Title

Evidence for the acquisition of a proteorhodopsin-like rhodopsin by a chrysophyte-infecting giant virus

Authors

Petra Byl*[1, 2, 3], Christopher R. Schvarcz[1, 2], Julie Thomy[1, 2], Qian Li[1, 2], Cori B. Williams[1], Kurt LaButti[3], Frederik Schulz[3], Kyle F. Edwards[2], Grieg F. Steward*[1, 2]

Affiliations

[1]Daniel K. Inouye Center for Microbial Oceanography: Research and Education, School of Ocean and Earth Science and Technology (SOEST), University of Hawai‘i Mānoa, 1950 East-West Road, Honolulu, HI 96822.

[2]Department of Oceanography, School of Ocean and Earth Science and Technology (SOEST), University of Hawai‘i Mānoa, 1000 Pope Road, Honolulu, HI 96822.

[3]DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720.

*Corresponding author(s): pkbyl@hawaii.edu; grieg@hawaii.edu

Abstract

Chrysophytes are widespread nanoflagellate protists in aquatic ecosystems with diverse trophic strategies ranging from phototrophy to heterotrophy. Here, we report the first isolation of a chrysophyte-infecting virus, Chrysophyceae clade-H virus SA1, referred to hereafter as ChrysoHV, on a sympatric mixotrophic chrysophyte. Both host and virus were isolated from surface waters in the subtropical North Pacific. The ChrysoHV capsid (291 ± 42 nm diameter) is associated with a loose, sac-like membrane that extends its effective diameter (716 ± 115 nm), and presents a long, thin tail extending 1,202 (± 243) nm. This unique morphology has not been previously observed in other isolates or environmental surveys. The sequenced, assembled genome is 1.19 Mbp, and phylogenetic analysis places ChrysoHV as the third cultivated member of the *Aliimimivirinae* subfamily in the *Mimiviridae* family of giant viruses. ChrysoHV is the first cultivated member of *Aliimimivirinae* to encode a rhodopsin, which has high phylogenetic relatedness and structural similarity to proteorhodopsin sequences retrieved from the same subtropical research site. The most closely related sequences from cultivated organisms derive from bacteria in the genus *Pelagibacter*. This suggests that ChrysoHV acquired the proteorhodopsin-like protein through lateral gene transfer from the chrysophyte's bacterial prey.

Key words

Nucleocytoviricota, horizontal gene transfer (HGT), mixotrophy

47 **Main text**

48 Giant viruses are large double-stranded DNA viruses in the phylum *Nucleocytoviricota* [1]. They
49 display extensive genomic repertoires, including viral homologs of cellular genes, among which
50 rhodopsin homologs are commonly found [2]. These rhodopsin homologs could augment or
51 compensate for host functions compromised during infection [3-5] or interfere with host defenses
52 [6]. Viral rhodopsin homologs have been experimentally confirmed to convert light energy into
53 electrochemical potential across multiple clades: viral rhodopsin groups I and II [3, 7, 8],
54 channelrhodopsin [4], and heliorhodopsin [6]. Furthermore, these rhodopsin homologs are
55 globally distributed in marine habitats, and their abundance decreases over depth with the
56 attenuation of sunlight [3].

57 Although giant virus rhodopsin homologs are widespread and potentially functional, only a few
58 rhodopsin-encoding giant viruses have been isolated and remain in culture with their native host
59 [5, 6, 9]. Here, we report the isolation of the giant virus ChrysoHV (Fig. 1A-E), which encodes a
60 proteorhodopsin-like protein (Fig. 2A). ChrysoHV was isolated with a co-occurring mixotrophic
61 chrysophyte from an open ocean research site in the North Pacific Subtropical Gyre (Station
62 ALOHA: 22° 45' N, 158° W; Supplementary Information).

