



Article Incomplete LPS Core-Specific Felix01-Like Virus vB_EcoM_VpaE1

Eugenijus Šimoliūnas¹, Monika Vilkaitytė¹, Laura Kaliniene¹, Aurelija Zajančkauskaitė¹, Algirdas Kaupinis², Juozas Staniulis³, Mindaugas Valius², Rolandas Meškys¹ and Lidija Truncaitė^{1,*}

Received: 8 October 2015; Accepted: 18 November 2015; Published: 27 November 2015 Academic Editor: Rob Lavigne

- ¹ Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Vilnius University, Vilnius LT-08662, Lithuania; eugenijus.simoliunas@bchi.vu.lt (E.Š.); m.vilke@gmail.com (M.V.); laura.kaliniene@bchi.vu.lt (L.K.); aurelija.zajanckauskaite@bchi.vu.lt (A.Z.); rolandas.meskys@bchi.vu.lt (R.M.)
- ² Proteomics Centre, Institute of Biochemistry, Vilnius University, Vilnius LT-08662, Lithuania; algirdas.kaupinis@gf.vu.lt (A.K.); mindaugas.valius@bchi.vu.lt (M.V.)
- ³ Laboratory of Plant Viruses, Institute of Botany, Nature Research Centre, Vilnius LT-08412, Lithuania; juozas.staniulis@botanika.lt
- * Correspondence: lidija.truncaite@bchi.vu.lt; Tel.: +370-5-2729146

Abstract: Bacteriophages represent a valuable source for studying the mechanisms underlying virus-host interactions. A better understanding of the host-specificity of viruses at the molecular level can promote various phage applications, including bacterial diagnostics, antimicrobial therapeutics, and improve methods in molecular biology. In this study, we describe the isolation and characterization of a novel coliphage, vB_EcoM_VpaE1, which has different host specificity than its relatives. Morphology studies, coupled with the results of genomic and proteomic analyses, indicate that vB_EcoM_VpaE1 belongs to the newly proposed genus *Felix01likevirus* in the family Myoviridae. The genus Felix01likevirus comprises a group of highly similar phages that infect O-antigen-expressing Salmonella and Escherichia coli (E. coli) strains. Phage vB_EcoM_VpaE1 differs from the rest of Felix01-like viruses, since it infects O-antigen-deficient E. coli strains with an incomplete core lipopolysaccharide (LPS). We show that vB_EcoM_VpaE1 can infect mutants of E. coli that contain various truncations in their LPS, and can even recognize LPS that is truncated down to the inner-core oligosaccharide, showing potential for the control of rough E. coli strains, which usually emerge as resistant mutants upon infection by O-Ag-specific phages. Furthermore, VpaE1 can replicate in a wide temperature range from 9 to 49 °C, suggesting that this virus is well adapted to harsh environmental conditions. Since the structural proteins of such phages tend to be rather robust, the receptor-recognizing proteins of VpaE1 are an attractive tool for application in glycan analysis, bacterial diagnostics and antimicrobial therapeutics.

Keywords: *E. coli* lipopolysaccharide; inner core oligosacharide; bacteriophage vB_EcoM_VpaE1 host-range; genus *Felix01likevirus*; PACS; J0101

1. Introduction

Attachment to cell surface receptors is the first step in phage infection. Phages of Gram-negative bacteria use a variety of cell-associated structures including pili, flagella, outer membrane proteins (Omp) or lipopolysaccharides (LPS) for host cell recognition [1–3]. LPS is a major constituent of the bacterial outer membrane (OM). Natural *Escherichia* and *Salmonella* isolates usually contain LPS that consists of: (i) the lipid A, a hydrophobic moiety of the molecule that anchors LPS to the OM; (ii) a

phosphorylated, nonrepetitive hetero-oligosaccharide known as the core oligosaccharide (core OS); and (iii) the O-antigen side chain polysaccharide (O-PS or O-Ag) that is a polymer of defined repeat units attached to the core OS, and is usually used for serotyping [4]. The inner part of the core OS is highly conserved within the Gram-negative bacteria, whereas the outer core is more diverse, and the O-Ag is the most variable moiety [5,6].

As mentioned previously, a number of bacteriophages use LPS as a receptor for attachment to the host cell. In turn, as the response to a strong selective pressure imposed by a phage infection, bacteria may evade phage attachment by altering the structure of this complex glycolipid through mutations in genes involved in LPS biosynthesis [7–9]. As a result, phage-resistant bacterial strains expressing truncated core LPS can emerge. However, mutation to phage resistance is usually accompanied by pleiotropic fitness costs [10,11]. Indeed, it has been reported that "rough" (O-Ag deficient), and "deep rough" (O-Ag deficient, truncated core OS) LPS mutants are less efficient competitors for growth-limiting resources than strains with intact "smooth" LPS [4,12–14]. Although costly in terms of reduced growth rate, such a strategy is often employed by bacteria in response to viral attack during phage treatment. In this regard, while the use of incomplete LPS-specific phages as auxiliary components of phage cocktails has proved beneficial in biocontrol of pathogenic bacteria [15–17], it is surprising that only a very limited number of reports on the isolation and characterization of such viruses have been published thus far.

In this study, we describe incomplete core LPS-specific coliphage vB_EcoM_VpaE1 (subsequently, VpaE1) that has been isolated from its natural habitat. Based on the results presented here, VpaE1 belongs to the newly proposed genus *Felix01likevirus*, represented by a well-studied *Salmonella* phage Felix01 [18,19] that infects numerous *Salmonella* strains and has been often used as a diagnostic and therapeutic agent [20,21]. Felix01 has long been considered as a singleton within the family *Myoviridae* [22]. However, during the last few years, a dozen of its close relatives, active mainly against *E. coli* [23–27] or *Salmonella* [28–32] have been isolated and sequenced. In addition, more distant Felix01 relatives that infect *Erwinia* [33], *Citrobacter* [34] or *Pseudomonas* [35] have been described. All Felix01-related *Salmonella* or *Escherichia coli* (*E. coli*) phages were isolated on strains with smooth LPS.

Bacteriophage VpaE1 described in this study differs from other characterized Felix01-like viruses by its specificity towards deep rough strains of *E. coli*. By using core LPS mutant strains of *E. coli* K-12 [36], we show that VpaE1 can even use LPS that is truncated down to the inner core OS. Moreover, we demonstrate that VpaE1 can replicate in a wide temperature range from 9 to 49 °C, suggesting that this phage is well adapted to harsh environmental conditions. As a coliphage with unique physiological characteristics, VpaE1 is an attractive model for studying virus-host co-evolution as well as investigating carbohydrate-protein interactions, a research area that has attracted considerable attention in recent years [37]. In addition, to combat the emergence of smooth/rough LPS-specific phage-resistant pathogenic *E. coli* strains, VpaE1 may be used in phage cocktails.

2. Materials and Methods

2.1. Bacterial Strains and Culture Conditions

Bacterial strains used for host range analysis and phage propagation are listed in Tables 1 and 2 respectively. Bacteria were cultivated in LB medium, in the temperature range of 7–49 °C. When needed, LB medium was supplemented with either 25 μ g·mL⁻¹ kanamycin (Kan, Sigma-Aldrich, St. Louis, MO, USA) or 12.5 mg·mL⁻¹ tetracycline (Tet, Sigma-Aldrich).

2.2. Phage Techniques

Bacteriophage VpaE1 was isolated from avian feces collected at a farm (Vievis, Lithuania) in 2008. Initially, samples were soaked in PB buffer (70 mM NaCl, 10 mM MgSO₄, 50 mM Na₂HPO₄,

30 mM KH₂PO₄) and centrifuged at 4000× *g* for 20 min before filtering through 0.45 µm pore-size filters (Sartorius, Goettingen, Germany). Decimal dilutions of the sample were then added to the liquid cultures of *E. coli*, mixed with 2.5 mL of 0.5% (w/v) soft agar and poured over the 1.2% LB agar plates that were incubated at 37 °C overnight. Then, phage suspensions in LB medium were prepared from single plaques. Decimal dilutions of suspensions were plated on the lawns of *E. coli* B derivatives and incubated in the temperature range of 7 to 49 °C. Thus, phage VpaE1 was selected for further analysis based on its plaque morphology (uniform clear plaques on the lawns of *E. coli* B strains) and a wide temperature range.

Bacteriophage VpaE1 was propagated from a single plaque suspension that was added to the exponentially growing liquid culture of *E. coli* BL21. Then, the culture was incubated at 37 °C with vigorous shaking until complete lysis (~2 h). Phage particles were collected from the lysate by centrifugation at $16,000 \times g$ for 1 h at 8 °C. The resulting pellet was soaked in 1 mL of SM buffer (100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-HCl, pH 7.5) containing 0.1 mL CHCl₃ and DNase I (2 units per mL; Thermo Fisher Scientific, Vilnius, Lithuania) at 4 °C overnight. Next day, the bacterial debris was removed by centrifugation at $5000 \times g$ for 10 min at room temperature. The supernatant was then subjected to CsCl density gradient centrifugation. The band with the highest opalescence was collected and dialyzed against SM buffer. Purified phage particles were used for TEM, host-range analyses and DNA isolation. Phage stock solution was stored at 4 °C.

The best temperature for VpaE1 propagation was determined by assessing plaque morphology on the lawns of *E. coli* BL21, after the plates were incubated in the temperature range from 9 °C to 49 °C.

2.3. Transmission Electron Microscopy (TEM)

Phage suspension was diluted to approximately 10^{11} PFU/mL with distilled water; 5 µL of the sample was deposited directly on the carbon-coated nitrocellulose grid. Excess liquid was drained with filter paper before staining with two successive drops of 2% uranyl acetate (pH 4.5), dried and examined in Morgagni 268(D) transmission electron microscope (Field Electron and Ion Co., Hillsboro, OR, USA). VpaE1 virions were measured (30 virions in total) after calibration with catalase crystals and/or T4 phage particles.

2.4. Host Range Analysis

The host range of VpaE1 was investigated using both spot test assay and determination of the efficiency of plating (EOP) as described by Kutter [38].

2.5. Adsorption Assays

2.5.1. Assay of Adsorbtion Kinetics

VpaE1 adsorption tests were carried out using *E. coli* BL21 (reference host), *E. coli* BW25113 (VpaE1-resistant negative control), and VpaE1-susceptible mutants of *E. coli* BW25113 from Keio collection [36]. The assay was carried out in triplicate, as described by Kropinski *et al.* [39] with minor modifications. The mid-log-phase cultures of the aforementioned *E. coli* strains (OD₆₀₀ of 0.5) were mixed with phage suspension to give a MOI (multiplicity of infection) of 0.1 and incubated at 37 °C. Samples (0.1 mL) were taken at different time intervals during the time period of 15 min. and diluted with LB medium. Phage-cell complexes were removed by treatment with chloroform (1/10 volume) and centrifugation at 5000× *g* for 5 min. The titer of the unadsorbed phage in the supernatant was then determined by the double-layer agar method.

2.5.2. Periodate and Proteinase K Treatments

To clarify whether the receptor for VpaE1 is a protein or a carbohydrate, VpaE1 adsorption to the proteinase K- or periodate-treated *E. coli* cells was tested by phage pull-down assay essentially as described by Kiljunen *et al.* [40]. For this, cultures of *E. coli* BL21 and BW25113, as well as BW25113

mutants $\Delta waag$ and $\Delta waap$ were grown in LB medium at 37 °C with vigorous shaking until mid-log (OD₆₀₀ of 0.5). For proteinase K (Thermo Fisher Scientific) treatment, 2 mL of each culture was collected by centrifugation at 14,000 × *g* for 1 min and the pellet was suspended in 2 mL of LB medium. The bacterial suspension was divided into two equal parts: one of these was incubated at 37 °C for 2 h (control sample), while the other was treated with Proteinase K (0.2 mg/mL) at 37 °C for 2 h.

For the periodate treatment, 2 mL of each culture was harvested by centrifugation, as described above. The pellet was then suspended in 1 mL of sodium acetate (50 mM, adjusted to pH 5.2) or 1 mL sodium acetate containing 100 mM of NaIO₄ (BDH Chemicals Ltd, Poole, England) and incubated at 37 °C for 2 h in the dark.

Treated cells were washed three times with 1 mL of LB medium. After the final wash, the cells were suspended in 1 mL of LB medium. Then, the A_{600} of the bacterial suspension was adjusted to ~0.5, and the phage adsorption assay was carried out. Phage adsorption was assessed by pull-down assay, where centrifugation removes phage particles bound to the bacterial cells. For this, treated bacterial suspensions were incubated with VpaE1 (MOI = 0.1) for 10 min at 37 °C, then cell-phage complexes were removed by centrifugation for 5 min at 5000× *g*, and the number of unadsorbed phage particles remaining in the supernatant was determined in triplicate by the double-layer agar method using *E. coli* BL21. The adsorption ratio was calculated from comparison to a negative control in LB medium without cells.

