

Article

Molecular Insights into Bacteriophage Evolution toward Its Host

Marina de Leeuw^{1,2,†} , Maayan Baron^{1,†}, Oshrit Ben David¹ and Ariel Kushmaro^{1,3,*} 

¹ Avram and Stella Goldstein-Goren Department of Biotechnology Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Be'er Sheva 8410501, Israel; deleeuw@post.bgu.ac.il (M.d.L.); m2baron@health.ucsd.edu (M.B.); oshritbd@bgu.ac.il (O.B.D.)

² The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

³ The Ilse Katz Center for Meso and Nanoscale Science and Technology, Ben-Gurion University of the Negev, Be'er Sheva 8410501, Israel

* Correspondence: arielkus@bgu.ac.il

† These authors contributed equally to this work.

Received: 31 July 2020; Accepted: 29 September 2020; Published: 6 October 2020



Abstract: Bacteriophages (phages), viruses that infect bacteria, are considered to be highly host-specific. To add to the knowledge about the evolution and development of bacteriophage speciation toward its host, we conducted a 21-day experiment with the broad host-range bacteriophage *Aquamicrobium* phage P14. We incubated the phage, which was previously isolated and enriched with the Alphaproteobacteria *Aquamicrobium* H14, with the Betaproteobacteria *Alcaligenaceae* H5. During the experiment, we observed an increase in the phage's predation efficacy towards *Alcaligenaceae* H5. Furthermore, genome analysis and the comparison of the bacteriophage's whole genome indicated that rather than being scattered evenly along the genome, mutations occur in specific regions. In total, 67% of the mutations with a frequency higher than 30% were located in genes that encode tail proteins, which are essential for host recognition and attachment. As control, we incubated the phage with the Alphaproteobacteria *Aquamicrobium* H8. In both experiments, most of the mutations appeared in the gene encoding the tail fiber protein. However, mutations in the gene encoding the tail tubular protein B were only observed when the phage was incubated with *Alcaligenaceae* H5. This highlights the phage's tail as a key player in its adaptation to different hosts. We conclude that mutations in the phage's genome were mainly located in tail-related regions. Further investigation is needed to fully characterize the adaptation mechanisms of the *Aquamicrobium* phage P14.

Keywords: bacteriophage; bacteria; evolution; mutation; speciation; tubular protein B

1. Introduction

Bacteriophages, or phages, are viruses that infect bacteria [1]. Consequently, phages have a marked influence on microbial populations [2–4]. Several studies in recent years have investigated the effects of phages on their host populations in natural environments [5,6], in the food industry [7–11], in agriculture [12], for medical research [13] and in wastewater treatment processes [14–18]. The high level of interest in phages is due to the potential of phage research not only to answer basic evolutionary questions, but also to lead to practical applications, e.g., fighting pathogens [7,8,19] and improving industrial processes [1,3,20].

Phage–host interaction is exceptional compared to other predator–prey associations. Phages are considered highly specific to their hosts, and their replication depends on the host cell's machinery [21]. The first stage in the phage's replication cycle is its attachment to the bacterium's cell surface,

during which a receptor on the tail or capsid of the phage recognizes receptor molecules on the host cell's surface [22]. The ability of a phage to recognize and attach to those receptors is one of the factors that influences its host range [23]. After phage attachment to the bacterial host, it injects its genetic material into the cell. The subsequent stages in phage replication depend on the specific infection pathway [24].

To resist phage predation, bacteria exploit defense mechanisms that interfere with the phage adsorption or replication steps of the infection process [25]. Concurrently, the phage evolves novel mechanisms to enter the host's cell and reproduce [26–28]. Both bacteria and phage undergo genetic changes that are favored by increased fitness and selection. The selection pressure that acts on one of the players (bacteria or phage) triggers a corresponding selection pressure on the other [29].

Previous studies have shown that mutations in viral genomes may help shift the parasite's host range [30]. Ferris et al. [31] showed that random mutations in the host attachment gene of the RNA virus, bacteriophage $\phi 6$, enabled the phage to infect an additional strain of *Pseudomonas syringae*. Similarly, substitutions in the major capsid gene of the bacteriophage phiX174 were responsible for the phage's adaptation to *Salmonella* and its decreased ability to infect *Escherichia coli* [32]. These experiments clearly demonstrated the abilities of phages to adapt to new hosts that are members of the same family as their original host.

In this study, we examined the ability of *Aquamicrobium* phage P14—isolated previously from a full-scale membrane bioreactor used to treat chemical industry wastewater, and propagated using *Aquamicrobium* H14—to adapt to the bacterium *Alcaligenaceae* H5 (isolated from the same reactor). Changes in the efficacy of phage predation toward *Alcaligenaceae* H5 were measured during a 21-day experiment, in which *Aquamicrobium* H14 served as the reference host. In addition, we incubated the phage with *Aquamicrobium* H8 (related to *Aquamicrobium* H14 and isolated from the same reactor). Samples from both experiments at different time points were sequenced and changes in the phage genome were monitored. The results demonstrate that rather than being evenly distributed across the phage genome, the mutations tend to cluster in tail-related regions, thus leading to increased phage predation efficacy toward *Alcaligenaceae* H5.

2. Materials and Methods

2.1. System Description

Both the phage and the bacteria samples were obtained previously from a full-scale membrane bioreactor used to treat chemical industry wastewater [15]. The phage, *Aquamicrobium* phage P14 (KX660669), is a broad host-range phage that was previously isolated using the bacteria *Aquamicrobium* H14 (GQ254284) as host. In contrast to most phages, *Aquamicrobium* phage P14 has the unusual ability to infect bacteria from different classes [33]. The phage is able to form plaques when plated with Alphaproteobacteria (*Aquamicrobium* H8 (GQ254278) and *Aquamicrobium* H14) and Betaproteobacteria (*Alcaligenaceae* H5 (GQ254275), *Alcaligenaceae* H13 (GQ254283) and *Alcaligenaceae* H17 (GQ254287)).

To examine the development of host specificity, we designed a phage-bacteria experiment. For this experiment we used *Aquamicrobium* phage P14 and two host bacteria: *Alcaligenaceae* H5 and, as the control for comparing the mutations in the genome, *Aquamicrobium* H8, which belongs to the same family as *Aquamicrobium* H14.

2.1.1. Preparation

Aquamicrobium phage P14 was incubated with *Aquamicrobium* H14 for two days at 30 °C and constant shaking using 10 mL Luria–Bertani (LB) broth. The mixture of *Aquamicrobium* phage P14 with *Aquamicrobium* H14 was then filtered through a 0.22 μm Durapore PVDF membrane (Durapore® PVDF membrane, Merck Millipore, Billerica, MA, USA). In parallel, *Alcaligenaceae* H5 and *Aquamicrobium* H8 were incubated in LB broth for two days in separate test tubes (30 °C, constant shaking).

