

1 **Genetic tool development in marine protists:**

2 **Emerging model organisms for experimental cell biology**

3

4 Drahomíra Faktorová^{1,#,*}, R. Ellen R. Nisbet^{2,#}, José A. Fernández Robledo^{3,#}, Elena
5 Casacuberta^{4,#}, Lisa Sudek^{5,#}, Andrew E. Allen^{6,7}, Manuel Ares Jr.⁸, Cristina Aresté⁴, Cecilia
6 Balestreri⁹, Adrian C. Barbrook², Patrick Beardslee¹⁰, Sara Bender¹¹, David S. Booth¹²,
7 François-Yves Bouget¹³, Chris Bowler¹⁴, Susana A. Breglia¹⁵, Colin Brownlee⁹, Gertraud
8 Burger¹⁶, Heriberto Cerutti¹⁰, Rachele Cesaroni¹⁷, Miguel A. Chiurillo¹⁸, Thomas Clemente¹⁰,
9 Duncan B. Coles³, Jackie L. Collier¹⁹, Elizabeth C. Cooney²⁰, Kathryn Coyne²¹, Roberto
10 Docampo¹⁸, Christopher L. Dupont⁷, Virginia Edgcomb²², Elin Einarsson², Pía A.
11 Elustondo^{15,\$}, Fernan Federici²³, Veronica Freire-Beneitez^{24,25}, Nastasia J. Freyria³, Kodai
12 Fukuda²⁶, Paulo A. García²⁷, Peter R. Girguis²⁸, Fatma Gomaa²⁸, Sebastian G. Gornik²⁹, Jian
13 Guo^{5,8}, Vladimír Hampl³⁰, Yutaka Hanawa³¹, Esteban R. Haro-Contreras¹⁵, Elisabeth
14 Hehenberger²⁰, Andrea Highfield⁹, Yoshihisa Hirakawa³¹, Amanda Hopes³², Christopher J.
15 Howe², Ian Hu², Jorge Ibañez²³, Nicholas A.T. Irwin²⁰, Yuu Ishii³³, Natalia Ewa Janowicz³⁰,
16 Adam C. Jones¹¹, Ambar Kachale¹, Konomi Fujimura-Kamada³⁴, Binnypreet Kaur¹, Jonathan
17 Z. Kaye¹¹, Eleanna Kazana^{24,25}, Patrick J. Keeling²⁰, Nicole King¹², Lawrence A.
18 Klobutcher³⁵, Noelia Lander¹⁸, Imen Lassadi², Zhuhong Li¹⁸, Senjie Lin³⁵, Jean-Claude
19 Lozano¹³, Fulei Luan¹⁰, Shinichiro Maruyama³³, Tamara Matute²³, Cristina Miceli³⁶, Jun
20 Minagawa^{34,37}, Mark Moosburner^{6,7}, Sebastián R. Najle^{4,38}, Deepak Nanjappa²¹, Isabel C.
21 Nimmo², Luke Noble^{39,**}, Anna M.G. Novák Vanclová³⁰, Mariusz Nowacki¹⁷, Isaac Nuñez²³,
22 Arnab Pain^{40,41}, Angela Piersanti³⁶, Sandra Pucciarelli³⁶, Jan Pyrih^{1,24,30}, Joshua S. Rest⁴²,
23 Mariana Rius¹⁹, Deborah Robertson⁴³, Albane Ruaud^{23,***}, Iñaki Ruiz-Trillo^{4,44,45}, Monika A.
24 Sigg¹², Pamela A. Silver^{46,47}, Claudio H. Slamovits¹⁵, G. Jason Smith⁴⁸, Brittany N.
25 Sprecher³⁵, Rowena Stern⁹, Estienne Swart^{17,&}, Anastasios Tsaousis^{24,25}, Lev Tsypin^{49,50},

26 Aaron Turkewitz⁴⁹, Jernej Turnšek^{6,7,46,47}, Matus Valach¹⁶, Valérie Vergé¹³, Peter von
27 Dassow^{23,51}, Tobias von der Haar²⁴, Ross F. Waller², Lu Wang⁵², Xiaoxue Wen¹⁰, Glen
28 Wheeler⁹, April Woods⁴⁸, Huan Zhang³⁵, Thomas Mock^{32,*}, Alexandra Z. Worden^{5,53,*} &
29 Julius Lukeš^{1,*}

30

31 #Equal contribution

32 *Corresponding authors: dranov@paru.cas.cz (DF); t.mock@uea.ac.uk (TM);

33 azworden@geomar.de (AZW); jula@paru.cas.cz (JL)

34

35

36 ¹*Institute of Parasitology, Biology Centre, Czech Academy of Sciences and Faculty of
37 Sciences, University of South Bohemia, České Budějovice, Czech Republic*

38 ²*Department of Biochemistry, University of Cambridge, Cambridge, UK*

39 ³*Bigelow Laboratory for Ocean Sciences, East Boothbay, USA*

40 ⁴*Institut de Biologia Evolutiva, CSIC-Universitat Pompeu Fabra, Barcelona, Spain*

41 ⁵*Monterey Bay Aquarium Research Institute, Moss Landing, USA*

42 ⁶*Integrative Oceanography Division, Scripps Institution of Oceanography, University of
43 California, San Diego, USA*

44 ⁷*Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, USA*

45 ⁸*Molecular, Cell and Developmental Biology, University of California, Santa Cruz, USA*

46 ⁹*The Marine Biological Association, Plymouth and School of Ocean and Earth Sciences,
47 University of Southampton, UK*

48 ¹⁰*School of Biological Sciences, University of Nebraska, Lincoln, USA*

49 ¹¹*Gordon and Betty Moore Foundation, Palo Alto, USA*

50 ¹²*Department of Molecular and Cell Biology, University of California, Berkeley, USA*

- 51 ¹³*Sorbonne Université, CNRS UMR7621, Observatoire Océanologique, Banyul sur Mer,*
52 *France*
- 53 ¹⁴*Institute de Biologie de l'ENS, Département de biologie, École Normale Supérieure,*
54 *CNRS, INSERM, Paris, France*
- 55 ¹⁵*Centre for Comparative Genomics and Evolutionary Bioinformatics, Dalhousie*
56 *University, Halifax, Canada*
- 57 ¹⁶*Department of Biochemistry and Robert-Cedergren Centre for Bioinformatics and*
58 *Genomics, Université de Montréal, Montreal, Canada*
- 59 ¹⁷*Institute of Cell Biology, University of Bern, Bern, Switzerland*
- 60 ¹⁸*Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, USA*
- 61 ¹⁹*School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, USA*
- 62 ²⁰*Department of Botany, University of British Columbia, Vancouver, Canada*
- 63 ²¹*University of Delaware College of Earth, Ocean and Environment, Lewes, USA*
- 64 ²²*Woods Hole Oceanographic Institution, Woods Hole, MA, USA*
- 65 ²³*Facultad Ciencias Biológicas, Pontificia Universidad Católica de Chile, Fondo de*
66 *Desarrollo de Areas Prioritarias, Center for Genome Regulation and Millennium*
67 *Institute for Integrative Biology (iBio), Santiago de Chile, Chile*
- 68 ²⁴*School of Biosciences, University of Kent, Canterbury, Kent, UK*
- 69 ²⁵*Laboratory of Molecular and Evolutionary Parasitology, University of Kent, UK*
- 70 ²⁶*Graduate School of Life and Environmental Sciences, University of Tsukuba, Ibaraki,*
71 *Japan*
- 72 ²⁷*Department of Mechanical Engineering, Massachusetts Institute of Technology, Boston,*
73 *USA*
- 74 ²⁸*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge,*
75 *USA*

- 76 ²⁹*Centre for Organismal Studies, University of Heidelberg, Heidelberg, Germany*
- 77 ³⁰*Department of Parasitology, Faculty of Science, Charles University, BIOCEV, Vestec,*
78 *Czech Republic*
- 79 ³¹*Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan*
- 80 ³²*School of Environmental Sciences, University of East Anglia, Norwich, UK*
- 81 ³³*Graduate School of Life Sciences, Tohoku University, Sendai, Miyagi, Japan*
- 82 ³⁴*Division of Environmental Photobiology, National Institute for Basic Biology, Okazaki,*
83 *Aichi, Japan*
- 84 ³⁵*Department of Marine Sciences, University of Connecticut, Groton, USA*
- 85 ³⁶*School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy*
- 86 ³⁷*Department of Basic Biology, School of Life Science, Graduate University for Advanced*
87 *Studies, Okazaki, Aichi, Japan*
- 88 ³⁸*Instituto de Biología Molecular y Celular de Rosario, CONICET, and Facultad de*
89 *Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario,*
90 *Argentina*
- 91 ³⁹*Center for Genomics and Systems Biology, New York University, New York, USA*
- 92 ⁴⁰*Biological and Environmental Science and Engineering Division, King Abdullah*
93 *University of Science and Technology, Thuwal, Jeddah, Saudi Arabia*
- 94 ⁴¹*Center for Zoonosis Control, Global Institution for Collaborative Research and*
95 *Education, Hokkaido University, Sapporo, Japan*
- 96 ⁴²*Department of Ecology and Evolution, Stony Brook University, Stony Brook, USA*
- 97 ⁴³*Lasry Center for Biosciences, Clark University, Worcester, USA*
- 98 ⁴⁴*Departament de Genètica Microbiologia i Estadística, Universitat de Barcelona,*
99 *Barcelona, Catalonia, Spain*
- 100 ⁴⁵*Catalan Institution for Research and Advanced Studies, Barcelona, Catalonia, Spain*

101 ⁴⁶*Department of Systems Biology, Harvard Medical School, Boston, USA*

102 ⁴⁷*Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, USA*

103 ⁴⁸*Department of Environmental Biotechnology, Moss Landing Marine Laboratories, Moss*
104 *Landing, USA*

105 ⁴⁹*Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago,*
106 *USA*

107 ⁵⁰*Department of Biology, California Institute of Technology, Pasadena, USA*

108 ⁵¹*Instituto Milenio de Oceanografía de Chile, Concepción, Chile*

109 ⁵²*Institute of Oceanography, Minjiang University, Fuzhou, China*

110 ⁵³*Ocean EcoSystems Biology Unit, Marine Ecology Division, Helmholtz Centre for Ocean*
111 *Research, Kiel, Germany*

112

113

114 [§]*Present address: AGADA Biosciences Inc., Halifax, Canada*

115 ^{**}*Present address: Institute de Biologie de l'ENS, Département de biologie, École Normale*
116 *Supérieure, CNRS, INSERM, Paris, France*

117 ^{***}*Present address: Department of Microbiome Science, Max Planck Institute for*
118 *Developmental Biology, Tübingen, Germany*

119 [&]*Present address: Max Planck Institute of Developmental Biology, Tübingen, Germany*

120

121

122 **ABSTRACT**

123 Marine microbial eukaryotes underpin the largest food web on the planet and influence global
124 biogeochemical cycles that maintain habitability. They are also remarkably diverse and
125 provide insights into evolution, including the origins of complex life forms, as revealed

126 through genome analyses. However, their genetic tractability has been limited to a few
127 species that do not represent the broader diversity of eukaryotic life or some of the most
128 environmentally relevant taxa. Here, we report on genetic systems developed as a community
129 resource for experimental cell biology of aquatic protists from across the eukaryotic tree and
130 primarily from marine habitats. We present evidence for foreign DNA delivery and
131 expression in 14 species never before transformed, report on the advancement of genetic
132 systems in 7 species, review of an already published transformation protocol in 1 species and
133 discuss why the transformation of 17 additional species has not been achieved yet. For all
134 protists studied in this community effort, we outline our methods, constructs, and genome-
135 editing approaches in the context of published systems. The reported breakthroughs on
136 genetic manipulation position the community to dissect cellular mechanisms from a breadth
137 of protists, which will collectively provide insights into ancestral eukaryotic lifeforms,
138 protein diversification and evolution of cellular pathways.

139

140 **INTRODUCTION**

141 The ocean represents the largest continuous planetary ecosystem, hosting an enormous
142 variety of organisms. These range from some of the largest creatures on Earth to a vast
143 microscopic biota including unicellular eukaryotes (protists). Despite their small size, protists
144 play key roles in marine biogeochemical cycles and harbor tremendous evolutionary
145 diversity¹⁻³. Notwithstanding their significance for understanding the evolution of life on
146 Earth and their role in marine food webs, as well as driving biogeochemical cycles to
147 maintain habitability, little is known about their cell biology including reproduction,
148 metabolism, and signalling⁴. Most of the biological information available is based on
149 comparison of proteins from cultured genome-sequenced species to homologs in genetically
150 tractable model taxa, such as yeast⁵⁻⁹. A major impediment to understanding the cell biology

151 of these diverse eukaryotes is that protocols for genetic modification are only available for a
152 small number of species that represent neither the most ecologically relevant protists nor the
153 breadth of eukaryotic diversity.