2.6. Phage DNA Isolation and Analysis

Aliquots of phage suspension $(10^{11}-10^{12} \text{ PFU/mL})$ were subjected to phenol/chloroform extraction and ethanol precipitation. Isolated phage DNA was subsequently used for restriction analysis, or was subjected to genome sequencing. Restriction digestion was performed with Bsu15I, XbaI, HhaI, Csp6I, EcoRV, and CfrI restriction endonucleases (Thermo Fisher Scientific) according to the supplier's recommendations. DNA fragments were separated by electrophoresis in a 0.8% agarose gel containing ethidium bromide. Restriction analysis was performed in triplicate to confirm the results.

2.7. Genome Sequencing, Assembly and Annotation

The complete genome sequence of VpaE1 was determined using Illumina DNA sequencing technology at BaseClear, the Netherlands. Open reading frames (ORFs) were predicted with Glimmer v2.02 [41] and Geneious v5.5.6. [42]. Analysis of the genome sequence was performed using the Fasta-Protein, Fasta-Nucleotide, Fasta-Genome, BLAST2, PSI-Search, Transeq, ClustalW2, (European Bioinformatics Institute, Hinxton, UK), programs, and tRNAscan-SE 1.21 [43] was used to search for tRNAs. The complete genome sequence of vB_EcoM_VpaE1 was deposited in the EMBL nucleotide sequence database under accession number KM657822.

2.8. Phylogenetic Analysis

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 5 [44]. The neighbor-joining tree based on the whole-genome sequence alignment was constructed using Geneious v5.5.6 [42], Genome sequences used for alignment can be found at NCBI: Felix_01 (AF320576); Mushroom (KP143762); FO1a (JF461087); vB_EcoM_AYO145A (KR014248); UAB_Phi87 (JN225449); HY02 (KM092515); O157_phage_12 (KP869110); O157_phage_11 (KP869109); WV8 (EU877232); EC6 (JX560968); JH2 (KF055347); HB-2014 (KP010413); O157_phage_1 (KP869100); phiEa21-4 (FQ482083); and Michonne (KT001916).

2.9. Analysis of VpaE1 Virion Structural Proteins

2.9.1. SDS-PAGE

CsCl-purified phage particles (~10¹⁰) resuspended in a buffer containing 60 mM Tris-HCl (pH 6.8), 1% SDS (w/v), 1% 2-mercaptoethanol (v/v), 10% glycerol (v/v) and 0.01% bromophenol blue (w/v) were boiled for 3 min and separated on 12% SDS PAGE following the method described by Laemmli [45]. Protein bands were visualized by staining with PageBlue Protein Staining Solution (Thermo Fisher Scientific).

2.9.2. Filter-Aided Protein Sample Preparation (FASP) for Mass Spectrometry Analysis

CsCl-purified phage particles were concentrated on Amicon Ultra-0.5 mL 30 kDa centrifugal filter unit and were denatured in 8 M urea, 100 mM DTT solution with continuous rotation at 800 rpm in the temperature controlled shaker for 3 hours at 37 °C. Trypsin digestion was performed according to a modified FASP protocol [46,47]. Briefly, phage particles were washed with buffer containing 8 M urea. The proteins were alkylated using iodoacetamide. Buffer was exchanged by washing twice with 50 mM NH₄HCO₃, and proteins digested overnight with TPCK Trypsin 20233 (Thermo Scientific, Rockford, IL, USA). Then the peptides were recovered by centrifugation and washed twice with 50% CH₃CN. Afterwards, the samples were combined, acidified, lyophilized, redissolved in 0.1% formic acid and analyzed by mass spectrometry.

2.9.3. Liquid Chromatography and Mass Spectrometry

Liquid chromatography (LC) separation of trypsin cleaved peptides was performed with a nanoAcquity UPLC system (Waters Corporation, Wilmslow, UK). Peptides were loaded on a reversed-phase trap column PST C18, $100A^{\circ}$ 5 µm, 180 µm × 20 mm (Waters Corporation) at a flow-rate of 15 µL/min using loading buffer of 0.1% formic acid and subsequently separated on HSS-T3 C18 1.8 µm, 75 µm × 250 mm analytical column (Waters Corporation) in 60 min linear gradient (A: 0.1% formic acid, B: 100% CH₃CN and 0.1% formic acid) at a flow rate of 300 nL per min. The analytical column temperature was kept at 40 °C.

The nano-LC was coupled online through a nanoESI 7 cm length, 10 mm tip emitter (New Objective, Woburn, MA, USA) with HDMS Synapt G2 mass spectrometer (Waters Corporation). Data were acquired using Masslynx version 4.1 software (Waters Corporation) in positive ion mode. LC-MS data were collected using data independent acquisition (DIA) mode in combination with online ion mobility separations. The trap and transfer collision energy for high-energy scans in HDMS mode was ramped from 4 to 5 eV and from 27 to 50 eV. For both analyses, the mass range was set to 50–2000 Da with a scan time set to 0.9 s. A reference compound [Glu1]-Fibrinopeptide B (Waters Corporation) was infused continuously (500 fmol/ μ L at flow rate 500 nL per min) and scanned every 30 s for online mass spectrometer calibration purposes.

2.9.4. Data Processing, Search and Analysis

Raw data files were processed and searched using ProteinLynx Global SERVER (PLGS) version 2.5.3 (Waters Corporation). The following parameters were used to generate peak lists: (i) minimum intensity for precursors was set to 150 counts; (ii) minimum intensity for fragment ions was set to 50 counts; (iii) intensity was set to 500 counts. Processed data was analyzed using trypsin as the cleavage protease, one missed cleavage was allowed and fixed modification was set to carbamidomethylation of cysteines, variable modification was set to oxidation of methionine. Minimum identification criteria included 1 fragment ion per peptide, 3 fragment ions per protein and minimum of 2 peptides per protein. The false discovery rate (FDR) for peptide and protein identification was determined based on the search of a reversed database, which was generated automatically using PLGS when global false discovery rate was set to 4%. The database comprising

all possible VpaE1 ORFs (>50 codons) was generated using Geneious v5.5.6 [42], and was used for the identification of VpaE1 structural proteins.

3. Results

3.1. Phage Isolation and Analysis of VpaE1 Plaque and Virion Morphology

Bacteriophage VpaE1 was isolated in Lithuania while screening for phages with distinct physiological properties [48–51]. In the course of this screening, no natural bacterial hosts were collected; therefore *E. coli* strain BL21 was used as a reference host for the isolation and propagation of VpaE1.

To investigate the effect of temperature on VpaE1 plaque formation, diluted phage suspensions were plated on *E. coli* BL21, and the plates were incubated at 9, 37 and 47 °C. As seen in Figure 1, at 37 °C the plaques produced by VpaE1 (~1.5 mm in diameter) were the most clear and uniform, therefore the subsequent phage propagation and adsorption experiments were carried out at this temperature.



Figure 1. The morphology of VpaE1 plaques produced on the lawns of *E. coli* BL21 at different temperatures. Plaques are shown after: (**A**) five days of incubation at 9 °C; (**B**) 24 h of incubation at 37 °C; and (**C**) 24 h of incubation at 47 °C.

It is worth mentioning that bacteriophage VpaE1 was also capable of multiplying at 49 $^{\circ}$ C, yet the plaques formed at this temperature were too small to be photographed.

When the content of a single plaque was added to a 10 mL of the exponentially growing liquid culture in LB medium (OD₆₀₀ of 0.5) lysis occurred within ~2 h. At 37 °C, *E. coli* BL21 was capable of producing phage titers of 10^{11} – 10^{12} pfu mL⁻¹. The morphology of CsCl-purified VpaE1 particles was analyzed using transmission electron microscopy (TEM) (Figure 2).



Figure 2. Electron micrographs of bacteriophage VpaE1. (**A** and **B.2**) CsCl density gradient-purified VpaE1 and (**B.1**) T4 particles; (**C**) VpaE1 particles with extended (**C.1**) and contracted (**C.2**) tail; and (**D**) close-up view of VpaE1 particle with tail fibers.

Electron micrographs showed that VpaE1 belongs to the family *Myoviridae* and is characterized by an isometric head of 77 nm in apical diameter and a necked contractile tail of 120 by 16 nm (in the extended form) with long tail fibers connected to the baseplate (Figure 2). In addition, the general appearance of VpaE1 particle is similar to that of other characterized Felix01-like viruses [23,33].

3.2. The Host Range of VpaE1

To investigate the host range of VpaE1, a quantitative spot dilution test was employed. As seen in Table 1, those *E. coli* strains that express O-antigen (STEC group) were completely resistant to VpaE1, as were both *Salmonella* and *Klebsiella* strains used in this study. Moreover, the test revealed that *E. coli* K-12 derivatives were also resistant to VpaE1, and only *E. coli* B strains were sensitive to the phage tested.

Strain	Spot Test	Source/Reference
E. coli B strains		
B^E	+	Lindsay W. Black
BL21	+	Novagen
BL21(DE3)	+	Novagen
B40	+	Lindsay W. Black
BE-BS	+	Kenneth N. Kreuzer/[52]
E. coli K12 strains		
MG1655	_	Edita Sužiedėlienė
K803	_	Lindsay W. Black
XL1blue	_	Stratagene
CR63	_	Kenneth N. Kreuzer/[52]
BW25113	_	[53]
E. coli STEC representatives		
O157:H7	_	Edita Sužiedėlienė
O103:H2	_	Edita Sužiedėlienė
O111:H8	_	Edita Sužiedėlienė
O145:H28	_	Edita Sužiedėlienė
O26:H11	_	Edita Sužiedėlienė
Other enterobacteria		
Salmonella Typhimurium LT2	_	Jaunius Urbonavičius
Klebsiella pneumoniae KV-3	_	[54]

Table 1. Host Range of vB_EcoM_VpaE1.

+: clear plaques were observed after incubation at 37 $^{\circ}$ C; -: no plaques.

3.3. Annotation and Overview of VpaE1 Genome

Since many of the genomes of Felix01-related phages have already been examined in detail and published, herein is provided only a brief overview of VpaE1 genome. Bacteriophage VpaE1 has a linear ds DNA genome of 88,403 bp that is similar in size to those of other Felix01-related viruses whose genome sizes range between 84 and 93 kb [19,23,24,27,29,32,33,35]. As is the case with other virulent phages, the VpaE1 genome has an average mol% G+C content of 38.9%, which differs significantly from that of the host (~50%). According to the literature, the drop in G+C content in combination with various forms of DNA modification is known to protect the phage DNA from cleavage by restriction enzymes of the host [55,56].

The restriction digestion analysis *in silico* revealed that the DNA of VpaE1 has no recognition sites for many common restriction enzymes (such as BamHI, XhoI, SaII, KpnI, SacI, PstI and even MboI, which recognizes GATC), a feature that is usually observed in phages with unmodified DNA. To test for the digestibility of VpaE1 DNA, 6 restriction enzymes with different sensitivity to DNA methylation were used, namely CfrI (affected by Dcm/CpG methylation), Bsu15I (sensitivity to Dam/CpG methylation), XbaI (sensitivity to Dam methylation), HhaI (affected by CG methylation), Eco32I and Csp6I (both not affected by Dcm, CpG or Dam methylation). Since all enzymes

digested the DNA of VpaE1 (Figure 3), restriction analysis indicated that VpaE1 genome contains no modified nucleotides.



Figure 3. Restriction digestion patterns of phage VpaE1 genomic DNA (gDNA). Lane M represents GeneRuler[™] 1 kb DNA Ladder (Thermo Fisher Scientific).

This is further supported by the results from bioinformatics analysis, which revealed that no characterized enzymes implicated in DNA modification have homologues in VpaE1. Bioinformatics analysis indicated that the genome of VpaE1 encodes 132 putative open reading frames (ORFs) and 20 tRNAs (Figure 4).

All VpaE1 ORFs had homologues in other sequenced members of *Felix01likevirus* (Table S1). As was the case with other members of *Felix01likevirus*, nearly 70% of VpaE1 ORFs encoded hypothetical proteins with no known functions. Thus, through the examination of homology search results, a putative function could be assigned to only 38 VpaE1 ORFs (Figure 4). In addition, MS/MS analysis allowed for the identification of 11 virion proteins, which were encoded by VpaE1 ORFs with homology to various hypothetical proteins only.