2.1.2. The Experiment

At the beginning of the experiment, 1 mL of each bacteria sample was added to 10 mL of fresh LB broth (bacteria concentration: $2.33 \times 10^7 \pm 0.39 \times 10^7$ CFU/mL for *Alcaligenaceae* H5 and $2.73 \times 10^6 \pm 1.63 \times 10^6$ CFU/mL for *Aquamicrobium* H8). Then, 100 μ L of *Aquamicrobium* phage P14 (7.72×10^8 PFU/mL, counted with *Aquamicrobium* H14) was added to each test tube that contained the *Alcaligenaceae* H5 or *Aquamicrobium* H8 bacteria. See Figure 1 for a schematic presentation of the experimental setup.

The experiment was conducted once and continued for 21 days under batch conditions (30 °C, constant shaking), during which the medium was refreshed every 3–4 days by transferring 1 mL from each test tube to a new test tube containing 10 mL fresh medium. Approximately every six days, live counts (colony forming units, CFUs) were used to assess the bacterial concentrations, while plaque forming unit (PFU) assays were performed using the top agar procedure to enumerate the phages. Throughout the experiment, 16S rRNA gene sequencing was used to verify bacterial identity and to screen for contamination.

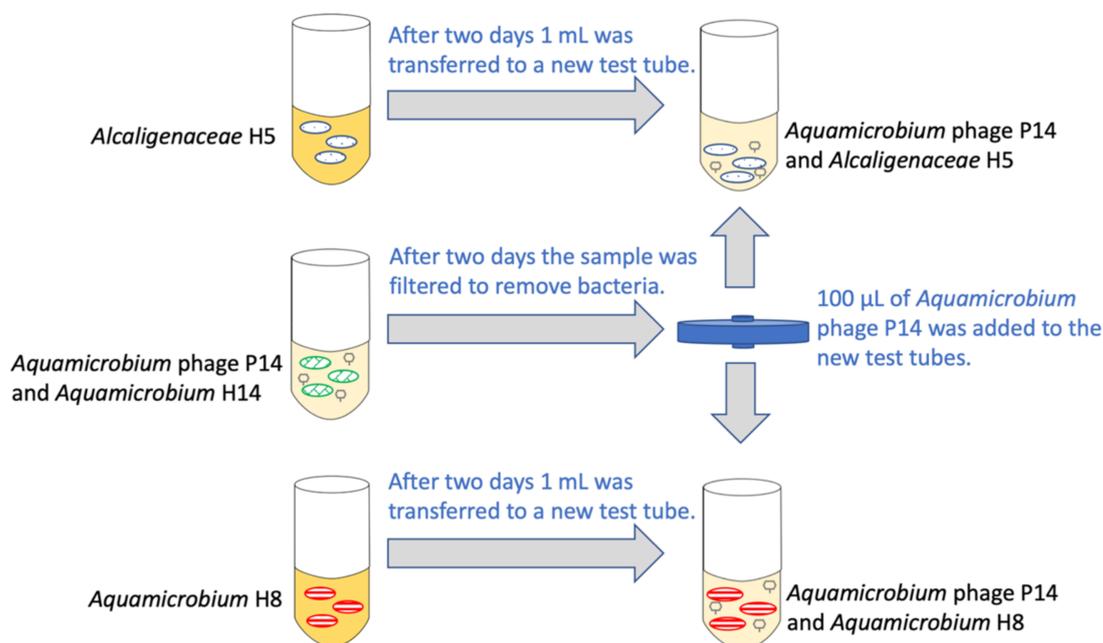


Figure 1. Schematic representation of the experimental setting.

2.2. Sampling

Samples containing both bacteria and phage were taken every three to four days and stored at -80 °C in 20% glycerol for future analysis. Phage samples were obtained by centrifuging the biomass at $10,000\times g$ for 5 min. The supernatant was filtered through a $0.22 \mu\text{m}$ Durapore PVDF membrane (Durapore® PVDF membrane, Merck Millipore, Billerica, MA, USA) and kept frozen (-80 °C) until analysis (DNA extraction and sequencing).

2.3. Quantification of Bacteriophage and Bacteria

Viable bacterial counts were determined on days 0, 6, 13 and 20 by plating 20 μ L of serial 10-fold diluted triplicate samples on LB agar plates. The plates were incubated at 30 °C for 48 h until the formation of colonies. Phages were quantified by counting PFUs on soft agar. Briefly, 100 μ L quantities of serial 10-fold diluted phage samples were mixed with either *Aquamicrobium* H14 or *Alcaligenaceae* H5 in soft LB agar (0.7% agar in LB medium, 45 °C). The soft agar was then plated on warm LB agar plates (30 °C) and the plates were incubated for 48 h at 30 °C until plaques formed. Plaque counts and morphologies were documented.

2.4. DNA Extraction and Sequencing

Samples (1.5 mL) from the test tubes were filtered through a 0.22 µm syringe filter (Durapore® PVDF membrane, Merck Millipore, Billerica, MA, USA) and then concentrated 10-fold using Amicon® Ultra-0.5 3K Centrifugal Filters (Merck KGaA, Darmstadt, Germany). The DNA was extracted from the concentrated 0.15 mL samples using the UltraClean Microbial DNA Isolation Kit by MO BIO (Carlsbad, CA, USA) following the manufacturer's protocol with two modifications: the first step of the DNA extraction, which concentrates bacterial cells by centrifugation, was omitted, and only 25 µL of elution solution was used. The purified DNA was sequenced on a MiSeq Illumina machine (San Diego, CA, USA) with 150 bp paired-end reads. The raw sequences used in this research are available in the NCBI's Sequence Read Archive (SRA) under accessions SRX827032 and SRX7412192.

2.5. DNA Sequence Analysis, Bioinformatics and Comparative Analysis

Sequence analysis was performed using the CLC Genomics Workbench 12.0. The raw paired-end sequence reads were trimmed (ambiguous limit of 2, quality score limit of 0.05), and reads shorter than 15 nucleotides were discarded. To locate variants, reads from each sample were mapped to a reference genome previously assembled for *Aquamicrobium* phage P14 [33], GenBank database [34] accession KX660669. The mapping was performed using the default parameters (length fraction 0.5, and similarity fraction 0.8) and variants were detected using the quality variant detection function, in which only those variants with frequencies higher than 10% were annotated. Only SNPs with a forward/reverse balance greater than 0.05 and an average base quality of more than 20 were used, as was done previously [35]. The gene sequences with mutations were translated into protein sequences using the translate tool ExpASy (SIB Swiss Institute of Bioinformatics) [36].

3. Results and Discussion

Aquamicrobium phage P14 is a phage belonging to the *Autographiviridae* family [33]. It has icosahedral symmetry, it measures approximately 50 nm in length, it has a short tail, and its genome is 40,551 bp long. The phage used in this study was previously isolated from an industrial wastewater treatment plant [15] and enriched using *Aquamicrobium* H14 as the host. In contrast to most phages, however, it can infect bacteria from different classes. We therefore decided to explore this ability more rigorously by incubating the *Aquamicrobium* phage P14 with *Alcaligenaceae* H5 (which belongs to a different class than *Aquamicrobium* H14) and, as a control for comparing the mutations in the genome, with *Aquamicrobium* H8 (which belongs to the same family as the bacteria that was used during the phage enrichment stage). Since the experiment was performed once, more biological replicates are needed in order to strengthen our conclusions.