63 We tested the host range of ChrysoHV against eight protist isolates from Stn. ALOHA, resulting
64 in no lysis for six of the eight isolates (Supplementary Fig. S1, S2A,C; Supplementary
65 Information). Both Chrysophyceae clade-H representatives UHM3500 and UHM3501 lysed
66 when challenged with ChrysoHV, but a greater fraction of UHM3501 cells lysed and more
67 virions were produced (Supplementary Fig. S2B), even though virions produced per cell (burst
68 size) were similar for each host strain (22.0-25.5). Lower mortality for UHM3500 may be due to
69 resistant subpopulations, host immune responses, or virophage integrations, which are

70 widespread in uncultivated members of Chrysophyceae clade-H [10]. The virion particles
71 produced from the two chrysophytes shared the same notable morphological features, *i.e.*, an
72 icosahedral capsid (291 ± 42 nm diameter [mean ± SD]) associated with a loose membrane (716
73 ± 115 nm), and a long tail (1,202 ± 243 nm length) ending in tail fibers (206 ± 72 nm length)
74 (Fig. 1A-C).
75 The ChrysoHV genome was fully assembled as a putatively linear DNA sequence of 1.19 Mbp
76 encoding 1,138 predicted genes with a G+C content of 23.42% (Fig. 1B; Supplementary
77 Information). The ChrysoHV genome encodes genes of diverse putative function, including the
78 ribosomal proteins eL40 and S27a, making it the second known eukaryote-infecting virus to
79 encode a ribosomal protein (Supplementary Table S1) [5]. ChrysoHV also encodes cytochrome
80 b5 and eight other genes involved with Fe use or transport (Supplementary Table S1). Notably,
81 the genome includes a region of light-sensitive proteins, encompassing a proteorhodopsin-like
82 protein surrounded by two photolyases (Supplementary Table S1, Fig. 1D, Fig. 2C). The
83 positioning and shared orientation of the photolyases around the proteorhodopsin-like protein
84 implies that it may sit on the same operon and function like the heliorhodopsin in *Trichococcus*
85 *flocculiformis*, which was proposed to bind to nearby photolyases to enhance DNA repair [11].
86 ChrysoHV is the third cultivated isolate in the *Aliimimivirinae* subfamily of *Mimiviridae* (Fig.
87 1E); other known isolates include CroV BV-PW1 [12], infecting a heterotrophic nanoflagellate,
88 and CV-XW01 [13], infecting a green alga. In this subfamily only two other representatives,
89 cPacV-1605 and GVMAG-M-3300013131-43, encode rhodopsin, which belong to the viral
90 rhodopsin group I clade (Fig. 2A). However, 40 other *Aliimimivirinae* representatives encode at
91 least one photolyase homolog (Fig. 1E) in the (6-4) photolyase or DASH cryptochrome clades.
92 Only ChrysoHV and cPacV-1605 encode both rhodopsin and photolyase homologs, but the

