



**HAL**  
open science

## **Life cycle and morphogenetic differentiation in heteromorphic cell types of a cosmopolitan marine microalga**

Laurie Bousquet, Shai Fainsod, Johan Decelle, Omer Murik, Fabien Chevalier, Benoit Gallet, Rachel Templin, Yannick Schwab, Yoav Avrahami, Gil Koplovitz, et al.

### ► **To cite this version:**

Laurie Bousquet, Shai Fainsod, Johan Decelle, Omer Murik, Fabien Chevalier, et al.. Life cycle and morphogenetic differentiation in heteromorphic cell types of a cosmopolitan marine microalga. 2024. <hal-04789385>

**HAL Id: hal-04789385**

**<https://hal.science/hal-04789385v1>**

Preprint submitted on 18 Nov 2024

**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 CC BY-ND 4.0 - Attribution - No Derivative Works - International License

1 **Life Cycle and Morphogenetic Differentiation in Heteromorphic Cell Types of a**  
2 **Cosmopolitan Marine Microalga**

3

4 Laurie Bousquet<sup>1,2</sup>, Shai Fainsod<sup>1,2,°</sup>, Johan Decelle<sup>3</sup>, Omer Murik<sup>4</sup>, Fabien Chevalier<sup>3</sup>, Benoit  
5 Gallet<sup>3</sup>, Rachel Templin<sup>3</sup>, Yannick Schwab<sup>3</sup>, Yoav Avrahami<sup>1,2</sup>, Gil Koplovitz<sup>1,2</sup>, Chuan Ku<sup>5</sup>,  
6 Miguel J. Frada<sup>1,2</sup>

7

8 <sup>1</sup>Department of Ecology, Evolution and Behavior, Silberman Institute of Life Sciences, The  
9 Hebrew University of Jerusalem, Jerusalem, Israel; <sup>2</sup>The Interuniversity Institute for Marine  
10 Sciences in Eilat, Eilat, Israel; <sup>3</sup>Laboratoire Physiologie Cellulaire et Végétale, Univ.  
11 Grenoble Alpes, CNRS, CEA, INRAE, IRIG-DBSCI-LPCV, 38000 Grenoble, France;  
12 <sup>4</sup>Translation Genomics Lab and Medical Genetics Institute, Shaare Zedek Medical Center,  
13 93722, Jerusalem, Israel; <sup>5</sup>Institute of Plant and Microbial Biology, Academia Sinica, Taipei,  
14 Taiwan.

15 <sup>°</sup>Current address: Faculty of Biology, Technion – Israel Institute of Technology, Haifa  
16 3200003, Israel.

17 Author for correspondence

18 Miguel J. Frada

19 Email: [miguel.frada@mail.huji.ac.il](mailto:miguel.frada@mail.huji.ac.il)

20

21

## 22 **Summary**

- 23 • *Gephyrocapsa huxleyi* is a prevalent, bloom-forming phytoplankton species in the  
24 oceans. It exhibits a complex haplo-diplontic life cycle, featuring a diploid-calcified  
25 phase, a haploid phase, and a third 'decoupled' phase produced during viral  
26 infection. Decoupled cells display a haploid-like phenotype, but are diploid.
- 27 • Here, we investigated the fate of decoupled cells during culture observations and we  
28 compared the transcriptome profiling and the cellular architecture in three dimensions  
29 of the three cell types.
- 30 • We found that decoupled cells can revert to the calcified form in the absence of viral  
31 pressure, revealing the transient nature of this cell type. Ultrastructural analyses  
32 showed distinct nuclear organisation with variations in chromatin volume.  
33 Transcriptomic analyses revealed gene expression patterns specific to each life phase.  
34 These included multiple regulatory functions in chromatin remodelling, broader  
35 epigenetic mechanisms and life cycling, which likely contributed to cell differentiation.  
36 Finally, the exploration of available host-virus transcriptomes supports life cycle  
37 transition during viral infection.
- 38 • This study provides cellular and molecular foundations for nuclear remodelling and  
39 cell differentiation in coccolithophores and the identification of gene markers for  
40 studying coccolithophore life cycles in natural populations.

41 **Key words:** Cell differentiation, Coccolithophore, Chromatin organization, Epigenetic  
42 regulation, Life cycle, Transcriptome, Viral infection.

43

## 44 **Introduction**

45 Eukaryotes exhibit complex life cycles comprising multiple phases interlinked by  
46 sexual mechanisms (syngamy and meiosis). They involve gene recombination,  
47 variations in ploidy levels (haploid, 1n and diploid, 2n), as well as asexual  
48 mechanisms of cell differentiation. These transitions are crucial for the  
49 developmental and reproductive cycles of organisms, as well as for the adaptation to  
50 environmental changes through the differentiation of physiologically fit or resistant  
51 phenotypes (Mable & Otto, 1998; Von Dassow & Montresor, 2011).

52 Coccolithophores are a prevalent group of marine unicellular algae that produce  
53 distinct CaCO<sub>3</sub> plates, coccoliths, coating the cell surface (Monteiro *et al.*, 2016).  
54 Due to photosynthesis and calcification, these microorganisms play a pivotal role in  
55 the global carbon cycle (Balch, 2018). Coccolithophores have complex sexual  
56 haplodiplontic life cycles, composed of 1n- and calcified 2n-cell phases, each  
57 capable to independent cell division and expressing dissimilar morphology  
58 (heteromorphic). Such life cycles are proposed to widen niche-space and facilitate  
59 the adaptation to environmental variability (Houdan *et al.*, 2004; Frada *et al.*, 2019;  
60 de Vries *et al.*, 2021).

61 *Gephyrocapsa huxleyi* (formerly *Emiliana huxleyi*, (Bendif *et al.*, 2023)), is the most  
62 prevalent coccolithophore in the ocean and forms extensive seasonal blooms in high  
63 latitude regions, impacting marine ecology and biogeochemistry (Balch, 2018). The  
64 haplodiplontic, heteromorphic life cycle of *G. huxleyi* features diploid calcified cells  
65 (2n-calcified) that are non-motile and haploid non-calcified, biflagellated cells, coated  
66 by simpler organic scales (1n-flagellated) (Klaveness, 1972; Green *et al.*, 1996).  
67 Although 1n cells have been detected in 2n cultures, controlled manipulation of  
68 sexual transitions (syngamy or meiosis) has not yet been achieved. Transcriptome  
69 analyses have identified vast differences between diploid and haploid cells,  
70 particularly in transcripts linked to calcification, membrane trafficking and motility  
71 genes expressed in 2n-calcified cells and 1n-flagellated cells, respectively (von  
72 Dassow *et al.*, 2009; Rokitta *et al.*, 2011; Rokitta & Rost, 2012). However, despite  
73 extensive research on the ecological and biogeochemical roles of the 2n-calcified  
74 cells, the ecology of 1n-flagellated cells remains poorly understood (Frada *et al.*,  
75 2012).

76 A major influence on *G. huxleyi* blooms are specific lytic viruses (*Emiliania huxleyi*-  
77 specific viruses, EhV) that infect diploid cells and drive bloom decay (e.g. Bratbak *et al.*,  
78 1993; Vardi *et al.*, 2012). Strikingly, culture experiments have demonstrated that  
79 1n cells are resistant to EhV (Frada *et al.*, 2008). In addition, during infection of 2n-  
80 calcified cells, cells displaying a haploid-like phenotype (biflagellated, organic scale-  
81 bearing) and resistant to EhV are produced during viral infections (Frada *et al.* 2017).  
82 These cells are however diploid, thus referred to as the 'decoupled' cell type (2n-  
83 flagellated) (Frada *et al.*, 2017), whose production resembles apospory reported in  
84 plants and algae, where diploid gametophytes (typically haploid) develop without  
85 meiosis (Coelho *et al.*, 2007). Such differentiation of 2n-flagellated cells may confer  
86 an escape strategy from EhV, potentially setting the stage for post-bloom recovery  
87 (Frada *et al.*, 2017).

88 Life cycle transitions typically involve layers of mechanisms, notably epigenetic  
89 processes. These include chemical modifications of chromatin via the addition or  
90 removal of chemical moieties on histone tails (histone post-translational  
91 modifications, hPTMs) driving chromatin ultrastructural remodelling and gene  
92 expression regulation (Lawrence *et al.*, 2016). Some hPTMs and associated proteins  
93 are implicated in life cycle regulation. Notably, the Polycomb Repressive Complex 2  
94 (PRC2) mediates the silencing of key three-amino acids loop extension (TALE)  
95 homeodomain transcription factors (KNOX- and BELL-like) that in turn, regulate  
96 haploid-to-diploid transitions in land plants (Sakakibara *et al.*, 2013; Horst *et al.*,  
97 2016; Dierschke *et al.*, 2021) and in multiple algae ( e.g., Lee *et al.*, 2008; Thangavel  
98 & Nayar, 2018; Arun *et al.*, 2019; Mikami *et al.*, 2019; Hirooka *et al.*, 2022).

99 Specificities in these mechanisms are associated to different species. However,  
100 alterations of these life cycle regulators can result in the decoupling of ploidy and life  
101 phase generation (Schmidt, 2020), underlying either apospory (defined above) or  
102 apogamy where instead haploid sporophytes (typically diploid) develop without  
103 gamete fusion (e.g., Zhao *et al.*, 2001; Sakakibara *et al.*, 2013; Coelho *et al.*)

104 Here, we surveyed the fate of 2n-flagellated (decoupled) cells in culture of the  
105 cosmopolitan microalga *G. huxleyi*. Moreover, we examined their relation to 2n-  
106 calcified and 1n-flagellated forms, from morphological and gene expression  
107 perspectives using 3D-microscopy (FIB-SEM, Focused ion beam scanning electron  
108 microscopy) and transcriptome analyses. Finally, we also probed published *G.*

109 *huxleyi*-EhV transcriptomes to examine 2n-flagellated cell differentiation during viral  
110 infection. Our results provide novel insights on the regulatory mechanisms driving life  
111 cycle differentiation and response to viruses in *G. huxleyi*.

112

## 113 **Material and Methods**

114 Algal strains, maintenance and viral infections:

115 We used *Gephyrocapsa huxleyi* strains RCC1216 (2n-calcified, Houdan *et al.*, 2005),  
116 LC5-11A (1n-flagellated), and LC5-12A (2n-flagellated) in this study. LC5-11A was  
117 isolated from RCC1216 via single-cell pipetting following a spontaneous life cycle  
118 change in culture, 6 months prior the study. LC5-12A and other 2n-flagellated strains  
119 were isolated from RCC1216 during an EhV 201 infection (Schroeder *et al.*, 2002) 4  
120 to 6 months prior the study, as described in Frada *et al.*, 2017. Cultures, maintained  
121 in K/2 (-Tris, -Si) medium (Keller *et al.*, 1987), were not axenic. Genome sizes were  
122 validated by flow cytometry of isolated nuclei (von Dassow *et al.*, 2009). Ten  
123 additional decoupled strains were isolated from 3 independent RCC1216 infection  
124 experiments with EhV 201. Over 2 years, routine inspection of decoupled cultures  
125 was performed using optical microscopy. Sensitivity to EhV201 infection was  
126 assessed during the exponential growth phase with a multiplicity of infection of 0.2.  
127 Cell and viral concentrations were determined by flow cytometry (Attune NxT Life  
128 Technology).

129 Algal growth for transcriptome analyses:

130 The 3 strains were grown in triplicate in 4L borosilicate vessels, under gentle air-  
131 bubbling at 18°C, illuminated with 80  $\mu\text{mol photons m}^{-2}\text{s}^{-1}$  from warm-white LED  
132 panels on a 14:10h light/dark cycle. Cultures were initiated at  $5-8 \times 10^4$  cells  $\text{mL}^{-1}$   
133 and monitored for 48h by flow cytometry. The maximum photochemical quantum  
134 yields of PSII (Fv/Fm), were measured by fluorescence induction and relaxation  
135 system (Satlantic Inc., Halifax, NS, Canada; Gorbunov and Falkowski, 2005) (Table  
136 S1). Four hours after dawn, 200 mL of each culture was harvested and pelleted by  
137 centrifugation (8000 g, 10 min, 4°C). Cell pellets were flash-frozen in liquid N<sub>2</sub> and  
138 stored at -80°C until RNA extraction.

139 RNA isolation, libraries preparation and sequencing:

140 RNA was extracted from cell pellets using the RNeasy Plant Kit (Qiagen) and treated  
141 with DNase (Turbo DNA-free kit, Ambion). RNA concentration and integrity were  
142 measured by Nanodrop (Thermo Fisher Scientific) and TapeStation (Agilent  
143 Genomics). cDNA libraries were prepared according to the Illumina RNA sample  
144 preparation kit. mRNA was isolated using poly(T) oligo-attached magnetic beads,  
145 fragmented into 200-500 bp pieces and RNA fragments were reverse transcribed  
146 with SuperScript II reverse transcriptase (Life Technology). Adapters ligation,  
147 amplification and purifications preceded the sequencing of the cDNA libraries on an  
148 Illumina NextSeq platform into 75-bp single-end reads.

