynamics. *ISME J.* **2016**, *10*, 437–449. [CrossRef] [PubMed]

24. Chuvochina, M.S.; Marie, D.; Chevaillier, S.; Petit, J.R.; Normand, P.; Alekhina, I.A.; Bulat, S.A. Community variability of bacteria in alpine snow (Mont Blanc) containing Saharan dust deposition and their snow colonisation potential. *Microbes Environ.* **2011**, *26*, 237–247. [CrossRef] [PubMed]

25. Bäckhed, F.; Ley, R.E.; Sonnenburg, J.L.; Peterson, D.A.; Gordon, J.I. Host-bacterial mutualism in the human intestine. *Science* **2005**, *307*, 1915–1920. [CrossRef] [PubMed]

26. Allen, L.Z.; Allen, E.E.; Badger, J.H.; McCrow, J.P.; Paulsen, I.T.; Elbourne, L.D.H.; Thiagarajan, M.; Rusch, D.B.; Nealson, K.H.; Williamson, S.J.; et al. Influence of nutrients and currents on the genomic composition of microbes across an upwelling mosaic. *ISME J.* **2012**, *6*, 1403–1414. [CrossRef]

27. Kirchman, D.L. The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol. Ecol.* **2002**, *39*, 91–100. [CrossRef]

28. Pinhassi, J.; Sala, M.M.; Havskum, H.; Peters, F.; Guadayol, O.; Malits, A.; Marrase, C. Changes in bacterioplankton composition under different phytoplankton regimens. *Appl. Environ. Microbiol.* **2004**, *70*, 6753–6766. [CrossRef]

29. Fernandez-Gomez, B.; Richter, M.; Schuler, M.; Pinhassi, J.; Acinas, S.G.; Gonzalez, J.M.; Pedros-Alio, C. Ecology of marine Bacteroidetes: A comparative genomics approach. *ISME J.* **2013**, *7*, 1026–1037. [CrossRef]

30. Rodrigues, R.A.L.; Andrade, A.; Boratto, P.V.M.; Trindade, G.S.; Kroon, E.G.; Abrahao, J.S. An anthropocentric view of the virosphere-host relationship. *Front. Microbiol.* **2017**, *8*, 1673. [CrossRef]

31. McBride, M.J. The family *Flavobacteriaceae*. In *The Prokaryotes*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 643–676.

32. Borriss, M.; Lombardot, T.; Glockner, F.O.; Becher, D.; Albrecht, D.; Schweder, T. Genome and proteome characterization of the psychrophilic *Flavobacterium* bacteriophage 11b. *Extremophiles* **2007**, *11*, 95–104. [CrossRef]

33. Kang, I.; Jang, H.; Cho, J.C. Complete genome sequences of two *Persicivirga* bacteriophages, P12024S and P12024L. *J. Virol.* **2012**, *86*, 8907–8908. [CrossRef] [PubMed]

34. Kang, I.; Kang, D.; Cho, J.C. Complete genome sequence of *Croceibacter* bacteriophage P2559S. *J. Virol.* **2012**, *86*, 8912–8913. [CrossRef] [PubMed]

35. Laanto, E.; Mäntynen, S.; De Colibus, L.; Marjakangas, J.; Gillum, A.; Stuart, D.I.; Ravantti, J.J.; Huiskonen, J.T.; Sundberg, L.R. Virus found in a boreal lake links ssDNA and dsDNA viruses. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8378–8383. [CrossRef] [PubMed]

36. Holmfeldt, K.; Middelboe, M.; Nybroe, O.; Riemann, L. Large variabilities in host strain susceptibility and phage host range govern interactions between lytic marine phages and their *Flavobacterium* hosts. *Appl. Environ. Microbiol.* **2007**, *73*, 6730–6739. [CrossRef] [PubMed]

37. Holmfeldt, K.; Solonenko, N.; Shah, M.; Corrier, K.; Riemann, L.; Verberkmoes, N.C.; Sullivan, M.B. Twelve previously unknown phage genera are ubiquitous in global oceans. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12798–12803. [CrossRef]

