

1 **Unmasking cellular response of a bloom-forming alga to virus infection by**
2 **resolving expression profiling at a single-cell level**

3 Shilo Rosenwasser^{1,5*}, Miguel J. Frada^{1,4}, David Pilzer², Ron Rotkopf³ and Assaf Vardi^{1*}

4 ¹Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot,
5 7610001, Israel

6 ²Genomic Technologies Unit, Weizmann Institute of Science, Rehovot, 7610001, Israel.

7 ³Bioinformatics and Biological Computing Unit, Weizmann Institute of Science, Rehovot, Israel.

8 ⁴The Interuniversity Institute for Marine Sciences, Eilat, Israel & Department of Ecology,
9 Evolution and Behavior, Silberman Institute of Life Sciences, The Hebrew University of
10 Jerusalem, Israel

11 ⁵Present address: The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture,
12 The Hebrew University, Rehovot 7610001, Israel.

13 *Corresponding author: shilo.rosenwaser@mail.huji.ac.il, assaf.vardi@weizmann.ac.il

14 **Abstract:**

15 The interaction between *Emiliania huxleyi*, a bloom-forming alga, and its specific large virus
16 (EhV), is one of the most ecologically important algal-virus model system. Infection by EhV
17 depends on profound rewiring of metabolic network to supply essential building blocks for viral
18 assembly. Despite the clear evidence for the modulation of the host metabolic and signaling
19 pathways during infection, there is a major bottleneck to accurately discern between viral hijacking
20 strategies and host defense responses. Here we uncovered cell-to-cell heterogeneity of isogenic
21 cell population by host and virus gene expression profiling at a single cell level. This approach
22 enabled mapping of cells into newly defined infection states. Clustering of cells based on their
23 infection state and examination of host gene expression profiles unmasked a yet unrecognized
24 early phase in host response that occurs prior to viral expression. This early response includes the
25 induction of genes involved in cell fate regulation, ROS metabolism and sphingolipid catabolism.
26 Upregulation of various host metabolic genes coincided with late-viral gene and their expression
27 was suppressed in stationary phase cells which exhibited compromised infection. We propose that

28 resolving host-virus arms race at a single-cell level will provide important mechanistic insights
29 into viral life cycles and will uncover host defense strategies.

30

31 **Introduction**

32 Marine viruses are recognized as major ecological and evolutionary drivers and have immense
33 impact on the community structure and the flow of nutrients through marine microbial food webs
34 [1-5]. The cosmopolitan coccolithophore *Emiliania huxleyi* (Prymnesiophyceae, Haptophyta) is a
35 widespread unicellular eukaryotic alga, responsible for large oceanic blooms [6, 7]. Its intricate
36 calcite exoskeleton accounts for ~1/3 of the total marine CaCO₃ production [8]. *E. huxleyi* is also
37 a key producer of dimethyl sulfide [9], a bioactive gas with a significant climate-regulating role
38 that seemingly enhances cloud formation [10]. Therefore, the fate of these blooms may have a
39 critical impact on carbon and sulfur biogeochemical cycles. *E. huxleyi* spring blooms are
40 frequently terminated as a consequence of infection by a specific large dsDNA virus (*E. huxleyi*
41 virus, EhV) [11, 12]. The availability of genomic and transcriptomic data and a suite of host
42 isolates with a range of susceptibilities to various EhV strains, makes the *E. huxleyi*-EhV a
43 trackable host-pathogen model system with important ecological significance [13-19].

44 Recent studies demonstrated that viruses significantly alter the cellular metabolism of their
45 host either by rewiring of host-encoded metabolic networks, or by introducing virus-encoded
46 auxiliary metabolic genes (vAMG) which convert the infected host cell into an alternate cellular
47 entity (the virocell [20]) with novel metabolic capabilities [21-26]. A combined transcriptomic and
48 metabolomic approach taken during *E. huxleyi*-EhV interaction revealed major and rapid
49 transcriptome remodeling targeted towards *de novo* fatty acid synthesis [18] fueled by glycolytic
50 fluxes, to support viral assembly and the high demand for viral internal lipid membranes [27, 28].
51 Lipidomic analysis of infected *E. huxleyi* host and purified EhV virions further revealed a large
52 fraction of highly saturated triacylglycerols (TAGs) that accumulated uniquely within distinct lipid
53 droplets as a result of virus-induced lipid remodeling [26]. The EhV genome encodes for a unique
54 vAMG pathway for sphingolipid biosynthesis, never detected before in any other viral genome.
55 Biochemical characterization of EhV-encoded serine palmitoyl-CoA transferase (SPT), a key
56 enzyme in the sphingolipid biosynthetic pathway, revealed its unique substrate specificity which
57 resulted in the production of virus-specific glycosphingolipids (vGSLs) composed of unusual

58 hydroxylated C17 sphingoid-bases [29]. These viral-specific sphingolipids are essential for viral
59 assembly and infectivity and can induce host programmed cell death (PCD) during the lytic phase
60 of infection [14, 30]. Indeed, EhV can trigger hallmarks of PCD, including production of reactive
61 oxygen species (ROS), induction of caspase activity, metacaspase expression, changes in
62 ultrastructure features and compromised membrane integrity [31-33].

63 The high metabolic demand for building blocks required to support synthesis, replication
64 and assembly of large viruses with high burst size as EhV [33-35], point to high dependence of
65 viruses on their host metabolic state for optimal replication [20, 36]. Consequently, heterogeneity
66 in host metabolic states as a result of complex interactions between nutrient availability and stress
67 conditions may affect the infection dynamics. However, almost all of our current understanding of
68 the molecular mechanisms that govern host-virus interactions in the ocean, is derived from
69 experiments carried out at the population level, assuming synchrony and uniformity of the cell
70 populations and neglecting any heterogeneity. Additionally, averaging the phenotypes of a whole
71 population hinders the investigation of essential life cycle strategies to evade viral infection that
72 can be induced only by rare subpopulations[37]. Understanding microbial interactions at a single-
73 cell resolution is an emerging theme in microbiology. It enables the detection of complex
74 heterogeneity within microbial populations and has been instrumental to identify novel strategies
75 for acclimation to stress [38, 39]. The recent advancement of sensitive technologies to detect gene
76 expression from low input-RNA allows quantification of heterogeneity among cells by analyzing
77 gene expression at the single cell level [40, 41]. High-throughput profiling of single-cell gene
78 expression patterns in mammals and plant cells led to the discovery of new cell types, detection
79 of rare cell subtypes, and provides better definition and cataloging of developmental phases in high
80 resolution [42-46]. Importantly, the role of cell-to-cell communication and variability in
81 controlling infection outcomes has only been recently demonstrated in cells of the mammalian
82 immune system in response bacterial pathogens [47-50]. Cell-to-cell variability in host response
83 to viral infection was observed in several mammalian viruses and was attributed to several factors,
84 including intrinsic noise (e.g. stochasticity of biochemical interactions involved in the infection
85 process), the number of viral genomes initiating the infection process and the specific cell-state
86 before the infection [50-53].

87 Recently, simultaneous detection of host and pathogen gene expression profile was
88 suggested as a powerful tool used to gain a better understanding of the molecular mechanisms

89 underlying the infection process and to identify host resistance responses [20, 54-56]. However,
90 the existence of cell-to-cell variability during infection suggest that key events in host response
91 are masked by conventional bulk cell expression profiling and that detection of gene expression
92 on single cell resolution may uncover hidden host responses.

93 Here, we quantified the dynamics of host and virus gene expression profiles of individual
94 cells during infection of *E. huxleyi* populations. We provide strong evidence for heterogeneity
95 within the population and discern between cells at different infection states based on their viral
96 gene expression signatures. We unravel an unrecognized phase of early host response that
97 preceded viral gene expression within infected cells. We suggest that examining host and virus
98 gene expression profiles at the single cell resolution allows to infer the temporal dynamic of the
99 infection process, thereby it serve as an attractive approach to decipher the molecular mechanism
100 underlying host-virus interaction.

