ANNUAL REVIEWS

Annual Review of Virology

Integrating Viral Metagenomics into an Ecological Framework

Pacifica Sommers,[1,*] Anushila Chatterjee,[2,*] Arvind Varsani,[3,4] and Gareth Trubl[5]

[1]Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, Boulder, Colorado 80309, USA

[2]Department of Immunology and Microbiology, University of Colorado School of Medicine, Aurora, Colorado 80045, USA

[3]The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, Arizona 85287, USA; email: arvind.varsani@asu.edu

[4]Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Observatory 7925, South Africa

[5]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

Annu. Rev. Virol. 2021. 8:133–58

First published as a Review in Advance on May 25, 2021

The *Annual Review of Virology* is online at virology.annualreviews.org

https://doi.org/10.1146/annurev-virology-010421-053015

*These authors contributed equally to this article

Keywords

metagenomics, viromics, ecology, virology, high-throughput sequencing

Abstract

Viral metagenomics has expanded our knowledge of the ecology of uncultured viruses, within both environmental (e.g., terrestrial and aquatic) and host-associated (e.g., plants and animals, including humans) contexts. Here, we emphasize the implementation of an ecological framework in viral metagenomic studies to address questions in virology rarely considered ecological, which can change our perception of viruses and how they interact with their surroundings. An ecological framework explicitly considers diverse variants of viruses in populations that make up communities of interacting viruses, with ecosystem-level effects. It provides a structure for the study of the diversity, distributions, dynamics, and interactions of viruses with one another, hosts, and the ecosystem, including interactions with abiotic factors. An ecological framework in viral metagenomics stands poised to broadly expand our knowledge in basic and applied virology. We highlight specific fundamental research needs to capitalize on its potential and advance the field.

1. INTRODUCTION

Viral metagenomics has revolutionized the field of virology by providing culture-independent methods to detect and characterize the vast diversity of viruses that cannot be cultured and isolated in a high-throughput manner (1–4). Metagenomics is the study of nucleic acid sequences from a collection of entities in a sample, with viral metagenomics focusing on the nucleic acids of viruses within the sample. These high-throughput approaches have advanced our understanding of viral ecology by providing insights into the diversity, abundance, and functional potential of viruses associated with different environments (5, 6). Moreover, applying an ecological framework to the use of viral metagenomics in areas of virology traditionally considered nonecological holds enormous promise for advancing virology and understanding the global roles viruses play. In this review, we (*a*) describe the importance of an ecological framework for viral metagenomics; (*b*) review its implications for a suite of specific tools and approaches in viral metagenomics, including their limitations; (*c*) highlight the often-overlooked interactions of viruses with other viruses; and (*d*) address applications and research needs of viral metagenomics within an ecological framework.

Metagenomics: sequence-dependent and cultivation-independent characterization of genetic material isolated from a sample

Viral ecology: the study of interactions of viruses with their abiotic and biotic environments, including with their hosts, and with other viruses

Ecological framework: the diversity, distributions, dynamics, and interactions of biological entities within ecosystems containing communities comprising diverse populations of variable individuals

Ecology: the study of interactions of organisms with their abiotic and biotic environment

Virome: a metagenome that requires additional sample processing steps (e.g., size fractionation) to remove microbes and concentrate virus-like particles before sequencing

Bacteriophage (phage): virus that infects a bacterium

2. THE IMPORTANCE OF AN ECOLOGICAL FRAMEWORK FOR VIRAL METAGENOMICS

Ecology is the study of how organisms interact with their abiotic and biotic environment, and it seeks to explain the diversity, distributions, and dynamics of biological systems (**Figure 1**). Viral metagenomics is revealing the vast diversity and effect of viruses in ecosystems such as soils, oceans, and even glaciers (6–11). Within these ecosystems, the effect of viruses on one another and on the diversity and abundance of their hosts and the hosts' competitors, as well as the viral effects on hosts' metabolic functions, can now be studied beyond the limited virus-host pairs that can be cultured (7, 12, 13). The indirect effects of viral infections on biogeochemical cycling (14), the effect on hosts of viruses' mere presence even without infection (15), and the cascading effects of viral infection on the entire community their hosts interact with (16) are beginning to be explored and promise to advance virology in a range of settings, from environmental to medical. The range of hosts that viruses can infect and the virus-virus interactions (VVIs) enhancing or preventing coinfection of the same host are also being explored through viral metagenomics (17, 18). Monitoring viruses present in ecosystems with the potential to expand their host range to infect humans or domesticated animals, known as spillover, has already built links from environmental to medical virology (19, 20).

Advances in metagenomic approaches have been instrumental in the discovery of diverse viruses of host-associated microbial communities (21–24). A pioneering discovery in this context is the identification of cross-assembly phage (commonly known as crAssphage) and its relatives that comprise up to 90% of the human gut virome reads and co-occur with the *Bacteroidetes* phylum found in the gut (23, 25). The diversity and dynamics of viral community composition have been associated with outcomes such as the ability of opportunistic bacterial pathogens, such as *Clostridium difficile* and vancomycin-resistant *Enterococcus*, to establish in the gut (26, 27), acquired immunodeficiency syndrome (AIDS)-associated disease manifestation and progression (28), and the severity of respiratory tract infections (29, 30). These effects could be the result of multiple types of ecological interactions between bacteriophages (phage), which are predominant members of the commensal microbiota (31–33). For example, they could result from direct parasitism, from lysing their hosts and thereby altering community composition, or through modulating genetic diversity and functionality of native bacterial communities through horizontal gene transfer (reviewed in 34). Metagenomic approaches can be applied to better characterize the mechanistic

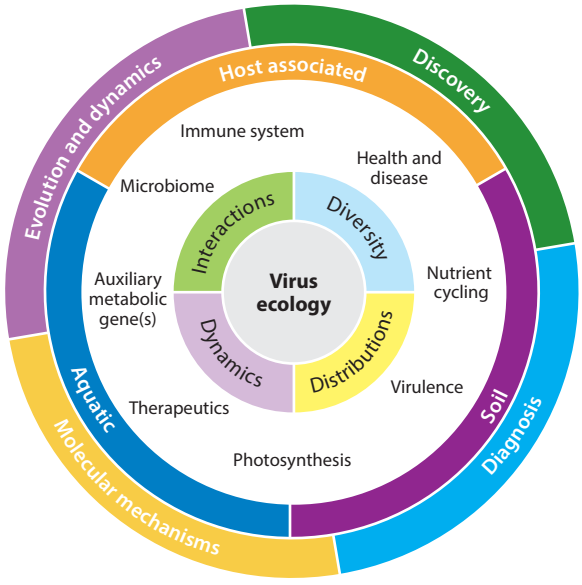

Figure 1

The study of viruses within ecological frameworks via metagenomic approaches can provide unique insights into the ecology of viruses, specifically their diversity, distribution, and dynamics, as well as their interactions with their hosts, other viruses, and the abiotic and biotic environment. The interactions of viruses with their hosts are often studied within specific contexts ranging from health and disease to biogeochemical processes (examples are in the *white layer*). These can all take place in different types of ecosystems, including soil, aquatic, and within plants or animals that create an ecosystem for their associated microbial communities. The pursuit of these ecological interactions across all ecosystems leads to advances in discovery and diagnosis of viruses and understanding of their molecular mechanisms and their resulting evolutionary dynamics.

ecological interactions leading to these outcomes and to take advantage of them in therapeutic contexts.

An ecological framework recognizing the importance of localized habitats within the larger host and between individual hosts is critical for recognizing subtleties of context-dependent interactions. Because different anatomic sites within the human body create distinct habitats and consequently are associated with distinct viral populations (35), interactions between viruses and mammalian components within each habitat likely have unique effects. Additionally, factors such as diet (32), gender (36), and age (37) influence the types and proportions of phages associated with a particular habitat or ecological niche. Crosstalk between native viruses and the immune system of mammalian hosts affects immune development and modulation (38). The dynamics and diversity of specific viruses can provide important benefits, such as therapeutically administered phages to treat multidrug-resistant bacteria (39).

Viral metagenomics is also advancing the understanding of viral molecular mechanisms. Phage-bacteria interaction studies can provide detailed insight into how phages influence the structure, function, and evolution of microbial communities (40); act as agents of lateral gene transfer (41); and can be exploited in clinical (42, 43) and industrial settings (44). Viral metagenomics, either independently or in combination with other high-throughput technologies, such as single-cell genomics or metaproteomics, have enabled more comprehensive cataloging of virus-host associations in the context of diverse microbial communities (6, 18). These studies serve as primers to

design more robust methods to explore the diversity of infection mechanisms adopted by viruses and assess how these newly discovered virus-host interactions can affect cellular life.

Virus-like particle (VLP): particle of virus size (less than 0.45 μm) with nucleic acid but identity is unknown as defined in viral ecology

Viral metagenomics of various ecosystems is helping to fill the viral sequence space in public databases (1–3), which builds a foundation of characterized viral diversity and evolutionary relationships to improve viral diagnosis, understand pathogenesis, and identify therapeutic and prophylactic interventions. Community standards for the minimum information needed about an uncultivated virus-derived genome sequence further increase the power of these databases by providing ecological context (45). Determination of near-complete or complete genomes enables a broader understanding of viral evolution (e.g., 46, 47) and allows classification of these viruses in a rather dynamic taxonomic framework (47, 48). This foundation of knowledge about viral diversity and evolutionary relationships has helped to inform the identification of and response to novel human pathogens, such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (49).

3. IMPLICATIONS OF AN ECOLOGICAL FRAMEWORK FOR THE TOOLS AND APPROACHES OF VIRAL METAGENOMICS

An ecological framework for interpreting viral metagenomic data conceptualizes multiple populations of diverse viruses interacting with variable ranges of potential host diversity and accounts for interactions among viruses via infected hosts in an environment and with ecosystem-level consequences (**Figure 2**). The diversity, distributions, and dynamics of viruses and their hosts can complicate the interpretations of metagenomic data. Viral metagenomics requires specific bioinformatic approaches (as reviewed in 50) because viral genomes pose unique challenges in metagenomics. The array of tools and approaches for viral metagenomics was recently reviewed elsewhere (50, 51), but we summarize key advantages and limitations of cutting-edge techniques specifically within an ecological framework. We broadly divide these tools into two groups: (*a*) approaches for addressing viral diversity and distributions, primarily focused on identification and detection of viruses, and (*b*) approaches for addressing viral dynamics and interactions, which focus on specifying hosts and quantifying the abundance and activity of viruses and their hosts.

3.1. Metagenomic Approaches to Viral Diversity and Their Distributions

The study of viral diversity and virus distribution requires assembly of viral genomes and detection of their presence. Identification through metagenomics requires assembling genomes, while detection of specific viral genomes can be accomplished by using marker genes for certain lineages [e.g., RNA-dependent RNA polymerase (52)] or viral hallmark genes and viral motifs (3, 53, 54). Genomes can be assembled using multiple approaches, including either short-read or long-read sequencing of extracted double-stranded (ds) DNA, single-stranded (ss) DNA, or DNA reverse transcribed from RNA (complementary DNA). These nucleic acids can be obtained directly from the sample or extracted from virus-like particles (VLPs) (see definition in 55) obtained through filtration or centrifugation to isolate viruses. Each approach has various advantages and limitations for different goals in the detection and identification of viruses.