3.1. Changes in Bacterial and Bacteriophage Concentrations over Time

During the experiment, CFUs and PFUs were used to assess the bacterial concentration and to enumerate the phages (Table 1). Interestingly, bacterial concentrations did not decrease significantly during the experiment. This indicates that the bacteria gained some resistance to the phages, probably as a result of coevolution occurring along the experiment [27,37]. Another possible explanation for the relatively stable bacterial concentrations could be pseudolysogeny [38–40]. However, no significant morphological changes were observed in the bacterial colonies.

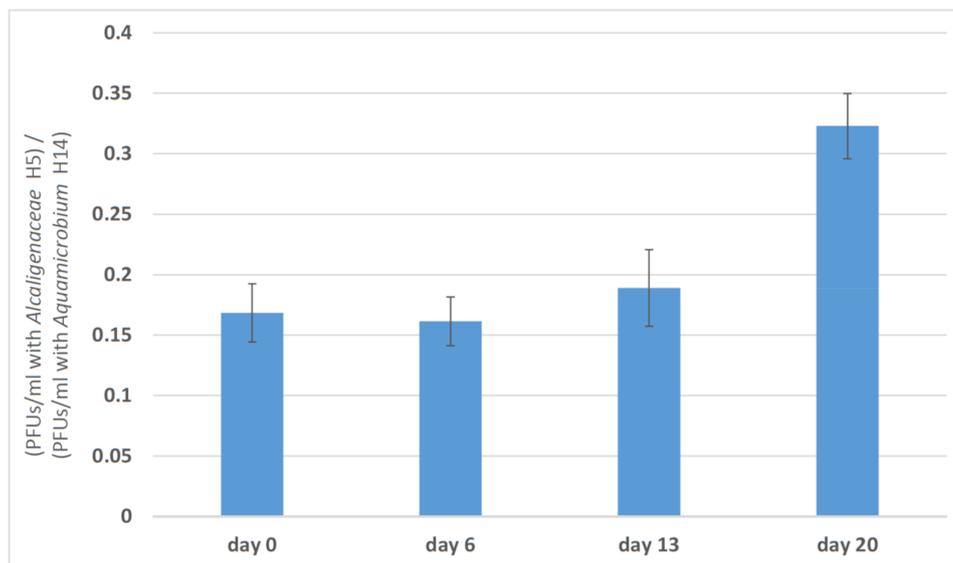
Table 1. *Alcaligenaceae* H5 and *Aquamicrobium* phage P14 concentrations during the experiment (phage concentration was determined using *Alcaligenaceae* H5 for the plaque assay).

Day	Bacteria Concentration (CFU/mL)	Phage Concentration (PFU/mL)
0	$2.3 \times 10^7 \pm 3.9 \times 10^6$	$1.3 \times 10^6 \pm 6.5 \times 10^5$
6	$1.2 \times 10^7 \pm 2.0 \times 10^6$	$1.9 \times 10^7 \pm 1.9 \times 10^6$
13	$9.7 \times 10^6 \pm 1.6 \times 10^6$	$3.1 \times 10^7 \pm 1.4 \times 10^6$
20	$2.9 \times 10^7 \pm 1.3 \times 10^6$	$1.3 \times 10^6 \pm 9.8 \times 10^5$

Curiously, the phage population remained marginally stable during the first two weeks of incubation with *Alcaligenaceae* H5. While refreshing the medium every several days could lead to a decrease in the phage concentration, the concentration calculated using *Alcaligenaceae* H5 for the plaque assay remained stable until day 13, indicating that the phages were able to propagate. Only on day 20 did the phage concentration decline; possible explanations for this may include phage inability to infect the host, in combination with natural decay and the dilution caused by the refreshment of the medium.

3.2. Changes in Phage Efficacy

Phage predation efficacy toward *Alcaligenaceae* H5, defined by the ratio of the PFUs per ml obtained for the *Alcaligenaceae* H5 to the PFU measurement obtained with the original host *Aquamicrobium* H14, increased substantially on day 20 (Figure 2). The ability of a phage to increase its efficacy toward one of its hosts following incubation with it was demonstrated earlier for ϕ X174 [32] and for bacteriophage T7 [41]. We did not determine whether the change in the infection ratio is due to an increase in phage predation efficacy toward *Alcaligenaceae* H5 or to a decrease in that efficacy toward *Aquamicrobium* H14. It should be noted, however, that even after 20 days, the phage PFU concentration enumerated by plating with *Aquamicrobium* H14 was higher than that found when the phage was plated with *Alcaligenaceae* H5.

**Figure 2.** Changes in phage efficacy during the course of the experiment. Each column represents the ratio, at different time points, between the PFUs/mL of *Aquamicrobium* phage P14 when enumerated with *Alcaligenaceae* H5 (three technical replicates) and when enumerated with the *Aquamicrobium* H14 (three technical replicates). The results are shown as the mean \pm standard deviation.

3.3. Mapping Reads to the Genome of *Aquamicrobium* Phage P14

Reads from each sample were trimmed and then mapped to the reference genome of *Aquamicrobium* phage P14, which is 40,551 bp long. Most of the samples were successfully mapped. For the *Aquamicrobium* phage P14 that was incubated with *Alcaligenaceae* H5, 637,884 reads were mapped on day 3, 174,701 reads were mapped on day 9, 230,898 reads were mapped on day 15 and 164,871 reads were mapped on day 21. For the *Aquamicrobium* phage P14 that was incubated with *Aquamicrobium* H8, 89,965 reads were mapped on day 3, 155,520 reads were mapped on day 9, 201,079 reads were mapped on day 15, and 1,513,299 reads were mapped on day 21. Unfortunately, the sample at the beginning of the experiment contained a short DNA sequence which belongs to *Aquamicrobium* H14 (accession: MW015988). Therefore, only 1183 out of more than 3.5 million reads were successfully mapped (~0.03% of the reads) to the *Aquamicrobium* phage P14. Thus, we will focus mainly on mutations which occurred after day 3, mutations that became dominant along the experiment and mutations that were detected when the phage was incubated with one bacteria but not with the other.

3.4. Mutations in the Phage Genome

The distribution of variants along the genome of *Aquamicrobium* phage P14 during incubation with *Alcaligenaceae* H5 is given in Figure 3, in which the variance frequency threshold was 10% (i.e., a minimum of 10% of the reads mapped to a specific spot had a mutation at that spot). When using a 30% variance frequency threshold, a total of 18 mutations were detected along the phage genome during the experiment (Table 2). From among these mutations, one was only detected on day 3, two were only detected on day 9 and five were only detected on day 21. All other 10 mutations were present in at least two time points.