154 The development of genetic tools requires reliable information about gene
155 organization and regulation of the emergent model species. Over the last decade, some of this
156 information has become available through genome^{5,6,8,10} and transcriptome sequencing
157 initiatives^{7,9,11,12} resulting in nearly 120 million unigenes from protists¹³. Insights from these
158 studies have enabled the phylogenetically-informed approach⁷ used for selecting and
159 developing key marine protists into model systems in the Environmental Model Systems
160 (EMS) Project presented herein. Forty-one scientific groups took part in the EMS Project, a
161 collaborative effort resulting in the development of genetic tools that significantly expand the
162 number of eukaryotic lineages, which can be manipulated, and which encompass multiple
163 ecologically important marine protists.

164 Here, we summarize detailed methodological achievements by this collaborative
165 effort and analyse results to provide a synthetic ‘Transformation Roadmap’ for creating new
166 microeukaryotic model systems. Although the organisms reported here are diverse, the paths
167 to overcome difficulties share similarities, highlighting the importance of building a well-
168 connected community to overcome technical challenges and accelerate the development of
169 genetic tools. The 14 emerging model species presented herein, and the collective set of
170 genetic tools from the overall collaborative project, will not only extend our knowledge of
171 cell biology, evolution and functional biodiversity, but also serve as platforms to advance
172 microbial biotechnology.

173

174 **RESULTS**

175 *Overview of taxa in the EMS Initiative*

176 Taxa were selected from multiple eukaryotic supergroups^{1,7} to maximize the potential to
177 compare fundamental aspects of cellular biology and to evaluate the numerous unigenes with
178 unknown functions found in marine protists (**Fig. 1**). Prior to the EMS initiative, reproducible
179 transformation of marine protists was limited to only a few species such as *Thalassiosira*
180 *pseudonana*, *Phaeodactylum tricornutum*, and *Ostreococcus tauri* (**Suppl. Table 1**). The
181 EMS initiative included 39 species, specifically, 6 Archaeplastids, 2 Haptophytes, 2
182 Rhizarians, 9 Stramenopiles, 12 Alveolates, 4 Discobans, and 4 Opisthokonts (**Fig. 1**). Most
183 of these taxa were isolated from coastal habitats, the focus area of several major culture
184 collections⁷. More than 50% of the selected species are considered photoautotrophs, with
185 another 35% divided between heterotrophic osmotrophs and phagotrophs, the remainder
186 being predatory mixotrophs. Almost 20% of the chosen species are symbionts and/or
187 parasites of marine plants or animals, 5% are associated with detritus, and several are
188 responsible for harmful algal blooms (**Suppl. Table 2**).

189 While some transformation systems for some protists have been developed in the
190 past^{17,18}, the challenge for this initiative was to develop novel genetic tools for species which
191 not only require different cultivation conditions but are also phenotypically extremely
192 diverse. It should be noted that not all major lineages could be pursued. For example,
193 amoebozoans did not feature in the initiative, in part because they tend to be most important
194 in soils, at least based on current knowledge, and manipulation systems exist for members of
195 this eukaryotic supergroup, in particular *Dictyostelium discoideum*¹⁴. Below we summarize
196 the taxa interrogated and overall EMS initiative outcomes (**Fig. 1**), provide evidence and
197 detailed protocols for 14 taxa, for which no transformation systems have been previously
198 reported (Category A), and 7 taxa for which existing protocols^{15,17,18,25,40,44,45,51} were
199 advanced (Category B) (**Figs. 2 and 3; Table 1**). We also review an already published
200 transformation protocol⁵⁶ in 1 species (Category C), and we discuss the unsuccessful

201 transformation of 17 additional taxa, covering important phylogenetic positions in the
202 eukaryotic tree of life (**Fig. 1**). Finally, we synthesize our findings in a roadmap for the
203 development of transformation systems in protists (**Fig. 4**).

204

205 **Archaeplastids**. Prasinophytes are important marine green algae distributed from polar to
206 tropical regions. They form a sister group to chlorophyte algae, and together, these two
207 groups branch adjacent to land plants, collectively comprising the Viridiplantae, which are
208 part of the Archaeplastida^{1,8} (**Fig. 1**). Genome sequences are available for the
209 picoprasinophytes (<3 µm cell diameter) tested herein, specifically, *Micromonas commoda*,
210 *Micromonas pusilla*, *Ostreococcus lucimarinus* and *Bathycoccus prasinus*. As part of the
211 EMS initiative, we report on the first genetic systems for *Bathycoccus*, a scaled, non-motile
212 genus, and *Micromonas*, a motile, naked genus with larger genomes than *Bathycoccus* and
213 *Ostreococcus*⁸. We also report on the first genetic systems for *Tetraselmis striata* and
214 *Ostreococcus lucimarinus*. The latter was transformed based on an adapted homologous
215 recombination system for *Ostreococcus tauri*¹⁹.

216 *O. lucimarinus* (RCC802) and *B. prasinus* (RCC4222) were transformed using
217 protocols adapted from *O. tauri*²⁰. Briefly, using electroporation for transfer of exogenous
218 genes, *O. lucimarinus* was transformed using a DNA fragment encoding the *O. tauri* high-
219 affinity phosphate transporter (HAPT) gene fused to a luciferase gene and a KanMX
220 selection marker (**Table 1; Suppl. Table 3**), which resulted in transient luciferase expression
221 24 h after electroporation (**Table 1; Fig. 3a**). After 2 weeks of growth in low melting agarose
222 plates containing G418 (1 mg/ml), 480 colonies were obtained, picked, and grown in artificial
223 seawater with the antibiotic neomycin. Of these, 76 displayed luminescence ≥ 2.5 fold above
224 background (80 Relative Luminescence Units, RLU), with widely variable levels (200 to

225 31020 RLU), likely reflecting either the site of integration and/or the number of integrated
226 genes (**Fig. 3a; Suppl. Fig. 1; Suppl. sheet 1**).

227 The *O. tauri* construct did not work in *B. prasinos*, while the use of the *B. prasinos*
228 histone H4 and high affinity phosphate transporter sequences in an otherwise identical
229 construct and conditions was successful. Although luciferase expression was not detected 24
230 h after electroporation, 48 G418-resistant colonies were obtained 2 weeks later, 20 being
231 luminescent when grown in liquid medium. Analysis of 14 resistant transformants revealed
232 that the luciferase sequence was integrated into the genome of 5 clones that were
233 luminescent, and one non-luminescent clone (**Fig. 3b; Suppl. sheet 1**), suggesting that the
234 chromatin context at integration sites in the latter was not favourable to luciferase expression.

235 Although transformation methods successful for *Bathycoccus* and *Ostreococcus* failed
236 in *Micromonas*, Lonza nucleofection was successful with *M. commoda* (CCMP2709) (**Table**
237 **1; Fig. 3c**) using 2 codon-optimized plasmids, one encoding the luciferase gene (NanoLuc,
238 Promega) flanked by an exogenous promoter and terminator sequence from the 5'- and 3'-
239 untranslated regions (UTRs) of histone H3 in *M. polaris*⁸, and the other encoding an eGFP
240 gene flanked by endogenous promoter and terminator sequences from the ribosomal protein
241 S9. Acclimated mid-exponential *M. commoda* cells grown in L1 medium at 21 °C were spun
242 at 5000 x *g* for 10 min, the pellet was resuspended in Buffer SF (Lonza) premixed with
243 carrier DNA (pUC19) and the plasmid, and 3×10^7 cells were used per reaction. After
244 applying the EW-113 pulse, 100 µl of ice-cold recovery buffer (10 mM HEPES-KOH, pH
245 7.5; 530 mM sorbitol; 4.7% [w/v] PEG 8000) was added to each well and incubated for 5 min
246 at room temperature. Each reaction was then transferred into 2 ml L1, placed at 21 °C and
247 light was increased stepwise over 72 h. Sensitivities to antibiotics were established (**Suppl.**
248 **Table 3**). Constructs did not include a selectable marker, as we aimed to introduce and
249 express foreign DNA while developing conditions suitable for transfection that supported

250 robust growth in this cell wall-lacking protist (**Table 1**). Transformants revealed a
251 significantly higher level of eGFP fluorescence than wild type cells, with 1.3% of the
252 population showing fluorescence per cell 45-fold higher than both the non-transformed
253 portion of the culture and the wild type cells (**Fig. 3c; Suppl. sheet 1**). Additionally, the RLU
254 was 1500-times higher than controls when using the luciferase-bearing construct, such that
255 multiple experiments with both plasmids confirmed expression of exogenous genes in *M.*
256 *commoda*.

257 *T. striata* (KAS-836) has been successfully transformed using microprojectile
258 bombardment (**Suppl. Fig. 2**). Two selectable marker genes were tested, consisting of a
259 putative promoter and 5' UTR sequences from the *T. striata* actin gene and either the coding
260 sequences of the *Streptoalloteichus hindustanus* bleomycin gene (conferring resistance to
261 zeocin) or the *Streptomyces hygroscopicus* bar gene (conferring resistance to glufosinate)
262 (**Table 1; Suppl. Fig. 2**). The terminator sequence was obtained from the *T. striata*
263 glyceraldehyde-3-phosphate dehydrogenase gene. Linearized plasmids were coated on gold
264 particles and introduced into *T. striata* cells by using the PDS-1000/He Particle Delivery
265 System (BioRad). Transformants were successfully selected on half-strength f/2 (seawater-
266 distilled water in a 1:1 volume ratio) agar plates containing either 150 µg/ml zeocin or 150
267 µg/ml glufosinate.

268

269 **Haptophytes (incertae sedis)**. Haptophytes are a group of photosynthetic protists that are
270 abundant in marine environments and include the major calcifying lineage, the
271 coccolithophores. Genome sequences are available for *Emiliana huxleyi*⁶ and
272 *Chrysochromulina tobin*²¹, and there is one report of nuclear transformation of a calcifying
273 coccolithophore species²². Here, as part of the EMS initiative, a stable nuclear transformation
274 system was developed for *Isochrysis galbana*, a species that lacks coccoliths, but represents

275 an important feedstock for shellfish aquaculture²³. However, transformation of *Emiliana*
276 *huxleyi*, which is the most prominent and ecologically significant coccolithophore, has not
277 been achieved yet²⁴.

278 *I. galbana* (CCMP1323) was transformed by biolistic bombardment with the pIgNAT
279 vector, which contains nourseothricin *N*-acetyltransferase (NAT, for nourseothricin
280 resistance) driven by the promoter and terminator of Hsp70 from *E. huxleyi* (CCMP1516).
281 Twenty four hours after bombardment, cells were transferred to liquid f/2 medium at 50%
282 salinity containing 80 µg/ml nourseothricin (NTC) and left to grow for 2-3 weeks to select for
283 transformants (**Table 1**). The presence of NAT in NTC-resistant cells was verified by PCR
284 and RT-PCR (**Fig. 3d, Suppl. Fig. 3**) and the sequence verified. To confirm NTC resistance
285 was a stable phenotype, cells were sub-cultured every 2-4 weeks at progressively higher NTC
286 concentrations (up to 150 µg/ml) in the above-mentioned media. Cells remained resistant to
287 NTC for approximately 6 months, as confirmed by PCR screening to identify the presence of
288 the NAT gene.

289
290 **Rhizarians.** Rhizarians include diverse non-photosynthetic protists, as well as the
291 photosynthetic chlorarachniophytes that acquired a plastid *via* secondary endosymbiosis of a
292 green alga. Uniquely, they represent an intermediate stage of the endosymbiotic process,
293 since their plastids still harbor a relict nucleus (nucleomorph)⁵. Here, we report on an
294 advanced transformation protocol for the chlorarachniophyte *Amorphochlora* (*Lotharella*)
295 *amoebiformis* for which low-efficiency transient transformation has previously been achieved
296 using particle bombardment²⁵.

