

# Molecular Characterization of Podoviral Bacteriophages Virulent for *Clostridium perfringens* and Their Comparison with Members of the *Picovirinae*

Nikolay V. Volozhantsev<sup>1\*</sup>, Brian B. Oakley<sup>2</sup>, Cesar A. Morales<sup>2</sup>, Vladimir V. Verevkin<sup>1</sup>, Vasily A. Bannov<sup>1</sup>, Valentina M. Krasilnikova<sup>1</sup>, Anastasia V. Popova<sup>1</sup>, Eugeni L. Zhilenkov<sup>1</sup>, Johnna K. Garrish<sup>2</sup>, Kathleen M. Schegg<sup>3</sup>, Rebekah Woolsey<sup>3</sup>, David R. Quilici<sup>3</sup>, J. Eric Line<sup>2</sup>, Kelli L. Hiatt<sup>2</sup>, Gregory R. Siragusa<sup>4</sup>, Edward A. Svetoch<sup>1</sup>, Bruce S. Seal<sup>2\*</sup>

**1** State Research Center for Applied Microbiology and Biotechnology, Obolensk, Moscow region, Russian Federation, **2** Poultry Microbiology Safety Research Unit, Richard B. Russell Agricultural Research Center, Agricultural Research Service, USDA, Athens, Georgia, United States of America, **3** Nevada Proteomics Center, University of Nevada, Reno, Nevada, United States of America, **4** Danisco/DuPont, Waukesha, Wisconsin, United States of America

## Abstract

*Clostridium perfringens* is a Gram-positive, spore-forming anaerobic bacterium responsible for human food-borne disease as well as non-food-borne human, animal and poultry diseases. Because bacteriophages or their gene products could be applied to control bacterial diseases in a species-specific manner, they are potential important alternatives to antibiotics. Consequently, poultry intestinal material, soil, sewage and poultry processing drainage water were screened for virulent bacteriophages that lysed *C. perfringens*. Two bacteriophages, designated  $\Phi$ CPV4 and  $\Phi$ ZP2, were isolated in the Moscow Region of the Russian Federation while another closely related virus, named  $\Phi$ CP7R, was isolated in the southeastern USA. The viruses were identified as members of the order *Caudovirales* in the family *Podoviridae* with short, non-contractile tails of the C1 morphotype. The genomes of the three bacteriophages were 17.972, 18.078 and 18.397 kbp respectively; encoding twenty-six to twenty-eight ORF's with inverted terminal repeats and an average GC content of 34.6%. Structural proteins identified by mass spectrometry in the purified  $\Phi$ CP7R virion included a pre-neck/appendage with putative lyase activity, major head, tail, connector/upper collar, lower collar and a structural protein with putative lysozyme-peptidase activity. All three podoviral bacteriophage genomes encoded a predicted N-acetylmuramoyl-L-alanine amidase and a putative stage V sporulation protein. Each putative amidase contained a predicted bacterial SH3 domain at the C-terminal end of the protein, presumably involved with binding the *C. perfringens* cell wall. The predicted DNA polymerase type B protein sequences were closely related to other members of the *Podoviridae* including *Bacillus* phage  $\Phi$ 29. Whole-genome comparisons supported this relationship, but also indicated that the Russian and USA viruses may be unique members of the sub-family *Picovirinae*.

**Citation:** Volozhantsev NV, Oakley BB, Morales CA, Verevkin VV, Bannov VA, et al. (2012) Molecular Characterization of Podoviral Bacteriophages Virulent for *Clostridium perfringens* and Their Comparison with Members of the *Picovirinae*. PLoS ONE 7(5): e38283. doi:10.1371/journal.pone.0038283

**Editor:** Earl G. Brown, University of Ottawa, Canada

**Received:** March 8, 2012; **Accepted:** May 2, 2012; **Published:** May 29, 2012

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Funding:** Support for the research was provided by the U.S. Department of Agriculture, Agricultural Research Service (CRIS project #6612-32000-060), the Russian Federation State Research Center for Applied Microbiology and Biotechnology, and the U.S. Department of State (DoS) via the International Science and Technology Center (ISTC grant #3108) administered by the ARS, USDA Office of International Research Programs (OIRP). The Nevada Proteomics Center operates under the auspices of NIH Grant #P20 RR-016464 from the INBRE Program of the National Center for Research Resources with instrumentation acquired through NIH grant #1S10 RR023587-01A2 awarded to the University of Nevada, Reno. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have read the journal's policy and have no conflicts: Greg Siragusa was an ARS-USDA employee when he isolated the bacteriophages and subsequently left USDA for employment with Danisco. There are currently no products in development with Danisco regarding this project. Consequently, this does not alter the authors' adherence to all PLoS ONE policies on sharing data and materials.

\* E-mail: nikvol@obolensk.org (NV); bruce.seal@ars.usda.gov (BS)

## Introduction

*Clostridium perfringens*, a Gram-positive, spore forming, anaerobic bacterium commonly present in the intestines of humans and animals, is classified into one of five types (A, B, C, D, or E) based on toxin production or presence of the toxin-encoding genes [1,2]. Human clinical symptoms of infection and histolytic pathogenesis are closely associated with the *C. perfringens* enterotoxin (CPE) produced by type A strains. Human food poisoning and gastroenteritis occurs if sufficient numbers of vegetative CPE-positive *C. perfringens* cells are ingested from contaminated food.

The vegetative cells may pass from the stomach to the intestinal tract where during sporulation, CPE is released into the intestinal lumen [3–5]. CPE-positive type A *C. perfringens* has been implicated in antibiotic-associated and sporadic diarrhea in humans that may also be food-related [6]. In addition to the association of CPE-positive type A strains with food-borne disease in humans, *C. perfringens* type A strains that produce the alpha toxin can cause necrotic enteritis and the subclinical form of infection in poultry [7].

Increased worldwide concerns over antimicrobial resistance (AMR) of zoonotic bacteria potentially circulating among food-

producing animals, including poultry [8], has resulted in a heightened public and scientific realization that antibiotic use by humans and in food animals selects for development of AMR among food-borne bacteria [9]. Because of these concerns, sub-therapeutic use of antibiotics as growth promoters was discontinued in the European Union [10] and it is hypothesized that human food-borne and poultry illnesses associated with the bacterium *C. perfringens* may increase as a consequence [7]. Historically, bacteriophages have been used to combat bacterial infections and, recently, there is renewed interest in utilizing live-phages or phage gene products as promising alternative antimicrobials to widely used antibiotics, including the control of food-borne pathogens [11–14]. In 2006, the U.S. Food and Drug Administration approved its first mixture of bacteriophages as food additives for use in processing plant spray applications onto ready-to-eat meat and poultry products to protect consumers from *Listeria monocytogenes* [15]. Zimmer et al. [16] isolated two temperate phages ( $\Phi$ 3626 and  $\Phi$ 8533) from lysogenic *C. perfringens* cultures and subsequently expressed a phage-specific enzyme identified as a murein hydrolase [17]. Another endolysin, a muramidase, was cloned and expressed from the episomal  $\Phi$ SM101 genome [18].

Given the potential use of lytic bacteriophages and/or their lytic enzymes for medical, veterinary and bio-industrial applications, our laboratories have been screening poultry intestinal material, soil, sewage and poultry processing drainage water for obligately lytic bacteriophages from *C. perfringens* [19]. Bacteriophage genomes of viruses previously isolated from broiler chicken offal washes (O) and poultry feces (F) in the USA, designated  $\Phi$ CP39O and  $\Phi$ CP26F, respectively, produced clear plaques on host strains and were classified as members of the family *Siphoviridae* in the order *Caudovirales* [20]. Bacteriophages lytic for *C. perfringens* classified in the family *Podoviridae* were isolated from broiler intestinal contents in the Russian Federation, one designated  $\Phi$ CPV1 [21], and another from raw sewage in the USA,  $\Phi$ CP24R [22]. Herein we report the molecular characterization of three unique, closely related Russian and US podoviruses that infect *C. perfringens* and present phylogenetic comparisons with members of the sub-family *Picovirinae*.