149 Transcriptome curation, assembly and annotation:

150 FastQC ([www.bioinformatics.babraham.ac.uk](http://www.bioinformatics.babraham.ac.uk)) and Trimmomatic (Bolger *et al.*, 2014)  
151 were used for quality assessment and adapter trimming. Reads from the 9 *G. huxleyi*  
152 libraries (triplicates for RCC1216, LC5-10A, LC5-12A) were aligned to the *G. huxleyi*  
153 RCC1217 genome (Kao *et al.*, 2024), a haploid strain isolated from RCC1216.  
154 Alignment was performed with STAR v2.7.10a (Dobin *et al.*, 2013). Transcripts were  
155 reconstructed and merged into a synthetic life phase transcriptome using StringTie  
156 v2.2.1 (Pertea *et al.*, 2015), Cuffmerge v2.2.1 (Trapnell *et al.*, 2012), and gffread  
157 v0.12.7 (Pertea & Pertea, 2020), yielding 60930 transcripts. Transcriptome  
158 completeness was evaluated with BUSCO v5.2.2 (Manni *et al.*, 2021) against the  
159 eukaryota\_odb10 database. Redundancy was reduced with CD-HIT-EST v4.8.1 (Fu  
160 *et al.*, 2012), resulting in 56092 non-redundant transcripts. Coding sequences were  
161 identified with TransDecoder v5.5.0. Functional annotation combined InterProScan  
162 v5.52-86.0, Diamond blastP v2.0.14.152 against the nr database and EggNOG  
163 mapper (Huerta-Cepas *et al.*, 2019; Blum *et al.*, 2021; Buchfink *et al.*, 2021). Results  
164 were consolidated in OmicsBox, and Gene Ontology annotations were retrieved  
165 using Blast2GO (Götz *et al.*, 2008). Parameters for each software are detailed in  
166 Methods S1. Life phase transcriptome information is provided in Methods S2.

167 Histones and chromatin associated transcripts were examined using two  
168 complementary strategies: (1) identification of close homologs through Blast  
169 similarity searches against local databases, and (2) detection of conserved protein  
170 domains with HMMER v3.3.2. Only shared hits were retained for subsequent  
171 analyses. Databases were assembled from Thiriet-Rupert *et al.*, 2016; Grau-Bové *et*

172 *al.*, 2022 and <https://plantfdb.gao-lab.org>. Details are provided in Methods S3.

173 Transcriptome Quantification, Differential Expression and Clustering:

174 RSEM v1.3.3 (Li & Dewey, 2011) with the parameter `--bowtie2` was used for gene-  
175 level abundance estimation. Features with fewer than 3 raw counts in at least 3  
176 samples were filtered out. Transcripts per Million (TPM) normalization was applied,  
177 and features with a mean expression  $> 0.01$  TPM were considered expressed.  
178 Differentially expressed transcripts were identified using DESeq2 (Love *et al.*, 2014),  
179 with those showing a log2fold change  $> |2|$  and FDR  $< 0.05$  clustered using K-  
180 means. The optimal number of clusters was determined through complementary  
181 statistical methods and visual inspection for homogeneity and separation.

182 Enrichment analysis was performed in R (v4.2.2; R Core Team 2021) using the  
183 hypergeometric (HG) test. Null p-values were adjusted by adding an epsilon value  
184 (smallest non-zero p-value divided by 1000) and applying the Benjamini-Hochberg  
185 (BH) correction (Benjamini & Hochberg, 1995). Data analysis and visualization were  
186 conducted in R, Unix, Python. Software parameter settings are detailed in Methods  
187 S1. Validation of differential expression between cell types was performed by qRT-  
188 PCR using various life phase-specific genes (Methods S4, Table S2).

189 Analyses of available *G. huxleyi* transcriptomes during viral infection:

190 Data from studies on the transcriptomic dynamics of *G. huxleyi* during viral infection  
191 (Feldmesser *et al.*, 2014; Rosenwasser *et al.*, 2014) were used to examine the  
192 expression patterns of specific life phase transcripts. *G. huxleyi* CCMP2090 was  
193 infected with lytic viral strain EhV 201 at a multiplicity of infection of  $\sim 1:1$ . Samples  
194 were collected at 1h and 24h post-infection, with non-infected cultures as controls,  
195 as indicated in the original studies. In the present study, life phase transcripts were  
196 blasted against CCMP2090 transcripts to identify homologs, retaining those with  
197 100% identities over 80% alignment length. We focused on "flagellated-specific"  
198 transcripts (expressed exclusively in flagellated forms - haploid and decoupled cells).  
199 Abundance estimations from CCMP2090 were used to track the expression of these  
200 transcripts over the infection timeline (see Table S3 for functional annotation and  
201 clustering information of CCMP2090 homologs to flagellated transcripts).

202

203 Focused Ion Beam-Scanning Electron Microscopy (FIB-SEM): sample preparation  
204 and analysis:

205 *G. huxleyi* life stages were maintained under similar conditions as described above  
206 and harvested during exponential growth. Cells were centrifuged at 3000 g for 2 min,  
207 cryo-fixed using high-pressure freezing (HPM100, Leica), and freeze-substituted (EM  
208 ASF2, Leica) as in Decelle *et al.*, 2019; Uwizeye *et al.*, 2021 and Gallet *et al.*, 2024.  
209 Samples were embedded in 100% resin, mounted on SEM stubs, gold sputter-  
210 coated (Quorum Q150RS; 180 s at 30 mA), and imaged with a FIB-SEM  
211 (Crossbeam 540, Carl Zeiss). Atlas3D software (Fibics Inc. and Carl Zeiss) was used  
212 for sample preparation and 3D acquisitions, with SEM images recorded at an 8 nm  
213 pixel size. Segmentation of organelles (plastids, mitochondria, nucleus) and cellular  
214 compartments was performed using 3D Slicer software (Oyama *et al.*, 2014) using a  
215 manually-curated, semi-automatic pixel clustering mode (Decelle *et al.*, 2021;  
216 Uwizeye *et al.*, 2021). Colors were assigned to segmented regions, and image  
217 intensity thresholds were adjusted. Morphometric analyses were conducted with the  
218 "segmentStatistics" module in 3D Slicer, converting values to  $\mu\text{m}^3$  or  $\mu\text{m}^2$  based on  
219 an 8 nm voxel size. Five diploid-calcified cells, five haploid cells, six 2n-flagellated  
220 cells (non-calcified), and four reverted calcifying cells were analyzed. Further  
221 methodological details are provided in Methods S5.

222

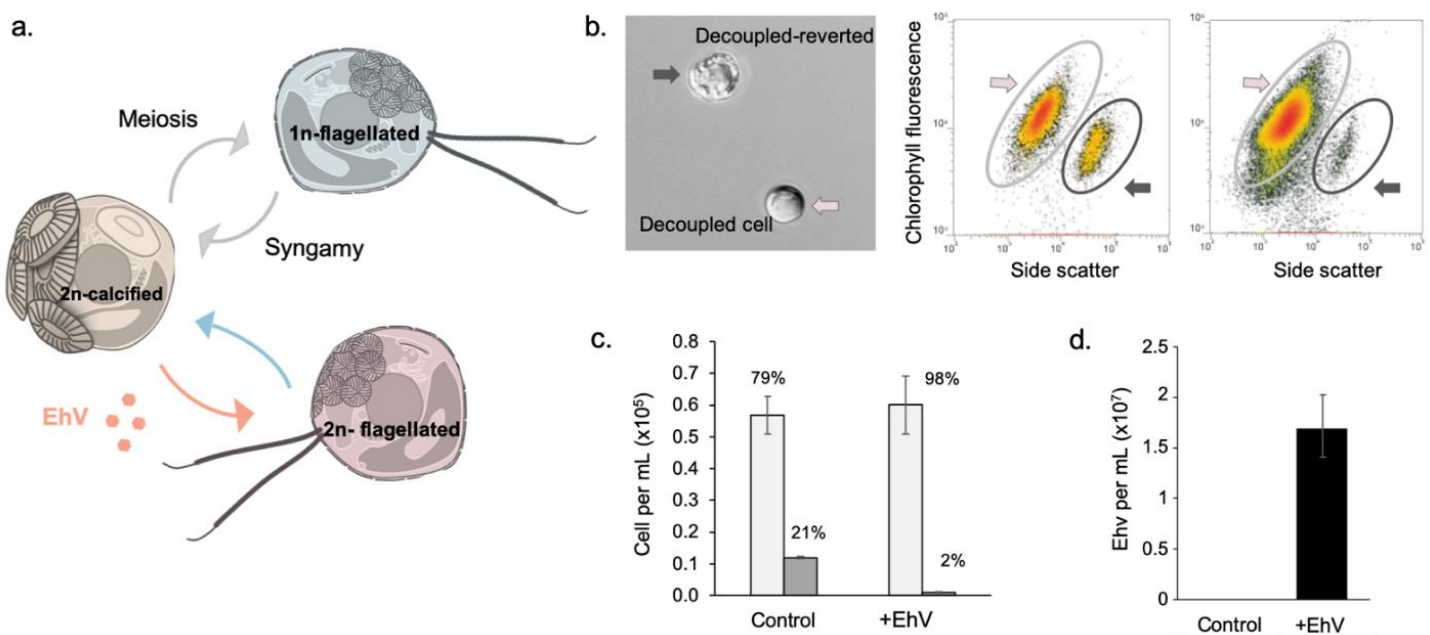
## 223 **Results**

224 The life cycle of *G. huxleyi* includes three cell types representing distinct phases: a  
225 2n-calcified (RCC1216), 1n-flagellated (LC5-10A) and 2n-flagellated (LC4-12A) cell  
226 type (Fig. 1a). Nuclear DNA content analyses by flow cytometry confirmed the ploidy  
227 of the three cell-types (Fig. S1a). The fate of decoupled cells and the morphogenetic  
228 differentiation between the three life cycle cell types was examined in our study.

229 Life cycle reversal: detection of calcified cells in 2n-flagellated cultures:

230 A total of eleven 2n-flagellated (decoupled) strains (LC4-12A-G and LC5-12A-D)  
231 were isolated during EhV -infections of RCC1216. Between 9 to 12 months of  
232 isolation, calcified cells were detected by microscopy in all 2n-flagellated cultures. As  
233 example, in LC4-12A, the fraction of calcified cells accounted for 21% of the total

234 cells as determined by flow cytometry (Fig. 1b, c). The new calcified cells are  
235 referred to as reverted cells. Subsequent addition of the virus EhV 201 to all 2n-  
236 flagellated cultures lead to the selective removal of reverted cells by lysis, but not 2n-  
237 flagellated cells, and resulted in viral production (Fig. 1b, c and d). The decoupled  
238 strain LC4-12A was randomly used in ensuing detailed 3D-microscopy and  
239 transcriptome analyses to further examine the characteristics of the 2n-flagellated life  
240 phase.



242 **Figure 1** *Gephyrocapsa huxleyi* life cycle, reversion and response to viral infection. (a)  
243 Schematic illustration of the life cycle of *G. huxleyi* and role of virus on life cycle transitions.  
244 Sexual reproduction (meiosis and syngamy) drives the alternation between haploid,  
245 uncalcified, biflagellated cells (1n-flagellated, grey) and diploid calcified cells (2n-calcified,  
246 orange). During infection by specific *Emiliana huxleyi* viruses (EhV), 2n-calcified cells can  
247 produce 2n-flagellated cells (red) that are phenotypically similar to 1n-flagellated cells. This  
248 phenotype is reversible in absence of viral pressure (see results). (b) Microscopy and flow  
249 cytometry depicting the emergence of calcified cells in 2n-decoupled cultures in the absence  
250 of EhV. Left panel: light microscopy images of a 2n-flagellated (light grey arrow) and  
251 reverted (dark grey arrow) cells, co-isolated in a clonal culture of 2n-flagellated cells in  
252 absence of viral pressure. Central and right panels: Gating of uncalcified, bi-flagellated (light  
253 grey ellipses) and calcified populations (reverted cells, dark grey ellipses) (chlorophyll  
254 autofluorescence/side scatter) in a clonal culture of 2n-flagellated cells. (c) Removal of  
255 reverted cells in 2n-flagellated cultures following EhV infections (120 h post-infection). The  
256 fraction (%) of 2n-flagellated and reverted cells is indicated (n = 3). (d) EhV production in 2n-

257 flagellated cultures with reverted cells ( $n = 3$ ). In c and d, non-infected controls were used for  
258 comparison.

