Viruses of Eukaryotic Algae: Diversity, Methods for Detection, and Future Directions

Samantha R. Coy [1] , **Eric R. Gann** [1] , **Helena L. Pound** [1] , **Steven M. Short** [2] and **Steven W. Wilhelm** [1,*]

[1] The Department of Microbiology, The University of Tennessee, Knoxville, TN 37996, USA; srose16@tennessee.edu (S.R.C.); egann@tennessee.edu (E.R.G.); hpound@tennessee.edu (H.L.P.)
[2] The Department of Biology, The University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada; steven.short@utoronto.ca
[*] Correspondence: wilhelm@utk.edu; Tel.: +1-865-974-0665

Abstract: The scope for ecological studies of eukaryotic algal viruses has greatly improved with the development of molecular and bioinformatic approaches that do not require algal cultures. Here, we review the history and perceived future opportunities for research on eukaryotic algal viruses. We begin with a summary of the 65 eukaryotic algal viruses that are presently in culture collections, with emphasis on shared evolutionary traits (e.g., conserved core genes) of each known viral type. We then describe how core genes have been used to enable molecular detection of viruses in the environment, ranging from PCR-based amplification to community scale "-omics" approaches. Special attention is given to recent studies that have employed network-analyses of -omics data to predict virus-host relationships, from which a general bioinformatics pipeline is described for this type of approach. Finally, we conclude with acknowledgement of how the field of aquatic virology is adapting to these advances, and highlight the need to properly characterize new virus-host systems that may be isolated using preliminary molecular surveys. Researchers can approach this work using lessons learned from the *Chlorella* virus system, which is not only the best characterized algal-virus system, but is also responsible for much of the foundation in the field of aquatic virology.

1. Introduction

Viruses infecting eukaryotic algae are extremely diverse. They have been reported with DNA or RNA genomes in various architectures (linear, circular, double-stranded, single-stranded, segmented) and sizes (4.4 to 638kb) [1]. Some viruses accomplish infection with just a few viral genes at their disposal, while others maintain a gene arsenal nearly 100 times that size. Viruses infecting algae influence large ecological and biogeochemical processes. They direct the evolution of hosts through predator-prey selection and genetic exchange, consequently influencing algal fitness, population dynamics, and ultimately, microbial community structure. Infection can also alter the composition and distribution of organic matter in the environment (a process referred to as the aquatic "viral shunt" [2]) and influence particle size-distribution, nutrient cycling, and biological system activity (e.g., respiration [3]). While algal viruses are important members in many aquatic environments, their contribution to these processes at the global scale primarily arises when they infect and lyse abundant bloom-forming algae. This includes harmful bloom formers and ecosystem scale specialists like coccolithophores that form blooms large enough to be observed from outer space [4].

It is a relatively recent realization that algal viruses are ecologically significant. In fact, the whole history of algal virus research has occurred primarily in just the last half century (Figure 1). While there have been sporadic observations of virus infection of algae cultures since the early 1970s [5,6], the importance of algal viruses in natural systems was brought into the limelight by a series of observations of virus-like-particles associated with important bloom-forming algae [7–9]. These findings inspired questions about the identity and evolutionary relationships within these virus-host systems. Such questions, however, required viruses to be isolated and genetically characterized.
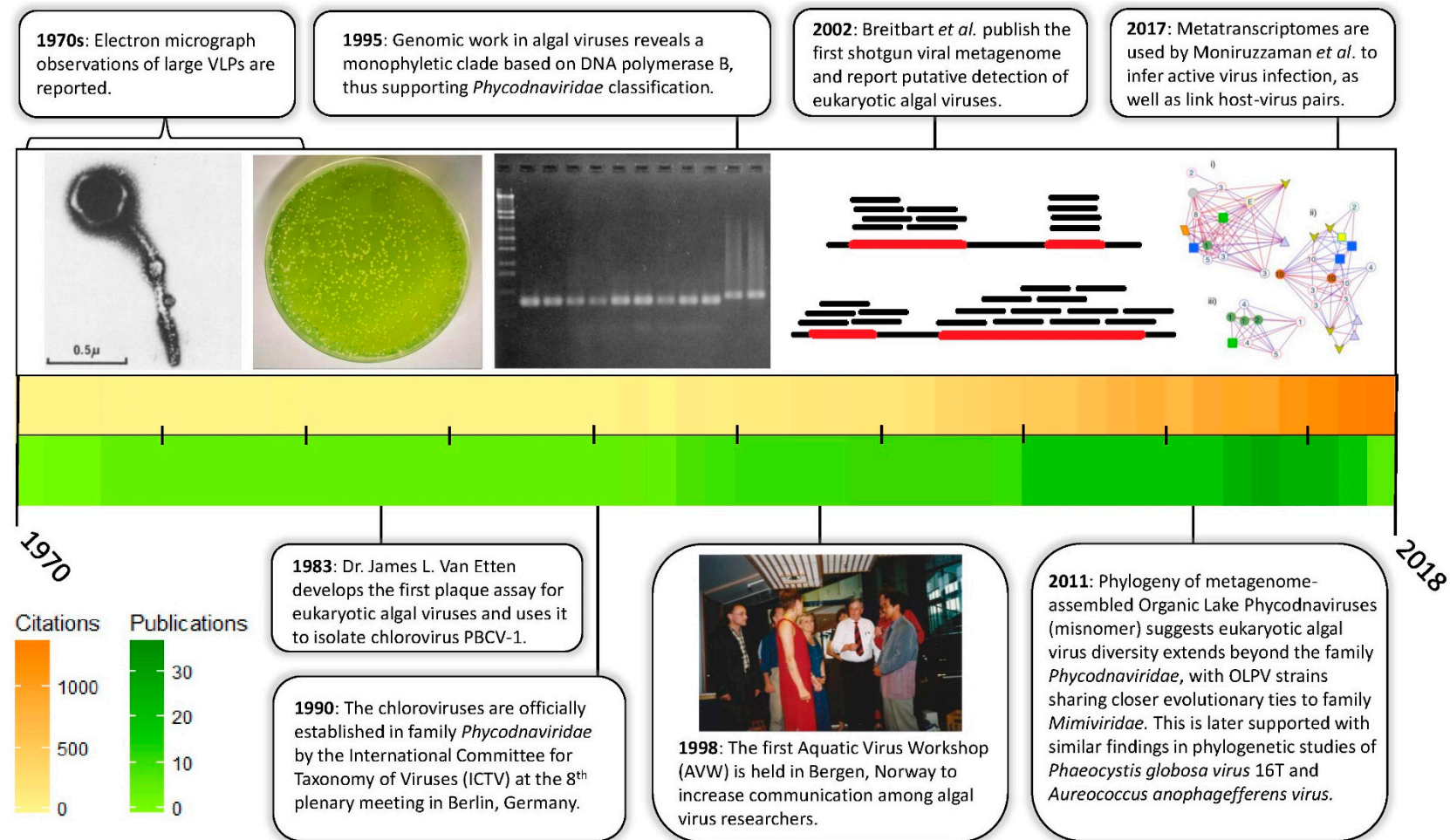
Figure 1. Timeline of eukaryotic algal virus research with important milestones highlighted. Colored bars represent the annual citations and publications generated from a Web of Science Citation Report using the field tag TS = (algal virus) for all databases. The search was conducted on 8 May 2018 at 11:00 a.m. Citation Report results were visualized as heatmaps using custom R scripts. Electron micrograph image [10] and electrophoretic gel [11] reprinted by permission. Network analysis [12] reprinted under authority of Creative Commons.

One of the first algal-virus systems to achieve "model" status were the double-stranded DNA (dsDNA) viruses that infect the unicellular, ex-symbiotic, green alga *Chlorella* [13]. The *Chlorella* virus-host model system remains the best characterized of all the algae-virus models, with genomes [14–18], transcriptomes [19,20], and proteomes [14] documented in the literature. Indeed, it was sequencing of the DNA polymerase B (*polB*) genes from *Chlorella* viruses PBCV-1 and NY-2A [21], and later from *Micromonas pusilla virus* SP1, that revealed a conserved amino acid sequence distinct from other known polB protein sequences. This observation enabled the development of degenerate PCR primers that selectively amplified these algal-virus *polB* genes [11,22]. The sequences of these PCR amplicons supported a unique monophyletic viral clade, now recognized as the family *Phycodnaviridae* of the Nucleocytoplasmic Large dsDNA Viruses (NCLDV). For a while the *Phycodnaviridae* was thought to be home to all of the large dsDNA algal viruses: perhaps even dominating the overall algal virus community. This perspective changed when sequencing of new isolates demonstrated that their "core" genes were more closely related to genes from the protist-infecting "giant viruses" of family *Mimiviridae* [23–25]. In general, algal-infecting viruses are recognized as members of one of these two families, though future work may challenge the monophyletic nature of these groups. For example, clustering of the *Phycodnaviridae* is at times disrupted when homologs from other cellular or viral families are included in phylogenetic reconstructions [26–28]

There have also been increasing reports of single-stranded DNA (ssDNA) viruses, mostly infecting diatoms, RNA viruses (Table 1) [29,30], and even parasites of these large algal viruses known as virophage [31,32]. The most informative reports on these systems have come from metagenomic and metatranscriptomic datasets that can detect the presence and activity of a wide range of DNA and RNA viruses. In turn, the known diversity of eukaryotic algal viruses has greatly expanded, at times even yielding putative full-length viral genome assemblies [12]. Perhaps most promising is the possibility of predicting virus-host relationships in silico [12,33,34], whereas traditional methods have relied on virus isolation from a relative few cultivated algae. Shotgun -omics further create the opportunity to identify virus-host pairs from environmental data and place them in semi-quantitative ecological context. Indeed, these studies may even serve as preliminary assessments of the future cultivation requirements for isolating new virus-host systems. This burgeoning scientific frontier necessitates a review on the known diversity of eukaryotic algal viruses, the molecular toolkit available for in situ studies on their ecology, and the direction aquatic virology is taking to adapt to these methodologies.

2. Diversity of Cultured Virus-Host Systems

The diversity of algal viruses mirrors that of their hosts, bearing in mind that the name "algae" does not denote a common evolutionary relationship. Indeed, algae have been observed in freshwater, marine, and terrestrial systems, in unicellular, colonial, or multicellular forms, and in disparate taxonomic lineages. Nevertheless, the diversity of algae can be depicted using an existing taxonomic framework that includes seven "supergroups" consisting of Excavata, Amoebozoa, Opisthokonta, Archaeplastida, the SAR group (Stramenophila, Alveolata, and Rhizaria), and a series of non-delineated, "cryptic" organisms collectively referred to as the *Incerta sedis* [35]. Beyond this framework, the manner in which certain taxa are placed within eukaryotic phylogeny varies in the literature and is a subject of ongoing scientific debate. We adapted the schematic phylogeny presented by the TARA Oceans group [36] to illustrate the diversity of marine eukaryotic plankton, their relative abundance based on TARA Oceans 18S rDNA gene surveys, and lineage association with viruses that have been isolated and are maintained in lab cultures (Figure 2 and Table 1). This framework demonstrates that marine eukaryotic algae are known to occupy all but the Amoebozoa and Opisthokonta supergroups. Algae-infecting viruses have been isolated using hosts spanning almost all abundant planktonic lineages, though many are single systems or instances without genomic information to define viral phylogenetic placement (e.g., TampV). Although *Pyramimonadales* and *Raphidophyceae* were not abundant in the TARA Oceans 18S dataset, select species in these groups are known bloom-formers [37–40] making the available algal-virus system for these lineages ecologically