38. Dutilh, B.E.; Noriko, C.W.; McNair, K.; Sanchez, S.E.; Silva, G.G.Z.; Boling, L.; Barr, J.J.; Speth, D.R.; Seguritan, V.; Aziz, R.K.; et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* **2014**, *5*, 4498. [CrossRef]

39. Yutin, N.; Makarova, K.S.; Gussow, A.B.; Krupovic, M.; Segall, A.; Edwards, R.A.; Koonin, E.V. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat. Microbiol.* **2018**, *3*, 38–46. [CrossRef]

40. Shkoporov, A.N.; Khokhlova, E.V.; Fitzgerald, C.B.; Stockdale, S.R.; Draper, L.A.; Ross, R.P.; Hill, C. Phi CrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides intestinalis*. *Nat. Commun.* **2018**, *9*, 4781. [CrossRef]

41. Riemann, L.; Grossart, H.P. Elevated lytic phage production as a consequence of particle colonization by a marine *Flavobacterium* (*Cellulophaga* sp.). *Microb. Ecol.* **2008**, *56*, 505–512. [CrossRef]

42. Dang, V.T.; Howard-Varona, C.; Schwenck, S.; Sullivan, M.B. Variably lytic infection dynamics of large *Bacteroidetes* podovirus phi38:1 against two *Cellulophaga baltica* host strains. *Environ. Microbiol.* **2015**, *17*, 4659–4671. [CrossRef]

43. Holmfeldt, K.; Howard-Varona, C.; Solonenko, N.; Sullivan, M.B. Contrasting genomic patterns and infection strategies of two co-existing *Bacteroidetes* podovirus genera. *Environ. Microbiol.* **2014**, *16*, 2501–2513. [CrossRef]

44. Castillo, D.; Higuera, G.; Villa, M.; Middelboe, M.; Dalsgaard, I.; Madsen, L.; Espejo, R.T. Diversity of *Flavobacterium psychrophilum* and the potential use of its phages for protection against bacterial cold water disease in salmonids. *J. Fish. Dis.* **2012**, *35*, 193–201. [CrossRef] [PubMed]

45. Christiansen, R.H.; Dalsgaard, I.; Middelboe, M.; Lauritsen, A.H.; Madsen, L. Detection and quantification of *Flavobacterium psychrophilum*-specific bacteriophages in vivo in rainbow trout upon oral administration: Implications for disease control in aquaculture. *Appl. Environ. Microbiol.* **2014**, *80*, 7683–7693. [CrossRef]

46. Stenholm, A.R.; Dalsgaard, I.; Middelboe, M. Isolation and characterization of bacteriophages infecting the fish pathogen *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* **2008**, *74*, 4070–4078. [CrossRef]

47. Castillo, D.; Espejo, R.; Middelboe, M. Genomic structure of bacteriophage 6H and its distribution as prophage in *Flavobacterium psychrophilum* strains. *FEMS Microbiol. Lett.* **2014**, *351*, 51–58. [CrossRef]

48. Castillo, D.; Middelboe, M. Genomic diversity of bacteriophages infecting the fish pathogen *Flavobacterium psychrophilum*. *FEMS Microbiol. Lett.* **2016**, *363*, fnw272. [CrossRef]

49. Duchaud, E.; Boussaha, M.; Loux, V.; Bernardet, J.F.; Michel, C.; Kerouault, B.; Mondot, S.; Nicolas, P.; Bossy, R.; Caron, C.; et al. Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*. *Nat. Biotechnol.* **2007**, *25*, 763–769. [CrossRef]

50. Castillo, D.; Christiansen, R.H.; Dalsgaard, I.; Madsen, L.; Espejo, R.; Middelboe, M. Comparative genome analysis provides insights into the pathogenicity of *Flavobacterium psychrophilum*. *PLoS ONE* **2016**, *11*, e0152515. [CrossRef] [PubMed]

51. Castillo, D.; Christiansen, R.H.; Dalsgaard, I.; Madsen, L.; Middelboe, M. Bacteriophage resistance mechanisms in the fish pathogen *Flavobacterium psychrophilum*: Linking genomic mutations to changes in bacterial virulence factors. *Appl. Environ. Microbiol.* **2015**, *81*, 1157–1167. [CrossRef] [PubMed]

52. Kauffman, K.M.; Polz, M.F. Streamlining standard bacteriophage methods for higher throughput. *MethodsX* **2018**, *5*, 159–172. [CrossRef]

53. Wickham, H. *Tidyverse: Easily install and load 'tidyverse' package*; Version 1.3.0; RStudio Inc.: Boston, MA, USA, 2017.