101

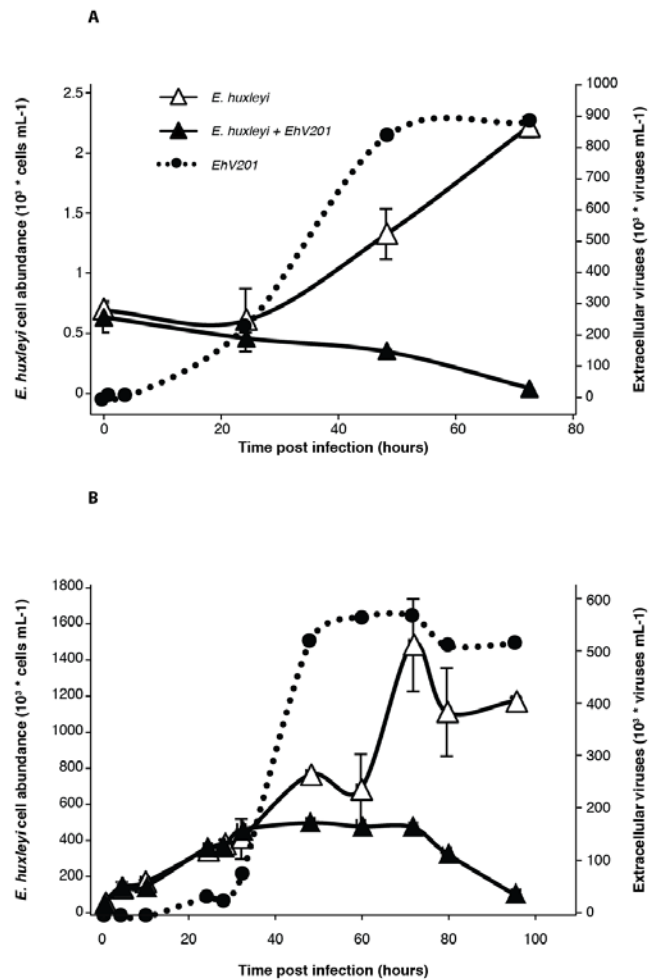
102 **Results and Discussion:**

103 To examine the variability within infected *E. huxleyi* cells, we measured the expression levels of
104 selected host and viral genes over the course of infection at a single-cell resolution. In total, 491
105 cells were isolated during infection of two *E. huxleyi* host strains at different phases: a non-
106 calcified strain CCMP2090, at 0, 2, 4, 24 hours post infection (hpi), and the calcified RCC1216
107 strain at 0, 6, 24, 48 hpi. These two host-virus systems exhibit different infection dynamics
108 reflected by faster production of viruses and host cell lysis in CCMP2090 (Figure 1A and B).

109 We used the C1 single-cell Auto Prep System to sort and extract RNA from single *E. huxleyi* cells
110 during viral infection by EhV201). The presence of a single cell captured in an individual isolation
111 chamber was confirmed by microscopic inspection of emitted chlorophyll auto-fluorescence
112 (Figure 2A). In order to detect variability in viral infection states, we conducted simultaneous
113 measurements of expression profiles of host and virus genes at a single-cell level by using
114 multiplexed qPCR. Viral genes were selected based on their temporal expression pattern during
115 different phases of infection (early, mid and late) [18, 57]. We examined the expression levels of
116 host genes that encode proteins which are involved in several metabolic pathways that were highly
117 remodeled during infection, including primary metabolism (glycolysis, fatty acid biosynthesis),
118 sphingolipid and terpenoid metabolism, autophagy and antioxidant genes [18, 26, 32, 33]. In

119 addition, we examined the expression of host genes associated with life cycle, meiosis and PCD
120 [31] that exhibited induction during infection [58], (see Supplemental Table 1 for primers list).

121



122 **Figure 1: Infection dynamics of *E. huxleyi* by its specific virus EhV.** *E. huxleyi* CCMP2090 (A) and RCC1216
123 (B) cultures were infected by the EhV201 lytic virus and compared with uninfected control cells. Host cell
124 abundance and production of extracellular viruses were monitored using flow-cytometry. (mean \pm SD, n = 3, at
125 least 6000 cells were measured at each time point).

122

123 To test the sensitivity in detection of gene expression measurement on a single cell level,
124 we spiked-in, to each C1 well, a set of External RNA Controls Consortium (ERCC) molecules that
125 span a wide range of RNA concentrations (from ~0.5 to ~100 molecules per well). We
126 subsequently quantified their concentration using similar qPCR amplification setup as used for the
127 host and virus genes. Pairwise correlation between spike concentrations and Et (Et=30-Ct) values
128 obtained from the qPCR was >0.98 (Pearson correlation coefficient, p-value= $4.2 \cdot 10^{-12}$, Figure
129 2B). We found a highly sensitive level of detection with 40% probability to detect an RNA spike
130 that is at a level of 1 molecule per sample (Figure 2C), similar to the detection level reported for

131 mammalian cells [59]. Mean expression of viral and host genes in all examined cells were found
132 to be 11.8 ± 4.0 and 6.96 ± 2.5 (Et values \pm SD), respectively (Figure 2D).

