Article

New Viruses Infecting Hyperthermophilic Bacterium *Thermus thermophilus*

Matvey Kolesnik [1], Constantine Pavlov [1], Alina Demkina [1], Aleksei Samolygo [1], Karyna Karneyeva [1], Anna Trofimova [1], Olga S. Sokolova [2,3], Andrei V. Moiseenko [3], Maria Kirsanova [1] and Konstantin Severinov [1,4,*]

1   Institute of Gene Biology Russian Academy of Sciences, 119334 Moscow, Russia; matveykolesnik@gmail.com (M.K.); annat@genebiology.ru (A.T.)
2   Faculty of Biology, MSU-BIT University, Shenzhen 518172, China
3   Faculty of Biology, Lomonosov Moscow State University, 119234 Moscow, Russia
4   Waksman Institute, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA
*   Correspondence: severik@waksman.rutgers.edu

Abstract: Highly diverse phages infecting thermophilic bacteria of the *Thermus* genus have been isolated over the years from hot springs around the world. Many of these phages are unique, rely on highly unusual developmental strategies, and encode novel enzymes. The variety of *Thermus* phages is clearly undersampled, as evidenced, for example, by a paucity of phage-matching spacers in *Thermus* CRISPR arrays. Using water samples collected from hot springs in the Kunashir Island from the Kuril archipelago and from the Tsaishi and Nokalakevi districts in the Republic of Georgia, we isolated several distinct phages infecting laboratory strains of *Thermus thermophilus*. Genomic sequence analysis of 11 phages revealed both close relatives of previously described *Thermus* phages isolated from geographically distant sites, as well as phages with very limited similarity to earlier isolates. Comparative analysis allowed us to predict several accessory phage genes whose products may be involved in host defense/interviral warfare, including a putative Type V CRISPR-*cas* system.

Keywords: bacteriophage; CRISPR-Cas system; mobile genetic elements

1. Introduction

Terrestrial hot springs are inhabited by diverse microorganisms capable of thriving at high temperatures. Thermophilic bacteria of the *Thermus* genus are widespread in hot springs with temperatures above 60 °C and neutral pH, in self-heating compost piles, as well as in industrial water heating systems [1]. Historically, representatives of the genus have been an important source of thermostable enzymes, most notably, the Taq (*Thermus aquaticus*) DNA polymerase widely used for PCR. In addition, *Thermus* proteins have been a popular object of crystallographic studies because of their robust structures [2].

A number of immunity mechanisms of *Thermus*, including multiple CRISPR-Cas [3] and restriction—modification [2] systems, and prokaryotic argonaute proteins [4] have been described. The abundance of these systems implies that there exist multiple phages and mobile genetic elements that use *Thermus* as a host. Indeed, several highly diverse *Thermus* phages have been isolated over the years. Many of these phages are unique, rely on highly unusual developmental strategies, and encode novel enzymes [5]. The variety of *Thermus* phages is clearly undersampled [6]. Additional searches for *Thermus* phages, particularly at sites which have not been sampled before, will likely result in novel/highly divergent viruses whose studies could benefit both basic science and biotechnology. Analysis of relatives of already known *Thermus* phages can also be highly informative. In the genomes of closely related bacteriophages, the conserved gene order is sometimes disrupted by highly variable regions encoding the so-called "moron genes" [7]. Numerous novel defense systems have been discovered in such regions [8]. Presumably, these defense systems contribute to local adaptations of phages to host strains and/or competing viruses.

Studies of such adaptations can be particularly interesting in the case of *Thermus* and other thermophiles whose habitats are highly fragmented and often separated from each other by very large distances.

In this work, we isolated several *Thermus* bacteriophages from hot springs in the Kunashir Island, Russia and from the South Caucasus mountains in the Republic of Georgia. Kunashir is a volcanic island belonging to the Great Kuril Chain. It is formed by four active volcanoes and has a large geothermal system including hot streams, lakes, and ponds with water temperatures ranging from 30 to 100 °C and pH varying from highly acidic (~1) to neutral (~8) [9]. The microbial communities of Kunashir hot springs were investigated previously using metabarcoding and metagenomic approaches and *Thermus* sequences were found in several hot springs with neutral water [10–12]. Yet, these studies did not assess viral diversity and, to the best of our knowledge, no phages infecting *Thermus* have been isolated on Kunashir. Another region enriched with geothermal springs is South Caucasus with a number of >60 °C hot springs located in the Republic of Georgia [13]. The microbiomes of Georgian hot springs were studied only using the metabarcoding approach and abundant *Thermus* sequences were detected [14]. Yet, no *Thermus* phages from this area are known. Using water samples collected from hot springs in Kunashir and from the Tsaishi and Nokalakevi districts in Georgia during expeditions organized in 2023, we isolated several distinct phages infecting laboratory *Thermus thermophilus* strains. Genomic sequence analysis of 11 new phages revealed both close relatives of previously described *Thermus* viruses isolated from geographically distant sites, as well as phages that could not be taxonomically placed using tools such as *tax_myPHAGE* (Table S1). Comparative analysis allowed us to predict several accessory phage genes whose products may be involved in host defense/interviral warfare, including a putative Type V CRISPR-*cas* system that may be involved in competition with *Thermus* mobile genetic elements.

2. Material and Methods

2.1. Phage Isolation

Environmental samples were placed into sterile 50 mL tubes and stored at room temperature for 3–4 days until arrival at the laboratory, where they were further stored at +4 °C. Cultures of *T. thermophilus* HB8 or HB27 strains were grown in an orbital shaker (180 rpm) at 70 °C in *Thermus* broth (TB) (8 g/L tryptone, 2 g/L yeast extract, 4 g/L NaCl, 1 mM $MgCl_2$, and 1 mM $CaCl_2$ dissolved in 'Vittel' mineral water. Sterile $MgCl_2$ and $CaCl_2$ stock solutions were added to the autoclaved medium. For cultivation on solid medium, 2% agarized TB plates were used, accompanied with 0.4% top agar when necessary. Petri dishes were incubated overnight at 65 °C.

A total of 5 mL of TB medium were inoculated with a 100 μL aliquot of overnight culture of one of the *Thermus* strains and growth proceeded until OD_{600} reached ~0.4. In total, 0.2–0.5 mL of environmental sample was added and incubation was continued overnight at 70 °C with vigorous agitation. To isolate individual phage plaques, 1 mL of enrichment culture was centrifuged at 15,000× *g* for 15 min at 4 °C, and 100 μL aliquots of supernatant were added to 15 mL of melted soft (0.4%) TB agar supplemented with 300 μL of freshly grown *T. thermophilus* HB8 or HB27 cultures (OD_{600}~0.4). Mixtures were shaken to ensure homogeneity, poured over 2% TB agar plates, and incubated overnight at 65 °C. Individual plaques were picked with pipette tips and purified by several passages on the host *Thermus* strain.

For phage propagation, host cultures grown until OD_{600}~0.2 were inoculated with a single phage plaque and incubated for 3 h. Cultures were centrifuged at 15,000× *g* for 30 min at 4 °C and supernatants were stored at 4 °C for further use. Phage titers obtained lysates that were measured by serial 10-fold dilutions drop tests on host-inoculated top agar. Lysates with titers less than 10^7 plaque forming units per mL were precipitated by overnight incubation on ice with 10% PEG-8000 and 1M NaCl with further centrifugation at 15,000× *g* for 30 min at 4 °C. The pellets were resuspended in 500 μL of SM buffer (50 mM Tris*HCl pH 7.5, 100 mM NaCl, 8 mM $MgSO_4$).

2.2. DNA Extraction

In total, 450 µL of phage lysates or concentrates were supplied with 1 µL of solutions of DNase I and RNase A (10 mg/mL) and incubated at 30 °C for 1 h. Next, 50 µL of buffer (5% SDS, 100 Tris, 10 mM EDTA, pH 8.0) and 20 µL of proteinase K (20 mg/mL) were added followed by incubation at 56 °C for 1 h. The DNA was extracted using phenol-chloroform extraction technique, precipitated with ethanol, and dissolved in nuclease-free water.

Phage DNA size and integrity were evaluated by electrophoretic separation in 1% agarose gel. Samples of genomic DNA of different sizes were digested with various Type II restriction enzymes and those showing clearly distinct digestion patterns were used for sequencing.