93 genome architecture of cPacV-1605 places the rhodopsin (CDS 205) away from the photolyase
94 cluster (CDS 131, 134) [14]. Given the mismatched distribution of rhodopsin and photolyase
95 homologs within the *Aliimimivirinae* subfamily, the ChrysoHV proteorhodopsin-like protein may
96 represent a unique adaptation to high light in the surface waters of the subtropical ocean.
97 Phylogenetic analysis of ChrysoHV's rhodopsin placed it within the proteorhodopsin clade (Fig.
98 2A), and ChrysoHV may have acquired the proteorhodopsin-like rhodopsin from the
99 chrysophyte host's putative bacterial prey, *Pelagibacter* (Fig. 2A). The sequence alignment of
100 transmembrane domain helix-III (TM3) indicates that the ChrysoHV rhodopsin is spectrally
101 tuned to blue-light (glutamine, Q; Fig. 2A), which is the dominant spectral tuning amino acid in
102 proteorhodopsins from Stn. ALOHA at 25-200 m depth [15]. While ChrysoHV lacks the
103 conserved DTE proton pumping motif found in most proteorhodopsins, it possesses an ETL
104 motif (Fig. 2A), which has been detected in proteorhodopsins from deep water at Stn. ALOHA
105 (Fig. 2A) [15]. Compared to rhodopsins with existing functional characterization, the predicted
106 structure of the ChrysoHV rhodopsin was most similar to the blue-light absorbing
107 proteorhodopsin from uncultured Gamma-proteobacterium clone HOT 75m4 (Fig. 2B)
108 confirmed to have ion (H^+) pumping functionality [16]. Although the amino acid identity was not
109 the highest (35.9%), this HOT 75m4 proteorhodopsin was also retrieved from Stn. ALOHA [16].
110 ChrysoHV does not encode β-carotene, precursor to retinal, the light-absorbing chromophore in
111 proteorhodopsin, but the retinal binding site (lysine, K) in transmembrane domain helix VII
112 (TM7) was conserved (Fig. 2A). Further investigation is needed to confirm whether the
113 ChrysoHV proteorhodopsin-like protein acts as an ion pump.
114 Analysis of phylotypes related to ChrysoHV (≥ 90% polB amino acid identity) from the TARA
115 Oceans survey [17] revealed greatest prevalence in the subtropical surface ocean (Supplementary

116 Fig. S3). In this environment, diverse microorganisms from proteobacteria to microalgae utilize
117 rhodopsins to supplement their energetic needs [15, 18]. SAR11 bacteria, of which *Pelagibacter*
118 is an abundant subclade, account for 25% of all plankton [19], and are a likely prey source for
119 mixotrophic algae such as Chrysophyceae clade-H sp. UHM3500 and UHM3501. The discovery
120 of a proteorhodopsin-like protein encoded by a giant virus supports the conceptual model that
121 horizontal gene transfer is driven by the co-mingling of infectious virions and ingested prey
122 within the same host cell [20]. By leveraging naturally co-occurring and experimentally
123 accessible virus-host pairs, this study lays the groundwork for future investigations into how the
124 virocell environment of predatory protists influences gene acquisition in giant viruses.
125 **Figures**