Viruses 2015, 7, 6163–6181

DE01 (dl)	500	1,000	1,500	2,000	2,500	3,000	3,500	4,000	4,500	5,000	5,500
	7)				ORFO	04 (putative PseT RF03 (hypotheti	polynucleotide (5'-kinase/3'-pho	osphatase) Ol	RF07 (hypotheti	cal protein)
00	6,500	7,000	7,500	8,000	8,500	9,000	9,500	ORF05 (I 10,000	hypothetical pro 10,500	otein) ORF06 (11,000	(hypothetical pro 11,500
	RF	0 OF	F13 (putative)	ohosphatase)	OR	-14 OR	ORF	RF1	ORF19		ORF21 (hyp
2 000	12 500	13 000	13 50	0 14.000	14.50	15.00	0 15 500	16.00	0 16.50	0 17.00	17.50
F23 (h	vnothetical pro	tein)	10,90	ORE26c	(hypothetical r	aratein) ORF27	c (hypothetical r	protein) -		ORE31c (hypo	thetical protein)-
22 (hyr	othetical prote	() ORF24 (h		ORF25 (hypothe	etical protein)	ORE28c (hyr	othetical protein	ORE29c (b)	vpothetical prot	ein) ORE30c	
18,00)0 18,	500 19,	ροο 19	,500 20,	,000 2	0,500 21	,000 21	,500 22	2,000 2:	2,500 23	3,000 23
ODES	Da //h	DE32 - /hum	ORF3	ORF36c		ORF38c (lysin)	E2De (autotion to	il anatain)		0	
UKES	2c (ny) < C	окгозе (пур		<u> </u>	<u>URF3/C</u>		rose (putative ta			41C (nyp	
24	4,000	24,500	25,000	25,500	26,000	26,500	27,000	27,500		28,500	
	ORF45 (hypoth	netical protein)						ORF46 (hypo	othetical proteir)	
44 (Hyp D0	30,000	30,500	31,000	31,500	32,000	32,500	33,000	33,500	34,000	34,500	35,000
	tR	VA-Cys	URFSU (termin	ase, large subur	11()	ORF51 (pu	tative structural	protein)			
47 (hyp	oothetical prote	ein) ORF48 (h	ORF49 (hypot) ypothetical pro	netical protein) tein)			ORF52	(hypothetical j	protein) ORFS	ORF54 (head	protein) maturation prote
500	36,000	36,500	37,000	37,500	38,000	38,500	39,000	39,500	40,000	40,500	41,000
	n		E (major canaj	(protoin)					61 (structural i	protein)	
	ODEEE (at	ORF5	o (major capsi		EEO (humathati		ODEC) (humatical	a set a in)	ODECO /h	on other time I would
54 (hea	ORF55 (str ad maturation	ORF5 uctural protein) protease)	o (major capsi	ORI ORF57 (h	F59 (hypotheti ypothetical pro	ical protein)	ORF60	0 (hypothetical tein)	protein)	ORF62 (h)	ypothetical prote
F54 (hea 41,500	ORF55 (str ad maturation 42,00	ORF5 uctural protein) protease) 0 42,50	0 43,0	ORF ORF57 (h 00 43,50	F59 (hypotheti ypothetical pro)0 44,(ical protein) otein) ORF58 (200 44,5	ORF60 hypothetical prot 00 45,00	0 (hypothetical tein))0 45,5	protein) 00 46,0	ORF62 (h)	ypothetical prote
F54 (hea 41,500	ORF55 (str ad maturation) 42,00	ORF5 uctural protein) protease) 0 42,50 ORF65 (tail tape	0 43,0 -measure prote	ORI ORF57 (h 00 43,50	F59 (hypotheti ypothetical pro 00 44,(ical protein) otein) ORF58 (000 44,5	ORF60 hypothetical prot 00 45,00	D (hypothetical iein) D 45,5	protein) 00 46,	ORF62 (h)	ypothetical prote
-54 (hea 41,500 -62 (hyp	ORF55 (str ad maturation) 42,00 contetical prote	ORF5 uctural protein) protease) 0 42,50 ORF65 (tail tape ORF64 (hypoth- ein) ORF63 (h	0 43,0 -measure prote etical protein) ypothetical prot	OR ORF57 (h 00 43,50 iin)	F59 (hypotheti ypothetical pro 00 44.(ORF66 (ical protein) otein) ORF58 (000 44,5 (hypothetical pro	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h	D (hypothetical tein))0 45,5 hypothetical pro RF68 (hypotheti	protein) 00 46,ç itein) tical protein)	ORF62 (h)	ypothetical prote
54 (hea 41,500 62 (hyp 47,5	ORF55 (str ad maturation 42,00 Doothetical prote	ORF5 uctural protein) protease) 0 42,50 0 0RF65 (tail tape 0RF64 (hypoth- sin) 0RF63 (h ,000 48	0 43,0 -measure prote etical protein) ypothetical pro	ORI ORF57 (h 00 43,50 iin) tein) 9,000 49	F59 (hypotheti ypothetical pro 00 44 (ORF66 0	ical protein), tein) ORF58 (000 44,5 (hypothetical pro 50,000 5	ORF60 hypothetical prot 00 45,00 tein) ORF67 (H OI 0,500 51	D (hypothetical tein) 00 45,5 hypothetical pro RF68 (hypothet 1,000 5	protein) 00 46.(itein) tical protein) (1,500 5	ORF62 (h) 000 46,5 ORF70 (h) ORF69 (basepla 52,000 5	ypothetical prote 500 47.0 ypothetical prote te assembly pro 52.500 53
54 (hea 41,500 62 (hyp 47,5	ORF55 (str ad maturation 42,00 Contentical prote 500 48 RF71 (baseplat	ORF5 uctural protein) orotease) 0 42.50 ORF65 (tail tape ORF64 (hypoth- in) ORF63 (h ,000 48 e component)	0 (trajor capsion 0 43,0 -measure prote etical protein) ypothetical prot ;500 4	ORF57 (h ORF57 (h 300 43,50 in) tein) 9,000 45	F59 (hypotheti ypothetical pro 00 44.0 ORF66 0	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 (ORF74 (tail	ORF6(hypothetical prot 00 45,00 tein) ORF67 (h O 0,500 51 fiber protein)	0 (hypothetical ein) 10 45,5 hypothetical pro RF68 (hypothet 1,000 5 ORF	protein) 00 46.ç itein) iical protein) (1,500 § 75 (tail fiber pro	ORF62 (h) 000 46.5 ORF70 (h) ORF69 (basepla 52,000 5 stein)	ypothetical prote 500 47 (2 ypothetical prote te assembly pro 52,500 53
62 (hea 62 (hyp 47,5 0R RF70 (hy	ORF55 (str ad maturation j 42,00 oothetical prote 500 48 RF71 (baseplat ypothetical pro	ORF5 uctural protein) orotease) 0 42.50 ORF65 (tail tape ORF64 (hypoth in) ORF63 (h 000 48 e component) otein)	0 (Inajor Capara -measure prote etical protein) ypothetical pro .500 4 ORF OF	ORF57 (h ORF57 (h 00 43,50 in) tein) 9,000 45 73 (hypothetical RF72 (hypothetical	F59 (hypothetic ypothetical pro 00 44.0 ORF66 (9,500 (protein)	ical protein) otein) ORF58 (000 44,5 (hypothetical pro 50,000 5 ORF74 (tail	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h Ol 0,500 51 fiber protein)	0 (hypothetical lein) 10 45.5 hypothetical pro RF68 (hypothet 1,000 5 ORF	protein) 00 46., tein) tical protein) (1.500 £ 75 (tail fiber pro	ORF62 (h) 000 46.5 ORF70 (h) 0RF70 (h) 0RF69 (basepla 52,000 £ 0tein)	ypothetical prote i00 47.0 ypothetical prote ite assembly pro 52,500 53
54 (hea 41,500 62 (hyp 47,5 0R RF70 (hy 5	OFF55 (str ad maturation 42,00 bothetical prote 500 48 F71 (baseplat ypothetical prot 53,500	ORF5 uctural protein) orrotease) 0 42,50 VRF65 (tail tape QRF64 (hypoth- inin) QRF63 (h QR00 48 e component) tein) 54,000 40	0 43.0 -measure prote etical protein) ypothetical protein) (500 4 CRF ORF 54.500	ORF57 (h ORF57 (h 00 43.50 in) tein) 9,000 45 9,000 45 73 (hypothetical 1772 (hypothetical 55,000	F59 (hypothetic ypothetical pro 0 44. ORF66 0 9,500 4 protein)- al protein) 55,500	ical protein) otein) ORF58 (000 44,5 (hypothetical pro 50,000 5 ORF74 (tail 56,000	ORF60 hypothetical prot 00 45,00 tein) ORF67 (0,500 51 fiber protein) 56,500	0 (hypothetical lein) 10 45.5 hypothetical pro RF68 (hypotheti 1,000 5 0RF 57.000	protein) 00 46.(ticin) tical protein) (1.500 5 75 (tail fiber pro	ORF62 (h) ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 £ otein) 58,000	ypothetical prote i00 47.0 ypothetical prote te assembly pro 52.500 53 58.500
54 (hea 41,500 62 (hyp 47,5 8F70 (hy 5	ORF55 (str ad maturation 42,00 contestical prote 500 48 CF71 (baseplat ypothetical prot 53,500 - ORF75 0.	ORF5 uctural protein) 0 42,50 0 42,50 0 42,50 0 42,50 0 26,764 0 47,50 0 48 0 48 0 48 0 48 0 48 0 54,000 5 (tail fiber prote 0 000000000000000000000000000000000000	0 43,0 -measure prote etical protein) ypothetical prot ;500 4 ORF OF 54,500	ORF57 (h ORF57 (h 00 43,50 in) 9,000 45 73 (hypothetical RF72 (hypothetic 55,000	F59 (hypothetic ypothetical pro 0 44., ORF66 1 9,500 4 protein) 55,500 55,500	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 CRF74 (tail 56,000 79c (dihydrofolat OR ORF81c	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 01 0,500 51 fiber protein) 56,500 e reductase) C c (h, ORF82.	0 (hypothetical lein) 10 45,5 hypothetical pro RF68 (hypotheti 0,000 5 0RF 57,000	protein) 00 46.0 ticin) tical protein) (1 1.500 5 75 (tail fiber pro 57.500 riptional regulat	ORF62 (h) 000 46.5 ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 £ 000 58,000 000 000 000 000 000 000 000 000 000	ypothetical prote i00 47.0 ypothetical prote tel assembly pro 52,500 53 58,500 REBGC () ORE
54 (hea 41,500 62 (hyp 47,5 (hyp (F70 (hyp) 5	ORF55 (str ad maturation 42,00 cothetical prote 500 48 RF71 (baseplat ypothetical prot 53,500 ORF72 O	ORF5 uctural protein) orrotease) 0 42,50 VRF65 (tail tape QRF64 (hypoth- sin) QRF63 (h QRF00 48 e component) tein) 54,000 5 5 (tail fiber prote QRF7	0 43.0 -measure prote etical protein) ypothetical pro ,500 4 ORF OF 54.500 in) ORF78c	ORF57 (h ORF57 (h 00 43,50 iin) 9,000 45 73 (hypothetical 73 (hypothetical 75,000 (thymidylate syn	F59 (hypothetic ypothetical pro- 00 44.0 ORF66 0 9.500 9 protein)- al protein) 55.500 thase) ORF3	ical protein) otein) ORF58 (000 44,5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat ORF81c 62,000	ORF60 hypothetical prot 00 45,00 tein) ORF67 (fr Ol 0,500 51 fiber protein) 56,500 e reductase) C c (h ORF82.	0 (hypothetical lein) 10 45.5 10 45.5	protein) 00 46.(tein) tical protein) (1.500 5 75 (tail fiber pro 57.500 riptional regulat ORF85 63.500	ORF62 (h) ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 c otelin) 58,000 ory protein)	ypothetical prote 300 47.0 ypothetical prote te assembly pro 52.500 53 58.500 RF86c () ORF 64.500
54 (hez 62 (hyp 47,5 F70 (h) Ε 6 6 6 6 6 6 6 6 6 6 6 6 6	ORF55 (str ad maturation 42,00 contestical prote 500 48 CF71 (baseplat ypothetical prot 53,500 ORF75 0 59,500	ORF5 uctural protein) 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 48 e component) 48 e component) 48 54.000 5 50 (tail fiber prote ORFF77 60.000 ORF89c	0 43,0 -measure prote etical protein) ypothetical prot ,500 4 ORF OF 54,500 in) ORF78c 60,500 (DNA ligase)	ORF57 (h ORF57 (h 00 43,50 in) 9,000 45 73 (hypothetical RF72 (hypothetic 55,000 10 10 10 10 10 10 10 10 10 10 10 10	F59 (hypothetic ypothetical pro 0 44.0 ORF66 0 9.500 9 protein) 55.500 61.500	ical protein) otein) ORF58 (0 000 44.5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat OR. 0RF81c 62,000	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 0) 0,500 51 fiber protein) 56,500 e reductase) C0 c (h ORF82. 62,500	0 (hypothetical lein) 10 45,5 10 45,5	protein) 00 46.0 tein) tical protein) (1 1.500 5 75 (tail fiber pro 57.500 fiptional regulat 0RF85 63.500 INA polymerasu	ORF62 (h) ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 £ tetin) 58,000 ory protein) 64,000 e)	ypothetical prote i00 47 c ypothetical prote te assembly pro 52,500 5: 58,500 RF86c () ORF 64,500