Genomic DNA of inoviruses was extracted from infected cells after 3 h of culture inoculation with a phage plaque. Overnight cultures were diluted 1:50 and grown until OD_{600} ~0.4 when the separate plaques were added. The DNA was extracted with Thermo Scientific GeneJET Plasmid Miniprep Kit following the standard protocol. DNA concentrations were measured by Qubit 1X dsDNA High Sensitivity (HS) assay kit, and DNA purity was evaluated by NanoDrop 2000C spectrophotometer (Thermo Fisher Scientific, Waltham, MD, USA).

2.3. NGS Library Preparation and Sequencing

Libraries were constructed from 1 µg of genomic DNA using MGI Easy PCR-Free Library Prep Set in accordance with User Manual v1.1 (MGI Tech Co., Shenzhen, China) and cleaned up with the provided DNA Clean Beads. Quality control was performed using Qubit ssDNA Assay Kit (Thermo Fisher Scientific, Waltham, MD, USA) and 4150 TapeStation System (Agilent Technologies Inc., Santa Clara, CA, USA). Barcoded libraries were sequenced using DNBSEQ-G400 (MGI Tech Co., Ltd., Shenzhen, China) in 2 × 150 bp PE mode.

2.4. Computational Analysis of Viral Sequences

Read quality was assessed with FastQC v0.12.1 [15]. Adapter removing and quality trimming was performed with fastp v0.23.4 [16]. Genomes were assembled with SPAdes v3.15.3 embedded in the shovill v1.1.0 pipeline [17]. Phage genome annotation was performed with pharokka v1.4.0 [18] with further manual curation. Protein sequences were searched with hhpred and InterProScan against PDB_mmCIF70_8_Mar and Inter-Pro databases [19,20]. To detect IN93-like prophages integrated into *Thermus* genomes, homologs of conserved IN93 major capsid proteins (MCPs) were searched in the NCBI nr protein database. Next, 20 kbp genomic regions flanking genes IN93 MCPs homologs were reannotated using the pharokka tool [18]. To determine the borders of integrated prophages, 200 bp nucleotide sequences encoded downstream of predicted transposase gene (expected to contain the attachment site) were searched with BLASTN over the entire length of the predicted prophage region to detect target duplication sites flanking the prophage.

To assign bacteriophage genomes to taxonomy groups, the tax_*myPHAGE* pipeline was used [21].

To build graphical alignments, protein sequences of bacteriophages were aligned with the blastn tool with e-value cutoff 1×10^{-6}; the alignments with annotations were rendered with the pygenomeviz package [22].

2.5. Electron Microscopy

The samples for the transmission electron microscopy were prepared with a negative staining technique. In total, 3 µL of samples were applied to the glow-discharged carbon support film (EMCN, Chicago, IL, USA) and stained with two droplets of 1% uranyl acetate for 5 and 25 s, respectively. The images were acquired with a JEM-2100 200 kV transmission electron microscope (JEOL, Akishima, Japan) equipped with an Orius camera (Gatan, Pleasanton, CA, USA).

3. Results

3.1. Sample Collection and Processing

At the Kunashir Island, samples were collected from the Valentina and Stolbovskiy springs located at the foot of the Mendeleev volcano. The Valentina springs represent a chain of small hot puddles and rills along the banks of the Valentina stream. The Stolbovskiy hot springs encompass two hot rills located close to each other and flowing into the Zmeiny stream [10]. Both Valentina and Stolbovskiy hot springs are located in a deciduous forest. Their waters are littered with fallen leaves and woods, which could serve as a substrate for heterotrophic thermophilic organisms. In Georgia, samples were collected from two hot springs that flow from artificial wells drilled in the 1970s during prospecting for coal deposits. The Tsaishi spring gives rise to a hot rill that flows through a grassy meadow into the Dzhumi river. The Nokalakevi spring includes two wells located next to each other and jointly supplying a rill flowing through a meadow into the Tekhuri river. This place is popular among local people and tourists and thus is a subject of a heavy anthropogenic load. In both Tsaishi and Nokalakevi, samples were collected from rills flowing on the ground. All samples represented mixtures of water and sediment (sand, soil, detritus). Water temperature and approximate pH were measured during sample collection and are listed in Supplementary Table S2.

To enrich for bacteriophages, aliquots of environmental samples were added to liquid cultures of *T. thermophilus* strains HB8 and HB27. These two strains differ from each other by the set of defensive systems they harbor (Supplementary Table S3) [23]. After overnight cultivation at 70 °C, serial dilutions of culture supernatants were plated on lawns of each strain. Plaques of different sizes and/or opacity were purified and used to prepare phage lysates from which DNA was isolated (see Section 2). DNA from 10 isolates with distinct restriction digestion patterns was sequenced on the DNBSEQ-G400 platform with paired reads. The genomic assemblies yielded circular contigs suggesting that they represented complete viral genomes, either circular or with terminally redundant ends. Below, we present an overview of the genomes of the newly isolated viruses.

3.2. Filamentous Phages from Kunashir and South Caucasus

Two phages from the Tsaishi and Valentina hot springs produced tiny (less than 0.5 mm) turbid plaques and did not induce noticeable lysis upon inoculation of *T. thermophilus* cultures. Cultures infected with both phages contained thin ~800 nm filamentous particles with one round and another bud-like end (marked with black arrows on Figure 1A, where a representative image of culture infected with the Tsaishi isolate is presented). Similar bud-like structures were previously observed in *Escherichia coli* filamentous phages and considered to be adsorption complexes [24,25].

The DNA extracted from cell-free supernatants of infected cultures (and, therefore, containing virions) could not be sequenced with our standard "tagmentation" procedure designed for sequencing libraries preparation from double-stranded DNA. However, cells infected with both phages contained small DNA molecules that (i) were absent from uninfected cells and (ii) could be extracted from infected cells using plasmid purification kits. We, therefore, assume that the phages have single-stranded DNA in their virions and that extractable DNA from infected host cells represents replicative forms.

Assembly of sequenced extracted circular DNA yielded circular contigs of 6773 (Tsaishi isolate) and 6680 (Valentina isolate) bp. We will refer to the corresponding phages as Zuza8 and Zuza27. The two phages demonstrate strict host specificity: Zuza8 produces plaques only on *T. thermophilus* HB8 lawns, while Zuza27—only on HB27. Alignments of phage-encoded proteins revealed that both phages are similar to each other and to previously published genomes of *T. thermophilus* phages φOH16 (Genbank ID: LC210520.1) and φOH3 (Genbank ID: NC_045425.1) isolated on Kyushu Island in Japan. The latter phage was studied in some detail and its virion components were annotated using proteometric approaches [26]. The genomes of all four phages share the same order of predicted open reading frames (ORFs). φOH3 is distinct from the other three phages due to the absence of

an IS110-family transposase gene (Figure 1B). Since in their alignable parts the φOH3 and φOH16 genomes are highly similar at the nucleotide level (ANI~97%), and the Georgian phages, which are less similar to both each other and to the Japanese phages, contain the IS110-family transposase gene, φOH3 must be derived from a φOH16-like phage through transposase gene loss.

Integrated proviruses containing the entire set of homologs of Zuza8/Zuza27/φOH16 protein coding genes are present in the genomes of several *Thermus* isolates collected around the world (Honshu Island, Japan; Yellowstone Park, USA; Tibet; China; Kunashir Island, Russia), with the endonuclease gene located at one end and the IS110-family transposase gene—at the other end (Figure S1). The same organization was previously observed in proviruses of *Neisseria* filamentous phage Nf, which is thought to rely on its IS110-family transposase for integration and excision [27]. Thus, it is likely that Zuza8, Zuza27, and φOH16 also integrate into the genomes of their hosts in the same way. Lysogenization mediated by the IS110-family transposase, however, is not responsible for the apparently chronic character of host infection, since φOH3 forms turbid plaques [26]. Like with other inoviruses, the turbid plaques are likely produced by reducing the growth of infected cells.