297 *A. amoebiformis* (CCMP2058) cells (1×10^7) were resuspended in 100 µl of Gene
298 Pulse Electroporation Buffer (BioRad) with 20 to 50 µg of the reporter plasmid encoding
299 eGFP-RubisCO fusion protein under the control of the native *rbcS1* promoter and subjected

300 to electroporation (**Table 1**). Cells were immediately transferred to fresh ESM medium and
301 incubated for 24 h. Transformation efficiency was estimated by the fraction of cells
302 expressing eGFP, resulting in 0.03-0.1% efficiency, as enumerated by microscopy, showing
303 an efficiency up to 1,000-fold higher than the previous study²⁵ (**Table 1**). Stable
304 transformants were generated by manual isolation using a micropipette, and a transformed
305 line has maintained eGFP fluorescence for at least 10 months without antibiotic selection
306 (**Fig. 2; Fig. 3e**).

307

308 **Stramenopiles.** Stramenopiles are a diverse lineage with many important photoautotrophic,
309 mixotrophic and heterotrophic taxa. As the most studied class in this lineage, diatoms
310 (Bacillariophyceae) were early targets^{26,18} for the development of reverse genetics tool²⁴.
311 Diatoms are estimated to contribute approximately 20% of annual carbon fixation²⁷ and, like
312 several other algal lineages, are used in bioengineering applications and biofuels²⁸. The
313 transformation protocol for the Antarctic diatom *Fragilariopsis cylindrus*²⁹ represents the
314 first transformation system for a cold-adapted eukaryote. A first transformation protocol has
315 also been developed for *Pseudo-nitzschia multiseries*, a toxin-producing diatom³⁰. Here we
316 also present work for non-diatom stramenopiles, including the first transformation protocol
317 for the Eustigmatophyte *Nannochloropsis oceanica*, and an alternative protocol for the
318 Labyrinthulomycete *Aurantiochytrium limacinum*³¹, both of which are used for
319 biotechnological applications. Furthermore, we report on advances for CRISPR/Cas-driven
320 gene knockouts in *P. tricornutum*^{15,16} and a more efficient bacterial conjugation system for *T.*
321 *pseudonana*¹⁵.

322 Microparticle bombardment was used on *F. cylindrus* (CCMP1102) which was
323 grown, processed and maintained at 4 °C in 24 hours light. Exponential phase cells (5×10^7
324 cells) were harvested onto a 1.2 µm membrane filter (Millipore) which was then placed on an

325 1.5% agar Aquil plate for bombardment with beads coated with a plasmid containing zeocin
326 resistance and eGFP, both controlled by an endogenous fucoxanthin chlorophyll *a/c* binding
327 protein (FCP) promoter and terminator (**Table 1; Suppl. Table 3**)³². Transformation was
328 performed using 0.7 μm tungsten particles and the biolistic particle delivery system
329 PDS1000/He (BioRad). Rupture discs for 1,350 and 1,550 pounds per square inch (psi) gave
330 the highest colony numbers (efficiencies of 20.7 colony forming units (cfu)/ 10^8 cells and 30
331 cfu/ 10^8 cells), respectively. Following bombardment, the filter was turned upside down and
332 left to recover (24 h) on the plate, then cells were rinsed from the plate/filter and spread
333 across five 0.8% agar Aquil plates with 100 $\mu\text{g}/\text{ml}$ zeocin. Colonies appeared 3 to 5 weeks
334 later. PCR on genomic DNA showed that 100% and 60% of colonies screened positive for
335 the zeocin resistance and eGFP genes, respectively. Confirmed by FACS and microscopy,
336 eGFP was localized to the cytosol and was distinguishable from plastid autofluorescence
337 (**Fig. 2**). Additional confirmation by PCR and RT-PCR (**Fig. 3f**) revealed that the *ShBle* gene
338 for zeocin resistance and the gene encoding eGFP were present in the genomes of
339 transformants after multiple transfers (> 10) two years later, indicating long-term stability.

340 Bacterial conjugation methods were improved in *T. pseudonana* (CCMP1335) using
341 the silaffin precursor TpSil3p (**Table 1**) as the target gene. TpSil3p was fused to eGFP
342 flanked by an FCP promoter and terminator, cloned into a pTpPuc3 episomal backbone, and
343 transformed into mobilization plasmid-containing EPI300 *E. coli* cells (Lucigen). The donor
344 cells were grown in SOC medium at 37 °C until OD₆₀₀ of 0.3–0.4, centrifuged and
345 resuspended in 267 μl SOC medium. Next, 200 μl donor cells were mixed with 4×10^7 *T.*
346 *pseudonana* cells, co-cultured on pre-dried 1% agar plates, dark incubated at 30 °C for 90
347 min, then at 18 °C in constant light for 4 h, followed by selection in 0.25% agar pour plates
348 containing 100 $\mu\text{g}/\text{ml}$ NTC. Colonies were observed after 2 weeks, inoculated into 300 μl L1
349 medium and supplemented with 200 $\mu\text{g}/\text{ml}$ NTC to reduce the number of false positives.

350 Positive transformants were identified by colony PCR screening (**Suppl. Fig. 4**) and
351 epifluorescence microscopy (**Fig. 2**).

352 The diatom *Pseudo-nitzschia multiseriis* (15093C) and other members of this genus
353 form buoyant linear chains with overlapping cell tips during active growth, and were
354 uncondusive to punctate colony formation on agar, where their growth is generally poor. To
355 address this challenge, a low-gelation-temperature agarose seawater medium (LGTA) was
356 developed to facilitate growth, antibiotic selection, and cell recovery. *P. multiseriis* exhibited
357 growth inhibition at relatively low concentrations under NTC, formaldehyde, and zeocin
358 (**Suppl. Table 3**). Biolistic transformation of two other *Pseudo-nitzschia* species had been
359 demonstrated at low efficiency³³. To complement this approach and explore potentially
360 higher efficiency methods for transformation with diatom episomal plasmids, we modified
361 the existing conjugation-based method¹⁵. The published conjugation protocol was modified to
362 enhance *P. multiseriis* post-conjugation viability by reducing SOC content. An episomal
363 version of the Pm_actP_egfp_actT expression cassette was transfected into *E. coli*
364 EPI300+pTAMOB and used for conjugation (**Table 1**). After 48 h in L1 medium, cells were
365 plated in LGTA and eGFP-positive cells were observed 7 days later (**Fig. 2**). PCR revealed
366 the presence of plasmids in all eGFP positive colonies (**Suppl. Fig. 5**). Similarly, conjugation
367 with the episome pPtPUC3 (bleomycin selection marker)-containing bacterial donors was
368 followed under zeocin selection (200 µg/ml). After 7 days, only viable cells (based on bright
369 chlorophyll fluorescence) contained the episome, as confirmed by PCR. Propagation of
370 transformants after the first medium transfer (under selection) has so far been unsuccessful.

371 Stable transformation of *A. limacinum* (ATCC MYA-1381) was achieved by knock-in
372 of a resistance cassette composed of the bleomycin-resistance gene (*ShBle*) driven by 1.3 kb
373 promoter and 1.0 kb terminator regions of the endogenous glyceraldehyde-3-phosphate
374 dehydrogenase gene carried in a pUC19-based plasmid (18GZG) along with the native 18S

375 rRNA gene, and by knock-in of a similar construct containing a yeGFP:shble fusion (**Suppl.**
376 **Fig. 6**). Approximately 1×10^8 cells were electroporated (**Suppl. Table 4**), adapting the
377 electroporation protocol used for *Schizochytrium*³⁴. The highest transformation efficiency was
378 achieved using 1 μ g of linearized 18GZG plasmid with 2 pulses, resulting in a time constant
379 of ~ 5 ms (**Table 1**). Expression of the fusion protein was confirmed by both the zeocin-
380 resistance phenotype and the detection of eGFP (**Fig. 2**). Six 18GZG transformants derived
381 from uncut and linearized plasmids were examined in detail. All maintained antibiotic
382 resistance throughout 13 serial transfers, first in selective, and subsequently in non-selective
383 media, and then again in selective medium. Integration of the plasmid into the genome was
384 confirmed by PCR as well as by Southern blots using a digoxigenin-labeled ShBle gene
385 probe, showing that 4 transformants had integrations by single homologous recombination,
386 while in 2 transformants, additional copies of the antibiotic resistance cassette were
387 integrated by non-homologous recombination elsewhere in the genome (**Suppl. Fig. 6**).

388 Electroporation of *N. oceanica* (CCMP1779) was optimized based on observation of
389 cells treated with fluorescein-conjugated 2000 kDa dextran and subsequent survival (**Table**
390 **1**). A sorbitol concentration of 800 mM and electroporation at between 5 and 9 kV/cm
391 resulted in highest cell recovery. These conditions were used during introduction of plasmids
392 containing the gene for the blue fluorescent reporter mTagBFP2 under control of the
393 cytomegalovirus (CMV), the cauliflower Mosaic Virus 35S, or the VCP1 promoter
394 previously described from *Nannochloropsis* sp.³⁵. Transient expression of blue fluorescence
395 (compared to cells electroporated simultaneously under the same conditions but without
396 plasmid) appeared within 2 h, lasted for at least 24 h, and disappeared by 48 h in subsets of
397 cells electroporated with mTagBFP2 under the control of CMV (**Suppl. Fig. 7**). The transient
398 transformation was more effective when a linearized plasmid was used compared to a circular

399 plasmid (**Table 1**). VCP1 did not induce blue fluorescence with a circular plasmid, while 35S
400 gave inconsistent results with either circularized or linearized plasmids.

401 For *P. tricornutum* (CCAP1055/1), we adapted the CRISPR/Cas9 system¹⁶ for
402 multiplexed targeted mutagenesis. Bacterial conjugation¹⁵ was used to deliver an episome
403 that contained a Cas9 cassette and 2 single-guide RNA (sgRNA) expression cassettes
404 designed to excise a 38-bp domain from the coding region of a nuclear-encoded, chloroplastic
405 glutamate synthase (Phatr3_J24739) and introduce an in-frame stop codon after strand
406 ligation (**Table 1**). The GoldenGate assembly was used to clone 2 expression cassettes
407 carrying sgRNAs into a *P. tricornutum* episome that contained a Cas9-2A-ShBle expression
408 cassette and the centromeric region CenArsHis (**Suppl. Fig. 8**). After their addition to a *P.*
409 *tricornutum* culture, plates were incubated in a growth chamber under standard growth
410 conditions for 2 days and transformed *P. tricornutum* colonies began to appear after 14 days.
411 Only colonies maintaining Cas9-2A-ShBle sequence on the delivered episome were able to
412 grow on selection plates because Cas9 and the antibiotic-resistant gene (ShBle) were
413 transcriptionally fused by the 2A peptide³⁶ (**Suppl. Fig. 8**). Gel electrophoresis migration and
414 sequencing of the genomic target loci confirmed the 38-bp excision and premature stop
415 codon (**Fig. 3g**).

416

417 ***Alveolates***. This species-rich and diverse group is comprised of ciliates, apicomplexans, and
418 dinoflagellates (**Fig. 1**). As a link between apicomplexan parasites and dinoflagellate algae,
419 perkinsids are key for understanding the evolution of parasitism, and also have potential
420 biomedical applications. Techniques currently exist for transformation of only a small
421 number of ciliates, perkinsids and apicomplexans^{37,38}. Here, we present a first transformation
422 protocol for *Karlodinium veneficum* (CCMP1975), a mixotroph (combining photosynthetic
423 and phagotrophic nutrition) that produces fish-killing karlotoxins³⁹, and first evidence for

424 DNA delivery without gene expression in *Oxyrrhis marina* (CCMP 1788/CCMP 1795) and
425 *Crypthecodinium cohnii* (CCMP 316), see **Suppl. Results, Table 1 and Supp. Fig. 17)** .
426 Additionally, we report on improved transformation systems for *Perkinsus marinus*
427 (PRA240) and *Amphidinium carterae* (CCMP1314) chloroplast, published recently as part of
428 EMS initiative⁴⁰.

429 *K. veneficum* (CCMP1975) was transformed based on electroporation and cloning the
430 selectable marker gene aminoglycoside 3'-phosphotransferase (*nptII* / *neo*; note that *nptII* /
431 *neo* is used synonymously with amino 3'-glycosyl phosphotransferase gene (*aph(3')*)
432 conferring resistance to kanamycin, neomycin, paromomycin, ribostamycin, butirosin and
433 gentamicin B) into the backbone of the dinoflagellate-specific expression vector DinoIII-
434 *neo*⁴¹, which confers resistance to neomycin and kanamycin (**Table 1**). In brief, DinoIII-*neo*
435 was linearized and electroporated using Nucleofector (Lonza). The preprogrammed
436 Nucleofector optimization pulse codes, buffer SF/Solution I (Lonza), and 2 µg/µl of
437 linearized DinoIII-*neo* were used. Electroporated cells were selected under 150 µg/ml
438 kanamycin 3 days post-electroporation. Fresh seawater with kanamycin was added every 2
439 weeks to the cultures and new subcultures were inoculated monthly. After three months,
440 DNA and RNA were isolated from the resistant cultures as previously reported^{42,43} and
441 cDNA was synthesized using random hexamers. Out of 16 transformations, two cell lines
442 (CA-137, DS-138) showed stable growth under kanamycin selection. CA-137 developed
443 dense cultures after 3 months, and the resistance gene was detected in both DNA and RNA
444 through nested PCRs and RT-PCRs, respectively (**Suppl. Fig. 9, Fig. 3h**).