259

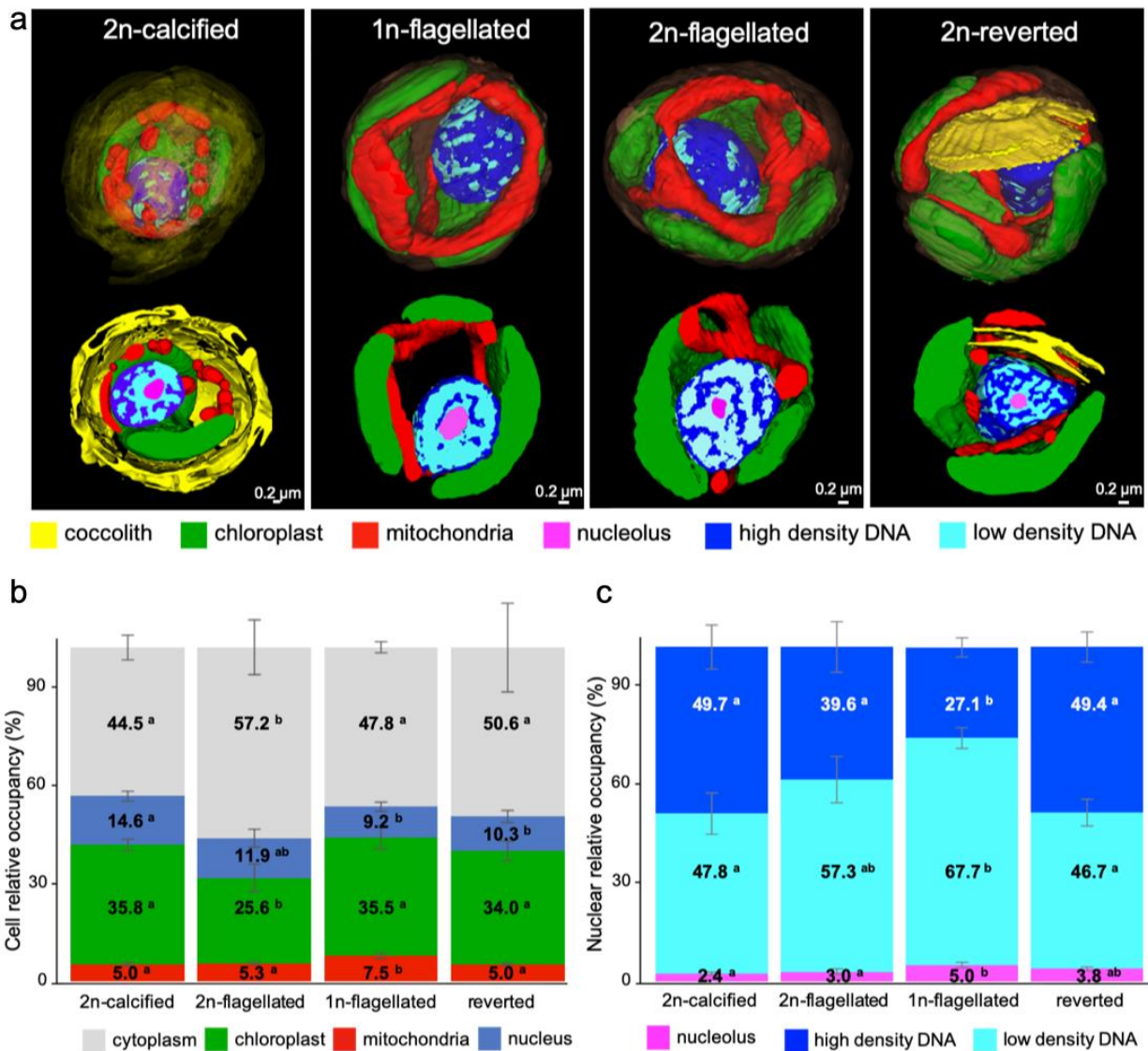
260 Cell and nuclear ultrastructural differences between cell types:

261 The cellular ultrastructure of 2n-calcified, 1n-flagellated, 2n-flagellated and reverted  
262 calcifying cells were analysed by FIB-SEM. Cells from each life phase were  
263 segmented and volumetrics of organelles (chloroplast, mitochondria and nuclear  
264 compartments, such as electron-dense chromatin and nucleolus) were measured.  
265 (Fig. 2a).

266 FIB-SEM results showed that cell volume was higher in 1n-flagellated and 2n-  
267 flagellated (data not shown). Strikingly, reverted calcifying cells were 3-fold more  
268 voluminous compared to 2n-calcified cells and coccolith formation was observed in  
269 the vacuole in some cells (fig.2a). In 2n-calcified and 1n-flagellated cells,  
270 chloroplasts occupied a similar cellular volume ( $35.8\% \pm 1.7\%$  and  $35.5\% \pm 3.4\%$  of  
271 the cell, respectively). Despite their higher cellular volume, chloroplasts in reverted  
272 cells accounted for similar proportions ( $34.0\% \pm 2.9\%$ ). By contrast, in 2n-flagellated  
273 cells chloroplast only represented  $25.6\% \pm 4.1\%$  of the cellular volume. Mitochondria  
274 were largest in 1n-flagellated cells, occupying  $7.5\% \pm 0.8\%$  of the cell volume,  
275 compared to about 5 % in other cell types (2n-flagellated, reverted and 2n-calcified  
276 cells).

277 The nucleus was proportionally larger in 2n-calcified cells ( $14.6\% \pm 1.5\%$  of cell  
278 volume) compared to  $11.9\% \pm 2.7\%$  in 2n-flagellated cells,  $10.3\% \pm 1.8\%$  in reverted  
279 cells, and  $9.2\% \pm 1.4\%$  in 1n cells (fig.2b). Within the nucleus, variations in electron  
280 density patterns were detected, namely high and low density DNA (chromatin) and  
281 nucleolus (Fig. 2 a,c). We therefore assessed the volume occupancy of these  
282 nuclear compartments (% of the nuclear volume) (fig. 2c). The 2n-calcified and  
283 reverted cells exhibited similar density patterns, with an average occupancy of ~49%  
284 for high-density chromatin, ~47% of low-density chromatin and between 2.4% to 3%  
285 for the nucleolus. In contrast, haploid cells (1n-flagellated) displayed an average  
286 occupancy of 27% for high-density chromatin (two-fold less than 2n calcified cells),  
287 nearly 68% for low-density chromatin and about 5% for nucleolus. Finally, 2n-  
288 flagellated cells displayed an intermediate profile, with average occupancies of

289 nearly 40% for high-density chromatin, ~57% for low-density chromatin and 3% for  
 290 the nucleolus.



291

292 **Figure 2 Ultrastructural differentiation of the different life stages of *Gephyrocapsa***  
 293 ***huxleyi*.** (a) 3D reconstitution with FIB-SEM (focused ion beam scanning electron  
 294 microscopy) of 2n-calcified, 1n-flagellated, 2n-flagellated (decoupled) and reverted  
 295 (calcifying) cells. Chloroplasts, mitochondria, nuclear areas and coccoliths are painted in  
 296 different colors. The scale bar represents 2μm. 1n and 2n: haploid and diploid respectively  
 297 (b) Relative proportions (% of cell volume) of cytoplasm, chloroplast, mitochondria and  
 298 nucleus in the four cell types, were calculated from the 3D reconstructions. Significant  
 299 differences ( $p < 0.05$ ) identified by Tukey HSD tests are indicated with different letters above  
 300 the bars (a,b, ab).(c) Relative proportions (% of nuclear occupancy) of high-density and low-  
 301 density areas of the nucleus and the nucleolus in the four cell types. Significant differences

302 ( $p < 0.05$ ) identified by Tukey HSD tests are indicated with different letters above the bars  
303 (a,b, ab). Statistical analysis was performed using ANOVA type II and Tukey HSD tests to  
304 identify significant differences across conditions for each organelle type. Measurements  
305 were based on five 2n-calcified cells, five 1n-flagellated cells, six 2n-flagellated cells, and  
306 four reverted cells.

307

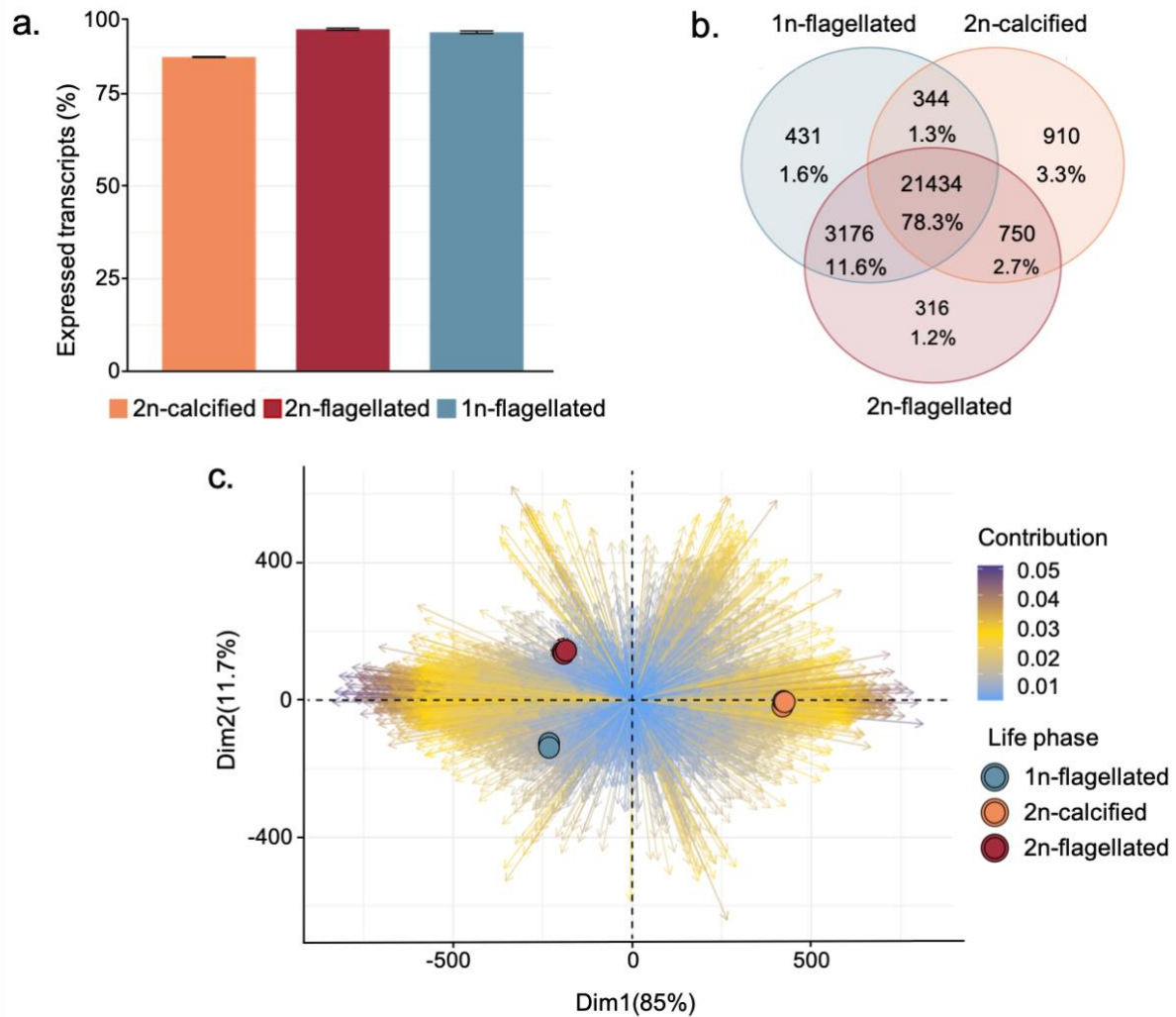
308 Cell growth parameters and global transcriptome features:

309 Gene diversity and expression differences between life phases were examined by  
310 transcriptome analyses. We note that transcriptome analyses were undertaken prior  
311 to the production of reverted cells in LC4-12A, thus solely 2n-calcified, 1n-flagellated  
312 and 2n-flagellated cells are included. Inspection by light microscopy validated that  
313 each culture contained only one cell type. Cells were grown in replete medium and  
314 harvested during exponential growth (Fig. S1b and Table S1). Variations in growth  
315 rates were detected one day prior to sampling for the transcriptomes. The 2n-  
316 calcified cells displayed higher growth rates ( $1.33 \pm 0.04 \text{ day}^{-1}$ ), followed by 1n-  
317 flagellated cells ( $0.72 \pm 0.02 \text{ day}^{-1}$ ) and then 2n-flagellated cell with lower growth rate  
318 dynamics ( $0.47 \pm 0.04 \text{ day}^{-1}$ ). All cultures maintained high photosynthetic quantum  
319 yield of photosystem II (Fv/Fm) throughout cell growth. At the time of biomass  
320 harvesting Fv/Fm equalled  $\sim 0.80$  in all strains (Fig. S1c).

321 To evaluate transcriptional profiles, we defined as transcriptome richness the ratio of  
322 expressed transcripts to the total number of transcripts within this 'synthetic life  
323 phase transcriptome'. It is important to note that we set a conservative expression  
324 threshold ( $> 0.01 \text{ TPM}$ ) to define a transcript as expressed or not. Therefore, we do  
325 not rule out the presence, in this core of transcripts, of sequences presenting only  
326 residual expression levels. Transcripts with less than 3 raw counts across the 9  
327 libraries were filtered-out, yielding 27350 expressed transcripts. Among these, an  
328 average of 23307 (85.2%) were expressed in the 2n-calcified cells, 25263 (92.4%) in  
329 the 1n-flagellated cells and 25543 (93.4%) in the 2n-flagellated ones (Fig. 3a),  
330 aligning with the observed nuclear patterns. Irrespective of expression magnitude or  
331 category-specific differential expression,  $>78\%$  (21434) of gene-transcripts were  
332 common to all three life phases (Fig. 3b). A larger overlap was detected between  
333 flagellated cell types (1n-flagellated and 2n-flagellated, 11.6%, 3176) than between  
334 either of the diploid (2n-calcified and 2n-flagellated ;2.7%, 750) or the non-decoupled

335 stages (1n-flagellated and 2n-calcified, 1.3%, 344). Only a small subset of transcripts  
336 was unique to each phase: 1.6% (431) in 1n-flagellated, 3.3% (910) in 2n-calcified  
337 cells, and only 1.2% (316) in the 2n-flagellated cells. Independent analyses by RT-  
338 qPCR on a set of 19 gene-transcripts validated the specificity of expression of the  
339 targeted transcripts (Table S2).