φOH3 infects the HB8 but not the HB27 strain [26]. The φOH3 and Zuza27 attachment proteins are homologous over their entire lengths. The N-terminal domain of the Zuza8 attachment protein could not be aligned by BLASTP to the corresponding parts of φOH3 and Zuza27 homologs (Figure 1B). However, Clustal Omega alignment revealed similarity in this region as well (Figure S2). The nature of host specificity of these viruses requires further investigation.
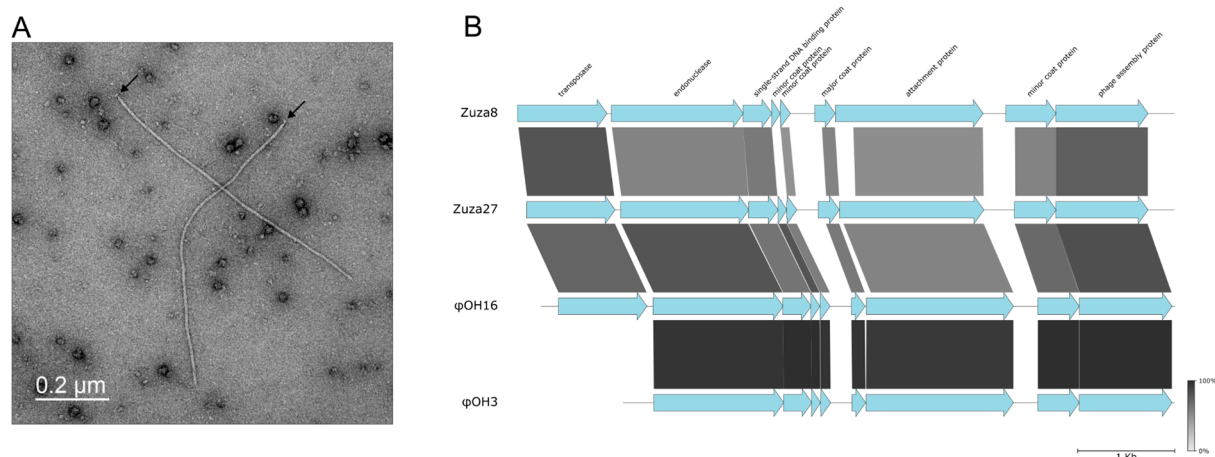

Figure 1. Filamentous phages from Tsaishi and Valentina hot springs. (**A**) Zuza27 virions in infected *T. thermophilus* HB27 cultures visualized with transmission electron microscopy with negative staining. (**B**) Graphical alignment of the genomes of Zuza8 (from Tsaishi hot spring in Georgia), Zuza27 (from Valentina hot spring in Kunashir), and φOH16/φOH3 phages from Kyushu Island in Japan. ORFs are represented with arrows that indicate the direction of transcription. Zuza8 and Zuza27 ORFs whose products' functions could be predicted by homology are colored blue; the predicted functions are listed at the top. The homologous gene segments shared between the viruses are connected with bars of different shades of gray based on the amino acid sequence identity levels shown in the bar at bottom right.

3.3. Lalka Phages—A New Group of Thermus Myoviruses

Samples from Valentina hot springs yielded three distinct lytic phages, Lalka8 (plaques on HB8 only) and Lalka27a/Lalka27b (plaque on HB27). Genomic DNA sequencing and assembly yielded circular contigs of 72,273 (Lalka8), 71,251 (Lalka27a), and 69,092 (Lalka27b) bp. Sequence analysis revealed a high level of similarity between the three genomes (ANI > 98%). Transmission electron microscopy with negative staining of lysates of *T. thermophilus* HB8 culture infected with Lalka8 revealed typical myovirus-like virions with isometric icosahe-

dral heads (~80 nm in diameter) with thick ca. 100 nm tails. Most virions were associated with spherical objects (Figure 2A, left), which likely represent extracellular membrane vesicles released by *T. thermophilus* cells [28]. In some cases, rods coming out from thick sheaths of phage virions tails were visible (Figure 2A, right).
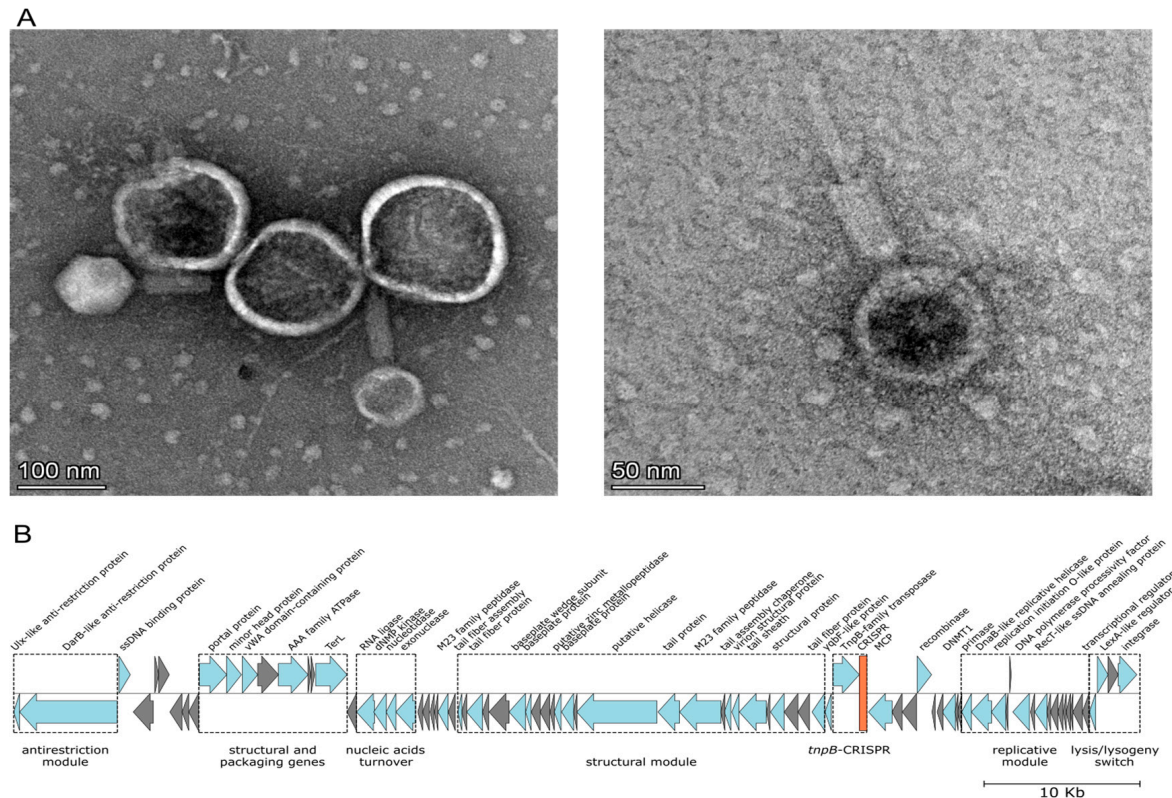

Figure 2. The Lalka8 phage from the Valentina hot spring in Kunashir. (**A**) Lalka8 virions visualized with transmission electron microscopy with negative staining. (**B**) A graphical map of the Lalka8 genome. See Figure 1B legend for details. Groups of genes encoding proteins of similar functions form modules indicated with dotted line rectangles; a putative CRISPR array is shown in red.

The virion morphology suggests that Lalka8, as well as the closely related Lalka27a and Lalka27b have linear double stranded DNA genomes. The circular topology of assembled contigs, therefore, likely results from redundant genomic termini. A search for homologous nucleotide sequences in NCBI nr/nt database (July 2024) with blastn revealed no similarity to previously isolated *Thermus* viruses except for a segment encoding a methyltransferase of phage phiMa, previously isolated from El Tatio hot springs in Chile [4]. The corresponding proteins are 72% similar (DNMT in Figure S3). At the level of amino acid sequences, some Lalka phage proteins do show 30–65% identity to several other phiMa-encoded proteins (Figure S3).

Both Lalka8 and Lalka27a encode integrase genes followed by downstream sequences matching the 3′ end segments of *Thermus* alanine tRNA gene (Figure S4). In phiMa, the integrase gene is also present, followed by a sequence matching the 3′ end segment of *Thermus* isoleucine tRNA gene. Consistently, in *T. thermophilus* SNM6-6 from Senami (Honshu Island, Japan) and in a metagenome assembled genome (MAG) of *T. antranikianii* RBS10-92 from Kemerovo, Russia, very close relatives of phiMa are integrated at the isoleucine tRNA genes (Figure S5) [29,30]. It thus follows that Lalka8 and Lalka27a are capable of integration into the genomes of their hosts using the alanine tRNA gene as an attachment site. Use of different tRNA genes as attachment sites by related phages should allow to avoid interphage competition during lysogenization [31].