54 (hes 41,500 47,5 0R (F70 (h) 5 00	ORF55 (str ad maturation 42,00 oothetical prote 500 48 CF71 (baseplat ypothetical prot 53,500 ORF7(0 59,500	ORF5 uctural protein) orotease) 0 42.50 VRF65 (tail tape QRF64 (hypoth- in) QRF63 (h QRF64 (hypoth- in) QRF63 (h 0 48 e component) 48 54.000 5 54.000 60.000 QRF77 60.000 QRF89c QRF88c	0 43,0 -measure prote etical protein) ypothetical prot 500 4 0RF 0F 54,500 in) ORF78c 60,500 (DNA ligase) (hypothetical p	ORF57 (h ORF57 (h 00 43,50 in) tein) 9,000 45 73 (hypothetical RF72 (hypothetic 55,000 (thymidylate syn 61,000 ORF92c rotein)	F59 (hypothetic ypothetical pro 0 44.0 ORF66 0 9,500 9 protein)- al protein) 55,500 61,500 61,500 c (DNA polyme	ical protein) otein) ORF58 (000 44,5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat ORF81c 62,000 erase) ORF93c	ORF60 hypothetical prot 00 45,00 tein) ORF67 (t 00 0,500 51 fiber protein) 56,500 e reductase) C (h. ORF82. 62,500 e (endonuclease	0 (hypothetical lein) 10 45.5 10 45.5 10 45.5 10 45.5 10 45.5 10 45.5 10 45.5 10 45.5 10 45.5 10 0 10 10 10 0 0 10 10 0 0 10 10 0 0 10 10 0 0 0	protein) 00 46.0 tein) tical protein) (1 1.500 57 75 (tail fiber pro 57,500 riptional regulat (ORF85) 63.500 INA polymerase othetical protei	ORF62 (h) ORF60 (h) ORF70 (h) ORF69 (basepla 52,000 e) 64,000 e) n)	ypothetical prote i00 47.0 ypothetical prote te assembly pro 58,500 58,500 RFB6c () ORF 64,500
54 (hes 41,500 62 (hyp 47,5 67 67 67 67 67 67 67 67 67 67 67 67 67	ORF55 (str ad maturation 42,00 contestical prote 500 48 CF71 (baseplat ypothetical prot 53,500 ORF75 0 59,500	ORF5 uctural proteins) 0 42,50 0 42,50 0 42,50 0 42,50 0 42,50 0 42,50 0 46,000 48 48 0 48 0 00 48 000 48 000 54,000 48 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 000 0 00	0 43.0 measure prote- etical protein) ypothetical pro- 54.500 4 0RF 0F 54.500 60.500 (DNA ligase) (hypothetical pro- 0RF78c 60.500 (DNA ligase) 0RF90c (66.500	ORF57 (h ORF57 (h 00 43,50 in) 9,000 45 9,000 45 73 (hypothetical 773 (hypothetical)	F59 (hypothetic ypothetical pro 0 44.5 ORF66 0 9,500 9 protein)- al protein) 55,500 thase) ORF 61,500 c (DNA polyme tein) ORF911 67,500	ical protein) otein) ORF58 (000 44.5 000 44.5 000 50,000 5 0000 5 000000 5 0000 5 00000 5 0000000000	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h O) 0,500 511 56,500 e reductase) C c (h. ORF62. 62,500 c (endonuclease rotein) 68,500	0 (hypothetical lein) 0 45,5 0 45,5 0 45,5 0 45,5 0 45,5 0 0 45 0 0 10 0 0 0 0	protein) 00 46.0 tein) tical protein) (1 1.500 5 75 (tail fiber protein) 57.500 riptional regulat ORF85 63.500 INA polymerase othetical protein 69.500	ORF62 (h) ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 £ otein) 58,000 ory protein) 64,000 e) n) 70,000	ypothetical prote 300 47.0 ypothetical prote te assembly pro 58,500 358,500 RFB6c () ORF 64,500 0 70,500
54 (hes 41,500 62 (hyp 47,5 CR 20 20 87 c (hy 87 c (hy 87 c (hy 87 c (hy) 87 c (hy) 900	ORF55 (str ad maturation 42,00 oothetical prote 500 48 F71 (baseplat ypothetical pro 53,500 ORF74 O 59,500	ORF5 uctural protein) 0 42.50 0 42.50 0 42.50 0 42.50 0 27.64 0 7.65 0 7.65 0 48 0 48 0 0.00 5 (tail fiber prote 0 0.00	0 43,0 -measure prote- etical protein) ypothetical prot- ;500 4 ORF OF 54,500 in) ORF78c 60,500 (DNA ligase) (hypothetical p ORF90c (66,500	ORF57 (h ORF57 (h 00 43,50 in) 9,000 45 73 (hypothetical 272 (hypothetic 55,000 173 (hypothetic 55,000 174 (hypothetical 275,000 174 (hypothetical 275,000 174 (hypothetical 276 (hypothetical 277 (hypothetical 279 (hypothetical 270 (hypothetical 2	F59 (hypothetic ypothetical pro- 00 44., ORF66 (9,500 5 protein)- al protein) 55,500 61,500 61,500 c (DNA polyme tein) ORF91(67,500 00 c (DNA polyme	ical protein) otein) ORF58 (0 000 444,5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat OR ORF81c 62,000 carase) ORF93c c (hypothetical pro 68,000 08F100 (DNA print)	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 0) 0,500 51 fiber protein) 56,500 e reductase) C0 c (h ORF82, 62,500 c (endonuclease) c (endonuclease) c (endonuclease)	0 (hypothetical lein) 10 45.5 10 45.5	protein) 00 46.0 tein) tical protein) (1 1.500 £ 75 (tail fiber protein) 57.500 riptional regulat ORF85 63.500 INA polymerase othetical protein 69.500 proteinial proteini	ORF62 (h) ORF70 (h) ORF70 (h) ORF70 (basepla 52,000 £ 58,000 orry protein) 0 64,000 e) n) 0 70,000	ypothetical prote i00 47 c i00 47 c i0 47 c i0 50 5 i0 58,500 i0 70,500 i0 70,500
54 (hes 41,500 62 (hyp 47,5 0 CR RF70 (hy ε ε ε ε ε ε ε ε ε ο α 00	ORF55 (str ad maturation 42,00 000000000000000000000000000000000	ORF5 uctural protein) orotease) 0 42,50 VRF55 (tail tape QRF64 (hypoth- sin) QRF63 (h QRF64 (hypoth- sin) QRF63 (h QRF65 (tail tape QRF63 (h Component) S4,000 54,000 5 Correst QRF72 60,000 QRF89c Correst QRF89c S5(b) G6,000 (se) QRF89c	0 43,0 -measure prote etical protein) ypothetical pro- .500 4 SORF ORF OF 54,500 in) ORF78c 60,500 (DNA ligase) (hypothetical p ORF90c (66,500 (DNA ligase) (hypothetical p	ORF57 (h ORF57 (h 00 43,50 in) itein) 9,000 45 73 (hypothetical RF72 (hypothetic 55,000 (thymidylate syn 61,000 ORF92c rotein) GRF92 (dMMP kin	F59 (hypothetic ypothetical pro 0 44.0 ORF66 0 9,500 9 protein)- 1 protein) 55,500 61,	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat OR ORF81c 62,000 c (hypothetical pro 50,000 CRF93 (hypothetical pro 50,000 CRF9	ORF60 hypothetical prot 00 45,00 tein) ORF67 (f 00 0,500 51 fiber protein) 56,500 e reductase) C0 (n. ORF62, 62,500 c (endonuclease) rotein) 68,500 asse/helicase) cal protein)	0 (hypothetical lein) 10 45.5 10 45.5 1,000 5 1,000 5 57,000 0RF63 (hypothetical pro- 0RF63 (hypothetical pro- 0RF63 (hypothetical pro- 0RF95 (hypothetical pro- 0RF	protein) 00 46.0 tein) tical protein) (1 1.500 5 75 (tail fiber protein) 57.500 75 (tail fiber protein) 63.500 INA polymerase 69.500 pothetical protein	ORF62 (h) ORF60 (h) ORF69 (basepla 52,000 c stein) 58,000 64,000 e) n) 0 70,000 cim)	ypothetical prote i00 47.0 ypothetical prote te assembly pro- 58.500 RF86c () ORF 64.500 0 70.500
F54 (hes 41,500 47,5 Cor RF70 (h) ε ε ε ε ε ε ε α α α α α α α α α α α α	ORF55 (str ad maturation 42,00 contestical prote 500 48 CF71 (baseplat ypothetical prote 53,500 ORF75 0 59,500 DNA polymera 5 (hypothetica 0 71,50	ORF51 uctural proteins) 0 42,50 0 42,50 0 42,50 0 42,50 0 42,50 0 42,50 0 42,50 0 42,50 0 48 0 0 0 48 e component) 48 ottein) 54,000 54,000 60,000 ORF89c ORF88c 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 72,0	0 43.0 -measure prote- etical protein) ypothetical pro- .500 4 ORF OF 54.500 (DNA ligase) (hypothetical pro- ORF78c 60.500 (DNA ligase) (hypothetical pro- ORF90c (66.500 ORF90c (66.500 ORF90c (66.700 ORF90c (66.700 ORF90c (66.700 ORF90c (66.700 ORF90c (60.700 ORF90c (00.700 ORF90c (00.700) ORF90c	ORF57 (h ORF57 (h 00 43,50 iin) 10 43,50 iin) 10 43,50 10 44,50 10 44,50 100000000000000000000000000000000000	F59 (hypothetic ypothetical pro 0 44., ORF66 0 9,500 9 protein) 55.500 thase) ORF31 61.500 c (DNA polyme tein) ORF910 67.500 ase) 0 0 73	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat OR. ORF91c 62,000 erase) ORF93c c (hypothetical p 0 68,000 Er100 (DNA prim RF99 (hypotheti ,500 74.)	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 01 0,500 51 fiber protein) 56,500 e reductase) C c (h ORF82. 62,500 c (endonuclease rotein) 68,500 asse/helicase) cal protein) 00 00 00 00 00 00 00 00 00 0	0 (hypothetical lein) 0 45,5 0 45,5 0 45,5 0 45,5 0 45,5 0 45,5 0 0 45 0 0 45 0 0 10 0 0 0 10 0 0 0 10 0 0 0 10 0 0 0 10 0 0 0 0	protein) 00 46.0 tein) tical protein) (1 1.500 5 75 (tail fiber protein) 57.500 riptional regulat . ORF85 63.500 INA polymerasi 63.500 INA polymerasi 69.500 pothetical protein 69.500 pothetical protein 69.500 pothetical protein 69.500 00 75.500 00 75.5000 00 75.5000 00 75.5000 00 75.5000 00 75.50000 00 75.5000 00 75.50000 00 75.50000 00 75.50000 00 75.50000 00 75.500000 00 75.5000000000000000000000000000000000	ORF62 (h; ORF60 (h; ORF69 (basepla 52,000 c otein) 58,000 ory protein) 64,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 70,000 e) 70,000 70,000 e) 70,0000 70,0000 70,0000 70,0000 70,00000000	ypothetical prote i00 47.0 ypothetical prote te assembly pro- 58,500 CRFB6c () ORF 64,500 0 70,500 (hypothetical pro- (hypothetical pro- 0 76,500
54 (hea 41,500 62 (hyp 47,5 6 6 6 7 7 7 1,000 6 7 1,000 6 7 1,000 6 7 1,000 6 7 1,000 6 7 1,000 6 7 1,000 7 1,000 6 7 1,000 1,0000 1,00000 1,00000000	OFF55 (str ad maturation 42,00 42,00 00 00 00 00 00 00 00 00 00	ORF5 uctural protein) 0 42.50 0 42.50 0 42.50 0 42.50 0 742.50 0 42.50 0 742.50 0 48 e component) 48 e component) 48 c (tail fiber protein) 54,000 5 (tail fiber protein) 0RF89c 0 0RF88c tein) 66,000 ise) 0RF00 uprotein) 0RF	0 43,0 -measure prote- etical protein) ypothetical prot ;500 4 ORF OF 54,500 in) ORF78c 60,500 (DNA ligase) (hypothetical p ORF90c (66,500 OI 97 (hypothetical 00 72,5	ORF57 (h ORF57 (h 00 43.50 in) tein) 9,000 45 73 (hypothetical RF72 (hypothetic 55,000 (hypothetical RF72 (hypothetic 55,000 (hypothetical prot 61,000 ORF92c rotein) 67,000 RF98 (dNMP kin al protein) 500 73,0	F59 (hypothetic ypothetical pro- 00 44., ORF66 0 9,500 9 protein)	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat OR ORF91c 62,000 erase) ORF93c c (hypothetical pro 56,000 0 ORF91c 62,000 0 68,000 RF199 (hypotheti 5,500 74.)	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 0) 0,500 51 fiber protein) 56,500 e reductase) C0 c (h ORF82 62,500 c (endonuclease) c (endonuclease)	0 (hypothetical lein) 10 45.5 10 45.	protein) 00 46.0 tein) tical protein) (1 1.500 £ 75 (tail fiber protein) 57.500 riptional regulat ORF85 63.500 INA polymerase othetical protein 69.500 pothetical protein pothetical protein (0) 75.00 pothetical protein (0) 75.00 (0) 75.00	ORF62 (h) ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 £ 58,000 ory protein) 0 64,000 e) n) 0 70,000 ein) ORF103 (5,00 76,)	ypothetical prote i00 47.0 ypothetical prote te assembly pro 58,500 58,500 64,500 0 70,500 (hypothetical pro (hypothetical pro 0,00 76,
