



11-17-2023

A Novel Subcluster of Closely Related *Bacillus* Phages with Distinct Tail Fiber/Lysin Gene Combinations

Rachel E. Loney
Duke University

Veronique A. Delesalle
Gettysburg College

Brianne E. Chaudhry
LGC Clinical Diagnostics

Megan Czerpak
Gettysburg College

Alexandra A. Guffey
Gettysburg College

See next page for additional authors

Follow this and additional works at: <https://cupola.gettysburg.edu/biofac>



Part of the [Biology Commons](#), [Genetics and Genomics Commons](#), and the [Virology Commons](#)

Share feedback about the accessibility of this item.

Recommended Citation

Loney, R.E.; Delesalle, V.A.; Chaudry, B.E.; Czerpak, M.; Guffey, A.A.; Goubet-McCall, L.; McCarty, M.; Strine, M.S.; Tanke, N.T.; Vill, A.C.; et al. A Novel Subcluster of Closely Related *Bacillus* Phages with Distinct Tail Fiber/Lysin Gene Combinations. *Viruses* 2023, *15*, 2267. <https://doi.org/10.3390/v15112267>

This open access article is brought to you by The Cupola: Scholarship at Gettysburg College. It has been accepted for inclusion by an authorized administrator of The Cupola. For more information, please contact cupola@gettysburg.edu.

A Novel Subcluster of Closely Related Bacillus Phages with Distinct Tail Fiber/Lysin Gene Combinations

Abstract

Bacteriophages (phages) are the most numerous entities on Earth, but we have only scratched the surface of describing phage diversity. We isolated seven *Bacillus subtilis* phages from desert soil in the southwest United States and then sequenced and characterized their genomes. Comparative analyses revealed high nucleotide and amino acid similarity between these seven phages, which constitute a novel subcluster. Interestingly, the tail fiber and lysin genes of these phages seem to come from different origins and carry out slightly different functions. These genes were likely acquired by this subcluster of phages via horizontal gene transfer. In conjunction with host range assays, our data suggest that these phages are adapting to hosts with different cell walls.

Keywords

bacteriophage, *Bacillus subtilis*, comparative genomics, horizontal gene transfer, functional annotation, tail fiber/lysine gene combinations

Disciplines

Biology | Genetics and Genomics | Microbiology | Virology

Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

Authors

Rachel E. Loney, Veronique A. Delesalle, Brianne E. Chaudhry, Megan Czerpak, Alexandra A. Guffey, Leo Goubet-McCall, Michael McCarty, Madison S. Strine, Natalie T. Tanke, Albert C. Vill, and Gregory P. Krukonis

Article

A Novel Subcluster of Closely Related *Bacillus* Phages with Distinct Tail Fiber/Lysin Gene Combinations

Rachel E. Loney^{1,*}, Véronique A. Delesalle^{2,*}, Brianne E. Chaudry³, Megan Czerpak², Alexandra A. Guffey⁴, Leo Goubet-McCall⁵, Michael McCarty², Madison S. Strine⁶, Natalie T. Tanke⁷, Albert C. Vill⁸ and Greg P. Krukoniš⁹

¹ University Program in Genetics and Genomics, School of Medicine, Duke University, Durham, NC 27708, USA

² Department of Biology, Gettysburg College, 300 N Washington St., Gettysburg, PA 17325, USA; czerme01@gettysburg.edu (M.C.); mccami02@gettysburg.edu (M.M.)

³ LGC Clinical Diagnostics, 910 Clopper Rd., Gaithersburg, MD 20874, USA; tomkobri12@gmail.com

⁴ Janssen Scientific Affairs, LLC, 200 Tournament Dr., Horsham, PA 19044, USA; aagesen22696@gmail.com

⁵ Department of Biology, The Pennsylvania State University, 201 Huck Life Sciences Building, University Park, PA 16802, USA; ljg5555@psu.edu

⁶ Department of Immunobiology, Yale School of Medicine, 333 Cedar St., New Haven, CT 06510, USA; madison.strine@yale.edu

⁷ Department of Cell Biology and Physiology, University of North Carolina, Chapel Hill, NC 27599, USA; ntanke22@live.unc.edu

⁸ Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA; albert.vill@yale.edu

⁹ Department of Biology, Angelo State University, Cavness Science Building 101, ASU Station #10890, San Angelo, TX 76909, USA; greg.krukoniš@angelo.edu

* Correspondence: rachel.loney@duke.edu (R.E.L.); delesall@gettysburg.edu (V.A.D.)

Abstract: Bacteriophages (phages) are the most numerous entities on Earth, but we have only scratched the surface of describing phage diversity. We isolated seven *Bacillus subtilis* phages from desert soil in the southwest United States and then sequenced and characterized their genomes. Comparative analyses revealed high nucleotide and amino acid similarity between these seven phages, which constitute a novel subcluster. Interestingly, the tail fiber and lysin genes of these phages seem to come from different origins and carry out slightly different functions. These genes were likely acquired by this subcluster of phages via horizontal gene transfer. In conjunction with host range assays, our data suggest that these phages are adapting to hosts with different cell walls.

Keywords: bacteriophage; *Bacillus subtilis*; comparative genomics; horizontal gene transfer; functional annotation; tail fiber/lysine gene combinations



Citation: Loney, R.E.; Delesalle, V.A.; Chaudry, B.E.; Czerpak, M.; Guffey, A.A.; Goubet-McCall, L.; McCarty, M.; Strine, M.S.; Tanke, N.T.; Vill, A.C.; et al. A Novel Subcluster of Closely Related *Bacillus* Phages with Distinct Tail Fiber/Lysin Gene Combinations. *Viruses* **2023**, *15*, 2267. <https://doi.org/10.3390/v15112267>

Academic Editor: René Kallies

Received: 14 October 2023

Revised: 11 November 2023

Accepted: 14 November 2023

Published: 17 November 2023



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

1. Introduction

The past several decades have seen a renewed interest in bacteriophages (phages) from both ecological and medical perspectives. Phages enact strong selective pressure on host bacterial species [1,2] and thus play a large role in shaping microbial ecology [1,3–5]. Phages also play a role in biogeochemical cycling by lysing their hosts and shunting host material back into the pool of organic matter [6] and by altering the gene expression and therefore the metabolism of their hosts [7,8]. Phages are also known to impact human health via their role in shaping the gut microbiome [9–11], and phage therapy has become an attractive option as more and more human pathogens acquire antibiotic resistance [12].

However, phages largely remain the “dark matter” of the biosphere [13]. With an estimated 10^{31} bacteriophage particles in the world [14], we have barely scratched the surface of the vast phage diversity that exists around us. Aside from the sheer quantity of phages that have not yet been isolated, within the phages that we have sequenced, there is a

plethora of genes with unknown functions [15–17]. Isolating and sequencing more phages will allow us to more fully describe phage genetic diversity, and will provide a better framework for understanding how different genes and functions impact phage ecology. Expanding our knowledge of phage genetics will allow for advancements in identifying and utilizing new phage genes that attack bacteria, uncovering how phages can manipulate bacterial gene expression and metabolism, and further understanding the co-evolutionary dynamics between phages and their hosts.

To date, most phage research has focused on Actinophages [17–19], phages found in aquatic environments [6], and lab-adapted phages (such as Lambda, T phages, or SPP1) [20]. Our goal has been to document the diversity of phages that are able to lyse “wild” bacterial isolates that have experienced little lab domestication. This will allow for a better picture of phage diversity as it exists in nature. To that end, we have focused on phages that can lyse various strains of *Bacillus subtilis* [20–22]. *B. subtilis* is a highly diverse Gram-positive bacterium typically found in soil [23], and its ability to form endospores facilitates its successful survival in arid desert soil [24–26]. In addition to being a model organism, *B. subtilis* is of interest for many reasons including its capacity for promoting plant growth [27] and its role as a symbiont in the digestive tracts of animal hosts [28,29].

Here, we describe seven phages that we isolated from desert soil from the southwest United States and which constitute a novel subcluster sharing little genetic similarity with previously described phages. As seen in many other phage genomes [16,17,30–34], our phages show evidence of having acquired many genes through horizontal gene transfer (HGT). While many of these are small genes with unknown functions, there is some indication that these phages are adapting to differences in host cell walls.