Upstream of Lalka8, Lalka27a, and phiMa integrase genes, there is a co-transcribed LexA-like transcription regulator gene and a gene encoding another transcriptional regula-

tor that is transcribed in the opposite direction and that precedes a module or replication genes (Figures 2 and S3, see also below). These divergent transcription regulator genes loci resemble the lysis–lysogeny switches observed in numerous phages including λ, P2, and 186 [32,33]. The presence of LexA-like transcription regulators suggests that the lysogeny-to-lysis transition in Lalka8 and Lalka27a could be induced by cellular response associated with DNA damage. Lalka27b lacks the integrase gene and the alanine tRNA gene attachment site but encodes the LexA-like protein and the transcription regulator (Figure S3). Thus, Lalka27b should be incapable of integration into the host genome. However, we cannot exclude a possibility that, similar to *E. coli* phage P1, Lalka27b may be maintained in lysogenic cells in a plasmid form [34].

In phage P1, virion-packed linear cyclically permuted DNA with redundant termini undergoes circularization with the assessment of tyrosine site-specific recombinase Cre [35]. All three Lalka phages, but not phiMa, encode tyrosine XerD-like recombinases (Figures 2 and S3), which may perform this function.

The Lalka genomes do not encode recognizable DNA polymerases, suggesting that they rely on host enzymes for replication. However, they encode several auxiliary replication proteins, namely, a RecT-like protein, a DNA polymerase processivity factor (β-clamp), a replication initiation O-like protein, a replicative helicase, and primase. Among these, only a homolog or RecT is present in the phiMa genome. Lalka replication-related genes, along with several genes with unidentified functions, cluster downstream of a transcriptional regulator gene and likely form a single operon. Outside this module, Lalka phages encode ssDNA-binding proteins which could also participate in phage DNA replication. The products of another Lalka phages module, an exonuclease, a nucleotidase, a dNMP kinase, and an RNA ligase, are probably involved in nucleic acids turnover (Figure 2). However, the RNA ligase genes have been partially lost from the Lalka27a/Lalka27b genomes and only short ORFs homologous to the C-terminal part of the Lalka8 RNA ligase remain (Figure S3).

All three Lalka phages encode large (2074 aa) proteins that contain methyltransferase and helicase domains and are similar with the DarB protein of phage P1. Similar to P1, the genes of DarB-like protein genes are followed by short genes encoding Ulx homologs (Figures 2 and S3) [36]. The DarB and Ulx proteins are packaged in P1 virions and are required for defense against Type I restriction–modification systems of the host [36,37]. We hypothesize that DarB and Ulx homologs are also packaged in Lalka virions and overcome Type I restriction–modification systems that some *Thermus* isolates are known to harbor.

The three Lalka phages harbor short CRISPR-like arrays containing similar 29 bp repeats separated by 32–33 bp spacers. Since the consensus Lalka repeat sequence (GT-TYYYCCCGTGCGCGCATGGAGGCACCG) does not match known *Thermus* CRISPR repeat sequences [4], it is unlikely that phages rely on host Cas proteins. Curiously, the putative Lalka CRISPR arrays are located next to TnpB-like transposase genes (Figures 2B and S6). Some TnpB proteins are RNA-guided DNA nucleases [38,39]. Phylogenetic analysis suggests that TnpB proteins have been recruited as effectors of Type V CRISPR-Cas systems on numerous independent occasions [40]. The Lalka phage-encoded *tnpB*-CRISPR loci may thus be a novel thermophilic Type V CRISPR-Cas system. However, Lalka phages do not encode homologs of *cas1*/*cas2* genes which are essential for acquisition of new spacers.

Lalka8, Lalka27a, and Lalka27b encode DNA methyltransferases (DNMT), which, as mentioned above, are highly similar to the phiMa homolog. Lalka27b harbors an additional gene encoding a diverged (AAI 34% to Lalka DNMT) methyltransferase DNMTb. Lalka27a, in its turn, encodes DNMTa, which is barely related to either DNMTs or DNMTb. However, the DNMTa gene is highly similar (ANI > 90%) to an orphan DNA methyltransferase encoded in the genomes of many *T. thermophilus* isolates. In *E. coli* phage P1, viral DNA methylation mediated by either host-encoded Dam or phage-encoded Dmt methylase is essential for efficient production of infectious phage progeny [41]. Since the DNMT proteins are conserved among Lalka phages and a distant phiMa phage, they may be essential for the life cycle of these phages. In contrast, the functions of auxiliary methyltransferases

DNMTa and DNMTb are likely nonessential. These enzymes may modify phage DNA to allow evasion from restriction–modification systems of specific host strains [42].

3.4. Cumys, a Phage from Kunashir Closely Related to the phiMa Phage from Chile

Another phage from Valentina hot springs, Cumys, was isolated on *T. thermophilus* HB8. It is very close (ANI 96%) to phage phiMa. Despite the high degree of overall similarity, there is a ~1900 bp region encoding a DarTG toxin-antitoxin system in phiMa that is absent in Cumys (Figure 3A). DarTG loci are also present in phiMa-like prophages found in the genomes of some *Thermus* isolates (Figure S5). In *E. coli*, DarTG protects cells from phage infection [43]. It is thus possible that *Thermus* cells lysogenized by phiMa-like phages are also protected from infection by other phages due to *darTG* loci.

Figure 3. Phage Cumys from Valentina hot springs. (**A**) A graphical map of the Cumys genome. See Figure 1B legend for details. (**B**) Loci similar to Cumys radical SAM (rSAM) protein gene module are present in diverse phages and prophages.

Both Cumys and phiMa harbor a locus encoding a radical SAM (rSAM) protein and two proteins of unknown functions, of which one contains a DUF483 domain (Figure 3A, rSAM module). Similar loci are found in prophages of several bacterial genomes, implying a functional connection between the three genes (Figure 3B). According to the InterPro database [44], there exist fusions of DUF483 domains with methyltransferase or rSAM domains, which supports the functional association between the Cumys/phiMa rSAM module genes. Curiously, homologs of Cumys/phiMa rSAM and DUF483-domain proteins are also detected in several cyanobacterial (e.g., *Geitlerinema*) genomes in a context of genes involved in DNA modification with deazapurines [45] (Figure 3B). Although *Thermus*

phages lack genes involved in deazapurine DNA modification, the rSAM locus may be involved in the introduction of other, yet unknown, DNA modifications.

3.5. Riverbug, a Phage from Kunashir, Closely Related to a Prophage of T. brockianus SNM4-1 Isolated from Japan

The Valentina hot springs yielded another phage, named Riverbug, which infects *T. thermophilus* HB8. This phage is very close (ANI 93%) to a prophage of *T. brockianus* SNM4-1 from the Senami hot spring on Honshu Island in Japan [30]. However, the prophage contains a ~2800 bp region that is absent from the Riverbug genome. This region harbors a transposase gene. Riverbug, in its turn, contains a ~1700 bp region that is absent from the *T. brockianus* SNM4-1 prophage and encodes two genes of unknown functions (Figure 4).
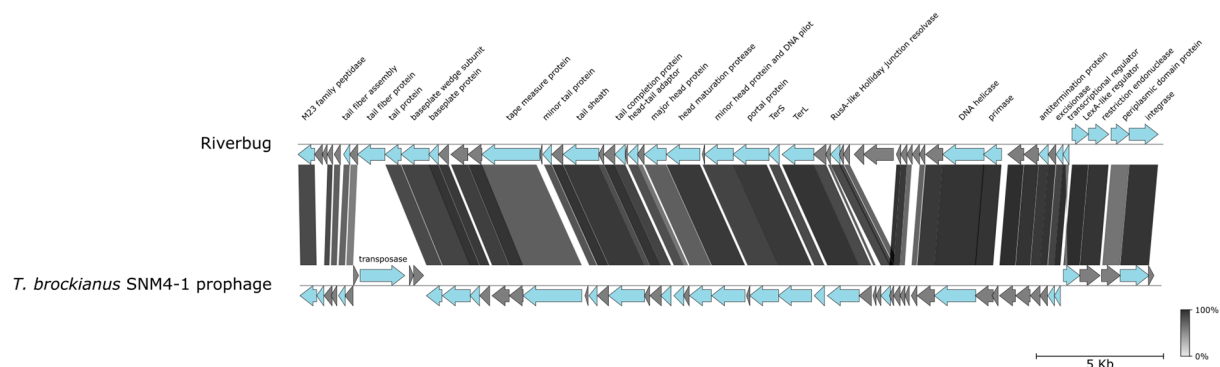
Figure 4. Phage Riverbug from Valentina hot springs. Graphical alignment of the Riverbug genome with the *T. brockianus* SNM4-1 prophage. See Figure 1B legend for details.

