

RESEARCH ARTICLE

Bacillus-infecting bacteriophage Izhevsk harbors thermostable endolysin with broad range specificity

Anna V. Skorynina¹ , Emma G. Pilgrimova¹ , Olesya A. Kazantseva¹, Vladislav A. Kulyabin , Svetlana D. Baicher¹, Natalya A. Ryabova², Andrey M. Shadrin ^{1*}

1 G.K. Skryabin Institute of Biochemistry and Physiology of Microorganisms, Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences, Federal Research Center, Pushchino, Russia, **2** Institute of Protein Research RAS, Pushchino, Russia

 These authors contributed equally to this work.

* andrey2010s@gmail.com, andrey.shadrin@ibpm.ru



OPEN ACCESS

Citation: Skorynina AV, Pilgrimova EG, Kazantseva OA, Kulyabin VA, Baicher SD, Ryabova NA, et al. (2020) Bacillus-infecting bacteriophage Izhevsk harbors thermostable endolysin with broad range specificity. PLoS ONE 15(11): e0242657. <https://doi.org/10.1371/journal.pone.0242657>

Editor: Kornelius Zeth, Roskilde Universitet, DENMARK

Received: June 23, 2020

Accepted: November 9, 2020

Published: November 24, 2020

Copyright: © 2020 Skorynina et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The genome sequence of phage Izhevsk is available from the NCBI GenBank database (accession number MT254578).

Funding: The reported study was funded by Russian Foundation for Basic Research (RFBR) according to the research project № 19-04-00300 (A.M.S., E.G.P., O.A.K., V.A.K.) and by RFBR and Moscow Oblast according to the research project 17-44-500067 (A.M.S., E.G.P., O.A.K., V.A.K.). The funders had no role in study design, data collection

Abstract

Several bacterial species belonging to the *Bacillus cereus* group are known to be causative agents of food poisoning and severe human diseases. Bacteriophages and their lytic enzymes called endolysins have been widely shown to provide for a supplemental or primary means of treating bacterial infections. In this work we present a new broad-host-range phage Izhevsk, which infects the members of the *Bacillus cereus* group. Transmission electron microscopy, genome sequencing and comparative analyses revealed that Izhevsk is a temperate phage with *Siphoviridae* morphology and belongs to the same genus as the previously described but taxonomically unclassified bacteriophages Tsamsa and Diildio. The Ply57 endolysin of Izhevsk phage has broad-spectrum activity against *B. cereus sensu lato*. The thermolability of Ply57 is higher than that of the PlyG of Wβ phage. This work contributes to our current understanding of phage biodiversity and may be useful for further development of efficient antimicrobials aimed at diagnosing and treating infectious diseases and food contaminations caused by the *Bacillus cereus* group of bacteria.

Introduction

In the early 2000s, the growing availability of sequencing technologies and the increasing emergence rate of multi-resistant pathogens served as the main factors which triggered a revived interest in phage research worldwide. Bacteriophages and their constituents are now widely used in medical diagnostics and food microbiology, and their applicability continues to expand in the treatment of human infections. Despite being undeniably limited compared to conventional antibacterial therapies, they are gaining more and more recognition. Many reports have been published describing phages and their bacteriolytic enzymes—natural endolysins or engineered derivatives—successfully used to prevent contamination of food products [1,2], as well as investigated in rodent-based models of human infections [3,4]. Endolysins are phage-encoded peptidoglycan-degrading proteins required for host cell wall lysis and

and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

possessing different enzymatic activities such as glycosidase, amidase, endopeptidase and lytic transglycosylase [5]. In the case of phages infecting Gram-positive bacteria, endolysins consist of one or more catalytic domains and a cell wall binding domain [5,6], offering wide opportunities for combining different molecules and generating proteins targeted to specific substrates.

Bacillus cereus sensu stricto is a Gram-positive spore-forming bacterium widespread in various environmental niches. It is also a member of the *Bacillus cereus* group of bacteria encompassing 18 recognized species [7] including *Bacillus anthracis*, a severe human pathogen. Many reports have shown that the members of the *Bacillus cereus* group commonly contaminate milk and dairy products [8–10], presenting potential threats because of their ability to grow at standard refrigerator temperatures and produce toxins and spores. Bacteriophages infecting bacteria of the *Bacillus cereus* group were among the first to be detected. Long before the term “bacteriophage” was proposed by D’Herelle, Nikolai Fedorovich Gamaleya observed antibacterial activity while working with *Bacillus anthracis* [11].

As one may notice from the information given in the “Taxonomy” section on the International Committee on the Taxonomy of Viruses (ICTV) website, *Bacillus*-infecting phages are now among the most numerous taxonomically classified bacteriophages of the order *Caudovirales* together with the viruses infecting *Enterobacteriaceae*, *Mycobacterium* and *Staphylococcus*. They constitute many genera of the *Myoviridae*, *Siphoviridae*, *Podoviridae* and *Herelleviridae* families and are also represented by a considerable number of taxonomically unclassified phages deposited in public databases.

The impressive variety of morphological features, genome characteristics, lifestyles, and lysogenic strategies have been reported for *Bacillus*-infecting bacteriophages [12], indicating that they are widespread, diverse and therefore of significant interest.

In this study, we describe the newly isolated temperate phage Izhevsk and characterize its thermostable broad-spectrum endolysin Ply57. The analysis of the Izhevsk genome has shown that the phage can be assigned to the same genus with the previously described and closely related *Bacillus*-infecting phages Tsamsa and Diildio, but the lytic range of Izhevsk phage is much wider than that of the previously described phages.

Materials and methods

Bacterial strains and growth conditions

Bacterial strains were obtained from the All-Russian Collection of Microorganisms (VKM) and cultivated in Lysogeny broth (LB) and on LB agar (1.5% and 0.5% for top layer) with 10 mM of CaCl₂ and 10 mM of MgCl₂ at 30°C.

Phage isolation and propagation

Phage Izhevsk was isolated from a soil sample collected in Izhevsk, Udmurtia, Russian Federation, and propagated on the sensitive strain *B. cereus* VKM B-370. For phage isolation, one gram of the soil sample was added to the cell culture at the optical density (OD₅₉₀) of 0.2 and the mixture was incubated until the optical density decreased. The lysate was filtered through 0.22-μm filters.

Serial dilutions of the phage suspension were prepared and added to 3 ml of 0.5% LB agar with 50 μL of the *B. cereus* VKM B-370 culture (OD₅₉₀ of 0.35); then, after 15 sec vortexing, the mixture was poured into previously prepared LB agar plates (1.5% agar) and incubated overnight at 37°C.

Phage purification, propagation and PEG 8000 (polyethylene glycol) precipitation were performed as described previously [13]. Then 3 ml of the high-titer preparation were used for

CsCl centrifugation in a preformed CsCl density gradient (1.3 g/ml, 1.4 g/ml, 1.5 g/ml, 1.6 g/ml, 1.7 g/ml) to prepare the final purified phage suspension, which was used later on for transmission electron microscopy.

Phage DNA sequencing

Phage DNA was extracted using the standard phenol-chloroform extraction protocol described by Sambrook et al. [14] and then sequenced using Illumina with the TruSeq library preparation technology. The genomic sequence was assembled *de novo* using SPAdes v.3.11.1 software [15]. Open reading frames (ORFs) were identified with RASTtk [16] and manually assigned the putative functions using BLAST (NCBI) [17,18] and HHpred [19]. tRNAscan-SE [20] were used for tRNA gene searching. The circular genome map was visualized with CGView [21].

Host range determination

Host range test was performed as described previously [13] using 34 strains of the *B. cereus* group as well as five other *Bacillus* strains and one *Enterococcus* strain. 2 μ L of phage suspension ($\geq 10^6$ p.f.u./mL) was dripped onto agar plate with each tested strain. Plates were incubated at 37°C for 48h. Lysis was monitored after 16, 30 and 48 h of incubation.

Transmission electron microscopy

Phage suspension applied onto 400 mesh carbon-formvar coated copper grids was negatively stained with 1% uranyl acetate and subsequently analyzed using a JEM-100C (JEOL, Japan) transmission electron microscope at 80 kV accelerating voltage. Images were taken on Kodak film SO-163 (Kodak, Cat. #74144, Hatfield, PA, USA) with 45000x magnification.