340 Principal Component Analysis (PCA) delineated the variance between the three life  
341 phases (Fig. 3c), with PC1 accounting for 85% of the variation, separating flagellated  
342 (haploid and decoupled) from 2n-calcified cells. PC2 encompassed 11.7% of the  
343 variation, with 1n-flagellated and 2n-flagellated cells well separated. The top 50  
344 transcripts predominantly influencing PC1 were mostly associated to a morphotype:  
345 34 unique to flagellated forms, two to calcified, nine to diploid, and two commonly  
346 detected across all conditions, albeit with expression differing by more than tenfold  
347 between flagellated and calcified states (see Table S4). Among these, 34 transcripts  
348 annotated with KEGG, KOG or GO terms, were linked to energy metabolism  
349 (including nitrogen), amino acid and inorganic metabolism, as well as genetic  
350 information processing (transcription machinery, chaperone and folding catalysts,  
351 chromosomal proteins, and the ubiquitin system). In contrast, the transcripts central  
352 to PC2 demonstrated different specificities: two associated with flagellated  
353 morphotypes, 19 with diploid-specific expressions, 13 with decoupled cells, and ten  
354 as core transcripts with significant differential gene expression (DGE) across  
355 conditions. Additionally, two transcripts were exclusive to the haploid phase, and four  
356 were associated with non-decoupled phases (Table S4). These transcripts were  
357 related to genetic information processing, including transcription, RNA processing  
358 and modification, replication, recombination, and post-transcriptional modifications,  
359 as well as lipid biosynthesis and photosynthesis.



360

361 **Figure 3 Global properties and differentiation of the transcriptomes of life cycle**  
 362 **phases of *Gephyrocapsa huxleyi*.** (a) Transcriptome richness of the 2n-calcified, 2n-  
 363 flagellated and 1n-flagellated. Richness was defined as the ratio of expressed transcripts to  
 364 the total number of transcripts. A transcript was considered expressed if > 0.01 TPM. (b)  
 365 Venn diagram of transcript specificity across life phases, in the 1n-flagellated, 2n-calcified  
 366 and 2n-flagellated. Number of transcripts in each category, and corresponding relative  
 367 proportions (%) are provided. (c) Principal Component Analysis (PCA) of transcriptional  
 368 expression profiles and relative separation between different life phases along first 2  
 369 principal components (PCs). Each point represents a sample, colored according to its life  
 370 phase. Arrows represent the contribution of individual transcripts to the PCA, with the color  
 371 gradient indicating the magnitude of their contribution (blue: low contribution, purple: high  
 372 contribution). x and y axes: standardized values of the PCs (arbitrary units).

373

374

375 Transcriptional differentiation and functional enrichments:

376 We identified transcriptional differences between 2n-calcified, 1n-flagellated and 2n-  
377 flagellated cells. Abundance estimations were performed at the gene level and  
378 23048 transcripts (84.3%) were statistically differentially expressed (False Discovery  
379 Rate (FDR) < 0.05). Of these, 9114 transcripts exhibited at least a fourfold change in  
380 expression in one pairwise comparison (log<sub>2</sub> fold-change) (Table S5). Cluster  
381 analysis with k-means resulted in 10 clusters with unique patterns of expression (Fig.  
382 4, Table S6) examined for functional enrichments (GO, KEGG and COG). In order to  
383 reduce terms redundancy of the presented Biological Processes (GO- BP, Fig. 4),  
384 significantly enriched terms were grouped under their direct GO ancestor term. Refer  
385 Table S7 for complete results concerning GO-molecular functions, KEGG and COG.  
386 Below are described the enrichments identified in each cluster across all annotation  
387 systems.

388 Clusters 1 and 5, accounted for over half (52%) of the differentially expressed genes  
389 and contained the most influential transcripts in the PCA displayed in Fig. 3c (and  
390 Supporting Information Tables S4, 6). These clusters, exhibited expression patterns  
391 specific to morphotypes (calcified vs non-calcified, flagellated cell types). Cluster 1  
392 (3097 transcripts) was overexpressed in flagellated forms (1n and 2n), featuring  
393 enrichments in cytoskeleton and flagellar processes (cilium or flagellum dependent  
394 cell motility, microtubule and cilium formation, microtubule-based transport, dynein  
395 and catenin binding), as well as metabolism and transport pathways such as  
396 carbohydrates (glycosaminoglycan, fucose), amino-acids and nucleotides and DNA  
397 mediated transposition and DMS<sup>P</sup>-dethiomethylase activity (Table S7). By contrast,  
398 cluster 5 (1644 transcripts) with transcripts overexpressed in calcified cells, featured  
399 enrichments for transport and metabolism of inorganic ions (NO<sub>3</sub><sup>-</sup>, NO<sub>2</sub><sup>-</sup>), as well as  
400 RNA transcription, processing and modification and in cell cycle control and  
401 chromosome partitioning (Table S7). Necrotic cell death, regulation of reactive  
402 oxygen species as well as sialic acid transferase were also associated to cluster 5.

403 Clusters 4 (900 transcripts) and 10 (823 transcripts) represented transcripts  
404 overexpressed in flagellated cells (1n and 2n). Cluster 4 was associated with  
405 coenzyme and nucleotide transport and metabolism (Supporting Information Table  
406 8). Cluster 10 (with higher expression in 2n-flagellated than in 1n-flagellated cells)

407 was enriched in cytoskeleton and cellular architecture organization (actin filament,  
408 calmodulin binding), in lipid and secondary metabolism associated to energy  
409 production and conversion (glucose metabolism), calcium transmembrane transport  
410 as well as RNA processing, modifications and chromatin structure and dynamic  
411 (Table S7).

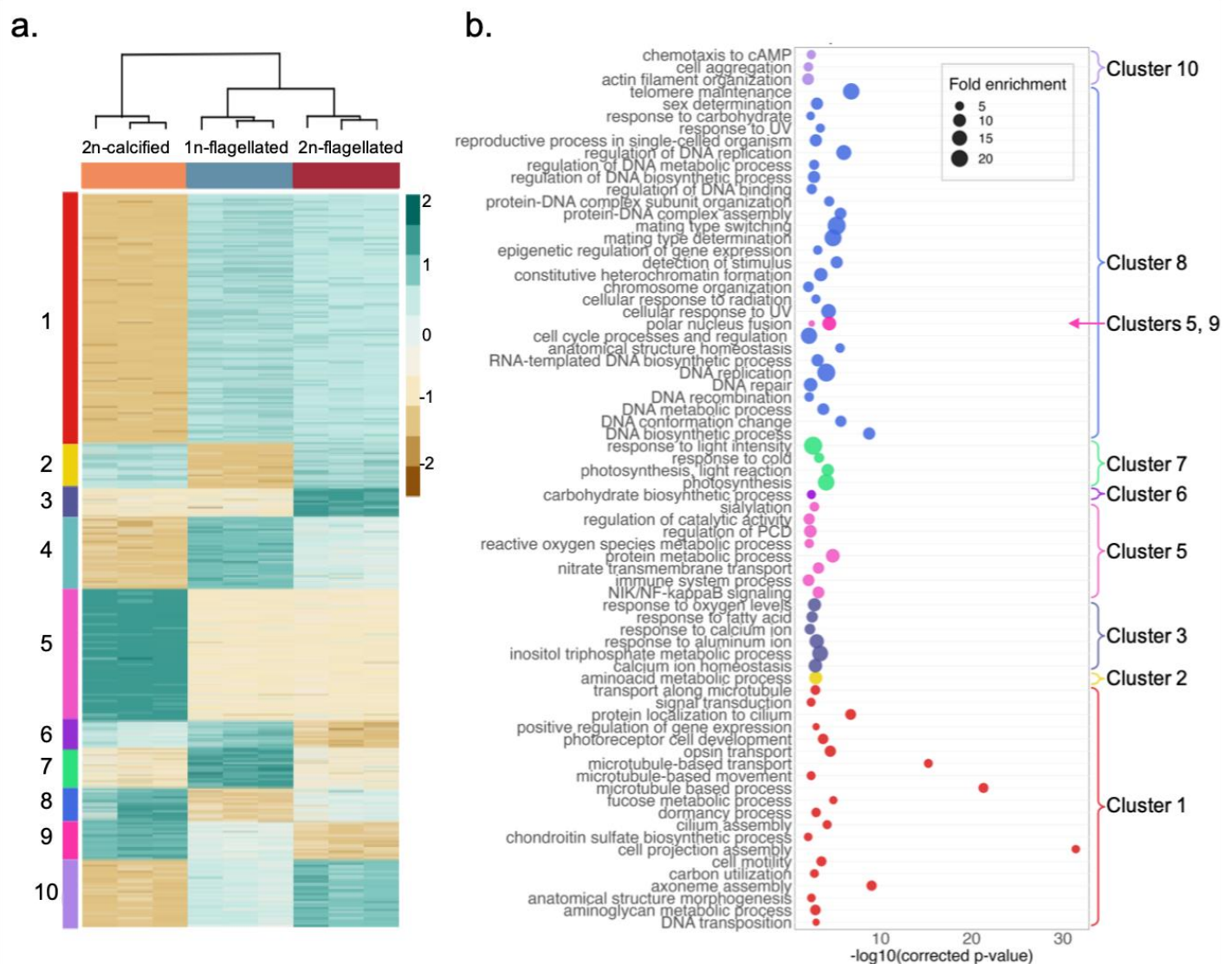
412 Clusters 2 (569 transcripts) and 8 (414 transcripts) comprised transcripts  
413 overexpressed in 2n-calcified and 2n-flagellated. Cluster 2, was associated with  
414 eukaryotic clusters of orthologous proteins (COGs) related to transcription,  
415 processing and modification of RNA and signal transduction mechanisms (Table S7).  
416 Cluster 8 (414 transcripts), with more elevated expression in the 2n-calcified phase,  
417 was enriched in the DNA-related functions replication, recombination and repair  
418 (DNA replication, post replication repair), as well as control of the cell cycle, cell  
419 division and chromosome partitioning (including control of mitotic cell cycle, nuclear  
420 chromosome segregation and sister chromatin cohesion). Processes related to the  
421 structure of the chromatin and its dynamic were also enriched in cluster 8 (chromatin  
422 binding, DNA conformation change, chromosome condensation) as well as mating  
423 type determination, switching and reproductive process in single-cell organism) and  
424 response to various stimuli (UV, carbohydrates).

425 Clusters 6 (337 transcripts) and 9 (481 transcripts), represented non-decoupled cell  
426 transcripts, overexpressed in both 1n-flagellated and 2n-calcified cells. Cluster 6 was  
427 associated with carbohydrate biosynthesis and organic cyclic compound binding. It  
428 was additionally associated to intracellular trafficking and DNA replication,  
429 recombination and repair and post-translational modifications (Table S7). Cluster 9  
430 (more prevalent in 2n-calcified cells), displayed enrichments in carbohydrate  
431 metabolism, polar nucleus fusion and cinnamyl and coniferyl-alcohol  
432 dehydrogenases activity (Table S7).

433 Cluster 3 (353 transcripts) was the only group specific to 2n-flagellated (decoupled)  
434 cells. It was associated to transcription, replication, recombination, repair and  
435 chromatin structure (Table S7), as well as phosphatidyl inositol binding, response  
436 and regulation to calcium ion and aluminum.

437 Finally, cluster 7 (496 transcripts) was haploid specific and included enrichments in  
438 photosynthesis-related functions (chlorophyll and pigment binding) and processes,

439 as well as signal transduction, energy production and conversion, carbohydrate and  
 440 lipid metabolism and intracellular trafficking and vesicular transport.



442 **Figure 4 Expression profiles of transcripts in the life phases of *Gephyrocapsa huxleyi*.**

443 (a) K-mean (10) clustering of transcripts with significant variations of expression (log<sub>2</sub> fold

444 change > 2) between life phases. Expression is standardized by transcript (z-scale). Dark

445 green, high expression levels; brown, low expression levels. b. Most significantly enriched

446 GO terms (Biological Processes), (hypergeometric test, adjusted p value < 0.05) in the

447 clusters presented in a. Terms were grouped under parental term to reduce redundancy and

448 facilitate visualization. For a complete list of enriched GO terms in each cluster, see

449 Supporting information table S8.

450

451

452 Expression of histone, chromatin associated transcripts and transcription factors:  
453 FIB-SEM observations, highlighting marked differences in nuclear ultrastructure  
454 between life cycle phases, stimulated the bioinformatic examination of genetic  
455 functions potentially involved in chromatin remodeling and likely implicated in life  
456 cycling. Alongside, several functional enrichments, particularly those related to  
457 general DNA replication and repair, chromosome conformation and regulation of  
458 gene expression, underscored the potential role of epigenetic regulation in  
459 differentiating life phases. Thus, we zoomed on the transcripts associated with the  
460 general chromatin machinery: histones, readers, writers, remodelers, chaperones, as  
461 well as transcription factors (TFs) (Fig. 5).