2. Materials and Methods

Soil was collected from various locations in the southwest United States (Table 1). At each collection site, the GPS coordinates were recorded (Table 1), and soil was collected into a sterile 50 mL conical tube. To isolate phages from each sample, 1 g of soil was added to 20 mL of LB broth containing 10 mM CaCl₂, 10 mM MgCl₂, and 1 mM MnCl₂ (LB2), incubated for 4 h at 37 °C shaking at 250 rpm, and stored at 4 °C after filtering (0.22 µm). A sample of filtrates was then plated on *B. subtilis* strains: either on wild strains T89-06 (also called S89-6 or T89-6), T89-20 (also called S89-20), or TT123, all isolated by Istock and colleagues [35,36] or on the lab-adapted strain 168 [23] which was acquired from the Bacillus Genetic Stock Center. Each of these *B. subtilis* strains are genetically distinct from each other and are kept as spore stocks in our lab (described in [20]). Phages were isolated by single-plaque purifying three times on lawns of the isolation host. Phages were successfully recovered from 1 g of soil from each collection site. Two phages were isolated from 1 g of soil taken from Death Valley, while only one phage was isolated from the other soil samples. High-titer lysates (HTL) were prepared by flooding multiple webby plates with LB broth and filtering the lysate (0.22 µm). DNA was extracted from phage 268TH004 and the phages from Big Bend following the phagesdb phenol–chloroform extraction procedure [37] and from the other phages with the Promega Wizard DNA Clean-Up System.

Phage DNA was sent to North Carolina State University’s Genomic Science Laboratory where libraries were prepared with the Illumina Truseq Nano DNA library prep kit and sequenced on Illumina MiSeq (v3 150 SE flow cell). Upon receiving sequencing data, reads were aligned and assembled using GS de novo assembler v2.9 [38], and quality was verified with ConSeq v29 [39]. Each genome was assembled as a single contig with >1000× coverage. Genome ends were determined with PhageTerm [40].

Finished sequences were annotated within DNAMaster v5.22.22 [41]. Putative genes were called with Glimmer v3.0 and Genemark v2.5 [42,43]. Predicted protein functions were called with BLAST v2.12 (if E-value < 10⁻⁵) [44] and HHPred (if probability > 85%, coverage > 50%, and E-values < 10⁻⁵) [45]. The presence of tRNA genes was confirmed with Aragorn [46]. Default settings were used in all programs, and all genome annotations have been submitted to NCBI (accession numbers can be found in Table 1). For select

genes of interest, predicted domain functions were called with HHPred and with NCBI's Conserved Domain Database (CDD) using RPS-BLAST [47] via Phamerator [48].

Table 1. Basic genomic characteristics of seven novel *B. subtilis* phages. ORF = Open Reading Frame signifying either known or putative gene; * = includes one Asn-GTT tRNA gene; NP = National Park. Genome size includes 839 bp direct terminal repeats. Phages 019DV002 and 019DV004 were isolated from the same gram of soil.

| Phage | Isolation Strain | Genome Size (bp) | % GC | No. of ORFs | Collecting Site and Date | GPS Coordinates | Accession |
|-----------|------------------|------------------|------|-------------|-----------------------------------|-------------------------------|------------|
| 019DV002 | T89-06 | 52,749 | 42.6 | 84 * | Death Valley NP, Ca, May 2014 | 36°27'51.4" N, 117°14'08.6" W | MN176220 |
| 019DV004 | T89-06 | 52,749 | 42.6 | 84 * | | | MN176221 |
| 056SW001B | T89-20 | 52,692 | 42.6 | 83 * | Saguaro West NP, AZ, May 2014 | 32°16'12.7" N, 111°12'20.6" W | MN176230 |
| 268TH004 | TT123 | 52,834 | 42.4 | 80 * | Tumamoc Hill, Tucson, AZ May 2016 | 32°13'04.9" N, 111°00'12.9" W | MW394467.1 |
| 274BB002 | TT123 | 52,372 | 42.5 | 81 * | Big Bend NP, TX, June 2016 | 29°15'54.6" N, 103°07'24.5" W | MZ501264 |
| 276BB001 | TT123 | 52,370 | 42.5 | 81 * | | 29°18'16.6" N, 103°15'52.9" W | MN176231 |
| 280BB001 | 168 | 52,371 | 42.5 | 81 * | | 29°18'29.0" N, 103°28'42.6" W | MN176232 |

A Phamerator database was constructed with 312 *Bacillus* phages with complete genomes in NCBI as of 1 May 2022 (*Bacillus* v5—<https://phamerator.org>) [48]. This database groups gene products together into protein families with related sequences (“phams”) using the PhamMMseqs pipeline, and phages that share at least 35% of their phams are described as belonging to the same cluster [49]. Genome maps were created with Phamerator to illustrate the relationships between phage genomes using both amino acid and nucleotide similarity. Additional genomic comparisons were made by (1) calculating the average amino acid identity (AAI) and average nucleotide identity (ANI) using *enve-omics* lab algorithms [50], (2) performing BLASTn (nr/nt database) and BLASTp (nr database) searches (performed July 2023) either against the complete database or against the database for tailed phages (taxid: 10699, 10662, and 10744), or (3) aligning amino acid sequences of homologous genes using MUSCLE [51] in MEGAX (default parameters) [52]. Alignment figures were made with ESPript3.0 [53].

Representative brightfield TEM images were taken of phage 019DV002. First, 1 mL of 019DV002 HTL was centrifuged at 10,000 × *g* for 1 h at 4 °C. TEM grids (Ted Pella, Inc., Redding, CA, USA, Prod. #01822) were treated with the concentrated phage particles, washed briefly with ddH₂O, and then stained with 2% phosphotungstic acid. Images were taken at 130,000 × magnification using a Philips CM1000 TEM.

We tested the phages' ability to lyse various *B. subtilis* strains using a microtiter plate-based assay [54] with an MOI of 0.1 using 9.1 × 10⁷ PFU/mL and 9.1 × 10⁸ CFU/mL. Phage and bacteria were plated in 220 µL of LB2 broth in a 96-well plate, and each combination of phage and bacteria was replicated five times. Wells with bacteria but without phage served as controls. Phages were tested against seven strains: two lab strains 168 (BGSCID 1A1) and W23SR (BGSCID 2A3) [55] and five genetically distinct wild strains in our lab collection originally from [35,36]. Plates were incubated for 20 h at 37 °C with auto-mixing for 10 s every 5 min, and absorbance at 600 nm was measured every 30 min. For a conservative estimate, a phage was said to impact bacterial growth if at least 4 of the 5 replicates showed lower absorbance readings than the control bacteria. To complement the plate assay data, spot testing phages on each host was also performed by spotting 10⁴ PFU on a lawn of each host and confirming presence or absence of plaques.

3. Results

3.1. Basic Genomic Data

Seven *B. subtilis* phages were isolated from soil collected in the southwest United States (Table 1). 280BB001 was isolated on the lab strain *B. subtilis* 168, and the other six phages were isolated on wild strains of *B. subtilis* (Table 1). These phages have double-stranded DNA genomes with 839 bp direct terminal repeats and a GC content (42.4–42.6%) quite close to *B. subtilis*'s GC content of 43.5% [56] (Table 1). The genomes range from 52,370 to 52,834 bp in length and contain 80–84 putative protein-coding genes (Table 1, Figures 1 and 2). All seven phages have the same Asn-GTT tRNA gene, located just before the scaffolding gene (Figure 1), and no other tRNAs. Each phage makes clear plaques on *B. subtilis* lawns, indicating that they are lytic phages. This conclusion is supported by the lack of integrase protein-coding genes in their genome annotations. Both genome organization and electron microscopy (Figure 3) support that these phages are siphoviruses [16].