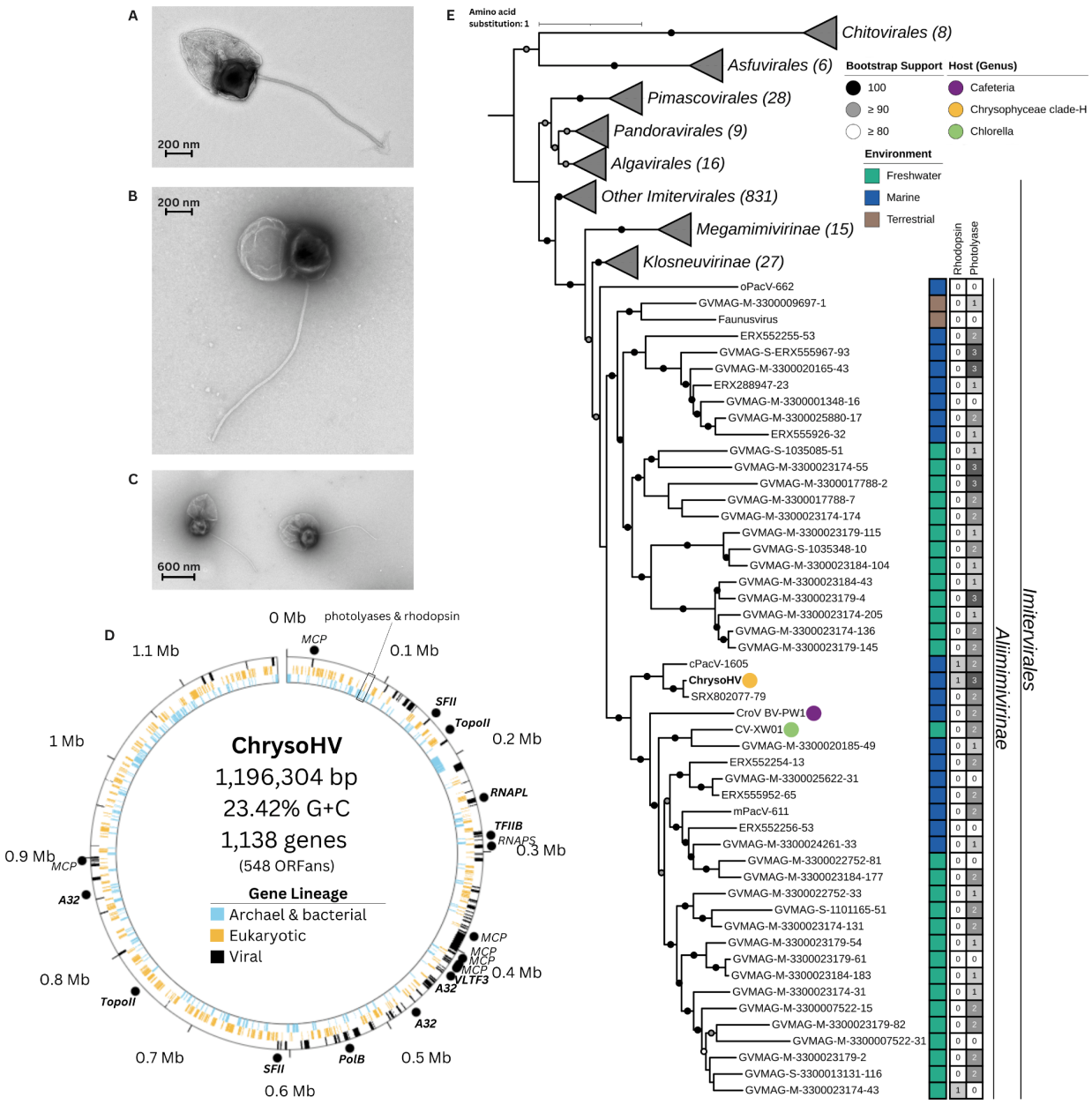
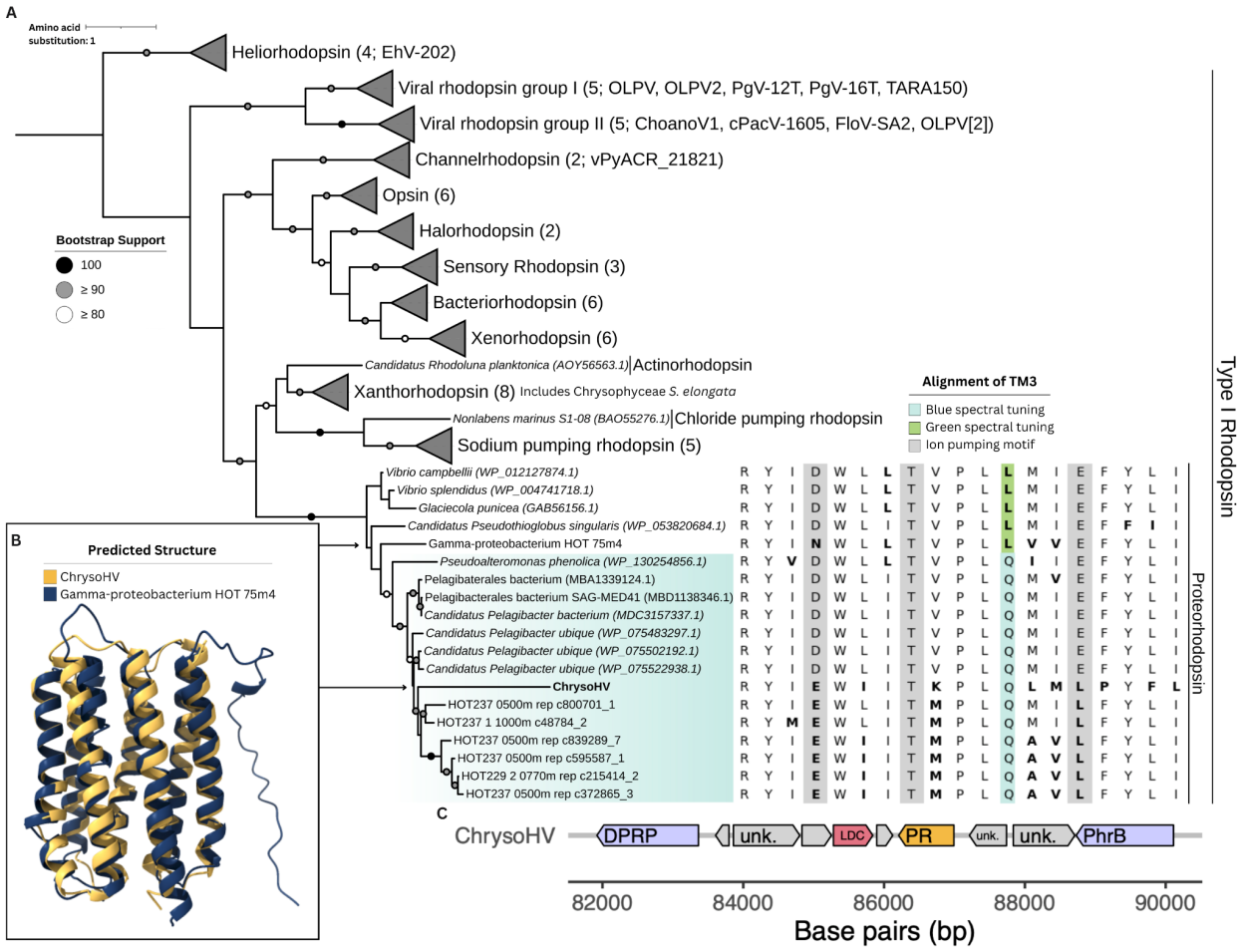