Endolysin gene cloning and enzyme purification

The genes of the Ply57 and PlyG endolysins were PCR-amplified using Herculase II Fusion DNA Polymerase (Agilent Technologies, Cat. # 600677–51) and the pairs of oligonucleotides 5′-tatatccatggCTATTTTCAGTAAGACAAAAA-3′ plus 5′-tatagcggccgcGTCTTG TACGAAACGCAC-3′ and 5′-tatatccatggAAATTCAGAAAAAATTAGTTGACCCA-3′ plus 5′-tatagcggccgcTTTAACTTCATACCACCA-3′, respectively. Total DNA of *B. cereus* W-strain, harboring W β prophage, was used as a template for PlyG cloning. The PCR products were cloned into the pET33 expression vector between the NcoI and EagI sites resulting in the AAALHHHHHH amino acid extension at the C-terminus of the recombinant proteins. For Ply57 and PlyG production, *E. coli* BL21(DE3) transformed with the respective pET33-based plasmid was grown on an orbital shaking platform at 37°C up to OD₅₉₀ = 0.5 and then 50 μ M IPTG was added to induce recombinant protein synthesis. The temperature was changed to 30°C and the incubation continued for the next 12 hours. Bacterial cells were precipitated by centrifugation and resuspended in 40 mL of ice-cold Buffer A [40 mM Tris-HCl pH 8.0; 0.5 M NaCl; 5% glycerol]. The cell suspension was sonicated thrice for 30 seconds using a Sonyprep disintegrator (Amplitude 4 and medium power). The sonicated cells were centrifuged for 90 min at 9000 g, 4°C, filtered through 0.45 μ m Millipore syringe filter, and loaded on a 5 mL Ni-chelating column (GE Healthcare). Purification was performed using imidazole elution as recommended by the manufacturer. The purification quality was confirmed by SDS polyacrylamide gel electrophoresis and the selected fractions were dialyzed against the dialysis buffer [40 mM Tris-HCl (pH 8.0); 0.5 M NaCl; 5% glycerol], which was refreshed three times. After the dialysis, protein concentrations were determined by measuring

adsorption at $\lambda = 280$ nm (extinction coefficient = 1) using NanoPhotometer Pearl P-360, and then glycerol and DTT were added to the final concentrations of 50% and 10 mM, respectively. The enzymes were stored at -20°C . The final concentration of the purified enzymes was 100–150 μM .

Endolysin specificity spot test

5 μl of Ply57 (100 μM) was dripped onto Brain Heart Infusion (BHI) agar plates with 4-hour lawns of each host strain, the Ply57 storage buffer was used as a control. The plates were incubated at 37°C for 16–48 hours. When a lysis zone was detected after incubation, host strain was considered sensitive.

Endolysin activity turbidimetry assays

The bacteriolytic activity of the purified Ply57 endolysin was determined via turbidity reduction assay on fresh exponentially growing bacterial cells as described previously [22], as well as on stationary phase cells, with minor modifications. Bacterial cells were precipitated, resuspended in 20 mM Tris-HCl (pH 8.0) buffer to the OD590 of approximately 1.0, and aliquoted into a 96-well plate (190 μl per each well). The purified endolysin was diluted in 20 mM Tris-HCl buffer to obtain 20 μM working solution. Ten microliters of the working solution were added to the bacterial cells resulting in the final volume of 200 μl with 1 μM of endolysin. The same amount of storage buffer was added to the control microplate well without endolysin. The plate was incubated at 37°C in the microplate spectrophotometer FilterMax F5 (Molecular Devices) with OD590 being measured every minute. The tested strains were considered sensitive during the exponential growth when OD590 decreased to the values lower than 80% of the control optical density (in the cultures without endolysin) after 30-min incubation. The strains were considered sensitive during the stationary phase when OD590 decreased to the values lower than 80% of the control optical density (in the cultures without endolysin) after 1 h incubation. Each experiment was performed in at least three replicates.

Thermostability assay was performed as described earlier by Heselpoth et al. [23] with several modifications. Ply57 and PlyG were incubated at 55°C in PBS (40 mM sodium phosphate pH 7.4; 150 mM NaCl) without reducing agents. Aliquots were taken after 10, 30, 60, 90 and 120 min and placed on ice bath until 120 min incubation was complete. The residual bacteriolytic activity was determined on exponentially growing *B. cereus* ATCC 4342 (also known as RSVF1) cells in PBS using the microplate spectrophotometer. The final concentrations of Ply57 and PlyG in the reactions were 1.5 μM and the reaction mixture volume was 200 μl . The endolysin half-inactivation time was estimated using the exponential trendline equation for the dependence of the residual bacteriolytic activity from incubation time (R^2 value >0.98). The average values and standard deviations were calculated based on at least three replicates.

Comparative genomics

To find the related phages, we performed a BLASTN search using the whole genome sequence of Izhevsk as the query. The linear comparison diagram showing the similarity between phage genomes was produced with EasyFig [24]. The number of proteins shared by phages was computed using the GET_HOMOLOGUES software [16] with the COGtriangles algorithm [17,18] (-t 0 -C 50 -e). VICTOR [25] was used for phylogenetic inference from the whole proteome sequences of Izhevsk and the closest viruses using the Genome-BLAST Distance Phylogeny method [26] with the settings recommended for prokaryotic viruses [25]. The resulting tree was rooted at the midpoint [27] and visualized with FigTree v1.4.4 [28].

Accession number

The genome sequence of phage Izhevsk is available in the GenBank under accession number MT254578.

Results

Isolation, host range and lytic spectrum

Phage Izhevsk was isolated from soil samples collected in Izhevsk, Udmurtia, Russian Federation, in 2016. The host range was determined on 34 bacterial strains of 3 different species belonging to the *Bacillus cereus sensu lato* group (Table 1). The phage was capable of forming plaques on the lawns of 33 strains. On the propagating host strain *B. cereus* VKM B-370, Izhevsk produced turbid plaques with an approximate diameter of 1 mm (S1 Fig).

Transmission electron microscopy

TEM analysis revealed that Izhevsk possesses a long flexible non-contractile tail with an approximate length of 430 nm and a non-elongated capsid, approximately 80 nm in diameter (Fig 1), the two features characteristic of the *Siphoviridae* B1 morphotype [29]. The tail contains about 120 visible striations (disk-like structures) (Fig 1B). At the end of the tail, the base-plate structure is visible with extending fiber-like structures of up to 30 nm (Fig 1C–1F), through which some phage particles are attached to one another (Fig 1A–1C).

Genome organization

A single circular contig with the length of 168350 bp, the average coverage of 68x and the GC-content of 34.3% was generated from sequencing reads. A simple commonly used method to predict phage genome termini is to find coverage deviations when aligning reads on the assembled genomic sequence [30]. No regions deviating ≥ 2 times from the average coverage were presented on read coverage diagram (not shown), so it was complicated to find the genome termini by this approach. The putative ends positions were roughly located taking into account the significant similarity between the Izhevsk genomic sequence and that of the previously described bacteriophage vB_BanS_Tsamsa with 284-bp DTR regions determined by the authors [31]. The postulate was then proven by electrophoresis of the Izhevsk DNA digested with the PacI and AfeI restriction enzymes (Fig 2).

The genome contains 256 predicted protein-encoding ORFs; 80 of them (31.25%) were functionally assigned. Like many phages with large genomes, the Izhevsk genes can be classified into several functional groups (Fig 3), with most of them scattered over more than half of the genome and not necessarily forming visible 'gene modules' characteristic of phages with smaller genomes.

Among the best-annotated groups of DNA packaging, structural and host lysis-related genes, we have identified those encoding small and large terminase subunits, portal protein, putative prohead protease, major capsid protein, head completion protein, putative assembly chaperone, tail tube protein, tail completion protein, tape measure protein, YiiX-like amidase, putative tail endopeptidase, putative tail fiber protein, putative receptor-binding protein, endosialidase, holin and endolysin with the N-acetylmuramoyl-L-alanine amidase activity.

Within the cluster of tail genes and in other parts of the genome, we have identified three genes (ORFs 208, 237, 80) encoding the Xer family recombinases known to participate in resolving the dimeric forms of circular DNA molecules such as bacterial chromosomes and plasmids [32]. There have also been several reports on the ability of these proteins to enable integration of mobile elements into the *Vibrio cholerae* chromosome [33–35]. Three of the

Table 1. The host range of phage Izhevsk and the lytic spectrum of its endolysin determined on 34 different *Bacillus cereus sensu lato* strains.

No.	Species	Strain	Phage Lysis	Endolysin lysis		
				Exp. phase*	Stat. phase	Spot test
1	<i>B. cereus</i>	VKM B-13	+	+++	+	+
2	<i>B. cereus</i>	VKM B-370	+	+++	-	+
3	<i>B. cereus</i>	VKM B-373	+	+++	-	+
4	<i>B. cereus</i>	VKM B-374	+	+++	+	+
5	<i>B. cereus</i>	VKM B-383	+	-	+	+
6	<i>B. cereus</i>	VKM B-445	+	++	+	+
7	<i>B. cereus</i>	VKM B-473	+	+	-	+
8	<i>B. cereus</i>	VKM B-491	+	+++	+	+
9	<i>B. cereus</i>	VKM B-504 ^T	+	++	+	+
10	<i>B. cereus</i>	VKM B-681	+	++	+	+
11	<i>B. cereus</i>	VKM B-682	+	+++	+	+
12	<i>B. cereus</i>	VKM B-683	+	+++	+	+
13	<i>B. cereus</i>	VKM B-684	+	+++	+	+
14	<i>B. cereus</i>	VKM B-686	+	+++	+	+
15	<i>B. cereus</i>	VKM B-688	+	+	+	+
16	<i>B. cereus</i>	VKM B-771	+	+	+	+
17	<i>B. cereus</i>	VKM B-810	+	++	+	+
18	<i>B. cereus</i>	VKM B-812	-	++	+	+
19	<i>B. cereus</i>	ATCC 4342	+	+++	+	+
20	<i>B. cereus</i>	ATCC 14893	+	++	+	-
21	<i>B. thuringiensis</i>	VKM B-83	+	+++	+	+
22	<i>B. thuringiensis</i>	VKM B-84	+	+++	+	+
23	<i>B. thuringiensis</i>	VKM B-85	+	+++	+	+
24	<i>B. thuringiensis</i>	VKM B-440	+	-	-	+
25	<i>B. thuringiensis</i>	VKM B-443	+	++	+	+
26	<i>B. thuringiensis</i>	VKM B-446	+	++	+	+
27	<i>B. thuringiensis</i>	VKM B-447	+	+++	+	+
28	<i>B. thuringiensis</i>	VKM B-450	+	-	-	+
29	<i>B. thuringiensis</i>	VKM B-453	+	-	-	+
30	<i>B. thuringiensis</i>	VKM B-454	+	-	-	+
31	<i>B. thuringiensis</i>	VKM B-1555	+	+	+	+
32	<i>B. thuringiensis</i>	VKM B-1557	+	++	+	+
33	<i>B. thuringiensis</i>	ATCC 35646	+	-	+	+
34	<i>B. weihenstephanensis</i>	KBAB4	+	+++	+	+
	<i>B. flexus</i>		-	++	+	-
	<i>B. pumilus</i>		-	-	-	-
	<i>B. subtilis</i>	168	-	+	-	+
	<i>B. subtilis</i>	WB 800n	-	-	-	-
	<i>B. megaterium</i>	MS941	-	++	+	+
	<i>E. faecium</i>	FS86	-	-	-	-

* Results recorded as “+++”, complete or near to complete lysis; “++”, moderate lysis; “+”, slight lysis; and “-”, no detectable lysis.