445 We improved the transformation protocol^{44,45} of *Perkinsus marinus*, a pathogen of
446 marine mollusks, fish, and amphibians⁴⁶, based on adapting the *A. carterae* transformation
447 system⁴⁷. In brief, a newly formulated transformation 3R buffer (200 mM Na₂HPO₄; 70 mM
448 NaH₂PO₄; 15 mM KCl; 1.5 mM CaCl₂; 150 mM HEPES-KOH, pH 7.3) was used to reduce

449 the cost of electroporation. We co-expressed 2 genes and efficiently selected transient and
450 stable transformants using FACS (**Figs. 2 and 3i, Suppl. Fig. 10, Table 1**). In addition, we
451 established the integration profile of ectopic DNA once introduced into the *P. marinus*
452 genome. We did not see evidence of integration through homologous recombination and
453 observed a propensity for plasmid fragmentation and integration within transposable elements
454 sites. An optimized alternative protocol for transformation using glass bead abrasion was also
455 developed. In brief, 5×10^7 cells were resuspended in 330 μ l of fresh ATCC Medium 1886
456 and were mixed with 5.0 μ g of linearized and circular [1:1] plasmid and 300 μ l of glass beads
457 (Sigma) in a 1.5 ml tube, vortexed for 30 s at maximum speed, and cells in 500 μ l of culture
458 medium were transferred to 6-well plates in a final volume of 3 ml. Two versions of the
459 previously published Moe gene promoter were tested⁴⁴. Whereas the 1.0 kb promoter version
460 induced expression after 2 or 3 days, the truncated version (0.5 kb) took 7 days for expression
461 to be detected. Resistance genes to bleomycin, blasticidin and puromycin have all been
462 shown to confer resistance to transformed *P. marinus*; however, selection regimes are still
463 relatively slow and inefficient, indicating further room for improvement⁴⁵.

464 We also report a new vector for the transformation of the *A. carterae* chloroplast, a
465 core, photosynthetic dinoflagellate. *A. carterae*, like other dinoflagellates with a peridinin-
466 containing chloroplast, contains a fragmented chloroplast genome made up of multiple
467 plasmid-like minicircles. The previous transformation protocols made use of this, to
468 introduce two vectors based on the *psbA* minicircle⁴⁰. Here, we show that other minicircles
469 are also suitable for use as vectors. We created a new artificial minicircle, using the *atpB*
470 minicircle as a backbone, but replacing the *atpB* gene with a codon-optimized
471 chloramphenicol acetyl transferase. This circular vector was introduced by biolistics to *A.*
472 *carterae* (**Suppl. Fig. 11**). Following selection with chloramphenicol, we were able to detect
473 transcription of the chloramphenicol acetyl transferase gene via RT-PCR (**Fig. 3j**). This result

474 suggests that all of the 20 or so minicircles in the dinoflagellate chloroplast genome would be
475 suitable for use as artificial minicircles, thus providing a large pool of potential vectors.

476 For *O. marina*, a basal-branching phagotroph that lacks photosynthetic plastids and *C.*
477 *cohnii*, a heterotroph used in food supplements, we achieved DNA delivery (**Table 1, Suppl.**
478 **Fig. 17**) but not the expression of the delivered genes yet.

479

480 ***Discobans***. This diverse group, recently split into Discoba and Metamonada⁴⁸, includes
481 heterotrophs, photoautotrophs, predatory mixotrophs, as well parasites. Discobans include
482 parasitic kinetoplastids with clinical significance, such as *Trypanosoma brucei*, *T. cruzi* and
483 *Leishmania* spp., for which efficient transformation protocols are available⁴⁹. However, such
484 protocols are missing for aquatic species. Here, we describe the first available transformation
485 protocols for the kinetoplastid *Bodo saltans* and the heterolobosean *Naegleria gruberi*. The
486 former was isolated from a lake, but identical 18S rRNA sequence have been reported from
487 the marine environment⁵⁰. The latter is a fresh water protist that represents a model organism
488 for closely related marine heterolobosean amoebas. Furthermore, we provide advanced
489 methods that build on published EMS results⁵¹ for the diplomemid *Diplonema papillatum*.

490 *B. saltans* (ATCC 30904) was transformed with a plasmid containing a cassette
491 designed to fuse an endogenous EF-1 α gene with eGFP for C-terminal tagging. This cassette
492 includes downstream of the eGFP, a *B. saltans* tubulin intergenic region followed by the
493 selectable marker gene *nptII/neo*, conferring resistance to neomycin. EF-1 α genes exist in
494 tandem repeats. The homologous regions that flank the cassette were chosen as targets for
495 inducing homology-directed repair, however they target only one copy of the gene. As
496 transcription in *B. saltans* is polycistronic⁵², insertion of the tubulin intergenic region into the
497 plasmid is essential for polyadenylation of the EF1- α /GFP fusion and *trans*-splicing of the
498 *nptII/neo* gene. Square-wave electroporation (Nepa21) was used with a poring pulse of

499 250V (25 ms) and 5 transfer pulses of 60V (99 ms) in the presence of Cytomix buffer (120
500 mM KCl; 0.15 mM CaCl₂; 10 mM KH₂PO₄; 2 mM EGTA; 5 mM MgCl₂; 25 mM HEPES-
501 KOH, pH 7.6). Selection of transfected cells began with 2 µg/ml of neomycin added 24 h
502 after electroporation, and this concentration was gradually increased over 2 weeks to 5 µg/ml
503 (**Table 1**). Cells were washed and subcultured into fresh selection medium every 4 days, and
504 neomycin-resistant cells emerged 7 to 9 days post-electroporation. The eGFP signal was
505 detected 2 days post-electroporation, albeit with low intensity. This may be due to the
506 inefficient translation of eGFP since it has not been codon-optimized for *B. saltans* (**Fig. 2**).
507 Genotyping analysis 9 months post-transfection confirmed the presence of the *nptII* / *neo*
508 gene and at least partial plasmid sequence (**Fig. 3k**; **Suppl. Fig. 12**). However, plasmid
509 integration into the *B. saltans* genome through homologous recombination is still
510 unconfirmed. This suggests either off-target plasmid integration or that the plasmid is
511 maintained episomally.

512 For *N. gruberi* (ATCC 30224) two plasmids were designed. The first one carried the
513 hygromycin B resistance gene (*hph*) with an actin promoter and terminator, along with an
514 HA-tagged eGFP driven by the ubiquitin promoter and terminator. The second plasmid
515 carried the *nptII* / *neo* gene instead. For each individual circular plasmid, 4 µg was
516 electroporated (**Table 1**). About 48 h after electroporation, dead cells were removed from the
517 suspension and viable cells were washed with PBS. Afterwards, 300 µg/ml of hygromycin B
518 or 700 µg/ml of neomycin was added to the fresh media. One to 4 weeks later, several
519 resistant clones were recovered and expression of eGFP and/or hygromycin was confirmed
520 by Western blotting (**Suppl. Fig. 13**). Expression of eGFP was observed by epifluorescence
521 microscopy (**Fig. 2**; **Suppl. Fig. 13**) with ~80% of transformants maintaining hygromycin B
522 or neomycin resistance in addition to expressing eGFP.

523 *D. papillatum* (ATCC 50162) was transformed by electroporation using 3 µg of *SwaI*-
524 linearised fragment (cut from p57-V5+NeoR plasmid) containing V5-tagged *nptII* / *neo* gene
525 flanked by partial regulatory sequences derived from the hexokinase gene of the kinetoplastid
526 *Blastocrithidia* (strain p57) (**Table 1**) using a published protocol⁵¹. About 18 h after
527 electroporation, 75 µg/ml G418 was added to the medium and after 2 weeks 7 neomycin-
528 resistant clones were recovered. Transcription of *nptII* / *neo* was verified in 4 clones by RT-
529 PCR (**Suppl. Fig. 14**) and the expression of the tagged *nptII* / *neo* protein was confirmed in 2
530 clones by Western blotting using α-V5 antibody (**Fig. 4I**).

531

532 *Opisthokonts*

533 The opisthokont clade Holozoa includes animals and their closest unicellular relatives
534 Choanoflagellates, Filastereans, Ichthyosporeans, and Corallochytreans. The establishment of
535 genetic tools in non-metazoan holozoans promises to help illuminate the cellular and genetic
536 foundations of animal multicellularity⁹. Genomic and transcriptomic data are available for
537 multiple representatives characterized by diverse cell morphologies, some of which can even
538 form multicellular structures^{9,12}. Until recently, only transient transformations have been
539 achieved for some opisthokonts such as the filasterean *Capsaspora owczarzaki*⁵³, the
540 ichthyosporean *Creolimax fragrantissima*⁵⁴ and the choanoflagellate *Salpingoeca rosetta*⁵⁵.
541 Through the EMS initiative, we report on the first evidence for transient transformation of the
542 ichthyosporean *Abeoforma whisleri*, isolated from the digestive tract of mussels, and review
543 on recently published stable transformation protocol for *S. rosetta* achieved by using the
544 selectable marker gene (puromycin N-acetyl-transferase - *PAC*) (**Fig. 2**)⁵⁶.

545 All *A. whisleri* life stages are highly sensitive to a variety of methods for
546 transformation. However, we developed a Lonza 4D-nucleofection-based protocol using 16-
547 well strips, wherein PBS-washed cells were resuspended in 20 µl of buffer P3 (Lonza)

548 containing 40 µg of carrier plasmid (empty pUC19) and 1-5 µg of the reporter plasmid (*A.*
549 *whisleri* H2B fused to mVenus fluorescent protein, mVFP) (**Table 1**), and subjected to code
550 EN-138 (Lonza). Immediately after the pulse, cells were recovered by adding 80 µl of marine
551 broth (Gibco) prior to plating in 12-well culture plates previously filled with 1 ml marine
552 broth. After 24 h, ~1% of the culture was transformed based on the fraction of cells
553 expressing mVFP in the nucleus (**Figs. 2 and 3m**).

554 *Microbial eukaryotes in natural planktonic communities*

555 Model organisms are typically selected based on criteria, such as relative ease of isolation
556 and asexual cultivation in the laboratory. These attributes may not correlate with the capacity
557 for uptake and expression of the exogenous DNA, so we explored whether such propensity
558 was common or rare in natural marine planktonic pico- and nanoeukaryote communities in a
559 culture-independent manner. Microbial plankton from natural seawater was concentrated and
560 electroporated with the plasmids containing mTagBFP2 under the control of CMV or 35S
561 promoters (see **Suppl. Results**). In most trials, blue fluorescent cells were very rare if
562 detected at all (compared to control samples). However, in one natural community tested a
563 particular picophytoeukaryote population was present where up to 50% transiently expressed
564 blue fluorescence when the CMV promoter was used (**Suppl. Fig. 15**). This suggests it might
565 be possible to selectively culture eukaryotic microorganisms based on capacity to express
566 exogenous DNA.

567

568 **Discussion**

569 Marine organisms play essential roles in global biogeochemical cycles and produce
570 approximately half of the Earth's oxygen^{1,27}. Decades of research by marine biologists,
571 ecologists, protistologists, and oceanographers have contributed to an increasingly coherent

572 picture of the oceanic ecosystem. These studies highlight the diversity of ocean life, including
573 the protistan component^{2,3}. Remarkable strides have also been made in developing an
574 overview of the genomes and predicted proteomes of these protists¹³. However, without
575 genetic manipulation systems, these taxa have remained an untapped resource for providing
576 deeper insights into their cell biology, with potentially valuable outcomes for evolutionary
577 studies, nanotechnology, biotechnology, medicine, and pharmacology.