## Results and Discussion

### Isolation of bacteriophages virulent for *C. perfringens* and virion morphology

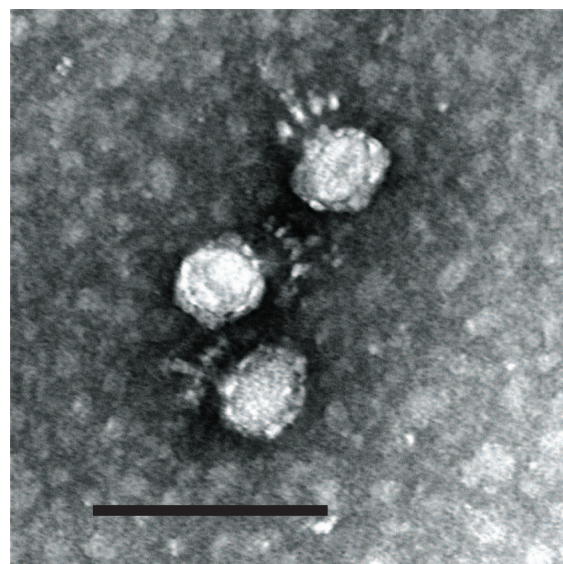
Bacteriophages were isolated by screening for lysis of forty-six *C. perfringens* strains [23] utilizing filter sterilized samples obtained from raw sewage or poultry (intestinal material and feces), soil and processing drainage water [20–22]. Bacterial viruses capable of lysing strains of the bacterium were identified by spot-testing the filter sterilized concentrated samples and titration of bacteriophages on susceptible *C. perfringens* strains. Individual bacteriophages from the USA were designated by their host strain (i.e.,  $\Phi$ CP7R) while Russian bacteriophages were named numerically (i.e.,  $\Phi$ CpV4 or  $\Phi$ ZP2). Each bacteriophage isolate only replicated in its respective host strain producing clear, 3 mm plaques with no evidence of background bacterial growth (Fig. S1). No other clostridial species assayed for lysis supported replication of these bacteriophages that were routinely obligately lytic for *C. perfringens* with no evidence of lysogeny. The titers of each bacteriophage lysate were approximately  $2 \times 10^8$  PFU/ml. Lytic bacteriophage preparations were initially characterized morphologically by electron microscopy (Fig. 1). Two short-tailed bacteriophages isolated from in the Moscow Region of the Russian Federation were designated  $\Phi$ CPV4 (from poultry waste) and  $\Phi$ ZP2 (from poultry feces), while one virus from the southeastern U.S., isolated

from raw sewage, was named  $\Phi$ CP7R. Each of the bacteriophages were morphologically equivalent with head diameters of approximately 40–42 nm, tail lengths of 35–38 nm, a basal plate with diameters of 30–32 nm that had short protrusions, and collars with a diameter of 27–28 nm. The advanced tail rods did not show thickening in the distal portion. Structurally, all three viruses can be considered members of the order *Caudovirales* in the family *Podoviridae*, characterized by short, non-contractile tails of the C1 morphotype [24].

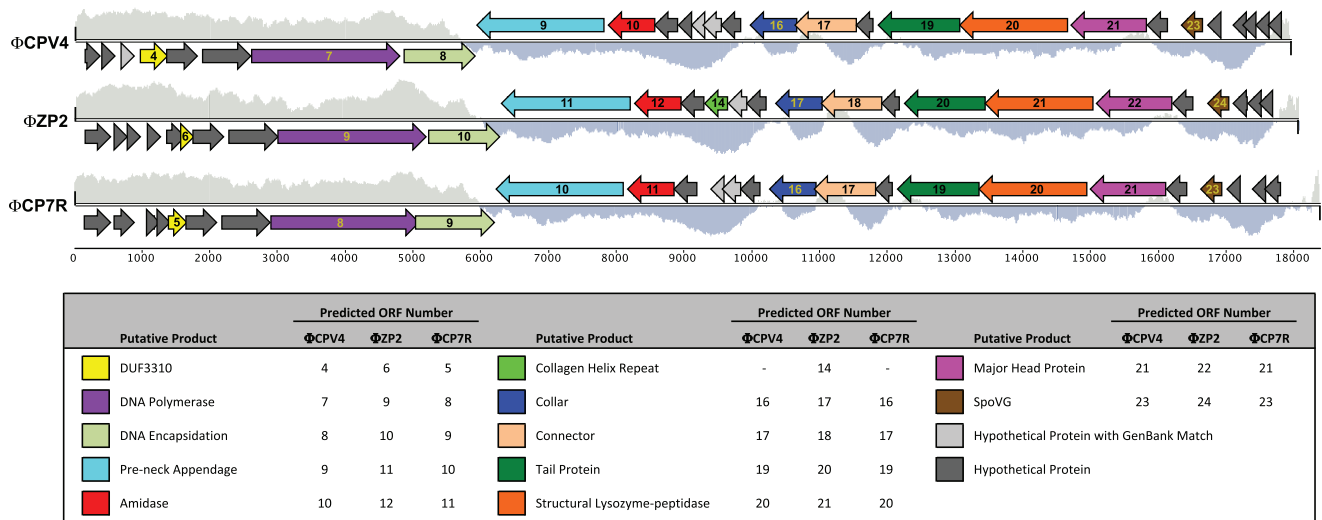
### Bacteriophage Genome Structures and Predicted Non-structural Protein ORF's

The genomes of  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R were 17.972, 18.078 and 18.397 kbp, respectively, with an average GC content of 34.6%. Genome sequence identity was 88% for  $\Phi$ ZP2 as compared to genome sequences of  $\Phi$ CPV4 or  $\Phi$ CP7R. Sequence identity was 95% between the  $\Phi$ CPV4 and  $\Phi$ CP7R genomes even though the bacteriophages were isolated from two disparate geographical regions, Russia and the USA. There were 26, 27 and 28 open reading frames (ORF) identified in the genomes of  $\Phi$ CP7R,  $\Phi$ ZP2 and  $\Phi$ CPV4, respectively (Fig. 2; Tables S1,S2,S3). In addition to the previously described morphological characteristics, the identified ORFs revealed protein similarity to previously described podoviral proteins [21,22] also suggests that the clostridial phages reported herein belong to the *Podoviridae* family. Following alignment of the genomes in Mauve, the principle regions of genomic dissimilarity were in the inverted terminal repeat (ITR) regions at both ends (Fig. S2). The ITRs of the  $\Phi$ CPV4 genome were 28 nucleotide pairs in length, while the ITRs of the  $\Phi$ ZP2 genome consisted of 30 nucleotide pairs and the ITRs of  $\Phi$ CP7R were 25 nucleotide pairs (Fig. 3), which are much longer than the ITRs of the  $\Phi$ 29 genome [25]. The small genome sizes and presence of ITRs are characteristics belonging to the subfamily *Picovirinae* [26]. PHACTS analysis [27] predicted that the isolated phages were virulent, which supports our culture observations.

The predicted proteins (Fig. 2; Tables S1,S2,S3) of the *C. perfringens* bacteriophages had demonstrated homology to other



**Figure 1. Electron Micrograph of Bacteriophage  $\Phi$ CP7R.** Scale bar represents 100 nm. doi:10.1371/journal.pone.0038283.g001



**Figure 2. Comparative Genome Maps for Podoviral Bacteriophages Isolated from the Russian Federation and the USA.** The %GC plot displays regions (500-bp window) above and below the average GC content. Open reading frames (ORFs) are depicted as arrows in the predicted direction of transcription and identified putative protein domains are listed in the legend with their respective ORF designation. doi:10.1371/journal.pone.0038283.g002

known bacteriophage proteins including their DNA polymerases. The predicted Type B polymerases of all three phages were 727 amino acids in length and ΦCPV4 and ΦCP7R were more closely related (99% sequence similarity) than either one compared to the ΦZP2 (96% similar to ΦCPV4 and ΦCP7R). BLAST analyses revealed the three phage proteins were most closely related in sequence to the ΦCPV1 [GenBank ADR30478; 36%] and ΦCP24R [AEW47837; 34%] polymerases with overall similarity of 24% to other DNA polymerases of the *Bacillus* Φ29-family [25,28]. After the ΦCP24R and ΦCPV1 polymerases, the most

#### ΦCPV4 Inverted Terminal Repeats:

```

1 TTTTGTATATAGGAACATTTTAACGACA      28
  |||
17972 AAAACATATATCCTTGAAAATTGCTGT 17945

```

#### ΦZP2 Inverted Terminal Repeats:

```

1 TTTTGTATATAGGAACATTTTAACGACATA     30
  |||
18078 AAAACATATATCCTTGAAAATTGCTGTAT 18049

```

#### ΦCP7R Inverted Terminal Repeats:

```

1 ATAGGAACATTTTAACGACACACTC         25
  |||
18397 TATCCTTGAAAATTGCTGTGTGAG 18373

```

**Figure 3. Inverted Terminal Repeat Genomic Nucleotide Sequences for Bacteriophages ΦCPV4, ΦZP2 and ΦCP7R.** Inverted terminal repeat (ITR) sequences were identified by whole-genome analysis utilizing the EMBOS 6.3.1: palindrome program. doi:10.1371/journal.pone.0038283.g003

closely related polymerases were from the *Bacillus* phages M2 [AAA32368], Nf [ACH57069], GA-1 [CAA65712], B103 [CAA67649] and the *Actinomyces* phage Av-1 [ABR67671]. The predicted DNA polymerase proteins contained a consensus sequence of Dx<sub>2</sub>SSYP rather than the Dx<sub>2</sub>SLYP with the intervening x<sub>2</sub> being IN rather than VN as found in the Φ29 polymerase motif A. Also, the two Asp (D) residues that participate in metal binding required for catalysis were located in the conserved motif C sequence identical to the Φ29 polymerase [25,29]. The Φ29-like viruses' genes encoding the terminal protein (TP) have been reported as adjacent to the *pol* gene [25]. Although no putative TP was identified by BLAST analysis of our podoviral genomes, the gene adjacent to the *pol* gene in our podoviruses encoded a protein of a predicted similar structure and physical characteristics to the Φ29 TP. However, a conserved threonine (T) was present in the putative terminal proteins encoded by our phage genomes rather than the serine (S232) in the Φ29 TP [28]. The existence of a terminal protein was further supported by the detection of a protein at the genomic terminal fragments by alteration of their relative mobility following protease and restriction enzyme digestion of ΦCPV4 DNA (Fig. S3).