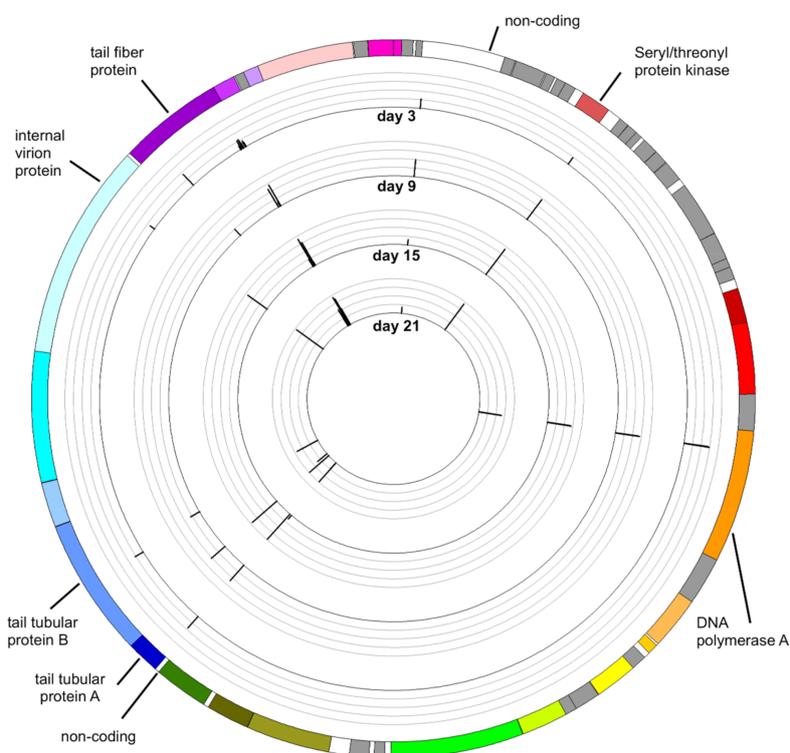


Figure 3. Variance from the reference genome of *Aquamicrobium* phage P14 over the course of the experiment. By each position at which the frequency of a mutation was higher than 10% of the mapped reads, a bar is shown to indicate the frequency (0–100%, grey line every 25%). Genes encoding hypothetical proteins are marked in grey while genes encoding proteins with an identified function are marked in color. The image was produced using Circos [42].

3.4.1. Mutation in the Gene Encoding Seryl-Threonyl Protein Kinase

One notable mutation was seen in the gene encoding seryl-threonyl protein kinase, in which two nucleotides were deleted at positions 4099 and 4100. These deletions interfered with the stop codon, and instead of a glutamic acid at the end of the protein's sequence, 20 amino acids were added to the C-terminus: VTCSINVRAWSCFCGCRTRGV. It is not clear, however, what influence, if any, this addition has. A few matches to the new C-terminus were found using blastp (protein to protein) [43]. None of them were significant, however, since the sequence is relatively short.

The mutation was observed when the *Aquamicrobium* phage P14 was incubated with *Alcaligenaceae* H5, on days 9, 15 and 21 (see Table 2). It was also observed when the phage was incubated with *Aquamicrobium* H8 (on days 3 and 9), but in this case, the frequency of the mutation was significantly lower, and on day 21 it was found in only 13.15% of the mapped reads (see Table S2). It is possible that this mutation occurred during the preparation of the experiment (from the beginning of the experiment, only four reads were mapped to this location, and one of them had the same mutation). Interestingly, the frequency of this mutation increased only during incubation with *Alcaligenaceae* H5, although this could be a coincidence and further investigation is needed.

3.4.2. Mutation in the Gene Encoding DNA Polymerase A

Two mutations were annotated in the gene encoding DNA polymerase. In both cases, a guanine was replaced by thymine (locations 11,077 and 11,086). As a result, glycine₁₃₄ and glycine₁₃₇, which have non-polar side chains, became valines, which have a hydrophobic side chain. Although such a switch may have marked effects on protein structure and function, we do not have enough data to support such a claim. Unfortunately, only two reads from the sample at the beginning of the experiment were mapped to this area, and both of them had thymine in locations 11,077 and 11,086. In addition, in both the experimental (incubation with *Alcaligenaceae* H5) and control (incubation with *Aquamicrobium* H8) samples, about two-thirds of the mapped reads had this mutation, and no significant increase or decrease was observed in the mutation's frequency over the course of the experiment.

Notably, these two mutations tend to appear together. For example, in 87% of the reads in which thymine is in position 11,077, a thymine is also in location 11,086, and 92% of the reads with a thymine in location 11,086 (same sample) also have a thymine in location 11,077. We used Phyre2 to predict the 3D structure based on the known structures of DNA polymerase proteins [44]. As can be seen in Figure 4, the affected amino acids are glycine₁₃₄ and glycine₁₃₇, both of which are replaced by a valine. According to the 3D structure prediction, these amino acids are located in a segment of a loop in which the distance between the alpha carbons of the glycines is 7.3 Å.

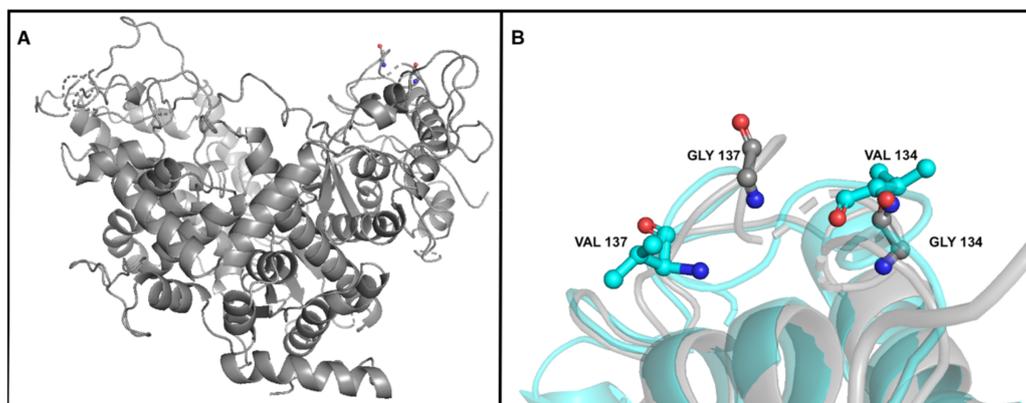


Figure 4. The predicted structure of the DNA polymerase A constructed using Phyre2 [44]. (A) The whole structure of the protein; Glycines 134 and 137 can be seen in the upper right corner. (B) Focus on amino acids 134 and 137. Grey: the structure with Glycines; Cyan: the structure with Valines.

3.4.3. Mutation in the Genes Encoding the Tail Tubular Proteins

One mutation in location 24,958 was marked in the gene encoding the tail tubular protein A. In this case, a cytosine was replaced by a thymine. This mutation caused the replacement of a proline₂₁, an amino acid with a non-polar side chain, with a leucine, which has a hydrophobic side chain. Interestingly, the frequency of this mutation changed along the timeline of the experiment. When the *Aquamicrobium* phage P14 was incubated with *Alcaligenaceae* H5 it increased from 44.52% on day 3 to 59.76% on day 9, and 98.11% on day 15. Then, it decreased on day 21 to 73.25%. However, during phage incubation with *Aquamicrobium* H8, this mutation was also observed (on days 3 (12.45%), 9 (34.76%) and 15 (79.93%)). Therefore, it is possible that the mutation already occurred during the preparation of the experiment.

