



HAL
open science

A Persistent Giant Algal Virus, with a Unique Morphology, Encodes an Unprecedented Number of Genes Involved in Energy Metabolism

Romain Blanc-Mathieu, Håkon Dahle, Antje Hofgaard, David Brandt, Hiroki Ban, Jörn Kalinowski, Hiroyuki Ogata, Ruth-Anne Sandaa

► **To cite this version:**

Romain Blanc-Mathieu, Håkon Dahle, Antje Hofgaard, David Brandt, Hiroki Ban, et al.. A Persistent Giant Algal Virus, with a Unique Morphology, Encodes an Unprecedented Number of Genes Involved in Energy Metabolism. *Journal of Virology*, 2021, 95 (8), 10.1128/JVI.02446-20 . hal-03229969

HAL Id: hal-03229969

<https://hal.science/hal-03229969>

Submitted on 20 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 **A persistent giant algal virus, with a unique morphology, encodes an**
2 **unprecedented number of genes involved in energy metabolism**

3

4 Romain Blanc-Mathieu^{1,2}, Håkon Dahle³, Antje Hofgaard⁴, David Brandt⁵, Hiroki
5 Ban¹, Jörn Kalinowski⁵, Hiroyuki Ogata¹ and Ruth-Anne Sandaa^{6*}

6

7 1: Institute for Chemical Research, Kyoto University, Gokasho, Uji, 611-0011, Japan

8 2: Laboratoire de Physiologie Cellulaire & Végétale, CEA, Univ. Grenoble Alpes,
9 CNRS, INRA, IRIG, Grenoble, France

10 3: Department of Biological Sciences and K.G. Jebsen Center for Deep Sea Research,
11 University of Bergen, Bergen, Norway

12 4: Department of Biosciences, University of Oslo, Norway

13 5: Center for Biotechnology, Universität Bielefeld, Bielefeld, 33615, Germany

14 6: Department of Biological Sciences, University of Bergen, Bergen, Norway

15 *Corresponding author: Ruth-Anne Sandaa, +47 55584646, ruth.sandaa@uib.no

16 Abstract

17 Viruses have long been viewed as entities possessing extremely limited metabolic
18 capacities. Over the last decade, however, this view has been challenged, as metabolic
19 genes have been identified in viruses possessing large genomes and virions—the
20 synthesis of which is energetically demanding. Here, we unveil peculiar phenotypic
21 and genomic features of *Prymnesium kappa* virus RF01 (PkV RF01), a giant virus of
22 the *Mimiviridae* family. We found that this virus encodes an unprecedented number of
23 proteins involved in energy metabolism, such as all four succinate dehydrogenase
24 (SDH) subunits (A–D) as well as key enzymes in the β -oxidation pathway. The *SDHA*
25 gene was transcribed upon infection, indicating that the viral SDH is actively used by
26 the virus—potentially to modulate its host’s energy metabolism. We detected
27 orthologous *SDHA* and *SDHB* genes in numerous genome fragments from
28 uncultivated marine *Mimiviridae* viruses, which suggests that the viral SDH is
29 widespread in oceans. PkV RF01 was less virulent compared with other cultured
30 prymnesioviruses, a phenomenon possibly linked to the metabolic capacity of this
31 virus and suggestive of relatively long co-evolution with its hosts. It also has a unique
32 morphology, compared to other characterized viruses in the *Mimiviridae* family.
33 Finally, we found that PkV RF01 is the only alga-infecting *Mimiviridae* virus
34 encoding two aminoacyl-tRNA synthetases and enzymes corresponding to an entire
35 base-excision repair pathway, as seen in heterotroph-infecting *Mimiviridae*. These
36 *Mimiviridae* encoded-enzymes were found to be monophyletic and branching at the
37 root of the eukaryotic tree of life. This placement suggests that the last common
38 ancestor of *Mimiviridae* was endowed with a large, complex genome prior to the
39 divergence of known extant eukaryotes.

40 **Importance**

41 Viruses on Earth are tremendously diverse in terms of morphology, functionality, and
42 genomic composition. Over the last decade, the conceptual gap separating viruses and
43 cellular life has tightened because of the detection of metabolic genes in viral
44 genomes that express complex virus phenotypes upon infection. Here, we describe
45 *Prymnesium kappa* virus RF01, a large alga-infecting virus with a unique
46 morphology, an atypical infection profile, and an unprecedented number of genes
47 involved in energy metabolism (such as the tricarboxylic (TCA) cycle and the β -
48 oxidation pathway). Moreover, we show that the gene corresponding to one of these
49 enzymes (the succinate dehydrogenase subunit A) is transcribed during infection and
50 is widespread among marine viruses. This discovery provides evidence that a virus
51 has the potential to actively regulate energy metabolism with its own gene.

52

53 **Key words:** algal viruses, *Mimiviridae*, persistent, co-evolution, metabolism, energy
54 production, succinate dehydrogenase, β -oxidation and aminoacyl-tRNA synthetases

55 **Introduction**

56 In their essay “Varieties of Living Things: Life at the Intersection of Lineage and
57 Metabolism,” Dupré and O’Malley proposed to address Schrödinger’s question
58 “What is Life?” by “*describing a spectrum of biological entities that illustrates why*
59 *no sharp dividing line between living and non-living things is likely to be useful*” (1).
60 Microbiologists have contributed considerably to this descriptive effort, both by
61 reporting the existence of viruses endowed with genes coding for functions once
62 thought to be exclusive to cellular life and by concomitantly proposing that actively
63 infecting viruses are a “living form” (2–4). Genes encoding elements for
64 photosynthesis (5, 6), carbon metabolism (7), and nitrogen- (8) and sulfur-cycling (9)
65 have been found in bacterial viruses, where they are used to maintain or augment
66 cellular processes during infection and to redirect energy and resources towards viral
67 production (8, 10, 11). Genes for protein synthesis, including translation initiation,
68 elongation, and termination, and a range of aminoacyl-tRNA synthetases have been
69 found in *Mimiviridae*, a group of giant viruses infecting single-celled eukaryotes (12–
70 14). *Mimiviridae* and other large DNA viruses, including some bacterial viruses, also
71 have tRNA genes (15, 16). Ribosomal proteins have recently been reported in viral
72 genomes derived from metagenomes (17). Genes involved in other metabolic
73 processes, such as fermentation (18), glycosylation (19), photosynthesis (20), and
74 rhodopsin (21), are encoded in *Mimiviridae* and other related large eukaryotic DNA
75 viruses. Metabolic genes are frequently observed within virus genomes (20, 22, 23);
76 although they represent a tiny fraction of the viral gene pool, these genes have the
77 potential to dramatically modify the phenotype of an actively infected cell and alter
78 the ecological role of the host (7, 24, 25). The infected host in this state has been
79 referred to as a virocell (2). One might expect that the interplay between viral genes

80 and host genes in virocells would become increasingly fine-tuned and complex during
81 prolonged virus–host co-evolution, which also typically leads to lower virulence.
82 Much of the complexity of virocells may still be undetected, as most *Mimiviridae*
83 isolated with their natural host (mostly algae) are highly virulent, with several
84 involved in rapid algal bloom termination events (26).

85 Viruses of the *Mimiviridae* family are known to infect heterotrophic and
86 autotrophic microbial eukaryotes. This divide is also reflected in the phylogeny of
87 these viruses, some of which are classified into two proposed sub-families:
88 “Megavirinae” and “Mesomimivirinae” (27). The former contains viruses with
89 genomes larger than 1 Mbp, all isolated from Amoebozoa, while the latter includes
90 viruses with smaller genomes isolated from haptophyte algae of class
91 Prymnesiophyceae. Several *Mimiviridae* members outside these two groups have
92 been characterized to some extent as well, namely, viruses isolated from heterotrophs
93 (*Cafeteria roenbergensis* virus, CroV; *Bodo saltans* virus, BsV; Choano virus),
94 autotrophs (*Aureococcus anophagefferens* virus, AaV; Tetraselmis virus 1, TetV;
95 *Pyramimonas orientalis* virus, PoV; *Prymnesium kappa* virus RF01, PkV RF01), a
96 metazoan (Namao virus), and metagenomes (Klosneuviruses). The Mesomimivirinae
97 sub-family includes viruses infecting bloom-forming hosts, such as *Phaeocystis*
98 *pouchetii*, *Phaeocystis globosa*, and *Prymnesium parvum* (PpV, PgV Group I, and
99 PpDVAV, respectively) (28–30); it also includes several viruses infecting *Haptolina*
100 *ericina* and *Prymnesium kappa*, which normally do not form massive blooms but are
101 present at low densities in seawater year round (31). In marine environments, viruses
102 infecting low-density and non-bloom-forming algae may be the most common virus–
103 host systems—that is, low-density hosts (non-blooming) and viruses that appear to
104 have co-evolved in response to host growth strategy. Thus far, the only known

105 representatives of such viruses are *Prymnesium kappa* viruses RF01 (PkV RF01) and
106 RF02 (PkV RF02), *Haptolina ericina* virus RF02 (HeV RF02), and
107 *Chrysochromulina ericina* virus (CeV 01B, infecting *Haptolina ericina*) (32, 33).
108 Together with PgV, all of these viruses, except for PkV RF01, belong to the sub-
109 family Mesomimivirinae on the basis of their monophyletic relationship and, in the
110 case of PgV and CeV, a shared genomic similarity (27). In contrast, phylogenetic
111 analysis of two partially sequenced marker genes has placed PkV RF01 deep inside
112 the *Mimiviridae* clade, and characterization of its life cycle has revealed an atypical
113 infection profile (33). Here, we report new phenotypic features as well as new viral
114 functions inferred from analysis of the genome sequence of PkV RF01. We found that
115 this virus has a unique morphology, is less virulent than most other alga-infecting
116 viruses and possesses an unprecedented number of energy-generating genes. We
117 uncovered clues suggesting that members of *Mimiviridae* that potentially modulate
118 the metabolism of their hosts are widespread in the ocean. Our findings of peculiar
119 genomic features in a persistent virus provide new insights on virus–host coevolution
120 and may stimulate further advances in modeling the history of their interaction.

121 **Results and Discussion**

122 **PkV RF01 has an atypical morphology**

123 The icosahedral PkV RF01 particle is approximately 400 nm in diameter (Fig. 1A-B).
124 Beneath the capsid, several convoluted inner membranes fill approximately 66% of
125 the interior. Treatment of chloroform can be used to identify possible functions of
126 lipid membranes, as it acts to remove lipid molecules that might be essential for
127 successful infection (34). Some algal viruses in the NCLDV group are sensitive to
128 chloroform (30, 35, 36) with the suggestions that lipid containing inner or outer

129 membranes are involved in the infection process (35, 37). In our experiment,
130 chloroform treatment of PkV RF01 drastically reduced the infectivity of the virus.
131 (Fig. 1C). As no outer membrane was detected by cryo-electron tomography, the
132 sensitivity to chloroform might be linked to lipid components in either the capsid or
133 the inner convoluted membranes. Internal lipid-containing membranes have been
134 detected in several icosahedral-shaped double-stranded DNA viruses, including algal
135 viruses belonging to families *Phycodnaviridae* and *Mimiviridae*, mimiviruses, and
136 various bacteriophages (38–43). In all of these viruses, the inner membranes are
137 suggested to play a role in the release of the viral nucleoprotein core or genome by
138 fusing with the host plasma membrane (40, 42, 43). Inner membranes in currently
139 described NCLDVs more or less adopt the icosahedral morphology defined by the
140 outer layer of capsomers (44, 45). We detected several convoluted inner membranes
141 in PkV RF01 that do not follow the structure of the capsid. To our knowledge, this
142 structural inconsistency has not been previously detected in any double-stranded
143 DNA viruses, which calls for further investigation to understand the assembly process
144 of PkV RF01 and how it enters its host. Another striking feature of the PkV RF01
145 virion is an internal rod-shaped core (ca. 55 nm in diameter), which is filled with
146 dense material and positioned in the center of the virus particle. Similar features have
147 been observed in TEM images of large virus-like particles (VLPs) (300–700 nm)
148 occurring in waste vacuoles of phaeodarian radiolarians collected from different
149 oceans (46) and in zoospores of the green alga *Chlorococcus minutum* (47). To our
150 knowledge, however, these features have not been described in isolated viruses thus
151 far.

152 **PkV RF01 has an atypical infection strategy**

153 Only 2% of the total PkV RF01 viral particles produced during infection of *Haptolina*
154 *ericina* UiO028 (He UiO028) were infectious (able to produce progeny) (Table 1).
155 This infectivity was much lower than that of the other two prymnesioviruses, HeV
156 RF02 and PkV RF02, which produced 13% and 44% of infectious progeny
157 respectively (Table 1). The portion of infectious particles of PkV RF01 is low also
158 when compared to other algal viruses (48, 49). In addition, the latent period of PkV
159 RF01 was previously reported to be longer (ca. 24–32 h, (33)) in comparison with
160 other prymnesioviruses (28, 29, 32, 33) and it has been demonstrated that PkV RF01
161 is also able to infect multi-species (33), that is another unusual trait among algal
162 viruses (26).

163 The hosts of PkV RF01, PkV RF02, and HeV RF02 all belong to order the
164 Prymnesiales, whose members are normally present in low abundance but co-occur
165 year round (*K*-strategists) (50). PkV RF01, PkV RF02, and HeV RF02 are less
166 virulent, as shown in the present study, and have longer latent periods compared with
167 viruses infecting bloom-forming haptophytes (*r*-strategists). Two of these viruses
168 (PkV RF01 and HeV RF02) are also able to infect multi species (generalists) (33).
169 Longer replication time and reduced virulence, as hosts becomes scarce, increases the
170 chances of vertical transmission rather than horizontal transmission of a virus. As
171 vertical parent-to-offspring transmission depends on host reproduction, it has been
172 argued that such transmission should select for reduced virulence because the virus
173 depend on host survival and reproduction for its transmission (51, 52). High
174 virulence, on the other hand, may be supported by large, dense host populations, as
175 e.g. algal blooms, because high host densities ensure successful horizontal
176 transmission of viral progeny to new hosts (51, 53). Viruses infecting the recurrent

177 bloom-forming haptophytes, *Phaeocystis pouchetii* virus (PpV), and *Phaeocystis*
178 *globosa* virus (PgV), are indeed highly virulent with between 60%–100% of virus
179 particles produced being infectious, resulting in rapid lysis of their hosts (48, 54).
180 Broad host range might also increase the chance of transmission in an environment
181 with low host abundances (*K*-strategists). Such strategy requires a tradeoff whereby
182 the virus decreases its opportunity of transmission by evolving longer replication
183 times, higher decay rates and reduced infectivity (discussed in (55, 56)). This fits
184 well with our two multi-species infecting haptophyte viruses, PkV RF01 and
185 HeV RF02, that have reduced proportions of infectious particles and longer
186 replication times (33), relative to other haptophyte viruses with restricted host ranges
187 (specialists) like e.g. the *Emiliania huxleyi* virus (EhV), PpV and PgV.

188 The balance between fitness traits, such as virulence, latent period and host
189 range, and tradeoffs is the result of the adaptive evolution between viruses and their
190 hosts, resulting in relationships spanning from acute to stable coexistence
191 (persistence). In the ocean, persistent relationships—such as between PkV RF01 and
192 its hosts—seem to be most common among viruses infecting unicellular algae; this
193 has been demonstrated by several metabarcoding studies revealing the persistence of
194 dominance of viral OTUs over several months (57, 58). The atypical infection
195 strategy of PkV RF01 evokes a persistent nature, different than the vast majority of
196 other so far characterized algal viruses.

