

Review

RNA Viruses in Aquatic Ecosystems through the Lens of Ecological Genomics and Transcriptomics

Sandra Kolundžija ¹, Dong-Qiang Cheng ²  and Federico M. Lauro ^{1,2,*}

¹ Asian School of the Environment, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore; sandra001@e.ntu.edu.sg

² Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore; dongqiang.cheng@ntu.edu.sg

* Correspondence: flauro@ntu.edu.sg

Abstract: Massive amounts of data from nucleic acid sequencing have changed our perspective about diversity and dynamics of marine viral communities. Here, we summarize recent metatranscriptomic and metaviromic studies targeting predominantly RNA viral communities. The analysis of RNA viromes reaffirms the abundance of lytic (+) ssRNA viruses of the order Picornavirales, but also reveals other (+) ssRNA viruses, including RNA bacteriophages, as important constituents of extracellular RNA viral communities. Sequencing of dsRNA suggests unknown diversity of dsRNA viruses. Environmental metatranscriptomes capture the dynamics of ssDNA, dsDNA, ssRNA, and dsRNA viruses simultaneously, unravelling the full complexity of viral dynamics in the marine environment. RNA viruses are prevalent in large size fractions of environmental metatranscriptomes, actively infect marine unicellular eukaryotes larger than 3 µm, and can outnumber bacteriophages during phytoplankton blooms. DNA and RNA viruses change abundance on hourly timescales, implying viral control on a daily temporal basis. Metatranscriptomes of cultured protists host a diverse community of ssRNA and dsRNA viruses, often with multipartite genomes and possibly persistent intracellular lifestyles. We posit that RNA viral communities might be more diverse and complex than formerly anticipated and that the influence they exert on community composition and global carbon flows in aquatic ecosystems may be underestimated.

Keywords: marine RNA virus; viral metatranscriptomics; metaviromics; dsRNA sequencing; viral ecology; viral diversity



Citation: Kolundžija, S.; Cheng, D.-Q.; Lauro, F.M. RNA Viruses in Aquatic Ecosystems through the Lens of Ecological Genomics and Transcriptomics. *Viruses* **2022**, *14*, 702. <https://doi.org/10.3390/v14040702>

Academic Editor: Jens H. Kuhn

Received: 1 March 2022

Accepted: 23 March 2022

Published: 28 March 2022

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



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

1. Introduction

Viruses are vital parts of the marine ecosystem. They are major contributors to processes such as remineralization of carbon and nutrients and continuously regulate the diversity and abundance of complex microbial communities in the ocean, consisting of bacteria, archaea, and unicellular eukaryotes, protists [1–4]. Plankton protist communities largely consist of unicellular eukaryotic algae (also known as eukaryotic phytoplankton), heterotrophic flagellates and amoebas [5]. Viruses themselves are incredibly complex when it comes to genome type, replication, infection mode, morphology, and host preference. They infect cellular organisms in all three domains of life, as well as other viruses, and a single species can be infected by multiple viruses. In aquatic environments, eukaryotic phytoplankton is known to host very diverse viral communities [6,7]. Among aquatic bacteriophages, double-stranded (ds) DNA viruses are the most extensively studied and were thought to dominate prokaryotic virosphere [8]. Reports of single-stranded (ss) RNA and single-stranded (ss) DNA bacteriophages in environmental metaviromic and metatranscriptomic datasets [9–13] challenged this idea and suggested that bacteriophages have equally diverse genome architecture as eukaryotic viruses. On the basis of their genome architecture and replication strategy, viruses were classified in seven groups, now widely

known as the Baltimore classes, a classification still used in parallel with the official viral taxonomy [14,15]. However, the molecular taxonomy has shown that these groups are not monophyletic, and for RNA viruses, the transition between single-stranded and double-stranded genome types has occurred multiple times during evolutionary history [16].

All RNA viruses are a part of the newly established viral realm, *Riboviria*. Within this realm, kingdom *Pararnavirinae* consists of RNA viruses with the reverse transcriptase (RT) and kingdom *Orthornavirinae* contains RNA viruses with the RNA-dependent RNA polymerase (RdRp). Kingdom *Orthornavirinae* splits into five phyla on the basis of the phylogenetic analysis of RNA-dependent RNA polymerase (Figure 1) [14,16]. Phylum *Lenarviricota* includes ssRNA with positive (+) genome polarity, such as RNA bacteriophages from family *Leviviridae* and their eukaryote-infecting descendants, *Narnaviridae* and *Mitoviridae*, simple RNA viruses without a capsid that are transmitted vertically. Phylum *Pisuviricota* is sometimes called Picornavirus supergroup, and it includes five orders of (+) ssRNA viruses and one order of dsRNA viruses with bipartite genomes, *Durnavirales*. Phylum *Kitrinoviricota* includes only (+) ssRNA viruses infecting eukaryotes, mostly terrestrial plants, and animals. Phylum *Duplornaviricota*, the fourth branch, consists of dsRNA viruses infecting both eukaryotes and prokaryotes. Finally, the fifth branch forms phylum *Negarnaviricota* and includes all known RNA viruses with negative (−) genome polarity [14,16].

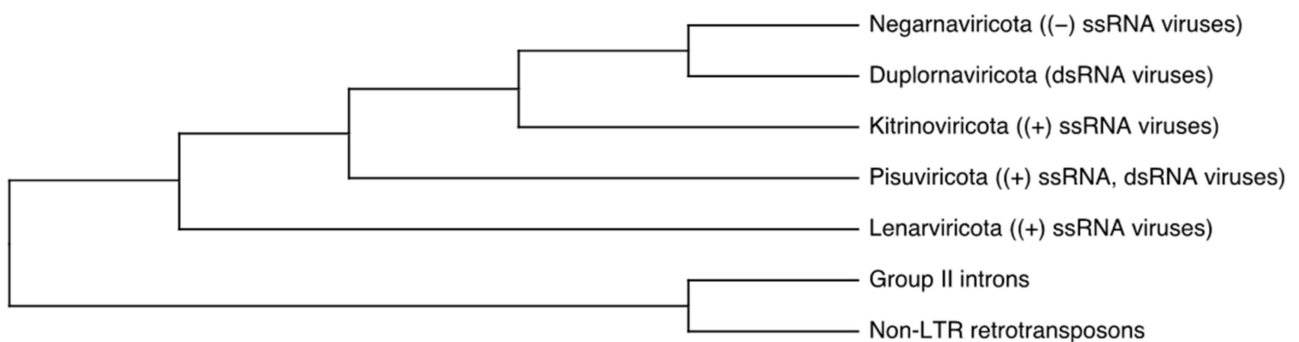


Figure 1. Phylogenetic analysis of RNA-dependent RNA polymerase (RdRp) of RNA viruses with reverse transcriptase (RT) used as an outgroup to root the tree. Five major branches have been assigned a phylum rank by International Committee on Taxonomy of Viruses (ICTV): Branch 1 = *Lenarviricota*, Branch 2 = *Pisuviricota*, Branch 3 = *Kitrinoviricota*, Branch 4 = *Duplornaviricota*, and Branch 5 = *Negarnaviricota*. Adapted from [16].

RNA viruses in the aquatic environments may infect a wide range of hosts with important consequences. RNA viral pathogens of marine mammals, seabirds, fish, crustaceans, and bivalves can decimate the populations of wild and farmed marine animals with large economic and ecological impacts [17]. RNA viruses have important ecological roles in regulating the structure of protist plankton communities in the oceans. Viral infection of diatoms and copepods, which are mostly infected with RNA viruses, can impact carbon remineralization via the “virus shunt” and carbon export via the “viral shuttle”, thereby altering the efficiency of the biological carbon pump (Figure 2) [1,4,18,19]. The analysis of eukaryotic virus composition showed that eukaryotic viral composition can predict carbon sequestration [18], possibly by affecting the flux through the viral shunt and viral shuttle. Yet, despite their fundamental roles and the enormous diversity of protistan species, only 12 RNA viruses and 41 DNA viruses [6,7,20] have been isolated from protistan hosts due to cultivation challenges.

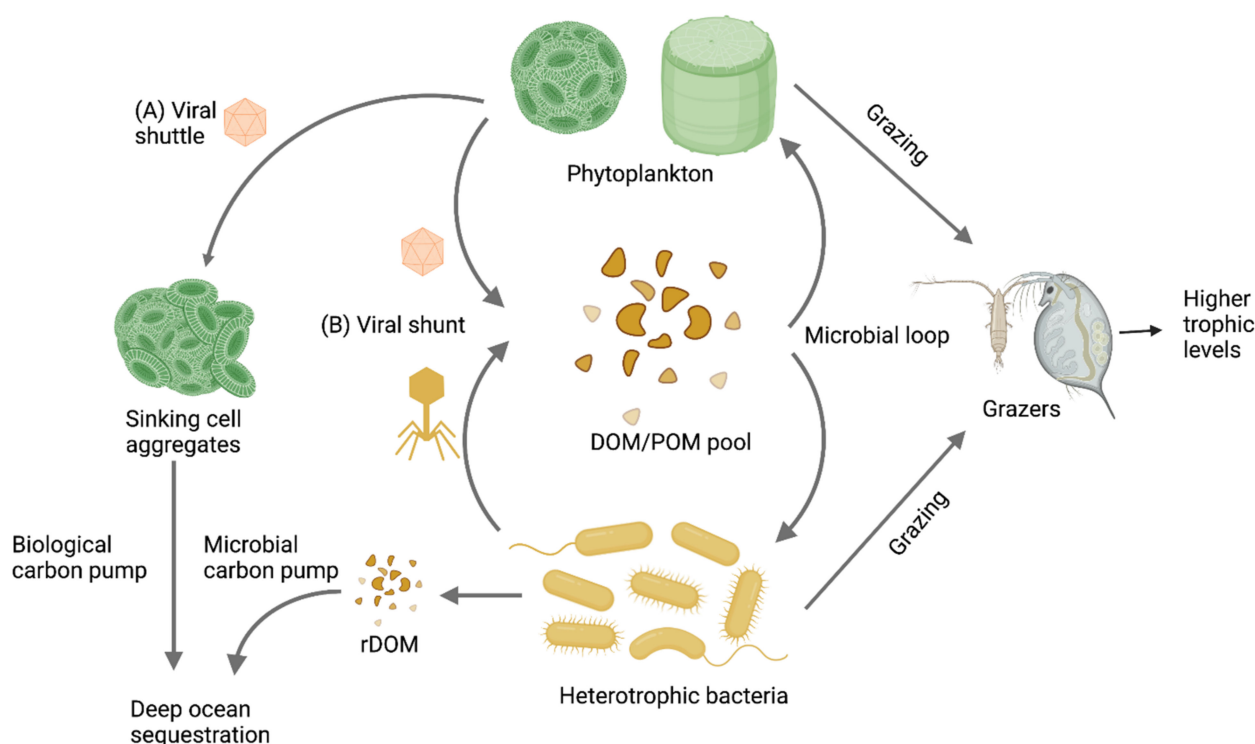


Figure 2. Marine viruses simultaneously control two processes in the carbon cycle: (A) the viral shuttle and (B) the viral shunt. In the viral shuttle, viral lysis of phytoplankton cells produces sticky aggregates with negative buoyancy that enhance the biological carbon pump by sequestering carbon in the deep ocean. In the viral shunt, viral lysis of cells has the opposite effect—it diverts the organic matter into a dissolved organic matter (DOM) pool that is rapidly and continuously recycled in the surface waters, preventing its sequestration or uptake by higher trophic levels. Adapted from [4]. Created in BioRender.com.