The ΦCPV4, ΦZP2 and ΦCP7R genomes examined by our laboratories each encoded a putative podovirus DNA encapsidation protein (Fig. 2 and Tables S1,S2,S3). This protein binds to packaging RNA (pRNA) and catalyzes the *in vivo* and *in vitro* genome-encapsidation reaction [30]. These genes encoded a predicted protein of 347 amino acids that was most similar by BLAST to the encapsidation proteins of *Bacillus* phage PZA [AAA88493], *Bacillus* phage Nf [ACH57084] and *Bacillus* phage Φ29 [AAA88348]. This protein functions as a pRNA dependent ATPase providing the energy for encapsidation of the viral DNA into the mature capsid [31,32]. Two ATP binding motifs occurred in the Φ29 encapsidation protein [33] that were most likely located at residues 34 through 40 for the A motif and residues 267 through 273 for the B motif among the podoviruses reported herein. These motifs were also reported in the encapsidation proteins encoded by the genomes of *Bacillus* phages B103 and GA-1 [25]. Interestingly, the podoviral encapsidation proteins reported during these investigations were more distantly related to DNA encapsidation

proteins of our previously reported clostridial  $\Phi$ CP24R [AEW47838] and  $\Phi$ CPV1 [ADR30484] relative to the *Bacillus*  $\Phi$ 29 encapsidation protein [YP\_002004545].

Other non-structural proteins encoded by the podoviral genomes included a potential stage V sporulation protein G (SpoVG; pfam 04026) that is reportedly essential for sporulation and specific to stage V sporulation in *Bacillus megaterium* and *B. subtilis* [34]. This protein is also involved in the regulation of septum location during cell envelope and outer membrane biogenesis [35]. The SpoVG-like protein encoded by our podoviral genomes were most similar to the regulatory protein SpoVG reported in *Borrelia hermsii* [YP\_001884203], *Borrelia turicatae* [YP\_945770] and *Stigmatella aurantiaca* [YP\_003955432], but were not detected as encoded by the genomes of our previously reported clostridial bacteriophages [19–21]. The only viral protein with any sequence similarity was a putative minor capsid protein 4 of the temperate *Streptococcus* phage MM1 [NP\_150171]. The *C. perfringens* encoded SpoVG [NP\_563407] is composed of only 90 residues compared to our putative podoviral SpoVG proteins of 103 residues and only the first seven N-terminal amino acids were similar between the two proteins. Although the N-terminal sequences were conserved, the functionally essential Proline-63 residue conserved among the SpoVG proteins [35] from *B. subtilis* [CAA44242], *B. hermsii* [YP\_001884203] and *C. perfringens* [NP\_563407] was missing in the podoviral proteins (F for P in the phage proteins) and it is unknown what role these viral proteins may play in *C. perfringens* biology or during a viral infection.

The three bacteriophages reported encoded a predicted N-acetylmuramoyl-L-alanine amidase (pfam01510), which includes zinc amidases (EC: 3.5.1.28) that cleave the amide bond between N-acetylmuramoyl and L-amino acids in bacterial cell walls (preferentially: D-lactyl-L-Ala). The amidase gene was preceded in the bacteriophage genomes by a gene encoding a predicted pre-neck appendage structural protein (Fig. 2; Tables S1,S2,S3). Similar to the siphoviral *C. perfringens* bacteriophages previously reported from our laboratories, the holin gene is most likely downstream of the lysin gene. This placement is unique to the other clostridial bacteriophages [19], including the podoviruses we recently reported [21,22]. BLAST analysis of the predicted amidases from  $\Phi$ ZP2,  $\Phi$ CPV4 and  $\Phi$ CP7R revealed similarity to a prophage LambdaCh01-like amidase encoded by *C. botulinum* isolates [YP\_002863851; YP\_001782467; YP\_001392135], *C. sporogenes* [ZP\_02994334] and *C. tetani* [NP\_783826]. The only closely related bacteriophage proteins were to a hypothetical protein BA3\_0024 from a *Thalassomonas* phage BA3 [YP\_001552293], a PlyM32 from an uncultured bacteriophage [ADF97557], an EBPR siphovirus 4 [AEI71112] and several putative lysozymes from *Vibrio* phages [ADX87661; YP\_249586; YP\_003347926; ADX87518].

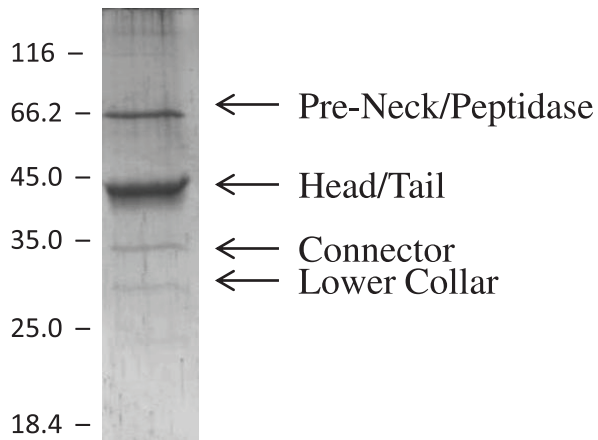
Residues 15 through 130 of the predicted podoviral amidase showed similarity to peptidoglycan recognition proteins (PGRPs), which are pattern recognition receptors that bind and can hydrolyze peptidoglycan (PGNs) of bacterial cell walls [36]. This portion of the protein also contains the substrate binding and putative enzymatically active site, as well as the potential Zn-binding residues. Residues 160 through 225 represented the bacterial Src homology 3 (SH3) domain (pfam08239; superfamily cl02551) that bind to target proteins through sequences containing proline and hydrophobic amino acids. The cell wall targeting-SH3 domains are associated with other hydrolases such as the phage associated cysteine, histidine-dependent amidohydrolase/peptidases (CHAPs), but does not contain the N-terminal 'F-[IV]-R' motif common to staphylococcal bacteriophage CHAPs [37,38]. Interestingly, the C-terminal SH3 portion of the bacteriophage lysins

aligned most closely with bacterial CHAP domains, specifically a glycosyl hydrolase, family 25-protein from *C. perfringens* [e.g., ZP\_02642156 and seventeen other related sequences]. Since the gene upstream of the amidase encoded a predicted structural protein, synteny with similar bacteriophage genomes suggested the gene downstream from the amidase presumably encoded a holin [19]. However, the small 111 residue gene product does not have a potential transmembrane domain and is predicted to be hydrophilic. The only predicted peptide encoded by the genomes with characteristics of a holin [39,40], with a single transmembrane domain, is a gene located between the ORF's encoding the phage connector or upper collar protein and a putative tail protein (Fig. 2; Tables S1,S2,S3). The peptide was predicted to have a single transmembrane domain from residues 10 through 32 and also as a potential signal peptide.

### Virion structural proteins

The bacteriophage structural proteins were identified by purifying virions by isopycnic gradient centrifugation through CsCl followed by SDS-PAGE and mass spectrometric analyses of the individual proteins (Fig. 4, Table 1 and Table S4). The principle structural proteins identified included the pre-neck/appendage protein, a putative bacteriophage structural protein, the predicted head protein, a tail protein, the upper collar also known as the connector protein, and a lower collar protein (Fig. 4, Table 1). The virion head of  $\Phi$ 29 consists of a major capsid or head protein that is attached to the tail by a connector or upper collar protein [41]. The head proteins from our podoviral phages were predicted to be 370 residues with a size of 42 kDa, similar to podoviral proteins encoded by viruses previously reported from our laboratories,  $\Phi$ CP24R [AEW47842] and  $\Phi$ CPV1 [ADR30492], that belong to a major bacteriophage head protein superfamily [PHA00144]. The proteins were also similar to the head protein of *Actinomyces naeslundii* phage Av-1 [YP\_001333662] and *Bacillus* phage  $\Phi$ 29 [YP\_002004536]. The connector (portal) or upper collar protein was predicted to be 301 amino acids in length with a size of 34.7 kDa that demonstrated 97% similarity among the three phages. This protein [pfam05352] is the central component of a rotary motor that packages bacteriophage genomic dsDNA into pre-formed proheads [42]. The connector proteins of  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R were most similar to those of our previously reported  $\Phi$ CP24R [AEW47848] and  $\Phi$ CPV1 [ADR30486] proteins and comparable to the respective proteins of  $\Phi$ 29-like viruses *Bacillus* phages Nf [ACH57078], B103 [NP\_690644] and  $\Phi$ 29 [YP\_002004539].

