



Review

# How Metagenomics Has Transformed Our Understanding of Bacteriophages in Microbiome Research

Laura K. Inglis \* and Robert A. Edwards

Flinders Accelerator for Microbiome Exploration, College of Science and Engineering, Flinders University, Bedford Park, SA 5042, Australia

\* Correspondence: ingl0055@flinders.edu.au

**Abstract:** The microbiome is an essential part of most ecosystems. It was originally studied mostly through culturing but relatively few microbes can be cultured, so much of the microbiome was left unexplored. The emergence of metagenomic sequencing techniques changed that and allowed the study of microbiomes from all sorts of habitats. Metagenomic sequencing also allowed for a more thorough exploration of prophages, viruses that integrate into bacterial genomes, and how they benefit their hosts. One issue with using open-access metagenomic data is that sequences added to databases often have little to no metadata to work with, so finding enough sequences can be difficult. Many metagenomes have been manually curated but this is a time-consuming process and relies heavily on the uploader to be accurate and thorough when filling in metadata fields and the curators to be working with the same ontologies. Using algorithms to automatically sort metagenomes based on either the taxonomic profile or the functional profile may be a viable solution to the issues with manually curated metagenomes, but it requires that the algorithm is trained on carefully curated datasets and using the most informative profile possible in order to minimize errors.

**Keywords:** metagenomics; bacteriophage; microbiome



**Citation:** Inglis, L.K.; Edwards, R.A. How Metagenomics Has Transformed Our Understanding of Bacteriophages in Microbiome Research. *Microorganisms* **2022**, *10*, 1671. <https://doi.org/10.3390/microorganisms10081671>

Academic Editor: Grzegorz Wegrzyn

Received: 11 June 2022

Accepted: 16 August 2022

Published: 19 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The microbiome is the microbial component of an ecosystem. It includes bacteria, archaea, and viruses and is an essential part of most ecosystems. It can greatly influence human and environmental health, and it can be heavily impacted by human activities.

Microbiomes are usually very diverse, and a microbiome dominated by a few species, a dysbiosis, is often a cause or symptom of a disease state [1,2]. Similar to macro-scale food webs, bacteria and viruses operate under complex parasitic and/or symbiotic relationships. Studying these dynamics can provide insight into disease and provide potential solutions for pressing issues such as antimicrobial resistance.

Historical approaches to studying the microbiome relied on in vitro culture as the first step in isolating a workable sample. However, only a small fraction of microbes could be cultivated, leading to a massive underestimation of the diversity of the microbiome. This source of error is known as the great plate count anomaly [3]. Currently, the main approach to studying the microbiome is through metagenomic sequencing, which does not require laboratory culture.

This review will briefly summarise the history of metagenomics and describe microbiome research made possible through the availability of metagenomic sequencing. Then, it will explore potential approaches for categorising metagenomes based on their isolation environment and compare the pros and cons of manual versus algorithmic methods for this research. Finally, the review will consider how metagenomic sequencing has improved research into bacteriophages, particularly for determining how these viruses operate within microbiomes.

## 2. A Brief Overview of Metagenomics

The term metagenome refers to the sum of the genomes found in a tested sample and was first coined in 1998 [4]. Outside of examining the few microbes that could be cultured, the first attempts to examine the wider metagenome were performed by isolating total DNA from fresh samples, then cloning large DNA fragments into plasmids or bacterial artificial chromosomes (BACs) maintained in microbes such as *Escherichia coli*. High-throughput screening was then used to examine the chemical diversity produced by the clones [4] or, later, sequencing the cloned fragments [5]. The first published metagenomes were sampled from acid mine drainage biofilms in early 2000 [6]. Since then, the methods for analysing metagenomes have been consistently improving to the current day. Now, we perform metagenomic sequencing using massively parallel sequencing or deep sequencing. This involves sequencing millions of small fragments of DNA, and then recreating the genome by connecting the fragment sequences using bioinformatics analyses [7].

There are two main methods for microbiome sequencing: amplicon sequencing and whole genome sequencing. Amplicon sequencing detects only the target gene, and usually focuses on the highly conserved 16S gene because of its ubiquity in bacterial genomes. This approach is relatively cheap and simple; however, it is limited in scope: the common 16S gene primers only target some bacteria, and depending on what primers, reference databases, and bioinformatics settings are used different genera can be underrepresented or missed entirely [8]. Viruses, archaea, and Eukarya go undetected by this method. Whole genome sequencing, as the name implies, sequences all the DNA in a sample. This approach has the benefit of being able to detect viruses and has also been shown to be more accurate in detecting bacteria to the species level [9].

Technological improvements have led to metagenomic sequencing becoming cheaper and more accessible, and it is now being used in many areas of research, from environmental to medical sciences. Since metagenomic sequencing was first applied to diagnose infection in a human patient in 2014 [10], its use as a clinical tool has slowly increased [11,12]. The fact that metagenomic sequencing targets all a sample's genetic material at once means the approach has many advantages over traditional diagnostic methods, which are generally limited to microbes that are well studied [13–18] and to testing for only a couple of potential pathogens at a time.

Metagenomic sequencing is also useful for monitoring pathogens in non-human settings. Recently it has been applied to monitor sewage for the SARS-CoV-2 virus [19–21], and has proved to be an effective monitoring tool, detecting SARS-CoV-2 in a community before any patients tested positive [22]. Metagenomic sequencing is also applied more broadly to monitor pathogens in human communities, including both bacterial and viral pathogens [23], and to examine or monitor antibiotic resistance genes [18,24].

The wider microbiome is also explored with metagenomics [1,25,26]. Only 1% of microbes can be cultured [2], limiting our capacity to understand the entire microbiome using traditional methods. Metagenomics has enabled access to the remaining 99% that was almost entirely unknown to us 20 years ago [25] and may be used to quickly screen microbiomes for useful functions [27]. For example, these previously uncharacterized microbes could be important indicators of host health [1,25,26] or reveal opportunities to develop new antimicrobials [2,28,29].

The Human Microbiome Project began with the intention of documenting the microbiome of healthy humans [25,30], but that is proving to be a challenging endeavour as the taxonomic composition of the human microbiome varies between people [31].

While metagenomic sequencing does find many more species than culturing methods, it does have its own issues and biases. As there are not comprehensive reference genomes for every species, and creating reference genomes using cultured microbes cannot capture the natural variation found in the field, assembly and binning software do not always correctly reconstruct the genome. Repeat regions and sequences that have a significantly different GC content than the rest of the genome are frequently not binned by common software [32].

### 3. Curating Metagenomes

Data curation is essential when working with metagenomes sourced from online databanks. Curation in this context refers to collecting the data and sorting it into useable categories. Metagenomes can either be manually curated by the researchers looking to use the data, or by an algorithm trained to sort the metagenomes based on different features of the sequences.

Ontologies are systems for categorising data. Numerous ontologies designed for many different types of data can be found on the online repository [www.bioontology.org](http://www.bioontology.org) (accessed on 1 April 2022) [33]. One ontology for categorising metagenomic data is the Functional Ontology Assignments for Metagenomes (FOAM) which uses Hidden Markov Models to classify gene functions [34]. Some other biological ontologies include the Biological Collections Ontology [35], or Interlinking Ontology for Biological Concepts [36]. One of the better-known ontologies for categorising biological samples is the environment ontology (ENVO). ENVO uses a directed acyclic graph (DAG) to categorise metagenomes based on the environment from which samples are sourced [37].