<https://doi.org/10.1371/journal.pone.0242657.t001>

genes located near ORF237 encode the Replic_Relax superfamily protein (ORF238), ParM-like protein (ORF247) and putative ParG-like protein (ORF246), the common components of plasmid relaxation and partitioning systems. Temperate *Bacillus* phages possess diverse lifestyles

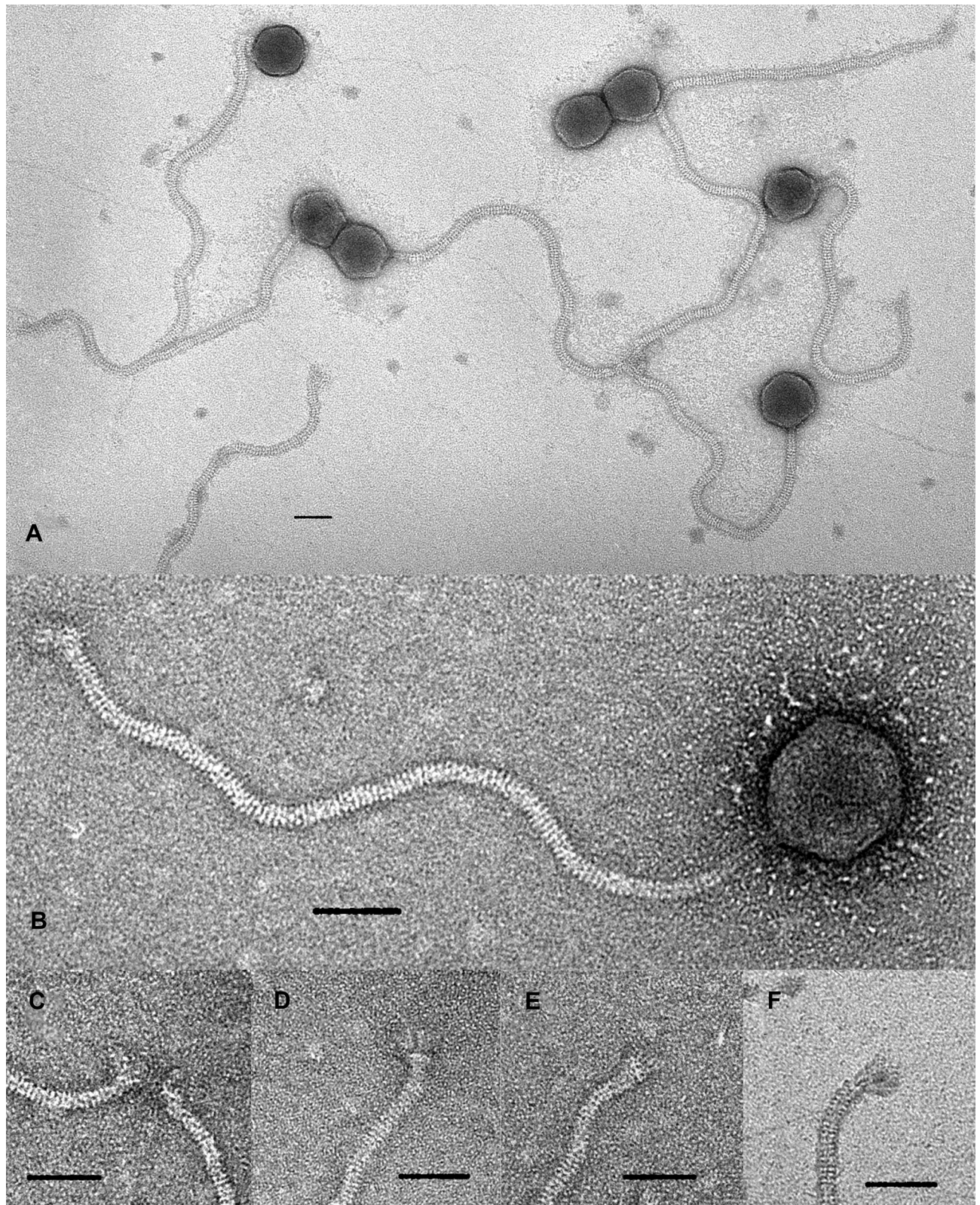


Fig 1. Transmission electron microscopy of *Bacillus* phage Izhevsk particles negatively stained with 1% (w/v) uranyl acetate. Scale bars are 50 nm. A. The bacteriophage particles. B. Closeup view of a single phage particle. C-F. Baseplate and fiber-like structures, closeup view.

<https://doi.org/10.1371/journal.pone.0242657.g001>

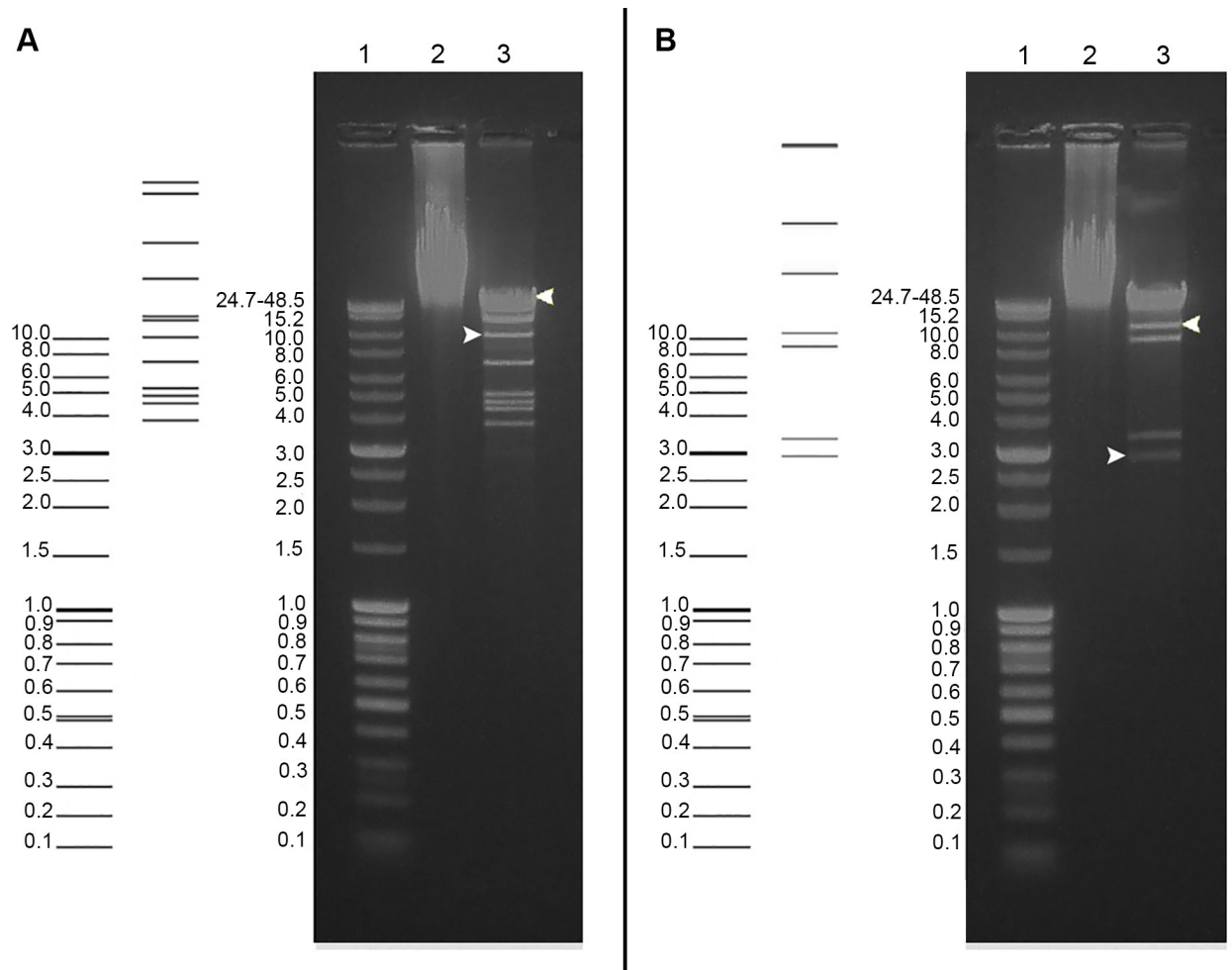


Fig 2. Phage DNA digested with restriction endonucleases PacI (A) and AfeI (B) *in silico* and *in vitro* (the left and right side of each image, respectively). Lanes (for each image): 1 –molecular weight markers (SibEnzyme, mix of Cat.# M29 and M33), 2 –intact phage DNA, 3 –digested DNA. The lengths of the fragments predicted for PacI: 33909, 32075, 24289, 18753, 13056, 12356, 9779, 6893, 4963, 4570, 4274, 3721; for AfeI: 54176, 40885, 27581, 19892, 10860, 8996, 3337, 2920. Terminal fragments are underlined here in the text and indicated by the arrows on the electrophoregrams.

