

Review

# Host Range and Coding Potential of Eukaryotic Giant Viruses

Tsu-Wang Sun <sup>1,2</sup>, Chia-Ling Yang <sup>1</sup>, Tzu-Tong Kao <sup>1</sup>, Tzu-Haw Wang <sup>1</sup>, Ming-Wei Lai <sup>1</sup>  
and Chuan Ku <sup>1,2,\*</sup>

<sup>1</sup> Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan; r09b48007@ntu.edu.tw (T.-W.S.); as0190267@gate.sinica.edu.tw (C.-L.Y.); tzutongkao@gate.sinica.edu.tw (T.-T.K.); tzuhawwang@gate.sinica.edu.tw (T.-H.W.); mwlai@gate.sinica.edu.tw (M.-W.L.)

<sup>2</sup> Genome and Systems Biology Degree Program, National Taiwan University and Academia Sinica, Taipei 10617, Taiwan

\* Correspondence: chuanku@gate.sinica.edu.tw

Received: 7 November 2020; Accepted: 19 November 2020; Published: 21 November 2020



**Abstract:** Giant viruses are a group of eukaryotic double-stranded DNA viruses with large virion and genome size that challenged the traditional view of virus. Newly isolated strains and sequenced genomes in the last two decades have substantially advanced our knowledge of their host diversity, gene functions, and evolutionary history. Giant viruses are now known to infect hosts from all major supergroups in the eukaryotic tree of life, which predominantly comprises microbial organisms. The seven well-recognized viral clades (taxonomic families) have drastically different host range. *Mimiviridae* and *Phycodnaviridae*, both with notable intrafamilial genome variation and high abundance in environmental samples, have members that infect the most diverse eukaryotic lineages. Laboratory experiments and comparative genomics have shed light on the unprecedented functional potential of giant viruses, encoding proteins for genetic information flow, energy metabolism, synthesis of biomolecules, membrane transport, and sensing that allow for sophisticated control of intracellular conditions and cell-environment interactions. Evolutionary genomics can illuminate how current and past hosts shape viral gene repertoires, although it becomes more obscure with divergent sequences and deep phylogenies. Continued works to characterize giant viruses from marine and other environments will further contribute to our understanding of their host range, coding potential, and virus-host coevolution.

**Keywords:** Nucleo-Cytoplasmic Large DNA Viruses (NCLDV); algae; protists; cophylogeny; host switch; auxiliary genes; virus-encoded metabolism; gene repertoire; genome evolution; lateral gene transfers

## 1. Introduction

The Nucleo-Cytoplasmic Large DNA Viruses (NCLDVs) are a group of double-stranded DNA viruses of eukaryotes that was established two decades ago [1]. Analyses of few widely distributed core genes suggest a monophyletic origin of NCLDVs [1,2], which have been formally named *Nucleocytoviricota* within the virus kingdom *Bamfordvirae* (realm *Varidnaviria*) by the ICTV [3,4]. For their extraordinary genome size (up to 2.8 Mb [5]) and virion size (up to 2.5  $\mu\text{m}$  [6]) among all viruses, NCLDVs are commonly known as giant viruses [7–11], which reflects their distinction from traditionally defined viruses [12]. Although arbitrary thresholds can be applied to draw a line between *large* and *giant* viruses, it is now clear that NCLDV genomes larger than a certain size (e.g., 500 [13] or 300 [14] kb) have multiple evolutionary origins and that their size distribution forms a continuum

with considerable variation both within and across families [13,15]. Still, little is known about the mechanisms that underpin the evolution and maintenance of giant virus genome diversity.

Recent advances in the biology of giant viruses have brought to the fore their expanded host range and coding potential, as shown in numerous studies based on isolation and cultivation, genomic and functional characterization, or environmental metagenomics. Giant viruses have been found in a myriad of eukaryotes previously unknown as hosts, gradually filling the gaps of giant virus hosts in the eukaryotic tree of life. Newly sequenced genomes often contain genes that have rarely or never been found in viruses, while the majority of genes in giant viruses do not even have homolog matches in sequence databases. The virus-encoded genes that are otherwise characteristic of cellular organisms could play crucial roles in manipulating the metabolism of infected cells, converting them into virocells [16,17]. By considering NCLDVs—giant viruses—as a whole, here we briefly summarize and highlight recent findings that have revolutionized our view of their host range and coding potential, with particular emphasis on the evolutionary implications for giant viral genomes.

## 2. Giant Viruses Infect Every Major Eukaryotic Lineage

### 2.1. The Founding Members of NCLDVs

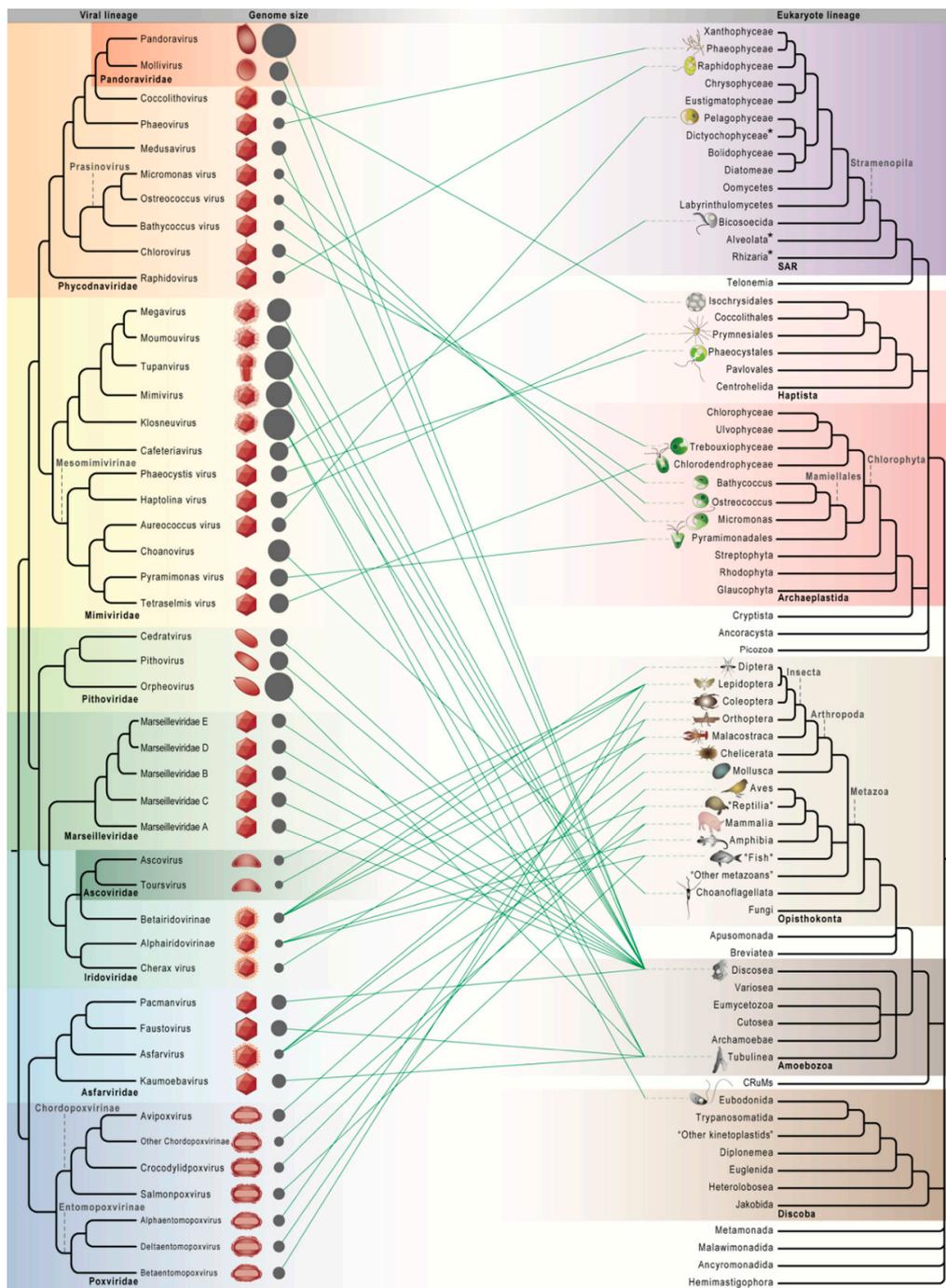
Our current knowledge of the host range of NCLDVs (Figure 1 and Table S1) is dramatically different from when NCLDVs were first recognized through comparative genomics and comprised by only four pre-existing families [1]: three families of animal (vertebrate and arthropod) viruses—*Poxviridae* (e.g., variola and vaccinia viruses), *Asfarviridae* (only African swine fever virus [ASFV]), and *Iridoviridae* (vertebrate and insect viruses, incl. *Ascoviridae* [18])—and the alga-infecting *Phycodnaviridae*, including chlorovirus of *Chlorella* (Trebouxiophyceae, Chlorophyta) and phaeovirus of the multicellular brown alga *Ectocarpus* (Phaeophyceae, Stramenopila). These first known hosts of NCLDVs were clearly biased toward human, livestock, and other multicellular organisms. In fact, only a few unicellular hosts of giant viruses had been reported before the establishment of NCLDV in 2001. Unlike the better studied chlorovirus [19], genome sequences of marine viruses infecting the green microalga *Micromonas* (Mamiellales, Chlorophyta) [20], the bloom-forming coccolithophore *Emiliania* (Isochrysidales, Haptista) [21], the heterotrophic flagellate *Cafeteria* (Bicosoecida, Stramenopila; originally misidentified as *Bodo*) [22], and the harmful bloom-forming alga *Heterosigma* (Raphidophyceae, Stramenopila) [23] became available only after 2001 [9,24–26], which eventually supported their grouping with other NCLDVs.

### 2.2. The Age of Discovery for Giant Viruses Infecting Microbial Eukaryotes

With the advancement in genome sequencing and environmental microbiology, NCLDVs have been expanding substantially in terms of known host diversity. A major shift to discovering giant viruses in microbial eukaryotes or protists, which comprise the vast majority of eukaryotic lineages [27], began with the recognition of mimivirus, a bacterium-like pathogen in *Acanthamoeba* (Discosea, Amoebozoa) and the first giant virus with an Mb-sized genome [8]. As the first non-photosynthetic protist host of giant viruses, *Acanthamoeba* (ubiquitous free-living amoebae and opportunistic human pathogens [28]) had since then become the model organisms for isolating giant viruses from a variety of environmental samples, leading to the discovery of marseillevirus [29], megavirus [30], moomouvirus [31], pandoravirus [5], pithovirus [32], mollivirus [33], cedratvirus [34], pacmanvirus [35], tupanvirus [36], and medusavirus [37]. A distantly related amoebal lineage, *Vermamoeba* (Tubulinea, Amoebozoa), has been used since 2015 as an alternative host system for viruses that cannot replicate in *Acanthamoeba*, including faustovirus [38], kaumoebavirus [39], orpheovirus [40], and yasminevirus (a klosneuvirus) [41]. Giant viruses isolated with this approach have given rise to the proposal of four new families—*Mimiviridae*, *Marseilleviridae*, *Pandoraviridae*, and *Pithoviridae*—and expanded some pre-existing families (Figure 1). These commonly recognized NCLDV families vary greatly in their virion morphology, genome characteristics, and replication cycle (Figure 1 and Table 1).