Routine methods for assessing viral abundance in aquatic environments, such as epifluorescence microscopy and flow cytometry, currently lack the sensitivity to detect the small genomes of RNA viruses [21,22]. Consequently, the abundance of RNA viruses in the viroplankton is presently unknown. Yet, there is some evidence that RNA viruses may exceed DNA viruses in abundance at times [23,24]. Marine RNA viruses are small in size, have a lytic cycle, and have very high burst sizes comparing to the DNA viruses, releasing 1000 to 10,000 particles per infected cell, which could affect their abundances [17,25]. Taken together, these results may indicate that abundances and ecological significance of RNA viruses may be underestimated in aquatic ecosystems.

Traditionally, RNA metaviromics has been the key method to study diversity and dynamics of aquatic RNA virus communities. A comprehensive review summarized genemarker studies targeting the RNA-dependent RNA polymerase gene and the first RNA metaviromics studies revealing consistent dominance of picorna-like RNA viruses in the marine environments [25]. Knowledge about the protist-infecting RNA viruses isolated through culturing was recently summarized by Sadeghi et al. [20]. Metatranscriptome sequencing greatly expanded known RNA viral diversity associated with invertebrate holobionts, bacteria, plants, and eukaryotes in the soil [9,26–29]. More recently, a growing number of studies have successfully explored the potential of metatranscriptome mining for RNA virus detection and discovery in aquatic ecosystems. To our knowledge, no reviews have systematically documented this progress from a metatranscriptomic point of view. Here, we describe the most commonly used meta-omic methodological approaches employed in aquatic RNA virology and examine their strengths and limitations. Throughout this review, we refer to meta-omic approaches as viromics, metatranscriptomics, and viral dsRNA sequencing.

We explore in detail 15 recent studies conducted in aquatic environments and 4 studies focusing on individual protists and macroalgae, in culture and in natural environments, and discuss how they have broadened our understanding of RNA viral diversity. The focus of this review is on RNA viruses, but if other viral groups were concurrently detected in a referenced study, they are discussed as well to get a more comprehensive overview of the viruses present in a specific environment. Finally, we address the remaining challenges and provide suggestions for future explorations of the ecological relevance of RNA viruses in aquatic ecosystems.

2. Methodological Challenges in the Study of Aquatic RNA Viral Communities

Two commonly used approaches for studying viral diversity in aquatic environments are metaviromics (i.e., sequencing of enriched virus particles) and data mining of environmental metagenomes and metatranscriptomes for virus discovery (Figure 3). These approaches are so successful that uncultivated viral genomes of DNA viruses now greatly outnumber cultivated viral isolates in the public databases [30].

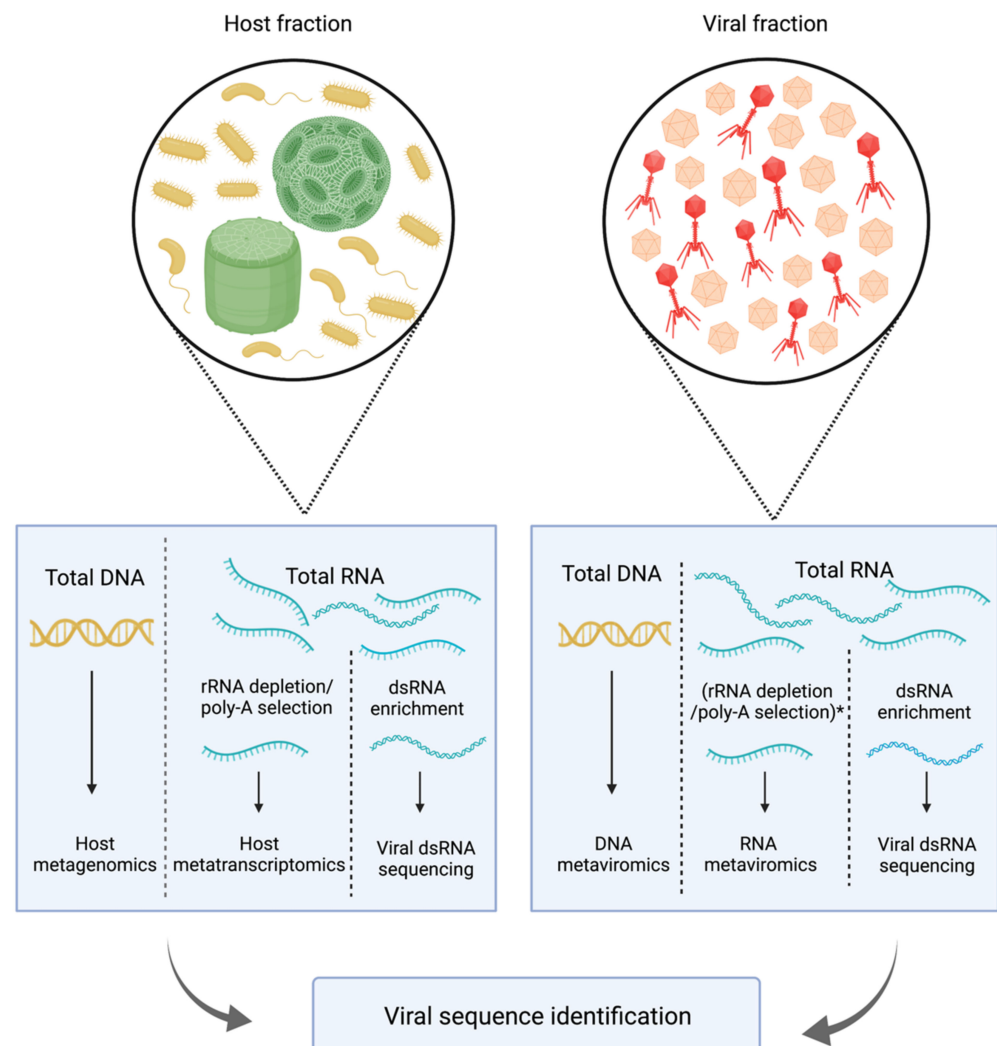


Figure 3. Most commonly used meta-omic sequencing approaches in marine virology. Host fraction ($>0.22 \mu\text{m}$) metagenomes and metatranscriptomes are mined for viruses, or DNA or RNA is extracted from the enriched viral fraction ($<0.22 \mu\text{m}$). * rRNA depletion or poly-A selection for RNA in the viral fraction is typically not performed due to low yields. An alternative approach that can enrich for replicative forms of RNA viruses inside the cells or for extracellular dsRNA viruses is dsRNA sequencing. Adapted from [30]. Created with BioRender.com.

Metaviromics (or viral metagenomics) is the method of sequencing extracellular total virus-like particles of either DNA or RNA viruses (Table 1; Figure 3). Enrichment of extracellular, free-floating virus-like particles is achieved through four steps: (i) removal of host cells and their nucleic acids by filtration through 0.22 μm filters, (ii) concentration of the viral fraction via tangential flow filtration [31] or iron flocculation [32], (iii) gradient centrifugation and/or nuclease treatment to eliminate contaminating non-encapsidated (host) nucleic acids [33], and (iv) nucleic acid extraction from virus-like particles [34]. Due to technical challenges, such as filtering large volumes of water and low nucleic acid yields after the extraction, studies using this approach typically target either DNA or RNA viruses. An advantage of the viral enrichment process is an increased proportion of viral reads resulting in improved genome assemblies that enable sensitive detection even of low-abundance viruses [33,34]. Limitations of this method include cellular contamination of the viral metagenomes, which can hinder the data analysis [35], and the fact that the filtration process will systematically remove viruses larger than 0.22 μm , such as giant viruses or long filamentous viruses such as archaeal *Lipothrixviridae* or bacterial *Inoviridae* that can be up to 3.5 μm long [13,30,36]. The latter technical limitation should not affect extracellular RNA viruses since the particle size is consistently around 30 nm, but the removal of the cellular fraction will exclude RNA viruses that do not possess a capsid, do not go through an extracellular state, and are vertically transmitted [37–39]. Conventional extraction and sequencing library preparation methods may not suit all genome types, limiting the diversity of viruses that can be detected. For example, ssDNA viruses require a special extraction method and sequencing approach [10] and are therefore systematically omitted from any study targeting dsDNA viruses, and dsRNA viruses may be underrepresented in the RNA viromes due to inefficient conversion to cDNA during the preparation of the sequencing library [40].

Table 1. Advantages and shortcomings of the meta-omic sequencing approaches used in marine viral ecology studies. Adapted from [28].

Method	Virus Nucleic Acid Detected	Shortcomings	Advantages
Metaviromics or viral metagenomics	RNA or DNA viral genomes in the extracellular stage	RNA viruses targeted separately than DNA viruses Needs special enrichment for dsRNA and ssDNA Large DNA viruses are filtered out High-burst-size viruses can be overrepresented	Enriched for viral sequences, better assembly
Metatranscriptomics	Transcripts of (+) and (–) ssRNA, dsRNA, ssDNA, dsDNA	High background of nonviral sequences Potentially fragmented assemblies Misses low-titre viruses Does not distinguish between (+) ssRNA viral genome and transcripts	Captures all types of DNA and RNA viruses simultaneously Captures active infection (for DNA viruses) Can capture RNA viruses without capsids
dsRNA sequencing	ssRNA as replicative intermediate dsRNA genomes	Not as effective for (–) ssRNA and DNA viruses Cellular metatranscriptomes are removed in the enrichment process	Enriched for RNA viruses Can be used for detection of both extracellular (<0.22 μm fractions) and intracellular RNA viruses

Cells of single-celled eukaryotes and prokaryotes from the environment are typically collected onto filters of different pore size, with 0.22 μm filter being the smallest filter size used. Typically, extracellular DNA and RNA viruses will pass through these filters, though aggregates or particle-associated extracellular viruses can still be retained. Viral sequences can be retrieved from both metagenomes and metatranscriptomes derived from the cellular fraction [18,41–51].

Metatranscriptomics, which is the sequencing of RNA transcripts from environmental microbial communities, is well suited to study the active viral infection in diverse viral groups with structurally different genomes (Table 1; Figure 3). As a part of the infection cycle, all seven Baltimore classes of viruses produce mRNA intermediate inside infected cells and can be simultaneously captured by metatranscriptomics [52,53]. The additional benefit of this approach is detection of RNA viruses that lack capsid and extracellular stages and are therefore absent from the viral particle fraction, such as *Narnaviridae*, *Hypoviridae*, and *Endornaviridae* [37–39]. The main preparations steps include (i) RNA extraction from cells collected on a filter, (ii) DNase treatment to remove the contaminating cellular DNA, and (iii) cellular mRNA enrichment by depletion of highly abundant ribosomal RNAs or by selection of polyadenylated RNA before the library preparation and sequencing. Ribosomal RNA is typically captured by complementary probes bound to magnetic beads, while polyadenylated RNA is typically selected by binding to oligo-dT probes [52,53]. The removal of ribosomal RNA is a less selective process because it retains all nonpolyadenylated and adenylated cellular RNA. In the context of viruses, this means that with rRNA depletion, RNA viruses that lack a poly-A tail and fragmented or degraded viral RNA genomes can be included in the sample, while during poly-A selection, only RNA viruses (and cellular mRNA) with intact poly-A tail will be retained. This is supported by a recent study demonstrating that rRNA-depleted libraries produced more viral reads, longer contigs, and greater viral richness compared to poly-A-selected libraries [47]. While these results will require additional experimental confirmation as different filter sizes were used for the two treatments, potentially biasing the data, similar results were observed in metatranscriptomes of marine invertebrates [54]. Limitations of this method include a high abundance of rRNA reads in the metatranscriptomes of cellular fractions even after the enrichment steps [55,56] and the fact that the majority of non-rRNA sequences in the cellular metatranscriptomes are of host origin, while viral reads typically constitute less than 1% of the sequenced reads (Table 2 and the references there).