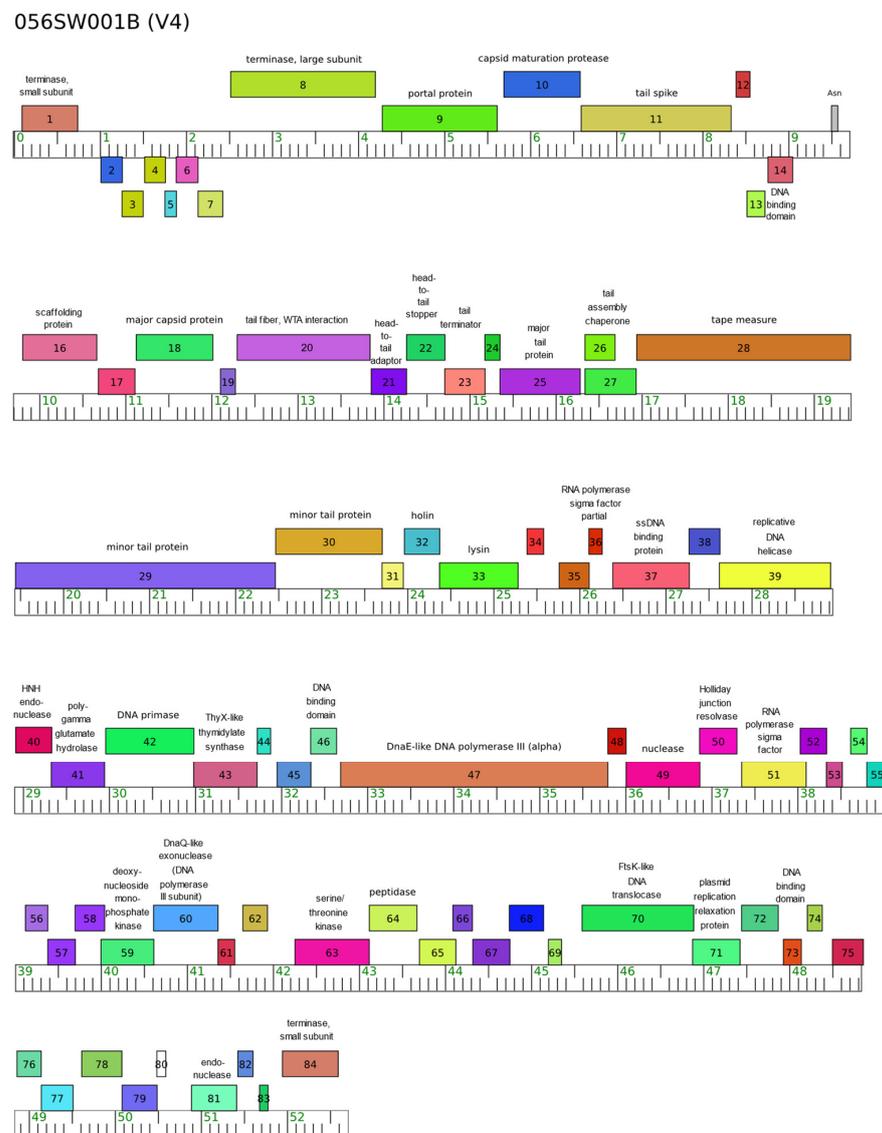


Figure 1. Genome map of 056SW001B, chosen because it shares phams with some but not all of the other phages in this subcluster. The ruler shows genome length (in kilobases) with forward and reverse genes shown above and below the ruler, respectively. Functions or putative functions are listed above genes. Due to direct terminal repeats, gene 84 is a duplication of gene 1 (terminase small subunit). Map was created using Phamerator [48].

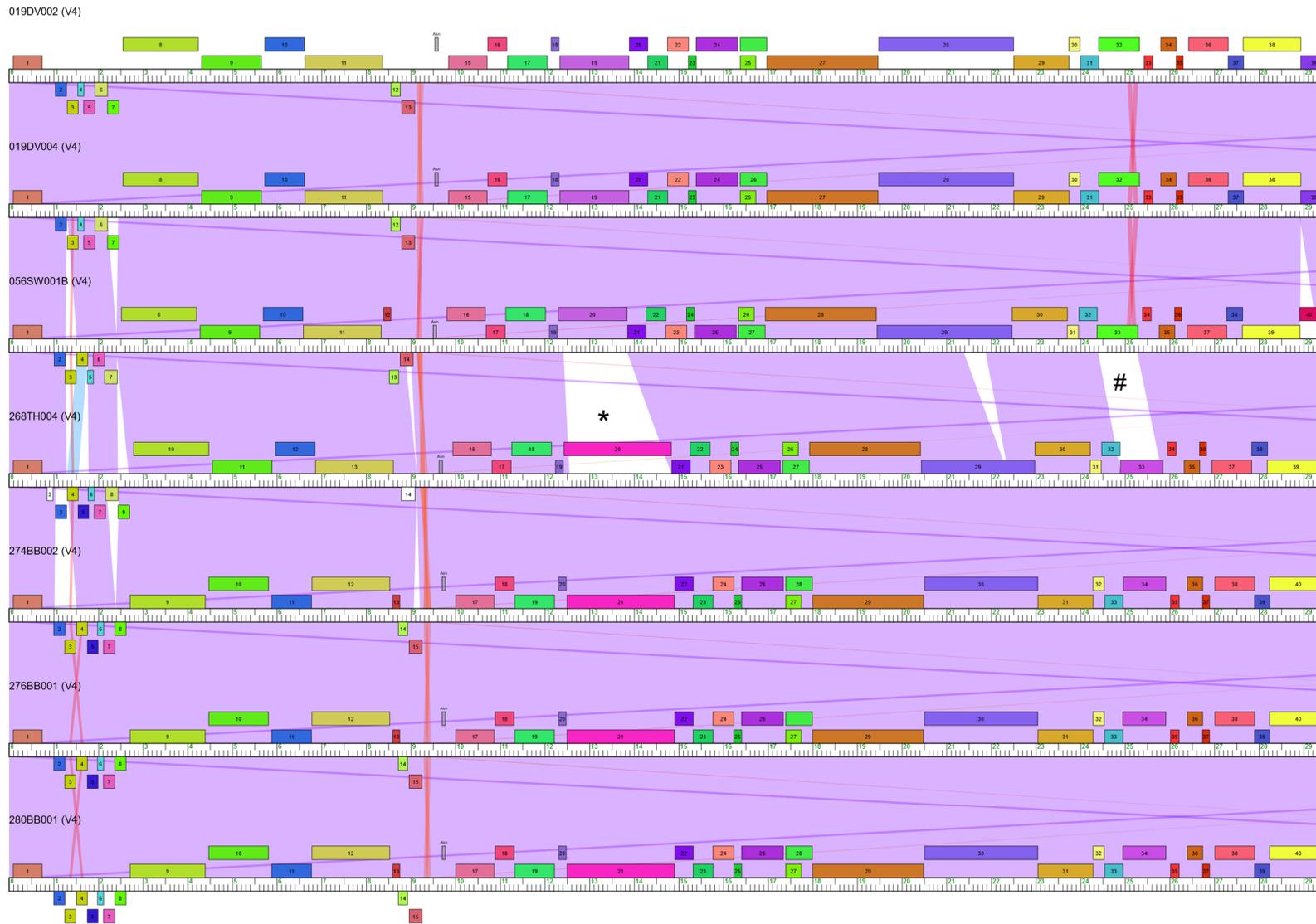


Figure 2. Cont.

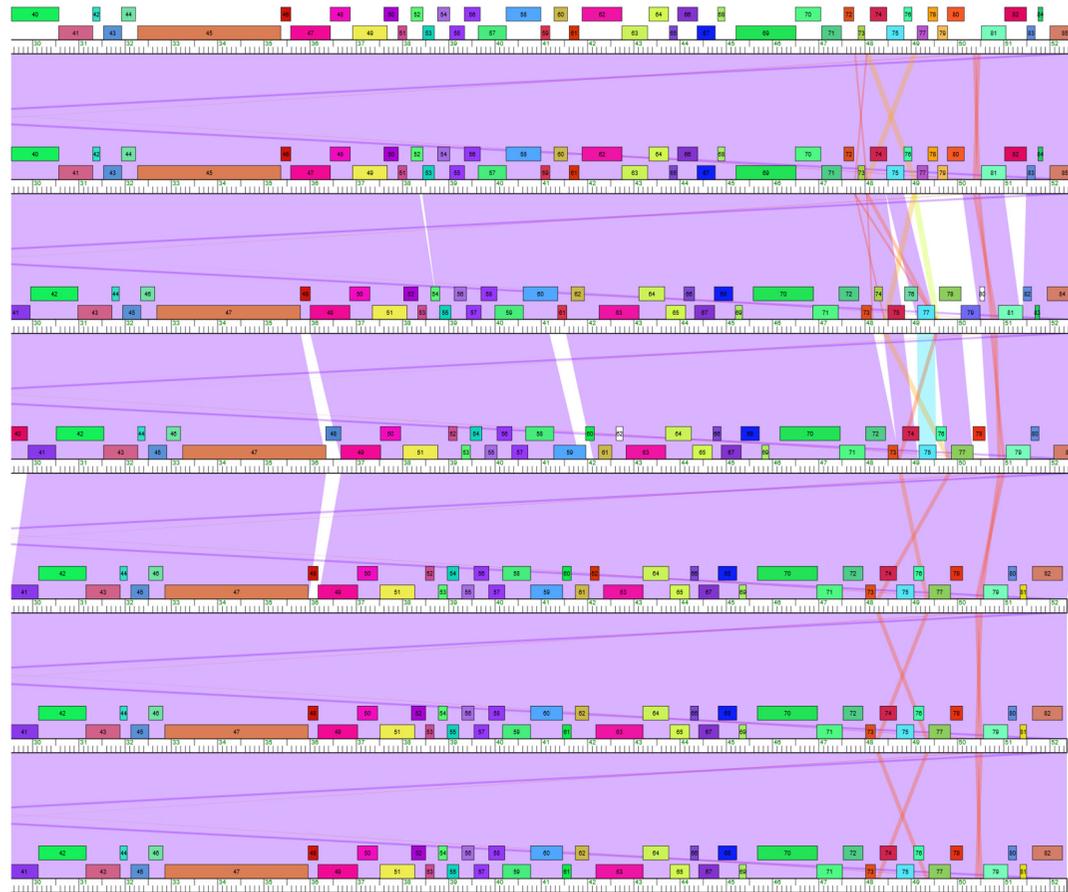


Figure 2. Whole genome maps of the V4 phages, created with Phamerator [48]. Nucleotide sequence similarity between adjacent genomes is indicated by shading from lavender (most similar) to red (least similar). A lack of shading shows that there is no similarity based on local alignment E score greater than 10^{-4} . * marks the WTA-tail fiber/peptidase phams, # marks the lysin/endolysin phams. A total of 96 phams were identified among these seven phages.