<https://doi.org/10.1371/journal.pone.0242657.g002>

and can integrate their DNA into host chromosomes and plasmids or replicate without integration as circular or linear plasmids [12]. Phages with circular plasmid lysogenic stages usually feature the Xer family recombinases and ParMRC or similar segregation systems [36,37]. Thus, taking into account the presence of plasmid replication and plasmid segregation-related set of genes, it seems possible that Izhevsk can exist as a circular plasmid prophage in the host cytoplasm while not necessarily integrating its DNA into the host chromosome.

The replication and recombination-related genes of Izhevsk encode the putative lambda O-type replication initiator, RecD-like DNA helicase, single-stranded DNA-binding protein, DnaB-type helicase, DnaG-type primase, RecJ-type DNA exonuclease, DNA gyrase subunit A, DNA gyrase subunit B, DNA ligase, DNA polymerase III alpha subunit and RuvC-type Holliday junction resolvase.

Another unusually large gene group highlighted in purple on the genome map includes the genes encoding products involved in different processes of nucleic acid metabolism such as DNA nucleotide synthesis, repair and tRNA splicing: deoxyribonucleoside kinase, Appr-1”p-

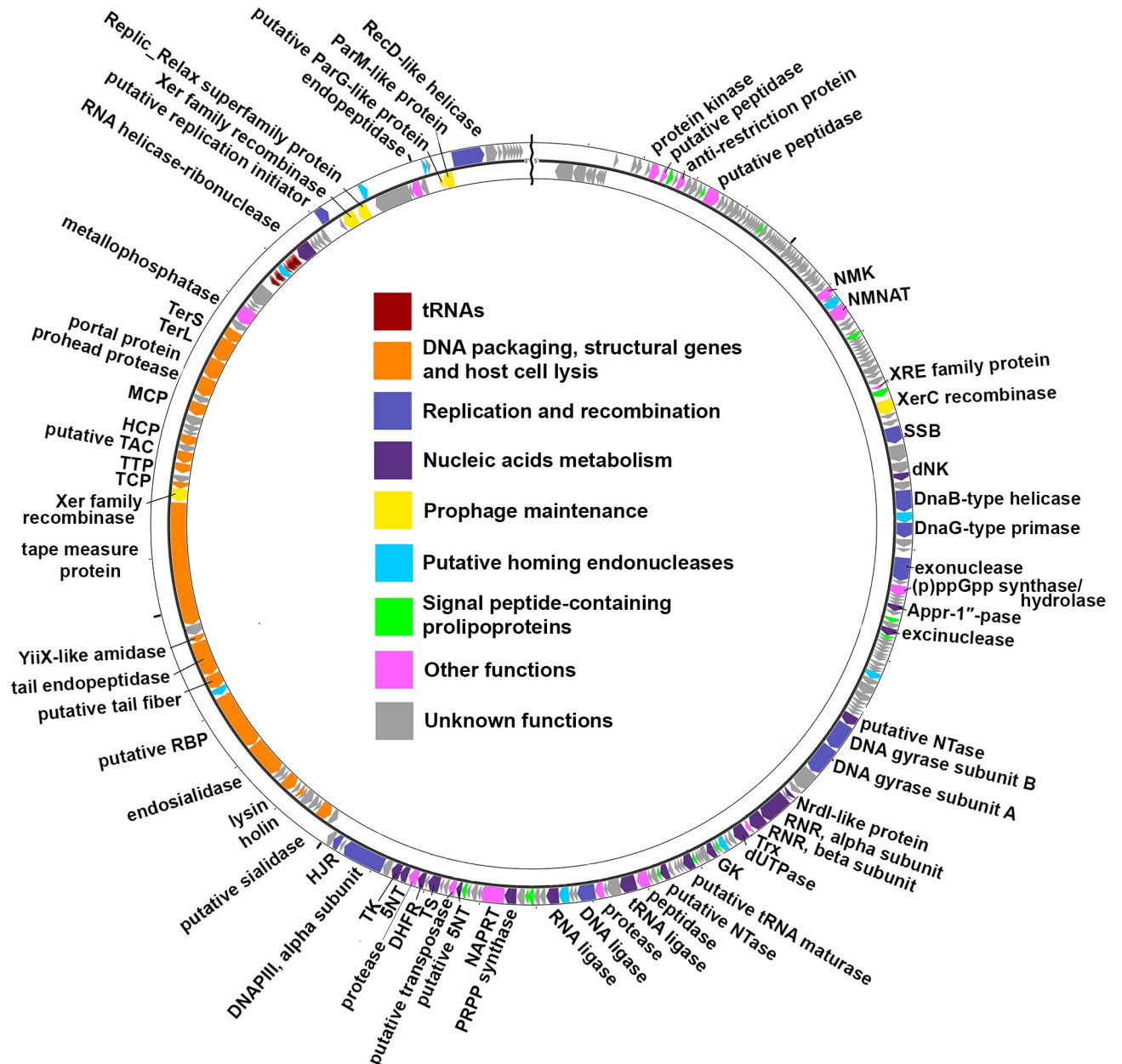


Fig 3. The Izhevsk genome map. Functionally assigned ORFs are highlighted according to their general functions (see the legend). Abbreviations: NMK—nicotinamide mononucleotide transporter, NMNAT—nicotinamide mononucleotide adenyltransferase, SSB—single-stranded DNA-binding protein, dNK—deoxyribonucleoside kinase, NTase—nucleotidyltransferase, RNR—Ribonucleotide reductase, Trx—thioredoxin, GK—guanylate kinase, PRPP synthase—Phosphoribosyl diphosphate synthase, NAPRT—Nicotinate Phosphoribosyltransferase, 5NT—5'-Nucleotidase, TS—Thymidylate synthase, DHFR—Dihydrofolate reductase, TK—thymidine kinase, DNAP—DNA polymerase, HJR—Holliday junction resolvase, RBP—receptor-binding protein, TCP—tail completion protein, TTP—tail tube protein, TAC—tail assembly chaperone, HCP—head completion protein, MCP—major capsid protein, TerL—terminase, large subunit, TerS—terminase, small subunit.

<https://doi.org/10.1371/journal.pone.0242657.g003>

processing protein, excision endonuclease, putative nucleotidyltransferases, NrdI-like protein, alpha and beta subunits of ribonucleotide reductase, dUTP diphosphatase, guanylate kinase, putative tRNA maturase, tRNA ligase, RNA ligase, phosphoribosyl pyrophosphate synthetase, putative 5'-nucleotidases, thymidine kinase and RNA helicase-ribonuclease.

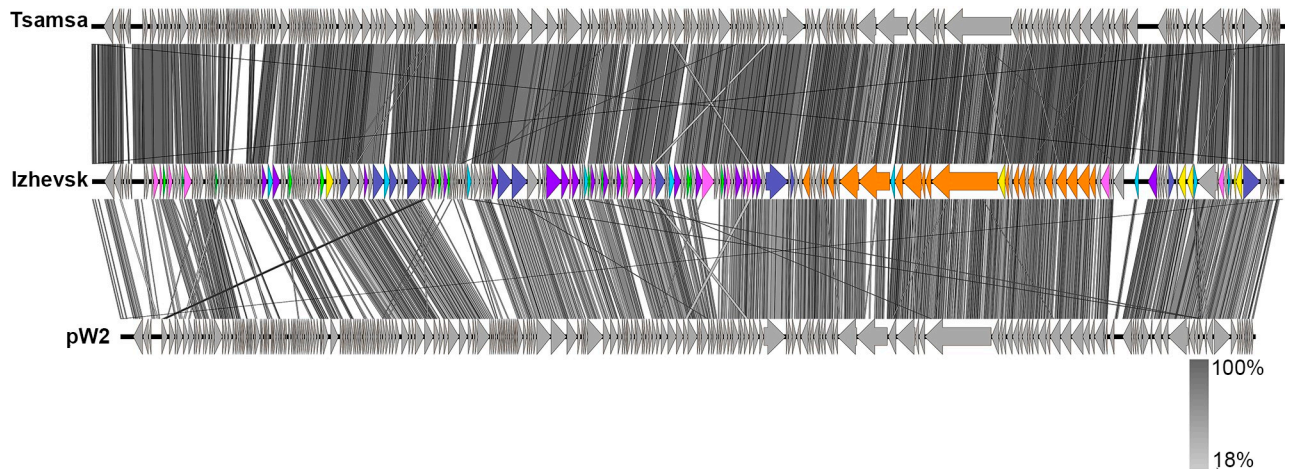


Fig 4. Pairwise TBLASTX comparisons of the Izhevsk genome and the related genomes visualized with Easyfig. The Izhevsk genome scheme coloring corresponds to Fig 3. Grey areas between the genome maps indicate the level of identity (see the legend).