54. R Core Team. *R: A language and environment for statistical computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

55. RStudio Team. *RStudio: Integrated development for R*; RStudio Inc.: Boston, MA, USA, 2015.

56. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]

57. Andrews, S. *FastQC: A quality control tool for high throughput sequence data*; Version 0.7.2; The Babraham Institute: Cambridge, UK, 2010.

58. 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.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef] [PubMed]

59. Simpson, J.T.; Wong, K.; Jackman, S.D.; Schein, J.E.; Jones, S.J.M.; Birol, I. ABySS: A parallel assembler for short read sequence data. *Genome Res.* **2009**, *19*, 1117–1123. [CrossRef] [PubMed]

60. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [CrossRef] [PubMed]

61. Tatusova, T.; DiCuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E.P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostell, J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **2016**, *44*, 6614–6624. [CrossRef]

62. McNair, K.; Zhou, C.; Dinsdale, E.A.; Souza, B.; Edwards, R.A. PHANOTATE: A novel approach to gene identification in phage genomes. *Bioinformatics* **2019**, *35*, 4537–4542. [CrossRef]

63. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [CrossRef]

64. Eddy, S.R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **2011**, *7*, e1002195. [CrossRef]

65. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [CrossRef]

66. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [CrossRef]

67. Ågren, J.; Sundström, A.; Håfström, T.; Segerman, B. Gegenees: Fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLoS ONE* **2012**, *7*, e39107. [CrossRef] [PubMed]

68. Pritchard, L.; Glover, R.H.; Humphris, S.; Elphinstone, J.G.; Toth, I.K. Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods UK* **2016**, *8*, 12–24. [CrossRef]

69. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [CrossRef] [PubMed]

70. Chaumeil, P.A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **2019**. [CrossRef] [PubMed]

71. Parks, D.H.; Chuvochina, M.; Chaumeil, P.-A.; Rinke, C.; Mussig, A.J.; Hugenholtz, P. Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy. *bioRxiv* **2019**. [CrossRef]

72. Huson, D.H.; Scornavacca, C. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Syst. Bio.* **2012**, *61*, 1061–1067. [CrossRef] [PubMed]

73. Meier-Kolthoff, J.P.; Göker, M. VICTOR: Genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* **2017**, *33*, 3396–3404. [CrossRef]

74. 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 Bioinform.* **2013**, *14*, 60. [CrossRef]

75. 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]

76. Göker, M.; Garcia-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]

77. Nilsson, E.; Li, K.; Fridlund, J.; Šulčius, S.; Bunse, C.; Karlsson, C.M.G.; Lindh, M.; Lundin, D.; Pinhassi, J.; Holmfeldt, K. Genomic and seasonal variations among aquatic phages infecting the Baltic Sea gammaproteobacterium *Rheinheimera* sp. strain BAL341. *Appl. Environ. Microbiol.* **2019**, *85*, e01003–e01019. [CrossRef] [PubMed]

78. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef] [PubMed]

79. Roux, S.; Emerson, J.B.; Eloe-Fadrosh, E.A.; Sullivan, M.B. Benchmarking viromics: An in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. *PeerJ* **2017**, *5*, e3817. [CrossRef] [PubMed]

80. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]

81. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [CrossRef]

82. Hugerth, L.W.; Larsson, J.; Alneberg, J.; Lindh, M.V.; Legrand, C.; Pinhassi, J.; Andersson, A.F. Metagenome-assembled genomes uncover a global brackish microbiome. *Genome Biol.* **2015**, *16*, 279. [CrossRef]

83. Adriaenssens, E.M.; Brister, J.R. How to name and classify your phage: An informal guide. *Viruses* **2017**, *9*, 70. [CrossRef]

84. Kropinski, A.M.; Prangishvili, D.; Lavigne, R. Position paper: The creation of a rational scheme for the nomenclature of viruses of *Bacteria* and *Archaea*. *Environ. Microbiol.* **2009**, *11*, 2775–2777. [CrossRef]