Five mutations were noted during the experiment in the gene encoding the tail tubular protein B. One of them first appeared on day 9 and remained throughout the experiment (location 25,728: thymine replaced by cytosine). Each of the other four mutations was only observed once; on day 3, a mutation was annotated on location 26,835 (27.03% of mapped reads), on day 9 on location 26,976 (29.73% of mapped reads) and on day 21 mutations were annotated on locations 25,923 (35.97% of mapped reads) and 27,164 (67.64% of mapped reads). It is possible that some of these mutations were responsible for the observed changes in phage predation efficacy toward *Alcaligenaceae* H5 over the course of the experiment. The influence of the mutations (those with frequencies higher than 30%) on the protein sequence is shown in Figure 5. Remarkably, no mutations were seen in this gene when the phage was incubated with *Aquamicrobium* H8.

Tubular proteins A and B form a tubular structure surrounded by tail fibers [45]. Tubular protein A is responsible for the attachment of the fibers, while tail tubular B forms the end of the tail [45]. In addition to their structural function, recent studies suggest that the tail tubular proteins have a role in host recognition and attachment to its surface [46,47]. Furthermore, the tail tubular protein A of the *Klebsiella pneumoniae* phage KP32 was found to exhibit lytic activity towards exopolysaccharide [48]. These studies highlight the possibility that the tail tubular proteins of *Aquamicrobium* phage P14 have multiple functions, and that changes in their amino acids can influence not only the injection of viral DNA into the host, but also the attachment to the host's receptors.

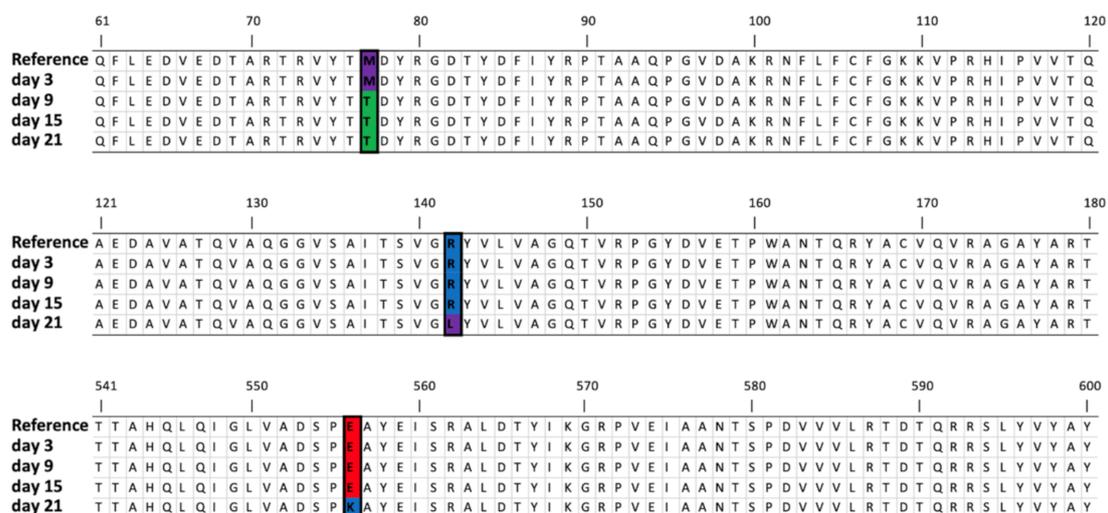


Figure 5. Amino acid sequences (61–180 and 541–600) of the tail tubular protein B and the mutations that appeared during the experiment (colored by polarity). Note: for simplicity, all mutations that occurred with more than 30% frequency at a certain time point are shown on the same protein sequence; this does not necessarily mean, however, that all the mutations occurred on the same specific protein sequence.

3.4.4. Mutations in Genes Encoding Internal Virion Proteins

Two mutations were observed in genes encoding internal virion proteins. In one mutation, marked at position 34,414, a thymine was replaced by guanine on day 3 in 15.65% of the reads. By day 21, this mutation became dominant and 95.04% of the mapped reads had a guanine in this location. This dominance suggests that the mutation gave the phage some advantage over phages that lacked it. While mutations in the internal virion protein may seem irrelevant to phage predation abilities, research conducted on bacteriophage T7 demonstrated the opposite. In the case of bacteriophage T7, the internal virion proteins are believed to play a major role in the ejection of the viral genome from the phage capsid into the bacterial cell [49]. In addition, mutations in these proteins were shown to have important implications for phage predation abilities [50].

Since only three reads were mapped to this location at the beginning of the experiment (none of them had the mutation), the possibility that this mutation occurred during phage preparation cannot be dismissed. Nonetheless, this mutation was not detected in any of the samples when the phage was incubated with *Aquamicrobium* H8. Thus, even if it occurred during phage preparation, it only became dominant when the phage was incubated with *Alcaligenaceae* H5.

A mutation in another internal virion protein (location 30,484) appeared only when the phage was incubated with *Aquamicrobium* H8. In this case, the frequency of the mutation did not increase over the course of the experiment, peaking with a frequency of 34.96% of the reads on day 3, then falling to 10.40% of the reads on day 9 and 12.88% of the reads on day 21.

3.4.5. Mutations in the Gene Encoding the Tail Fiber Protein

In total, 20 mutations were annotated in the gene encoding the tail fiber protein when *Aquamicrobium* phage P14 was incubated with *Alcaligenaceae* H5 (10% frequency threshold, Table S1). Out of them, only nine were observed in more than 30% of the mapped reads and will be discussed further.

One mutation, which was only present on day 3 (in location 35,714) and when the phage was incubated with *Aquamicrobium* H8 (the control), was probably introduced during the experiment preparation. Another mutation (in location 37,165) was present 9 days after the beginning of the experiment, but was not detected in the following samples. Three mutations emerged on day 15 and were still present at high frequencies on day 21, and another three emerged only on day 21 (see Table 2).

Remarkably, the nucleotide in location 37,092 (originally thymine) was replaced in the majority of mapped reads by guanine on day 9, by adenine on day 15 and again by guanine on day 21. Each mutation resulted in a different amino acid in location 609 of the phage's tail fiber protein. However, in both cases the Isoleucine (an amino acid with a hydrophobic side chain) was replaced by an amino acid with a positively charged side chain. On days 9 and 21, the Isoleucine was replaced by Arginine, while on day 15 it was replaced by Lysine. It is possible that a positively charged side chain in this location enhances the phage's ability to attach to *Alcaligenaceae* H5. Nevertheless, further investigation is needed in order to confirm this hypothesis.

Figure 6 shows an alignment of the translated protein sequences. All nine mutations were nonsynonymous, and eight of them influenced the translated amino acid sequence near the C-terminus. This end of the protein sequence does not have any blast [51] alignment (hits). In contrast, the N-terminus has a blastx [43] match (E-value: 1.28×10^{-11}) to the phage T7 tail fiber protein superfamily (pfam03906). This is the point at which the tail fiber protein of the bacteriophage T7 attaches to the phage's tail, while the C-terminus is the side that recognizes and attaches to the host [52].