Table 2. Basic geographic, ecosystem, and sampling information about the aquatic metaviromic (turquoise) studies, metatranscriptomic (coral) studies, and dsRNA sequencing (yellow) reviewed in this paper. Asterisk * denotes a temporal study in multiple locations.

Region	Sampling Location	Site Characteristics	Sampling Scheme	Temporal/ Spatial	Host or Viral Fractions Collected (Host-Fractionated/Unfractionated)	Relative Abundance of Viral Reads in the Metatranscriptome	Reference
Global	Multiple locations	Deep sea	133 locations	Spatial	Only viral fraction (RNA)	-	[57]
Polar	West Antarctic Peninsula	Highly productive coastal area	1 location	Temporal	Only viral fraction (RNA)	-	[24]
Polar and temperate	Multiple locations	Multiple site characteristics	11 locations	Spatial	Only viral fraction (RNA)	-	[58]
Temperate	Jericho Pier, Georgia Strait, Canada	Highly productive coastal area	2 locations	Spatial	Only viral fraction (RNA)	-	[59]
Temperate	Yangshan Harbour, Shanghai, China	Brackish coastal area	1 location	Spatial	Only viral fraction (RNA and DNA)	-	[11]
Subtropical	Kane'ohe Bay, Hawai'i	Coastal area	1 location	Spatial	Only viral fraction (RNA)	-	[60]
Subpolar	Honshu, Japan Jamstec cruise	Coastal and pelagic	5 locations	Spatial	Host: Unfractionated >0.22 µm dsRNA + ssRNA metatranscriptomes Viral fraction (RNA) dsRNA + ssRNA viromes	0.1% of ssRNA metatranscriptomes 1.3–36.6% of dsRNA metatranscriptomes	[56]
Polar	Chile Bay, Antarctica	Highly productive coastal area	1 location 2 samples, 3 weeks apart	Temporal	Host: Fractionated 8 µm–0.22 µm fraction	0.04–0.05% (rRNA depletion)	[48]
Temperate	Baltic sea Lake Torneträsk	Eutrophic, mostly brackish	11 location (2 depths)	Spatial	Host: Fractionated 200–3.0 µm 3.0–0.8 µm 0.8–0.1 µm + viral fraction (DNA)	3.2% (separate rRNA depletion and poly-A selection libraries)	[50]
Temperate	Narragansett Bay (NB) Quantuck Bay (QB), USA	Eutrophic (bloom) coastal	2 locations NB - 5 samples during 4 weeks QB - 3 samples within a week	Temporal *	Host: Fractionated QB 5–0.22 µm NB >5 µm	0.043–2.4% (poly-A selection)	[44]
Temperate	Lake Tai, China	Hypereutrophic (bloom) lake	9 locations 1 × monthly for 5 months	Temporal *	Host: Unfractionated >0.22 µm	0.02% (rRNA depletion)	[46]
Temperate	Owasco Lake Seneca Lake Cayuga Lake, USA	Mezotrophic to eutrophic lakes	3 locations 1 × monthly for 10 months	Temporal *	Host: Fractionated 5–0.22 µm	0.6% (rRNA depletion)	[45]

Table 2. Cont.

Region	Sampling Location	Site Characteristics	Sampling Scheme	Temporal/Spatial	Host or Viral Fractions Collected (Host-Fractionated/Unfractionated)	Relative Abundance of Viral Reads in the Metatranscriptome	Reference
Temperate	California Current, USA	Oligotrophic, with upwelling	1 location Every 4 h for 60 h	Temporal	Host: Fractionated >5 μm 5–0.22 μm	Not reported	[49]
Temperate	Quantuck Bay (QB) Tiana Beach (TB), USA	Eutrophic (bloom) coastal	2 locations QB 1 \times weekly for 10 weeks TB—1 \times weekly for 8 weeks	Temporal *	Host: Unfractionated >0.22 μm (for rRNA reduction) >1 μm (for poly-A selection)	0.33–0.53% in rRNA depleted libraries 0.02–0.023% in poly-A selected libraries	[47]
Global	Multiple locations	Multiple site characteristics	68 locations 2 depths	Spatial	Host: Fractionated 2000–180 μm 180–20 μm 20–5 μm 5–0.8 μm	0.0006% to 0.4% (poly-A selection)	[18,51]

Additional complexity in metatranscriptomics arises from viral genes being transcribed sequentially and in mono- or polycistronic transcription units, which can lead to uneven genome coverage [52]. The high background of nonviral sequences, combined with the heterogeneity of viral expression, also leads to more fragmented genome assemblies and is the biggest limitation of this approach [61]. Furthermore, for RNA viruses whose genome serves as mRNA, it is not possible to distinguish the viral genome from viral transcripts in the metatranscriptomic datasets.

Standard methods for reverse transcription used in RNA sequencing are inefficient in transcribing dsRNA molecules as the presence of complementary strand will prevent the binding of random primers. Therefore, compared to ssRNA viruses, the dsRNA viral lineages may be underrepresented in both metatranscriptomic and viral RNA metagenomic datasets [40]. An approach that might alleviate this problem is dsRNA sequencing (Table 1; Figure 3). Presence of long dsRNA in cells is a unique hallmark of viral infection. Double-stranded viral RNA in cells originates from genomes of dsRNA viruses or from replicative intermediates of ssRNA viruses, with the exception of retroviruses. For DNA viruses, bidirectional transcription (sense and anti-sense direction) of overlapping regions can lead to complementary transcripts that will form a dsRNA duplex. However, only dsRNA and (+) ssRNA viruses produce dsRNA in amounts detectable by immunofluorescence [62]. Separation of viral dsRNA from total RNA via cellulose chromatography or more recently through immune-based capture, followed by dsRNA sequencing, has been widely used for virus detection and identification in plants and fungi and has led to the discovery of many novel viruses [29,63,64]. Immuno-based capture of dsRNA with anti-dsRNA antibodies was used to pull down multiple (+) ssRNA viral genomes with high recovery (31–74% viral reads) from singular plant samples, including the near-complete genome of an unknown virus [63].

Recently, a novel dsRNA sequencing approach, fragmented and primer-ligated dsRNA sequencing (FLDS), has been developed and applied to marine environmental samples, demonstrating the method's potential to capture RNA viral diversity that previously went undetected [56]. After conventional total RNA extraction from either viral concentrates or cells, extracted total RNA is enriched for dsRNA via cellulose chromatography. Primers are ligated to the ends of fragmented dsRNA, which is then denatured, and reverse transcription into cDNA with complementary primers is initiated from the ends of the molecules. RNA is degraded with RNase and a complementary DNA strand synthesized by PCR. Using the FLDS method, on average, 11.3–36.6% of reads obtained from the cellular fraction may be identified as viral as opposed to 0.1% using the classic RNA sequencing approach with rRNA depletion. Similar trends were observed for the viral fraction, where viral reads comprised 35% of total reads with the FLDS method, and only 5% with classic RNA sequencing. A very low number of reads were assigned to rRNA [56]. Overall, virus enrichment through dsRNA purification might be an interesting avenue to explore for research of aquatic RNA viral communities, followed by either standard dsRNA sequencing used in plant virus detection [64] or the FLDS dsRNA sequencing method [56].

In synthesis, there is currently no method that will comprehensively capture all viral diversity. The comparative study of RNA viromes and metatranscriptomes has highlighted tremendous differences between the fractions as well as between different methodological approaches, which will be further discussed below [56]. This observation is not unique to RNA viruses because, for example, the genomes of uncultivated DNA viruses (>10 kb) in the Tara Oceans dataset recovered from virus-enriched fractions and by metagenome mining exhibit huge differences in diversity with an average of 75% and 90% unique sequences in the microbial and viral fractions, respectively [30]. This illustrates that every detection method comes with its biases and using only one approach may show an incomplete or biased snapshot of natural viral communities, underscoring the importance of using multiple approaches.

3. Environmental RNA Metaviromics: Lytic Positive-Sense ssRNA Viruses Dominate Pelagic but Not Benthic RNA Viral Assemblages

The first environmental viral metagenomic surveys consistently showed that most sequences in the RNA viromes were unknown. Among known sequences, typically more than 90% of RNA viral sequences identified as (+) ssRNA viruses from order *Picornavirales* in temperate, tropical, and polar oceans, often clustering with known RNA viruses that infect diatoms. Very few sequences of dsRNA viruses were typically recovered, and RNA phages or negative-sense RNA viruses were absent. Only five RNA metaviromic studies in the water column have been performed to date, predominantly in coastal productive environments, with only one study having an extensive geographical distribution [11,24,58–60]. Recently, the first RNA metavirome study focusing on deep sea sediments was published as a preprint (Figure 4; Table 2) [57].

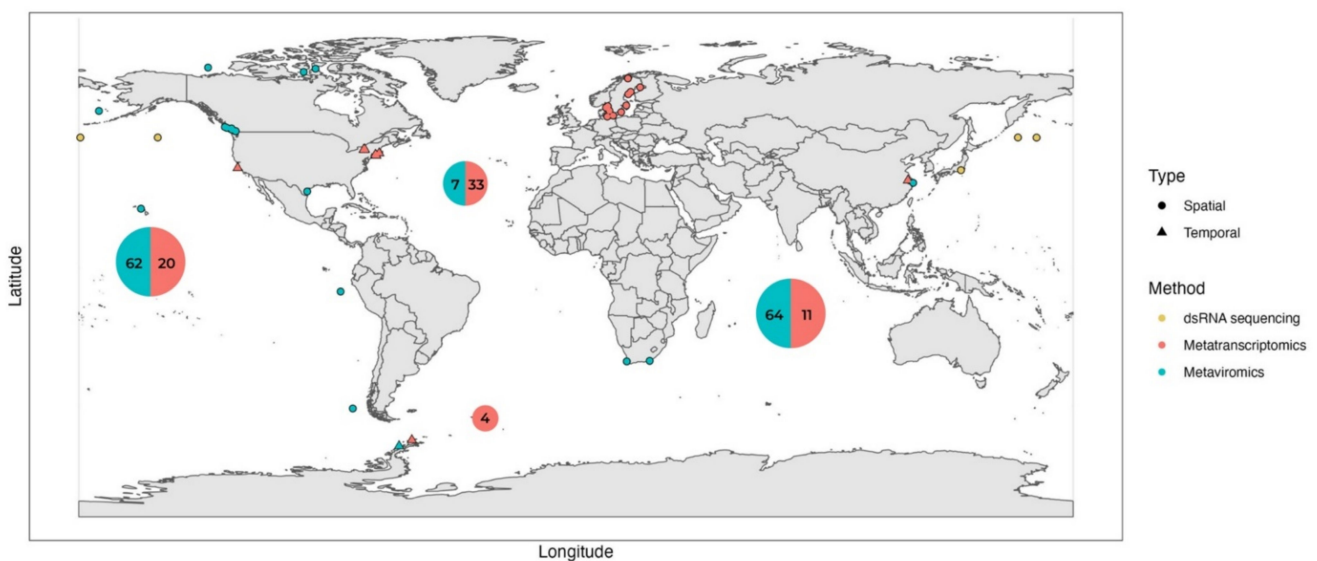


Figure 4. Geographic locations of aquatic metaviromic (blue), metatranscriptomic (coral), and dsRNA sequencing (yellow) studies focusing on RNA viruses reviewed in this paper. Spatially focused studies are represented with circles, and temporally focused studies are represented with triangles. Two global surveys with extensive geographic coverage, Dayang No. 1 cruises and Tara Oceans Expedition [51,57], are represented by large bubbles in each of the five oceans. The size of the bubble is proportional to the total number of samples collected from each ocean, with numbers of benthic virome samples from the Dayang No. 1 cruises [57] indicated in blue, and number of pelagic metatranscriptome samples from the Tara Oceans Expedition [51] in coral. Samples from the Mediterranean Sea were included in the counts of the Atlantic Ocean.