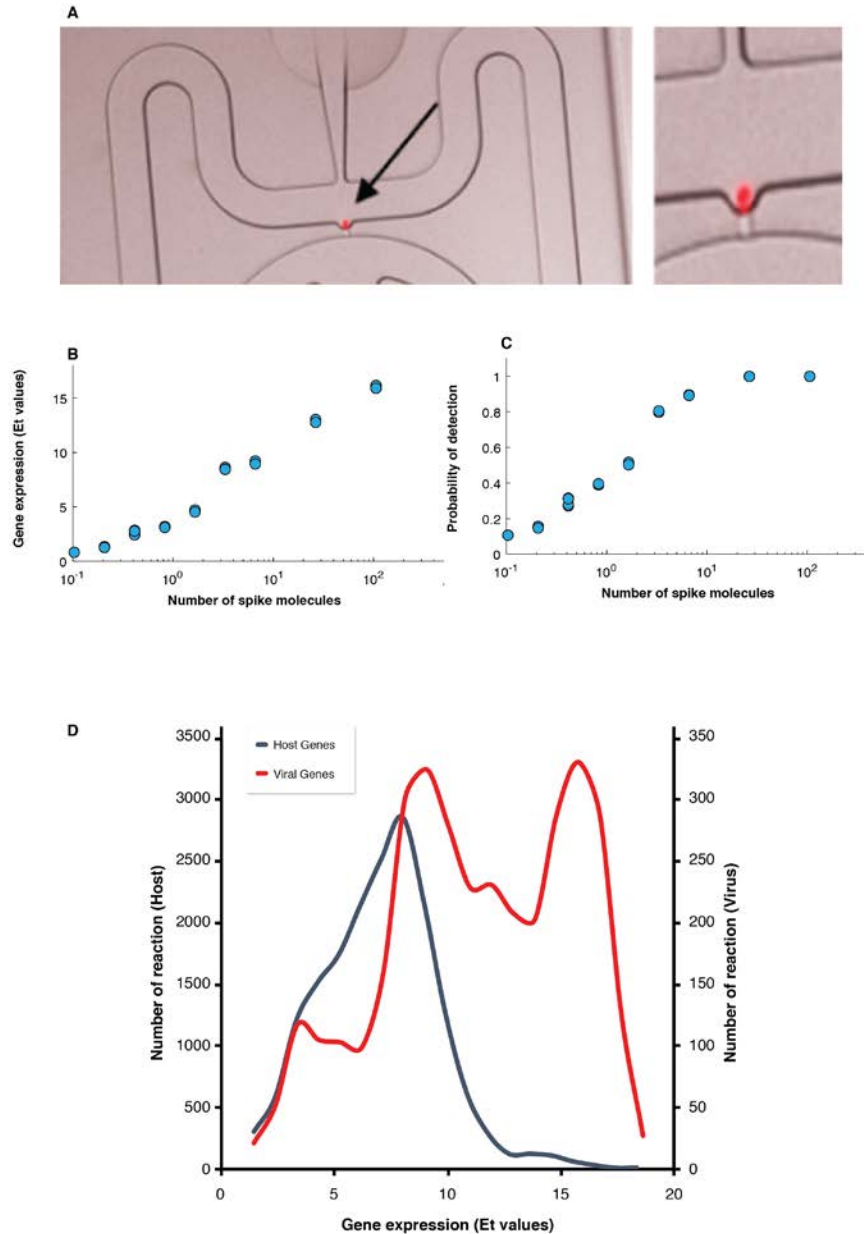
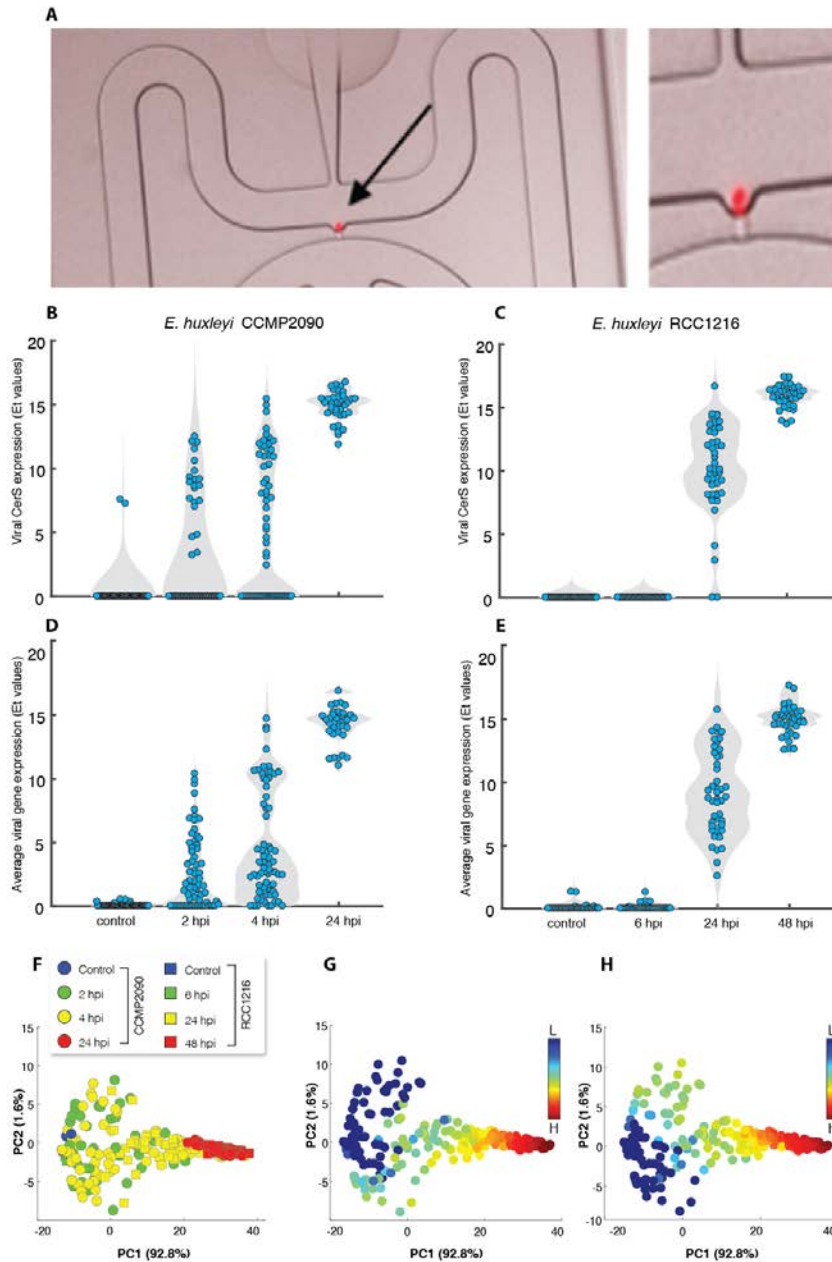


Figure 2. Host and virus gene expression profiling at a single cell level. (A) Automated microfluidic capture of a single *E. huxleyi* cell in the C1 chip (red: chlorophyll autofluorescence, indicated by a black arrow), the image on the right is a zoom into the image of a single cell. (B,C) Examination of detection level of single-cell gene expression analysis. A set of ERCC RNA molecules were spiked to each C1 well and their level was determined using multiplex qPCR. (B) The fraction of wells with positive qPCR reaction ($C_t < 30$) for each examined spike. (C) The correlation between the average level of expression (Et) value and the number of spike molecule. (D) Distribution of host and virus genes expression among individual cells. The average expression values of host and viral genes among isolated single cells was calculated and the distribution is presented.

134 We detected a high variability between expression profiles of viral genes within the same
135 infected population. For example, heterogeneity in the expression levels of virus-encoded
136 ceramide synthase (*vCerS*, EPVG_00014), a key enzyme in sphingolipid biosynthesis [18, 29] was
137 detected during early phase of infection (2 and 4 hpi of CCMP2090, Figure 3A). Similar results
138 were obtained for the average expression of 10 viral genes (Figure 3C). Heterogeneity in viral
139 expression was also detected in infected RCC1216 cells at 24 hpi (Figure 3B and D), corroborating
140 the delayed dynamics of lytic infection as compared with CCMP2090 (Figure 1). At the onset of
141 viral lytic phase (24 hpi in CCMP2090 and 48 hpi in RCC1216), all of the examined cells showed
142 high viral gene expression (Figure 3A-D), suggesting that viruses eventually infected all of the
143 examined host cells. Nevertheless, we cannot exclude the existence of a rare subpopulation that
144 did not express viral genes.

145 Principal component analysis (PCA) of viral gene expression among individual host cells
146 showed that infected cells are distributed across a continuum of viral expression levels (Figure
147 3E). All viral genes had positive and similar coefficients and contributed to the separation along
148 the PC1 component, while they had either positive or negative coefficient values for the PC2
149 component. For example, a negative coefficient value ($r = -0.43$) was found for the viral RNA
150 polymerase gene (EPVG_00062) which was previously reported to be expressed at early phases
151 of infection [18, 57], while a positive value ($r = 0.50$) was found for a viral gene (EPVG_00010)
152 that is known to be expressed at late phases of infection. Accordingly, cells with high PC2 levels
153 expressed EPVG_00010 and not EPVG_00062, while cells with low PC2 values exhibited the
154 opposite trend (Figure 3F and G). In order to validate the ability of single-cell expression profiling
155 to separate between cells at different viral infection states, we analyzed an extended set of 25 viral
156 genes associated with the early, mid and late phases of infection in 343 individual cells
157 (Supplemental Table 1). Again, we observed differential contribution of viral genes associated
158 with early, mid and late phases of infection to the separation along PC2 (Supplemental Figure 1).
159 These results demonstrate that PC1, which captures >90% of the variability of viral gene
160 expression and is highly correlated to the average viral infection level ($r = 0.99$, Pearson linear
161 correlation), reflected the intensity of viral infection. Accordingly, we used the score value of PC1
162 as an index for the level of expression of viral genes in each individual cell and termed it “infection
163 index”.