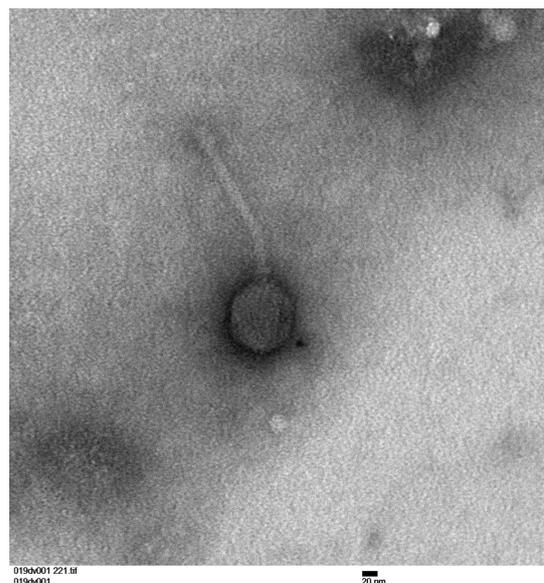


Figure 3. Transmission electron microscopy image of phage 019DV002. Image taken at 130,000x using a Phillips CM1000 TEM.

These phages show high average nucleotide and amino acid identity (ANI and AAI, respectively) when compared to each other, ranging from 95.4% (ANI) or 91.7% (AAI) to 100% (Table 2). Based on their high similarity to each other and using the Phamerator clustering program, these seven phages have been classified as belonging to their own subcluster, which is designated as V4. The phages isolated from the same locations are quite similar to each other with 100% ANI between the Death Valley (DV) phages and $\geq 99.9\%$ ANI between the Big Bend (BB) phages. The DV phages differ by only two non-synonymous SNPs in one gene (gene 19, tail fiber). 274BB002 differs from the other BB phages with two 1 bp intergenic insertions (a G at position 838 and a C at position 51,533, both present in 274BB002 only). 280BB001 differs from the other BB phages with one 1 bp intergenic indel (8575 in 280BB001: 16 vs. 15 T). Finally, 276BB001 and 280BB001 also differ by one non-synonymous SNP (base 33,057) in gene 47 (DnaE-like DNA polymerase III alpha). Although the DV phages were isolated from the same gram of soil, the BB phages were isolated from soil samples between 14 and 35 km apart from each other. Whole genome BLASTn searches show that the next best hit to these seven phages, besides each other, is *Bacillus pseudocaliphilus* phage vB_BpsS-36 from subcluster V2 (coverage = 29%, identity = 69%).

Table 2. Average nucleotide identity (ANI, above the diagonal) and average amino acid identity (AAI, below the diagonal) for the V4 phages. Data presented as percentage and calculated using algorithms from [50].

| | 019DV002 | 019DV004 | 056SW001B | 268TH004 | 274BB002 | 276BB001 | 280BB001 |
|-----------|----------|----------|-----------|----------|----------|----------|----------|
| 019DV002 | --- | 100 | 96.7 | 95.4 | 96.2 | 96.2 | 96.2 |
| 019DV004 | 100 | --- | 96.7 | 95.4 | 96.2 | 96.2 | 96.2 |
| 056SW001B | 94.6 | 94.8 | --- | 95.4 | 96.4 | 96.4 | 96.4 |
| 268TH004 | 91.2 | 92.3 | 93.2 | --- | 98.1 | 98.1 | 98.1 |
| 274BB002 | 92.3 | 92.7 | 92.4 | 95.5 | --- | 100 | 99.9 |
| 276BB001 | 92.4 | 92.8 | 92.4 | 95.3 | 100 | 100 | 100 |
| 280BB001 | 92.4 | 92.8 | 92.4 | 94.9 | 100 | 100 | --- |

Among the V4 phages, we have identified 96 phams (“phamilies” of related protein sequences [18,48]). Of those phams, 64 are shared by all seven phages, with 19 phams that are unique to the V4 subcluster (based on the 312 phages in the *Bacillus* Phamerator database). Only two of the unique phams shared by all V4 phages have functional annotations: a tail assembly chaperone and a partial RNA polymerase sigma factor. Of the remaining 32 phams shared by six or fewer V4 phages, 15 are exclusive to the V4 subcluster. These 15 phams include 4 orphans (phams with a single phage member): 3 from 268TH004, and 1 from 056SW001B. None of these orphans have functional annotations.

3.2. Horizontal Gene Transfer in Subcluster V4 Evolution

Despite the high pairwise AAls across phages in this subcluster, a third of the phams that we identified (32 out of 96) are not shared by all V4 phages. We see evidence of HGT in how these 32 phams are distributed across the V4 phage genomes. Most of these phams are fairly small, which is unsurprising given that smaller genes may be more readily exchanged [17]. These phams tend to have no known function, and tend to be inserted between genes that are conserved across the V4 phages (Figure 2).

However, there are two instances in the V4 genomes where a large gene with a functional annotation has been fully replaced by a different gene. First, where the DV (gene 19) and Saguario West (SW) (gene 20) phages have a tail fiber gene, the Tumamoc Hill (TH) (gene 20) and BB (gene 21) phages have a gene that has been annotated as a peptidase but likely functions as a tail fiber (Figure 2, marked with *). For clarity, the DV and SW gene will be referred to as the “WTA-tail fiber” (wall teichoic acid–tail fiber), and the TH and BB gene will be referred to as the “peptidase.” The second instance involves the same phages with different lysin genes in the DV (gene 32) and SW (gene 33) phages versus the TH (gene 33) and BB (gene 34) phages (where the lysin gene product is more specifically called a LysM-like endolysin) (Figure 2, marked with #). For clarity, the DV and SW gene will be referred to as the “lysin” and the TH and BB gene will be referred to as the “endolysin.” Outside of the V4 phages, there are no phages who share these same pairs of phams in the current *Bacillus* Phamerator database; the WTA–tail fiber is only present in V4 phages, and the endolysin/peptidase combination is only seen in V4 phages.

3.2.1. Tail Fiber

The WTA-tail fiber and peptidase genes likely carry out similar functions, but it appears that they come from different origins. The WTA-tail fiber gene has sequence similarity to the receptor binding protein (RBP) of *Staphylococcus aureus* siphophage ϕ 11 [57]. The ϕ 11 RBP assembles into a homotrimer that is required for ϕ 11 adsorption onto wall teichoic acids (WTA) [57], which are a key component of Gram-positive bacterial cell walls and known to be receptors for phage adsorption [58–61]. Interestingly, this is the gene where the two SNPs in the DV phages are found (at positions 12,866 and 13,757). In the TH and BB phages, the peptidase gene contains a C-terminal G2 peptidase domain that has been identified as the tail spike protein in a number of bacteriophages [62], supporting the assumption that the peptidase likely performs the same function as the DV and SW phage tail fiber.

The WTA-tail fiber pham is only present in the V4 subcluster, while the peptidase pham is found in 15 other *Bacillus* phages: subclusters AA, B1, D2, and U. Subcluster AA is composed of two phages that our lab isolated from Tumamoc Hill, which is the isolation site for 268TH004 [22]. In the V4 subcluster, the WTA-tail fiber/peptidase gene is located between the major capsid protein (MCP) and the head-to-tail adaptor (HTA). However, the other V phages do not share gene synteny here, and they either contain smaller genes that lack a structural function (subclusters V1, V2, V5) or have no gene separating the MCP and HTA (subcluster V3). Additionally, the next best BLASTp hits to the WTA-tail fiber are all to teichoic acid biosynthesis proteins in various *Bacillus* species (identity \leq 43%) rather than to other phages. Similarly, the next best BLASTp hits to the peptidase are to peptidase G2 autoproteolytic cleavage domain-containing proteins in various *Bacillaceae* species

(identity $\leq 68\%$) rather than other phages. The break in gene synteny among V phages, the sharing of the peptidase pham with phages in a different cluster that are geographically near 268TH004, and the BLASTp results suggest that these genes were acquired by the V4 phages through HGT. We note, however, that the first ~50 amino acid residues are highly conserved between the WTA-tail fiber and peptidase (Figure 4). This region is upstream of the first functional domain of either protein. It is possible that this region was present in the V4 common ancestor and was not recombined with the rest of the gene, although there is still no notable homology between this region and the other V subclusters.