-54 (hes 41,500 -62 (hyp; 47,5 - 	ORF55 (str ad maturation 42,00 000000000000000000000000000000000	ORF5 uctural protein) orotease) 0 42,50 VRF55 (tail tape QRF64 (hypoth- sin) QRF64 (hypoth- sin) QRF00 48 e component) 54,000 5 (tail fiber prote ORF89c ORF89c CORF77 60,000 vspinotein) ORF89c ORF89c ORF89c ORF89c ORF00 vspinotein) ORF2,00	0 43,0 -measure prote- etical protein) ypothetical pro- i,500 4 0RF 0F 54,500 in) ORF78c 60,500 (DNA ligase) (Nypothetical pro- ORF90c (66,500 0 0 0 0 0 0 0 0 0 0 0 0	OR ORF57 (h 00 43,50 in) tein) 9,000 45 73 (hypothetical RF72 (hypothetic 55,000 (thymidylate syn 61,000 ORF92c rotein) € 00 F72c 67,000 RF98 (dNMP kin al protein) 500 73,0 cal protein) €	F59 (hypothetic ypothetical pro- 00 44., ORF66 0 9,500 9 protein)- 31 protein) 55,500 0 61,500 0 61,50	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 73c (dihydrofolat OR ORF81c 62,000 erase) ORF93c c (hypothetical pro 58,000 RF100 (DNA prin RF99 (hypotheti 500 74.1	ORF60 hypothetical prot 00 45,00 tein) ORF67 (fr 0,500 51 fiber protein) 56,500 c reductase) CC (h ORF62. (endonuclease) CC total protein) 68,500 cal protein) cal protein) 000 74,5	0 (hypothetical lein) 10 45.5 1,000 45.5 1,000 5 57,000 0RF68 (hypotheti 1,000 5 57,000 0RF95c (D 0RF95c (D 0RF95c (D 0RF95c (L) 0RF95c (D 0RF95c (L) 0RF95c (L)	protein) 00 46.0 tein) tical protein) (1 1.500 5 75 (tail fiber protein) 57.500 riptional regulat ORF85 63.500 INA polymerase othetical protein 69.500 pothetical protein 1.000 75, thetical protein otein) ide trial	ORF62 (h) ORF60 (h) ORF69 (basepla 52,000 c ory protein) 64,000 e) n) 0 70,000 ein) 0 RF103 (500 76,)	ypothetical prote i00 47.0 ypothetical prote te assembly pro 58.500 58.500 RF86c () ORF 64.500 0 70.500 (hypothetical pro (hypothetical pro
-54 (hes 41,500 47,5 (OR RF70 (h) ε ε ε ε ε ε ε α α α α α α α α α α α α	ORF55 (str ad maturation 42,00 42,00 0000000000000000000000000000000000	ORF51 uctural protein) 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 48 0 0.00 48 0.00 0 48 0 0.00 54,000 5 54,000 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 72.00 protein) 0.000 0.000 0.000 0.000 72.00 protein) 0.000 0.000 0.000 0.000 0.000	0 43.0 -measure prote- etical protein) ypothetical pro- .500 4 ORF OF 54.500 .00 .00 .00 .00 .00 .00 .00	ORF57 (h ORF57 (h 00 43,50 in) 9,000 45 73 (hypothetical R72 (hypothetical R72 (hypothetical R72 (hypothetical R72 (hypothetical S5,000 ORF92 (h) 00 RF98 (d) NAD synthetase 78,500 7	F59 (hypothetic ypothetical pro- 00 44., ORF66 (9,500 9 protein)	ical protein) otein) ORF58 (000 44.5 (hypothetical pro 50,000 5 ORF74 (tail 56,000 79c (dihydrofolat OR ORF91c 62,000 erase) ORF93c c (hypothetical prot 50,000 74.1 500 74.1 hypothetical prot 50,000 74.1	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 01 0,500 511 fiber protein) 56,500 e reductase) CC (h ORF62. 62,500 c (endonuclease) rotein) 68,500 cal protein) 00 74,5 ORF109 (ein) 80,000 8	0 (hypothetical lein) 0 45,5 0 45,5 0 45,5 0 0 45,5 0 0 45,5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	protein) 00 46.0 tein) tical protein) 0 1,500 5 75 (tail fiber pro- 57,500 riptional regulat 0RF85 63,500 INA polymerasi 69,500 NA polymerasi 69,500 pothetical protein 00 75, thetical protein 00 75, thetical protein 1,000 75	ORF62 (h) ORF60 (basepla 52,000 6 otein) 58,000 ory protein) 64,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) ein) 0 70,000 e	ypothetical prote i00 47,0 ypothetical prote te assembly pro 58,500 58,500 RFB6c () ORF 64,500 0 70,500 (hypothetical pro pothetical prote 64,500 0 70,500 (hypothetical pro asse, alpha subu 82,000 82
-54 (hes 41,500 -62 (hyp; 47,5 - 	ORF55 (str ad maturation 42,00 42,00 000000000000000000000000000	ORF5 uctural protein) 0 42.50 0 42.50 0 42.50 0 42.50 0 750 0 42.50 0 42.50 0 42.50 0 7.50	0 43,0 -measure prote- etical protein) ypothetical prot- i,500 4 - 0RF 0RF 0RF 0RF 0RF 0RF 0RF 0RF	ORF57 (h ORF57 (h 00 43,50 in) tein) 9,000 45 73 (hypothetical 172 (hypothetical 172 (hypothetical 172 (hypothetical 173 (hypothetical 174 (hypothetical 174 (hypothetical 175,000 000 000 000 000 000 000 000	F59 (hypothetic ypothetical pro- 00 44., ORF66 (9,500 9 protein)- al protein) 55,500 61,500 61,500 61,500 61,500 61,500 61,500 61,500 61,500 67,500 67,500 00 73 00 73	ical protein) otein) ORF58 (000 44,5 000 44,5 000 55,000 5 000 55 000 55 000 55 000 75 000 75 0000 75 000 75 0000	ORF60 hypothetical prot 00 45,00 tein) ORF67 (from the tein) 0,500 51 fiber protein) 56,500 e reductase) CC (h ORF62 (endonuclease) CC (endonuclease) CC (endonuclease) CC (all protein) 68,500 all protein) CC 000 74,5 ORF109 (sim) S0,000 ORF117 (MobE CORF117 (MobE)	0 (hypothetical lein) 10 45.5 10 45.5 10 45.5 1,000 5 0 RF68 (hypothetical pro RF68 (hypothetical pro 0 RF63c (transcr 0 0 RF95c (D 0 0 RF95c (D 0 0 RF95c (D 0 0 RF95c (L) 0	protein) 00 46.0 tein) tical protein) 0 1,500 5 75 (tail fiber pro- 57,500 riptional regulat ORF85 63,500 iNA polymerase othetical protein 69,500 pothetical protein pothetical protein thetical protein otein) nucleoside tripl 81,000 ing endonucles	ORF62 (h) ORF60 (h) ORF70 (h) ORF69 (basepla 52,000 64,000 e) n) 0 70,000 ein) ORF103 (500 76,) hosphate reduct 81,500 ase) 0	ypothetical prote i00 47.0 ypothetical prote te assembly pro 58,500 58,500 64,500 0 70,500 (hypothetical pro 0 70,500 (hypothetical pro 0 70,500 0 70,500
54 (hee 41,500 62 (hyp; 47,5 600 6870 (hy 5,000 68750 (hy 5,000 600 600 600 600 600 600 600 600 600	ORF55 (str ad maturation 42,00 vothetical prote ioothetical prote i	ORF5 uctural protein) orotease) 0 42,50 VRF55 (tail tape QRF64 (hypoth- in) QRF64 (hypoth- in) QRF64 (hypoth- in) QRF64 (hypoth- in) Stain QRF64 (hypoth- in) 54,000 48 CORF00 48 CORF77 60,000 ORF77 60,000 ORF88c 00 Iprotein) ORF88c Iprotein) ORF 00 72,0 protein) ORF OO 72,0 Iprotein) ORF Stain 00 T,500 7	0 43.0 -measure prote- etical protein) ypothetical prot- .500 4 ORF OF 54.500 (DNA ligase) (hypothetical pro- ORF78c 60.500 (DNA ligase) (hypothetical pro- ORF78c 60.500 (DNA ligase) 97 (hypothetical pro- ORF70c (66.500 ORF70c (66.500 ORF70c (60.500 ORF70c (00.500 ORF70c (00.500) ORF7	ORF57 (h ORF57 (h 00 43,50 in) tein) 9,000 45 73 (hypothetical F72 (hypothetical F72 (hypothetical F72 (hypothetical F72 (hypothetical F72 (hypothetical F73 (hypothetical F74 (hypothetical	F59 (hypothetic ypothetical pro- 00 44., ORF66 0 9,500 9 protein)- al protein) 55,500 thase) ORF10 (1,500 c (DNA polyme tein) ORF910 67,500 c (DNA polyme tein) ORF910 67,500 00 73	ical protein) otein) ORF58 (000 44.5 (hypothetical pro- 50,000 5 66,000 79c (dihydrofolat ORF74 (tail 56,000 79c (dihydrofolat ORF93c 62,000 62,000 63,000 84,000 84,000 85,000 74,1 10,000	ORF60 hypothetical prot 00 45,00 00 60,500 56,500 61,00 66,500 62,500 62,500 68,500 68,500 68,500 1ase/helicase 68,500 1ase/helicase 68,500 000 74,5 000 74,5 000 74,5 000 74,5 000 74,5 000 74,5 000 74,5 000 74,5	0 (hypothetical lein) 10 45.5 1,000 45.5 1,000 5 57,000 0RF68 (hypothetical 0RF95c (0 0RF95c (0 0 0RF95c (0 0 0RF95c (0 0 0RF95c (0 0 0 0 0 0 0 0 0 0 0 0 0 0	protein) 00 46.0 tein) tical protein) (1 1.500 5 75 (tail fiber protein) 57.500 75 (tail fiber protein) 63.500 INA polymerasu 63.500 INA polymerasu 69.500 pothetical protein 69.500 pothetical protein 1.000 75, thetical protein totein) nucleoside tripl 81.000 00E122 (protein) 00E122 (protein)	ORF62 (h) ORF60 (h) ORF69 (basepla 52,000 c ory protein) 64,000 e) n) 0 70,000 e) 0 70,000	ypothetical prote i00 47,0 ypothetical prote te assembly pro 58,500 58,500 64,500 0 70,500 (hypothetical pro 64,500 0 70,500 (hypothetical pro 64,500 0 70,500 0 70,500 10 70,500 0 70,500 0,
F54 (hes 41,500 62 (hyr 47,5 60 60 60 60 71,000 60 71,000 60 71,000 60 71,0000 71,0000 71,0000 71,00000 71,0000 71,0000 71,0000000000	ORF55 (str ad maturation 42,00 oothetical prote 500 48 F71 (baseplat ypothetical prote 53,500 ORF7 ORF7 59,500 DNA polymera 5 (hypothetical 0 71,50 5 (hypothetical 0 71,50 5 (hypothetical 0 71,50 5 (hypothetical 0 71,50 5 (hypothetical 0 71,50 5 (hypothetical 0 71,50 5 (hypothetical 5 (hypothet	ORF5 uctural protein) 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 42.50 0 48 0 48 e component) train) 54.000 5 (tail fiber prote 0 0RF68c tein) 66.000 0 72.0 vclease) 0 protein) 0RF 00 72.0 vclease) 7.500 riphosphate red 83.500	0 43,0 -measure prote- etical protein) ypothetical prot- is,500 4 ORF OF 54,500 in) ORF78c 60,500 (DNA ligase) (hypothetical pro- ORF90c (66,500 ORF90c (00 ORF90c (66,500 ORF90c (66,500 ORF90c (66,500 ORF90c (60,500 ORF90c (00) ORF90c	ORF57 (h ORF57 (h 00 43.50 in) tein) 9,000 45 73 (hypothetical 73 (hypothetical 73 (hypothetical 74 (hypothetical 75,000 00RF92c 14 (hypothetical 76,000 00RF92c	F59 (hypothetical pro- ypothetical pro- 00 44.9 ORF66 (9,500 protein) 55,500 sal protein) 55,500 61,500 9 (DNA polymer 67,500 (c) PA polymer 61,500 (c) PA polymer 61	ical protein) otein) ORF58 (000 44.5 000 44.5 000 55 000 74 1 000 55 000 54 000 54 0000 54 000 54 0000 54 00000000000000000	ORF60 hypothetical prot 00 45,00 tein) ORF67 (h 01 0,500 51 fiber protein) 56,500 e reductase) CC c (h ORF82. 62,500 c (endonuclease) c (endo	0 (hypothetical tein) 10 45.51 10 45.551 10 45.51 10 45.55 10 45.55	protein) 00 46.0 tein) tical protein) 0 1,500 5 75 (tail fiber protein) 57,500 riptional regulat 0 ORF85 63,500 INA polymerasis othetical protein 69,500 INA polymerasis othetical protein 69,500 0 Tothetical protein 0 Tothe	ORF62 (h) ORF70 (h) ORF70 (h) ORF70 (h) ORF69 (basepla 52,000 £ 58,000 ory protein) 0 64,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) n) 0 70,000 e) 0 70,000 e	ypothetical prote i00 47.0 ypothetical prote te assembly pro 58,500 58,500 64,500 64,500 0 70,500 (hypothetical pro 0 70,500 (hypothetical pro 0 70,500 10 70,500