Because “environment” is a vaguely defined concept that is often made up of many smaller factors, ENVO uses multiple descriptors in its categorization process. For example, the surface of a stone at the bottom of a lake is both a rock and a lake environment. Depending on factors such as the depth of the lake and the water clarity, the microbiome of the rock surface may vary significantly. While the capability to factor in such details is important, categories can end up being split incredibly finely. In extreme cases, any ontology that attempts to account for all environmental details risks becoming overly complex and difficult to work with in terms of manual sorting of samples. Many ontologies avoid this by being more specialised for different research fields such as marine environments [38]. Training an algorithm to sort samples into the finer categories of an ontology would alleviate this difficulty. However, automatic curation has its own set of issues that need to be considered, as discussed below.

### 4. Issues with Curating Metagenomes

Manually curated genomes are susceptible to human error and rely on the data generator to provide sufficient and accurate information while uploading the sequences. As a result, manual curations often lack critical information and contain mistakes in the metadata provided.

Over the last few years, the number of metagenomes uploaded to databases such as the Sequence Read Archive (SRA) or MGnify has increased exponentially [39,40]. While this does provide more data for anyone to use the amounts of low-quality, contaminated, or mislabelled sequences have also increased. This adds to the amount of work required to manually curate enough genomes. While automated curation of metagenomes could address these issues, this approach requires that the categories have little overlap in the data being analysed by the algorithm, which can prove difficult. The datasets used to train such an algorithm would also need to be carefully curated to avoid contaminated or low-quality sequences, or sequences collected using very different methods. For example, in the SRA database, both amplicon sequencing and whole-genome shotgun sequencing can be categorised as metagenomic data [41]. This can be an issue as they provide different information that could cause issues for an algorithm trying to compare them. An algorithm can successfully differentiate between the two types of data; however, as the partition engine, PARTIE, has sorted many sequences from the SRA into amplicon and whole genome sequencing datasets [41]. Any attempts to automatically curate metagenomic data from the SRA database will need to take the broad definition of metagenomic data into account when collecting the training data.

### 5. Categorising Metagenomes

Distinguishing distinct environment categories for microbiome ontologies is difficult for a number of reasons. For human microbiomes, sequences vary greatly between

individuals [42–44], between different locations on a single individual [43], and even within the same location on a single individual over time [42,43,45].

The metagenome of an external environment also changes over time [46] or in response to minor variations [47] in features such as water depth [5]. Furthermore, large variations within environments may warrant subdividing the environmental category further. For example, previous attempts to test the ability of algorithms to assign metagenomes to environments had high rates of incorrect assignments in some categories, particularly the “human respiratory” category [48]. This category included the lungs, mouth, nose, throat, and a few dental plaque samples. The large error makes sense, as both anatomically and biologically the mouth and the lungs are entirely different environments and cultivate different microbiomes [45].

Depending on how specific the ontology needs to be, separate categories for healthy and unhealthy sample environments should also be considered, as the microbiomes of healthy and unhealthy humans and animals can vary significantly [43,49,50]. However, unhealthy microbiomes can also blur the lines between closely linked environments as opportunistic microbes take the chance to colonise other environments. For example, *Rothia mucilaginosa*, a common oral microbe, has been shown to infect the lungs of cystic fibrosis patients [51,52].

## 6. Using AI to Curate Metagenomes

Most existing genomic databases do not require uploaders to provide comprehensive sample information (metadata) during the submission process. This results in many sequences being submitted to databases with little to no context. Manual curation often relies on this contextual information. The impacts of these missing details can be broad, as they translate into reducing the pool of sequences that can be used for other projects. For example, a water sample without metadata detailing the depth at which the sample was taken will be unable to be used to answer a question about the difference in microbiomes at various water depths as it cannot be assigned to any relevant category.

Artificial intelligence (AI) algorithms hold considerable promise in metagenomics, including for curation. For example, differences between metagenomes could be used to automatically and accurately curate based on the environment from which samples are isolated. However, this sort of decision-making activity will only happen once enough existing data is categorised and labelled appropriately.

It is also important to consider hierarchy of labelling: for example, how should the choice of labelling categories be managed for samples collected from different places on the same person? We previously used a random forest algorithm to sort metagenomes into categories based on their environmental source with 78.5% accuracy [48]. There is potential to improve that accuracy by either refining the training data or categories, and/or by changing the type of information the algorithm uses to curate the metagenomes.

Automated curation could radically change how researchers do microbiome research. Sequences that are unusable for large scale data integrations due to insufficient metadata could become available if a pipeline for automatic curation became available.

## 7. Functional Diversity vs. Taxonomic Diversity

Different types of data can be used for curation. For example, in previous work by this lab the taxonomic profile of metagenomes was used to curate them [48]. This approach is tempered because taxonomic make-up can vary greatly in different variations of the same environment. For example, the human microbiome varies greatly between people and can be influenced by diet [44], weight [42], medication [53], and many other factors [44,54].

Another approach to curation is to use the functional profile. These genes are often relatively conserved across microbiomes within similar environments, with more similarity in the relative abundance of functional genes compared to bacterial phyla in the microbiome of different environments [31,42,55]. It is possible a ‘core microbiome’ of functional genes that is similar across groups of environments may exist.

Overall, while taxonomic differences between people’s microbiomes can differentiate between samples on a personal scale, use of the functional profile for larger-scale metagenome curation could be more accurate. However, further research is required to compare the two curation methods to consolidate this hypothesis.

## 8. Phages and the Microbiome

Access to metagenomics has transformed how we study the microbiome, including research focusing on bacteriophages. Commonly referred to as phages, these viruses infect bacteria and are an important part of almost every microbiome. Individual phages often have a very limited host range, although almost every species of bacteria have phages that prey on them [56]. There are two main categories of phages, virulent phages, and temperate phages.

Virulent phages infect bacteria and hijack their replication machinery to make more copies of themselves, before releasing the new virions into the environment, often by lysing the bacterial host [57]. These lytic viruses are thought to kill approximately 20% of the ocean’s microbial biomass each day [58]. The targeted application of lytic phages to treat bacterial infections in humans and animals can be used as an alternative or to augment antibiotics and has seen recent success in compassionate-use cases [59–65]. Not all virulent phages kill their hosts however, some filamentous phages are extruded from their hosts without killing them and may help their bacterial hosts. For example, phage Pf4 helps its *Pseudomonas aeruginosa* host with biofilm maintenance [66,67].

Temperate phages switch between lytic and lysogenic lifecycles in response to environmental conditions and/or external stresses that induce the phage. When in a lysogenic cycle they either integrate into the bacterial host’s DNA and replicate with the host’s chromosome, or are maintained extrachromosomally as plasmids. Referred to as prophages, they can remain within hosts for thousands of generations [68].

