UNIVERSITY OF LATVIA

Faculty of Biology

ŅIKITA ZRELOVS

EXPANSION OF THE KNOWN TAILED BACTERIOPHAGE DIVERSITY ONE PHAGE AT A TIME

Summary of Doctoral Thesis

Submitted for the degree of Doctor of Biology

Subfield of Molecular Biology

Rīga, 2023

Latvijas Biomedicīnas
pētījumu un studiju centrs
biomedicīnas pētījumi un izglītība no gēniem līdz cilvēkam

The research comprising the results of this doctoral thesis was carried out at the Latvian Biomedical Research and Study Centre from 2016 to 2023.

This work was supported by:

- Patron **SIA Mikrotīkls** through the "Excellence Scholarship in Bioinformatics" (2019-2020) and "Mikrotīkls doctoral scholarship in the field of exact and medical sciences" (2020-2023) administered by the **University of Latvia Foundation**.
- Patrons **Kristīna and Laimonis Laimiņš** through the "Laimiņš Family Scholarship in Microbiology" (2018-2019) administered by the **University of Latvia Foundation**.
- **Latvian Council of Science**, project no. lzp-2020/2–0373 "Characterization of DNA bacteriophages isolated from the insect intestine microflora" awarded to Dr. A. Dišlers.
- **European Social Fund**, project "Strengthening of the Capacity of Doctoral Studies at the University of Latvia within the Framework of the New Doctoral Model", identification No. 8.2.2.0/20/I/006.

This thesis is in the form of a collection of scientific articles in Biology, subfield - Molecular Biology.

Supervisors: Dr. biol. Andris Dišlers and Dr. biol. Andris Kazāks

Reviewers:
1) Assoc. Prof., Dr. Andrew Millard, University of Leicester
2) Prof., Dr. hab. biol. Indriķis Muižnieks, University of Latvia
3) Prof., Dr. med. Juta Kroiča, Rīga Stradiņš University

The thesis will be defended at the public session of the Doctoral Committee of Biology, University of Latvia, at 12:00 on 14th February, 2024, at the White Hall of the Latvian Biomedical Research and Study Centre, Rātsupītes street 1 k-1, Riga, Latvia.

The thesis is available at the Library of the University of Latvia, Kalpaka bulvāris 4.
This thesis is accepted for the commencement of the degree of Doctor of Biology (PhD) on ______________, 2024, by the Doctoral Committee of Biology, University of Latvia.

Chairman of the Doctoral Committee: _________________/ Prof., Dr. biol. Kaspars Tārs /
Secretary of the Doctoral Committee: _________________/ Dr. biol. Vita Rovīte /

ABSTRACT

Bacteriophages (phages) - the viruses of bacteria, are the most abundant and genetically diverse biological entities on Earth. Since their discovery more than 100 years ago, phages and their genome-encoded products have shown immense potential for practical applications to eliminate unwanted bacteria in medicine, food industry, aquaculture, and agriculture. Phage research efforts can also be considered the cornerstone of molecular biology, and many of the well-known and widely used techniques and common molecular biology lab products originate from phages. Due to their genome relative simplicity, it was also for the phages to kickstart the nucleic acid sequencing era. Currently, genomes of thousands of isolated bacteriophages have been completely elucidated, yet the incessant isolation of novel phages unlike any other regarding their genomic sequences and proteome contents proves that we have, indeed, just “scratched the surface” of the existing phage diversity. Although the meta-omics approaches allow for the mining of novel sequences of viral origin at unprecedented throughput, phages still withhold many secrets that make it hard to gain in-depth phenotypical and functional insights from their genomic sequences or parts of thereof alone. Moreover, current knowledge on phage diversity is highly biased towards bacterial hosts of healthcare or economic importance, whereas many of the currently recognized bacterial genera do not yet have any known phages associated with them at all, without any reason to assume such phages do not exist. Thus, it is in our best interest, to gain a better understanding of the phage phenomenon, to isolate and sequence not only as many phages as possible but also to make sure to do so from a variety of ecological niches and with particular attention to less commonplace hosts. In this work, I describe how the expansion of the known phage diversity is being approached in our lab “one phage at a time” and show that “straying off the beaten path” often results in the isolation of novel bacteriophages that are highly divergent from the so far cultured phage diversity and might serve as the sole representatives of the novel phage species and genera.

Keywords: virology, bacteriophages, bacteriophage diversity, isolation and characterization of phages, *Caudoviricetes*, dsDNA phages, whole-genome sequencing, genome annotation, comparative genomics, insect-associated phages

TABLE OF CONTENTS

ABBREVIATIONS

aa	amino acid
aSD	anti-Shine-Dalgarno sequence
bp	base pair
CDD	Conserved Domain Database
cos	cohesive sequence
DNA	deoxyribonucleic acid
DNAP	DNA polymerase
dsDNA	double-stranded DNA
dsRNA	double-stranded RNA
DUF	domain of unknown function
gp	gene product
ICTV	International Committee on Taxonomy of Viruses
INSDC	The International Nucleotide Sequence Database Collaboration
LDTR	long direct terminal repeats
MAG	metagenome-assembled genome
MCP	major capsid protein
ML	maximum likelihood
MSA	multiple sequence alignment
MWCO	molecular weight cut-off
NAGE	native agarose gel electrophoresis
NCBI	National Center for Biotechnology Information
NGS	next-generation sequencing
NJ	neighbor-joining
ORF	open reading frame
pac	packaging series initiation site
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PFU	plaque forming units
RNA	ribonucleic acid
RNAP	RNA polymerase
rRNA	ribosomal RNA
SD	Shine-Dalgarno sequence
SDS	sodium dodecyl sulfate
SDTR	short direct terminal repeats
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
TEM	transmission electron microscopy
TerL	terminase large subunit
TerS	terminase small subunit
TMP	tape measure protein
tRNA	transfer RNA
WGS	whole genome sequencing

INTRODUCTION

Bacteriophages or, shortly, phages are obligate bacterial intracellular parasites – viruses of bacteria, that can replicate themselves only by utilizing resources of the infected host cell. Bacteriophages are often credited with being the most genetically diverse and abundant biological entities on Earth that can be encountered in any environment in which the bacteria thrive. The estimate of "more than 10 to the power of 31" is an often-cited "staple number" of virus-like particle count in the biosphere that was derived from the observation that around 10 times more tailed phage particles than prokaryotic cells can be observed in the environmental samples by direct counting.

Each second countless numbers of bacteria are infected by phages all around us, and even on and within us regardless of wherever we may be at the time. Given their propensity to the horizontal gene exchange exemplified by the mosaic nature of the genomes of many studied phages, these infections allow phages to diversify not only by exploring the available sequence space vertically but also by abruptly changing fragments of their genomes that might span multiple genes. Furthermore, being mobile genetic elements in their essence, phages greatly facilitate the horizontal evolution of bacteria via transduction, whereas some temperate phages can even provide their hosts with additional fitness benefits via lysogenization. As a result of these events, an immense diversity of bacteriophages is expected to exist in nature, yet only a minuscule fraction of this diversity has been thoroughly characterized to be "understood", despite the vast impact phages have on their surrounding environments that cumulatively sum up to affecting the biosphere as a whole.

Properties bacteriophages exhibit promptly allowed to propose their practical application possibilities to solve pressing healthcare problems almost immediately after their discovery at the beginning of the 20th century. Moreover, it was for the phages, owing to their perceived simplicity, to kickstart the era of molecular biology and give rise to a plethora of discoveries resulting in the development and approbation of numerous widespread techniques and materials that quickly snowballed our global understanding of life. Yet, curiously, most of the existing phage-derived or phage-aided solutions or applications for a very long time originated from merely a handful of well-studied phages.

As most of the existing phage diversity is still untapped, we can expect numerous novel creative use cases to stem from bacteriophages in the future. To facilitate such innovation, as well as to gain a better understanding of the phage phenomenon and its complete implications, however, we must never be content with understanding several selected bacteriophages in the most elaborate detail, but we should also strive to advance our knowledge on as many phages as possible. Although one may argue that the in-depth studies of a particular phage might be extrapolated to many other phages with a high degree of confidence, the sheer number of exclusions to many previously set "phage biology rules" surfacing recently in the peer-reviewed literature warrants caution at trying to define any universal rules governing the peculiarities of large groups of bacteriophages in the future.

Even though it is lucrative to shift the phage diversity studies to metaviromics approaches that allow for an unmatched throughput for the acquisition of novel phage genetic material from the whole environment for further description, these approaches, while being very valuable in the ecological context and allowing formulation and testing of interesting hypotheses, are intrinsically limited to the stretches of nucleotides that serve as a study object. At the same time, laborious culture-based phage diversity studies, while having a multitude of times lower throughput, open extensive possibilities for studies of behavioral traits of a particular novel phage with very high confidence about the acquired results, thus having the ability to generate novel genotype-phenotype associations, as well as to document features of the phage impossible to arrive at from the sequence data alone.

In my thesis, I present several examples of how our largely culture-based studies of phage-host pairs isolated from the relatively understudied environments allowed for the expansion of the known phage diversity.

The aim of my thesis research was to expand the acknowledged bacteriophage diversity by characterizing novel tailed bacteriophages.

This aim was attained by performing the following tasks:

- Isolation and identification of phage-host pairs from globally understudied environments.
- Acquisition of the purified high-titer samples of the newly isolated bacteriophages.
- Transmission electron microscopy of the newly isolated phage virions to determine their morphological features.
- Elucidation of novel phage complete genome sequences and their functional annotation using comparative genomics approaches.
- Placement of the newly isolated phages within the context of the so far isolated phage diversity.

1. LITERATURE OVERVIEW

Viruses of bacteria were independently discovered in the early 20th century by Frederick William Twort (Twort, 1915) and Félix d'Hérelle (D'Herelle, 1917), with the latter introducing the name "bacteriophage" to describe a presumed "anti-microbe" causing "clear spots" within the bacterial lawns he observed. This discovery kickstarted a whole new wondrous bacteriophage research direction that has had its ups and downs but undeniably has made an immense impact on shaping biology as we know it today (Letarov, 2020).

1.1. Bacteriophage abundance

The total number of prokaryotes on Earth was estimated to be 4-6 × 10^{30} cells, and there is considerable variation in cell counts across different habitats (Whitman *et al.*, 1998). Obviously, wherever there is life – there will be viruses. Assuming that direct counts demonstrated approximately 10 times more tailed phage particles than cells, the total virus particle count on Earth was estimated to be on the order of 10^{31} (Hendrix *et al.*, 1999; Mushegian, 2020). This makes bacteriophages the most abundant biological entities on our planet. Moreover, bacteriophages are expected to be present in every environment where bacteria can be found, which makes them, essentially, omnipresent in nearly all ecological niches. For example, in marine environments, a milliliter of water contains from $\sim10^3$ to $\sim10^6$ microbes and $\sim10^5$ to $>10^7$ virus particles, with concentrations of both microbes and viruses generally decreasing with depth (Wigington *et al.*, 2016). In soils, viral abundance per gram dry weight ranges from approximately 10^3 in hot desert soils to more than 10^9 in agricultural and wetland soils, with soil viral abundances being significantly positively correlated with bacterial abundances (Williamson *et al.*, 2017). While in the human fecal filtrate, virus-like particle counts were around 10^8 per milliliter (Hoyles *et al.*, 2014). However, "The great plate count anomaly" (discrepancy between high bacteria counts observed microscopically versus low counts of colony forming units observable after cultivation attempts; (Staley, 1985)) implies that most of the biological diversity hidden in the microworld, including phages, was effectively hidden from researchers until the advent of metagenomics and metaviromics (Bragg and Tyson, 2014; Handelsman, 2004; Handelsman *et al.*, 1998).

1.2. Bacteriophage taxonomy

The official taxonomy of viruses infecting bacteria is under the oversight of the International Committee on Taxonomy of Viruses (ICTV). More specifically, currently ICTV bacterial viruses subcommittee (BVS) serves as the intendant of the phage taxonomy. Although any person can propose either updates or refinement of the phage taxonomy by filling in a template for taxonomic proposal in detail, for the changes to become part of the officially recognized phage taxonomy, the proposal needs to be ratified by the majority vote of the ICTV members (Lefkowitz *et al.*, 2018).

The International Code of Virus Classification and Nomenclature (ICTV, 2021) states that: "Viruses *sensu stricto* are defined operationally by the ICTV as a type of MGEs that encode at least one protein that is a major component of the virion encasing the nucleic acid of the respective MGE and therefore the gene encoding the major virion protein itself; or MGEs that are clearly demonstrable to be members of a line of evolutionary descent of such major virion protein-encoding entities. Any monophyletic group of MGEs that originates from a virion protein-encoding ancestor should be classified as a group of viruses". While the currently adopted definition for virus species sounds as follows: "a species is a monophyletic group of mobile genetic elements (MGEs) whose properties can be distinguished from those of other species by multiple criteria", and these distinguishing criteria are defined by the respective ICTV study groups (ICTV, 2021).

The taxons recognized by ICTV for virus classification purposes today are as follows (from less to more specific; 15 ranks in total, with species being the most specific taxon): realm ("*-viria*"), subrealm ("*-vira*"), kingdom ("*-virae*"), subkingdom ("*-virites*"), phylum ("*-viricota*"), subphylum ("*-viricotina*"), class ("*-viricetes*"), subclass ("*-viricetidae*"), order ("*-virales*"), suborder ("*-virineae*"), family ("*-viridae*"), subfamily ("*-virinae*"), genus ("*-virus*"), subgenus ("*-virus*"), and species (no particular suffix). Obviously, virus taxonomy is in active flux, and in the current era of (meta)genomics, the demarcation of particular viruses to taxa is starting to be driven almost exclusively by their genomic sequences and the information encoded within these genomes with a strong focus on the shared evolutionary descent (monophyly) so that the virus taxonomy reflects the evolutionary history of viruses (Simmonds *et al.*, 2023). Due to the taxonomy of viruses, especially phages, still being actively shaped and reworked, in its current state it is very common to find viruses classified at the lower taxonomic levels (e.g., species, genus) but lacking assignment to many of the higher-level taxa (*incertae sedis*). For tailed phages from class *Caudoviricetes* (previously known as order *Caudovirales*), classification at the ranks of species and genera seems to have coalesced towards a single mechanistic criterion of pairwise intergenomic similarities between the representatives (Turner *et al.*, 2021). Based on this criterion, phages representing the same species should demonstrate intergenomic similarity exceeding 95% (95% identical at the nucleotide level over their full genome length, tested reciprocally), whereas for genus intergenomic similarity of more than 70% is defined as a demarcation criterion. However, investigation of the pangenomes and monophyly in the phylogenies of the signature genes within the appropriate context are ought to be carried out in "edge-cases" (Moraru *et al.*, 2020; Turner *et al.*, 2021). Demarcation of tailed phages into many of the higher taxa (apart from realm *Duplodnaviria*, kingdom *Heunggongvirae*, phylum *Uroviricota*, and class *Caudoviricetes* in which all the tailed phages currently seem to fall into just on the basis of having a tail), however, is not as straightforward, and the universal guidelines are either absent or very vague.

Bacterial viruses recognized by ICTV in the official taxonomy Master Species List v37 ratified in 2022 were represented by 3601 species, 1199 genera, 98 subfamilies, 47 families, and 4 orders (Turner *et al.*, 2023). Whereas inspection of the ICTV Virus Metadata Resource based on the Master Species List 38 v1 (which was uploaded to the ICTV website on 25th of April 2023) filtered to include only bacteriophages shows that currently viruses of bacteria are already represented by at least 4 different recognized realms, 6 kingdoms, 7 phyla, 8 classes, 10 orders, 60 families, 111 subfamilies, 1809 genera, and 5024 species. This clearly demonstrates how the focus on genome-based criteria for phage classification has facilitated a rapid expansion of the phage taxonomy that is currently still experiencing major changes.

General characteristics and selected contemporary taxonomical placement for some of the most well-studied bacteriophage species whose representatives are used as important examples to highlight the diversity of phages throughout my thesis are found in Appendix 1.

1.3. Bacteriophage diversity

ICTV-ratified taxonomy, however, cannot keep up with encompassing all of the thus far available phage diversity, that is steadily expanding. Although this expansion is mainly achieved through metagenome or metavirome-derived sequences, most of such sequences are expected to be of lackluster quality regarding the completeness of viral genomes due to intrinsic limitations accompanying the road from a sample to contigs, thus, the inclusion of most such sequences in taxonomy is highly undesirable at most of the ranks (Roux *et al.*, 2019). For example, IMG/VR v4 database of uncultivated virus genomes lists more than 15 million virus genomes and genome fragments clustered into 8.7 million operational taxonomic units (231 408 with at least one "high-quality" representative; (Camargo *et al.*, 2022)). At the same time, the amount of cultured virus isolated (including phages) that get their complete or near-complete genomes sequenced and deposited in public biological sequence repositories continues to grow as well due to an increase in the availability of the NGS to a wider range of researchers. Importantly, only the cultured phage complete genomes represent the diversity of phages readily available "at hand" for further studies on the bacteriophage phenomena in connection to establishing novel genotype-phenotype associations and driving phage-based practical application development, ultimately advancing the amount of information deducible from the uncultured phage genomic sequences or parts of thereof. Such complete or near-complete genomes of cultured viruses can be found in the International Nucleotide Sequence Database Collaboration repositories (DNA Data Bank of Japan (Mashima *et al.*, 2017), EMBL-EBI European Nucleotide Archive (Leinonen *et al.*, 2011), and NCBI GenBank (Sayers *et al.*, 2019)), although they are becoming increasingly permissive to metagenome-assembled sequences as well, which makes it non-trivial to evaluate the extent of how many of such genomes have the corresponding isolated phages are actually available in labs or culture collections around the world. Moreover, the quality of the sequences and the associated annotations making it into such databases are often very lackluster for what the submitting authors refer to as the "complete genome" of a given phage, with obviously non-sensical entries slipping through frequently. Importantly, although the definition of what is a bacteriophage implies that only viruses infecting bacteria should be considered phages, the moniker "phage" is still being incorrectly attributed to viruses of archaea as well due to some overlaps in the properties of some phages and archaeal viruses, and the fact that the first viruses infecting archaea were isolated prior to recognition of archaea as a domain of line (Abedon and Murray, 2013; Pietilä *et al.*, 2014; Woese *et al.*, 1990). Thus, while *sensu stricto* "phage" is only correctly applicable to viruses infecting bacteria, it is not unusual to see the incorrect *sensu lato* interpretation that consolidates both viruses of bacteria and viruses of archaea under the term "phage".

Currently, the best effort to systematically follow the expansion of the diversity of high-quality phage (including archaeal viruses under this definition) genomes publicly available is represented by the "INfrastructure for a PHAge REference Database" or, shortly, INPHARED (Cook *et al.*, 2021). Although it also includes metagenome-assembled genomes of uncultured viruses despite initially being presented as a resource aggregating genomes of cultured phages, as well as some very questionable entries tend to slip through as well, the authors update the phage list monthly (at least so far), and allow the community to extend the "exclusion list" of accession numbers for sequences that are clearly not "complete genomes" of phages. The INPHARED release from the 1st of August of 2023 listed 25,903 complete or near-complete genome entries of bacteriophages and viruses of archaea and the associated metadata (20,506 entries excluding redundant RefSeq accessions). These viruses, however, were isolated using representatives of only up to different 302 bacterial or archaeal genera. Out of these 302 host genera, 103 have only a single associated virus, whereas 205 – less than ten viruses infecting them. Moreover, there is a high bias towards phages isolated on hosts representing only a few of these genera. Approximately half of the entries with at least some host-associated metadata correspond to genomes of the phages isolated using only eight bacterial genera representatives: *Mycobacterium* spp. – 2272, *Escherichia* spp. – 1738, *Pseudomonas* spp. – 954, *Vibrio* spp. – 908, *Salmonella* spp. – 847, *Klebsiella* spp. – 840, *Streptococcus* spp. – 737, *Gordonia* spp. – 616 (out of 17,789/20,506 non-RefSeq accessions from the corresponding INPHARED release having at least some host indication). Approximately 82.09% or 16,834 of the 20,506 viruses listed are thought to be tailed dsDNA viruses from class *Caudoviricetes*. This being said, INPHARED, probably, currently serves as the best take at the acknowledgment of bacteriophage diversity based on complete or near-complete genomes mostly originating from cultured bacteriophages *sensu lato* (Cook *et al.*, 2021).

1.4. Bacteriophage infections process

Based on their infection cycle differences, generally, bacteriophages can be categorized as being either virulent (also known as "lytic") or temperate (also known as "lysogenic"). In the case of virulent phages (e.g., *Tequatrovirus T4* (Miller *et al.*, 2003)) the only expected outcome of their successful lytic cycle is the death of the host cell and release of the phage progeny into the surrounding environment. On the other hand, temperate phages (e.g., *Lambdavirus lambda* (Casjens and Hendrix, 2015)) make a lysis-lysogeny decision upon the successful transfer of their genetic material within the appropriate host cell. Initiation of the lysogenic cycle results in the phage becoming a part of the host genome for an unspecified amount of time (either through the integration of phage genetic material in the host's chromosome, or persistence in the form of a circular or linear episome, as is the case of *Punavirus P1* (Łobocka *et al.*, 2004, p. 1) and *Ravinvirus N15* (Ravin, 2011), respectively). This implies that the integrated phage – a state known as "prophage", shuts down the expression of most of its genes, and is able to replicate as part of the bacterial genome without both forming any progeny virions and leading to the demise of its host. The prophages, however, are also capable of being activated ("induced") either spontaneously or by unfavorable conditions of physical (e.g., UV radiation) and chemical (e.g., Mytomicin C) nature, which lead to the initiation of the lytic cycle that mimics the events happening upon the infection of the host cell by a strictly virulent phage (Raya and H'bert, 2009; Xu *et al.*, 2018).

However, there are also other infection cycle varieties, such as chronic infection cycle filamentous and pleomorphic phages use – progeny bacteriophage production and release into the surrounding environment without either causing the demise of the host-bacterium or integrating into the host's genome (Mäntynen *et al.*, 2021). Additionally, a phenomenon where phage nucleic acids are just inactively residing in cytoplasm after the intrusion into the host cell and are being passed asymmetrically to one of the daughter cells upon host cell division termed "pseudolysogeny" has also been documented for several bacteriophages (frequently associated with starvation of the host cell (Ripp and Miller, 1998, 1997)), albeit it eventually resolves in either lytic infection or proper lysogenezitation of the host.

All three, lytic, temperate, and chronic phages, nevertheless, have conceptually the same infection process initiation – phage virion should identify a suitable host and transfer its genome within the respective cell. At a finer resolution, this can be described in a two correlating step manner: (1) recognition of the cognate cell surface receptor(s) by structural components of the respective phage virion and adsorption to the host and (2) penetration of the host cell barriers and the viral genetic material entry into the cytoplasm of the host (Rakhuba *et al.*, 2010).

1.4.1. Recognition of the host cell by bacteriophages

Through constant co-evolution together with their hosts for at least several billion years since the last universal common ancestor emerged (Krupovic *et al.*, 2020), phages have developed numerous strategies to be capable of identifying the appropriate hosts while passively diffusing throughout the environment and randomly colliding with all sorts of matter, including a variety of different bacteria. Currently, multiple components of either the bacteria cell wall or its protruding structures, such as flagella and pili, were shown to serve as receptors for phage adsorption (Bertozzi Silva *et al.*, 2016). The receptor repertoire various phages can use, expectedly, differs between Gram-positive and Gram-negative hosts. For tailed Gram-negative host phages, structures of smooth and rough lipopolysaccharide varieties, as well as outer membrane porins and other membrane transport proteins serve as adsorption targets, whereas tailed Gram-positive host phages utilize teichoic acid (substituent groups or backbone) and peptidoglycan components as their receptors in addition to cell wall proteins (Nobrega *et al.*, 2018).

Importantly, the majority of the tailed phages are thought to first reversibly adsorb to a "primary" receptor, before binding to a "secondary" receptor irreversibly and ejecting the content of the virion inside the cell, and a degree of cognate receptor "preferences" for different phage lineages was previously demonstrated (Maffei *et al.*, 2021). While the appropriate receptors on the cell surface might not be immediately accessible for a phage to initiate adsorption in the case of capsular- or extracellular-polysaccharide-producing bacterial hosts, phages of such hosts encode a diversity of either virion-associated or soluble polysaccharide depolymerases that allow for the degradation of such carbohydrate barriers for the exposure of the sought adsorption receptors (Latka *et al.*, 2017).

1.4.2. Delivery of the phage genome into the cytoplasm of the host

After the successful adsorption to the terminal receptor, a bacterial virus must next enter the host cytoplasm to begin replicating itself. Given the vast diversity of both bacterial hosts and their phages, the entry into the cytoplasm differs greatly among different phage groups as well.