197 **PkV RF01 has the largest genome among algal viruses**

198 The genome of PkV RF01 was assembled as a linear DNA sequence of 1,421,182 bp.
199 This size is more than twice that of the genome of TetV, which means that PkV RF01
200 has the largest reported genome of any virus infecting a photosynthetic organism (Fig.
201 2A). Evidence for the linear structure of this genome is the presence of ~5-kbp

202 terminal inverted repeats. Despite being phylogenetically more closely related to alga-
203 infecting *Mimiviridae*, the genome size of PkV RF01 is in the range of heterotroph-
204 infecting *Mimiviridae*. The overall G+C content of PkV RF01 is 22.8%, which is low
205 compared with other *Mimiviridae* (23%–41%). Similar to other *Mimiviridae*, the
206 average G+C content of PkV RF01 in intergenic regions is relatively low, 17.8%.
207 This lower G+C content may reflect an ongoing loss of G and C nucleotides, more
208 prevalent in non-coding than coding regions because of weaker background selection
209 in non-coding regions. The genome of PkV RF01 is predicted to contain 1,161 genes
210 comprising 1,121 protein-coding DNA sequences (CDSs) and 40 tRNA genes
211 corresponding to 13 amino acids. Most tRNA genes (30 out of 40) are clustered in
212 three genomic regions that lack predicted CDSs, a feature also observed in other
213 *Mimiviridae*. For example, all tRNAs of TetV ($n = 10$) and CroV ($n = 22$) are encoded
214 consecutively on the same strand (18, 59). The average CDS length is 1,046 bp
215 (minimum: 297; maximum: 1,493). Intergenic regions average 217 bp in length, with
216 a cumulative sum of 244,005 bp, which corresponds to a gene density of 82.8%.

217 Of the 1,121 predicted CDSs, 641 (57%) exhibited sequence similarities
218 (BLASTP E -value conservative cutoff of 1×10^{-5}) to protein sequences in the
219 UniRef90 database (Fig. 2A). Among them, 165 were most similar to *Mimiviridae*.
220 Curiously, among the CDSs most similar to *Mimiviridae*, sixty were closest to
221 ChoanoVirus which was isolated from choanoflagellates cultures, followed by
222 Mesomimivirinae ($n = 49$) and Klosneuvirinae ($n = 30$) (Fig. 2B). Among the 181
223 closest homologs found in eukaryotic organisms 23 were haptophytes. A sequence-
224 based homology search of corrected nanopore reads and scaffolds composing the
225 initial assembly against *Lavidaviridae* proteomes (BLASTX; matrix: BLOSUM45, E -
226 value $< 1 \times 10^{-5}$) yielded no significant alignments against any major or minor

227 *Lavidaviridae* capsid proteins, which suggests that virophages were absent from the
228 sample used for sequencing.

229 A previous analysis of PkV RF01 family-B DNA polymerase (PolB) and the
230 major capsid protein (MCP) placed this virus in the family *Mimiviridae* (33). We also
231 recently reported that the PkV RF01 genome has additional NCLDV core genes, such
232 as A32-like virion packing ATPase (NCVOG0249) and RNAPol (RNA pol subunit I
233 [NCVOG0274] and subunit II [NCVOG0271]), and orthologous genes that are
234 specific to *Mimiviridae*, namely, MutS7 (NCVOG2626) and asparagine synthase
235 (AsnS, NCVOG0061) (60). Phylogenetic reconstruction using five NCLDV core
236 genes confirmed the deep branching of PkV RF01 within the *Mimiviridae* family and
237 suggested that PkV RF01, along with ChoanoV1, TetV and AaV, is more closely
238 related to Mesomimivirinae than to Megavirinae (Fig. 3A). In support of this
239 evolutionary relationship, PkV RF01 has an additional copy of the second largest
240 RNA polymerase subunit gene (*rpb2*). This *rpb2* duplication is shared with all other
241 *Mimiviridae* infecting algae, including Mesomimivirinae members, AaV (whose
242 second copy is very short), and TetV and was previously proposed as a useful feature
243 to discriminate between the two main clades (autotroph versus heterotroph-infecting
244 viruses) within the *Mimiviridae* family (27). This additional *rpb2* copy is not found in
245 other *Mimiviridae* to the exception of ChoanoV1 whose genome was derived from a
246 single cell metagenome in choanoflagellates cultures. Phylogenetic analysis indicates
247 that these two *rpb2* copies were present in the ancestor of alga-infecting *Mimiviridae*
248 and ChoanoV1 (Fig. 3B). In agreement with the five NCLDV core genes phylogeny,
249 it suggests that PkV RF01 and ChoanoV1, although evolutionarily distant, are more
250 related with each other compared to any other *Mimiviridae*.

251 Out of 1,121 predicted protein-coding genes in the genome of PkV RF01, only
252 about a third could be annotated with some functional description based on their
253 sequence homology with characterized proteins. Such a small percentage is typical of
254 divergent eukaryotic viruses detected for the first time. A total of 339 proteins (30%)
255 showed significant sequence similarity with proteins in the Cluster of Orthologous
256 Gene (COG) database (61) (Fig. 4). The distribution of COG functions associated
257 with these hits was dominated by “Posttranslational modification, protein turnover,
258 chaperones” (43 proteins) and “Cell wall/membrane/envelope biogenesis” (42
259 proteins), which is approximately two times more proteins than in other *Mimiviridae*
260 members except for Tupanvirus. Among other well-represented categories, numbers
261 of proteins in “Replication, recombination and repair” (36 proteins) and
262 “Transcription” (23 proteins) were similar to those of other *Mimiviridae*, while the
263 categories of “Translation, ribosomal structure and biogenesis” (25 proteins) and
264 “Amino acid transport and metabolism” (20 proteins) were respectively in the same
265 range or higher than those of heterotroph-infecting *Mimiviridae* (mimiviruses, BsV,
266 and CroV). Interestingly, 24, 17, and 9 PkV RF01 proteins were respectively assigned
267 to the categories of “Lipid transport and metabolism”, “Carbohydrates transport and
268 metabolism,” and “Energy production and conservation,” all much higher compared
269 with other *Mimiviridae* viruses (Fig. 5).

270 Similar to other *Mimiviridae*, PkV RF01 encodes several genes involved in
271 DNA repair, transcription, and translation. Notably, this virus has the full set of
272 enzymes required for the base excision repair (BER) pathway, which is also the case
273 for all *Mimiviridae* members except for those with smaller genomes (PgV, CeV, and
274 AaV). PkV RF01 BER enzymes are closer (i.e., have a greater alignment score) to
275 heterotroph-infecting *Mimiviridae* than to cellular homologs, thus suggesting that this

276 pathway was present in the last common ancestor of *Mimiviridae*. According to a
277 previous phylogenetic analysis, *Mimiviridae* BER enzymes are monophyletic with
278 regard to *Mimiviridae* and have not recently been acquired from eukaryotes (62).

279 Unlike alga-infecting *Mimiviridae*, PkV RF01 encodes two amino-acyl tRNA
280 synthetases (aaRS): an isoleucyl-tRNA synthetase (IleRS; ORF 480) and an
281 asparaginyl-tRNA synthetase (AsnRS; ORF 764). Both of these synthetases are found
282 in most lineages of heterotroph-infecting *Mimiviridae* (AsnRS is missing from CroV
283 and BsV, and IleRS is missing from *Mimivirus* lineage A). Phylogenetic analyses of
284 these two proteins revealed a deep branching of viral homologs, which formed a
285 monophyletic clade well separated from cellular homologs (Fig. 6).

286 **A viral-encoded succinate dehydrogenase and energy production genes**

287 We found six predicted protein-coding genes (ORFs 893 to 900) related to energy
288 production in an 8,026-bp region (Fig. 7A). Four ORFs (ORFs 893 and 898–900)
289 were predicted to code for all four subunits (SDHA, D, C, and B) of a functional
290 succinate dehydrogenase (SDH, or Electron Transport Chain Complex II) of the
291 oxidative phosphorylation pathway (Fig. 7B). In eukaryotes, all four subunits of this
292 enzyme are encoded in the nuclear genome. This enzyme acts in the mitochondrial
293 respiratory chain and participates in both the TCA cycle and the respiratory electron
294 transfer chain. In the TCA cycle, this succinate dehydrogenase oxidizes succinate to
295 fumarate, while its activity in the inner mitochondrial membrane involves the
296 reduction of a FAD cofactor followed by electron transfer through three Fe–S centers
297 to ubiquinone (Fig. 7C).

298 *SDH* genes have recently been reported in viral genomes assembled from
299 environmental samples for which functional experiments cannot be done (63). In a
300 RT-PCR experiment using primers specific for the PkV RF01 gene for SDHA

301 (hereafter, *vSDHA*), we detected transcripts of this gene in samples collected 24, 72,
302 and 96 h post infection (Fig. 8). The *vSDHA* primers were tested on an uninfected
303 culture to ensure that only the viral version of the *SDHA* gene was amplified (Fig. 9).
304 The MCP gene of PkV RF01 was used both for protocol optimization and later as an
305 internal positive control (Fig. 10). Although the transcription of the viral *SDHA*
306 suggests that the viral SDH is functional, we can only speculate on the possible role
307 of this enzyme during infection. One possibility is that the viral SDH sustains the
308 carbohydrate metabolism of infected cells (i.e., virocells) to supply building blocks of
309 viral particles such as amino acids and to support proper replication of this large virus.
310 Another possibility is that PkV RF01 uses its SDH as a part of an arms race with its
311 host to turn on the TCA cycle after the host had turned it off to counter viral
312 replication, or more simply to boost the energy metabolism of the virocells to
313 augment the fitness of the host and/or to maximize virus production efficiency.

314 The discovery of the viral SDH prompted us to search for other potential viral-
315 encoded *SDHA* and *SDHB* homologs in marine metagenomes. These two subunits
316 (*SDHA* and *SDHB*) form the catalytic core containing the redox cofactors that
317 participate in electron transfer to ubiquinone; they are thus more conserved than
318 *SDHC* and *SDHD* subunits. To test for the presence of this viral SDH in other viruses,
319 we searched for *vSDHA* and *B* in marine metagenomes of the *Tara* Oceans
320 expedition. The 50 most-similar and non-redundant *SDHA* and *B* sequences predicted
321 from 101 *Tara* Oceans genome fragments were most likely derived from *Mimiviridae*
322 viruses (Fig. 11). Indeed, out of 1,113 genes predicted from these 101 genome
323 fragments, 681 were annotated at some taxonomic level, of which 449 were predicted
324 to be cellular and 157 viral. Of the 157 viral genes, 146 and 130 had their last
325 common ancestor in *Mimiviridae* and *Mesomimivirinae*, respectively. A total of 32 of

326 the 101-genome fragments contained at least one gene predicted to be of *Mimiviridae*
327 origin, and the larger the genome fragment, the more *Mimiviridae* genes it was found
328 to encode (Fig. 11A). Functional analysis indicated that 12 of the 1,113 predicted
329 genes were NCLDV hallmark genes (encoding five VLTF3s, two capsid proteins, two
330 PCNAs, two helicases, and one PolB). The high proportion of unknown genes and
331 genes annotated as *Mimiviridae* in the 101 *Tara* Oceans genome fragments encoding
332 SDHA or SDHB strongly suggests that these fragments belong to *Mimiviridae*
333 viruses. This finding demonstrates that the presence of SDH is not restricted to PkV
334 RF01 and is arguably widespread among marine *Mimiviridae*. According to
335 phylogenetic analyses of cellular and viral SDHA and SDHB, the viral homologs
336 form a monophyletic group that branches deeply within eukaryotic lineages (Fig.
337 11B-C). Long-branch attraction bias could generate such topologies but, as explained
338 above for the IleRS and AsnRS, it is more likely that the viral SDHA and SDHB were
339 acquired at an early stage in the radiation of eukaryotic lineages. The transcription of
340 *vSDHA* and its occurrence in marine environments calls for further investigation to
341 understand the biological role and co-evolutionary significance of this viral SDH.

342 Other genes related to energy production were detected in the 8,026 bp-long
343 region. ORF 894 and ORF 896, respectively corresponding to cytochrome *c* (CytC)
344 and cytochrome b6-f complex iron-sulfur (Cyt b6-f) subunits, showed high sequence
345 conservation with *Chrysochromulina* sp. CCMP291 proteins (78% and 59% amino
346 acid [aa] identities, respectively). CytC is a short protein (~100 aa) involved in the
347 oxidative phosphorylation pathway, where it accommodates the transfer of electrons
348 between the coenzymes Q-cytochrome *c* reductase (complex III) and cytochrome *c*
349 oxidase (complex IV). The presence of Cyt b6-f between oxidative phosphorylation
350 genes is puzzling because the cytochrome b6-f complex is involved in photosynthesis.

351 The core of the chloroplast b6f complex, however, is similar to the analogous
352 respiratory cytochrome bc(1) complex. The other two predicted ORFs in this region
353 are similar to ubiquinone biosynthesis protein UbiB (ORF 895) or contain a NAD-
354 binding domain and a Fe-S cluster (ORF 897) and may thus be associated with
355 electron transport as well. ORF 897 has two distant (25%–31% aa identity) homologs
356 in the PkV RF01 genome (ORF 456 and ORF 625).

357 Some other genes were predicted to encode enzymes involved in pyruvate
358 metabolism. ORF 79 has sequence homology with L-lactate dehydrogenases; it might
359 thus catalyze the conversion of lactate to pyruvate, an intermediary compound serving
360 as a starting point for several major metabolic pathways, such as glycolysis,
361 gluconeogenesis, and the TCA cycle. ORF 727 was predicted to code for an
362 isochorismate hydrolase that also produces pyruvate from isochorismate. ORF 24 and
363 ORF 726 share sequence homology with phosphoenolpyruvate synthase and a partial
364 pyruvate kinase, respectively. The former catalyzes the conversion of pyruvate to
365 phosphoenolpyruvate (PEP), while the latter catalyzes the reverse reaction. Formation
366 of PEP is an initial step in gluconeogenesis.

367 **A nearly complete viral-encoded β -oxidation pathway**

368 In this study, 22 predicted genes were inferred to code for proteins involved in lipid
369 synthesis or degradation, including key enzymes of the β -oxidation pathway (Table
370 2). Several genes were predicted to code for lipase-like proteins (ORFs 386, 481, 635,
371 653, and 690), including a triacylglycerol lipase (ORF 386) that can break down
372 triacylglycerol into glycerol and fatty acids. Glycerol and fatty acids can be used as a
373 starting point for ATP production—by glycolysis and β -oxidation, respectively. In the
374 β -oxidation pathway, fatty acids are fully oxidized to produce acetyl-CoA, which can
375 then enter the TCA cycle to yield NADH and FADH₂; these latter two products can

376 funnel through to the electron transport chain to produce ATP (Fig. 7C). Each β -
377 oxidation cycle itself also produces NADH and FADH₂ cofactors. We found that PkV
378 RF01 encodes key β -oxidation enzymes. First, two distantly related ORFs (ORF 142
379 and ORF 904 sharing 22% aa identity) have sequence homology with a long-chain
380 fatty acyl-CoA synthetase. This enzyme catalyzes the formation of fatty acyl-CoA in
381 the cytosol. Fatty acyl-CoA can be imported to mitochondria using a (carnitine) CoA-
382 transferase also encoded in PkV RF01 (ORF 33). Once in the mitochondrial matrix,
383 fatty acyl-CoA serves as a substrate on which an acyl-CoA dehydrogenase (ORF
384 1046) oxidizes the fatty acyl-CoA and reduces a FAD cofactor to produce a FADH₂
385 cofactor. We identified a 2,4-dienoyl-CoA reductase (ORF 30) that may facilitate the
386 next oxidation step to produce a NADH cofactor. FADH₂ and NADH molecules
387 produced by a β -oxidation cycle can both be oxidized in the electron transport chain
388 to generate ATP. The enzymes involved in the two intermediate steps following each
389 oxidation, either an enoyl-CoA hydratase or a β -ketothiolase, were not detected in our
390 analysis.