<https://doi.org/10.1371/journal.pone.0242657.g004>

In total, nine genes were predicted to encode HNH- and GIY-YIG-family endonucleases and were assigned the putative homing endonuclease function based on their absence in the respective positions in the genomes of related phages, Tsamsa and pW2 (Fig 4), which may be considered evidence of their potential mobility [38]. Homing endonucleases are mobile elements which are common in bacterial and large phage genomes and promote their own spread between related genomes by the double-strand break repair pathway [39]. All the putative homing endonucleases of Izhevsk are free-standing, although several phages possess intron-encoded homing endonucleases [40,41].

Among the rest of the genes not assigned to any of the main modules, the most noticeable is ORF17 encoding putative DNA restriction enzyme inhibitor similar to T7 protein Ocr, mimicking B-form DNA and inhibiting the restriction-modification activity of the Type I EcoKI complex [42]. Also, the NAD metabolism-related genes have been identified encoding: PnuC-like nicotinamide mononucleotide transporter (ORF59), NadR-type nicotinamide nucleotide adenyltransferase (ORF62) and nicotinate phosphoribosyltransferase (ORF169). In addition to the extended gene baggage that seems to include not only the minimum set needed for phage functions but also the genes beneficial for the host metabolism, Izhevsk possesses a tRNA array consisting of 16 tRNA and 2 pseudo-tRNA genes, predicted with tRNAscan-SE [20].

Altogether 13 ORFs were predicted to encode lipoprotein signal peptides. Although first described long ago [43], phage-encoded lipoproteins are still considered rare, probably because signal sequence-prediction tools are not generally included in the annotation pipelines for bacteriophage genomes. The presence of lipoprotein-encoding ORFs is mostly characteristic of temperate phages, where these molecules have been shown to contribute to superinfection exclusion, lysis, host virulence [44] and lysis-lysogeny switch [45].

The related phage genomes were found by BLASTN search with the whole genome sequence of Izhevsk against the NCBI nr database (taxid:10239) and the *Bacillus* Phage Database (BPD) (<http://bacillus.phagesdb.org/>). Eight complete phage genomes, including previously described phages Tsamsa [31] and PBC2 [22], were downloaded from GenBank and another closely related phage named Diildio was downloaded from BPD.

Whole-genome comparison based on the translated sequences of all predicted ORFs was performed with VICTOR (formula D6), and the resulting phylogram shown in Fig 5 illustrates

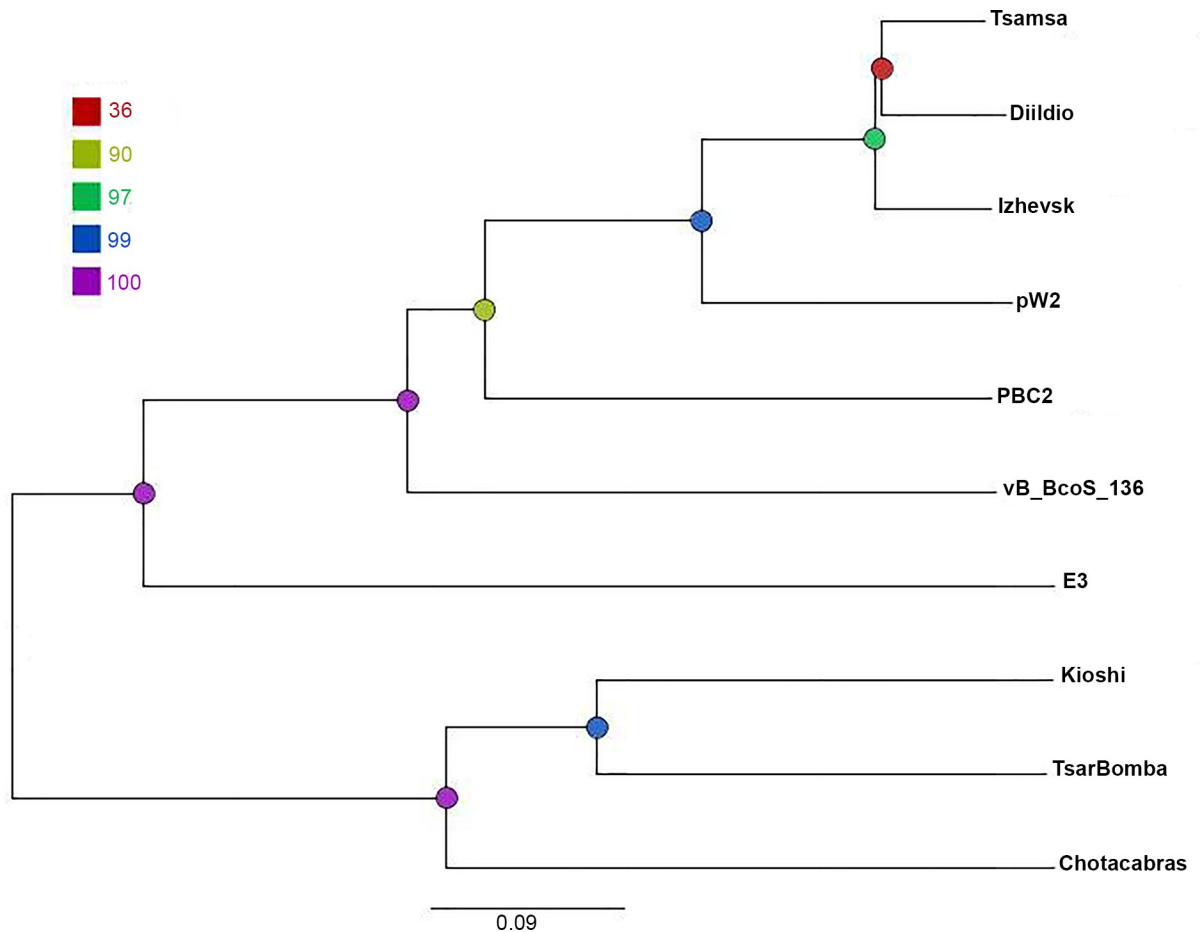


Fig 5. Genome-BLAST Distance Phylogeny (GBDP) phylogram based on the amino-acid sequences of whole proteomes and inferred using the formula D6. The nodes are colored according to the legend representing GBDP pseudo-bootstrap support values from 100 replications. The branch lengths are scaled in terms of the GBDP distance formula D6.

<https://doi.org/10.1371/journal.pone.0242657.g005>

the evolutionary relationship between Izhevsk and the closest viruses. The branching order is consistent with the percentile amount of proteins shared by the five most closely related phages with Izhevsk (Table 2).

Table 2. The properties of Izhevsk and the most closely related genomes.

Name	Database accession number	Genome length	ORFs (protein-coding)	GC-content	BLASTN nucleotide identity to Izhevsk, %*	Proteins shared with Izhevsk, number (%)**	Number of tRNAs	Number of proteins shared with the phage by all of the previous phages of the table.
Izhevsk	MT254578.1	168638	256	34.3			18	
Tsamsa	NC_023007.1	168876	272	34.3	88.6	220 (80.88)	19	220
Diildio		171148	267	34.3	87.5	214 (80,15)	18	204
pW2	MK288021.1	160627	255	32.5	57.2	154 (60.39)	4	150
PBC2	KT070867.1	168689	251	34.4	42.6	128 (51)	17	103
vB_BcoS_136	MH884508.1	160590	239	32.2	29.2	101 (42,26)	17	82

*Determined using BLASTN compared to Izh57 (multiplying % coverage by % identity)

**Determined using GET_HOMOLOGUES (COGtriangles algorithm, -t 0, -C 50, -e).

<https://doi.org/10.1371/journal.pone.0242657.t002>

Specificity range, homologs and thermostability of Ply57

ORF194 and ORF196 of phage Izhevsk were annotated as holin and endolysin, respectively. Like most endolysins from the phages infecting Gram-positive bacteria, Ply57 has a modular organization and consists of three domains: N-terminal enzymatic active domain (EAD) and two C-terminal cell wall binding domains (CBD). BLASTp analysis revealed several highly related endolysins including those of phages vB_BanS-Tsamsa (96.81% aa identity), pW2 (84.76%), PBC2 (80.88%), and bg2 (79.94%) (Fig 6). In contrast to endolysins of PBC2 and bg2 phages, Ply57, like endolysins of Tsamsa and pW2 phages, contains a short linker part connecting the EAD and CBD domains.