Phages with contractile tails, as exemplified by *Tequatrovirus T4* myophages, breach the host-cell envelope by initiating their baseplate conformational change that results in the contraction of the tail sheath that, in turn, drives the tail tube towards the cell membrane accomplishing the puncturing of the cell wall with the help of baseplate central spike protein. After this, a tail spike dissociates from the tail tube to open the tail tube channel, with the tail-associated exolysin further locally degrading the peptidoglycan allowing for the tail tube to protrude until the plasma membrane to finish the tail contraction. Although the tail tube interactions with the inner (plasma) membrane happening after are still not fully understood, they are thought to involve fusion with the T4 ejection nanomachine to create a transmembrane channel, ultimately allowing for the viral DNA translocation into the host cytoplasm through the tail tube (Hu *et al.*, 2015; Leiman and Shneider, 2012).

Phages with long, flexible non-contractile tails, as exemplified by *Lambdavirus lambda* siphophages, similarly, undergo significant conformational change that brings the bottom of the tail tip into the contact with cell surface upon recognition of the appropriate by receptor by tail tip complex-associated fibers. Subsequently, the tail tape measure protein that might have a relevant enzymatic activity to degrade the envelopes of the host cell is forced out of the tail tube, puncturing the outer membrane and likely forming the channel for DNA passage into the cell interior (Davidson *et al.*, 2012).

Phages with short stubby non-contractile tails, as exemplified by *Teseptimavirus T7*podophages, use the inner capsid proteins to form a channel perforating the outer membrane of the cell, degrading the peptidoglycan, and, ultimately, extending through the inner membrane as well. Tail proteins might have the depolymerizing activities necessary to pass through the host cell's outer barriers. The phage DNA is subsequently passed into the host cell through this channel. Details of the processes happening between the successful recognition of the primary receptor on the surface of the cell and before the genetic material of the phage is already inside the cell are still very poorly understood for podoviruses infecting Gram-positive hosts (Casjens and Molineux, 2012).

What forces drive the ejection of the dsDNA from the capsid of tailed phages into the cytoplasm of the host is still a matter of debate (Molineux and Panja, 2013).

1.4.3. Expression of the tailed phage genome

Despite usually having comparatively large genomes (for viruses), some of which might even approach the genomes of simplest prokaryotes lengthwise (e.g., megaphages with genomes >540 kbp versus *Mycoplasma genitalium* G37 580.1 kbp), and relatively complex structure of the virions, phage genomes are expressed in quite straightforward fashion with unexpectedly simple regulation.

Owing to their modular organizational structure, generally, phage genomes have discernible gene modules corresponding to the functionally connected proteins that are expected to simultaneously act at a particular time in the bacteriophage infection cycle and, thus, are being expressed at the same time post-infection. Studies on multiple tailed dsDNA phages allowed to propose a temporal resolution of tailed dsDNA phage gene expression patterns to discern early, middle, and late gene classes (Table 1), which usually encode proteins involved in overtaking the host systems (early), proteins involved in phage genome replication (middle), and viral particle morphogenesis and release (late). For some phages, pre-early, early, and late gene classes are outlined. However, it was also noted that functionally analogous genes might be expressed in different temporal groups for different phage-host systems (e.g., see Figure 2 from (Clokie *et al.*, 2020)).

Interestingly, transcription strategies also tend to differ greatly between dsDNA phages. For example, phages such as *Lambdavirus lambda* (Casjens and Hendrix, 2015) and *Tequintavirus T5* (Wang *et al.*, 2005) completely rely on the host RNA polymerase (RNAP) for the expression of their (pre-early, early, and late) genes. The gene expression regulation of such phages is thought to be accomplished using mainly transcriptional regulators (repressors, activators, antitermination factors, etc.). *Tequatrovirus T4* also uses host RNAP for the transcription of early, middle genes, and late genes, with an important distinction that the respective enzyme gets modified by a phage, and its late genes are even transcribed using a host RNAP that is modified with a phage sigma factor-like protein (Luke *et al.*, 2002; Tabib-Salazar *et al.*, 2019). *Okubovirus SPO1* uses the native RNA polymerase of the host for its early genes, while the middle and late gene regions are subsequently transcribed also using an RNA polymerase modified by a phage-encoded sigma factor (Losick and Pero, 1981; Stewart *et al.*, 2009). *Teseptimavirus T7*, similarly, uses host RNAP for the early gene transcription, while the middle and late genes are then transcribed by the phages' own RNAP polymerase (Dunn *et al.*, 1983). Things get even more interesting with the *Phikzvirus phiKZ*, whose genome encodes two DNA-dependent RNAPs – virion and non-virion RNAP. That way, the early gene region of the phiKZ genome is transcribed using virion RNAP that is injected into the host alongside the DNA, whereas the middle and late gene regions – using phage-encoded non-virion RNAP (Ceyssens *et al.*, 2014; Mesyanzhinov *et al.*, 2002). *Enquatrovirus N4* also encodes virion and non-virion RNAPs, however, it also does not shy away from using the RNAP of the host. In phage N4 infection, the early gene region is transcribed using virion RNAP, middle genes – using phage-encoded non-virion RNAP, whereas late genes, amusingly, are transcribed using RNA polymerase of the host cell (Cho *et al.*, 1995; Zivin *et al.*, 1981). This change and/or modulation of RNAPs to identify different promoters seems to be a recurring ingenious stunt among many different phages that helps to achieve temporal resolution of gene expression.

Gene temporal group	**Functions**	**Examples**
Early	Neutralization of the host defense systems	Host restriction enzyme inhibitors
	Halting of macromolecule synthesis by the host	Modification of the host cell enzymes; degradation of the host DNA
	Superinfection exclusion	Blocking infection by other (usually related) phages
	Insurance of the middle gene transcription	phage RNA-polymerase
Middle	Replication of the phage genome	Primosome/replisome proteins
	Metabolic modification of the host	tRNA pool adaptation; unusual dNTP synthesis; phage chaperone synthesis
	Insurance of the late gene transcription	Transcription and sigma factor synthesis
Late	Morphogenesis of progeny phage particles	Virion structural components, DNA packaging motor
	Release of phage progeny into the environment via cell lysis	Lysis genes (endolysin, holin, spanin(s))

Table 1. Generalized temporal resolution of gene functional groups expressed by tailed dsDNA phages and the associated examples. Adapted from (Letarov, 2019). Note that the presence or absence of particular phage-encoded proteins is highly varied among phage groups.

1.4.4. Replication of the tailed phage genome

Despite that all the so far known tailed bacteriophages have a linear double-stranded DNA genome packaged within their virions, a diversity of possible tailed phage genome replication strategies and the resultant forms of the replicated DNA was previously demonstrated. These include theta (θ) replication, rolling circle or sigma (σ) replication, replication using recombinational initiation, replication involving hybridization of terminal repeats, replication using replicative transposition, protein-primed replication, replication with the regeneration of telomeres (Letarov, 2019) (Table 2).

Some bacteriophage genomes circularize when reaching the cytoplasm of the host and rely mostly on the host protein machinery for the replisome assembly and action (e.g., lambda), while the other phage genomes stay in a linear form upon arrival into the cytoplasm and encode all the components of the replication and recombination machinery (e.g., T4), although some of these components can be substituted with the host counterparts *in vivo*.

Even though there seems to be a great deal of diversity regarding both the mechanisms (Weigel and Seitz, 2006), and the end products of phage genome replication that shall, later on in the infection cycle, serve as substrates for packaging into the proheads of the progeny virions to-be, the genome packaging machinery itself is rather conserved among the tailed bacteriophages.

Phage	**DNA molecule inside the virion**	**Replication strategy**	**Form of the replicated DNA**
Lambda	Cohesive ends (12-base 5' cos)	Theta switching to sigma (rolling circle)	Linear concatemers
T4	Terminal redundancy (circularly permuted molecules)	Recombinational initiation	Branched concatemers
T7	Direct terminal repeats (160-base direct terminal repeats)	Hybridization of terminal repeats	Linear concatemers
Mu	Unique host DNA segments attached at both ends (different lengths of left-flanking and right-flanking host sequences)	Replicative transposition	Genome copies integrated into multiple sites of the host chromosome
phi29	Protein (gp3) covalently attached at 5' ends	Protein-primed	Monomers with gp3 covalently attached at 5' ends
N15	Cohesive ends (12-base 5' cos)	Regeneration of telomeres (prophage state)	Linear monomers with covalently closed ends – telomeres (prophage state)
P2	Cohesive ends (19-base 5’ cos)	Modified rolling circle	Circular monomers

Table 2. Examples of the diversity of replication strategies employed by tailed dsDNA bacteriophages. Inside the virions genomes of all tailed dsDNA phages known so far are linear.

1.4.5. Tailed phage virion morphogenesis

Morphogenesis genes of tailed dsDNA bacteriophages are usually expressed late in the infection cycle. In many phages (e.g., lambda and lambda-like siphoviruses), morphogenetic genes represent a single operon under the regulation of a sole promoter (e.g., promoter $P_{R'}$ in phage lambda). Assembly or morphogenesis of tailed phage progeny virions in essence represents combining the phage virion structural feature intermediates (e.g., capsids, tails, tail fibers) that assemble in the cytoplasm of the host independently of each other in case of sipho- and myophages. While the details of the assembly process might be notably different even within the phages representing the same phage group, it is believed that several common principles observed when combining the results of studies of several tailed model phages that have garnered a lot of interest for historical reasons (e.g., T4 myophage, Lambda siphophage, T7 podophage, etc.) might be extended in a generalized manner to rather broad spectra of phages according to their morphotypes, regardless of the individual phage differences in capsid sizes and symmetries, tail lengths, etc. (Aksyuk and Rossmann, 2011).

Generally, the following steps take place during tailed phage virion morphogenesis:
1. Proteins involved in virion morphogenesis are synthesized under the control of late genes.
2.1.1. Initiation of the prohead assembly on the portal protein complex with the help of scaffolding protein.
2.1.2. Proteolysis of the scaffolding protein or major capsid protein delta domains, and cleavage of other proteins involved in prohead morphogenesis, followed by the removal of the cleaved unnecessary fragments from within the prohead.
2.1.3. ATP-dependent phage DNA packaging inside the proheads and their expansion.
2.1.4. Sealing of the expanded proheads with the DNA packaged inside by the head closure and "neck" proteins.
2.1.5. In the case of podophages, tail proteins cooperatively assemble on the filled capsid without forming structures independently of the head.
2.2.1. Baseplate assembly (for sipho- and myophages).
2.2.2. Tail tape measure protein-tail chaperone complex connection to the assembled baseplate.
2.2.3. "Growth" of the tail along the tail tape measure protein (polymerization of tail tube protein) with the displacement of chaperones, if any.
2.2.4. Tail assembly completion by tail terminator protein connection (coupled with tail sheath polymerization along the tail tube for myophages).
3. Assembly of capsid-tail from preassembled capsids with DNA packaged inside and tail structures (sipho- and myophages).
4. Connection of fibers to the capsid-tail structures to finish sipho- and myophage virion assembly.

1.4.6. Tailed phage progeny release into the environment

The last step of the bacteriophage lytic infection cycle is expected to end in the release of phage progeny into the surrounding environment and give rise to new infections within a population of susceptible host cells. To do this, a multitude of the newly formed virions need to get out of the host cell. In the case of tailed phages, this step ends in death and disintegration of the host cell – the so-called cell lysis, which is achieved by the controlled action of a few phage-encoded lysis proteins.

The lysis cassette of tailed bacteriophages is usually represented at least by genes encoding for an endolysin – one of the enzymes with peptidoglycan-degrading activity (e.g., hydrolases such as muramidases, glucosaminidases, amidases, peptidases, as well as transglycosylases (Vázquez and Briers, 2023)), and a holin – a protein responsible for breaching the integrity of the host cell membrane (Tran *et al.*, 2005). Although simple endolysin-holin lysis cassettes might suffice for phages infecting Gram-positive hosts, phages infecting Gram-negative hosts also ought to possess genes encoding either a unimolecular spanin or both inner and outer spanins (Kongari *et al.*, 2018). Lysis genes are generally expressed late in the infection cycle to warrant sufficient build-up of the phage progeny inside the host prior to its lysis, which very frequently happens way before the resources of the host are depleted.

1.5. Tailed bacteriophage genomics

The immense genomic diversity tailed phages demonstrate can undoubtedly be attributed to their propensity to horizontal gene exchange between both the different phages infecting the same host cell, and even between phages and the host (e.g., intact or cryptic prophage regions). This tendency to participate in recombinational events has resulted in one of the main tailed phage genome features – genomic "mosaicism", where stretches of the same genomic nucleotide sequence tend to have different evolutionary origins (Brüssow *et al.*, 2004). Such mosaicism is usually easy to spot when comparing the genomes of very closely related phages – almost identical stretches of DNA can be interspersed by unrelated sequences. Moreover, nearly identical stretches of DNA in related phages can suddenly differ by the presence of an "extra" sequence squeezed within one of the genomes while being flanked by homologous regions. This observation of high genomic mosaicism allowed to offer a concept of phages having access to the “repository” of shared genetic material (Hendrix *et al.*, 1999). Moreover, it also authorized speculations that, in natural communities, speciation by classical mutation accumulation and further selection might have a secondary role in creating the immense diversity of the phage genotypes when compared to horizontal gene transfer that is able to rapidly shift large segments of the phage genomic DNA giving rise to novel phage genotypes (Bellas *et al.*, 2020).

Comparative genomics studies demonstrated that horizontal gene transfer events resulting in the exchange of gene blocks are occurring not only between very closely related phages but also among significantly different phages. Throughout the years such studies further extended the previously postulated phage "modular evolution theory" (Botstein, 1980) and allowed further refinement of the definition of a module in this regard. For instance, *Lambdavirus lambda* and *Lederbergvirus P22* are infecting hosts representing different genera (*Escherichia coli* and *Salmonella typhimurium*, respectively), and even have different virion morphologies (lambda is a siphovirus, whereas P22 is a podovirus), yet they share similar genomic organization and gene order despite having little nucleotide sequence homology (Ackermann, 2015). In experiments creating chimeric phage lambda-P22 hybrids, where one of the genes from the first bacteriophage was replaced by a homolog from the other phage, the ability to infect and replicate was retained in a lot of cases (Botstein and Herskowitz, 1974). This has helped to define the concept of a "block" or "module" - any genomic sequence responsible for a specific biological function(s) that can effectively be replaced by another phage sequence with an analogous function(s) with the retention of the ability to infect and replicate for the hybrid phage progeny. This means that a foreign module must be functionally compatible with other recipient phage modules. Currently, it is known that such an exchange can occur between genome segments of very different lengths and contexts, starting from short specific protein functional domain-encoding sequences up to the entire operons comprising multiple genes (Veesler and Cambillau, 2011).

As only chimeric phages able to reproduce will persist in any given environment for longer periods of time, comparative genomics using the wealth of phage genome sequences elucidated from cultured phages isolated from natural environments, which are available in the public biological sequence repositories, reveals that not all the phage genes or gene modules exhibit equal mobility between different genomes (Hatfull and Hendrix, 2011). Moreover, lateral gene flux seems to have many confounding variables, such as donor/recipient genome enrichment in genes associated with recombination, phage lineage, as well as host identity (where the gene transfer events take place), among others (Mavrich and Hatfull, 2017; Moura de Sousa *et al.*, 2021). Based on the homologous gene representation in different phage genomes within a wider group, it was proposed that phage genes can be somewhat reliably divided into two groups based on their conceptually differing mobility patterns between the genomes - "core" and "accessory" genome, the latter also frequently being referred to as "hyperplastic regions" (Cazares *et al.*, 2014; Comeau *et al.*, 2007; Hatfull and Hendrix, 2011). In this context, the "core" genome comprises genes or gene groups whose homologs are found in all the genomes representing a specific phage taxonomic group. Usually, “core” genome modules are responsible for a particular functional pathway in the phage infection cycle where congruence between the involved gene products is especially important for their efficient interactions (Petrov *et al.*, 2010). Functions for many of such cornerstone genes have already been elucidated using model phages and reasonably extended for larger phage groups based on genome similarity and synteny, as well as resultant protein homology. Unfavorable replacement of or defects in the "core" genes are thought to either preclude or greatly hamper phage reproduction abilities, leading to selection against such genotypes in the population (Abedon, 2009). The "core" genes are usually responsible for vital functions such as virion structural feature morphogenesis, phage genome replication, and phage genome packaging into procapsids (Comeau *et al.*, 2007). Contrary to that, the hyperplastic “accessory" or “peripheral” genome is thought to be directly responsible for the adaptation of a given phage to a particular ecological niche. Products of such genes that seem to be frequent targets of horizontal flux between even the distantly related phage genomes are proteins allowing to counteract host defense mechanisms, modify host cell phenotypes via providing additional functions, facilitate recombination, etc. (Moura de Sousa *et al.*, 2021). The "accessory" genome contents, however, expectedly exhibit significant variability among phages and, thus, most of the genes found in such hypervariable regions of the phage genomes are proteins with functions yet unknown (Hatfull and Hendrix, 2011). Many of such “accessory” genes encoding “hypothetical proteins” are thought to be characteristic of a relatively limited range of phages, with some of them being so far found only in a single or few previously studied biological objects, lacking any other homologs in public biological sequence repositories rife with prokaryotic or their virus genomes. Usually, "accessory" gene clusters are located outside of or between the "core" gene clusters, and the genes in these peripheral clusters were proposed to be considerably shorter (Hatfull and Hendrix, 2011).

Although it was shown that horizontal gene exchange can happen between phages of different lifestyles, and even between rather distantly related phages (possibly due to illegitimate recombination mechanisms), temperate-to-temperate phage horizontal gene transfer is the most frequent class of events when speaking about phage lifestyles in this context (Moura de Sousa *et al.*, 2021). Temperate phages are also notoriously known to harbor genetic elements referred to as "morons" (derived from "more on" (Juhala *et al.*, 2000)). These elements (“morons”) are classified as part of the "accessory" genome, with an important distinction that morons are not essential for the successful lytic cycle of the phage harboring them (they may play no role in a lytic cycle whatsoever (Cumby *et al.*, 2012)). Usually, there is no apparent benefit from the moron for the phage itself *sensu stricto*. Yet, morons can independently (having all the necessary elements such as promoter, ORF, and terminator as part of the “moron”, as per the original definition (Juhala *et al.*, 2000); the extended definition views morons as “non-conserved genes in phage genomes that do not have a phage function” (Taylor *et al.*, 2019)) express products that give a competitive advantage to the host cell after it gets lysogenized by a moron-carrying phage. In a particular community, products of morons can confer advantages to lysogens in terms of fitness/survival when compared to bacteria of the same species that do not harbor the same moron-containing prophage (e.g., superinfection exclusion, virulence factors, antibiotic resistance (Brüssow *et al.*, 2004; Taylor *et al.*, 2019)).

Reconstruction of precise phage evolutionary relationships is hampered by the fact that no single universal marker gene having both well-conserved and variable regions (such as 16S rRNA in gene in prokaryotes) exists in all phages. While some phage proteins such as major capsid protein or terminase large subunit show considerably higher evolutionary conservation than most other phage proteins, they are still so diverged that a reconstruction of their complete evolutionary history based on their primary structure (sequence of amino acids) alignments gets very unreliable towards the root (Liu *et al.*, 2006; Rohwer and Edwards, 2002). The mosaic nature of tailed phage genomes complicates the use of whole genome alignment-based phylogenetic tree reconstruction to clarify the evolutionary relationships between phages, especially in the cases when phages being compared are not very closely related. Even relatively closely related phage colinear genomes of comparable lengths demonstrating overall synteny may be composed of different mosaic "building blocks" with completely different evolutionary histories (Hatfull, 2018). Moreover, the impact of horizontal gene flux that constantly shapes the phage genomes has prompted some researchers to propose a demonstration of the evolutionary relationships of phages in the form of a network, rather than a tree, which would allow to accommodate for horizontal gene exchange as opposed to only a divergent evolution (Lima-Mendez *et al.*, 2008). Even in the case of conserved, vertically diverging regions, the redundancy of the genetic code coupled with very high coding capacities hints that comparison of protein amino acid sequences might give more reliable results, as the respective nucleotide sequences are prone to a relatively quick erosion of information about evolutionary history. The phylogenetic reconstruction of evolutionary relationships using either concatenated amino acid sequences of “core” genome products, or amino acid sequences of independent “core” genome products chosen to serve as markers, however, is generally possible for reconstructing plausible vertical evolutionary histories of related phage groups on the levels below families. Nevertheless, a combination of approaches is generally advisable for more robust conclusions about the relatedness of novel phages to the ones characterized genomically before, and finding a place for a novel phage within the current taxonomy of viruses (Simmonds *et al.*, 2023).

Bacteriophages belonging to the class *Caudoviricetes* (tailed phages with double-stranded DNA serving as the genome) exhibit a variety of termini types of their genomic DNA molecules that always reside within the capsid in a linear form due to constraints of the portal protein channel through which the DNA is eventually ejected (Casjens and Gilcrease, 2009). The diversity of phage genome termini types is a consequence of differences in the mechanisms of processing the replicated phage genome packaging substrates during virion morphogenesis. Currently, at least seven different types of known tailed phage genome molecule physical termini can be outlined: cohesive ends (*cos*) with 5' or 3' overhangs, exact direct terminal repeats whose sequence differs among the phage virion population (circularly permuted genomes), exact short or long direct terminal repeats (SDTR - up to several hundred bp, LDTR – up to several thousand bp, respectively) that are the same in all the virions comprising phage population, containing host DNA sequences flanking the phage genome, or covalently attached terminal proteins (Casjens and Gilcrease, 2009). In tailed phages, the processing of replicated DNA during virion morphogenesis (packaging into proheads) is governed by a "DNA-packaging motor" comprising a terminase complex bound to a portal ring (Casjens, 2011). Terminase complex usually consists of two phage-encoded subunits with distinct functions - the small terminase subunit (TerS – dsDNA-binding activity) is responsible for recognizing the phage DNA, while the large terminase subunit (TerL - ATPase and nuclease activities) is involved in attaching to the phage head precursor and filling it with genetic material, as well as cleaving the DNA (if necessary) after a single packaging event is complete (Rao and Feiss, 2008). TerL aa sequence phylogeny reconstruction was proposed as a method allowing for effective prediction of the packaging strategy and, consequently, genome physical termini types newly isolated phages use, which might guide their further experimental validation (Casjens *et al.*, 2005). Identification of the TerL gene in newly sequenced and assembled phage genome is usually not hard due to the high conservation of this protein relative to most other phage-encoded proteins (Casjens, 2003), and its use in multiple sequence alignment alongside TerL sequences originating from other phages (for which the packaging strategy or physical DNA ends have been previously experimentally determined) allows to build a phylogenetic tree enabling to hypothetically infer the expected type of phage genome termini and a packaging strategy phage utilizes (Casjens and Gilcrease, 2009). This TerL phylogeny reconstruction method is based upon the observation that, given enough phylogenetic context, terminase large subunit protein sequences from phages utilizing the same packaging strategy usually fall within well-supported distinct clades represented by phages with matching physical genome molecule termini types (Casjens *et al.*, 2005; Casjens and Gilcrease, 2009; Merrill *et al.*, 2016).

Conserved core gene order across even divergent phage genomes is especially noticeable in the genes comprising the morphogenesis module of the temperate phages, which were noted to be often arranged in the following transcription order: TerS – TerL – portal protein – prohead protease – scaffolding protein – major capsid protein and other capsid proteins – head and tail connector proteins – tail morphogenesis proteins – tail length determining (tape measure) protein – tail tip/baseplate proteins – tail fiber proteins (Casjens, 2003). Variations in this order of genes are more commonly observed in large (long-genome) lytic phages, such as T4-like phages (Casjens, 2003).

Speaking of tailed phage genes, the opinions from various researchers generally converge on the following (as far as generalization of such an immense diversity can go):

I) Tailed phage genomes show very high coding sequence density, as a consequence, long non-coding regions between ORFs are rare, and adjacent ORFs might have short overlaps.

II) In the majority of cases, only a single reading frame and one strand (either forward or reverse) are used for a protein-coding gene in any given stretch of DNA.

III) Functionally related gene groups are usually co-transcribed in one direction, and switches in gene orientations (forward to reverse or reverse to forward) are infrequent.

IV) Shine-Dalgarno sequences are expected upstream of the start codons for most ORFs, and start codons might differ from the canonical ATG (e.g., GTG, CTG, TTG).

V) Short protein-coding genes are rare, most of the ORFs are >50 codons long (>150 bp).

VI) Programmed translational frameshifts and introns are rare but possible.

VII) Up to several transfer RNAs might be encoded (especially by phages with large genomes), although lack of any tRNA genes is not uncommon.