3.1.1. Considerations for whole sample nucleic acid or virus-like particle–derived nucleic acid sequencing. When it is time to decide on an approach for identifying the diversity and distribution of viruses in an environment, the first consideration is whether to focus on the whole sample or only the VLPs. The entire sample will contain far more nucleic acids from the bacterial, archaeal, and eukaryotic cells than viral nucleic acids, making it difficult to assemble and detect

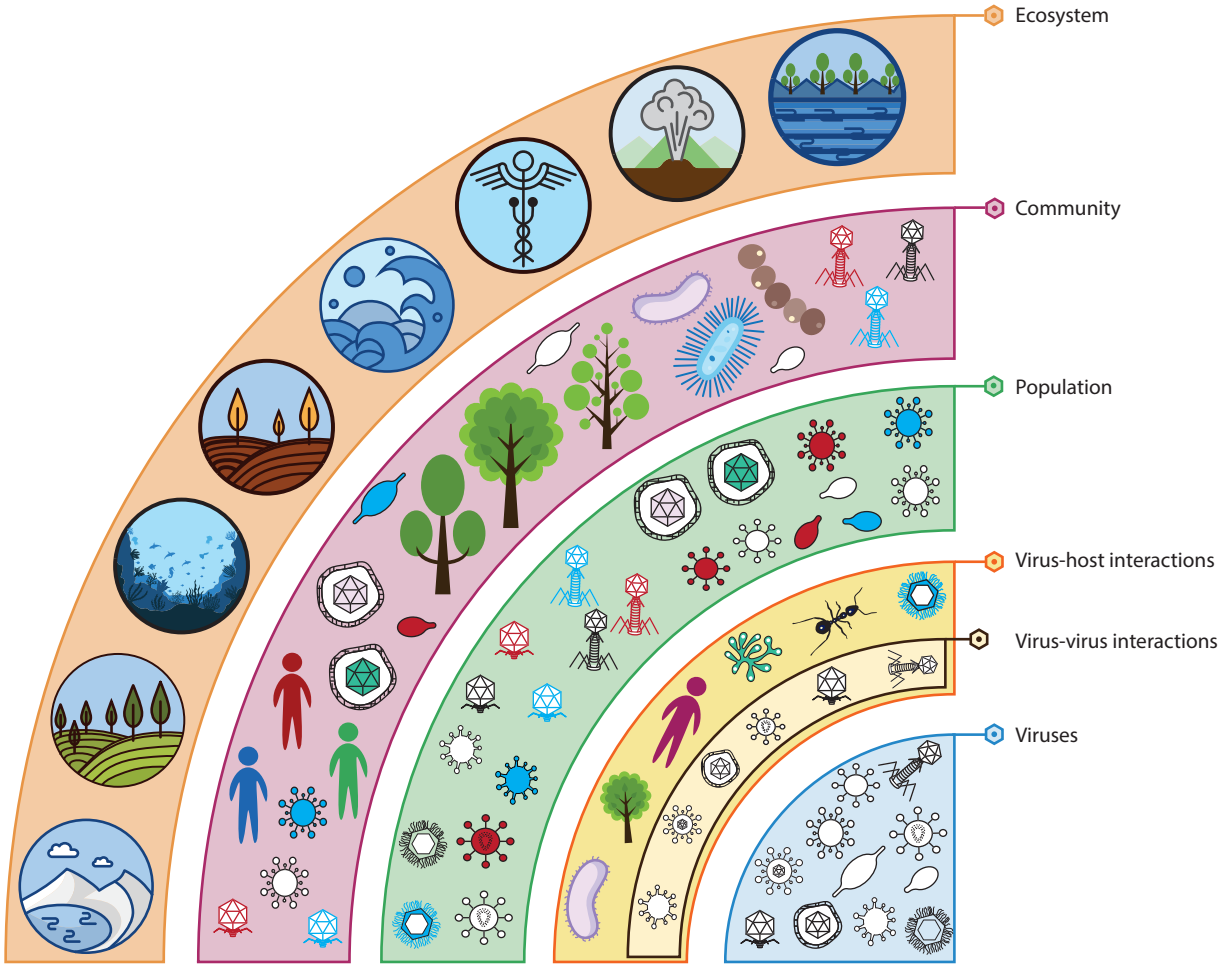

Figure 2

Viruses interact at various levels and thus can be studied from individual entities to larger-scale populations and communities through viral metagenomics. Hence, viral metagenomics is facilitating the study of viruses within an ecological framework, including studying viral diversity and distributions across ecosystems, and viral dynamics across variable ranges of potential host diversity and in a multitude of environments where they can have positive or negative effects.

rarer viral genomes (51, 56, 57). Therefore, to obtain these viral genomes, VLPs often need to be separated and concentrated before sequencing.

Methods for separating and concentrating VLPs vary depending on the type of sample (58), with approaches having been optimized for seawater (4), soil (59, 60), fecal matter (61, 62), and plant (63) or animal tissues (64). These methods often rely on size-dependent filtration or centrifugation. Often samples will first be centrifuged or filtered through a 0.45-μm filter to remove large particles and eukaryotic cells, followed by a 0.2-μm filter to remove most cells, before a final VLP purification step, such as a cesium chloride (CsCl) density gradient, or DNase/RNase, OptiPrep™, or sucrose density gradient centrifugation to cleave nonencapsidated nucleic acid (58, 65). It is important to note that some of these methods may affect the virions (58). Finally, VLPs are concentrated through filtration (e.g., tangential flow filtration), ultracentrifugation, or

precipitation (e.g., using polyethylene glycol). VLPs can be collected directly on a 0.02-μm filter (66) or by first using ferrous chloride to flocculate viral particles and collecting them on a larger pore filter (67). The range of viral sizes of interest and the sample type will determine which protocols will yield the most viral genomes.

Viral single amplified genome (vSAG): sequencing of virus genome following single viral particle sorting and multiple displacement amplification

Metatranscriptomics: uses RNA sequencing technology to capture gene expression profiles or transcripts from a sample

Alternatively, a single VLP can be separated (68) for genome amplification and sequencing [viral single amplified genome (vSAG)] (69, 70). In this method, VLPs are sorted into droplets or agarose beads and applied to multiwell plates using flow cytometry. Extracted DNA is amplified, most often with multiple displacement amplification, and sequenced; for RNA viruses, a transcriptomic approach can be used. The advantages of vSAGs lie in assembly and assignment of contigs to the correct viral genome, especially in the case of multicomponent, segmented, and satellite viruses, which can be difficult to distinguish bioinformatically from chimeras as sequencing and assembly artifacts in the scrambled puzzle of viral community short reads (71, 72). vSAGs also permit assembly of genomes with many variants present in the environment that otherwise challenge assembly algorithms and can be overlooked, especially in the case of highly abundant and cosmopolitan viruses with many variants (73). The disadvantages lie in the additional cumbersome steps and expense in sequencing each genome instead of thousands of viral operational taxonomic units (vOTUs) in one single sample. However, de novo assembled viral genomes obtained from vSAGs can be mapped against metagenomic sequence data to determine the relative abundance and ubiquity of these viruses across multiple samples or environments (74, 75).

Although sequencing the nucleic acid from VLPs (as opposed to the whole sample) has the advantage of obtaining greater coverage of viral genomes and detecting less abundant viral genomes, it can miss certain viral genomes, including viruses that are nonencapsidated in hosts, such as retroviruses and proviruses integrated into the hosts' genomes or persisting as replicons in their hosts. Additionally, any viruses in the process of infecting a host when the sample is run through a filter would be missed because the host they were adsorbed to or inside of would be filtered out. Metagenomics or metatranscriptomics of the entire sample can help address these weaknesses, especially when used in combination with sequencing of the VLPs. Which approach is taken depends a great deal on the question being asked, whether the diversity and identity of the entire viral community is of interest, or whether only a specific subset of viruses or viral states is important.

3.1.2. Considerations for which nucleic acid(s) to sequence. The second consideration, after whether to sequence nucleic acids from purified VLPs or a whole sample, is which type(s) of nucleic acid to extract and amplify for sequencing. On one end of the spectrum, eukaryote-infecting viruses appear to be dominated by RNA viruses, particularly in the case of plant and fungi-infecting viruses (see 76); on the other end of the spectrum, viruses of bacteria and archaea are dominated by dsDNA viruses (see 47). These divisions have led to a schism in which viruses are detected and characterized. To a newcomer interested in characterizing viruses in their environment of interest, the dominant hosts would direct whether they optimize protocols for detecting dsDNA or ssDNA or RNA, further biasing our understanding of the virosphere. Below, we describe how some of these biases came about and how we could overcome them to gain a more integrated understanding of the full ecological spectrum of viral communities (**Figure 3**).

Bacteria and archaea make up the vast majority of living organisms on Earth, and as such their viruses make up the largest proportion of the virosphere. To date, phages with dsDNA genomes are over-represented in the literature because most viral databases primarily consist of cultured viruses, which for certain environments are largely represented by medically relevant dsDNA phages, and most metagenomic preparation protocols exclusively target dsDNA (60). These challenges preclude our ability to detect other viruses if their genes are unknown. Recently, standards were put forth to improve the reporting of uncultivated viral genomes in public databases (45),

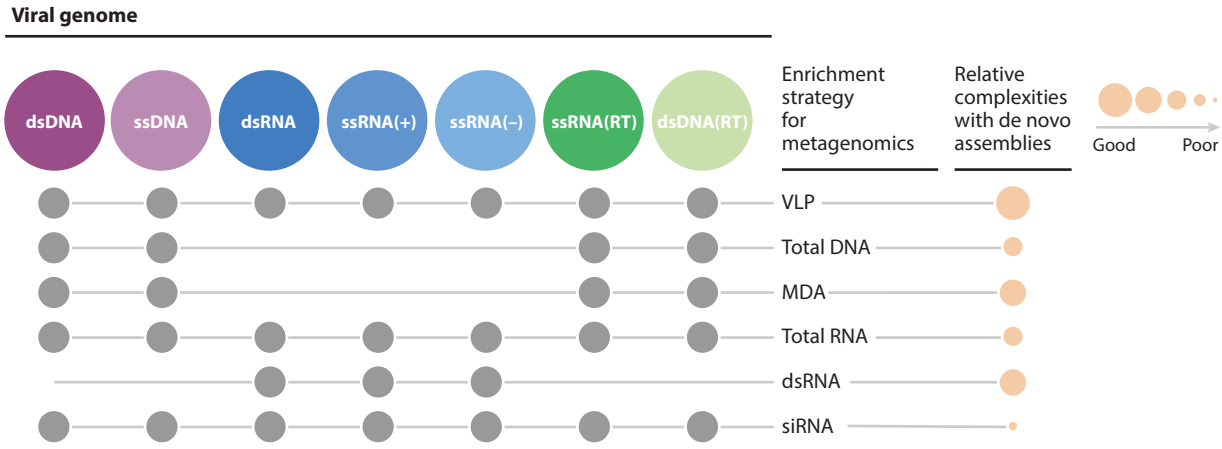

Figure 3

A brief summary of the various approaches that can be used for enrichment of viral nucleic acid for metagenomic studies based on the genome type of the viruses, i.e., double-stranded (ds) DNA, single-stranded (ss) DNA, dsRNA, ssRNA positive sense (+), ssRNA negative sense (−), ssRNA reverse-transcribing (RT), and dsDNA (RT). It is important to note that even though virus-like particle (VLP)-based approaches enable the targeting of all viral genome types, there exist various viral-like mobile elements (including satellites) and viruses [such as mitoviruses (198)] that do not encode capsid proteins. Furthermore, retro-transcribing viruses and temperate phages that have integrated into their host genomes or plasmids may not be detectable using the VLP-based approach. Multiple displacement amplification (MDA) is commonly used to increase the concentration of nucleic acid and preferentially amplifies circular molecules, be they single stranded or double stranded. Hence, MDA has been used extensively for identification and characterization of complete genomes of circular DNA viruses. All these approaches have their own issues when it comes to the de novo assemblies, and a crude scale of the relative complexity of assembly is provided. Here, good implies relatively straightforward and feasible methods are available, while poor reflects viral types for which significant complexities challenge de novo assembly. For example, small interfering RNA (siRNA)-based de novo assemblies that have been used for known viruses may not work well for novel viruses, as no comparative scaffolds are available and siRNAs are not produced uniformly across the genome. With the total RNA approach, one needs to be mindful of spliced transcripts. Another important consideration is the library preparation kits, and this is particularly important when it comes to ssDNA viruses extracted from VLPs, as the majority of kits rely on dsDNA in the library preparation workflow. To overcome this issue, one would need to amplify DNA using MDA or use a library kit that can quantitatively amplify both ssDNA and dsDNA (e.g., Accel-NGS 1S Plus DNA Library Kit, Swift Biosciences).

promoting transparency and a better assessment of methods bias (77), allowing for better identification of key variables across data sets.