Figure 1. (A) Transmission electron microscopy (TEM) negative stain micrograph of ChrysoHV propagated on UHM3501. (B, C) TEM negative stain micrograph of ChrysoHV propagated on UHM3500. (D) Genome plot of ChrysoHV with track placement and color indicating gene lineage by domain: inner to outer tracks show prokaryote (blue), eukaryote (yellow), and virus (black). Phylogenetically informative giant virus marker genes are represented as black dots on the outside track and show gene name. The genes used in the *Nucleocytoviricota* phylogenetic

analysis are bolded. (E) Maximum Likelihood (ML) phylogeny of the *Nucleocytoviricota* phylum based on a concatenated alignment of translated marker genes (*TopoII*, *RNAPL*, *TFIIB*, *VLTFB*, *A32*, *PolB* and *SFII*) using model LG+I+F+G4 with ultrafast bootstrap supports (-bb 1000) shown by node symbol color– white (≥ 80%), gray (≥ 90%), and black (100%) with no node symbol for supports < 80%. The tree is rooted with *Chitovirales* as the outgroup, and the scale bar represents one amino acid substitution per site. The frequency of light-sensitive proteins such as rhodopsin and photolyases (phrB) and (DRPR) present in the *Aliimimivirinae* subfamily are annotated in the grayscale matrix and the environment the virus was collected from is shown in jade (freshwater), blue (marine), and brown (terrestrial).