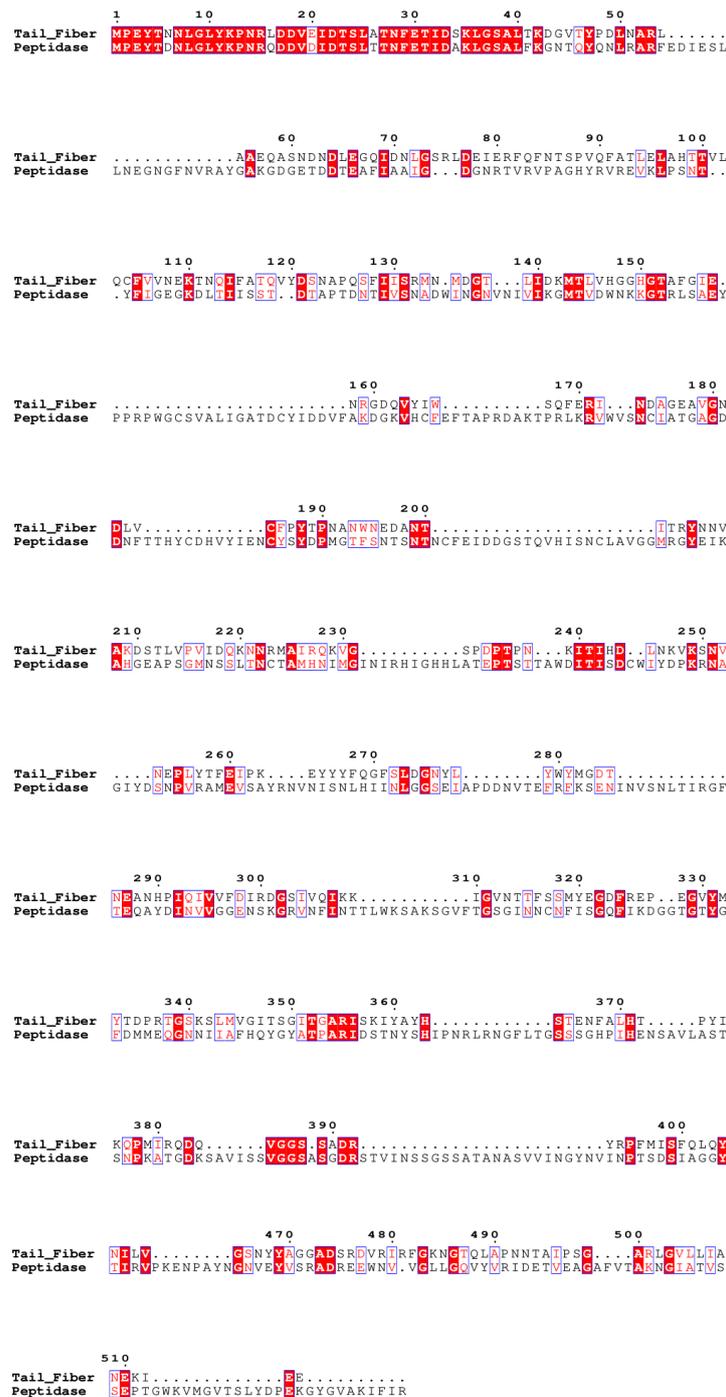


Figure 4. MUSCLE alignment of 056SW001B gene 20 (Tail_Fiber) and 268TH004 gene 20 (Peptidase), made with [52,53].

3.2.2. Lysin

As is typical for Gram-positive bacteriophages [63], the V4 lysin proteins each have an N-terminal catalytic domain and a C-terminal domain that binds to the bacterial cell wall. The DV and SW lysin's N-terminal domain is a muramidase in the GH25 family and is predicted to cleave the 1,4-beta-linkages between the N-acetylmuramic acid and N-acetyl-D-glucosamine residues, which construct the peptidoglycan backbone [64]. The TH and BB endolysin's N-terminal domain is an N-acetylmuramoyl-L-alanine amidase predicted to break the amide bond between N-acetylmuramoyl and L-amino acids, which is a key component of the cross-linked peptide chains of peptidoglycan [64]. Both genes have a C-terminal LysM domain which is predicted to bind to peptidoglycan.

MUSCLE alignment of the lysin and endolysin genes reveals low homology between them in the N-terminal region (Figure 5). The N-terminal domains of the lysin and endolysin (identified by the CDD as residues 10-183 in the alignment) share only 17% identity in their amino acid sequences (30 residues out of 173). The C-terminal domains of these genes (identified by the CDD as residues 261-303) share 44% identity (19 residues out of 43). Phage lysins can retain their C-terminal domains while swapping out new N-terminal domains in order to cleave different peptidoglycan bonds [19,63], and we predict that the differences between the V4 lysin and endolysin are a product of one of these recombination events.

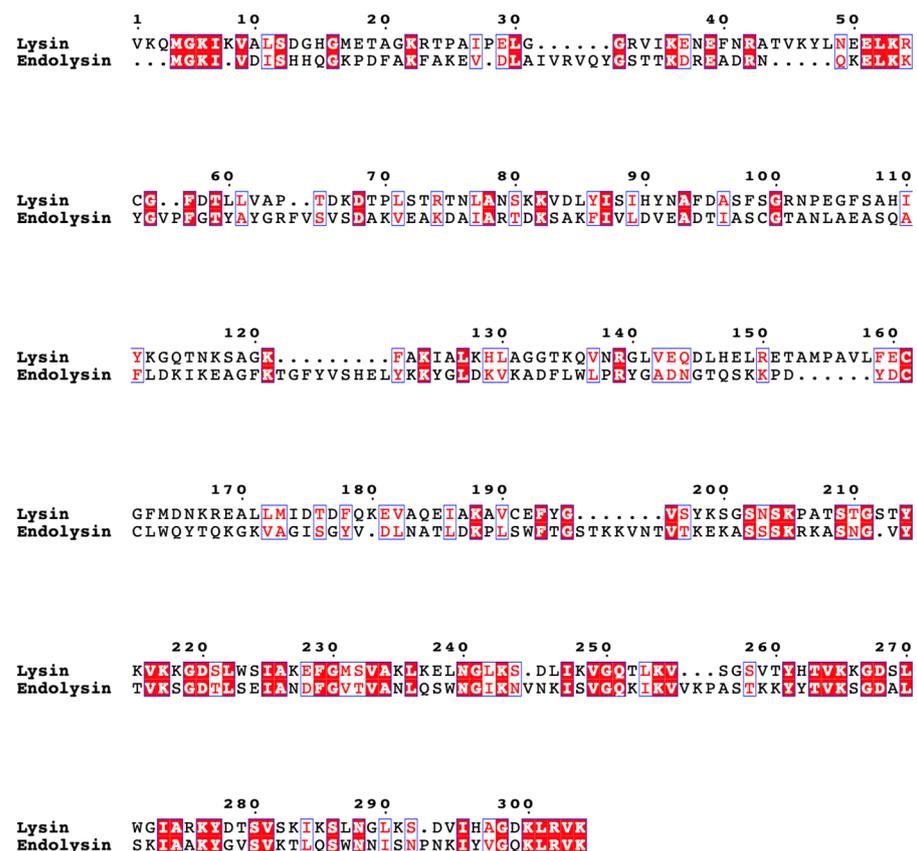


Figure 5. MUSCLE alignment of 056SW001B gene 33 (Lysin) and 268TH004 gene 33 (Endolysin), made with [52,53].

Overall, 13 other *Bacillus* phages outside of the V4 subcluster share the endolysin pham (across subclusters A5, J, X, Y2), and 33 other *Bacillus* phages (subclusters B2, B3, B4, E, G, V1, Y1, and 6 singletons which do not have cluster assignments) and 1 *Geobacillus* phage (subcluster V3, phage GBK1) share the lysin pham. There is gene synteny between the cluster V phages, as the location of the lysin is conserved between the V1 and V4 phages. Interestingly, while the endolysin has multiple BLASTp hits to other *Bacillus* phages, the

lysin's best BLASTp hits (after the V4 phages) are to various *Bacillaceae* species (coverage \leq 83%, identity \leq 73%) rather than to other phages. The differences in N-terminal domain function, low homology between the N-terminal domains, and BLASTp results suggest that the V4 phages acquired at least the lysin/endolysin N-terminal domains, if not the whole gene, from HGT.