Many viruses infecting animals, plants, and fungi have RNA genomes and thus are not detected by metagenomic methods. During replication in their corresponding hosts, RNA viruses generate dsRNA intermediates that can be purified based on their binding properties to silica (78). Additionally, antibodies have been identified and experimentally used for detection and purification of dsRNA (79). Metatranscriptomic approaches that target total RNA have also been successful at detecting RNA viruses, although ribosomal (r) RNA, transfer RNA, and the messenger (m) RNA of the host can dominate the raw sequence data, limiting the resolution of RNA viruses (52). Host-specific ribodepletion is commonly implemented during library preparation to address this issue by removing host rRNA. However, the ribodepletion products are available only for certain organisms. As a final approach to detecting RNA viruses, viral genomes have been assembled via sequencing of purified small interfering RNA 21–24 nucleotides (80, 81), commonly produced as a defense to pathogen infection in plants and insects, although in most cases viral genomes as scaffolds have been used.

High-throughput sequencing: the collection of constantly evolving nucleic acid sequencing approaches that yields a large volume of sequence data (including short- and long-read sequencing)

Single amplified genome (SAG): genome sequence derived from the amplification of the genome of a single cell; generally, the cells are sorted by flow cytometry or microfluidic approaches

3.1.3. Considerations for short- or long-read sequencing. Finally, the length of reads obtained through metagenomic sequencing has advantages and disadvantages for addressing questions about the diversity and distribution of viruses. High-throughput sequencing of short reads (100–250 bp) is the most common and cheapest sequencing platform and has the most optimized computational tools to detect and characterize viral genomes. Over the past 5 years, there has been tremendous effort to improve long-read sequencing for both quality and throughput, with some platforms yielding more than 1,500-kb-length reads (see 82). The advantages of long reads include (*a*) picking up taxa that are otherwise missed by short reads, which are sometimes rarer vOTUs but can also be highly abundant ones; (*b*) identifying hypervariable regions in genomes; (*c*) identifying recombinants within a population; and (*d*) enabling better assemblies of complete viral genomes (83).

3.2. Metagenomic Approaches to Viral Dynamics and Interactions

Metagenomic approaches can be used to study dynamics of viral and host populations temporally and spatially but can also be used to hypothesize interactions from genetic signatures in a single snapshot sample. However, viral metagenomics provides unique opportunities and limitations in identifying the interactions and dynamics of viruses.

One key viral interaction that can be determined from metagenomic analysis of a single time point sample is host prediction. This can be done bioinformatically [e.g., matching a viral genome with CRISPR sequences in a host's genome (6, 84)] or by sequencing the amplified genomic material of infected single cells. Host single amplified genomes (SAGs) also have the opportunity to reveal coinfection by multiple viruses, another key viral interaction, and provide a powerful opportunity for identifying the host(s) of viruses, especially when the cells are sorted by some method of discerning infection status.

Although SAG sequencing can reveal the interaction partners, the dynamics of those interactions require knowledge about activity rates and changes in absolute abundance. Metagenomic data from a single time point, and particularly the metagenomic analysis of nucleic acids that have been amplified, cannot be used as a reliable representation of absolute abundance in the sample and cannot distinguish activity of hosts or replication of viruses. Furthermore, with improvements in single-cell and single-virus genomic approaches, we are certain to see more identification of virus-host interactions (see review in 75).

Although replicative and transcriptional activities of viruses cannot be determined from a metagenomic sample alone, metagenomics can be paired with various methods to investigate viral dynamics (summarized in **Table 1**). One approach is to have a paired metatranscriptome from the same sample. This approach has the benefit of characterizing both DNA and RNA viruses and gene expression levels from detected viral genomes. While methods for obtaining metatranscriptomes have come a long way, there are still some issues (85), especially with using gene expression profiles to infer viral activity. Gene expression does not always lead to translation; some genes are highly conserved or undergo horizontal gene transfer so frequently that they cannot be assigned to a particular virus (86), and there are many host mechanisms to stop viral infection after gene expression [e.g., abortive infection (87)]. Additionally, variables that affect a virus's infection efficiency, including the virus's burst size and the amount of redirection of the host's metabolism that is required, can alter gene expression profiles (88, 89).

Stable isotope probing (SIP) combined with metagenomics has become a popular technique to characterize specific microbes (or link individuals to specific functions) and the viruses that infect them (90–93) or to label organisms and viruses active in a sample (94). This technique involves incubating samples with isotopically labeled substrate(s), separating labeled and unlabeled nucleic

Table 1 Tools available to pair with metagenomics to overcome limitations in metagenomic methods

Application	Benefit(s)	Challenge(s)/limitation(s)	Reference(s)
SIP	Substrate-specific viral influence; activity; increased resolution; virus-host linking	Costly (enriched substrate and sequencing of multiple fractions); ~2 μg of nucleic acid required	93, 94, 199, 200
BONCAT	In situ virion production; burst size; activity; virus-host linking	No detection of viruses undergoing lysogenic infection; often paired with fluorescence, which cannot distinguish between a virus and other virus-sized entities; broad protein labeling	102, 201
NanoSIMS/NanoSIP	Morphological analysis; stoichiometry; quantification	Hard to characterize smaller or tailed viruses (e.g., phage)	102–104
Paired metagenome-metatranscriptome	DNA and RNA viruses; gene expression levels	rRNA consumes RNA signal; mRNA production does not always result in translation to proteins	7
Single-cell polony	Virus-host pairing	Virus sequence-specific probes required	105–107
PhageFISH	Intracellular phage infection dynamics; detection of free phages; quantification relative per cell phage DNA copy number	Virus- and host sequence-specific DNA probes required; absolute quantification	108, 109
Virocell-FISH	Both host and virus activity	Virus- and host-specific mRNA probes required; activity is inferred from gene expression	110
Single-cell viral tagging	Virus-host pairing; isolates and plaque formation not required; can capture a diversity of viruses infecting a host in a mixed community sample	Targeted viruses need to be fluorescently labeled; sufficient labeling is needed to overcome noise associated with sorting; cannot be used to describe infection dynamics	111, 112
Hi-C	Host predictions; may be exploited to predict virus coinfections	Misses lytic phages that rapidly destroy host cells; viral genome must interact with the host chromosome	113, 114
Single-cell Hi-C	Virus-host genome interactions (epigenetics)	Epigenetic signature dependent on the state of viral replication; viral genome must interact with the host chromosome	116, 117

Abbreviations: BONCAT, bioorthogonal noncanonical amino acid tagging; FISH, fluorescent in situ hybridization; Hi-C, high-throughput chromosome conformation capture; mRNA, messenger RNA; rRNA, ribosomal RNA; SIMS, secondary ion mass spectroscopy; SIP, stable isotope probing.

acid via CsCl-density centrifugation, and sequencing the different fractions. Heavy-water ($H_2{}^{18}O$) SIP-metagenomics has recently gained a lot of attention because water is a universal substrate and incorporation of ^{18}O into newly synthesized DNA allows taxon-specific microbial growth and mortality rates to be calculated (95). Downsides to this method include the expensive cost of isotopically labeled substrate, the amount needed to ensure nucleic acid is sufficiently labeled, a high amount of nucleic acid required for input, and the cost and effort associated with sequencing multiple fractions as well as an unlabeled control sample to provide enough resolution for ecological inference (96).

Temperate bacteriophage: a virus that can follow either a productive (lytic) or a quiescent (lysogenic) infection; in the lysogenic state, the viral nucleic acid can be integrated into the host DNA or in the case of *Escherichia* virus P1 maintained as a replicon in the cytoplasm

Another strategy is to combine metagenomics with bioorthogonal noncanonical amino acid tagging (BONCAT). Like with SIP, samples are incubated with a substrate, but in this case, it is a synthetic amino acid that is amenable to azide-alkyne click chemistry (97). The two most widely used synthetic amino acids are azidohomoalanine and homopropargylglycine, which both replace methionine during translation (98) and can permit the attachment of another substrate, such as a fluorophore, enabling time-resolved analysis of protein synthesis with epifluorescence microscopy (99) of single cells or in complex communities. A benefit of this method over SIP is the capability to quantify virion production; however, major drawbacks include missing temperate bacteriophage in the lysogenic cycle due to a virion not being produced and the broad labeling of all proteins in a sample. To overcome these challenges, a different substrate could be used for the click chemistry (100), or it can be combined with SIP and nanoscale secondary ion mass spectroscopy (NanoSIMS).

NanoSIP is a powerful technique that brings together NanoSIMS with SIP and could be used to identify and quantify metabolic activities and track their source and fate (101). These methods have been applied to viruses to directly quantify virus-to-host ratios, track virus activity and biogeochemical influence, and illuminate indirect viral effects (such as organomineralization) and they can also be done in tandem with BONCAT to quantify virion production (102–104).

In contrast to the previously described methods, the single-cell polony method allows for direct quantification of viral infection (105–107). This method pairs flow-cytometric sorting and PCR-based polony technique to simultaneously screen thousands of taxonomically resolved individual cells for intracellular viral DNA, enabling sensitive, high-throughput, and direct quantification of infection by viruses. Given that this is a probe-based method, viral lineages that share common sequence regions can be detected with the same probe set. The single-cell polony method thus relies on some prior knowledge of the viral sequence and does not detect RNA viruses.

Other probe-based techniques include variations of fluorescent in situ hybridization (FISH), phageFISH and Virocell-FISH, which both use two different probes tagged with fluorophores to study the virus and the host infection dynamics. PhageFISH enables the detection of replicating and encapsidated phage DNA using dsDNA probes, and at the same time the host can be identified and quantified based on rRNA probes. This approach allows for single-cell measurement (108, 109). Virocell-FISH, however, is based on fluorescently labeled mRNA of a gene of the host and virus, and thus gene expression dynamics are studied using a high-throughput imaging flow cytometer. This approach allows for the study of viral infection dynamics via distinct transcriptional states at a single-virocell level (110).

Viral tagging is an approach that links fluorescently labeled viruses to specific host cells via high-throughput flow sorting and characterizes host-associated viral diversity with high-throughput sequencing (111, 112). This method has the added benefit of physically linking viruses to a host, like culture-dependent methods, but it is not limited to isolates or viruses that can form plaques. This approach cannot be used to characterize infection dynamics, as hosts can become labeled during infection and remain labeled even if infection is terminated or virion production and host lysis do not occur.

Finally, high-throughput chromosome conformation capture (Hi-C) is another promising approach that has been recently applied to gain insights into virus-host interactions in the gastrointestinal tract of mice and humans (113, 114). The Hi-C method combines proximity-based fixation and high-throughput sequencing to capture the 3D architecture of chromosomes and uses conformational signatures to estimate close physical proximity between DNA fragments (115). Although this approach enables characterization of prophages and slow-growing lytic phages, it potentially leaves out highly virulent phages that rapidly kill their hosts. Therefore, sequencing of the VLPs is still required to generate an exhaustive inventory of viral genomes present in the sample.

Single-cell Hi-C allows for the chromosome-viral genome interaction studies at a single-cell level and thus can be used to study epigenetics (116, 117).

The full potential of these techniques remains underutilized and untapped. Further research developing these techniques should include diverse environmental samples, experiments incorporating multiple inducing agents to track microbial lysis and new virion production from proviruses, and more exploration of combining these techniques and long-read sequencing.