informative. Viruses have also been isolated on important non-planktonic species, such as brown and red macroalgae (*Phaeophyceae* and *Rhodophyceae*). Abundant lineages without an algae-infecting virus include photosynthetic *Dictyochophyceae*, the Prasino Clade 7 group, the Chryso/Synuro group, and the Apicomplexans—though some of the highly represented lineages could be attributed to non-photosynthetic members. Establishing well characterized host-virus systems in these lineages could be very useful for bloom-forming algae of these lineages. For example, it would be appealing to isolate a *Pseudochattonella* (*Dictyochophyceae*) infecting virus, as the host alga is responsible for fish kills. In 2016 *Pseudochattonella* was responsible for a massive fish kill in Peru amounting to an economic loss of ~$800 million dollars [41]. In another interesting, albeit more complicated example, survival of the red-tide, bloom-forming ciliate *Mesodinium rubrum* depends on ingestion of photosynthetic cryptophytes to obtain necessary organelles (e.g., plastid, mitochondria, nucleus) [42]. Viruses infecting cryptophyte prey may compete with this grazer, thus serving as an important control on the frequency and duration of red tides. Such broad trophic effects have been shown in studies on *Emiliania huxleyi*, where viral-infected cells are ingested by zooplankton at different rates than non-infected cells [43,44].
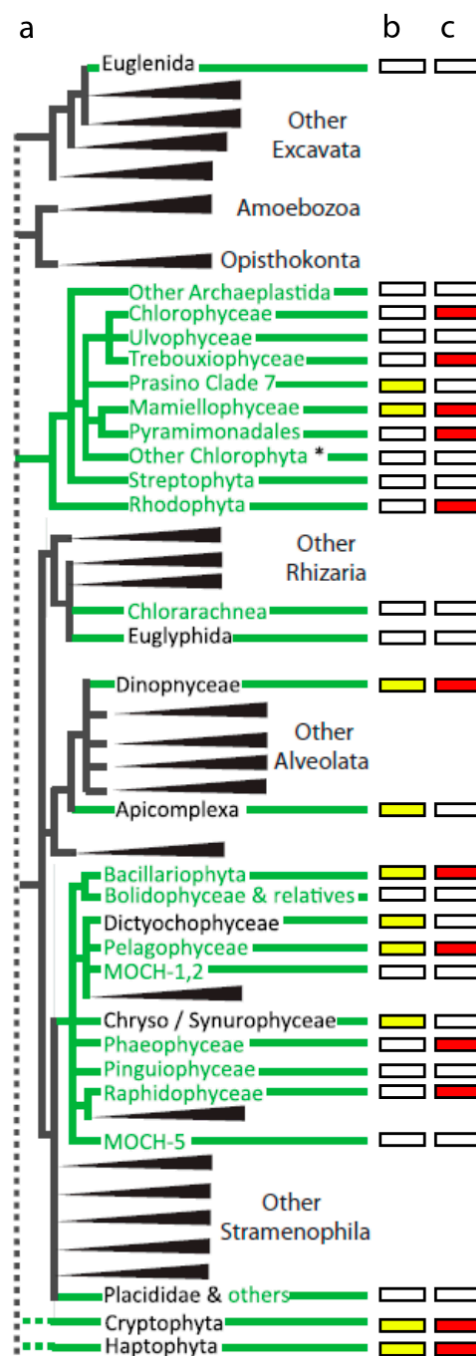
Figure 2. (**a**) Schematic phylogeny adapted from de Vargas et al. demonstrating known virus-interactions with eukaryotic alga lineages. The phylogeny was originally constructed on recognized eukaryotic plankton lineages that were detected in TARA Oceans datasets, which included hits to all aquatic algal containing lineages. We collapsed the original tree to highlight these lineages in the context of their current phylogenetic placement. Green lines denote lineages with photosynthetic algal representatives, whereas the text color indicates whether all or only some representatives are phototrophic-green or black text, respectively; (**b**) Yellow boxes denote the top ten most abundant, planktonic, phototroph-associated lineages based on 18S rDNA surveyed in the TARA Oceans study. Asterisks denote lineages that were artificially grouped for simplicity, and their full descriptions can be found at http://taraoceans.sb-roscoff.fr/EukDiv/; (**c**) Red boxes denote algal-lineages that have an isolated algae-infecting virus in culture collection, though these are not all marine systems. The virus isolates are listed in Table 1.

Table 1. Algal viruses currently in culture collection.

Host Algae	Type	Size (kbp or knt)	Code	References
Chlorophyceaea				
Tetraselmis spp.	dsDNA	668	TetV	Schvarcz et al., 2018 [45]
Tetraselmis striata	dsDNA	31	Tsv-N1	Pagarete et al., 2015 [46]
Trebouxiophyceae				
Chlorella variabilis NC64A	dsDNA	287–369	PBCV-1	Jeanniard et al., 2013 [17]
Chlorella variabilis Syngen 2-3	dsDNA	327	OSy-NE5	Quispe et al., 2017 [18]
Chlorella heliozoae SAG 3.83	dsDNA	288–327	ATCV-1	Jeanniard et al., 2013 [17]
Micratinium conductrix Pbi	dsDNA	302–329	CVM	Jeanniard et al., 2013 [17]
Mamiellophyceae				
Ostreococcus lucimarinus	dsDNA	182–196	OlV1	Derelle et al., 2015 [47]
Ostreococcus tauri	dsDNA	184–192	OtV5	Weynberg et al., 2011 [48]
Ostreococcus mediterraneus	dsDNA	193	OmV1	Derelle et al., 2015 [47]
Bathycoccus sp. RCC1105	dsDNA	187–198	BpV	Moreau et al., 2010 [49]
Micromonas pusilla CCMP1545	dsDNA	186–195	MpV-02T	Martinez Martinez et al., 2015 [50]
Micromonas pusilla LAC38	dsDNA	173–205	MpV1	Finke et al., 2017 [51]
Micromonas pusilla LAC38	dsRNA	25.5	MpRV	Brussaard et al., 2004 [52]
Micromonas polaris	dsDNA	191–205	MpoV	Maat et al., 2017 [53]
Pyramimonadales				
Pyramimonas orientalis	dsDNA	560	PoV	Sandaa et al., 2001 [54]
Rhodophyta				
Chondrus crispus	dsRNA	6	CcV	Rousvoal et al., 2016 [55]
Dinophyceaea				
Heterocapsa circularisquama	dsDNA	356	HcDNAV	Ogata et al., 2009 [56]
Heterocapsa circularisquama	ssRNA	4.4	HcRNAV	Tomaru et al., 2004 [57]
Heterocapsa pygmea	dsDNA	ND	HpygDNAV	Kim et al., 2012 [58]
Gymnodinium mikimotoi	ND	ND	GM6/GM7	Onji et al., 2003 [59]
Bacillariophyta				
Chaetoceros cf. *gracilise*	ND	ND	CspNIV	Bettarel et al., 2005 [60]
Chaetoceros salsugineum	ssDNA	6	CsalDNAV*	Nagasaki et al., 2005 [61]
Chaetoceros setoensis	ssDNA	5.8	CsetDNAV*	Tomaru et al., 2013 [62]
Chaetoceros socialis f. *radians*	ssRNA	9.4	CsfrRNAV	Tomaru et al., 2009b [63]
Chaetoceros lorenzianus	ssDNA	5.8	ClorDNAV*	Tomaru et al., 2011 [64]
Chaetoceros tenuissimus	ssDNA	5.6	CtenDNAV-I*	Tomaru et al., 2011 [65]
Chaetoceros tenuissimus	ssDNA	5.6	CtenDNAV-II*	Kimura and Tomaru 2015 [66]

Table 1. *Cont.*

Host Algae	Type	Size (kbp or knt)	Code	References
Chaetoceros tenuissimus	ssRNA	9.4	CtenRNAV	Shirai et al., 2008 [67]
Chaetoceros tenuissimus, Chaetoceros spp.	ssRNA	9.6	CtenRNAV-II	Kimura and Tomaru 2015 [66]
Chaetoceros spp. SS628-11	ssDNA	5.5	Csp07DNAV*	Kimura et al., 2013 [68]
Chaetoceros spp. TG07-C28	ssDNA	ND	Csp05DNAV	Toyoda et al., 2012 [69]
Chaetoceros debilis	ssDNA	ND	CdebDNAV	Tomaru et al., 2008 [70]
Chaetoceros sp. SS08-C03	ssRNA	9.4	Csp03RNAV	Tomaru et al., 2013 [71]
Chaetoceros cf. *wighamii*	ssDNA	7-8	CwNIV	Eissler et al., 2009 [72]
Asterionellopsis glacialis	ssRNA	9.5	AglaRNAV	Tomaru et al., 2012 [73]
Thalassionema nitzschioides	ssDNA	5.5	TnitDNAV	Tomaru et al., 2012 [73]
Rhizosolenia setigera	ssRNA	11.2	RsetRNAV	Nagasaki et al., 2004 [74]
Skeletonema costatum	ND	ND	ScosV	Kim et al., 2015 [75]
Stephanopyxis palmeriana	ND	ND	SpalV	Kim et al., 2015 [76]
Pelagophyceae				
Aureococcus anophagefferens	dsDNA	370	AaV	Moniruzzaman et al., 2014 [23]
Phaeophyceae				
Ectocarpus fasciculatus	dsDNA	340	EfasV	Kapp et al., 1997 [77]
Ectocarpus siliculosus	dsDNA	320	EsV	Kapp et al., 1997 [77]
Feldmannia irregularis	dsDNA	180	FirrV	Kapp et al., 1997 [77]
Feldmannia simplex	dsDNA	220	FlexV	Kapp et al., 1997 [77]
Feldmannia species	dsDNA	170	FsV	Henry and Meints 1992 [78]
Hincksia hinckiae	dsDNA	240	HincV	Kapp et al., 1997 [77]
Myriotrichia clavaeformis	dsDNA	320	MclaV	Kapp et al., 1997 [77]
Pilayella littoralis	dsDNA	280	PlitV	Maier et al., 1998 [79]
Raphidophyceae				
Heterosigma akashiwo	dsDNA	ND	HaV	Nagasaki et al., 1997 [80]
Heterosigma akashiwo	dsDNA	180	O1s1	Lawrence et al., 2006 [81]
Heterosigma akashiwo	ssRNA	9.1	HaRNAV	Tai et al., 2003 [82]
Heterosigma akashiwo	ND	ND	HaNIV	Lawrence et al., 2001 [83]
Haptophyta				
Emiliania huxleyi	dsDNA	415	EhV	Castberg et al., 2002 [84]
Phaeocystis globosa	dsDNA	466	PgV-16T (Group I)	Baudoux et al., 2005 [85]
Phaeocystis globosa	dsDNA	177	PgV-03T (Group II)	Baudoux et al., 2005 [85]
Phaeocystis globosa	dsDNA	176	PgV-102P	Wilson et al., 2006 [86]
Phaeocystis pouchetii	dsDNA	485	PpV	Jacobsen et al., 1996 [87]

Table 1. *Cont.*

Host Algae	Type	Size (kbp or knt)	Code	References
Chrysochromulina brevifilum, *Chrysochromulina strobilus*	dsDNA	ND	CbV	Suttle and Chan 1995 [88]
Chrysochromulina ericina	dsDNA	510	CeV	Sandaa et al., 2001 [54]
Chrysochromulina parva	dsDNA	485	CpV	Mirza et al., 2015 [89]
Haptolina ericina, Prymnesium kappa	dsDNA	530	HeV-RF02	Johannessen et al., 2015 [90]
Prymnesium kappa, Haptolina ericina	dsDNA	ND	PkV-RF01	Johannessen et al., 2015 [90]
Prymnesium kappa	dsDNA	507	PkV-RF02	Johannessen et al., 2015 [90]
Prymnesium parvum	dsDNA	ND	PpDNAV	Wagstaff et al., 2017 [91]
Cryptophyta				
Teleaulax amphioxeia	ND	ND	TampV	Nagasaki et al., 2009 [92]

Table 1. Summary of all reported eukaryotic algal viruses that have been isolated. A range of genome sizes (kbp or knt) represents multiple virus strains associated with the same host species, and in this case, only the type virus is reported under the code column. Asteriks denote original names for some of the diatom ssDNA viruses, which have since been renamed and placed into genera of the family *Bacilladnaviridae* (*Chaetoceros setoensis* DNA virus = *Diatodnavirus*; *Chaetoceros salsugineum* DNA virus 1 = *Chaetoceros protobacilladnavirus* 1; *Chaetoceros* sp. DNA virus 7 = *Chaetoceros protobacilladnavirus* 2; *Chaetoceros lorenzianus* DNA virus = *Chaetoceros protobacilladnavirus* 3; *Chaetoceros tenuissimus* DNA viruses type I and II = *Chaetoceros protobacilladnavirus* 4). ND = Not detected or reported.

Eukaryotic algal viruses in culture collections have been isolated from ~60 alga species (Table 1). Most of these are lytic, dsDNA viruses of the NCLDV group with a narrow, known host-range. The abundance of NCLDVs would imply that these are an ecologically relevant algal-virus type in the virus community, but whether or not these are the dominating type is unclear. This would certainly contrast with plant viromes which are dominated by RNA viruses. It is also possible that NCLDVs are more easily detected and isolated, thus explaining why only dsDNA viruses have been isolated from water samples that putatively contained other types of viruses. For example, electron micrographs of bloom-associated *Emiliania huxleyi* cells have been observed to simultaneously contain both small (50–60 nm) and large (185–200 nm) intracellular VLPs [93]. Similar observations been made in *Pyramimonas orientalis* [94], but currently only one type of dsDNA virus has been isolated for this algae [54]. It is possible that these viruses compete for algal infection, but they may also represent a case of virus-infecting virophage that are already known to co-occur with *Mimiviridae* [95,96], and perhaps even *Phycodnaviridae* [31,33] viruses. Observations of co-occurring viruses are not limited to microscopy either; network analysis of metatranscriptomic data has linked the brown alga *Aureococcus anophagefferens* to its known dsDNA virus AaV as well as to uncharacterized ssDNA viruses [12], although the mechanism of this linkage (either direct, or via a co-occurring microbial host of the virus) remains elusive. In short, algae may be infected by many types of viruses, potentially at the same time, and the numerically dominant virus type may not always represent that which is in the culture collection.