391 Most of these genes have no homologs in reference viral genomes, and, to our
392 knowledge, this is the first report of a virus possessing proteins directly involved in
393 lipid-based energy production. By diverting host lipid machinery, interactions of
394 viruses with lipids or lipid based-structures have long been known to have structural
395 or signaling roles at different stages of the virus life cycle, such as entry, genome
396 replication, morphogenesis, and exit (64–66). More recently, several studies on
397 human viruses (two herpesviruses and one RNA virus) have shown that the metabolic
398 state of an infected cell can be shifted toward energy generation to support viral
399 replication (65). These studies have highlighted the increasing abundance—up to 48 h
400 after HCV infection—of enzymes involved in β -oxidation, amino acid catabolism,

401 and the TCA cycle (67) and an increase in cellular β -oxidation following the release
402 of free fatty acids caused by Dengue virus-induced autophagy (68). Among algal
403 viruses, EhV remodels the transcription of host lipid genes for fatty acid synthesis to
404 support viral assembly (69) and also to generate triacylglycerols stored in the virion
405 and available as an energy pool in later infection phases (70). Besides diverting the
406 host metabolism, EhV encodes seven proteins involved in the sphingolipid
407 biosynthesis pathway (71). This pathway produces a viral sphingolipid that is a
408 central component of EhV lipid membranes and that can also act as a signaling lipid
409 and induce programmed cell death during the lytic infection phase (72). EhV also
410 encodes a triglyceride lipase (with detectable homology to predicted PkV RF01
411 lipases ORF 635 and ORF653) that is highly expressed during late infection
412 concomitantly with significant up-regulation of host β -oxidation genes (69). These
413 examples and our observations of several genes involved in β -oxidation clearly show
414 that viruses can introduce new metabolism-related genes, sometimes representing
415 entire pathways, into the host, most likely to satisfy the high metabolic requirement of
416 these giant viruses.

417 **High representation of glycosyltransferases**

418 Compared with other viruses, PkV RF01 was found to encode an unusually high
419 number of glycosyltransferases (GTs) as well as other carbohydrate-active enzymes.
420 Automated annotation of GTs (and other carbohydrate-active enzymes) in reference
421 viral proteomes using dbCAN2 (73) revealed that the largest number of GT domains
422 was encoded by PkV RF01 ($n = 48$), followed by CeV ($n = 13$), *Mimivirus* members,
423 and CroV and AaV ($n = 8-10$) (Fig. 12). We uncovered 48 GT domains encoded in
424 40 ORFs, 8 of which were predicted to encode more than one GT domain. These
425 domains correspond to 16 different GT families. Most domains were inferred to be

426 functional, as 31 out of 48 covered at least 70% of the dbCAN2 reference domain,
427 with coverage ranging from 44% to 99%. GTs were found scattered across the
428 genome of PkV RF01 but with some local clustering (Fig. 2A), the latter indicating
429 possible involvement in the same pathway. GT32 was the most represented domain,
430 with 11 proteins (as annotated by dbCAN2) and potentially three additional proteins
431 (ORFs 40, 84, and 861). Eight proteins possessed a GT25 domain that can catalyze
432 the transfer of various sugars onto a growing lipopolysaccharide chain during its
433 biosynthesis. Among these eight predicted ORFs, four contained an additional non-
434 overlapping GT domain (two GT2s, one GT6, and one GT60). Functional analyses of
435 GTs in mimiviruses (or in related *Paramecium bursaria* Chlorella viruses) have
436 demonstrated that some of these enzymes are functional, being able to modify viral
437 collagen-like proteins (74) and polymerize sugars (75). Conservation between PkV
438 RF01 GTs and functionally characterized GTs in viruses and cells is absent or
439 extremely low, which precludes any predictions as to the specific roles of these
440 enzymes in the PkV RF01 life cycle. Nevertheless, this putative glycosylation-
441 conducive autonomy possibly allows the virus to infect a variety of hosts, as the virus
442 can modify its own glycans, which are used for host recognition, independently of the
443 host system (76). In alpha-, flavi-, and herpes-viruses, fusion is mediated by viral
444 glycoproteins (40).

445 Other carbohydrate-active enzymes in the PkV RF01 genome include seven
446 glycoside hydrolases (GHs), four carbohydrate esterases (CEs), one polysaccharide
447 lyase (PL), one carbohydrate-binding module (CBM), and a putative sugar
448 fermentation stimulation protein A (ORF 1003) possibly involved in maltose
449 metabolism. These numbers are not excessively high compared with other viruses.
450 Other detected ORFs were homologous to enzymes involved in carbohydrate

451 transport and metabolism, notably a transketolase (ORF 528) involved in the pentose
452 phosphate pathway in all organisms and in the Calvin cycle of photosynthetic
453 organisms. Finally, we detected a 6-phosphofructo-2-kinase/fructose-2,6-
454 biphosphatase 2 (ORF 539) and a mannose-1-phosphate
455 guanylyltransferase/mannose-6-phosphate isomerase (ORF 836) respectively involved
456 in fructose and mannose metabolism.

457 **Conclusions**

458 The haptophyte virus PkV RF01 has been previously shown to have a longer
459 replication cycle and a broader host range compared with other prymnesioviruses and
460 most other algal viruses. Here, we revealed that PkV RF01 has atypical virion
461 morphology and that infections yield several orders of magnitude fewer infectious
462 particles than other tested prymnesioviruses. In-depth phylogenetic analysis using
463 genes conserved in NCLDVs confirmed that PkV RF01 belongs to *Mimiviridae* but is
464 deeply diverged from existing members, although closer to alga-infecting
465 *Mimiviridae* than heterotroph-infecting ones. Unlike other alga-infecting *Mimiviridae*,
466 however, PkV RF01 has a large genome (1.4 Mb) and contains genes coding for two
467 aminoacyl-tRNA synthetases and the complete BER pathway. All these features are
468 conserved in most heterotroph-infecting *Mimiviridae* and therefore must have been
469 lost in other alga-infecting *Mimiviridae*. This outlier virus features an
470 unprecedentedly high number of genes involved in energy metabolism and
471 glycosylation machinery that may enable its longer replication cycle and broader host
472 range compared with other algal viruses. These genomic and phenotypic features are
473 suggestive of a persistent infection behavior that probably evolved in response to the
474 host growth strategy. Because of nutrient limitations, these persistent systems of slow-

475 growing but ubiquitous hosts with less virulent viruses may represent the most
476 common type of virocells in oceans.

477 **Materials and Methods**

478 **Culturing and infection**

479 All algal host cultures were grown in liquid IMR/2 medium consisting of 70% aged
480 seawater, 30% distilled water (25 PSU), and additional selenite (10 nM final
481 concentration). The cultures were kept at 14°C and partially synchronized using a
482 14:10 h light: dark cycle with irradiance of 100 $\mu\text{mol photons m}^{-2} \text{s}^{-2}$ supplied by
483 white fluorescent tubes. Viruses were produced by adding freshly produced viral
484 lysate (ca. 2×10^8 VLP/mL), propagated three time on the host before added to
485 exponentially growing host cultures (ca. 5×10^5 cells/mL) in a ratio of 1:10 volume.
486 Infection was followed by flow cytometry (FCM) (77, 78) for 72 h by counting viral
487 particles and host cells, as described in (33). Burst size was calculated as the number
488 of viral particles released from each host cell, estimated from the total number of host
489 cells pre-infection and the total number of VLPs produced during the infection cycle
490 (33).

491 **Infectious progeny**

492 The percentage of viral infectious progeny was determined by comparing the most
493 probable number (MPN; endpoint dilution (78)) and flow cytometric total counts of
494 viral particles produced during infection. The number of infectious particles released
495 in a burst was determined based on the percentage of viral infectivity produced during
496 the infection cycle and the burst size. Infectivity was tested using *Haptolina ericina*
497 UiO028 as a host, and also compared with two other prymnesioviruses, HeV RF02

498 and PkV RF02 (33), propagated on He UiO028 and *Prymnesium kappa* RCC3423,
499 respectively.

500 Briefly, 10× dilution were prepared from fresh viral lysate and added to
501 exponentially growing host cells in 96-well microtiter plates (eight replicates for each
502 dilution). The plates were incubated for 7 days under normal incubation conditions.
503 Cell lysis was measured by monitoring *in situ* fluorescence on a plate reader
504 (PerkinElmer EnSpire™ 2300 Multilabel Reader) at 460/680 nm. Numbers of
505 infectious particles were estimated from the proportion of lysed wells using the
506 MPN_ver4.xls excel spreadsheet from (79).

507 **Sensitivity to chloroform**

508 The effect of chloroform on infectivity, used to infer the presence of a lipid membrane
509 or lipid molecules in the capsid, was tested by adding 50% (v/v) chloroform to PkV
510 RF01 lysate. After mixing, the chloroform phase was separated from the solution by
511 centrifugation at 4,000 *g* for 5 min. The tubes were incubated at 37°C for 2 h with the
512 lids open to allow evaporation of any remaining chloroform.

513 Triplicates of exponentially growing He UiO028 cells ($1,6 \times 10^5$ cells /mL)
514 were incubated with 1:10 volumes of chloroform-treated viruses (ca. 2×10^8
515 VLP/mL). The incubation was followed for 7 days by counting host cells by FCM
516 (78). Host cells in chloroform-treated or untreated medium at the same ratio used with
517 the viral lysate were used as controls. Virus propagation was confirmed in lysed
518 cultures by FCM.

519 **Cryo-electron tomography**

520 A small drop of concentrated PkV RF01 (8×10^9) was deposited on a glow-discharged,
521 200-mesh copper grid with holey carbon film (R2/1 Cu 200, Quantifoil Micro Tools

522 GmbH, Germany). The sample was blotted with filter paper and immediately plunge
523 frozen in liquid ethane. Grids were transferred under liquid nitrogen to a cryo-transfer
524 tomography holder (Fishione Instruments, USA) and inserted in a 200-kV
525 transmission electron microscope (Thermo Scientific Talos F200C) equipped with a
526 Ceta 16M camera. Tilt series were recorded at 45,000× magnification and $-7\ \mu\text{m}$
527 defocus between -60° to 60° in 2° increments. Finally, reconstruction, segmentation,
528 and visualization of the tomograms was performed with IMOD v4.9 software (80).

529 **Purification of viral particles and DNA isolation**

530 Exponentially growing He UiO028 cultures (2 L) were infected with 20 mL of PkV
531 RF01 and inspected visually for lysis. An uninfected culture (100 mL) was used as a
532 control. Lysed algal cultures were checked for viruses by FCM counting. Lysed
533 cultures were first centrifuged to remove algal debris and some bacteria (5,500 rpm
534 for 15 min). Viruses were then pelleted by ultracentrifugation at 25,000 rpm in a
535 Beckman Coulter Optima L90K ultracentrifuge for 2 h. The pellets were resuspended
536 in SM buffer (0.1 M NaCl, 8 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 50 mM Tris-HCl, and 0.005%
537 glycerin). Viral particles were further purified by Optiprep gradient centrifugation
538 (81). Fractions were checked for viruses by FCM and for infectivity by infection of
539 He UiO028.

540 Isolation of high-quality DNA for sequencing was done by following the
541 protocol of (82) with some modifications. Viral particles were disrupted by one
542 round of heating to 90°C for 2 min and then chilling on ice for 2 min. Disodium
543 ethylenediaminetetraacetic acid and proteinase K at a final concentration of 20 mM
544 and $100\ \mu\text{g mL}^{-1}$, respectively, were then added before incubation of the samples for
545 10 min at 55°C . Sodium dodecyl sulfate at a final concentration of 0.5% (w/v) was
546 subsequently added, and samples were incubated for an additional 1 h at 55°C .

547 Double-stranded DNA was then purified from the lysates using a Zymo Genomic
548 DNA Clean & Concentrator Kit-10 (Zymo Research, Irvine, CA, USA) according to
549 the manufacturer's protocols. To avoid shearing DNA, gentle pipetting and mixing
550 (accomplished by turning the tubes instead of vortexing) were performed in all steps.

551 **Genome assembly**

552 Isolated DNA from PkV RF01 was subjected to Illumina TruSeq PCR-free library
553 preparation (insert size 350 bp). The generated library was sequenced on an Illumina
554 MiSeq instrument in paired-end mode (2 × 300 bp) to yield approximately 1.9 million
555 reads, which corresponds to about 400× coverage. Reads were assembled into 2,498
556 contigs of 500 bp or more with a total assembly size of 4.75 Mb using Newbler (83).
557 In addition, a ligation-based 1D² nanopore library (LSK-308) was constructed and
558 sequenced using an Oxford Nanopore MinION Mk1b device and a FLO-MIN107
559 flow cell, which resulted in 825 long reads with an N50 of 13.6 kb and a total of 9.89
560 Mb. To improve the assembly, short-read contigs were manually bridged with the
561 long reads. Manual assembly using Consed (84) yielded a linear genome sequence of
562 1.4 Mb with inverted terminal repeats. After assembly, the consensus was polished
563 using Nanopolish (85) and Pilon (86).

564 **Phylogenetic analyses**

565 **Five core genes, SDHA, and SDHB**

566 The phylogenetic position of PkV RF01 was inferred from concatenated protein
567 alignments of five core nucleocytoplasmic virus orthologous genes (NCVOGs) (87):
568 D5-like helicase-primase (NCVOG0023), DNA polymerase elongation subunit family
569 B (NCVOG0038), DNA or RNA helicases of superfamily II (NCVOG0076),
570 packaging ATPase (NCVOG0249), and Poxvirus Late Transcription Factor VLTF3-

571 like (NCVOG0262). Sequences were obtained from the NCVOG database
572 (<ftp.ncbi.nlm.nih.gov/pub/wolf/COGs/NCVOG/>) (88). Additional sequences were
573 obtained from genomes retrieved from GenBank and annotated with HMMER v3.12b
574 using the `hmmsearch` (89) command with hidden Markov models available in Schulz
575 et al. (2017) (13). Sequences from each NCVOG were aligned independently using
576 MAFFT L-INS-i (90). The alignments were trimmed with trimAl v1.2 in *gapyout*
577 mode (91) prior to concatenation using a custom Python script. Bayesian phylogenetic
578 trees were inferred with PhyloBayes 1.7 (92) using the CAT model and a GTR
579 substitution matrix. Four chains were run for 34,500–35,500 generations. The *bpcomp*
580 command was used to check for convergence and stop when *maxdiff* = 0.3. One chain
581 was discarded, and a consensus tree was constructed using the remaining three chains.

582

583 For phylogenetic analyses of succinate dehydrogenase subunits, top hits of PkV RF01
584 SDHA and SDHB were retrieved from UniProt (<https://www.uniprot.org/>) using
585 online PHMMR searches (<https://www.ebi.ac.uk/Tools/hmmer/search/phmmer>) and
586 also from the *Tara* Oceans project using online BLASTP searches ([http://tara-](http://tara-oceans.mio.osupytheas.fr/ocean-gene-atlas/)
587 [oceans.mio.osupytheas.fr/ocean-gene-atlas/](http://tara-oceans.mio.osupytheas.fr/ocean-gene-atlas/)) (Villar et al., 2018). Alignments
588 generated with MAFFT L-INS-i were filtered with trimAl in *gapyout* mode.