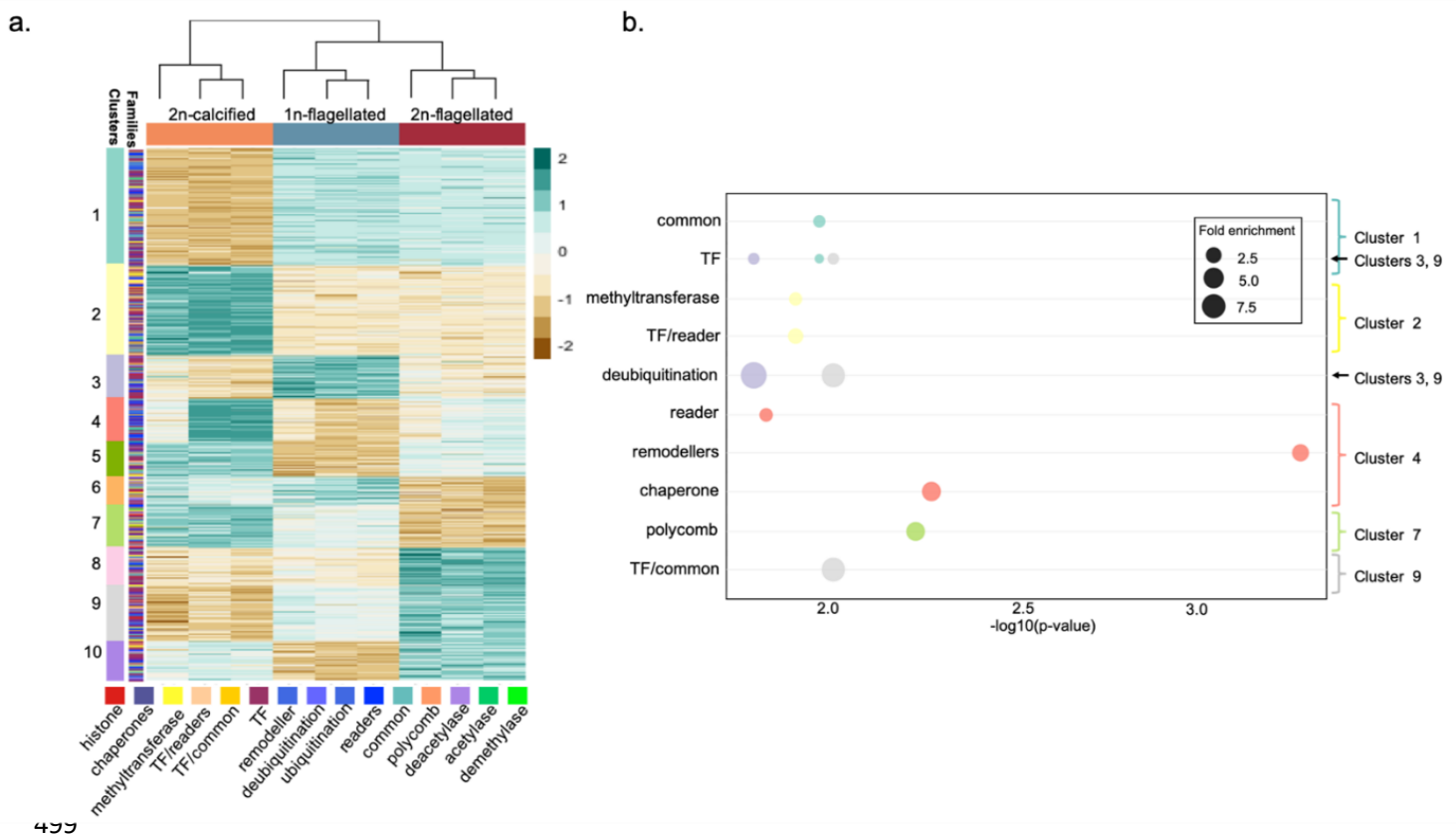
462 Ten clusters were defined, each with highly heterogeneous gene family composition.  
463 Clusters 1 and 2 contained more than a third of the transcripts, segregating cells by  
464 their morphotypes. Cluster 1 (146 transcripts), primarily expressed in flagellated  
465 cells, was enriched in TFs and protein domains commonly associated with the  
466 chromatin machinery ('common'), though no specific family was identified. Cluster 2  
467 (111 transcripts), highly expressed in calcified cells, was enriched in chromatin  
468 specificmethyltransferases, TFs, and chromatin readers.

469 The other clusters were considerably smaller and associated to specific life phases.  
470 Cluster 9 (69 transcripts), expressed in both flagellated conditions, was more  
471 abundant in the 2n-decoupled cells compared to the 1n-calcified cells. This cluster  
472 was associated with TFs, proteins involved in deubiquitination processes, and  
473 'common' domains. Clusters 4, 5, and 10 (54, 43, and 47 transcripts, respectively)  
474 were expressed in diploid cells. Cluster 4, characterized by clear heterogeneity  
475 among its replicates and higher expression in diploids, was enriched in readers,  
476 chaperones, and remodelers. Clusters 5 and 10, more abundant in the 2n-calcified  
477 and 2n-flagellated cells respectively, had no significant enrichments. Cluster 3 (53  
478 transcripts) was the only group characterized by an overexpression exclusive of the  
479 1n-flagellated cells and was associated with TFs and deubiquitination proteins.  
480 Cluster 8 (47 transcripts), the only group overexpressed in the decoupled cells, was  
481 not associated with any significant enrichments. Finally, clusters 6 and 7 (35 and 53  
482 transcripts, respectively) were more expressed in non-decoupled cells. Cluster 6,  
483 with higher expression in 1n-flagellated cells, was not associated with significant

484 enrichments. Cluster 7, with transcripts more expressed in the 2n-calcified than in  
 485 the 1n-flagellated cells, was associated with the Polycomb protein complex.

486 Twenty-two transcripts were associated to canonical histones or their variants in our  
 487 dataset: 10 histones H2A, 2 H2B, 11 histones H3, 1 H4, 1 transcript had homologies  
 488 with the linker histone H1 and 3 transcripts were associated to the histone family  
 489 without a clear identification (Table S5). 5 histones H2A were highly expressed in all  
 490 cells and 2 of them were strongly depleted in the 2n-flagellated and 1n-flagellated  
 491 respectively. One H2B was over 50% more expressed in 2n-calcified cells. Four  
 492 histones H3 were abundant and strongly overexpressed in one life phase. One was  
 493 expressed only in the haploid cells and 1 was expressed in the diploid phases and  
 494 more abundant in the 2n-flagellated. Two other histones H3 were overexpressed in  
 495 the 2n-calcified cells compared to the 2 other conditions. (Table S5). Finally, histone  
 496 H1 was also more detected in 2n cells compared to the 1n cells (Table S5).

497 One TALE-HD TF was identified and expressed only in the haploid cells This  
 498 specificity was validated by RT-qPCR (Table S2).



499

500 **Figure 5 Expression profiles of chromatin associated transcripts related to epigenetic**  
 501 **functions in the life phases of *Gephyrocapsa huxleyi*.** (a) K-mean (10) clustering of

502 transcripts associated to statistically significant variations of expression (FDR < 0.05,  
503 log2fold change >|0.5|) between life phases. Expression is standardized by transcript (z-  
504 scale). Dark green, high expression levels; brown, low expression levels. (b) Significantly  
505 enriched epigenetic families, (hypergeometric test, adjusted p value < 0.05) in clusters  
506 presented in (a). “common”: transcripts for which only one protein domain, commonly found  
507 associated to chromatin-related proteins and processes, could be identified. In the lack of  
508 other identifications, transcripts family could not be determined (conserved domains: Dpy-30,  
509 Kelch, WD-40, zf-C2H2 and zf-CXXC).

510

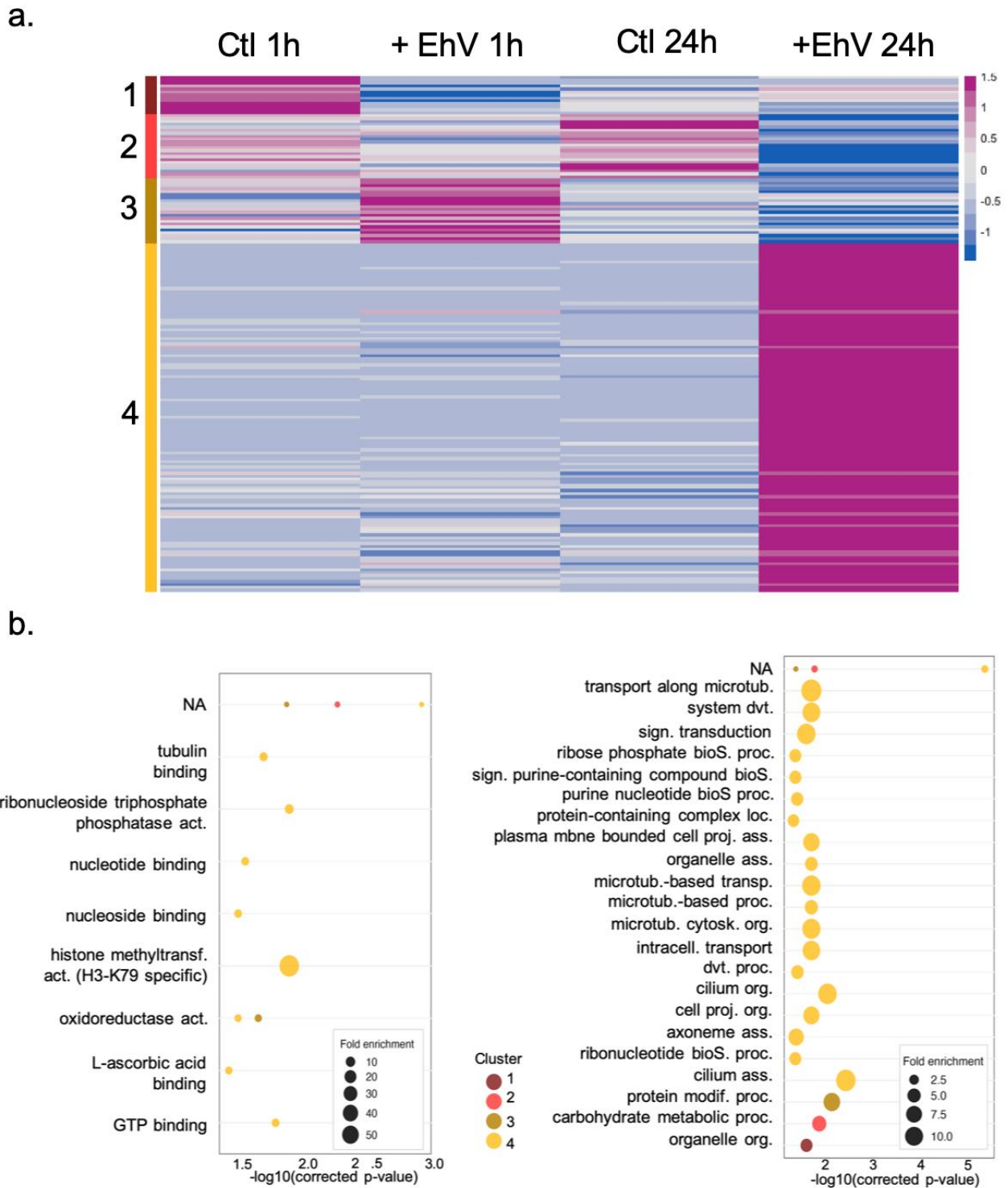
511 Flagellated cell-specific expression during viral infection:

512 Earlier work showed that decoupled cells are produced during Ehv infection (Frada  
513 et al. 2017). To test life cycle transition during viral infection and gain insight on  
514 mechanisms of 2n-flagellated production, we re-examined published transcriptomes  
515 of the diploid *G. huxleyi* strain CCMP2090 interacting with EhV201 (1 and 24 hr post  
516 infection) (Rosenwasser *et al.*, 2014). We targeted transcripts that demonstrated a  
517 high homology level (see methods) with flagellated-specific transcripts identified in  
518 our work.

519 We identified 176 gene homologs in this dataset. Through k-means clustering, four  
520 co-expressed groups were identified and functional enrichments were examined  
521 (Fig. 6a; see Table S3 for complete list of clustered transcripts, associated  
522 annotations and Table S7 for enrichments). Clusters 1 and 2 grouped transcripts  
523 with higher expression levels in control (non-infected) conditions. Cluster 1 (13  
524 transcripts) was found to be enriched in organelle organisation process and to  
525 general metabolism annotations, as well as with protein families involved in genetic  
526 information processing (mRNA biogenesis, chaperones and folding catalysts).