Figure 4. Functional genome map of phage VpaE1. Functions are assigned according to the characterized ORFs in NCBI database and/or MS/MS analysis. The color code is as follows: yellow—DNA replication, recombination, repair and packaging; brown—transcription, translation, nucleotide metabolism; blue—structural proteins; green—lysis; purple—phage-host interaction; and grey—ORFs of unknown function black—tRNA.

3.4. Structural Proteins and Virion Proteome

In-gel and in-solution digestion are two common methods used for the analysis of phage structural proteins [57,58]. However, to identify the structural proteins of VpaE1, a recently improved filter-aided sample preparation (FASP) method [46,47] was used. Since following this procedure, most of the contaminants and detergents are being effectively removed from the biological samples. Prior to LC-MS/MS, a quality of phage suspension was verified by SDS PAGE (Figure 5A). Then, reversed-phase nano-liquid chromatography directly coupled with LC-MS/MS analysis was performed that allowed experimental identification of 22 structural proteins of VpaE1 (Figure 5B), including 11 that had no reliable homologues to annotated structural proteins in other organisms.

Α		B					
		No.	Anotation	mW (Da)	Putative function	Felix01 homolog	Identity aa %
LD.	M VpaE1	1.	VpaE1 gp56	41451	major capsid protein	gp58	99
KDa		2.	VpaE1_gp51	55741	structural protein	gp53	99
170		3.	VpaE1_gp55	13686	structural protein	gp57	92
130-		4.	VpaE1_gp65	80304	tail tape-measure protein	gp67	99
100-		5.	VpaE1_gp70	15590	hypothetical protein	gp72	100
70—		6.	VpaE1_gp52	18320	hypothetical protein	gp54	98
==		7.	VpaE1_gp54	48344	head maturation protease	gp56**	99
22—		8.	VpaE1_gp71	53095	baseplate component	gp73	97
40—		9.	VpaE1_gp53	11673	hypothetical protein	gp55	99
	A Company of the	10.	VpaE1_gp39	31595	putative tail protein	gp36	93
35-	-	11.	VpaE1_gp62	16184	hypothetical protein	gp64	99
	· · · · · · · ·	12.	VpaE1_gp75	89749	tail fiber protein	gp77	65
25 -		13.	VpaE1_gp72	31652	hypothetical protein	gp74	96
		14.	VpaE1_gp67	13189	hypothetical protein	gp69	98
		15.	VpaE1_gp69	23119	baseplate assembly protein	gp71*	99
		16.	VpaE1_gp47	28169	hypothetical protein	gp49	98
15-		17.	VpaE1_gp66	28716	hypothetical protein	gp68	98
		18.	VpaE1_gp57	16947	hypothetical protein	gp59	100
	-	19.	VpaE1_gp58	17661	hypothetical protein	gp60**	100
		20.	VpaE1_gp61	48863	structural protein	gp63	97
10 —	-	21.	VpaE1_gp74	41040	tail fiber protein	gp76*	76
		22.	VpaE1_gp68	37102	hypothetical protein	gp70**	99
		23.	VpaE1_gp25**		hypothetical protein	gp23	63

Notes: ** not found in the virion by MS; * not reliably found in the virion by MS.

Figure 5. Analysis of protein composition of VpaE1 virion: (**A**) SDS-PAGE analysis of phage VpaE1 particles; and (**B**) the list of VpaE1 structural proteins identified by MS in comparison with those in the virion of Felix01 identified previously [19].

All identified structural proteins of VpaE1, except the putative tail fiber protein VpaE1_gp75, shared high sequence similarities with their corresponding counterparts in Felix01. Unexpectedly, MS/MS failed to detect a homologue of Felix01_gp23 among the virion proteins of VpaE1. In contrast, three proteins that were detected in VpaE1 (namely, VpaE1_gp54, VpaE1_gp58 and VpaE1_gp68) had not been identified within the particle of Felix01 [19].

3.5. Phylogenetic Relatedness

To assess the evolutionary relationships between VpaE1 and other Felix01-related phages, a comparative whole genome sequence analysis was performed. For this, complete genome sequences of 14 Felix01-like viruses that infect *E. coli* or *Salmonella* strains as well as those of *Erwinia* phage phiEa21-4 and *Citrobacter* phage Michonne were aligned with the genome of VpaE1. As seen in

Figure 6, the phylogenetic analysis revealed that Felix01-like viruses form five separate but closely related groups.



Figure 6. Neighbor-joining tree based on the whole-genome sequence alignment. The numbers at the nodes indicate the bootstrap probabilities.

Overall, the branching order of groups is somewhat consistent with the defined host range phenotypes seen in *Felix01likevirus*. One such group comprises six *E. coli* O157:H7-infecting phages (namely, WV8, EC6, HY02, ECTP1, ECTP11 and ECTP 12), suggesting that these viruses are likely adapted to avoid specific restriction-modification systems that vary considerably between different lineages of *E. coli* [59–61]. Bacteriophage JH2, which is the only Felix01-like virus capable of infecting *E. coli* ETEC (enterotoxigenic) serotype O8:H9 [25], together with *Citrobacter* phage Michonne forms a separate divergent branch on the tree. Bacteriophage VpaE1, which is most closely related to *E. coli* O145-serotype-specific phage vB_EcoM_AYO145A isolated in Canada [27], occupies a separate branch together with 5 other Felix01-like viruses that infect *Salmonella* or *Escherichia*.

3.6. Identification of VpaE1 Receptor

The results presented in Table 1 suggest that the ability of VpaE1 to adsorb to the host cell is likely influenced by the LPS structure. To confirm this, bacteriophage VpaE1 was tested against a set of *E. coli* K-12 strains that are mutant in the biosynthesis pathway of LPS [36]. Initially, a spotting assay was employed (Table 2).

Then, when the clearly modified infections were observed, determination of the efficiency of plating (EOP) was performed using *E. coli* K-12 BW25113 and other VpaE1-resistant strains from the Keio collection as a negative control (Figure 7A). In the course of this experiment, the ability of VpaE1 to infect different *E. coli* B derivatives was also examined.

E. coli K-12 BW25113 Mutants	Spot Test	Source/Reference
ΔwaaC	-	Keio collection/[36]
ΔwaaF	_	Keio collection/[36]
ΔwaaG	+	Keio collection/[36]
ΔwaaP	+	Keio collection/[36]
ΔwaaY	_	Keio collection/[36]
ΔwaaQ	_	Keio collection/[36]
ΔwaaO	_	Keio collection/[36]
∆waaR	+	Keio collection/[36]
ΔwaaB	\pm	Keio collection/[36]
ΔwaaZ	±	Keio collection/[36]

Table 2. Susceptibility of E. coli BW25113 Mutant Strains to VpaE1.

+: clear plaques were observed after incubation at 37 $^{\circ}$ C; \pm : turbid plaques or no plaques at high dilutions.

Α		
Strain	Predicted LPS Composition	EOP
E. coli BL21	Inner core OS + GlcI + GlcII	1.00
BL21(DE3)	Inner core OS + GlcI + GlcII	1.19 ± 0.05
B40	Inner core OS + GlcI + GlcII	1.05 ± 0.02
BE-BS	Inner core OS + GlcI + GlcII	0.41 ± 0.05
BE	Inner core OS + GlcI + GlcII	0.98 ± 0.03
E. coli K-12 BW25113	Intact core OS	-
E. coli K-12 BW25113 mutants:		
$\Delta waaC$	Hep-deficient	-
$\Delta waaF$	Kdo + HepI	-
$\Delta waaG$	Inner core OS, HepI lacks P	1.64 ± 0.01
$\Delta waaP$	Intact core OS, HepIII-deficient, HepI and HepII lacks P	1.43 ± 0.04
$\Delta waaY$	Intact core OS, HepII lacks P	-
$\Delta waaQ$	Intact core OS, HepIII-deficient, HepII lacks P	-
ΔwaaO	Inner core OS + GlcI	-
$\Delta waaR$	Inner core OS + GlcI + GlcII	$1.30\pm0.02^*$
$\Delta waaB$	Intact core OS, Gal-deficient	$0.36 \pm 0.03^{**}$
$\Delta waaZ$	Intact core OS, KdoIII-deficient	$0.51 \pm 0.15^{**}$



Notes: * small plaques; ** very small plaques.



Figure 7. Plating efficiency of VpaE1 on different *E. coli* strains at 37 °C (**A**); and schematic representation of the LPS structure of *E. coli* K-12 (**B**), adapted based on the data from [4,12,13,62–66]. Residues shown in black are shared between *E. coli* K-12 and *E. coli* B LPS.

As seen in Figure 7, in the case of *E. coli* K-12 BW25113 $\Delta waaG$, a significantly higher EOP of VpaE1 was observed than on the lawns of the reference host BL21. The same effect was seen when *E. coli* K-12 BW25113 $\Delta waaR$ was used, although the plaques formed by VpaE1 on this particular strain were relatively smaller than those on BL21. Therefore, when LPS was truncated even down to the inner core OS, bacteriophage VpaE1 showed higher infectivity towards mutant *E. coli* K-12 strains rather than B derivatives. In addition, the increase in the EOP was also observed when VpaE1 was grown on *E. coli* K-12 $\Delta waaP$, a mutant strain with an impaired inner core region (unphosphorylated core, no HepIII added) [12]. Notably, VpaE1 could also infect *E. coli* K $\Delta waaB$ and $\Delta waaZ$ (lack Gal and KdoIII residues, respectively), while both $\Delta waaF$ and $\Delta waaC$ (the inner core OS truncated down to KdoI and KdoI-HepI, respectively) mutants were resistant to VpaE1 infection. These results suggest that the core residues of the inner core OS are important for VpaE1 adsorption. Unexpectedly, *E. coli* B strains differed in their susceptibility to VpaE1 as well. As seen in Figure 7, the EOP of VpaE1 on *E. coli* BE-BS was reduced more than two-fold, as compared to its EOP on other *E. coli* B-derivatives. However, this puzzling phenomenon needs to be investigated in more depth and is left for future research.

At first glance, it seems that VpaE1 adsorb better when HepI lacks a P moiety, however, our results show that *E. coli* K-12 BW25113 mutants $\Delta waaR$, $\Delta waaB$ and $\Delta waaZ$ as well as *E. coli* B strains—that all possess phosphorylated HepI—are all susceptible to VpaE1. Therefore, the additional studies are needed to elucidate the precise mechanism of VpaE1 host-cell recognition.

To confirm that VpaE1 failed to form plaques on *E. coli* BW25113 due to its inability to adsorb to K-12-type LPS, the adsorption assay was performed using several phage-resistant and phage-susceptible strains. The results presented in Figure 8A indicated that bacteriophage VpaE1 indeed failed to adsorb to *E. coli* K-12 BW25113, whereas in the case of *E. coli* BL21, BW25113 $\Delta waaG$ and $\Delta waaP$ the adsorption of the phage was unaffected.



Figure 8. VpaE1 adsorption assays: (**A**) VpaE1 adsorption kinetics; and (**B**) effect of proteinase K, sodium acetate and periodate treatment on the ability of VpaE1 to adsorb to *E. coli* BL21, BW25113 $\Delta waaG$ and $\Delta waaP$.

To test whether the receptor for VpaE1 is a carbohydrate but not a protein, bacteria were treated with proteinase K or periodate prior to the phage adsorption assay (Figure 8B). Sodium periodate degrades carbohydrates, whereas proteinase K that has a broad substrate specificity degrades cell surface proteins. As seen in Figure 8, the proteinase K treatment had no effect on VpaE1 adsorption, whereas the incubation of *E. coli* cells in the presence of 100 mM periodate reduced VpaE1 adsorption by ~80%. These results indicate that a receptor for bacteriophage VpaE1 is carbohydrate of incomplete core LPS.

4. Discussion

Here we present bacteriophage VpaE1 whose virion morphology as well as genome characteristics are reminiscent of those found in the members of the newly proposed genus Felix01likevirus, which comprises 14 fully sequenced Salmonella and E. coli phages. The genomes of Felix01-related phages exhibit a high degree of homology, and VpaE1 is no exception. Genome alignment comparison has shown remarkable similarity (in most cases, >84% nucleotide sequence identity) between the genome of VpaE1 and other sequenced Felix01-like viruses. Moreover, almost all functionally assigned VpaE1 genes, except for gene 75 coding putative tail fiber protein, share >90% amino acid sequence identity with those from the genomes of other Felix01-like viruses, suggesting that the core genome of *Felix01likevirus* is highly conserved. As is the case with Felix01 and its relatives, bacteriophage VpaE1 contains numerous genes involved in DNA replication and metabolism suggesting that Felix01-related viruses, like many other lytic phages, direct their own replication machinery. Also, the genome of VpaE1 codes for 22 apparently functional tRNAs (a feature that has been also documented in other Felix01-like viruses) that may have a positive impact on translation of phage-derived mRNA. Bacteriophage VpaE1 and its relatives do not code for their own RNA polymerase, nor do they code for proteins required for modification of the host RNAP. Hence, in the case of Felix01-like viruses, the question on how transcription is temporally regulated (especially at early stages of infection) remains unanswered.