Like endolysins of closely related phages Tsamsa [31] and PBC2 [22], Ply57 shows a broad range of bacteriolytic activity against the collection of laboratory strains. Taking together the results of turbidimetry reduction and spot test assays, all of the 34 *B. cereus* group strains, as well as *B. megaterium* MS941, *B. subtilis* 168 and *B. flexus* IPP1, are sensitive to Ply57 (Table 1). Turbidimetry reduction assay has shown that two strains (*B. cereus* VKM B-383 and ATCC 35646) are not sensitive to Ply57 in the exponential growth phase, but become sensitive to Ply57 in the stationary phase. *B. cereus* VKM B-370, VKM B-373 and VKM B-473 strains are sensitive to Ply57 during exponential growth, but resistant in the stationary phase. *B. thuringiensis* VKM B-440, VKM B-453, VKM B-454 strains are not sensitive in both exponential

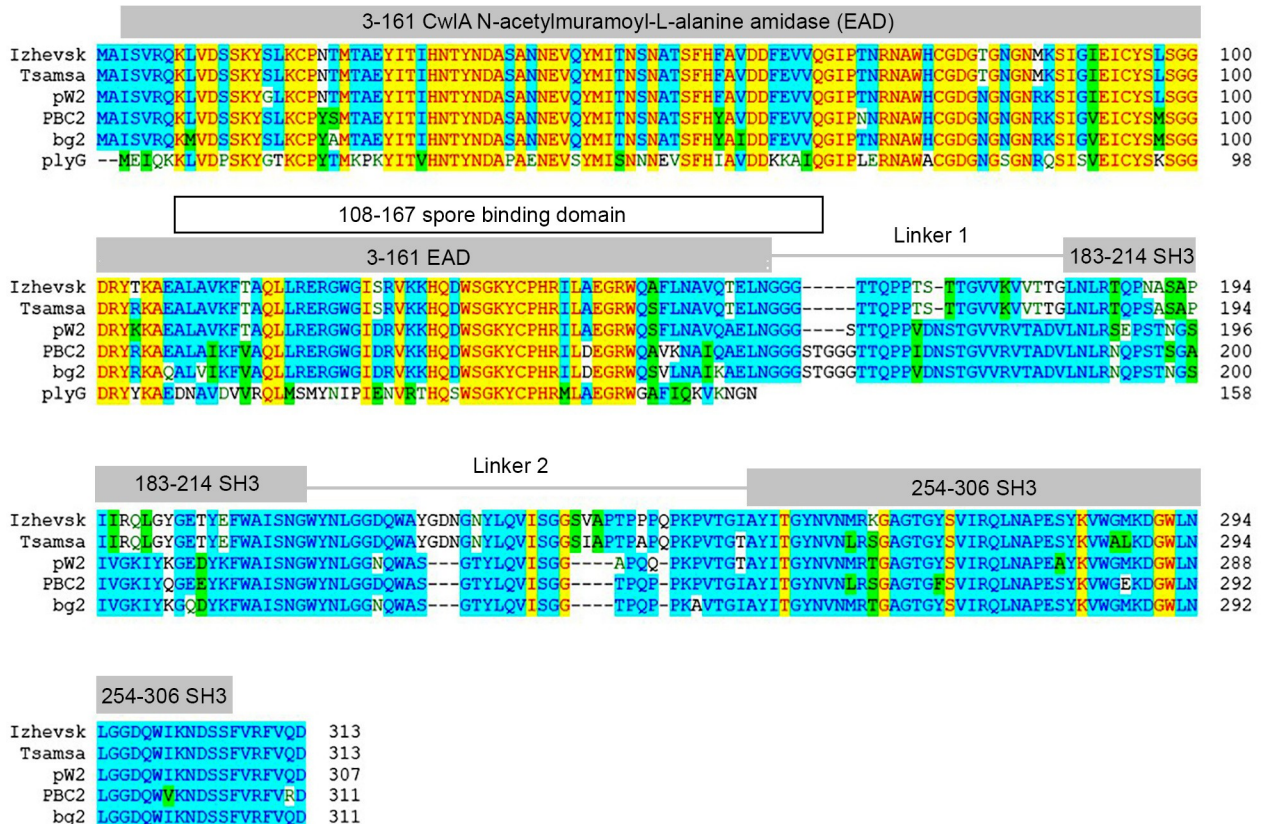


Fig 6. Multiple amino acid sequence alignment of Ply57 (ORF196) from bacteriophage Izhevsk with endolysins from vB_BanS-Tsamsa (YP_008873459.1), pW2 (AZU98917.1), PBC2 (AKQ08512.1), bg2 (ABX56142.1) and the enzymatically active domain of PlyG from *Bacillus phage gamma* (2L47_A). Conservative amino acids are highlighted in blue; amino acids matching the residues of the PlyG endolysin are yellow; synonymous substitutions are green. Conservative domain regions revealed with DELTA-BLAST are indicated above the aligned sequences.

<https://doi.org/10.1371/journal.pone.0242657.g006>

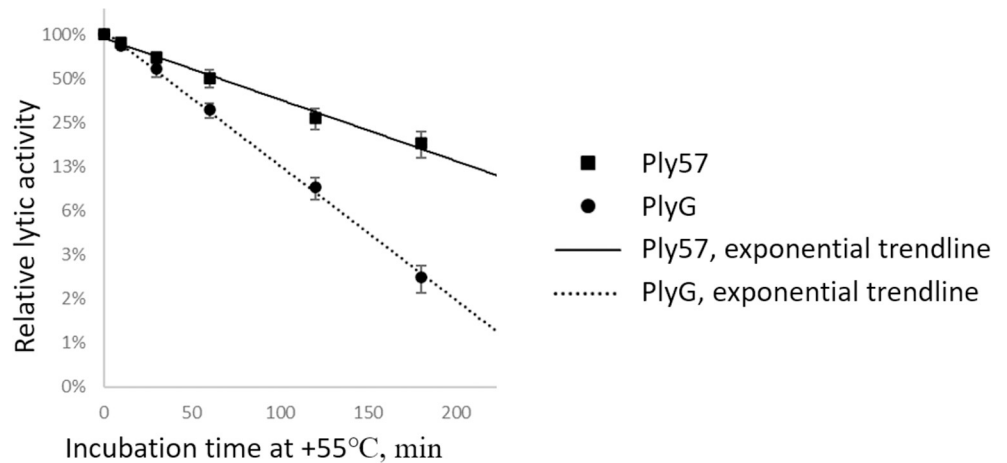


Fig 7. The thermal kinetic inactivation half-life experiment. Error bars indicate standard deviation values based on four replicates.

<https://doi.org/10.1371/journal.pone.0242657.g007>

and stationary growth phases. However, all of the nine strains mentioned above were sensitive to Ply57 in spot assay. On the other hand, *B. cereus* ATCC 14893 was not sensitive to Ply57 in spot test, but appeared to be sensitive in turbidimetry reduction assay (Table 1). Apparently, the bacteriolytic spectrum of endolysins must be tested using several different approaches to exclude false-negative results, which may be due to changes in the peptidoglycan structure during bacterial culture growth as well as changes in growth conditions [46].

The N-terminal (aa residues 1–165) EAD of Ply57 shows a 53% identity to the EAD of the PlyG endolysin from phage Wβ, the best-characterized endolysin from the phages infecting the *B. cereus* group of bacteria [23]. Since the stability of bacteriolytic enzymes is one of the most essential criteria for the development of antimicrobial compounds, we have determined the relative thermostability of Ply57 compared to that of PlyG. For this purpose, the enzymes were incubated at 55°C and aliquots were removed at different time intervals. The residual bacteriolytic activity is exponentially dependent on the incubation time at 55°C, with the R^2 values of 0.98 for Ply57 and 0.99 for PlyG (Fig 7). The thermostability assay has shown that Ply57 is two-fold more stable than PlyG under the experimental conditions. The time required for half-inactivation of the Ply57 bacteriolytic activity was 65 ± 9 min whereas the half-inactivation time for PlyG was found to be 35 ± 3 min.

Discussion

The temperate *Bacillus* phage Izhevsk isolated and characterized in this work belongs to a distinct group of viruses within the *Siphoviridae* family along with the previously described phages Tsamsa [31] and PBC2 [22] and three other *Bacillus*-infecting viruses: Diildio, pW2 and vB_BcoS_136.

There have been many genome-based approaches proposed, aimed at establishing a universal scheme of phage taxonomical classification including those using concatenated phylogeny, shared protein content and pairwise genomic distances [25,47,48]. The nucleotide identity level of 95% is now used by the International Committee on Taxonomy of Viruses (ICTV) as a strict criterion for species demarcation, whereas the boundaries of higher taxa are less conservative.

In several extensive works, the phages of the order *Caudovirales* were assigned to one genus on the basis of shared protein content (>40%), similar morphology, whole-genome

collinearity and shared key features [49–51]. Our comparative analyses have revealed that the group containing the four most closely related phages: Izhevsk, Tsamsa, Diildio and pW2, possesses 150 shared proteins (Table 2), i.e., 57.2% on average (in terms of the total number of proteins encoded in each of the genomes).

Although not all of these phages have been described in detail, their genomes have similar gene orders with highly identical DNA packaging, structural and replication genes (Fig 4) and possess the homologs of all six Izhevsk genes (ORFs 80, 208, 237, 238, 246, 247) predicted to have prophage-related functions, suggesting these are all temperate bacteriophages with the *Siphoviridae* morphology using short DTR packaging strategy. However, there are only 4 tRNA-coding genes in the genome of pW2, which differs significantly from the Izhevsk, Tsamsa and Diildio genomes harboring 18–19 tRNA genes.

The length and GC-content of the pW2 genome are two more characteristics deviating much stronger away from Izhevsk, Tsamsa and Diildio compared to more distantly related PBC2, which shares much less genes with the rest of the phages (Table 2). Thus, while considering the available set of genomes we could not decide whether pW2 and PBC2 can be placed into the same genus with Izhevsk, Tsamsa and Diildio; yet, the latter three phages can definitely be classified into a single genus in conformity with the accepted principles.