Sociovirology: the study of ecological and evolutionary aspects of virus-virus interactions occurring both within and among hosts

4. VIRUS-VIRUS INTERACTIONS

Viral metagenomics is helping to bring the social lives of viruses to light. Social interactions between viruses can alter genetic architecture, phenotypic characteristics, reproduction strategy, infection dynamics, and evolution of viruses, which influence how viruses interact with their hosts and ultimately the ecosystems. Although, the interactions of viruses with other viruses can have important consequences for viral ecology, this key interaction type is often overlooked within an ecological framework. Here, we highlight the role of VVIs, including the new field of sociovirology, in driving the diversity, distribution, and dynamics of viruses, their hosts, and their broader ecosystems.

VVI can be defined as the changes in the infectivity, lifestyle, development, and persistence during concurrent or sequential infection of the host by two or more viruses (118) (**Figure 4**). Interactions between coinfecting viruses were observed as early as the 1940s; however, the landmark study that put VVI on the map was performed by Turner and Chao in the late 1990s (119). The

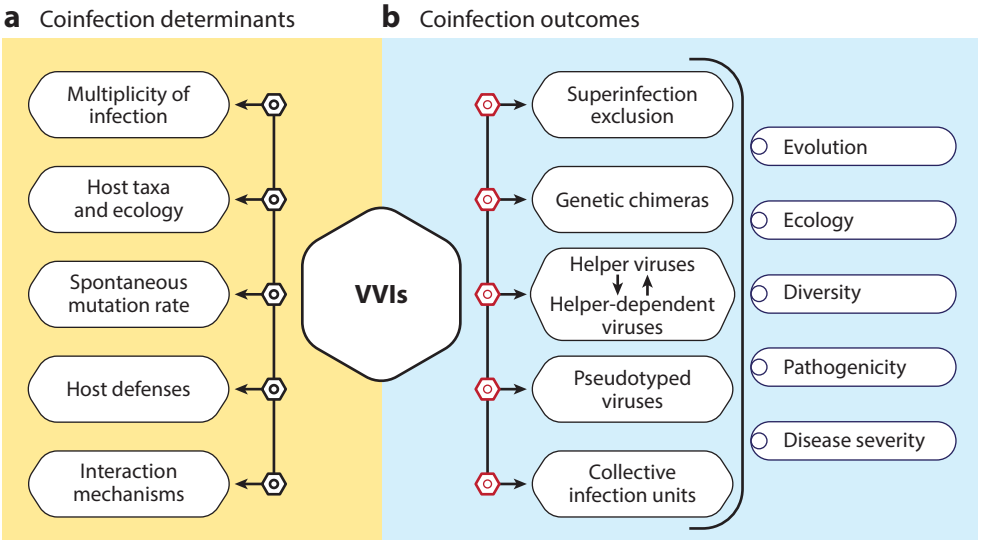

Figure 4

Summary of factors governing virus-virus interactions (VVIs) and potential outcomes of these interactions. Panel *a* (*yellow*) indicates important variables that affect VVIs during coinfections. High spontaneous mutation rate and multiplicity of infection favor the emergence of genetically diverse viral populations and enable exchange of public goods between virus variants. Host environments, including host habitat, tissue damage that compromises host barriers, and altered expression of viral receptors, are strong predictors of VVIs. While the presence of quiescently integrated proviruses within host genomes strongly inhibited coinfections, interactions between double-stranded DNA and single-stranded DNA phages have been suggested to promote coinfections. Additionally, acquisition of adaptive immunological memory to the primary virus may alter the outcomes of subsequent infections by a second virus. Panel *b* (*blue*) highlights the immediate and broad-scale effects of VVIs.

authors showed that an RNA phage infecting *Pseudomonas phaseolicola*, called phage ϕ6, adopts a cooperative strategy that resembles the prisoner's dilemma. In this scenario, high phage ϕ6 coinfection rates boost evolution of selfish phage ϕ6 variants with reduced fitness that rely on gene products of coinfecting genotypes for propagation. Such individual sacrifices and evolution of a cooperative breeding system allow viruses to essentially chaperone useful genes into successive generations and favor kin selection.

The interactions of viruses include cooperation, communication, and conflict, such as competition. Competitive VVIs can range from exploitative competition to interference. Some viruses infecting bacteria, plants, and mammals induce protection against subsequent infection by a related virus, a phenomenon termed superinfection exclusion (120). Cooperation and conflict go hand in hand when helper-dependent viruses (121–125) interact with their helper viruses. Helper-dependent viruses lack the structural components or replication machinery required for the formation of fully infectious particles, and hence they are obligatorily reliant on another virus, termed helper virus, for propagation. However, helper-dependent and helper virus pairs can interfere with one another to alter viral infection dynamics (122), pathogenicity (126, 127), and evolutionary trajectory (128). In some viral populations, defective interfering particles (DIPs) harboring large lethal deletions can propagate only in the presence of a fully functional wild-type virus (129). In the presence of a large number of total infecting particles, DIPs have competitive advantage over the wild-type variants and attain high population densities by quickly replicating their short genomes (129). Influenza virus DIPs that out-replicate fully infectious homologous viruses are regularly generated during human infections, and it has been suggested that they are cotransmitted with the infectious particles (130). Accumulation of DIPs during influenza infection can alter host immune response and has been correlated with reduced pathogenicity of influenza virus. Recent breakthroughs in high-throughput sequencing shed light on the enormous genetic diversity of influenza DIPs and revealed viral strain-specific patterns of DIP formation (130, 131). These studies set the stage to further probe molecular events underlying DIP formation across diverse influenza strains and study the contributions of distinct DIP populations to disease severity.

Investigators have sampled viral communities from insect vectors to characterize helper DNA viruses and associated satellite molecules that target plants (132) and even other viruses. Unlike DIPs that rely on wild-type virus variants for successful propagation, helper-dependent satellite viruses are incapable of completing their infection cycle without the assistance of a helper virus belonging to a different virus family. Metagenomic studies have vastly amplified our inventory of nucleocytoplasmic large DNA viruses (NCLDVs) and smaller satellite viruses that parasitize NCLDVs, called virophages, and provided insights on the gene repertoire, cognate NCLDVs, and distinct abundance profiles of virophages associated with unique environments (133–135). From an ecological perspective, virophages can have profound implications in microbial nutrient cycling, often referred to as the microbial loop. Predator-prey simulation models indicate that the presence of virophages regulates helper virus-algal host dynamics and alters carbon flux through the microbial loop in aquatic ecosystems (136). In parallel, several metagenomic and single-cell genomic studies have facilitated the discovery of novel NCLDVs and provided insight into their genetic diversity, biology, evolution, and ecology (133, 136–140).

Mounting evidence suggests that cooperation within and between viruses is also prevalent. Studies show that two independently replicating yet homologous viruses can recombine (phenotypically or genotypically) during coinfection to produce hybrid viral particles that have increased viral fitness and thereby deleterious consequences for human health (120). Others have demonstrated that viruses acquired traits that facilitated their collective transmission and enabled them to function as infection units (141). Some phages have also adapted distinct strategies to eject multiple phage genomes into the same bacterial host (142). A notable characteristic of filamentous

coliphage f1, for example, is their ability to encapsulate multiple phage genome copies within a single capsid. Additionally, factors such as temperature, pH, and salt concentrations can influence formation of phage aggregates and thereby favor interactions among virions during infection. In other instances, beneficial cooperative behaviors have been shown to guide infection cycle choices of phages (143) or allow phages to evade bacterial adaptive immune systems (144, 145). However, implications of collective behavior of phages on population fitness and evolution under a given environmental context are open avenues for future exploration.

Recent advances with metagenomic techniques have enabled characterization of the frequency, extent, and ecology of VVIs. Evolutionary analysis of microphage genomes assembled from public virome sequences indicates that coinfection events may have primed genetic exchange between ssDNA or dsDNA viruses (146). Metagenomic studies that uncovered chimeric RNA-DNA hybrid viruses imply that systematic coinfections have played crucial roles in the ecology and evolution of these viruses (147, 148). For example, evaluation of 13,103 viruses from 6,564 microbial hosts in three distinct large-scale microbial data sets predicted that host ecology and VVIs strongly regulate frequency and extent of viral coinfections (149) (**Figure 4**). Collectively, these reports indicate that VVI is a major driving force for viral evolution and coinfection is a widespread phenomenon, and that VVIs should be emphasized in the study of viral ecology across many ecosystems.

Sociovirology offers a conceptual foundation to decipher ecological and evolutionary consequences of interactions among and between viruses (118). Coinfecting viruses may exchange genetic material or proteins, complement growth and boost pathogenicity, or cheat and interfere with infection. Such interplays between coinfecting viruses can have profound consequences on the evolution of viruses and health of their hosts. The field of sociovirology is still in its infancy, and deploying high-throughput tools to untangle complex VVIs could reshape the fundamentals of virology and help the design of next-generation therapeutics.

Deep-sequencing studies have only further highlighted the pervasiveness of VVIs. While investigating virus-host relationships in publicly available sequence data, Roux et al. (150) discovered that virus coinfections are common, although the exchange of genetic information is less prevalent. Single-cell sequencing and viral metagenomic approaches confirm that coinfection is common across diverse environments (18, 151). Díaz-Muñoz (149) exploited multiple large-scale data sets of virus-host interactions to pinpoint biotic and abiotic factors that regulate viral coinfections. Collectively, these studies and a recent perspective article (17) highlight how high-throughput technologies that integrate molecular and evolutionary approaches can be employed to gain insights into the social life of viruses. These social interactions, and their implications for the diversity, distribution, and dynamics of viruses, should be accounted for in the analysis of viral metagenomics, as it is increasingly integrated into an ecological framework.

5. APPLICATIONS AND RESEARCH NEEDS OF VIRAL METAGENOMICS WITHIN AN ECOLOGICAL FRAMEWORK

Beyond advancing the field of virology, viral metagenomics has yielded important applications for medicine, agriculture, and conservation biology, and it will continue to advance applications as critical basic research needs are met. We describe areas of fundamental research needs within viral metagenomics highlighted by an ecological framework, including in VVIs, that will advance viral ecology in environmental and human-relevant contexts.

5.1. Applications of Viral Metagenomics Within an Ecological Framework

The steady decline of antibiotic susceptibility among pathogens, coupled with the plummeting rate of new antibiotic discovery, has resulted in new attention to the application of virulent phages

for pathogen control. Viral metagenomic approaches have been used to identify the phages and screen for potential virulence or antimicrobial resistance genes in a commercial phage cocktail targeting *Escherichia coli/Proteus* infections (152). In another study, Fujimoto et al. (153) employed viral metagenomics to reveal associations between intestinal phages and their hosts and identified novel phage-encoded antibacterial enzymes that specifically control opportunistic pathogens, such as *C. difficile*. However, the development of effective phage cocktails has been hampered in part by the diversity of phages in nature, limiting replicability and replication of studies (154, 155). A systematic and high-throughput search for effective phage cocktails using metagenomics to characterize the diversity and potential hosts of phages is required to develop phage-based alternatives to widespread antibiotic use (154). Considering that viral genomes encode a repertoire of proteins with therapeutic and/or diagnostic potential (156), culture-independent viral metagenomic tools are poised to explore novel viral genes and discover novel phage-derived proteins with clinical and agricultural applications.

Given the complex nature of phage-bacterial interactions, more consideration is needed for the evolutionary dynamics and consequences of phage resistance in the target bacteria (155). Reports indicate that spatial heterogeneity (157–159) and the presence of alternate hosts in the neighboring microbiota (160) can alter the trajectory of phage-host coevolution dynamics and have significant implications for the outcome of phage therapy. Another area that deserves further attention is the effect of phages on nontarget bystander bacteria present in the microbiome. Conflicting views exist on the potential effect of phage predation on microbiota; one study indicates that phage infection drives compositional shifts in the murine gut microbiota (161), whereas another group claims that phages cause negligible fluctuations in phylogenetic compositions of the microbiota (162). Recent in vitro work provides direct evidence of how phage infection of a clinically relevant opportunistic pathogen triggers an antibacterial response and consequently restricts the growth of neighboring bacterial species that are not susceptible to phage infection (16). Together, these studies set the stage to address two key outstanding questions in the field: (*a*) how the evolution of phage resistance in a bacterial species affects its competitive interactions for resources with other species with similar or overlapping ecological niches and (*b*) whether phage-induced changes in host metabolism result in alterations of the commensal bacterial and phage communities.