164



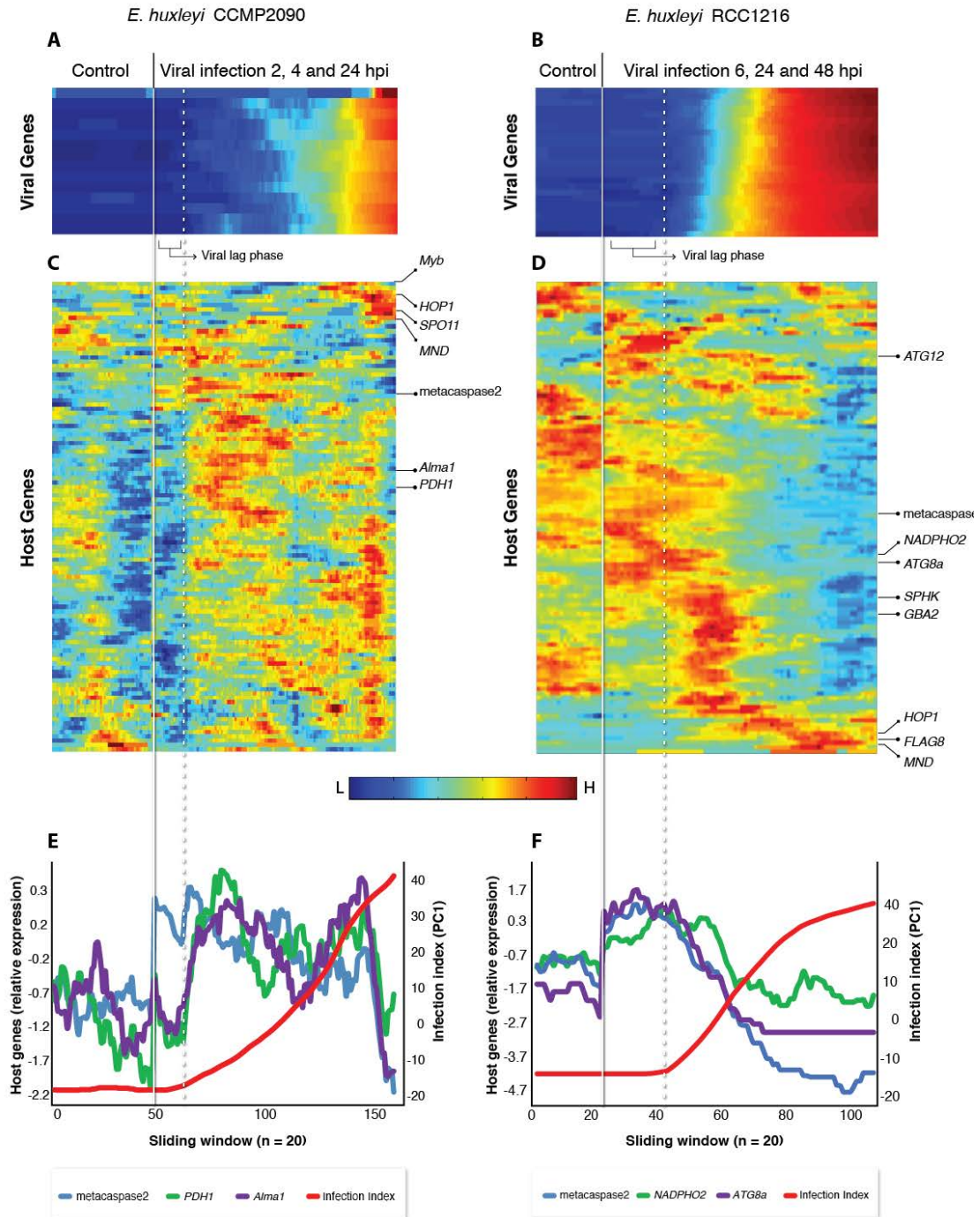
165

Figure 3. Single-cell analysis of infected population unmasking heterogeneity in viral gene expression profiles.

(A, B) Violin plots of the expression value (Et) of viral dihydroceramide synthase (vCerS, EPVG_14, Gene bank: AET97902.1) at different hours post infection (hpi) of CCMP2090 (A) and RCC1216 (B) cells infected by EhV201. (C, D) Violin plots of the mean expression value of 10 viral genes at different times post infection of CCMP2090 (C) and RCC1216 (D) with EhV201. (E) Principal component analysis (PCA) plots of gene expression profiles of 10 viral genes derived from 491 individual *E. huxleyi* cells that were isolated from infected CCMP2090 (circles) and RCC1216 (squares) cultures at different hpi. (F, G) The same PCA plots as in (E) with overlay, by a color code, representing the expression level of viral genes (Et values) that are associated with early-mid (F) and late (G) phases of viral infection (EPVG_00062 and EPVG_00010, respectively).

166 We further realized that averaging host phenotypes over the course of infection might
167 hinder our ability to observe the initial response of the host to viral infection and that single-cell
168 analysis could significantly increase the resolution for sensitive detection of host response at this
169 early stage of infection. We therefore re-ordered infected cells based on their viral infection index,
170 rather than the actual time of infection (i.e. hpi), resulting in “pseudotemporal” hierarchy of single
171 cells. Intriguingly, we unmasked a fraction of cells that were exposed to the virus but did not
172 exhibit any detectable expression of viral genes. These cells had similar infection index values as
173 control cells, with PC1 values < -10 . We found that 33/179 (17%) and 40/126 (31.7%) of infected
174 cells of CCMP2090 and RCC1216, respectively, were at this distinct “lag phase” of viral infection.
175 These individual cells were analyzed for their respective host gene expression levels based on a
176 sliding window approach, as it is less sensitive to technical noise, typically observed in single cell
177 data. We also used a statistical model to test for genes that are differentially expressed at these
178 early stages of viral infection. This model incorporates the two types of heterogeneity that usually
179 appear in single cell data, namely, the percentage of cells expressing a gene in a given population
180 (e.g. Et value > 0) and the variability in expression levels in cells exhibiting positive expression
181 values [60]. Up-regulation of several host genes in infected cells of both strains was detected prior
182 to viral expression (Figure 4A-F and supplemental Table 3). An intriguing example is the
183 metacaspase-2 gene ($p=0.000018$) which was previously suggested to be induced and recruited
184 during EhV lytic phase and activation of *E. huxleyi* PCD [31]. We also found early induction of
185 triosephosphate isomerase (*TPI*, $p=0.0022$) and phospholipid:diacylglycerol acyltransferase
186 (*PDAT*, $p = 0.013$) in CCMP2090, which are involved in glycolysis and TAG biosynthesis, as well
187 as down-regulation of phosphomevalonate kinase (*PMVK*, $p = 0.034$) which is involved in
188 isoprenoid and sterol biosynthesis (supplemental Table 3). A similar analysis in RCC1216
189 uncovered the early induction of genes involved in sphingolipid catabolism (ceramidase-2 and
190 ceramide glycosyltransferase1, $p = 0.047$ and $p = 0.033$, respectively) and autophagy (*ATG8a*,
191 *ATG8b* and *ATG12*, $p = 0.000002$, $p = 0.003$ and $p = 0.002$, respectively). Since both autophagy
192 [33] and de novo sphingolipid biosynthesis [18, 29] are essential for EhV life cycle, early induction
193 of these pathways may serve as an effective viral strategy to prime optimal infection. Alternatively,
194 this phase of early host response prior to viral gene expression may represent a newly unrecognized
195 phase of immediate host anti-viral defense response. For example, we detected up-regulation of
196 NADPH oxidase (*NADPHO2*, $p = 0.032$) that was demonstrated to mediate diverse host-pathogen

197 recognition [61] and is an enzymatic source for superoxide production which may serve as an early
198 anti-viral mechanism. Similarly, induction of *Alma1* that catalyzes the production of DMS and
199 acrylate from DMSP [9] may serve as an essential antioxidant due to activation of oxidative burst
200 during EhV infection [32, 62]. At the late stages of infection (infection index >10), we observed
201 induction of several meiosis-related genes, including *HOP1* and *MND* in both CCMP2090 and
202 RCC1216, and two *SPO11* homologues and *MYB* in CCMP2090 (Figure 4C and D). These results
203 are in agreement with previous studies that suggested a phenotype switch of *E. huxleyi* to evade
204 viral infection [37] and propose the induction of meiosis-related genes as part of transcriptomic
205 reprogramming of highly infected cells [58].
206



207

Figure 4. Host-virus co-expression patterns across viral infection states. Cells were re-ordered based on their infection index to reconstruct pseudotemporal separation of the infection process. (A-D) Clustrogram representation of the average expression value of viral (a, b) and host (c, d) genes across the infection process of CCMP2090 (A, C) and RCC1216 (B, D) using a sliding window approach (window size = 20 cells). (E, F) plots of expression of selected host genes along the viral infection index (PC1) in the sliding windows of 20 cells for CCMP2090 (E) and RCC1216 (F).