578 The results of the EMS initiative were enhanced by the research teams having met and
579 conversed over a 3-year period. This facilitated identification and optimization of the steps
580 required to create new model systems in aquatic protists, ultimately culminating in our
581 roadmap for establishing new model organisms (**Fig. 4**). Successes and failures with
582 selectable markers, transformation conditions, and reporters were qualitatively compared
583 across species (**Suppl. Tables 3, 4 and 5**), and efforts were detailed using consistent
584 terminology (**Table 1; Figs. 2 and 3**). An important aspect of establishing a roadmap is to
585 also incorporate information on partially successful (e.g. DNA delivered) or failed
586 approaches as done herein (**Suppl. Results**).

587 For some of the selected species, the first step was to identify cultivation conditions
588 for robust growth in the laboratory to either generate high cell densities or large culture
589 volumes for obtaining sufficient biomass required for a variety of molecular biology
590 methods. Unlike established microbial model species, cultivation of marine protists can be
591 challenging especially under axenic conditions and for predatory taxa that require co-
592 cultivation with their prey. Nevertheless, 13 out of 35 species have been made axenic prior to
593 the development of transformation protocols. For the remaining species, we were unable to
594 remove bacteria and therefore had to make sure that the transformation signals were coming
595 from the targeted protist rather than contaminants (**Suppl. Table 2**). Subsequent steps
596 included the identification of suitable antibiotics and their corresponding selectable markers

597 (Table 1; Suppl. Table 3), conditions for introducing exogenous DNA (Table 1; Suppl.
598 Table 4), and selection of promoter and terminator sequences for designing transformation
599 vectors (Table 1; Suppl. Table 5; Suppl. Doc. 1).

600 As exemplified in the new systems provided herein (Figs. 2 and 3; Table 1), a variety
601 of methods were used to test whether exogenous DNA was integrated into the genome or
602 maintained as a plasmid, and whether the introduced genes were expressed. Approaches to
603 show the former included inverse PCR, Southern blotting and whole genome sequencing,
604 whereas approaches to demonstrate the latter included various combinations of PCR, RT-
605 PCR, Western blotting, epifluorescence microscopy, fluorescence-activated cell sorting
606 (FACS), antibody-based methods, and/or growth assays in the presence of antibiotics to
607 confirm transcription and translation of introduced selection and reporter genes (e.g., eGFP,
608 YFP, mCherry). For fluorescent markers, it was first ensured that the wild type, or
609 manipulated controls cells, had no signals conflicting with the marker (e.g. Fig. 2, Fig. 3C),
610 an important step because photosynthetic protists contain chlorophyll and other
611 autofluorescent pigments. Overall transformation outcomes for each species were parsed into
612 three groups according to the level of success or lack thereof (A = first transformation
613 protocol for a given species; B = advanced protocol based on prior work; C = published
614 protocol based on the EMS initiative) and are discussed below according to their
615 phylogenetic position (Fig. 1).

616 Our studies did not result in a universally applicable protocol, likely because
617 transformability and a range of other key conditions varied greatly across taxa and
618 approaches. Factors influencing outcomes include intrinsic features of the genome (e.g.,
619 presence/absence of homologous recombination, extrachromosomal elements, genome size,
620 level of ploidy), as well as morphology and structural features of the cell. In general,
621 electroporation proved the most common method for introducing exogenous DNA stably into

622 the cell. This approach was utilized for naked cells and protoplasts, yet frequently also
623 worked, albeit with lower efficiency, on cells protected by cell walls. Linearized plasmids
624 were most effective for delivery, and 5' and 3' UTRs-containing promoters of highly
625 expressed endogenous genes provided the strongest expression of selective reporters and
626 markers. The collaborative network facilitated many of the studies undertaken in this
627 initiative. The names and contact details of all co-authors assigned to particular species are
628 listed in **Suppl. Table 6**. Lessons from our integrated approach for instance include examples
629 such as the adoption of the highly sensitive reporter *nanoluc* (*Nluc*) gene by several teams for
630 initial rapid screening to find optimal conditions for transformation. If successful, teams
631 usually continued with fluorescence-based or other methods, as well as a hypertonic recovery
632 buffer. Furthermore, large amounts of carrier DNA usually facilitated successful initial
633 transformations (e.g. *M. commoda*, *A. whisleri*) or improved existing protocols (*S. rosetta*⁵⁵).

634 As outlined above, the global synthesis of approaches developed, or failed, herein,
635 and in published EMS studies provides a 'Transformation Roadmap' to guide future efforts to
636 establish or improve new and emergent model organisms (**Fig. 4**). However, it also allows us
637 to identify lineages that were particularly problematic – in that despite efforts of several
638 teams, successful expression of foreign DNA, whether transient or stable, was not achieved.
639 Here, dinoflagellates and the coccolithophore *E. huxleyi* were particularly recalcitrant to our
640 efforts and often, even if it appeared that DNA was delivered (**Suppl. Table 5**), expression of
641 encoded genes was not confirmed, and the number of transfected cells was very low. For
642 example, in the case of *C. cohnii*, DNA appeared ultimately to be delivered using Lonza
643 based approaches. However, after testing of multiple methods and plasmids (PmMOE:GFP-
644 11, UB-GFP, EF-GFP, PAY and PAYCO) there were no positive outcomes. Furthermore,
645 attempts to use the successful system developed for *A. carterae* for transforming its relative

646 *Symbiodinium microadriaticum* were unsuccessful⁵⁷. These and other lineages such as
647 coccolithophores, warrant concerted future efforts.

648 The combination of new results presented herein and others arising from the EMS
649 initiative⁵⁸ significantly expand the segment of extant eukaryotic diversity amenable to
650 reverse genetics approaches. Out of the 39 microbial eukaryotes selected for the initiative,
651 exogenous DNA was delivered and expressed in more than 50% of them. This high rate of
652 success testifies to the benefits of taking on high-risk research in a large, open, and
653 collaborative context⁵⁸. Moreover, the new systems reported herein open a wealth of
654 opportunities for exploring functional differences between members of relatively conserved
655 protein families shared across eukaryotes, or even domains of life. Collectively, these novel
656 and recently described protistan transformation systems enable us for the first time to shed
657 light on the function of species-specific genes which likely reflect key adaptations to specific
658 niches in dynamic ocean habitats.

659

660 **ACKNOWLEDGEMENTS**

661 We thank M. Salisbury and D. Lacono for assistance, C. Poirier and M. Hamilton (Monterey
662 Bay Aquarium Research Institute/GEOMAR) for FACS analysis; and V. K. Nagarajan, M.
663 Accerbi, and P. J. Green (University of Delaware) who carried out *Agrobacterium* studies in
664 *Heterosigma akashiwo*; and N. Kraeva, C. Bianchi and V. Yurchenko (University of Ostrava)
665 for the help with designing the p57-V5+NeoR construct. We are also grateful to the
666 protocols.io team for their support and thank four anonymous reviewers for their constructive
667 criticisms of our manuscript. This collaborative effort was supported by the Gordon and Betty
668 Moore Foundation EMS Program of the Marine Microbiology Initiative and other forms of
669 grant support to the participating laboratories.

670

671 **AUTHOR CONTRIBUTIONS**

672 The project was conceived and designed by A.C.J., J.Z.K., S.B., D.F., J.L., R.E.R.N.,
673 J.A.F.R., E.C., L.S., A.Z.W., T.M., A.E.A., F.Y.B, C.B, Ch.B., H.C., T.C., J.L.C., K.C.,
674 C.L.D., V.E., V.H., Y.H., C.J.H., P.J.K., N.K., S.L., C.M., J.M., I.R.T., P.A.S., C.H.S.,
675 G.J.S., A.T., P.V.D., A.T. and R.F.W. Data analysis was carried out by M.A.J., C.A., C.B.,
676 A.C.B., P.B., D.S.B., S.A.B., A.B., G.B., R.C., M.A.C., D.B.C., L.C., R.D., E.E., P.A.E.,
677 F.F., V.F.B., N.J.F., K.F., P.A.G., P.R.G., F.G., S.G., J.G., Y.H., E.R.H.C., E.H., A.Hi.,
678 A.Ho., I.H., J.I., N.A.T.I., Y.I., N.E.J., A.K., K.F.K., B.K., E.K., L.A.K., N.L., I.L., Z.L.,
679 J.C.L., F.L., S.M., T.M., M.M., S.R.N., D.N., I.C.N., L.N., A.M.G.N.V., M.N., I.N., A.Pa.,
680 A.Pi., S.P., J.P., J.S.R., M.R., D.R., A.R., M.A.S., E.S., B.S., R.S., T.V.H., L.T., J.T., M.V.,
681 V.V., L.W., X.W., G.W., A.W. and H.Z. The manuscript was written by D.F., R.E.R.N.,
682 J.A.F.R., E.C., L.S., T.M., A.Z.W. and J.L. with input from all authors.

683

684 **COMPETING INTERESTS**

685 The authors declare no competing interests.

686

687 **REFERENCES**

688

689

- 690 1. Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E. &
691 Keeling, P. J. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles
692 of microbes. *Science* **347**, 1257594 (2015).
- 693 2. de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney,
694 C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.
695 M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák,
696 A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco,
697 E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-

- 698 Lewis, S., Tara Oceans Coordinators., Acinas, S. G., Bork, P., Bowler, C., Gorsky, G.,
699 Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki,
700 M. E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P. & Karsenti,
701 E. Eukaryotic plankton diversity in the sunlit global ocean. *Science* **348**, 1261605 (2015).
- 702 3. Duarte, C. M., Regaudie-de-Gioux, A., Arrieta, J. M., Delgado-Huertas, A. & Agustí, S.
703 The oligotrophic ocean is heterotrophic. *Annu. Rev. Mar. Sci.* **5**, 551-569 (2013).
- 704 4. Collier, J. L. & Rest, J. S. Swimming, gliding, and rolling toward the mainstream: cell
705 biology of marine protists. *Mol. Biol. Cell.* **30**, 1245-1248 (2019).
- 706 5. Curtis, B. A., Tanifuji, G. & Burki, F. Algal genomes reveal evolutionary mosaicism and
707 the fate of nucleomorphs. *Nature* **492**, 59–65 (2012).
- 708 6. Read, B. A., Kegel, J., Klute, M. J., Kuo, A., Lefebvre, S. C., Maumus, F., Mayer, C.,
709 Miller, J., Monier, A., Salamov, A., Young, J., Aguilar, M., Claverie, J. M., Frickenhaus,
710 S., Gonzalez, K., Herman, E. K., Lin, Y. C., Napier, J., Ogata, H., Sarno, A. F., Shmutz,
711 J., Schroeder, D., de Vargas, C., Verret, F., von Dassow, P., Valentin, K., Van de Peer, Y.,
712 Wheeler, G., *Emiliana huxleyi* Annotation Consortium, Dacks, J. B., Delwiche, C. F.,
713 Dyhrman, S. T., Glöckner, G., John, U., Richards, T., Worden, A. Z, Zhang, X. &
714 Grigoriev, IV. Pan genome of the phytoplankton *Emiliana* underpins its global
715 distribution. *Nature* **499**, 209-213 (2013).
- 716 7. Keeling P. J., Burki, F., Wilcox, H. M., Allam, B. & Allen, E. E. The Marine Microbial
717 Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional
718 diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* **12**,
719 e1001889 (2014).

- 720 8. van Baren, M. J., Bachy, C., Reistetter, E. N., Purvine, S. O., Grimwood, J., Sudek, S., Yu,
721 H., Poirier, C., Deerinck, T. J., Kuo, A., Grigoriev, IV., Wong, C. H., Smith, R. D.,
722 Callister, S. J., Wei, C. L., Schmutz, J. & Worden, A. Z. Evidence-based green algal
723 genomics reveals marine diversity and ancestral characteristics of land plants. *BMC*
724 *Genomics* **17**, 1–22 (2016).
- 725 9. Richter, D. J., Fozouni, P., Eisen, M. & King, N. Gene family innovation, conservation
726 and loss on the animal stem lineage. *eLife* **7**, 1–43 (2018).
- 727 10. Strassert, J. F. H., Karnkowska, A., Hehenberger, E., Del Campo, J., Kolisko, M.,
728 Okamoto, N., Burki, F., Janouškovec, J., Poirier, C., Leonard, G., Hallam, S. J., Richards,
729 T. A., Worden, A. Z., Santoro, A. E. & Keeling, P. J. Single cell genomics of uncultured
730 marine alveolates shows paraphyly of basal dinoflagellates. *ISME J.* **12**, 304-308 (2018).
- 731 11. Joseph, S. J., Fernández-Robledo, J. A., Gardner, M. J., El-Sayed, N. M., Kuo, C. H.,
732 Schott, E. J., Wang, H., Kissinger, J. C. & Vasta, G. R. The Alveolate *Perkinsus marinus*:
733 Biological insights from EST gene discovery. *BMC Genomics* **11**, 228 (2010).
- 734 12. Grau-Bové, X., Torruella, G., Donachie, S., Suga, H., Leonard, G., Richards, T. A. & Ruiz-
735 Trillo, I. Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife* **6**,
736 e26036 (2017).
- 737 13. Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A. & Seeleuthner, Y. A global atlas of
738 eukaryotic genes. *Nat. Commun.* **9**, 373 (2018).
- 739 14. Paschke, P., Knecht, D. A., Williams, T. D., Thomason, P. A., Insall, R. H., Chubb, J. R.,
740 Kay, R. R. & Veltman, D. M. Genetic engineering of *Dictyostelium discoideum* cells based
741 on selection and growth on bacteria. *J. Vis. Exp.* **143**, e58981 (2019).