3.2.3. Other Genes

In addition to the above two pairs of genes, there is an indel in the minor tail protein that further segregates the V4 phages. All of the V4 phages share the same pham for the minor tail protein (gene 29 in 056SW001B), but this gene is 480 bp longer in the DV and SW phages than it is in the TH and BB phages (Figure 2). The non-conserved indel region contains a domain in the carbohydrate-binding module (CBM) 4_9 family.

The WTA-tail fiber/peptidase and lysin/endolysin genes account for 4 of the 32 phams that are only present in some, but not all, of the V4 phages. Of the remaining 28 phams, only 2 have functional annotations. First, all of the V4 phages except for 268TH004 share a gene that codes for a DNA-binding domain (gene 14 in 056SW001B) nestled within the structural cassette between the tail spike and the scaffolding protein (Figure 1). Given the presence of other genes coding for DNA binding domains in the genome, including one shared by all V4 phages (gene 46 in 056SW001B), it is unclear how the lack of this pham may impact 268TH004's life cycle or fitness. Second, the BB phages all lack a gene coding for an HNH endonuclease present in the other V4 phages. This pham's location is not perfectly conserved across the other four phages. It is located in the middle of the replication cassette in the SW and TH phages (gene 40), but in the DV phages it is located at the end of the genome among genes with no known function (gene 82). Homing endonucleases are often found as free-standing genes inserted amongst areas of the genome that are otherwise conserved between related phages [65], so their presence in only some of the V4 phages is unsurprising.

The remaining 26 phams that are only present in some of the V4 phages have no known function, making it difficult to draw any conclusions about the fitness implications of their presence or absence. We do note that while all V4 phages have the same complement of genes involved in DNA metabolism (see Figure 1 for functions), several of these genes of unknown function are scattered in the replication cassette. First, after the DnaE-like DNA polymerase III (alpha) (gene 47 in 056SW001B), each V4 phage has a small gene of unknown function, but 268TH004's gene is in a different pham than the rest. The pham found in 268TH004 is also found in a cluster W3 myovirus, which was also isolated from soil in the southwest United States by our lab (phage 035JT004 from Joshua Tree National Park, CA, USA). Second, after the DnaQ-like exonuclease (DNA polymerase III subunit, gene 60 in 056SW001B), the DV and SW phages share one pham of unknown function, while the TH and BB phages share a different pham of unknown function (Figure 2). Both of these phams are only present in the V4 phages.

3.3. Host Range

We tested the host range of our phages against seven *B. subtilis* strains, including five wild strains kept by the lab and the domesticated strains 168 and W23. We chose to assess host range using a microtiter plate-based assay, where a depression or delay in the bacterial growth curve indicates that phage are killing the host. Laboratory methods of determining host range can be inconsistent or only show an incomplete measure of a phage's infective ability [4,66]. To counter this, we opted for a conservative measure of host range where we judged a phage as able to infect a particular host if phage presence depressed bacterial population growth compared to the no-phage controls in at least four out of five replicates. These observations were also complemented by spot test assays on bacterial lawns.

All of the V4 phages except for 268TH004 and 274BB002 were able to infect 168 (Table 3). All V4 phages were able to lyse W23 with the caveat that 276BB001 was able to plaque on a lawn of W23 but unable to substantially depress the bacterial population in liquid

media (Table 3). For four of the wild strains, the V4 phages largely had the same infective capability. However, only the DV and SW phages were able to lyse the fifth wild strain, T89-30 (Table 3). This split follows the same division in the V4 phages as the long/short minor tail protein, WTA-tail fiber/peptidase, and lysin/endolysin phams described above.

Table 3. Host range data for the V4 phages on *B. subtilis* hosts. Y = the phage can infect the host, N = the phage cannot infect the host. Y* = the effect of the phage on bacteria growth was small but noticeable. N** = spot testing data suggest the phage can lyse the host on solid media, but this was undetectable in liquid assay.

| Phage | Bacterial Strain | | | | | | |
|-----------|------------------|-----|-------|--------|--------|--------|--------|
| | Domesticated | | | Wild | | | |
| | 168 | W23 | TG115 | T89-05 | T89-06 | T89-20 | T89-30 |
| 019DV002 | Y* | Y | N | Y | Y | Y | Y |
| 019DV004 | Y* | Y | N | Y | Y | Y | N** |
| 056SW001B | Y | Y | N | Y | Y | Y | Y |
| 268TH004 | N | Y | N | Y | Y | Y | N |
| 274BB002 | N | Y | N | Y | Y | Y | N |
| 276BB001 | Y | N** | N | Y | N** | Y | N |
| 280BB001 | Y | Y | N | Y | Y | Y | N |

4. Discussion

Here, we described seven recently isolated *Bacillus subtilis* phages from the southwest United States. Given that relatively few *Bacillus* phages have been described [20], it is not surprising that isolating these phages allowed us to describe a novel subcluster. More surprisingly, we were able to identify highly related phages belonging to the same subcluster across a broad geographic range, as some of our collection sites were ~1500 km apart. Microbial dispersal is of great ecological significance [67], and the isolation of three nearly identical phages 14–35 km apart underscores the dispersal capacity of phages.

Despite the high ANI/AAI among these phages, only two thirds of the phams that we identified are conserved across all seven phages. The non-conserved phams tend to be small and have no known function, which is a common finding in other phages as well [15,17]. Some studies suggest that the gene products of small, non-conserved phage genes may act as a “molecular splint” and bind to host proteins to inhibit or modify their function, as has been described with some small T4 genes [15], and that may well be the case with some of these V4 phams. Further advances in protein modeling or the crystallization of these small proteins may give additional insight into the putative functions of these genes. Comparative analyses may also identify gene functions as more phage and host bacteria genomes are sequenced and explored.

Our data provide additional evidence that phage genomes are mosaic with HGT as a key player in phage genome evolution [16,17,30–34]. The V4 phages largely differ by the presence/absence of small genes or regions, which is a common finding in phage genomes. It is possible that these regions were acquired or shared by the V4 phages through HGT, although further computational studies would be required to determine their origin. We also observed two sets of functional genes that were acquired either entirely or modularly through HGT, which were notable given their large size.

The division in the V4 phages between the WTA-tail fiber/peptidase and lysin/endolysin phams, as well as the indel of a functional domain in the minor tail protein, all suggest that these phages are adapting to differences in the host cell walls that they encounter. The differences in the minor tail protein and the tail fibers suggest that these phages may bind to different cell wall substrates, although our comparative analyses cannot be conclusive in this regard. The division between the lysin and endolysin genes

suggests that the phages use different strategies for lysing host cell walls. We do note that the enzymatic domains contained in these phams are very common in phages [19] and break very conserved and common bonds in peptidoglycan [68]. Nevertheless, the fact that phages with the long minor tail, WTA-tail fiber, and lysin gene are unanimously able to infect host T89-30 while the other V4 phages cannot indicates that these genes do play a role in determining host range. Perhaps these genetic differences speak to the changing availability of certain hosts between California, Arizona, and Texas, or a shift in the structure of host cell walls due to differing abiotic conditions across that geographical range. We would need a better understanding of how *B. subtilis* cell walls vary by strain and environment to fully understand how the different V4 tail fiber and lysin gene products impact V4 phage ecology. Additionally, further work to characterize host genomes, such as by determining the presence/absence of certain phage receptors or differences in cell wall composition (e.g., type of cell wall teichoic acids), will enhance our framework for understanding V4 phage ecology and evolution.

Phage diversity remains the “dark matter” of the biosphere with relatively very few sequenced phage genomes compared to the number of sequenced bacterial genomes [13]. As the phage community continues to identify and describe new phage isolates, we will undoubtedly characterize additional novel phage clusters and gain a fuller picture of bacterial and viral ecology [14]. Integrating host range data with genomic data may help us to map additional gene functions. Future studies comparing data generated by different host range methods would aid us in linking genomic features to functional host range effects especially when combined with studies determining at what stage of the life cycle (e.g., restriction, adaptive immunity, abortive infection) some phage/strain combinations are prevented from productive infection. Once a sufficient number of *Bacillus* strains and associated phages have been isolated and annotated, future work integrating what we know of phage and host bacteria diversity and ecology into a comprehensive review will give us much-needed insights into how phages adapt to their hosts in a natural environment.