Considering the diversity of physicochemical parameters possible in the microcosms coupled with the variability in the make-up of resident biological communities and countless possible interactions within their members, there is not much of a surprise that the dominant fraction of most tailed phage genomes encodes “hypothetical proteins” of unknown function (Lima-Mendez *et al.*, 2007). Yet, the specialization of bacteriophages for specific ecological niches makes an assumption of the accessory pangenome of known and yet-to-be-discovered bacteriophages serving as an almost inexhaustible reservoir of potentially useful products. To harness the full potential of these products, it was proposed that researchers should not only focus on continuous studies of several selected individual model phage gene products' structural and functional aspects but should also aim to acquire data from novel bacteriophages inhabiting diverse and undersampled ecological niches, as well as infecting “uncommon” host organisms (Brüssow and Hendrix, 2002; Hatfull, 2015).

1.6. Tailed bacteriophage and tailed phage-derived product practical applications

The central role of bacteriophages as model organisms or bacteriophage-encoded enzymes as part of the toolkit used for the development of molecular biology and biotechnology fields cannot be underestimated. Moreover, many of the products derived from bacteriophages in the previous decades are still very frequently used in laboratories even today. However, despite the ever-growing number of completely sequenced bacteriophage genomes available in the public biological sequence repositories, most of the widely used bacteriophage gene products were originally derived from a surprisingly small number of phages that were collectively extensively studied (e.g., T4, T7, lambda, etc., (Schoenfeld *et al.*, 2010)). Given the diversity of phages that is constantly expanding with reports of bizarre objects, as well as recent technological advances allowing for higher experimental throughputs, this points towards almost limitless possibilities for further discovery of new phage products to be ingeniously used in the future.

Arguably the most important practical application of phages, however, is their deliberate usage for the removal of unwanted bacteria from different environments, including in medicinal practice – known as “phage therapy”, and this research area is truly experiencing a “renaissance” garnering a lot of attention from the researchers worldwide.

1.6.1. Phage therapy

Given the fact that multiple bacterial pathogens are becoming increasingly resistant to different antibiotic types available to mankind, coupled with the rapid pace at which the emergence and spread of multidrug-resistant strains are occurring, several experts conclude that the world is rapidly forging toward the "post-antibiotic" era (with some considering such era to have already begun). As a result of constant clinical and scientific observations that undeniably show the trajectory of the MDR issue only getting worse, antibiotic resistance is now considered by the World Health Organization to represent one of the biggest threats to global health, food security, as well as to the overall world development (World Health Organization, 2021, 2020). As a consequence of the pressure such resistant strains exert on the healthcare systems worldwide, the scientific community is urged to direct their attention towards developing and implementing alternative methods for combating pathogenic bacteria in a “race against time” manner, while important pathogens have not yet acquired multi-resistance to all known classes of antibiotics (De Oliveira *et al.*, 2020). The therapeutic potential of bacteriophages for medicinal use started to be considered immediately after their discovery at the beginning of the 20th century (Fruciano and Bourne, 2007). Thus, attempts to use bacteriophages as therapeutic antibacterial agents were made even before the discovery of the first antibiotics (starting from A. Flemming describing bacteria-killing “mould juice” that he named penicillin in 1928 (Tan and Tatsumura, 2015)). However, the interest in phage therapy diminished fairly quickly in most of the world with the advent of antibiotics, only to get globally rekindled as antibiotic resistance started to be recognized as a looming threat (Chanishvili, 2012; Gordillo Altamirano and Barr, 2019).

Compared to antibiotics, phage therapy has several advantages in the treatment of bacterial infections (as outlined by Loc-Carrillo and Abedon): suitable lytic phages guarantee the killing of target pathogenic bacteria in case of productive infection; phage preparations can themselves naturally increase the "dosage" depending on the number of host cells available in the environment as they replicate through multiple rounds of productive phage infection; due to their pronounced host specificity, phages generally do not significantly disturb the normal microflora or disrupt it minimally; given the diversity of bacteriophages, it is relatively easy to find a suitable phage or phage mixture for specific bacterial targets; certain phages have demonstrated the ability to clear bacterial biofilms; purified phages pose little to no risk to human health; phages can kill antibiotic-resistant bacteria; among other pros (Loc-Carrillo and Abedon, 2011). However, it's important to recognize that bacteria can develop resistance to phages as well (Labrie *et al.*, 2010). When a susceptible bacterial population is exposed to the infecting phage population for a prolonged period of time, resistance to that specific phage (that is sometimes extensible to a broader related phage group) can, obviously, emerge (Luria and Delbrück, 1943). This may result in such phage-resistant cells eventually overtaking the niche of the susceptible host as a result of phage pressure removing susceptible host cells, possibly rendering therapy a failure (Oechslin, 2018). Therefore, for therapeutic or biocontrol purposes, it might be necessary to seek a combination of “genetically safe” and “well-characterized” phages (“phage cocktail”) for each specific pathogen strain spectrum (Pirnay *et al.*, 2015). Ideally, such a cocktail should comprise phages genetically distinct from each other, as well as utilizing different receptors for cell surface attachment (Abedon *et al.*, 2021).

Currently, many renowned experts in the field of phage therapy consider the optimal situation to involve tailoring individual therapeutic cocktails to each patient's pathogenic bacterial strains, drawing from well-characterized phage collections at institutions like the Eliava Institute in Georgia (Kutateladze and Adamia, 2008), Queen Astrid Military Hospital in Belgium (Djebara *et al.*, 2019), the Hirszfeld Institute of Immunology and Experimental Therapy of the Polish Academy of Sciences (Żaczek *et al.*, 2020) - respected phage therapy centers/units with profound experience in phage therapy. However, the cost of this personalized approach (*sur-mesure*) is considerably higher while the “throughput” is lower compared to using pre-formulated commercially readily available polyvalent "universal" (*prêt-à-porter*) phage cocktails, (e.g., phage cocktails by "Microgen", universal phage cocktails from Eliava, “BENE phage”, etc.), which also undeniably have their use, especially in the case of particular attention to such cocktail timely reformulation according to the empirical and epidemiological data (Abedon *et al.*, 2021; Pirnay *et al.*, 2018, 2011). Even though personalized phage therapy is also prone to resistance emergence and phage immune neutralization, a series of 100 consecutive phage therapy cases representing 114 difficult-to-treat infections of different etiology and types recently reported by Pirnay and colleagues shows great promise (Pirnay *et al.*, 2023). Notably, in 69.3% (79/114) of the treated infections, bacteriophages were administered in combination with standard-of-care antibiotics, and 77.2% (88/114) of the infections have shown clinical improvement after personalized phage therapy either with or without concomitant antibiotic usage, and the eradication of targeted bacteria was observed in 61.3% (65/106) of the infections. The same authors coin the phrase “Bacteriophages Can Make Antibiotics Great Again” and propose a phage therapy paradigm shift with personalized phages being used as an additive to standard-of-care antibiotics, which has led to the improved eradication rates of targeted bacteria in a considered patient population (Pirnay *et al.*, 2023).

1.6.2. Genetically modified phages for therapy

Despite the fact that there has been a number of clinical studies that considered natural lytic phage therapy (genetically unmodified and not combined with antibiotics) as a promising method for treating diseases caused by pathogenic bacteria (the effectiveness of such phage therapy approach has been demonstrated in treating skin ulcers, skin purulent diseases, infections caused by methicillin-resistant *Staphylococcus aureus*, wound and burn prophylaxis, eye infections, gastrointestinal diseases, middle ear inflammation, respiratory and urogenital tract infections, as well as sepsis cases, among others (e.g., (Abedon *et al.*, 2011; Slopek *et al.*, 1987; Weber-Dabrowska *et al.*, 2001)), natural phages can always be improved further. For example, although strictly lytic bacteriophages are obviously inherently more suitable natural agents for phage therapy and biocontrol than temperate phages, and their use in therapy has a much longer history, recently there have been studies emerging on naturally temperate polyvalent bacteriophages that have been genetically engineered to produce lytic derivatives. These derivatives, however, should lack the genes required for initiating and maintaining the temperate life cycle (Dedrick *et al.*, 2019a). A few years ago, a clinical case report was published where the successful treatment of a patient's *Mycobacterium abscessus* infection was attributed mainly to a therapeutic cocktail of bacteriophages containing lytic derivatives targeting the *Mycobacterium* strain causing infection, which is the first case of genetically-modified phage use in therapy (Dedrick *et al.*, 2019b). Even more recently, Gencay and colleagues have described their extensive subfamily *Tevenvirinae* phage engineering efforts to create a combination of tail-fiber-engineered and CRISPR-Cas-armed phages that would target a diverse spectrum of *E. coli* strains (Gencay *et al.*, 2023). As a result of this study, a development candidate called SNIPR001 comprising four genetically modified bacteriophages that have shown enhanced bacteria-killing potency and target spectrum compared to their ancestral natural phages has recently entered clinical development (ClinicalTrials.gov ID NCT05277350), (Gencay *et al.*, 2023). Similarly, fresh phage-engineering efforts from Du and colleagues have resulted in another approach to the extension of the natural phage capabilities for the elimination of target bacteria they term “heterologous effector phage therapeutics” (Du *et al.*, 2023). In their approach, engineered phages express additionally introduced effector genes such as bacteriocins and cell wall hydrolases during the infection. The products of these effector genes are then released into the environment to function as secondary pathogen-specific antimicrobials upon host cell lysis, which complements and enhances the phage-mediated killing resulting in a “two-pronged attack” (Du *et al.*, 2023).

1.6.3. Phages in the food industry

Bacteriophages can also be used as pathogenic bacteria biocontrol agents in the food industry, allowing the reduction of contamination with disease-causing bacterial strains during the production, processing, and/or storage of the final product. Similar to other viruses, phages are constantly consumed with the food comprising the human diet, and in the studies conducted so far, no direct evidence of bacteriophage harmfulness to mammalian cells was demonstrated to the best of my knowledge. Several relatively recent review articles summarized the outcomes of several studies where various types of human food were treated with specific phages or their mixtures, which significantly reduced the colony-forming unit counts of foodborne bacterial pathogens highly undesirable in, yet commonly associated with, the food industry settings (e.g., *E. coli*, *Campylobacter* sp., *Salmonella* sp., *Listeria monocytogenes*, *S. aureus* among others, (Kazi and Annapure, 2016; Moye *et al.*, 2018; Połaska and Sokołowska, 2019; Sillankorva *et al.*, 2012)). Currently, several of the phage products are already approved for food safety applications, and as phage biocontrol is becoming increasingly more accepted throughout the world for effective elimination or significant reduction of pathogens, we can expect more such products to become available in the near future (Moye *et al.*, 2018).

1.6.4. Phages in plant protection

The use of bacteriophages for plant protection in agriculture is also being continuously explored to determine the extent of their capabilities for the prevention of bacterial infections in crops, as well as fruits and vegetables. Considering the ever-increasing global human population, it is clear that food production should increase proportionally to satisfy future food needs and provide food security for humanity, in which crops, fruits, and vegetables play an important role (FAO, 2018). Around 20 years ago, about 10% of the entire plant production was estimated to be lost due to plant diseases caused by plant pathogens, among which bacteria play a significant role (Buttimer *et al.*, 2017; Strange and Scott, 2005). The "major" phytopathogenic bacteria causing the most losses in plant production were previously outlined as representatives of the following bacterial genera: *Pseudomonas*, *Ralstonia*, *Agrobacterium*, *Xanthomonas*, *Erwinia*, *Xylella*, *Pectobacterium*, and *Dickeya* (Mansfield *et al.*, 2012).

Results of the first attempts to use bacteriophages against phytopathogenic bacteria in agriculture were documented as early as 1924 (Hemstreet and Mallmann, 1924), just a few years after the discovery of bacteriophages (Jones *et al.*, 2012). In nearly a century since these early efforts, numerous studies have been conducted to demonstrate the usefulness of lytic bacteriophages in controlling phytopathogenic bacteria (Buttimer *et al.*, 2017; Holtappels *et al.*, 2021). Despite that several technical and legal inconveniences hinder the extensive use of bacteriophages in the agro-food industry, currently, several bacteriophage-containing antibacterial agents are already commercially available on the market for farmers in both the USA and Europe (e.g., "Agriphage" in the US, "Erwiphage" in Hungary, "Biolyse" in the UK) and these can be used either individually or in combination with other methods to protect plants from the detrimental effects of phytopathogenic bacteria (Fernández *et al.*, 2018). Additionally, many of the now limiting technical shortcomings of phage applicability for agricultural use are being actively solved, and an increase in phage-based product application in this sector is predicted owing to the lack of phage toxicity compared to less environmentally friendly non-biological alternatives without significant differences in effectiveness (Buttimer *et al.*, 2017; Holtappels *et al.*, 2021).

1.6.5. Phages in poultry and livestock production

The concepts behind "phage therapy" are also well translatable into veterinary medicine in connection with livestock and poultry, where quality demands for the end product to be used therapeutically are lower than those in humans in most legislative frameworks. Unsurprisingly, the livestock and poultry production sector, which is considered one of the most serious AMR development and leakage hotspots (Berendonk *et al.*, 2015; Kunhikannan *et al.*, 2021), also seeks effective and environmentally friendly alternatives to be used for combating bacterial pathogens as antibiotic efficacy is steadily waning (Svircev *et al.*, 2018). The phage research in connection to poultry was pioneered by none other than d'Herelle himself. In his note presented in the year 1919, d'Herelle described efforts of isolating bacteriophages that would be effective against the etiological agent of the avian typhoid outbreak that was ravaging several poultry farms (D'Herelle, 1919). Recent reviews that summarized outcomes of phage therapy in livestock, poultry, and even companion animals, conclude that phages are able to significantly reduce pathogen counts, however, the efficacy varies greatly among targeted bacteria and complexity and location of the infection sites, as well as delivery method (Desiree *et al.*, 2021; Ferriol-González and Domingo-Calap, 2021; Gigante and Atterbury, 2019; Loponte *et al.*, 2021; Mosimann *et al.*, 2021).

1.6.6. Phages in aquaculture

Phages are also being actively researched regarding their practical application possibilities for the removal of unwanted bacteria from aquaculture environments, both as a prophylaxis and a treatment option. Aquaculture is currently one of the fastest-growing food industry sectors where the development is powered by both expansion and intensification of production (Edwards, 2015; "Overview of EU aquaculture (fish farming)", 2023; Zhang *et al.*, 2022). The intensification, however, is not only causing stress in fish, which depresses their immunity, but is also often approached without adequate attention to the maintenance of water quality, general hygiene, and biosecurity measures (Bedane *et al.*, 2022; Hoseinifar *et al.*, 2020; Jia *et al.*, 2017; Ssekyanzi *et al.*, 2023). This makes aquaculture one of the sectors especially prone to bacterial disease occurrence, and these diseases are being overcome by the rather reckless use of antibiotics. This makes intensive aquaculture, especially fish farming, not only one of the sectors suffering from but also further propagating antibiotic resistance (Preena *et al.*, 2020; Reverter *et al.*, 2020; Schar *et al.*, 2021). The aquaculture setting seems especially welcoming to the use of bacteriophages due to their ability to spread via passive diffusion and eventually encounter a susceptible host (Culot *et al.*, 2019). Multiple studies investigating phage efficacy in decreasing the symptoms of bacterial diseases in aquaculture settings have been undertaken to this date (Schulz *et al.*, 2022).

While the laboratory trials tend to demonstrate the efficacy of phage therapy in aquaculture, moving up to the industrial scale use might be problematic in terms of both production and feasibility of their application (e.g., injecting a multitude of fishes with phage cocktails involves much more labor than using feed-pellets soaked in phages) (Culot *et al.*, 2019). Additionally, many more pathosystems need to be explored regarding confounding variables (e.g., species and age of the animal, disease condition, optimal phage or phage cocktail dose evaluation, etc.) and different administration routes to better understand the extent of phage therapy applicability in each particular case, eventually allowing to standardize the practices for specific use cases (Gon Choudhury *et al.*, 2017). Nevertheless, as of today, several phage-based biocontrol solutions for bacterial pathogens in aquaculture are either already available (e.g., five bacteriophage-based feed additive "BAFADOR" to control *Aeromonas* spp. and *Pseudomonas* spp.; phage-based biocontrol product "CUSTUSYRS" against *Yersinia ruckerii*, albeit the latter currently available only in Norway) or are being actively developed (see Table 2 from (Pereira *et al.*, 2022)).

1.6.7. Endolysins

Another solution for bacterial control that is being actively explored and developed for practical applications involves the use of bacteriophage endolysins (phage-encoded peptidoglycan degrading enzymes, see section 3.6 of my thesis for details) which were shown to kill the unwanted bacteria rapidly and efficiently when applied exogenously (Abdelrahman *et al.*, 2021; Murray *et al.*, 2021; Nelson *et al.*, 2001). Historically, endolysins were only used for controlling Gram-positive pathogens, as their peptidoglycan layer is not protected by an outer membrane, unlike in Gram-negative bacteria, where it hinders the enzyme's ability to access its substrate effectively (Schmelcher *et al.*, 2012). However, several studies have shown that certain compounds can render the outer membrane of Gram-negative bacteria more permeable by disrupting its structure, the so-called "outer membrane permeabilizers" (Vaara, 1992). The corpus of work advancing the understanding of outer membrane permeabilizers has inspired bioengineering efforts towards the "improvement" of phage-encoded endolysins with the goal of making them suitable for controlling Gram-negative pathogens as well. This has resulted in attempts of either using endolysins with outer membrane permeabilizing agents and/or the modification of endolysins with lipopolysaccharide-destabilizing peptides, ultimately allowing the creation of artificial proteins with strong antimicrobial activity against Gram-negative bacteria, known as "Artilysins" (Briers *et al.*, 2014; Briers and Lavigne, 2015; Gerstmans *et al.*, 2018).

2. MATERIALS AND METHODS

2.1. Sampling and initial sample processing to obtain host-phage pairs

As sampling for most of the research presented in this thesis was not carried out within a funded project but was rather driven purely by the contagious curiosity of Dr. Dišlers (and later that of myself), there was no standardized sample collection procedure.

The processing of the samples, nevertheless, was rather standardized and dependent on the sample type, for the work presented herein – almost exclusively dead insect specimens collected and processed by Dr. Dišlers long before I joined the lab. The samples for phage isolation were locally collected and were either in the form of a single insect specimen (Nocturne116 and "Three Piglet" phages), multiple specimens of the same species (Mimir87 and Mecenats66), or even an assortment of different species (Midgardsormr38). Apart from the case with phage Midgardsormr38 host, the environmental strains of bacteria that subsequently served as indicator cultures were first isolated and identified from the same insect material (different sample for host isolation used in the case of Midgardsormr38).

The collected insect samples were processed by crushing the deceased insects with a pestle and a mortar, with further resuspension in a volume of physiological saline and incubation at +4 °C overnight that allows for large particles to sediment. Afterward, small aliquots of the supernatant were spread on

Petri dishes containing simple agarized media (either "CY" (g/L: casamino acids — 6, yeast extract — 3, NaCl — 3, agar — 15) or "LB" (g/L: tryptone—10, yeast extract—5, NaCl—10, agar — 15) prepared on the spot with distilled water) and incubated for a period of up to several days at RT. Morphologically differing colonies were next picked from the plates and subcultured a few times to obtain pure environmental strains of yet unidentified bacteria. These bacterial isolates were further propagated in liquid LB media and screened for their ability to form uniform bacterial lawn in the double agar overlays using 1.5 % agar media for the bottom layer and 0.7% LB agar for the top layer, seeding the top layer agar only with the indicator cultures (Kropinski *et al.*, 2009).

To obtain cell-free viral fraction, the crushed insect suspension was usually clarified by centrifugation of benchtop centrifuge and/or filtered through syringe filters (e.g., 0.45 µm pore size syringe filter). Afterward, the aliquots of supernatant/filtrate were used for double agar overlays with the previously obtained bacterial isolates that served as indicator cultures (lawn). In the case of negative colony (plaque) presence after overnight incubation, individual plaques of differing morphology were picked and resuspended in small volumes of either buffer or physiological saline and subcultured for a few passages.

2.2. Isolated phage propagation, concentration, and purification

Purified plaques were used for scale-up propagation of phages through both liquid cultures and confluent lysis double agar overlay plates. Propagation in liquid culture involved infection of liquid bacterial culture during the early log phase (as determined by optical density changes) using the purified phage plaque material to give a multiplicity of infection of around 1 phage per host cell was routinely attempted for any given host-phage pair. However, in case of lackluster yields after several attempts while varying some growth conditions, no further actions to optimize yields from liquid cultures were undertaken, and the efforts were commonly shifted to propagation using phage extraction from the soft agar layers of multiple double agar overlay assays demonstrating confluent lysis of the lawn.

In both propagation workflows, resulting lysates were most commonly sedimented by centrifugation and the supernatant was filtered through a syringe filter to obtain cell-free phage filtrate. Filtrates were next concentrated using ultracentrifugation (e.g., up to 70,000 × g for 1 h using JA-30.50 Ti or 70 Ti rotors (Beckman Coulter)). The resulting pellets were resuspended in either a few mL of supernatant or buffer, and around 2 mL of concentrated phage sample were layered on top of ~11.5 mL CsCl solution (0.6-0.65 g CsCl per mL of buffer) in each Ultra-Clear centrifuge tube (14 × 95 mm, Beckman Coulter). Subsequent ultracentrifugation was performed under 100,000 × g for 20 h in an SW 40 Ti rotor (Beckman Coulter) at +4 °C on the Beckman Optima L-100XP ultracentrifuge. The phage-containing zones from each tube were collected by pipetting and further desalted on NAP-25/Sephadex G-25 columns (Pharmacia, Uppsala, Sweden) using PBS as an exchange buffer. Alternatively, sometimes phages were purified by gel-filtration and ion exchange chromatography and concentrated on Amicon Ultra-15, 100 K MWCO filters (Merck; e.g., 3214× g on an Eppendorf 5810R centrifuge (Eppendorf) until the desired volume). Purified and concentrated phage samples then proceeded to long-term storage in dsDNA phage collection of the Latvian Biomedical Research and Study Centre, whereas some aliquots were immediately used for subsequent experimentation.

2.3. Identification of environmental bacterial isolates

Environmental bacterial isolates on which phages could be isolated were subjected to their identification by means of 16S rRNA gene sequencing and evaluation of the resulting sequence similarity to that of reference bacterial taxa.

First, the genomic DNA of the host was mostly isolated using commercially available gDNA extraction kits (e.g., MagaZorb® DNA Mini-Prep from Promega or, more commonly, Genomic DNA Clean & Concentrator-10 Kit from Zymo Research) as per manufacturer instructions from an overnight culture. Afterwards, PCR using conventional bacterial 16S rRNA gene 27F and 1492R primers (ordered at Metabion, Steinkirchen, Germany; 27F: 5'-AGAGTTTGATCMTGGCTCAG-3';

1492R: 5'-TACGGYTACCTTGTTACGACTT-3'; (Weisburg *et al.*, 1991)) was carried out. 16S rRNA PCR product was subjected to native agarose gel electrophoresis and the band corresponding to the expected ~1450 bp product was extracted from the gel using GeneJET Gel Extraction kit (Thermo Fisher Scientific). Two reactions for the Sanger-based sequencing (Sanger *et al.*, 1977) of near-complete 16S rRNA gene were prepared using the same 27F and 1492R primers independently according to the BigDye® Terminator v3.1 Cycle Sequencing Kit instruction, and the sequencing itself was performed via ABI PRISM 3130xl system (Thermo Fisher Scientific). Sometimes, genomic DNA was sequenced directly, omitting both the PCR and extraction of the purified product from NAGE or from the PCR product, without running NAGE.

Resulting read (from 27F and 1492R primers, respectively) chromatograms were manually inspected and and had their low-quality terminal trace regions trimmed in GeneStudio (v. 2.2.0.0.). Most of the time remaining read lengths were enough to assemble them into a contig favoring higher quality traces in the read overlap region in case of ambiguities of base calls during consensus calling (if the reads did not overlap, additional sequencing reactions were prepared to obtain higher quality reads). The resulting contig representing near-complete 16S rRNA gene sequence of a given bacterial isolate was queried against established bacterial 16S rRNA sequence databases (EzBioCloud (Yoon *et al.*, 2017), The Ribosomal Database Project (RDP;(Cole *et al.*, 2014)), and the 16S RefSeq Nucleotide sequence records (O'Leary *et al.*, 2016)) enabling to classify the isolate at least at the genus level based on the hits giving alignments with least differences. Most of the time 16S rRNA gene phylogeny (either Neigbor-joining or Maximum-Likelihood) was reconstructed using the appropriate representative sequences (e.g., up to several tens of validly named hits from EzBioCloud database) to better visualize the evolutionary links of the sequence from the isolate of interest within the context of closely related bacterial species.