**Table 1.** A list of representative giant viruses and their morphological and genomic features.

Family	Virus (Lineage)	Shape	Size (nm)	Genome Size (kb)	DNA	tRNAs	RNA Polymerase Subunits [54]	Replication Cycle	Reference
<i>Pandoraviridae</i> ( <i>Phycodnaviridae</i> )	<i>Pandoravirus salinus</i> (Pandoravirus)	Ovoid with a pore	1200 × 500	2770	Linear	3	4	Nuclear and cytoplasmic	[5,55,56]
<i>Phycodnaviridae</i>	<i>Paramecium bursaria Chlorella virus 1</i> (Chlorovirus)	Icosahedral with a spike	170	331	Linear	11	0	Nuclear and cytoplasmic	[7,19]
<i>Mimiviridae</i>	<i>Acanthamoeba polyphaga mimivirus</i> (Mimivirus)	Icosahedral (with fibers)	390 (630)	1181	Linear	6	9	Cytoplasmic	[55,57]
<i>Pithoviridae</i>	<i>Pithovirus sibericum</i> (Pithovirus)	Ovoid with a capped pore	1500 × 800	610	Circular	0	4	Cytoplasmic	[6,32,55]
<i>Marseilleviridae</i>	<i>Marseillevirus marseillevirus</i> (Marseilleviridae A)	Icosahedral	250	368	Circular	0	3	Cytoplasmic, involving the nucleus	[29,58,59]
<i>Ascoviridae</i> ( <i>Iridoviridae</i> )	<i>Spodoptera frugiperda</i> ascovirus 1a (Ascovirus)	Bacilliform or allantoid	400 × 130	157	Circular	0	4	Nuclear and cytoplasmic	[60,61]
<i>Iridoviridae</i>	Frog virus 3 (Alphairidovirinae)	Icosahedral	175	106	Linear	0	2	Nuclear and cytoplasmic	[18,62]
<i>Asfarviridae</i>	African swine fever virus BA71V (Asfarvirus)	Icosahedral	200	170	Linear	0	7	Nuclear and cytoplasmic	[63,64]
<i>Poxviridae</i>	Vaccinia virus (other Chordopoxvirinae)	Brick-shaped	310 × 240	195	Linear	0	9	Cytoplasmic	[65]



**Figure 1.** Cophylogenetic relationships between giant viruses and eukaryotes. Virus-host connections are mapped onto reference trees summarizing giant virus phylogenies based on up to 10 putatively vertically inherited core genes [13,14,37,42,43] and phylogenetic relationships across the eukaryotic tree of life [27,44–50]. The seven commonly delineated giant virus clades (seven families with two nested families) are divided into subfamilial lineages, each of which comprises one or more viral strains with known hosts (Table S1), with a schematic diagram of virion morphology, if known, and the average genome size across strains (proportional to the circle area; Table S1) shown. Names in double quotes correspond to non-monophyletic groupings. \* Giant viruses are also known to infect the SAR lineages dictyochophytes (Stramenopila) [51,52], dinoflagellates (Alveolata) [53], and chlorarachniophytes (Rhizaria) [52], but their phylogenetic positions in the virus reference tree are less certain.

Giant viruses have also been found in other diverse aquatic microbial eukaryotes, greatly broadening their known host diversity. These include heterotrophic flagellates *Bodo* (Kinetoplastida, Discoba; infected by a klosneuvirus) [66] and *Bicosta* (Choanoflagellata, Opisthokonta) [14] and marine microalgae *Haptolina* (previously *Chrysochromulina*) (Prymnesiales, Haptista) [67], *Pyramimonas* (Pyramimonadales, Chlorophyta) [67], *Phaeocystis globosa* (Phaeocystales, Haptista) [68], *Aureococcus* (Pelagophyceae, Stramenopila) [69], *Florenciella* and *Rhizochromulina* (Dictyochophyceae, Stramenopila) [51,52], and Chlorarachniophyceae (Cercozoa, Rhizaria) [52]. As alternatives to isolation and cultivation, single-virion genomics [70], single-cell genomics [14], and metagenomics [14,43,71] are especially useful when the host eukaryotes are unknown or cannot be purified and cultured. These methods not only generate sequences of individual viruses but can also infer their putative hosts, such as *Bicosta* of choanovirus [14] or cercozoans (Rhizaria) of wastewater klosneuviruses [71].

### 2.3. Increasingly Non-Algal Phycodnaviridae and Increasingly Non-Amoebal Mimiviridae

Genomic information from newly reported viruses has often challenged family delineations of NCLDV, in particular Phycodnaviridae and Mimiviridae, the two families with the most diverse host and genome size range (Figure 1). *Phycodnaviridae*, which literally means algal DNA viruses, originally encompassed only chlorovirus, and it was proposed that this family should include phaeovirus and *Micromonas* virus based on some common properties [7] despite the lack genome sequences at that time. These two viral lineages, as well as coccolithovirus, *Heterosigma* raphidovirus, and viruses of *Ostreococcus* and *Bathycoccus* (in the same order Mamiellales as *Micromonas*), turned out to be indeed closely related at the genomic level. By contrast, the story became more complicated for the other alga-infecting giant viruses. *Haptolina*, *Pyramimonas* and *Phaeocystis* viruses were suggested to be part of *Phycodnaviridae* [67,68], and so was *Aureococcus* virus despite some apparently contradictory molecular evidence [69]. It was only until phylogenetic and genomic analyses with mimivirids were conducted that it became clear that these viruses, along with the metagenomically discovered Organic Lake Phycodnaviruses (OLPVs) [72], more recently reported virus of *Tetraselmis* (Chlorodendrophyceae, Chlorophyta) [73], and choanovirus, are more closely related to mimivirids (incl. cafeteriavirus) [72,73]. It was further proposed that these viruses should form part of an extended *Mimiviridae* [11,13,42,74,75] or even the subfamily *Mesomimivirinae* within *Mimiviridae* [42,43,76] (Figure 1). *Mesomimivirinae* is certainly expanding as most newly reported alga-infecting viruses join this clade instead of *Phycodnaviridae*, such as the viruses from *Florenciella*, *Rhizochromulina*, a chlorarachniophyte, and *Prymnesium kappa* [51,52,77] (not shown in Figure 1). These, plus the viruses of heterotrophic flagellates *Cafeteria*, *Bodo*, and *Bicosta*, all transform *Mimiviridae* into a more non-amoebal virus family.

On the other hand, amoeba-infecting viruses have “invaded” other families. In addition to *Mimiviridae* and the two purely amoebal families, *Marseilleviridae* and *Pithoviridae*, viruses that infect *Acanthamoeba* or *Vermamoeba* have often been placed within the two NCLDV founding families *Asfarviridae* and *Phycodnaviridae* (Figure 1). *Asfarviridae*, with its name derived from ASFVs—notorious viruses that kill pigs and boars, can be transmitted through tick (Chelicerata, Metazoa) vectors, and have close relatives infecting abalones (Mollusca, Metazoa) [64,78] (Table S1)—has recently been joined by related amoebal viruses with genomes smaller than 500 kb [13,35,38,39] and possibly also a marine algal virus infecting *Heterocapsa* (Dinoflagellata, Alveolata) [53]. *Phycodnaviridae* has lost some alga-infecting members, but at the same time, it has been joined by amoebal viruses characterized by the largest known genomes—pandoravirus. Phylogenetic analyses based on core genes have strikingly and consistently nested *Pandoraviridae* (here incl. mollivirus) within *Phycodnaviridae* and as sister to the alga-infecting coccolithovirus with a much smaller genome [13,14] (Figure 1). Besides, the more recently discovered medusavirus could also be related to this subclade of *Phycodnaviridae* based on the phylogenies of major capsid protein [37] or 12 core genes [79] (Figure 1). With the extended host range, giant virus families that are cladistically defined based on core gene phylogenies are obviously not confined to their prototypic host, be it amoebal, algal, or swine.

#### 2.4. Undiscovered Virus–Host Relationships

To date, giant virus infection has been reported from each of the most taxon-rich and well established lineages (supergroups) of eukaryotes [27,80], including Opisthokonta, Archaeplastida, SAR (incl. each of the three subgroups Stramenopila, Alveolata, and Rhizaria), Haptista, Amoebozoa, and Discoba (part of the now non-monophyletic “Excavata” that also includes Metamonada and Malawimonada) (Figure 1). Eukaryotes from which giant viruses were first isolated tend to be more relevant or observable to humans, whereas more recently discovered giant viruses are mostly from relatively understudied protist lineages of which research has been greatly accelerated by high-throughput genome sequencing. In addition to the known virus–host relationships (Figure 1 and Table S1), several lines of evidence are pointing to the immense diversity of undiscovered viruses and their hosts. Individual genomes assembled from metagenomic sequencing data (i.e., metagenome-assembled genomes [MAGs]) have led to the discoveries of OLPVs from Antarctica [72], 16 giant viruses from soil [43], hundreds of aquatic MAGs related to *Mimiviridae* and *Phycodnaviridae* [17], and over 2000 giant virus MAGs from across the globe [81], which greatly outnumber the currently described members in giant virus families. Genome sequences of eukaryotes have also hinted at putative (past) associations between giant viruses and Streptophyta, Cryptista, Fungi, and many other eukaryotic lineages where giant viruses have not been isolated [75]. However, despite similar evidence for some land plants [82], or embryophytes (Streptophyta), the complete lack of giant virus reported from any plant could indeed represent a small gap of giant virus host range on the tree of eukaryotes. This could be explained by the fact that plant viruses usually take advantage of the plasmodesmata aperture to spread systemically and encode movement proteins for intercellular transport through plasmodesmata [83], which is unlikely for the large size of giant virus particles or genomic DNA.

### 3. Variation and Evolution of Host Range

At the level of individual virus, most giant viruses are known to infect only specific hosts. However, it is often uncertain whether the known hosts are the only and natural hosts. Because of the systematic isolation approach, many giant virus lineages are only known to infect *Acanthamoeba* or *Vermamoeba*, resulting in the pattern of multiple viruses connected to only one host (Figure 1). It could be that these widely occurring amoebae are indeed the natural, specific host of all those viruses, that they are simply more permissive lab hosts in which a wide range of viruses can be propagated, or that they are secondary hosts for those viruses with primary hosts and serve as the genomic melting pot [29] for various giant viruses. It should be noted that, with few exceptions such as tupanvirus [36], lab experiments have demonstrated that amoebal viruses can only replicate within *Vermamoeba* but not *Acanthamoeba* [38–41], or vice versa [84], suggesting there is still virus–host specificity between the two permissive hosts.