Through lysogenic conversion, host bacteria can gain a variety of potentially beneficial traits that were encoding by the prophage’s DNA. For example, many prophages provide their hosts with protection against infection by other phages [69,70]. Prophages can also provide genes conferring resistance to antimicrobial drugs, protection from immune responses, or production of toxins [71–74]. For example, a prophage encodes the toxin genes that enable *Vibrio cholerae* to cause the life-threatening disease cholera [75]. The symbiotic relationship between phage and host can be parasitic as well as mutualistic: prophages can be a source of Selfish Genetic Elements which, as far as we know, only increase the replication of their own DNA and have a neutral or detrimental parasitic relationship with the host [76].

## 9. Piggyback-the-Winner and Other Hypotheses

The mechanisms that trigger the lytic-lysogenic switch are most well known in phage lambda [77], but various studies on the abundance of free temperate phages in various environments have led to a few hypotheses about general environmental conditions that influence the change in phage lifecycles. Three popular hypotheses are the Piggyback-the-Winner (PTW) [78], Piggyback-the-Loser (PTL) (aka refugium hypothesis), and Kill-the-Winner (KTW) hypotheses [79]. Each hypothesis is characterized by different environmental conditions (Table 1).

**Table 1.** A brief summary of the different hypotheses for explaining lysogeny rates. Both Piggyback-the-Winner (PTW) and Kill-the-Winner hypotheses (KTW) occur in similar environmental conditions.

Lysogeny Hypothesis	Environmental Conditions	Level of Lysogeny
Kill-the-Winner	High energy, cells growing	Low
Piggyback-the-Winner	High energy, cells growing	High
Piggyback-the-Loser	Low energy, little growth	High

The Kill-the-Winner hypothesis is the oldest of the three and was first proposed in 2000 by Thingstad and looks at the problem from the perspective of the lytic lifecycle. KTW proposes that a lytic phage is more likely to come across a bacterium it can infect if that bacterium is more abundant. Therefore, abundant bacteria are encountering and being infected by lytic phages more often than bacteria that are not growing as robustly [79].

The Piggyback-the-Winner hypothesis was the next one proposed by Knowles et al. in 2016. In contrast to the KTW hypothesis, it suggests that while host bacterial populations are replicating quickly, the phage uses a lysogenic lifecycle to take advantage of the host's success [78]. This hypothesis also suggests the phage's lytic lifecycle is triggered when bacteria start to become stressed. They supported their hypothesis from data from coral reef microbiomes and compared that data to previous models as well as their own hypothesis.

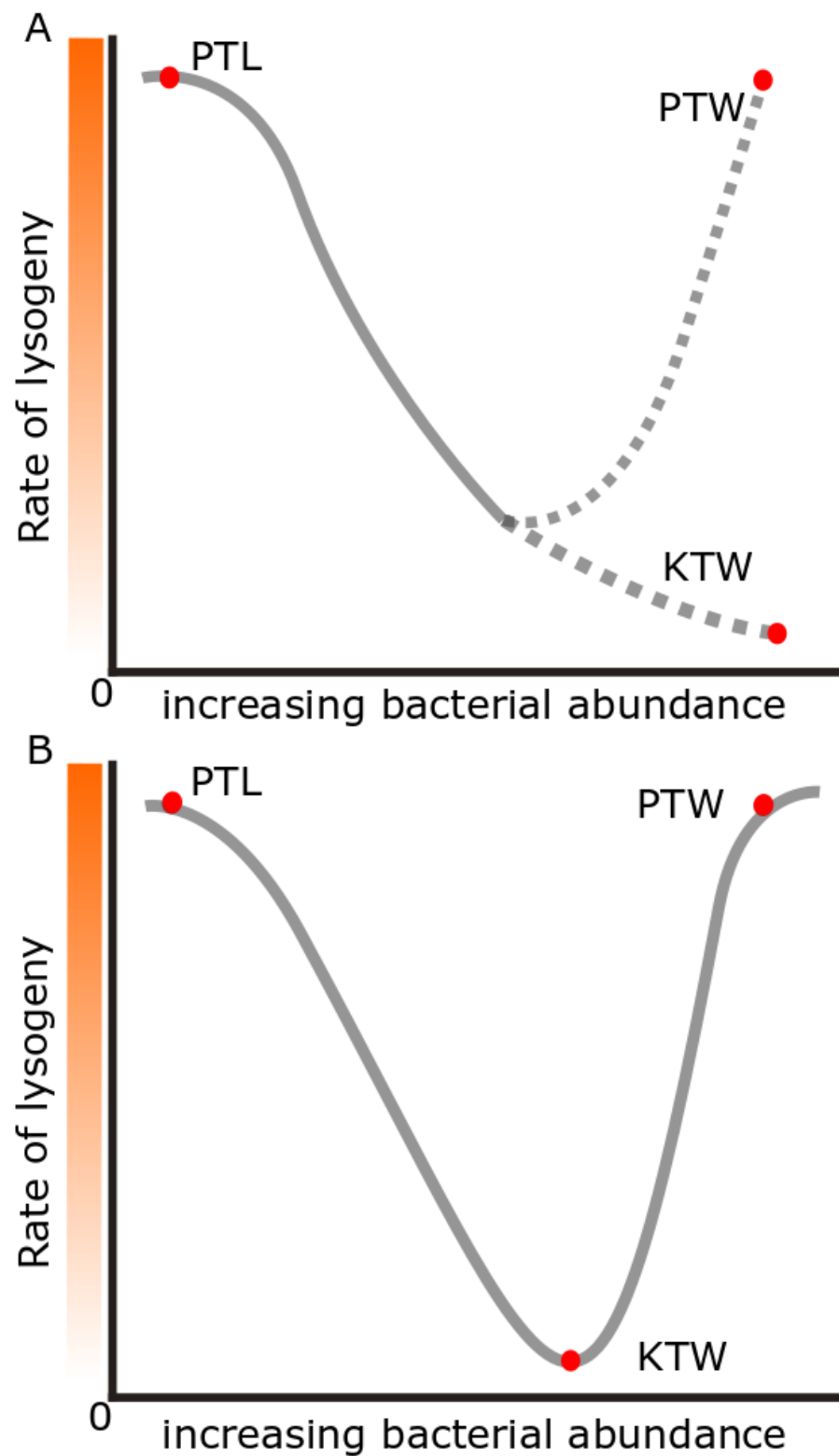
The third hypothesis is Piggyback-the-Loser (PTL) that suggests the opposite of the PTW hypothesis. The PTL hypothesis argues that when the host is growing slowly it is more beneficial to remain in a lysogenic state, as it is less likely that the phages produced through lysis will find a new host [80]. This hypothesis has also been called Piggyback-the-Persistent (PTP) [81] as the prophages were found in host species that persistently occurred at low abundance. PTP was first proposed in 2019 by Paterson et al. This study looked at the virus–microbe ratio in aquifer environments and found that the viral abundance was incredibly low. They were also one of the first groups to float the idea that the three hypotheses are competing strategies that become more or less prominent in different conditions [81].

Each of these hypotheses explains phage–host dynamics in studied environments. For example, PTL occurs in some regions of the deep sea or polar oceans [80,82] but PTW dynamics occur in some coral reef regions [83]. However, each hypothesis is an incomplete explanation of microbiome dynamics on its own. PTW and KTW also appear to be conflicting hypotheses (Figure 1) as they both have been shown to occur in high-energy environments but have the opposite results [79,80,83]. However, a recent study has taken a more in depth look at how the three hypotheses can coexist [80].