Author Contributions: Conceptualization, R.E.L., V.A.D. and G.P.K.; methodology, V.A.D., G.P.K.; validation, V.A.D. and G.P.K.; formal analysis, R.E.L. and V.A.D.; investigation, R.E.L., V.A.D., B.E.C., M.C., A.A.G., L.G.-M., M.M., M.S.S., N.T.T., A.C.V. and G.P.K.; resources, V.A.D. and G.P.K.; data curation, R.E.L., V.A.D., B.E.C., M.C., A.A.G., L.G.-M., M.M., M.S.S., N.T.T., A.C.V. and G.P.K.; writing—original draft preparation, R.E.L. and V.A.D.; writing—review and editing, R.E.L., V.A.D., L.G.-M., M.M., M.S.S., N.T.T., A.C.V. and G.P.K.; visualization, R.E.L. and V.A.D.; supervision, V.A.D. and G.P.K.; project administration, V.A.D. and G.P.K.; funding acquisition, V.A.D. and G.P.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant to Gettysburg College from the Howard Hughes Medical Institute through the Precollege and Undergraduate Science Education Program (grant number 52007540) and by Research and Professional Development grants from Gettysburg College to V.A.D.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All genome sequences have been deposited in the GenBank database of NCBI, and accession numbers are listed in Table 1.

Acknowledgments: We thank Steve Cresawn for developing and maintaining the Phamerator software. Both Steve and Lawrence (Owen) Abad helped build the Bacillus v5 Phamerator database. We also acknowledge the peer reviewers who provided insightful and helpful comments to improve our work.

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

References

1. Koskella, B.; Brockhurst, M.A. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **2014**, *38*, 916–931. [[CrossRef](#)] [[PubMed](#)]
2. Dy, R.L.; Richter, C.; Salmond, G.P.C.; Fineran, P.C. Remarkable Mechanisms in Microbes to Resist Phage Infections. *Annu. Rev. Virol.* **2014**, *1*, 307–331. [[CrossRef](#)] [[PubMed](#)]
3. Harrison, E.; Brockhurst, M.A. Ecological and Evolutionary Benefits of Temperate Phage: What Does or Doesn't Kill You Makes You Stronger. *BioEssays* **2017**, *39*, 1700112. [[CrossRef](#)] [[PubMed](#)]
4. Koskella, B.; Hernandez, C.A.; Wheatley, R.M. Understanding the Impacts of Bacteriophage Viruses: From Laboratory Evolution to Natural Ecosystems. *Annu. Rev. Virol.* **2022**, *9*, 57–78. [[CrossRef](#)]
5. Brown, T.L.; Charity, O.J.; Adriaenssens, E.M. Ecological and functional roles of bacteriophages in contrasting environments: Marine, terrestrial and human gut. *Curr. Opin. Microbiol.* **2022**, *70*, 102229. [[CrossRef](#)]
6. Suttle, C.A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Microbiol.* **2007**, *5*, 801–812. [[CrossRef](#)]
7. Puxty, R.J.; Millard, A.D.; Evans, D.J.; Scanlan, D.J. Viruses Inhibit CO₂ Fixation in the Most Abundant Phototrophs on Earth. *Curr. Biol.* **2016**, *26*, 1585–1589. [[CrossRef](#)]
8. Howard-Varona, C.; Lindback, M.M.; Bastien, G.E.; Solonenko, N.; Zayed, A.A.; Jang, H.; Andreopoulos, B.; Brewer, H.M.; Glavina Del Rio, T.; Adkins, J.N.; et al. Phage-specific metabolic reprogramming of virocells. *ISME J.* **2020**, *14*, 881–895. [[CrossRef](#)]
9. Borodovich, T.; Shkoporov, A.N.; Ross, R.P.; Hill, C. Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. *Gastroenterol. Rep.* **2022**, *10*, goac012. [[CrossRef](#)]
10. Ma, Y.; You, X.; Mai, G.; Tokuyasu, T.; Liu, C. A human gut phage catalog correlates the gut phageome with type 2 diabetes. *Microbiome* **2018**, *6*, 24. [[CrossRef](#)]
11. Norman, J.M.; Handley, S.A.; Baldrige, M.T.; Droit, L.; Liu, C.Y.; Keller, B.C.; Kambal, A.; Monaco, C.L.; Zhao, G.; Fleshner, P.; et al. Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. *Cell* **2015**, *160*, 447–460. [[CrossRef](#)] [[PubMed](#)]
12. Kortright, K.E.; Chan, B.K.; Koff, J.L.; Turner, P.E. Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria. *Cell Host Microbe* **2019**, *25*, 219–232. [[CrossRef](#)] [[PubMed](#)]
13. Hatfull, G.F. Dark Matter of the Biosphere: The Amazing World of Bacteriophage Diversity. *J. Virol.* **2015**, *89*, 8107–8110. [[CrossRef](#)] [[PubMed](#)]
14. Bibby, K. Improved Bacteriophage Genome Data is Necessary for Integrating Viral and Bacterial Ecology. *Microb. Ecol.* **2014**, *67*, 242–244. [[CrossRef](#)]
15. Comeau, A.M.; Hatfull, G.F.; Krisch, H.M.; Lindell, D.; Mann, N.H.; Prangishvili, D. Exploring the prokaryotic virosphere. *Res. Microbiol.* **2008**, *159*, 306–313. [[CrossRef](#)]
16. Hatfull, G.F. Bacteriophage genomics. *Curr. Opin. Microbiol.* **2008**, *11*, 447–453. [[CrossRef](#)]
17. Hatfull, G.F.; Jacobs-Sera, D.; Lawrence, J.G.; Pope, W.H.; Russell, D.A.; Ko, C.-C.; Weber, R.J.; Patel, M.C.; Germane, K.L.; Edgar, R.H.; et al. Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition, and Gene Size. *J. Mol. Biol.* **2010**, *397*, 119–143. [[CrossRef](#)]
18. Hatfull, G.F.; Pedulla, M.L.; Jacobs-Sera, D.; Cichon, P.M.; Foley, A.; Ford, M.E.; Gonda, R.M.; Houtz, J.M.; Hryckowian, A.J.; Kelchner, V.A.; et al. Exploring the Mycobacteriophage Metaproteome: Phage Genomics as an Educational Platform. *PLoS Genet.* **2006**, *2*, e92. [[CrossRef](#)]
19. Payne, K.M.; Hatfull, G.F. Mycobacteriophage Endolysins: Diverse and Modular Enzymes with Multiple Catalytic Activities. *PLoS ONE* **2012**, *7*, e34052. [[CrossRef](#)]
20. Delesalle, V.A.; Tomko, B.E.; Vill, A.C.; Lichty, K.B.; Krukoni, G.P. Forty Years without Family: Three Novel Bacteriophages with High Similarity to SPP1 Reveal Decades of Evolutionary Stasis since the Isolation of Their Famous Relative. *Viruses* **2022**, *14*, 2106. [[CrossRef](#)]
21. Vill, A.C.; Delesalle, V.A.; Tomko, B.E.; Lichty, K.B.; Strine, M.S.; Guffey, A.A.; Burton, E.A.; Tanke, N.T.; Krukoni, G.P. Comparative Genomics of Six Lytic *Bacillus subtilis* Phages from the Southwest United States. *PHAGE* **2022**, *3*, 171–178. [[CrossRef](#)] [[PubMed](#)]
22. Krukoni, G.P.; Kemp, A.K.; Storrie, K.F.; Chavira, V.R.; Lantrip, H.W.; Perez, V.D.; Reyes, D.A.; Truax, J.A.; Loney, R.; Delesalle, V.A. Complete Genome Sequences of Two Temperate *Bacillus subtilis* Phages Isolated at Tumamoc Hill Desert Laboratory. *Microbiol. Resour. Announc.* **2022**, *11*, e00455-22. [[CrossRef](#)] [[PubMed](#)]
23. Earl, A.M.; Losick, R.; Kolter, R. Ecology and genomics of *Bacillus subtilis*. *Trends Microbiol.* **2008**, *16*, 269–275. [[CrossRef](#)] [[PubMed](#)]
24. Chanal, A.; Chapon, V.; Benzerara, K.; Barakat, M.; Christen, R.; Achouak, W.; Barras, F.; Heulin, T. The desert of Tataouine: An extreme environment that hosts a wide diversity of microorganisms and radiotolerant bacteria. *Environ. Microbiol.* **2006**, *8*, 514–525. [[CrossRef](#)] [[PubMed](#)]
25. Prestel, E.; Salamitou, S.; DuBow, M.S. An examination of the bacteriophages and bacteria of the Namib desert. *J. Microbiol.* **2008**, *46*, 364–372. [[CrossRef](#)] [[PubMed](#)]
26. Prestel, E.; Regeard, C.; Salamitou, S.; Neveu, J.; DuBow, M.S. The bacteria and bacteriophages from a Mesquite Flats site of the Death Valley desert. *Antonie Van Leeuwenhoek* **2013**, *103*, 1329–1341. [[CrossRef](#)]