The largest tail protein was a predicted pre-neck/appendage protein that consisted of 624 amino acids for  $\Phi$ CP7R and  $\Phi$ CPV4. The predicted pre-neck/appendage for  $\Phi$ CPZP2 was a 631 residue protein with a relative mobility at approximately 70 kDa and most similar to proteins reported in previously discovered bacteriophages from the USA [AEW47853] and Russia [ADR30483]. The predicted protein contained a putative pectate lyase domain (residues 140 through 350) and these domains are found among bacterial endopygalactorunases involved with cell envelope and outer membrane biogenesis [pfam12708 superfamily protein]. The proteins are encoded in the genomes of *Bacillus*  $\Phi$ 29-like phages [AAA32285] and are appendages attached to the virion neck region [43,44]. In  $\Phi$ 29 this protein was reported to recognize glucosylated poly-teichoic acids and, thus, is presumably involved with host attachment, cell wall digestion and entry into the cell [44,45]. The lower collar proteins encoded by our three bacteriophages were 99% similar, consisting of 229 amino acids with a predicted size of 26.7 kDa and belong to the cl10184 superfamily. Again, these proteins from the bacteriophages



**Figure 4. Polyacrylamide Gel Electrophoresis of Virion Proteins from Purified Bacteriophage  $\Phi$ CP7R.** Bacteriophages were purified by isopycnic centrifugation through CsCl followed by electrophoresis in a 10% SDS-PAGE. Size markers are on the left and regions of proteins identified mass spectrometry are on the right side of the figure. doi:10.1371/journal.pone.0038283.g004

reported herein were most similar to our previously reported phages  $\Phi$ CP24R [AEW47853] and  $\Phi$ CPV1 [ADR30483], but are also related to the lower collar protein of *Bacillus* phage  $\Phi$ 29 [YP\_002004540].

The lower portion of the podovirus  $\Phi$ 29 tail below the collar protein consists of two proteins [46]. One tail protein encoded by the three bacteriophages reported was predicted to be 399 amino acids in length (46.2 kDa) with homology to the gp9 of the  $\Phi$ 29-like phages [PHA00380 tail protein; superfamily cl15539]. These proteins were most similar to our previously reported tail proteins of  $\Phi$ CP24R [AEW47845] and  $\Phi$ CPV1 [ADR30489] which were much smaller in size relative to homologous tail proteins such as those encoded by *Bacillus* phage GA-1 [NP\_073692; 612 aa], *Bacillus* phage B103 [NP\_690643; 598 aa] and *Bacillus* phage  $\Phi$ 29 [YP\_002004538; 599 aa]. Another structural protein with a predicted size of 529 amino acids (59.2 kDa) was revealed by BLAST analysis to be a multiple-domain protein consisting of a bacteriophage tail domain [PHA00380; PHA00965] (residues 1 through 150), a lysozyme domain [COG3772] (residues 200 through 350), and a peptidase M23 domain [pfam01551] (residues 400 through 500). This tail component was similar by BLAST analysis to proteins reported as a phage structural protein [YP\_696017] encoded in the genomes of *C. perfringens* isolates. The predicted protein was also similar to proteins reported in our  $\Phi$ CP24R [AEW47844] and  $\Phi$ CPV1 [ADR30490] bacteriophag-

es, as well as the Gp15 protein of *C. perfringens* phage  $\Phi$ 3626 [NP\_612844]. The homologous  $\Phi$ 29 gp13 protein is essential for assembly of the  $\Phi$ 29 virion [47] and has two domains related to lysozymes and metallo-endopeptidases [48]. By analogy the 59.2 kDa protein detected in our podoviruses may function similar to the  $\Phi$ 29 gp13 tail-associated, peptidoglycan-degrading enzyme essential for infection.

#### Phylogenetics of the Podoviral Bacteriophages from Russia and USA

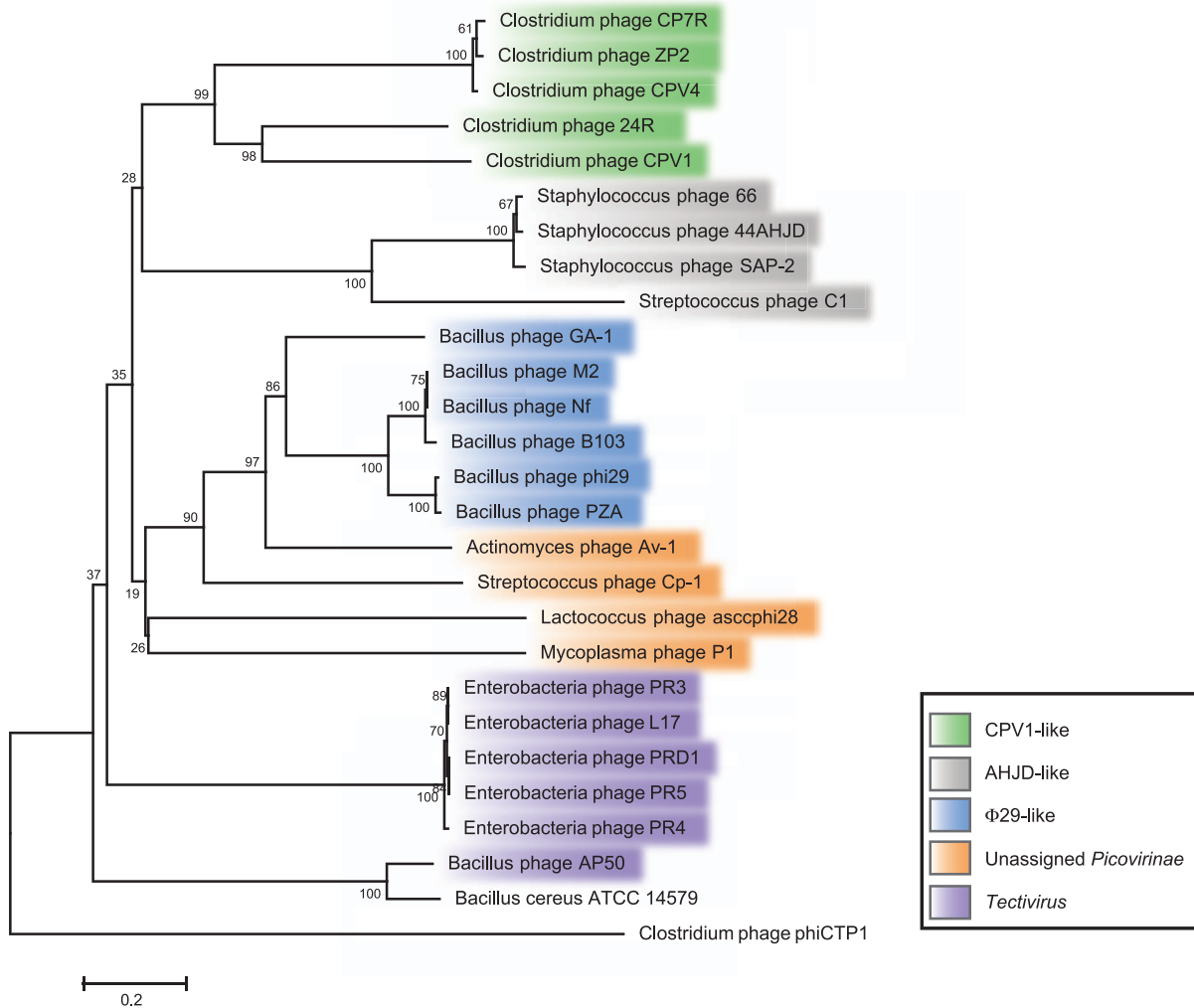
Based on the DNA polymerase predicted amino acid sequences,  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R were most closely related to our previously reported *C. perfringens* phages  $\Phi$ CP24R [AEW47837] and  $\Phi$ CPV1 [ADR30478] (Figure 5). Beyond the clostridial *Podoviridae*, the other similar bacteriophage polymerases included *Streptococcus* phage C1 [NP\_852013], as well as the *Staphylococcus* phages 66 [YP\_239478], 44AHJD [NP\_817305] and SAP-2 [YP\_001491534]. The *Lactococcus* phage accphi28 [YP\_001687520] and *Mycoplasma* phage P1 [NP\_064636] separated as a cluster with other closely related DNA polymerases of *Streptococcus* phage Cp-1 [NP\_044817] and *Actinomyces* phage Av-1 [YP\_001333659]. These were most similar to the *Bacillus* phages GA-1 [NP\_073685], M2 [AAA32368], Nf [ACH57069], B103 [NP\_690635], PZA [AAA88478] and  $\Phi$ 29 [YP\_002004529], representing pfam03175type-B DNA polymerases [25]. All of these bacteriophages are considered members of the podoviral sub-family *Picovirinae* and contain type-B DNA polymerases requiring a terminal protein to prime DNA synthesis [49–53]. The C-terminal portion, beyond residue 505, of our podoviral predicted polymerase proteins had sequence similarities to members of the *Tectiviridae* such as the enterobacterial phages PRD1 [NP\_040682], PR3 [AAX45563], PR4 [AAX45594], PR5 [AAX45625] and L17 [AAX45532] that infect Gram-negative bacterial hosts. These bacteriophages have a lipid membrane underneath the capsid and a protein covalently linked to the 5'-termini of their genomes for priming DNA replication [54]. Other polymerases more distantly related to those reported herein belonged to the *Bacillus cereus* ATCC 14579 [NP\_829893] and *Bacillus* phage AP50 [YP\_002302517], tentatively also considered members of the *Tectiviridae* [55].