In RNA metaviromics study of coastal, temperate waters of British Columbia, 98% of the identifiable reads represented (+) ssRNA viruses [59]. In the Jericho Pier library, viral sequences from the order *Picornavirales* were most abundant, mainly consisting of families *Marnaviridae* and *Dicistroviridae*. Two complete viral genomes (JP-A and JP-B) were phylogenetically similar to the diatom-infecting *Rhizolenia setigera* RNA virus RsRNAV and could have a protist (diatom) host. In the Strait of Georgia library, the majority of viral sequences fell into *Tombusviridae*, a family of viruses from the phylum *Kitrinoviricota* [59]. Tropical seawater viral RNA metagenomes of Kaneohe Bay, Hawaii again confirmed the dominance of the order *Picornavirales*, which encompassed 95% of the assignable reads, with a minority of reads being identified as dsRNA viruses. Similarly, the full genomes recovered from the dataset were phylogenetically related to diatom-infecting viruses, or viruses from families *Marnaviridae* and *Dicistroviridae*, consistent with the previous RNA metaviromic study in temperate waters [60].

The trend continued with the overwhelming majority of identifiable RNA viral sequences (97.8%) from polar RNA viral metagenomes collected during a spring diatom bloom in West Antarctica Peninsula [24] being classified as (+) ssRNA from the order *Picornavirales*. Out of five full genomes assembled from the dataset, three (PAL128, PAL156, and PAL473) clustered together with diatom-infecting RNA viral isolates or marine RdRp sequences, and two (PAL 438 and PAL_E4) contained highly divergent RdRps. There was a smaller proportion of viral sequences identified as *Marnaviridae* or *Dicistroviridae* compared to the tropical RNA viral metagenomes of Kaneohe Bay. Instead, the relative abundance of diatom-infecting RNA viruses from the genus *Bacilarnavirus* surpassed 90% in three samples or reached around 50%, with unclassified *Picornavirales* taking up the rest. The relative abundance of RNA viruses in the total viroplankton was estimated to range 2–79% in samples on the basis of nucleic acid content of gradient-purified fractions [24].

The largest spatially extensive study of RNA viral metagenomes in the water column to date explored the biogeography of the six marine picorna-like RNA viruses: three complete genomes of picorna-like viruses recovered in this metagenomic study itself [58], two complete picorna-like genomes recovered in a previous metagenomic study [59], and one genome originating from an isolate of HaRNAV, which is genetically more divergent and infects a raphidophyte. Out of the three full genomes recovered in this study, two of them have the highest similarity to PAL 156 genome and one has the highest similarity to PAL 128, the full-length genomes recovered in the temporal study of Antarctic diatom bloom [24] that are thought to be infecting diatoms. Reads from global metagenomic datasets comprising marine, freshwater, and reclaimed water RNA viromes were recruited to these genomes, revealing high geographical and temporal variation of RNA viruses. On a temporal scale, samples collected within a year at the same sample site, Jericho Pier in British Columbia, differ remarkably in distribution of six genomes. In 2013, raphidophyte-infecting HaRNAV was the most abundant virus at Jericho Pier, and in 2014, it was diatom-infecting BC-2 and BC-3 viruses, suggesting that the temporal distribution of RNA viruses in one location is controlled by biotic factors such as host availability. On a spatial scale, the diatom-infecting JP-A was the most abundant genome type off the coast of South Africa, and the raphidophyte-infecting HaRNAV genome dominated the samples from Bering Sea, Laguna Madre in the USA, and samples from Peru and Chile. In the samples from Queen Charlotte Strait and Johnstone Strait in Canada, reads were more equally distributed between diatom-infecting JB-A and JP-B and raphidophyte-infecting HaRNAV genomes. Plenty of viral sequences mapped to the six genomes with low amino acid identity, demonstrating ubiquity and richness of aquatic picorna-like viral communities worldwide. Even if the study did not explore the full diversity of RNA viruses at each sampling location, but instead focused on six selected picorna-like viruses, it demonstrated how dramatically the abundance of certain RNA viral groups can vary between locations [58].

More recently, a study recovered 4593 near-full-length RdRp sequences clustered within 2192 clusters at 75% amino acid identity (AAI), in one water sample from a single location in China [11]. Consistent with previous studies, no (+) ssRNA enveloped viruses or (−) ssRNA viruses have been detected and only six RNA viruses were identified as dsRNA viruses, supporting the idea that positive-sense RNA viruses are dominant in the marine environment. In contrast to previous studies, all three phyla of (+) ssRNA viruses were present in the samples, with the order *Picornavirales* constituting only around 26% of RdRp sequences in the sample. Within the phylum *Pisuviricota*, the largest number of RdRp sequences ($n = 854$) grouped with the aquatic picorna-like viruses (order *Picornavirales*) encompassing *Marnaviridae* and other protist-infecting viruses. *Protopotyviruses*, ancestors of complex plant-infecting *Potyviruses*, formed a new clade that was not previously detected in marine environments. Within the phylum *Kitrinoviricota*, many highly divergent RdRp sequences formed a new, phylogenetically well supported clade with previously “orphan” RdRPs of RNA viruses found in marine invertebrates and wastewater, with multiple new orders and families such as *Weiviruses*, *Yanvoiviruses*, *Zhaoviruses*, and *Shangaiiviruses*, significantly expanding the presence of this phylum in aquatic environ-

ments, previously limited to detection of *Tombusviridae* in one metagenomic library. Within the phylum *Lenarviricota*, a high number of sequences had similarity to the *Leviviridae*, (+) ssRNA bacteriophages, and the simple eukaryote-infecting *Ourmiaviridae*, documenting for the first time the presence of (+) ssRNA bacteriophages as members of marine RNA viral communities. Capsidless endogenous RNA viruses, such as *Mitoviridae* and *Narnaviridae*, were absent from the extracellular RNA virome as they would have been removed together with cells during filtration. The largest RdRp cluster was affiliated with the *Marnaviridae*, viruses of diatoms and stramenopiles and, together with the use of the alternative (ciliate) genetic code, might suggest that a considerable fraction of RNA viruses infect protists [11].

The high numerical abundance of viral particles and the tremendous impacts of viral infection on bacterial and archaeal communities in marine sediments have been well documented [65,66]. As with the studies of viruses in the water column, research has been predominantly focusing on dsDNA viromes [67–70], with reports of ssDNA [71] and RNA viruses [57] significantly lagging. A recent preprint by Zhang et al. [57] is the only available RNA metaviromic study in marine sediments, and with 133 samples and almost 3 billion of raw reads, is the largest marine RNA virome dataset to date. Sediments were collected between 2010 and 2018 in a wide range of benthic deep seafloor environments around the globe: hydrothermal vents, cold seeps, sea mounts, and ocean basins at depths between 1000 and 6000 m. The sequenced RNA viromes were searched for marker genes for both kingdoms *Orthornavirinae* and *Pararnavirinae*, RdRp genes, and RT genes, respectively. Most of the identifiable viral RNA sequences were classified as RT-containing *Retroviridae* and *Metaviridae*. Out of RNA viruses with the RdRp gene, the two most abundant families were surprisingly dsRNA viruses from families *Totiviridae* and *Cystoviridae*, in contrast with water column RNA viromes that are typically dominated with ssRNA viruses. Single-stranded RNA viruses from family *Leviviridae* were the most abundant ssRNA viruses detected, supporting the idea that RNA bacteriophages are important members of marine RNA viral communities in the water column as well as in sediments. However, most of the viral sequences could not be assigned to known viral groups, suggesting that the deep sea might represent a rich reservoir of unexplored viral diversity. Alternatively, this might just reflect the current limitations of the available software programs in assigning and classifying viral sequences from environmental samples [57].

It is still unclear what fraction of the large differences in the estimates of RNA viral diversity between different studies and different habitats is real rather than a result of methodological biases. RNA viruses from order *Picornavirales* have very high burst sizes [17], and it has been shown that viruses with higher burst sizes can be overrepresented in enriched samples [30]. Consistent with this, in the virus-enriched samples from a diatom bloom from polar waters, the diversity was highest in samples with low abundance of RNA viruses [24]. Similarly, in the virus-enriched samples from temperate waters, the reads in the two libraries were dominated by a few contigs: in Jericho Pier library, 66% of the library assembled in four contigs, and in Strait of Georgia library, 59% of reads assembled into one viral contig [59]. It is evident that the study by Wolf et al. [11] uncovered a staggering diversity of (+) ssRNA viruses in the aquatic environments, broadly expanding the typical range of viral diversity in the first studies. This study had much deeper sequencing depth, as well as a different library preparation method. In the other pelagic viromic studies [24,58–60], RNA was preprocessed by sequence-independent, single-primer amplification (SISPA) prior to sequencing, which can lead to overrepresentation of the dominant sequences [72]. More recently, the characterization of the benthic RNA virome by Zhang et al. [57] also had a greater sequencing depth than that of earlier studies, but the results might have been skewed by biases inherent with the use of a Phi29 multiple displacement amplification (MDA) method during library preparation [73,74]. Whether the detected differences in RNA viral diversity are due to potential amplification bias of SISPA libraries and MDA amplification or due to sequencing depth or some other factor will be clarified as more studies become available and published.

4. Environmental (Viral) Metatranscriptomics: Sampling with Size Fractionation Improves Metatranscriptome Resolution and Uncovers the Role of ssRNA Viruses

Metatranscriptomics can capture the active viral infection of diverse viral groups and has been increasingly used in studies of RNA virus community diversity and ecological dynamics in the last few years. Most studies were performed in coastal, productive environments, either on temporal or on spatial resolution (Figure 4) and employing different filtration or rRNA removal approaches. Typically, protist and bacterial communities are collected either as “whole” microbial fraction without prefiltration or separated into size fractions using serial filtration, which may allow for higher resolution (Table 2 and the references there).

One of the first studies that used environmental marine metatranscriptomes followed active infections of diverse viral groups (dsDNA, ssRNA, dsRNA) in two highly productive sites in the US, Narragansett Bay and Quantuck Bay. Seasonal diatom blooms occur in Narragansett Bay, and harmful brown tide blooms caused by unicellular eukaryotic algae *Aureococcus anophagefferens* take place in Quantuck Bay. Active infections of eukaryotic viruses were tracked during the initiation, peak, and demise of the bloom in Quantuck Bay and in the steady-state, non-bloom conditions in Narragansett Bay using marker gene approach [44]. Contigs were searched for marker genes, including the major capsid protein (MCP) for giant viruses, RdRp for RNA viruses, and the viral replicase for ssDNA viruses. Giant viruses from families *Mimiviridae* and *Phycodnaviridae* were continuously present in both locations, with two distinct dynamics. Some giant viruses were steadily expressed across sampling dates, and others exhibited sharp peaks in abundance, or a “bloom and bust” type of expression. In the non-giant virus community, between 62% and 74% of contigs in both bays were unclassified (+) ssRNA viruses in the order *Picornavirales*, followed by classified families *Picornaviridae*, *Secoviridae*, and *Dicistroviridae* from the same order. Few dsRNA viruses resembling *Totiviridae*, *Partitiviridae*, and *Hypoviridae* were detected. The viral community (excluding giant viruses) in Narragansett Bay was steadily dominated (>90%) by *Picornavirales*. In Quantuck Bay, both ssDNA and unclassified *Picornavirales* were present in the *Aureococcus anophagefferens* bloom samples in equal relative abundances, around 50%. Unclassified *Picornavirales* viruses took over (>90%) when bloom disintegrated, possibly infecting other protists in the bloom succession. Co-occurrence analysis between viruses and host marker genes found that most of the hosts might be phytoplankton or heterotrophic protist, but also fungi [44]. Another temporal study of the RNA viral community of three temperate lakes detected RNA viruses resembling families *Picornaviridae*, *Virgaviridae*, and *Reoviridae* with putative protist or invertebrate hosts. The low number of contigs retrieved might be due to low sequencing depth of 1–5 million reads per library [45].