589 Maximum-likelihood phylogenies were inferred with RAxML 8.2.9 (93) using the
590 PROTCATALG model and automatic bootstrapping with the following options: ‘-N
591 autoMRE -f a -n autoresult’. Phylogenetic trees of PkV RF01, SDHA, and SDHB
592 were visualized using iTOL (94).

593 **Rpb2, IleRS, and AsnRS**

594 To reconstruct a phylogenetic tree based on the second largest RNA polymerase
595 subunit, homologs were recruited by comparing Mimivirus Rpb2 against all proteins

596 of viruses and selected organisms in the KEGG database using the GenomeNet
597 BLASTP tool (<https://www.genome.jp/>). Organisms were manually selected from the
598 KEGG list to ensure broad taxonomic coverage of the tree of life. The retrieved amino
599 acid sequences were aligned using MAFFT-LINSI (90) and then trimmed using
600 trimAl (91) with the following parameters: ‘-resoverlap 0.5 -seqoverlap 70 -gt 0.8 -st
601 0.001 -cons 50’. The tree was reconstructed using FastTree (95) as implemented in
602 the GenomeNet TREE tool (<https://www.genome.jp/tools-bin/ete>). Isoleucine tRNA
603 synthase and aspartyl tRNA synthetase viral and cellular homologs were retrieved and
604 aligned in the same way. Trees were searched using PhyloBayes MPI (96) with the
605 non-homogeneous CAT+GTR model (97). For each protein three chains were run
606 until *maxdiff* parameter reach < 0.3 (0.27 for AsnRS and 0.16 for IleRS). One chain
607 was discarded for IleRS, and a consensus tree was constructed using the remaining
608 chains.

609 **Gene prediction and functional and taxonomic annotation**

610 GeneMarkS with the option ‘virus’ (98) predicted 1,121 open reading frames (ORFs)
611 in the fully assembled genome sequence of PkV RF01, while tRNAscan-SE (99)
612 predicted 41 tRNAs. PkV RF01 CDS amino acid sequences were searched against
613 Virus-Host DB (100), RefSeq (101), UniRef90 (102), and COG (61) databases using
614 BLASTP with an *E*-value of 1×10^{-5} as the significant similarity threshold and
615 against the Conserved Domain Database (103) using RPS-BLAST with an *E*-value
616 threshold of 1×10^{-2} . The 10 best hits for each database were compiled in a single file
617 and manually inspected to transfer annotations of subject sequences to our query. In
618 ambiguous cases, such as distant homologs (often seen in viral genomes) or unclear or
619 contradictory annotations of subject sequences, the query was searched against KEGG
620 genes (104) to allow extensive manual checking using GenomeNet tools

621 (<https://www.genome.jp/>; alignment quality, length comparison to canonical genes,
622 and links with KEGG orthology). We automatically annotated glycosyltransferases
623 (GTs) and other carbohydrate-active enzymes (glycoside hydrolases, GHs;
624 polysaccharide lyases, PLs; carbohydrate esterases, CEs; and auxiliary activities,
625 AAs) in PkV RF01 and all viral genomes in Virus-Host DB (as of June 2018) using
626 the *hmm* option of the dbCAN2 pipeline and its profile database (73). We retained hits
627 with *E*-values $< 1 \times 10^{-5}$ and domain coverage $> 35\%$, which corresponded to default
628 settings.

629 **Taxonomic and functional analysis of vSDHA homologs in OM-RGCv1**

630 We searched PkV RF01 SDHA and SDHB against OM-RGCv1 (105) using the
631 Ocean Gene Atlas (106) BLAST-based tool and kept the top 50 hits with significant
632 *E*-values for further analysis. We then collected genome fragments (contigs) encoding
633 these 50 SDHAs and 50 SDHBs by searching via BLASTN for identical hits over full
634 *SDHA* or *SDHB* lengths against *Tara* ocean assemblies (downloaded from EBI) used
635 to construct OM-RGCv1. We predicted ORFs in these genome fragments using
636 GeneMarkS. The resulting 1,113 amino acid sequences were functionally annotated
637 by searching against Pfam protein families (107) using profile HMM scan (108) and
638 also taxonomically using a last common ancestor strategy as in (109); in brief, protein
639 sequences were searched against a database composed of UniRef cells, MMETSP
640 (110) and Virus-Host DB (100) data using DIAMOND (111). Selected hits were then
641 used to derive the last common ancestor of the query using a NCBI taxonomic tree re-
642 wired to reflect the taxonomy of NCLDV.

643 **PCR and RT-PCR optimization**

644 We designed specific primers (Table 3) targeting a 256-bp region of the *mcp* gene to
645 use both as an internal control in the RT-PCR and to confirm that our protocols were
646 optimized. For each PCR, a negative control (sterile distilled H₂O) was included. PCR
647 amplifications were carried out in 50- μ L total volumes containing 1 μ L of template
648 using a DNA HotStarTaq Master Mix kit (Qiagen). The cycling protocol was as
649 follows: 15 min at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 59°C, and 30 s
650 at 72°C, with a final extension of 12 min at 72°C.

651

652 RT-PCRs were performed using the SuperScript III One-Step RT-PCR with Platinum
653 *Taq* DNA Polymerase system (Thermo Fisher). Cycling conditions were as follows:
654 16 min at 55°C and 2 min at 94°C, followed by 40 cycles of 15 s at 94°C, 30 s at
655 49°C, and 30 s at 68°C, and a final extension of 5 min at 68°C.

656

657 All PCR products were checked for the correct size on a 1.5% agarose gel stained
658 with GelRed (Biotium). PCR products were further checked by sequencing using
659 BigDye v3.1 (Thermo Fisher) for cycle sequencing (Sekvenseringslaboratoriet, UiB,
660 Norway).

661 **PCR amplification and RT-PCR analysis of *vSDHA***

662 To investigate whether the *vSDHA* gene is transcribed during infection, an infected
663 culture of He_UiO028 plus PkV RF01 as well as an uninfected He_UiO028 culture
664 (control) were set up as described above. Samples were collected at 24, 72, and 96 h
665 post infection from both cultures. RNA was extracted using an RNeasy Plus Universal
666 Mini kit (Qiagen), with gDNA removed in an extra step using a TURBO DNA-free
667 kit (Ambion).

668

669 Specific primers were designed to target a 150-bp region of the *vSDHA* gene (Table
670 3). For each PCR, two negative controls (sterile distilled H₂O and extracted DNA
671 from He028) were included. As positive controls for the transcription, we used
672 primers targeting the *mcp* gene (see above). As a positive PCR control, we used
673 genomic PkV RF01 DNA. PCR amplifications were conducted in 50- μ L total
674 volumes containing 1 μ L of template DNA using an ExTaq kit (Takara). The cycling
675 protocol was as follows: 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at
676 59°C, and 30 s at 72°C, with a final extension of 12 min extension at 72°C.

677

678 RT-PCRs were performed using a SuperScript III One-Step RT-PCR with Platinum
679 Taq DNA Polymerase system (Thermo Fisher). Cycling conditions were as follows:
680 16 min at 55°C and 2 min at 94°C, followed by 40 cycles of 15 s at 94°C, 30 s at
681 49°C, and 30 s at 68°C, with a final extension of 5 min at 68°C. PCR products were
682 checked as described above.

683 **Data availability**

684 Raw sequence reads and PkV RF01 genome sequence were deposited at the European
685 Bioinformatics Institute (EMBL-EBI) (<https://www.ebi.ac.uk>) under project name
686 PRJEB37450. The complete video records of a cryo-electron tomogram of a PkV
687 RF01 virion and sequence data as well as curated gene annotation table as reported in
688 this study are available at <https://github.com/RomainBlancMathieu/PkV-RF01>.

689 **Acknowledgements**

690 The recording of tilt series was performed with the help of Sebastian Schultz at the
691 Unit of Cellular Electron Microscopy, the Norwegian Radium Hospital. Initial
692 sequencing (MiSeq and Pacbio) of PkV RF01 total DNA was performed at the
693 Norwegian Sequencing Center (<https://www.sequencing.uio.no/>). We thank Hilde M.
694 K. Stabell and Solveig Siqveland, Department of Biological Sciences, University of
695 Bergen, Norway, for technical assistance with molecular biology experiments as well
696 as Christian Rückert, Bielefeld University, for support in manual finishing of genome
697 assembly and Minyue Fan, Kyoto University, for assistance in genes analysis. This

698 work was supported by the Research Council of Norway project entitled “Uncovering
699 the key players for regulation of phytoplankton function and structure: lesson to be
700 learned from algal virus-haptophyte coexistence” (VirVar, project number 294364 to
701 RAS). Additional funding was provided by the European Union Horizons 2020
702 research and innovation program, grant agreement no. 685778 (“Virus-X”) to RAS
703 and DB. This work was also supported by the Future Development Funding Program
704 of the Kyoto University Research Coordination Alliance. HO was supported by
705 JSPS/KAKENHI (No. 18H02279), and Scientific Research on Innovative Areas from
706 the Ministry of Education, Culture, Science, Sports and Technology (MEXT) of Japan
707 (Nos. 16H06429, 16K21723, 16H06437). The Super Computer System, Institute for
708 Chemical Research, Kyoto University, provided computational time. We thank
709 Barbara Goodson, from Edanz Group (www.edanzediting.com/ac), for editing the
710 English text of a draft of this manuscript.

711 **Competing interests**

712 Authors declare having no competing interests.

713 **References**

- 714 1. Dupré JO. 2009. Varieties of Living Things: Life at the Intersection of Lineage
715 and Metabolism. *Philosophy & Theory in Biology* 1.
- 716 2. Forterre P. 2012. Virocell Concept, The, p. a0023264. *In* John Wiley & Sons,
717 Ltd (ed.), eLS. John Wiley & Sons, Ltd, Chichester, UK.
- 718 3. Raoult D, Forterre P. 2008. Redefining viruses: lessons from Mimivirus. *Nat*
719 *Rev Microbiol* 6:315–319.
- 720 4. Claverie J-M. 2006. Viruses take center stage in cellular evolution. *Genome*
721 *Biology* 7:110.
- 722 5. Mann NH, Cook A, Millard A, Bailey S, Clokie M. 2003. Bacterial
723 photosynthesis genes in a virus. *Nature* 424:741–741.
- 724 6. Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW.
725 2004. Transfer of photosynthesis genes to and from Prochlorococcus viruses.
726 *PNAS* 101:11013–11018.
- 727 7. Hurwitz BL, Hallam SJ, Sullivan MB. 2013. Metabolic reprogramming by
728 viruses in the sunlit and dark ocean. *Genome Biol* 14:R123.
- 729 8. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, Chisholm
730 SW. 2011. Phage auxiliary metabolic genes and the redirection of cyanobacterial
731 host carbon metabolism. *Proc Natl Acad Sci USA* 108:E757-764.

- 732 9. Anantharaman K, Duhaime MB, Breier JA, Wendt KA, Toner BM, Dick GJ.
733 2014. Sulfur Oxidation Genes in Diverse Deep-Sea Viruses. *Science* 344:757–
734 760.
- 735 10. Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW. 2005.
736 Photosynthesis genes in marine viruses yield proteins during host infection.
737 *Nature* 438:86–89.
- 738 11. Fridman S, Flores-Urbe J, Larom S, Alalouf O, Liran O, Yacoby I, Salama F,
739 Bailleul B, Rappaport F, Ziv T, Sharon I, Cornejo-Castillo FM, Philosof A,
740 Dupont CL, Sánchez P, Acinas SG, Rohwer FL, Lindell D, Béjà O. 2017. A
741 myovirus encoding both photosystem I and II proteins enhances cyclic electron
742 flow in infected *Prochlorococcus* cells. *Nat Microbiol* 2:1350–1357.
- 743 12. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan
744 M, Claverie J-M. 2004. The 1.2-megabase genome sequence of Mimivirus.
745 *Science* 306:1344–1350.
- 746 13. Schulz F, Yutin N, Ivanova NN, Ortega DR, Lee TK, Vierheilig J, Daims H,
747 Horn M, Wagner M, Jensen GJ, Kyrpides NC, Koonin EV, Woyke T. 2017.
748 Giant viruses with an expanded complement of translation system components.
749 *Science* 356:82–85.
- 750 14. Abrahão J, Silva L, Silva LS, Khalil JYB, Rodrigues R, Arantes T, Assis F,
751 Boratto P, Andrade M, Kroon EG, Ribeiro B, Bergier I, Seligmann H, Ghigo E,
752 Colson P, Levasseur A, Kroemer G, Raoult D, Scola BL. 2018. Tailed giant
753 Tupanvirus possesses the most complete translational apparatus of the known
754 virosphere. *Nat Commun* 9:1–12.
- 755 15. Miller ES, Heidelberg JF, Eisen JA, Nelson WC, Durkin AS, Ciecko A,
756 Feldblyum TV, White O, Paulsen IT, Nierman WC, Lee J, Szczypinski B, Fraser
757 CM. 2003. Complete Genome Sequence of the Broad-Host-Range Vibriophage
758 KVP40: Comparative Genomics of a T4-Related Bacteriophage. *J Bacteriol*
759 185:5220–5233.
- 760 16. Yoshikawa G, Askora A, Blanc-Mathieu R, Kawasaki T, Li Y, Nakano M,
761 Ogata H, Yamada T. 2018. *Xanthomonas citri* jumbo phage XacN1 exhibits a
762 wide host range and high complement of tRNA genes. *Sci Rep* 8:4486.
- 763 17. Mizuno CM, Guyomar C, Roux S, Lavigne R, Rodriguez-Valera F, Sullivan
764 MB, Gillet R, Forterre P, Krupovic M. 2019. Numerous cultivated and
765 uncultivated viruses encode ribosomal proteins. *Nature Communications* 10:752.
- 766 18. Schvarcz CR, Steward GF. 2018. A giant virus infecting green algae encodes
767 key fermentation genes. *Virology* 518:423–433.
- 768 19. Piacente F, Gaglianone M, Laugieri ME, Tonetti MG. 2015. The Autonomous
769 Glycosylation of Large DNA Viruses. *International Journal of Molecular*
770 *Sciences* 16:29315–29328.