527 Cluster 2 (22 transcripts) was enriched in general metabolism pathways  
528 (carbohydrate, sugar, lipid). Its transcripts were also related to proteins involved in  
529 genetic information processing (transcription and translation machinery, chromatin  
530 structure). Cluster 3 (22 transcripts) comprised transcripts that were upregulated at  
531 1h post-infection but downregulated by 24h. This cluster was enriched in protein  
532 modification process, and its transcripts associated to the metabolism of amino  
533 acids, the ubiquitin system and cilium and its associated proteins. Cluster 4,  
534 containing 68% (119 transcripts) of the total clustered transcripts was markedly

535 upregulated 24h post-infection. It was enriched in orthologous groups and functions  
536 associated to the cell cycle control and division and partitioning of chromosomes, to  
537 signal transduction and cytoskeleton mechanisms (tubulin binding, microtubule  
538 cytoskeleton organisation), as well as post-translational modifications, chaperones  
539 and protein turnover (histone methyltransferase activity, nucleotide/nucleoside  
540 biosynthetic processes and binding) and intracellular trafficking and carbohydrate  
541 transport and metabolism.



542

543 **Figure 6 Expression profiles of “flagellated-specific” genes in the *Gephyrocapsa***  
 544 ***huxleyi* strain CCMP2090 during viral infection by EhV201 and non-infected control.**

545 (a). K-means clustering of homologs in CCMP2090 to flagellated transcripts in absence  
 546 (control, ctl) or presence (+EhV) of viral infection, 1h and 24h post infection. Expression is  
 547 standardized by transcript (z-scale). Dark pink; high expression level, Blue, low expression  
 548 level. (b) Significantly enriched GO terms (1left panel -molecular functions, right panel-

549 biological processes), (hypergeometric test, adjusted p value < 0.05) in cluster presented in  
550 (a). abbreviations: microtub.: microtubule, dvt.: development, sign.: signal, bioS.:  
551 biosynthetic, proc.: processes, loc.: localization, mbne.: membrane, proj.: projection, ass.:  
552 assembly, transp.: transport, cytosk.: cytoskeleton, org.: organization, intracell.: intracellular,  
553 modif.: modification

## 554 **Discussion**

555 The life cycle, encompassing all cellular stages, is a fundamental attribute of  
556 eukaryotes (Mable & Otto, 1998), underpinning their range of biological and  
557 ecological functions, and their capacity to adapt to environmental changes through  
558 cell differentiation. In marine phytoplankton, the understanding of life cycle versatility  
559 and regulation remains limited. Here, we examined the life cycle of the cosmopolitan,  
560 bloom-forming coccolithophore *G. huxleyi*, which includes 2n-calcified cells, 1n-  
561 flagellated cells and 2n-flagellated cells produced during viral infections (Frada *et al.*,  
562 2017). We specifically investigated the fate of the virus-resistant 2n-flagellated cells  
563 during culture observations and the morphogenetic differentiation between 2n-  
564 calcified, 1n-flagellated and 2n-flagellated cell types through ultrastructural and  
565 transcriptome analyses.

566 A key finding of this study is the ability of 2n-flagellated cells to revert back to a  
567 calcifying morphotype in the absence of viral pressure. Given the low proportion of  
568 newly calcifying cells (reverted) within the 2n-flagellated cultures (20%), it is  
569 reasonable to assume that only a small subset of cells undergoes flagellate-to-  
570 calcified switch. Reverted cells subsequently proliferate mitotically, alongside  
571 flagellated cells, becoming detectable by microscopy perhaps only weeks post-  
572 transition.

573 Morphological transformations into a resistant phenotype in response to viral  
574 infections, followed by a reversal to the original cell phenotype, have been observed  
575 in other protists. For example, certain diatom and raphidophyte species can produce  
576 viral-resistant cysts during infections, which germinate back to the vegetative cell-  
577 type in the absence of viruses (Dingman, Jennifer, 2015; Pelusi *et al.*, 2021).  
578 However, the role of viral infection in triggering life cycle transitions and the  
579 subsequent ability to reverse this process, has been to our knowledge uniquely  
580 described here, in *G. huxleyi*. The production of decoupled cells resembling

581 apospory observed in algae and plants (e.g., (Cock *et al.*, 2014) is thought to  
582 represent a natural escape strategy from EhV infection (Frada *et al.*, 2008, 2017).  
583 However, the reversal to a calcifying state suggests a potential cost for diploid cells  
584 in maintaining a haploid-like phenotype. The cost of viral resistance is a common  
585 phenomenon in bacteria and protists (e.g, Lenski & Levin, 1985; Tomaru *et al.*, 2009;  
586 Thomas *et al.*, 2011; Dingman, Jennifer, 2015). In *G. huxleyi*, as corroborated by a  
587 previous study (Frada *et al.*, 2017), 2n-flagellated (decoupled) cells display reduced  
588 growth rate and carrying capacity compared to both 2n-calcified and 1n-flagellated  
589 (haploid) cells. This may be explained by their inherently larger cell volume, revealed  
590 here with electron microscopy, and ensuing higher nutrient requirements for cell  
591 division. In addition, FIB-SEM measurements indicate a relative proportion of  
592 chloroplasts ~25% lower in the 2n-flagellated cells compared to other cell types,  
593 suggesting a reduced capacity to generate energy, while the mitochondrial volume in  
594 these cells, does not change. Thus, the phenotype reversion (2n-flagellated to 2n-  
595 calcified) likely restores the metabolic and physiological advantages.

596 We also observed that the reversion of 2n-flagellated cell in *G. huxleyi*, is associated  
597 with a loss of the viral resistance. Calcified reverted cells, once re-exposed to EhV,  
598 were eliminated from the cultures. Mechanisms underlying *G. huxleyi*'s sensitivity  
599 and resistance to EhV may involve a variety of metabolic processes, reflecting the  
600 complex host-virus 'arms-race' dynamics (Bidle & Vardi, 2011; Bidle, 2015). Previous  
601 work linked specific sphingolipid molecules (sialic acid sphingolipids, sGSL) to EhV  
602 sensitivity in diploid cells (Fulton *et al.*, 2014). sGSLs were proposed to play an  
603 intrinsic role in coccolith biosynthesis and serve as ligands for viral attachment and  
604 entry. These sGSLs are not biochemically detected in haploid cells, which are  
605 insensitive to EhV (Hunter *et al.*, 2015). Enrichment for sialylation transcripts (cluster  
606 5, Fig. 4) was observed only in diploid-calcified cells and not in flagellated cell types  
607 (Table S7). This suggests that sGSLs are likely absent in 2n-flagellated cells but may  
608 reappear in reverted cells, potentially restoring viral sensitivity.

609 Ultrastructural analyses enabled the quantitative assessment of significant  
610 differences in the nuclear architecture between cell-types. Notably, nuclei of 2n-  
611 calcified and reverted cells displayed nearly equal parts of low- and high-density  
612 chromatin, accounting for 47-49% of the nuclear volume, respectively. In contrast,  
613 nuclei of 1n-flagellated were predominantly of low-density chromatin (> 67%), while

614 the nuclei of 2n-flagellated cells displayed an intermediate phenotype (Fig. 2c).  
615 Specific approaches are needed to confirm the nature of high- and low-density areas  
616 ('Chromatin accessibility profiling methods', 2021). However, it is reasonable to  
617 assume that they correspond to heterochromatin (densely packed and usually  
618 transcriptionally inactive), and euchromatin (relaxed and generally associated with a  
619 more active transcriptional state of the DNA), respectively (Huisinga *et al.*, 2006). In  
620 eukaryotes, chromatin packaging is highly dynamic and central for spatiotemporal  
621 regulation of transcription in many cellular processes, including cell differentiation,  
622 development and life cycle in animals, plants and algae (Zheng & Xie, 2019; Vigneau  
623 & Borg, 2021). Rearrangements in chromatin ultrastructure is often mediated by  
624 extensive epigenetic reprogramming, involving hPTMs and the activity of TFs that  
625 modulate transcription landscapes (Morgan *et al.*, 2005; Thiriet-Rupert *et al.*, 2016;  
626 Vigneau & Borg, 2021). Thus, the observed nuclear ultrastructural patterns across  
627 the different phases of the life cycle in *G. huxleyi* through unprecedented volume  
628 quantification, suggest vast divergence of gene expression between cell types, with  
629 2n-flagellated cells maybe displaying an intermediary transcriptional profile.

630 Complementary comparative transcriptome analyses enabled to test this hypothesis,  
631 while gaining insights on wider functional attributes of 2n-calcified, 1n-flagellated and  
632 2n-flagellated cells during exponential growth. Conservative identification of common  
633 and phase specific transcripts highlighted that the strong majority of transcripts were  
634 shared between the three life phases (>78%), regardless of expression levels (Fig.  
635 3b). A more important overlap was observed between flagellated morphotypes  
636 (11.6%) compared to the diploid cells (2.7%) (Fig. 3b). Moreover, as revealed by the  
637 PCA (Fig.3c), the main factor of separation between samples (>85% of the variance)  
638 was attributed to the cell morphology (flagellated vs calcified), while only 11.7% of  
639 the variance equally separated 1n and 2n-flagellated from 2n-calcified(Fig. 3c). This  
640 separation might lay in the mechanisms of production of these cells, haploid cells  
641 being produced by meiosis while decoupled cells would originate from apospory like  
642 mechanisms. Thus, from a transcriptomic perspective, 2n-flagellate cells do not  
643 represent an intermediary level of expression but are in fact much closer to the  
644 haploid profile. Regulation of cell morphology, rather than the ploidy level itself, is the  
645 main factor of differentiation between cell types in *G. huxleyi*, as similarly detected in  
646 brown algae, displaying a haplodiplontic life cycle and apomictic variants (Arun *et al.*,

647 2019). Examination of DGE and enrichment profiles further highlighted the contrasts  
648 between flagellated and 2n-calcified cells. Namely, enrichments in functions related  
649 to cell motility were identified in flagellated cell types. By contrast, enrichments in  
650 functions related to the transport and metabolism of inorganic ions were associated  
651 in 2n-calcified cells. This can be associated with flagella biosynthesis, signalling, and  
652 with calcification, respectively, in broad agreement with previous transcriptome  
653 comparisons of 1n-flagellated and 2n-calcified cells (von Dassow *et al.*, 2009; Rokitta  
654 *et al.*, 2011).

655 Significant differences between cell types were also detected in functions associated  
656 to chromatin remodelling and transcriptional regulation (including transcription  
657 factors, TFs). At this level, morphotype is still the main differentiating factor of  
658 transcript expression (Fig.5). A majority of transcripts were either over-expressed in  
659 the calcified cells or in the flagellated cells, further emphasizing phenotype as the  
660 main manifestation of gene regulation in *G. huxleyi* life phases, independently of  
661 ploidy levels. We detected for example an over expression in the 2n-calcified cells  
662 associated to histones methyltransferases (Fig. 5a, cluster 2), which can align with  
663 the high chromatin compaction (Chen & Dent, 2014) detected in these cells (Fig. 2b).  
664 Additionally, ubiquitin removal (deubiquitination) enrichments were associated to  
665 transcripts with clear DE between cell types, a function associated with chromatin  
666 relaxation and increased transcriptional activity (Mikulski *et al.*, 2017). It fits overall  
667 patterns of chromatin, less dense in flagellated cells compared to 2n-calcified cells  
668 (Fig. 5a, clusters 3 and 9). Enrichments observed in cluster 4 (Fig. 5a) were also  
669 relevant. We observed an heterogeneous gradient of expression from 2n-calcified to  
670 2n-flagellated cells, in transcripts associated to chromatin "architects" (remodelers,  
671 readers, and chaperones) (Clapier & Cairns, 2009). suggesting intermediate level of  
672 expression and recruitment. We think it might further be associated to a transitional  
673 level of chromatin organization between the two types of diploid cells. We also  
674 identified enrichments for the Polycomb-Complex, associated with transcripts over-  
675 expressed in 2n-calcified cells compared to 1n-flagellated cells, a TALE-  
676 homeodomain TF (KNOX) specific to 1n cells and histone H3 variants with marked  
677 DGE between life phases. All these gene-functions are conserved across eukaryotes  
678 and play central roles in the regulation of life cycle transitions (Horst *et al.*, 2016;  
679 Mikami *et al.*, 2019; Dierschke *et al.*, 2021; Lee *et al.*). Polycomb associated

680 transcripts in our dataset, likely correspond to the Polycomb Repressive Complex 2  
681 (PRC2,(Grau-Bové *et al.*, 2022)). PRC2 mediates the trimethylation of histones 3  
682 (H3) on lysine 27 (H3K27me3), associated to the silencing of genes encoding TALE-  
683 HD TF (KNOX and BELL). These transcription factors are crucial for haploid-to-  
684 diploid transitions in plants and algae (Yelagandula *et al.*, 2014; Wollmann *et al.*,  
685 2017; Thangavel & Nayar, 2018; Loppin & Berger, 2020; Vigneau & Borg, 2021;  
686 Hirooka *et al.*, 2022) (Loppin & Berger 2020, Wollman *et al.* 2017, Yelagandula *et al.*  
687 2014, Vigneau and Borg, 2021). Overall, these results suggest the involvement of  
688 conserved key mechanisms in the life cycle differentiation in *G. huxleyi*.