This observation correlates well with the findings of previous studies with RNA and DNA viruses [53–56]. Tétart et al. (1996) and Le et al. (2013) demonstrated that changes in the C-terminus of the tail fiber protein lead to changes in the host range of the phage [55,56]. Thus, mutations in the C-terminus of its tail fiber protein confer on the phage the ability to attach to and infect new hosts, while the N-terminus, which attaches to the phage's tail, is conserved.

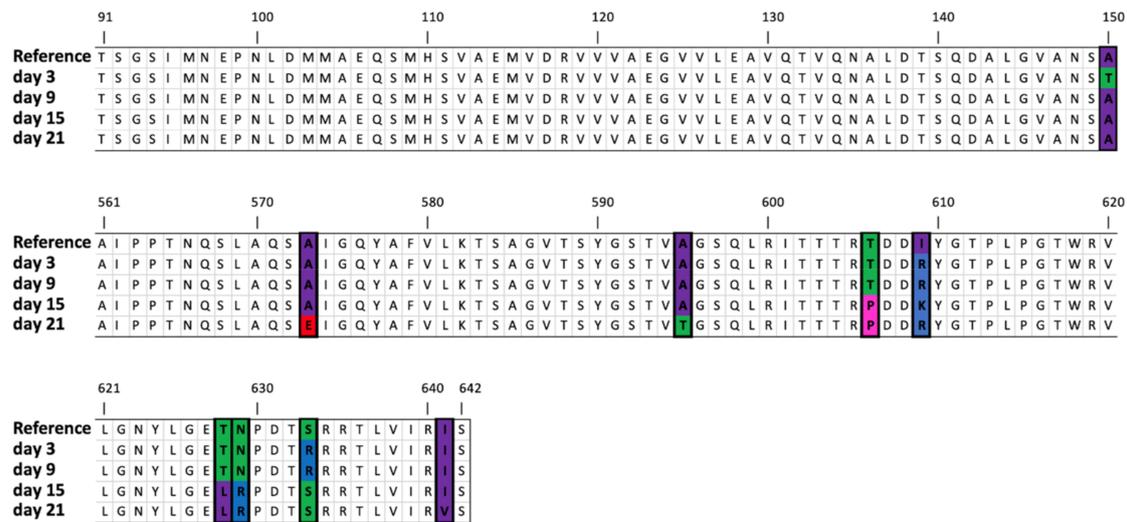


Figure 6. Amino acid sequence (91–150 and 561–642) of the tail fiber protein showing the mutations that appeared during the experiment (colored by polarity). Note: for simplicity, all mutations that occurred with more than 30% frequency at a certain time point are shown on the same protein sequence; this does not necessarily mean, however, that all the mutations occurred on the same specific protein sequence.

3.4.6. Mutations in Non-Coding Sequences

When using the 10% frequency threshold, two mutations in non-coding regions are evident (Figure 3). The mutation in location 24,876 appeared only when *Aquamicrobium* phage P14 was incubated with *Alcaligenaceae* H5 and only on day 15 (13.87% of the mapped reads). The other mutation was the insertion of a cytosine between nucleotides 570 and 571. This mutation appeared both when the phage was incubated with *Alcaligenaceae* H5 and when incubated with *Aquamicrobium* H8. Therefore, it is possible that it already occurred during the preparation of the experiment. Since the frequency of these mutations did not increase over time, it can be assumed that they did not confer on the phage any advantage.

3.4.7. Distribution of Variants along the Genome

Mutations were mainly observed in specific genes and were not evenly spread over the entire genome. Among the 15 mutations identified on day 21 (using a 30% frequency threshold), 7 were located in the gene encoding the tail fiber protein (46.7% of the mutations), which is only 1929 bp long (4.76% of the genome). Furthermore, one mutation was located in the gene encoding the tubular tail protein A, and three were identified in the gene encoding tubular tail protein B. Thus, 11 out of 15 mutations (73.3%) were located in tail-associated genes (Table 2).

To explore the tendencies of specific genes to mutate, the number of mutations observed in a gene (in the whole phage community at a certain time point) was divided by the gene's length. The average variances from the reference genome were 0.10, 0.20, 0.25 and 0.37 mutations per 1000 bp for days 3, 9, 15 and 21, respectively. The average variances for the tail fiber protein were 0.5, 1.0, 2.1 and 3.6 mutations per 1000 bp for days 3, 9, 15 and 21, respectively. Thus, the average variance for the gene encoding the tail fiber protein was significantly higher than that for the whole genome.

There are several possible explanations for this observation, and they raise important evolutionary questions. First, it is possible that while mutations are created randomly along the genome, some mutations, those that confer on the phage an advantage, become frequent. Thus, selection stress in the environment may dictate which genes show a tendency for mutation. On the other hand, it is also possible that the genes that enable the phage to shift its host range are more prone to mutation, and are not as sensitive to modification as other genes. These questions are also relevant to specific regions within a single gene, such as that of the gene encoding the tail fiber protein. In the case of this

gene, the mutations were mainly located in a region which encodes the area near the C-terminus of the tail fiber protein; an area which is used in host recognition.

The fact that most of the observed mutations occurred in tail-related genes correlates well with previously published work by Yosef et al. (2017) [57]. In their work, the researchers extended the host range of a T7 phage by designing hybrid particles that displayed various phage tail and tail fiber proteins. Similarly, the random mutations in the tail-related genes of *Aquamicrobium* phage P14 allow it to explore new infection strategies that are suited to *Alcaligenaceae* H5.

Interestingly, in the study by Crill et al. (2000) [32], ϕ X174 showed mutations in almost all of its genes as a result of switching hosts. However, more than half of the reversions (nucleotides that swapped back to the previous nucleotide due to incubation with the previous host) were spotted in the gene encoding the major capsid protein. Since ϕ X174 has no tail, the major capsid protein is responsible for phage interaction with the host cell membrane and receptors. Taken together, these findings and observations strengthen the conclusion that the phage's ability to rapidly alter the proteins responsible for host recognition allows it to adapt to new hosts.

4. Conclusions

In this study, the *Aquamicrobium* phage P14, a phage belonging to the *Autographiviridae* family, was grown with *Alcaligenaceae* H5 for 21 days. This experiment allowed us to monitor the mutations that occurred in the phage genome over time, and to observe changes in phage efficacy toward its host. Indeed, during the 21-day experiment, the phage efficacy toward the *Alcaligenaceae* H5 host increased. Furthermore, mutations in the genome were observed in specific regions, and mostly in the gene encoding the tail fiber protein, which is responsible for host recognition. The experiment was performed once and therefore more biological replicates are needed in order to confirm our conclusions. In addition, further investigation is needed to answer basic evolutionary questions about the tendency of specific genes to mutate, and the implications of those mutations for host shifts and speciation.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4915/12/10/1132/s1>, Table S1: Variants (with frequency above 10%) in the *Aquamicrobium* phage P14 genome over time when incubated with *Alcaligenaceae* H5, Table S2: Variants (with frequency above 10%) in the *Aquamicrobium* phage P14 genome over time when incubated with *Aquamicrobium* H8.