Beyond phage therapy, emerging views suggest that the diversity, dynamics, and interactions of phages influence mammalian health and disease by regulating the native bacteriome and through direct interactions with the human immune system (34, 163). Several recent reports advocated correlations between altered virome and human diseases such as inflammatory bowel diseases (164, 165), acute malnutrition (166), AIDS (28), autism spectrum disorders (167, 168), arthritic autoimmune diseases (169, 170), diabetes (171, 172), and cancer (173, 174). Although these disease-associated shifts in the phage communities imply that phages might be responsible for the co-occurring bacterial dysbiosis, direct evidence for such reciprocal relationships is missing (164). However, phage can carry auxiliary metabolic genes (AMGs) (175–177) or antibiotic-resistant markers (178) and directly influence the ecology and evolution of bacterial communities. Analyses of viral metagenomes from oropharyngeal swab samples resulted in the discovery of phage-derived virulence genes in these communities (179). Development of computational tools such as DRAM-v (Distilled and Refined Annotation of Metabolism in viral mode) that predict metabolic profiles of viromes provides an excellent avenue to exploring phage AMGs (180) and will vastly increase our knowledge of the functional potential of viromes.

Viral metagenomics also has significant applications within agricultural settings, where it has, for example, revealed the sensitivity of soil viral dynamics to spatial and temporal changes in factors such as plant presence (57). RNA sequencing moreover has shown the extent of infection and coinfection in economically important crop plants through viral metagenomics (181), and

potential risks and sources of infective plant viruses from irrigation with reclaimed wastewater (182). Beyond screening for the presence of pathogens, metagenomics has also been used to design phage-based products such as endolysins for application directly to food for enhancing its safety and preservation during transport and sale (155, 183).

Viral metagenomics can also be used to achieve goals in conservation biology, from the protection of individual species of concern to forecasting effects of climate change on soil carbon cycling and thereby improving projections of future climate change. For example, RNA viruses have been found in soil that appear to be actively replicating, lysing their hosts and releasing carbon to be taken up by other organisms (184). Further time series of metatranscriptomics and metagenomics on soils, including experiments under forecasted future conditions, are needed to understand how viral-mediated carbon cycling may change in the future and how phages in the rhizosphere could mediate plant responses to those changes (185). Changes in climate and habitat may also affect the likelihood of viral disease in wildlife populations that are critically endangered (186) or of high economic importance in ecosystem management (187). Sequence analyses of viral pathogens from these populations can shed light on the sources of infection spillover (20, 63, 188). Research into effective and affordable strategies for using viral metagenomics in a surveillance program could help protect vulnerable wildlife populations and integrate these data into community ecology (189) for better management of ecosystems.

5.2. Research Needs in Viral Metagenomics Highlighted by an Ecological Framework

An ecological framework highlights the need for specific approaches to be better developed and more widely adopted. These include improving our ability to characterize viral diversity and distributions across scales and the development of more quantitative methods for measuring dynamics and detecting interactions (summarized in **Table 2**).

Our ability to characterize viral diversity within and among populations is improving with the development of new approaches, although limitations remain for each approach. Problems in the assembly of short reads arise when similar variants of a viral species are present in the sample (71, 190). This diversity, however, may be ecologically relevant (11, 53) and could be better addressed through methods such as long-read sequencing, vSAGs, and in silico approaches to predict variant combinations in assembly (9, 11, 75, 83). These methods should especially be developed in more high-throughput pipelines for more accurately determining whether a genome is complete and which genes actually belong to that genome.

Additionally, viral metagenomic methods likely miss many of the rarer viral genomes due to small amounts of nucleic acids present in samples and difficulty in annotating short sequences (191). Methods such as long-read pipelines should be further developed to help detect those viral genomes because while rare, they could still play an ecologically important role, as evidenced by their ubiquity in some cases (192). Methods of extracting nucleic acids from VLPs alone improve the resolution of viral genomes relative to metagenomes obtained from the whole sample (57), and these methods continue to improve for low-biomass and difficult samples (10, 60, 193, 194). Large gaps remain in our knowledge of the distribution of viral diversity, VVIs and virus-host interactions, and the context-specific nature of viral dynamics across a wide range of habitats and environmental conditions. Ecological interactions are often dependent on the environmental context and need to be studied within a community ecology framework to understand how context shapes viral evolution (76, 189, 195). New databases of viral sequences from metagenomic studies such as the IMG/VR (1) are beginning to shed light on the importance of relatively understudied environments, as well as less-studied habitats within host-associated ecosystems such as the

Table 2 Fundamental research needs for advancing viral metagenomics highlighted by an ecological framework

Focus	Challenge(s)	Recent progress	Needs
Diversity	Detecting the rare virosphere	Nucleic acids extracted from VLPs to exclude hosts and deeper sequencing; methods targeting a subset of the community (e.g., SIP to label active members or degraders of a particular substrate)	Further resolution in high-throughput viral metagenomics for detecting rare species in less sample material; sampling different environmental conditions; using complementary methods of isolating VLPs for sequencing
	Variant heterogeneity preventing assembly	Deep sequencing; long-read sequencing; in silico approaches to closing viromic islands; single-virus sequencing	Wider application of combinations of sequencing approaches; further development of in silico approaches for assembling genomes with variant heterogeneity
	Determination of genome completeness, including segmented/multicomponent viruses	vSAGs; long-read sequencing; bioinformatics tools	More complete reference genomes in databases
	Prevalence and importance of various types of viruses (ssDNA, dsDNA, ssRNA, dsRNA, etc.)	Paired metagenomes with metatranscriptomes or viromes; paired DNA and RNA viromes; amplification methods selecting for both ssDNA and dsDNA	Wider application of complementary sequencing techniques, along with methods to determine absolute abundances and ecological effects
Distributions	Lack of environmental metadata	Minimum information standards and increased environmental metadata	Studies taking advantage of new database resources and designing experiments to determine environmental drivers of distributions
	Undersampled environments	Greater sampling of novel, extreme, and overlooked environments	Prioritization of environments to characterize; development of predictive models benchmarked with empirical data
	Uneven sampling across environments with different approaches	Paired DNA and RNA viromes	Wider application of multiple sequencing approaches; further development of unbiased quantitative sequencing approaches
Dynamics	Activity cannot be assessed from single metagenome alone	Time series and pairing metagenomes with measures of activity	Development of methods to apply SIP or other activity measures in situ across diverse environments
	Quantitative estimates of absolute abundance	Development of low-input sequencing methods to reduce the need for amplification; benchmarking of quantitative viral metagenomics with representative mock communities	Further optimization of low-input sequencing methods and statistical methods for quantifying uncertainty
	Context-dependent dynamics	Studies of comparative dynamics under differing conditions	Studies manipulating environmental conditions to compare dynamics and identify drivers of differences

(*Continued*)

Table 2 (*Continued*)

Focus	Challenge(s)	Recent progress	Needs
	Prevalence and timing of latent infections, including lysogeny	Estimates using induction of lytic phase; searches for proviruses in genomic data sets	Better resolution of proviruses and their prevalence and distributions through approaches such as SAGs and bioinformatic tools
	Linkages across scales, from populations to communities to ecosystems	Studies linking spillover of zoonotic viruses to humans or domesticated animals to land conversion	Development of in situ tractable model systems in which to link scales
Interactions	Characterizing and quantifying the host range of both cultured and uncultured viruses	CRISPR spacers; machine-learning predictive approaches; SAGs of host cells	Optimization of vSAG to infer phage-bacteria interaction networks across different habitats and determine host range of bioengineered phages
	Direct and indirect influences of viruses on host communities and ecosystem functions	Studies focusing on the characterization of phage-resistant bacteria during phage therapy; effect of phage on microbial community structure; indirect effects of viral infection or presence alone on host communities and ecosystem function	Broad application of high-throughput sequencing technologies to explore how the host and its surrounding community change in response to viruses; broadening research focus to include indirect and cascading interactions
	Characterizing and quantifying VVIs	Identification of environmental and biological predictors of these interactions; improved characterization of virus-virus communication (quorum sensing or arbitrium system); genetic diversity and strain-specific profiles of viral genetic variants within the strain	Implications of therapeutically administered coinfecting phages on (*a*) replication dynamics and evolution of the phages and (*b*) structure and function of the host-associated microbial community; effect of viral strains on VVIs and its epidemiological implications; studies of whether and how distinct variants affect evolution of the homologous wild-type virus

Abbreviations: ds, double-stranded; SAG, single amplified genome; SIP, stable isotope probing; ss, single-stranded; VLP, virus-like particle; vSAG, viral single amplified genome; VVI, virus-virus interaction.

human skin and vaginal mucosa. Further work on undersampled environments and modeling to predict and prioritize specific environments will improve our understanding of viral distributions and context-specific interactions.

Another key gap in our understanding of viral distributions is the mismatch between the spatial resolution at which virus-host interactions occur in structured environments, such as soil or human skin, and the resolution at which we are able to detect viruses and microbial hosts (196). Single-cell sequencing approaches are beginning to address this in some microbial mats (196, 197). Further work characterizing the viral distribution on scales difficult to sample will improve with continued developments in low-input sampling and sequencing techniques.

Viral metagenomic studies have moved from simply characterizing single metagenomes to seeking ecological inference about the interactions and dynamics of the viruses. The development

of methods to distinguish actively replicating populations from dormant or dead biomass (e.g., SIP) and their use in combination with viral metagenomics should be more widely adopted (for an example, see 94). Alternatively, metatranscriptomes paired with metagenomes can reveal not only RNA viruses but also transcriptional activity of DNA viruses, although the small amounts of mRNA and lack of correlation between transcription and actual translation and virion production do require care in interpreting such data.

Progress on addressing each of these areas has advanced rapidly in recent years, and we are optimistic these challenges can and will be addressed as bioinformatic tools continue to be developed and laboratory processing of samples continues to be optimized.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We apologize to any colleagues we could not cite due to limitations on the number of references. G.T.'s work conducted by the Lawrence Livermore National Laboratory (LLNL) was supported by the US Department of Energy (DOE) Office of Science, Office of Biological and Environmental Research Genomic Science program award SCW1632, LLNL LDRD 21-LW-060, and under the auspices of the DOE under contract DE-AC52-07NA27344.