Multiple metatranscriptome size fractions and a metagenome picoplankton size fraction from the Tara Ocean expedition dataset were screened for eukaryotic viruses using marker gene approach: PolB for giant viruses, RdRp for RNA viruses, and RepB for ssDNA viruses [18]. Interestingly, with 975 sequences, RdRp was the most abundant marker gene detected in the metatranscriptomic dataset, followed by 388 PolB sequences of giant viruses and 299 ssDNA viruses. An additional 3846 PolB marker gene sequences of giant viruses were detected in the metagenome picoplankton size fraction. Giant viruses were identified as *Mimiviridae* (mostly algae-infecting), *Phycodnaviridae*, and *Iridoviridae*. Major RNA viral groups in the dataset were *Picornavirales* as well as *Narnaviridae*, *Tombusviridae*, *Virgaviridae*, and dsRNA *Partitiviridae*. Giant viruses and RNA viruses that were positively correlated with carbon export efficiency were abundant in the oceanic regions with high carbon export efficiency, such as the Indian Ocean and the Mediterranean Sea. Hosts of these giant viruses are predicted to be picoplankton (*Mamiellales*) and haptophytes. Hosts of RNA viruses were either copepods or diatoms from the order *Chaetocerales*, both groups being important contributors to the biological carbon pump [18].

More recently, the active viral community within a *Microcystis*-dominated harmful algal bloom in a temperate Lake Tai in China was characterized using the same viral marker

genes approach [46]. As expected, most of viral transcripts (48%) present in the metatranscriptome assemblies consisted of bacteriophages infecting *Microcystis*, and RNA viruses constituted the remaining 42% of viral transcripts. Though very diverse, giant viruses take up only 8% of the transcripts. Surprisingly, less than 1% of detected viral transcripts was identified as ssDNA bacteriophages or dsDNA bacteriophages infecting heterotrophic bacteria, though *Microcystis* bloom is typically accompanied with an abundant heterotrophic bacterial community. Interestingly, this is one of the first metatranscriptome studies reporting substantial diversity and abundance of dsRNA viruses. Of the total RNA viral contigs, 28% were identified as dsRNA, mainly *Partitiviridae* and *Picobirnaviridae*, and 72% were ssRNA viruses of the *Picornavirales* order, largely *Marnaviridae* and *Discistroviridae* [46]. This notable increase of detected dsRNA viruses might result from using rRNA depletion rather than poly-A selection [47].

Both metagenomes and metatranscriptomes of the microbial community collected during a diatom bloom succession at Chile Bay in the West Antarctic Peninsula were mined for viruses [48]. Viral reads in the cellular metagenomes of the low-chlorophyll sample were dominated by dsDNA viruses: 82% of viral reads belonging to *Myoviridae*, 9% to giant viruses infecting eukaryotic phytoplankton from the family *Phycodnaviridae*, and 8% to the filamentous ssDNA bacteriophages from family *Inoviridae*. In high-chlorophyll cellular metagenome samples collected during the bloom, the proportion of giant viruses rose to 93% of the DNA reads and previously dominant *Myoviridae* dropped to 3%. In high-chlorophyll metatranscriptomes collected during the bloom, the proportion of (+) ssRNA viruses from family *Picornaviridae* increased to 38%, mostly composed of PAL E4 and PAL 156 viruses, both lytic viruses, and the latter closely related to diatom-infecting RNA viruses from the genus *Bacilarnavirus* [48]. A similar RNA viral community was detected in RNA viromes (<0.22 μm fraction) from Palmer Station at the West Antarctic Peninsula during a summer diatom bloom in 2016 [24]. Taken together, these metatranscriptomic studies show that during a bloom, viral communities that are typically dominated by prokaryotic viruses can shift to eukaryotic viruses, further supporting the ecological relevance of lytic DNA and RNA viruses in bloom disintegration. A metatranscriptome study of virus–host dynamics of the coast of California [49] provided evidence of RNA and DNA viruses tightly controlling protist communities even in non-bloom conditions. After serial filtration, the small fraction showed enrichment for bacteriophages of abundant bacterial groups, while a large fraction was enriched for viruses infecting large phytoplankton, the majority (74%) being dsDNA viruses. The large fraction (<5 μm) was also enriched in fungi, demonstrating fungal importance in the surface ocean as well as their potential as hosts, as they are known hosts of mostly dsRNA viruses. Metatranscriptomic reads from a 60 h sampling study with a 4 h resolution [49] were mapped to the reference genomes of selected RNA and DNA viruses and their host to resolve the diel dynamics of virus–host pairs. An interesting observation from this study was that dsDNA transcripts of dsDNA viruses and their putative host peaked at the same time, without a temporal lag. On the contrary, ssRNA viruses followed a typical “kill the winner” scenario, with the viral transcripts’ peak lagging after the putative host peak. Multiple peaks of viral transcripts of three selected ssRNA viruses and one dsRNA virus during the two sampling days indicated that multiple “boom and bust” cycles of RNA viruses might be happening within short temporal scales. For example, diatom-infecting *Chaetoceros* sp. RNA virus 02 exhibited only one peak, and heterotrophic labyrinthid protist-infecting *Aurantiochytrium* single-stranded RNA virus 01 peaked five times during the sampling period. This suggests that RNA viruses are important players in regulating the microbial host communities on hourly scales in steady-state systems not perturbed with phytoplankton blooms [49].

A spatial metatranscriptome study in the Baltic Sea at 11 locations spanning a salinity gradient in the brackish Baltic Sea and one freshwater lake identified a wide distribution of ssDNA and RNA populations and high transcriptional activity of fish viruses in different microbial size fractions [50]. Both ssRNA and dsRNA viruses infecting marine protists were detected throughout the dataset, as well as negative-sense RNA viruses from

Mononegavirales and a high proportion of *Retroviridae*-like sequences. At two high-salinity sites, DNA viruses infecting the protist *Ostreococcus* outnumbered the bacteriophages, amounting to 20%–40% of total viral transcripts, which indicated an ongoing infection. The sample collected in the lake Tornetrask contained virus transcripts mostly belonging to ssDNA bacteriophages from family *Microviridae*, and two samples had a large proportion of ssDNA viruses infecting pigs that could have originated from nearby animal farms. Collectively, these results provide a valuable snapshot into virus diversity at multiple locations and demonstrate that low-abundance viruses can rise to high abundance if conditions allow [50].

Studies using the size-fractionated approach show differential enrichment of viral groups in different size fractions [49–51]. In the Baltic Sea spatial metatranscriptome dataset and the Tara Ocean Expedition spatial metatranscriptome dataset, RNA viruses were more abundant in libraries originating from the larger cell fractions and are thought to be infecting large unicellular eukaryotes [50,51]. Giant viruses (including *Phycodnaviridae*) and bacteriophages dominated the viral portion of the Baltic Sea metatranscriptome (50–95% viral transcripts) [50]. Similar trends of large DNA viruses dominating viral portions of cellular metatranscriptomes were observed in the Tara Oceans expedition, with 86% of viral genes present originating from giant viruses [51]. In conclusion, these results suggest that size fractionation may improve the resolution and capture a greater diversity of less-abundant viruses. However, care should be taken to avoid using a defined size fraction without a clear understanding of the intended viral targets, because this might lead to more biased results compared to unfractionated samples.

5. Environmental dsRNA Sequencing: The Enrichment of dsRNA from Marine Samples Greatly Expands the Diversity of dsRNA Viruses

The earlier metatranscriptome studies described above recovered mostly ssRNA viruses and very few dsRNA viruses, in agreement with the diversity found in extracellular RNA viromes. A combined metatranscriptomics and RNA metaviromics study compared RNA viruses in the cellular and purified viral particle fractions in four pelagic sampling sites and one coastal sampling site off the coast of Japan [56]. For both cellular and viral fractions, the total RNA was separated by chromatography in ssRNA and dsRNA fractions. The double-stranded RNA fraction contains dsRNA genomes and double-stranded replicative intermediates of ssRNA genomes, and it is enriched for viruses compared to the rRNA-depleted ssRNA fraction that contains mostly cellular mRNA and in which viral ssRNA genomes or transcripts are in minority. Single-stranded RNA and dsRNA were sequenced separately, the latter using the newly developed method, fragmented and primer-ligated dsRNA sequencing (FLDS), resulting in four datasets for each sampling station: cellular ssRNA metatranscriptome, cellular dsRNA metatranscriptome, ssRNA virome, and dsRNA virome. A total of 1270 RNA viral contigs were recovered in the study, mostly originating from dsRNA-enriched metatranscriptomes and dsRNA-enriched viromes (>1000), underlining the efficiency of the dsRNA enrichment procedure compared with the traditional sequencing approach for obtaining viral reads. Due to low number of reads, ssRNA metatranscriptomes were excluded from the analysis. Viral families detected in dsRNA metatranscriptomes were predominantly endogenous ssRNA viruses from family *Narnaviridae*, and dsRNA viruses from families *Picobirnaviridae* and *Reoviridae*. The ssRNA extracellular virome of the coastal site was dominated by (+) ssRNA viruses from the order *Picornavirales*, as reported by other studies [24,58–60]. The other sampling sites had too few reads in this fraction for a meaningful comparison. Parallel analysis of the dsRNA virome from the same coastal sampling site revealed an almost equal number of reads belonging to *Picobirnaviridae* and *Reoviridae*, dsRNA viruses with a capsid and an extracellular phase. Many contigs resemble viruses with protist hosts such as diatom colony-associated dsRNA virus-1 (DCADSRV-1) [75] and *Micromonas pussila* reovirus [76]. There are two possible explanations for why FLDS sequencing detects an increased number of dsRNA viruses. The first one is that denaturation of dsRNA before transcription

allows for efficient transcription of dsRNA genomes, as they are not efficiently amplified during standard metatranscriptomic library preparation [40]. The second is that there is an amplification bias of the FDLS method, such as that shown for SISPA libraries, which are amplified on a similar principle. Whichever the case, these viruses are present in the aquatic environment and had not been previously detected with standard metatranscriptomic methods. This remains the only study that has explored dsRNA sequencing as a way of studying aquatic RNA viral communities, and it demonstrates the potential of the method to uncover novel RNA viruses and contribute to a clearer picture of the aquatic RNA viral diversity.