To date, there are four algal species that are known to be infected by diverse viruses comprised of different nucleic acid types. These include *Heterosigma akashiwo*, *Chaetoceros tenuissimus*, *Micromonas pusilla*, and *Heterocapsa circularisquama*, and in all cases the different virus types infect the same host strain [97]. The coexistence of *Heterosigma akashiwo* viruses HaRNAV and HaDNAV is especially intriguing given these viruses exhibit opposite infection dynamics; the RNA virus has a high viral production rate, but a slower lytic cycle, whereas the DNA virus quickly replicates but produces fewer particles [81]. It was hypothesized that coexistence could be maintained through variable host densities and viral decay rates, thus representing viruses that may have evolved as r- or k- strategists as has been proposed for *Heterocapsa* viruses [98], but is certainly not supported enough to be extrapolated as an explanation for all co-occurring viruses. Even virus isolates of the same nucleic acid type and species can exhibit considerable diversity. This can be extreme in some cases, where dsDNA viruses infecting the same algal host, which would be expected to cluster phylogenetically, are affiliated with NCLDV viral families *Mimiviridae* or *Phycodnaviridae* (e.g., *Phaeocystis globosa* Virus Groups I and Groups II [25,99]. It is possible that eukaryotic algae may commonly be infected by viruses of diverse replication strategies, and evolutionary histories, but the extent of this, as well as the factors that may allow this, needs more thorough investigation.

2.1. dsDNA Viruses Infecting Eukaryotic Algae

Most dsDNA viruses infecting algae are members of the NCLDV group, with the proposed exception of Tsv-N1 [46]. Algal-NCLDV viruses have large genomes that encode hundreds of protein coding genes. Their evolutionary relationship has been inferred by core genes conserved across NCLDVs [100], placing them into either the family *Phycodnaviridae* or as extended members of the family *Mimiviridae*. Algal viruses of the latter group have recently been given the proposed distinction of Mesomimivirinae [101], but for our purposes we will maintain the *Mimiviridae* description. The one exception to these two family assignments is HcDNAV, which shares closer similarity to the family *Asfarviridae* [56]. To date, the NCLDV core gene compliment has been reduced to just a few genes (e.g., D5R packaging ATPase, D13L major capsid protein, and B family DNA polymerase), implying that the genetic diversity is huge among this group. Indeed, a genomic comparison among *Phycodnaviridae* members PBCV-1 (Chloroviruses), EsV-1 (Phaeoviruses), and EhV-86 (Coccolithoviruses) yielded only 14 conserved homologs from a pool of ~1000 genes [102]. A more comprehensive look at these diverse genes can be found in genus-specific reviews of the *Phycodnaviridae* [17,47,51,103–105].

It is anticipated that any single algal host can be permissive to many closely related virus variants, whereby phylogenetic comparisons of their core genes will reveal distinct clades (e.g., *Micromonas pusilla* and *Chlorella variabilis* viruses) with differences in latent phases, burst sizes, and genome size [17]. In closely related viruses this is best resolved using concatenated alignments of marker protein sequences. At the same time, the origin of some of these genes is often attributed to gene transfer events. Many algal NCLDVs have acquired non-ancestral genes, but the majority of these appear to come from difference sources: Prasinoviruses acquire most of these from their host, Chlorovirus non-ancestral genes mostly derive from bacteria [106], and *Aureococcus anophagefferens* Virus (AaV) encodes a more even mixture of host, bacterial, archaeal, and viral genes [23]. At the same time, it is worth noting that the origin of some genes could be difficult to ascertain if only a limited subset of viral (and host) homologs have been sequenced and annotated in public databases. Regardless, it has been suggested that viruses whose hosts are in closer association with bacteria tend to encode more putative non-ancestral genes, and that these genes cluster near the terminal ends of the viral genome [107]. However, while the *Chlorella* algae is an endosymbiont of *Paramecium* that is certainly in close proximity to bacteria, the non-ancestral genes carried by the virus are evenly dispersed across its genome [17]. In contrast, AaV displays terminal clusters of non-ancestral genes [23], but its host is a free-living photo/osmotroph. In either case, the biological implication of such high viral gene diversity, and how it is generated, is unclear. It may help the virus acquire its specific needs for infection but has also been proposed to allow viruses to infect multiple hosts.

2.2. ssDNA Viruses Infecting Eukaryotic Algae

To date, the only ssDNA alga-infecting viruses that have been isolated are those which infect diatoms (Bacillariophyceae). In total, diatoms are a collective of an estimated 12,000–30,000 species, representing one of the most abundant phytoplankton groups in freshwater and marine environments [108]. Most diatom-virus systems currently in culture are those infecting the cosmopolitan genus *Chaetoceros*. These isometric virus particles are ~35 nm in diameter and house circular, ssDNA genomes ranging from ~5.5–6.0 kb [66]. The genomes generally encode four open reading frames consisting of an endonuclease (Rep), a major capsid protein, and two ORFs with unknown function. The capsid and replication initiating endonuclease are used in phylogenetic analyses. Three new members (whose genomes are ~4.5–4.7 kb) were recently reported from a de novo assembly of metagenomic reads from the mollusk *Amphibola crenata* and from sediment within an estuary in New Zealand [109]. Phylogenetic analysis of the capsid proteins suggest this gene is a recent acquisition from ssRNA viruses, which is interesting, though not without precedent [110,111]. These metagenome assembled viruses have resulted in the taxonomic reclassification of diatom viruses into the family *Bacilladnaviridae* that includes cultured diatom viruses noted in Table 1 with asterisks [112]. Many other ssDNA viruses are being detected in omics datasets [12], though resolving their specific host is an ongoing challenge.

2.3. RNA Viruses Infecting Eukaryotic Algae

Algae-infecting viruses with single (ss) and double-stranded (ds) RNA genomes have also been isolated and characterized, although most attention has been focused on the ssRNA isolates. Both virus groups encode an RNA-dependent RNA polymerase (RdRP), as well as proteases and helicases that can be used to infer distant evolutionary relationships. Most information on dsRNA algal viruses has been derived from the original isolation papers describing the evolutionary relationships of the isolates. MpRV, a dsRNA virus of *Micromonas pusilla*, forms its own genus within the family *Reoviridae* (unassigned order) and has been proposed to be the ancestral line of the *Reoviridae* based on its placement between clades that demonstrate turreted or non-turreted virions [113]. The other dsRNA virus isolate is *Chondrus crispus* virus (CcV), a toti-virus like entity. CcV represents an extraordinary case of a putative quasispecies virus that was accidentally discovered when a small band of dsRNA (~6 kb) was observed during host genomic preparation for sequencing [55]. Similar dsRNA bands

have been observed in extracts from all algal life phases, geographic locations, and in extracts from other red algae, though virus-like-particles and host lysis was not observed. The CcV system may represent either a latent or chronic (i.e., particle production below the limit of detection) viral infection that is ubiquitous among red algae, similar to known latent dsDNA viral infections of brown algae by Phaeoviruses [114]. Since both *Chondrus crispus* and *Micromonas pusilla* are ecologically important algae, characterization of their relationship with these viruses is important and perhaps reflective of a need to search for more dsRNA viruses associated with algae.

ssRNA viruses have received considerably more attention since their hosts are common marine phytoplankton with some species capable of forming harmful blooms [39,115,116]. Most of the alga-infecting ssRNA viruses are members of the order Picornavirales (Figure 3), with a few contradictions that are awaiting a taxonomic re-evaluation based on molecular data. The viruses infecting *Heterocapsa* and *Heterosigma* are the sole members of the families *Alvernaviridae* (unassigned order) and *Marnavirdiae* (order Picornavirales), respectively [109,112], while the genus *Bacillarnavirus* (order Picornavirales) includes formal members *Chaetoceros socialis forma radians* RNA virus, *Chaetoceros tenuissimus* RNA virus 01, and *Rhizosolenia setigera* RNA virus 01. Other diatom viruses Csp03RNAV, AglaRNAV, and CtenRNAV type II are putative members of *Bacillarnavirus* based on phylogenetic relationships of replicase or structural proteins [1]. The diatom viruses are generally thought to be highly species specific based on host-range experiments, with the exception of CtenRNAV type II which can infect four *Chaetoceros* sp. in addition to *Chaetoceros tenuissimus* [66]. These viruses and their hosts represent ecologically important systems that may reveal much on the persistence, co-existence, and competition of diatom viruses.
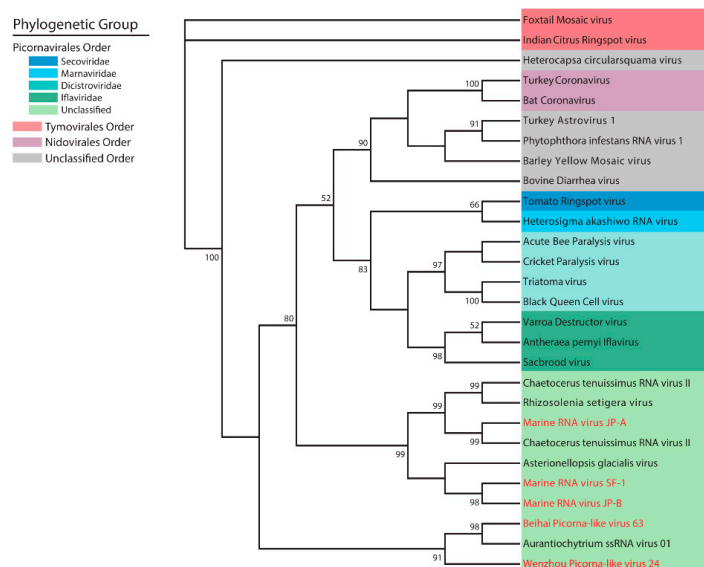

Figure 3. Diversity of single-stranded RNA viruses depicted based on phylogeny of RNA-dependent RNA polymerase (Rdrp NCBI CDD:01699) reference sequences downloaded from NCBI RefSeq database (see in Supplemental Table S1). Sequences were aligned and trimmed in Mega7 [117] and an unrooted maximum likelihood phylogeny was created using PhyML 3.0 with LG model [118]. Empirical equilibrium frequencies were used with aLRT SH-like statistics for branch support. Phylogenetic groups are color coded with algal viruses denoted by a star. Viral isolates from metagenomic assemblies are in red text.

3. Culture Independent Approaches: Expanding Known Diversity

3.1. PCR Applications for Estimating Viral Diversity and Dynamics

Developing algal-virus model systems in the lab can inform much on the biology and ecology of algal viruses, but dependence on these systems is a limiting step. The ability to determine viral

geographic distributions, population fluctuations, and diversity ultimately depends on analysis of environmental samples. Microscopic methods [119], flow cytometry [120–122], and infectivity assays (e.g., most probable number, plaque assay [13]) have been used to answer these questions, but these approaches lack taxonomic resolution and/or the relatively quick processing time that molecular techniques provide. To date, the principal molecular method for studying environmental algal viruses has been based on PCR amplification of conserved marker genes. Most of this work has focused on algal NCLDVs using polB [11] and the NCLDV major capsid protein (mcp) as gene targets [123]: subsets of this community have been further examined using primers that specifically target the extended, algal *Mimiviridae* major capsid protein (AMmcp) [124]. For reference, the potential amplification ranges of these primers are mapped against a phylogeny of sequenced virus isolates (Figure 4). There has been discussion on amplification bias of *polB* primers based on observations that environmental datasets tend to amplify prasinoviruses, even though these may be environmentally abundant viral types [98]. The gene amplified by this primer set has also been suggested to be a poor marker for resolving within algal virus genera. For example, there are two distinct groups of *Phaeocystis globosa* infecting viruses, and these groups phylogenetically cluster into different families [1]. Diversity may be better assessed using genome fluidity measurements of the pan-genome [125], but this would work better for describing viruses with full-genome sequences. Indeed, marker gene primer sets remain useful for elucidating environmental diversity of algal NCLDVs.
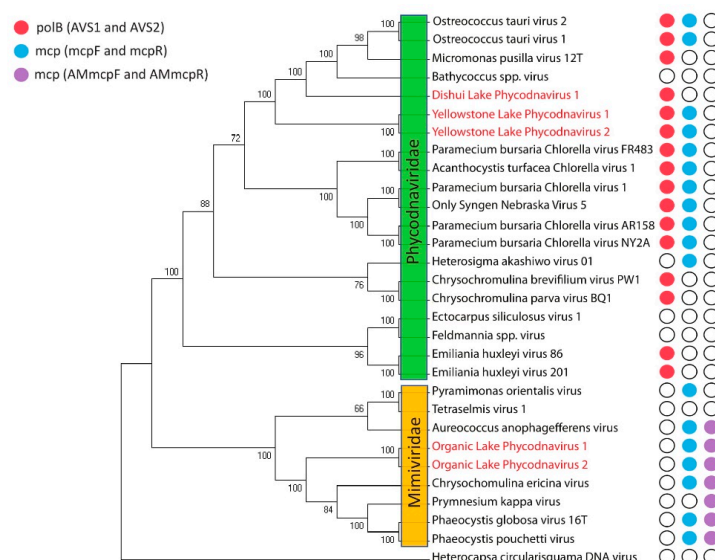

Figure 4. Phylogenetic tree depicting the evolutionary relationships of algal NCLDVs based on amino acid alignment (ClustalW) of the core gene, DNA polymerase B (see in Supplemental Table S2). The tree was built using the maximum likelihood method based on the JTT matrix-based model with 200 iterations in MEGA7 [117]. Viruses belong either to the family *Phycodnaviridae* or are recognized "extended members" of the family *Mimiviridae*. The recently discovered dinoflagellate infecting virus, *Heterocapsa circularisquama* DNA virus, was used to root the tree and shows little similarity to other algal NCLDVs despite being a large DNA virus. Viruses in red text denote metagenome assembled viral genomes, meaning their association with an alga host is putative. Colored dots to the right indicate the viruses can be putatively PCR amplified by the respective PCR primer set based on ≥90% match between each primer and its respective target binding site. This equates to ≤2 primer mismatches, which has been shown to be capable of producing a PCR reaction, albeit at lower efficiency (for RT-qPCR) [126]. The same study shows that three or more mismatches in the same primer completely inhibit a PCR reaction, and is an observation that aligns with failed PCR reactions reported for *Ectocarpus siliculosus* virus 1 and *Feldmannia* spp. virus [11]. PCR amplification predictions were done using motif searches in CLC Genomics and the software De-MetaST-BLAST [127].