At the family level, giant virus families show wide variation in the extent of host range. *Poxviridae* and *Iridoviridae* (incl. *Ascoviridae*) infect only animals (Metazoa, Opisthokonta) or more specifically only vertebrates and arthropods. Within each of them exist subclades with narrower host range, e.g., vertebrates (Chordopoxvirinae and Alphairidovirinae) or arthropods (Entomopoxvirinae and Ascoviridae–Betairidovirinae). The more recently established *Marseilleviridae* and *Pithoviridae* also have narrow host range (Amoebozoa). On the contrary, *Phycodnaviridae* is associated with four eukaryotic supergroups and *Mimiviridae* with six in total. This host diversity cannot be attributed to their numbers of described viruses, which are dwarfed by that of *Poxviridae* or *Iridoviridae* (Table S1). Neither can it be explained by their intrafamilial phylogenetic divergence in terms of core genes [13,42,43], which is obviously higher than among the isolated members of *Marseilleviridae* but comparable to that of *Poxviridae* or *Iridoviridae*. The major difference between wide– and narrow–host-range viral families is probably the extent of genomic variation (Figure 1). This is evidently greater within *Phycodnaviridae* and *Mimiviridae* than within *Poxviridae* or *Iridoviridae* with generally small genomes, implying plasticity and variability in genome content could be key to conquering a wider range of hosts.

Insights into virus-host specificity have also been gained from studies on closely related viruses. A cross-infection network between coccolithovirus and *Emiliania huxleyi* strains showed a nested host-virus interaction pattern where more resistant hosts are only infected by viruses with broader host range, suggesting strong coevolution in host-virus system [85]. At a larger scale, phylogenetic correspondence has been observed between three genera of Mamiellales (“Prasinophyceae”, Chlorophyta) and their prasinoviruses [44]. Some discrepancy (i.e., imperfect cophylogeny and non-monophyly of *Ostreococcus* viruses) does exist between the trees of viral DNA polymerase and algal ribosomal RNA genes [44], but this can also be seen in strictly vertically inherited symbiont-host system [86] and can be due to incomplete lineage sorting, choice of genes, or taxon sampling. A later phylogenetic analysis with 22 genes from fewer strains of prasinoviruses resolved *Ostreococcus* viruses as monophyletic [87]. Overall, the cophylogenetic pattern indicates long-term coevolution between Mamiellales and prasinoviruses, with either cospeciation or host-switching events [44]. In contrast to host variation between closely related viral strains, processes of host change or expansion involving phylogenetically distant eukaryotes still remain largely unknown.

#### 4. Functional Potential of Virus-Encoded Proteins

Most predicted genes in giant viruses have unknown functions, and many of them have no homolog match in sequence databases at all [55]. In addition to near-universal core genes fundamental in virus replication cycle (e.g., DNA polymerase, primase-helicase, major capsid protein, genome packaging ATPase, transcription factor VLTF3 [13,88]), the minority of genes with functional predictions and cellular homologs often show unprecedented occurrence in the viral world (Table 2). The expanded genome size of giant viruses paves the way for harboring a large variable portion of the genome encoding auxiliary metabolic genes [16] (virus-encoded metabolism) and genes with other functions. They can allow for finer modulation of metabolism, gene expression, and behaviors in diverse hosts, converting them into virocells [16,17] during infection and playing a key role in the virus-host interaction and genome evolution (Table 2).

**Table 2.** A glimpse of the functional diversity of protein-encoding genes in giant viruses.

COG Category	Function	Distribution	Putative LGT Source	Reference
<b>Cellular Processes and Signaling</b>				
[M] Cell wall/membrane/envelope biogenesis	Hyaluronan synthesis	Chlorovirus	-	[89]
	Fucose synthesis	Chlorovirus	-	[90]
	L-rhamnose synthesis	Chlorovirus; Mimivirus	Trebouxiophyceae; eukaryotes	[91]
	3-deoxy-D-manno-octulosonate synthesis	Cafeteriavirus	Phagocytosed bacteria in the <i>Cafeteria</i> host	[9]
	4-Amino-4,6-dideoxy-D-glucose (Viosamine) synthesis	Mimivirus	-	[92]
[O] Posttranslational modification, protein turnover, chaperones	Protein glycosylase	Chlorovirus; Mimivirus	-	[93,94]
	Prolyl 4-hydroxylase	Chlorovirus	-	[95]
	Sulfhydryl oxidase	Asfarvirus; Mimivirus	-	[96,97]
	Isomerization of peptide bonds (Cyclophilin)	Mimivirus	-	[98]
<b>Information Storage and Processing</b>				
[J] Translation, ribosomal structure, and biogenesis	Aminoacyl tRNA synthetase	<i>Mimiviridae</i> except <i>Mesomimivirinae</i> ; Pandoravirus; Orpheovirus	-	[13]
	Translation factors	<i>Mimiviridae</i> ; Chlorovirus; <i>Pandoraviridae</i> ; <i>Marseilleviridae</i> ; <i>Asfarviridae</i> ; <i>Alphairidovirinae</i> ; <i>Pithoviridae</i> ; <i>Alphaentomopoxvirus</i>	-	[13]
[K] Transcription	DNA-dependent RNA polymerase (DDRP) subunits	Most NCLDV's except many phycodnavirids	-	[37,54]
	Transcription elongation factors (TFIIS)	Most NCLDV's	-	[54]
	General transcription factors (TBP-like)	Some phycodnavirids and <i>Mesomimivirinae</i>	-	[54]
	General transcription factors (TFIIB-like)	<i>Asfarviridae</i> ; <i>Mimiviridae</i> ; <i>Marseilleviridae</i> ; <i>Pithovirus</i> ; <i>Prasinovirus</i>	-	[54]
[L] Replication, recombination, and repair	DNA Glycosylase	<i>Marseilleviridae</i> ; most mimivirids; <i>Poxviridae</i>	-	[99]
	NAD/ATP dependent DNA ligase	Most NCLDV families	-	[100]
<b>Metabolism</b>				
[C] Energy production and conversion	Tricarboxylic acid (TCA) cycle	<i>Mimiviridae</i> MAGs; <i>Prymnesium</i> virus; Pandoravirus	-	[17,77,101]
	Cellular fermentation	<i>Tetraselmis</i> virus	<i>Tetraselmis</i> or related chlorophytes	[73]
[C] Energy production and conversion; [T] Signal transduction mechanisms	Rhodopsin	<i>Phaeocystis</i> virus; Choanovirus; metagenomic contigs	-	[10,14,79,102]
[E] Amino acid transport and metabolism	Polyamine synthesis	Chlorovirus	-	[103]
	Amino acid synthesis	Prasinovirus	Chlorophyta or bacteria	[25]
[F] Nucleotide transport and metabolism	Nucleoside-diphosphate kinase	Mimivirus	-	[104]
[G] Carbohydrate transport and metabolism	UDP-N-acetylglucosamine synthesis	Chlorovirus; <i>Mimiviridae</i>	Bacteria; -	[105,106]
	UDP-2-acetamido-2,6-dideoxy-hexose synthesis	Megavirus	-	[107]
	Cell wall (polysaccharide) degradation	Chlorovirus; Mimivirus	Non-Chloroplastida sources; -	[108–113]
[I] Lipid transport and metabolism	Sphingolipid synthesis	Coccolithovirus	Isochrysidales	[24,114,115]
[P] Inorganic ion transport and metabolism	Potassium ion channel	Chlorovirus; Phaeovirus; Prasinovirus; many <i>Mesomimivirinae</i> viruses	-	[51,116–119]
	Aquaporin	Chlorovirus	-	[120]
	Sodium/phosphate symporter	Coccolithovirus; Prasinovirus	Isochrysidales; Mamiellales	[121,122]
	Potassium ion transporter	Chlorovirus	Trebouxiophyceae	[121,123]
	Ammonium transporter	<i>Ostreococcus</i> virus; <i>Haptolina</i> virus	<i>Ostreococcus</i> ; -	[87,121]
	Calcium-transporting ATPase	Chlorovirus	-	[124]
	Cu/Zn superoxide dismutase	Betaentomopoxvirus; Chlorovirus; Megavirus	-	[125–127]

#### 4.1. Information Storage and Flow

Giant viruses exert control over different levels of genetic information in a cell. In addition to their own DNA polymerase and ligase for genome replication, some giant viruses encode DNA glycosylase involved in base excision repair pathways that could potentially remove damages to their large genomes [77,99]. Most NCLDVs also encode DNA-dependent RNA polymerase (DDRP) subunits (Table 1) with architectural modifications that confer them higher speed and processivity than the cellular homologs [54]. Interestingly, in *Phycodnaviridae* (as defined in Figure 1), these genes are only found in coccolithovirus [24], pandoravirus [5], and mollivirus [33], which have the largest genomes within the family (Figure 1). Except medusavirus [37], all those phycodnavirids without DDRP genes infect algae (Figure 1), which is in sharp contrast to the alga-infecting mimivirids (in subfamily *Mesomimivirinae*; Figure 1) that have the most complete complement of DDRP subunits among giant viruses [54]. Giant viruses also have various transcription factors involved in basic transcriptional regulation (initiation, elongation, and termination) and expressional control of viral kinetic classes [9,54,128,129]. Some unknown genes could further rewire the entire cellular transcriptomes, such as differential shut-down of nucleus- and organelle-encoded transcripts [129]. Presence of genes for translational control is a major hallmark of giant viruses. Except ribosomal proteins or RNA, a wide range of translation system components can be encoded, including tRNAs (Table 1), aminoacyl tRNA synthetases, and translation factors [13,71] (Table 2). There can be extensive variation in the repertoire of these translational machinery genes even among closely related viruses, for example, klosneuviruses (*Mimiviridae*) where *Bodo* virus has completely lost all its tRNAs while some others have nearly all the translation machinery genes found in giant viruses [13,66,71]. Whereas informational genes generally comprise the essential and core components in genomes of cellular organisms, the extreme variation among viruses with similar hosts or close phylogenetic relationships further demonstrates the plasticity and variability of giant virus genomes.

#### 4.2. Energy Metabolism

The requirements for energy during genome replication, gene expression, and virus assembly make the control of energy metabolism a natural target of giant viruses. Such control can be transcriptional regulation of nucleus- and mitochondrion-encoded genes related to energy metabolism, as in coccolithovirus [129]. Recently reported genome sequences of environmental MAGs [17] or isolated viruses even encode their own genes related to glycolysis, tricarboxylic acid cycle, succinate dehydrogenase,  $\beta$ -oxidation, and photosynthesis [17,77,81,101]. Genes encoding enzymes in cellular fermentation, such as pyruvate formate-lyase, have been found in *Tetrasetelmis* virus infecting green algae [73], which have anaerobic energy metabolism in low-oxygen condition [130].

#### 4.3. Synthesis of Biomolecules

Giant viruses encode various proteins participating in the synthesis of different virion components, with notable examples in carbohydrate, lipid, and nucleotide metabolism. Chloroviruses have plenty of carbohydrate metabolic genes for synthesis of hyaluronan, nucleotide sugars, glycans, and glycoproteins (e.g., capsid proteins glycosylated with distinct glycan structures) [89–91,131]. Coccolithovirus encodes unique host-derived genes for making virus-specific glycosphingolipids that not only constitute the virion membranes but induce host programmed cell death [16,24,114,115,132]. To meet the demand of nucleotides for synthesis of genomic DNA and RNA transcripts, nucleoside-diphosphate kinase and reductase are encoded by multiple giant viruses for nucleotide synthesis and conversion [24,66,104], which can be coupled with induction of host pentose phosphate pathway to enlarge the pool of available nucleotides [132,133].