143 **Figure 2**. (A) Microbial rhodopsin ML phylogeny with heliorhodopsin as the outgroup, built
144 with model LG+F+R10. Select representatives of viral rhodopsin homologs and putative
145 chrysophyte rhodopsins are referred to next to their corresponding collapsed clade. The
146 proteorhodopsin (PR) clade is annotated with the alignment of transmembrane helix 3 (TM3)
147 outlining the ion-pumping motif (gray) and the spectral tuning amino acid (green or blue).
148 Residues that differ from the most common residue at that locus are in bold (B). The predicted
149 structure of ChrysoHV's proteorhodopsin-like protein conserved domain (yellow) and the HOT
150 75m4 Gamma-proteobacterium proteorhodopsin (blue). (C) Gene neighborhood of light-
151 sensitive proteins in ChrysoHV, including the proteorhodopsin (CDS 97) between the two
152 photolyases (CDS 91 and 100).
153

154 **Acknowledgments**

155 We thank the personnel of the Hawai‘i Ocean Time-series program (NSF award 12-60164) and
156 Research Vessel Kilo Moana for assistance with seawater collection. We thank the Biological
157 Electron Microscope Facility (BEMF) at the University of Hawai‘i Mānoa, especially Facility
158 Supervisors, Tina Weatherby and Dr. Orion Rivers, for their assistance with electron microscopy
159 procedures. A special thanks to US DOE Joint Genome Institute Scientific Project Manager,
160 Kerrie Barry, for overseeing DNA library preparation, quality control, and sequencing of
161 ChrysoHV.
162 We are grateful for the thoughtful feedback provided by Dr. Karen Selph (Department of
163 Oceanography, University of Hawai‘i Mānoa), and Dr. Alison Sherwood (School of Life
164 Sciences, University of Hawai‘i Mānoa).
165

Funding

This work was supported by the US National Science Foundation Graduate Research Fellowship Program (Grant No. 1842402 to PB), Division of Ocean Sciences (Grant No. 2129697 to GFS and KFE), and RII Track-2 FEC (Grant No. 1736030 to GFS and KFE). The work conducted by the US Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a Department of Energy (DOE) Office of Science User Facility, was supported by the Office of Science of the US DOE, operated under Contract No. DE-AC02-05CH11231.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability

ChrysoHV sequencing data is available through JGI IMG under Sequencing Project ID 1432190, and the genome assembly has been deposited to NCBI GenBank under accession ID PV764857. Additional data and scripts on gene alignments, phylogenies, genome annotation, protein structural predictions, host range challenges, and biogeography are available in the Dryad repository: http://datadryad.org/share/8duyiXJ3riDLgLbseVaWvUBnbOPOM4SkBw9WNBaRYA4. The virus and protist strains used in this study are available through the Marine Viral Ecology Laboratories at the University of Hawai‘i Mānoa upon request to KFE and GFS.

References

188 1. Aylward, F. O. *et al.* Taxonomic update for giant viruses in the order Imitervirales (phylum
189 Nucleocytoviricota). *Arch. Virol.* **168**, 283 (2023). 10.1371/journal.pbio.3001430
190 2. Moniruzzaman, M. *et al.* Virologs, viral mimicry, and virocell metabolism: the expanding
191 scale of cellular functions encoded in the complex genomes of giant viruses. *FEMS Microbiol.*
192 *Rev.* **47**, fuad053 (2023). 10.1093/femsre/fuad053
193 3. Needham, D. M. *et al.* A distinct lineage of giant viruses brings a rhodopsin photosystem to
194 unicellular marine predators. *Proc. Natl. Acad. Sci.* **116**, 20574–20583 (2019).
195 10.1073/pnas.1907517116
196 4. Rozenberg, A. *et al.* Lateral Gene Transfer of Anion-Conducting Channelrhodopsins between
197 Green Algae and Giant Viruses. *Curr. Biol.* **30**, 4910-4920.e5 (2020). 10.1016/j.cub.2020.09.056
198 5. Thomy, J., Schvarcz, C. R., McBeain, K. A., Edwards, K. F. & Steward, G. F. Eukaryotic
199 viruses encode the ribosomal protein eL40. *Npj Viruses* **2**, 51 (2024).
200 6. Hososhima, S. *et al.* Proton-transporting heliorhodopsins from marine giant viruses. *eLife* **11**,
201 e78416 (2022). 10.1038/s44298-024-00060-2
202 7. Bratanov, D. *et al.* Unique structure and function of viral rhodopsins. *Nat. Commun.* **10**, 4939
203 (2019). 10.1038/s41467-019-12718-0
204 8. Zabelskii, D. *et al.* Viral rhodopsins 1 are an unique family of light-gated cation channels. *Nat.*
205 *Commun.* **11**, 5707 (2020). 10.1038/s41467-020-19457-7
206 9. Brussaard, C. P. D., Bratbak, G., Baudoux, A. -C. & Ruardij, P. Phaeocystis and its interaction
207 with viruses. *Biogeochemistry* **83**, 201–215 (2007). 10.1007/s10533-007-9096-0
208 10. Castillo, Y. M. *et al.* Assessing the viral content of uncultured picoeukaryotes in the global-
209 ocean by single cell genomics. *Mol. Ecol.* **28**, 4272–4289 (2019). 10.1111/mec.15210