Silveira et al. propose that a combination of different mechanisms results in lysogeny being promoted at both very high and very low host abundances [80]. Their new hypothesis suggests that the effects of microbial density, community diversity, host–virus interactions, host metabolism, and the environment combine to either promote or repress lysis depending on which factors are more prominent. Overall, the main factors at play in determining whether a phage will be lytic or lysogenic are thought to be (1) chance of co-infections, and (2) host metabolism.

Co-infection is one of the best predictors for lysogenization [80] as it increases the abundance of phage repressors in the host cell, preventing lysis [84]. Slower host growth increases the window for co-infections to occur, which increases the potential for co-infections to promote lysogeny, supporting the PTL hypothesis. Conversely, hosts that are growing quickly have a shorter window for co-infection. However, eventually, the abundance of phages in the environment will increase to the point where co-infections happen quickly, leading to increased lysogeny at high host abundances which supports the PTW hypothesis. Co-infection is also affected by community diversity. When the community is less diverse phages are more likely to encounter bacteria that they can infect, which increases the probability of coinfections (compared to the lower probability of coming across an appropriate host if it is a relatively rare species).

Like co-infection, cell metabolism favours lysogeny at the extreme ends of the spectrum. In a low-energy environment where the hosts are growing slowly, cell starvation causes a signalling cascade. This increases the abundance of phage repressors [84], favouring lysogeny. When the hosts are not starving, they increase their stores of ATP. This allows the host proteases to degrade the phage repressors, and increase the chance of lysis [84]. However, in a very high-energy environment with limited oxygen, bacteria switch to a less efficient form of metabolism (e.g., fermentation) leading to low ATP/high NADPH levels. This prevents the host proteases from destroying the phage repressors and thereby limiting lysis.



**Figure 1.** Plotting the hypothetical rates of lysogeny between the PTL and KTW/PTW dynamics shows how conflicting the hypotheses appear to be (A). However, a recent hypothesis [80] suggests that the previous hypotheses can coexist and PTW overtakes KTW dynamics at high enough bacterial abundances (B).

As influencing factors, co-infection and metabolism work together to favour lysogeny at both very high and very low bacterial abundances. Silveira and colleagues suggest the rate of lysis peaks in environments with a bacterial concentration of  $10^6$  per millilitre [80].

They propose that starvation and large windows for co-infection to occur favour lysogeny in low energy environments, while high bacterial concentrations and inefficient metabolism favour lysogeny in higher energy environments.

Silveira et al. suggests approximate bacterial concentrations for different virus–host ratios. In theory, this means that this hypothesis could be tested by counting the prophages, viruses, and bacteria in different environmental metagenomes and modelling environmental carrying capacity from the number of prophages in each sample [68]. A further test is to experimentally manipulate the virus-to-microbe ratio (VMR), for example by adding phages or removing bacteria (e.g., by filtration).

## 10. Prophage Abundances in Different Environments

Environments vary greatly in the amount of bacterial growth they can sustain, called the carrying capacity. Each situation exists at a unique point on Silveira et al. model [80]. For example, there are likely to be relatively few prophages per cell in low-energy environments such as the deep ocean, though likely still enough to trigger lysogeny, because prophages take resources to maintain. In higher energy environments—such as the ocean’s surface or a coral reef—lytic growth is also the dominant lifestyle, so very few prophages would be expected in bacterial genomes. In highly productive environments, such as coastal waters and estuaries, the environments favour lysogeny. However, here, the energy cost to maintain the prophages is not as big of a limiting factor (because energy is plentiful), so the number of prophages per cell should theoretically be much greater than in low energy environments.

Numerous studies have examined phages across many different environments, including in various animals [85], in humans [86–89], from most aquatic environments [71,90–93], and from soil [94–96]. While many of these papers examine phages from single environments, there are comparatively few studies that compare phages from different environments in one analysis. Of these, many only compare a couple of environments at once [71,89,93,97]. As different prophage detection methods can produce slightly different results [98], comparing phages from environmental samples that have been collected and analysed with various methods could produce inaccurate results.

## 11. Identifying Prophages

The first step in analysing prophages is to identify their sequences in the bacterial genome. The approaches for doing so broadly separate into two main categories: experimental induction, or computational examination of a bacterial genome for prophage-like regions.

Induction involves causing lysis in the bacteria, which releases the phage into solution. It is one of the oldest methods for detecting phages and uses chemicals that signal to the phage to switch to a lytic lifecycle, destroying the host bacteria and releasing the phages into the solution. The issue with this method is that it can only be used on culturable bacteria and incomplete induction will cause the phage-host ratio to be skewed. Although we have discovered many chemical inducers, there are also plenty of prophages where we have not identified the signals they are sensing, and so cannot induce them.

Computational examination of bacterial genomes first requires that the bacterial DNA is sequenced, then the sequences can be examined for genes that belong to prophages either manually or by using programs such as PhiSpy [99] and many others [98]. The accuracy of these predictions is affected by the quality of the sequence assembly. As genomes are being assembled small, overlapping segments of DNA sequences are pieced together to form larger sequences, or contigs. A genome with fewer contigs is more complete than one with more contigs. A genome with many contigs generally has a lot more predicted prophages than genomes with few contigs and the rate of false positives increases as the number of contigs in a genome increases as it is more likely for a smaller segment of DNA to appear the same as a prophage by chance alone. A comparison of eight different prophage prediction tools showed that each package has its strengths and weaknesses [98]. Some sacrificed speed and efficiency for accuracy while others tried to find a balance. Some had good



precision but that necessitates an increased chance of false-positive results, while others increased recall performance which comes with more false positives. The pros and cons of each program need to be evaluated by users before deciding which is most appropriate for each project.

## 12. Conclusions

Metagenomic sequencing is useful for finding prophages from many different environmental conditions, but many genomes are added to databases without the inclusion of comprehensive metadata. Being able to automatically sort these sequences into an environmental ontology would allow for these sequences to be useful in future projects, but we need considerably more high-quality data to determine how best to sort these sequences.

Prophages play significant roles in the microbiomes of many species and in different environments. Phages can protect their host from deadly infections, and give their host access to beneficial genes such as antimicrobial resistance or toxin production. The way phages interact with their host changes depending on the environment. Researchers have conducted many studies on phages from various environments, and developed hypotheses regarding what factors influence survival strategies, such as the lytic/lysogenic decision. However, there is still much more to learn about how prophages interact with their hosts under different conditions.

Learning more about metagenomes and prophages could provide many insights into human and environmental health, while obtaining a better understanding of what a healthy microbiome should be may enable us to detect changes more quickly or accurately in microbiomes that could be a sign of disease.