Among closely related phages, bacteriophage Izhevsk has the broadest specificity range: it is capable of infecting 97% of the tested *B. cereus sensu lato* strains (Table 1), whereas the specificity range of other related phages is quite narrow. The PBC2 phage was able to infect only one out of 11 tested *B. cereus sensu lato* strains [22]. Bacteriophage Tsamsa lysed 15 out of 38 (39%) tested *B. cereus sensu lato* strains, six of them being *B. anthracis* strains [31].

The Izhevsk endolysin Ply57 is active when produced recombinantly and shows a broad lytic spectrum. Ply57 belongs to the N-acetylmuramoyl-L-alanine amidase family and is highly similar to several other *Bacillus* phage endolysins, including the endolysins of phages Tsamsa (96.81% aa identity), pW2 (84.76%), PBC2 (80.88%) bg2 (79.94%) and Wβ (53% for EAD only). It has been reported that PlyG EAD contains a spore binding domain (SBD) allowing this molecule to degrade *B. anthracis* spores upon germination, even when CBD is completely deleted [52]. Structurally, SBD is a part of EAD but harbors amino acids involved in amidase catalytic activity (PDB 2L47). PlyG SBD is highly specific to *B. anthracis* spores [52]. Another characterized endolysin (from phage PBC2) also harbors a conservative SBD, which is specific to *B. cereus* spores, in contrast to the PlyG SBD [22]. We suppose that Ply57 can be active against germinating spores, similar to PBC2 endolysin, because the amino acid sequences of these endolysins corresponding to SBDs share 87% similarity (85% identity) (Fig 6).

All three endolysins of the Izhevsk, PBC2 and Tsamsa bacteriophages demonstrate a broad range of activity against the *B. cereus* group [22, 31]. PBC2 endolysin was active against *B. subtilis*, *B. megaterium*, *B. circulans*, *B. licheniformis*, *B. pumilus*, *L. monocytogenes*, *C. perfringens* and *S. aureus* [22]. Ply57 was also active against one of the two *B. subtilis* strains tested, as well as against *B. megaterium* and *B. flexus*, but not active against *B. pumilus* and *Enterococcus faecium* (Table 1). Ply57 and PBC2 endolysin may have different specificity due to the structural differences within the proximal SH3 domain and the linker connecting this domain to EAD. It has been recently shown that the length and flexibility of the linker part might affect thermal stability and bactericidal properties of the endolysin [53]. The linker part of Ply57, as well as that of bacteriophage Tsamsa endolysin, is six amino acids shorter than the linker in PBC2 endolysin.

Taken together, the broad specificity of phage Izhevsk, the high thermostability and broad specificity of its endolysin, Ply57, these objects are a useful platform for developing antibacterial agents and pathogen control tools to prevent bacterial contaminations caused by *B. cereus* group bacteria, including *B. anthracis*.

Supporting information

S1 Fig. *Bacillus* phage Izhevsk plaque morphology on 0.75% w/v LB overlay.
(PDF)

S1 Raw images.
(PDF)

S1 Table. Annotation of *Bacillus* phage Izhevsk.
(PDF)

Author Contributions

Conceptualization: Andrey M. Shadrin.

Data curation: Anna V. Skorynina, Olesya A. Kazantseva, Natalya A. Ryabova, Andrey M. Shadrin.

Formal analysis: Anna V. Skorynina, Emma G. Piligrimova, Olesya A. Kazantseva, Andrey M. Shadrin.

Funding acquisition: Andrey M. Shadrin.

Investigation: Anna V. Skorynina, Emma G. Piligrimova, Olesya A. Kazantseva, Vladislav A. Kulyabin, Svetlana D. Baicher, Natalya A. Ryabova, Andrey M. Shadrin.

Methodology: Anna V. Skorynina, Emma G. Piligrimova, Olesya A. Kazantseva, Andrey M. Shadrin.

Project administration: Andrey M. Shadrin.

Resources: Andrey M. Shadrin.

Software: Olesya A. Kazantseva.

Supervision: Andrey M. Shadrin.

Validation: Anna V. Skorynina, Emma G. Piligrimova, Olesya A. Kazantseva, Andrey M. Shadrin.

Visualization: Emma G. Piligrimova, Olesya A. Kazantseva, Andrey M. Shadrin.

Writing – original draft: Emma G. Piligrimova, Andrey M. Shadrin.

Writing – review & editing: Andrey M. Shadrin.

References

1. García P, Martínez B, Rodríguez L, Rodríguez A. Synergy between the phage endolysin LysH5 and nisin to kill *Staphylococcus aureus* in pasteurized milk. *International journal of food microbiology*. 2010 Jul 15; 141(3):151–5. <https://doi.org/10.1016/j.ijfoodmicro.2010.04.029> PMID: 20537744
2. Guo T, Xin Y, Zhang C, Ouyang X, Kong J. The potential of the endolysin Lysdb from *Lactobacillus delbrueckii* phage for combating *Staphylococcus aureus* during cheese manufacture from raw milk. *Applied microbiology and biotechnology*. 2016 Apr 1; 100(8):3545–54. <https://doi.org/10.1007/s00253-015-7185-x> PMID: 26621799
3. Gondil VS, Harjai K, Chhibber S. Endolysins as emerging alternative therapeutic agents to counter drug-resistant infections. *International journal of antimicrobial agents*. 2020 Feb 1; 55(2):105844. <https://doi.org/10.1016/j.ijantimicag.2019.11.001> PMID: 31715257
4. Trigo G, Martins TG, Fraga AG, Longatto-Filho A, Castro AG, Azeredo J, et al. Phage therapy is effective against infection by *Mycobacterium ulcerans* in a murine footpad model. *PLoS neglected tropical diseases*. 2013 Apr; 7(4).

5. Nelson DC, Schmelcher M, Rodriguez-Rubio L, Klumpp J, Pritchard DG, Dong S, et al. Endolysins as antimicrobials. *Advances in virus research* 2012 Jan 1 (Vol. 83, pp. 299–365). Academic Press. <https://doi.org/10.1016/B978-0-12-394438-2.00007-4> PMID: 22748813
6. Loessner MJ. Bacteriophage endolysins—current state of research and applications. *Current opinion in microbiology*. 2005 Aug 1; 8(4):480–7. <https://doi.org/10.1016/j.mib.2005.06.002> PMID: 15979390
7. Carroll LM, Wiedmann M, Kovac J. Proposal of a taxonomic nomenclature for the *Bacillus cereus* group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. *Mbio*. 2020 Feb 25; 11(1). <https://doi.org/10.1128/mBio.00034-20> PMID: 32098810
8. Kumari S, Sarkar PK. Prevalence and characterization of *Bacillus cereus* group from various marketed dairy products in India. *Dairy science & technology*. 2014 Sep 1; 94(5):483–97.
9. Gundogan N, Avci E. Occurrence and antibiotic resistance of *Escherichia coli*, *Staphylococcus aureus* and *Bacillus cereus* in raw milk and dairy products in Turkey. *International journal of dairy technology*. 2014 Nov; 67(4):562–9.
10. Owusu-Kwarteng J, Wuni A, Akabanda F, Tano-Debrah K, Jespersen L. Prevalence, virulence factor genes and antibiotic resistance of *Bacillus cereus sensu lato* isolated from dairy farms and traditional dairy products. *BMC microbiology*. 2017 Dec; 17(1):65. <https://doi.org/10.1186/s12866-017-0975-9> PMID: 28288581
11. Bardell D. An 1898 report by Gamaleya for a lytic agent specific for *Bacillus anthracis*. *Journal of the history of medicine and allied sciences*. 1982 Apr 1; 37(2):222–5. <https://doi.org/10.1093/jhmas/xxxvii.2.222> PMID: 6806352
12. Gillis A, Mahillon J. Phages preying on *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*: past, present and future. *Viruses*. 2014 Jul; 6(7):2623–72. <https://doi.org/10.3390/v6072623> PMID: 25010767
13. Pilgrimova EG, Kazantseva OA, Nikulin NA, Shadrin AM. *Bacillus* Phage vB_BtS_B83 Previously Designated as a Plasmid May Represent a New *Siphoviridae* Genus. *Viruses*. 2019 Jul; 11(7):624. <https://doi.org/10.3390/v11070624> PMID: 31284652
14. Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: a laboratory manual*. Cold spring harbor laboratory press; 1989.
15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*. 2012 May 1; 19(5):455–77. <https://doi.org/10.1089/cmb.2012.0021> PMID: 22506599
16. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific reports*. 2015 Feb 10; 5:8365. <https://doi.org/10.1038/srep08365> PMID: 25666585
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of molecular biology*. 1990 Oct 5; 215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. BLAST: Basic local alignment search tool [Internet]. National Center for Biotechnology Information. U.S. National Library of Medicine; 1990 [cited 2020May22]. Available from: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
19. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of molecular biology*. 2018 Jul 20; 430(15):2237–43. <https://doi.org/10.1016/j.jmb.2017.12.007> PMID: 29258817
20. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic acids research*. 1997 Mar 1; 25(5):955–64. <https://doi.org/10.1093/nar/25.5.955> PMID: 9023104
21. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView. *Bioinformatics*. 2005 Feb 15; 21(4):537–9. <https://doi.org/10.1093/bioinformatics/bti054> PMID: 15479716
22. Kong M, Na H, Ha NC, Ryu S. LysPBC2, a novel endolysin harboring a *Bacillus cereus* spore binding domain. *Appl. Environ. Microbiol*. 2019 Mar 1; 85(5):e02462–18. <https://doi.org/10.1128/AEM.02462-18> PMID: 30552194
23. Heselpoth RD, Owens JM, Nelson DC. Quantitative analysis of the thermal stability of the gamma phage endolysin PlyG: a biophysical and kinetic approach to assaying therapeutic potential. *Virology*. 2015 Mar 1; 477:125–32. <https://doi.org/10.1016/j.virol.2014.11.003> PMID: 25432575
24. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics*. 2011 Apr 1; 27(7):1009–10. <https://doi.org/10.1093/bioinformatics/btr039> PMID: 21278367