689 Finally, we screened for homologs of flagellated-specific transcripts (expressed only  
690 in the 1n- and 2n-flagelled cells) in the *G. huxleyi* strain CCMP2090 during EhV  
691 infection (Rosenwasser *et al.*, 2014). A variety of transcripts mainly involved in  
692 metabolic functions were detected in control (non-infected) conditions, indicating that  
693 life cycle specificity of various functional genes can change between *G. huxleyi*  
694 strains, and likely as well between growth conditions and external environmental  
695 stimuli. However, most of the flagellated-cell homologs identified were over-  
696 expressed 24h post-infection, strongly supporting life cycle transition during EhV  
697 infections (Frada *et al.*, 2017). Among flagellated-cell homologs, we detected a wide  
698 set of motility-related genes and cell architecture, as well as signaling transduction  
699 mechanisms. We note that CCMP2090 does not form flagella, albeit a strong up-  
700 regulation of some flagellar functions upon infection (Frada *et al.*, 2017) . This  
701 morphological impairment likely results from the genome erosion of this strain and  
702 subsequent loss of functional flagella biosynthesis genes(von Dassow *et al.*, 2015).  
703 In any case, CCMP2090 can form viral resistant cells that display organic-body  
704 scales, only featured in flagellated cells (Frada *et al.*, 2017). We further detected the  
705 association of these flagellated specific transcripts with functions involved in gene  
706 expression, including a strong enrichment in H3K79 specific methyltransferase that  
707 regulates diverse cellular processes, such as development, reprogramming,  
708 differentiation, and proliferation (Feng *et al.*, 2010; Takahashi *et al.*, 2011; Tatum &  
709 Li, 2011). It is generally associated to active transcription and maintains chromatin  
710 accessibility for histone acetylation and transcription factor binding (Godfrey *et al.*,  
711 2019). Although limited, these lines of evidence further recall the involvement of the

712 upstream epigenetic regulation during the differentiation of 2n-flagellated cells in *G.*  
713 *huxleyi* during viral infection.

714 In conclusion, our study provides detailed comparative assessment of three life-cycle  
715 cell types in *G. huxleyi*, evidence for life cycle versatility during viral interactions and  
716 novel insight on chromatin remodeling and transcription regulators, which likely play  
717 a central role in mediating life cycle transitions. These results should provide the  
718 basis for further development of future studies on role epigenetic control during cell  
719 differentiation and life cycle in coccolithophores. Finally, one of the greatest gaps  
720 limiting comprehensive understanding about *G. huxleyi* is the lack of knowledge of  
721 the occurrence, distribution, and role of flagellated cells in natural populations and  
722 during blooms where viral infections are recurrent. The analyses presented in this  
723 study enable to refine useful gene markers for the three life phases of *G. huxleyi* that  
724 can be used to interrogate coccolithophore assemblages in the oceans.

725

## 726 **Acknowledgements**

727 This study was supported by the Israel Science Foundation (2921/20) attributed to  
728 MJF. JD was supported by CNRS and ATIP-Avenir program funding. We thank Prof.  
729 Aharon Kaplan for hosting SF and Prof. Assaf Vardi to enable access to his lab for  
730 molecular biology work. We thank The Life Sciences Core Facilities from The  
731 Weizmann Institute of Science for the preparation and transcriptome sequencing as  
732 well as Esther Feldmesser for support and feedback on bioinformatic analyses. We  
733 thank Guy Schoehn and Christine Moriscot, and the electron microscope facility at  
734 IBS, which is supported by the Rhône-Alpes Region, the Fondation Recherche  
735 Medicale (FRM), the fonds FEDER, the Center National de la Recherche  
736 Scientifique (CNRS), the CEA, the University of Grenoble, EMBL, and the GIS  
737 Infrastructures en Biologie Sante et Agronomie (IBISA). CK was supported by  
738 Academia Sinica (AS-CDA-110-L01) and NSTC, Taiwan (111-2611-M001-008-MY3).

739

## 740 **Author contributions**

741 Conceptualization: MJF, LB, SF; Methodology: MJF, JD, LB, SF, OM; Investigation:  
742 LB, SF, YA, GK, FC, BG, RT, YS, MJF; Formal Analysis: LB, YD, SF; Resources:

743 CK, FC, MJF; Data curation: LB, JD, BG, RT, YS; Visualization: LB, JD, RT, YS;  
744 Writing (original draft): LB, MJF; Writing (review & editing): LB, MJF, SF, JD and CK;  
745 Supervision: MJF, OM; Funding acquisition: MJF, CK.

#### 746 **Competing interests:**

747 All authors declare they have no competing interests.

#### 748 **References**

- 749 **Arun A, Coelho SM, Peters AF, Bourdareau S, Pérès L, Scornet D, Strittmatter M,**  
750 **Lipinska AP, Yao H, Godfroy O, et al. 2019.** Convergent recruitment of TALE  
751 homeodomain life cycle regulators to direct sporophyte development in land plants and  
752 brown algae (S McCormick and CS Hardtke, Eds.). *eLife* **8**: e43101.
- 753 **Balch WM. 2018.** The Ecology, Biogeochemistry, and Optical Properties of  
754 Coccolithophores. *Annual Review of Marine Science* **10**: 71–98.
- 755 **Bendif EM, Probert I, Archontikis OA, Young JR, Beaufort L, Rickaby RE, Filatov D.**  
756 **2023.** Rapid diversification underlying the global dominance of a cosmopolitan  
757 phytoplankton. *The ISME Journal* **17**: 630–640.
- 758 **Benjamini Y, Hochberg Y. 1995.** Controlling the False Discovery Rate: A Practical and  
759 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B*  
760 *(Methodological)* **57**: 289–300.
- 761 **Bidle KD. 2015.** The Molecular Ecophysiology of Programmed Cell Death in Marine  
762 Phytoplankton. *Annual Review of Marine Science* **7**: 341–375.
- 763 **Bidle KD, Vardi A. 2011.** A chemical arms race at sea mediates algal host–virus  
764 interactions. *Current Opinion in Microbiology* **14**: 449–457.
- 765 **Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G,**  
766 **Paysan-Lafosse T, Qureshi M, Raj S, et al. 2021.** The InterPro protein families and  
767 domains database: 20 years on. *Nucleic Acids Research* **49**: D344–D354.
- 768 **Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina  
769 sequence data. *Bioinformatics* **30**: 2114–2120.
- 770 **Bratbak G, Egge JK, Heldal M. 1993.** Viral mortality of the marine alga *Emiliania huxleyi*  
771 (Haptophyceae) and termination of algal blooms. *Marine Ecology Progress Series* **93**: 39–  
772 48.
- 773 **Buchfink B, Reuter K, Drost H-G. 2021.** Sensitive protein alignments at tree-of-life scale  
774 using DIAMOND. *Nature Methods* **18**: 366–368.
- 775 **Chen T, Dent SYR. 2014.** Chromatin modifiers and remodellers: regulators of cellular  
776 differentiation. *Nature Reviews Genetics* **15**: 93–106.
- 777 **Chromatin accessibility profiling methods. 2021.** *Nature Reviews Methods Primers* **1**: 11.
- 778 **Clapier CR, Cairns BR. 2009.** The Biology of Chromatin Remodeling Complexes. *Annual*  
779 *Review of Biochemistry* **78**: 273–304.
- 780 **Cock JM, Godfroy O, Macaisne N, Peters AF, Coelho SM. 2014.** Evolution and regulation  
781 of complex life cycles: a brown algal perspective. *Current Opinion in Plant Biology* **17**: 1–6.

- 782 **Coelho SM, Godfroy O, Arun A, Corguillé GL, Peters AF, Cock JM.** OUROBOROS is a  
783 master regulator of the gametophyte to sporophyte life cycle transition in the brown alga  
784 *Ectocarpus*. : 6.
- 785 **Coelho SM, Peters AF, Charrier B, Roze D, Destombe C, Valero M, Cock JM. 2007.**  
786 Complex life cycles of multicellular eukaryotes: New approaches based on the use of model  
787 organisms. *Gene* **406**: 152–170.
- 788 **von Dassow P, John U, Ogata H, Probert I, Bendif EM, Kegel JU, Audic S, Wincker P,**  
789 **Da Silva C, Claverie J-M, et al. 2015.** Life-cycle modification in open oceans accounts for  
790 genome variability in a cosmopolitan phytoplankton. *The ISME Journal* **9**: 1365–1377.
- 791 **von Dassow P, Ogata H, Probert I, Wincker P, Da Silva C, Audic S, Claverie J-M, de**  
792 **Vargas C. 2009.** Transcriptome analysis of functional differentiation between haploid and  
793 diploid cells of *Emiliana huxleyi*, a globally significant photosynthetic calcifying cell. *Genome*  
794 *Biology* **10**: R114.
- 795 **Decelle J, Stryhanyuk H, Gallet B, Veronesi G, Schmidt M, Balzano S, Marro S,**  
796 **Uwizeye C, Jouneau P-H, Lupette J, et al. 2019.** Algal Remodeling in a Ubiquitous  
797 Planktonic Photosymbiosis. *Current Biology* **29**: 968-978.e4.
- 798 **Decelle J, Veronesi G, LeKieffre C, Gallet B, Chevalier F, Stryhanyuk H, Marro S,**  
799 **Ravanel S, Tucoulou R, Schieber N, et al. 2021.** Subcellular architecture and metabolic  
800 connection in the planktonic photosymbiosis between *Collodaria* (radiolarians) and their  
801 microalgae. *Environmental Microbiology* **23**: 6569–6586.
- 802 **Dierschke T, Flores-Sandoval E, Rast-Somssich MI, Althoff F, Zachgo S, Bowman JL.**  
803 **2021.** Gamete expression of TALE class HD genes activates the diploid sporophyte program  
804 in *Marchantia polymorpha*. *eLife* **10**: e57088.
- 805 **Dingman, Jennifer. 2015.** Virus induced cell death, evasion and resistance in the harmful  
806 bloom-forming alga, *Heterosigma akashiwo*.
- 807 **Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,**  
808 **Gingeras TR. 2013.** STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
- 809 **Feldmesser E, Rosenwasser S, Vardi A, Ben-Dor S. 2014.** Improving transcriptome  
810 construction in non-model organisms: integrating manual and automated gene definition in  
811 *Emiliana huxleyi*. *BMC Genomics* **15**: 148.
- 812 **Feng S, Cokus SJ, Zhang X, Chen P-Y, Bostick M, Goll MG, Hetzel J, Jain J, Strauss**  
813 **SH, Halpern ME, et al. 2010.** Conservation and divergence of methylation patterning in  
814 plants and animals. *Proceedings of the National Academy of Sciences* **107**: 8689–8694.
- 815 **Frada MJ, Bendif EM, Keuter S, Probert I. 2019.** The private life of coccolithophores.  
816 *Perspectives in Phycology* **6**: 11–30.
- 817 **Frada MJ, Bidle KD, Probert I, de Vargas C. 2012.** In situ survey of life cycle phases of the  
818 coccolithophore *Emiliana huxleyi* (Haptophyta): *Emiliana huxleyi* life cycling. *Environmental*  
819 *Microbiology* **14**: 1558–1569.
- 820 **Frada M, Probert I, Allen MJ, Wilson WH, de Vargas C. 2008.** The “Cheshire Cat” escape  
821 strategy of the coccolithophore *Emiliana huxleyi* in response to viral infection. *Proceedings*  
822 *of the National Academy of Sciences* **105**: 15944–15949.
- 823 **Frada MJ, Rosenwasser S, Ben-Dor S, Shemi A, Sabanay H, Vardi A. 2017.**