**Author Contributions:** L.K.I. and R.A.E. conceived of this study and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by an award from US National Institutes of Health NIDDK RC2DK116713 and an award from the Australian Research Council DP220102915 to R.A.E.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mirsepasi-Lauridsen, H.C.; Vrankx, K.; Engberg, J.; Friis-Møller, A.; Brynskov, J.; Nordgaard-Lassen, I.; Petersen, A.M.; Kroghfelt, K.A. Disease-Specific Enteric Microbiome Dysbiosis in Inflammatory Bowel Disease. *Frontiers in Medicine*. *Front. Med.* **2018**, *5*, 304. [[CrossRef](#)] [[PubMed](#)]
2. de la Calle, F. Marine microbiome as source of natural products. *Microb. Biotechnol.* **2017**, *10*, 1293–1296. [[CrossRef](#)] [[PubMed](#)]
3. Staley, J.T.; Konopka, A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* **1985**, *39*, 321–346. [[CrossRef](#)] [[PubMed](#)]
4. 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*, R245–R249. [[CrossRef](#)]
5. DeLong, E.F.; Preston, C.M.; Mincer, T.; Rich, V.; Hallam, S.J.; Frigaard, N.-U.; Martinez, A.; Sullivan, M.B.; Edwards, R.; Brito, B.R.; et al. Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior. *Science* **2006**, *311*, 496–503. [[CrossRef](#)]
6. Tyson, G.W.; Chapman, J.; Hugenholtz, P.; Allen, E.E.; Ram, R.J.; Richardson, P.M.; Solovyev, V.V.; Rubin, E.M.; Rokhsar, D.S.; Banfield, J.F. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **2004**, *428*, 37–43. [[CrossRef](#)]
7. Papudeshi, B.; Haggerty, J.M.; Doane, M.; Morris, M.M.; Walsh, K.; Beattie, D.T.; Pande, D.; Zaeri, P.; Silva, G.G.Z.; Thompson, F.; et al. Optimizing and evaluating the reconstruction of Metagenome-assembled microbial genomes. *BMC Genom.* **2017**, *18*, 915. [[CrossRef](#)]
8. Abellan-Schneyder, I.; Machado Monica, S.; Reitmeier, S.; Sommer, A.; Sewald, Z.; Baumbach, J.; List, M.; Neuhaus, K.; Tringe Susannah, G. Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing. *mSphere* **2021**, *6*, e01202-20. [[CrossRef](#)]

9. Thompson, C.C.; Chimetto, L.; Edwards, R.A.; Swings, J.; Stackebrandt, E.; Thompson, F.L. Microbial genomic taxonomy. *BMC Genom.* **2013**, *14*, 913. [CrossRef]
10. Wilson, M.R.; Naccache, S.N.; Samayoa, E.; Biagtan, M.; Bashir, H.; Yu, G.; Salamat, S.M.; Somasekar, S.; Federman, S.; Miller, S.; et al. Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing. *N. Engl. J. Med.* **2014**, *370*, 2408–2417. [CrossRef]
11. Brown, J.R.; Bharucha, T.; Breuer, J. Encephalitis diagnosis using metagenomics: Application of next generation sequencing for undiagnosed cases. *J. Infect.* **2018**, *76*, 225–240. [CrossRef] [PubMed]
12. Lim, Y.W.; Evangelista, J.S.; Schmieder, R.; Bailey, B.; Haynes, M.; Furlan, M.; Maughan, H.; Edwards, R.; Rohwer, F.; Conrad, D.; et al. Clinical Insights from Metagenomic Analysis of Sputum Samples from Patients with Cystic Fibrosis. *J. Clin. Microbiol.* **2014**, *52*, 425–437. [CrossRef] [PubMed]
13. de Vries, J.J.C.; Brown, J.R.; Couto, N.; Beer, M.; Le Mercier, P.; Sidorov, I.; Papa, A.; Fischer, N.; Oude Munnink, B.B.; Rodriguez, C.; et al. Recommendations for the introduction of metagenomic next-generation sequencing in clinical virology, part II: Bioinformatic analysis and reporting. *J. Clin. Virol.* **2021**, *138*, 104812. [CrossRef] [PubMed]
14. Winter, A.K.; Hegde, S.T. The important role of serology for COVID-19 control. *Lancet Infect. Dis.* **2020**, *20*, 758–759. [CrossRef]
15. Chiu, C.Y.; Miller, S.A. Clinical metagenomics. *Nat. Rev. Genet.* **2019**, *20*, 341–355. [CrossRef]
16. Gupta, S.; Mortensen, M.S.; Schjørring, S.; Trivedi, U.; Vestergaard, G.; Stokholm, J.; Bisgaard, H.; Krogfelt, K.A.; Sørensen, S.J. Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. *Commun. Biol.* **2019**, *2*, 291. [CrossRef]
17. Hematian, A.; Sadeghifard, N.; Mohebi, R.; Taherikalani, M.; Nasrolahi, A.; Amraei, M.; Ghafourian, S. Traditional and Modern Cell Culture in Virus Diagnosis. *Osong Public Health Res. Perspect.* **2016**, *7*, 77–82. [CrossRef]
18. Schmieder, R.; Edwards, R. Insights into antibiotic resistance through metagenomic approaches. *Future Microbiol.* **2012**, *7*, 73–89. [CrossRef]
19. Crits-Christoph, A.; Kantor, R.S.; Olm, M.R.; Whitney, O.N.; Al-Shayeb, B.; Lou, Y.C.; Flamholz, A.; Kennedy, L.C.; Greenwald, H.; Hinkle, A.; et al. Genome Sequencing of Sewage Detects Regionally Prevalent SARS-CoV-2 Variants. *mBio* **2021**, *12*, e02703-20. [CrossRef]
20. COVID-19 Sewage Surveillance Program-COVID-19 (Coronavirus). 2022. Available online: <https://www.health.nsw.gov.au/Infectious/covid-19/Pages/sewage-surveillance.aspx> (accessed on 1 April 2022).
21. Landgraf, C.; Wang, L.Y.R.; Buchanan, C.; Wells, M.; Schonfeld, J.; Bessonov, K.; Ali, J.; Robert, E.; Nadon, C. Metagenomic sequencing of municipal wastewater provides a near-complete SARS-CoV-2 genome sequence identified as the B.1.1.7 variant of concern from a Canadian municipality concurrent with an outbreak. *MedRxiv* **2021**. [CrossRef]
22. Public Health Alert-New Venues of Concern-News. 2022. Available online: [https://www.health.nsw.gov.au/news/Pages/20210813\\_03.aspx](https://www.health.nsw.gov.au/news/Pages/20210813_03.aspx) (accessed on 1 April 2022).
23. Rothman, J.A.; Loveless, T.B.; Griffith, M.L.; Steele, J.A.; Griffith, J.F.; Whiteson, K.L. Metagenomics of Wastewater Influent from Southern California Wastewater Treatment Facilities in the Era of COVID-19. *Microbiol. Resour. Announc.* **2020**, *9*, e00907-20. [CrossRef] [PubMed]
24. Hendriksen, R.S.; Munk, P.; Njage, P.; van Bunnik, B.; McNally, L.; Lukjancenko, O.; Röder, T.; Nieuwenhuijse, D.; Pedersen, S.K.; Kjeldgaard, J.; et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **2019**, *10*, 1124. [CrossRef] [PubMed]
25. Cho, I.; Blaser, M.J. The human microbiome: At the interface of health and disease. *Nat. Rev. Genet.* **2012**, *13*, 260–270. [CrossRef] [PubMed]
26. Vandenkoornhuyse, P.; Quaiser, A.; Duhamel, M.; Le Van, A.; Dufresne, A. The importance of the microbiome of the plant holobiont. *New Phytol.* **2015**, *206*, 1196–1206. [CrossRef]
27. Dinsdale, E.A.; Edwards, R.A.; Hall, D.; Angly, F.; Breitbart, M.; Brulc, J.M.; Furlan, M.; Desnues, C.; Haynes, M.; Li, L.; et al. Functional metagenomic profiling of nine biomes. *Nature* **2008**, *452*, 629–632. [CrossRef]
28. Hover, B.M.; Kim, S.-H.; Katz, M.; Charlop-Powers, Z.; Owen, J.G.; Ternei, M.A.; Maniko, J.; Estrela, A.B.; Molina, H.; Park, S.; et al. Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat. Microbiol.* **2018**, *3*, 415–422. [CrossRef]
29. Tortorella, E.; Tedesco, P.; Palma Esposito, F.; January, G.G.; Fani, R.; Jaspars, M.; De Pascale, D. Antibiotics from Deep-Sea Microorganisms: Current Discoveries and Perspectives. *Mar. Drugs* **2018**, *16*, 355. [CrossRef]
30. Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The Human Microbiome Project. *Nature* **2007**, *449*, 804–810. [CrossRef]
31. Lloyd-Price, J.; Abu-Ali, G.; Huttenhower, C. The healthy human microbiome. *Genome Med.* **2016**, *8*, 51. [CrossRef]
32. Nelson, W.C.; Tully, B.J.; Mobberley, J.M. Biases in genomes reconstruction from metagenomic data. *PeerJ* **2020**, *8*, e10119. [CrossRef]
33. Welcome to the NCBO BioPortal | NCBO BioPortal. 2022. Available online: <https://bioportal.bioontology.org> (accessed on 1 April 2022).
34. Prestat, E.; David, M.M.; Hultman, J.; Taş, N.; Lamendella, R.; Dvornik, J.; Mackelprang, R.; Myrold, D.D.; Jumpponen, A.; Tringe, S.G.; et al. FOAM (Functional Ontology Assignments for Metagenomes): A Hidden Markov Model (HMM) database with environmental focus. *Nucleic Acids Res.* **2014**, *42*, e145. [CrossRef] [PubMed]