210 11. Shim, J. *et al.* Heliorhodopsin Helps Photolyase to Enhance the DNA Repair Capacity.
211 *Microbiol. Spectr.* **10**, e02215-22 (2022). 10.1128/spectrum.02215-22
212 12. Fischer, M. G., Allen, M. J., Wilson, W. H. & Suttle, C. A. Giant virus with a remarkable
213 complement of genes infects marine zooplankton. *Proc. Natl. Acad. Sci.* **107**, 19508–19513
214 (2010). 10.1073/pnas.1007615107
215 13. Sheng, Y., Wu, Z., Xu, S. & Wang, Y. Isolation and Identification of a Large Green Alga
216 Virus (*Chlorella* Virus XW01) of *Mimiviridae* and Its Virophage (*Chlorella* Virus Virophage
217 SW01) by Using Unicellular Green Algal Cultures. *J. Virol.* **96**, e02114-21 (2022).
218 10.1128/jvi.02114-21
219 14. Needham, D. M. *et al.* Targeted metagenomic recovery of four divergent viruses reveals
220 shared and distinctive characteristics of giant viruses of marine eukaryotes. *Philos. Trans. R. Soc.*
221 *B Biol. Sci.* **374**, 20190086 (2019). 10.1098/rstb.2019.0086
222 15. Olson, D. K., Yoshizawa, S., Boeuf, D., Iwasaki, W. & DeLong, E. F. Proteorhodopsin
223 variability and distribution in the North Pacific Subtropical Gyre. *ISME J.* **12**, 1047–1060
224 (2018). 10.1038/s41396-018-0074-4
225 16. Ran, T., *et al*. 2013. Cross-protomer interaction with the photoactive site in oligometric
226 proteorhodopsin complexes. *Acta Crystallogr* **69**, 1965–1980 (2013).
227 10.1107/S0907444913017575
228 17. Endo, H. *et al.* Biogeography of marine giant viruses reveals their interplay with eukaryotes
229 and ecological functions. *Nat. Ecol. Evol.* **4**, 1639–1649 (2020). 10.1038/s41559-020-01288-w
230 18. Strauss, J. *et al.* Plastid-localized xanthorhodopsin increases diatom biomass and ecosystem
231 productivity in iron-limited surface oceans. *Nat. Microbiol.* **8**, 2050–2066 (2023).
232 10.1038/s41564-023-01498-5

19. Giovannoni, S. J. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. *Annu. Rev. Mar. Sci.* **9**, 231–255 (2017). 10.1146/annurev-marine-010814-015934

20. Filée, J., Siguier, P. & Chandler, M. I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends Genet.* **23**, 10–15 (2007). 10.1016/j.tig.2006.11.002