- 824 Morphological switch to a resistant subpopulation in response to viral infection in the bloom-  
825 forming coccolithophore *Emiliana huxleyi* (JS Weitz, Ed.). *PLOS Pathogens* **13**: e1006775.
- 826 **Fu L, Niu B, Zhu Z, Wu S, Li W. 2012.** CD-HIT: accelerated for clustering the next-  
827 generation sequencing data. *Bioinformatics* **28**: 3150–3152.
- 828 **Fulton JM, Fredricks HF, Bidle KD, Vardi A, Kendrick BJ, DiTullio GR, Van Mooy BAS.**  
829 **2014.** Novel molecular determinants of viral susceptibility and resistance in the lipidome of  
830 *Emiliana huxleyi*. *Environmental Microbiology* **16**: 1137–1149.
- 831 **Gallet B, Moriscot C, Schoehn G, Decelle J. 2024.** Cryo-fixation and resin embedding of  
832 biological samples for electron microscopy and chemical imaging v1.
- 833 **Godfrey L, Crump NT, Thorne R, Lau I-J, Repapi E, Dimou D, Smith AL, Harman JR,**  
834 **Telenius JM, Oudelaar AM, et al. 2019.** DOT1L inhibition reveals a distinct subset of  
835 enhancers dependent on H3K79 methylation. *Nature Communications* **10**: 2803.
- 836 **Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M,**  
837 **Talón M, Dopazo J, Conesa A. 2008.** High-throughput functional annotation and data  
838 mining with the Blast2GO suite. *Nucleic Acids Research* **36**: 3420–3435.
- 839 **Grau-Bové X, Navarrete C, Chiva C, Pribasniq T, Antó M, Torruella G, Galindo LJ, Lang**  
840 **BF, Moreira D, López-García P, et al. 2022.** A phylogenetic and proteomic reconstruction  
841 of eukaryotic chromatin evolution. *Nature Ecology & Evolution* **6**: 1007–1023.
- 842 **Green JC, Course PA, Tarran GA. 1996.** The life-cycle of *Emiliana huxleyi*: A brief review  
843 and a study of relative ploidy levels analysed by flow cytometry. *Journal of Marine Systems*  
844 **9**: 33–44.
- 845 **Hirooka S, Itabashi T, Ichinose TM, Onuma R, Fujiwara T, Yamashita S, Jong LW,**  
846 **Tomita R, Iwane AH, Miyagishima S. 2022.** Life cycle and functional genomics of the  
847 unicellular red alga *Galdieria* for elucidating algal and plant evolution and industrial use.  
848 *Proceedings of the National Academy of Sciences* **119**: e2210665119.
- 849 **Horst NA, Katz A, Pereman I, Decker EL, Ohad N, Reski R. 2016.** A single homeobox  
850 gene triggers phase transition, embryogenesis and asexual reproduction. *Nature Plants* **2**.
- 851 **Houdan A, Billard C, Marie D, Not F, Sáez AG, Young JR, Probert I. 2004.**  
852 Holococcolithophore-heterococcolithophore (Haptophyta) life cycles: Flow cytometric  
853 analysis of relative ploidy levels. *Systematics and Biodiversity* **1**: 453–465.
- 854 **Houdan A, Probert I, Van Lenning K, Lefebvre S. 2005.** Comparison of photosynthetic  
855 responses in diploid and haploid life-cycle phases of *Emiliana huxleyi* (Prymnesiophyceae).  
856 *Marine Ecology Progress Series* **292**: 139–146.
- 857 **Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H,**  
858 **Mende DR, Letunic I, Rattei T, Jensen LJ, et al. 2019.** eggNOG 5.0: a hierarchical,  
859 functionally and phylogenetically annotated orthology resource based on 5090 organisms  
860 and 2502 viruses. *Nucleic Acids Research* **47**: D309–D314.
- 861 **Huisinga KL, Brower-Toland B, Elgin SCR. 2006.** The contradictory definitions of  
862 heterochromatin: transcription and silencing. *Chromosoma* **115**: 110–122.
- 863 **Hunter JE, Frada MJ, Fredricks HF, Vardi A, Van Mooy BAS. 2015.** Targeted and  
864 untargeted lipidomics of *Emiliana huxleyi* viral infection and life cycle phases highlights  
865 molecular biomarkers of infection, susceptibility, and ploidy. *Frontiers in Marine Science* **2**.

- 866 **Kao T-T, Lai M-W, Wang T-H, Yang C-L, Frada MJ, Ku C. 2024.** Haplotype-aware  
867 multiomics unveils the regulatory basis of haplodiplontic life-cycle differentiation in a  
868 cosmopolitan marine alga. : 2024.05.26.595999.
- 869 **Keller MD, Selvin RC, Claus W, Guillard RRL. 1987.** Media for the Culture of Oceanic  
870 Ultraphytoplankton1,2. *Journal of Phycology* **23**: 633–638.
- 871 **Klaveness D. 1972.** *Coccolithus huxleyi* (Lohm.) Kamptn II. The flagellate cell, aberrant cell  
872 types, vegetative propagation and life cycles. *British Phycological Journal* **7**: 309–318.
- 873 **Lawrence M, Daujat S, Schneider R. 2016.** Lateral Thinking: How Histone Modifications  
874 Regulate Gene Expression. *Trends in Genetics* **32**: 42–56.
- 875 **Lee J-H, Lin H, Joo S, Goodenough U. 2008.** Early Sexual Origins of Homeoprotein  
876 Heterodimerization and Evolution of the Plant KNOX/BELL Family. *Cell* **133**: 829–840.
- 877 **Lenski RE, Levin BR. 1985.** Constraints on the Coevolution of Bacteria and Virulent Phage:  
878 A Model, Some Experiments, and Predictions for Natural Communities. *The American*  
879 *Naturalist* **125**: 585–602.
- 880 **Li B, Dewey CN. 2011.** RSEM: accurate transcript quantification from RNA-Seq data with or  
881 without a reference genome. : 16.
- 882 **Loppin B, Berger F. 2020.** Histone Variants: The Nexus of Developmental Decisions and  
883 Epigenetic Memory. *Annual Review of Genetics* **54**: 121–149.
- 884 **Love MI, Huber W, Anders S. 2014.** Moderated estimation of fold change and dispersion  
885 for RNA-seq data with DESeq2. *Genome Biology* **15**.
- 886 **Mable BK, Otto SP. 1998.** The evolution of life cycles with haploid and diploid phases.  
887 *BioEssays* **20**: 453–462.
- 888 **Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021.** BUSCO Update: Novel  
889 and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for  
890 Scoring of Eukaryotic, Prokaryotic, and Viral Genomes (J Kelley, Ed.). *Molecular Biology and*  
891 *Evolution* **38**: 4647–4654.
- 892 **Mikami K, Li C, Irie R, Hama Y. 2019.** A unique life cycle transition in the red seaweed  
893 *Pyropia yezoensis* depends on apospory. *Communications Biology* **2**: 299.
- 894 **Mikulski P, Komarynets O, Fachinelli F, Weber APM, Schubert D. 2017.** Characterization  
895 of the Polycomb-Group Mark H3K27me3 in Unicellular Algae. *Frontiers in Plant Science* **8**:  
896 607.
- 897 **Monteiro FM, Bach LT, Brownlee C, Bown P, Rickaby REM, Poulton AJ, Tyrrell T,  
898 Beaufort L, Dutkiewicz S, Gibbs S, et al. 2016.** Why marine phytoplankton calcify. *Science*  
899 *Advances* **2**: e1501822.
- 900 **Morgan HD, Santos F, Green K, Dean W, Reik W. 2005.** Epigenetic reprogramming in  
901 mammals. *Human Molecular Genetics* **14**: R47–R58.
- 902 **Oyama R, Haba G, Kaido Y, Kanasugi T, Isurugi C, Kikuchi A, Sugiyama T, Jakab M,  
903 Pujol S, Kikinis R. 2014.** P03.15: Towards improved ultrasound based analysis and 3D  
904 visualisation of the fetal brain using 3DSlicer. *Ultrasound in Obstetrics & Gynecology* **44**:  
905 200–200.
- 906 **Pelusi A, De Luca P, Manfellotto F, Thamatrakoln K, Bidle KD, Montresor M. 2021.**

- 907 Virus-induced spore formation as a defense mechanism in marine diatoms. *New Phytologist*  
908 **229**: 2251–2259.
- 909 **Pertea G, Pertea M. 2020.** GFF Utilities: GffRead and GffCompare. *F1000Research* **9**: 304.
- 910 **Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015.**  
911 StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature*  
912 *Biotechnology* **33**: 290–295.
- 913 **Rokitta SD, de Nooijer LJ, Trimborn S, de Vargas C, Rost B, John U. 2011.**  
914 Transcriptome Analyses Reveal Differential Gene Expression Patterns Between the Life-  
915 Cycle Stages of *Emiliana Huxleyi* (haptophyta) and Reflect Specialization to Different  
916 Ecological Niches<sup>1</sup>. *Journal of Phycology* **47**: 829–838.
- 917 **Rokitta SD, Rost B. 2012.** Effects of CO<sub>2</sub> and their modulation by light in the life-cycle  
918 stages of the coccolithophore *Emiliana huxleyi*. *Limnology and Oceanography* **57**: 607–618.
- 919 **Rosenwasser S, Mausz MA, Schatz D, Sheyn U, Malitsky S, Aharoni A, Weinstock E,**  
920 **Tzfadia O, Ben-Dor S, Feldmesser E, et al. 2014.** Rewiring Host Lipid Metabolism by  
921 Large Viruses Determines the Fate of *Emiliana huxleyi*, a Bloom-Forming Alga in the  
922 Ocean. *The Plant Cell* **26**: 2689–2707.
- 923 **Sakakibara K, Ando S, Yip HK, Tamada Y, Hiwatashi Y, Murata T, Deguchi H, Hasebe**  
924 **M, Bowman JL. 2013.** KNOX2 Genes Regulate the Haploid-to-Diploid Morphological  
925 Transition in Land Plants. *Science* **339**: 1067–1070.
- 926 **Schmidt A. 2020.** Controlling Apomixis: Shared Features and Distinct Characteristics of  
927 Gene Regulation. *Genes* **11**: 329.
- 928 **Schroeder DC, Oke J, Malin G, Wilson WH. 2002.** Coccolithovirus (Phycodnaviridae):  
929 Characterisation of a new large dsDNA algal virus that infects *Emiliana huxleyi*. *Archives of*  
930 *Virology* **147**: 1685–1698.
- 931 **Takahashi Y-H, Schulze JM, Jackson J, Hentrich T, Seidel C, Jaspersen SL, Kobor MS,**  
932 **Shilatifard A. 2011.** Dot1 and Histone H3K79 Methylation in Natural Telomeric and HM  
933 Silencing. *Molecular Cell* **42**: 118–126.
- 934 **Tatum D, Li S. 2011.** Evidence That the Histone Methyltransferase Dot1 Mediates Global  
935 Genomic Repair by Methylating Histone H3 on Lysine 79. *Journal of Biological Chemistry*  
936 **286**: 17530–17535.
- 937 **Thangavel G, Nayar S. 2018.** A Survey of MIKC Type MADS-Box Genes in Non-seed  
938 Plants: Algae, Bryophytes, Lycophytes and Ferns. *Frontiers in Plant Science* **9**: 510.
- 939 **Thiriet-Rupert S, Carrier G, Chénais B, Trottier C, Bougaran G, Cadoret J-P, Schoefs**  
940 **B, Saint-Jean B. 2016.** Transcription factors in microalgae: genome-wide prediction and  
941 comparative analysis. *BMC Genomics* **17**: 282.
- 942 **Thomas R, Grimsley N, Escande M, Subirana L, Derelle E, Moreau H. 2011.** Acquisition  
943 and maintenance of resistance to viruses in eukaryotic phytoplankton populations: Viral  
944 resistance in *Mamiellales*. *Environmental Microbiology* **13**: 1412–1420.
- 945 **Tomaru Y, Mizumoto H, Nagasaki K. 2009.** Virus resistance in the toxic bloom-forming  
946 dinoflagellate *Heterocapsa circularisquama* to single-stranded RNA virus infection.  
947 *Environmental Microbiology* **11**: 2915–2923.
- 948 **Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL,**

- 949 **Rinn JL, Pachter L. 2012.** Differential gene and transcript expression analysis of RNA-seq  
950 experiments with TopHat and Cufflinks. *Nature Protocols* **7**: 562–578.
- 951 **Uwizeye C, Decelle J, Jouneau P-H, Flori S, Gallet B, Keck J-B, Bo DD, Moriscot C,**  
952 **Seydoux C, Chevalier F, et al. 2021.** Morphological bases of phytoplankton energy  
953 management and physiological responses unveiled by 3D subcellular imaging. *Nature*  
954 *Communications* **12**: 1049.
- 955 **Vardi A, Haramaty L, Van Mooy BAS, Fredricks HF, Kimmance SA, Larsen A, Bidle KD.**  
956 **2012.** Host-virus dynamics and subcellular controls of cell fate in a natural coccolithophore  
957 population. *Proceedings of the National Academy of Sciences* **109**: 19327–19332.
- 958 **Vigneau J, Borg M. 2021.** The epigenetic origin of life history transitions in plants and algae.  
959 *Plant Reproduction*.
- 960 **Von Dassow P, Montresor M. 2011.** Unveiling the mysteries of phytoplankton life cycles:  
961 patterns and opportunities behind complexity. *Journal of Plankton Research* **33**: 3–12.
- 962 **de Vries J, Monteiro F, Wheeler G, Poulton A, Godrijan J, Cerino F, Malinverno E,**  
963 **Langer G, Brownlee C. 2021.** Haplo-diplontic life cycle expands coccolithophore niche.  
964 *Biogeosciences* **18**: 1161–1184.
- 965 **Wollmann H, Stroud H, Yelagandula R, Tarutani Y, Jiang D, Jing L, Jamge B, Takeuchi**  
966 **H, Holec S, Nie X, et al. 2017.** The histone H3 variant H3.3 regulates gene body DNA  
967 methylation in *Arabidopsis thaliana*. *Genome Biology* **18**: 94.
- 968 **Yelagandula R, Stroud H, Holec S, Zhou K, Feng S, Zhong X, Muthurajan UM, Nie X,**  
969 **Kawashima T, Groth M, et al. 2014.** The Histone Variant H2A.W Defines Heterochromatin  
970 and Promotes Chromatin Condensation in *Arabidopsis*. *Cell* **158**: 98–109.
- 971 **Zhao H, Lu M, Singh R, Snell WJ. 2001.** Ectopic expression of a *Chlamydomonas* mt+-  
972 specific homeodomain protein in mt- gametes initiates zygote development without gamete  
973 fusion. *Genes & Development* **15**: 2767–2777.
- 974 **Zheng H, Xie W. 2019.** The role of 3D genome organization in development and cell  
975 differentiation. *Nature Reviews Molecular Cell Biology* **20**: 535–550.
- 976