A recent clone library of PCR amplicons generated using the two mcp primer sets demonstrates a wide diversity of algal viruses isolated from marine and freshwater environments [124]. This study also used PCR amplification to track the occurrence and dynamics of virus groups (defined by sequence clustering as operational taxonomic units, OTUs) over the course of a harmful brown-alga event. Biases aside, the approach used in that study has certainly expanded the known diversity of algal NCLDVs. It has also shown that cultured viral isolates are often distinct from environmental viruses, and that viruses are widely dispersed in the environment [123,124,128–131]. Another recent group of primer sets was developed by Wilson et al. that amplifies a putative algal-*Mimiviridae* specific mismatch repair gene (*MutS*) [132]. Novel groups of algal NCLDVs were detected in all of the samples tested, making this gene/primer set another potentially useful tool for studying virus diversity. RNA virus diversity has been assessed using primer sets targeting RNA dependent RNA polymerase (RdRP), a protein encoded by all RNA viruses [29,133]. This led to the discovery of a highly diverse super group of putative, marine, protist-infecting picorna-like viruses [133] that are consistently represented in metagenomic datasets [134]. Moreover, alignments of conserved regions of RdRP form clades that are congruent with virion structure, host, and epidemiology [29].

While diversity can be addressed with degenerate primer PCR amplification, one of the major drawbacks of this approach is that it is generally not suitable for quantitative measurements [135]. Indeed, degeneracies allow for biases in primer-binding and template amplification in mixed communities [136]. Use of more specific primer sets and quantitative PCR approaches can avoid this issue [137,138], but at the risk of not detecting closely related viruses. Even when using specific primer sets, recent duplications of marker genes can result in overestimation of viral abundances. One of the recent developments to overcome this is to spatially separate viruses and subject them to solid-phase, single-molecule PCR polony amplification [139]. Family specific degenerate primers amplify diverse members without the issue of competitive amplification, then categorize and quantify the amplicons using probes for virus group specific genes. Of course, this method is also dependent on prior sequence knowledge on the virus types of interest and has been validated only in cyanophage thus far, but it is certainly an appealing method for the study of eukaryotic alga infecting viruses. Another recently discovered application of PCR is its potential to link viruses and hosts. Microfluidics can be used to isolate infected single-cells that can then be subjected to simultaneous PCR detection of viral and host genes [140].

3.2. Using Omics Approaches to Estimate Virus Diversity and Dynamics

Because community scale genomics and transcriptomics are not dependent on target amplification, they are better suited for resolving viral diversity and can in some cases allow for the assembly of complete viral genomes. Though this is more readily accomplished in small RNA and DNA viruses [12,109,141], it has also been possible for some large dsDNA viruses and virophage [24,31,32]. This potential is so valuable that a proposal was recently submitted to the International Committee on the Taxonomy of Viruses (ICTV) for the inclusion of metagenomic-assembled viruses into the official classification scheme [142]. Not only was this approved, but it initiated a change in the primary approach ICTV uses for virus classification from phenotypic characterization based on viral isolates to molecular characterization based on viral DNA sequences. Since this time, metagenome assembled circular Rep-encoding single-stranded (CRESS) DNA viruses have been properly classified, including the *Bacilladnaviridae* [109], the putative vertebrate infecting *Smacoviridae* [143], and many more [112,144]. Some of the initial taxonomic classifications may also need to be reassessed in light of molecular methods, as classical taxonomy based on phenotype is not always congruent with phylogenetic clustering: The order Nidovirales may in fact belong to the Picornavirales.

While becoming more common, sequencing entire viral communities remains challenging and each experimental step must be considered in the context of existing biases and the project objectives. Virus particles have very low nucleic acid contents, necessitating amplification, concentration, or enrichment to obtain adequate sequencing depth. Simple approaches to do this involve concentration

of environmental samples via filtration [145] or chemical flocculation [146]. Virus enrichment can be done for specific viral types with some quantitative applications. For example, dsDNA can be quantitatively amplified using fusion PCR primers, and adaptase will quantitatively amplify both ssDNA and dsDNA viruses [147]. Rolling circle amplification can increase detection of circular viruses [109], and recombinant plant proteins that non-specifically bind dsRNA can select for dsRNA viruses [148]. There are also methods to separate DNA and RNA viruses for separate analyses using hydroxyapatite-mediated techniques [149]. One of the most appealing enrichment strategies recently used involves selection (*via* binding) of poly-A containing nucleic acid (i.e., mRNA) to focus on the active viral community [12]. This is a useful signal to distinguish virus particles from active infection, as the former will not produce an mRNA signal, though this excludes some (+) ssRNA viruses that have polyadenylated genomes independent of infection [150]. Though all of these methods are useful for improving detection, there are biases to be considered before making conclusions about viral abundances. These issues have been elucidated for sampling, extraction, and purification methods [151,152], but these studies are not comprehensive.

The viral sequences generated from any sequencing approach are subjected to a general analytical workflow involving quality filtering, assembly, annotation, and diversity analyses. Many tools are available to perform this bioinformatic workflow [153], but few of these are designed to complete the full workflow. Moreover, careful understanding of the sequence databases searched in each workflow is necessary to know whether biases exist for particular virus types. GenBank and the nt/nr databases are preferred as these are continually updated and contain information for all virus types; however, their large size can slow processing considerably. To overcome this, creating custom workflows using marker genes of interest can speed up processing time while maintaining the ability to detect diverse virus types.

An example of a bioinformatics workflow using a custom marker gene database to interpret NGS sequences (i.e., Illumina™ paired-end sequencing) is shown in Figure 5. First, reads must be preprocessed to remove contaminating adapter sequences and trim low-quality reads. The next step involves assembly of reads into larger contigs, followed by contig annotation using a database of known sequences and a homology or alignment search tool (BLAST, HMMER, Bowtie2, etc.). BLAST tools have commonly been used for this purpose in cellular organisms, and even in some virus studies [33], but may be less efficient for identifying novel virus homologs since they often have low pairwise sequence identities [154]. An alternative to using sequence alignments are Hidden Markov models (MMS), which score hits to protein domains. These analyses can be done with the search tool HMMER to create a marker gene database (HMM-build) that can be queried against assembled contigs [155]. Once viral contigs have been identified, the relevant gene hits can be extracted for post-processing (i.e., phylogenetic analysis). In many cases, especially when using small databases, it is useful to verify viral hits with a second similarity search of the extracted gene. Following verification, extracted viral hits can be placed unto an existing phylogenetic tree built with homologous reference sequences (e.g., pplacer [156]). Tree topology can be confirmed using a variety of other tree-building software (e.g., FastTree 2.1.7 [157], PhyML [118], RAxML [158], IQ-tree [159]) and methods (e.g., MrBayes for Bayesian tree-building [160]).
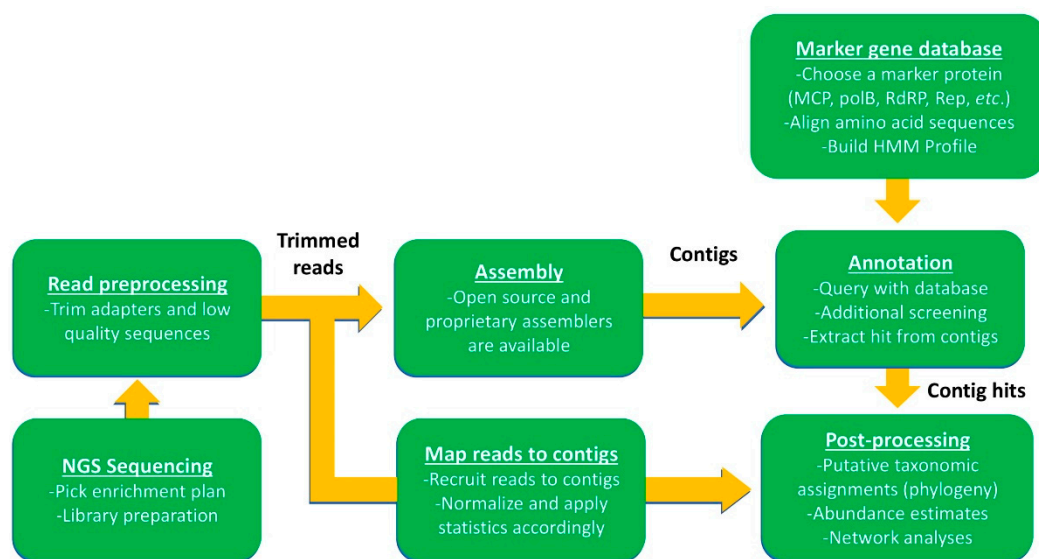
Figure 5. General bioinformatic pipeline using marker gene probing of community sequence data. This framework follows that used by Moniruzzaman et al., 2017 [12], where viral activity was assessed using marker gene detection from environmental mRNA. Though this framework was modeled off the cited study, it is flexible enough to incorporate both metagenomic and metatranscriptomic applications.

Information on virus abundance or activity can be inferred by mapping trimmed metagenomic or metatranscriptomic reads back to viral contigs normalized for between-sample comparisons (e.g., internal standards, library size, length, and reads per kilobase of transcript per million mapped reads [RPKM] values). However, there are some caveats to consider when examining environmental metatranscriptomes. Transcript abundance is not directly related to viral abundance for two reasons: First, biases are known to exist for highly transcriptionally active viruses, and second, single host organisms can support high viral loads. Moreover, virus metatranscriptomes can be contaminated with chimeras generated during assembly, remnant viral genes may be expressed from cells [161], and genomic duplications of marker genes could confound expression profiles. Some problems can be avoided with proper sampling and sequencing approaches mentioned previously, but others remain a significant obstacle for quantitative community analyses, though this has been resolved for bacteria-infecting viruses [147,162]. Until these confounding issues can be remedied and benchmarked for all viral types, they must be considered during the analysis of environmental data. A recent review by Nooij et al. provided a comprehensive description of workflows that have been produced for viromic analyses, including specific applications, classification biases, and open-source availability [153].

3.3. Other Downstream Applications of Omic Assemblies

Another enticing application of community sequence data is the potential to deduce biological interactions using co-occurrence or network analyses. This is a relatively new approach that was developed for microbiome communities but has the potential to identify novel virus-host pairs [163]. Two studies tracking the temporal dynamics of virus communities have been reported thus far [12,33]. From a metagenomics standpoint, these studies were striking because they generated putatively full-length Picornavirales and virophage genomes. Moreover, in the case of Moniruzzaman et al., 2017 [12] the viral genomes were generated from transcripts, indicating these virus genomes were actively expressed and were therefore, produced from infected cells. Beyond these exciting findings, each study used network analyses to link potential virus-host pairs. Clusters created from sequencing data collected over the course of a brown-tide bloom (*Aureococcus anophagefferens*) linked the brown alga to its known virus, AaV, demonstrating the ability to extract known relationships with this approach. Several other clusters were generated from the same study, including smaller networks of single virus-host pairs and expected associations between *Prasinophyceae* and *Phycodnaviridae*. Roux et al.,

2017 [33] focused on using networks to link virophage with giant NCLDV hosts and found strong specific associations with *Mimiviridae* and their extended alga-infecting members to drastically expand the diversity of known virophage hosts.

Altogether, predictions stemming from the studies noted above demonstrate how network analyses can generate testable hypotheses for future studies of algal virus-host interactions. By deducing sequences of virus-host pairs, one can attempt to confirm probable virus-host interactions. For example, a variation of fluorescent in-situ hybridization, deemed phageFISH, could be used to label virus and host genes in infected cells [164]. Additionally, networks predicting viruses of cultured algae could be followed up with virus tagging experiments [165]. It might even be worthwhile to use more than one network building approach to look at ecosystem structures. Weiss et al. used real and mock in silico data to benchmark eight methods used for bacterial network analyses and found that some methods generate drastically different outputs [166]. This is explained, in part, by differing strengths for detecting particular biological relationships (e.g., mutualism and commensalism) across different network approaches. It was also suggested that *p*-values of 0.001 should be used for high-precision network detection and rare OTUs should be removed prior to network construction.