6. Holobiont Metatranscriptomics: Marine Macroalgae and Cultured Marine Protists Reveal the Broad Distribution of Non-Lytic Strategies in Marine RNA Viruses

There have been several new viral discovery studies in metatranscriptomes of (mostly) axenic protist cultures that have discovered an array of small dsRNA and ssRNA viruses, often with multipartite genomes that coexist with the host and do not cause lysis of the host cell. These persistent infections can be chronic or latent (silent). The advantage of viral discovery in cultured metatranscriptomes over broad environmental metatranscriptomics surveys is that the host can be assigned with certainty and low-titre viruses, which are frequently missed by environmental metaviromics and metatranscriptomics, can be picked up.

The first report of RNA viruses in the cultured metatranscriptomes came from six species of phytoplankton belonging to two distant phylogenetic groups, green algae (*Chlorophyta*) and chlorarachniophytes [55]. They recovered 18 novel RNA viruses with very low AAI (27–38%). Most viruses were associated with the green algae *Ostreobium* sp., comprising dsRNA viruses falling into families *Partitiviridae* and (+) ssRNA *Mitoviridae*—non-encapsidated small RNA viruses with tiny genomes. Fewer viruses were detected within chlorarachniophytes, possibly due to the clade being divergent and poorly characterized itself, and the highly divergent RdRps might have gone undetected. Additionally, segmented (+) ssRNA viruses from family *Virgoviridae* were detected along with the first detection of a (–) ssRNA virus resembling family *Bunyaviridae* in phytoplankton culture, though the confirmation that the chlorarachniophyte algae are the true host requires additional work [55]. This came as a surprise because both environmental metatranscriptomes and virus isolates showed evidence that green algae are predominantly infected by giant DNA viruses and 85% of isolated RNA viruses were associated with diatoms [6,55].

Recently, the metatranscriptomes of cultured phytoplankton obtained from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) were screened for the presence of viral RNA-dependent RNA polymerase hallmark gene [77]. Thirty new RNA viral species with highly divergent RdRps were observed, on average only sharing 35% AAI with known RNA viruses. RNA viruses were observed for the first time in protist families such as *Haptophyceae* and *Chromeraceae*, *Xantophyceae*, and *Bolidophyceae*, which were previously thought to only host dsDNA viruses. One-third of the newly described viruses belong to (+) ssRNA non-encapsidated viruses from families *Narnaviridae* and *Mitoviridae*, known for persistent lifestyles [37]. The double-stranded RNA viruses discovered were dominated by viruses resembling *Totiviridae* that were previously considered to be fungal pathogens. A member of *Mononegavirales*, a negative-sense RNA virus, was detected in the *Pseudo-Nitzschia* culture, but again, host assignment requires additional experimental work [77].

Two studies screened diatom holobionts and marine macroalgae for RNA viruses with the FDLS dsRNA sequencing approach [75,78]. The first one unveiled 22 and the second six full-length RNA viral genomes, encompassing five families of dsRNA viruses (*Totiviridae*, *Picobirnaviridae*, *Cystoviridae*, *Partitiviridae*, and *Endornaviridae*) and four families of ssRNA viruses (*Flaviviridae*, *Narnaviridae*, *Virgaviridae*, and *Hypoviridae*). Typically, the viral sequences were very divergent, exhibiting less than 50% AAI to known viral RdRp sequences [75,78].

These metatranscriptome studies focused on single protist or macroalgae holobionts have demonstrated that (+) ssRNA, dsRNA, and potentially (−) ssRNA viruses are distributed across many diverse protist groups and went previously undetected in “bulk” environmental approaches. These viruses have small genomes (typically <5 kb), may be segmented, and some lack capsid proteins. Viral families *Amalgaviridae*, *Chrysoviridae*, *Totiviridae*, *Partitiviridae*, *Endornaviridae*, *Mitoviridae*, and *Narnaviridae* are all associated with persistent, asymptomatic infections in plants and fungi [29] and likely cause nonlytic, but persistent chronic infection in their marine hosts.

7. Recommendations for Future Studies

Several studies discussed in this review demonstrate the plethora of information available when different sequencing approaches are combined [48,50,56], but these studies are still very rare. It is of utmost importance to conduct more studies that integrate metagenomics, metatranscriptomics, and DNA/RNA metaviromics to fully understand the spectrum of the viral diversity and capture the ecological interactions between viruses and their host. Targeted enrichment of dsRNA viral nucleic acid followed by dsRNA sequencing and addition of long-read sequencing could also complement RNA diversity studies by capturing unknown RNA viruses and alleviate the biases of the current methods. To our knowledge, only one study compared RNA viruses in cellular and extracellular viral fractions simultaneously, and it added a novel perspective by demonstrating that both fractions contain a substantial number of dsRNA viruses [56]. A valuable addition to the molecular toolshed for the study of aquatic RNA viral communities could be the use of nanopore sequencing (Oxford Nanopore Technology, Oxford, UK). This technology can be used to sequence native DNA and RNA molecules as they pass through a nanopore; the nucleotide sequence is determined as a change in the current and is specific to the oligomer passing through the pore. Read length is limited only by fragment length, and with the right sample preparation, reads can reach lengths of megabases [79]. It was shown that the inclusion of long reads benefits contig length and captures more DNA viral diversity [80–82]. Incorporating long reads in the studies of RNA viral diversity may improve sequence assemblies and/or enable recovery of full-length RNA genomes from environmental samples in a single long read and therefore circumvent the need for assembly. This may help to identify novel viruses and strains. Sequencing of native RNA molecules could also help to overcome biases associated with reverse transcription and cDNA amplification [83]. The constraints of the long-read approach are the amount and quality of starting RNA material.

Despite the immense new RNA viral diversity uncovered with multi-omic approaches, it is evident that RNA viruses in aquatic environments are very sparsely sampled (Table 2; Figure 4). In addition to conducting diversity surveys, more ecology-focused studies are needed to understand the ecological patterns of RNA viruses in aquatic environments. There was only one RNA metaviromics temporal study conducted during Antarctic diatom bloom [24] and one larger spatial study [58], with no continuous temporal studies on a weekly, monthly, or yearly timescale. Metatranscriptomic datasets are more temporally oriented, typically following a bloom over a few weeks on a weekly sampling basis (Table 2). Short time-series, such as that employed by Kolody et al. [49], demonstrate that RNA viruses can be highly active and dynamic and have diurnal fluctuations, emphasizing the need for more higher-resolution studies.

Most metatranscriptome studies focus on productive areas during algal blooms (Table 2). Efforts to compare across multiple contexts (bloom vs. non-bloom) and in different biomes will undoubtedly lead to a more holistic and comprehensive understanding of RNA and DNA viral dynamics and impacts. This is a worthy endeavour, as the abundance and activity of RNA viruses, and of viruses in general, are intimately connected to the proper functioning of marine food webs, and future climate change will alter the viral-mediated control on biogeochemical cycling [84], with potentially unpredictable consequences.

Another area that deserves attention is standardization and validation of the pre-sequencing and sequencing sample preparation protocols. Enrichment methods, extraction kits, and library preparation can significantly affect the number of viral reads, and consequently the assembly, and skew the relative abundances of particular viral groups and diversity of the viruses present in the sample [47,85]. To ensure robust comparisons, protocols should be compared and validated with mock viral communities to quantify the efficiency and identify the biases of the methods, as previously done for the human virome [85–88]. Lastly, both metaviromics and metatranscriptomics depend on identifying homology with viral sequences present in the database, and highly divergent viral sequences with little or no sequence homology cannot be readily identified. Even in enriched samples, most reads have low similarity to known nucleotide or protein sequences. Therefore, novel bioinformatic approaches for detecting highly divergent RNA viruses are required to further extend the boundaries of our knowledge about RNA viral diversity and evolution [53].

8. Conclusions

Multi-omic sequencing approaches have advanced our understanding of RNA viral diversity and expanded our understanding of their ecological roles in aquatic ecosystems. Three common messages have emerged from the synthesis of all these studies: (1) lytic ssRNA viruses from the order *Picornavirales* are a stable component of the aquatic RNA viral communities, but are perhaps not as dominant as previously thought. (2) All three phyla of positive-sense ssRNA viruses are represented in comparable proportions in the aquatic ecosystem as well as the dsRNA viruses. Highly divergent RNA viruses are being discovered frequently, and are infecting protist hosts that were previously thought only to host DNA viruses. (3) Negative-sense ssRNA viruses, which are traditionally associated with multicellular animals and plants, were found for the first time in the metatranscriptomes of some single-celled eukaryotic phytoplankton cultures. The articles discussed above demonstrate significant progress in characterizing aquatic RNA viral communities, but also highlight how undersampled RNA viral communities are. However, as more data accumulate, we expect to paint a more comprehensive picture of RNA viral diversity, abundance, and host range across the aquatic environments.

Author Contributions: S.K., D.-Q.C. and F.M.L. contributed to all sections, drafted and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation, Prime’s Minister’s Office, Singapore under its Intra-CREATE Seed Collaboration Grant (Award NRF2018-ITS004-0001) and Competitive Research Programme (Award CRP21-2018-0005). Sandra Kolundzija is supported by a Singapore International Graduate Award (SINGA) from the Agency for Science, Technology & Research (A*STAR).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: We thank Winona Wijaya and Maja Mihinjac for their assistance with Figure 4.

Conflicts of Interest: The authors declare no competing interests.

References

1. Suttle, C.A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Microbiol.* **2007**, *5*, 801–812. [[CrossRef](#)] [[PubMed](#)]
2. Fuhrman, J. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541–548. [[CrossRef](#)] [[PubMed](#)]
3. Weitz, J.S.; Wilhelm, S.W. Ocean viruses and their effects on microbial communities and biogeochemical cycles. *F1000 Biol. Rep.* **2012**, *4*, 2–9. [[CrossRef](#)]
4. Zimmerman, A.E.; Howard-Varona, C.; Needham, D.M.; John, S.G.; Worden, A.Z.; Waldbauer, J.R.; Coleman, M.L. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* **2020**, *18*, 21–34. [[CrossRef](#)] [[PubMed](#)]
5. Burki, F.; Roger, A.J.; Brown, M.W.; Simpson, A.G. The New Tree of Eukaryotes. *Trends Ecol. Evol.* **2020**, *35*, 43–55. [[CrossRef](#)]