85. Sullivan, M.J.; Petty, N.K.; Beatson, S.A. Easyfig: A genome comparison visualizer. *Bioinformatics* **2011**, *27*, 1009–1010. [CrossRef]

86. Hatfull, G.F.; Jacobs-Sera, D.; Lawrence, J.G.; Pope, W.H.; Russell, D.A.; Ko, C.C.; Weber, R.J.; Patel, M.C.; Germane, K.L.; Edgar, R.H.; et al. Comparative genomic analysis of 60 mycobacteriophage genomes: Genome clustering, gene acquisition, and gene size. *J. Mol. Biol.* **2010**, *397*, 119–143. [CrossRef]

87. Sulcius, S.; Holmfeldt, K. Viruses of microorganisms in the Baltic Sea: Current state of research and perspectives. *Mar. Biol. Res.* **2016**, *12*, 115–124. [CrossRef]

88. Luhtanen, A.M.; Eronen-Rasimus, E.; Kaartokallio, H.; Rintala, J.M.; Autio, R.; Roine, E. Isolation and characterization of phage-host systems from the Baltic Sea ice. *Extremophiles* **2014**, *18*, 121–130. [CrossRef] [PubMed]

89. Koonin, E.V.; Senkevich, T.G.; Dolja, V.V. The ancient virus world and evolution of cells. *Biol. Direct.* **2006**, *1*, 29. [CrossRef] [PubMed]

90. Roux, S.; Hawley, A.K.; Beltran, M.T.; Scofield, M.; Schwientek, P.; Stepanauskas, R.; Woyke, T.; Hallam, S.J.; Sullivan, M.B. Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta- genomics. *eLife* **2014**, *3*, e03125. [CrossRef]

91. Breitbart, M. Marine viruses: Truth or dare. *Ann. Rev. Mar. Sci.* **2012**, *4*, 425–448. [CrossRef]

92. Wang, X.; Kim, Y.; Ma, Q.; Hong, S.H.; Pokusaeva, K.; Sturino, J.M.; Wood, T.K. Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* **2010**, *1*, 147. [CrossRef]

93. Lwoff, A. Lysogeny. *Bacteriol. Rev.* **1953**, *17*, 269–337. [CrossRef]

94. Alexeeva, S.; Martinez, J.A.G.; Spus, M.; Smid, E.J. Spontaneously induced prophages are abundant in a naturally evolved bacterial starter culture and deliver competitive advantage to the host. *BMC Microbiol.* **2018**, *18*. [CrossRef]

95. Nanda, A.M.; Thormann, K.; Frunzke, J. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *J. Bacteriol.* **2015**, *197*, 410–419. [CrossRef]

96. Labrie, S.J.; Samson, J.E.; Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **2010**, *8*, 317–327. [CrossRef]

97. Jiang, S.C.; Paul, J.H. Occurrence of lysogenic bacteria in marine microbial communities as determined by prophage induction. *Mar. Ecol. Prog. Ser.* **1996**, *142*, 27–38. [CrossRef]

98. Shan, J.; Korbsrisate, S.; Withatanung, P.; Adler, N.L.; Clokie, M.R.; Galyov, E.E. Temperature dependent bacteriophages of a tropical bacterial pathogen. *Front. Microbiol.* **2014**, *5*, 599. [CrossRef] [PubMed]

99. Koskella, B.; Meaden, S. Understanding bacteriophage specificity in natural microbial communities. *Viruses* **2013**, *5*, 806–823. [CrossRef]

100. Sullivan, M.B.; Waterbury, J.B.; Chisholm, S.W. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **2003**, *424*, 1047–1051. [CrossRef] [PubMed]

101. Arber, W. Host specificity of DNA produced by Escherichia coli: V. The role of methionine in the production of host specificity. *J. Mol. Biol.* **1965**, *11*, 247–256. [CrossRef]

102. Hattman, S.; Fukasawa, T. Host-induced modification of T-even phages due to defective glucosylation of their DNA. *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 297–300. [CrossRef]

103. Luria, S.E. Host-induced modifications of viruses. *Cold Spring Harb Symp Quant. Biol.* **1953**, *18*, 237–244. [CrossRef]