208

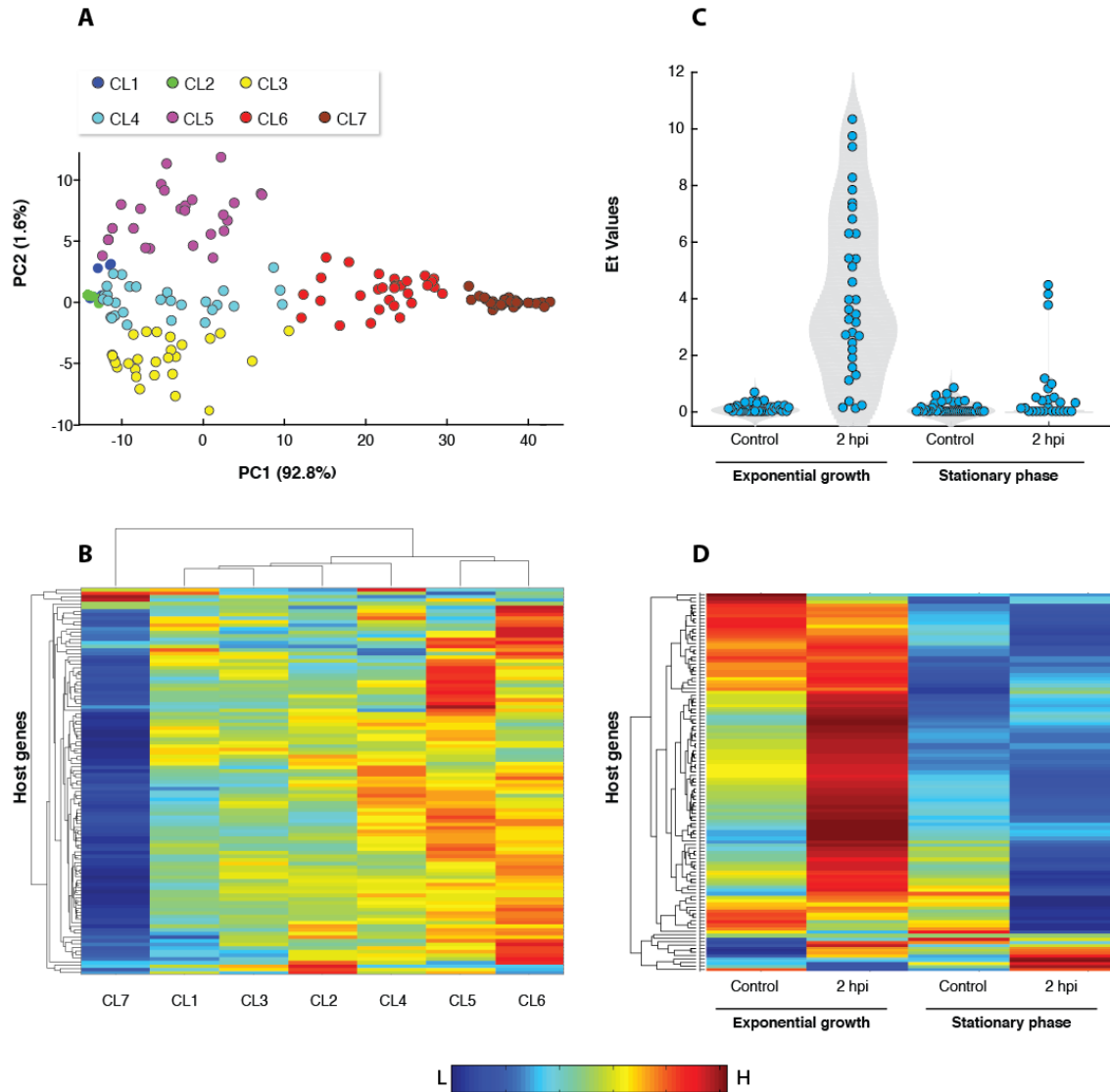


Figure 5. Viral expression is associated with induction of host metabolic genes at distinct phases of infection in exponential and stationary cultures. (A) PCA plots represent gene expression profiles of 10 viral genes derived from infected CCMP2090 culture at different hpi. Cells were clustered manually based on their infection index (PC1) and PC2 scores. (B) Clustogram representation of the of expression values of 109 host metabolic genes in the different clusters (defined in A). (C) Violin plots of the mean expression of viral genes in exponential and stationary phase CCMP2090 cells at 2 hpi and in uninfected cells (Control). (D) Clustogram representation of expression values of 109 host metabolic genes in exponential and stationary phase CCMP2090 cells at 2 hpi and in uninfected cells.

209

210 As described above single-cell analysis allowed to discern between cell expressing early
211 and late viral gene marker (Figure 3 and Supplemental Figure 1). To further characterize host gene
212 expression during different phases of infection, we manually clustered CCMP2090 cells according
213 to their infection index (PC1) and the expression of either early or late viral genes (PC2). This
214 analysis showed that induction of most of host metabolic genes occurred in cells that expressed

215 predominantly late viral genes (Figure 5A and B, CL5, $-10 < PC1 < 10$, $PC2 > 2.5$) and in cells with
216 moderate expression of viral genes (Figure 5A and B, CL6, $10 < PC1 < 28$). Down-regulation of
217 many host genes was found in cells exhibiting high viral expression (Figure 5A and B, CL7 ,
218 $PC1 > 28$), suggesting that these cells were at the final stages of infection. In order to further
219 characterize the link between optimal host metabolic state and efficient viral infection, we infected
220 CCMP2090 stationary culture and subjected single cells to dual gene expression analysis at 2 hpi
221 (Figure 5C and D). While most of the exponential growing cells exhibited viral expression, we
222 detected only moderate viral expression in 3/27 (11%) of the stationary phase cells (Figure 5C),
223 while the rest of the cells had viral expression patterns similar to uninfected cells (control). In
224 parallel, stationary phase cells (either control or infected) exhibited down-regulation of most of
225 the examined host metabolic genes, in contrast to their general up-regulation in infected
226 exponential phase cells (Figure 5D).

227 “Kill the Winner” is a key theory in microbial ecology which suggests that viruses shape
228 diversity of microbial populations by infecting the most dominant proliferative host [63]. We
229 propose that “Kill the Winner” may even act within isogenic populations based on the variability
230 in the metabolic state, which will lead to differential susceptibility to viral infection, forming
231 continuous host-virus co-existence [64]. It is possible that cell-to-cell heterogeneity in the
232 metabolic activity is shaped by the tradeoff between complex abiotic stress conditions (e.g. nutrient
233 availability [65-67] and light regime) and biotic interactions (e.g. pathogenicity or allelopathy),
234 and may result in differential susceptibility to viral infection in the marine environment.

235 **Conclusions**

236 The data presented here suggests detection of host and virus expression profiles on a single-
237 cell level as a novel approach to characterized host responses during viral infection in high
238 resolution which is commonly masked in whole population RNAseq approaches[68]. By applying
239 dual gene expression profiling during algal host-virus interactions, we uncovered an early host
240 transcriptional responses. This newly defined phase can result from either induction of host
241 resistance mechanism or derived from viral priming of host metabolic pathways. The new ability
242 to define distinct “infection states” on a pseudo-temporal manner can potentially provide valuable
243 information regarding the dynamics of active viral infection in “real time” also in the natural
244 environment. Clustering of individual cells based on their specific transcriptomic signatures will
245 uncover the relationship between host metabolic states and specific phenotypes associated with
246 differential levels of viral infection or modes of resistance in natural populations. *In situ*
247 quantification of the fraction of infected cells, their infection and metabolic states and the fraction
248 of resistant cells will provide important insights into the infection dynamics and may provide
249 fundamental understating of host-virus co-existence strategies in the ocean. Resolving host-virus
250 interaction on a single cell will provide novel sensitive biomarkers to assess the ecological impact
251 of marine viruses and their role in regulating the fate of algal blooms in the ocean.