3.6. Three P23-45-like Bacteriophages from Stolbovskiy Hot Springs

Samples from Stolbovskiy hot springs yielded three closely related (ANI > 93%) phages, P23-45_stolb1, P23-45_stolb2, and P23-45_stolb3. These phages are similar (ANI > 85%) to a long-tailed phage P23-45 previously isolated from hot springs at the Valley of Geysers in Kamchatka peninsula [46]. The P23-45 genome is 86,364 bp long and contains two direct terminal repeats [5]; the non-redundant genomic sequence is 84,201 bp. The non-redundant genomes of P23-45_stolb1, P23-45_stolb2, and P23-45_stolb3 are shorter by ~8–10 kbps (76,044, 75,309, and 74,147 bp, respectively). Alignment between the four genomes revealed that the decreased size of Stolbovskiy phages is due to deletions in three regions of the P23-45 genome (Figure S7). In P23-45, these variable regions (VRs) contain 24 short genes, which encode proteins with unknown functions, except for host RNA polymerase inhibitor gp76 [47] located in VR1, and a cytosine methyltransferase in VR3. While gp76 homologs are conserved in all P23-45-like phages, the cytosine methyltransferase gene is absent in Stolbovskiy-derived phages, suggesting that this protein is not essential. P23-45 and P23-45-like phages from Stolbovskiy hot springs encode conserved ribonucleotide reductase proteins. In P23-45_stolb1, the ribonucleotide reductase contains a 230 aa insertion representing an intein-containing homing endonuclease.

3.7. A South Caucasus Phage Related to Phage IN93 from Japan

A phage infecting *T. thermophilus* HB8 was isolated from the Nokalakevi hot springs located ca. 30 km from Tsaishi. This phage, Tsatsa, produced large, 3–4 mm, plaques and rapidly lysed liquid cell cultures. Sequencing and assembly yielded a circular 20,815 bp contig. Extensive similarities of Tsatsa proteins to those encoded by a previously described *T. thermophilus* phage IN93 (Genbank ID: NC_004462.2) isolated on Honshu Island in Japan [48], were observed. Genes encoding homologous proteins are colinearly located in the genomes of the two phages (Figure 5). Like Tsatsa, IN93 infects the HB8 but not the HB27 *T. thermophilus* strain [49].
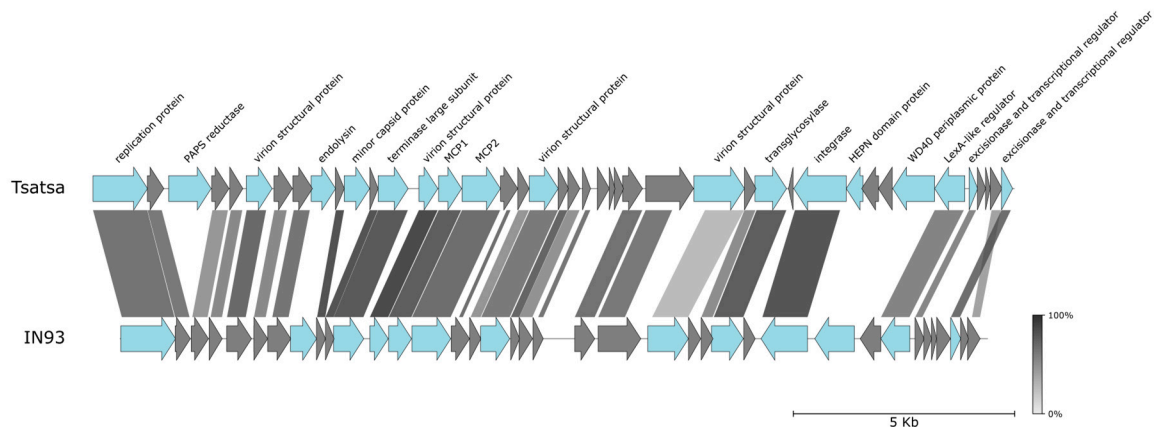
Figure 5. Phage Tsatsa from the Nokalakevi hot spring (Georgia). Graphical alignment of Tsatsa and IN93 genomes. See Figure 1B legend for details.

IN93 is a lysogenic phage; it integrates into the host genome through an attachment site located in the Ile-tRNA gene [49]. A sequence matching 45 bp of the 3′ terminal part of *T. thermophilus* Ile-tRNA is present in the Tsatsa genome immediately downstream of the integrase gene, implying that both phages integrate into the same site. Interestingly, though Tsatsa and IN93 are unrelated to phiMa-like phages, their integration sites are identical. Yet, the identity between the corresponding integrases is only ~40%.

Although Tsatsa and IN93 genomes share a collinear order of predicted protein coding genes, several differences exist. For example, in the Tsatsa genome, there is an insert between the homologs of adjacent IN93 genes *2* and *3* located close to the left end of the genome (Figure 5). The insert contains a gene encoding a protein similar to phosphoadenosine phosphosulfate (PAPS) reductases, enzymes involved in sulfur metabolism. PAPS reductase homologs were observed in various mobile genetic elements, including bacteriophage λ [50], marine bacteriophages [51], and archaeal viruses and plasmids [52–54]. A gene encoding a PAPS reductase homolog is located, in the same position as in the Tsatsa genome, in the genome of *Thermus* phage P23-77, a more distant relative of IN93 [55]. PAPS reductase homologs are also associated with Type IV-B CRISPR-*cas* [56] Type IV BREX [57], and phosphorothioation-based prokaryotic defense systems. In the latter, PAPS reductases are presumably involved in the DNA modification process [58]. Since PAPS reductases are homologous to diverse ATP pyrophosphatases, their functions may extend beyond sulfur metabolism [59]. Clearly, these interesting phage proteins deserve further investigation.

Comparison between the Tsatsa and IN93 genomes revealed a variable region located between genes encoding an integrase and a LexA-like protein (Figure 5). In IN93, this region contains two genes. The product of the first gene has no detectable homology with characterized proteins but contains a predicted N-terminal lipoprotein signal peptide. The second encodes a protein with an N-terminal domain resembling Type IV restriction endonucleases followed by a membrane-spanning segment and a predicted extracellular DUF4101 domain of unknown function. The corresponding region in Tsatsa contains four genes (Figure 6). The product of the first gene resembles periplasmic TolB proteins and contains six WD40 repeats and a predicted N-terminal signal peptide [60]. The fourth gene encodes a small (126 aa) HEPN-domain protein.

HEPN-domain nucleases are known to participate in numerous defense systems [61], while WD40 repeat proteins are involved in protein–protein interactions and signal transduction [62]. We propose that in both phages, this variable genomic region encodes defense systems. This prediction is almost certainly valid for IN93, given that Type IV restriction endonucleases are known to degrade modified DNA molecules and thus the enzyme may protect IN93 lysogens from bacteriophages with modified DNA [63]. To perform this function, the IN93 Type IV restriction endonuclease gene has to be expressed in the lysogenic state. Indeed, genes encoding the LexA-like regulator, the putative periplasmic protein,

the restriction endonuclease, and the integrase comprise an operon and are thought to be transcribed in the IN93 prophage [49].

Inspection of IN93-like proviruses retrieved from sequenced *Thermus* genomes revealed in many the presence of genes encoding restriction endonucleases, HEPN domain proteins similar to that of Tsatsa, as well as variable sets of neighboring genes encoding putative lipoproteins and proteins with predicted extracytoplasmic localization or transmembrane proteins containing predicted extracytoplasmic domains (Figure 6). We propose that the HEPN-domain nucleases are effectors of putative phage defense systems active in the lysogenic state. While the functions of other proteins are unknown, we hypothesize that they are involved in superinfection exclusion as some superinfection exclusion proteins were shown to be lipoproteins [64] or transmembrane proteins [65] expressed in the lysogenic stage. Indeed, in several phage genomes, genes encoding superinfection exclusion proteins are flanked with LexA-like repressor and integrase-encoding genes [66].