Simple NJ (Saitou and Nei, 1987) phylogeny reconstructions were performed using built-in MEGA utilities (Kumar *et al.*, 2016), and ML phylogeny reconstructions were performed using IQ-TREE (Nguyen *et al.*, 2015) with the best-fit substitution model being chosen automatically according to ModelFinder results (Kalyaanamoorthy *et al.*, 2017). My currently preferred multiple sequence aligner of choice is MAFFT (Katoh and Standley, 2013). Based on the resultant MSA quality, sometimes MSAs get trimmed before phylogeny building (e.g., removal of poorly aligned columns using Gblocks (Castresana, 2000)). In the case of NJ trees, multiple bootstrap test replicates are used to measure branch supports (Felsenstein, 1985), and in the case of ML trees - ultrafast bootstrap (UFBoot (Minh *et al.*, 2013)). Upon its discovery, FigTree (Rambaut, 2018) quickly became my preferred tree browser for the initial phylogenetic tree (as well as other tree-like structure) annotation and visualization.

2.4. Phage virion examination using transmission electron microscopy

Although phage samples subjected to TEM ranged from crude lysates to high-titer purified stocks according to our current immediate needs at the time, pure samples with sufficient infective virion concentration (>10^8 PFU/mL) to get multiple objects of interest in fields of view under smaller magnifications are always the desired input. This allows to quickly take multiple micrographs with a number of virions in the field of view from different parts of the grid which facilitates subsequent virion structural element dimension determination. Commonly, a few microliters of the purified sample with a high infective virion titer validated shortly before are allowed to adsorb on a Formvar/carbon-coated copper mesh grid for ~5 minutes, and are then negatively stained using 0.5% uranyl acetate and allowed to dry for several hours. Negatively stained samples are then examined using a JEM-1230 transmission electron microscope (JEOL) and the micrographs are being taken with a Morada 11 MegaPixel TEM CCD microscope-mounted camera via iTEM imaging software (Olympus). Given the different planes at which virions can lay on the grid, it is necessary to take pictures of multiple individual presumably intact virions and to do so at different magnifications.

For phage virion structural feature dimension measurements, I use ImageJ software (v1.52a; (Schneider *et al.*, 2012)) capabilities. After setting the pixel-to-nm ratio based on the scale bar from iTEM imaging software, dimensions such as capsid and tail length and width can be measured using either straight (for most of the virion structural features) or curved (for siphophage tail lengths) line utilities. Measurements of a particular feature dimension should be summarized across multiple virions, and preferably, be represented not only as an average but also be complemented with a quantitative measure of variation.

2.5. Phage whole genome sequencing and *de novo* assembly

To ensure the quality of the output data, phage DNA should be extracted from a highly purified and concentrated phage sample (although results from the concentrated phage lysate filtrates were most of the times of satisfactory quality as well; DNAse pretreatment of samples is common to remove any free DNA in the sample, although not strictly necessary). Before extracting the phage DNA using commercial kits such as Genomic DNA Clean & Concentrator-10 (Zymo Research), samples get incubated at +56 °C for 1 h with the addition of proteinase K (50 μg) and SDS (0.5% final concentration). To approximately quantify the obtained phage genomic DNA and evaluate its purity, the resulting sample aliquot first gets evaluated on a NanoDrop ND—1000 spectrophotometer (Thermo Fisher Scientific) and then diluted accordingly to verify specifically the dsDNA amount using a Qubit fluorometer (Invitrogen) dsDNA high-sensitivity quantification assay (Invitrogen).

Although during my undergraduate studies I started sequencing complete phage genomes using IonTorrent technology, I quickly switched to Illumina when I got an option to because of the added benefits of paired-end sequencing. Our current setup for cost-efficient sequencing of novel phages in small batches revolves around pooling up to 12 differently barcoded libraries using a 500-cycle MiSeq Reagent Kit v2 nano (Illumina). As input for each individual library, we use 200 ng of dsDNA that gets sonified (Covaris S220 focused-ultrasonicator with a target fragment length of 550 bp) to ensure random fragmentation of the phage genomic DNA (which is highly desirable for downstream processes to be discussed further in this thesis). Fragmented DNA next proceeds to the TruSeq DNA Nano Low Throughput Library Prep Kit (Illumina) as per the manufacturer's reference guide. Resultant libraries get quantified and qualitatively evaluated using an Agilent 2100 bioanalyzer (Agilent) with a High Sensitivity DNA kit (Agilent) and Qubit fluorometer (Invitrogen) dsDNA high-sensitivity quantification assay (Invitrogen) before being sequenced using the MiSeq system (Illumina).

My individual phage de novo assembly workflow involves multiple points of branching throughout the process. FastQC (Andrews, 2010) is used on raw reads, as well as after most manipulations involving read datasets (e.g., length and base quality trimming, normalization, and read error correction via tools from the BBMap package (Bushnell, 2014)) to allow quick visual inspection of the dataset changes in case of unexpected assembly results. Although I've tried multiple *de novo* assemblers throughout the years, my preferences shifted from MIRA 4 (Chevreux, Bastien *et al.*, 1999) to SPAdes (Bankevich *et al.*, 2012), and then to Unicycler (Wick *et al.*, 2017) (although I've yet to encounter a situation where short reads would not suffice for a high-quality assembly of a properly-prepared dsDNA phage library). I still tend to run several parallel assemblies using variations of different input datasets (e.g., untrimmed, trimmed, normalized, unnormalized, subset of a read dataset, etc.) and assemblers under different parameters, and later on compare the results from different assembly variants with QUAST (Gurevich *et al.*, 2013) and, occasionally, Bandage (Wick *et al.*, 2015) (if I don't get a single "circular" scaffold likely representing a complete phage genome, which is rare).

2.6. Assembled phage genome validation and finishing

With a firm belief that publicly shared complete phage genomes ought to be a faithful representation of the dsDNA molecule as seen within the capsid of a given phage, I have strived to make sure that my assemblies are complete from the first base pair to the last.

The circularity of the assembled genome scaffold is a good indication that the genome is complete, however, as described in the "Replication of the tailed phage genome" and "Tailed phage virion morphogenesis" sections of the literature review, all known tailed bacteriophages package linear dsDNA within their procapsids, albeit the nature of phage genome molecule physical termini tends to show considerable variation. Given that phage genomic library preparation for NGS in our workflow involves sonication as a means of fragmentation, the randomized nature of the generated library fragment starting and ending positions allows for the prediction of possible genome termini from the raw read pile-up (mapping) analyses onto the genome assembly. Read mapping patterns, and their inspection can also reveal regions with coverage dips and ambiguous base regions that might need closer manual inspection and verification, for example, using Sanger-based sequencing from the primers upstream of regions to be re-checked. However, a wonderful tool - “PhageTerm”, specifically developed for the task of performing automated analyses of read pile-up patterns (with an important caveat that the tool was designed for analyzing reads originating from the libraries prepared using randomized fragmentation of the input DNA) and starting/ending position coverage differences has shown great efficacy in correctly determining the most plausible packaging strategy phage employs, and, in case of defined molecule termini – even their sequence, and reorganizing the scaffold accordingly (Garneau *et al.*, 2017).

Additionally, I also look into the complete genomes of phages closely related to the ones I am working using public biological sequence repositories (e.g., BLASTN (Altschul *et al.*, 1990) search against NCBI Nucleotide (Sayers *et al.*, 2022) database restricted to “TaxID:10239” (Schoch *et al.*, 2020) corresponding to sequences of viral origin). In case of sufficiently high intergenomic sequence similarity (Query coverage percentage multiplied by identity percentage), peer-reviewed literature describing genome assembly and validation of the relevant hits is inspected, additionally revealing the place of the newly sequenced phages within the so far uncovered sequenced phage diversity. If there was an experimental “wet-lab” validation of the organization of the genome regarding the termini for a phage showing sufficiently high similarity to the query phage, regions of local sequence similarity are to be inspected. If there are alignments in the regions corresponding to the genome termini of such a phage – the newly sequenced phage genome can be assumed to begin in the same homologous sequence region and reorganized accordingly.

Another hint at the packaging type the phage newly sequenced phage employs (although not the exact sequence in case of exact termini) can be found from the TerL gene amino acid sequence phylogeny reconstruction within a sufficient diversity of TerL sequences corresponding to the phages for which the packaging strategies/genome termini sequences were determined experimentally. TerL aa sequences from phages employing the same packaging strategies tend to reliably cluster together in such phylogenies, allowing for a “guilt by association” approach if the TerL sequence from the newly sequenced phage falls within one of such well-defined clades. For the implementation of this approach, I tend to use a dataset of such “truth” sequences compiled by Merrill and colleagues (Merrill *et al.*, 2016) as a “core” for MSA to be generated, which gets expanded by additional manually selected “proven packaging strategy phage” sequences from the public databases for better resolution in case of need. Although this approach necessitates having at least an automatic ORF calling and respective product functional annotation done to allow for TerL aa sequence acquisition from the newly sequenced and *de novo* assembled phage of interest.

In the case of defined termini that are the same in each virion representing the phage population, it is relatively easy to validate or elucidate the genome termini completely in the wet lab, if there is at least some indication regarding the presumed location of the real termini from either read pile-up pattern analyses or sequence similarity to other phages. We approach this by designing custom primer pairs (ordered at Metabion) that ought to bind to sequences several hundred base pairs upstream of the expected genome termini and run a Sanger-based sequencing. As soon as our polymerase of choice reaches the physical end of a genome – a non-specific adenine gets added after the last specific base of the template (the so-called “glorious A”). This works well for bacteriophages that have cohesive 5’ sequences. For LDTR and SDTR phages the end of a terminal repeat, obviously, is not always marked by a non-specific base. For instance, if the length of the terminal repeat is longer than several hundred base pairs (to get it sequenced by the Sanger-based sequencing methodology we use), and the respective primer needs to be designed to bind within the repeat region, it will get amplified from both repeats when genomic DNA is used as a template (if we assume not to perform genomic DNA in-silico digestion-aided DNA restriction experiment, extract the band of choice from the restriction product NAGE, and approach sequencing of only the terminus containing fragment(s) of interest, that might be essential in some cases). In such cases, however, Sanger read chromatogram peak intensities will mark the end of a template that was read from the primer upstream of the terminus it was designed to run off from. A sudden drop of intensities that are expected to be at least two-fold less pronounced on average will begin at the base after the last base of the real sequence from the terminus of interest, while the read will continue from the second terminus into the genomic sequence of a phage, additionally, a specific single nucleotide polymorphism will be evident after the last base if the genomic sequence preceding the other repeat does not end at the same base as the non-specific base added by the polymerase. For 3' cos phages, ligation of the genomic DNA with further sequencing of the ligated region to elucidate the exact overhang sequence is necessary. In the case of headful phages, read pileups can be used to determine whether there is at least a single preferred packaging series initiation site (Headful *pac*-site utilizing phages), this then can get resolved by attempting restriction experiments to get the *pac*-containing fragment(s) and sequence it/them directly as well. Headful phages with seemingly random packaging series initiation sites will usually have uniform coverage of the genome in the raw read pile-ups, and then it is best to "open up" the pseudo circular contig around the region of the TerL gene, as is common in such cases as per convention (although some phage groups might have other conventions). In the case of exact termini employing phages, I work with the faithful representation genome sequence that is represented by having the respective termini at both ends of the genome as the respectively annotated sequence (although that might be viewed as redundant by some colleagues, this is the most faithful representation of the genome "as seen within the capsid"). In the case of headful packaging phages, terminal redundancy that shows sequence variation due to its circular permutation along the individual genomes from the phage virion population gets removed by me. For a detailed rationale behind the experimental design to determine the precise genome termini of tailed phages refer to (Casjens and Gilcrease, 2009).

Detailed implementation of (as well as interpretation of the results acquired from) the aforementioned approaches to investigate the genome organization and packaging strategy for a newly discovered phage is extensively presented in the sixth publication making up the result section of my thesis (Zrelovs *et al.*, 2022a), although also presented to an extent throughout all of the novel phage characterization papers I co-authored (including those that are not part of this thesis (Akmal *et al.*, 2022; Korniienko *et al.*, 2022)).

2.7. Novel bacteriophage genome annotation

Despite trying out numerous tools that provide phage auto-annotation capabilities, the quality of the resultant annotation always proved to be sub-par compared to the manually curated annotations our lab prefers by a margin. Usually, for novel phages from our lab at least two people are performing the tasks related to the same phage sequence manually curated annotations (e.g., either myself and Dr. Kazāks, or myself and a student of mine), the resultant genome annotations in the GenBank (*.gb) format are then compared by a small python in-house script I wrote to compare the called genomic feature absence/presence, as well as differences in the start codon positions for the same ORFs. This is followed by a discussion of the differences by both annotators, collectively reweighing the evidence for this or that call presence or function of an ORF product in case of incongruencies to prepare the submission-ready complete annotated phage genome to be sent to one of the INSDC databases (preferring NCBI's GenBank (Sayers *et al.*, 2019)) for consideration and eventual public release.

"DNAmaster" (https://phagesdb.org/DNAMaster/) is my current sequence explorer of choice for annotating novel phage genomes due to multiple relevant features being built in, although essentially any sequence explorer having basic functionality of translating nucleotide sequences and allowing to annotate genomic regions can suffice (Pope and Jacobs-Sera, 2018). As a first step, I am running automated ORF (Glimmer (Delcher *et al.*, 2007) and GeneMark (Besemer and Borodovsky, 2005)) and tRNA gene (ARAGORN (Laslett and Canback, 2004) and tRNAscan-SE (Lowe and Eddy, 1997)) prediction with the help of the respective tool implementations within the DNAmaster genome browser (v 5.23.6, at the time of writing). Four possible start codons (ATG, GTG, CTG, and TTG) are being allowed and only putative ORFs encoding a product of >30 amino acids (aa) in length are being considered further. When choosing a putative start codon for any given ORF, initial preference is usually given to the start codon corresponding to the longest possible ORF product, however, no long ORF overlaps and "gene within gene" arrangements are allowed for most genes (e.g., spanin genes for Gram-negative host phages is one of the most obvious exceptions). Next, for functional annotation, products of the predicted ORFs are queried against the NCBI conserved domain database (CDD) using NCBI conserved domain search (Marchler-Bauer *et al.*, 2011), against the non-redundant protein sequence database using BLASTp (Altschul *et al.*, 1990), against the Protein Data Bank (PDB), Pfam, UniProt-SwissProt-viral70 and NCBI CD databases using HHpred (Söding *et al.*, 2005). Additionally, TMHMM (Krogh *et al.*, 2001) and Phobius (Käll *et al.*, 2007) web servers are used in cases when the ORFs responsible for coding the expected lysis proteins were not identified by previous means. Start codons of ORFs are next corrected by comparison of the sequence with close homologs from other phages or bacteria plus inspection of the upstream regions of all the plausible alternative start codons for the presence of putative Shine-Dalgarno (SD) sequences. A change in the free energy (ΔG, kcal/mol) required to bring two strands of nucleotides together (putative SD-containing sequence upstream of the start codon and an aSD-containing 16S rRNA tail of the host species) is then calculated using "free_align.pl" script (Starmer *et al.*, 2006) ("helix-only" mode, assuming G-U wobble base pairing and hybridization taking place at +37 °C, considering sequences with ΔG of more than −3.60 kcal/mol to be subpar SD sequences).

2.8. Placing the novel phage within the context of the so far isolated phage diversity

Identification of related sequenced phages whose complete genomes are publicly available gets determined by querying the novel phage genome using BLASTN (Altschul *et al.*, 1990) search against NCBI Nucleotide (Sayers *et al.*, 2022) database restricted to "TaxID:10239" - viruses (Schoch *et al.*, 2020). A representative set of hit sequences gets selected to represent a wider context and the respective genomes are downloaded both in the form of *.gb and *.fasta files for further comparative analyses. Pairwise intergenomic sequence similarities/distances between all the phages in such dataset get calculated using VIRIDIC (Moraru *et al.*, 2020) under default parameters to identify approximate phage genus and species level clusters as per mechanistic ICTV-set criteria (more than ~70% intergenomic nucleotide sequence similarity means that phages belong to the same phage genus, more than ~95% phages represent different isolates of the same species (Turner *et al.*, 2021)).

Most intergenomically similar phages, ideally, with the standing in current phage taxonomy, and well-studied representatives from a larger group (if relevant) are next selected for comparison of respective genome organizations and synteny, as well as similarity of the underlying sequences. For this, genomic features present in the downloaded *.gb files are marked according to their functional group (e.g., ORFs coding for products involved in host cell lysis have different "color" identified added than ORFs involved in virion morphogenesis, etc.). These files are next used as input for EasyFig (Sullivan *et al.*, 2011) to visually compare the genome organizations and the underlying nucleotide and translated nucleotide sequences (with BLASTN and tBLASTx, respectively). Additionally, a comparison of genome organizations and visualization of ORF product similarities gets carried out using Clinker (Gilchrist and Chooi, 2021). For quick pangenome overview analyses to identify "core" genes homologous within a group of phages, I tend to use Roary (Page *et al.*, 2015) under different BLASTP protein similarity thresholds.

I try to also reconstruct individual phylogenetic trees for the selected phage marker proteins (usually a few proteins that are, ideally, functionally independent of each other; e.g., MCP, TerL, DNAP, endolysin). For this, usually, I do a BLASTP search against the non-redundant protein sequences of viral origin (TaxID:10239) under other default settings, to select both the top highest scoring hits, as well as more distantly related hits that are approximately the same length as a query. Subsequently, phylogenies get reconstructed as described in the last paragraph of the Materials and Methods subsection "Identification of environmental bacterial isolates".

Recently, I also started to employ proteome-content-based clustering of phages via vConTACT2 (Bin Jang *et al.*, 2019) with the result representation in a network format via Cytoscape (Shannon *et al.*, 2003).

3. RESULTS

3.1. Isolation and characterization of the novel *Virgibacillus*-infecting bacteriophage Mimir87

Highlights (Zrelovs *et al.*, 2020a)**:**

- A bacterial isolate identified as a strain of *Virgibacillus halotolerans* was isolated from dead worker honeybees alongside a siphophage shown to infect it which was subsequently named Mimir87.
- The complete genome of *Virgibacillus* phage Mimir87 comprises 48,016 base pairs and was predicted to encode 68 proteins, for half of which no functional assignment could be made at the time.
- Despite the Mimir87 genome encoding several proteins associated with the temperate lifestyle, no convincing lysogenic behavior of the phage was observed.
- Exact genome physical termini determination revealed that the Mimir87 genome molecule has the following 3' cohesive overhang sequence: 5'- CGGATGGGC-3'.
- *Virgibacillus* phage Mimir87 represents the first known tailed bacteriophage that was shown to infect a representative of the bacterial genus *Virgibacillus*, and the first *Virgibacillus* phage to have its complete genome elucidated.
- Based on the ICTV-set intergenomic similarity phage genus and species level demarcation criteria, Mimir87 could represent both a novel phage species and genus, as the sole founding member for both tentative taxa.
- The complete annotated genome of *Virgibacillus* phage Mimir87 was deposited to GenBank and is publicly available under the accession number MK560763.

3.2. Novel *Erwinia persicina* Infecting Phage Midgardsormr38 Within the Context of Temperate Erwinia Phages

Highlights (Zrelovs *et al.*, 2020b)**:**

- A bacterial isolate identified as a strain of *Erwinia persicina* was isolated from dead fruit flies, after, a siphophage able to infect this bacterial isolate was retrieved from locally collected mixed dead insects (houseflies, lady beetles, green lacewings) and was subsequently named Midgardsormr38.
- The complete genome of *Erwinia* phage Midgardsormr38 comprises 50,485 base pairs and was predicted to encode 93 proteins, for 38 of the ORFs (~40.9%) no functional assignment could be made at the time.
- Midgardsormr38 genome encoded multiple proteins associated with the temperate lifestyle, most of which seemed to be analogous to their counterparts from phage λ.
- Midgardsormr38 prophage-containing *E. persicina* host strain lysogens were unstable, spontaneous induction was observed relatively frequently under laboratory conditions.
- Midgardsormr38 genomes are likely packaged into the procapsids in a "headful" manner, resulting in the progeny virions having circularly permuted genomes with terminal redundancy of ~0.4% non-redundant genome length that is unique to each virion.
- Isolation of Midgardsormr38 and scarcity of the completely sequenced temperate *Erwinia* spp. phages prompted to perform analysis of the prophages found within the complete *Erwinia* genomes.
- Up to 28 prophage regions totaling 17.67% of the chromosome length (4,891,733 bp) were predicted in the genome of *E. tracheiphila* strain MDcuke (CP013970.1).
- A high false-positive rate among putative "intact" prophage regions called by the prediction algorithm used was noted after inspecting functional annotations of the ORF products within

such regions that revealed an absence of many essential phage proteins vital for the prophage to be induced and self-sufficiently go through the lytic cycle.

- *Erwinia* phage Midgardsormr38 represents the first cultured and sequenced temperate *E. persicina*-infecting phage.
- The complete annotated genome of *Erwinia* phage Midgardsormr38 was deposited to GenBank and is publicly available under the accession number MN602881.

P.S. *Erwinia* phage Midgardsormr38 was used in an ICTV taxonomy proposal "Create a new genus (*Midgardsormrvirus*) with a single species (*Caudoviricetes*)" under assigned code "2022.051B" in May 2022 by Turner D., Moraru C., Kropinski A.M. and is now a part of the official virus taxonomy as a single representative both the *Midgardsormrvirus midgardsormr38* phage species and the *Midgardsormrvirus* phage genus within the class *Caudoviricetes*, while yet remaining *incertae sedis* at the order and family levels.

3.3. Motley crew: Overview of the currently available phage diversity

Highlights (Zrelovs *et al.*, 2021a, 2020c)**:**

- As of June 2, 2020, the NCBI Nucleotide database, when filtered to show only complete bacteriophage genomes, listed 13,132 entries that represented individual phage complete genomes as stated by the respective submission authors (deduplication at the approximate mechanistic intergenomic nucleotide sequence similarity criterion for phage species of ~95% has resulted in 8,245 putative phage species genomes).
- At the time, complete phage genomes were scattered throughout 19 phage families, the dominant majority of which were tailed dsDNA-containing phages from the now obsolete order of *Caudovirales* (~93.6% of the ICTV-recognized or tentative species).
- Host-wise, the so far cultured and completely sequenced bacteriophage diversity was shown to be highly biased towards bacteria of healthcare or economic importance (Top 5 hosts – *Mycobacterium* spp., *Streptococcus* spp., *Escherichia* spp., *Pseudomonas* spp., *Salmonella* spp.).
- At the time, phages infecting only 219 different bacterial genera representatives (based on the submission-associated metadata) were completely sequenced, 74 of these genera had only a single completely sequenced phage known to infect them.
- After deduplication, the counts of genomes representing phage species had a distribution showing "gravitation" toward three size ranges: "small-sized" genomes (<25 kbp) were most frequently 10-15 kbp in length, "medium-sized" genomes (25–100 kbp) - 40-45 kbp in length, whereas "large-sized" (>100 kbp) genomes had a peak at the 150–175 kbp length range.
- Some serious yet easily avoidable issues that preclude the usability of the data "as is" for both summarizing overviews and single entry-oriented analyses, were noted throughout a number of the submitted and publicly available individual phage complete genome entries:
    1. Typing errors (e.g., Panteoa instead of *Pantoea,* Eschericha instead of *Escherichia*, *Mycobacterium* misspelled in multiple ways, etc.)
    2. Misuse of GenBank metadata qualifiers (e.g., "/host = " qualifier of phage complete genome entries used for organisms other than bacteria; presumably from metagenomic/metaviromic surveys of higher organisms)
    3. Lack of any functional annotation of the genomic sequence submitted to the public biological sequence repositories (e.g., no ORFs, tRNAs annotated) or suboptimal functional annotations (e.g., meaningless gene product functional assignments; missing annotations for highly conserved gene products; utilization of auto-annotation software without post-hoc sanity checks, etc.)
- Many of the currently recognized bacterial genera do not yet have the phage that infects its members described (while there is no reason to think that such phages do not exist).

For the expansion of the known phage diversity, traditional culture-based approaches that show an unambiguous phage-host association and allow for in-depth studies of particular phages should not be abandoned in favor of purely culture-free approaches.