Despite the aforementioned similarities, bacteriophage VpaE1 profoundly differs from the rest of Felix01-like viruses, since it is the only coliphage among *Felix01likevirus* that does not infect smooth LPS bacteria. All Felix 01-like *Salmonella* and *E. coli* phages characterized thus far have been isolated based on their ability to target a specific serotype of bacteria, and to date, no Felix01-related bacteriophages have been reported to infect O-Ag-deficient strains, much less O-Ag-deficient strains with a truncated core LPS. This could be at least partially due to the fact that such phages seem to be relatively rare in nature, since their rough or deep rough hosts, which usually emerge as a result of the selective pressure imposed by O-Ag/smooth LPS-specific phage infection [3], have been found to be not fit enough to survive the competition for natural resources with smooth strains, and are usually more susceptible to abiotic stress [11–13,62]. However short-lived, bacterial populations that produce mutant LPS occasionally do emerge in nature, and so do, apparently, mutant LPS-specific bacteriophages.

In the case of many tailed viruses, the distal part of the phage tail fibers is responsible for host cell recognition. As seen in Figure 5, bacteriophage VpaE1 harbors two putative tail fiber proteins, namely VpaE1_gp74 and VpaE1_gp75. Based on BLASTP analysis, both the C-terminus of VpaE1_gp75 (aa 584–780) and the C-terminal region of VpaE1_gp74 (aa 146–362) show homology to the long tail fiber adhesin of T4 (54% identity/E-value 2e-51 and 37% identity/E-value 9e-32, respectively). In T4, the long tail fiber adhesin is located in the C-terminal domain of gp37, and recognizes *E. coli* B LPS with at least one of the two terminal glucose (Glc) residues present [67] and OmpC found on *E. coli* K-12 but not B derivatives [68]. On the contrary, the host range of VpaE1 is limited to *E. coli* B strains only, since this phage does not use K-12-type LPS as a receptor. Although strikingly similar in genome size and gene content, *E. coli* B and K-12 derivatives mostly differ in the copy number and distribution of IS elements [65,69] that account for the majority of phenotypic differences between the strains [70,71]. According to the literature, *E. coli* B and K-12 derivatives are devoid of the O-antigen

since, in both groups, IS elements are inserted at the gene clusters for O-antigen biosynthesis (within *wbbL* in K-12 and *wbbD* in *E. coli* B). However, in the case of *E. coli* B derivatives, due to the IS1 insertion in *waaT*, the core part of LPS is further disrupted leading to the truncated core oligosaccharide [65].

Based on bioinformatics analysis, the tail fiber proteins of VpaE1 and their corresponding counterparts in other Felix01-related phages contain a conserved N-terminus (likely responsible for protein attachment to the phage particle) and a less-conserved C-terminal domain that may be responsible for the attachment of the virus to the host cell. Also, in the case of VpaE1_gp75 and its homologues, the region between the N- and C-terminal domains is highly variable in size and sequence.

Phylogenetic analysis based on the alignment of the amino acid sequences of putative tail fiber adhesins from various tailed viruses (Figure 9) revealed that the sequences of VpaE1_gp74 and vB_EcoM_AYO145A_gp077 are the most closely related, and both clustered on a separate divergent branch. The same was observed for the sequence of the C-terminal domain (238 aa) of gp75 of phage VpaE1 (VpaE1_gp75C).



Figure 9. Phylogenetic analysis. Neighbor-joining trees were constructed based on the amino acid sequence alignment of VpaE1_gp74 (**A**) and VpaE1_gp75C (**B**) with tail fiber proteins from various myoviruses. The numbers at the nodes indicate the bootstrap probabilities.

As was mentioned above, gp74 and gp75 of VpaE1 both share a region of sequence similarity with gp37 of T4 yet, based on the results of phylogenetic analysis, these three proteins are only distantly related. Nevertheless, considering that both phages recognize *E. coli* B LPS but differ in their ability to use proteins as receptors, the adhesins of T4 and VpaE1 offer a promising model for studying protein structure-function relationship.

The majority of bacteriophages that belong to the newly proposed genus *Felix01likevirus* are highly virulent viruses, and have long been used in diverse applications, such as phage-based identification of bacteria, biocontrol of pathogens, phage-therapy, *etc.* [20,21,28,30,31,72]. Nevertheless, as is the case with antibiotics, the effectiveness of phage applications is greatly limited by the emergence of phage-resistant bacteria that often express rough or deep rough LPS. To surmount this obstacle, the use of phage cocktails that target multiple bacterial receptors has been proposed. In this regard, the isolation of truncated LPS-specific bacteriophages is of great importance.

5. Conclusions

Bacteriophage VpaE1 belongs to the newly proposed genus, *Felix011ikevirus*, in the family *Myoviridae*. All Felix01-like viruses described to date are highly similar in genome structure and organization, and are adapted to infect O-Ag-expressing *Salmonella* or *Escherichia coli* strains. In contrast, phage VpaE1 presented here is the first Felix01-like virus that infects O-Ag-deficient *E. coli* strains with a truncated core LPS. Moreover, VpaE1 can replicate over a wide temperature range from 9 to 49 °C, suggesting that this phage is well adapted to harsh environmental conditions. Hence, bacteriophage VpaE1 shows potential for the development of phage-based applications.

Supplementary Materials: The following are available online at www.mdpi.com/link, Table S1: List of annotated VpaE1 ORFs, Table S2: Data on proteomic analysis of VpaE1 virion proteins.

Acknowledgments: We would like to thank Audrius Driukas for assistance in VpaE1 isolation. We thank Lindsay W. Black, Kenneth N. Kreuzer, Edita Sužiedėlienė and Jaunius Urbonavičius for bacterial strains. This research was funded by a grant (No. MIP-002/2014) from the Research Council of Lithuania.

Author Contributions: Eugenijus Šimoliūnas purified phage for TEM, and analyzed and annotated genome sequence; Monika Vilkaitytė performed adsorption tests and DNA restriction analysis; Laura Kalinienė analyzed phylogenetic relationships and wrote the paper; Aurelija Zajančkauskaitė performed host-range analysis; Algirdas Kaupinis performed LC MS/MS analysis; Juozas Staniulis executed TEM analysis; Mindaugas Valius contributed reagents and equipment and analyzed MS data; Rolandas Meškys contributed reagents and equipment, analyzed data, and advised on paper writing; and Lidija Truncaitė conceived and designed the experiments, analyzed data and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Rakhuba, D.V.; Kolomiets, E.I.; Dey, E.S.; Novik, G.I. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. *Pol. J. Microbiol.* **2010**, *59*, 145–155. [PubMed]
- Shin, H.; Lee, J.H.; Kim, H.; Choi, Y.; Heu, S.; Ryu, S. Receptor diversity and host interaction of bacteriophages infecting *Salmonella* enterica serovar Typhimurium. *PLoS ONE* 2012, 7, e43392. [CrossRef] [PubMed]
- Michel, A.; Clermont, O.; Denamur, E.; Tenaillon, O. Bacteriophage PhiX174's ecological niche and the flexibility of its *Escherichia coli* lipopolysaccharide receptor. *Appl. Environ. Microbiol.* 2010, 76, 7310–7313. [CrossRef] [PubMed]
- 4. Raetz, C.R.; Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **2002**, *71*, 635–700. [CrossRef] [PubMed]
- 5. Hull, S. *Escherichia coli* lipopolysaccharide in pathogenesis and virulence. In *Escherichia coli: Mechanisms of Virulence;* Sussman, M., Ed.; Cambridge University Press: Cambridge, UK, 1997; pp. 145–167.
- Amor, K.; Heinrichs, D.E.; Frirdich, E.; Ziebell, K.; Johnson, R.P.; Whitfield, C. Distribution of core oligosaccharide types in lipopolysaccharides from *Escherichia coli*. *Infect. Immun.* 2000, *68*, 1116–1124. [CrossRef] [PubMed]
- Kim, M.; Ryu, S. Spontaneous and transient defence against bacteriophage by phase-variable glucosylation of O-antigen in *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.* 2012, *86*, 411–425. [CrossRef] [PubMed]
- Labrie, S.J.; Samson, J.E.; Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 2010, 8, 317–327. [CrossRef] [PubMed]
- Perry, E.B.; Barrick, J.E.; Bohannan, B.J. The Molecular and genetic basis of repeatable coevolution between *Escherichia coli* and bacteriophage T3 in a laboratory microcosm. *PLoS ONE* 2015, *10*, e0130639. [CrossRef] [PubMed]
- 10. Brockhurst, M.A.; Rainey, P.B.; Buckling, A. The effect of spatial heterogeneity and parasites on the evolution of host diversity. *Proc. Biol. Sci.* 2004, 271, 107–111. [CrossRef] [PubMed]
- 11. Buckling, A.; Wei, Y.; Massey, R.C.; Brockhurst, M.A.; Hochberg, M.E. Antagonistic coevolution with parasites increases the cost of host deleterious mutations. *Proc. Biol. Sci.* **2006**, 273, 45–49. [CrossRef] [PubMed]
- 12. Yethon, J.A.; Heinrichs, D.E.; Monteiro, M.A.; Perry, M.B.; Whitfield, C. Involvement of *waaY*, *waaQ*, and *waaP* in the modification of *Escherichia coli* lipopolysaccharide and their role in the formation of a stable outer membrane. *J. Biol. Chem.* **1998**, 273, 26310–26316. [CrossRef] [PubMed]
- 13. Yethon, J.A.; Vinogradov, E.; Perry, M.B.; Whitfield, C. Mutation of the lipopolysaccharide core glycosyltransferase encoded by *waaG* destabilizes the outer membrane of *Escherichia coli* by interfering with core phosphorylation. *J. Bacteriol.* **2000**, *182*, 5620–5623. [CrossRef] [PubMed]
- Heinrichs, D.E.; Yethon, J.A.; Whitfield, C. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. *Mol. Microbiol.* **1998**, *30*, 221–232. [CrossRef] [PubMed]

- Marti, R.; Zurfluh, K.; Hagens, S.; Pianezzi, J.; Klumpp, J.; Loessner, M.J. Long tail fibres of the novel broad-host-range T-even bacteriophage S16 specifically recognize *Salmonella* OmpC. *Mol. Microbiol.* 2013, *87*, 818–834. [CrossRef] [PubMed]
- Kim, M.; Kim, S.; Park, B.; Ryu, S. Core lipopolysaccharide-specific phage SSU5 as an auxiliary component of a phage cocktail for *Salmonella* biocontrol. *Appl. Environ. Microbiol.* 2014, *80*, 1026–1034. [CrossRef] [PubMed]
- 17. Bao, H.; Zhang, P.; Zhang, H.; Zhou, Y.; Zhang, L.; Wang, R. Bio-control of *Salmonella enteritidis* in foods using bacteriophages. *Viruses* **2015**, *7*, 4836–4853. [CrossRef] [PubMed]
- Felix, A.; Callow, B.R. Typing of paratyphoid B bacilli by Vi bacteriophage. *Brit. Med. J.* 1943, 2, 127–130. [CrossRef] [PubMed]
- Whichard, J.M.; Weigt, L.A.; Borris, D.J.; Li, L.L.; Zhang, Q.; Kapur, V.; Pierson, F.W.; Lingohr, E.J.; She, Y.M.; Kropinski, A.M.; *et al.* Complete genomic sequence of bacteriophage Felix O1. *Viruses* 2010, *2*, 710–730. [CrossRef] [PubMed]
- 20. Hirsh, D.C.; Martin, L.D. Rapid detection of *Salmonella* spp. by using Felix-O1 bacteriophage and high-performance liquid chromatography. *Appl. Environ. Microbiol* **1983**, *45*, 260–264. [PubMed]
- Whichard, J.M.; Sriranganathan, N.; Pierson, F.W. Suppression of *Salmonella* growth by wild-type and large-plaque variants of bacteriophage Felix O1 in liquid culture and on chicken frankfurters. *J. Food Protect.* 2003, 66, 220–225.
- 22. Lavigne, R.; Darius, P.; Summer, E.J.; Seto, D.; Mahadevan, P.; Nilsson, A.S.; Ackermann, H.W.; Kropinski, A.M. Classification of *Myoviridae* bacteriophages using protein sequence similarity. *BMC Microbiol.* **2009**, *9*. [CrossRef] [PubMed]
- 23. Villegas, A.; She, Y.M.; Kropinski, A.M.; Lingohr, E.J.; Mazzocco, A.; Ojha, S.; Waddell, T.E.; Ackermann, H.W.; Moyles, D.M.; Ahmed, R.; *et al.* The genome and proteome of a virulent *Escherichia coli* O157:H7 bacteriophage closely resembling *Salmonella* phage Felix O1. *Virol. J.* **2009**, *6.* [CrossRef] [PubMed]
- 24. Tiwari, B.R.; Kim, J. Complete genome sequence of bacteriophage EC6, capable of lysing *Escherichia coli* O157:H7. *Genome Announc.* **2013**, *1*, e00085-12. [CrossRef] [PubMed]
- 25. Hoa, N.X.; Tang, F.; Bai, Q.; Shen, C.; Vo, T.T.; Lu, C. Isolation and characterization of bacteriophages against enterotoxigenic *Escherichia coli* strains causing swine diseases. *Afr. J. Microbiol. Res.* **2013**, *7*, 2787–2793.
- Cowley, L.A.; Beckett, S.J.; Chase-Topping, M.; Perry, N.; Dallman, T.J.; Gally, D.L.; Jenkins, C. Analysis of whole genome sequencing for the *Escherichia coli* O157:H7 typing phages. *BMC Genom.* 2015, *16*. [CrossRef] [PubMed]
- Wang, J.; Niu, Y.D.; Chen, J.; Anany, H.; Ackermann, H.W.; Johnson, R.P.; Ateba, C.N.; Stanford, K.; McAllister, T.A. Feces of feedlot cattle contain a diversity of bacteriophages that lyse non-O157 Shiga toxin-producing *Escherichia coli. Can. J. Microbiol.* 2015, *61*, 467–475. [CrossRef] [PubMed]
- Bardina, C.; Spricigo, D.A.; Cortes, P.; Llagostera, M. Significance of the bacteriophage treatment schedule in reducing *Salmonella* colonization of poultry. *Appl. Environ. Microbiol.* 2012, 78, 6600–6607. [CrossRef] [PubMed]
- Moreno Switt, A.I.; Orsi, R.H.; den Bakker, H.C.; Vongkamjan, K.; Altier, C.; Wiedmann, M. Genomic characterization provides new insight into *Salmonella* phage diversity. *BMC Genom.* 2013, 14. [CrossRef] [PubMed]
- 30. Spricigo, D.A.; Bardina, C.; Cortes, P.; Llagostera, M. Use of a bacteriophage cocktail to control *Salmonella* in food and the food industry. *Int. J. Food Microbiol.* **2013**, *165*, 169–174. [CrossRef] [PubMed]
- Colom, J.; Cano-Sarabia, M.; Otero, J.; Cortes, P.; Maspoch, D.; Llagostera, M. Liposome-encapsulated bacteriophages for enhanced oral phage therapy against *Salmonella* spp. *Appl. Environ. Microbiol.* 2015, *81*, 4841–4849. [CrossRef] [PubMed]
- 32. Tolen, T.N.; Xie, Y.; Hernandez, A.C.; Kuty Everett, G.F. Complete genome sequence of *Salmonella enterica* serovar Typhimurium myophage mushroom. *Genome Announc.* **2015**, *3*, e00154-15. [CrossRef] [PubMed]
- Lehman, S.M.; Kropinski, A.M.; Castle, A.J.; Svircev, A.M. Complete genome of the broad-host-range Erwinia amylovora phage phiEa21-4 and its relationship to Salmonella phage felix O1. Appl. Environ. Microbiol. 2009, 75, 2139–2147. [CrossRef] [PubMed]
- 34. Nguyen, Q.T.; Luna, A.J.; Hernandez, A.C.; Kuty Everett, G.F. Complete genome sequence of *Citrobacter freundii* myophage Moogle. *Genome Announc.* **2015**, *3*, e01426-14. [CrossRef] [PubMed]