Interestingly, in a completely unrelated bacteriophage Riverbug, as well as the Riverbug-like prophage from *T. brockianus* SNM4-1, a gene encoding a protein homologous to Type IIL restriction-and-modification enzymes and another encoding a protein with predicted transmembrane and extracytoplasmic domains is also located between a LexA-like repressor and integrase genes (Figure 6). It thus follows that such modules are capable of horizontal transfer between unrelated phages.
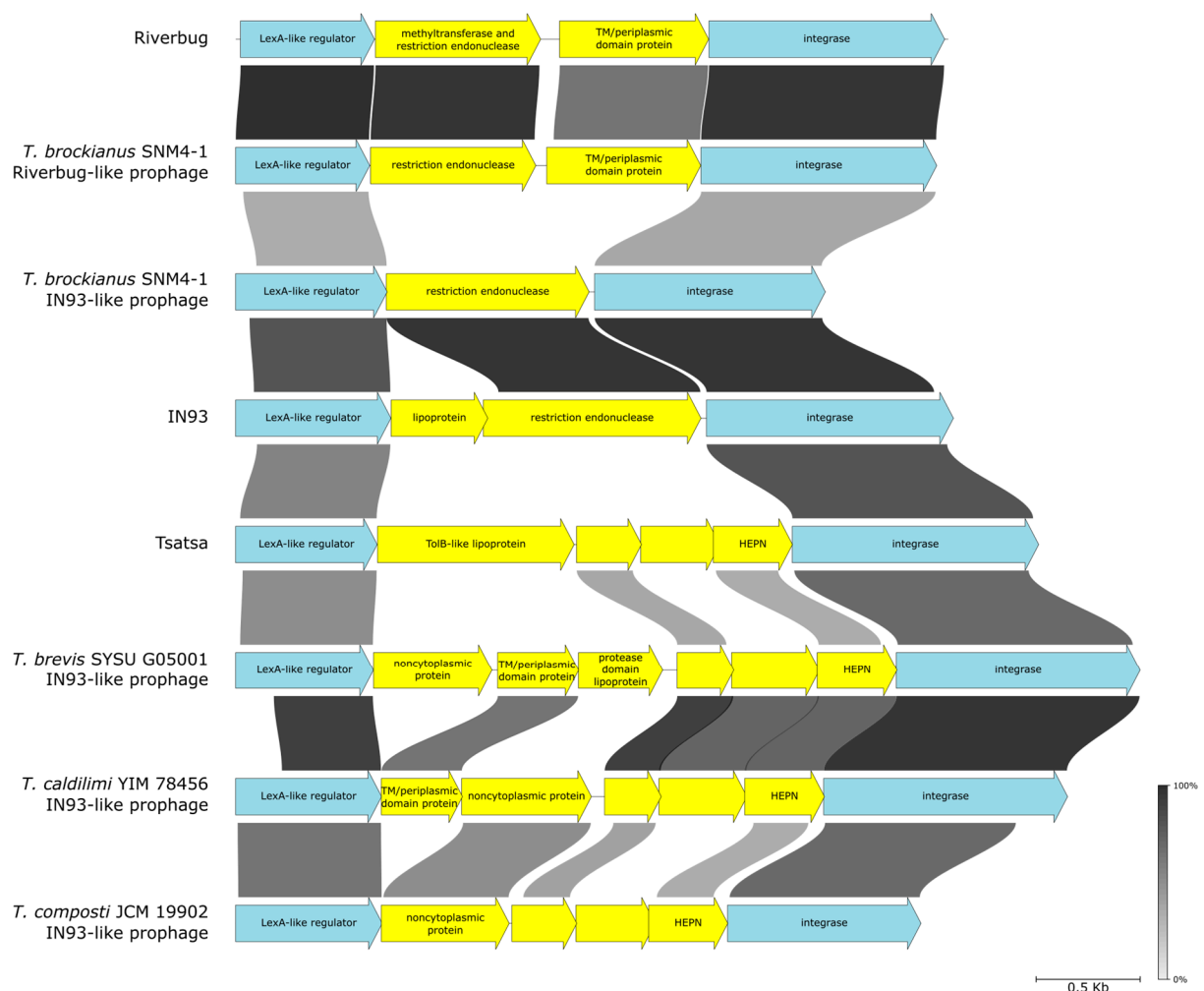
Figure 6. Putative phage-encoded defense systems located in regions likely expressed in a lysogenic state. Graphical alignment of diverse phage genome segments flanked by genes encoding a LexA-like regulator and an integrase. Genes constituting putative defense systems are colored yellow.

4. Conclusions

The impetus for this study came from a hypothesis that the variety of *Thermus* phages is undersampled. Indeed, the 10 *Thermus* phages presented here show remarkable diversity and include both novel (Lalka phages) and very close relatives of known viruses previously isolated from around the world. Phages from the latter group are globally distributed and it is interesting to determine how they spread given the highly fragmented nature of *Thermus* habitat. Even closely related viruses contain variable regions that harbor known or hypothetical defense systems. Functional analysis of these systems, including the putative novel CRISPR-Cas of Lalka phages, may provide valuable insights for both basic and applied science.

References

1. Brock, T.D.; Boylen, K.L. Presence of Thermophilic Bacteria in Laundry and Domestic Hot-Water Heaters. *Appl. Microbiol.* **1973**, *25*, 72–76. [CrossRef] [PubMed]
2. Cava, F.; Hidalgo, A.; Berenguer, J. *Thermus thermophilus* as Biological Model. *Extremophiles* **2009**, *13*, 213–231. [CrossRef]
3. Skowron, P.M.; Anton, B.P.; Czajkowska, E.; Zebrowska, J.; Sulecka, E.; Krefft, D.; Jezewska-Frackowiak, J.; Zolnierkiewicz, O.; Witkowska, M.; Morgan, R.D.; et al. The Third Restriction–Modification System from Thermus Aquaticus YT-1: Solving the Riddle of Two TaqII Specificities. *Nucleic Acids Res.* **2017**, *45*, 9005–9018. [CrossRef] [PubMed]
4. Swarts, D.C.; Jore, M.M.; Westra, E.R.; Zhu, Y.; Janssen, J.H.; Snijders, A.P.; Wang, Y.; Patel, D.J.; Berenguer, J.; Brouns, S.J.J.; et al. DNA-Guided DNA Interference by a Prokaryotic Argonaute. *Nature* **2014**, *507*, 258–261. [CrossRef] [PubMed]

5. Chaban, A.; Minakhin, L.; Goldobina, E.; Bae, B.; Hao, Y.; Borukhov, S.; Putzeys, L.; Boon, M.; Kabinger, F.; Lavigne, R.; et al. Tail-Tape-Fused Virion and Non-Virion RNA Polymerases of a Thermophilic Virus with an Extremely Long Tail. *Nat. Commun.* **2024**, *15*, 317. [CrossRef] [PubMed]

6. Lopatina, A.; Medvedeva, S.; Artamonova, D.; Kolesnik, M.; Sitnik, V.; Ispolatov, Y.; Severinov, K. Natural Diversity of CRISPR Spacers of *Thermus*: Evidence of Local Spacer Acquisition and Global Spacer Exchange. *Philos. Trans. R. Soc. B Biol. Sci.* **2019**, *374*, 20180092. [CrossRef]

7. Juhala, R.J.; Ford, M.E.; Duda, R.L.; Youlton, A.; Hatfull, G.F.; Hendrix, R.W. Genomic Sequences of Bacteriophages HK97 and HK022: Pervasive Genetic Mosaicism in the Lambdoid Bacteriophages 1 1Edited by M. Gottesman. *J. Mol. Biol.* **2000**, *299*, 27–51. [CrossRef]

8. Rousset, F.; Depardieu, F.; Miele, S.; Dowding, J.; Laval, A.-L.; Lieberman, E.; Garry, D.; Rocha, E.P.C.; Bernheim, A.; Bikard, D. Phages and Their Satellites Encode Hotspots of Antiviral Systems. *Cell Host Microbe* **2022**, *30*, 740–753.e5. [CrossRef]

9. Zharkov, R.V. Active Volcanoes and Thermal Springs of Kunashir Island (Russia). *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *459*, 022039. [CrossRef]