252 **Methods**

253 Culture growth and viral infection dynamics

254 Cells of the non- calcified CCMP2090 and calcifying RCC1216 *E. huxleyi* strains were cultured
255 in K/2 medium [69] and incubated at 18°C with a 16:8 h light— dark illumination cycle. A light
256 intensity of 100 $\mu\text{M photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ was provided by cool white LED lights. Experiments were
257 performed with exponential phase ($5\cdot 10^5$ - $1\cdot 10^6$ cells $\cdot\text{ml}^{-1}$) or stationary phase ($5\cdot 10^6$ cells $\cdot\text{ml}^{-1}$)
258 cultures. *E. huxleyi* virus EhV201 (lytic) used for this study was isolated originally in [12]. In
259 CCMP2090 experiments, *E. huxleyi* was infected with a 1:50 volumetric ratio of viral lysate to
260 culture (multiplicity of infection (MOI) of about 1:1 infectious viral particles per cell). In
261 RCC1216 experiments,
262 *E. huxleyi* was infected with a 1:1000 volumetric ratio of viral lysate to culture (MOI of about
263 1:0.2 infectious viral particles per cell). For single-cell analysis, *E. huxleyi* cells were
264 concentrated to $2.5\cdot 10^6$ cells $\cdot\text{ml}^{-1}$ by gentle centrifugation (3000 RPM, 3 min) prior to single-
265 cell isolation. To compare between viral infection in exponential and stationary phases,
266 stationary phase cells were diluted to similar concentration of exponential phases cells using
267 stationary conditioned medium ($5\cdot 10^5$ - $1\cdot 10^6$ cells $\cdot\text{ml}^{-1}$) and then infected by EhV. The growth
268 dynamics of *E. huxleyi* CCMP2090 strain and RCC1216 strain clones were monitored in
269 seawater-based K/2 medium in control conditions and in the presence of the lytic viral strain
270 EhV201.

271

272 Enumeration of cell and virus abundance

273 Cells were monitored and quantified using a Multisizer 4 Coulter counter (Beckman Coulter,
274 Nyon, Switzerland). For extracellular viral production, samples were filtered using 0.45 μM
275 PVDF filters (Millex-HV, Millipore). Filtrate was fixed with a final concentration of 0.5%
276 glutaraldehyde for 30 min at 4°C, then plunged into liquid nitrogen, and stored at -80°C until
277 analysis. After thawing, 2:75 ratio of fixed sample was stained with SYBER gold (Invitrogen)
278 prepared in Tris–EDTA buffer as instructed by the manufacturer (5 μl SYBER gold in 50 mL
279 Tris–EDTA), then incubated for 20 min at 80°C and cooled down to room temperature. Flow
280 cytometric analysis was performed with excitation at 488 nm and emission at 525 nm.

281

282 Single-Cell Quantitative RT-PCR

283 Single cells were captured on a C1 STA microfluidic array (5–10 μm cells) using the Fluidigm
284 C1 and imaged on IX71S1F-3-5 motorized inverted Olympus microscope (Tokyo, Japan) to
285 examine chlorophyll autofluorescence (ex:500/20 nm, em:650 nm LP). Only wells that
286 exhibited chlorophyll autofluorescence signal emitted from single cells were further analyzed.
287 External RNA Controls Consortium (ERCC) spikes were added to each well in a final dilution
288 of 1:40,000. Cells were lysed and pre-amplified cDNA was generated from each cell using the
289 Single Cells-to-CT Kit (Life Technologies). Pooled qPCR primers and Fluidigm STA reagents
290 were added according to manufacturer's recommendations. Preamplified cDNA was then used
291 for high-throughput qPCR measurement of each amplicon using a BioMark HD system. Briefly,
292 a 2.7 μl aliquot of each amplified cDNA was mixed with 3 μl of 2X SsoFast EvaGreen Supermix
293 with Low
294 ROX (Bio-Rad) and 0.3 μl of 20X DNA Binding Dye Sample Loading Reagent (Fluidigm), and
295 5 μl of each sample mix was then pipetted into one sample inlet in a
296 96.96 Dynamic Array IFC chip (Fluidigm). Individual qPCR primer pairs (50 μM , Supplemental
297 Table 1) in a 1.08 μl volume were mixed with 3 μl Assay Loading Reagent (Fluidigm) and 1.92
298 μl Low TE, and 5 μl of each mix was pipetted into one assay inlet in the same Dynamic Array
299 IFC chip. Subsequent sample/assay loading was performed with an IFC Controller HX
300 (Fluidigm) and qPCR was performed on the BioMark HD real- time PCR reader (Fluidigm)
301 following manufacturer's instructions using standard fast cycling conditions and melt-curve
302 analysis, generating an amplification curve for each gene of interest in each sample (cell). Data
303 was analyzed using Real-time PCR Analysis software (Fluidigm) with the following settings:
304 0.65 curve quality threshold, linear derivative baseline correction, automatic thresholding by
305 assay (gene), and manual melt curve exclusion. Cycle threshold (Ct) values for each reaction
306 were exported. As seen in other applications of this technology[60], the data had a bimodal
307 distribution with some cells ranging from 2.5 Ct to 30 Ct, and another set of cells with Ct >40.
308 Similar bimodal distribution was also observed for the ERCC spikes. Accordingly, we set the
309 minimal threshold level of detection to 30 Ct and calculated expression threshold values (Et) by
310 linear transformation of the data so that minimal Et was zero (30 Ct). For heat map visualization,
311 expression data was normalized by subtracting the mean of each gene and dividing it with its
312 standard deviation across cells. Single-cell PCR data was analyzed and displayed using
313 MATLAB (MathWorks). Additional statistical analyses were performed using The

314 SingleCellAssay R package [60]. Calculation of number of spike molecule per Fluidigm C1 well
315 was performed according to [59].

316

317

318 **Acknowledgments:** We wish thank Dr. Daniella Schatz and Guy Schleyer from the Vardi lab, Dr.
319 Roi Avraham from the Department of Biological Regulation at the Weizmann Institute of Science
320 and Dr. Noam Stern-Ginossar from the Department of Molecular Genetics at the Weizmann
321 Institute of Science for critical comments on the manuscript and fruitful discussion. We would also
322 like to thank Tal Bigdary from the Design, Photography and Printing Branch at the Weizmann
323 Institute of Science for assistance in designing the graphs for this manuscript. **Funding:** This
324 research was supported by the European Research Council (ERC) StG (INFOTROPHIC grant no.
325 280991) and CoG (VIROCELLSPHERE grant no. 681715) awarded to A.V. **Competing**
326 **interests:** The authors declare that they have no competing interests. **Author contributions:** S.R.
327 and A.V. conceived and designed the experiments and wrote the manuscript. S.R., M.J.F and D.P
328 conducted the single-cell experiments. R.R. performed single-cell statistical analysis. S.R analyzed
329 the single-cell expression data.

330

331

332 **References:**

- 333 1. Bergh O, Borsheim KY, Bratbak G, Heldal M: High abundance of viruses found in aquatic environments.
334 *Nature* 1989, 340:467-468.
- 335 2. Suttle CA: Marine viruses - major players in the global ecosystem. *Nat Rev Micro* 2007, 5:801-812.
- 336 3. Fuhrman JA: Marine viruses and their biogeochemical and ecological effects. *Nature* 1999, 399:541-548.
- 337 4. Wilhelm SW, Suttle CA: Viruses and nutrient cycles in the sea: viruses play critical roles in the structure
338 and function of aquatic food webs. *BioScience* 1999, 49:781-788.
- 339 5. Weitz JS, Stock CA, Wilhelm SW, Bourouiba L, Coleman ML, Buchan A, Follows MJ, Fuhrman JA, Jover
340 LF, Lennon JT, et al: A multitrophic model to quantify the effects of marine viruses on microbial food
341 webs and ecosystem processes. *ISME J* 2015, 9:1352-1364.
- 342 6. Holligan PM, Fernandez E, Aiken J, Balch WM, Boyd P, Burkill PH, Finch M, Groom SB, Malin G,
343 Muller K, et al: A biogeochemical study of the coccolithophore *Emiliana huxleyi*, in the North Atlantic.
344 *Global Biogeochem Cy* 1993, 7:879-900.
- 345 7. Taylor AR, Brownlee C, Wheeler G: Coccolithophore Cell Biology: Chalking Up Progress. *Ann Rev Mar*
346 *Sci* 2017, 9:283-310.
- 347 8. Iglesias-Rodriguez D, Halloran PR, Rickaby REM, Hall IR, Colmenero-Hidalgo E, Gittins JR, Green DRH,
348 Tyrrell T, Gibbs SJ, von Dassow P, et al: Phytoplankton calcification in a high-CO₂ World. *Science* 2008,
349 320:336-340.
- 350 9. Alcolombri U, Ben-Dor S, Feldmesser E, Levin Y, Tawfik DS, Vardi A: Identification of the algal
351 dimethyl sulfide-releasing enzyme: A missing link in the marine sulfur cycle. *Science* 2015, 348:1466-
352 1469.