LITERATURE CITED

1. Roux S, Paez-Espino D, Chen IA, Palaniappan K, Ratner A, et al. 2020. IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic Acids Res.* 49:D764–75
2. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, et al. 2016. Redefining the invertebrate RNA virosphere. *Nature* 540:539–43
3. Tisza MJ, Pastrana DV, Welch NL, Stewart B, Peretti A, et al. 2020. Discovery of several thousand highly diverse circular DNA viruses. *eLife* 9:e51971
4. Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, et al. 2002. Genomic analysis of uncultured marine viral communities. *PNAS* 99:14250–55
5. Brum JR, Ignacio-Espinoza JC, Roux S, Doulcier G, Acinas SG, et al. 2015. Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498
6. Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, et al. 2016. Uncovering Earth's virome. *Nature* 536:425–30
7. Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, et al. 2018. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* 3:870–80
8. Zhong ZP, Rapp JZ, Wainaina JM, Solonenko NE, Maughan H, et al. 2020. Viral ecogenomics of arctic cryopeg brine and sea ice. *mSystems* 5:e00246-20
9. Bellas CM, Schroeder DC, Edwards A, Barker G, Anesio AM. 2020. Flexible genes establish widespread bacteriophage pan-genomes in cryoconite hole ecosystems. *Nat. Commun.* 11:4403
10. Trubl G, Jang HB, Roux S, Emerson JB, Solonenko N, et al. 2018. Soil viruses are underexplored players in ecosystem carbon processing. *mSystems* 3:e00076-18
11. Gregory AC, Zayed AA, Conceicao-Neto N, Temperton B, Bolduc B, et al. 2019. Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177:1109–23.e14
12. Rosenwasser S, Ziv C, Creveld SGV, Vardi A. 2016. Virocell metabolism: metabolic innovations during host-virus interactions in the ocean. *Trends Microbiol.* 24:821–32

13. Zimmerman AE, Howard-Varona C, Needham DM, John SG, Worden AZ, et al. 2020. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* 18:21–34
14. Fuchsman CA, Palevsky HI, Widner B, Duffy M, Carlson MCG, et al. 2019. Cyanobacteria and cyanophage contributions to carbon and nitrogen cycling in an oligotrophic oxygen-deficient zone. *ISME J.* 13:2714–26
15. Secor PR, Sweere JM, Michaels LA, Malkovskiy AV, Lazzareschi D, et al. 2015. Filamentous bacteriophage promote biofilm assembly and function. *Cell Host Microbe* 18:549–59
16. Chatterjee A, Willett JLE, Dunny GM, Duerkop BA. 2021. Phage infection and sub-lethal antibiotic exposure mediate *Enterococcus faecalis* type VII secretion system dependent inhibition of bystander bacteria. *PLOS Genet.* 17:e1009204
17. Díaz-Muñoz SL. 2019. Uncovering virus-virus interactions by unifying approaches and harnessing high-throughput tools. *mSystems* 4(3):e00121-19
18. Munson-McGee JH, Peng S, Dewerff S, Stepanauskas R, Whitaker RJ, et al. 2018. A virus or more in (nearly) every cell: ubiquitous networks of virus-host interactions in extreme environments. *ISME J.* 12:1706–14
19. Temmam S, Chretien D, Bigot T, Dufour E, Petres S, et al. 2019. Monitoring silent spillovers before emergence: a pilot study at the tick/human interface in Thailand. *Front. Microbiol.* 10:2315
20. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, et al. 2017. Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* 15:502–10
21. Duerkop BA, Hooper LV. 2013. Resident viruses and their interactions with the immune system. *Nat. Immunol.* 14:654–59
22. Rascovan N, Duraisamy R, Desnues C. 2016. Metagenomics and the human virome in asymptomatic individuals. *Annu. Rev. Microbiol.* 70:125–41
23. Koonin EV, Yutin N. 2020. The crAss-like phage group: how metagenomics reshaped the human virome. *Trends Microbiol.* 28:349–59
24. Creasy A, Rosario K, Leigh BA, Dishaw LJ, Breitbart M. 2018. Unprecedented diversity of ssDNA phages from the family *Microviridae* detected within the gut of a protochordate model organism (*Ciona robusta*). *Viruses* 10(8):404
25. Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GG, et al. 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* 5:4498
26. Abt MC, Buffie CG, Susac B, Becattini S, Carter RA, et al. 2016. TLR-7 activation enhances IL-22-mediated colonization resistance against vancomycin-resistant enterococcus. *Sci. Transl. Med.* 8:327ra25
27. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, et al. 2017. Efficacy of sterile fecal filtrate transfer for treating patients with *Clostridium difficile* infection. *Gastroenterology* 152:799–811.e7
28. Monaco CL, Gootenberg DB, Zhao G, Handley SA, Ghebremichael MS, et al. 2016. Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. *Cell Host Microbe* 19:311–22
29. Lysholm F, Wetterbom A, Lindau C, Darban H, Bjerkner A, et al. 2012. Characterization of the viral microbiome in patients with severe lower respiratory tract infections, using metagenomic sequencing. *PLOS ONE* 7:e30875
30. Willner D, Furlan M, Haynes M, Schmieder R, Angly FE, et al. 2009. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. *PLOS ONE* 4:e7370
31. Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, et al. 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 466:334–38
32. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, et al. 2011. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* 21:1616–25
33. Pride DT, Salzman J, Haynes M, Rohwer F, Davis-Long C, et al. 2012. Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J.* 6:915–26
34. Chatterjee A, Duerkop BA. 2018. Beyond bacteria: bacteriophage-eukaryotic host interactions reveal emerging paradigms of health and disease. *Front. Microbiol.* 9:1394

35. Hannigan GD, Meisel JS, Tyldsley AS, Zheng Q, Hodkinson BP, et al. 2015. The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *mBio* 6:e01578-15
36. Abeles SR, Robles-Sikisaka R, Ly M, Lum AG, Salzman J, et al. 2014. Human oral viruses are personal, persistent and gender-consistent. *ISME J.* 8:1753–67
37. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. 2020. The gut virome database reveals age-dependent patterns of virome diversity in the human gut. *Cell Host Microbe* 28:724–40.e8
38. Kernbauer E, Ding Y, Cadwell K. 2014. An enteric virus can replace the beneficial function of commensal bacteria. *Nature* 516:94–98
39. Van Belleghem JD, Dabrowska K, Vaneechoutte M, Barr JJ, Bollyky PL. 2018. Interactions between bacteriophage, bacteria, and the mammalian immune system. *Viruses* 11:10
40. Koskella B, Lin DM, Buckling A, Thompson JN. 2012. The costs of evolving resistance in heterogeneous parasite environments. *Proc. R. Soc. B* 279:1896–903
41. Modi SR, Lee HH, Spina CS, Collins JJ. 2013. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499:219–22
42. Adler BA, Kazakov AE, Zhong C, Lui H, Kutter E, et al. 2021. The genetic basis of phage susceptibility, cross-resistance and host-range in *Salmonella*. bioRxiv 2020.04.27.058388. **https://doi.org/10.1101/2020.04.27.058388**
43. Chatterjee A, Willett JLE, Nguyen UT, Monogue B, Palmer KL, et al. 2020. Parallel genomics uncover novel enterococcal-bacteriophage interactions. *mBio* 11:e03120-19
44. Soffer N, Woolston J, Li M, Das C, Sulakvelidze A. 2017. Bacteriophage preparation lytic for *Shigella* significantly reduces *Shigella sonnei* contamination in various foods. *PLOS ONE* 12:e0175256
45. Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, et al. 2019. Minimum information about an uncultivated virus genome (MIUViG). *Nat. Biotechnol.* 37:29–37
46. Kazlauskas D, Varsani A, Koonin EV, Krupovic M. 2019. Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nat. Commun.* 10:3425
47. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, et al. 2020. Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.* 84:e00061-19
48. Simmonds P, Adams MJ, Benko M, Breitbart M, Brister JR, et al. 2017. Consensus statement: virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* 15:161–68
49. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, et al. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579:270–73
50. Pappas N, Roux S, Hölzer M, Lamkiewicz K, Mock F, et al. 2021. Virus bioinformatics. In *Encyclopedia of Virology*, ed. DH Bamford, M Zuckerman, pp. 124–32. Amsterdam: Academic. 4th ed.
51. Trubl G, Hyman P, Roux S, Abedon ST. 2020. Coming-of-age characterization of soil viruses: a user's guide to virus isolation, detection within metagenomes, and viromics. *Soil Syst*. 4:23
52. Zhang YZ, Chen YM, Wang W, Qin XC, Holmes EC. 2019. Expanding the RNA virosphere by unbiased metagenomics. *Annu. Rev. Virol.* 6:119–39
53. Guo J, Bolduc B, Zayed A, Varsani A, Dominguez-Huerta G, et al. 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 9:37
54. Tisza MJ, Belford AK, Domínguez-Huerta G, Bolduc B, Buck CB. 2021. Cenote-Taker 2 democratizes virus discovery and sequence annotation. *Virus Evol.* 7:veaa100
55. Hyman P, Trubl G, Abedon ST. 2021. Virus-like particle: evolving meanings in different disciplines. *Phage* 2:11–15
56. Sommers P, Fontenele RS, Kringen T, Kraberger S, Porazinska DL, et al. 2019. Single-stranded DNA viruses in Antarctic cryoconite holes. *Viruses* 11:1022
57. Santos-Medellin C, Zinke LA, ter Horst AM, Gelardi DL, Parikh SJ, Emerson JB. 2021. Viromes outperform total metagenomes in revealing the spatiotemporal patterns of agricultural soil viral communities. *ISME J.* **https://doi.org/10.1038/s41396-021-00897-y**
58. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. 2009. Laboratory procedures to generate viral metagenomes. *Nat. Protoc.* 4:470–83
59. Trubl G, Solonenko N, Chittick L, Solonenko SA, Rich VI, Sullivan MB. 2016. Optimization of viral resuspension methods for carbon-rich soils along a permafrost thaw gradient. *PeerJ* 4:e1999

60. Trubl G, Roux S, Solonenko N, Li YF, Bolduc B, et al. 2019. Towards optimized viral metagenomes for double-stranded and single-stranded DNA viruses from challenging soils. *PeerJ* 7:e7265
61. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, et al. 2003. Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* 185:6220–23
62. Hjelmso MH, Hellmer M, Fernandez-Cassi X, Timoneda N, Lukjancenko O, et al. 2017. Evaluation of methods for the concentration and extraction of viruses from sewage in the context of metagenomic sequencing. *PLOS ONE* 12:e0170199
63. Bernardo P, Charles-Dominique T, Barakat M, Ortet P, Fernandez E, et al. 2018. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *ISME J.* 12:173–84
64. Vega Thurber RL, Barott KL, Hall D, Liu H, Rodriguez-Mueller B, et al. 2008. Metagenomic analysis indicates that stressors induce production of herpes-like viruses in the coral *Porites compressa*. *PNAS* 105:18413–18
65. Zolotukhin S, Byrne BJ, Mason E, Zolotukhin I, Potter M, et al. 1999. Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.* 6:973–85
66. Sun G, Xiao J, Wang H, Gong C, Pan Y, et al. 2014. Efficient purification and concentration of viruses from a large body of high turbidity seawater. *MethodsX* 1:197–206
67. John SG, Mendez CB, Deng L, Poulos B, Kauffman AK, et al. 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ. Microbiol. Rep.* 3:195–202
68. Reichart NJ, Jay ZJ, Krukenberg V, Parker AE, Spietz RL, Hatzenpichler R. 2020. Activity-based cell sorting reveals responses of uncultured archaea and bacteria to substrate amendment. *ISME J.* 14:2851–61
69. Allen LZ, Ishoey T, Novotny MA, McLean JS, Lasken RS, Williamson SJ. 2011. Single virus genomics: a new tool for virus discovery. *PLOS ONE* 6:e17722
70. Martinez-Garcia M, Martinez-Hernandez F, Martínez JM. 2020. Single-virus genomics: studying uncultured viruses, one at a time. In *Encyclopedia of Virology*, ed. DH Bamford, M Zuckerman, pp. 184–90. Amsterdam: Academic. 4th ed.
71. Roux S, Emerson JB, Eloe-Fadrosh EA, Sullivan MB. 2017. Benchmarking viromics: an *in silico* evaluation of metagenome-enabled estimates of viral community composition and diversity. *PeerJ* 5:e3817
72. Varsani A, Lefeuvre P, Roumagnac P, Martin D. 2018. Notes on recombination and reassortment in multipartite/segmented viruses. *Curr. Opin. Virol.* 33:156–66
73. Martinez-Hernandez F, Fornas O, Lluesma Gomez M, Bolduc B, de la Cruz Pena MJ, et al. 2017. Single-virus genomics reveals hidden cosmopolitan and abundant viruses. *Nat. Commun.* 8:15892
74. de la Cruz Pena MJ, Martinez-Hernandez F, Garcia-Heredia I, Lluesma Gomez M, Fornas O, Martinez-Garcia M. 2018. Deciphering the human virome with single-virus genomics and metagenomics. *Viruses* 10(3):113
75. Martínez JM, Martinez-Hernandez F, Martinez-Garcia M. 2020. Single-virus genomics and beyond. *Nat. Rev. Microbiol.* 18:705–16
76. Lefeuvre P, Martin DP, Elena SF, Shepherd DN, Roumagnac P, Varsani A. 2019. Evolution and ecology of plant viruses. *Nat. Rev. Microbiol.* 17:632–44
77. Roux S, Solonenko NE, Dang VT, Poulos BT, Schwenck SM, et al. 2016. Towards quantitative viromics for both double-stranded and single-stranded DNA viruses. *PeerJ* 4:e2777
78. Okada R, Kiyota E, Moriyama H, Fukuhara T, Natsuaki T. 2015. A simple and rapid method to purify viral dsRNA from plant and fungal tissue. *J. Gen. Plant Pathol.* 81:103–7
79. Lukács N. 1994. Detection of virus infection in plants and differentiation between coexisting viruses by monoclonal antibodies to double-stranded RNA. *J. Virol. Methods* 47:255–72
80. Seguin J, Rajeswaran R, Malpica-Lopez N, Martin RR, Kasschau K, et al. 2014. *De novo* reconstruction of consensus master genomes of plant RNA and DNA viruses from siRNAs. *PLOS ONE* 9:e88513
81. Wu Q, Luo Y, Lu R, Lau N, Lai EC, et al. 2010. Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *PNAS* 107:1606–11
82. Logsdon GA, Vollger MR, Eichler EE. 2020. Long-read human genome sequencing and its applications. *Nat. Rev. Genet.* 21:597–614