- 35. Henry, M.; Bobay, L.M.; Chevallereau, A.; Saussereau, E.; Ceyssens, P.J.; Debarbieux, L. The search for therapeutic bacteriophages uncovers one new subfamily and two new genera of *Pseudomonas*-infecting *Myoviridae*. *PLoS ONE* **2015**, *10*, e0117163. [CrossRef] [PubMed]
- Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K.A.; Tomita, M.; Wanner, B.L.; Mori, H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2006, 2. [CrossRef] [PubMed]
- 37. Simpson, D.J.; Sacher, J.C.; Szymanski, C.M. Exploring the interactions between bacteriophage-encoded glycan binding proteins and carbohydrates. *Curr. Opin. Struct. Biol.* **2015**, *34*, 69–77. [CrossRef] [PubMed]
- 38. Kutter, E. Phage host range and efficiency of plating. In *Bacteriophages: Methods in Molecular Biology;* Clokie, M.R.J., Kropinski, A.M., Eds.; Humana Press: New York, NY, USA, 2009; Volume 1, pp. 141–149.
- Kropinski, A.M. Measurement of the bacteriophage inactivation kinetics with purified receptors. In Bacteriophages: Methods in Molecular Biology; Clokie, M.R.J., Kropinski, A.M., Eds.; Humana Press: New York, NY, USA, 2009; Volume 1, pp. 157–160.
- 40. Kiljunen, S.; Datta, N.; Dentovskaya, S.V.; Anisimov, A.P.; Knirel, Y.A.; Bengoechea, J.A.; Holst, O.; Skurnik, M. Identification of the lipopolysaccharide core of *Yersinia pestis* and *Yersinia pseudotuberculosis* as the receptor for bacteriophage phiA1122. *J. Bacteriol.* **2011**, *193*, 4963–4972. [CrossRef] [PubMed]
- 41. Delcher, A.L.; Harmon, D.; Kasif, S.; White, O.; Salzberg, S.L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **1999**, *27*, 4636–4641. [CrossRef] [PubMed]
- 42. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; *et al.* Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]
- 43. Schattner, P.; Brooks, A.N.; Lowe, T.M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* **2005**, *33*, W686–W689. [CrossRef] [PubMed]
- 44. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [CrossRef] [PubMed]
- 45. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, 227, 680–685. [CrossRef] [PubMed]
- 46. Feist, P.; Hummon, A.B. Proteomic challenges: Sample preparation techniques for microgram-quantity protein analysis from biological samples. *Int. J. Mol. Sci.* **2015**, *16*, 3537–3563. [CrossRef] [PubMed]
- 47. Wisniewski, J.R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Ant. Methods* **2009**, *6*, 359–362. [CrossRef] [PubMed]
- 48. Kaliniene, L.; Klausa, V.; Truncaite, L. Low-temperature T4-like coliphages vB_EcoM-VR5, vB_EcoM-VR7 and vB_EcoM-VR20. *Arch. Virol.* **2010**, *155*, 871–880. [CrossRef] [PubMed]
- 49. Kaliniene, L.; Klausa, V.; Zajanckauskaite, A.; Nivinskas, R.; Truncaite, L. Genome of low-temperature T4-related bacteriophage vB_EcoM-VR7. *Arch. Virol.* **2011**, *156*, 1913–1916. [CrossRef] [PubMed]
- 50. Kaliniene, L.; Zajanckauskaite, A.; Simoliunas, E.; Truncaite, L.; Meskys, R. Low-temperature bacterial viruses VR—A small but diverse group of *E. coli* phages. *Arch. Virol.* **2015**, *160*, 1367–1370. [CrossRef] [PubMed]
- Truncaite, L.; Simoliunas, E.; Zajanckauskaite, A.; Kaliniene, L.; Mankeviciute, R.; Staniulis, J.; Klausa, V.; Meskys, R. Bacteriophage vB_EcoM_FV3: A new member of "rV5-like viruses". *Arch. Virol.* 2012, 157, 2431–2435. [CrossRef] [PubMed]
- Selick, H.E.; Kreuzer, K.N.; Alberts, B.M. The bacteriophage T4 insertion/substitution vector system. A method for introducing site-specific mutations into the virus chromosome. *J. Biol. Chem.* 1988, 263, 11336–11347. [PubMed]
- 53. Grenier, F.; Matteau, D.; Baby, V.; Rodrigue, S. Complete genome sequence of *Escherichia coli* BW25113. *Genome Announc.* **2014**, *2*, e01038-14. [CrossRef] [PubMed]
- 54. Simoliunas, E.; Kaliniene, L.; Truncaite, L.; Klausa, V.; Zajanckauskaite, A.; Meskys, R. Genome of *Klebsiella* sp.-infecting bacteriophage vB_KleM_RaK2. *J. Virol.* **2012**, *86*. [CrossRef] [PubMed]
- 55. Hyman, P.; Abedon, S.T. Bacteriophage host range and bacterial resistance. In *Advances in Applied Microbiology*; Laskin, A.I., Sariaslani, S., Gadd, G.M., Eds.; Elsevier Inc. Academic Press: Burlington, MA, USA, 2010; Volume 70, pp. 217–248.

- 56. Dy, R.L.; Richter, C.; Salmond, G.P.; Fineran, P.C. Remarkable mechanisms in microbes to resist phage infections. *Annu. Rev. Virol.* 2014, *1*, 307–331. [CrossRef]
- 57. Lavigne, R.; Noben, J.P.; Hertveldt, K.; Ceyssens, P.J.; Briers, Y.; Dumont, D.; Roucourt, B.; Krylov, V.N.; Mesyanzhinov, V.V.; Robben, J.; *et al.* The structural proteome of *Pseudomonas aeruginosa* bacteriophage phiKMV. *Microbiology* **2006**, *152*, 529–534. [CrossRef] [PubMed]
- Lavigne, R.; Ceyssens, P.J.; Robben, J. Phage proteomics: Applications of mass spectrometry. In Bacteriophages: Methods in Molecular Biology; Clockie, M.R.J., Kropinski, A.M., Eds.; Humana Press: Totowa, NJ, USA, 2009; Volume 2, pp. 239–251.
- 59. Pupo, G.M.; Karaolis, D.K.; Lan, R.; Reeves, P.R. Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. *Infect. Immun.* **1997**, *65*, 2685–2692. [PubMed]
- 60. Reid, S.D.; Herbelin, C.J.; Bumbaugh, A.C.; Selander, R.K.; Whittam, T.S. Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* **2000**, *406*, 64–67. [PubMed]
- 61. Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; *et al.* Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* **2001**, *8*, 11–22. [CrossRef] [PubMed]
- 62. Wang, Z.; Wang, J.; Ren, G.; Li, Y.; Wang, X. Influence of core oligosaccharide of lipopolysaccharide to outer membrane behavior of *Escherichia coli*. *Mar. Drugs* **2015**, *13*, 3325–3339. [CrossRef] [PubMed]
- 63. Frirdich, E.; Lindner, B.; Holst, O.; Whitfield, C. Overexpression of the *waaZ* gene leads to modification of the structure of the inner core region of *Escherichia coli* lipopolysaccharide, truncation of the outer core, and reduction of the amount of O polysaccharide on the cell surface. *J. Bacteriol.* **2003**, *185*, 1659–1671. [CrossRef] [PubMed]
- 64. Taylor, C.M.; Goldrick, M.; Lord, L.; Roberts, I.S. Mutations in the *waaR* gene of *Escherichia coli* which disrupt lipopolysaccharide outer core biosynthesis affect cell surface retention of group 2 capsular polysaccharides. *J. Bacteriol.* **2006**, *188*, 1165–1168. [CrossRef] [PubMed]
- Jeong, H.; Barbe, V.; Lee, C.H.; Vallenet, D.; Yu, D.S.; Choi, S.H.; Couloux, A.; Lee, S.W.; Yoon, S.H.; Cattolico, L.; *et al.* Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J. Mol. Biol.* 2009, 394, 644–652. [CrossRef] [PubMed]
- 66. Qian, J.; Garrett, T.A.; Raetz, C.R. *In vitro* assembly of the outer core of the lipopolysaccharide from *Escherichia coli* K-12 and *Salmonella typhimurium*. *Biochemistry* **2014**, *53*, 1250–1262. [CrossRef] [PubMed]
- Kaszowska, M.; Niedziela, T.; Maciejewska, A.; Lukasiewicz, J.; Jachymek, W.; Lugowski, C. Core oligosaccharide of *Escherichia coli* B-the structure required for bacteriophage T4 recognition. *Carbohydr. Res.* 2015, 413, 51–54. [CrossRef] [PubMed]
- 68. Yu, F.; Mizushima, S. Roles of lipopolysaccharide and outer membrane protein OmpC of *Escherichia coli* K-12 in the receptor function for bacteriophage T4. *J. Bacteriol.* **1982**, *151*, 718–722. [PubMed]
- 69. Studier, F.W.; Daegelen, P.; Lenski, R.E.; Maslov, S.; Kim, J.F. Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the *E. coli* B and K-12 genomes. *J. Mol. Biol.* **2009**, *394*, 653–680. [CrossRef] [PubMed]
- Schneider, D.; Duperchy, E.; Depeyrot, J.; Coursange, E.; Lenski, R.E.; Blot, M. Genomic comparisons among *Escherichia coli* strains B, K-12, and O157: H7 using IS elements as molecular markers. *BMC Microbiol.* 2002, 2. [CrossRef]
- Yoon, S.H.; Han, M.J.; Jeong, H.; Lee, C.H.; Xia, X.X.; Lee, D.H.; Shim, J.H.; Lee, S.Y.; Oh, T.K.; Kim, J.F. Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. *Genome Biol.* 2012, 13. [CrossRef]
- 72. Niu, Y.D.; Johnson, R.P.; Xu, Y.; McAllister, T.A.; Sharma, R.; Louie, M.; Stanford, K. Host range and lytic capability of four bacteriophages against bovine and clinical human isolates of Shiga toxin-producing *Escherichia coli* O157:H7. *J. Appl. Microbiol.* **2009**, *107*, 646–656. [CrossRef] [PubMed]



© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).