35. Walls, R.L.; Deck, J.; Guralnick, R.; Baskauf, S.; Beaman, R.; Blum, S.; Bowers, S.; Buttigieg, P.L.; Davies, N.; Endresen, D.; et al. Semantics in Support of Biodiversity Knowledge Discovery: An Introduction to the Biological Collections Ontology and Related Ontologies. *PLoS ONE* **2014**, *9*, e89606. [[CrossRef](#)] [[PubMed](#)]
36. Kushida, T.; Kozaki, K.; Tateisi, Y.; Watanabe, K.; Masuda, T.; Matsumura, K.; Kawamura, T.; Takagi, T. Efficient Construction of a New Ontology for Life Sciences by Sub-classifying Related Terms in the Japan Science, Technology Agency Thesaurus. In Proceedings of the 8th International Conference on Biomedical Ontologies (ICBO), Newcastle upon Tyne, UK, 13–15 September 2017.
37. Buttigieg, P.L.; Morrison, N.; Smith, B.; Mungall, C.J.; Lewis, S.E.; The ENVO Consortium. The environment ontology: Contextualising biological and biomedical entities. *J. Biomed. Semant.* **2013**, *4*, 43. [[CrossRef](#)] [[PubMed](#)]
38. Blumberg, K.L.; Ponsero, A.J.; Bomhoff, M.; Wood-Charlson, E.M.; DeLong, E.F.; Hurwitz, B.L. Ontology-Enriched Specifications Enabling Findable, Accessible, Interoperable, and Reusable Marine Metagenomic Datasets in Cyberinfrastructure Systems. *Front. Microbiol.* **2021**, *12*, 765268. [[CrossRef](#)]
39. Mitchell, A.L.; Almeida, A.; Beracochea, M.; Boland, M.; Burgin, J.; Cochrane, G.; Crusoe, M.R.; Kale, V.; Potter, S.C.; Richardson, L.J.; et al. MGnify: The microbiome analysis resource in 2020. *Nucleic Acids Res.* **2020**, *48*, D570–D578. [[CrossRef](#)]
40. Available online: <https://www.ncbi.nlm.nih.gov/sra/docs/sragrowth/> (accessed on 1 April 2022).
41. Torres, P.J.; Edwards, R.A.; McNair, K.A. PARTIE: A partition engine to separate metagenomic and amplicon projects in the Sequence Read Archive. *Bioinformatics* **2017**, *33*, 2389–2391. [[CrossRef](#)]
42. Turnbaugh, P.J.; Hamady, M.; Yatsunenko, T.; Cantarel, B.L.; Duncan, A.; Ley, R.E.; Sogin, M.L.; Jones, W.J.; Roe, B.A.; Affourtit, J.P.; et al. A core gut microbiome in obese and lean twins. *Nature* **2009**, *457*, 480–484. [[CrossRef](#)]
43. Yi, H.; Yong, D.; Lee, K.; Cho, Y.-J.; Chun, J. Profiling bacterial community in upper respiratory tracts. *BMC Infect. Dis.* **2014**, *14*, 583. [[CrossRef](#)]
44. Lassalle, F.; Spagnoletti, M.; Fumagalli, M.; Shaw, L.; Dyble, M.; Walker, C.; Thomas, M.G.; Bamberg Migliano, A.; Balloux, F. Oral microbiomes from hunter-gatherers and traditional farmers reveal shifts in commensal balance and pathogen load linked to diet. *Mol. Ecol.* **2018**, *27*, 182–195. [[CrossRef](#)]
45. Man, W.H.; de Steenhuijsen Pijters, W.A.A.; Bogaert, D. The microbiota of the respiratory tract: Gatekeeper to respiratory health. *Nat. Rev. Microbiol.* **2017**, *15*, 259–270. [[CrossRef](#)]
46. Fernández, A.; Huang, S.; Seston, S.; Xing, J.; Hickey, R.; Criddle, C.; Tiedje, J. How Stable Is Stable? Function versus Community Composition. *Appl. Environ. Microbiol.* **1999**, *65*, 3697–3704. [[CrossRef](#)] [[PubMed](#)]
47. Mahoney, A.K.; Yin, C.; Hulbert, S.H. Community Structure, Species Variation, and Potential Functions of Rhizosphere-Associated Bacteria of Different Winter Wheat (*Triticum aestivum*) Cultivars. *Front. Plant Sci.* **2017**, *8*, 132. [[CrossRef](#)] [[PubMed](#)]
48. Burke, J.T. Machine-Learning Based Identification of Metagenomic Source Environments. Master's Thesis, San Diego State University, San Diego, CA, USA, 2019. Available online: <https://digitallibrarydev.sdsu.edu/islandora/object/sdsu%3A27581> (accessed on 1 April 2022).
49. Cuthbertson, L.; Walker, A.W.; Oliver, A.E.; Rogers, G.B.; Rivett, D.W.; Hampton, T.H.; Ashare, A.; Elborn, J.S.; De Soyza, A.; Carroll, M.P.; et al. Lung function and microbiota diversity in cystic fibrosis. *Microbiome* **2020**, *8*, 45. [[CrossRef](#)]
50. Yatera, K.; Noguchi, S.; Mukae, H. The microbiome in the lower respiratory tract. *Respir. Investig.* **2018**, *56*, 432–439. [[CrossRef](#)]
51. Lim, Y.W.; Schmieder, R.; Haynes, M.; Furlan, M.; Matthews, T.D.; Whiteson, K.; Poole, S.J.; Hayes, C.S.; Low, D.A.; Maughan, H.; et al. Mechanistic Model of *Rothia mucilaginosa* Adaptation toward Persistence in the CF Lung, Based on a Genome Reconstructed from Metagenomic Data. *PLoS ONE* **2013**, *8*, e64285. [[CrossRef](#)]
52. Willis, J.R.; Saus, E.; Iraola-Guzmán, S.; Cabello-Yeves, E.; Ksiezopolska, E.; Cozzuto, L.; Bejarano, L.A.; Andreu-Somavilla, N.; Alloza-Trabado, M.; Blanco, A.; et al. Citizen-science based study of the oral microbiome in Cystic fibrosis and matched controls reveals major differences in diversity and abundance of bacterial and fungal species. *J. Oral Microbiol.* **2021**, *13*, 1897328. [[CrossRef](#)]
53. Falony, G.; Joossens, M.; Vieira-Silva, S.; Wang, J.; Darzi, Y.; Faust, K.; Kurilshikov, A.; Bonder, M.J.; Valles-Colomer, M.; Vandeputte, D.; et al. Population-level analysis of gut microbiome variation. *Science* **2016**, *352*, 560–564. [[CrossRef](#)]
54. Camarillo-Guerrero, L.F.; Almeida, A.; Rangel-Pineros, G.; Finn, R.D.; Lawley, T.D. Massive expansion of human gut bacteriophage diversity. *Cell* **2021**, *184*, 1098–1109.e9. [[CrossRef](#)] [[PubMed](#)]
55. Rodriguez-Brito, B.; Li, L.; Wegley, L.; Furlan, M.; Angly, E.F.; Breitbart, M.; Buchanan, J.; Desnues, C.; Dinsdale, E.; Edwards, R.; et al. Viral and microbial community dynamics in four aquatic environments. *ISME J.* **2010**, *4*, 739–751. [[CrossRef](#)] [[PubMed](#)]
56. Aziz, R.K.; Dwivedi, B.; Akhter, S.; Breitbart, M.; Edwards, R.A. Multidimensional metrics for estimating phage abundance, distribution, gene density, and sequence coverage in metagenomes. *Front. Microbiol.* **2015**, *6*, 381. [[CrossRef](#)]
57. Howard-Varona, C.; Hargreaves, K.R.; Abedon, S.T.; Sullivan, M.B. Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. *ISME J.* **2017**, *11*, 1511–1520. [[CrossRef](#)] [[PubMed](#)]
58. Suttle, C.A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Genet.* **2007**, *5*, 801–812. [[CrossRef](#)] [[PubMed](#)]
59. Romero-Calle, D.; Guimarães Benevides, R.; Góes-Neto, A.; Billington, C. Bacteriophages as Alternatives to Antibiotics in Clinical Care. *Antibiotics* **2019**, *8*, 138. [[CrossRef](#)] [[PubMed](#)]
60. Tagliaferri, T.L.; Jansen, M.; Horz, H.-P. Fighting Pathogenic Bacteria on Two Fronts: Phages and Antibiotics as Combined Strategy. *Front. Cell. Infect. Microbiol.* **2019**, *9*, 22. [[CrossRef](#)]