#### 4.4. Membrane Transport and Sensing

Giant viruses not only take control of information flow, energy metabolism, and biosynthesis but can also alter interactions between the cell and the environment through membrane proteins. A variety of such proteins, including rhodopsins, channels, and transporters, are encoded in the genomes of *Mesomimivirinae* and alga-infecting phycodnavirids. Type-1 rhodopsin genes are found in OLPVs, *Phaeocystis* virus, and Choanovirus, where they pump protons as a light-dependent energy transfer system [10,14]. Choanovirus additionally possess biosynthesis genes for the rhodopsin chromophore, retinal, which are absent in *Phaeocystis* virus but present in its host *Phaeocystis* [14]. A newly discovered type of rhodopsins—heliorhodopsins—is encoded in coccolithovirus genomes [134], which could play a role in light sensing during virus-host interactions. Light-gated anion-conducting channelrhodopsins have recently been found to be encoded in metagenomic contigs of *Mesomimivirinae* and Phycodnaviridae, probably transferred from Pyramimonadales green algae, and could be used to change the host's swimming behavior in response to light [79]. Potassium channels are commonly encoded by algal viruses in Phycodnaviridae and *Mesomimivirinae* and in the chlorovirus-*Chlorella* system they cause membrane depolarization, decrease turgor pressure, and promote viral DNA ejection [51,116–119]. The potential function in other algal viruses could be to make the intracellular environment more favorable to virus-encoded proteins [51], which might also be achieved by the calcium transporting ATPase encoded in chlorovirus [124]. Furthermore, nutrient transporters, including ammonium [87] and sodium/phosphate [121,122] transporters, are commonly encoded in the genomes of algal giant viruses. Some of the aforementioned membrane proteins are brought into the virus-host system by viruses, while some are encoded in both the viral and host genomes. In the latter case, it is often found that the viral and host homologs have different activities or substrate affinities. For example, the ammonium transporter unique to one *Ostreococcus* virus shows higher uptake rate than the host counterpart at lower substrate concentrations and can potentially alter the nutrient uptake of the cell [87].

### 5. Evolution of Genome Content

#### 5.1. Expansive Evolution

Similar to cellular genomes, giant virus genomes undergo both expansive and reductive genome evolution. The increased genome size in giant viruses can be attributed to gene duplications, de novo gene origination, and lateral gene transfers (LGTs, or horizontal gene transfers [HGTs]) from cellular organisms or other viruses [13,55,56,135–137]. Among these sources, LGTs generally bring in more innovative functions to viral genomes. Their identification can provide insights into virus-host interactions, connections between viruses and their current or past hosts, and how hosts play a role in shaping viral genomes.

Laterally acquired genes in giant viruses largely fall into two categories: recently acquired genes from current hosts or related organisms [73,79,87,115] and anciently acquired, divergent viral homologs from an unknown source that sometimes form a clade with only viral and metagenomic sequences (e.g., type-1 rhodopsins [10,14]). It is notable that for some recent LGTs, viral genes demonstrate somewhat higher sequence divergence than their closest eukaryotic homologs in phylogenetic trees [73,79,87]. This might be due to generally higher evolutionary rates in giant viruses, though analyses of closely related marseilleviruses suggest they do not evolve faster than cells [138]. Alternatively, the transferred viral genes could be relieved from purifying selection, since host cells already have the same genes. This could allow viral homologs to acquire distinct functional properties that alter cellular behaviors upon infection [87]. Another interesting observation is the multiple independent acquisitions of the same gene across viral lineages with similar hosts, such as algae. Examples include potassium channels that have been repeatedly gained by viruses of marine and freshwater, unicellular and multicellular algae [51], and sodium/phosphate transporters

with at least three independent events in coccolithovirus, *Ostreococcus* virus, and *Bathycoccus* virus, respectively [121].

### 5.2. Reductive Evolution

Giant viruses with larger genome size can potentially better manipulate specific hosts in a variety of pathways and cellular processes, but there are clearly factors that limit their genome size or cause reductive genome evolution as in other parasitic entities. Random gene losses intrinsically lead to genomic reduction during viral evolution [139]. There is almost no limit to the genes that can be lost, even genes central to information processing. Largely speaking, the repertoires of translation-related genes are the most variable [66], followed by transcription-related genes and then by genes for DNA replication [13,55,139]. A 16% reduction in genome size was observed in mimivirus subcultured 150 times in axenic *Acanthamoeba* cultures, which was accompanied by marked changes in virion morphology [140]. This illustrates how hosts and environment can cause fast genome size changes in giant viruses. Substantial genome size variation between closely related strains [56] and sister viral lineages (Figure 1) also point to highly dynamic expansive and reductive evolution at work.

Viral genomes and gene contents can also be shaped by certain host factors such as host genome size. It was shown that the burst size of phytoplankton dsDNA viruses correlates with host-to-virus genome size ratio [141]. Host genome size as a limiting factor can partially explain why, for example, prasinoviruses of Mamiellales, which have the smallest cell and genome size in Chlorophyta, have some of the smallest genomes in Phycodnaviridae or among all alga-infecting giant viruses [142]. Compared with its sister group chlorovirus (Figure 1), the prasinovirus genomes could have undergone reduction in size during evolution.

### 5.3. Generalist Viruses and Genome Evolution

The host range of a virus is determined by its genome, including the encoded genes and their regulation. On the other hand, a host can shape the genomes of its viruses, selecting for those better adapted to the host. This apparently chicken-or-egg relationship poses the question on how viruses can jump between distantly related eukaryotic host lineages like those of *Phycodnaviridae* and *Mimiviridae*. Here, we approach this question by proposing that there exist specialist viruses infecting only a specific host lineage and generalist viruses which can replicate in multiple eukaryotic host lineages across supergroups. After a generalist virus acquires genes that aid in the infection of a specific host, it can become a more specialized virus or remain as a generalist with more successful infection of the specific host. With higher replication success, more specialized viruses gradually become the dominant virus of the specific host, which could be why most viral lineages with isolated members are only known to infect a specific eukaryote lineage. True generalist viruses that can infect eukaryotes from different supergroups are unknown to date, either because they are less abundant and more difficult to isolate or because we have not explored the entire host range of the isolated giant viruses. There could be a pool of generalist viruses in the environment that would have been detected in metagenomic sequencing, where MAGs are the most abundant from *Mimiviridae* and *Phycodnaviridae* [17,81]—the two families with the most divergent eukaryotic hosts (Figure 1 and Section 3). In line with this generalist hypothesis for host range variation, the multiple independent acquisitions of similar membrane transport genes in the two families (Section 5.1) could correspond to transitions from generalist to specialist viruses infecting different algal or protist lineages.

### 5.4. Origin of Giant Viruses and Their Families

An even more challenging question is what kind of hosts were infected by ancient NCLDVs, including the ancestors of NCLDVs and of each NCLDV family. Answers to this question would depend on our understanding of the genomic compositions of these ancient viruses. Although it can be still disputable [139], NCLDVs are generally believed to have evolved from a common ancestor. Phylogenomic and comparative genomic analyses suggest that the NCLDV common ancestor had a

small viral genome [13] rather than that of a cellular organism [143], but what this common ancestor was like and its relationship to cellular eukaryotes are much more debated. Based on the phylogeny of two DDRP subunits, an NCLDV-early hypothesis was proposed where both NCLDVs and the individual NCLDV families originated before the last eukaryotic common ancestor (LECA), which is close to 2 billion years ago [144], and had infected a lineage of “proto-eukaryotes” that led to LECA [88]. Given the archaeal [145] and bacterial [146] ancestry of eukaryotic genomes, such “proto-eukaryotes” would be more like prokaryotes than at least LECA and its descendants. However, no NCLDV-like infection in prokaryotic cells has been reported so far. As single-gene trees can be misleading in inferring the ancient past of eukaryotes [147], it is also disputable whether we could take the virus-cell DDRP tree at face value [148] and assume they have evolved in viral lineages without being lost, regained, or replaced, which are especially problematic when inferring deep viral phylogenies [139]. Besides, if the NCLDV families *Iridoviridae*, *Marseilleviridae*, and *Pithoviridae* had originated before LECA [88], it is difficult to imagine that today they have such confined host range across the eukaryotic tree of life (Figure 1).

On the contrary, an NCLDV-late view would suggest a eukaryotic host for the NCLDV common ancestor. The deep divergence of shared genes is not necessarily the actual divergence among the giant viruses themselves, because they, as compartments of genes, can acquire divergent genes from different domains of cellular life. There is no doubt that widely-occurring core genes strongly shape the biology of giant viruses and are phylogenetically related. Nevertheless, this does not mean that they all have been passed on, together and vertically, through the deep bifurcations as depicted in their concatenated gene tree. Under the NCLDV-late hypothesis, the association between giant viruses and eukaryotes could have taken place later than the LECA origin and then spread across major eukaryotic lineages, as what could happen within just a single NCLDV family (*Mimiviridae* and *Phycodnaviridae*). In contrast to the obscure deep inter-familial relationships, NCLDVs form more coherent groups at the family level, which is true in terms of the number of shared genes and the viral biology [13,55] and probably especially so for families with similar genome size and lower sequence divergence. For families of viruses with similar hosts (*Marseilleviridae*, *Poxviridae*, and *Iridoviridae*), their ancient hosts were most likely from the same host lineage, but it would be more difficult to infer the ancestral host of *Phycodnaviridae* and of *Mimiviridae*. Even more challenging is to understand and predict genome variation across giant viruses. Rampant occurrence of LGTs is known to cause extensive gene content variation in prokaryotes, even among strains with highly similar core gene sequences [149–151], such that the 3% most vertically inherited genes are not predictive of the rest of the genome [152]. With limited evidence for long-term verticality plus substantial genome size differences, little can be inferred about giant virus gene contents from a phylogenetic tree based on a concatenated alignment of 10 genes that are not devoid of conflicting signals. More comprehensive whole-genome analyses are needed to determine the factors affecting gene content evolution both within and across giant virus families.

## 6. Future Perspective

Giant viruses demonstrate at least three unconventional features. Their bacterial-sized virions and genomes defy the idea that viruses are small infectious entities. Identification of new giant viruses from global ecosystems revealed the enormous diversity of their hosts across the now better resolved tree of eukaryotes. The functional potential of their genomes revolutionized our knowledge of how viruses can manipulate the host to complete their replication cycle. In a way, giant viruses function like powerful, innovative, yet often lethal, plug-ins in the program of eukaryotic life. This group of ubiquitous and ecologically important viruses will continue to be a source of exciting findings. In addition to crucial endeavors to isolate new strains, infection assays, metagenomics, single-particle genomics, functional characterization, and virus-host genomic analyses are expected to shed light on their biology, natural host range, virus-host interactions, and genome evolution within and across families.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/12/11/1337/s1>. Table S1: Giant virus strains and their phylogenetic placement, genome information, and hosts.