Whole-genome analyses of publicly-available picovirus and tectivirus sequences demonstrated that the three phages sequenced here,  $\Phi$ CPV4,  $\Phi$ CP7R, and  $\Phi$ ZP2, were most closely related to each other and their closest relatives were two podoviruses ( $\Phi$ CPV1 and  $\Phi$ CP24R) previously sequenced by our laboratories (Fig. 6). The use of tetranucleotide frequencies as a metric of whole-genome relatedness has proven to be a valuable tool for comparative genomics [56,57] and provides a useful alternative to

**Table 1.  $\phi$ CP7R Virion Proteomics Data Summary as Determined by Mass Spectrometry Analyses.**

Protein name	Predicted size (kDa)	Number of unique peptides	Percent coverage
Pre-neck appendage	69.7	10	20.80%
Lysozyme-peptidase	59.2	4	11.30%
Tail protein	46.3	2	6.50%
Major head protein	41.8	18	70.50%
Connector Protein	34.7	6	23.30%
Collar Protein	26.7	15	78.60%

doi:10.1371/journal.pone.0038283.t001



**Figure 5. Phylogenetic Relationships of  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R among other related Bacteriophages based on the DNA Polymerase Protein.** Phylogenetic analysis of polymerase amino acid sequences were completed in MEGA5 utilizing MUSCLE for alignments with two thousand bootstrap replications and the *Clostridium* phage  $\Phi$ CTP1 polymerase outgroup. doi:10.1371/journal.pone.0038283.g005

phylogenetic reconstructions based on multiple sequence alignments (MSA). As applied here, the tetranucleotide approach provided an MSA-independent confirmation of the high similarities among  $\Phi$ CPV4,  $\Phi$ CP7R, and  $\Phi$ ZP2 with pairwise correlation coefficients of 0.98 ( $\Phi$ CPV4 vs.  $\Phi$ CP7R), 0.92 ( $\Phi$ ZP2 vs.  $\Phi$ CP7R), and 0.94 ( $\Phi$ ZP2 vs.  $\Phi$ CPV4; Fig. 6). Tetranucleotide comparisons were also consistent with our phylogenetic analyses of the DNA polymerases (Fig. 5), confirming the close relationships among *Staphylococcus* picovirinae, and also the *Bacillus* phages B103, GA-1, and  $\Phi$ 29. The five genomes sequenced by our laboratories belonged to a larger clade containing the *Bacillus* phages, including  $\Phi$ 29 (Fig. 7). The whole-genome phylogeny was slightly different than the polymerase phylogeny in which our phages were most closely related to *Staphylococcus* phages, although this was not strongly supported by bootstrap analysis (Fig. 5). Interestingly, queries of the genomes shown in Fig. 7 for putative paralogs showed only three potential gene duplications; the head proteins of  $\Phi$ CP24R (ORFs 10 and 11) which had been previously noted [22], hypothetical proteins of  $\Phi$ ZP2 (ORFs 13 and 14) and putative DNA polymerases in  $\Phi$ CPV1 (ORFs 5 and 6).

As a standard reference genome with important commercial applications [58], we were interested to determine proteins shared by  $\Phi$ 29 and the ‘pan-genome’ of  $\Phi$ CP7R,  $\Phi$ CP24R,  $\Phi$ CPV1,  $\Phi$ CPV4 and  $\Phi$ ZP2 (Fig. 8). Genes encoding for eight major bacteriophage proteins were shared by these six genomes that consisted of a DNA encapsidation protein, DNA polymerase, connector protein, lower collar protein, major head protein, peptidoglycan hydrolase, preneck appendage protein and a tail protein. These are all proteins encoded by the  $\Phi$ 29 family of lytic bacteriophages that have similar genome structures with a terminal protein and encode a Type B DNA polymerase typical of the *Picovirinae* [25].

### Concluding Remarks

Bacteriophages constitute the largest number of biological entities found on earth [59] and there is resurgent interest for utilizing either bacteriophages [60,61], or their gene products [12,14], to control bacterial infections that plague humans, animals and plants. Consequently, we have conducted collaborative research between the Russian Federation and USA to search for bacteriophages that clearly lyse the bacterium *Clostridium*



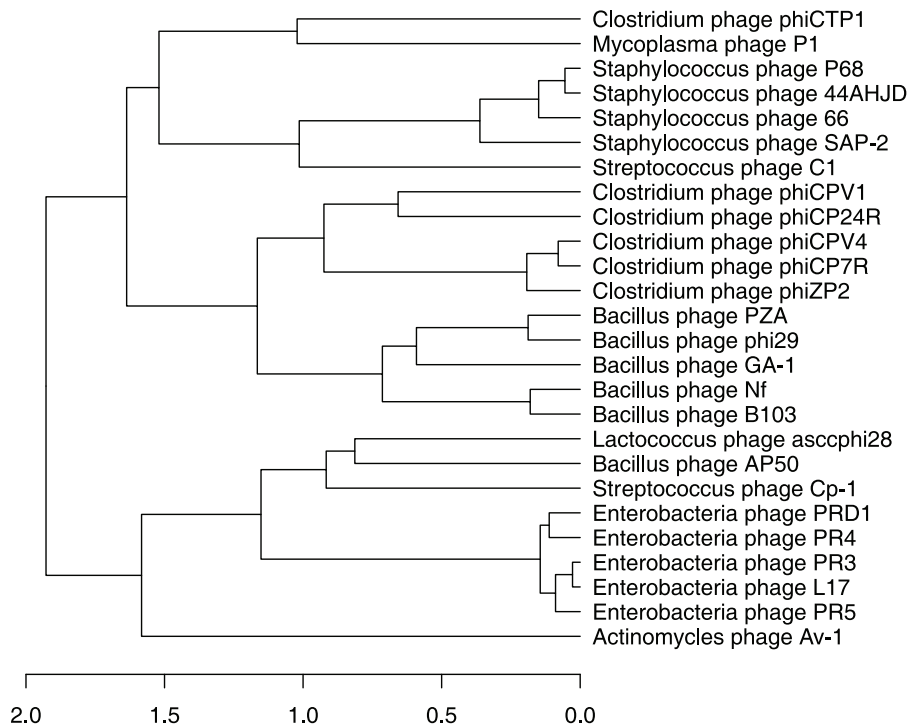
**Figure 6. Whole-genome Tetra-nucleotide Frequency Comparisons Relative to Bacteriophages  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R.** Lower panel shows normalized frequencies of 256 tetranucleotides for each genome, upper panel shows Pearson correlation coefficients. In lower panel, the most closely-related genomes are shown in blue and red. Genome comparisons with correlation coefficients  $>0.5$  are shown with blue points, r-squared values from simple linear regression  $>0.5$  are shown with red lines. Genome names are shown on diagonal axis. doi:10.1371/journal.pone.0038283.g006

*perfringens*. This pathogen causes a plethora of diseases in humans and animals that are both medically and agriculturally important [2]. Previously we reported on bacteriophages that were both members of the *Siphoviridae* [19,20] and *Podoviridae* [21,22] that met our criteria for producing clear plaques on the bacterium *C. perfringens*. We have now identified three more highly-related bacteriophages from disparate geographical regions virulent for the host *C. perfringens* and demonstrate their inclusion as unique members of the  $\Phi$ 29-like *Picovirinae*.

## Materials and Methods

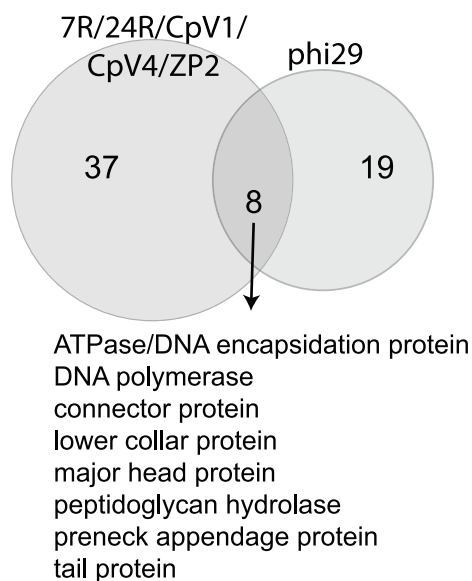
### Bacterial hosts, bacteriophage isolation and propagation

*C. perfringens* isolates utilized as hosts for propagation of bacteriophages were cultured in brain heart infusion (BHI) broth or on agar (Remel, Lenexa, KS) and characterized by standard methods including 16S rRNA-DNA sequence analysis [23,62]. Offal washes (O) and feces (F), obtained at chicken processing facilities, or raw sewage samples (R) were collected following verbal consent from poultry processing facility management or



**Figure 7. Whole-genome Similarities as Determined by Tetranucleotide Frequencies.** Dendrogram is based on a dissimilarity matrix for whole bacteriophage genomes constructed in R as described in the methods section.  
doi:10.1371/journal.pone.0038283.g007

personnel at the Athens Clarke County, GA, USA sewage treatment facility and under similar circumstances in the Russian Federation. Samples were clarified by low-speed centrifugation ( $5,000\times g$  for 20 min. at 5C) followed by filtration of the supernatant through cheesecloth, then by filtration through 0.45  $\mu\text{m}$  bottle filters (Corning Inc., Corning, NY). Bacteriophages