4. Conclusions

The opportunities for algal virus ecologists are at an all-time high. Bioinformatic tools are becoming more accessible to a wide variety of scientists through the creation of publicly available genomic databases and graphic interfaces that mediate interactions with traditional command-line software [167]. At the same time, researchers are increasing collaborations with one another by sharing methodologies in an interactive framework on *protocols.io* (e.g., Viral Ecology Research and Virtual Exchange network, or VERVE Net; https://www.protocols.io/groups/verve-net) and with cross-discipline collaborations fostered at research workshops funded by organizations like the *Gordon & Betty Moore Foundation* (GBMF) and the *Canadian Institute for Advanced Research* (CIFAR). The development of long-read sequencing methods, preemptively deemed "third-generation sequencing", may address many of the issues with short-read assembly and viral quantification. DNA barcoding has been suggested as a cheap, reliable method to quickly track virus populations, and has recently been shown to recapitulate general viral community structures using sample volumes no bigger than a cup of water [168]. New virus isolates can be discovered from sequencing of single aquatic viruses sorted by flow cytometry, [169], as closely related, hyper diverse viruses are suggested to be difficult to assemble from metagenomes [170]. Even better, isolation and sequencing of infected single-cells may allow for the identification of new virus-host systems. Network analyses of community sequence data predict ecological structures that may lead to the discovery and isolation of several new algal-virus systems, bringing the scientific community "full-circle" to studying these systems in the lab. In light of that, this exciting frontier cannot be appreciated without recognition of the early work done by some of the first aquatic virologists in the field.

James L. Van Etten, for whom this special issue is in honor of, has spent the last forty years laying the foundation for aquatic virology. Not only did he open doors for other algal virus researchers to join the field, but he has set the standard for characterizing the biology and ecology of isolated algal virus systems. Along with the genomic, transcriptomic, and proteomic work done in the *Chlorella* virus system, the Van Etten lab has also shown that *Chlorella* viruses are biochemically novel in multiple ways. Virion proteins are glycosylated using a unique viral encoded machinery [171,172], and the viral genomes can be methylated by a range of DNA methyltransferases [173]. Many of these enzymes are paired with a restriction endonuclease that recognizes the same nucleotide sequence to comprise a viral restriction-modification system that recycles host DNA for virus replication [174]. *Chlorella* viruses have the smallest potassium ion channels that function to depolarize host membranes, concomitantly inhibiting secondary infection and host metabolite transporters [175]. Within seven minutes post-infection, transcriptional activity begins to shift away from the infected host towards producing viral transcripts [20,176]. Along with the extensive biological studies on

this system, the Van Etten group has also established many important findings on their ecology. *Chlorella* viruses are ubiquitous in freshwaters across the globe, despite their hosts being sequestered as endosymbionts of *Paramecium bursaria* [6]. This inspired questions about their resistance to degradation as well as how viruses and hosts make contact with one another. Predatory activity on *Paramecium bursaria* catalyzes this contact by making the endosymbiotic algae available to *Chlorella* viruses in the environment [177,178]. Another group has shown that *Chlorella* viruses are more resistant to environmental degradation than other algal viruses, and can even overwinter under ice [179]. Collectively, these questions can be investigated in many types of algal viruses. Although dsDNA viruses certainly have the largest number of genes, even smaller DNA and RNA viruses must deal with many of the same selective pressures. Indeed, there are many lessons algal virus researchers can learn from the body of work produced by James L. Van Etten and his collaborators.

References

1. Short, S.M.; Staniewski, M.A.; Chaban, Y.V.; Long, A.M.; Wang, D. Diversity of viruses infecting eukaryotic algae. In *Viruses of Microorganisms*; Paul, H., Abedon, S.T., Eds.; Caister Academic Press: Poole, UK, 2018.
2. Wilhelm, S.W.; Suttle, C.A. Viruses and nutrient cycles in the sea-viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* **1999**, *49*, 781–788. [CrossRef]
3. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541–548. [CrossRef] [PubMed]
4. Holligan, P.M.; Viollier, M.; Harbour, D.S.; Camus, P.; Champagne-philippe, M. Satellite and ship studies of coccolithophore production along a continental-shelf edge. *Nature* **1983**, *304*, 339–342. [CrossRef]
5. Wilhelm, S.W.; Bird, J.T.; Bonifer, K.S.; Calfee, B.C.; Chen, T.; Coy, S.R.; Gainer, P.J.; Gann, E.R.; Heatherly, H.T.; Lee, J.; et al. A student's guide to giant viruses infecting small eukaryotes: From acanthamoeba to zooxanthellae. *Viruses* **2017**, *9*, 46. [CrossRef] [PubMed]
6. Van Etten, J.L.; Lane, L.C.; Meints, R.H. Viruses and virus-like particles of eukaroytic algae. *Microbiol. Rev.* **1991**, *55*, 586–620. [PubMed]
7. Bratbak, G.; Egge, J.K.; Heldal, M. Viral mortality of the marine alga emiliania-huxleyi (haptophyceae) and termination of algal blooms. *Mar. Ecol. Prog. Ser.* **1993**, *93*, 39–48. [CrossRef]
8. Nagasaki, K.; Ando, M.; Itakura, S.; Imai, I.; Ishida, Y. Viral mortality in the final stage of heterosigma-akashiwo (raphidophyceae) red tide. *J. Plankton Res.* **1994**, *16*, 1595–1599. [CrossRef]
9. Gastrich, M.D.; Anderson, O.R.; Benmayor, S.S.; Cosper, E.M. Ultrastructural analysis of viral infection in the brown-tide alga, *Aureococcus anophagefferens* (pelagophyceae). *Phycologia* **1998**, *37*, 300–306. [CrossRef]
10. Dodds, J.A.; Cole, A. Microscopy and biology of uronema-gigas, a filamentous eukaryotic green-alga, and its associated tailed virus-like particle. *Virology* **1980**, *100*, 156–165. [CrossRef]
11. Chen, F.; Suttle, C.A. Amplification of DNA-polymerase gene fragments from viruses infecting microalgae. *Appl. Environ. Microbiol.* **1995**, *61*, 1274–1278. [PubMed]

12. 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*. [CrossRef] [PubMed]

13. Van Etten, J.L.; Burbank, D.E.; Kuczmarski, D.; Meints, R.H. Virus-infection of culturable chlorella-like algae and development of a plaque assay. *Science* **1983**, *219*, 994–996. [CrossRef] [PubMed]

14. Dunigan, D.D.; Cerny, R.L.; Bauman, A.T.; Roach, J.C.; Lane, L.C.; Agarkova, I.V.; Wulser, K.; Yanai-Balser, G.M.; Gurnon, J.R.; Vitek, J.C.; et al. *Paramecium bursaria* chlorella virus 1 proteome reveals novel architectural and regulatory features of a giant virus. *J. Virol.* **2012**, *86*, 8821–8834. [CrossRef] [PubMed]

15. Fitzgerald, L.A.; Graves, M.V.; Li, X.; Feldblyum, T.; Nierman, W.C.; Van Etten, J.L. Sequence and annotation of the 369-kb NY-2a and the 345-kb AR158 viruses that infect *Chlorella* NC64Aa. *Virology* **2007**, *358*, 472–484. [CrossRef] [PubMed]

16. Fitzgerald, L.A.; Graves, M.V.; Li, X.; Feldblyum, T.; Hartigan, J.; Van Etten, J.L. Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi. *Virology* **2007**, *358*, 459–471. [CrossRef] [PubMed]

17. Jeanniard, A.; Dunigan, D.D.; Gurnon, J.R.; Agarkova, I.V.; Kang, M.; Vitek, J.; Duncan, G.; McClung, O.W.; Larsen, M.; Claverie, J.M.; et al. Towards defining the chloroviruses: A genomic journey through a genus of large DNA viruses. *BMC Genomics* **2013**, *14*. [CrossRef] [PubMed]

18. Quispe, C.F.; Esmael, A.; Sonderman, O.; McQuinn, M.; Agarkova, I.; Battan, M.; Duncan, G.A.; Dunigan, D.D.; Smith, T.P.L.; De Castro, C.; et al. Characterization of a new chlorovirus type with permissive and non-permissive features on phylogenetically related algal strains. *Virology* **2017**, *500*, 103–113. [CrossRef] [PubMed]

19. Yanai-Balser, G.M.; Duncan, G.A.; Eudy, J.D.; Wang, D.; Li, X.; Agarkova, I.V.; Dunigan, D.D.; Van Etten, J.L. Microarray analysis of *Paramecium bursaria* chlorella virus 1 transcription. *J. Virol.* **2010**, *84*, 532–542. [CrossRef] [PubMed]

20. Blanc, G.; Mozar, M.; Agarkova, I.V.; Gurnon, J.R.; Yanai-Balser, G.; Rowe, J.M.; Xia, Y.N.; Riethoven, J.J.; Dunigan, D.D.; Van Etten, J.L. Deep RNA sequencing reveals hidden features and dynamics of early gene transcription in *Paramecium bursaria* chlorella virus 1. *PLoS ONE* **2014**, *9*, e90989. [CrossRef] [PubMed]

21. Grabherr, R.; Strasser, P.; Vanetten, J.L. The DNA-polymerase gene from chlorella viruses PBCV-1 and NY-2A contains an intron with nuclear splicing sequences. *Virology* **1992**, *188*, 721–731. [CrossRef]

22. Chen, F.; Suttle, C.A.; Short, S.M. Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Appl. Environ. Microbiol.* **1996**, *62*, 2869–2874. [PubMed]

23. Moniruzzaman, M.; LeCleir, G.R.; Brown, C.M.; Gobler, C.J.; Bidle, K.D.; Wilson, W.H.; Wilhelm, S.W. Genome of brown tide virus (AaV), the little giant of the megaviridae, elucidates ncldv genome expansion and host-virus coevolution. *Virology* **2014**, *466*, 60–70. [CrossRef] [PubMed]

24. Zhang, W.J.; Zhou, J.L.; Liu, T.G.; Yu, Y.X.; Pan, Y.J.; Yan, S.L.; Wang, Y.J. Four novel algal virus genomes discovered from yellowstone lake metagenomes. *Sci. Rep.* **2015**, *5*. [CrossRef] [PubMed]

25. Santini, S.; Jeudy, S.; Bartoli, J.; Poirot, O.; Lescot, M.; Abergel, C.; Barbe, V.; Wommack, K.E.; Noordeloos, A.A.M.; Brussaard, C.P.D.; et al. Genome of *Phaeocystis globosa* virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10800–10805. [CrossRef] [PubMed]

26. Yutin, N.; Koonin, E.V. Hidden evolutionary complexity of nucleo-cytoplasmic large DNA viruses of eukaryotes. *Virol. J.* **2012**, *9*. [CrossRef] [PubMed]

27. Yutin, N.; Koonin, E.V. Pandoraviruses are highly derived Phycodnaviruses. *Biol. Direct* **2013**, *8*. [CrossRef] [PubMed]

28. Maruyama, F.; Ueki, S. Evolution and phylogeny of large DNA viruses, *Mimiviridae* and *Phycodnaviridae* including newly characterized *Heterosigma akashiwo* virus. *Front. Microbiol.* **2016**, *7*. [CrossRef] [PubMed]

29. Culley, A.I.; Lang, A.S.; Suttle, C.A. High diversity of unknown picorna-like viruses in the sea. *Nature* **2003**, *424*, 1054–1057. [CrossRef] [PubMed]

30. 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]

31. Yau, S.; Lauro, F.M.; DeMaere, M.Z.; Brown, M.V.; Thomas, T.; Raftery, M.J.; Andrews-Pfannkoch, C.; Lewis, M.; Hoffman, J.M.; Gibson, J.A.; et al. Virophage control of antarctic algal host-virus dynamics. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6163–6168. [CrossRef] [PubMed]

32. Gong, C.W.; Zhang, W.J.; Zhou, X.W.; Wang, H.M.; Sun, G.W.; Xiao, J.Z.; Pan, Y.J.; Yan, S.L.; Wang, Y.J. Novel virophages discovered in a freshwater lake in China. *Front. Microbiol.* **2016**, *7*. [CrossRef] [PubMed]

33. Roux, S.; Chan, L.K.; Egan, R.; Malmstrom, R.R.; McMahon, K.D.; Sullivan, M.B. Ecogenomics of virophages and their giant virus hosts assessed through time series metagenomics. *Nat. Commun.* **2017**, *8*. [CrossRef] [PubMed]

34. Allen, L.Z.; 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]

35. Burki, F. The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*. [CrossRef] [PubMed]

36. De Vargas, C.; Audic, S.; Henry, N.; Decelle, J.; Mahe, F.; Logares, R.; Lara, E.; Berney, C.; Le Bescot, N.; Probert, I.; et al. Eukaryotic plankton diversity in the sunlit ocean. *Science* **2015**, *348*, 1261605. [CrossRef] [PubMed]

37. Kataoka, T.; Yamaguchi, H.; Sato, M.; Watanabe, T.; Taniuchi, Y.; Kuwata, A.; Kawachi, M. Seasonal and geographical distribution of near-surface small photosynthetic eukaryotes in the western North Pacific determined by pyrosequencing of 18s rDNA. *FEMS Microbiol. Ecol.* **2017**, *93*. [CrossRef] [PubMed]

38. Gowing, M.M. Large viruses and infected microeukaryotes in ross sea summer pack ice habitats. *Mar. Biol.* **2003**, *142*, 1029–1040. [CrossRef]

39. Honjo, T. Overview on bloom dynamics and physiological ecology of Heterosigma-akashiwo. In *Toxic Phytoplankton Blooms in the Sea, Proceedings of the Fifth International Conference on Toxic Marine Phytoplankton, Newport, Rhode Island, USA, 28 October–1 November 1991*; Smayda, T.J., Shimizu, Y., Eds.; Elsevier: Philadelphia, PA, USA, 1993; Volume 3, pp. 33–41.