- 771 20. Schulz F, Roux S, Paez-Espino D, Jungbluth S, Walsh D, Denev VJ, McMahon
772 KD, Konstantinidis KT, Eloë-Fadrosh EA, Kyrpides N, Woyke T. 2020. Giant
773 virus diversity and host interactions through global metagenomics. *Nature* 1–7.
- 774 21. Needham DM, Yoshizawa S, Hosaka T, Poirier C, Choi CJ, Hehenberger E,
775 Irwin NAT, Wilken S, Yung C-M, Bachy C, Kurihara R, Nakajima Y, Kojima
776 K, Kimura-Someya T, Leonard G, Malmstrom RR, Mende DR, Olson DK, Sudo
777 Y, Sudek S, Richards TA, DeLong EF, Keeling PJ, Santoro AE, Shirouzu M,
778 Iwasaki W, Worden AZ. 2019. A distinct lineage of giant viruses brings a
779 rhodopsin photosystem to unicellular marine predators. *PNAS* 116:20574–
780 20583.
- 781 22. Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A, Poulos BT,
782 Solonenko N, Lara E, Poulain J, Pesant S, Kandels-Lewis S, Dimier C, Picheral
783 M, Searson S, Cruaud C, Alberti A, Duarte CM, Gasol JM, Vaqué D, Bork P,
784 Acinas SG, Wincker P, Sullivan MB. 2016. Ecogenomics and potential
785 biogeochemical impacts of globally abundant ocean viruses. *Nature* 537:689–
786 693.
- 787 23. Nishimura Y, Watai H, Honda T, Mihara T, Omae K, Roux S, Blanc-Mathieu R,
788 Yamamoto K, Hingamp P, Sako Y, Sullivan MB, Goto S, Ogata H, Yoshida T.
789 2017. Environmental Viral Genomes Shed New Light on Virus-Host
790 Interactions in the Ocean. *mSphere* 2.
- 791 24. Forterre P. 2013. The virocell concept and environmental microbiology. *ISME J*
792 7:233–236.
- 793 25. Rosenwasser S, Ziv C, Creveld SG van, Vardi A. 2016. Virocell Metabolism:
794 Metabolic Innovations During Host–Virus Interactions in the Ocean. *Trends in*
795 *Microbiology* 24:821–832.
- 796 26. Coy SR, Gann ER, Pound HL, Short SM, Wilhelm SW. 2018. Viruses of
797 Eukaryotic Algae: Diversity, Methods for Detection, and Future Directions.
798 *Viruses* 10:487.
- 799 27. Gallot-Lavallée L, Blanc G, Claverie J-M. 2017. Comparative Genomics of
800 Chrysochromulina Ericina Virus and Other Microalga-Infecting Large DNA
801 Viruses Highlights Their Intricate Evolutionary Relationship with the
802 Established Mimiviridae Family. *Journal of Virology* 91:e00230-17.
- 803 28. Jacobsen A, Bratbak G, Heldal M. 1996. Isolation and characterization of a virus
804 infecting phaeocystis pouchetii (prymnesiophyceae)1. *Journal of Phycology*
805 32:923–927.
- 806 29. Baudoux A-C, Brussaard CPD. 2005. Characterization of different viruses
807 infecting the marine harmful algal bloom species Phaeocystis globosa. *Virology*
808 341:80–90.
- 809 30. Wagstaff BA, Vladu IC, Barclay JE, Schroeder DC, Malin G, Field RA. 2017.
810 Isolation and Characterization of a Double Stranded DNA Megavirus Infecting
811 the Toxin-Producing Haptophyte Prymnesium parvum. *Viruses* 9.

- 812 31. Egge ES, Johannessen TV, Andersen T, Eikrem W, Bittner L, Larsen A, Sandaa
813 R-A, Edvardsen B. 2015. Seasonal diversity and dynamics of haptophytes in the
814 Skagerrak, Norway, explored by high-throughput sequencing. *Mol Ecol*
815 24:3026–3042.
- 816 32. Sandaa RA, Heldal M, Castberg T, Thyrhaug R, Bratbak G. 2001. Isolation and
817 characterization of two viruses with large genome size infecting
818 *Chrysochromulina ericina* (Prymnesiophyceae) and *Pyramimonas orientalis*
819 (Prasinophyceae). *Virology* 290:272–280.
- 820 33. Johannessen TV, Bratbak G, Larsen A, Ogata H, Egge ES, Edvardsen B, Eikrem
821 W, Sandaa R-A. 2015. Characterisation of three novel giant viruses reveals huge
822 diversity among viruses infecting Prymnesiales (Haptophyta). *Virology*
823 476:180–188.
- 824 34. Feldman HA, Wang SS. 1961. Sensitivity of Various Viruses to Chloroform.
825 *Proceedings of the Society for Experimental Biology and Medicine* 106:736–
826 738.
- 827 35. Martínez M, Boere A, Gilg I, van Lent J, Witte H, van Bleijswijk J, Brussaard C.
828 2015. New lipid envelope-containing dsDNA virus isolates infecting
829 *Micromonas pusilla* reveal a separate phylogenetic group. *Aquatic Microbial*
830 *Ecology* 74:17–28.
- 831 36. Mirza SF, Staniewski MA, Short CM, Long AM, Chaban YV, Short SM. 2015.
832 Isolation and characterization of a virus infecting the freshwater algae
833 *Chrysochromulina parva*. *Virology* 486:105–115.
- 834 37. Mackinder LCM, Worthy CA, Biggi G, Hall M, Ryan KP, Varsani A, Harper
835 GM, Wilson WH, Brownlee C, Schroeder DC. 2009. A unicellular algal virus,
836 *Emiliania huxleyi* virus 86, exploits an animal-like infection strategy. *Journal of*
837 *General Virology*, 90:2306–2316.
- 838 38. Huiskonen JT, Kivelä HM, Bamford DH, Butcher SJ. 2004. The PM2 virion has
839 a novel organization with an internal membrane and pentameric receptor binding
840 spikes. *Nat Struct Mol Biol* 11:850–856.
- 841 39. Yan X, Chipman PR, Castberg T, Bratbak G, Baker TS. 2005. The Marine Algal
842 Virus PpV01 Has an Icosahedral Capsid with T=219 Quasisymmetry. *J Virol*
843 79:9236–9243.
- 844 40. Huiskonen JT, Butcher SJ. 2007. Membrane-containing viruses with
845 icosahedrally symmetric capsids. *Curr Opin Struct Biol* 17:229–236.
- 846 41. King AM, Lefkowitz E, Adams MJ, Carstens EB. 2011. *Virus Taxonomy: Ninth*
847 *Report of the International Committee on Taxonomy of Viruses*. Elsevier.
- 848 42. Peralta B, Gil-Carton D, Castaño-Díez D, Bertin A, Boulogne C, Oksanen HM,
849 Bamford DH, Abrescia NGA. 2013. Mechanism of Membranous Tunnelling
850 Nanotube Formation in Viral Genome Delivery. *PLOS Biology* 11:e1001667.

- 851 43. Philippe C, Krupovic M, Jaomanjaka F, Claisse O, Petrel M, le Marrec C. 2018.
852 Bacteriophage GC1, a Novel Tectivirus Infecting *Gluconobacter Cerinus*, an
853 Acetic Acid Bacterium Associated with Wine-Making. *Viruses* 10.
- 854 44. Yan X, Yu Z, Zhang P, Battisti AJ, Chipman PR, Bajaj C, Bergoin M,
855 Rossmann MG, Baker TS. 2009. The Capsid Proteins of a Large, Icosahedral
856 dsDNA Virus. *J Mol Biol* 385:1287–1299.
- 857 45. Klose T, Rossmann MG. 2014. Structure of large dsDNA viruses. *Biol Chem*
858 395:711–719.
- 859 46. Gowing MM. 1993. Large virus-like particles from vacuoles of phaeodarian
860 radiolarians and from other marine samples. *Marine Ecology Progress Series*
861 101:33–43.
- 862 47. Gromov BV, Mamkaeva KA. 1981. A virus infection in the synchronized
863 population of the *Chlorococcum minutum* zoospores. *Algological Studies/Archiv*
864 *für Hydrobiologie, Supplement Volumes* 252–259.
- 865 48. Bratbak G, Jacobsen A, Heldal M, Nagasaki K, Thingstad F. 1998. Virus
866 production in *Phaeocystis pouchetii* and its relation to host cell growth and
867 nutrition. *Aquatic Microbial Ecology* 16:1–9.
- 868 49. Zimmerman AE, Bachy C, Ma X, Roux S, Jang HB, Sullivan MB, Waldbauer
869 JR, Worden AZ. 2019. Closely related viruses of the marine picoeukaryotic alga
870 *Ostreococcus lucimarinus* exhibit different ecological strategies. *Environ*
871 *Microbiol* 21:2148–2170.
- 872 50. Thomsen HA, Buck KR, Chavez FP. 1994. Haptophytes as components of
873 marine phytoplankton. *DTU Research Database* 187–208.
- 874 51. Berngruber TW, Froissart R, Choisy M, Gandon S. 2013. Evolution of Virulence
875 in Emerging Epidemics. *PLOS Pathogens* 9:e1003209.
- 876 52. Day T, Proulx SR. 2004. A General Theory for the Evolutionary Dynamics of
877 Virulence. *The American Naturalist* 163:E40–E63.
- 878 53. King AA, Shrestha S, Harvill ET, Bjørnstad ON. 2009. Evolution of Acute
879 Infections and the Invasion-Persistence Trade-Off. *The American Naturalist*
880 173:446–455.
- 881 54. Brussaard CPD, Bratbak G, Baudoux A-C, Ruardij P. 2007. *Phaeocystis* and its
882 interaction with viruses. *Biogeochemistry* 83:201–215.
- 883 55. Leggett HC, Buckling A, Long GH, Boots M. 2013. Generalism and the
884 evolution of parasite virulence. *Trends Ecol Evol (Amst)* 28:592–596.
- 885 56. Woolhouse MEJ, Taylor LH, Haydon DT. 2001. Population Biology of
886 Multihost Pathogens. *Science* 292:1109–1112.

- 887 57. Johannessen TV, Larsen A, Bratbak G, Pagarete A, Edvardsen B, Egge ED,
888 Sandaa R-A. 2017. Seasonal Dynamics of Haptophytes and dsDNA Algal
889 Viruses Suggest Complex Virus-Host Relationship. *Viruses* 9.
- 890 58. Gran-Stadniczeńko S, Krabberød AK, Sandaa R-A, Yau S, Egge E, Edvardsen
891 B. 2019. Seasonal Dynamics of Algae-Infecting Viruses and Their Inferred
892 Interactions with Protists. *Viruses* 11:1043.
- 893 59. Fischer MG, Allen MJ, Wilson WH, Suttle CA. 2010. Giant virus with a
894 remarkable complement of genes infects marine zooplankton. *Proc Natl Acad
895 Sci USA* 107:19508–19513.
- 896 60. Sandaa R-A, Dahle H, Brussaard CPD, Ogata H, Blanc-Mathieu R. Algal
897 viruses belonging to a subgroup within the Mimiviridae family *Encyclopedia of
898 Virology*, 4th ed. Bamford, D., M. Zuckerman, Elsevier, Academic.
- 899 61. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a
900 tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids
901 Res* 28:33–36.
- 902 62. Blanc-Mathieu R, Ogata H. 2016. DNA repair genes in the Megavirales
903 pangenome. *Current Opinion in Microbiology* 31:94–100.
- 904 63. Moniruzzaman M, Martinez-Gutierrez CA, Weinheimer AR, Aylward FO. 2020.
905 Dynamic genome evolution and complex virocell metabolism of globally-
906 distributed giant viruses. 1. *Nature Communications* 11:1710.
- 907 64. Ono A. 2010. Viruses and Lipids. *Viruses* 2:1236–1238.
- 908 65. Heaton NS, Randall G. 2011. Multifaceted roles for lipids in viral infection.
909 *Trends Microbiol* 19:368–375.
- 910 66. Lange PT, Lagunoff M, Tarakanova VL. 2019. Chewing the Fat: The Conserved
911 Ability of DNA Viruses to Hijack Cellular Lipid Metabolism. *Viruses* 11:119.
- 912 67. Diamond DL, Syder AJ, Jacobs JM, Sorensen CM, Walters K-A, Proll SC,
913 McDermott JE, Gritsenko MA, Zhang Q, Zhao R, Metz TO, Camp DG, Waters
914 KM, Smith RD, Rice CM, Katze MG. 2010. Temporal Proteome and Lipidome
915 Profiles Reveal Hepatitis C Virus-Associated Reprogramming of Hepatocellular
916 Metabolism and Bioenergetics. *PLoS Pathog* 6.
- 917 68. Heaton NS, Randall G. 2010. Dengue virus induced autophagy regulates lipid
918 metabolism. *Cell Host Microbe* 8:422–432.
- 919 69. Rosenwasser S, Mausz MA, Schatz D, Sheyn U, Malitsky S, Aharoni A,
920 Weinstock E, Tzfadia O, Ben-Dor S, Feldmesser E, Pohnert G, Vardi A. 2014.
921 Rewiring Host Lipid Metabolism by Large Viruses Determines the Fate of
922 *Emiliana huxleyi*, a Bloom-Forming Alga in the Ocean[C][W][OPEN]. *Plant
923 Cell* 26:2689–2707.
- 924 70. Malitsky S, Ziv C, Rosenwasser S, Zheng S, Schatz D, Porat Z, Ben-Dor S,
925 Aharoni A, Vardi A. 2016. Viral infection of the marine alga *Emiliana huxleyi*

- 926 triggers lipidome remodeling and induces the production of highly saturated
927 triacylglycerol. *New Phytologist* 210:88–96.
- 928 71. Wilson WH, Schroeder DC, Allen MJ, Holden MTG, Parkhill J, Barrell BG,
929 Churcher C, Hamlin N, Mungall K, Norbertczak H, Quail MA, Price C,
930 Rabinowitsch E, Walker D, Craigon M, Roy D, Ghazal P. 2005. Complete
931 Genome Sequence and Lytic Phase Transcription Profile of a Coccolithovirus.
932 *Science* 309:1090–1092.
- 933 72. Vardi A, Van Mooy BAS, Fredricks HF, Pependorf KJ, Ossolinski JE,
934 Haramaty L, Bidle KD. 2009. Viral glycosphingolipids induce lytic infection
935 and cell death in marine phytoplankton. *Science* 326:861–865.
- 936 73. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y.
937 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme
938 annotation. *Nucleic Acids Res* 46:W95–W101.
- 939 74. Luther KB, Hülsmeier AJ, Schegg B, Deuber SA, Raoult D, Hennet T. 2011.
940 Mimivirus collagen is modified by bifunctional lysyl hydroxylase and
941 glycosyltransferase enzyme. *J Biol Chem* 286:43701–43709.
- 942 75. Rommel AJ, Hülsmeier AJ, Jurt S, Hennet T. 2016. Giant mimivirus R707
943 encodes a glycogenin paralogue polymerizing glucose through α - and β -
944 glycosidic linkages. *Biochem J* 473:3451–3462.
- 945 76. Parakkottil Chothi M, Duncan GA, Armirotti A, Abergel C, Gurnon JR, Van
946 Etten JL, Bernardi C, Damonte G, Tonetti M. 2010. Identification of an l-
947 Rhamnose Synthetic Pathway in Two Nucleocytoplasmic Large DNA Viruses. *J*
948 *Virol* 84:8829–8838.
- 949 77. Marie D, Brussaard CPD, Thyrraug R, Bratbak G, Vaultot D. 1999. Enumeration
950 of Marine Viruses in Culture and Natural Samples by Flow Cytometry. *Appl*
951 *Environ Microbiol* 65:45–52.
- 952 78. Brussaard CPD. 2004. Optimization of procedures for counting viruses by flow
953 cytometry. *Appl Environ Microbiol* 70:1506–1513.
- 954 79. Jarvis B, Wilrich C, Wilrich P-T. 2010. Reconsideration of the derivation of
955 Most Probable Numbers, their standard deviations, confidence bounds and rarity
956 values. *J Appl Microbiol* 109:1660–1667.
- 957 80. Kremer JR, Mastrorade DN, McIntosh JR. 1996. Computer visualization of
958 three-dimensional image data using IMOD. *J Struct Biol* 116:71–76.
- 959 81. Lawrence JE, Steward GF. 2010. Purification of viruses by centrifugation, p.
960 166–181. *In* Wilhelm, S, Weinbauer, M, Suttle, C (eds.), *Manual of Aquatic*
961 *Viral Ecology*. American Society of Limnology and Oceanography.
- 962 82. Sandaa R-A, E. Storesund J, Olesin E, Lund Paulsen M, Larsen A, Bratbak G,
963 Ray JL. 2018. Seasonality Drives Microbial Community Structure, Shaping both
964 Eukaryotic and Prokaryotic Host–Viral Relationships in an Arctic Marine
965 Ecosystem. *Viruses* 10.