**Figure 8. Shared and Unique Genes present in Bacillus phage  $\Phi$ 29 versus Bacteriophages  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R.** Gene clustering was determined by uclust and blastp as described in the methods section.  
doi:10.1371/journal.pone.0038283.g008

from the USA were isolated on their respective host strain (Cp7), while Russian bacteriophages were isolated on *C. perfringens* strains ATCC3624 and strain 46 (provided by the Tarasevich State Institute of Standardization and Control of Biomedical Preparations, Russia). Bacterial viruses producing clear plaques on individual strains of *C. perfringens* were identified by spot-testing and titration on each strain-specific host [20,21,63]. Briefly, 10  $\mu\text{l}$  of filtrate was spotted onto lawns of *C. perfringens* and incubated for 16 hours, followed by visualization for areas of clearance indicating lysis of the bacterium. Several clostridial species including *C. absoum*, *C. acetobutylicum*, *C. beijerinckii*, *C. novyi*, *C. rubrum*, *C. sordelli*, *C. sporogenes*, *C. tetani*, and *C. tetanomorphum* [62] were spot-tested for lytic activity. The *C. perfringens*-specific bacteriophages were propagated by plating with low-melt agar using the *C. perfringens* host bacterium cultured at 37C in the Anaero Pack<sup>TM</sup> (Mitsubishi Gas Chemical Co., Japan) system with AnaeroGen (OXOID Ltd., Basingstoke, England) sachets [20,23]. The bacteriophage isolates were subjected to three rounds of plaque purification and suspended in TMGS (10 mM Tris, pH 8, 10 mM Mg<sup>++</sup>, 0.55% NaCl, 0.1% gelatin) at an average titer of  $2\times 10^8$  pfu/ml. Bacteriophages  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R were propagated utilizing a plate lysis method [64] under anaerobic conditions.

#### Purification of bacteriophages, genomic DNA purification and electron microscopy

Following plate lysis [64] in anaerobic chambers, bacteriophage genomic DNA was purified using the Qiagen<sup>TM</sup> Lambda Phage DNA isolation protocol. Additionally, bacteriophages were purified from plate lysates by centrifugation at  $2,000\times g$  for 20 minutes to remove bacterial debris and low-melt agarose. The clarified supernatant was centrifuged at  $103,800\times g$  for 90 minutes followed by suspension of the phage pellet in 1 ml TBS



(20 mM Tris, 500 mM NaCl at pH 7.5) followed by purification using cesium chloride equilibrium gradient centrifugation [65]. The bacteriophage bands were extracted from the gradient, diluted in TBS and concentrated by centrifugation at 105,000× g in a Beckman™ JS 24.15 rotor for 90 minutes followed by suspension in TBS. Purity of bacteriophages was examined by electron microscopy [66]. The bacteriophage pellets were also subjected to proteinase K (20 ug/ml) digestion in the presence of 0.1% sarcosyl and 0.2 M EDTA followed by phenol-chloroform extraction and ethanol precipitation to obtain genomic DNA [65].

### Molecular cloning, sequencing, annotation of genomic DNA and phylogenetic analyses

Following purification of bacteriophage genomic DNA, the nucleic acid was subjected to spectrophotometer readings at 260/280 nm and restriction enzyme digestion followed by agarose gel electrophoresis [65]. Sequencing of the bacteriophage genomes was completed by MWG Biotech, Inc High Point, NC USA and pyrosequencing [67]. For Sanger sequencing, phage DNA was sheared using a nebulizer, blunt-end repaired and dephosphorylated [65]. DNA fragments of desired size (1 to 4 kb) were ligated into pSmart (Lucigen™) for propagation in *E. coli* following transformation. Clones were sequenced such that approximately 14-fold redundancy was obtained for the genome that included primer-walking to fill gaps [68]. Molecular cloning was also completed using the restriction enzymes *HindIII*, *EcoRI*, *EcoRV*, *AclI* and *ClaI* (New England Biolabs, Ipswich, MA) to cleave phage DNA followed by treatment with *Taq* polymerase [69] and cloning [70] into the TOPO TA vector (Invitrogen™, Carlsbad, CA). Additionally, end-repair and G-tailing was completed for cloning restriction enzyme fragments into pSmart vectors (Lucigen™, Middleton, WI) for nucleotide sequencing. Double stranded nucleotide sequencing reactions using fluorescent labeled dideoxynucleotide terminators were also completed and sequences determined using an automated sequencer (Applied Biosystems Inc., Foster City, CA) [71].

Nucleotide sequence assembly, editing, analysis, prediction of amino acid sequences and alignments were conducted using the MacVector 7.2™ (Accelrys, San Diego, CA) and DNASTAR™ (Madison, WI) software. The European Molecular Biology Open Software Site was accessed to identify the inverted terminal repeats using EMBOSS 6.3.1: palindrome [72]. Protein-encoding genes (ORFs) were predicted using GeneMark.hmm for prokaryotes version 2.4 (<http://opal.biology.gatech.edu/GeneMark>) [73] and SoftBerry FGENE SB (<http://linux1.softberry.com/berry.phtml>; Mount Kisco, NY, USA) programs. Final genome sequences were also submitted to the IMG/ER pipeline for gene predictions and initial annotation [74]. The computational approach Phage Classification Tool Set (PHACTS) was developed to classify phages as to whether the lifestyle of a phage, described by its proteome, is virulent or temperate by using the known phage genomes in the PHANTOME database. PHACTS utilizes known genomes to find similarities in the unknown phage genome and was used to predict whether our bacteriophages were virulent or temperate [27]. Predicted ORFs were searched for similarity to proteins in databases by BLAST analyses [75] at the NCBI website (<http://www.ncbi.nlm.nih.gov>) as well as the conserved domain database [76] algorithms with NCBI accession numbers and domain designations reported in the results for similar proteins. Putative promoters were analyzed by using Martin Reese's neural network prediction program at [http://www.fruitfly.org/seq\\_tools/promoter.html](http://www.fruitfly.org/seq_tools/promoter.html) and BPROM (Softberry, Inc., Mount Kisco, NY, USA) at its website <http://linux1.softberry.com/berry.phtml>. Potential transcriptional terminators were assessed using

the software programs TransTerm at the Nano+Bio-Center (<http://nbc3.biologie.uni-kl.de>) and FindTerm (Softberry, Inc., Mount Kisco, NY, USA) at the web site <http://linux1.softberry.com/berry.phtml>. The amino acid sequences of the phage ORF gene products were analyzed for helical transmembrane domains by using the prediction program TMHMM version 2.0 [77] at the website <http://www.cbs.dtu.dk> and by Dense Alignment Surface method [78] at <http://www.sbc.su.se/~melen/TMHMMfix/> and <http://phobius.sbc.su.se/>.

Multiple genome alignments were generated by Mauve software to provide for comparative analyses of multiple bacteriophage genomes [79,80]. Phylogenetic analyses of polymerase amino acid sequences were completed in MEGA5 [81] utilizing MUSCLE for alignments with two thousand bootstrap replications [82] and an outgroup [83,84]. Tetra-nucleotide distributions for clostridial phage genomes and correlation coefficients between genomes were calculated with TETRA [56]. Correlation coefficients were transformed to a dissimilarity matrix for tree construction using the hierarchical clustering algorithm hclust in R [85], which was also used to generate dendrograms and visualize tetra-nucleotide distributions. Additional whole-genome comparisons were performed by reciprocal blastp with an e-value cutoff of 0.001 to identify core and accessory proteins between the clostridial phages sequenced here and several reference phage genomes. The number of genes shared among the five clostridial phages sequenced by our laboratories was determined by UCLUST [86] which identifies similarity clusters based on all-versus-all sequence comparisons. Blastp with an evalue cutoff of 0.001 was used to confirm genes shared between the clostridial phages sequenced here and the reference phage genome of  $\Phi$ 29.

### Preparation of purified virions for gel electrophoresis and identification of purified bacteriophage proteins by mass spectrometry

Following purification of bacteriophages by isopycnic centrifugation in CsCl the bacteriophage bands were dialyzed against TBS and the dialyzed bacteriophage preparation was centrifuged to pellet the viruses [20]. The virus pellet was suspended in electrophoresis buffer followed by SDS-PAGE [87]. Selected bands were digested using a previously described protocol with some modifications [88,89]. Samples were washed twice with 25 mM ammonium bicarbonate (ABC) and 100% acetonitrile (ACN), reduced and alkylated using 10 mM dithiothreitol and 100 mM iodoacetamide then incubated with 75 ng sequencing grade modified porcine trypsin (Promega, Fitchburg WI) in 25 mM ABC overnight at 37C. Peptides were first separated by a Paradigm Multi-Dimensional Liquid Chromatography (MDLC) instrument (Michrom Bioresources Inc., Auburn, CA) with a Magic C18AQ 3  $\mu$  200 Å (0.2×50 mm) column, (Michrom Bioresources Inc.) using a ZORBAX 300SB-C18 5  $\mu$  (5×0.3 mm) trap (Agilent Technologies, Santa Clara, CA). The flow rate was 4  $\mu$ l/min and the solvent gradient was from 5% B (5 min) to 45% B over 90 min, then 80% B (1 min). Solvent A was 0.1% aqueous formic acid and solvent B contained 0.1% formic acid in ACN. Eluted peptides were analyzed using a LTQ-Orbitrap XL (ThermoElectron, Bremen Germany) equipped with a Captive Spray source (Michrom Bioresources Inc.) using Xcalibur v2.0.7. The MS was operated in data-dependent mode switching between Orbitrap-MS and LTQ-MS/MS. Full scan MS spectra ( $m/z$  300–1800) were acquired in the positive ion mode with resolution of 60,000 in profile mode. The five most intense data-dependent peaks were subjected to MS/MS using collision-induced dissociation with a minimum signal of 2,000 and isolation width of 3.0 with normalized 35.0 collision energy. Ions already

selected were dynamic excluded for 30 seconds after a repeat count of 2 with a repeat duration of 10 seconds. A reject mass list was used which included known background ions and trypsin fragments.