**Author Contributions:** Conceptualization, C.K.; summarization and visualization of phylogenies and host range, T.-W.S., C.-L.Y., and T.-T.K.; compilation of gene functions, T.-H.W.; writing—original draft preparation, C.K. and M.-W.L.; writing—review and editing, M.-W.L., T.-T.K., T.-W.S., C.-L.Y., T.-H.W., and C.K.; supervision, C.K.; project administration, C.K. and M.-W.L.; funding acquisition, C.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Science and Technology, Taiwan, grant numbers MOST 108-2311-B-001-040-MY3 (C.K.), MOST 109-2811-B-001-508, and MOST 109-2811-B-001-598 (C.K., T.-T.K.), and by the Institute of Plant and Microbial Biology, Academia Sinica, Taiwan. The APC was funded by the Institute of Plant and Microbial Biology, Academia Sinica, Taiwan.

**Acknowledgments:** We thank the editor and two anonymous reviewers for their valuable suggestions and comments on an earlier version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- 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](#)]
- Iyer, L.M.; Balaji, S.; Koonin, E.V.; Aravind, L. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res.* **2006**, *117*, 156–184. [[CrossRef](#)] [[PubMed](#)]
- Koonin, E.V.; Dolja, V.V.; Krupovic, M.; Varsani, A.; Wolf, Y.I.; Yutin, N.; Zerbini, F.M.; Kuhn, J.H. Global Organization and Proposed Megataxonomy of the Virus World. *Microbiol. Mol. Biol. Rev.* **2020**, *84*, e00061-19. [[CrossRef](#)] [[PubMed](#)]
- International Committee on Taxonomy of Viruses Executive Committee. The new scope of virus taxonomy: Partitioning the virosphere into 15 hierarchical ranks. *Nat. Microbiol.* **2020**, *5*, 668–674. [[CrossRef](#)] [[PubMed](#)]
- Philippe, N.; Legendre, M.; Doutre, G.; Couté, Y.; Poirot, O.; Lescot, M.; Arslan, D.; Seltzer, V.; Bertaux, L.; Bruley, C.; et al. Pandoraviruses: Amoeba Viruses with Genomes Up to 2.5 Mb Reaching That of Parasitic Eukaryotes. *Science* **2013**, *341*, 281–286. [[CrossRef](#)] [[PubMed](#)]
- Okamoto, K.; Miyazaki, N.; Song, C.; Maia, F.R.N.C.; Reddy, H.K.N.; Abergel, C.; Claverie, J.M.; Hajdu, J.; Svenda, M.; Murata, K. Structural variability and complexity of the giant *Pithovirus sibericum* particle revealed by high-voltage electron cryo-tomography and energy-filtered electron cryo-microscopy. *Sci. Rep.* **2017**, *7*, 13291. [[CrossRef](#)] [[PubMed](#)]
- Van Etten, J.L.; Meints, R.H. Giant viruses infecting algae. *Annu. Rev. Microbiol.* **1999**, *53*, 447–494. [[CrossRef](#)]
- La Scola, B.; Audic, S.; Robert, C.; Jungang, L.; De Lamballerie, X.; Drancourt, M.; Birtles, R.; Claverie, J.M.; Raoult, D. A giant virus in amoebae. *Science* **2003**, *299*, 2033. [[CrossRef](#)]
- Fischer, M.G.; Allen, M.J.; Wilson, W.H.; Suttle, C.A. Giant virus with a remarkable complement of genes infects marine zooplankton. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19508–19513. [[CrossRef](#)]
- Yutin, N.; Koonin, E.V. Proteorhodopsin genes in giant viruses. *Biol. Direct* **2012**, *7*, 34. [[CrossRef](#)]
- Wilhelm, S.W.; Coy, S.R.; Gann, E.R.; Moniruzzaman, M.; Stough, J.M.A. Standing on the shoulders of giant viruses: Five lessons learned about large viruses infecting small eukaryotes and the opportunities they create. *PLoS Pathog.* **2016**, *12*, e1005752. [[CrossRef](#)] [[PubMed](#)]
- Claverie, J.-M.; Abergel, C. Giant viruses: The difficult breaking of multiple epistemological barriers. *Stud. Hist. Philos. Biol. Biomed. Sci.* **2016**, *59*, 89–99. [[CrossRef](#)] [[PubMed](#)]
- Koonin, E.V.; Yutin, N. Multiple evolutionary origins of giant viruses. *F1000Research* **2018**, *7*, 1840. [[CrossRef](#)] [[PubMed](#)]
- Needham, D.M.; Yoshizawa, S.; Hosaka, T.; Poirier, C.; Choi, C.J.; Hehenberger, E.; Irwin, N.A.T.T.; Wilken, S.; Yung, C.-M.M.; Bachy, C.; et al. A distinct lineage of giant viruses brings a rhodopsin photosystem to unicellular marine predators. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 20574–20583. [[CrossRef](#)]
- Brandes, N.; Linial, M. Giant Viruses—Big Surprises. *Viruses* **2019**, *11*, 404. [[CrossRef](#)]
- Rosenwasser, S.; Ziv, C.; van Creveld, S.G.; Vardi, A. Virocell metabolism: Metabolic innovations during host–virus interactions in the ocean. *Trends Microbiol.* **2016**, *24*, 821–832. [[CrossRef](#)]

17. Moniruzzaman, M.; Martinez-Gutierrez, C.A.; Weinheimer, A.R.; Aylward, F.O. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nat. Commun.* **2020**, *11*, 1710. [[CrossRef](#)]
18. Chinchar, V.G.; Waltzek, T.B.; Subramaniam, K. Ranaviruses and other members of the family Iridoviridae: Their place in the virosphere. *Virology* **2017**, *511*, 259–271. [[CrossRef](#)]
19. Van Etten, J.L.; Agarkova, I.V.; Dunigan, D.D. Chloroviruses. *Viruses* **2020**, *12*, 20. [[CrossRef](#)]
20. Mayer, J.A.; Taylor, F.J.R. A virus which lyses the marine nanoflagellate *Micromonas pusilla*. *Nature* **1979**, *281*, 299–301. [[CrossRef](#)]
21. 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](#)]
22. Garza, D.R.; Suttle, C.A. Large double-stranded DNA viruses which cause the lysis of a marine heterotrophic nanoflagellate (*Bodo* sp.) occur in natural marine viral communities. *Aquat. Microb. Ecol.* **1995**, *9*, 203–210. [[CrossRef](#)]
23. 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](#)]
24. Wilson, W.H. Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science* **2005**, *309*, 1090–1092. [[CrossRef](#)] [[PubMed](#)]
25. 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](#)]
26. Maruyama, F.; Ueki, S. Evolution and phylogeny of large DNA viruses, *Mimiviridae* and *Phycodnaviridae* including newly characterized *Heterosigma akashiwo* Virus. *Front. Microbiol.* **2016**, *7*, 1942. [[CrossRef](#)]
27. Adl, S.M.; Bass, D.; Lane, C.E.; Lukeš, J.; Schoch, C.L.; Smirnov, A.; Agatha, S.; Berney, C.; Brown, M.W.; Burki, F.; et al. Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* **2019**, *66*, 4–119. [[CrossRef](#)]
28. Marciano-Cabral, F.; Cabral, G. *Acanthamoeba* spp. as agents of disease in humans. *Clin. Microbiol. Rev.* **2003**, *16*, 273–307. [[CrossRef](#)]
29. Boyer, M.; Yutin, N.; Pagnier, I.; Barrassi, L.; Fournous, G.; Espinosa, L.; Robert, C.; Azza, S.; Sun, S.; Rossmann, M.G.; et al. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 21848–21853. [[CrossRef](#)]
30. Arslan, D.; Legendre, M.; Seltzer, V.; Abergel, C.; Claverie, J.M. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17486–17491. [[CrossRef](#)]
31. Yoosuf, N.; Yutin, N.; Colson, P.; Shabalina, S.A.; Pagnier, I.; Robert, C.; Azza, S.; Klose, T.; Wong, J.; Rossmann, M.G.; et al. Related giant viruses in distant locations and different habitats: *Acanthamoeba polyphaga* *moumouvirus* represents a third lineage of the *Mimiviridae* that is close to the *Megavirus* lineage. *Genome Biol. Evol.* **2012**, *4*, 1324–1330. [[CrossRef](#)] [[PubMed](#)]
32. Legendre, M.; Bartoli, J.; Shmakova, L.; Jeudy, S.; Labadie, K.; Adrait, A.; Lescot, M.; Poirot, O.; Bertaux, L.; Bruley, C.; et al. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4274–4279. [[CrossRef](#)] [[PubMed](#)]
33. Legendre, M.; Lartigue, A.; Bertaux, L.; Jeudy, S.; Bartoli, J.; Lescot, M.; Alempic, J.M.; Ramus, C.; Bruley, C.; Labadie, K.; et al. In-depth study of *Mollivirus sibericum*, a new 30,000-yold giant virus infecting *Acanthamoeba*. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5327–E5335. [[CrossRef](#)] [[PubMed](#)]
34. Andreani, J.; Aherfi, S.; Bou Khalil, J.; Di Pinto, F.; Bitam, I.; Raoult, D.; Colson, P.; La Scola, B. Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. *Viruses* **2016**, *8*, 300. [[CrossRef](#)] [[PubMed](#)]
35. Andreani, J.; Khalil, J.Y.B.; Sevvana, M.; Benamar, S.; Di Pinto, F.; Bitam, I.; Colson, P.; Klose, T.; Rossmann, M.G.; Raoult, D.; et al. Pacmanvirus, a new giant icosahedral virus at the crossroads between *Asfarviridae* and faustoviruses. *J. Virol.* **2017**, *91*, 1–11. [[CrossRef](#)]
36. Abrahão, J.; Silva, L.; Silva, L.S.; Khalil, J.Y.B.; Rodrigues, R.; Arantes, T.; Assis, F.; Boratto, P.; Andrade, M.; Kroon, E.G.; et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **2018**, *9*, 749. [[CrossRef](#)]