25. Meier-Kolthoff JP, Göker M. VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics*. 2017 Nov 1; 33(21):3396–404. <https://doi.org/10.1093/bioinformatics/btx440> PMID: 29036289
26. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC bioinformatics*. 2013 Dec 1; 14(1):60. <https://doi.org/10.1186/1471-2105-14-60> PMID: 23432962
27. Farris JS. Estimating phylogenetic trees from distance matrices. *The American Naturalist*. 1972 Sep 1; 106(951):645–68.
28. Rambaut A. A graphical viewer of phylogenetic trees 487 and a program for producing publication-ready figures [Internet]. FigTree. [cited 2020May22]. Available from: <http://tree.bio.ed.ac.uk/software/figtree/>
29. Ackermann HW. Frequency of morphological phage descriptions in the year 2000. *Archives of virology*. 2001 May 1; 146(5):843–57. <https://doi.org/10.1007/s007050170120> PMID: 11448025
30. Russell DA. Sequencing, assembling, and finishing complete bacteriophage genomes. In *Bacteriophages 2018* (pp. 109–125). Humana Press, New York, NY.
31. Ganz HH, Law C, Schmuki M, Eichenseher F, Calendar R, Loessner MJ, et al. Novel giant siphovirus from *Bacillus anthracis* features unusual genome characteristics. *PLoS One*. 2014; 9(1). <https://doi.org/10.1371/journal.pone.0085972> PMID: 24475065
32. Blakely G, May G, McCulloch R, Arciszewska LK, Burke M, Lovett ST, et al. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell*. 1993 Oct 22; 75(2):351–61. [https://doi.org/10.1016/0092-8674\(93\)80076-q](https://doi.org/10.1016/0092-8674(93)80076-q) PMID: 8402918
33. Huber KE, Waldor MK. Filamentous phage integration requires the host recombinases XerC and XerD. *Nature*. 2002 Jun; 417(6889):656–9. <https://doi.org/10.1038/nature00782> PMID: 12050668
34. Campos J, Martínez E, Izquierdo Y, Fando R. VEJ ϕ . A novel filamentous phage of *Vibrio cholerae* able to transduce the cholera toxin genes. *Microbiology*. 2010 Jan 1; 156(1):108–15. <https://doi.org/10.1099/mic.0.032235-0> PMID: 19833774
35. Midonet C, Das B, Paly E, Barre FX. XerD-mediated FtsK-independent integration of TLC ϕ into the *Vibrio cholerae* genome. *Proceedings of the National Academy of Sciences*. 2014 Nov 25; 111(47):16848–53. <https://doi.org/10.1073/pnas.1404047111> PMID: 25385643
36. Yuan Y, Peng Q, Wu D, Kou Z, Wu Y, Liu P, et al. Effects of actin-like proteins encoded by two *Bacillus pumilus* phages on unstable lysogeny, revealed by genomic analysis. *Appl. Environ. Microbiol*. 2015 Jan 1; 81(1):339–50. <https://doi.org/10.1128/AEM.02889-14> PMID: 25344242
37. Sakaguchi Y, Hayashi T, Kurokawa K, Nakayama K, Oshima K, Fujinaga Y, et al. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. *Proceedings of the National Academy of Sciences*. 2005 Nov 29; 102(48):17472–7. <https://doi.org/10.1073/pnas.0505503102> PMID: 16287978
38. Edgell DR. Free-standing homing endonucleases of T-even phage: freeloaders or functionaries? *Homing endonucleases and Inteins 2005* (pp. 147–160). Springer, Berlin, Heidelberg.
39. Chevalier BS, Stoddard BL. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic acids research*. 2001 Sep 15; 29(18):3757–74. <https://doi.org/10.1093/nar/29.18.3757> PMID: 11557808
40. Bell-Pedersen D, Quirk S, Clyman J, Belfort M. Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. *Nucleic Acids Research*. 1990 Jul 11; 18(13):3763–70. <https://doi.org/10.1093/nar/18.13.3763> PMID: 2165250
41. Landthaler M, Lau NC, Shub DA. Group I intron homing in *Bacillus* phages SPO1 and SP82: a gene conversion event initiated by a nicking homing endonuclease. *Journal of bacteriology*. 2004 Jul 1; 186(13):4307–14. <https://doi.org/10.1128/JB.186.13.4307-4314.2004> PMID: 15205433
42. Tock MR, Dryden DT. The biology of restriction and anti-restriction. *Current opinion in microbiology*. 2005 Aug 1; 8(4):466–72. <https://doi.org/10.1016/j.mib.2005.06.003> PMID: 15979932
43. Reeve JN, Shaw JE. Lambda encodes an outer membrane protein: the *lom* gene. *Molecular and General Genetics MGG*. 1979 Jan 1; 172(3):243–8. <https://doi.org/10.1007/BF00271723> PMID: 45607
44. Sun X, Göhler A, Heller KJ, Neve H. The *ltp* gene of temperate *Streptococcus thermophilus* phage TP-J34 confers superinfection exclusion to *Streptococcus thermophilus* and *Lactococcus lactis*. *Virology*. 2006 Jun 20; 350(1):146–57. <https://doi.org/10.1016/j.virol.2006.03.001> PMID: 16643978
45. Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, et al. Communication between viruses guides lysis–lysogeny decisions. *Nature*. 2017 Jan; 541(7638):488–93. <https://doi.org/10.1038/nature21049> PMID: 28099413

46. Humann J, Lenz LL. Bacterial peptidoglycan-degrading enzymes and their impact on host muropeptide detection. *Journal of innate immunity*. 2009; 1(2):88–97. <https://doi.org/10.1159/000181181> PMID: [19319201](https://pubmed.ncbi.nlm.nih.gov/19319201/)
47. Low SJ, Džunková M, Chaumeil PA, Parks DH, Hugenholtz P. Evaluation of a concatenated protein phylogeny for classification of tailed double-stranded DNA viruses belonging to the order Caudovirales. *Nature microbiology*. 2019 Aug; 4(8):1306–15. <https://doi.org/10.1038/s41564-019-0448-z> PMID: [31110365](https://pubmed.ncbi.nlm.nih.gov/31110365/)
48. Rohwer F, Edwards R. The Phage Proteomic Tree: a genome-based taxonomy for phage. *Journal of bacteriology*. 2002 Aug 15; 184(16):4529–35. <https://doi.org/10.1128/jb.184.16.4529-4535.2002> PMID: [12142423](https://pubmed.ncbi.nlm.nih.gov/12142423/)
49. Adriaenssens EM, Edwards R, Nash JH, Mahadevan P, Seto D, Ackermann HW, et al. Integration of genomic and proteomic analyses in the classification of the *Siphoviridae* family. *Virology*. 2015 Mar 1; 477:144–54. <https://doi.org/10.1016/j.virol.2014.10.016> PMID: [25466308](https://pubmed.ncbi.nlm.nih.gov/25466308/)
50. Lavigne R, Darius P, Summer EJ, Seto D, Mahadevan P, Nilsson AS, et al. Classification of *Myoviridae* bacteriophages using protein sequence similarity. *BMC microbiology*. 2009 Dec 1; 9(1):224. <https://doi.org/10.1186/1471-2180-9-224> PMID: [19857251](https://pubmed.ncbi.nlm.nih.gov/19857251/)
51. Lavigne R, Seto D, Mahadevan P, Ackermann HW, Kropinski AM. Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTP-based tools. *Research in microbiology*. 2008 Jun 1; 159(5):406–14. <https://doi.org/10.1016/j.resmic.2008.03.005> PMID: [18555669](https://pubmed.ncbi.nlm.nih.gov/18555669/)
52. Yang H, Wang DB, Dong Q, Zhang Z, Cui Z, Deng J, et al. Existence of separate domains in lysin PlyG for recognizing *Bacillus anthracis* spores and vegetative cells. *Antimicrobial Agents and Chemotherapy*. 2012 Oct 1; 56(10):5031–9. <https://doi.org/10.1128/AAC.00891-12> PMID: [22802245](https://pubmed.ncbi.nlm.nih.gov/22802245/)
53. Yang H, Luo D, Etobayeva I, Li X, Gong Y, Wang S, et al. Linker editing of pneumococcal lysin ClyJ conveys improved bactericidal activity. *Antimicrobial Agents and Chemotherapy*. 2020 Jan 27; 64(2). <https://doi.org/10.1128/AAC.01610-19> PMID: [31767724](https://pubmed.ncbi.nlm.nih.gov/31767724/)