27. Radhakrishnan, R.; Hashem, A.; Abd_Allah, E.F. Bacillus: A Biological Tool for Crop Improvement through Bio-Molecular Changes in Adverse Environments. *Front. Physiol.* **2017**, *8*, 667. [CrossRef]
28. Viju, N.; Punitha, S.M.J.; Satheesh, S. Antibiofilm activity of symbiotic Bacillus species associated with marine gastropods. *Ann. Microbiol.* **2020**, *70*, 11. [CrossRef]
29. Melara, E.G.; Avellaneda, M.C.; Valdivi , M.; Garc a-Hern andez, Y.; Aroche, R.; Mart nez, Y. Probiotics: Symbiotic Relationship with the Animal Host. *Animals* **2022**, *12*, 719. [CrossRef]
30. Lawrence, J.G.; Hatfull, G.F.; Hendrix, R.W. Imbroglios of Viral Taxonomy: Genetic Exchange and Failings of Phenetic Approaches. *J. Bacteriol.* **2002**, *184*, 4891–4905. [CrossRef]
31. Pedulla, M.L.; Ford, M.E.; Houtz, J.M.; Karthikeyan, T.; Wadsworth, C.; Lewis, J.A.; Jacobs-Sera, D.; Falbo, J.; Gross, J.; Pannunzio, N.R.; et al. Origins of Highly Mosaic Mycobacteriophage Genomes. *Cell* **2003**, *113*, 171–182. [CrossRef] [PubMed]
32. Brussow, H.; Desiere, F. Comparative phage genomics and the evolution of Siphoviridae: Insights from dairy phages. *Mol. Microbiol.* **2001**, *39*, 213–223. [CrossRef] [PubMed]
33. Belcaid, M.; Bergeron, A.; Poisson, G. Mosaic Graphs and Comparative Genomics in Phage Communities. *J. Comput. Biol.* **2010**, *17*, 1315–1326. [CrossRef] [PubMed]
34. Mavrich, T.N.; Hatfull, G.F. Bacteriophage evolution differs by host, lifestyle and genome. *Nat. Microbiol.* **2017**, *2*, 17112. [CrossRef]
35. Duncan, K.E.; Ferguson, N.; Kimura, K.; Zhou, X.; Istock, C.A. Fine-scale genetic and phenotypic structure in natural populations of *Bacillus subtilis* and *Bacillus licheniformis*: Implications for bacterial evolution and speciation. *Evolution* **1994**, *48*, 2002–2025. [CrossRef]
36. Istock, C.A.; Ferguson, N.; Istock, N.L.; Duncan, K.E. Geographical diversity of genomic lineages in *Bacillus subtilis* (Ehrenberg) Cohn sensu lato. *Org. Divers. Evol.* **2001**, *1*, 179–191. [CrossRef]
37. The Actinobacteriophage Database | Protocols for Extraction. Available online: <https://phagesdb.org/workflow/Extraction/> (accessed on 19 August 2023).
38. Russell, D.A. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. In *Bacteriophages*; Clokie, M.R.J., Kropinski, A.M., Lavigne, R., Eds.; Methods in Molecular Biology; Springer: New York, NY, USA, 2018; Volume 1681, pp. 109–125. ISBN 978-1-4939-7341-5. [CrossRef]
39. Gordon, D.; Green, P. Consed: A graphical editor for next-generation sequencing. *Bioinformatics* **2013**, *29*, 2936–2937. [CrossRef]
40. Garneau, J.R.; Depardieu, F.; Fortier, L.-C.; Bikard, D.; Monot, M. PhageTerm: A tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Sci. Rep.* **2017**, *7*, 8292. [CrossRef]
41. Lawrence, J.G. DNAMaster. Available online: <http://cobamide2.bio.pitt.edu/computer.htm> (accessed on 16 September 2023).
42. Delcher, A. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **1999**, *27*, 4636–4641. [CrossRef]
43. Lukashin, A. GeneMark.hmm: New solutions for gene finding. *Nucleic Acids Res.* **1998**, *26*, 1107–1115. [CrossRef]
44. Altschu, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef] [PubMed]
45. Soding, J.; Biegert, A.; Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **2005**, *33*, W244–W248. [CrossRef] [PubMed]
46. Laslett, D. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **2004**, *32*, 11–16. [CrossRef]
47. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [CrossRef] [PubMed]
48. Cresawn, S.G.; Bogel, M.; Day, N.; Jacobs-Sera, D.; Hendrix, R.W.; Hatfull, G.F. Phamerator: A bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinform.* **2011**, *12*, 395. [CrossRef]
49. Gauthier, C.H.; Cresawn, S.G.; Hatfull, G.F. PhaMMseqs: A new pipeline for constructing phage gene families using MMseqs2. *G3 GenesGenomesGenetics* **2022**, *12*, jkac233. [CrossRef] [PubMed]
50. Rodriguez-R, L.M.; Konstantinidis, K.T. The Enveomics Collection: A Toolbox for Specialized Analyses of Microbial Genomes and Metagenomes. *PeerJ Prepr* **2016**. Available online: <https://peerj.com/preprints/1900v1> (accessed on 16 September 2023).
51. Edgar, R.C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **2004**, *5*, 113. [CrossRef]
52. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
53. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [CrossRef]
54. Xie, Y.; Wahab, L.; Gill, J. Development and Validation of a Microtiter Plate-Based Assay for Determination of Bacteriophage Host Range and Virulence. *Viruses* **2018**, *10*, 189. [CrossRef]
55. Bacillus Genetic Stock Center. Available online: <https://bgsc.org/> (accessed on 15 September 2023).
56. Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessi eres, P.; Bolotin, A.; Borchert, S.; et al. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **1997**, *390*, 249–256. [CrossRef] [PubMed]

57. Koç, C.; Xia, G.; Kühner, P.; Spinelli, S.; Roussel, A.; Cambillau, C.; Stehle, T. Structure of the host-recognition device of *Staphylococcus aureus* phage ϕ 11. *Sci. Rep.* **2016**, *6*, 27581. [[CrossRef](#)] [[PubMed](#)]
58. Neuhaus, F.C.; Baddiley, J. A Continuum of Anionic Charge: Structures and Functions of D-Alanyl-Teichoic Acids in Gram-Positive Bacteria. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 686–723. [[CrossRef](#)]
59. Brown, S.; Santa Maria, J.P.; Walker, S. Wall Teichoic Acids of Gram-Positive Bacteria. *Annu. Rev. Microbiol.* **2013**, *67*, 313–336. [[CrossRef](#)]
60. Xia, G.; Corrigan, R.M.; Winstel, V.; Goerke, C.; Gründling, A.; Peschel, A. Wall Teichoic Acid-Dependent Adsorption of Staphylococcal Siphovirus and Myovirus. *J. Bacteriol.* **2011**, *193*, 4006–4009. [[CrossRef](#)] [[PubMed](#)]
61. Baptista, C.; Santos, M.A.; São-José, C. Phage SPP1 Reversible Adsorption to *Bacillus subtilis* Cell Wall Teichoic Acids Accelerates Virus Recognition of Membrane Receptor YueB. *J. Bacteriol.* **2008**, *190*, 4989–4996. [[CrossRef](#)]
62. Schulz, E.C.; Ficner, R. Knitting and snipping: Chaperones in β -helix folding. *Curr. Opin. Struct. Biol.* **2011**, *21*, 232–239. [[CrossRef](#)]
63. Fischetti, V.A. Bacteriophage lysins as effective antibacterials. *Curr. Opin. Microbiol.* **2008**, *11*, 393–400. [[CrossRef](#)]
64. Vollmer, W.; Blanot, D.; De Pedro, M.A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **2008**, *32*, 149–167. [[CrossRef](#)]
65. Edgell, D.R.; Gibb, E.A.; Belfort, M. Mobile DNA elements in T4 and related phages. *Viol. J.* **2010**, *7*, 290. [[CrossRef](#)] [[PubMed](#)]
66. Hyman, P.; Abedon, S.T. Bacteriophage Host Range and Bacterial Resistance. In *Advances in Applied Microbiology*; Elsevier: Amsterdam, The Netherlands, 2010; Volume 70, pp. 217–248. ISBN 978-0-12-380991-9. [[CrossRef](#)]
67. Choudoir, M.J.; DeAngelis, K.M. A framework for integrating microbial dispersal modes into soil ecosystem ecology. *iScience* **2022**, *25*, 103887. [[CrossRef](#)] [[PubMed](#)]
68. Smith, T.J.; Blackman, S.A.; Foster, S.J. Autolysins of *Bacillus subtilis*: Multiple enzymes with multiple functions. *Microbiology* **2000**, *146*, 249–262. [[CrossRef](#)] [[PubMed](#)]

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