3.4. Genome Characterization of Nocturne116, Novel *Lactococcus lactis*-Infecting Phage Isolated from Moth

Highlights (Zrelovs *et al.*, 2021b)**:**

- A bacterial isolate identified as a strain of *Lactococcus lactis* was isolated from a deceased local moth specimen alongside a siphophage (subsequently named Nocturne116) that was able to infect this bacterial strain.
- The complete genome of *Lactococcus* phage Nocturne116 comprises 25,554 base pairs and was predicted to encode 52 proteins, for 37 of the ORFs (~71.2%) no functional assignment could be made at the time.
- Exact genome physical termini determination revealed that the Nocturne116 genome molecule has the following 3' cohesive overhang sequence: 5'-CGCAGTAACT-3'.
- Intact virions of Nocturne116 demonstrate uncommon siphovirus morphology with a prolate capsid (59.1 ± 1.5 nm × 40.1 ± 1.2 nm) to which a non-contractile tail is attached (112.0 ± 8.1 nm long and 9.5 ± 0.8 nm wide).
- *Lactococcus* phage Nocturne116 was most similar to the *Lactococcus* phage Q54 (albeit demonstrating only ~24.5% intergenomic nucleotide sequence similarity) which had a comparable genome size, similar genome organization, exactly the same 3' cohesive genome terminal overhang sequence, and similar virion feature dimensions).
- The proteome of Nocturne116 seemed rather interesting: 22 out of 52 predicted proteins had no homologs available in the public amino acid sequence databases, and another 15 ORF product homologs could be found only in the proteome of Lactococcus phage Q54.
- Despite the genus *Lactococcus* being one of the most popular host genera for cultured phages completely sequenced (nearly 350 phage genome entries available at the time of writing), *Lactococcus* phage Nocturne116 tentatively represents a novel phage genus as its sole representative as of now.
- The so far cultured and completely sequenced *Lactococcus* phage diversity was highly biased towards representatives of just two phage genera – *Skunavirus* and *Ceduovirus*.
- A detailed description of the Nocturne116 genome annotation process using freely available tools, as well as elaboration on the rationale behind assigning its ORF-encoded products with a function, was given for a phage that had no close relatives described yet, which made it an uneasy annotation target even employing a manual supervised annotation approach.
- The complete annotated genome of *Lactococcus* phage Nocturne116 was deposited to GenBank and is publicly available under the accession number MW791312.

3.5. PVJ1 Is Not the First Tailed Temperate Phage Infecting Bacteria from Genus *Psychrobacillus*. Comment on Liu et al. Isolation and Characterization of the First Temperate Virus Infecting *Psychrobacillus* from Marine Sediments. Viruses 2022, 14, 108

Highlights (Zrelovs *et al.*, 2022b)**:**

- To our knowledge, temperate siphophages Perkons (MT325768) and Spoks (MT410774) are the first phages capable of infecting a bacterium belonging to the genus *Psychrobacillus* that had their complete annotated genomes publicly available.
- Even though not formally published in the peer-reviewed literature yet, the process of isolation and characterization of these phages has been a major part of my MSc. thesis on the isolation and characterization of phages for bacterial genera, which had no known phages able to infect them

yet, so as to expand the known cultured phage diversity, defended at the University of Latvia in summer 2020.

- Despite the fact that there is a large fraction of complete annotated phage genome entries publicly available that still do not have an article or at least a genome announcement from the peer-reviewed literature linked to them for a variety of reasons (including *Psychrobacillus* phages Perkons and Spoks), it is incorrect to ignore the existence of such entries when analyzing the place of any newly isolated phage within the context of known phages, regardless of the phages not being mentioned anywhere in the peer-reviewed literature.

3.6. *Morganella* Phage Mecenats66 Utilizes an Evolutionarily Distinct Subtype of Headful Genome Packaging with a Preferred Packaging Initiation Site

Highlights (Zrelovs *et al.*, 2022a)**:**

- A bacterial isolate identified as a strain representing the bacterial genus *Morganella* was isolated from dead worker honeybees, after, a myovirus (subsequently named Mecenats66) that was able to infect this bacterial strain was isolated from the same material.
- The complete genome of *Morganella* phage Mecenats66 was 86,193 base pair-long and was predicted to encode 123 proteins, for 84 of the ORFs (~68.3%) remained without any plausible functional annotation for their product at the time.
- From the proteome of phage Mecenats66, 52 putative product amino acid sequences had no BLASTp hits to proteins encoded by other publicly available phage genomes.
- Phage particles of Mecenats66 showed myovirus morphology with an icosahedral capsid of 75.6 ± 3.9 nm in diameter attached to a 142.1 ± 3.7 nm long and 19.7 ± 1.7 nm wide contractile tail (dimensions of an uncontracted tail are provided) that ends with a “lush” basal plate.
- *Morganella* phage Mecenats66 demonstrated remarkable intergenomic nucleotide sequence distance to any of the bacteriophages sequenced at the time (less than 3% intergenomic similarity to any viral entry in GenBank).
- Phage Mecenats66 seems to be the ninth cultured virus capable of infecting a representative of the bacterial genus *Morganella*.
- Genome physical termini scrutinization performed for *Morganella* phage Mecenats66 suggested that it may employ an evolutionarily distinct headful genome packaging strategy with a preferred packaging initiation site.
- As there seem to be clusters of sequences showing plausible MRCAs that seem to be evolutionarily quite distant from any of the experimentally verified packaging strategy TerL/terminase sequence-containing clades, elucidation of the genome termini, at least for some of the phages falling within those “unknown packaging strategy” clades, might further refine the applicability of TerL-phylogeny based approaches for the packaging strategy prediction of novel phages.
- The complete annotated genome of *Morganella* phage Mecenats66 was deposited to GenBank and is publicly available under the accession number OL614104.

3.7. Three Phages One Host: Isolation and Characterization of *Pantoea agglomerans* Phages from a Grasshopper Specimen

Highlights (Zrelovs *et al.*, 2023)**:**

- A bacterial isolate identified as a strain of *Pantoea agglomerans* was isolated from the crushed dead specimen of an unidentified Latvian grasshopper species.
- Three phage plaque lines demonstrating different appearances on the lawn of the same host in double agar overlay and presumably representing distinct phages were recovered from the same source material.

- WGS revealed that the isolated phages, indeed, are distinct of each other – the phages were subsequently named taking inspiration from the Latvian rendition of the "Three Piglets" fable (Nifs112, Nufs112, and Nafs113):
  - Podovirus Nifs112 has a 46,202 bp genome (including the 296 bp short direct terminal repeats) with G+C content of 50.2%, comprising 59 ORFs, with 28 of them (~47.5%) encoding "hypothetical proteins".
  - Podovirus Nufs112 has a 45,951 bp genome (including the 410 bp short direct terminal repeats) with G+C content of 47.7%, comprising 67 ORFs, with 36 of them (~52.2%) encoding "hypothetical proteins".
  - Myovirus Nafs113 is thought to employ a headful packaging strategy resulting in circularly permuted terminally redundant genomes inside the population of virions, non-redundant genome of Nafs113 is 75,899 bp long and boasts a G+C content of 54.1% and hosts 130 ORFs, 91 of which (70%) encode "hypothetical proteins".
- TEM allowed to determine features of the aforementioned phage virions:
  - Podophage Nifs112 has an icosahedral capsid of 57.2 ± 2.4 nm in diameter to which an 11.2 ± 1.6 nm short tail is attached.
  - Podophage Nufs112 has an icosahedral capsid of 54.4 ± 1.8 nm in diameter to which a tail of 12.8 ± 2 nm is attached.
  - Although very similar in their virion feature dimensions, the tails of both isolated podophages had different appearances with Nifs112 having seemingly slimmer tails in contrast to the "chunky stubs" of Nufs112.
  - Myophage Nafs113 was determined to have a less common morphotype with an elongated head of approximately 3:1 length (121.4 ± 6.8 nm) to width (43.4 ± 3.8 nm) ratio, to which an 88.5 ± 3.2 nm long contractile tail with a width of 15.4 ± 0.6 nm (uncontracted state) was attached.
- "Three piglet phages" had different degrees of novelty within the context of the so far uncovered phage diversity:
  - *Pantoea* phage Nifs112 might represent a novel species within the phage family *Autographiviridae* subfamily *Molineuxvirinae*, genus *Eracentumvirus*.
  - *Pantoea* phage Nufs112 is distinct enough to serve as an exemplar isolate for a tentative novel species that would represent a new phage genus within the family *Autographivirdae*.
  - *Pantoea* phage Nafs113 is very distantly related to any of the so-far completely sequenced phages that are publicly available, except for the *Pantoea* phage vB_PagM_LIET2, suggesting that Nafs113 should be included within the genus *Lietduovirus*, and even proposing to treat it as an isolate of the species *Lietduovirus LIET2* for practicality reasons at the time being despite it falling below the mechanistic intergenomic similarity species demarcation criterion.

4. DISCUSSION

In the era of metagenomics and metaviromics, one should always remember that sequence data alone is worth only as much as the information that can be deduced from it. It is nearly exclusively to the previous wet lab experimental work that we are currently able to interpret such data in at least partially meaningful ways. Despite all the hype around metaviromics, so many of its current limitations are to be solved using exactly wet-lab, including culture-based, approaches. The positive thing, however, is that such data will constantly grow in its value step-by-step, with every novel phage isolated, with every protein "domain of unknown function" having its function elucidated, with fewer and fewer "hypothetical proteins" remaining in the global phage panproteome. Re-analysis of the metaviromes generated up to this day will undoubtedly provide a lot more reliable insight in the future, but if, and only if, the viruses now represented as merely a sequence will get further attention in the wet lab.

The current pitfalls of culture-independent phage diversity research (including the inability to answer a fundamental question for characterization of any virus – "who serves as the host?" reliably) clearly demonstrate that culture-dependent methods are indispensable for advancing our current knowledge about bacteriophages. Thus, I believe that the term "effective phage diversity" might be necessary to introduce for description only of the completely sequenced bacteriophages that are available to researchers in the form of actual isolates accompanied by the isolation host strain, which can be used for further experimentation, rather than a phage genomic sequence alone, regardless of its completion. The modest results expanding the phage diversity acquired thus far in our laboratory (only a small part of which are published to this date in peer-reviewed literature) demonstrate that at least some of the biases in effective tailed phage diversity can be relatively easily overcome. A pattern that novel phages are actively isolated for representatives of bacterial genera, species, or even particular strains of healthcare and/or economic importance is understandable in light of such phage potential practical applications that might add up to the solution to the pressing challenges pathogenic bacteria put forward. However, this creates obvious biases in the effective phage diversity, resulting in a multitude of times higher saturation of such diversity for phages infecting select hosts. At the same time, an overwhelming number of recognized bacteria have no phages known to infect them at all, while there is completely no reason to assume such phages do not exist. Interestingly, the effective diversity of phages infecting some of the non-negligible (opportunistic) bacterial pathogens is also currently either very limited (e.g., *Morganella* (Liu *et al.*, 2016), *Tsukamurella* (Safaei *et al.*, 2018), *Brevundimonas* (Ryan and Pembroke, 2018), *Kluyvera* (Lee *et al.*, 2019)) or lacking at all (e.g., *Alloscardovia*, *Capnocytophaga*, number of intracellular pathogens such as *Rickettsia*, etc. (Vouga and Greub, 2016)) to the best of my knowledge. Such bacteria, obviously, look like very appropriate targets to enroll as hosts for phage-hunting campaigns (although the difficulty of microbiological handling is probably the reason that explains the lack of extensive data on phages against at least some of the mentioned bacteria (Śliwka *et al.*, 2022)). The absence of appropriate conditions allowing to efficiently work with strictly anaerobic bacteria and their phages in many labs is also one of the limiting factors that creates the observed effective phage diversity biases. I believe that efforts on isolating bacteriophages for less commonplace hosts, even if such hosts are not of prime economic or healthcare importance to be pursued in light of applied research, are still of utmost importance to advance the global knowledge about the virosphere. Additionally, I feel that the leading public biological sequence repositories (e.g., GenBank) should be stricter regarding the quality control of the entries being submitted into them (including metadata-wise) and better disambiguate between the metagenome-assembled genomes (MAGs) of phages and the complete genomes derived from actual phage isolates that were cultured to enable more precise and robust evaluations of the effective phage diversity without the necessity for an extensive manual curation from the user side with every database update. Moreover, the "effective phage diversity" is, in my opinion, an idea for a very useful database waiting for its developer/creator. While INPHARED is extremely useful as an aggregator database for qualitative phage genomes, to my liking, the *sensu lato* interpretation of the term bacteriophage with extension to viruses of archaea, as well as permissiveness to metagenome-assembled genomes, leaves room for further improvement of the concept behind it to attempt the creation of a manually curated regularly updated database comprising strictly cultured phage complete genome sequences (maintaining an inclusion, rather than exclusion list).

Working in a purely curiosity-driven fashion without having to worry about reaching the promised “key performance indicators” for yet another project is highly satisfying and allows a great degree of freedom to play around in the lab, although comes with a downside of having very limited resources for such leisure “side projects” in terms of both allocated funds and time available. Most of the phages isolated in our lab for a long time, including all those mentioned within this thesis (Appendix 2), are a result of the productive leisure work of Dr. A. Dišlers, who happened to eventually create a Latvian Biomedical Research and Study Centre’s tailed phage collection now comprising many tens of phages infecting very different environmental bacterial isolates. Many of the phages from this collection demonstrate a very high degree of divergence from any of the phages representing effective phage diversity available to the researchers so far. This allows to outline several important thoughts that are assumed to be behind the ability to rather consistently recover phages that are divergent from the ones already known, and to do so in a cost-efficient manner: I) Uncommon sources hide undersampled biological diversity; II) Hosts come first; III) If there is a bacterium, there will be a phage infecting it. One may argue that such an approach is bizarre and akin to “stamp collecting”, yet both “stamp collecting” and isolation of phages in such a fashion are valid hobbies. As can be seen from the acquired results, such a genuine curiosity-driven approach to the expansion of effective phage diversity without any particular preformulated aim goes a long way toward better understanding bacteriophages. Each isolated phage that gets isolated and characterized means that there is one less phage out there in the virosphere to “cut down” for the extension of the effective phage diversity.

Guided by these principles we were able to isolate, completely sequence and annotate, as well as publicly release, to our knowledge, the first genomes of tailed phages infecting *Virgibacillus* spp. (Zrelovs *et al.*, 2020a), *Psychrobacillus* spp. (Zrelovs, 2020; Zrelovs *et al.*, 2022b), *Sporosarcina* spp. (Zrelovs, 2020), *Cedecea* spp. (Jonikāne, 2022). Additionally, many other intergenomically distinct or otherwise interesting tailed phages were isolated, completely sequenced and publicly released, focusing mostly on working with less commonplace hosts retrieved from the environment. I believe that such a curiosity-driven approach to novel phage isolation and characterization is also perfect for engaging undergraduate students in what becomes their first real research endeavors with useful outcomes of their research that might further spark ideas about building upon the acquired results. Throughout the years I’ve had the pleasure to serve as an advisor and lead several undergraduate student phage research projects that have resulted in successfully defended Bachelor theses that had their own highlights:

- Karīna Švānberga focused on the isolation of *Aeromonas salmonicida* phages from wastewater samples, successfully retrieving and characterizing a distinct *Autographiviridae* podophage (JELG-KS1, ON604651) despite around 200 *Aeromonas* phages being isolated prior (Švānberga, 2022). This stimulated us to find local collaboration partners and pursue funding for a pilot project aimed evaluation of the possibilities to develop a phage cocktail prototype against local *Aeromonas* spp. strains from aquaculture settings, and now, in the second year of her Master’s research project that I am supervising, Karīna managed to isolate more than 30 phages using 20 different *Aeromonas* strains. These phages were sequenced and most of them seem genetically safe to use in biocontrol, however, they are yet to undergo further microbiological characterization to evaluate their potential for such applications.
- Agneta Jonikāne looked for phages infecting an environmental *Cedecea davisae* strain and characterized them, which resulted in the public availability of the first complete genome for a phage infecting *Cedecea* spp. (*Autographiviridae* podophage Yanou, ON568193) (Jonikāne, 2022).
- Beāte Galvāne worked with soil sampled in the yard of our institute and managed to isolate the first *Salasmaviridae* phage for our lab which also happens to be the first known *Viridibacillus*-infecting phage (*Salasmaviridae* podophage Skaradils, the complete annotated genome not yet submitted to public biological sequence repositories) (Galvāne, 2023).

The ideal case to advance our labs vision of phage diversity research would definitely be to bring at least some of such activities to the university classrooms/labs as a dedicated lecture/practical course that would span several terms and would be based on teaching wider undergraduate audiences by performing actual research involving novel phages (e.g., getting inspiration from the massively successful SEA-PHAGES program but focusing on many unusual hosts rather than a few selected ones (Hanauer *et al.*, 2017)).

Generation of novel genotype-phenotype associations is highly sought after for many bacterial genera and their phages to fully unleash the potential of bacteriophages as effective antimicrobials. Moreover, systematic data on different hosts and their phages is of importance also for the development and perfecting of the metaviromics analysis approach utility, as training models based on such data will allow for better predictions of particular phage features from the sequence data alone. Focusing down on a single less commonplace host genus representatives and retrieving a substantial diversity of both the respective bacterial strains and phages infecting them are of higher value than describing merely a single/few isolated phage-host strain pairs. Systematic studies of less commonplace hosts and their phages are necessary to better understand both the hosts and the viruses generally, yet such efforts of "going one's own way" mean building up from way less knowledge already available on a particular system in contrast to the model systems actively researched in this regard (e.g., mycobacteriophages, coliphages, *Pseudomonas* phages, *Klebsiella* phages).

The creation of DNA bacteriophage collection at the Latvian Biomedical Research and Study Centre (BMC) by Dr. Dišlers dates back to the year 2006. Today this collection consists of more than 70 different bacteriophage isolates from various environmental sources that are already described to a different degree throughout the years of investigation at LV BMC. The majority of the isolates in the collection have already been sequenced, and their annotated genomes are being systematically deposited to the publicly available biological sequence repositories. The most interesting isolates are being profoundly characterized and their descriptions are being published in relevant scientific journals of the field. A plethora of unique bacteriophages that represent novel phage genera were isolated during these studies and have aided in the expansion of the globally known effective phage diversity. The phages presented within this thesis (Appendix 2) clearly demonstrate that very distinct phages can be readily retrieved from different environments along with their hosts, although these represent only a small fraction of the phages stored in our collection.

Before I joined the Latvian Biomedical Research and Study Centre, tailed phage WGS and their genomic characterization aspects were almost absent in Latvia (single phage Enc34 sequenced and the genome publicly released accompanied by the genome announcement (Kazaks *et al.*, 2012)), and characterization of the phages from the collection of our institute was largely confined to microbiological assays and TEM. I take great pride in the fact that both my long-term co-supervisors Dr. Dišlers and Dr. Kazāks eventually entrusted me with co-developing and further perfecting the dsDNA phage NGS, *de novo* assembly, and genomic characterization workflow we currently use to better acknowledge the phage diversity in the surrounding environments. Currently, complete and annotated genomes of only 16 novel dsDNA tailed phages that went through our lab are publicly available at INSDC databases, whereas at least 40 more phages from our lab were additionally sequenced but were not yet annotated in detail and deposited publicly. With the only funding for these activities at the time coming from a year-long project "Characterization of DNA bacteriophages isolated from the insect intestine microflora" awarded to Dr. Dišlers (lzp-2020/2-0373). I am very happy that already at this stage my phage genomics expertise has managed to gain some attention of colleagues from abroad, with whom I eventually had the pleasure of collaborating, co-writing, and co-publishing articles on the phages they've isolated and I've helped to sequence and/or genomically characterize (Akmal *et al.*, 2022; Korniienko *et al.*, 2022). I genuinely enjoy sharing or applying my expertise when it can be of use and hope for more such collaborations in the future. I, however, feel very ashamed of the fact that, despite clear indications that many of our novel phages seem to meet the criteria to represent founding members of novel phage genera and/or species, I never managed to find time for the preparation of the official taxonomy proposals to be sent to the ICTV BVS. Preparation of such proposals is definitely a job that needs to be done to increase the impact our findings might have within the wider research community.

The importance of novel unique phages is also hidden in the fact that their genomes represent a "vault" of hypothetical proteins with a yet unknown function. Thus, in addition to genomic studies of phages and their diversity, we have also begun the structural-functional elucidation of the biological role of the selected proteins found in this "vault" to gain a better understanding of their role in phage and their host lifecycles. That way a colleague Elīna Černooka was able to determine the structure and function of several interesting phage proteins, with the findings being expandible to their conserved counterparts from other biological objects (e.g., DUF2815-containing hypothetical protein from phage Enc34 function as a single-stranded DNA binding protein, which is extendible to homologs from other phages (Cernooka *et al.*, 2017)). Later on, also spearheaded by E. Černooka, a structure of an unusual endolysin encoded by phage Enc34 was solved (Cernooka *et al.*, 2022). In addition to the fundamental research on bacteriophage biology, recently, attempts to evaluate the potential of the practical application of studied phages or their genome-derived products have been started. Currently, our lab is also in the middle of a three-year project devoted to "Molecular design of DNA phage-derived endolysins for targeting Gram-negative bacteria" (lzp-2021/1-0050, awarded to Dr.biol. Tatjana Kazāka). Within this project we strive to improve the lytic potential of a novel type of globular endolysin from our phage Enc34 (Cernooka *et al.*, 2022), as well as several selected modular endolysins of other phages, by modifying their activity to allow lysis of gram-negative hosts when applied from without.

Building upon the results of genomic characterization of the novel bacteriophages isolated from Antarctic ice-free soil samples during my MSc. studies (Zrelovs, 2020), we have recently managed to secure a three-year project "Structural and functional studies of proteome encoded by three novel temperate DNA bacteriophages isolated from Antarctic soil" (lzp-2022/1-0111, awarded to Dr. biol. A. Kazāks). Pangenome of these three phages (*Psychrobacillus* phages Perkons and Spoks presented as part of this thesis, and *Sporosarcina phage* Lietuvens not included in this thesis) encoded totally for 370 predicted proteins for 239 of which the exact function could not be predicted by similarity searches through the public biological sequence databases. Among these "dark matter" proteins there is a considerable number of hypothetical proteins and proteins with domains of unknown function (DUFs) that show a degree of conservation and are found in proteomes of other phages or even their hosts. Thus, we aim to solve the structures and attempt to determine the functions for several such proteins, which would allow us to extend the function to a broader range of homologous proteins. Additionally, we plan to resolve the temporal transcription pattern of these phages upon infection of the host by implementation of the phage+host system transcriptome sequencing at different times post-infection (reasonable sampling times to be determined by one-step growth curves). Currently, such RNA sequencing activities seem interesting due to the fact that only several tens of phages seem to have temporally well-resolved transcriptomes available (Clokie *et al.*, 2020).

In addition to the plans related to the already ongoing projects, we attempt to further pursue basic dsDNA phage research funding to allow for the characterization of other insect-associated bacteria and phages – insects represent a very undersampled source from which many more interesting phages are expected to be relatively easily retrieved. Speaking of applied direction, we will likely try to find ways to extend our investigations of phages infecting *Aeromonas* isolated from aquaculture settings to not only try to evaluate the possibilities for the creation of a locally relevant phage-based biocontrol product prototype (phage cocktail) but also pursue in-depth investigations of the phage-host interactions that underpin success or failure of such applications.

As can be seen from the information above, what was kickstarted by my teacher and thesis co-supervisor Dr. Andris Dišlers from scratch and happened in a curiosity-driven fashion without any dedicated funding for a long time, has resulted in the establishment of a dsDNA phage research direction novel within Latvia. In 2021 our efforts received a diploma of appreciation from the Latvian Academy of Sciences for the "Development of a novel DNA bacteriophage isolation and characterization research direction in Latvia" (awarded to Mg.biol. Ņikita Zrelovs, Mg.biol. Elīna Černooka, Dr.biol. Jānis Rūmnieks, Dr.biol. Andris Kazāks, Dr.biol. Andris Dišlers). Although largely enthusiasm-driven, the collection of the dsDNA phages from the Latvian Biomedical Research and Study Centre is expected to incessantly get expanded by the novel phage host pairs, and the ready availability of interesting phages inevitably leads to new ideas of what do to with them further down the line after isolation and characterization.

The broader the effective phage diversity readily available at hand – the more opportunities there are for follow–up research potentially leading to valuable results of broader interest.

CONCLUSIONS

1. There is a great bias regarding the effective phage diversity towards hosts of healthcare and/or economic importance, many bacteria have no known phages able to infect them.

2. *Virgibacillus* siphophage Mimir87 isolated from honeybees is the first completely sequenced (MK560763) tailed bacteriophage that infects a member of the bacterial genus *Virgibacillus* and can be considered for the creation of a novel phage genus as its sole representative.

3. *Erwinia* siphophage Midgardsormr38 isolated from mixed insects represents the first completely sequenced (MN602881) temperate *E. persicina*-infecting phage, as well as is a sole representative of *Midgardsormrvirus* phage genus.

4. Despite the genus *Lactococcus* being one of the most popular host genera for cultured phages completely sequenced (nearly 350 phage genome entries available at the time of writing), *Lactococcus* siphophage Nocturne116 (MW791312) isolated from moth can be considered for the creation of a novel phage genus as its sole representative.

5. Temperate siphophages Perkons (MT325768) and Spoks (MT410774) isolated from Antarctic ice-free soil samples are the first phages shown to be capable of infecting a bacterium from the genus *Psychrobacillus* that had their complete annotated genomes publicly available, both can be considered for the creation of novel phage genera as their sole representatives.

6. *Morganella* myophage Mecenats66 (OL614104) isolated from honeybees utilizes an evolutionarily distinct subtype of headful packaging and seems to be the ninth cultured virus capable of infecting bacteria from the genus *Morganella*, Mecenats66 can be considered for the creation of a novel phage genus as its sole representative.