37. Yoshikawa, G.; Blanc-Mathieu, R.; Song, C.; Kayama, Y.; Mochizuki, T.; Murata, K.; Ogata, H.; Takemurad, M.; Blanc-Mathieu, R.; Song, C.; et al. Medusavirus, a novel large DNA virus discovered from hot spring water. *J. Virol.* **2019**, *93*, 1–25. [[CrossRef](#)]
38. Reteno, D.G.; Benamar, S.; Khalil, J.B.; Andreani, J.; Armstrong, N.; Klose, T.; Rossmann, M.; Colson, P.; Raoult, D.; La Scola, B. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol.* **2015**, *89*, 6585–6594. [[CrossRef](#)]
39. Bajrai, L.H.; Benamar, S.; Azhar, E.I.; Robert, C.; Levasseur, A.; Raoult, D.; La Scola, B. Kaumoebavirus, a new virus that clusters with Faustoviruses and *Asfarviridae*. *Viruses* **2016**, *8*, 278. [[CrossRef](#)]
40. Andreani, J.; Khalil, J.Y.B.; Baptiste, E.; Hasni, I.; Michelle, C.; Raoult, D.; Levasseur, A.; La Scola, B. Orpheovirus IHUMI-LCC2: A new virus among the giant viruses. *Front. Microbiol.* **2018**, *8*, 2643. [[CrossRef](#)]
41. Bajrai, L.H.; Mougari, S.; Andreani, J.; Baptiste, E.; Delerce, J.; Raoult, D.; Azhar, E.I.; La Scola, B.; Levasseur, A. Isolation of Yasminevirus, the first member of Klosneuvirinae isolated in coculture with *Vermamoeba vermiformis*, demonstrates an extended arsenal of translational apparatus components. *J. Virol.* **2019**, *94*, e01534-19. [[CrossRef](#)] [[PubMed](#)]
42. Needham, D.M.; Poirier, C.; Hehenberger, E.; Jiménez, V.; Swalwell, J.E.; Santoro, A.E.; Worden, A.Z. Targeted metagenomic recovery of four divergent viruses reveals shared and distinctive characteristics of giant viruses of marine eukaryotes. *Philos. Trans. R. Soc. B Biol. Sci.* **2019**, *374*, 20190086. [[CrossRef](#)] [[PubMed](#)]
43. Schulz, F.; Alteio, L.; Goudeau, D.; Ryan, E.M.; Yu, F.B.; Malmstrom, R.R.; Blanchard, J.; Woyke, T. Hidden diversity of soil giant viruses. *Nat. Commun.* **2018**, *9*, 4881. [[CrossRef](#)] [[PubMed](#)]
44. Bellec, L.; Clerissi, C.; Edern, R.; Foulon, E.; Simon, N.; Grimsley, N.; Desdevises, Y. Cophylogenetic interactions between marine viruses and eukaryotic picophytoplankton. *BMC Evol. Biol.* **2014**, *14*, 59. [[CrossRef](#)]
45. Derelle, R.; López-García, P.; Timpano, H.; Moreira, D. A phylogenomic framework to study the diversity and evolution of stramenopiles (=heterokonts). *Mol. Biol. Evol.* **2016**, *33*, 2890–2898. [[CrossRef](#)]
46. Ševčíková, T.; Yurchenko, T.; Fawley, K.P.; Amaral, R.; Strnad, H.; Santos, L.M.A.; Fawley, M.W.; Eliáš, M. Plastid genomes and proteins illuminate the evolution of eustigmatophyte algae and their bacterial endosymbionts. *Genome Biol. Evol.* **2019**, *11*, 362–379. [[CrossRef](#)] [[PubMed](#)]
47. Fučíková, K.; Leliaert, F.; Cooper, E.D.; Škaloud, P.; D'Hondt, S.; De Clerck, O.; Gurgel, C.F.D.; Lewis, L.A.; Lewis, P.O.; Lopez-Bautista, J.M.; et al. New phylogenetic hypotheses for the core Chlorophyta based on chloroplast sequence data. *Front. Ecol. Evol.* **2014**, *2*, 63.
48. Burki, F.; Roger, A.J.; Brown, M.W.; Simpson, A.G.B. The new tree of eukaryotes. *Trends Ecol. Evol.* **2020**, *35*, 43–55. [[CrossRef](#)]
49. Misof, B.; Liu, S.; Meusemann, K.; Peters, R.S.; Donath, A.; Mayer, C.; Frandsen, P.B.; Ware, J.; Flouri, T.; Beutel, R.G.; et al. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **2014**, *346*, 763–767. [[CrossRef](#)]
50. Kang, S.; Tice, A.K.; Spiegel, F.W.; Silberman, J.D.; Pánek, T.; Čepička, I.; Kostka, M.; Kosakyan, A.; Alcântara, D.M.C.; Roger, A.J.; et al. Between a pod and a hard test: The deep evolution of amoebae. *Mol. Biol. Evol.* **2017**, *34*, 2258–2270. [[CrossRef](#)]
51. Kukovetz, K.; Hertel, B.; Schvarcz, C.R.; Saponaro, A.; Manthey, M.; Burk, U.; Greiner, T.; Steward, G.F.; Van Etten, J.L.; Moroni, A.; et al. A functional K<sup>+</sup> channel from Tetraselmis virus 1, a member of the *Mimiviridae*. *Viruses* **2020**, *12*, 1107. [[CrossRef](#)] [[PubMed](#)]
52. Schvarcz, C.R. Cultivation and Characterization of Viruses Infecting Eukaryotic Phytoplankton from the Tropical North Pacific Ocean. Ph.D. Thesis, University of Hawai'i at Manoa, Honolulu, HI, USA, 2018.
53. 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](#)]
54. Mirzakhanyan, Y.; Gershon, P.D. Multisubunit DNA-dependent RNA polymerases from vaccinia virus and other nucleocytoplasmic large-DNA viruses: Impressions from the age of structure. *Microbiol. Mol. Biol. Rev.* **2017**, *81*, e00010-17. [[CrossRef](#)] [[PubMed](#)]
55. Abergel, C.; Legendre, M.; Claverie, J.M. The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol. Rev.* **2015**, *39*, 779–796. [[CrossRef](#)] [[PubMed](#)]

56. Legendre, M.; Fabre, E.; Poirot, O.; Jeudy, S.; Lartigue, A.; Alempic, J.-M.M.; Beucher, L.; Philippe, N.N.; Bertaux, L.; Christo-Foroux, E.; et al. Diversity and evolution of the emerging *Pandoraviridae* family. *Nat. Commun.* **2018**, *9*, 2285. [[CrossRef](#)] [[PubMed](#)]
57. Colson, P.; La Scola, B.; Levasseur, A.; Caetano-Anollés, G.; Raoult, D. Mimivirus: Leading the way in the discovery of giant viruses of amoebae. *Nat. Rev. Microbiol.* **2017**, *15*, 243–254. [[CrossRef](#)] [[PubMed](#)]
58. Blanca, L.; Christo-Foroux, E.; Rigou, S.; Legendre, M. Comparative Analysis of the Circular and Highly Asymmetrical Marseilleviridae Genomes. *Viruses* **2020**, *12*. [[CrossRef](#)]
59. Fabre, E.; Jeudy, S.; Santini, S.; Legendre, M.; Trauchessec, M.; Couté, Y.; Claverie, J.M.; Abergel, C. Noumeavirus replication relies on a transient remote control of the host nucleus. *Nat. Commun.* **2017**, *8*, 1–12. [[CrossRef](#)]
60. Federici, B.A.; Vlaskovits, J.M.; Hamm, J.J. Comparative study of virion structure, protein composition and genomic DNA of three ascovirus isolates. *J. Gen. Virol.* **1990**, *71*, 1661–1668. [[CrossRef](#)]
61. Bideshi, D.K.; Demattei, M.-V.; Rouleux-Bonnin, F.; Stasiak, K.; Tan, Y.; Bigot, S.; Bigot, Y.; Federici, B.A. Genomic Sequence of Spodoptera frugiperda Ascovirus 1a, an Enveloped, Double-Stranded DNA Insect Virus That Manipulates Apoptosis for Viral Reproduction. *J. Virol.* **2006**, *80*, 11791–11805. [[CrossRef](#)]
62. Majji, S.; Thodima, V.; Sample, R.; Whitley, D.; Deng, Y.; Mao, J.; Chinchar, V.G. Transcriptome analysis of Frog virus 3, the type species of the genus Ranavirus, family Iridoviridae. *Virology* **2009**, *391*, 293–303. [[CrossRef](#)] [[PubMed](#)]
63. Alonso, C.; Borca, M.; Dixon, L.; Revilla, Y.; Rodriguez, F.; Escribano, J.M. ICTV virus taxonomy profile: Asfarviridae. *J. Gen. Virol.* **2018**, *99*, 613–614. [[CrossRef](#)] [[PubMed](#)]
64. Galindo, I.; Alonso, C. African swine fever virus: A review. *Viruses* **2017**, *9*, 103. [[CrossRef](#)] [[PubMed](#)]
65. Condit, R.C.; Moussatche, N.; Traktman, P. In A Nutshell: Structure and Assembly of the Vaccinia Virion. *Adv. Virus Res.* **2006**, *65*, 31–124.
66. Deeg, C.M.; Chow, C.E.T.; Suttle, C.A. The kinetoplastid-infecting Bodo saltans virus (Bsv), a window into the most abundant giant viruses in the sea. *Elife* **2018**, *7*, e33014. [[CrossRef](#)]
67. 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](#)]
68. 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](#)]
69. Rowe, J.M.; Dunlap, J.R.; Gobler, C.J.; Anderson, O.R.; Gastrich, M.D.; Wilhelm, S.W. Isolation of a non-phage-like lytic virus infecting *Aureococcus anophagefferens*. *J. Phycol.* **2008**, *44*, 71–76. [[CrossRef](#)]
70. Wilson, W.H.; Gilg, I.C.; Moniruzzaman, M.; Field, E.K.; Koren, S.; LeCleir, G.R.; Martínez Martínez, J.; Poulton, N.J.; Swan, B.K.; Stepanauskas, R.; et al. Genomic exploration of individual giant ocean viruses. *ISME J.* **2017**, *11*, 1736–1745. [[CrossRef](#)]
71. Schulz, F.; Yutin, N.; Ivanova, N.N.; Ortega, D.R.; Lee, T.K.; Vierheilig, J.; Daims, H.; Horn, M.; Wagner, M.; Jensen, G.J.; et al. Giant viruses with an expanded complement of translation system components. *Science* **2017**, *356*, 82–85. [[CrossRef](#)]
72. 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](#)]
73. Schvarcz, C.R.; Steward, G.F. A giant virus infecting green algae encodes key fermentation genes. *Virology* **2018**, *518*, 423–433. [[CrossRef](#)] [[PubMed](#)]
74. Yutin, N.; Colson, P.; Raoult, D.; Koonin, E.V. *Mimiviridae*: Clusters of orthologous genes, reconstruction of gene repertoire evolution and proposed expansion of the giant virus family. *Virol. J.* **2013**, *10*, 106. [[CrossRef](#)] [[PubMed](#)]
75. Gallot-Lavallée, L.; Blanc, G. A glimpse of nucleo-cytoplasmic large DNA virus biodiversity through the eukaryotic genomics window. *Viruses* **2017**, *9*, 17. [[CrossRef](#)]
76. Gallot-Lavallée, 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*, e00230-17. [[CrossRef](#)]