- 966 83. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bembien LA, Berka J,
967 Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV,
968 Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI,
969 Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz
970 SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna
971 MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT,
972 Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt
973 KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg
974 JM. 2005. Genome sequencing in microfabricated high-density picolitre
975 reactors. *Nature* 437:376–380.
- 976 84. Gordon D, Green P. 2013. Consed: a graphical editor for next-generation
977 sequencing. *Bioinformatics* 29:2936–2937.
- 978 85. Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled
979 de novo using only nanopore sequencing data. *Nat Methods* 12:733–735.
- 980 86. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo
981 CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: An Integrated Tool
982 for Comprehensive Microbial Variant Detection and Genome Assembly
983 Improvement. *PLOS ONE* 9:e112963.
- 984 87. Yutin N, Wolf YI, Raoult D, Koonin EV. 2009. Eukaryotic large nucleo-
985 cytoplasmic DNA viruses: Clusters of orthologous genes and reconstruction of
986 viral genome evolution. *Virology Journal* 6:223.
- 987 88. Yutin N, Wolf YI, Koonin EV. 2014. Origin of giant viruses from smaller DNA
988 viruses not from a fourth domain of cellular life. *Virology* 466–467:38–52.
- 989 89. Eddy SR. 2011. Accelerated Profile HMM Searches. *PLOS Computational
990 Biology* 7:e1002195.
- 991 90. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software
992 version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–
993 780.
- 994 91. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for
995 automated alignment trimming in large-scale phylogenetic analyses.
996 *Bioinformatics* 25:1972–1973.
- 997 92. Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software
998 package for phylogenetic reconstruction and molecular dating. *Bioinformatics*
999 25:2286–2288.
- 1000 93. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and
1001 post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
- 1002 94. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the
1003 display and annotation of phylogenetic and other trees. *Nucleic Acids Res*
1004 44:W242–W245.

- 1005 95. Price MN, Dehal PS, Arkin AP. 2009. FastTree: Computing Large Minimum
1006 Evolution Trees with Profiles instead of a Distance Matrix. *Mol Biol Evol*
1007 26:1641–1650.
- 1008 96. Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes MPI:
1009 phylogenetic reconstruction with infinite mixtures of profiles in a parallel
1010 environment. *Syst Biol* 62:611–615.
- 1011 97. Lartillot N, Philippe H. 2004. A Bayesian Mixture Model for Across-Site
1012 Heterogeneities in the Amino-Acid Replacement Process. *Mol Biol Evol*
1013 21:1095–1109.
- 1014 98. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training
1015 method for prediction of gene starts in microbial genomes. Implications for
1016 finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29:2607–2618.
- 1017 99. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and
1018 context for analysis of transfer RNA genes. *Nucleic Acids Res* 44:W54–57.
- 1019 100. Mihara T, Nishimura Y, Shimizu Y, Nishiyama H, Yoshikawa G, Uehara H,
1020 Hingamp P, Goto S, Ogata H. 2016. Linking Virus Genomes with Host
1021 Taxonomy. *Viruses* 8:66.
- 1022 101. Pruitt KD, Tatusova T, Maglott DR. 2007. NCBI reference sequences (RefSeq):
1023 a curated non-redundant sequence database of genomes, transcripts and proteins.
1024 *Nucleic Acids Res* 35:D61–D65.
- 1025 102. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH. 2015. UniRef clusters: a
1026 comprehensive and scalable alternative for improving sequence similarity
1027 searches. *Bioinformatics* 31:926–932.
- 1028 103. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY,
1029 Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song
1030 JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015.
1031 CDD: NCBI’s conserved domain database. *Nucleic Acids Res* 43:D222–D226.
- 1032 104. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a
1033 reference resource for gene and protein annotation. *Nucleic Acids Res* 44:D457–
1034 D462.
- 1035 105. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G,
1036 Djahanschiri B, Zeller G, Mende DR, Alberti A, Cornejo-Castillo FM, Costea
1037 PI, Cruaud C, d’Ovidio F, Engelen S, Ferrera I, Gasol JM, Guidi L, Hildebrand
1038 F, Kokoszka F, Lepoivre C, Lima-Mendez G, Poulain J, Poulos BT, Royo-
1039 Llonch M, Sarmiento H, Vieira-Silva S, Dimier C, Picheral M, Searson S,
1040 Kandels-Lewis S, Tara Oceans coordinators, Bowler C, de Vargas C, Gorsky G,
1041 Grimsley N, Hingamp P, Iudicone D, Jaillon O, Not F, Ogata H, Pesant S,
1042 Speich S, Stemmann L, Sullivan MB, Weissenbach J, Wincker P, Karsenti E,
1043 Raes J, Acinas SG, Bork P. 2015. Ocean plankton. Structure and function of the
1044 global ocean microbiome. *Science* 348:1261359.

- 1045 106. Villar E, Vannier T, Vernet C, Lescot M, Cuenca M, Alexandre A, Bachelerie
1046 P, Rosnet T, Pelletier E, Sunagawa S, Hingamp P. 2018. The Ocean Gene Atlas:
1047 exploring the biogeography of plankton genes online. *Nucleic Acids Res*
1048 46:W289–W295.
- 1049 107. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M,
1050 Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L,
1051 Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in
1052 2019. *Nucleic Acids Res* 47:D427–D432.
- 1053 108. Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* 14:755–763.
- 1054 109. Carradec Q, Pelletier E, Silva CD, Alberti A, Seeleuthner Y, Blanc-Mathieu R,
1055 Lima-Mendez G, Rocha F, Tirichine L, Labadie K, Kirilovsky A, Bertrand A,
1056 Engelen S, Madoui M-A, Méheust R, Poulain J, Romac S, Richter DJ,
1057 Yoshikawa G, Dimier C, Kandels-Lewis S, Picheral M, Searson S, Jaillon O,
1058 Aury J-M, Karsenti E, Sullivan MB, Sunagawa S, Bork P, Not F, Hingamp P,
1059 Raes J, Guidi L, Ogata H, Vargas C de, Iudicone D, Bowler C, Wincker P. 2018.
1060 A global ocean atlas of eukaryotic genes. *Nature Communications* 9:373.
- 1061 110. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA,
1062 Armbrust EV, Archibald JM, Bharti AK, Bell CJ, Beszteri B, Bidle KD,
1063 Cameron CT, Campbell L, Caron DA, Cattolico RA, Collier JL, Coyne K, Davy
1064 SK, Deschamps P, Dyhrman ST, Edvardsen B, Gates RD, Gobler CJ,
1065 Greenwood SJ, Guida SM, Jacobi JL, Jakobsen KS, James ER, Jenkins B, John
1066 U, Johnson MD, Juhl AR, Kamp A, Katz LA, Kiene R, Kudryavtsev A, Leander
1067 BS, Lin S, Lovejoy C, Lynn D, Marchetti A, McManus G, Nedelcu AM,
1068 Menden-Deuer S, Miceli C, Mock T, Montresor M, Moran MA, Murray S,
1069 Nadathur G, Nagai S, Ngam PB, Palenik B, Pawlowski J, Petroni G, Piganeau
1070 G, Posewitz MC, Rengefors K, Romano G, Rumpho ME, Ryneerson T,
1071 Schilling KB, Schroeder DC, Simpson AGB, Slamovits CH, Smith DR, Smith
1072 GJ, Smith SR, Sosik HM, Stief P, Theriot E, Twary SN, Umale PE, Vaultot D,
1073 Wawrik B, Wheeler GL, Wilson WH, Xu Y, Zingone A, Worden AZ. 2014. The
1074 Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP):
1075 Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through
1076 Transcriptome Sequencing. *PLOS Biology* 12:e1001889.
- 1077 111. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using
1078 DIAMOND. *Nat Methods* 12:59–60.
- 1079

1080 **Figure legends**

1081 **FIG 1** PkV RF01 morphology and reduced viral infectivity under chloroform
1082 treatment. (A) Screen shot of a cryo-electron tomogram of a PkV RF01 virion. (B)
1083 Composite image of 61 cryo-electron tomograms (-60 to 60° , imaged every 2°).
1084 Purple, capsid; green, inner membrane consisting of multiple irregular, convoluted
1085 membranes; blue, internal rod-shaped core filled with dense material. The full set of
1086 records is available on GitHub (see [Data availability](#) section). Scale bar, 100 nm. (C)
1087 Reduction of PkV RF01 infectivity with chloroform. Experiments were set up in
1088 triplicate, and host cells were counted by flow cytometry. Chloroform-treated PkV
1089 RF01 was added to exponentially growing He UiO028 cells in a 1:10 volume ratio.
1090 Controls were He UiO028 cells incubated with chloroform-treated medium (Control
1091 A), untreated PkV RF01 (Control B), and untreated medium (Control C). SDs are
1092 indicated with error bars.

1093

1094 **FIG 2** Structure and gene taxonomic composition of the PkV RF01 genome sequence.
1095 (A) Rhizome and genomic features of the PkV RF01 genome. As illustrated by the
1096 rhizome (inner part of the figure), ORFans comprise the largest set of PkV RF01
1097 genes, and a substantial portion (15%) have their best BLAST hits (in UniRef90)
1098 against “*Mimiviridae*.” Colors indicate taxonomic origin. Intergenic regions are white.
1099 Percentage hits per taxonomic group higher than 5% of total genes are indicated. In
1100 the outermost ring, rectangles indicate the positions of glycosyltransferases (white),
1101 lipid-related enzymes (black), and succinate dehydrogenase genes (red), and the
1102 numbers correspond to *Mimiviridae* key enzymes (1 and 3: DNA-directed RNA
1103 polymerase II subunits 1 and 2, respectively; 2: DNA mismatch repair protein MutS7;
1104 4: Packaging ATPase; 5: VLTF3, 6: Major capsid protein; 7: Eukaryotic translation

1105 initiation factor 4E; 8: Asparagine synthase; 9: DNA polymerase family B). The ring
1106 adjacent to the outermost ring shows GC skew over a 10-KB window. (B) Taxonomic
1107 breakdown of 180 genes with best hits to virus genes. Mega, Megavirinae; AaV,
1108 Aureococcus anophagefferens virus; TetV, Tetraselmis virus 1; PoV, Pyramimonas
1109 orientalis virus.

1110

1111 **FIG 3** Phylogenetic evidence for PkV RF01 as a distant relative of
1112 “Mesomimivirinae.” (A) Bayesian phylogenetic tree of NCLDVs reconstructed from
1113 a concatenated alignment of five core nucleocytoplasmic virus orthologous genes.
1114 Values at branches are posterior probabilities support. The tree was rooted using
1115 *Poxviridae* as outgroup. The scale bar indicates substitutions per site. (B) Maximum
1116 likelihood phylogenetic tree of cellular and NCLDV DNA-directed RNA polymerase
1117 subunit beta (RPB2). Values at branches are Shimodaira-Hasegawa-like local support.
1118

1119 **FIG 4** COG functional distribution of 339 proteins encoded by PkV RF01.

1120

1121 **FIG 5** Comparative COG functional distribution among Mimiviridae members. COG
1122 sequences were automatically searched against the proteomes of each virus using
1123 BLASTP with an E-value of 1×10^{-5} as the significant similarity threshold.

1124

1125 **FIG 6** Bayesian phylogenetic trees of two viral amino-acyl tRNA synthetases and
1126 their cellular homologs. (A) Isoleucine tRNA synthetases. (B) Aspartyl tRNA
1127 synthetases. Branches supported by posterior probability (PP) values >70% are
1128 indicated by circles whose diameters are proportional to the PP value.

1129

1130 **FIG 7** Genes in PkV RF01 predicted to encode enzymes of oxidative phosphorylation
1131 and β -oxidation pathways. (A) Gene organization in the succinate dehydrogenase-
1132 containing region. (B) Schematic representation of the canonical enzymatic complex
1133 II in the mitochondrial membrane. (C) Location of succinate dehydrogenase in the
1134 TCA cycle and electron transport chain as known in plants and a schematic
1135 reconstruction of the PkV RF01-encoded β -oxidation metabolic pathway.

1136

1137 **FIG 8** The viral SDHA gene is transcribed during infection. Gels of PCR and RT-
1138 PCR in combination with a TURBO DNA-free kit. Samples were taken 24, 72, and 96
1139 h after infection. (A) PCR with *vSDHA*-specific primers was used to check for the
1140 presence of genomic DNA after RNA isolation treated with 1x and 2x DNase, in the
1141 upper and lower panels respectively. P, positive control (PKV RF01 genomic DNA);
1142 N, negative control (*sdH₂O*). (B) RT-PCR of RNA samples using *vSDHA*-specific
1143 primers. M, DNA marker (MassRuler DNA Ladder Mix, Thermo Fisher, 80 to 10,000
1144 bp).

1145

1146 **FIG 9** PCR optimization and confirmation of the SDHA gene in the PkV RF01
1147 genome. (A–B) Results of PCR with SDHA primers using genomic PkV RF01 DNA
1148 (A) and genomic He UiO028 DNA (B) as templates. Lanes 1 and 9, DNA ladder; 2–
1149 7, optimization of the PCR annealing temperature from 55°C (2) to 60°C (7); 8,
1150 negative control (*sdH₂O*).

1151

1152 **FIG 10** PCR and RT-PCR optimization using an internal control gene (*mcp*).
1153 PCR and RT-PCR were carried out after removal of genomic DNA using a TURBO
1154 DNA-free kit. Samples were taken 24, 72, and 96 h after infection. Two different

1155 protocols, both provided in the TURBO DNA-free kit manual, were used to optimize
1156 the reactions. (A) PCR check for the presence of genomic DNA after RNA isolation
1157 treated with 1x and 2x DNase, in the upper and lower panels respectively. P, positive
1158 control (Pkv RF01 genomic DNA); N, negative control (sdH20). (B) Result of RT-
1159 PCR of samples harvested 24, 72 and 96 h post infection. M, DNA marker
1160 (MassRuler DNA Ladder Mix, Thermo Fisher, 80 to 10,000 bp).

1161

1162 **FIG 11** Origin of Pkv RF01 SDHA and SDHB and their most similar homologs in
1163 *Tara* Oceans metagenomes. (A) Taxonomy of genes predicted in *Tara* Oceans
1164 metagenome assembled-genome fragments encoding the 50 SDHAs and SDHBs most
1165 similar to Pkv RF01 genes (for genome fragments having at least five predicted
1166 genes). (B and C) Phylogenetic trees of viral and cellular SDHAs (B) and SDHBs (C).
1167 Clades in green contain Pkv RF01 SDHA or SDHB and their 50 most similar hits
1168 identified in *Tara* Oceans metagenomes (predicted to be *Mimiviridae* homologs from
1169 A). Red, eukaryotic phyla; black, unclassified eukaryotes. Trees are rooted with
1170 Proteobacteria and Firmicutes homologs (not shown). Circles indicate branches with
1171 posterior probability support $\geq 50\%$.

1172

1173 **FIG 12** Comparative distribution of glycosyltransferase domains among viruses.

1174

1175 **Tables**

1176 **TABLE 1** Infection parameters of *Prymnesium kappa* viruses RF01 and RF02 and
 1177 *Haptolina ericina* virus RF02.