7. Three distinct *Pantoea* phages - podophages Nifs112 (OK570184) and Nufs112 (OK570185), as well as myophage Nafs113 (OK570184), were co-isolated from a single grasshopper specimen but only Nufs112 can be considered for the creation of a novel phage genus as its sole representative.

THESES

I. Metaviromics approaches have been instrumental in acknowledging the true extent of viral diversity in the virosphere, but the current extent of usability of sequence data alone makes culture-dependent methods indispensable for advancing our current knowledge about bacteriophages.

II. It is in our best interest to gain a better understanding of the phage phenomenon to isolate and sequence not only as many phages as possible but also to make sure to do so from a variety of ecological niches and with particular attention to less commonplace hosts.

III. Utilization of bacterial strains recovered from the undersampled environments as phage indicator cultures allows for a rather consistent isolation of bacteriophages highly divergent from the so far available effective phage diversity.

APPROBATION OF RESEARCH

Publications

I. **Zrelovs, N.**; Cernooka, E.; Dislers, A.; Kazaks, A. Isolation and Characterization of the Novel *Virgibacillus*-Infecting Bacteriophage Mimir87. *Archives of Virology* **2020**, 165, 737–741, doi:10.1007/s00705-019-04516-2

II. **Zrelovs, N.**; Dislers, A.; Kazaks, A. Novel *Erwinia Persicina* Infecting Phage Midgardsormr38 Within the Context of Temperate *Erwinia* Phages. *Frontiers in Microbiology* **2020**, 11, doi:10.3389/fmicb.2020.01245

III. **Zrelovs, N.**; Dislers, A.; Kazaks, A. Motley Crew: Overview of the Currently Available Phage Diversity. *Frontiers in Microbiology* **2020**, 11, doi:10.3389/fmicb.2020.579452

IV. **Zrelovs, N.**; Dislers, A.; Kazaks, A. Genome Characterization of Nocturne116, Novel *Lactococcus lactis*-infecting Phage Isolated from Moth. *Microorganisms* **2021**, doi:10.3390/microorganisms9071540.

V. **Zrelovs, N.**; Lamsters, K.; Karuss, J.; Krievans, M.; Dislers, A.; Kazaks, A. PVJ1 Is Not the First Tailed Temperate Phage Infecting Bacteria from Genus *Psychrobacillus*. Comment on Liu et al. Isolation and Characterization of the First Temperate Virus Infecting *Psychrobacillus* from Marine Sediments. Viruses 2022, 14, 108. *Viruses* **2022**, 14, doi:10.3390/v14030495.

VI. **Zrelovs, N.**; Jansons, J.; Dislers, A.; Kazaks, A. *Morganella* Phage Mecenats66 Utilizes an Evolutionarily Distinct Subtype of Headful Genome Packaging with a Preferred Packaging Initiation Site. *Microorganisms* **2022**, 10, 1799, doi:10.3390/microorganisms10091799

VII. **Zrelovs, N.**; Jansons, J.; Kazaka, T.; Kazaks, A.; Dislers, A. Three Phages One Host: Isolation and Characterization of *Pantoea agglomerans* Phages from a Grasshopper Specimen. *Int. J. Mol. Sci.* **2023**, 24, 1820. https://doi.org/10.3390/ijms24031820

International conferences

I. **Zrelovs, N.**, Cernooka, E., Akopjana, I., Bogans, J., Priedeslaipa, A., Rumnieks, J., Dislers, A., Kazaka, T., Kazaks, A. (Poster presentation). Studies of conserved proteins of unknown function from the proteome of three Antarctic soil phages. *Viruses of Microbes 2023*, July 3-7, **2023**, Tbilisi (Georgia).

II. **Zrelovs, N.**, Dislers, A., Kazaks, A. (Poster presentation). Genomic GC% Content Relationships Between Bacteriophages and Their Hosts. *International Virus Bioinformatics Meeting 2023*, May 24-26, **2023**, Valencia (Spain).

III. **Zrelovs, N.**, Dislers, A., Kazaks, A. (Poster presentation). Genomic GC% Content Relationships Between Bacteriophages and Their Hosts. *VIII Baltic Genetics Congress*, March 22-24, **2023**, Kaunas (Lithuania).

IV. **Zrelovs, N.**, Lamsters, K., Karuss, J., Krievans, M., Jansons, J., Kazaka, T., Dislers, A., Kazaks, A. (Poster presentation). Isolation and genomic characterization of novel bacteriophages from Antarctic ice-free soil samples. *Viruses of Microbes 2022*, July 18-22, **2022**, Guimarães (Portugal).

V. **Zrelovs, N.**, Dislers, A., Kazaks, A. (Poster presentation). Genomic characterization of three Pantoea agglomerans infecting phages isolated from a single insect specimen. *24th Biennial Evergreen Phage Meeting*, August 2-5, **2021**, Olympia, WA, USA (participated online).

VI. **Zrelovs, N.**, Cernooka, E., Dislers, A., Kazaks, A. (Oral presentation). Genomic and proteomic studies of novel bacteriophages from Latvian Biomedical Research and Study Centre's phage collection. *9th International conference 'Bioresources and Viruses'*, Sep 9- 11, **2019**, Kyiv (Ukraine).

VII. **Zrelovs, N.**, Cernooka, E., Lamsters, K., Karuss, J., Krievans, M., Dislers, A., Kazaks, A. (Oral presentation). Isolation and characterization of novel bacteriophages from Antarctic soil samples. *9th International conference 'Bioresources and Viruses'*, September 9-11, **2019**, Kyiv (Ukraine).

VIII. **Zrelovs, N.**, Dislers, A., Kazaks, A. (Poster presentation). Genomic characterization of novel Bacillaceae bacteriophage Mimir87. *The 7th Baltic Genetics Congress*, October 24–27, **2018**, Riga (Latvia).

IX. **Zrelovs, N.**, Dislers, A., Kazaks, A. (Poster presentation). Genomic characterization of two novel Enterobacteriaceae bacteriophages. *Viruses of Microbes 2018,* July 9-13, **2018**, Wrocław (Poland).

ACKNOWLEDGEMENTS

My most sincere gratitude goes to my long-term co-supervisors **Dr. Andris Dišlers** and **Dr. Andris Kazāks**. First of all, for sharing all their relevant expertise and "know-how" regarding micro- and molecular biology techniques used in phage research. Secondly, for never saying "no" without a proper reason to anything during what is now around seven years since I joined the BMC during my undergrad studies. Thirdly, for being supportive of most of my ideas, thus, allowing for a great extent of academic freedom in both performing as well as proposing research activities.

I am also thankful to **Prof. Jānis Kloviņš**, **Dr. Inna Iņaškina**, **Dr. Dāvids Fridmanis**, and **Dr. Ina Baļķe** for pushing me to further develop my skillset and/or finding ways to make it feasible for me to continue pursuing career in science in a country that has one of the lowest Research and Development expenditures as % of the GDP in the EU during the periods when we were unlucky with the funding situation within our own research group.

Another thank you goes to my students, especially **BSc. Karīna Švānberga**, for her helpful hands and collaboration in performing the research activities on a scale that would not be possible without her hard work. I dedicate this thesis to my students - former, current, and future.

I am much obliged to **Dr. Juris Jansons** for bearing with all the insane amounts of varying quality samples we tend to provide for transmission electron microscopy, as well as **MSc. Gunta Reseviča** for her aid in ultracentrifugation.

Another cheer goes to all the other LV BMC personnel I've had the pleasure to learn from, discuss, or collaborate with, especially to the **members of the Prof. Kaspars Tārs** lab hosting me, as well as the **Latvian Biomedical Research and Study Centre's "Genome Centre"** and **Bioinformatics core facility** members.

Special thanks go to **my wife Anna** and **my dog Haru** for mostly being successful in keeping me away from work-related activities during off-work hours despite there nearly always being results to continue building upon at any given time.

Additionally, I would like to appreciate **my family** who consistently remained supportive regardless of whether they actually understood what I was up to research-wise.

I would not be able to focus on my research full-time without receiving financial support in the form of scholarships from the University of Latvia Foundation, thus, my closing thank you goes to the **University of Latvia Foundation** and the **patrons** that established the scholarships I was lucky enough to receive throughout my master's and doctoral degree studies. These scholarships were instrumental in allowing me to pursue a career in science and I doubt that I would make it this far otherwise, for this I am very grateful.

REFERENCES

Abdelrahman, F., Easwaran, M., Daramola, O.I., Ragab, S., Lynch, S., Oduselu, T.J., Khan, F.M., Ayobami, A., Adnan, F., Torrents, E., Sanmukh, S., El-Shibiny, A. Phage-Encoded Endolysins. *Antibiotics*, **2021**, 10(2), 124. doi:10.3390/antibiotics10020124

Abedon, S.T. Phage evolution and ecology. *Adv. Appl. Microbiol.*, **2009**, 67, 1–45. doi:10.1016/S0065-2164(08)01001-0

Abedon, S.T., Danis-Wlodarczyk, K.M., Wozniak, D.J. Phage Cocktail Development for Bacteriophage Therapy: Toward Improving Spectrum of Activity Breadth and Depth. *Pharmaceuticals*, **2021**, 14(10), 1019. doi:10.3390/ph14101019

Abedon, S.T., Kuhl, S.J., Blasdel, B.G., Kutter, E.M. Phage treatment of human infections. *Bacteriophage*, **2011**, 1(2), 66–85. doi:10.4161/bact.1.2.15845

Abedon, S.T., Murray, K.L. Archaeal viruses, not archaeal phages: An archaeological dig. *Archaea*, **2013**, 2013. doi:10.1155/2013/251245

Ackermann, H.-W. The lambda - P22 problem. *Bacteriophage*, **2015**, 5(1), e1017084. doi:10.1080/21597081.2015.1017084

Akmal, M., Nishiki, I., Zrelovs, N., Yoshida, T. Complete genome sequence of a novel lytic bacteriophage, PLG-II, specific for Lactococcus garvieae serotype II strains that are pathogenic to fish. *Arch. Virol.*, **2022**, 167(11), 2331–2335. doi:10.1007/s00705-022-05568-7

Aksyuk, A.A., Rossmann, M.G. Bacteriophage assembly. *Viruses*, **2011**, 3(3), 172–203. doi:10.3390/v3030172

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.*, **1990**, 215(3), 403–410. doi:10.1016/S0022-2836(05)80360-2

Andrews, S. *FastQC - A quality control tool for high throughput sequence data.* [online]. *Babraham Bioinforma.*, **2010**. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.*, **2012**, 19(5), 455–477. doi:10.1089/cmb.2012.0021

Bedane, T.D., Agga, G.E., Gutema, F.D. Hygienic assessment of fish handling practices along production and supply chain and its public health implications in Central Oromia, Ethiopia. *Sci. Rep.*, **2022**, 12(1), 13910. doi:10.1038/s41598-022-17671-5

Bellas, C.M., Schroeder, D.C., Edwards, A., Barker, G., Anesio, A.M. Flexible genes establish widespread bacteriophage pan-genomes in cryoconite hole ecosystems. *Nat. Commun.*, **2020**, 11(1), 4403. doi:10.1038/s41467-020-18236-8

Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., Pons, M.-N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., Martinez, J.L. Tackling antibiotic resistance: the environmental framework. *Nat. Rev. Microbiol.*, **2015**, 13(5), 310–317. doi:10.1038/nrmicro3439

Bertozzi Silva, J., Storms, Z., Sauvageau, D. Host receptors for bacteriophage adsorption. *FEMS Microbiol. Lett.*, **2016**, 363(4), fnw002. doi:10.1093/femsle/fnw002

Besemer, J., Borodovsky, M. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.*, **2005**, 33, 451–454. doi:10.1093/nar/gki487

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister, J.R., Kropinski, A.M., Krupovic, M., Lavigne, R., Turner, D., Sullivan, M.B. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.*, **2019**, 37(6), 632–639. doi:10.1038/s41587-019-0100-8

Botstein, D. a Theory of Modular Evolution for Bacteriophages. *Ann. N. Y. Acad. Sci.*, **1980**, 354(1), 484–491. doi:10.1111/j.1749-6632.1980.tb27987.x

Botstein, D., Herskowitz, I. Properties of hybrids between Salmonella phage P22 and coliphage λ. *Nature*, **1974**, 251(5476), 584–589. doi:10.1038/251584a0

Bragg, L., Tyson, G.W. Metagenomics Using Next-Generation Sequencing*, in: Paulsen, I., Holmes, A. (Eds.), Environmental Microbiology. Methods in Molecular Biology (Methods and Protocols)*, **2014**. Humana Press, Totowa, NJ, pages 183–201, ISBN: 978-1-62703-711-2. doi:10.1007/978-1-62703-712-9_15

Briers, Y., Lavigne, R. Breaking barriers: Expansion of the use of endolysins as novel antibacterials against Gram-negative bacteria. *Future Microbiol.*, **2015**, 10(3), 377–390. doi:10.2217/FMB.15.8

Briers, Y., Walmagh, M., Puyenbroeck, V.V., Cornelissen, A., Cenens, W., Aertsen, A., Oliveira, H., Azeredo, J., Verween, G., Pirnay, J.-P., Miller, S., Volckaert, G., Lavigne, R. Engineered Endolysin-Based "Artilysins" To Combat Multidrug-Resistant Gram-Negative Pathogens. *mBio*, **2014**, 5(4), e01379-14. doi:10.1128/mBio.01379-14

Brüssow, H., Canchaya, C., Hardt, W.-D. Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiol. Mol. Biol. Rev.*, **2004**, 68(3), 560–602. doi:10.1128/MMBR.68.3.560-602.2004

Brüssow, H., Hendrix, R.W. Phage Genomics: Small is beautiful. *Cell*, **2002**, 108(1), 13–16. doi:10.1016/S0092-8674(01)00637-7

Bushnell, B. BBMap : A Fast, Accurate, Splice-Aware Aligner. *Conf. 9th Annu. Genomics Energy Environ. Meet. Walnut Creek CA March 17-20 2014*, **2014**.

Buttimer, C., Mcauliffe, O., Ross, R.P., Hill, C., Mahony, J.O., Coffey, A., Abedon, S.T., Chan, B.K. Bacteriophages and Bacterial Plant Diseases, **2017**, 8(January), 1–15. doi:10.3389/fmicb.2017.00034

Camargo, A.P., Nayfach, S., Chen, I.-M.A., Palaniappan, K., Ratner, A., Chu, K., Ritter, S.J., Reddy, T.B.K., Mukherjee, S., Schulz, F., Call, L., Neches, R.Y., Woyke, T., Ivanova, N.N., Eloe-Fadrosh, E.A., Kyrpides, N.C., Roux, S. IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. *Nucleic Acids Res.*, **2022**, 51(D1), D733–D743. doi:10.1093/nar/gkac1037

Casjens, S. Prophages and bacterial genomics: What have we learned so far?. *Mol. Microbiol.*, **2003**, 49(2), 277–300. doi:10.1046/j.1365-2958.2003.03580.x

Casjens, S.R. The DNA-packaging nanomotor of tailed bacteriophages. *Nat. Rev. Microbiol.*, **2011**, 9(9), 647–657. doi:10.1038/nrmicro2632

Casjens, S.R., Gilcrease, E.B. Determining DNA Packaging Strategy by Analysis of the Termini of the Chromosomes in Tailed-Bacteriophage Virions*, in: Bacteriophages: Methods and Protocols, Volume 2: Molecular and Applied Aspects*, **2009**, pages 91–111. doi:10.1007/978-1-60327-565-1_7

Casjens, S.R., Gilcrease, E.B., Winn-Stapley, D.A., Schicklmaier, P., Schmieger, H., Pedulla, M.L., Ford, M.E., Houtz, J.M., Hatfull, G.F., Hendrix, R.W. The generalized transducing Salmonella bacteriophage ES18: Complete genome sequence and DNA packaging strategy. *J. Bacteriol.*, **2005**, 187(3), 1091–1104. doi:10.1128/JB.187.3.1091-1104.2005

Casjens, S.R., Hendrix, R.W. Bacteriophage lambda: Early pioneer and still relevant. *Virology*, **2015**, 479–480, 310–330. doi:10.1016/j.virol.2015.02.010

Casjens, S.R., Molineux, I.J. Short noncontractile tail machines: adsorption and DNA delivery by podoviruses. *Adv. Exp. Med. Biol.*, **2012**, 726, 143–179. doi:10.1007/978-1-4614-0980-9_7

Castresana, J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Mol. Biol. Evol.*, **2000**, 17(4), 540–552. doi:10.1093/oxfordjournals.molbev.a026334

Cazares, A., Mendoza-Hernández, G., Guarneros, G. Core and accessory genome architecture in a group of Pseudomonas aeruginosa Mu-like phages. *BMC Genomics*, **2014**, 15(1), 1146. doi:10.1186/1471-2164-15-1146

Cernooka, E., Rumnieks, J., Tars, K., Kazaks, A. Structural Basis for DNA Recognition of a Single-stranded DNA-binding Protein from Enterobacter Phage Enc34. *Sci. Rep.*, **2017**, 7(1), 1–10. doi:10.1038/s41598-017-15774-y

Cernooka, E., Rumnieks, J., Zrelovs, N., Tars, K., Kazaks, A. Diversity of the lysozyme fold: structure of the catalytic domain from an unusual endolysin encoded by phage Enc34. *Sci. Rep.*, **2022**, 12(1), 1–11. doi:10.1038/s41598-022-08765-1

Ceyssens, P.-J., Minakhin, L., Van den Bossche, A., Yakunina, M., Klimuk, E., Blasdel, B., De Smet, J., Noben, J.-P., Bläsi, U., Severinov, K., Lavigne, R. Development of Giant Bacteriophage ϕKZ Is Independent of the Host Transcription Apparatus. *J. Virol.*, **2014**, 88(18), 10501–10510. doi:10.1128/JVI.01347-14

Chanishvili, N. Chapter 1 - Phage Therapy—History from Twort and d'Herelle Through Soviet Experience to Current Approaches*, in: Łobocka, M., Szybalski, W. (Eds.), Advances in Virus Research, Bacteriophages, Part B*, **2012**. Academic Press, pages 3–40. doi:10.1016/B978-0-12-394438-2.00001-3

Chevreux, Bastien, Wetter, Thomas, Suhai, Sandor. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information.. *Comput. Sci. Biol. Proc. Ger. Conf. Bioinforma. GCB*, **1999**, 99(1995), 45–56.

Cho, N.Y., Choi, M., Rothman-Denes, L.B. The bacteriophage N4-coded single-stranded DNA-binding protein (N4SSB) is the transcriptional activator of Escherichia coli RNA polymerase at N4 late promoters. *J. Mol. Biol.*, **1995**, 246(4), 461–471. doi:10.1006/jmbi.1994.0098

Clokie, M.R.J., Blasdel, B.G., Demars, B.O.L., Sicheritz-Pontén, T. Rethinking Phage Ecology by Rooting it Within an Established Plant Framework. *Phage*, **2020**, 1(3), 121–136. doi:10.1089/phage.2020.0015

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.*, **2014**. doi:10.1093/nar/gkt1244

Comeau, A.M., Bertrand, C., Letarov, A., Tétart, F., Krisch, H.M. Modular architecture of the T4 phage superfamily: A conserved core genome and a plastic periphery. *Virology*, **2007**, 362(2), 384–396. doi:10.1016/j.virol.2006.12.031

Cook, R., Brown, N., Redgwell, T., Rihtman, B., Barnes, M., Clokie, M., Stekel, D.J., Hobman, J., Jones, M.A., Millard, A. INfrastructure for a PHAge REference Database: Identification of Large-Scale Biases in the Current Collection of Cultured Phage Genomes. *PHAGE*, **2021**, 2(4), 214–223. doi:10.1089/phage.2021.0007

Culot, A., Grosset, N., Gautier, M. Overcoming the challenges of phage therapy for industrial aquaculture: A review. *Aquaculture*, **2019**, 513, 734423. doi:10.1016/j.aquaculture.2019.734423

Cumby, N., Davidson, A.R., Maxwell, K.L. The moron comes of age. *Bacteriophage*, **2012**, 2(4), e23146. doi:10.4161/bact.23146

Davidson, A.R., Cardarelli, L., Pell, L.G., Radford, D.R., Maxwell, K.L. Long Noncontractile Tail Machines of Bacteriophages*, in: Rossmann, M.G., Rao, V.B. (Eds.), Viral Molecular Machines,* *Advances in Experimental Medicine and Biology*, **2012**. Springer US, Boston, MA, pages 115–142, ISBN: 978-1-4614-0979-3. doi:10.1007/978-1-4614-0980-9_6

De Oliveira, D.M.P., Forde, B.M., Kidd, T.J., Harris, P.N.A., Schembri, M.A., Beatson, S.A., Paterson, D.L., Walker, M.J. Antimicrobial Resistance in ESKAPE Pathogens. *Clin. Microbiol. Rev.*, **2020**, 33(3), e00181-19. doi:10.1128/CMR.00181-19

Dedrick, R.M., Guerrero Bustamante, C.A., Garlena, R.A., Pinches, R.S., Cornely, K., Hatfull, G.F. Mycobacteriophage ZoeJ: A broad host-range close relative of mycobacteriophage TM4. *Tuberculosis*, **2019a**, 115(December 2018), 14–23. doi:10.1016/j.tube.2019.01.002

Dedrick, R.M., Guerrero-Bustamante, C.A., Garlena, R.A., Russell, D.A., Ford, K., Harris, K., Gilmour, K.C., Soothill, J., Jacobs-Sera, D., Schooley, R.T., Hatfull, G.F., Spencer, H. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. *Nat. Med.*, **2019b**, 25(5), 730–733. doi:10.1038/s41591-019-0437-z

Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*, **2007**. doi:10.1093/bioinformatics/btm009

Desiree, K., Mosimann, S., Ebner, P. Efficacy of phage therapy in pigs: systematic review and meta-analysis. *J. Anim. Sci.*, **2021**, 99(7), skab157. doi:10.1093/jas/skab157

D’Herelle, F. Sur le rôle du microbe bactériophage dans la typhose aviaire. *Comptes Rendus Académie Sci.*, **1919**, 169, 932–934.

D’Herelle, F. Sur un microbe invisible antagoniste des bacilles dysentériques. *Comptes Rendus Académie Sci.*, **1917**, 165, 373–375.

Djebara, S., Maussen, C., De Vos, D., Merabishvili, M., Damanet, B., Pang, K.W., De Leenheer, P., Strachinaru, I., Soentjens, P., Pirnay, J.-P. Processing Phage Therapy Requests in a Brussels Military Hospital: Lessons Identified. *Viruses*, **2019**, 11(3), 265. doi:10.3390/v11030265

Du, J., Meile, S., Baggenstos, J., Jäggi, T., Piffaretti, P., Hunold, L., Matter, C.I., Leitner, L., Kessler, T.M., Loessner, M.J., Kilcher, S., Dunne, M. Enhancing bacteriophage therapeutics through in situ production and release of heterologous antimicrobial effectors. *Nat. Commun.*, **2023**, 14(1), 1–10. doi:10.1038/s41467-023-39612-0

Dunn, J.J., Studier, F.W., Gottesman, M. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.*, **1983**, 166(4), 477–535. doi:10.1016/S0022-2836(83)80282-4

Edwards, P. Aquaculture environment interactions: Past, present and likely future trends. *Aquaculture, Research for the Next 40 Years of Sustainable Global Aquaculture*, **2015**, 447, 2–14. doi:10.1016/j.aquaculture.2015.02.001

FAO. The future of food and agriculture – Alternative pathways to 2050, **2018**. Food and Agriculture Organization of the United Nations, Rome, pages 228.

Felsenstein, J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP.. *Evol. Int. J. Org. Evol.*, **1985**, 39(4), 783–791. doi:10.1111/j.1558-5646.1985.tb00420.x

Fernández, L., Gutiérrez, D., Rodríguez, A., García, P. Application of Bacteriophages in the Agro-Food Sector: A Long Way Toward Approval. *Front. Cell. Infect. Microbiol.*, **2018**, 8(296). doi:10.3389/fcimb.2018.00296

Ferriol-González, C., Domingo-Calap, P. Phage Therapy in Livestock and Companion Animals. *Antibiotics*, **2021**, 10(5), 559. doi:10.3390/antibiotics10050559

Fruciano, D.E., Bourne, S. Phage as an antimicrobial agent: d’Herelle’s heretical theories and their role in the decline of phage prophylaxis in the West. *Can. J. Infect. Dis. Med. Microbiol.*, **2007**, 18(1), 19–26. doi:10.1155/2007/976850

Galvāne, B. No augsnes jaunizdalīto Viridibacillus un Bacillus bakteriofāgu raksturošana [in Latvian; Characterization of Viridibacillus and Bacillus bacteriophages isolated from soil], **2023**. Univeristy of Latvia, 79 pages.