- 742 15. Karas, B. J., Diner, R. E., Lefebvre, S. C., McQuaid, J., Phillips, A. P., Noddings, C. M.,
743 Brunson, J. K., Valas, R. E., Deerinck, T. J., Jablanovic, J., Gillard, J. T., Beeri, K.,
744 Ellisman, M. H., Glass, J. I., Hutchison, C. A. 3rd., Smith, H. O., Venter, J. C., Allen, A.
745 E., Dupont, C. L. & Weyman, P. D. Designer diatom episomes delivered by bacterial
746 conjugation. *Nat. Commun.* **6**, 6925 (2015).
- 747 16. Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M. & Winge, P. A. CRISPR/Cas9
748 system adapted for gene editing in marine algae. *Sci. Rep.* **6**, 24951 (2016).
- 749 17. Hopes, A., Nekrasov, V., Kamoun, S. & Mock, T. Editing of the urease gene by CRISPR-
750 Cas in the diatom *Thalassiosira pseudonana*. *Plant Methods* **12**, 49 (2016).
- 751 18. Apt, K. E., Kroth-Pancic, P. G. & Grossman, A. R. Stable nuclear transformation of the
752 diatom *Phaeodactylum tricornutum*. *Mol. Gen. Genet.* **252**, 572-579 (1996).
- 753 19. Lozano, J. C., Schatt, P., Botebol, H., Vergé, V., Lesuisse, E., Blain, S., Carré, I. A. &
754 Bouget, F. Y. Efficient gene targeting and removal of foreign DNA by homologous
755 recombination in the picoeukaryote *Ostreococcus*. *Plant J.* **78**, 1073-1083 (2014).
- 756 20. Van Ooijen, G., Knox, K., Kis, K., Bouget, F. Y. & Millar, A. J. Genomic transformation
757 of the picoeukaryote *Ostreococcus tauri*. *J. Vis. Exp.* **65**, e4074 (2012).
- 758 21. Hovde, B. T., Deodato, C. R., Hunsperger, H. M., Ryken, S. A., Yost, W., Jha, R. K.,
759 Patterson, J., Monnat, R. J. Jr., Barlow, S. B., Starkenburg, S. R. & Cattolico, R. A.
760 Genome sequence and transcriptome analyses of *Chrysochromulina tobin*: Metabolic tools
761 for enhanced algal fitness in the prominent order Prymnesiales (Haptophyceae). *PLoS*
762 *Genet.* **11**, e1005469 (2015).

- 763 22. Endo, H., Yoshida, M., Uji, T., Saga, N., Inoue, K. & Nagasawa, H. Stable nuclear
764 transformation system for the coccolithophorid alga *Pleurochrysis carterae*. *Sci. Rep.* **6**,
765 22252 (2016).
- 766 23. Dörner, J., Carbonell, P., Pino, S. & Farias, A. Variation of fatty acids in *Isochrysis*
767 *galbana* (T-Iso) and *Tetraselmis suecica*, cultured under different nitrate availabilities. *J.*
768 *Fish. Aquacult.* **5**, 1-3 (2014).
- 769 24. Velmurugan, N. & Deka, D. Transformation techniques for metabolic engineering of
770 diatoms and haptophytes: current state and prospects. *Appl Microbiol Biotechnol.* **102**,
771 4255-4267 (2018).
- 772 25. Hirakawa, Y., Kofuji, R. & Ishida, K. Transient transformation of a chlorarachniophyte
773 alga, *Lotharella amoebiformis* (chlorarachniophyceae), with uidA and egfp reporter genes.
774 *J. Phycol.* **44**, 814–20 (2008).
- 775 26. Dunahay, T. G., Jarvis, E. E. & Roessler, P. G. Genetic transformation of the diatoms
776 *Cyclotella cryptica* and *Navicula saprophila*. *J. Phycol.* **31**, 1004–1012 (1995).
- 777 27. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of
778 the biosphere: Integrating terrestrial and oceanic components. *Science* **281**, 237-240
779 (1998).
- 780 28. Mishra, M., Arukha, A. P., Bashir, T., Yadav, D. & Prasad, G. B. K. S. All new faces of
781 diatoms: Potential source of nanomaterials and beyond. *Front. Microbiol.* **8**, 1239 (2017).
- 782 29. Mock, T., Otilar, R. P., Strauss, J., McMullan, M., Paajanen, P., Schmutz, J., Salamov, A.,
783 Sanges, R., Toseland, A., Ward, B. J., Allen, A. E., Dupont, C. L., Frickenhaus, S.,
784 Maumus, F., Veluchamy, A., Wu, T., Barry, K. W., Falciatore, A., Ferrante, M. I.,
785 Fortunato, A. E., Glöckner, G., Gruber, A., Hipkin, R., Janech, M. G., Kroth, P. G., Leese,

- 786 F., Lindquist, E. A., Lyon, B. R., Martin, J., Mayer, C., Parker, M., Quesneville, H.,
787 Raymond, J. A., Uhlig, C., Valas, R. E., Valentin, K. U., Worden, A. Z., Armbrust, E. V.,
788 Clark, M. D., Bowler, C., Green, B. R., Moulton, V., van Oosterhout, C. & Grigoriev, IV.
789 Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**,
790 536–540 (2017).
- 791 30. Brunson, J. K., McKinnie, S. M. K., Chekan, J. R., McCrow, J. P., Miles, Z. D., Bertrand,
792 E. M., Bielinski, V. A., Luhavaya, H., Oborník, M., Smith, G. J., Hutchins D. A., Allen,
793 A. E. & Moore, B. S. Biosynthesis of the neurotoxin domoic acid in a bloom-forming
794 diatom. *Science* **361**, 1356-1358 (2018).
- 795 31. Sakaguchi K., Matsuda T., Kobayashi T., Ohara J., Hamaguchi R., Abe E., Nagano N.,
796 Hayashi M., Ueda M., Honda D., Okita Y., Taoka Y., Sugimoto S., Okino N. & Ito M.
797 Versatile transformation system that is applicable to both multiple transgene expression
798 and gene targeting for thraustochytrids. *Applied and Environmental Microbiology* **78**:
799 3193-3202 (2012).
- 800 32. Kroth, P. G. Genetic Transformation; A tool to study protein targeting in diatoms. *Methods*
801 *Mol. Biol.* **390**, 257–267 (2007).
- 802 33. Sabatino, V., Russo, M. T., Patil, S., d'Ippolito, G., Fontana, A. & Ferrante, M. I.
803 Establishment of genetic transformation in the sexually reproducing diatoms *Pseudo-*
804 *nitzschia multistriata* and *Pseudo-nitzschia arenysensis* and inheritance of the transgene.
805 *Marine Biotech.* **17**, 452-462 (2015).
- 806 34. Ono, K., Aki, T. & Kawamoto, S. Method for introducing a gene into labyrinthulomycota.
807 US Patent 7,888,123 (2011).

- 808 35. Kilian, O., Benemann, C. S., Niyogi, K. K. & Vick, B. High-efficiency homologous
809 recombination in the oil-producing alga *Nannochloropsis sp.* *Proc. Natl. Acad. Sci. USA*
810 **108**, 21265-21269 (2011).
- 811 36. Duda, K., Lonowski, L. A., Kofoed-Nielsen, M., Ibarra, A., Delay, C. M., Kang, Q., Yang,
812 Z., Pruett-Miller, S. M., Bennett, E. P., Wandall, H. H., Davis, G. D., Hansen, S. H. &
813 Frödin, M. High-efficiency genome editing via 2A-coupled co-expression of fluorescent
814 proteins and zinc finger nucleases of CRISPR/Cas9 nuclease pairs. *Nucl. Acids. Res.* **42**, e84
815 (2014).
- 816 37. Donald, R. G. & Roos, D. S. Stable molecular transformation of *Toxoplasma gondii*: a
817 selectable dihydrofolate reductase-thymidylate synthase marker based on drug-resistance
818 mutations in malaria. *Proc. Natl. Acad. Sci. USA* **90**, 11703-11707 (1993).
- 819 38. Rug, M. & Maier, A. G. Transfection of *Plasmodium falciparum*. *Methods Mol. Biol.* **923**,
820 75-98 (2013).
- 821 39. Barbrook, A. C., Howe, C. & Nisbet, R. Breaking up is hard to do: the complexity of the
822 dinoflagellate chloroplast genome. *Perspect. Phycol.*
823 <https://doi.org/10.1127/pip/2018/0084> (2018).
- 824 40. Nimmo, I. C., Barbrook, A. C., Chen, J. E., Geisler, K., Smith, A. G., Aranda, A., Purton,
825 P., Waller, R. F., Nisbet, R. E. R. & Howe, C. J. Genetic transformation of the
826 dinoflagellate chloroplast. *eLife* **8**, e45292 (2019).
- 827 41. Sprecher, B. N., Zhang, H. & Lin, S. Nuclear gene transformation in a dinoflagellate.
828 *bioRxiv* 602821 (2019).

- 829 42. Zhang, H., Campbell, D. A., Sturm, N. R., Rosenblad, M. A., Dungan, C. F. & Lin, S.
830 Signal recognition particle RNA in dinoflagellates and the perkinsid *Perkinsus marinus*.
831 *Protist* **164**, 748-761 (2013).
- 832 43. Lin, S., Zhang, H. & Dubois, A. Low abundance distribution of *Pfiesteria piscicida* in
833 Pacific and Western Atlantic as detected by mtDNA-18S rDNA real-time polymerase
834 chain reaction. *J. Plankton Res.* **28**, 667-681 (2006).
- 835 44. Fernández Robledo, J. A., Lin, Z. & Vasta, G. R. Transfection of the protozoan parasite
836 *Perkinsus marinus*. *Mol. Biochem. Parasitol.* **157**, 44-53 (2008).
- 837 45. Sakamoto, H., Hirakawa, Y., Ishida, K. I., Keeling, P. J., Kita, K. & Matsuzaki M.
838 Puromycin selection for stable transfectants of the oyster-infecting parasite *Perkinsus*
839 *marinus*. *Parasitol. Int.* **69**, 13-16 (2018).
- 840 46. Chambouvet, A., Gower, D. J., Jirků, M., Yabsley, M. J., Davis, A. K., Leonard, G.,
841 Maguire, F., Doherty-Bone, T. M., Bittencourt-Silva, G. B., Wilkinson, M. & Richards,
842 T. A. Cryptic infection of a broad taxonomic and geographic diversity of tadpoles by
843 Perkinsea protists. *Proc. Natl. Acad. Sci. USA* **112**, E4743-4751 (2015).
- 844 47. Ten Lohuis, M. R. & Miller, D. J. Genetic transformation of dinoflagellates (*Amphidinium*
845 and *Symbiodinium*): expression of GUS in microalgae using heterologous promoter
846 constructs. *Plant J.* **13**, 427-435, (1998).
- 847 48. Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., Agatha, S., Berney,
848 C., Brown, M. W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., Del Campo, J.,
849 Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A. A., Hoppenrath,
850 M., James, T. Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A.,
851 Lahr, D. J. G., Lara, E., LeGall, L., Lynn, D. H., Mann, D. G., Massana, R., Mitchell, E.