Viral species and hosts	Infectious progeny/mL (MPN)	Host cells/mL (FCM) ^a	Total VLP/mL (FCM)	Burst size (VLP) ^b	Infectivity (%) ^c	Infectious particles in a burst ^d
PkV RF01 (He UiO028)	2.9x10 ⁶ (± 0.2)	4.9x10 ⁵	1.8x10 ⁸ (±0.9)	363	2	6
PkV RF02 (Pk RCC3423)	2.2x10 ⁸ (± 0.2)	4.6x10 ⁵	5.0x10 ⁸ (±0.1)	1093	44	483
HeV RF02 (He UiO028)	5.8x10 ⁷ (±0.2)	4.9 x 10 ⁵	4.4 x10 ⁸ (±0.0)	907	13	119

1178 VLP, virus-like particle; MPN, most probable number; FCM, flow cytometry.

1179 ^aMeasurement performed in duplicates

1180 ^bThe number of viral particles released from each host cell, estimated from the total number of host cells pre-infection and the total number of VLPs produced during the infection cycle.

1181 ^cEstimated as the percentage of infectious progeny of all VLPs produced during the infection cycle.

1182 ^dNumber of infectious particles released per host cell.

1184

1185 **TABLE 2** Gene related to lipid metabolism.

ORF	Annotation	KEGG orthology	Pathway
30	2,4-dienoyl-CoA reductase, mitochondrial [EC:1.3.1.34]	K13236	Beta oxidation
33	Putative CoA-transferase	NS	Beta oxidation
121	glycerophosphoryl diester phosphodiesterase	K01126	Glycerophospholipids metabolisms
138	Fatty acid synthase (FASN)	K00665	Fatty acid biosynthesis
142	Long-chain-fatty-acid--CoA ligase ACSBG [EC:6.2.1.3]	K15013	Fatty acid degradation /biosynthesis / Beta Oxidation
175	Acetyl-CoA carboxylase / biotin carboxylase 1 [EC:6.4.1.2 6.3.4.14 2.1.3.15]	K11262	Fatty acid biosynthesis
236	Glutaryl-CoA dehydrogenase [EC:1.3.8.6]	K00252	Fatty acid degradation
293	Lysophospholipase like	NS	NS
357	Lysophospholipase like	NS	NS
386	Triacylglycerol lipase [EC:3.1.1.3]	K01046	Glycerolipid metabolism
481	Lipase like	NS	NS
635	Lipase-like	NS	NS
653	Lipase-like	NS	NS
690	Lipase-like	NS	NS
774	Lysophospholipid Acyltransferases [EC:2.3.1.22]	K14457	Glycerolipid metabolism
694	Lipase esterase (Carbohydrate esterase CE10)	NS	NS
695	Lipase esterase (Carbohydrate esterase CE10)	NS	NS
886	Stearoyl-CoA desaturase (Delta-9 desaturase) [EC:1.14.19.1]	K00507	Biosynthesis of unsaturated fatty acids
902	Fatty acid synthase (FASN)	K00665	Fatty acid biosynthesis

904	Long-chain-fatty-acid--CoA ligase ACSBG [EC:6.2.1.3]	K15013	Fatty acid degradation /biosynthesis / Beta Oxidation
1016	Cyclopropane-fatty-acyl-phospholipid synthase [EC:2.1.1.79]	k00574	NS
1046	Acyl-CoA dehydrogenase	K06445	Fatty acid degradation / Beta oxidation

1186

1187

1188 **TABLE 3** Forward and reverse PCR primers for amplification of vSDHA and MCP

1189

Primer name	Sequence (5' - 3')	PCR product (bp)
vSDHA-F1	ATGTGCCGAGAAGCTCCTAA	154
vSDHA-R1	CTGCACAGGCTGTTTCGATAA	
PkV-RF01-MCP-F	GATGAACCTTGCCCACAACCT	256
PkV-RF01-MCP-F	GTGCATGGTACGTTTTTCGTG	

1190

Figure 1

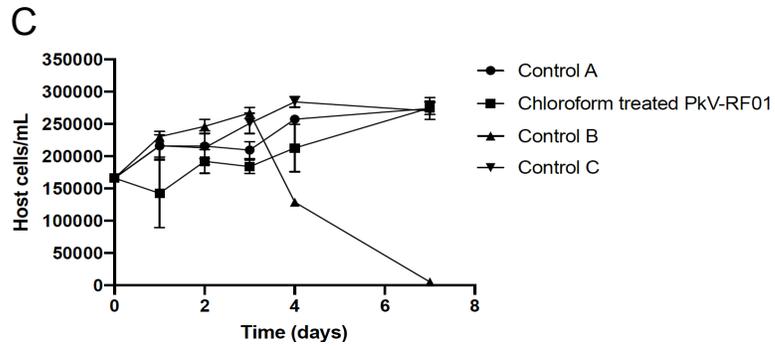
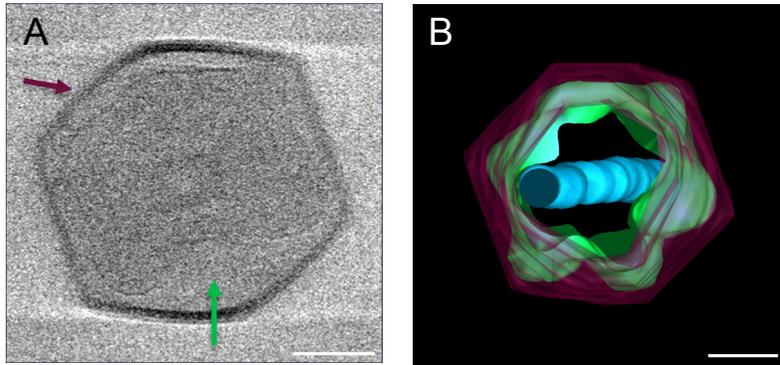