Garneau, J.R., Depardieu, F., Fortier, L.C., Bikard, D., Monot, M. PhageTerm: A tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Sci. Rep.*, **2017**, 7(1), 1–10. doi:10.1038/s41598-017-07910-5

Gencay, Y.E., Jasinskytė, D., Robert, C., Semsey, S., Martínez, V., Petersen, A.Ø., Brunner, K., de Santiago Torio, A., Salazar, A., Turcu, I.C., Eriksen, M.K., Koval, L., Takos, A., Pascal, R., Schou, T.S., Bayer, L., Bryde, T., Johansen, K.C., Bak, E.G., Smrekar, F., Doyle, T.B., Satlin, M.J., Gram, A., Carvalho, J., Jessen, L., Hallström, B., Hink, J., Damholt, B., Troy, A., Grove, M., Clube, J., Grøndahl, C., Haaber, J.K., van der Helm, E., Zdravkovic, M., Sommer, M.O.A. Engineered phage with antibacterial CRISPR–Cas selectively reduce E. coli burden in mice. *Nat. Biotechnol.*, **2023**, 1–10. doi:10.1038/s41587-023-01759-y

Gerstmans, H., Criel, B., Briers, Y. Synthetic biology of modular endolysins. *Biotechnol. Adv., Prospects in Biotechnology*, **2018**, 36(3), 624–640. doi:10.1016/j.biotechadv.2017.12.009

Gigante, A., Atterbury, R.J. Veterinary use of bacteriophage therapy in intensively-reared livestock. *Virol. J.*, **2019**, 16(1), 155. doi:10.1186/s12985-019-1260-3

Gilchrist, C.L.M., Chooi, Y.H. Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics*, **2021**, 37(16), 2473–2475. doi:10.1093/bioinformatics/btab007

Gon Choudhury, T., Tharabenahalli Nagaraju, V., Gita, S., Paria, A., Parhi, J. Advances in Bacteriophage Research for Bacterial Disease Control in Aquaculture. *Rev. Fish. Sci. Aquac.*, **2017**, 25(2), 113–125. doi:10.1080/23308249.2016.1241977

Gordillo Altamirano, F.L., Barr, J.J. Phage Therapy in the Postantibiotic Era. *Clin. Microbiol. Rev.*, **2019**, 32(2), 10.1128/cmr.00066-18. doi:10.1128/cmr.00066-18

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, **2013**, 29(8), 1072–1075. doi:10.1093/bioinformatics/btt086

Hanauer, D.I., Graham, M.J., SEA-PHAGES, Betancur, L., Bobrownicki, A., Cresawn, S.G., Garlena, R.A., Jacobs-Sera, D., Kaufmann, N., Pope, W.H., Russell, D.A., Jacobs, W.R., Sivanathan, V., Asai, D.J., Hatfull, G.F. An inclusive Research Education Community (iREC): Impact of the SEA-PHAGES program on research outcomes and student learning. *Proc. Natl. Acad. Sci.*, **2017**, 114(51), 13531–13536. doi:10.1073/pnas.1718188115

Handelsman, J. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev.*, **2004**, 68(4), 669–685. doi:10.1128/MBR.68.4.669–685.2004

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.*, **1998**, 5(10), R245–R249. doi:10.1016/S1074-5521(98)90108-9

Hatfull, G.F. Mycobacteriophages. *Microbiol. Spectr.*, **2018**, 6(5), 10.1128/microbiolspec.gpp3-0026–2018. doi:10.1128/microbiolspec.gpp3-0026-2018

Hatfull, G.F. Dark Matter of the Biosphere: the Amazing World of Bacteriophage Diversity. *J. Virol.*, **2015**, 89(16), 8107–8110. doi:10.1128/jvi.01340-15

Hatfull, G.F., Hendrix, R.W. Bacteriophages and their genomes. *Curr. Opin. Virol.*, **2011**, 1(4), 298–303. doi:10.1016/j.coviro.2011.06.009

Hemstreet, C., Mallmann, W.L. Isolation of an inhibitory substance from plants. *J. Agric. Res.*, **1924**, XXVIII(6), 599–602.

Hendrix, R.W., Smith, M.C.M., Burns, R.N., Ford, M.E., Hatfull, G.F. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl. Acad. Sci. U. S. A.*, **1999**, 96(5), 2192–2197. doi:10.1073/pnas.96.5.2192

Holtappels, D., Fortuna, K., Lavigne, R., Wagemans, J. The future of phage biocontrol in integrated plant protection for sustainable crop production. *Curr. Opin. Biotechnol., Systems Biology • Nanobiotechnology*, **2021**, 68, 60–71. doi:10.1016/j.copbio.2020.08.016

Hoseinifar, S.H., Sun, Y.-Z., Zhou, Z., Van Doan, H., Davies, S.J., Harikrishnan, R. Boosting Immune Function and Disease Bio-Control Through Environment-Friendly and Sustainable Approaches in Finfish Aquaculture: Herbal Therapy Scenarios. *Rev. Fish. Sci. Aquac.*, **2020**, 28(3), 303–321. doi:10.1080/23308249.2020.1731420

Hoyles, L., McCartney, A.L., Neve, H., Gibson, G.R., Sanderson, J.D., Heller, K.J., van Sinderen, D. Characterization of virus-like particles associated with the human faecal and caecal microbiota. *Res. Microbiol.*, **2014**, 165(10), 803–812. doi:10.1016/j.resmic.2014.10.006

Hu, B., Margolin, W., Molineux, I.J., Liu, J. Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl. Acad. Sci. U. S. A.*, **2015**, 112(35), E4919–E4928. doi:10.1073/pnas.1501064112

ICTV. *The International Code of Virus Classification and Nomenclature (ICVCN)* [online], **2021**. Available from: https://ictv.global/about/code [accessed 31 August 2023].

Jia, B., St-Hilaire, S., Singh, K., Gardner, I.A. Biosecurity knowledge, attitudes and practices of farmers culturing yellow catfish (Pelteobagrus fulvidraco) in Guangdong and Zhejiang provinces, China. *Aquaculture*, **2017**, 471, 146–156. doi:10.1016/j.aquaculture.2017.01.016

Jones, J.B., Vallad, G.E., Iriarte, F.B., Obradović, A., Wernsing, M.H., Jackson, L.E., Balogh, B., Hong, J.C., Momol, M.T. Considerations for using bacteriophages for plant disease control. *Bacteriophage*, **2012**, 2(4), 208–214. doi:10.4161/bact.23857

Jonikāne, A. Jaunizdalīto Cedecea davisae inficējošo bakteriofāgu izolēšana un raksturošana [in Latvian; Isolation and characterization of novel Cedecea davisae-infecting bacteriophages], **2022**. Univeristy of Latvia, 81 pages.

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. *J. Mol. Biol.*, **2000**, 299(1), 27–51. doi:10.1006/jmbi.2000.3729

Käll, L., Krogh, A., Sonnhammer, E.L.L. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.*, **2007**, 35(Web Server issue), W429–W432. doi:10.1093/nar/gkm256

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods*, **2017**. doi:10.1038/nmeth.4285

Katoh, K., Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.*, **2013**. doi:10.1093/molbev/mst010

Kazaks, A., Dislers, A., Lipowsky, G., Nikolajeva, V., Tars, K. Complete Genome Sequence of the Enterobacter cancerogenus Bacteriophage Enc34. *J. Virol.*, **2012**, 86(20), 11403–11404. doi:10.1128/JVI.01954-12

Kazi, M., Annapure, U.S. Bacteriophage biocontrol of foodborne pathogens. *J. Food Sci. Technol.*, **2016**, 53(3), 1355–1362. doi:10.1007/s13197-015-1996-8

Kongari, R., Rajaure, M., Cahill, J., Rasche, E., Mijalis, E., Berry, J., Young, R. Phage spanins: diversity, topological dynamics and gene convergence. *BMC Bioinformatics*, **2018**, 19(1), 326. doi:10.1186/s12859-018-2342-8

Korniienko, N., Kharina, A., Zrelovs, N., Jindˇ, B. Isolation and Characterization of Two Lytic Phages Efficient Against Phytopathogenic Bacteria From Pseudomonas and Xanthomonas Genera. *Frontier*, **2022**, 13(April). doi:10.3389/fmicb.2022.853593

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *J. Mol. Biol.*, **2001**, 305(3), 567–580. doi:10.1006/jmbi.2000.4315

Kropinski, A.M., Mazzocco, A., Waddell, T.E., Johnson, R.P. Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay. Bacteriophages*, in: Bacteriophages. Methods in Molecular Biology.*, **2009**. Humana Press, pages 69–76, ISBN: 978-1-58829-682-5. doi:10.1007/978-1-60327-164-6

Krupovic, M., Dolja, V.V., Koonin, E.V. The LUCA and its complex virome. *Nat. Rev. Microbiol.*, **2020**, 18(11), 661–670. doi:10.1038/s41579-020-0408-x

Kumar, S., Stecher, G., Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.*, **2016**, 33(7), 1870–1874. doi:10.1093/molbev/msw054

Kunhikannan, S., Thomas, C.J., Franks, A.E., Mahadevaiah, S., Kumar, S., Petrovski, S. Environmental hotspots for antibiotic resistance genes. *MicrobiologyOpen*, **2021**, 10(3), e1197. doi:10.1002/mbo3.1197

Kutateladze, M., Adamia, R. Phage therapy experience at the Eliava Institute. *Médecine Mal. Infect., Numéro spécial CEMI 2007 et 2008*, **2008**, 38(8), 426–430. doi:10.1016/j.medmal.2008.06.023

Labrie, S.J., Samson, J.E., Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, **2010**, 8(5), 317–327. doi:10.1038/nrmicro2315

Laslett, D., Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.*, **2004**, 32(1), 11–16. doi:10.1093/nar/gkh152

Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y., Drulis-Kawa, Z. Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl. Microbiol. Biotechnol.*, **2017**, 101(8), 3103–3119. doi:10.1007/s00253-017-8224-6

Lee, J., Hwang, Joo Hee., Jo, D.S., Lee, H.S., Hwang, Jeong Hwan. Kluyvera ascorbata as a pathogen in adults and children: Clinical features and antibiotic susceptibilities in a single center study. *Jpn. J. Infect. Dis.*, **2019**, 72(3), 142–148. doi:10.7883/yoken.JJID.2018.375

Lefkowitz, E.J., Dempsey, D.M., Hendrickson, R.C., Orton, R.J., Siddell, S.G., Smith, D.B. Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.*, **2018**, 46(D1), D708–D717. doi:10.1093/nar/gkx932

Leiman, P.G., Shneider, M.M. Contractile Tail Machines of Bacteriophages*, in: Rossmann, M.G., Rao, V.B. (Eds.), Viral Molecular Machines, Advances in Experimental Medicine and Biology*, **2012**. Springer US, Boston, MA, pages 93–114, ISBN: 978-1-4614-0979-3. doi:10.1007/978-1-4614-0980-9_5

Leinonen, R., Akhtar, R., Birney, E., Bower, L., Cerdeno-Tárraga, A., Cheng, Y., Cleland, I., Faruque, N., Goodgame, N., Gibson, R., Hoad, G., Jang, M., Pakseresht, N., Plaister, S., Radhakrishnan, R., Reddy, K., Sobhany, S., Ten Hoopen, P., Vaughan, R., Zalunin, V., Cochrane, G. The European Nucleotide Archive. *Nucleic Acids Res.*, **2011**, 39(Database issue), D28–D31. doi:10.1093/nar/gkq967

Letarov, A.V. History of Early Bacteriophage Research and Emergence of Key Concepts in Virology. *Biochem. Biokhimiia*, **2020**, 85(9), 1093–1010. doi:10.1134/S0006297920090096

Letarov, A.V. Современные концепции биологии бактериофагов [In Russian; Modern Concepts in Bacteriophage Biology], **2019**. ДеЛи, Moscow, 384 pages, ISBN: 978-5-604-27124-7.

Lima-Mendez, G., Toussaint, A., Leplae, R. Analysis of the phage sequence space: The benefit of structured information. *Virology*, **2007**, 365(2), 241–249. doi:10.1016/j.virol.2007.03.047

Lima-Mendez, G., Van Helden, J., Toussaint, A., Leplae, R. Reticulate Representation of Evolutionary and Functional Relationships between Phage Genomes. *Mol. Biol. Evol.*, **2008**, 25(4), 762–777. doi:10.1093/molbev/msn023

Liu, H., Zhu, J., Hu, Q., Rao, X. Morganella morganii, a non-negligent opportunistic pathogen. *Int. J. Infect. Dis.*, **2016**, 50, 10–17. doi:10.1016/j.ijid.2016.07.006

Liu, J., Glazko, G., Mushegian, A. Protein repertoire of double-stranded DNA bacteriophages. *Virus Res., Comparative Genomics and Evolution of Complex Viruses*, **2006**, 117(1), 68–80. doi:10.1016/j.virusres.2006.01.015

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(21), 7032–7068. doi:10.1128/JB.186.21.7032-7068.2004

Loc-Carrillo, C., Abedon, S.T. Pros and cons of phage therapy. *Bacteriophage*, **2011**, 1(2), 111–114. doi:10.4161/bact.1.2.14590

Loponte, R., Pagnini, U., Iovane, G., Pisanelli, G. Phage Therapy in Veterinary Medicine. *Antibiotics*, **2021**, 10(4), 421. doi:10.3390/antibiotics10040421

Losick, R., Pero, J. Cascades of Sigma factors. *Cell*, **1981**, 25(3), 582–584. doi:10.1016/0092-8674(81)90164-1

Lowe, T.M., Eddy, S.R. tRNAscan-SE : a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, **1997**, 25(5), 955–964. doi:https://doi.org/10.1093/nar/25.5.955

Luke, K., Radek, A., Liu, X., Campbell, J., Uzan, M., Haselkorn, R., Kogan, Y. Microarray Analysis of Gene Expression during Bacteriophage T4 Infection. *Virology*, **2002**, 299(2), 182–191. doi:10.1006/viro.2002.1409

Luria, S.E., Delbrück, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*, **1943**, 28(6), 491–511. doi:10.1093/genetics/28.6.491

Maffei, E., Shaidullina, A., Burkolter, M., Heyer, Y., Estermann, F., Druelle, V., Sauer, P., Willi, L., Michaelis, S., Hilbi, H., Thaler, D.S., Harms, A. Systematic exploration of Escherichia coli phage–host interactions with the BASEL phage collection. *PLOS Biol.*, **2021**, 19(11), e3001424. doi:10.1371/journal.pbio.3001424

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., Toth, I., Salmond, G., Foster, G.D. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.*, **2012**, 13(6), 614–629. doi:10.1111/j.1364-3703.2012.00804.x

Mäntynen, S., Laanto, E., Oksanen, H.M., Poranen, M.M., Díaz-Muñoz, S.L. Black box of phage–bacterium interactions: exploring alternative phage infection strategies. *Open Biol.*, **2021**, 11(9), 210188. doi:10.1098/rsob.210188

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.*, **2011**, 39(SUPPL. 1), 225–229. doi:10.1093/nar/gkq1189

Mashima, J., Kodama, Y., Fujisawa, T., Katayama, T., Okuda, Y., Kaminuma, E., Ogasawara, O., Okubo, K., Nakamura, Y., Takagi, T. DNA Data Bank of Japan. *Nucleic Acids Res.*, **2017**, 45(D1), D25–D31. doi:10.1093/nar/gkw1001

Mavrich, T.N., Hatfull, G.F. Bacteriophage evolution differs by host, lifestyle and genome. *Nat. Microbiol.*, **2017**, 2(9), 1–9. doi:10.1038/nmicrobiol.2017.112

Merrill, B.D., Ward, A.T., Grose, J.H., Hope, S. Software-based analysis of bacteriophage genomes, physical ends, and packaging strategies. *BMC Genomics*, **2016**, 17(1), 1–16. doi:10.1186/s12864-016-3018-2

Mesyanzhinov, V.V., Robben, J., Grymonprez, B., Kostyuchenko, V.A., Bourkaltseva, M.V., Sykilinda, N.N., Krylov, V.N., Volckaert, G. The genome of bacteriophage φKZ of Pseudomonas aeruginosa11Edited by M. Gottesman. *J. Mol. Biol.*, **2002**, 317(1), 1–19. doi:10.1006/jmbi.2001.5396

Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., Rüger, W. Bacteriophage T4 Genome. *Microbiol. Mol. Biol. Rev.*, **2003**, 67(1), 86–156. doi:10.1128/MMBR.67.1.86-156.2003

Minh, B.Q., Nguyen, M.A.T., Von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.*, **2013**. doi:10.1093/molbev/mst024

Molineux, I.J., Panja, D. Popping the cork: mechanisms of phage genome ejection. *Nat. Rev. Microbiol.*, **2013**, 11(3), 194–204. doi:10.1038/nrmicro2988

Moraru, C., Varsani, A., Kropinski, A.M. VIRIDIC — A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses*, **2020**, 12, 1268. doi:10.3390/v12111268

Mosimann, S., Desiree, K., Ebner, P. Efficacy of phage therapy in poultry: a systematic review and meta-analysis. *Poult. Sci.*, **2021**, 100(12), 101472. doi:10.1016/j.psj.2021.101472

Moura de Sousa, J.A., Pfeifer, E., Touchon, M., Rocha, E.P.C. Causes and Consequences of Bacteriophage Diversification via Genetic Exchanges across Lifestyles and Bacterial Taxa. *Mol. Biol. Evol.*, **2021**, 38(6), 2497–2512. doi:10.1093/molbev/msab044

Moye, Z.D., Woolston, J., Sulakvelidze, A. Bacteriophage Applications for Food Production and Processing. *Viruses*, **2018**, 10(4), 205. doi:10.3390/v10040205

Murray, E., Draper, L.A., Ross, R.P., Hill, C. The Advantages and Challenges of Using Endolysins in a Clinical Setting. *Viruses*, **2021**, 13(4), 680. doi:10.3390/v13040680

Mushegian, A.R. Are There 1031 Virus Particles on Earth, or More, or Fewer?. *J. Bacteriol.*, **2020**, 202(9), e00052-20. doi:10.1128/JB.00052-20

Nelson, D., Loomis, L., Fischetti, V.A. Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc. Natl. Acad. Sci. U. S. A.*, **2001**, 98(7), 4107–4112. doi:10.1073/pnas.061038398

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.*, **2015**. doi:10.1093/molbev/msu300

Nobrega, F.L., Vlot, M., de Jonge, P.A., Dreesens, L.L., Beaumont, H.J.E., Lavigne, R., Dutilh, B.E., Brouns, S.J.J. Targeting mechanisms of tailed bacteriophages. *Nat. Rev. Microbiol.*, **2018**, 16(12), 760–773. doi:10.1038/s41579-018-0070-8

Oechslin, F. Resistance Development to Bacteriophages Occurring during Bacteriophage Therapy. *Viruses*, **2018**, 10(7), 351. doi:10.3390/v10070351

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C.M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A.R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D., Pruitt, K.D. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.*, **2016**. doi:10.1093/nar/gkv1189

"Overview of EU aquaculture (fish farming)." , **2023**. Available from: https://oceans-and-fisheries.ec.europa.eu/ocean/blue-economy/aquaculture/overview-eu-aquaculture-fish-farming_en [accessed 3 April 2023].

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, **2015**, 31(22), 3691–3693. doi:10.1093/bioinformatics/btv421

Pereira, C., Duarte, J., Costa, P., Braz, M., Almeida, A. Bacteriophages in the Control of Aeromonas sp. in Aquaculture Systems: An Integrative View. *Antibiotics*, **2022**, 11(2), 163. doi:10.3390/antibiotics11020163

Petrov, V.M., Ratnayaka, S., Nolan, J.M., Miller, E.S., Karam, J.D. Genomes of the T4-related bacteriophages as windows on microbial genome evolution. *Virol. J.*, **2010**, 7(1), 292. doi:10.1186/1743-422X-7-292

Pietilä, M.K., Demina, T.A., Atanasova, N.S., Oksanen, H.M., Bamford, D.H. Archaeal viruses and bacteriophages: comparisons and contrasts. *Trends Microbiol.*, **2014**, 22(6), 334–344. doi:10.1016/j.tim.2014.02.007

Pirnay, J.-P., Blasdel, B.G., Bretaudeau, L., Buckling, A., Chanishvili, N., Clark, J.R., Corte-Real, S., Debarbieux, L., Dublanchet, A., De Vos, D., Gabard, J., Garcia, M., Goderdzishvili, M., Górski, A., Hardcastle, J., Huys, I., Kutter, E., Lavigne, R., Merabishvili, M., Olchawa, E., Parikka, K.J., Patey, O., Pouilot, F., Resch, G., Rohde, C., Scheres, J., Skurnik, M., Vaneechoutte, M., Van Parys, L., Verbeken, G., Zizi, M., Van den Eede, G. Quality and safety requirements for sustainable phage therapy products. *Pharm. Res.*, **2015**, 32(7), 2173–2179. doi:10.1007/s11095-014-1617-7

Pirnay, J.-P., De Vos, D., Verbeken, G., Merabishvili, M., Chanishvili, N., Vaneechoutte, M., Zizi, M., Laire, G., Lavigne, R., Huys, I., Van Den Mooter, G., Buckling, A., Debarbieux, L., Pouillot, F., Azeredo, J., Kutter, E., Dublanchet, A., Górski, A., Adamia, R. The Phage Therapy Paradigm: Prêt-à-Porter or Sur-mesure?. *Pharm. Res.*, **2011**, 28(4), 934–937. doi:10.1007/s11095-010-0313-5

Pirnay, J.-P., Djebara, S., Steurs, G., Griselain, J., Cochez, C., Soir, S.D., Glonti, T., Spiessens, A., Berghe, E.V., Green, S., Wagemans, J., Lood, C., Schrevens, E., Chanishvili, N., Kutateladze, M., Jode, M. de., Ceyssens, P.-J., Draye, J.-P., Verbeken, G., Vos, D.D., Rose, T., Onsea, J., Nieuwenhuyse, B.V., Providers, B.T., Donors, B., Soentjens, P., Lavigne, R., Merabishvili, M. Retrospective, observational analysis of the first one hundred consecutive cases of personalized bacteriophage therapy of difficult-to-treat infections facilitated by a Belgian consortium, **2023**. doi:10.1101/2023.08.28.23294728

Pirnay, J.-P., Verbeken, G., Ceyssens, P.-J., Huys, I., De Vos, D., Ameloot, C., Fauconnier, A. The Magistral Phage. *Viruses*, **2018**, 10(2), 64. doi:10.3390/v10020064

Połaska, M., Sokołowska, B. Bacteriophages—a new hope or a huge problem in the food industry. *AIMS Microbiol.*, **2019**, 5(4), 324–346. doi:10.3934/microbiol.2019.4.324

Pope, W.H., Jacobs-Sera, D. Annotation of Bacteriophage Genome Sequences Using DNA Master: An Overview*, in: Clokie, M.R.J., Kropinski, A.M., Lavigne, R. (Eds.), Bacteriophages, Methods in Molecular Biology*, **2018**. Springer New York, New York, NY, pages 217–229, ISBN: 978-1-4939-7341-5. doi:10.1007/978-1-4939-7343-9_16

Preena, P.G., Swaminathan, T.R., Kumar, V.J.R., Singh, I.S.B. Antimicrobial resistance in aquaculture: a crisis for concern. *Biologia (Bratisl.)*, **2020**, 75(9), 1497–1517. doi:10.2478/s11756-020-00456-4

Rakhuba, D.V., Kolomiets, E.I., Dey, E.S., Novik, G.I. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. *Pol. J. Microbiol.*, **2010**, 59(3), 145–155.