40. Karosiene, J.; Kasperoviciene, J.; Koreiviene, J.; Savadova, K.; Vitonyte, I. Factors promoting persistence of the bloom-forming *Gonyostomum semen* in temperate lakes. *Limnologica* **2016**, *60*, 51–58. [CrossRef]

41. Leon-Munoz, J.; Urbina, M.A.; Garreaud, R.; Iriarte, J.L. Hydroclimatic conditions trigger record harmful algal bloom in western Patagonia (summer 2016). *Sci. Rep.* **2018**, *8*, 1330. [CrossRef] [PubMed]

42. Gustafson, D.E.; Stoecker, D.K.; Johnson, M.D.; Van Heukelem, W.F.; Sneider, K. Cryptophyte algae are robbed of their organelles by the marine ciliate *Mesodinium rubrum*. *Nature* **2000**, *405*, 1049–1052. [CrossRef] [PubMed]

43. Vermont, A.I.; Martinez, J.M.; Waller, J.D.; Gilg, I.C.; Leavitt, A.H.; Floge, S.A.; Archer, S.D.; Wilson, W.H.; Fields, D.M. Virus infection of *Emiliania huxleyi* deters grazing by the copepod *Acartia tonsa*. *J. Plankton Res.* **2016**, *38*, 1194–1205. [CrossRef]

44. Evans, C.; Wilson, W.H. Preferential grazing of *Oxyrrhis marina* on virus-infected *Emiliania huxleyi*. *Limnol. Oceanogr.* **2008**, *53*, 2035–2040. [CrossRef]

45. Schvarcz, C.R.; Steward, G.F. A giant virus infecting green algae encodes key fermentation genes. *Virology* **2018**, *518*, 423–433. [CrossRef] [PubMed]

46. Pagarete, A.; Grebert, T.; Stepanova, O.; Sandaa, R.A.; Bratbak, G. Tsv-N1: A novel DNA algal virus that infects *Tetraselmis striata*. *Viruses* **2015**, *7*, 3937–3953. [CrossRef] [PubMed]

47. Derelle, E.; Monier, A.; Cooke, R.; Worden, A.Z.; Grimsley, N.H.; Moreau, H. Diversity of viruses infecting the green microalga *Oostreococcus lucimarinus*. *J. Virol.* **2015**, *89*, 5812–5821. [CrossRef] [PubMed]

48. Weynberg, K.D.; Allen, M.J.; Gilg, I.C.; Scanlan, D.J.; Wilson, W.H. Genome sequence of *Ostreococcus tauri* virus OtV-2 throws light on the role of picoeukaryote niche separation in the ocean. *J. Virol.* **2011**, *85*, 4520–4529. [CrossRef] [PubMed]

49. Moreau, H.; Piganeau, G.; Desdevises, Y.; Cooke, R.; Derelle, E.; Grimsley, N. Marine prasinovirus genomes show low evolutionary divergence and acquisition of protein metabolism genes by horizontal gene transfer. *J. Virol.* **2010**, *84*, 12555–12563. [CrossRef] [PubMed]

50. Martinez, J.M.; Boere, A.; Gilg, L.; van Lent, J.W.M.; Witte, H.J.; van Bleijswijk, J.D.L.; Brussaard, C.P.D. New lipid envelope-containing dsdna virus isolates infecting *Micromonas pusilla* reveal a separate phylogenetic group. *Aquat. Microb. Ecol.* **2015**, *74*, 17–28. [CrossRef]

51. Finke, J.F.; Winget, D.M.; Chan, A.M.; Suttle, C.A. Variation in the genetic repertoire of viruses infecting *Micromonas pusilla* reflects horizontal gene transfer and links to their environmental distribution. *Viruses* **2017**, *9*, 116. [CrossRef] [PubMed]

52. Brussaard, C.P.D.; Noordeloos, A.A.M.; Sandaa, R.A.; Heldal, M.; Bratbak, G. Discovery of a dsRNA virus infecting the marine photosynthetic protist *Micromonas pusilla*. *Virology* **2004**, *319*, 280–291. [CrossRef] [PubMed]

53. Maat, D.S.; Biggs, T.; Evans, C.; van Bleijswijk, J.D.L.; van der Wel, N.N.; Dutilh, B.E.; Brussaard, C.P.D. Characterization and temperature dependence of Arctic *Micromonas polaris* viruses. *Viruses* **2017**, *9*. [CrossRef] [PubMed]

54. Sandaa, R.A.; Heldal, M.; Castberg, T.; Thyrhaug, R.; Bratbak, G. Isolation and characterization of two viruses with large genome size infecting *Chrysochromulina ericina* (prymnesiophyceae) and *Pyramimonas orientalis* (prasinophyceae). *Virology* **2001**, *290*, 272–280. [CrossRef] [PubMed]

55. Rousvoal, S.; Bouyer, B.; Lopez-Cristoffanini, C.; Boyen, C.; Collen, J. Mutant swarms of a totivirus-like entities are present in the red macroalga *Chondrus crispus* and have been partially transferred to the nuclear genome. *J. Phycol.* **2016**, *52*, 493–504. [CrossRef] [PubMed]

56. Ogata, H.; Toyoda, K.; Tomaru, Y.; Nakayama, N.; Shirai, Y.; Claverie, J.M.; Nagasaki, K. Remarkable sequence similarity between the dinoflagellate-infecting marine girus and the terrestrial pathogen african swine fever virus. *Virol. J.* **2009**, *6*, 178. [CrossRef] [PubMed]

57. Tomaru, Y.; Katanozaka, N.; Nishida, K.; Shirai, Y.; Tarutani, K.; Yamaguchi, M.; Nagasaki, K. Isolation and characterization of two distinct types of HcRNAV, a single-stranded RNA virus infecting the bivalve-killing microalga *Heterocapsa circularisquama*. *Aquat. Microb. Ecol.* **2004**, *34*, 207–218. [CrossRef]

58. Kim, J.; Kim, C.H.; Takano, Y.; Jang, I.K.; Kim, S.W.; Choi, T.J. Isolation and physiological characterization of a new algicidal virus infecting the harmful dinoflagellate *Heterocapsa pygmaea*. *Plant Pathol. J.* **2012**, *28*, 433–438. [CrossRef]

59. Onji, M.; Nakano, S.; Suzuki, S. Virus-like particles suppress growth of the red-tide-forming marine dinoflagellate *Gymnodinium mikimotoi*. *Mar. Biotechnol.* **2003**, *5*, 435–442. [PubMed]

60. Bettarel, Y.; Kan, J.; Wang, K.; Williamson, K.E.; Cooney, S.; Ribblett, S.; Chen, F.; Wommack, K.E.; Coats, D.W. Isolation and preliminary characterisation of a small nuclear inclusion virus infecting the diatom *Chaetoceros* cf. *gracilis*. *Aquat. Microb. Ecol.* **2005**, *40*, 103–114. [CrossRef]

61. Nagasaki, K.; Tomaru, Y.; Takao, Y.; Nishida, K.; Shirai, Y.; Suzuki, H.; Nagumo, T. Previously unknown virus infects marine diatom. *Appl. Environ. Microbiol.* **2005**, *71*, 3528–3535. [CrossRef] [PubMed]

62. Tomaru, Y.; Toyoda, K.; Suzuki, H.; Nagumo, T.; Kimura, K.; Takao, Y. New single-stranded DNA virus with a unique genomic structure that infects marine diatom *Chaetoceros setoensis*. *Sci. Rep.* **2013**, *3*. [CrossRef] [PubMed]

63. Tomaru, Y.; Takao, Y.; Suzuki, H.; Nagumo, T.; Nagasaki, K. Isolation and characterization of a single-stranded RNA virus infecting the bloom-forming diatom *Chaetoceros socialis*. *Appl. Environ. Microbiol.* **2009**, *75*, 2375–2381. [CrossRef] [PubMed]

64. Tomaru, Y.; Takao, Y.; Suzuki, H.; Nagumo, T.; Koike, K.; Nagasaki, K. Isolation and characterization of a single-stranded DNA virus infecting *Chaetoceros lorenzianus* Grunow. *Appl. Environ. Microbiol.* **2011**, *77*, 5285–5293. [CrossRef] [PubMed]

65. Tomaru, Y.; Shirai, Y.; Toyoda, K.; Nagasaki, K. Isolation and characterisation of a single-stranded DNA virus infecting the marine planktonic diatom *Chaetoceros tenuissimus*. *Aquat. Microb. Ecol.* **2011**, *64*, 175–184. [CrossRef]

66. Kimura, K.; Tomarua, Y. Discovery of two novel viruses expands the diversity of single-stranded DNA and single-stranded RNA viruses infecting a cosmopolitan marine diatom. *Appl. Environ. Microbiol.* **2015**, *81*, 1120–1131. [CrossRef] [PubMed]

67. Shirai, Y.; Tomaru, Y.; Takao, Y.; Suzuki, H.; Nagumo, T.; Nagasaki, K. Isolation and characterization of a single-stranded RNA virus infecting the marine planktonic diatom *Chaetoceros tenuissimus* Meunier. *Appl. Environ. Microbiol.* **2008**, *74*, 4022–4027. [CrossRef] [PubMed]

68. Kimura, K.; Tomaru, Y. Isolation and characterization of a single-stranded DNA virus infecting the marine diatom *Chaetoceros* sp. strain SS628-11 isolated from western Japan. *PLoS ONE* **2013**, *8*, e82013. [CrossRef] [PubMed]

69. Toyoda, K.; Kimura, K.; Hata, N.; Nakayama, N.; Nagasaki, K.; Tomaru, Y. Isolation and characterization of a single-stranded DNA virus infecting the marine planktonic diatom *Chaetoceros* sp (strain TG07-C28). *Plankton Benthos. Res.* **2012**, *7*, 20–28. [CrossRef]

70. Tomaru, Y.; Shirai, Y.; Suzuki, H.; Nagumo, T.; Nagasaki, K. Isolation and characterization of a new single-stranded DNA virus infecting the cosmopolitan marine diatom *Chaetoceros dehilis*. *Aquat. Microb. Ecol.* **2008**, *50*, 103–112. [CrossRef]

71. Tomaru, Y.; Toyoda, K.; Kimura, K.; Takao, Y.; Sakurada, K.; Nakayama, N.; Nagasaki, K. Isolation and characterization of a single-stranded RNA virus that infects the marine planktonic diatom *Chaetoceros* sp. (SS08-c03). *Phycol. Res.* **2013**, *61*, 27–36. [CrossRef]

72. Eissler, Y.; Wang, K.; Chen, F.; Wommack, K.E.; Coats, D.W. Ultrastructural characterization of the lytic cycle of an intracellular virus infecting the diatom *Chaetoceros cf wighamii* (Bacillariophyceae) from Chesapeake Bay, USA. *J. Phycol.* **2009**, *45*, 787–797. [CrossRef] [PubMed]

73. Tomaru, Y.; Toyoda, K.; Kimura, K.; Hata, N.; Yoshida, M.; Nagasaki, K. First evidence for the existence of pennate diatom viruses. *ISME J.* **2012**, *6*, 1445–1448. [CrossRef] [PubMed]

74. Nagasaki, K.; Tomaru, Y.; Katanozaka, N.; Shirai, Y.; Nishida, K.; Itakura, S.; Yamaguchi, M. Isolation and characterization of a novel single-stranded RNA virus infecting the bloom-forming diatom *Rhizosolenia setigera*. *Appl. Environ. Microbiol.* **2004**, *70*, 704–711. [CrossRef] [PubMed]

75. Kim, J.; Kim, C.H.; Youn, S.H.; Choi, T.J. Isolation and physiological characterization of a novel algicidal virus infecting the marine diatom *Skeletonema costatum*. *Plant Pathol. J.* **2015**, *31*, 186–191. [CrossRef] [PubMed]