Figure 2

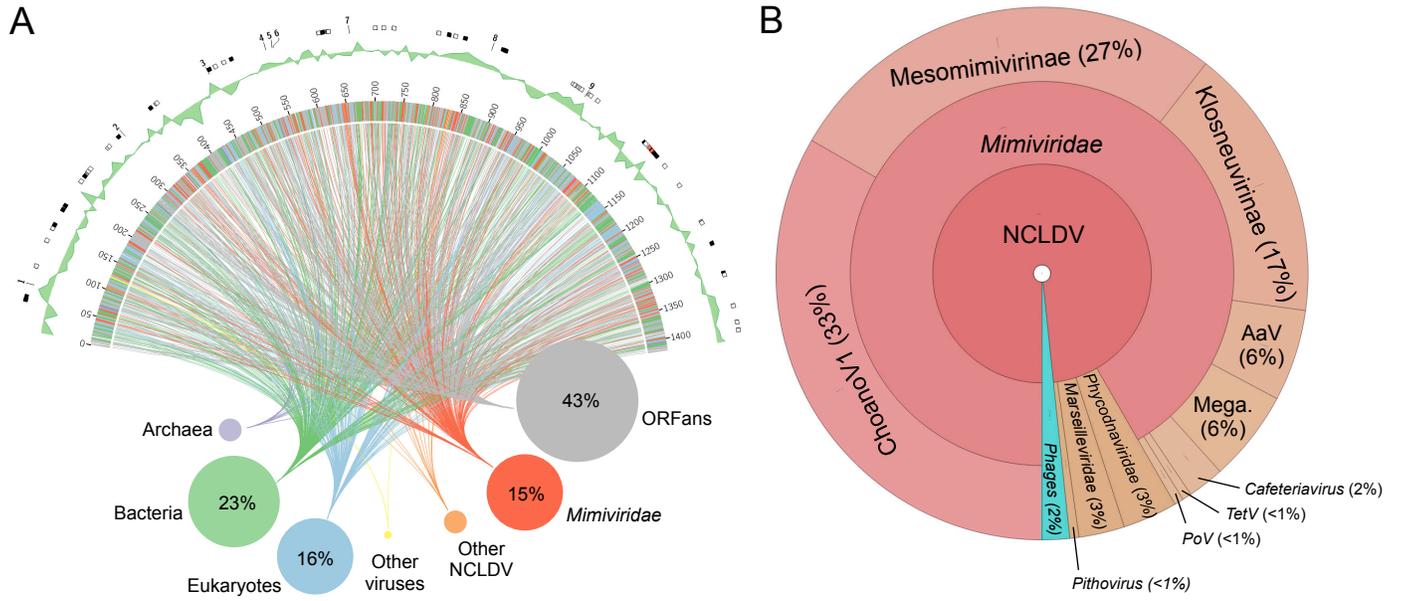


Figure 3

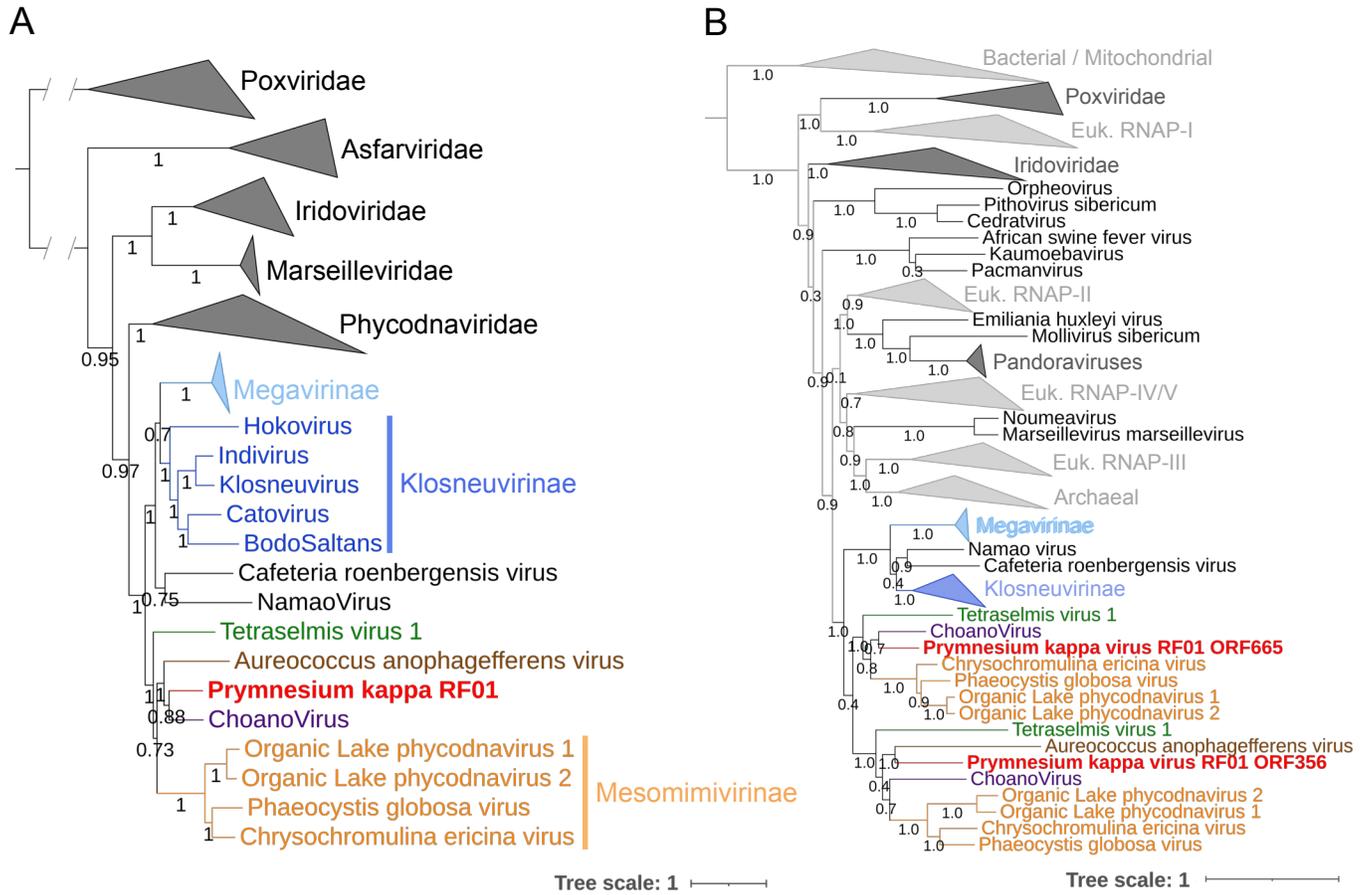


Figure 4

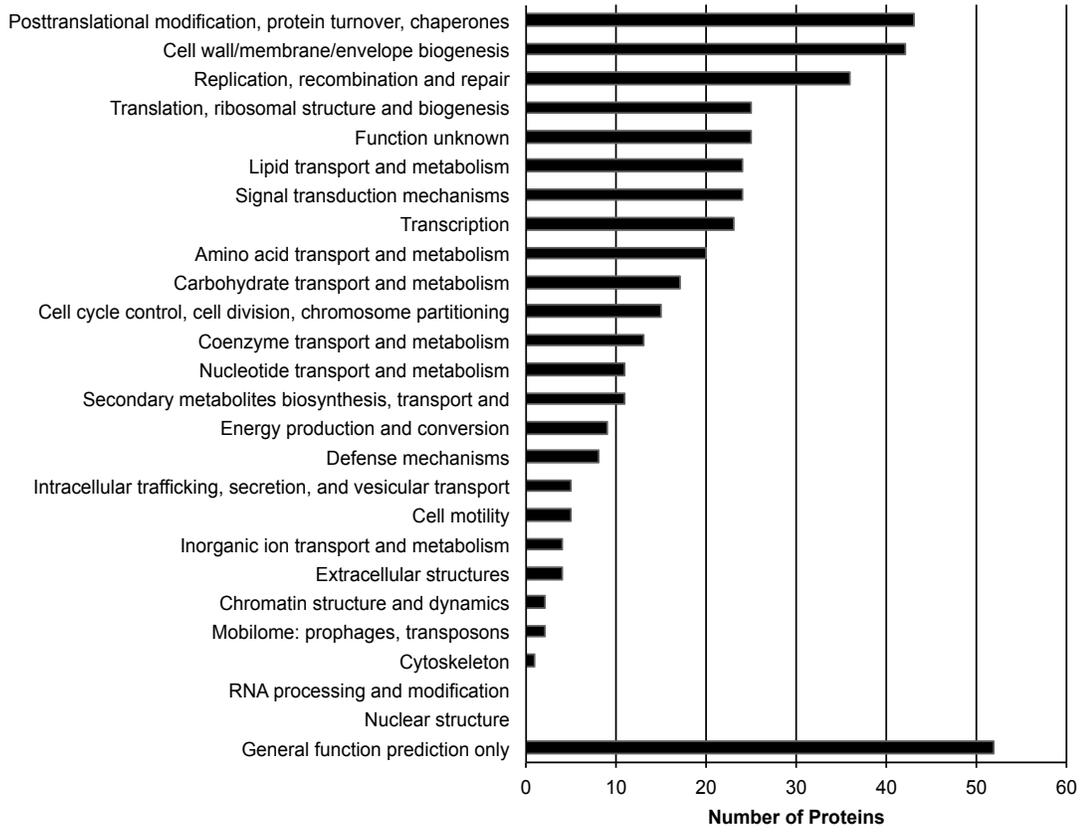


Figure 5

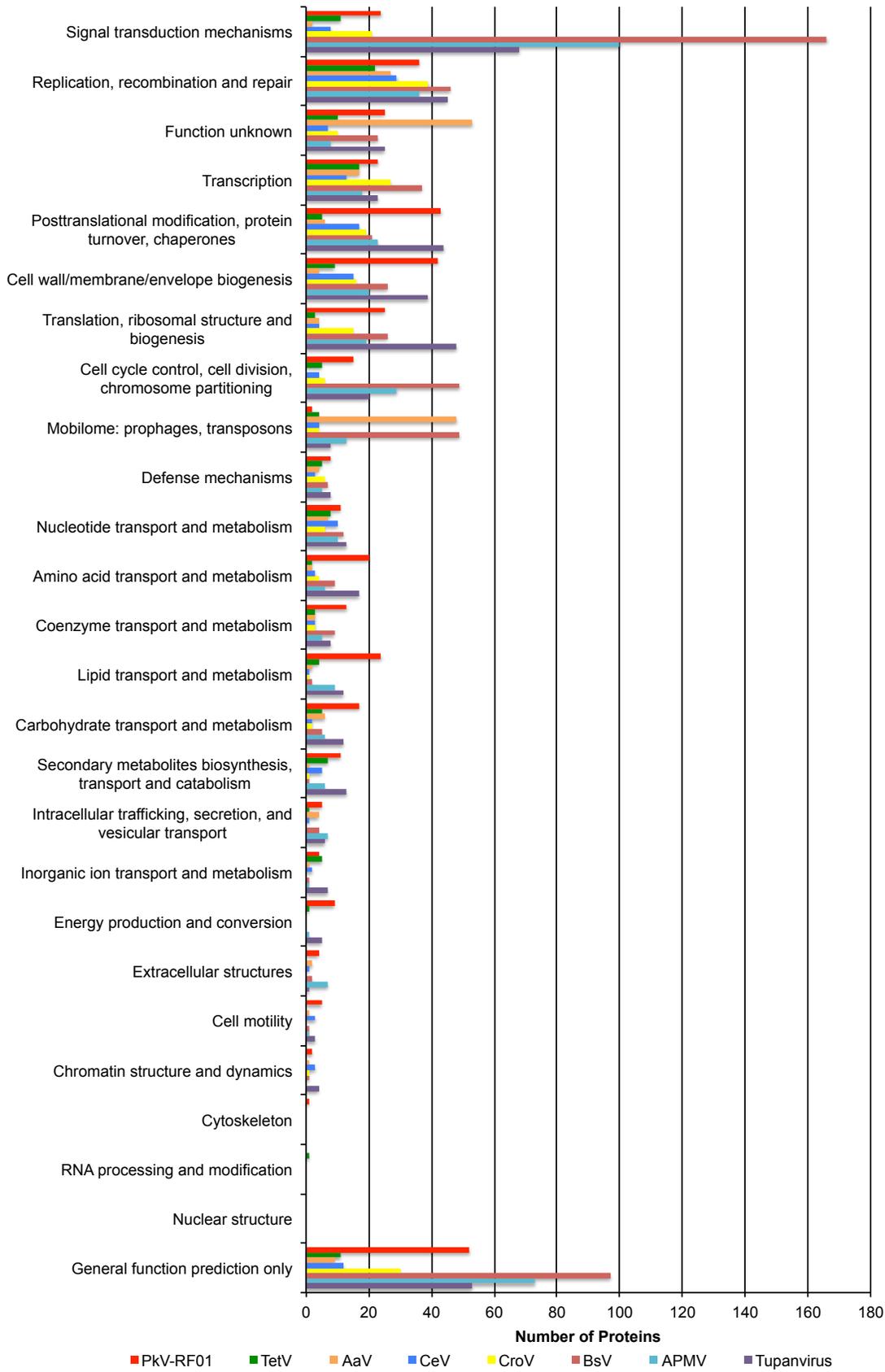


Figure 7

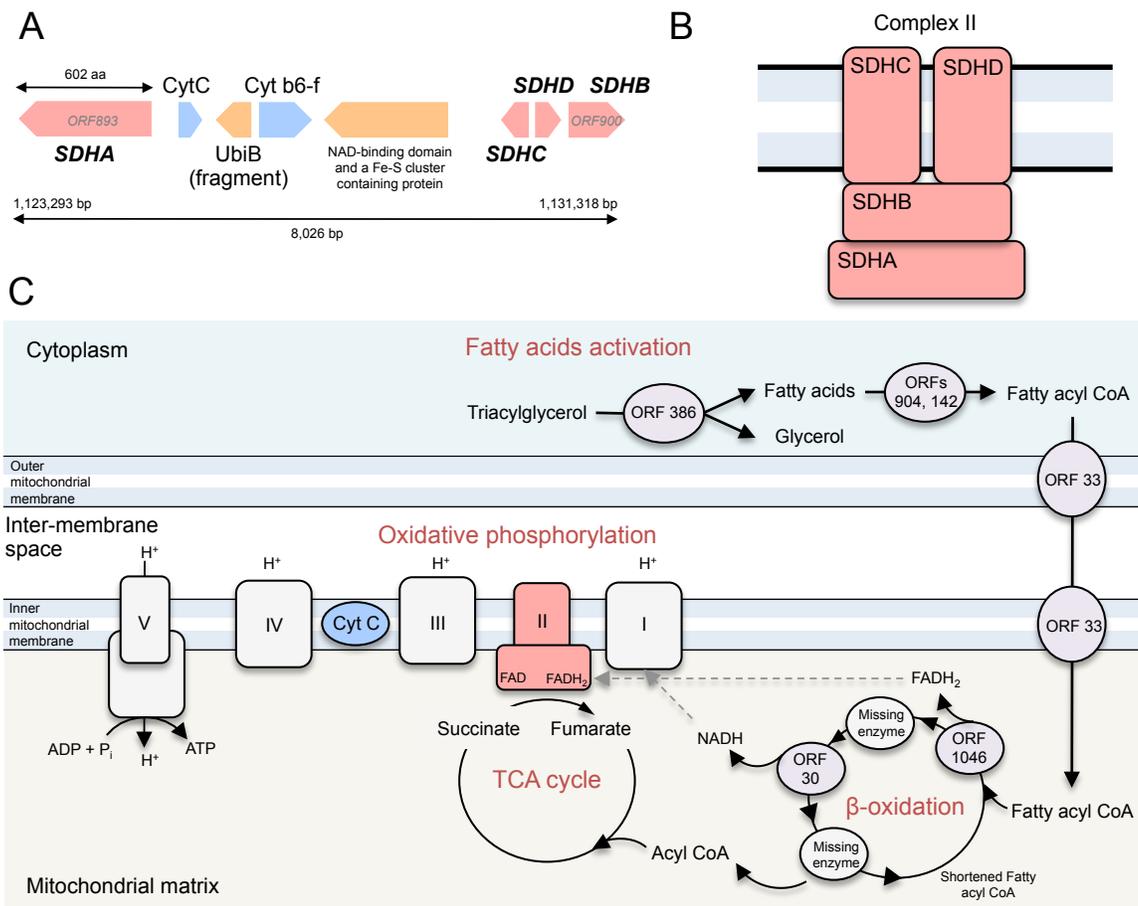


Figure 8

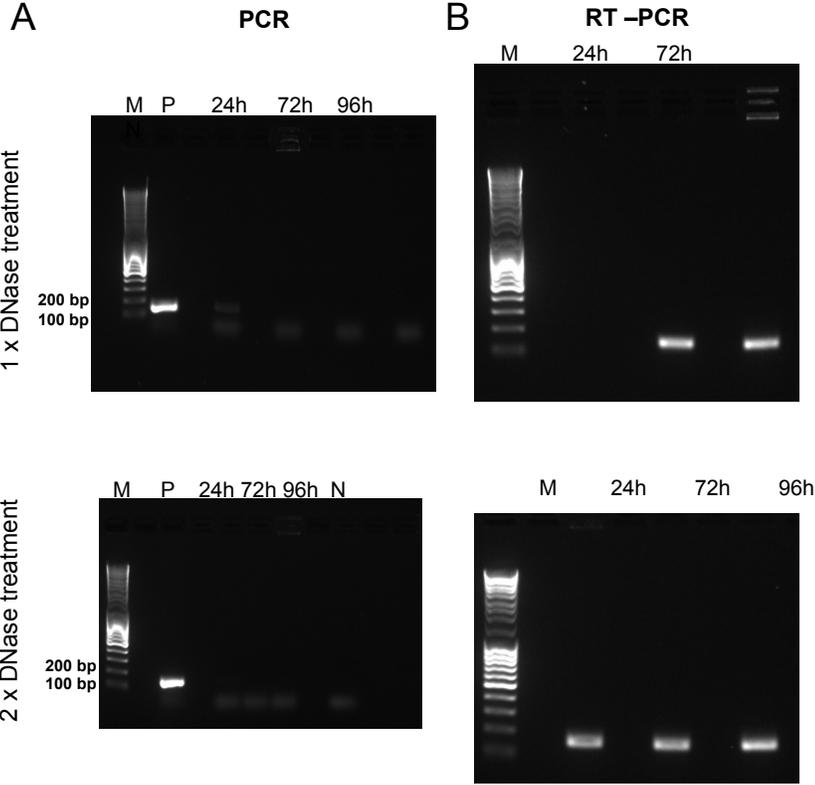


Figure 9

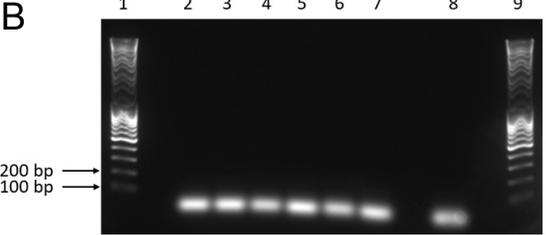
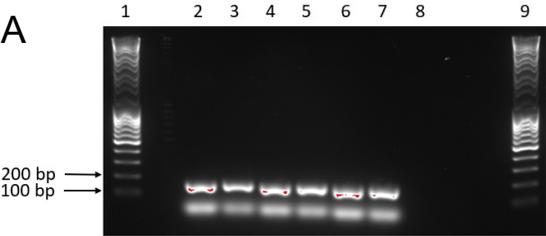


Figure 10

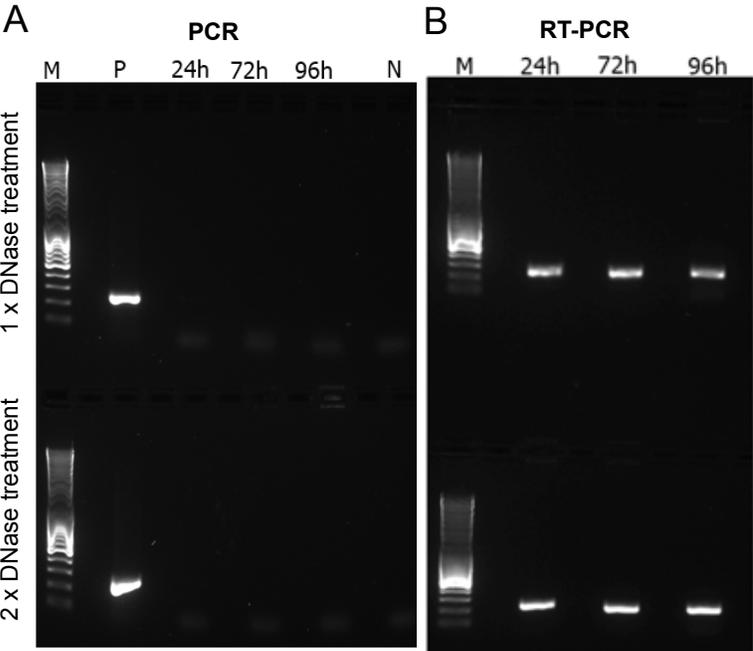


Figure 11

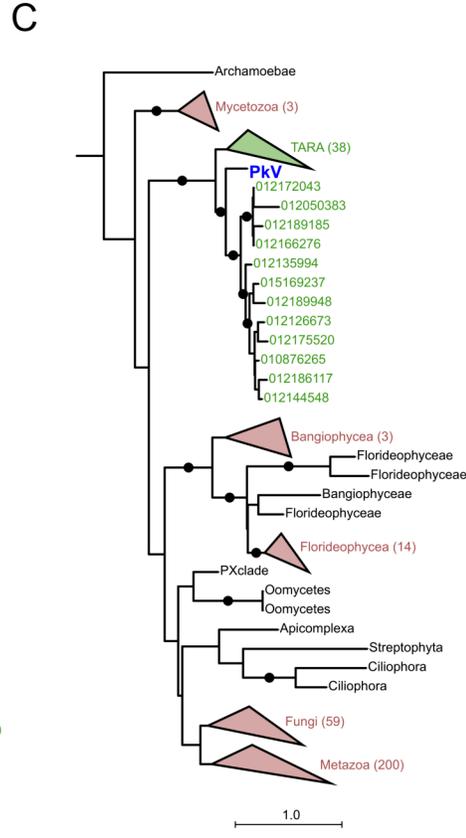
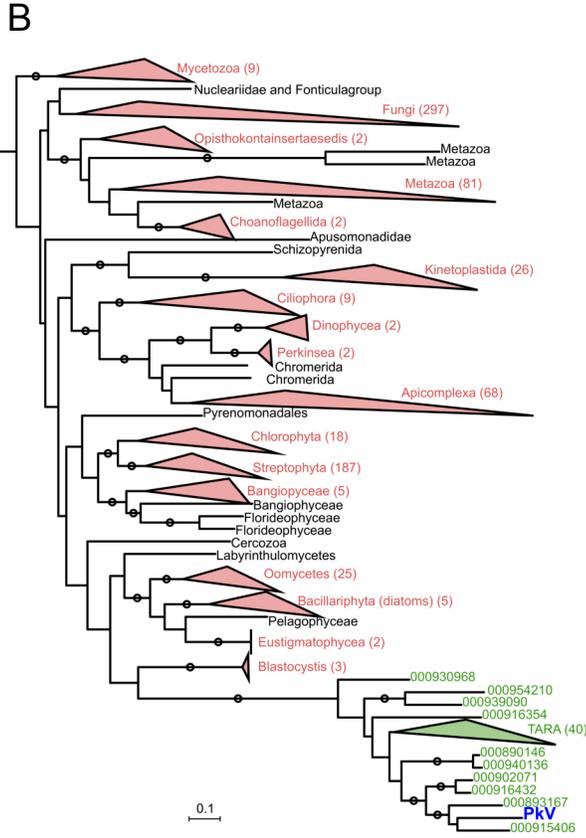
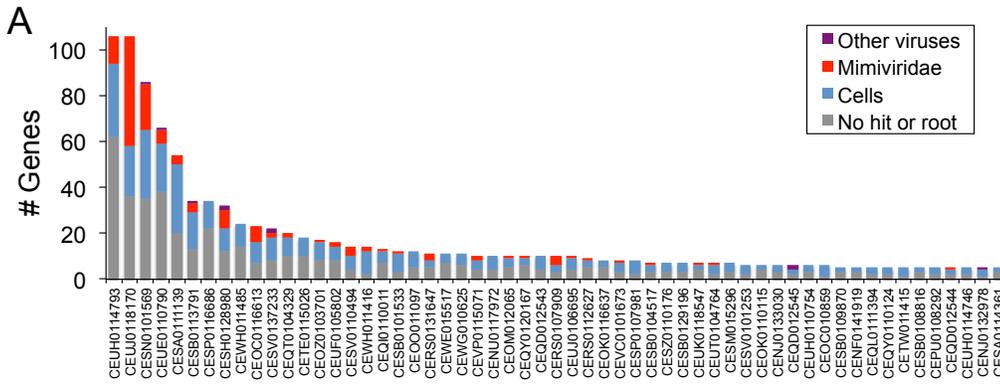


Figure 12