10. Malygina, A.; Balkin, A.; Polyakova, E.; Stefanov, S.; Potekhin, A.; Gogoleva, N. Taxonomic Diversity of the Microbial Biofilms Collected along the Thermal Streams on Kunashir Island. *Ecologies* **2023**, *4*, 106–123. [CrossRef]

11. Rozanov, A.S.; Korzhuk, A.V.; Shekhovtsov, S.V.; Vasiliev, G.V.; Peltek, S.E. Microorganisms of Two Thermal Pools on Kunashir Island, Russia. *Biology* **2021**, *10*, 924. [CrossRef] [PubMed]

12. Rozanov, A.S.; Korzhuk, A.V.; Shlyakhtun, V.N.; Shipova, A.A.; Peltek, S.E. Metagenomic Data on Prokaryotic Diversity of Kunashir Island Geothermal Spring. *Data Brief* **2020**, *32*, 106244. [CrossRef]

13. Buachidse, I.M.; Buachidse, G.I.; Shaorshadse, M.P. Thermal Waters of Georgia. *Geothermics* **1970**, *2*, 1092–1101. [CrossRef]

14. Geliashvili, N.; Jaiani, E.; Tediashvili, M.; Birkeland, N.-K. Taxonomic Characteristics of Dominant Microbial Communities in Hot Spring Sediments in Western Georgia. In *Microbial Communities and Their Interactions in the Extreme Environment*; Egamberdieva, D., Birkeland, N.-K., Li, W.-J., Panosyan, H., Eds.; Microorganisms for Sustainability; Springer: Singapore, 2021; Volume 32, pp. 87–96, ISBN 9789811637308.

15. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 2 August 2024).

16. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef] [PubMed]

17. Seemann, T. Assemble Bacterial Isolate Genomes from Illumina Paired-End Reads. Available online: https://github.com/tseemann/shovill (accessed on 2 August 2024).

18. Bouras, G.; Nepal, R.; Houtak, G.; Psaltis, A.J.; Wormald, P.-J.; Vreugde, S. Pharokka: A Fast Scalable Bacteriophage Annotation Tool. *Bioinformatics* **2023**, *39*, btac776. [CrossRef] [PubMed]

19. Zimmermann, L.; Stephens, A.; Nam, S.-Z.; Rau, D.; Kübler, J.; Lozajic, M.; Gabler, F.; Söding, J.; Lupas, A.N.; Alva, V. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at Its Core. *J. Mol. Biol.* **2018**, *430*, 2237–2243. [CrossRef] [PubMed]

20. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-Scale Protein Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240. [CrossRef]

21. Millard, A.; Denise, R.; Lestido, M.; Thomas, M.; Turner, D.; Turner, D.; Sicheritz-Pontén, T. taxmyPHAGE: Automated Taxonomy of dsDNA Phage Genomes at the Genus and Species Level. *bioRxiv* **2024**, 2024.08.09.606593. [CrossRef]

22. Shimoyama, Y. pyGenomeViz: A Genome Visualization Python Package for Comparative Genomics. Available online: https://moshi4.github.io/pyGenomeViz/ (accessed on 2 August 2024).

23. Payne, L.J.; Meaden, S.; Mestre, M.R.; Palmer, C.; Toro, N.; Fineran, P.C.; Jackson, S.A. PADLOC: A Web Server for the Identification of Antiviral Defence Systems in Microbial Genomes. *Nucleic Acids Res.* **2022**, *50*, W541–W550. [CrossRef]

24. Gray, C.W.; Brown, R.S.; Marvin, D.A. Adsorption Complex of Filamentous Fd Virus. *J. Mol. Biol.* **1981**, *146*, 621–627. [CrossRef]

25. Conners, R.; León-Quezada, R.I.; McLaren, M.; Bennett, N.J.; Daum, B.; Rakonjac, J.; Gold, V.A.M. Cryo-Electron Microscopy of the F1 Filamentous Phage Reveals Insights into Viral Infection and Assembly. *Nat. Commun.* **2023**, *14*, 2724. [CrossRef] [PubMed]

26. Nagayoshi, Y.; Kumagae, K.; Mori, K.; Tashiro, K.; Nakamura, A.; Fujino, Y.; Hiromasa, Y.; Iwamoto, T.; Kuhara, S.; Ohshima, T.; et al. Physiological Properties and Genome Structure of the Hyperthermophilic Filamentous Phage φOH3 Which Infects *Thermus thermophilus* HB8. *Front. Microbiol.* **2016**, *7*, 50. [CrossRef] [PubMed]

27. Kawai, M.; Uchiyama, I.; Kobayashi, I. Genome Comparison In Silico in Neisseria Suggests Integration of Filamentous Bacteriophages by Their Own Transposase. *DNA Res.* **2005**, *12*, 389–401. [CrossRef] [PubMed]

28. Blesa, A.; Berenguer, J. Contribution of Vesicle-Protected Extracellular DNA to Horizontal Gene Transfer in *Thermus* spp. *Int. Microbiol.* **2015**, *18*, 177–187. [CrossRef]

29. Kadnikov, V.V.; Mardanov, A.V.; Beletsky, A.V.; Karnachuk, O.V.; Ravin, N.V. Metagenomic Analysis of the Microbial Community in the Underground Coal Fire Area (Kemerovo Region, Russia) Revealed Predominance of Thermophilic Members of the Phyla Deinococcus-Thermus, Aquificae, and Firmicutes. *Microbiology* **2021**, *90*, 578–587. [CrossRef]

30. Miyazaki, K. Complete Genome Sequences of *Thermus* Strains Isolated from Senami Hot Spring in Japan. *Microbiol. Resour. Announc.* **2022**, *11*, e00354-22. [CrossRef]

31. Williams, K.P. Integration Sites for Genetic Elements in Prokaryotic tRNA and tmRNA Genes: Sublocation Preference of Integrase Subfamilies. *Nucleic Acids Res.* **2002**, *30*, 866–875. [CrossRef]

32. Christie, G.E.; Calendar, R. Bacteriophage P2. *Bacteriophage* **2016**, *6*, e1145782. [CrossRef]

33. Casjens, S.R.; Hendrix, R.W. Bacteriophage Lambda: Early Pioneer and Still Relevant. *Virology* **2015**, *479–480*, 310–330. [CrossRef]

34. Łobocka, M.B.; Rose, D.J.; Plunkett, G.; Rusin, M.; Samojedny, A.; Lehnherr, H.; Yarmolinsky, M.B.; Blattner, F.R. Genome of Bacteriophage P1. *J. Bacteriol.* **2004**, *186*, 7032–7068. [CrossRef]

35. Hochman, L.; Segev, N.; Sternberg, N.; Cohen, G. Site-Specific Recombinational Circularization of Bacteriophage P1 DNA. *Virology* **1983**, *131*, 11–17. [CrossRef] [PubMed]

36. Piya, D.; Vara, L.; Russell, W.K.; Young, R.; Gill, J.J. The Multicomponent Antirestriction System of Phage P1 Is Linked to Capsid Morphogenesis: P1 Antirestriction Is Linked to Morphogenesis. *Mol. Microbiol.* **2017**, *105*, 399–412. [CrossRef] [PubMed]

37. Iida, S.; Streiff, M.B.; Bickle, T.A.; Arber, W. Two DNA Antirestriction Systems of Bacteriophage P1, darA, and darB: Characterization of darA− Phages. *Virology* **1987**, *157*, 156–166. [CrossRef] [PubMed]

38. Karvelis, T.; Druteika, G.; Bigelyte, G.; Budre, K.; Zedaveinyte, R.; Silanskas, A.; Kazlauskas, D.; Venclovas, Č.; Siksnys, V. Transposon-Associated TnpB Is a Programmable RNA-Guided DNA Endonuclease. *Nature* **2021**, *599*, 692–696. [CrossRef]

39. Altae-Tran, H.; Kannan, S.; Demircioglu, F.E.; Oshiro, R.; Nety, S.P.; McKay, L.J.; Dlakić, M.; Inskeep, W.P.; Makarova, K.S.; Macrae, R.K.; et al. The Widespread IS200/IS605 Transposon Family Encodes Diverse Programmable RNA-Guided Endonucleases. *Science* **2021**, *374*, 57–65. [CrossRef]