- 353 10. Simo R: Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and
354 evolutionary links. *Trends Ecol Evol* 2001, 16:287-294.
- 355 11. Bratbak G, Egge J, Heldal M: Viral mortality of the marine alga *Emiliana huxleyi* (Haptophyceae) and the
356 termination of the algal bloom. *Mar Ecol Prog Ser* 1993, 93:39-48.
- 357 12. Schroeder DC, Oke J, Malin G, Wilson WH: Coccolithovirus (*Phycodnaviridae*): characterisation of a new
358 large dsDNA algal virus that infects *Emiliana huxleyi*. *Arch Virol* 2002, 147:1685-1698.
- 359 13. Read BA, Kegel J, Klute MJ, Kuo A, Lefebvre SC, Maumus F, Mayer C, Miller J, Monier A, Salamov A,
360 et al: Pan genome of the phytoplankton *Emiliana* underpins its global distribution. *Nature* 2013, 499:209-
361 213.
- 362 14. Vardi A, Haramaty L, Van Mooy BA, Fredricks HF, Kimmance SA, Larsen A, Bidle KD: Host-virus
363 dynamics and subcellular controls of cell fate in a natural coccolithophore population. *Proc Natl Acad Sci*
364 *USA* 2012, 109:19327-19332.
- 365 15. Bidle KD, Vardi A: A chemical arms race at sea mediates algal host-virus interactions. *Curr Opin*
366 *Microbiol* 2011, 14:449-457.
- 367 16. Wilson WH, Schroeder DC, Allen MJ, Holden MTG, Parkhill J, Barrell BG, Churcher C, Hamlin N,
368 Mungall K, Norbertczak H, et al: Complete Genome Sequence and Lytic Phase Transcription Profile of a
369 *Coccolithovirus*. *Science* 2005, 309:1090-1092.
- 370 17. Feldmesser E, Rosenwasser S, Vardi A, Ben-Dor S: Improving transcriptome construction in non-model
371 organisms: integrating manual and automated gene definition in *Emiliana huxleyi*. *BMC Genomics* 2014,
372 15:148-163.
- 373 18. Rosenwasser S, Mausz MA, Schatz D, Sheyn U, Malitsky S, Aharoni A, Weinstock E, Tzfadia O, Ben-Dor
374 S, Feldmesser E, et al: Rewiring host lipid metabolism by large viruses determines the fate of *emiliana*
375 *huxleyi*, a bloom-forming alga in the ocean. *Plant Cell* 2014, 26:2689-2707.
- 376 19. Zhang X, Gamarra J, Castro S, Carrasco E, Hernandez A, Mock T, Hadaegh AR, Read BA:
377 Characterization of the small RNA transcriptome of the marine coccolithophorid, *Emiliana huxleyi*. *PLOS*
378 *ONE* 2016, 11:e0154279.
- 379 20. Rosenwasser S, Ziv C, Creveld SGv, Vardi A: Virocell Metabolism: Metabolic Innovations During Host-
380 Virus Interactions in the Ocean. *Trends Microbiol* 2016, 24:821-832.
- 381 21. Ankrah NYD, May AL, Middleton JL, Jones DR, Hadden MK, Gooding JR, LeClerc GR, Wilhelm SW,
382 Campagna SR, Buchan A: Phage infection of an environmentally relevant marine bacterium alters host
383 metabolism and lysate composition. *ISME J* 2014, 8:1089-1100.
- 384 22. Enav H, Mandel-Gutfreund Y, Beja O: Comparative metagenomic analyses reveal viral-induced shifts of
385 host metabolism towards nucleotide biosynthesis. *Microbiome* 2014, 2:9.
- 386 23. Hurwitz BL, Hallam SJ, Sullivan MB: Metabolic reprogramming by viruses in the sunlit and dark ocean.
387 *Genome Biol* 2013, 14:R123.
- 388 24. De Smet J, Zimmermann M, Kogadeeva M, Ceysens P-J, Vermaelen W, Blasdel B, Bin Jang H, Sauer U,
389 Lavigne R: High coverage metabolomics analysis reveals phage-specific alterations to *Pseudomonas*
390 *aeruginosa* physiology during infection. *ISME J* 2016.
- 391 25. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, Chisholm SW: Phage auxiliary
392 metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci USA*
393 2011, 108:E757-764.
- 394 26. Malitsky S, Ziv C, Rosenwasser S, Zheng S, Schatz D, Porat Z, Ben-Dor S, Aharoni A, Vardi A: Viral
395 infection of the marine alga *Emiliana huxleyi* triggers lipidome remodeling and induces the production of
396 highly saturated triacylglycerol. *New Phytol* 2016, 210:88-96.
- 397 27. Lehahn Y, Koren I, Schatz D, Frada M, Sheyn U, Boss E, Efrati S, Rudich Y, Trainic M, Sharoni S, et al:
398 Decoupling Physical from Biological Processes to Assess the Impact of Viruses on a Mesoscale Algal
399 Bloom. *Curr Biol* 2014, 24:2041-2046.
- 400 28. Mackinder LC, Worthy CA, Biggi G, Hall M, Ryan KP, Varsani A, Harper GM, Wilson WH, Brownlee C,
401 Schroeder DC: A unicellular algal virus, *Emiliana huxleyi* virus 86, exploits an animal-like infection
402 strategy. *J Gen Virol* 2009, 90:2306-2316.
- 403 29. Ziv C, Malitsky S, Othman A, Ben-Dor S, Wei Y, Zheng S, Aharoni A, Hornemann T, Vardi A: Viral
404 serine palmitoyltransferase induces metabolic switch in sphingolipid biosynthesis and is required for
405 infection of a marine alga. *Proc Natl Acad Sci USA* 2016, 113: E1907-E1916.
- 406 30. Vardi A, Van Mooy BA, Fredricks HF, Pependorf KJ, Ossolinski JE, Haramaty L, Bidle KD: Viral
407 glycosphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* 2009, 326:861-
408 865.