- 852 A. D., Morrow, C., Park, J. S., Pawlowski, J. W., Powell, M. J., Richter, D. J., Rueckert,
853 S., Shadwick, L., Shimano, S., Spiegel, F. W., Torruella, G., Youssef, N., Zlatogursky, V.
854 & Zhang, Q. Revision to the classification, nomenclature and diversity of eukaryotes. *J.*
855 *Euk. Microbiol.* **66**, 4-119 (2019).
- 856 49. Matthews, K. R. 25 years of African trypanosome research: From description to molecular
857 dissection and new drug discovery. *Mol. Biochem. Parasitol.* **200**, 30-40 (2015).
- 858 50. Jackson, A. P., Quail, M. A. & Berriman, M. Insights into the genome sequence of a free-
859 living Kinetoplastid: *Bodo saltans* (Kinetoplastida: Euglenozoa). *BMC Genomics* **9**, 594
860 (2008).
- 861 51. Kaur, B., Valach, M., Peña-Díaz, P., Moreira, S., Keeling, P. J., Burger, G., Lukeš, J. &
862 Faktorová, D. Transformation of *Diplonema papillatum*, the type species of the highly
863 diverse and abundant marine micro-eukaryotes Diplonemida (Euglenozoa). *Env.*
864 *Microbiol.* **20**, 1030-1040 (2018).
- 865 52. Opperdoes, F. R., Butenko, A., Flegontov, P., Yurchenko, V. & Lukeš, J. Comparative
866 metabolism of free-living *Bodo saltans* and parasitic trypanosomatids. *J. Eukaryot.*
867 *Microbiol.* **63**, 657-678 (2016).
- 868 53. Parra-Acero, H., Ros-Rocher, N., Perez-Posada, A., Kożyczkowska, A., Sánchez-Pons, N.,
869 Nakata, A., Suga, H., Najle, S. R. & Ruiz-Trillo, I. Transfection of *Capsaspora owczarzaki*,
870 a close unicellular relative of animals. *Development* **145**, 162107 (2018).
- 871 54. Suga, H. & Ruiz-Trillo, I. Development of ichthyosporean sheds light on the origin of
872 metazoan multicellularity. *Dev. Biol.* **377**, 284-292 (2013).
- 873 55. Booth, D., Middleton, H. & King, N. Choanoflagellate transfection illuminates their cell
874 biology and the ancestry of animal septins. *Mol. Biol. Cell.* **29**, 3026-38 (2018).

- 875 56. Wetzel, L. A., Levin, T. C., Hulett, R. E., Chan, D., King, G. A., Aldayafleh, R., Booth, D.
876 S., Sigg, M. A. & King, N. Predicted glycosyltransferases promote development and
877 prevent spurious cell clumping in the choanoflagellate *S. rosetta*. *eLife* **7**, e41482 (2018).
- 878 57. Chen, J. E., Barbrook, A. C., Cui, G., Howe, C. J. & Aranda, M. The genetic intractability
879 of *Symbiodinium microadriaticum* to standard algal transformation methods. *PLoS ONE*
880 **14**, e0211936 (2019).
- 881 58. Waller, R. F., Cleves, P. A., Rubio-Brotons, M., Woods, A., Bender, S. J., Edgcomb, V.,
882 Gann, E. R., Jones, A. C., Teytelman, L., von Dassow, P., Wilhelm, S. W. & Collier, J. L.
883 Strength in numbers: collaborative science for new experimental model systems. *PLoS*
884 *Biol.* **16**, e2006333 (2018).

885

886

887 **FIGURE LEGENDS**

888

889 **Fig. 1. | Phylogenetic relationships and transformation status of marine protists.** A
890 phylogenetically-informed approach was used to select protists for concerted genetic
891 manipulation efforts. A schematic view of the eukaryotic tree of life with effigies of main
892 representatives. Colour-coordinated species we have attempted to genetically modify, are
893 listed below. Current transformability status is schematized in circles indicating: DNA
894 delivered and shown to be expressed (yellow; for details see text and Table 1); DNA
895 delivered, but no expression seen (grey); no successful transformation achieved despite
896 efforts (blue). The details of transformation of species that belong to “DNA delivered” and
897 “Not achieved yet” categories are described in the Supplementary Data. Delivery methods
898 and construct types are shown pictorially. Overall, protocols and specific reagents are

899 available to transfect 22 protist species belonging to 7 eukaryotic supergroups, for 14 species
900 we show first successful attempt of either stable or transient transformation, for 7 species an
901 alternative transformation or improvement of existing transformation protocol is shown and
902 for 1 species we review an already published or existing protocol.

903

904 **Fig. 2. | Epifluorescence micrographs of transformed marine protists**

905 Transformants and wild-type cell lines of 10 selected protist species. Colored boxes behind
906 species names refer to phylogenetic supergroup assignments given in Fig. 1. For each of these
907 fluorescence micrographs details of transformation protocol as well as a second form of
908 evidence are described in the Results and summarized in Table 1.

909 Scale bars are as follows: 10 μm for *A. amoebiformis*, *T. pseudonana*, *A. limacinum*, *B.*

910 *saltans*, *N. gruberi*, *A. whisleri*, and *S. rosetta*; 15 μm for *P. marinus*; 20 μm for *F. cylindrus*;

911 100 μm for *P. multiseriis*.

912

913 **Fig. 3. | Various methods were used to demonstrate successful transformation in** 914 **different species.**

915 Luminescence, Fluorescence (by FACS and epifluorescence), Western blot, RT-PCR or
916 sequencing (in case of Cas9-induced excision by CRISPR) were used to verify the expression
917 of introduced constructs in three Archaeplastids (**a, b, c**), one Haptophytan (**d**), one Rhizarian
918 (**e**), two Stramenopiles (**f, g**), three Alveolates (**h, i, j**), two Discobans (**k, l**) and one
919 Opisthokont (**m**). Note that *nptII* / *neo* is used synonymously with amino 3'-glycosyl
920 phosphotransferase gene (*aph(3')*) conferring resistance to kanamycin, neomycin,
921 paromomycin, ribostamycin, butirosin and gentamicin B. (**a**) *In vivo* luminescence of 8 *O.*
922 *lucimarinus* G418-resistant transformants depicted as relative luminescence unit (RLU) per 5
923 s and a corresponding gel showing PCR amplification from DNA of transformants of the

924 whole *pH4:KanMX pHAPT:luc* transgene. A band of the expected size was amplified in all
925 transformants and absent in WT (control cells electroporated without constructs) and the PCR
926 negative control (H₂O). **(b)** *In vivo* luminescence of 14 *B. prasinus* transformants resistant to
927 G418 depicted as RLU per 5 s and a corresponding gel showing PCR amplification from
928 DNA of transformants of *pH4:KanMx* (primers 1 and 2) and *pHAPT:luc* sequences (primers
929 3 and 4). **(c)** FACS analysis of *M. commoda* cells in two treatments: controls (wild type [WT]
930 or “no-pulse”, in which *eGFP* constructs were added but no electroporation pulse applied),
931 and in the treatment to which the pulse was applied (“Transformed”). Note that lower panels
932 include only the population of healthy cells selected using the depicted gates in upper panels.
933 No compensation was used and PMT voltages were identical. Bar graphs show the mean and
934 standard deviation of eGFP fluorescence from biological triplicates in the same experiment,
935 analyzed as control (WT, i.e., “no-pulse” on cultures with constructs present) and pulsed
936 (EW-113 treatment, with cells in the latter treatment analyzed as non-transformed, all eGFP
937 and high eGFP (cells for which eGFP fluorescence was an order of magnitude higher than
938 controls). Transformation efficiencies also reflect mean and standard deviations of biological
939 triplicates. Images show natural chlorophyll fluorescence (red, top left), eGFP (green, top
940 right) from a transformed cell (indicated by white arrow in all panels), with DAPI DNA
941 staining (blue, bottom left), and the overlay of all three (bottom right) demonstrating the
942 eGFP fluorescence is localized to the nucleus as designed. The scale bar represents 5 μ m. **(d)**
943 Expression of nourseothricin N-acetyl transferase (*nat*) in pIgNAT transformed *I. galbana*,
944 from cells exhibiting resistance to nourseothricin. Expression of the *nat* transgene in WT
945 (lane 3) and transformants (lanes 1 and 2) of *I. galbana* verified by RT-PCR. RT- negative
946 and no template controls are shown in lanes 4 and 5, respectively, and 50 bp ladder (Bioline,
947 lane 6), positive control PCR of genomic DNA from transformed cells (lane 7). **(e)** Western
948 blot of expressed eGFP in *A. amoebiformis*, with total proteins extracted from transformants

949 (TF) and WT cells using α -GFP monoclonal antibody (Takara; 1:1,000), and an α -mouse
950 horseradish peroxidase (HRP)-coupled antibody (GE Healthcare; 1:10,000). (f) Expression of
951 *eGFP*, *ShBle* and *Rhodopsin* (endogenous gene used as control) in WT (lane 2) and
952 transformants (1350 FC2: lane 4, 1100 FC2: lane 6) of *F. cylindrus* verified by RT-PCR. RT-
953 negative controls are shown in lanes 3, 5 and 7. The NEB 100 kb ladder is shown in lane 1.
954 (g) Cas9-induced excision of a 38-bp region in the *cGOGAT* gene (Phatr3_J24739). The
955 sequence alignments from 10 *P. tricornutum* mutants (24739 KO-1 to KO-10) and two WT
956 cell lines are shown. The two sgRNA target loci (g24739-A and g24739-B) are shown by
957 green bars and include both the 20-nt sequence loci and the 3-nt PAM (protospacer-adjacent
958 motif) sequence. The Cas9 double-stranded break sites are represented by the red rectangle.
959 The predicted excision genotype is shown and represents the excision by dashes; the 38-bp
960 excision is followed by an in-frame TGA stop codon. The excision/mutagenesis frequency is
961 shown in the table below. (h) RT-PCR of the *nptII/neo* gene in kanamycin resistant *K.*
962 *veneficum* transformants. No template negative control (lane 1); RT- negative controls of WT
963 and transformed cells (lanes 2 and 4, respectively); cDNA from WT cells grown without
964 antibiotics (lane 3), cDNA from transformants grown under kanamycin selection (lane 5),
965 PCR positive control using the original vector DNA (DinoIII-*neo*, lane 6), GeneRuler DNA
966 ladder (lane 7). Nothing was loaded in lane 8. cDNA libraries were made with 200 ng of
967 RNA from WT and CA-137 transformant and all PCRs were nested. (i) Western blot of
968 expressed MOE-GFP in *P. marinus* transformants. Cells subjected to electroporation without
969 the plasmid present (MOCK) and WT were included as controls. Polyclonal rabbit α -GFP
970 antibody (Invitrogen; 1:1,000) and secondary goat α -rabbit HRP-coupled antibody
971 (Invitrogen; 1:10,000) were used for visualization. Polyclonal rabbit α -histone H3 antibody
972 (Invitrogen; 1:1,000) was used as a loading control. (j) RT-PCR verification of *A. carterae*
973 transformants showing transcription from an artificial minicircle based on *atpB*. HyperLadder

974 1 kb plus (lane 1), positive control PCR with artificial minicircle template (lane 2), negative
975 control with no template (lane 3), RT-PCR against chloramphenicol acetyltransferase (*cat*)
976 transcript from artificial minicircle (RT+, lane 4), RT-PCR against *cat* transcript from
977 artificial minicircle (RT- negative control, lane 5). (k) RT-PCR confirming expression of
978 *nptII/neo* gene expression in *B. saltans* transformants. GeneRuler 1 kb DNA Ladder (lane
979 1), *nptII/neo* expressed in *B. saltans* cells transformed with EF-1 alpha plasmid (lane 2);
980 RT- negative control from transformed *B. saltans* cells (lane 3), PCR for the EF -1 alpha
981 plasmid on RNA after DNase treatment (to verify absence of DNA, lane 4), PCR positive
982 control using the EF- 1 alpha plasmid DNA (lane 5), negative control with no template (lane
983 6). (l) Western blot of *D. papillatum* WT and C5 and C4 transformants that express the V5-
984 tagged *nptII/neo* gene. Monoclonal mouse α -V5 antibody (Invitrogen; 1:2,000) and
985 secondary α -mouse HRP-coupled antibody (Sigma; 1:1,000) were used. V5-tagged-
986 mNeonGreen *Trypanosoma brucei* cells served as a positive control and mouse α -alpha-
987 tubulin antibody (Sigma; 1:5,000) was used as a loading control. (m) For top and bottom
988 gels, RT-PCR of *A. whisleri* cells transfected with either 1 μ g pAwhi_H2Bvenus vector plus
989 10 mg carrier DNA (pUC19) (lanes 1-3), carrier DNA only (lane 4) or without DNA (lane 5),
990 and pAwhi_H2Bvenus used as positive control (lane 7). A 1 kb DNA ladder was used (lane
991 6).