76. Kim, J.; Yoon, S.H.; Choi, T.J. Isolation and physiological characterization of a novel virus infecting *Stephanopyxis palmeriana* (Bacillariophyta). *Algae* **2015**, *30*, 81–87. [CrossRef]

77. Kapp, M.; Knippers, R.; Mueller, D.G. New members of a group of DNA viruses infecting brown algae. *Phycol. Res.* **1997**, *45*, 85–90. [CrossRef]

78. Henry, E.C.; Meints, R.H. A persistent virus-infection in *Feldmannia* (Phaeophyceae). *J. Phycol.* **1992**, *28*, 517–526. [CrossRef]

79. Maier, I.; Wolf, S.; Delaroque, N.; Muller, D.G.; Kawai, H. A DNA virus infecting the marine brown alga *Pilayella littoralis* (Ectocarpales, Phaeophyceae) in culture. *Eur. J. Phycol.* **1998**, *33*, 213–220. [CrossRef]

80. Nagasaki, K.; Yamaguchi, M. Isolation of a virus infectious to the harmful bloom causing microalga *Heterosigma akashiwo* (Raphidophyceae). *Aquat. Microb. Ecol.* **1997**, *13*, 135–140. [CrossRef]

81. Lawrence, J.E.; Brussaard, C.P.D.; Suttle, C.A. Virus-specific responses of *Heterosigma akashiwo* to infection. *Appl. Environ. Microbiol.* **2006**, *72*, 7829–7834. [CrossRef] [PubMed]

82. Tai, V.; Lawrence, J.E.; Lang, A.S.; Chan, A.M.; Culley, A.I.; Suttle, C.A. Characterization of HaRNAV, a single-stranded RNA virus causing lysis of Heterosigma akashiwo (Raphidophyceae). *J. Phycol.* **2003**, *39*, 343–352. [CrossRef]

83. Lawrence, J.E.; Chan, A.M.; Suttle, C.A. A novel virus (HaNIV) causes lysis of the toxic bloom-forming alga *Heterosigma akashiwo* (Raphidophyceae). *J. Phycol.* **2001**, *37*, 216–222. [CrossRef]

84. Castberg, T.; Thyrhaug, R.; Larsen, A.; Sandaa, R.A.; Heldal, M.; Van Etten, J.L.; Bratbak, G. Isolation and characterization of a virus that infects *Emiliania huxleyi* (haptophyta). *J. Phycol.* **2002**, *38*, 767–774. [CrossRef]

85. Baudoux, A.C.; Brussaard, C.P.D. Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virology* **2005**, *341*, 80–90. [CrossRef] [PubMed]

86. Wilson, W.H.; Schroeder, D.C.; Ho, J.; Canty, M. Phylogenetic analysis of PgV-102P, a new virus from the english channel that infects phaeocystis globosa. *J. Marine Biol. Assoc. UK* **2006**, *86*, 485–490. [CrossRef]

87. Jacobsen, A.; Bratbak, G.; Heldal, M. Isolation and characterization of a virus infecting *Phaeocystis pouchetii* (Prymnesiophyceae). *J. Phycol.* **1996**, *32*, 923–927. [CrossRef]

88. Suttle, C.A.; Chan, A.M. Viruses infecting the marine *Prymnesiophyte chrysochromulina* spp.: Isolation, preliminary characterization, and natural-abundance. *Mar. Ecol. Progr. Ser.* **1995**, *118*, 275–282. [CrossRef]

89. Mirza, S.F.; Staniewski, M.A.; Short, C.M.; Long, A.M.; Chaban, Y.V.; Short, S.M. Isolation and characterization of a virus infecting the freshwater algae *Chrysochromulina parva*. *Virology* **2015**, *486*, 105–115. [CrossRef] [PubMed]

90. Johannessen, T.V.; Bratbak, G.; Larsen, A.; Ogata, H.; Egge, E.S.; Edvardsen, B.; Eikrem, W.; Sandaa, R.A. Characterisation of three novel giant viruses reveals huge diversity among viruses infecting *Prymnesiales* (Haptophyta). *Virology* **2015**, *476*, 180–188. [CrossRef] [PubMed]

91. Wagstaff, B.A.; Vladu, I.C.; Barclay, J.E.; Schroeder, D.C.; Malin, G.; Field, R.A. Isolation and characterization of a double stranded DNA megavirus infecting the toxin-producing haptophyte *Prymnesium parvum*. *Viruses* **2017**, *9*, 40. [CrossRef] [PubMed]

92. Nagasaki, K.; Kim, J.-J.; Tomaru, Y.; Takao, Y.; Nagai, S. Isolation and characterization of a novel virus infecting *Teleaulax amphioxeia* (Cryptophyceae). *Plankton Benthos Res.* **2009**, *4*, 122–124. [CrossRef]

93. Brussaard, C.P.D.; Kempers, R.S.; Kop, A.J.; Riegman, R.; Heldal, M. Virus-like particles in a summer bloom of *Emiliania huxleyi* in the North Sea. *Aquat. Microb. Ecol.* **1996**, *10*, 105–113. [CrossRef]

94. Moestrup, H.T.; Thomsen, H.A. An ultrastructural study of the flagellate *Pyramimonas orientalis* with particular emphasis on golgi apparatus activity and the flagellar apparatus. *Protoplasma* **1974**, *81*, 247–269. [CrossRef]

95. La Scola, B.; Desnues, C.; Pagnier, I.; Robert, C.; Barrassi, L.; Fournous, G.; Merchat, M.; Suzan-Monti, M.; Forterre, P.; Koonin, E.; et al. The virophage as a unique parasite of the giant mimivirus. *Nature* **2008**, *455*, 100–104. [CrossRef] [PubMed]

96. Fischer, M.G.; Suttle, C.A. A virophage at the origin of large DNA transposons. *Science* **2011**, *332*, 231–234. [CrossRef] [PubMed]

97. Mojica, K.D.A.; Brussaard, C.P.D. Factors affecting virus dynamics and microbial host-virus interactions in marine environments. *FEMS Microbiol. Ecol.* **2014**, *89*, 495–515. [CrossRef] [PubMed]

98. Short, S.M. The ecology of viruses that infect eukaryotic algae. *Environ. Microbiol.* **2012**, *14*, 2253–2271. [CrossRef] [PubMed]

99. Brussaard, C.P.D.; Short, S.M.; Frederickson, C.M.; Suttle, C.A. Isolation and phylogenetic analysis of novel viruses infecting the phytoplankton *Phaeocystis globosa* (Prymnesiophyceae). *Appl. Environ. Microbiol.* **2004**, *70*, 3700–3705. [CrossRef] [PubMed]

100. Iyer, L.M.; Aravind, L.; Koonin, E.V. Common origin of four diverse families of large eukaryotic DNA viruses. *J. Virol.* **2001**, *75*, 11720–11734. [CrossRef] [PubMed]

101. Gallot-Lavallee, L.; Blanc, G.; Claverie, J.M. Comparative genomics of *Chrysochromulina ericina* virus and other microalga-infecting large DNA viruses highlights their intricate evolutionary relationship with the established *Mimiviridae* family. *J. Virol.* **2017**, *91*. [CrossRef] [PubMed]

102. Allen, M.J.; Schroeder, D.C.; Holden, M.T.G.; Wilson, W.H. Evolutionary history of the Coccolithoviridae. *Mol. Biol. Evol.* **2006**, *23*, 86–92. [CrossRef] [PubMed]

103. Dunigan, D.D.; Fitzgerald, L.A.; Van Etten, J.L. Phycodnaviruses: A peek at genetic diversity. *Virus Res.* **2006**, *117*, 119–132. [CrossRef] [PubMed]

104. Nissimov, J.I.; Pagarete, A.; Ma, F.; Cody, S.; Dunigan, D.D.; Kimmance, S.A.; Allen, M.J. Coccolithoviruses: A review of cross-kingdom genomic thievery and metabolic thuggery. *Viruses* **2017**, *9*, 52. [CrossRef] [PubMed]

105. Clerissi, C.; Grimsley, N.; Ogata, H.; Hingamp, P.; Poulain, J.; Desdevises, Y. Unveiling of the diversity of Prasinoviruses (*Phycodnaviridae*) in marine samples by using high-throughput sequencing analyses of PCR-amplified DNA polymerase and major capsid protein genes. *Appl. Environ. Microbiol.* **2014**, *80*, 3150–3160. [CrossRef] [PubMed]

106. Filee, J. Genomic comparison of closely related giant viruses supports an accordion-like model of evolution. *Front. Microbiol.* **2015**, *6*. [CrossRef] [PubMed]

107. Filee, J.; Pouget, N.; Chandler, M. Phylogenetic evidence for extensive lateral acquisition of cellular genes by nucleocytoplasmic large DNA viruses. *BMC Evol. Biol.* **2008**, *8*. [CrossRef] [PubMed]

108. Malviya, S.; Scalco, E.; Audic, S.; Vincenta, F.; Veluchamy, A.; Poulain, J.; Wincker, P.; Iudicone, D.; de Vargas, C.; Bittner, L.; et al. Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1516–E1525. [CrossRef] [PubMed]

109. Kazlauskas, D.; Dayaram, A.; Kraberger, S.; Goldstien, S.; Varsani, A.; Krupovic, M. Evolutionary history of ssDNA Bacilladnaviruses features horizontal acquisition of the capsid gene from ssRNA Nodaviruses. *Virology* **2017**, *504*, 114–121. [CrossRef] [PubMed]

110. Diemer, G.S.; Stedman, K.M. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol. Direct* **2012**, *7*, 13. [CrossRef] [PubMed]

111. Krupovic, M.; Koonin, E.V. Evolution of eukaryotic single-stranded DNA viruses of the Bidnaviridae family from genes of four other groups of widely different viruses. *Sci. Rep.* **2014**, *4*, 5347. [CrossRef] [PubMed]

112. King, A.M.Q.; Lefkowitz, E.J.; Mushegian, A.R.; Adams, M.J.; Dutilh, B.E.; Gorbalenya, A.E.; Harrach, B.; Harrison, R.L.; Junglen, S.; Knowles, N.J.; et al. Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2018). *Arch. Virol.* **2018**. [CrossRef] [PubMed]

113. 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] [PubMed]

114. Delaroque, N.; Maier, I.; Knippers, R.; Muller, D.G. Persistent virus integration into the genome of its algal host, *Ectocarpus siliculosus* (Phaeophyceae). *J. Gen. Virol.* **1999**, *80*, 1367–1370. [CrossRef] [PubMed]

115. Nagai, K.; Matsuyama, Y.; Uchida, T.; Yamaguchi, M.; Ishimura, M.; Nishimura, A.; Akamatsu, S.; Honjo, T. Toxicity and LD (50) levels of the red tide dinoflagellate *Heterocapsa circularisquama* on juvenile pearl oysters. *Aquaculture* **1996**, *144*, 149–154. [CrossRef]

116. Nagasaki, K. Dinoflagellates, diatoms, and their viruses. *J. Microbiol.* **2008**, *46*, 235–243. [CrossRef] [PubMed]

117. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]

118. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef] [PubMed]

119. Noble, R.T.; Fuhrman, J.A. Use of Sybr green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquat. Microb. Ecol.* **1998**, *14*, 113–118. [CrossRef]

120. Brussaard, C.P.D.; Marie, D.; Bratbak, G. Flow cytometric detection of viruses. *J. Virol. Methods* **2000**, *85*, 175–182. [CrossRef]

121. Brussaard, C.P.D. Optimization of procedures for counting viruses by flow cytometry. *Appl. Environ. Microbiol.* **2004**, *70*, 1506–1513. [CrossRef] [PubMed]

122. Marie, D.; Brussaard, C.P.D.; Thyrhaug, R.; Bratbak, G.; Vaulot, D. Enumeration of marine viruses in culture and natural samples by flow cytometry. *Appl. Environ. Microbiol.* **1999**, *65*, 45–52.