- 409 31. Bidle KD, Haramaty L, Barcelos e Ramos J, Falkowski P: Viral activation and recruitment of metacaspases
410 in the unicellular coccolithophore, *Emiliania huxleyi*. *Proc Natl Acad Sci USA* 2007, 104:6049-6054.
- 411 32. Sheyn U, Rosenwasser S, Ben-Dor S, Porat Z, Vardi A: Modulation of host ROS metabolism is essential
412 for viral infection of a bloom forming coccolithophore in the ocean. *ISME J* 2016, 10:1742–1754.
- 413 33. Schatz D, Shemi A, Rosenwasser S, Sabanay H, Wolf SG, Ben-Dor S, Vardi A: Hijacking of an autophagy-
414 like process is critical for the life cycle of a DNA virus infecting oceanic algal blooms. *New Phytol* 2014,
415 204:854–863.
- 416 34. Cheng Y-S, Labavitch J, VanderGheynst JS: Organic and Inorganic Nitrogen Impact *Chlorella variabilis*
417 Productivity and Host Quality for Viral Production and Cell Lysis. *Appl Biochem Biotechnol* 2015,
418 176:467-479.
- 419 35. Colson P, De Lamballerie X, Yutin N, Asgari S, Bigot Y, Bideshi DK, Cheng X-W, Federici BA, Van
420 Etten JL, Koonin EV, et al: “Megavirales”, a proposed new order for eukaryotic nucleocytoplasmic large
421 DNA viruses. *Arch Virol* 2013, 158:2517-2521.
- 422 36. Forterre P: To be or not to be alive: How recent discoveries challenge the traditional definitions of viruses
423 and life. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of*
424 *Biological and Biomedical Sciences* 2016, 59:100-108.
- 425 37. Frada M, Probert I, Allen MJ, Wilson WH, de Vargas C: The “Cheshire Cat” escape strategy of the
426 coccolithophore *Emiliania huxleyi* in response to viral infection. *Proc Natl Acad Sci USA* 2008, 105:15944-
427 15949.
- 428 38. Kashtan N, Roggensack SE, Rodrigue S, Thompson JW, Biller SJ, Coe A, Ding H, Martinen P,
429 Malmstrom RR, Stocker R, et al: Single-Cell Genomics Reveals Hundreds of Coexisting Subpopulations in
430 Wild *Prochlorococcus*. *Science* 2014, 344:416-420.
- 431 39. Yoon HS, Price DC, Stepanauskas R, Rajah VD, Sieracki ME, Wilson WH, Yang EC, Duffy S,
432 Bhattacharya D: Single-Cell Genomics Reveals Organismal Interactions in Uncultivated Marine Protists.
433 *Science* 2011, 332:714-717.
- 434 40. Liu S, Trapnell C: *Single-cell transcriptome sequencing: recent advances and remaining challenges*
435 *[version 1; referees: 2 approved]*. F1000Research; 2016.
- 436 41. Guillaume-Gentil O, Grindberg RV, Kooger R, Dorwling-Carter L, Martinez V, Ossola D, Pilhofer M,
437 Zambelli T, Vorholt JA: Tunable Single-Cell Extraction for Molecular Analyses. *Cell*, 166:506-516.
- 438 42. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, et al:
439 mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Meth* 2009, 6:377-382.
- 440 43. Hashimshony T, Wagner F, Sher N, Yanai I: CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear
441 Amplification. *Cell Reports* 2012, 2:666-673.
- 442 44. Grun D, Lyubimova A, Kester L, Wiebrands K, Basak O, Sasaki N, Clevers H, van Oudenaarden A:
443 Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 2015, 525:251-255.
- 444 45. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, La Manno G, Juréus A, Marques S, Munguba
445 H, He L, Betsholtz C, et al: Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-
446 seq. *Science* 2015, 347:1138-1142.
- 447 46. Efroni I, Mello A, Nawy T, Ip P-L, Rahni R, DelRose N, Powers A, Satija R, Birnbaum Kenneth D: Root
448 Regeneration Triggers an Embryo-like Sequence Guided by Hormonal Interactions. *Cell* 2016, 165:1721-
449 1733.
- 450 47. Avraham R, Haseley N, Brown D, Penaranda C, Jijon HB, Trombetta JJ, Satija R, Shalek AK, Xavier RJ,
451 Regev A, Hung DT: Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses.
452 *Cell*, 163:523.
- 453 48. Shalek AK, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, Chen P, Gertner RS, Gaublomme JT, Yosef
454 N, et al: Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 2014,
455 510:363-369.
- 456 49. Saliba A-E, Li L, Westermann AJ, Appenzeller S, Stapels DAC, Schulte LN, Helaine S, Vogel J: Single-
457 cell RNA-seq ties macrophage polarization to growth rate of intracellular *Salmonella*. *Nature Microbiol*
458 2016, 2:16206.
- 459 50. Patil S, Fribourg M, Ge Y, Batish M, Tyagi S, Hayot F, Sealfon SC: Single-cell analysis shows that
460 paracrine signaling by first responder cells shapes the interferon- β response to viral infection. *Sci Signal*
461 2015, 8:ra16-ra16.
- 462 51. Snijder B, Sacher R, Ramo P, Damm E-M, Liberali P, Pelkmans L: Population context determines cell-to-
463 cell variability in endocytosis and virus infection. *Nature* 2009, 461:520-523.

- 464 52. Heldt FS, Kupke SY, Dorl S, Reichl U, Frensing T: Single-cell analysis and stochastic modelling unveil
465 large cell-to-cell variability in influenza A virus infection. *Nat Comm* 2015, 6:8938.
- 466 53. Cohen EM, Kobiler O: Gene Expression Correlates with the Number of Herpes Viral Genomes Initiating
467 Infection in Single Cells. *PLOS Pathog* 2016, 12:e1006082.
- 468 54. Westermann AJ, Gorski SA, Vogel J: Dual RNA-seq of pathogen and host. *Nat Rev Micro* 2012, 10:618-
469 630.
- 470 55. Rosani U, Varotto L, Domeneghetti S, Arcangeli G, Pallavicini A, Venier P: Dual analysis of host and
471 pathogen transcriptomes in ostreid herpesvirus 1-positive *Crassostrea gigas*. *Environ Microbiol* 2015,
472 17:4200-4212.
- 473 56. Nuss AM, Beckstette M, Pimenova M, Schmöhl C, Opitz W, Pisano F, Heroven AK, Dersch P: Tissue dual
474 RNA-seq allows fast discovery of infection-specific functions and riboregulators shaping host–pathogen
475 transcriptomes. *Proc Natl Acad Sci USA* 2017, 114:E791-E800.
- 476 57. Allen MJ, Forster T, Schroeder DC, Hall M, Roy D, Ghazal P, Wilson WH: Locus-Specific Gene
477 Expression Pattern Suggests a Unique Propagation Strategy for a Giant Algal Virus. *J Virol* 2006, 80:7699-
478 7705.
- 479 58. Frada MJ, Rosenwasser S, Ben-Dor S, Shemi A, Sabanay H, Vardi A: Morphological switch and
480 development of a resistant subpopulation in response to viral infection in a bloom-forming marine
481 microalgae. 2017.
- 482 59. Wu AR, Neff NF, Kalisky T, Dalerba P, Treutlein B, Rothenberg ME, Mburu FM, Mantalas GL, Sim S,
483 Clarke MF, Quake SR: Quantitative assessment of single-cell RNA-sequencing methods. *Nat Meth* 2014,
484 11:41-46.
- 485 60. McDavid A, Finak G, Chattopadhyay PK, Dominguez M, Lamoreaux L, Ma SS, Roederer M, Gottardo R:
486 Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments.
487 *Bioinformatics* 2013, 29:461-467.
- 488 61. Torres MA, Dangl JL, Jones JDG: Arabidopsis gp91phox homologues AtrbohD and AtrbohF are required
489 for accumulation of reactive oxygen intermediates in the plant defense response. *Proc Natl Acad Sci USA*
490 2002, 99:517-522.
- 491 62. Evans C, Malin G, Wilson WH, Liss PS: Infectious titres of *Emiliana huxleyi* virus 86 are reduced by
492 exposure to millimolar dimethyl sulfide and acrylic acid. *Limnol Oceanog* 2006, 51:2468-2471.
- 493 63. Thingstad TF: Elements of a theory for the mechanisms controlling abundance, diversity, and
494 biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanog* 2000, 45:1320-1328.
- 495 64. Thyraug R, Larsen A, Thingstad FT, Bratbak G: Stable coexistence in marine algal host-virus systems.
496 *Mar Ecol Prog Ser* 2003, 254:27–35.
- 497 65. Schreiber F, Littmann S, Lavik G, Escrig S, Meibom A, Kuypers MMM, Ackermann M: Phenotypic
498 heterogeneity driven by nutrient limitation promotes growth in fluctuating environments. *Nat Microbiol*
499 2016, 1:16055.
- 500 66. Martínez JM, Schroeder DC, Larsen A, Bratbak G, Wilson WH: Molecular Dynamics of *Emiliana huxleyi*
501 and Cooccurring Viruses during Two Separate Mesocosm Studies. *Appl Environ Microbiol* 2007, 73:554-
502 562.
- 503 67. Maat DS, Crawford KJ, Timmermans KR, Brussaard CPD: Elevated CO₂ and Phosphate Limitation Favor
504 *Micromonas pusilla* through Stimulated Growth and Reduced Viral Impact. *Appl Environ Microbiol* 2014,
505 80:3119-3127.
- 506 68. Westermann AJ, Barquist L, Vogel J: Resolving host–pathogen interactions by dual RNA-seq. *PLOS*
507 *Pathogens* 2017, 13:e1006033.
- 508 69. Keller MD, Selvin RC, Claus W, Guillard RRL: Media for the culture of oceanic ultraphytoplankton. *J*
509 *Phycol* 1987, 23:633-638.