40. Altae-Tran, H.; Shmakov, S.A.; Makarova, K.S.; Wolf, Y.I.; Kannan, S.; Zhang, F.; Koonin, E.V. Diversity, Evolution, and Classification of the RNA-Guided Nucleases TnpB and Cas12. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2308224120. [CrossRef]

41. Sternberg, N.; Coulby, J. Cleavage of the Bacteriophage P1 Packaging Site (Pac) Is Regulated by Adenine Methylation. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 8070–8074. [CrossRef]

42. Samson, J.E.; Magadán, A.H.; Sabri, M.; Moineau, S. Revenge of the Phages: Defeating Bacterial Defences. *Nat. Rev. Microbiol.* **2013**, *11*, 675–687. [CrossRef]

43. LeRoux, M.; Srikant, S.; Teodoro, G.I.C.; Zhang, T.; Littlehale, M.L.; Doron, S.; Badiee, M.; Leung, A.K.L.; Sorek, R.; Laub, M.T. The DarTG Toxin-Antitoxin System Provides Phage Defence by ADP-Ribosylating Viral DNA. *Nat. Microbiol.* **2022**, *7*, 1028–1040. [CrossRef]

44. Paysan-Lafosse, T.; Blum, M.; Chuguransky, S.; Grego, T.; Pinto, B.L.; Salazar, G.A.; Bileschi, M.L.; Bork, P.; Bridge, A.; Colwell, L.; et al. InterPro in 2022. *Nucleic Acids Res.* **2023**, *51*, D418–D427. [CrossRef]

45. Hutinet, G.; Swarjo, M.A.; De Crécy-Lagard, V. Deazaguanine Derivatives, Examples of Crosstalk between RNA and DNA Modification Pathways. *RNA Biol.* **2017**, *14*, 1175–1184. [CrossRef] [PubMed]

46. Yu, M.X.; Slater, M.R.; Ackermann, H.-W. Isolation and Characterization of Thermus Bacteriophages. *Arch. Virol.* **2006**, *151*, 663–679. [CrossRef] [PubMed]

47. Ooi, W.-Y.; Murayama, Y.; Mekler, V.; Minakhin, L.; Severinov, K.; Yokoyama, S.; Sekine, S. A Thermus Phage Protein Inhibits Host RNA Polymerase by Preventing Template DNA Strand Loading during Open Promoter Complex Formation. *Nucleic Acids Res.* **2018**, *46*, 431–441. [CrossRef] [PubMed]

48. Matsushital, I.; Yamashital, N.; Yokota, A. Isolation and Characterization of Bacteriophage Induced from a New Isolate of Thermus Aquaticus. *Microbiol. Cult. Collect.* **1995**, *11*, 133–138.

49. Matsushita, I.; Yanase, H. The Genomic Structure of Thermus Bacteriophage IN93. *J. Biochem.* **2009**, *146*, 775–785. [CrossRef]

50. Thomason, L.C.; Court, D.L. Study of Ren, RexA, and RexB Functions Provides Insight Into the Complex Interaction Between Bacteriophage λ and Its Host, *Escherichia coli*. *PHAGE* **2022**, *3*, 153–164. [CrossRef]

51. Mara, P.; Vik, D.; Pachiadaki, M.G.; Suter, E.A.; Poulos, B.; Taylor, G.T.; Sullivan, M.B.; Edgcomb, V.P. Viral Elements and Their Potential Influence on Microbial Processes along the Permanently Stratified Cariaco Basin Redoxcline. *ISME J.* **2020**, *14*, 3079–3092. [CrossRef]

52. Medvedeva, S.; Sun, J.; Yutin, N.; Koonin, E.V.; Nunoura, T.; Rinke, C.; Krupovic, M. Three Families of Asgard Archaeal Viruses Identified in Metagenome-Assembled Genomes. *Nat. Microbiol.* **2022**, *7*, 962–973. [CrossRef]

53. Medvedeva, S.; Brandt, D.; Cvirkaite-Krupovic, V.; Liu, Y.; Severinov, K.; Ishino, S.; Ishino, Y.; Prangishvili, D.; Kalinowski, J.; Krupovic, M. New Insights into the Diversity and Evolution of the Archaeal Mobilome from Three Complete Genomes of *Saccharolobus Shibatae*. *Environ. Microbiol.* **2021**, *23*, 4612–4630. [CrossRef]

54. Rahlff, J.; Turzynski, V.; Esser, S.P.; Monsees, I.; Bornemann, T.L.V.; Figueroa-Gonzalez, P.A.; Schulz, F.; Woyke, T.; Klingl, A.; Moraru, C.; et al. Lytic Archaeal Viruses Infect Abundant Primary Producers in Earth's Crust. *Nat. Commun.* **2021**, *12*, 4642. [CrossRef]

55. Pawlowski, A.; Rissanen, I.; Bamford, J.K.H.; Krupovic, M.; Jalasvuori, M. Gammasphaerolipovirus, a Newly Proposed Bacteriophage Genus, Unifies Viruses of Halophilic Archaea and Thermophilic Bacteria within the Novel Family Sphaerolipoviridae. *Arch. Virol.* **2014**, *159*, 1541–1554. [CrossRef] [PubMed]

56. Shmakov, S.A.; Makarova, K.S.; Wolf, Y.I.; Severinov, K.V.; Koonin, E.V. Systematic Prediction of Genes Functionally Linked to CRISPR-Cas Systems by Gene Neighborhood Analysis. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5307–E5316. [CrossRef]

57. Goldfarb, T.; Sberro, H.; Weinstock, E.; Cohen, O.; Doron, S.; Charpak-Amikam, Y.; Afik, S.; Ofir, G.; Sorek, R. BREX Is a Novel Phage Resistance System Widespread in Microbial Genomes. *EMBO J.* **2015**, *34*, 169–183. [CrossRef]

58. Wang, L.; Jiang, S.; Deng, Z.; Dedon, P.C.; Chen, S. DNA Phosphorothioate Modification—A New Multi-Functional Epigenetic System in Bacteria. *FEMS Microbiol. Rev.* **2019**, *43*, 109–122. [CrossRef] [PubMed]

59. Bork, P.; Koonin, E.V. A P-loop-like Motif in a Widespread ATP Pyrophosphatase Domain: Implications for the Evolution of Sequence Motifs and Enzyme Activity. *Proteins Struct. Funct. Bioinform.* **1994**, *20*, 347–355. [CrossRef]

60. Isnard, M.; Rigal, A.; Lazzaroni, J.C.; Lazdunski, C.; Lloubes, R. Maturation and Localization of the TolB Protein Required for Colicin Import. *J. Bacteriol.* **1994**, *176*, 6392–6396. [CrossRef]

61. Pillon, M.C.; Gordon, J.; Frazier, M.N.; Stanley, R.E. HEPN RNases—An Emerging Class of Functionally Distinct RNA Processing and Degradation Enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2021**, *56*, 88–108. [CrossRef] [PubMed]

62. Stirnimann, C.U.; Petsalaki, E.; Russell, R.B.; Müller, C.W. WD40 Proteins Propel Cellular Networks. *Trends Biochem. Sci.* **2010**, *35*, 565–574. [CrossRef]

63. Bair, C.L.; Black, L.W. A Type IV Modification Dependent Restriction Nuclease That Targets Glucosylated Hydroxymethyl Cytosine Modified DNAs. *J. Mol. Biol.* **2007**, *366*, 768–778. [CrossRef]

64. Sun, X.; Göhler, A.; Heller, K.J.; Neve, H. The Ltp Gene of Temperate *Streptococcus thermophilus* Phage TP-J34 Confers Superinfection Exclusion to *Streptococcus thermophilus* and *Lactococcus lactis*. *Virology* **2006**, *350*, 146–157. [CrossRef]

65. Cumby, N.; Edwards, A.M.; Davidson, A.R.; Maxwell, K.L. The Bacteriophage HK97 Gp15 Moron Element Encodes a Novel Superinfection Exclusion Protein. *J. Bacteriol.* **2012**, *194*, 5012–5019. [CrossRef] [PubMed]

66. McGrath, S.; Fitzgerald, G.F.; Sinderen, D.V. Identification and Characterization of Phage-resistance Genes in Temperate Lactococcal Bacteriophages. *Mol. Microbiol.* **2002**, *43*, 509–520. [CrossRef] [PubMed]