The MS/MS data were extracted using Sorcerer v3.5 (Sage-N Research, Milpitas CA). Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Sequest (Thermo Fisher Scientific, San Jose, CA, version v.27, rev. 11). The search was performed using a combined version of all *Clostridium* spp. entries out of NCBI and in-house bacteriophage sequences (data from these investigations) with a random, concatenated decoy database added (including 529,702 entries) assuming the digestion enzyme trypsin. Sequest was searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 10 ppm. Iodoacetamide derivative of cysteine was specified in Sequest as a fixed modification, oxidation of methionine was specified in Sequest as variable. Scaffold (version Scaffold\_3.3.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the PeptideProphet algorithm [90]. Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned using the Protein Prophet algorithm [91]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. The False Discovery Rate (FDR) was calculated by Scaffold using an empirical method by Kall et al. [92].

## Supporting Information

**Figure S1 Representative Plaque Morphology.** Bacteriophage  $\Phi$ CP7R was propagated in the host *Clostridium perfringens* Cp7 for titration and plaques were photographed with an AlphaImagerHP. (EPS)

**Figure S2 Whole Genome similarity Comparison for Bacteriophages  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R.** The bacteriophage genomes were aligned in Mauve for comparative genomics to determine overall sequence similarity and if any rearrangement,

segmental duplication, gain, or loss created a mosaic pattern of homology.

(EPS)

**Figure S3 Restriction Endonuclease Maps for the Genomes of Bacteriophages  $\Phi$ CPV4,  $\Phi$ ZP2 and  $\Phi$ CP7R and Terminal Protein Detection.** A) *EcoRV* and *Eco91I* maps for all three bacteriophages based on genome nucleotide sequences. B) Electropherogram of  $\Phi$ CPV4 genomic fragments following digestion with protease K and *EcoRV* (lane 1) and *EcoRV* without protease digestion (lane 2). Arrows indicate terminal genome fragments.

(EPS)

**Table S1 Open Reading Frames for the  $\Phi$ CPV4 Genome.**

(XLS)

**Table S2 Open Reading Frames for the  $\Phi$ ZP2 Genome.**

(XLS)

**Table S3 Open Reading Frames for the  $\Phi$ CP7R Genome.**

(XLS)

**Table S4 Purified Bacteriophage Protein Mass Spectrometry Proteomics Data.**

(XLSX)

## Acknowledgments

The continuing support of Melanie Peterson at ARS-OIRP and Patrick Russo at ISTC is greatly appreciated. We thank Mary Ard in the Department of Pathology, College of Veterinary Medicine at the University of Georgia, Athens, USA for electron microscopy. GenBank NCBI accession numbers for the genomes were assigned as JQ729990 for  $\Phi$ CP7R, JQ729991 for  $\Phi$ CPV4 and JQ729992 for  $\Phi$ ZP2. NOTE: Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

## Author Contributions

Conceived and designed the experiments: NV ES GS BS. Performed the experiments: VV VB VK AP EZ JG KS RW DQ JL KH. Analyzed the data: NV BO CM BS. Contributed reagents/materials/analysis tools: NV BO CM. Wrote the paper: BS NV.

## References

- Smedley JG, 3rd, McClane BA (2004) Fine mapping of the N-terminal cytotoxicity region of *Clostridium perfringens* enterotoxin by site-directed mutagenesis. *Infect Immun* 72: 6914–6923.
- Sawires YS, Songer JG (2006) *Clostridium perfringens*: insight into virulence evolution and population structure. *Anaerobe* 12: 23–43.
- Rood JI (1998) Virulence genes of *Clostridium perfringens*. *Annual Reviews of Microbiology* 52: 333–360.
- McClane B (2001) *Clostridium perfringens*. In: Doyle M, ed. *Food Microbiology: Fundamentals and Frontiers*. Second ed. Washington DC: ASM Press. pp 351–372.
- Sarker MR, Carman RJ, McClane BA (1999) Inactivation of the gene (*cpe*) encoding *Clostridium perfringens* enterotoxin eliminates the ability of two *cpe*-positive *C. perfringens* type A human gastrointestinal disease isolates to affect rabbit ileal loops. *Mol Microbiol* 33: 946–958.
- Lindstrom M, Heikinheimo A, Lahti P, Korkeala H (2011) Novel insights into the epidemiology of *Clostridium perfringens* type A food poisoning. *Food Microbiol* 28: 192–198.
- Van Immerseel F, De Buck J, Pasmans F, Huyghebaert G, Haesebrouck F, et al. (2004) *Clostridium perfringens* in poultry: an emerging threat for animal and public health. *Avian Pathol* 33: 537–549.
- Gyles CL (2008) Antimicrobial resistance in selected bacteria from poultry. *Anim Health Res Rev* 9: 149–158.
- DuPont HL (2007) The growing threat of foodborne bacterial enteropathogens of animal origin. *Clin Infect Dis* 45: 1353–1361.
- Castanon JI (2007) History of the use of antibiotic as growth promoters in European poultry feeds. *Poult Sci* 86: 2466–2471.
- Liu J, Dehbi M, Moeck G, Arhin F, Bauda P, et al. (2004) Antimicrobial drug discovery through bacteriophage genomics. *Nat Biotechnol* 22: 185–191.
- Fischetti VA (2008) Bacteriophage lysins as effective antibacterials. *Curr Opin Microbiol* 11: 393–400.
- Kutateladze M, Adamia R (2010) Bacteriophages as potential new therapeutics to replace or supplement antibiotics. *Trends Biotechnol* 28: 591–595.
- Fenton M, Ross P, McAuliffe O, O'Mahony J, Coffey A (2010) Recombinant bacteriophage lysins as antibacterials. *Bioeng Bugs* 1: 9–16.
- Bren L (2007) Bacteria-eating virus approved as food additive. *FDA Consum* 41: 20–22.
- Zimmer M, Scherer S, Loessner MJ (2002) Genomic analysis of *Clostridium perfringens* bacteriophage phi3626, which integrates into *guaA* and possibly affects sporulation. *J Bacteriol* 184: 4359–4368.
- Zimmer M, Vukov N, Scherer S, Loessner MJ (2002) The murein hydrolase of the bacteriophage phi3626 dual lysis system is active against all tested *Clostridium perfringens* strains. *Appl Environ Microbiol* 68: 5311–5317.
- Nariya H, Miyata S, Tamai E, Sekiya H, Maki J, et al. (2011) Identification and characterization of a putative endolysin encoded by episomal phage phiSM101 of *Clostridium perfringens*. *Appl Microbiol Biotechnol* 90: 1973–1979.
- Oakley BB, Talundzic E, Morales CA, Hiatt KL, Siragusa GR, et al. (2011) Comparative genomics of four closely related *Clostridium perfringens* bacterio-