6. Coy, S.R.; Gann, E.R.; Pound, H.L.; Short, S.M.; Wilhelm, W.S. Viruses of eukaryotic algae: Diversity, methods for detection, and future directions. *Viruses* **2018**, *10*, 487. [[CrossRef](#)]
7. Short, S.M.; Staniewski, M.A.; Chaban, Y.V.; Long, A.M.; Wang, D. Diversity of viruses infecting eukaryotic algae. *Curr. Issues Mol. Biol.* **2020**, *39*, 29–62. [[CrossRef](#)]
8. Nasir, A.; Forterre, P.; Kim, K.M.; Caetano-Anollés, G. The distribution and impact of viral lineages in domains of life. *Front. Microbiol.* **2014**, *5*, 194. [[CrossRef](#)]
9. Callanan, J.; Stockdale, S.R.; Shkoporov, A.; Draper, L.A.; Ross, R.P.; Hill, C. Expansion of known ssRNA phage genomes: From tens to over a thousand. *Sci. Adv.* **2020**, *6*, eaay5981. [[CrossRef](#)]
10. Labonté, J.M.; Suttle, C.A. Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J.* **2013**, *7*, 2169–2177. [[CrossRef](#)]
11. Wolf, Y.I.; Silas, S.; Wang, Y.; Wu, S.; Bocek, M.; Kazlauskas, D.; Krupovic, M.; Fire, A.; Dolja, V.V.; Koonin, E.V. Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. *Nat. Microbiol.* **2020**, *5*, 1262–1270. [[CrossRef](#)] [[PubMed](#)]
12. Roux, S.; Krupovic, M.; Poulet, A.; Debross, D.; Enault, F. Evolution and diversity of the *Microviridae* viral family through a collection of 81 new complete genomes assembled from virome reads. *PLoS ONE* **2012**, *7*, e40418. [[CrossRef](#)] [[PubMed](#)]
13. Roux, S.; Krupovic, M.; Daly, R.A.; Borges, A.L.; Nayfach, S.; Schulz, F.; Sharrar, A.; Matheus Carnevali, P.B.; Cheng, J.F.; Ivanova, N.N.; et al. Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nat. Microbiol.* **2019**, *4*, 1895–1906. [[CrossRef](#)] [[PubMed](#)]
14. Koonin, E.V.; Dolja, V.V.; Krupovic, M.; Varsani, A.; Wolf, Y.I.; Yutin, N.; Zerbini, F.M.; Kuhn, J.H. Global Organization and Proposed Megataxonomy of the Virus World. *Microbiol. Mol. Biol. Rev.* **2020**, *84*, e00061-19. [[CrossRef](#)]
15. Koonin, E.V.; Krupovic, M.; Agol, V.I. The Baltimore Classification of Viruses 50 Years Later: How Does It Stand in the Light of Virus Evolution? *Microbiol. Mol. Biol. Rev.* **2021**, *85*, e0005321. [[CrossRef](#)]
16. Wolf, Y.I.; Kazlauskas, D.; Iranzo, J.; Lucía-Sanz, A.; Kuhn, J.H.; Krupovic, M.; Dolja, V.V.; Koonin, E.V. Origins and evolution of the global RNA virome. *MBio* **2018**, *9*, e02329-18. [[CrossRef](#)]
17. Lang, S.; Rise, M.L.; Culley, A.I.; Steward, G.F. RNA viruses in the sea. *FEMS Microbiol. Rev.* **2009**, *33*, 295–323. [[CrossRef](#)]
18. Kaneko, H.; Blanc-Mathieu, R.; Endo, H.; Chaffron, S.; Delmont, T.O.; Gaia, M.; Henry, N.; Hernández-Velázquez, R.; Nguyen, C.H.; Mamitsuka, H.; et al. Eukaryotic virus composition can predict the efficiency of carbon export in the global ocean. *iScience* **2021**, *24*, 102002. [[CrossRef](#)]
19. Weinbauer, M.G. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* **2004**, *28*, 127–181. [[CrossRef](#)]
20. Sadeghi, M.; Tomaru, Y.; Ahola, T. RNA Viruses in Aquatic Unicellular Eukaryotes. *Viruses* **2021**, *13*, 362. [[CrossRef](#)]
21. Tomaru, Y.; Nagasaki, K. Flow cytometric detection and enumeration of DNA and RNA viruses infecting marine eukaryotic microalgae. *J. Oceanogr.* **2007**, *63*, 215–221. [[CrossRef](#)]
22. Patel, A.; Noble, R.T.; Steele, J.A.; Schwabach, M.S.; Hewson, I.; Fuhrman, J.A. Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. *Nat. Protoc.* **2007**, *2*, 269–276. [[CrossRef](#)] [[PubMed](#)]
23. Steward, G.F.; Culley, A.I.; Mueller, J.A.; Wood-Charlson, E.M.; Belcaid, M.; Poisson, G. Are we missing half of the viruses in the ocean? *ISME J.* **2013**, *7*, 672–679. [[CrossRef](#)] [[PubMed](#)]
24. Miranda, J.A.; Culley, A.I.; Schwarcz, C.R.; Steward, G.F. RNA viruses as major contributors to Antarctic viroplankton. *Environ. Microbiol.* **2016**, *18*, 3714–3727. [[CrossRef](#)]
25. Culley, A. New insight into the RNA aquatic virosphere via viromics. *Virus Res.* **2017**, *244*, 84–89. [[CrossRef](#)]
26. Starr, E.P.; Nuccio, E.E.; Pett-Ridge, J.; Banfield, J.F.; Firestone, M.K. Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 25900–25908. [[CrossRef](#)]
27. Shi, M.; Lin, X.D.; Tian, J.H.; Chen, L.J.; Chen, X.; Li, C.X.; Qin, X.C.; Li, J.; Cao, J.P.; Eden, J.S.; et al. Redefining the invertebrate RNA virosphere. *Nature* **2016**, *540*, 539–543. [[CrossRef](#)]
28. Roossinck, M.J.; Martin, D.P.; Roumagnac, P. Plant virus metagenomics: Advances in virus discovery. *Phytopathology* **2015**, *105*, 716–727. [[CrossRef](#)]
29. Roossinck, M.J. Evolutionary and ecological links between plant and fungal viruses. *New Phytol.* **2019**, *221*, 86–92. [[CrossRef](#)]
30. Roux, S.; Adriaenssens, E.M.; Dutilh, B.E.; Koonin, E.V.; Kropinski, A.M.; Krupovic, M.; Kuhn, J.H.; Lavigne, R.; Brister, J.R.; Varsani, A.; et al. Minimum information about an uncultivated virus genome (MIUVIG). *Nat. Biotechnol.* **2019**, *37*, 29–37. [[CrossRef](#)]
31. Suttle, C.A.; Chan, A.M.; Cottrell, M.T. Use of ultrafiltration to isolate viruses from seawater which are pathogens of marine phytoplankton. *Appl. Environ. Microbiol.* **1991**, *57*, 721–726. [[CrossRef](#)] [[PubMed](#)]
32. John, S.G.; Mendez, C.B.; Deng, L.; Poulos, B.; Kauffman, A.K.M.; Kern, S.; Brum, J.; Polz, M.F.; Boyle, E.A.; Sullivan, M.B. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ. Microbiol. Rep.* **2011**, *3*, 195–202. [[CrossRef](#)]
33. Hurwitz, B.L.; Deng, L.; Poulos, B.T.; Sullivan, M.B. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. *Environ. Microbiol.* **2013**, *15*, 1428–1440. [[CrossRef](#)] [[PubMed](#)]
34. Thurber, R.V.; Haynes, M.; Breitbart, M.; Wegley, L.; Rohwer, F. Laboratory procedures to generate viral metagenomes. *Nat. Protoc.* **2009**, *4*, 470–483. [[CrossRef](#)] [[PubMed](#)]

35. Roux, S.; Krupovic, M.; Debroas, D.; Forterre, P.; Enault, F. Assessment of viral community functional potential from viral metagenomes may be hampered by contamination with cellular sequences. *Open Biol.* **2013**, *3*, 130160. [[CrossRef](#)] [[PubMed](#)]
36. Prangishvili, D.; Bamford, D.H.; Forterre, P.; Iranzo, J.; Koonin, E.V.; Krupovic, M. The enigmatic archaeal virosphere. *Nat. Rev. Microbiol.* **2017**, *15*, 724–739. [[CrossRef](#)] [[PubMed](#)]
37. Hillman, B.I.; Cai, G. *The Family Narnaviridae. Simplest of RNA Viruses*, 1st ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2013; Chapter 6, pp. 149–176.
38. Valverde, R.A.; Khalifa, M.E.; Okada, R.; Fukuhara, T.; Sabanadzovic, S. ICTV virus taxonomy profile: *Endornaviridae*. *J. Gen. Virol.* **2019**, *100*, 1024–1025. [[CrossRef](#)]
39. Suzuki, N.; Ghabrial, S.A.; Kim, K.H.; Pearson, M.; Marzano, S.Y.L.; Yaegashi, H.; Xie, J.; Guo, L.; Kondo, H.; Koloniuk, I.; et al. ICTV virus taxonomy profile: *Hypoviridae*. *J. Gen. Virol.* **2018**, *99*, 615–616. [[CrossRef](#)]
40. Wilcox, H.; Delwart, E.; Díaz-Muñoz, S.L. Next-generation sequencing of dsRNA is greatly improved by treatment with the inexpensive denaturing reagent DMSO. *Microb. Genom.* **2019**, *5*, e000315. [[CrossRef](#)]
41. Endo, H.; Blanc-Mathieu, R.; Li, Y.; Salazar, G.; Henry, N.; Labadie, K.; de Vargas, C.; Sullivan, M.B.; Bowler, C.; Wincker, P.; et al. Biogeography of marine giant viruses reveals their interplay with eukaryotes and ecological functions. *Nat. Ecol. Evol.* **2020**, *4*, 1639–1649. [[CrossRef](#)]
42. Sieradzki, E.T.; Ignacio-Espinoza, J.C.; Needham, D.M.; Fichot, E.B.; Fuhrman, J.A. Dynamic marine viral infections and major contribution to photosynthetic processes shown by spatiotemporal picoplankton metatranscriptomes. *Nat. Commun.* **2019**, *10*, 1169. [[CrossRef](#)]
43. Ha, D.; Moniruzzaman, M.; Aylward, F.O. High Transcriptional Activity and Diverse Functional Repertoires of Hundreds of Giant Viruses in a Coastal Marine System. *mSystems* **2021**, *6*, e00293-21. [[CrossRef](#)] [[PubMed](#)]
44. Moniruzzaman, M.; Wurch, L.L.; Alexander, H.; Dyhrman, S.T.; Gobler, C.J.; Wilhelm, S.W. Virus-host relationships of marine single-celled eukaryotes resolved from metatranscriptomics. *Nat. Commun.* **2017**, *8*, 16054. [[CrossRef](#)] [[PubMed](#)]
45. Hewson, I.; Bistolas, K.S.; Button, J.B.; Jackson, E.W. Occurrence and seasonal dynamics of RNA viral genotypes in three contrasting temperate lakes. *PLoS ONE* **2018**, *13*, e0194419. [[CrossRef](#)]
46. Pound, H.L.; Gann, E.R.; Tang, X.; Krausfeldt, L.E.; Huff, M.; Staton, M.E.; Talmy, D.; Wilhelm, S.W. The ‘Neglected Viruses’ of Taihu: Abundant Transcripts for Viruses Infecting Eukaryotes and Their Potential Role in Phytoplankton Succession. *Front. Microbiol.* **2020**, *11*, 338. [[CrossRef](#)] [[PubMed](#)]
47. Gann, E.R.; Kang, Y.; Dyhrman, S.T.; Gobler, C.J.; Wilhelm, S.W. Metatranscriptome Library Preparation Influences Analyses of Viral Community Activity During a Brown Tide Bloom. *Front. Microbiol.* **2021**, *12*, 664189. [[CrossRef](#)] [[PubMed](#)]
48. Alarcón-Schumacher, T.; Guajardo-Leiva, S.; Antón, J.; Llewellyn, C.A. Elucidating Viral Communities During a Phytoplankton Bloom on the West Antarctic Peninsula. *Front. Microbiol.* **2019**, *10*, 1014. [[CrossRef](#)]
49. Kolody, B.C.; McCrow, J.P.; Allen, L.Z.; Aylward, F.O.; Fontanez, K.M.; Moustafa, A.; Moniruzzaman, M.; Chavez, F.P.; Scholin, C.A.; Allen, E.E.; et al. Diel transcriptional response of a California Current plankton microbiome to light, low iron, and enduring viral infection. *ISME J.* **2019**, *13*, 2817–2833. [[CrossRef](#)]
50. Zeigler Allen, L.; McCrow, J.P.; Ininbergs, K.; Dupont, C.L.; Badger, J.H.; Hoffman, J.M.; Ekman, M.; Allen, A.E.; Bergman, B.; Venter, J.C. The Baltic Sea Virome: Diversity and Transcriptional Activity of DNA and RNA Viruses. *mSystems* **2017**, *2*, e00125-16. [[CrossRef](#)]
51. Carradec, Q.; Pelletier, E.; Da Silva, C.; Alberti, A.; Seeleuthner, Y.; Blanc-Mathieu, R.; Lima-Mendez, G.; Rocha, F.; Tirichine, L.; Labadie, K.; et al. A global ocean atlas of eukaryotic genes. *Nat. Commun.* **2018**, *9*, 373. [[CrossRef](#)]
52. Depledge, D.P.; Mohr, I.; Wilson, A.C. Going the Distance: Optimizing RNA-Seq Strategies for Transcriptomic Analysis of Complex Viral Genomes. *J. Virol.* **2019**, *93*, e01342-18. [[CrossRef](#)] [[PubMed](#)]
53. Zhang, Y.Z.; Chen, Y.M.; Wang, W.; Qin, X.C.; Holmes, E.C. Expanding the RNA Virosphere by Unbiased Metagenomics. *Annu. Rev. Virol.* **2019**, *6*, 119–139. [[CrossRef](#)] [[PubMed](#)]
54. Rosani, U.; Shapiro, M.; Venier, P.; Allam, B. A needle in a haystack: Tracing bivalve-associated viruses in high-throughput transcriptomic data. *Viruses* **2019**, *11*, 205. [[CrossRef](#)] [[PubMed](#)]
55. Charon, J.; Marcelino, V.R.; Wetherbee, R.; Verbruggen, H.; Holmes, E.C. Metatranscriptomic identification of diverse and divergent RNA viruses in green and chlorarachniophyte algae cultures. *Viruses* **2020**, *12*, 1180. [[CrossRef](#)]
56. Urayama, S.I.; Takaki, Y.; Nishi, S.; Yoshida-Takashima, Y.; Deguchi, S.; Takai, K.; Nunoura, T. Unveiling the RNA virosphere associated with marine microorganisms. *Mol. Ecol. Resour.* **2018**, *18*, 1444–1455. [[CrossRef](#)]
57. Zhang, X. Environmental viromes reveal global virosphere of deep-sea RNA viruses. *Res. Sq.* **2022**, *in press*. [[CrossRef](#)]
58. Vlok, M.; Lang, A.S.; Suttle, C.A. Marine RNA Virus Quasispecies Are Distributed throughout the Oceans. *mSphere* **2019**, *4*, e00157-19. [[CrossRef](#)]
59. Culley, I.; Lang, A.S.; Suttle, C.A. Metagenomic analysis of coastal RNA virus communities. *Science* **2006**, *312*, 1795–1798. [[CrossRef](#)]
60. Culley, I.; Mueller, J.A.; Belcaid, M.; Wood-Charlson, M.E.; Poisson, G.; Steward, G.F. The Characterization of RNA Viruses in Tropical Seawater Using Targeted PCR and Metagenomics. *mBio* **2014**, *5*, e01210-14. [[CrossRef](#)]
61. Shi, M.; Zhang, Y.Z.; Holmes, E.C. Meta-transcriptomics and the evolutionary biology of RNA viruses. *Virus Res.* **2018**, *243*, 83–90. [[CrossRef](#)]