83. Zablocki O, Michelsen M, Burris M, Solonenko N, Warwick-Dugdale J, et al. 2020. VirION2: a short- and long-read sequencing and informatics workflow to study the genomic diversity of viruses in nature. *PeerJ* 9:e11088
84. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. 2016. Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiol. Rev.* 40:258–72
85. Bikel S, Valdez-Lara A, Cornejo-Granados F, Rico K, Canizales-Quinteros S, et al. 2015. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Comput. Struct. Biotechnol. J.* 13:390–401
86. Sieradzki ET, Ignacio-Espinoza JC, Needham DM, Fichot EB, Fuhrman JA. 2019. Dynamic marine viral infections and major contribution to photosynthetic processes shown by spatiotemporal picoplankton metatranscriptomes. *Nat. Commun.* 10:1169
87. Rostol JT, Marraffini L. 2019. (Ph)ighting phages: how bacteria resist their parasites. *Cell Host Microbe* 25:184–94
88. Howard-Varona C, Hargreaves KR, Solonenko NE, Markillie LM, White RA 3rd, et al. 2018. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *ISME J.* 12:1605–18
89. Howard-Varona C, Lindback MM, Bastien GE, Solonenko N, Zayed AA, et al. 2020. Phage-specific metabolic reprogramming of virocells. *ISME J.* 14:881–95
90. Starr EP, Shi S, Blazewicz SJ, Probst AJ, Herman DJ, et al. 2018. Stable isotope informed genome-resolved metagenomics reveals that Saccharibacteria utilize microbially-processed plant-derived carbon. *Microbiome* 6:122
91. Barnett SE, Buckley DH. 2020. Simulating metagenomic stable isotope probing datasets with MetaSIPSim. *BMC Bioinform.* 21:37
92. Dumont MG, García MH. 2019. *Stable Isotope Probing*. New York: Humana
93. Haig SJ, Schirmer M, D'Amore R, Gibbs J, Davies RL, et al. 2015. Stable-isotope probing and metagenomics reveal predation by protozoa drives *E. coli* removal in slow sand filters. *ISME J.* 9:797–808
94. Trubl G, Kimbrel JA, Liquet-Gonzalez J, Nuccio EE, Weber PK, et al. 2021. Ecology of active viruses and their bacterial hosts in frozen Arctic peat soil revealed with $H_2{}^{18}O$ stable isotope probing metagenomics. bioRxiv 2021.01.25.428156. **https://doi.org/10.1101/2021.01.25.428156**
95. Blazewicz SJ, Hungate BA, Koch BJ, Nuccio EE, Morrissey E, et al. 2020. Taxon-specific microbial growth and mortality patterns reveal distinct temporal population responses to rewetting in a California grassland soil. *ISME J.* 14:1520–32
96. Sieradzki ET, Koch BJ, Greenlon A, Sachdeva R, Malmstrom RR, et al. 2020. Measurement error and resolution in quantitative stable isotope probing: implications for experimental design. *mSystems* 5:e00151-20
97. Kolb HC, Finn MG, Sharpless KB. 2001. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem.* 40:2004–21
98. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. 2002. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *PNAS* 99:19–24
99. Samo TJ, Smriga S, Malfatti F, Sherwood BP, Azam F. 2014. Corrigendum: broad distribution and high proportion of protein synthesis active marine bacteria revealed by click chemistry at the single cell level. *Front. Mar. Sci.* 1:60
100. Muller TG, Sakin V, Muller B. 2019. A spotlight on viruses—application of click chemistry to visualize virus-cell interactions. *Molecules* 24:481
101. Pett-Ridge J, Weber PK. 2012. NanoSIP: NanoSIMS applications for microbial biology. *Methods Mol. Biol.* 881:375–408
102. Pasulka AL, Thamatrakoln K, Kopf SH, Guan Y, Poulos B, et al. 2018. Interrogating marine virus-host interactions and elemental transfer with BONCAT and nanoSIMS-based methods. *Environ. Microbiol.* 20:671–92
103. Gates SD, Condit RC, Moussatche N, Stewart BJ, Malkin AJ, Weber PK. 2018. High initial sputter rate found for vaccinia virions using isotopic labeling, NanoSIMS, and AFM. *Anal. Chem.* 90:1613–20
104. Pacton M, Wacey D, Corinaldesi C, Tangherlini M, Kilburn MR, et al. 2014. Viruses as new agents of organomineralization in the geological record. *Nat. Commun.* 5:4298

105. Baran N, Goldin S, Maidanik I, Lindell D. 2018. Quantification of diverse virus populations in the environment using the polony method. *Nat. Microbiol.* 3:62–72
106. Castillo YM, Mangot JF, Benites LF, Logares R, Kuronishi M, et al. 2019. Assessing the viral content of uncultured picoeukaryotes in the global-ocean by single cell genomics. *Mol. Ecol.* 28:4272–89
107. Mruwat N, Carlson MCG, Goldin S, Ribalet F, Kirzner S, et al. 2021. A single-cell polony method reveals low levels of infected *Prochlorococcus* in oligotrophic waters despite high cyanophage abundances. *ISME J.* 15:41–54
108. Allers E, Moraru C, Duhaime MB, Beneze E, Solonenko N, et al. 2013. Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. *Environ. Microbiol.* 15:2306–18
109. Barrero-Canosa J, Moraru C. 2019. PhageFISH for monitoring phage infections at single cell level. *Methods Mol. Biol.* 1898:1–26
110. Vincent F, Sheyn U, Porat Z, Vardi A. 2021. Visualizing active viral infection reveals diverse cell fates in synchronized algal bloom demise. *PNAS* 118:e2021586118
111. Deng L, Ignacio-Espinoza JC, Gregory AC, Poulos BT, Weitz JS, et al. 2014. Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature* 513:242–45
112. Dzunkova M, Low SJ, Daly JN, Deng L, Rinke C, Hugenholtz P. 2019. Defining the human gut host-phage network through single-cell viral tagging. *Nat. Microbiol.* 4:2192–203
113. Marbouty M, Baudry L, Cournac A, Koszul R. 2017. Scaffolding bacterial genomes and probing host-virus interactions in gut microbiome by proximity ligation (chromosome capture) assay. *Sci. Adv.* 3:e1602105
114. Marbouty M, Thierry A, Milluit GA, Koszul R. 2021. MetaHiC phage-bacteria infection network reveals active cycling phages of the healthy human gut. *eLife* 10:e60608
115. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326:289–93
116. Kim KD, Tanizawa H, De Leo A, Vladimirova O, Kossenkov A, et al. 2020. Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus. *Nat. Commun.* 11:877
117. Nagano T, Lubling Y, Yaffe E, Wingett SW, Dean W, et al. 2015. Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat. Protoc.* 10:1986–2003
118. Díaz-Muñoz SL, Sanjuan R, West S. 2017. Sociovirology: conflict, cooperation, and communication among viruses. *Cell Host Microbe* 22:437–41
119. Turner PE, Chao L. 1999. Prisoner's dilemma in an RNA virus. *Nature* 398:441–43
120. DaPalma T, Doonan BP, Trager NM, Kasman LM. 2010. A systematic approach to virus-virus interactions. *Virus Res.* 149:1–9
121. Novick RP, Christie GE, Penades JR. 2010. The phage-related chromosomal islands of Gram-positive bacteria. *Nat. Rev. Microbiol.* 8:541–51
122. Penades JR, Christie GE. 2015. The phage-inducible chromosomal islands: a family of highly evolved molecular parasites. *Annu. Rev. Virol.* 2:181–201
123. Ryu W-S. 2017. Subviral agents and prions. In *Molecular Virology of Human Pathogenic Viruses*, ed. W-S Ryu, pp. 277–88. San Diego, CA: Elsevier
124. Krupovic M. 2020. Plant satellite viruses (*Albetovirus*, *Aumaivirus*, *Papanivirus*, *Virtovirus*). In *Encyclopedia of Virology*, ed. DH Bamford, M Zuckerman, pp. 581–85. Amsterdam: Academic. 4th ed.
125. Desnues C, Raoult D. 2010. Inside the lifestyle of the virophage. *Intervirology* 53:293–303
126. Alves C, Branco C, Cunha C. 2013. Hepatitis delta virus: a peculiar virus. *Adv. Virol.* 2013:560105
127. Gnanasekaran P, KishoreKumar R, Bhattacharyya D, Vinoth Kumar R, Chakraborty S. 2019. Multi-faceted role of geminivirus associated betasatellite in pathogenesis. *Mol. Plant Pathol.* 20:1019–33
128. Frigols B, Quiles-Puchalt N, Mir-Sanchis I, Donderis J, Elena SF, et al. 2015. Virus satellites drive viral evolution and ecology. *PLOS Genet.* 11:e1005609
129. Vignuzzi M, Lopez CB. 2019. Defective viral genomes are key drivers of the virus-host interaction. *Nat. Microbiol.* 4:1075–87
130. Brooke CB. 2017. Population diversity and collective interactions during influenza virus infection. *J. Virol.* 91:e01164-17