77. Blanc-Mathieu, R.; Dahle, H.; Hofgaard, A.; Brandt, D.; Ogata, H.; Sandaa, R.-A. The genome of a persistent giant algal virus encodes an unprecedented number of genes involved in energy metabolism. *bioRxiv* **2020**. [[CrossRef](#)]
78. Matsuyama, T.; Takano, T.; Nishiki, I.; Fujiwara, A.; Kiryu, I.; Inada, M.; Sakai, T.; Terashima, S.; Matsuura, Y.; Isowa, K.; et al. A novel Asfarvirus-like virus identified as a potential cause of mass mortality of abalone. *Sci. Rep.* **2020**, *10*, 4620. [[CrossRef](#)]
79. Rozenberg, A.; Oppermann, J.; Wietek, J.; Fernandez Lahore, R.G.; Sandaa, R.-A.; Bratbak, G.; Hegemann, P.; Beja, O. Lateral gene transfer of anion-conducting channelrhodopsins between green algae and giant viruses. *Curr. Biol.* **2020**, *30*, in press.
80. Keeling, P.J.; Burki, F. Progress towards the Tree of Eukaryotes. *Curr. Biol.* **2019**, *29*, R808–R817. [[CrossRef](#)]
81. Schulz, F.; Roux, S.; Paez-Espino, D.; Jungbluth, S.; Walsh, D.A.; Denef, V.J.; McMahon, K.D.; Konstantinidis, K.T.; Eloe-Fadrosh, E.A.; Kyrpides, N.C.; et al. Giant virus diversity and host interactions through global metagenomics. *Nature* **2020**, *578*, 432–436. [[CrossRef](#)]
82. Maumus, F.; Epert, A.; Nogu e, F.; Blanc, G. Plant genomes enclose footprints of past infections by giant virus relatives. *Nat. Commun.* **2014**, *5*, 4268. [[CrossRef](#)] [[PubMed](#)]
83. Heinlein, M. Plasmodesmata: Channels for viruses on the move. *Methods Mol. Biol.* **2014**, *1217*, 25–52.
84. Dornas, F.P.; Khalil, J.Y.B.; Pagnier, I.; Raoult, D.; Abrah o, J.; La Scola, B. Isolation of new Brazilian giant viruses from environmental samples using a panel of protozoa. *Front. Microbiol.* **2015**, *6*, 1086. [[CrossRef](#)] [[PubMed](#)]
85. Ruiz, E.; Oosterhof, M.; Sandaa, R.A.; Larsen, A.; Pagarete, A. Emerging interaction patterns in the *Emiliania huxleyi*-EhV system. *Viruses* **2017**, *9*, 61. [[CrossRef](#)]
86. Ku, C.; Hu, J.-M. Phylogenetic and cophylogenetic analyses of the leaf-nodule symbiosis in *Ardisia* Subgenus *Crispardisia* (Myrsinaceae): Evidence from nuclear and chloroplast markers and bacterial *rrn* operons. *Int. J. Plant Sci.* **2014**, *175*, 92–109. [[CrossRef](#)]
87. Monier, A.; Chambouvet, A.; Milner, D.S.; Attah, V.; Terrado, R.; Lovejoy, C.; Moreau, H.; Santoro, A.E.; Derelle,  .; Richards, T.A. Host-derived viral transporter protein for nitrogen uptake in infected marine phytoplankton. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E7489–E7498. [[CrossRef](#)]
88. Guglielmini, J.; Woo, A.C.; Krupovic, M.; Forterre, P.; Gaia, M. Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 19585–19592. [[CrossRef](#)]
89. DeAngelis, P.L.; Jing, W.; Graves, M.V.; Burbank, D.E.; Van Etten, J.L. Hyaluronan synthase of chlorella virus PBCV-1. *Science* **1997**, *278*, 1800–1803. [[CrossRef](#)]
90. Tonetti, M.; Zanardi, D.; Gurnon, J.R.; Fruscione, F.; Armirotti, A.; Damonte, G.; Sturla, L.; De Flora, A.; Van Etten, J.L. *Paramecium bursaria* chlorella virus 1 encodes two enzymes involved in the biosynthesis of GDP-L-fucose and GDP-D-rhamnose. *J. Biol. Chem.* **2003**, *278*, 21559–21565. [[CrossRef](#)]
91. Chothi, M.P.; Duncan, G.A.; Armirotti, A.; Abergel, C.; Gurnon, J.R.; Van Etten, J.L.; Bernardi, C.; Damonte, G.; Tonetti, M. Identification of an L-rhamnose synthetic pathway in two nucleocytoplasmic large DNA viruses. *J. Virol.* **2010**, *84*, 8829–8838. [[CrossRef](#)]
92. Piacente, F.; Marin, M.; Molinaro, A.; De Castro, C.; Seltzer, V.; Salis, A.; Damonte, G.; Bernardi, C.; Claverie, J.M.; Abergel, C.; et al. Giant DNA virus Mimivirus encodes pathway for biosynthesis of unusual sugar 4-amino-4,6-dideoxy-D-glucose (Viosamine). *J. Biol. Chem.* **2012**, *287*, 3009–3018. [[CrossRef](#)] [[PubMed](#)]
93. Graves, M.V.; Bernadt, C.T.; Cerny, R.; Van Etten, J.L. Molecular and genetic evidence for a virus-encoded glycosyltransferase involved in protein glycosylation. *Virology* **2001**, *285*, 332–345. [[CrossRef](#)] [[PubMed](#)]
94. H ulsmeier, A.J.; Hennet, T. O-Linked glycosylation in *Acanthamoeba polyphaga* mimivirus. *Glycobiology* **2014**, *24*, 703–714. [[CrossRef](#)] [[PubMed](#)]
95. Eriksson, M.; Myllyharju, J.; Tu, H.; Hellman, M.; Kivirikko, K.I. Evidence for 4-hydroxyproline in viral proteins. Characterization of a viral prolyl 4-hydroxylase and its peptide substrates. *J. Biol. Chem.* **1999**, *274*, 22131–22134. [[CrossRef](#)] [[PubMed](#)]
96. Rodr guez, I.; Redrejo-Rodr guez, M.; Rodr guez, J.M.; Alejo, A.; Salas, J.; Salas, M.L. African swine fever virus pB119L protein is a flavin adenine dinucleotide-linked sulfhydryl oxidase. *J. Virol.* **2006**, *80*, 3157–3166. [[CrossRef](#)] [[PubMed](#)]
97. Hakim, M.; Ezerina, D.; Alon, A.; Vonshak, O.; Fass, D. Exploring ORFan domains in giant viruses: Structure of mimivirus sulfhydryl oxidase R596. *PLoS ONE* **2012**, *7*, e50649. [[CrossRef](#)] [[PubMed](#)]

98. Thai, V.; Renesto, P.; Fowler, C.A.; Brown, D.J.; Davis, T.; Gu, W.; Pollock, D.D.; Kern, D.; Raoult, D.; Eisenmesser, E.Z. Structural, biochemical, and in vivo characterization of the first virally encoded cyclophilin from the mimivirus. *J. Mol. Biol.* **2008**, *378*, 71–86. [[CrossRef](#)]
99. Redrejo-Rodríguez, M.; Salas, M.L. Repair of base damage and genome maintenance in the nucleo-cytoplasmic large DNA viruses. *Virus Res.* **2014**, *179*, 12–25. [[CrossRef](#)]
100. Yutin, N.; Koonin, E.V. Evolution of DNA ligases of nucleo-cytoplasmic large DNA viruses of eukaryotes: A case of hidden complexity. *Biol. Direct* **2009**, *4*, 51. [[CrossRef](#)]
101. Aherfi, S.; Brahim Belhaouari, D.; Pinault, L.; Decloquement, P.; Abrahao, J.; Colson, P.; Levasseur, A.; Lamb, D.C.; Chabriere, E.; Raoult, D.; et al. Tricarboxylic acid cycle and proton gradient in *Pandoravirus massiliensis*: Is it still a virus? *bioRxiv* **2020**. [[CrossRef](#)]
102. Zabelskii, D.; Alekseev, A.; Kovalev, K.; Oliveira, A.-S.; Balandin, T.; Soloviov, D.; Bratanov, D.; Volkov, D.; Vaganova, S.; Astashkin, R.; et al. Viral channelrhodopsins: Calcium-dependent Na<sup>+</sup>/K<sup>+</sup> selective light-gated channels. *bioRxiv* **2020**. [[CrossRef](#)]
103. Baumann, S.; Sander, A.; Gurnon, J.R.; Yanai-Balser, G.M.; Van Etten, J.L.; Piotrowski, M. Chlorella viruses contain genes encoding a complete polyamine biosynthetic pathway. *Virology* **2007**, *360*, 209–217. [[CrossRef](#)] [[PubMed](#)]
104. Jeudy, S.; Lartigue, A.; Claverie, J.-M.; Abergel, C. Dissecting the unique nucleotide specificity of mimivirus nucleoside diphosphate kinase. *J. Virol.* **2009**, *83*, 7142–7150. [[CrossRef](#)] [[PubMed](#)]
105. Landstein, D.; Graves, M.V.; Burbank, D.E.; Deangelis, P.; Van Etten, J.L. Chlorella virus PBCV-1 encodes functional glutamine: Fructose-6- phosphate amidotransferase and UDP-glucose dehydrogenase enzymes. *Virology* **1998**, *250*, 388–396. [[CrossRef](#)]
106. Piacente, F.; Bernardi, C.; Marin, M.; Blanc, G.; Abergel, C.; Tonetti, M.G. Characterization of a UDP-N-acetylglucosamine biosynthetic pathway encoded by the giant DNA virus mimivirus. *Glycobiology* **2014**, *24*, 51–61. [[CrossRef](#)]
107. Piacente, F.; De Castro, C.; Jeudy, S.; Molinaro, A.; Salis, A.; Damonte, G.; Bernardi, C.; Abergel, C.; Tonetti, M.G. Giant virus *Megavirus chilensis* encodes the biosynthetic pathway for uncommon acetamido sugars. *J. Biol. Chem.* **2014**, *289*, 24428–24439. [[CrossRef](#)]
108. Sun, L.; Adams, B.; Gurnon, J.R.; Ye, Y.; Van Etten, J.L. Characterization of two chitinase genes and one chitosanase gene encoded by Chlorella virus PBCV-1. *Virology* **1999**, *263*, 376–387. [[CrossRef](#)]
109. Sun, L.; Gurnon, J.R.; Adams, B.J.; Graves, M.V.; Van Etten, J.L. Characterization of a beta-1,3-glucanase encoded by chlorella virus PBCV-1. *Virology* **2000**, *276*, 27–36. [[CrossRef](#)]
110. Suda, K.; Tanji, Y.; Hori, K.; Unno, H. Evidence for a novel chlorella virus-encoded alginate lyase. *FEMS Microbiol. Lett.* **1999**, *180*, 45–53. [[CrossRef](#)]
111. Sugimoto, I.; Hiramatsu, S.; Murakami, D.; Fujie, M.; Usami, S.; Yamada, T. Algal-lytic activities encoded by Chlorella virus CVK2. *Virology* **2000**, *277*, 119–126. [[CrossRef](#)]
112. Blanc, G.; Duncan, G.; Agarkova, I.; Borodovsky, M.; Gurnon, J.; Kuo, A.; Lindquist, E.; Lucas, S.; Pangilinan, J.; Polle, J.; et al. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* **2010**, *22*, 2943–2955. [[CrossRef](#)] [[PubMed](#)]
113. Klose, T.; Herbst, D.A.; Zhu, H.; Max, J.P.; Kenttämää, H.I.; Rossmann, M.G. A mimivirus enzyme that participates in viral entry. *Structure* **2015**, *23*, 1058–1065. [[CrossRef](#)] [[PubMed](#)]
114. Vardi, A.; Van Mooy, B.A.S.; Fredricks, H.F.; Pependorf, K.J.; Ossolinski, J.E.; Haramaty, L.; Bidle, K.D. Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* **2009**, *326*, 861–865. [[CrossRef](#)] [[PubMed](#)]
115. Monier, A.; Pagarete, A.; De Vargas, C.; Allen, M.J.; Read, B.; Claverie, J.M.; Ogata, H. Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. *Genome Res.* **2009**, *19*, 1441–1449. [[CrossRef](#)] [[PubMed](#)]
116. Plugge, B.; Gazzarrini, S.; Nelson, M.; Cerana, R.; Van Etten, J.L.; Derst, C.; DiFrancesco, D.; Moroni, A.; Thiel, G. A potassium channel protein encoded by Chlorella virus PBCV-1. *Science* **2000**, *287*, 1641–1644. [[CrossRef](#)] [[PubMed](#)]
117. Cock, J.M.; Sterck, L.; Rouzé, P.; Scornet, D.; Allen, A.E.; Amoutzias, G.; Anthouard, V.; Artiguenave, F.; Aury, J.M.; Badger, J.H.; et al. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* **2010**, *465*, 617–621. [[CrossRef](#)]