Rambaut, A. *FigTree v. 1.4.4* [online]. *http://tree.bio.ed.ac.uk/software/figtree/*, **2018**. Available from: http://tree.bio.ed.ac.uk/software/figtree/

Rao, V.B., Feiss, M. The bacteriophage DNA packaging motor. *Annu. Rev. Genet.*, **2008**, 42, 647–681. doi:10.1146/annurev.genet.42.110807.091545

Ravin, N.V. N15: the linear phage-plasmid. *Plasmid*, **2011**, 65(2), 102–109. doi:10.1016/j.plasmid.2010.12.004

Raya, R.R., H'bert, E.M. Isolation of Phage via Induction of Lysogens*, in: Clokie, M.R.J., Kropinski, A.M. (Eds.), Bacteriophages, Methods in Molecular Biology*, **2009**. Humana Press, Totowa, NJ, pages 23–32, ISBN: 978-1-58829-682-5. doi:10.1007/978-1-60327-164-6_3

Reverter, M., Sarter, S., Caruso, D., Avarre, J.-C., Combe, M., Pepey, E., Pouyaud, L., Vega-Heredía, S., de Verdal, H., Gozlan, R.E. Aquaculture at the crossroads of global warming and antimicrobial resistance. *Nat. Commun.*, **2020**, 11(1), 1870. doi:10.1038/s41467-020-15735-6

Ripp, S., Miller, R.V. Dynamics of the pseudolysogenic response in slowly growing cells of Pseudomonas aeruginosa. *Microbiology*, **1998**, 144(8), 2225–2232. doi:10.1099/00221287-144-8-2225

Ripp, S., Miller, R.V. The role of pseudolysogeny in bacteriophage-host interactions in a natural freshwater environment. *Microbiology*, **1997**, 143(6), 2065–2070. doi:10.1099/00221287-143-6-2065

Rohwer, F., Edwards, R. The Phage Proteomic Tree: a Genome-Based Taxonomy for Phage. *J. Bacteriol.*, **2002**, 184(16), 4529–4535. doi:10.1128/JB.184.16.4529-4535.2002

Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E.V., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Lavigne, R., Brister, J.R., Varsani, A., Amid, C., Aziz, R.K., Bordenstein, S.R., Bork, P., Breitbart, M., Cochrane, G.R., Daly, R.A., Desnues, C., Duhaime, M.B., Emerson, J.B., Enault, F., Fuhrman, J.A., Hingamp, P., Hugenholtz, P., Hurwitz, B.L., Ivanova, N.N., Labonté, J.M., Lee, K.-B., Malmstrom, R.R., Martinez-Garcia, M., Mizrachi, I.K., Ogata, H., Páez-Espino, D., Petit, M.-A., Putonti, C., Rattei, T., Reyes, A., Rodriguez-Valera, F., Rosario, K., Schriml, L., Schulz, F., Steward, G.F., Sullivan, M.B., Sunagawa, S., Suttle, C.A., Temperton, B., Tringe, S.G., Thurber, R.V., Webster, N.S., Whiteson, K.L., Wilhelm, S.W., Wommack, K.E., Woyke, T., Wrighton, K.C., Yilmaz, P., Yoshida, T., Young, M.J., Yutin, N., Allen, L.Z., Kyrpides, N.C., Eloe-Fadrosh, E.A. Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat. Biotechnol.*, **2019**, 37(1), 29–37. doi:10.1038/nbt.4306

Ryan, M.P., Pembroke, J.T. Brevundimonas spp: Emerging global opportunistic pathogens. *Virulence*, **2018**, 9(1), 480–493. doi:10.1080/21505594.2017.1419116

Safaei, S., Fatahi-Bafghi, M., Pouresmaeil, O. Role of Tsukamurella species in human infections: first literature review. *New Microbes New Infect.*, **2018**, 22, 6–12. doi:10.1016/j.nmni.2017.10.002

Saitou, N., Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees.. *Mol. Biol. Evol.*, **1987**, 4(4), 406–425. doi:10.1093/oxfordjournals.molbev.a040454

Sanger, F., Nicklen, S., Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.*, **1977**, 74(12), 5463–5467. doi:10.1073/pnas.74.12.5463

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B.W., Pruitt, K.D., Sherry, S.T. Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **2022**, 50(D1), D20–D26. doi:10.1093/nar/gkab1112

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K.D., Karsch-Mizrachi, I. GenBank. *Nucleic Acids Res.*, **2019**, 47(D1), D94–D99. doi:10.1093/nar/gky989

Schar, D., Zhao, C., Wang, Y., Larsson, D.G.J., Gilbert, M., Van Boeckel, T.P. Twenty-year trends in antimicrobial resistance from aquaculture and fisheries in Asia. *Nat. Commun.*, **2021**, 12(1), 5384. doi:10.1038/s41467-021-25655-8

Schmelcher, M., Donovan, D.M., Loessner, M.J. Bacteriophage endolysins as novel antimicrobials. *Future Microbiol.*, **2012**, 7(10), 1147–1171. doi:10.2217/fmb.12.97

Schneider, C.A., Rasband, W.S., Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **2012**, 9(7), 671–675. doi:10.1038/nmeth.2089

Schoch, C.L., Ciufo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Leipe, D., McVeigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J.P., Sun, L., Turner, S., Karsch-Mizrachi, I. NCBI Taxonomy: A comprehensive update on curation, resources and tools. *Database*, **2020**, 2020(2), 1–21. doi:10.1093/database/baaa062

Schoenfeld, T., Liles, M., Wommack, K.E., Polson, S.W., Godiska, R., Mead, D. Functional viral metagenomics and the next generation of molecular tools. *Trends Microbiol.*, **2010**, 18(1), 20–29. doi:10.1016/j.tim.2009.10.001

Schulz, P., Pajdak-Czaus, J., Siwicki, A.K. In Vivo Bacteriophages' Application for the Prevention and Therapy of Aquaculture Animals–Chosen Aspects. *Animals*, **2022**, 12(10), 1233. doi:10.3390/ani12101233

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **2003**, 13(11), 2498–2504. doi:10.1101/gr.1239303

Sillankorva, S.M., Oliveira, H., Azeredo, J. Bacteriophages and Their Role in Food Safety. *Int. J. Microbiol.*, **2012**, 2012, 863945. doi:10.1155/2012/863945

Simmonds, P., Adriaenssens, E.M., Zerbini, F.M., Abrescia, N.G.A., Aiewsakun, P., Alfenas-Zerbini, P., Bao, Y., Barylski, J., Drosten, C., Duffy, S., Duprex, W.P., Dutilh, B.E., Elena, S.F., García, M.L., Junglen, S., Katzourakis, A., Koonin, E.V., Krupovic, M., Kuhn, J.H., Lambert, A.J., Lefkowitz, E.J., Łobocka, M., Lood, C., Mahony, J., Meier-Kolthoff, J.P., Mushegian, A.R., Oksanen, H.M., Poranen, M.M., Reyes-Muñoz, A., Robertson, D.L., Roux, S., Rubino, L., Sabanadzovic, S., Siddell, S., Skern, T., Smith, D.B., Sullivan, M.B., Suzuki, N., Turner, D., Doorslaer, K.V., Vandamme, A.-M., Varsani, A., Vasilakis, N. Four principles to establish a universal virus taxonomy. *PLOS Biol.*, **2023**, 21(2), e3001922. doi:10.1371/journal.pbio.3001922

Śliwka, P., Ochocka, M., Skaradzińska, A. Applications of bacteriophages against intracellular bacteria. *Crit. Rev. Microbiol.*, **2022**, 48(2), 222–239. doi:10.1080/1040841X.2021.1960481

Slopek, S., Weber-Dabrowska, B., Dabrowski, M., Kucharewicz-Krukowska, A. Results of bacteriophage treatment of suppurative bacterial infections in the years 1981-1986.. *Arch. Immunol. Ther. Exp. (Warsz.)*, **1987**, 35(5), 569–583.

Söding, J., Biegert, A., Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, **2005**. doi:10.1093/nar/gki408

Ssekyanzi, A., Nevejan, N., Kabbiri, R., Wesana, J., Stappen, G.V. Knowledge, Attitudes, and Practices of Fish Farmers Regarding Water Quality and Its Management in the Rwenzori Region of Uganda. *Water*, **2023**, 15(1), 42. doi:10.3390/w15010042

Staley, J. Measurement of In Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. *Annu. Rev. Microbiol.*, **1985**, 39(1), 321–346. doi:10.1146/annurev.micro.39.1.321

Starmer, J., Stomp, A., Vouk, M., Bitzer, D. Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Comput. Biol.*, **2006**, 2(5), 454–466. doi:10.1371/journal.pcbi.0020057

Stewart, C.R., Casjens, S.R., Cresawn, S.G., Houtz, J.M., Smith, A.L., Ford, M.E., Peebles, C.L., Hatfull, G.F., Hendrix, R.W., Huang, W.M., Pedulla, M.L. The Genome of Bacillus subtilis Bacteriophage SPO1. *J. Mol. Biol.*, **2009**, 388(1), 48–70. doi:10.1016/j.jmb.2009.03.009

Strange, R.N., Scott, P.R. Plant disease: a threat to global food security. *Annu. Rev. Phytopathol.*, **2005**, 43, 83–116. doi:10.1146/annurev.phyto.43.113004.133839

Sullivan, M.J., Petty, N.K., Beatson, S.A. Easyfig: A genome comparison visualizer. *Bioinformatics*, **2011**, 27(7), 1009–1010. doi:10.1093/bioinformatics/btr039

Švānberga, K. No Latvijas notekūdeņiem jaunizdalīto Aeromonas salmonicida bakteriofāgu raksturošana [in Latvian; Characterization of novel Aeromonas salmonicida bacteriophages isolated from Latvian wastewaters], **2022**. Univeristy of Latvia, 122 pages.

Svircev, A., Roach, D., Castle, A. Framing the Future with Bacteriophages in Agriculture. *Viruses*, **2018**, 10(5), 218. doi:10.3390/v10050218

Tabib-Salazar, A., Mulvenna, N., Severinov, K., Matthews, S.J., Wigneshweraraj, S. Xenogeneic Regulation of the Bacterial Transcription Machinery. *J. Mol. Biol., RNA polymerase reaches 60*, **2019**, 431(20), 4078–4092. doi:10.1016/j.jmb.2019.02.008

Tan, S.Y., Tatsumura, Y. Alexander Fleming (1881–1955): Discoverer of penicillin. *Singapore Med. J.*, **2015**, 56(7), 366–367. doi:10.11622/smedj.2015105

Taylor, V.L., Fitzpatrick, A.D., Islam, Z., Maxwell, K.L. The Diverse Impacts of Phage Morons on Bacterial Fitness and Virulence. *Adv. Virus Res.*, **2019**, 103, 1–31. doi:10.1016/bs.aivir.2018.08.001

Tran, T.A.T., Struck, D.K., Young, R. Periplasmic Domains Define Holin-Antiholin Interactions in T4 Lysis Inhibition. *J. Bacteriol.*, **2005**, 187(19), 6631–6640. doi:10.1128/JB.187.19.6631-6640.2005

Turner, D., Kropinski, A.M., Adriaenssens, E.M. A Roadmap for Genome-Based Phage Taxonomy. *Viruses*, **2021**, 13(3), 506. doi:10.3390/v13030506

Turner, D., Shkoporov, A.N., Lood, C., Millard, A.D., Dutilh, B.E., Alfenas-Zerbini, P., van Zyl, L.J., Aziz, R.K., Oksanen, H.M., Poranen, M.M., Kropinski, A.M., Barylski, J., Brister, J.R., Chanisvili, N., Edwards, R.A., Enault, F., Gillis, A., Knezevic, P., Krupovic, M., Kurtböke, I., Kushkina, A., Lavigne, R., Lehman, S., Lobocka, M., Moraru, C., Moreno Switt, A., Morozova, V., Nakavuma, J., Reyes Muñoz, A., Rūmnieks, J., Sarkar, B.L., Sullivan, M.B., Uchiyama, J., Wittmann, J., Yigang, T., Adriaenssens, E.M. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch. Virol.*, **2023**, 168(2), 1–9. doi:10.1007/s00705-022-05694-2

Twort, F.W. AN INVESTIGATION ON THE NATURE OF ULTRA-MICROSCOPIC VIRUSES.. *The Lancet*, **1915**, 186(4814), 1241–1243. doi:10.1016/S0140-6736(01)20383-3

Vaara, M. Agents that increase the permeability of the outer membrane. *Microbiol. Rev.*, **1992**, 56(3), 395–411. doi:10.1128/mr.56.3.395-411.1992

Vázquez, R., Briers, Y. What's in a Name? An Overview of the Proliferating Nomenclature in the Field of Phage Lysins. *Cells*, **2023**, 12(15), 2016. doi:10.3390/cells12152016

Veesler, D., Cambillau, C. A Common Evolutionary Origin for Tailed-Bacteriophage Functional Modules and Bacterial Machineries. *Microbiol. Mol. Biol. Rev.*, **2011**, 75(3), 423–433. doi:10.1128/mmbr.00014-11

Vouga, M., Greub, G. Emerging bacterial pathogens: The past and beyond. *Clin. Microbiol. Infect.*, **2016**, 22(1), 12–21. doi:10.1016/j.cmi.2015.10.010

Wang, Jianbin., Jiang, Y., Vincent, M., Sun, Y., Yu, H., Wang, Jing., Bao, Q., Kong, H., Hu, S. Complete genome sequence of bacteriophage T5. *Virology*, **2005**, 332(1), 45–65. doi:10.1016/j.virol.2004.10.049

Weber-Dabrowska, B., Mulczyk, M., Görski, A. Bacteriophage therapy for infections in cancer patients. *Clin. Appl. Immunol. Rev.*, **2001**, 1(3–4), 131–134. doi:10.1016/S1529-1049(01)00015-0

Weigel, C., Seitz, H. Bacteriophage replication modules. *FEMS Microbiol. Rev.*, **2006**, 30(3), 321–381. doi:10.1111/j.1574-6976.2006.00015.x

Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.*, **1991**. doi:10.1128/jb.173.2.697-703.1991

Whitman, W.B., Coleman, D.C., Wiebe, W.J. Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci. U. S. A.*, **1998**, 95(12), 6578–6583. doi:10.1073/pnas.95.12.6578

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.*, **2017**, 13(6), 1–22. doi:10.1371/journal.pcbi.1005595

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*, **2015**, 31(20), 3350–3352. doi:10.1093/bioinformatics/btv383

Wigington, C.H., Sonderegger, D., Brussaard, C.P.D., Buchan, A., Finke, J.F., Fuhrman, J.A., Lennon, J.T., Middelboe, M., Suttle, C.A., Stock, C., Wilson, W.H., Wommack, K.E., Wilhelm, S.W., Weitz, J.S. Re-examination of the relationship between marine virus and microbial cell abundances. *Nat. Microbiol.*, **2016**, 1(3), 1–9. doi:10.1038/nmicrobiol.2015.24

Williamson, K.E., Fuhrmann, J.J., Wommack, K.E., Radosevich, M. Viruses in Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory. *Annu. Rev. Virol.*, **2017**, 4(1), 201–219. doi:10.1146/annurev-virology-101416-041639

Woese, C.R., Kandler, O., Wheelis, M.L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.. *Proc. Natl. Acad. Sci.*, **1990**, 87, 4576–4579. doi:10.1073/pnas.87.12.4576

World Health Organization. *Antimicrobial resistance* [online], **2021**. Available from: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance [accessed 28 August 2023].

World Health Organization. *Antibiotic resistance* [online], **2020**. Available from: https://www.who.int/news-room/fact-sheets/detail/antibiotic-resistance [accessed 28 August 2023].

Xu, J., Kiesel, B., Kallies, R., Jiang, F.-L., Liu, Y., Maskow, T. A fast and reliable method for monitoring of prophage-activating chemicals. *Microb. Biotechnol.*, **2018**, 11(6), 1112–1120. doi:10.1111/1751-7915.13042

Yoon, S.H., Ha, S.M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.*, **2017**. doi:10.1099/ijsem.0.001755

Żaczek, M., Weber-Dąbrowska, B., Międzybrodzki, R., Łusiak-Szelachowska, M., Górski, A. Phage Therapy in Poland – a Centennial Journey to the First Ethically Approved Treatment Facility in Europe. *Front. Microbiol.*, **2020**, 11(1056). doi:10.3389/fmicb.2020.01056

Zhang, W., Belton, B., Edwards, P., Henriksson, P.J.G., Little, D.C., Newton, R., Troell, M. Aquaculture will continue to depend more on land than sea. *Nature*, **2022**, 603(7900), E2–E4. doi:10.1038/s41586-021-04331-3

Zivin, R., Zehring, W., Rothman-Denes, L.B. Transcriptional map of bacteriophage N4: Location and polarity of N4 RNAs. *J. Mol. Biol.*, **1981**, 152(2), 335–356. doi:10.1016/0022-2836(81)90246-1

Zrelovs, N. Trīs no antarktiskās ledus-brīvas augsnes jaunizdalīto Caudovirales kārtas bakteriofāgu raksturošana [in Latvian; Characterization of three novel bacteriophages from order Caudovirales isolated from Antarctic ice-free soils], **2020**. Univeristy of Latvia, 126 pages.

Zrelovs, N., Cernooka, E., Dislers, A., Kazaks, A. Isolation and characterization of the novel Virgibacillus-infecting bacteriophage Mimir87. *Arch. Virol.*, **2020a**, 165(3), 737–741. doi:10.1007/s00705-019-04516-2

Zrelovs, N., Dislers, A., Kazaks, A. Corrigendum: Motley Crew: Overview of the Currently Available Phage Diversity. *Front. Microbiol.*, **2021a**, 11.

Zrelovs, N., Dislers, A., Kazaks, A. Genome Characterization of Nocturne116, Novel Lactococcus lactis-Infecting Phage Isolated from Moth. *Microorganisms*, **2021b**, 9(7), 1540. doi:10.3390/microorganisms9071540

Zrelovs, N., Dislers, A., Kazaks, A. Novel Erwinia persicina Infecting Phage Midgardsormr38 Within the Context of Temperate Erwinia Phages. *Front. Microbiol.*, **2020b**, 11(June). doi:10.3389/fmicb.2020.01245

Zrelovs, N., Dislers, A., Kazaks, A. Motley Crew: Overview of the Currently Available Phage Diversity. *Front. Microbiol.*, **2020c**, 11(October). doi:10.3389/fmicb.2020.579452

Zrelovs, N., Jansons, J., Dislers, A., Kazaks, A. Morganella Phage Mecenats66 Utilizes an Evolutionarily Distinct Subtype of Headful Genome Packaging with a Preferred Packaging Initiation Site. *Microorganisms*, **2022a**, 10(9), 1799. doi:10.3390/microorganisms10091799

Zrelovs, N., Jansons, J., Kazaka, T., Kazaks, A., Dislers, A. Three Phages One Host: Isolation and Characterization of Pantoea agglomerans Phages from a Grasshopper Specimen. *Int. J. Mol. Sci.*, **2023**, 24(3), 1820. doi:10.3390/ijms24031820

Zrelovs, N., Lamsters, K., Karuss, J., Krievans, M., Dislers, A., Kazaks, A. PVJ1 Is Not the First Tailed Temperate Phage Infecting Bacteria from Genus Psychrobacillus. Comment on Liu et al. Isolation and Characterization of the First Temperate Virus Infecting Psychrobacillus from Marine Sediments. Viruses 2022, 14, 108. *Viruses*, **2022b**, 14(3). doi:10.3390/v14030495

APPENDICES

Virion morphology	Membrane	Tail	Genetic material	Exemplar isolate genome (b or bp)	Realm	Kingdom	Phylum	Class	Order	Family	Selected species example(s)
Icosahedral, with internal membrane	+	-	linear dsDNA	14,927; 14,935	*Varidnaviria*	*Bamfordvirae*	*Preplasmiviricota*	*Tectiliviricetes*	*Kalamavirales*	*Tectiviridae*	*Alphatectivirus PRD1;* *Betatectivirus Bam35*
Icosahedral, with internal membrane	+	-	circular dsDNA	10,079	*Varidnaviria*	*Bamfordvirae*	*Preplasmiviricota*	*Tectiliviricetes*	*Vinavirales*	*Corticoviridae*	*Corticovirus PM2*
Icosahedral, with internal membrane	+	-	circular ssDNA	9,174	*Varidnaviria*	*Bamfordvirae*	*Preplasmiviricota*	*Ainoaviricetes*	*Lautamovirales*	*Finnlakeviridae*	*Finnlakevirus FLiP*
Enveloped	+	-	tripartite dsRNA	L: 6,374; M: 4,063; S: 2,948	*Riboviria*	*Orthornavirae*	*Duplornaviricota*	*Vidaverviricetes*	*Mindivirales*	*Cystoviridae*	*Cystovirus phi6*
Enveloped, pleomorphic	+	-	circular dsDNA	11,965	*incertae sedis*	*incertae sedis*	*incertae sedis*	*incertae sedis*	*incertae sedis*	*Plasmaviridae*	*Plasmavirus L2*
Spherical	-	-	(+)ssRNA	3,569	*Riboviria*	*Orthornavirae*	*Lenarviricota*	*Leviviricetes*	*Norzivirales*	*Fiersviridae*	*Emesvirus zinderi*
Filamentous	-	-	circular ssDNA	6,407	*Monodnaviria*	*Loebvirae*	*Hofneiviricota*	*Faserviricetes*	*Tubulavirales*	*Inoviridae*	*Inovirus M13*
Icosahedral	-	-	circular ssDNA	5,386	*Monodnaviria*	*Sangervirae*	*Phixviricota*	*Malgrandaviricetes*	*Petitvirales*	*Microviridae*	*Sinsheimervirus phiX174*
Prolate icosahedron, with tail	-	Long, contractile	linear dsDNA	168,903	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*Straboviridae*	*Tequatrovirus T4*
Icosahedral, with tail	-	Long, non-contractile	linear dsDNA	48,502	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*incertae sedis*	*Lambdavirus lambda*
Icosahedral, with tail	-	Short, non-contractile	linear dsDNA	39,937	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*Autographiviridae*	*Teseptimavirus T7*
Icosahedral, with tail	-	Long, contractile	linear dsDNA	36,717	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*incertae sedis*	*Muvirus mu*
Prolate icosahedron, with tail	-	Short, non-contractile	linear dsDNA	19,282	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*Salasmaviridae*	*Salasvirus phi29*
Icosahedral, with tail	-	Long, non-contractile	linear dsDNA	46,375	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*incertae sedis*	*Ravinvirus N15*
Icosahedral, with tail	-	Long, contractile	linear dsDNA	33,593	*Duplodnaviria*	*Heunggongvirae*	*Uroviricota*	*Caudoviricetes*	*incertae sedis*	*Peduoviridae*	*Peduovirus P2*

Appendix 1. General characteristics and selected taxonomy for some of the most well-studied bacteriophage species used as examples throughout this thesis.

Accession	**Phage**	**Genome length (bp)**	**Genome termini**	**GC content (%)**	**ORF count**	**Hypothetical proteins (%)**	**Morphology**	**Genus**	**Host**	**Similarity to the best total scoring BLASTN hit against the isolated *Caudoviricetes***
OK570184.1	*Pantoea* phage Nifs112	46,202	Short direct terminal repeats (296 bp)	50.21	59	47.46%	Podophage	*Eracentumvirus**	*Pantoea agglomerans*	*Erwinia phage* vB_EamP-S2 (NC_047917.1; identity ~88%)
OK570185.1	*Pantoea* phage Nufs112	45,951	Short direct terminal repeats (410 bp)	47.70	67	52.24%	Podophage	Gen. nov.*	*Pantoea agglomerans*	*Klebsiella* phage 6939 (OL362271.1; identity ~44%)
OK570186.1	*Pantoea* phage Nafs113	75,899	Circularly permuted and terminally redundant, no *pac* site discernible	54.07	130	70.00%	Myophage	*Lietduovirus**	*Pantoea agglomerans*	*Pantoea* phage vB_PagM_LIET2 (NC_048751.1; identity ~90%)
OL614104.1	*Morganella* phage Mecenats66	86,193	Circularly permuted and terminally redundant, *pac* site discernible	49.07	123	68.29%	Myophage	Gen. nov.*	*Morganella* sp.	N/A; no reasonable similarity to any isolated *Caudoviricetes* (identities <5%)
MW791312.1	*Lactococcus* phage Nocturne116	25,554	3' cos overhang (5'-CGCAGTAACT-3')	37.99	52	71.15%	Siphophage	Gen. nov.*	*Lactococcus lactis*	*Lactococcus* phage Q54 (NC_008364.1; identity ~25%)
MT410774.1	*Psychrobacillus* phage Spoks	36,472	3' cos overhang (5'-CGGTAGGGGGA-3')	35.83	52	40.38%	Siphophage	Gen. nov.*	*Psychrobacillus psychrodurans*	*Virgibacillus* phage Mimir87 (MK560763.1; identity ~11%)
MT325768.1	*Psychrobacillus* phage Perkons	136,811	Short direct terminal repeats (264 bp)	30.59	217	75.58%	Siphophage	Gen. nov.*	*Psychrobacillus psychrodurans*	N/A; no reasonable similarity to any isolated *Caudoviricetes* (identities <5%)
MN602881.1	*Erwinia* phage Midgardsormr38	50,485	Circularly permuted and terminally redundant, no *pac* site discernible	50.86	93	40.86%	Siphophage	*Midgardsormrvirus*	*Erwinia persicina*	*Erwinia* phage vB_EhrS_49 (NC_048197.1; identity ~20%)
MK560763.1	*Virgibacillus* phage Mimir87	48,016	3' cos overhang (5'-CGGATGGGC-3')	36.86	68	45.59%	Siphophage	Gen. nov.*	*Virgibacillus halotolerans*	*Halomonas* phage YPHTV-1 (ON854455.1; identity ~12%)

Appendix 2. Selected features of the novel tailed bacteriophages isolated within our lab and characterized within this thesis. Asterisk (*) in the genus column indicates our tentative designations, as only Midgardsormr38 is part of the official ICTV-recognized phage taxonomy as of now.