131. Alnaji FG, Holmes JR, Rendon G, Vera JC, Fields CJ, et al. 2019. Sequencing framework for the sensitive detection and precise mapping of defective interfering particle-associated deletions across influenza A and B viruses. *J. Virol.* 93:e00354-19
132. Ng TF, Duffy S, Polston JE, Bixby E, Vallad GE, Breitbart M. 2011. Exploring the diversity of plant DNA viruses and their satellites using vector-enabled metagenomics on whiteflies. *PLOS ONE* 6:e19050
133. Roux S, Chan LK, Egan R, Malmstrom RR, McMahon KD, Sullivan MB. 2017. Ecogenomics of virophages and their giant virus hosts assessed through time series metagenomics. *Nat. Commun.* 8:858
134. Zhou J, Zhang W, Yan S, Xiao J, Zhang Y, et al. 2013. Diversity of virophages in metagenomic data sets. *J. Virol.* 87:4225–36
135. Paez-Espino D, Zhou J, Roux S, Nayfach S, Pavlopoulos GA, et al. 2019. Diversity, evolution, and classification of virophages uncovered through global metagenomics. *Microbiome* 7:157
136. Yau S, Lauro FM, DeMaere MZ, Brown MV, Thomas T, et al. 2011. Virophage control of antarctic algal host-virus dynamics. *PNAS* 108:6163–68
137. Schulz F, Roux S, Paez-Espino D, Jungbluth S, Walsh DA, et al. 2020. Giant virus diversity and host interactions through global metagenomics. *Nature* 578:432–36
138. Schulz F, Alteio L, Goudeau D, Ryan EM, Yu FB, et al. 2018. Hidden diversity of soil giant viruses. *Nat. Commun.* 9:4881
139. Zhang W, Zhou J, Liu T, Yu Y, Pan Y, et al. 2015. Four novel algal virus genomes discovered from Yellowstone Lake metagenomes. *Sci. Rep.* 5:15131
140. Backstrom D, Yutin N, Jorgensen SL, Dharamshi J, Homa F, et al. 2019. Virus genomes from deep sea sediments expand the ocean megavirome and support independent origins of viral gigantism. *mBio* 10:e02497–18
141. Leeks A, Sanjuan R, West SA. 2019. The evolution of collective infectious units in viruses. *Virus Res.* 265:94–101
142. Domingo-Calap P, Mora-Quilis L, Sanjuan R. 2020. Social bacteriophages. *Microorganisms* 8:533
143. Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, et al. 2017. Communication between viruses guides lysis-lysogeny decisions. *Nature* 541:488–93
144. Borges AL, Zhang JY, Rollins MF, Osuna BA, Wiedenheft B, Bondy-Denomy J. 2018. Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. *Cell* 174:917–25.e10
145. Landsberger M, Gandon S, Meaden S, Rollie C, Chevallereau A, et al. 2018. Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *Cell* 174:908–16.e12
146. Roux S, Krupovic M, Poulet A, Debroas D, Enault F. 2012. Evolution and diversity of the *Microviridae* viral family through a collection of 81 new complete genomes assembled from virome reads. *PLOS ONE* 7:e40418
147. Diemer GS, Stedman KM. 2012. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol. Direct.* 7:13
148. de la Higuera I, Kasun GW, Torrance EL, Pratt AA, Maluenda A, et al. 2020. Unveiling crucivirus diversity by mining metagenomic data. *mBio* 11:e01410–20
149. Díaz-Muñoz SL. 2017. Viral coinfection is shaped by host ecology and virus-virus interactions across diverse microbial taxa and environments. *Virus Evol.* 3:vex011
150. Roux S, Hallam SJ, Woyke T, Sullivan MB. 2015. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *eLife* 4:e08490
151. Labonte JM, Swan BK, Poulos B, Luo H, Koren S, et al. 2015. Single-cell genomics-based analysis of virus-host interactions in marine surface bacterioplankton. *ISME J.* 9:2386–99
152. McCallin S, Alam Sarker S, Barretto C, Sultana S, Berger B, et al. 2013. Safety analysis of a Russian phage cocktail: from metagenomic analysis to oral application in healthy human subjects. *Virology* 443:187–96
153. Fujimoto K, Kimura Y, Shimohigoshi M, Satoh T, Sato S, et al. 2020. Metagenome data on intestinal phage-bacteria associations aids the development of phage therapy against pathobionts. *Cell Host Microbe* 28:380–89.e9
154. Kahn LH, Bergeron G, Bourassa MW, De Vegt B, Gill J, et al. 2019. From farm management to bacteriophage therapy: strategies to reduce antibiotic use in animal agriculture. *Ann. N. Y. Acad. Sci.* 1441:31–39
155. Fernandez L, Gutierrez D, Rodriguez A, Garcia P. 2018. Application of bacteriophages in the agro-food sector: a long way toward approval. *Front. Cell Infect. Microbiol.* 8:296

156. Schoenfeld T, Liles M, Wommack KE, Polson SW, Godiska R, Mead D. 2010. Functional viral metagenomics and the next generation of molecular tools. *Trends Microbiol.* 18:20–29
157. Yang H, Schmitt-Wagner D, Stingl U, Brune A. 2005. Niche heterogeneity determines bacterial community structure in the termite gut (*Reticulitermes santonensis*). *Environ. Microbiol.* 7:916–32
158. Heilmann S, Sneppen K, Krishna S. 2012. Coexistence of phage and bacteria on the boundary of self-organized refuges. *PNAS* 109:12828–33
159. Brockhurst MA, Buckling A, Rainey PB. 2006. Spatial heterogeneity and the stability of host-parasite coexistence. *J. Evol. Biol.* 19:374–79
160. De Sordi L, Khanna V, Debarbieux L. 2017. The gut microbiota facilitates drifts in the genetic diversity and infectivity of bacterial viruses. *Cell Host Microbe* 22:801–8.e3
161. Galtier M, De Sordi L, Maura D, Arachchi H, Volant S, et al. 2016. Bacteriophages to reduce gut carriage of antibiotic resistant uropathogens with low impact on microbiota composition. *Environ. Microbiol.* 18:2237–45
162. Hsu BB, Gibson TE, Yeliseyev V, Liu Q, Lyon L, et al. 2019. Dynamic modulation of the gut microbiota and metabolome by bacteriophages in a mouse model. *Cell Host Microbe* 25:803–14.e5
163. Keen EC, Dantas G. 2018. Close encounters of three kinds: bacteriophages, commensal bacteria, and host immunity. *Trends Microbiol.* 26:943–54
164. Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, et al. 2015. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 160:447–60
165. Wang W, Jovel J, Halloran B, Wine E, Patterson J, et al. 2015. Metagenomic analysis of microbiome in colon tissue from subjects with inflammatory bowel diseases reveals interplay of viruses and bacteria. *Inflamm. Bowel Dis.* 21:1419–27
166. Reyes A, Blanton LV, Cao S, Zhao G, Manary M, et al. 2015. Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. *PNAS* 112:11941–46
167. Kang DW, Adams JB, Gregory AC, Borody T, Chittick L, et al. 2017. Microbiota transfer therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome* 5:10
168. Lim ES, Wang D, Holtz LR. 2016. The bacterial microbiome and virome milestones of infant development. *Trends Microbiol.* 24:801–10
169. Lee JY, Mannaa M, Kim Y, Kim J, Kim GT, Seo YS. 2019. Comparative analysis of fecal microbiota composition between rheumatoid arthritis and osteoarthritis patients. *Genes* 10:748
170. Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, et al. 2013. Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife* 2:e01202
171. Ma Y, You X, Mai G, Tokuyasu T, Liu C. 2018. A human gut phage catalog correlates the gut phageome with type 2 diabetes. *Microbiome* 6:24
172. Tetz G, Brown SM, Hao Y, Tetz V. 2019. Type 1 diabetes: an association between autoimmunity, the dynamics of gut amyloid-producing *E. coli* and their phages. *Sci. Rep.* 9:9685
173. Minot SS, Willis AD. 2019. Clustering co-abundant genes identifies components of the gut microbiome that are reproducibly associated with colorectal cancer and inflammatory bowel disease. *Microbiome* 7:110
174. Yu AI, Zhao L, Eaton KA, Ho S, Chen J, et al. 2020. Gut microbiota modulate CD8 T cell responses to influence colitis-associated tumorigenesis. *Cell Rep.* 31:107471
175. Anantharaman K, Duhaime MB, Breier JA, Wendt KA, Toner BM, Dick GJ. 2014. Sulfur oxidation genes in diverse deep-sea viruses. *Science* 344:757–60
176. Sullivan MB, Coleman ML, Weigele P, Rohwer F, Chisholm SW. 2005. Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLOS Biol.* 3:e144
177. Campbell DE, Ly LK, Ridlon JM, Hsiao A, Whitaker RJ, Degnan PH. 2020. Infection with bacteroides phage BV01 alters the host transcriptome and bile acid metabolism in a common human gut microbe. *Cell Rep.* 32:108142
178. Abeles SR, Ly M, Santiago-Rodriguez TM, Pride DT. 2015. Effects of long term antibiotic therapy on human oral and fecal viromes. *PLOS ONE* 10:e0134941
179. Willner D, Furlan M, Schmieder R, Grasis JA, Pride DT, et al. 2011. Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. *PNAS* 108(Suppl. 1):4547–53
180. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, et al. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* 48:8883–900

181. Hily JM, Candresse T, Garcia S, Vigne E, Tanniere M, et al. 2018. High-throughput sequencing and the viromic study of grapevine leaves: from the detection of grapevine-infecting viruses to the description of a new environmental *Tymovirales* member. *Front. Microbiol.* 9:1782
182. Bacnik K, Kutnjak D, Pecman A, Mehle N, Tusek Znidaric M, et al. 2020. Viromics and infectivity analysis reveal the release of infective plant viruses from wastewater into the environment. *Water Res.* 177:115628
183. Chang Y. 2020. Bacteriophage-derived endolysins applied as potent biocontrol agents to enhance food safety. *Microorganisms* 8:724
184. Starr EP, Nuccio EE, Pett-Ridge J, Banfield JF, Firestone MK. 2019. Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *PNAS* 116:25900–8
185. Pratama AA, Terpstra J, de Oliveria ALM, Salles JF. 2020. The role of rhizosphere bacteriophages in plant health. *Trends Microbiol.* 28:709–18
186. Dunay E, Apakupakul K, Leard S, Palmer JL, Deem SL. 2018. Pathogen transmission from humans to great apes is a growing threat to primate conservation. *EcoHealth* 15:148–62
187. Gustafson LL, Creekmore LH, Snekvik KR, Ferguson JA, Warg JV, et al. 2018. A systematic surveillance programme for infectious salmon anaemia virus supports its absence in the Pacific Northwest of the United States. *J. Fish Dis.* 41:337–46
188. Sokolow SH, Nova N, Pepin KM, Peel AJ, Pulliam JRC, et al. 2019. Ecological interventions to prevent and manage zoonotic pathogen spillover. *Philos. Trans. R. Soc. B* 374:20180342
189. Johnson PT, de Roode JC, Fenton A. 2015. Why infectious disease research needs community ecology. *Science* 349:1259504
190. Mizuno CM, Ghai R, Rodriguez-Valera F. 2014. Evidence for metaviromic islands in marine phages. *Front. Microbiol.* 5:27
191. Pascoal F, Costa R, Magalhaes C. 2020. The microbial rare biosphere: current concepts, methods and ecological principles. *FEMS Microbiol. Ecol.* 97:fiaa227
192. Holmfeldt K, Solonenko N, Shah M, Corrier K, Riemann L, et al. 2013. Twelve previously unknown phage genera are ubiquitous in global oceans. *PNAS* 110:12798–803
193. Duhaime MB, Deng L, Poulos BT, Sullivan MB. 2012. Towards quantitative metagenomics of wild viruses and other ultra-low concentration DNA samples: a rigorous assessment and optimization of the linker amplification method. *Environ. Microbiol.* 14:2526–37
194. Duhaime MB, Sullivan MB. 2012. Ocean viruses: rigorously evaluating the metagenomic sample-to-sequence pipeline. *Virology* 434:181–86
195. French RK, Holmes EC. 2020. An ecosystems perspective on virus evolution and emergence. *Trends Microbiol.* 28:165–75
196. Fierer N. 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* 15:579–90
197. Jarett JK, Dzunkova M, Schulz F, Roux S, Paez-Espino D, et al. 2020. Insights into the dynamics between viruses and their hosts in a hot spring microbial mat. *ISME J.* 14:2527–41
198. Hillman BI, Cai G. 2013. The family *Narnaviridae*: simplest of RNA viruses. *Adv. Virus Res.* 86:149–76
199. Sieradzki ET, Morando M, Fuhrman JA. 2021. Metagenomics and quantitative stable isotope probing offer insights into metabolism of polycyclic aromatic hydrocarbon degraders in chronically polluted seawater. *mSystems* 6:e00245-21
200. Starr EP, Shi S, Blazewicz SJ, Koch BJ, Probst AJ, et al. 2020. Stable isotope informed genome-resolved metagenomics uncovers potential trophic interactions in rhizosphere soil. bioRxiv 2020.08.21.262063. **https://doi.org/10.1101/2020.08.21.262063**
201. Berjón-Otero M, Duponchel S, Hackl T, Fischer M. 2020. Visualization of giant virus particles using BONCAT labeling and STED microscopy. bioRxiv 2020.07.14.202192. **https://doi.org/10.1101/2020.07.14.202192**