61. Hatful, G.F.; Dedrick, R.M.; Schooley, R.T. Phage Therapy for Antibiotic-Resistant Bacterial Infections. *Annu. Rev. Med.* **2022**, *73*, 197–211. [[CrossRef](#)]
62. Khalid, A.; Lin, R.C.Y.; Iredell, J.R. A Phage Therapy Guide for Clinicians and Basic Scientists: Background and Highlighting Applications for Developing Countries. *Front. Microbiol.* **2020**, *11*, 599906. [[CrossRef](#)]
63. Luong, T.; Salabarria, A.-C.; Edwards, R.A.; Roach, D.R. Standardized bacteriophage purification for personalized phage therapy. *Nat. Protoc.* **2020**, *15*, 2867–2890. [[CrossRef](#)]
64. Schooley, R.T.; Biswas, B.; Gill, J.J.; Hernandez-Morales, A.; Lancaster, J.; Lessor, L.; Barr, J.J.; Reed, S.L.; Rohwer, F.; Benler, S.; et al. Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant *Acinetobacter baumannii* Infection. *Antimicrob. Agents Chemother.* **2017**, *61*, e00954-17. [[CrossRef](#)]
65. 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.; et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat. Med.* **2019**, *25*, 730–733. [[CrossRef](#)]
66. Hay, I.D.; Lithgow, T. Filamentous phages: Masters of a microbial sharing economy. *EMBO Rep.* **2019**, *20*, e47247. [[CrossRef](#)]
67. Rice, S.A.; Tan, C.H.; Mikkelsen, P.J.; Kung, V.; Woo, J.; Tay, M.; Hauser, A.; McDougald, D.; Webb, J.S.; Kjelleberg, S. The biofilm life cycle and virulence of *Pseudomonas aeruginosa* are dependent on a filamentous prophage. *ISME J.* **2009**, *3*, 271–282. [[CrossRef](#)] [[PubMed](#)]
68. Clokie, M.R.; Millard, A.D.; Letarov, A.V.; Heaphy, S. Phages in nature. *Bacteriophage* **2011**, *1*, 31–45. [[CrossRef](#)] [[PubMed](#)]
69. Zinder, N.D. Lysogenization and superinfection immunity in *Salmonella*. *Virology* **1958**, *5*, 291–326. [[CrossRef](#)]
70. Mavrich, T.N.; Hatfull, G.F. Evolution of Superinfection Immunity in Cluster A *Mycobacteriophages*. *mBio* **2019**, *10*, e00971-19. [[CrossRef](#)] [[PubMed](#)]
71. Castillo, D.; Kauffman, K.; Hussain, F.; Kalatzis, P.; Rørbo, N.; Polz, M.F.; Middelboe, M. Widespread distribution of prophage-encoded virulence factors in marine *Vibrio* communities. *Sci. Rep.* **2018**, *8*, 9973. [[CrossRef](#)] [[PubMed](#)]
72. Sweere, J.M.; Van Bellegghem, J.D.; Ishak, H.; Bach, M.S.; Popescu, M.; Sunkari, V.; Kaber, G.; Manasherob, R.; Suh, G.A.; Cao, X.; et al. Bacteriophage trigger antiviral immunity and prevent clearance of bacterial infection. *Science* **2019**, *363*, aat9691. [[CrossRef](#)]
73. Bielaszewska, M.; Idelevich Evgeny, A.; Zhang, W.; Bauwens, A.; Schaumburg, F.; Mellmann, A.; Peters, G.; Karch, H. Effects of Antibiotics on Shiga Toxin 2 Production and Bacteriophage Induction by Epidemic *Escherichia coli* O104:H4 Strain. *Antimicrob. Agents Chemother.* **2012**, *56*, 3277–3282. [[CrossRef](#)]
74. von Wintersdorff, C.J.H.; Penders, J.; van Niekerk, J.M.; Mills, N.D.; Majumder, S.; van Alphen, L.B.; Savelkoul, P.H.M.; Wolfs, P.F.G. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front. Microbiol.* **2016**, *7*, 173. [[CrossRef](#)]
75. Waldor, M.K.; Mekalanos, J.J. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **1996**, *272*, 1910–1914. [[CrossRef](#)]
76. Argen, J.A.; Clark, A.G. Selfish genetic elements. *PLoS Genet.* **2018**, *14*, e1007700. [[CrossRef](#)]
77. Ptashne, M. *A Genetic Switch: Phage Lambda Revisited*; CSHL Press: Cold Spring Harbor, NY, USA, 2004.
78. Knowles, B.; Silveira, C.B.; Bailey, B.A.; Barott, K.; Cantu, V.A.; Cobián-Güemes, A.G.; Coutinho, F.H.; Dinsdale, E.A.; Felts, B.; Furby, K.A.; et al. Lytic to temperate switching of viral communities. *Nature* **2016**, *531*, 466–470. [[CrossRef](#)] [[PubMed](#)]
79. Thingstad, T.F. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol. Oceanogr.* **2000**, *45*, 1320–1328. [[CrossRef](#)]
80. Silveira, C.B.; Luque, A.; Rohwer, F. The landscape of lysogeny across microbial community density, diversity and energetics. *Environ. Microbiol.* **2021**, *23*, 4098–4111. [[CrossRef](#)] [[PubMed](#)]
81. Paterson, J.S.; Smith, R.J.; McKerral, J.C.; Dann, L.M.; Launer, E.; Goonan, P.; Kleinig, T.; Fuhrman, J.A.; Mitchell, J.G. A hydrocarbon-contaminated aquifer reveals a Piggyback-the-Persistent viral strategy. *FEMS Microbiol. Ecol.* **2019**, *95*, fiz116. [[CrossRef](#)]
82. Brum, J.R.; Hurwitz, B.L.; Schofield, O.; Ducklow, H.W.; Sullivan, M.B. Seasonal time bombs: Dominant temperate viruses affect Southern Ocean microbial dynamics. *ISME J.* **2016**, *10*, 437–449. [[CrossRef](#)]
83. Silveira, C.B.; Rohwer, F.L. Piggyback-the-Winner in host-associated microbial communities. *NPJ Biofilms Microbiomes* **2016**, *2*, 16010. [[CrossRef](#)]
84. Cheng, H.H.; Muhlrad, P.J.; Hoyt, M.A.; Echols, H. Cleavage of the cII protein of phage lambda by purified HflA protease: Control of the switch between lysis and lysogeny. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7882–7886. [[CrossRef](#)]
85. Kim, M.S.; Bae, J.W. Lysogeny is prevalent and widely distributed in the murine gut microbiota. *ISME J.* **2018**, *12*, 1127–1141. [[CrossRef](#)]
86. Cazares, D.; Cazares, A.; Figueroa, W.; Guarneros, G.; Edwards Robert, A.; Vinuesa, P.; Bordenstein, S.; Svircev, A. A Novel Group of Promiscuous Podophages Infecting Diverse Gammaproteobacteria from River Communities Exhibits Dynamic Intergenous Host Adaptation. *mSystems* **2021**, *6*, e00773-20. [[CrossRef](#)]
87. Dutilh, B.E.; Cassman, N.; McNair, K.; Sanchez, S.E.; Silva, G.G.Z.; Boling, L.; Barr, J.; Speth, D.; Seguritan, V.; Aziz, R.; et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* **2014**, *5*, 4498. [[CrossRef](#)]