- phages reveals variable evolution among core genes with therapeutic potential. *BMC Genomics* 12: 282.
20. Seal BS, Fouts DE, Simmons M, Garrish JK, Kuntz RL, et al. (2011) *Clostridium perfringens* bacteriophages PhiCP390 and PhiCP26F: genomic organization and proteomic analysis of the virions. *Arch Virol* 156: 25–35.
  21. Volozhantsev NV, Verevkin VV, Bannov VA, Krasniknikova VM, Myakinina VP, et al. (2011) The genome sequence and proteome of bacteriophage PhiCPV1 virulent for *Clostridium perfringens*. *Virus Res* 155: 433–439.
  22. Morales CA, Oakley BB, Garrish JK, Siragusa GR, Ard MB, et al. (2012) Complete genome sequence of the podoviral bacteriophage PhiCP24R, which is virulent for *Clostridium perfringens*. *Arch Virol*.
  23. Siragusa GR, Danyluk MD, Hiatt KL, Wise MG, Craven SE (2006) Molecular subtyping of poultry-associated type A *Clostridium perfringens* isolates by repetitive-element PCR. *J Clin Microbiol* 44: 1065–1073.
  24. Ackermann HW (1998) Tailed bacteriophages: the order *Caudovirales*. *Adv Virus Res* 51: 135–201.
  25. Meijer WJ, Horcajadas JA, Salas M (2001) Phi29 family of phages. *Microbiol Mol Biol Rev* 65: 261–287.
  26. Lavigne R, Seto D, Mahadevan P, Ackermann H-W, Kropinski AM (2008) Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTP-based tools. *Research in Microbiology* 159: 406–414.
  27. McNair K, Bailey BA, Edwards RA (2012) PHACTS, a computational approach to classifying the lifestyle of phages. *Bioinformatics* 28: 614–618.
  28. Salas M, Blanco L, Lazaro JM, de Vega M (2008) The bacteriophage phi29 DNA polymerase. *IUBMB Life* 60: 82–85.
  29. Perez-Arnaiz P, Lazaro JM, Salas M, de Vega M (2010) phi29 DNA polymerase active site: role of residue Val250 as metal-dNTP complex ligand and in protein-primed initiation. *J Mol Biol* 395: 223–233.
  30. Garvey KJ, Saedi MS, Ito J (1985) The complete sequence of *Bacillus* phage phi 29 gene 16: a protein required for the genome encapsidation reaction. *Gene* 40: 311–316.
  31. Guo P, Peterson C, Anderson D (1987) Initiation events in *in-vitro* packaging of bacteriophage phi 29 DNA-gp3. *J Mol Biol* 197: 219–228.
  32. Guo PX, Erickson S, Anderson D (1987) A small viral RNA is required for *in vitro* packaging of bacteriophage phi 29 DNA. *Science* 236: 690–694.
  33. Guo P, Peterson C, Anderson D (1987) Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage phi 29. *J Mol Biol* 197: 229–236.
  34. Hudspeth DS, Vary PS (1992) spoVG sequence of *Bacillus megaterium* and *Bacillus subtilis*. *Biochim Biophys Acta* 1130: 229–231.
  35. Matsuno K, Sonenshein AL (1999) Role of SpoVG in asymmetric septation in *Bacillus subtilis*. *J Bacteriol* 181: 3392–3401.
  36. Liepinsh E, Genereux C, Dehareng D, Joris B, Otting G (2003) NMR structure of *Citrobacter freundii* AmpD, comparison with bacteriophage T7 lysozyme and homology with PGRP domains. *J Mol Biol* 327: 833–842.
  37. Porter CJ, Schuch R, Pelzek AJ, Buckle AM, McGowan S, et al. (2007) The 1.6 Å crystal structure of the catalytic domain of PlyB, a bacteriophage lysin active against *Bacillus anthracis*. *J Mol Biol* 366: 540–550.
  38. Zou Y, Hou C (2010) Systematic analysis of an amidase domain CHAP in 12 *Staphylococcus aureus* genomes and 44 staphylococcal phage genomes. *Comput Biol Chem* 34: 251–257.
  39. Wang IN, Smith DL, Young R (2000) Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* 54: 799–825.
  40. Young R (2002) Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* 4: 21–36.
  41. Morais MC, Choi KH, Koti JS, Chipman PR, Anderson DL, et al. (2005) Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol Cell* 18: 149–159.
  42. Simpson AA, Tao Y, Leiman PG, Badasso MO, He Y, et al. (2000) Structure of the bacteriophage phi29 DNA packaging motor. *Nature* 408: 745–750.
  43. Vlcek C, Paces V (1986) Nucleotide sequence of the late region of *Bacillus* phage phi 29 completes the 19,285-bp sequence of phi 29 genome. Comparison with the homologous sequence of phage PZA. *Gene* 46: 215–225.
  44. Xiang Y, Leiman PG, Li L, Grimes S, Anderson DL, et al. (2009) Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tail spike. *Mol Cell* 34: 375–386.
  45. Villanueva N, Salas M (1981) Adsorption of bacteriophage phi 29 to *Bacillus subtilis* through the neck appendages of the viral particle. *J Virol* 38: 15–19.
  46. Xiang Y, Morais MC, Battisti AJ, Grimes S, Jardine PJ, et al. (2006) Structural changes of bacteriophage phi29 upon DNA packaging and release. *EMBO J* 25: 5229–5239.
  47. Garcia JA, Carrascosa JL, Salas M (1983) Assembly of the tail protein of the *Bacillus subtilis* phage phi 29. *Virology* 125: 18–30.
  48. Xiang Y, Morais MC, Cohen DN, Bowman VD, Anderson DL, et al. (2008) Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage phi29 tail. *Proc Natl Acad Sci U S A* 105: 9552–9557.
  49. Martin AC, Lopez R, Garcia P (1996) Analysis of the complete nucleotide sequence and functional organization of the genome of *Streptococcus pneumoniae* bacteriophage Cp-1. *J Virol* 70: 3678–3687.
  50. Vybiral D, Takac M, Loessner M, Witte A, von Ahsen U, et al. (2003) Complete nucleotide sequence and molecular characterization of two lytic *Staphylococcus aureus* phages: 44AHJD and P68. *FEMS Microbiol Lett* 219: 275–283.
  51. Nelson D, Schuch R, Zhu S, Tscherne DM, Fischetti VA (2003) Genomic sequence of C1, the first streptococcal phage. *J Bacteriol* 185: 3325–3332.
  52. Kwan T, Liu J, DuBow M, Gros P, Pelletier J (2005) The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc Natl Acad Sci U S A* 102: 5174–5179.
  53. Delisle AL, Barcak GJ, Guo M (2006) Isolation and expression of the lysis genes of *Actinomyces naeslundii* phage Av-1. *Appl Environ Microbiol* 72: 1110–1117.
  54. Saren AM, Ravanti JJ, Benson SD, Burnett RM, Paulin L, et al. (2005) A snapshot of viral evolution from genome analysis of the tectiviridae family. *J Mol Biol* 350: 427–440.
  55. Sozhmannan S, McKinstry M, Lentz SM, Jalasvuori M, McAfee F, et al. (2008) Molecular characterization of a variant of *Bacillus anthracis*-specific phage AP50 with improved bacteriolytic activity. *Appl Environ Microbiol* 74: 6792–6796.
  56. Teeling H, Waldmann J, Lombardot T, Bauer M, Glockner FO (2004) TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinformatics* 5: 163.
  57. Richter M, Rosello-Mora R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106: 19126–19131.
  58. Dean FB, Nelson JR, Giesler TL, Lasken RS (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res* 11: 1095–1099.
  59. Edwards RA, Rohwer F (2005) Viral metagenomics. *Nat Rev Microbiol* 3: 504–510.
  60. Burrows B, Harper DR, Anderson J, McConville M, Enright MC (2011) Bacteriophage therapy: potential uses in the control of antibiotic-resistant pathogens. *Expert Rev Anti Infect Ther* 9: 775–785.
  61. Chan BK, Abedon ST (2012) Phage therapy pharmacology phage cocktails. *Adv Appl Microbiol* 78: 1–23.
  62. Wise MG, Siragusa GR (2005) Quantitative detection of *Clostridium perfringens* in the broiler fowl gastrointestinal tract by real-time PCR. *Appl Environ Microbiol* 71: 3911–3916.
  63. Smith HW (1959) The bacteriophages of *Clostridium perfringens*. *J Gen Microbiol* 21: 622–630.
  64. Helms C, Graham MY, Dutchik JE, Olson MV (1985) A new method for purifying lambda DNA from phage lysates. *DNA* 4: 39–49.
  65. Sambrook J, Maniatis T, Fritsch EF (1989) Molecular cloning: A laboratory manual. Cold Spring Harbor, NY: Cold Spring Laboratory Press.
  66. Ackermann HW (2009) Basic phage electron microscopy. *Methods Mol Biol* 501: 113–126.
  67. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.
  68. Fouts DE, Rasko DA, Cer RZ, Jiang L, Fedorova NB, et al. (2006) Sequencing *Bacillus anthracis* typing phages gamma and cherry reveals a common ancestry. *J Bacteriol* 188: 3402–3408.
  69. Lewis JG, Chang GJ, Lanciotti RS, Trent DW (1992) Direct sequencing of large flavivirus PCR products for analysis of genome variation and molecular epidemiological investigations. *J Virol Methods* 38: 11–23.
  70. Mead DA, Pey NK, Hernstadt C, Marciel RA, Smith LM (1991) A universal method for the direct cloning of PCR amplified nucleic acid. *Biotechnology (N Y)* 9: 657–663.
  71. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, et al. (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* 321: 674–679.
  72. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European molecular biology open software suite. *Trends in Genetics* 16: 276–277.
  73. Lukashin AV, Borodovsky M (1998) GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* 26: 1107–1115.
  74. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, et al. (2009) IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25: 2271–2278.
  75. Schaffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, et al. (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res* 29: 2994–3005.
  76. Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, et al. (2007) CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res* 35: D237–240.
  77. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567–580.
  78. Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng* 10: 673–676.
  79. Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14: 1394–1403.
  80. Darling AE, Mau B, Perna NT (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5: e11147.
  81. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
  82. Hedges SB (1992) The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies. *Mol Biol Evol* 9: 366–369.
  83. Smith AB (1994) Rooting molecular trees - Problems and strategies. *Biological Journal of the Linnean Society* 51: 279–292.

84. Huelsenbeck JP, Bollback JP, Levine AM (2002) Inferring the root of a phylogenetic tree. *Syst Biol* 51: 32–43.
85. RDC: T (2009) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
86. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461.
87. Hames BD (1981) An introduction to polyacrylamide gel electrophoresis. In: Hames BD, Rickwood D, eds. *Gel Electrophoresis of Proteins: A Practical Approach*. Oxford: IRL Press. pp 1–91.
88. Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal Biochem* 203: 173–179.
89. Finchout EJ, Lee KH (2003) Comparison of automated in-gel digest methods for femtomole level samples. *Electrophoresis* 24: 3508–3516.
90. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74: 5383–5392.
91. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75: 4646–4658.
92. Kall L, Storey JD, MacCoss MJ, Noble WS (2008) Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. *J Proteome Res* 7: 29–34.