62. Weber, F.; Wagner, V.; Rasmussen, S.B.; Hartmann, R.; Paludan, S.R. Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses. *J. Virol.* **2006**, *80*, 5059–5064. [[CrossRef](#)] [[PubMed](#)]
63. Blouin, G.; Ross, H.A.; Hobson-Peters, J.; O'Brien, C.A.; Warren, B.; MacDiarmid, R. A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies. *Mol. Ecol. Resour.* **2016**, *16*, 1255–1263. [[CrossRef](#)] [[PubMed](#)]
64. Marais, A.; Faure, C.; Bergey, B.; Candresse, T. Viral Double-Stranded RNAs (dsRNAs) from Plants: Alternative Nucleic Acid Substrates for High-Throughput Sequencing. *Methods Mol. Biol.* **2018**, *1746*, 45–53. [[CrossRef](#)] [[PubMed](#)]
65. Danovaro, R.; Corinaldesi, C.; Rastelli, E.; Dell'Anno, A. Towards a better quantitative assessment of the relevance of deep-sea viruses, Bacteria and Archaea in the functioning of the ocean seafloor. *Aquat. Microb. Ecol.* **2015**, *75*, 81–90. [[CrossRef](#)]
66. Danovaro, R.; Dell'Anno, A.; Corinaldesi, C.; Rastelli, E.; Cavicchioli, R.; Krupovic, M.; Noble, R.T.; Nunoura, T.; Prangishvili, D. Virus-mediated archaeal hecatomb in the deep seafloor. *Sci. Adv.* **2016**, *2*, e1600492. [[CrossRef](#)] [[PubMed](#)]
67. Yoshida, M.; Takaki, Y.; Eitoku, M.; Nunoura, T.; Takai, K. Metagenomic Analysis of Viral Communities in (Hado)Pelagic Sediments. *PLoS ONE* **2013**, *8*, e57271. [[CrossRef](#)]
68. De Corte, D.; Martínez, J.M.; Cretoiu, M.S.; Takaki, Y.; Nunoura, T.; Sintes, E.; Herndl, G.J.; Yokokawa, T. Viral communities in the global deep ocean conveyor belt assessed by targeted viromics. *Front. Microbiol.* **2019**, *10*, 1801. [[CrossRef](#)]
69. Li, Z.; Pan, D.; Wei, G.; Pi, W.; Zhang, C.; Wang, J.H.; Peng, Y.; Zhang, L.; Wang, Y.; Hubert, C.R.; et al. Deep sea sediments associated with cold seeps are a subsurface reservoir of viral diversity. *ISME J.* **2021**, *15*, 2366–2378. [[CrossRef](#)]
70. Corinaldesi, C.; Tangherlini, M.; Dell'Anno, A. From virus isolation to metagenome generation for investigating viral diversity in deep-sea sediments. *Sci. Rep.* **2017**, *7*, 8355. [[CrossRef](#)]
71. Yoshida, M.; Mochizuki, T.; Urayama, S.I.; Yoshida-Takashima, Y.; Nishi, S.; Hirai, M.; Nomaki, H.; Takaki, Y.; Nunoura, T.; Takai, K. Quantitative viral community DNA analysis reveals the dominance of single-stranded DNA viruses in offshore upper bathyal sediment from Tohoku, Japan. *Front. Microbiol.* **2018**, *9*, 75. [[CrossRef](#)]
72. Karlsson, E.; Belák, S.; Granberg, F. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. *Biosecur. Bioterror.* **2013**, *11*, S227–S234. [[CrossRef](#)] [[PubMed](#)]
73. Yilmaz, S.; Allgaier, M.; Hugenholtz, P. Multiple displacement amplification compromises quantitative analysis of metagenomes. *Nat. Methods* **2010**, *7*, 943–944. [[CrossRef](#)] [[PubMed](#)]
74. Marine, R.; McCarren, C.; Vorrassane, V.; Nasko, D.; Crowgey, E.; Polson, S.W.; Wommack, K.E. Caught in the middle with multiple displacement amplification: The myth of pooling for avoiding multiple displacement amplification bias in a metagenome. *Microbiome* **2014**, *2*, 3. [[CrossRef](#)] [[PubMed](#)]
75. Urayama, S.I.; Takaki, Y.; Nunoura, T. FLDS: A comprehensive DSRNA sequencing method for intracellular RNA virus surveillance. *Microbes Environ.* **2016**, *31*, 33–40. [[CrossRef](#)] [[PubMed](#)]
76. Attoui, H.; Jaafar, F.M.; Belhouchet, M.; de Micco, P.; de Lamballerie, X.; Brussaard, C.P.D. *Micromonas pusilla* reovirus: A new member of the family *Reoviridae* assigned to a novel proposed genus (*Mimoreovirus*). *J. Gen. Virol.* **2006**, *87*, 1375–1383. [[CrossRef](#)]
77. Charon, J.; Murray, S.; Holmes, E.C. Revealing RNA virus diversity and evolution in unicellular algae transcriptomes. *Virus Evol.* **2021**, *7*, veab070. [[CrossRef](#)]
78. Chiba, Y.; Tomaru, Y.; Shimabukuro, H.; Kimura, K.; Hirai, M.; Takaki, Y.; Hagiwara, D.; Nunoura, T.; Urayama, S.I. Viral RNA genomes identified from marine macroalgae and a diatom. *Microbes Environ.* **2020**, *35*, ME20016. [[CrossRef](#)]
79. Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **2020**, *21*, 30. [[CrossRef](#)]
80. Overholt, W.A.; Hölzer, M.; Geesink, P.; Diezel, C.; Marz, M.; Küsel, K. Inclusion of Oxford Nanopore long reads improves all microbial and viral metagenome-assembled genomes from a complex aquifer system. *Environ. Microbiol.* **2020**, *22*, 4000–4013. [[CrossRef](#)]
81. Warwick-Dugdale, J.; Solonenko, N.; Moore, K.; Chittick, L.; Gregory, A.C.; Allen, M.J.; Sullivan, M.B.; Temperton, B. Long-read viral metagenomics captures abundant and microdiverse viral populations and their niche-defining genomic islands. *PeerJ* **2019**, *7*, e6800. [[CrossRef](#)]
82. Beaulaurier, J.; Luo, E.; Eppley, J.M.; Den Uyl, P.; Dai, X.; Burger, A.; Turner, D.J.; Pendelton, M.; Juul, S.; Harrington, E.; et al. Assembly-free single-molecule sequencing recovers complete virus genomes from natural microbial communities. *Genome Res.* **2020**, *30*, 437–446. [[CrossRef](#)] [[PubMed](#)]
83. Kozarewa, I.; Ning, Z.; Quail, M.A.; Sanders, M.J.; Berriman, M.; Turner, D.J. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat. Methods* **2009**, *6*, 291–295. [[CrossRef](#)] [[PubMed](#)]
84. Danovaro, R.; Corinaldesi, C.; Dell'Anno, A.; Fuhrman, J.A.; Middelburg, J.J.; Noble, R.T.; Suttle, C.A. Marine viruses and global climate change. *FEMS Microbiol. Rev.* **2011**, *35*, 993–1034. [[CrossRef](#)]
85. Hjelmsø, M.H.; Hellmér, M.; Fernandez-Cassi, X.; Timoneda, N.; Lukjancenko, O.; Seidel, M.; Elsässer, D.; Aarestrup, F.M.; Löfström, C.; Bofill-Mas, S.; et al. Evaluation of methods for the concentration and extraction of viruses from sewage in the context of metagenomic sequencing. *PLoS ONE* **2017**, *12*, e0170199. [[CrossRef](#)]

86. Conceição-Neto, N.; Zeller, M.; Lefrère, H.; De Bruyn, P.; Beller, L.; Deboutte, W.; Yinda, C.K.; Lavigne, R.; Maes, P.; Ranst, M.V.; et al. Modular approach to customise sample preparation procedures for viral metagenomics: A reproducible protocol for virome analysis. *Sci. Rep.* **2015**, *5*, 16532. [[CrossRef](#)]
87. Kohl, C.; Brinkmann, A.; Dabrowski, P.W.; Radonić, A.; Nitsche, A.; Kurth, A. Protocol for metagenomic virus detection in clinical specimens. *Emerg. Infect. Dis.* **2015**, *21*, 48–57. [[CrossRef](#)]
88. Rosseel, T.; Ozhelvaci, O.; Freimanis, G.; Van Borm, S. Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. *J. Virol. Meth.* **2015**, *222*, 72–78. [[CrossRef](#)]