Author Contributions: Conceptualization: M.B., M.d.L. and A.K.; Formal analysis: M.d.L.; Funding acquisition: A.K.; Investigation: M.B. and O.B.D.; Writing—original draft: M.B., M.d.L., and A.K.; Writing—review and editing: M.B., M.d.L., O.B.D. and A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Deutsche Forschungsgemeinschaft: CH 731/2-1; Ne'ot Hovav Council, Israel N/A; Israeli Ministry of Science, Technology and Space 3-9724; the Kreitman Foundation N/A; Israeli Council for Higher Education N/A; the Rieger Foundation N/A; Novo Nordisk Fonden NNF10CC1016517; Novo Nordisk Fonden NNF14OC0009473.

Acknowledgments: This work was supported by the German Research Foundation DFG (grant number CH 731/2-1) and the Ne'ot Hovav Council, Israel. We would like to thank the Kreitman Foundation, the Israeli Ministry of Science, Technology and Space (3-9724), the Israeli Council for Higher Education, the Rieger Foundation and the Novo Nordisk Foundation (NNF10CC1016517; NNF14OC0009473) for their support to M.d.L.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors 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. Withey, S.; Cartmell, E.; Avery, L.M.; Stephenson, T. Bacteriophages—Potential for application in wastewater treatment processes. *Sci. Total Environ.* **2005**, *339*, 1–18. [[CrossRef](#)] [[PubMed](#)]
2. Suttle, C.A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Microbiol.* **2007**, *5*, 801–812. [[CrossRef](#)] [[PubMed](#)]
3. Johnke, J.; Cohen, Y.; de Leeuw, M.; Kushmaro, A.; Jurkevitch, E.; Chatzinotas, A. Multiple micro-predators controlling bacterial communities in the environment. *Curr. Opin. Biotechnol.* **2014**, *27*, 185–190. [[CrossRef](#)]

4. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541–548. [[CrossRef](#)]
5. Jacquet, S.; Miki, T.; Noble, R.; Peduzzi, P.; Wilhelm, S. Viruses in aquatic ecosystems: Important advancements of the last 20 years and prospects for the future in the field of microbial oceanography and limnology. *Adv. Oceanogr. Limnol.* **2010**, *1*, 97–141. [[CrossRef](#)]
6. Sime-Ngando, T. Environmental bacteriophages: Viruses of microbes in aquatic ecosystems. *Front. Microbiol.* **2014**, *5*. [[CrossRef](#)]
7. Guenther, S.; Huwlyer, D.; Richard, S.; Loessner, M.J. Virulent bacteriophage for efficient biocontrol of *Listeria monocytogenes* in ready-to-eat foods. *Appl. Environ. Microbiol.* **2009**, *75*, 93–100. [[CrossRef](#)]
8. Kocharunchitt, C.; Ross, T.; McNeil, D.L. Use of bacteriophages as biocontrol agents to control *Salmonella* associated with seed sprouts. *Int. J. Food Microbiol.* **2009**, *128*, 453–459. [[CrossRef](#)]
9. Sillankorva, S.M.; Oliveira, H.; Azeredo, J. Bacteriophages and their role in food safety. *Int. J. Microbiol.* **2012**, *2012*, 863945. [[CrossRef](#)]
10. Carlton, R.M.; Noordman, W.H.; Biswas, B.; De Meester, E.D.; Loessner, M.J. Bacteriophage P100 for control of *Listeria monocytogenes* in foods: Genome sequence, bioinformatic analyses, oral toxicity study, and application. *Regul. Toxicol. Pharmacol.* **2005**, *43*, 301–312. [[CrossRef](#)]
11. Hagens, S.; Loessner, M.J. Application of bacteriophages for detection and control of foodborne pathogens. *Appl. Microbiol. Biotechnol.* **2007**, *76*, 513–519. [[CrossRef](#)] [[PubMed](#)]
12. Bhunchoth, A.; Phironrit, N.; Leksomboon, C.; Chatchawankanphanich, O.; Kotera, S.; Narulita, E.; Kawasaki, T.; Fujie, M.; Yamada, T. Isolation of *Ralstonia solanacearum*-infecting bacteriophages from tomato fields in Chiang Mai, Thailand, and their experimental use as biocontrol agents. *J. Appl. Microbiol.* **2015**, *118*, 1023–1033. [[CrossRef](#)] [[PubMed](#)]
13. Weber-Dabrowska, B.; Jończyk-Matysiak, E.; Zaczek, M.; Łobocka, M.; Łusiak-Szelachowska, M.; Górski, A. Bacteriophage procurement for therapeutic purposes. *Front. Microbiol.* **2016**, *7*, 1177. [[CrossRef](#)] [[PubMed](#)]
14. Wu, Q.; Liu, W.T. Determination of virus abundance, diversity and distribution in a municipal wastewater treatment plant. *Water Res.* **2009**, *43*, 1101–1109. [[CrossRef](#)] [[PubMed](#)]
15. Shapiro, O.H.; Kushmaro, A.; Brenner, A. Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater. *ISME J.* **2010**, *4*, 327–336. [[CrossRef](#)] [[PubMed](#)]
16. Shapiro, O.H.; Kushmaro, A. Bacteriophage ecology in environmental biotechnology processes. *Curr. Opin. Biotechnol.* **2011**, *22*, 449–455. [[CrossRef](#)]
17. Khan, M.A.; Satoh, H.; Katayama, H.; Kurisu, F.; Mino, T. Bacteriophages isolated from activated sludge processes and their polyvalency. *Water Res.* **2002**, *36*, 3364–3370. [[CrossRef](#)]
18. Otawa, K.; Lee, S.H.; Yamazoe, A.; Onuki, M.; Satoh, H.; Mino, T. Abundance, diversity, and dynamics of viruses on microorganisms in activated sludge processes. *Microb. Ecol.* **2007**, *53*, 143–152. [[CrossRef](#)]
19. Petrovski, S.; Seviour, R.J.; Tillett, D. Characterization of the genome of the polyvalent lytic bacteriophage GTE2, which has potential for biocontrol of *Gordonia*-, *Rhodococcus*-, and *Nocardia*-stabilized foams in activated sludge plants. *Appl. Environ. Microbiol.* **2011**, *77*, 3923–3929. [[CrossRef](#)]
20. Thomas, J.A.; Soddell, J.A.; Kurtböke, D.I. Fighting foam with phages? *Water Sci. Technol.* **2002**, *46*, 511–518. [[CrossRef](#)]
21. Koskella, B.; Meaden, S. Understanding bacteriophage specificity in natural microbial communities. *Viruses* **2013**, *5*, 806–823. [[CrossRef](#)] [[PubMed](#)]
22. Baranowski, E.; Ruiz-Jarabo, C.M.; Domingo, E. Evolution of cell recognition by viruses. *Science* **2001**, *292*, 1102–1105. [[CrossRef](#)] [[PubMed](#)]
23. Hyman, P.; Abedon, S.T. Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.* **2010**, *70*, 217–248. [[PubMed](#)]
24. Weinbauer, M.G. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* **2004**, *28*, 127–181. [[CrossRef](#)] [[PubMed](#)]
25. Labrie, S.J.; Samson, J.E.; Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **2010**, *8*, 317–327. [[CrossRef](#)] [[PubMed](#)]
26. Buckling, A.; Rainey, P.B. Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. R. Soc. B Biol. Sci.* **2002**, *269*, 931–936. [[CrossRef](#)]
27. Gómez, P.; Buckling, A. Bacteria-phage antagonistic coevolution in soil. *Science* **2011**, *332*, 106–109. [[CrossRef](#)]