123. Larsen, J.B.; Larsen, A.; Bratbak, G.; Sandaa, R.A. Phylogenetic analysis of members of the *Phycodnaviridae* virus family, using amplified fragments of the major capsid protein gene. *Appl. Environ. Microbiol.* **2008**, *74*, 3048–3057. [CrossRef] [PubMed]

124. Moniruzzaman, M.; Gann, E.R.; LeCleir, G.R.; Kang, Y.; Gobler, C.J.; Wilhelm, S.W. Diversity and dynamics of algal megaviridae members during a harmful brown tide caused by the pelagophyte, *Aureococcus anophagefferens*. *FEMS Microbiol. Ecol.* **2016**, *92*. [CrossRef] [PubMed]

125. Kislyuk, A.O.; Haegeman, B.; Bergman, N.H.; Weitz, J.S. Genomic fluidity: An integrative view of gene diversity within microbial populations. *BMC Genomics* **2011**, *12*, 10. [CrossRef] [PubMed]

126. Lefever, S.; Pattyn, F.; Hellemans, J.; Vandesompele, J. Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clin. Chem.* **2013**, *59*, 1470–1480. [CrossRef] [PubMed]

127. Gulvik, C.A.; Effler, T.C.; Wilhelm, S.W.; Buchan, A. De-metast-blast: A tool for the validation of degenerate primer sets and data mining of publicly available metagenomes. *PLoS ONE* **2012**, *7*, e50362. [CrossRef] [PubMed]

128. Short, S.M.; Suttle, C.A. Sequence analysis of marine virus communities reveals that groups of related algal viruses are widely distributed in nature. *Appl. Environ. Microbiol.* **2002**, *68*, 1290–1296. [CrossRef] [PubMed]

129. Short, S.M.; Short, C.M. Diversity of algal viruses in various North American freshwater environments. *Aquat. Microb. Ecol.* **2008**, *51*, 13–21. [CrossRef]

130. Clasen, J.L.; Suttle, C.A. Identification of freshwater *Phycodnaviridae* and their potential phytoplankton hosts, using DNA *pol* sequence fragments and a genetic-distance analysis. *Appl. Environ. Microbiol.* **2009**, *75*, 991–997. [CrossRef] [PubMed]

131. Rowe, J.M.; Fabre, M.F.; Gobena, D.; Wilson, W.H.; Wilhelm, S.W. Application of the major capsid protein as a marker of the phylogenetic diversity of *Emiliania huxleyi* viruses. *FEMS Microbiol. Ecol.* **2011**, *76*, 373–380. [CrossRef] [PubMed]

132. Wilson, W.H.; Gilg, I.C.; Duarte, A.; Ogata, H. Development of DNA mismatch repair gene, *muts*, as a diagnostic marker for detection and phylogenetic analysis of algal megaviruses. *Virology* **2014**, *466*, 123–128. [CrossRef] [PubMed]

133. Culley, A.I.; Steward, G.F. New genera of RNA viruses in subtropical seawater, inferred from polymerase gene sequences. *Appl. Environ. Microbiol.* **2007**, *73*, 5937–5944. [CrossRef] [PubMed]

134. Culley, A. New insight into the RNA aquatic virosphere via viromics. *Virus Res.* **2018**, *244*, 84–89. [CrossRef] [PubMed]

135. Sullivan, M.B. Viromes, not gene markers, for studying double-stranded DNA virus communities. *J. Virol.* **2015**, *89*, 2459–2461. [CrossRef] [PubMed]

136. Polz, M.F.; Cavanaugh, C.M. Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* **1998**, *64*, 3724–3730. [PubMed]

137. Short, S.M.; Short, C.M. Quantitative PCR reveals transient and persistent algal viruses in Lake Ontario, Canada. *Environ. Microbiol.* **2009**, *11*, 2639–2648. [CrossRef] [PubMed]

138. Short, C.M.; Rusanova, O.; Short, S.M. Quantification of virus genes provides evidence for seed-bank populations of phycodnaviruses in Lake Ontario, Canada. *ISME J.* **2011**, *5*, 810–821. [CrossRef] [PubMed]

139. Baran, N.; Goldin, S.; Maidanik, I.; Lindell, D. Quantification of diverse virus populations in the environment using the polony method. *Nat. Microbiol.* **2018**, *3*, 62–72. [CrossRef] [PubMed]

140. Tadmor, A.D.; Ottesen, E.A.; Leadbetter, J.R.; Phillips, R. Probing individual environmental bacteria for viruses by using microfluidic digital PCR. *Science* **2011**, *333*, 58–62. [CrossRef] [PubMed]

141. Culley, A.I.; Mueller, J.A.; Belcaid, M.; Wood-Charlson, E.M.; Poisson, G.; Steward, G.F. The characterization of RNA viruses in tropical seawater using targeted PCR and metagenomics. *mBio* **2014**, *5*, e01210-14. [CrossRef] [PubMed]

142. Simmonds, P.; Adams, M.J.; Benko, M.; Breitbart, M.; Brister, J.R.; Carstens, E.B.; Davison, A.J.; Delwart, E.; Gorbalenya, A.E.; Harrach, B.; et al. Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* **2017**, *15*, 161–168. [CrossRef] [PubMed]

143. Varsani, A.; Krupovic, M. Smacoviridae: A new family of animal-associated single-stranded DNA viruses. *Arch. Virol.* **2018**, *163*, 2005–2015. [CrossRef] [PubMed]

144. Adams, M.J.; Lefkowitz, E.J.; King, A.M.Q.; Harrach, B.; Harrison, R.L.; Knowles, N.J.; Kropinski, A.M.; Krupovic, M.; Kuhn, J.H.; Mushegian, A.R.; et al. Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2017). *Arch. Virol.* **2017**, *162*, 2505–2538. [CrossRef] [PubMed]

145. Wommack, K.E.; Hill, R.T.; Colwell, R.R. A simple method for the concentration of viruses from natural water samples. *J. Virol. Methods* **1995**, *22*, 57–67. [CrossRef]

146. 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] [PubMed]

147. Roux, S.; Solonenko, N.E.; Dang, V.T.; Poulos, B.T.; Schwenk, S.M.; Goldsmith, D.B.; Coleman, M.L.; Breitbare, M.; Sullivan, M.B. Towards quantitative viromics for both double-stranded and single-stranded DNA viruses. *PeerJ* **2016**, *4*. [CrossRef] [PubMed]

148. Kobayashi, K.; Tomita, R.; Sakamoto, M. Recombinant plant dsRNA-binding protein as an effective tool for the isolation of viral replicative form dsRNA and universal detection of RNA viruses. *J. Gen. Plant Pathol.* **2009**, *75*, 87–91. [CrossRef]

149. Andrews-Pfannkoch, C.; Fadrosh, D.W.; Thorpe, J.; Williamson, S.J. Hydroxyapatite-mediated separation of double-stranded DNA, single-stranded DNA, and RNA genomes from natural viral assemblages. *Appl. Environ. Microbiol.* **2010**, *76*, 5039–5045. [CrossRef] [PubMed]

150. Shatkin, A.J. Animal RNA viruses-genome structure and function. *Annu. Rev. Biochem.* **1974**, *43*, 643–665. [CrossRef] [PubMed]

151. Steward, G.F.; Culley, A. Extraction and purification of nucleic acids from viruses. In *Manual of Aquatic Viral Ecology*; ASLO: Waco, TX, USA, 2010; pp. 154–165.

152. 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]

153. Nooij, S.; Schmitz, D.; Vennema, H.; Kroneman, A.; Koopmans, M.P.G. Overview of virus metagenomic classification methods and their biological applications. *Front. Microbiol.* **2018**, *9*, 21. [CrossRef] [PubMed]

154. Skewes-Cox, P.; Sharpton, T.J.; Pollard, K.S.; DeRisi, J.L. Profile Hidden Markov Models for the detection of viruses within metagenomic sequence data. *PLoS ONE* **2014**, *9*, e105067. [CrossRef] [PubMed]

155. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **2013**, *41*, e121. [CrossRef] [PubMed]

156. Matsen, F.A.; Kodner, R.B.; Armbrust, E.V. Pplacer: Linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinform.* **2010**, *11*. [CrossRef] [PubMed]

157. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650. [CrossRef] [PubMed]

158. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef] [PubMed]

159. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-tree: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef] [PubMed]

160. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **2001**, *17*, 754–755. [CrossRef] [PubMed]

161. Gallot-Lavallee, L.; Blanc, G. A glimpse of nucleo-cytoplasmic large DNA virus biodiversity through the eukaryotic genomicswindow. *Viruses* **2017**, *9*, 17. [CrossRef] [PubMed]

162. Roux, S.; Emerson, J.B.; Eloe-Fadrosh, E.A.; Sullivan, M.B. Benchmarking viromics: An in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. *PeerJ* **2017**, *5*. [CrossRef] [PubMed]

163. Sullivan, M.B.; Weitz, J.S.; Wilhelm, S.W. Viral ecology comes of age. *Environ. Microbiol. Rep.* **2017**, *9*, 33–35. [CrossRef] [PubMed]

164. Allers, E.; Moraru, C.; Duhaime, M.B.; Beneze, E.; Solonenko, N.; Barrero-Canosa, J.; Amann, R.; Sullivan, M.B. Single-cell and population level viral infection dynamics revealed by phagefish, a method to visualize intracellular and free viruses. *Environ. Microbiol.* **2013**, *15*, 2306–2318. [CrossRef] [PubMed]

165. Deng, L.; Ignacio-Espinoza, J.C.; Gregory, A.C.; Poulos, B.T.; Weitz, J.S.; Hugenholtz, P.; Sullivan, M.B. Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature* **2014**, *513*, 242–245. [CrossRef] [PubMed]

166. Weiss, S.; Van Treuren, W.; Lozupone, C.; Faust, K.; Friedman, J.; Deng, Y.; Xia, L.C.; Xu, Z.Z.; Ursell, L.; Alm, E.J.; et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J.* **2016**, *10*, 1669–1681. [CrossRef] [PubMed]

167. Bolduc, B.; Youens-Clark, K.; Roux, S.; Hurwitz, B.L.; Sullivan, M.B. Ivirus: Facilitating new insights in viral ecology with software and community data sets imbedded in a cyberinfrastructure. *ISME J.* **2017**, *11*, 7–14. [CrossRef] [PubMed]

168. Flaviani, F.; Schroeder, D.C.; Balestreri, C.; Schroeder, J.L.; Moore, K.; Paszkiewicz, K.; Pfaff, M.C.; Rybicki, E.P. A pelagic microbiome (viruses to protists) from a small cup of seawater. *Viruses* **2017**, *9*, 47. [CrossRef] [PubMed]

169. Wilson, W.H.; Gilg, I.C.; Moniruzzaman, M.; Field, E.K.; Koren, S.; LeCleir, G.R.; Martinez, J.M.; Poulton, N.J.; Swan, B.K.; Stepanauskas, R.; et al. Genomic exploration of individual giant ocean viruses. *ISME J.* **2017**, *11*, 1736–1745. [CrossRef] [PubMed]

170. Martinez-Hernandez, F.; Fornas, O.; Gomez, M.L.; Bolduc, B.; de la Cruz Pena, M.J.; Martinez, J.M.; Anton, J.; Gasol, J.M.; Rosselli, R.; Rodriguez-Valera, F.; et al. Single-virus genomics reveals hidden cosmopolitan and abundant viruses. *Nat. Commun.* **2017**, *8*, 15892. [CrossRef] [PubMed]

171. Wang, I.N.; Li, Y.; Que, Q.D.; Bhattacharya, M.; Lane, L.C.; Chaney, W.G.; Vanetten, J.L. Evidence for virus-encoded glycosylation specificity. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3840–3844. [CrossRef] [PubMed]

172. Van Etten, J.L.; Gurnon, J.R.; Yanai-Balser, G.M.; Dunigan, D.D.; Graves, M.V. *Chlorella* viruses encode most, if not all, of the machinery to glycosylate their glycoproteins independent of the endoplasmic reticulum and golgi. *Biochim. Biophys. Acta* **2010**, *1800*, 152–159. [CrossRef] [PubMed]

173. Van Etten, J.L.; Schuster, A.M.; Girton, L.; Burbank, D.E.; Swinton, D.; Hattman, S. DNA methylation of viruses infecting a eukaryotic *Chlorella*-like green alga. *Nucleic Acids Res.* **1985**, *13*, 3471–3478. [CrossRef] [PubMed]

174. Agarkova, I.V.; Dunigan, D.D.; Van Etten, J.L. Virion-associated restriction endonucleases of chloroviruses. *J. Virol.* **2006**, *80*, 8114–8123. [CrossRef] [PubMed]

175. Agarkova, I.; Dunigan, D.; Gurnon, J.; Greiner, T.; Barres, J.; Thiel, G.; Van Etten, J.L. Chlorovirus-mediated membrane depolarization of *Chlorella* alters secondary active transport of solutes. *J. Virol.* **2008**, *82*, 12181–12190. [CrossRef] [PubMed]

176. Rowe, J.M.; Jeanniard, A.; Gurnon, J.R.; Xia, Y.N.; Dunigan, D.D.; Van Etten, J.L.; Blanc, G. Global analysis of *Chlorella variabilis* NC64A mRNA profiles during the early phase of *Paramecium bursaria* chlorella virus-1 infection. *PLoS ONE* **2014**, *9*, e90988. [CrossRef] [PubMed]

177. DeLong, J.P.; Al-Ameeli, Z.; Duncan, G.; Van Etten, J.L.; Dunigan, D.D. Predators catalyze an increase in chloroviruses by foraging on the symbiotic hosts of zoochlorellae. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13780–13784. [CrossRef] [PubMed]

178. DeLong, J.P.; Al-Ameeli, Z.; Lyon, S.; Van Etten, J.L.; Dunigan, D.D. Size-dependent catalysis of chlorovirus population growth by a messy feeding predator. *Microb. Ecol.* **2018**, *75*, 847–853. [CrossRef] [PubMed]

179. Long, A.M.; Short, S.M. Seasonal determinations of algal virus decay rates reveal overwintering in a temperate freshwater pond. *ISME J.* **2016**, *10*, 1602–1612. [CrossRef] [PubMed]