118. Thiel, G.; Moroni, A.; Blanc, G.; Van Etten, J.L. Potassium ion channels: Could they have evolved from viruses? *Plant Physiol.* **2013**, *162*, 1215–1224. [[CrossRef](#)]
119. Siotto, F.; Martin, C.; Rauh, O.; Van Etten, J.L.; Schroeder, I.; Moroni, A.; Thiel, G. Viruses infecting marine picoplankton encode functional potassium ion channels. *Virology* **2014**, *466–467*, 103–111. [[CrossRef](#)]
120. Gazzarrini, S.; Kang, M.; Epimashko, S.; Van Etten, J.L.; Dainty, J.; Thiel, G.; Moroni, A. Chlorella virus MT325 encodes water and potassium channels that interact synergistically. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5355–5360. [[CrossRef](#)]
121. Greiner, T.; Moroni, A.; Van Etten, J.L.; Thiel, G. Genes for membrane transport proteins: Not so rare in viruses. *Viruses* **2018**, *10*, 456. [[CrossRef](#)]
122. Monier, A.; Welsh, R.M.; Gentemann, C.; Weinstock, G.; Sodergren, E.; Armbrust, E.V.; Eisen, J.A.; Worden, A.Z. Phosphate transporters in marine phytoplankton and their viruses: Cross-domain commonalities in viral-host gene exchanges. *Environ. Microbiol.* **2012**, *14*, 162–176. [[CrossRef](#)] [[PubMed](#)]
123. Greiner, T.; Ramos, J.; Alvarez, M.C.; Gurnon, J.R.; Kang, M.; Van Etten, J.L.; Moroni, A.; Thiel, G. Functional HAK/KUP/KT-like potassium transporter encoded by chlorella viruses. *Plant J.* **2011**, *68*, 977–986. [[CrossRef](#)] [[PubMed](#)]
124. Bonza, M.C.; Martin, H.; Kang, M.; Lewis, G.; Greiner, T.; Giacometti, S.; Van Etten, J.L.; De Michelis, M.I.; Thiel, G.; Moroni, A. A functional calcium-transporting ATPase encoded by chlorella viruses. *J. Gen. Virol.* **2010**, *91*, 2620–2629. [[CrossRef](#)] [[PubMed](#)]
125. Becker, M.N.; Greenleaf, W.B.; Ostrov, D.A.; Moyer, R.W. *Amsacta moorei* entomopoxvirus expresses an active superoxide dismutase. *J. Virol.* **2004**, *78*, 10265–10275. [[CrossRef](#)]
126. Kang, M.; Duncan, G.A.; Kuszynski, C.; Oyler, G.; Zheng, J.; Becker, D.F.; Van Etten, J.L. Chlorovirus PBCV-1 encodes an active copper-zinc superoxide dismutase. *J. Virol.* **2014**, *88*, 12541–12550. [[CrossRef](#)]
127. Lartigue, A.; Burlat, B.; Coutard, B.; Chaspoul, F.; Claverie, J.-M.; Abergel, C. The megavirus chilensis Cu,Zn-superoxide dismutase: The first viral structure of a typical cellular copper chaperone-independent hyperstable dimeric enzyme. *J. Virol.* **2015**, *89*, 824–832. [[CrossRef](#)]
128. Legendre, M.; Audic, S.; Poirot, O.; Hingamp, P.; Seltzer, V.; Byrne, D.; Lartigue, A.; Lescot, M.; Bernadac, A.; Poulain, J.; et al. mRNA deep sequencing reveals 75 new genes and a complex transcriptional landscape in Mimivirus. *Genome Res.* **2010**, *20*, 664–674. [[CrossRef](#)]
129. Ku, C.; Sheyn, U.; Seb e-Pedr s, A.; Ben-Dor, S.; Schatz, D.; Tanay, A.; Rosenwasser, S.; Vardi, A. A single-cell view on alga-virus interactions reveals sequential transcriptional programs and infection states. *Sci. Adv.* **2020**, *6*, eaba4137. [[CrossRef](#)]
130. M ller, M.; Mentel, M.; van Hellemond, J.J.; Henze, K.; Woehle, C.; Gould, S.B.; Yu, R.-Y.; van der Giezen, M.; Tielens, A.G.M.; Martin, W.F. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 444–495. [[CrossRef](#)]
131. Van Etten, J.; Agarkova, I.; Dunigan, D.; Tonetti, M.; De Castro, C.; Duncan, G.; Van Etten, J.L.; Agarkova, I.; Dunigan, D.D.; Tonetti, M.; et al. Chloroviruses Have a Sweet Tooth. *Viruses* **2017**, *9*, 88. [[CrossRef](#)]
132. Rosenwasser, S.; Mausz, M.A.; Schatz, D.; Sheyn, U.; Malitsky, S.; Aharoni, A.; Weinstock, E.; Tzfadia, O.; Ben-Dor, S.; Feldmesser, E.; et al. Rewiring Host Lipid Metabolism by Large Viruses Determines the Fate of *Emiliania huxleyi*, a Bloom-Forming Alga in the Ocean. *Plant Cell* **2014**, *26*, 2689–2707. [[CrossRef](#)] [[PubMed](#)]
133. Thamatrakoln, K.; Talmy, D.; Haramaty, L.; Maniscalco, C.; Latham, J.; Knowles, B.; Natale, F.; Coolen, M.J.L.; Follows, M.J.; Bidle, K.D. Light regulation of coccolithophore host-virus interactions. *New Phytol.* **2019**, *221*, 1289–1302. [[CrossRef](#)] [[PubMed](#)]
134. Pushkarev, A.; Inoue, K.; Larom, S.; Flores-uribe, J.; Singh, M.; Konno, M.; Tomida, S.; Filosof, A.; Sharon, I.; Yutin, N.; et al. A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* **2018**, *558*, 595–599. [[CrossRef](#)] [[PubMed](#)]
135. 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](#)]
136. Fil e, J.; Pouget, N.; Chandler, M. Phylogenetic evidence for extensive lateral acquisition of cellular genes by Nucleocytoplasmic large DNA viruses. *BMC Evol. Biol.* **2008**, *8*, 320. [[CrossRef](#)]

137. Christo-Foroux, E.; Alempic, J.; Lartigue, A.; Santini, S.; Labadie, K.; Legendre, M.; Abergel, C.; Claverie, J. Characterization of *Mollivirus kamchatka*, the First Modern Representative of the Proposed *Molliviridae* Family of Giant Viruses. *J. Virol.* **2020**, *94*, 1–16. [[CrossRef](#)]
138. Doutre, G.; Philippe, N.; Abergel, C.; Claverie, J.-M. Genome Analysis of the First Marseilleviridae Representative from Australia Indicates that Most of Its Genes Contribute to Virus Fitness. *J. Virol.* **2014**, *88*, 14340–14349. [[CrossRef](#)]
139. Claverie, J.M. Fundamental difficulties prevent the reconstruction of the deep phylogeny of viruses. *Viruses* **2020**, *12*, 1130. [[CrossRef](#)]
140. Boyer, M.; Azza, S.; Barrassi, L.; Klose, T.; Campocasso, A.; Pagnier, I.; Fournous, G.; Borg, A.; Robert, C.; Zhang, X.; et al. Mimivirus shows dramatic genome reduction after intraamoebal culture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10296–10301. [[CrossRef](#)]
141. Edwards, K.F.; Steward, G.F. Host traits drive viral life histories across phytoplankton viruses. *Am. Nat.* **2018**, *191*, 566–581. [[CrossRef](#)]
142. Weynberg, K.D.; Allen, M.J.; Wilson, W.H. Marine prasinoviruses and their tiny plankton hosts: A review. *Viruses* **2017**, *9*, 43. [[CrossRef](#)] [[PubMed](#)]
143. Yutin, N.; Wolf, Y.I.; Koonin, E.V. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology* **2014**, *466–467*, 38–52. [[CrossRef](#)] [[PubMed](#)]
144. Betts, H.C.; Puttick, M.N.; Clark, J.W.; Williams, T.A.; Donoghue, P.C.J.; Pisani, D. Integrated genomic and fossil evidence illuminates life’s early evolution and eukaryote origin. *Nat. Ecol. Evol.* **2018**, *2*, 1556–1562. [[CrossRef](#)] [[PubMed](#)]
145. Imachi, H.; Nobu, M.K.; Nakahara, N.; Morono, Y.; Ogawara, M.; Takaki, Y.; Takano, Y.; Uematsu, K.; Ikuta, T.; Ito, M.; et al. Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature* **2020**, *577*, 519–525. [[CrossRef](#)] [[PubMed](#)]
146. Brueckner, J.; Martin, W.F. Bacterial genes outnumber archaeal genes in eukaryotic genomes. *Genome Biol. Evol.* **2020**, *12*, 282–292. [[CrossRef](#)] [[PubMed](#)]
147. Ku, C.; Nelson-Sathi, S.; Roettger, M.; Garg, S.; Hazkani-Covo, E.; Martin, W.F. Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10139–10146. [[CrossRef](#)] [[PubMed](#)]
148. Ku, C.; Sun, T. Did giant and large dsDNA viruses originate before their eukaryotic hosts? *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 2747–2748. [[CrossRef](#)]
149. Ochman, H.; Lawrence, J.G.; Groisman, E.A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **2000**, *405*, 299–304. [[CrossRef](#)]
150. Popa, O.; Dagan, T. Trends and barriers to lateral gene transfer in prokaryotes. *Curr. Opin. Microbiol.* **2011**, *14*, 615–623. [[CrossRef](#)]
151. McInerney, J.O.; McNally, A.; O’Connell, M.J. Why prokaryotes have pangenomes. *Nat. Microbiol.* **2017**, *2*, 17040. [[CrossRef](#)]
152. Nagies, F.S.P.; Brueckner, J.; Tria, F.D.K.; Martin, W.F. A spectrum of verticality across genes. *PLoS Genet.* **2020**, *16*, e1009200. [[CrossRef](#)] [[PubMed](#)]

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



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).