28. Pal, C.; Maciá, M.D.; Oliver, A.; Schachar, I.; Buckling, A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **2007**, *450*, 1079–1081. [[CrossRef](#)]
29. Woolhouse, M.E.J.; Webster, J.P.; Domingo, E.; Charlesworth, B.; Levin, B.R. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat. Genet.* **2002**, *32*, 569–577. [[CrossRef](#)]
30. Longdon, B.; Brockhurst, M.A.; Russell, C.A.; Welch, J.J.; Jiggins, F.M. The evolution and genetics of virus host shifts. *PLoS Pathog.* **2014**, *10*, e1004395. [[CrossRef](#)]
31. Ferris, M.T.; Joyce, P.; Burch, C.L. High frequency of mutations that expand the host range of an RNA virus. *Genetics* **2007**, *176*, 1013–1022. [[CrossRef](#)] [[PubMed](#)]
32. Crill, W.D.; Wichman, H.A.; Bull, J.J. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **2000**, *154*, 27–37. [[PubMed](#)]
33. De Leeuw, M.; Baron, M.; Brenner, A.; Kushmaro, A. Genome analysis of a novel broad host range proteobacteria phage isolated from a bioreactor treating industrial wastewater. *Genes* **2017**, *8*, 40. [[CrossRef](#)] [[PubMed](#)]
34. Benson, D.A.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2015**, *43*, D30–D35. [[CrossRef](#)] [[PubMed](#)]
35. Tang, N.; San Clemente, H.; Roy, S.; Bécard, G.; Zhao, B.; Roux, C. A Survey of the gene repertoire of *Gigaspora rosea* unravels conserved features among glomeromycota for obligate biotrophy. *Front. Microbiol.* **2016**, *7*, 233. [[CrossRef](#)]
36. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [[CrossRef](#)]
37. 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](#)]
38. Ripp, S.; Miller, R.V. The role of pseudolysogeny in bacteriophage-host interactions in a natural freshwater environment. *Microbiology* **1997**, *143*, 2065–2070. [[CrossRef](#)]
39. Glowacka-Rutkowska, A.; Gozdek, A.; Empel, J.; Gawor, J.; Zuchniewicz, K.; Kozinska, A.; Debski, J.; Gromadka, R.; Lobočka, M. The ability of lytic staphylococcal podovirus vB-SauP-phiAGO1.3 to coexist in equilibrium with its host facilitates the selection of host mutants of attenuated virulence but does not preclude the phage antistaphylococcal activity in a nematode infection model. *Front. Microbiol.* **2019**, *10*, 3227.
40. Latino, L.; Midoux, C.; Hauck, Y.; Vergnaud, G.; Pourcel, C. Pseudolysogeny and sequential mutations build multiresistance to virulent bacteriophages in *Pseudomonas aeruginosa*. *Microbiology* **2016**, *162*, 748–763. [[CrossRef](#)]
41. Yosef, I.; Edgar, R.; Levy, A.; Amitai, G.; Sorek, R.; Munitz, A.; Qimron, U. Natural selection underlies apparent stress-induced mutagenesis in a bacteriophage infection model. *Nat. Microbiol.* **2016**, *1*, 16047. [[CrossRef](#)] [[PubMed](#)]
42. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645. [[CrossRef](#)] [[PubMed](#)]
43. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [[CrossRef](#)] [[PubMed](#)]
44. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)] [[PubMed](#)]
45. Cuervo, A.; Pulido-Cid, M.; Chagoyen, M.; Arranz, R.; González-García, V.A.; Garcia-Doval, C.; Castón, J.R.; Valpuesta, J.M.; Van Raaij, M.J.; Martín-Benito, J.; et al. Structural Characterization of the Bacteriophage T7 Tail Machinery. *J. Biol. Chem.* **2013**, *288*, 26290–26299. [[CrossRef](#)] [[PubMed](#)]
46. Hu, M.; Zhang, H.; Gu, D.; Ma, Y.; Zhou, X. Identification of a novel bacterial receptor that binds tail tubular proteins and mediates phage infection of *Vibrio parahaemolyticus*. *Emerg. Microbes Infect.* **2020**, *9*, 855–867. [[CrossRef](#)] [[PubMed](#)]
47. Kemp, P.; Garcia, L.R.; Molineux, I.J. Changes in bacteriophage T7 virion structure at the initiation of infection. *Virology* **2005**, *340*, 307–317. [[CrossRef](#)]
48. Pyra, A.; Brzozowska, E.; Pawlik, K.; Gamian, A.; Dauter, M.; Dauter, Z. Tail tubular protein A: A dual-function tail protein of *Klebsiella pneumoniae* bacteriophage KP. *Sci. Rep.* **2017**, *7*, 1–9. [[CrossRef](#)]
49. Hu, B.; Margolin, W.; Molineux, I.J.; Liu, J. The bacteriophage T7 virion undergoes extensive structural remodeling during infection. *Science* **2013**, *339*, 576–579. [[CrossRef](#)]

50. Chang, C.Y.; Kemp, P.; Molineux, I.J. Gp15 and gp16 cooperate in translocating bacteriophage T7 DNA into the infected cell. *Virology* **2010**, *398*, 176–186. [[CrossRef](#)]
51. Altschul, 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](#)]
52. Garcia-Doval, C.; Van Raaij, M.J. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9390–9395. [[CrossRef](#)] [[PubMed](#)]
53. De Sordi, L.; Khanna, V.; Debarbieux, L. The gut microbiota facilitates drifts in the genetic diversity and infectivity of bacterial viruses. *Cell Host Microbe* **2017**, *22*, 801–808. [[CrossRef](#)] [[PubMed](#)]
54. Duffy, S.; Burch, C.L.; Turner, P.E. Evolution of host specificity drives reproductive isolation among RNA viruses. *Evolution (N. Y.)* **2007**, *61*, 2614–2622. [[CrossRef](#)] [[PubMed](#)]
55. Tétart, F.; Repoila, F.; Monod, C.; Krisch, H.M. Bacteriophage T4 host range is expanded by duplications of a small domain of the tail fiber adhesin. *J. Mol. Biol.* **1996**, *258*, 726–731. [[CrossRef](#)] [[PubMed](#)]
56. Le, S.; He, X.; Tan, Y.; Huang, G.; Zhang, L.; Lux, R.; Shi, W.; Hu, F. Mapping the tail fiber as the receptor binding protein responsible for differential host specificity of pseudomonas aeruginosa bacteriophages PaP1 and JG. *PLoS ONE* **2013**, *8*, e68562. [[CrossRef](#)]
57. Yosef, I.; Goren, M.G.; Globus, R.; Molshanski-Mor, S.; Qimron, U. Extending the host range of bacteriophage particles for DNA transduction. *Mol. Cell* **2017**, *66*, 721–728. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).