88. Edwards, R.A.; Vega, A.A.; Norman, H.M.; Ohaeri, M.; Levi, K.; Dinsdale, E.A.; Cinek, O.; Aziz, R.K.; McNair, K.; Barr, J.J.; et al. Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat. Microbiol.* **2019**, *4*, 1727–1736. [[CrossRef](#)] [[PubMed](#)]
89. Oh, J.-H.; Lin, X.B.; Zhang, S.; Tollenaar Stephanie, L.; Özçam, M.; Dunphy, C.; Walter, J.; van Pijkeren, J.-P.; Kivisaar, M. Prophages in *Lactobacillus reuteri* Are Associated with Fitness Trade-Offs but Can Increase Competitiveness in the Gut Ecosystem. *Appl. Environ. Microbiol.* **2019**, *86*, e01922-19. [[CrossRef](#)] [[PubMed](#)]
90. Motlagh, A.M.; Bhattacharjee, A.S.; Coutinho, F.H.; Dutilh, B.E.; Casjens, S.R.; Goel, R.K. Insights of Phage-Host Interaction in Hypersaline Ecosystem through Metagenomics Analyses. *Front. Microbiol.* **2017**, *8*, 352. [[CrossRef](#)] [[PubMed](#)]
91. McDaniel, L.; Breitbart, M.; Mobberley, J.; Long, A.; Haynes, M.; Rohwer, F.; Paul, J.H. Metagenomic Analysis of Lysogeny in Tampa Bay: Implications for Prophage Gene Expression. *PLoS ONE* **2008**, *3*, e3263. [[CrossRef](#)]
92. McDaniel, L.D.; Rosario, K.; Breitbart, M.; Paul, J.H. Comparative metagenomics: Natural populations of induced prophages demonstrate highly unique, lower diversity viral sequences. *Environ. Microbiol.* **2014**, *16*, 570–585. [[CrossRef](#)]
93. Parmar, K.; Dafale, N.; Pal, R.; Tikariha, H.; Purohit, H. An Insight into Phage Diversity at Environmental Habitats using Comparative Metagenomics Approach. *Curr. Microbiol.* **2018**, *75*, 132–141. [[CrossRef](#)]
94. Stokar-Avihail, A.; Tal, N.; Erez, Z.; Lopatina, A.; Sorek, R. Widespread Utilization of Peptide Communication in Phages Infecting Soil and Pathogenic Bacteria. *Cell Host Microbe* **2019**, *25*, 746–755.e5. [[CrossRef](#)]
95. Gómez, P.; Bennie, J.; Gaston, K.J.; Buckling, A. The Impact of Resource Availability on Bacterial Resistance to Phages in Soil. *PLoS ONE* **2015**, *10*, e0123752. [[CrossRef](#)]
96. Braga, L.P.; Spor, A.; Kot, W.; Breuil, M.-C.; Hansen, L.H.; Setubal, J.C.; Philippot, L. Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. *Microbiome* **2020**, *8*, 52. [[CrossRef](#)]
97. Rezaei Javan, R.; Ramos-Sevillano, E.; Akter, A.; Brown, J.; Brueggemann, A.B. Prophages and satellite prophages are widespread in *Streptococcus* and may play a role in pneumococcal pathogenesis. *Nat. Commun.* **2019**, *10*, 4852. [[CrossRef](#)]
98. Roach, M.; McNair, K.; Giles, S.; Inglis, L.; Pargin, E.; Decewicz, P.; Edwards, R. Philympics 2021: Prophage Predictions Perplex Programs. *bioRxiv* **2021**. [[CrossRef](#)]
99. Akhter, S.; Aziz, R.K.; Edwards, R.A. PhiSpy: A novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.* **2012**, *40*, e126. [[CrossRef](#)] [[PubMed](#)]