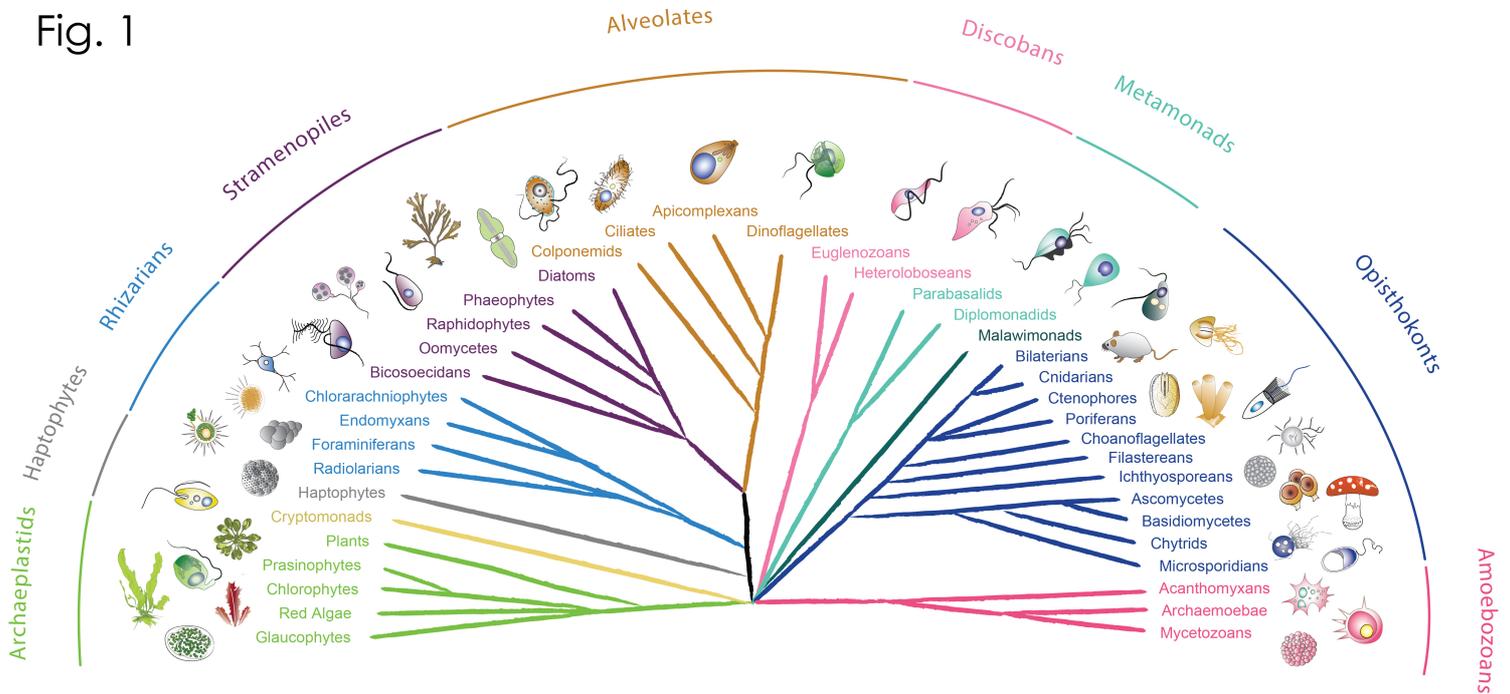
992

993 **Fig. 4. | ‘Transformation Roadmap’ for the creation of genetically tractable protists.**

994 (a) **Vector design and construction.** The three main goals for transformation of protists or
995 other taxa are: 1) to determine an appropriate selection marker, 2) to design a suitable
996 construct and 3) to find an appropriate transformation method for delivering DNA. Examples
997 from this research exhibited different levels of success many with DNA delivered and
998 expressed but ranging to DNA delivered without confirmation of gene expression or DNA

999 delivery not yet achieved (see **Fig. 1**). **(b) Transformation approaches.** Different symbols
1000 represent methods (e.g. chemical, physical or biological) for introducing DNA/RNA/protein
1001 into a living cell. **(c) Protocol.** Key methodological steps for successful transformation are
1002 listed in an abbreviated form (for particular examples, see table 1 and text).

Fig. 1



Archaeplastids

<i>Ostreococcus lucimarinus</i>	● A	○	⚡	R
<i>Bathycoccus prasinos</i>	● A	○	⚡	R
<i>Micromonas commoda</i>	● A	○	⚡	R
<i>Micromonas pusilla</i>	● A	○	⚡	✗ ✗
<i>Tetraselmis striata</i>	● A	○	⚡	R
<i>Pyramimonas parkeae</i>	● A	○	⚡	✗ ✗ ✗

Haptophytans

<i>Isochrysis galbana</i>	● A	○	⚡	R
<i>Emiliania huxleyi</i>	● A	○	⚡	R

Rhizarians

<i>Amorphochlora amoebiformis</i>	● B ²⁵	○	⚡	R
<i>Bigelowiella natans</i>	● A	○	⚡	✗ ✗

Stramenopiles

<i>Fragilariopsis cylindrus</i>	● A	○	⚡	R
<i>Thalassiosira pseudonana</i>	● B ^{15,17}	○	⚡	R
<i>Seminavis robusta</i>	● A	○	⚡	✗ ✗
<i>Pseudo-nitzschia multiseriata</i>	● A	○	⚡	R
<i>Heterosigma akashiwo</i>	● A	○	⚡	✗ ✗ ✗
<i>Aurantiochytrium limacinum</i>	● B ³¹	○	⚡	R
<i>Caecitellus</i> sp.	● A	○	⚡	R
<i>Nannochloropsis oceanica</i>	● A	○	⚡	R
<i>Phaeodactylum tricomutum</i>	● B ^{15,18}	○	⚡	R

Alveolates

<i>Euplotes crassus</i>	●	○	⚡	✗ ✗ ✗
<i>Euplotes focardii</i>	●	○	⚡	✗ ✗ ✗
<i>Chromera velia</i>	●	○	⚡	✗ ✗ ✗
<i>Perkinsus marinus</i>	● B ^{44,45}	○	⚡	R
<i>Oxyrrhis marina</i>	● A	○	⚡	✗ ✗ ✗
<i>Hematodinium</i> sp.	●	○	⚡	✗ ✗ ✗
<i>Fugacium (Symbiodinium) kawagutii</i>	●	○	⚡	✗ ✗
<i>Alexandrium catenella</i>	●	○	⚡	✗ ✗

<i>Brevium (Symbiodinium) sp.</i>	●	○	⚡	✗ ✗
<i>Cryptothecodinium cohnii</i>	● A	○	⚡	✗ ✗ ✗
<i>Amphidinium carterae</i> (chloroplast)	● B ⁴⁰	○	⚡	✗ ✗ ✗
<i>Karlodinium veneficum</i>	● A	○	⚡	R

Discobans

<i>Bodo saltans</i>	● A	○	⚡	R
<i>Diplonema papillatum</i>	● B ⁵¹	○	⚡	R
<i>Eutreptiella gymnastica</i>	● A	○	⚡	✗ ✗ ✗
<i>Naegleria gruberi</i>	● A	○	⚡	R

Opisthokonts

<i>Pirum gemmata</i>	●	○	⚡	✗ ✗ ✗
<i>Sphaeroforma arctica</i>	●	○	⚡	✗ ✗ ✗
<i>Abeoforma whisleri</i>	● A	○	⚡	R
<i>Salpingoeca rosetta</i>	● C ^{55,56}	○	⚡	R

Natural Communities

Pico-Eukaryotes	● A	○	⚡	
-----------------	-----	---	---	--

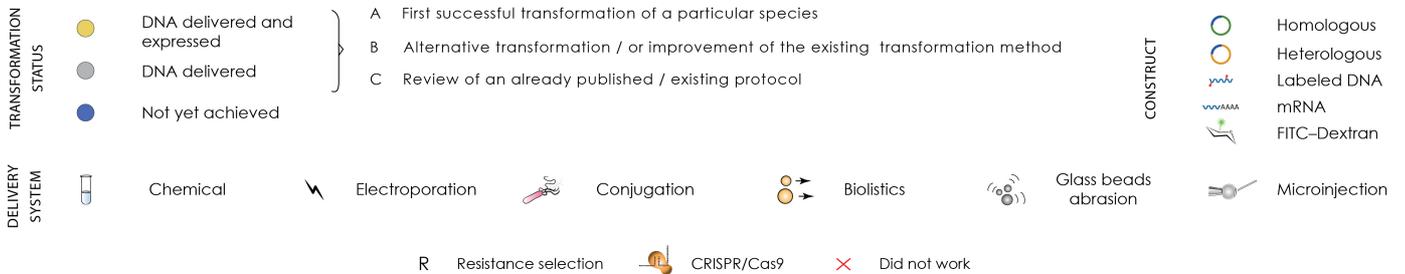


Fig. 2

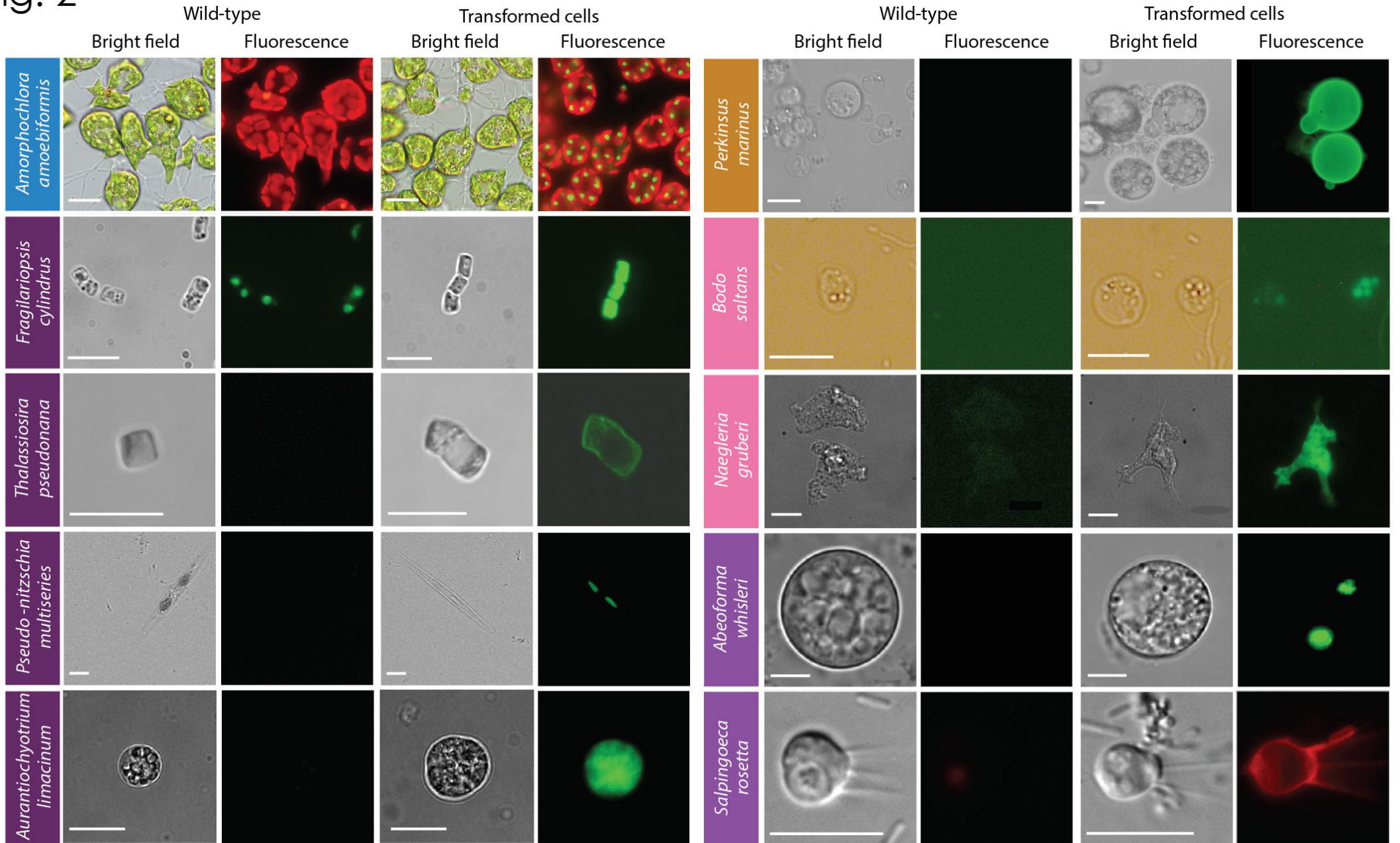


Fig. 3

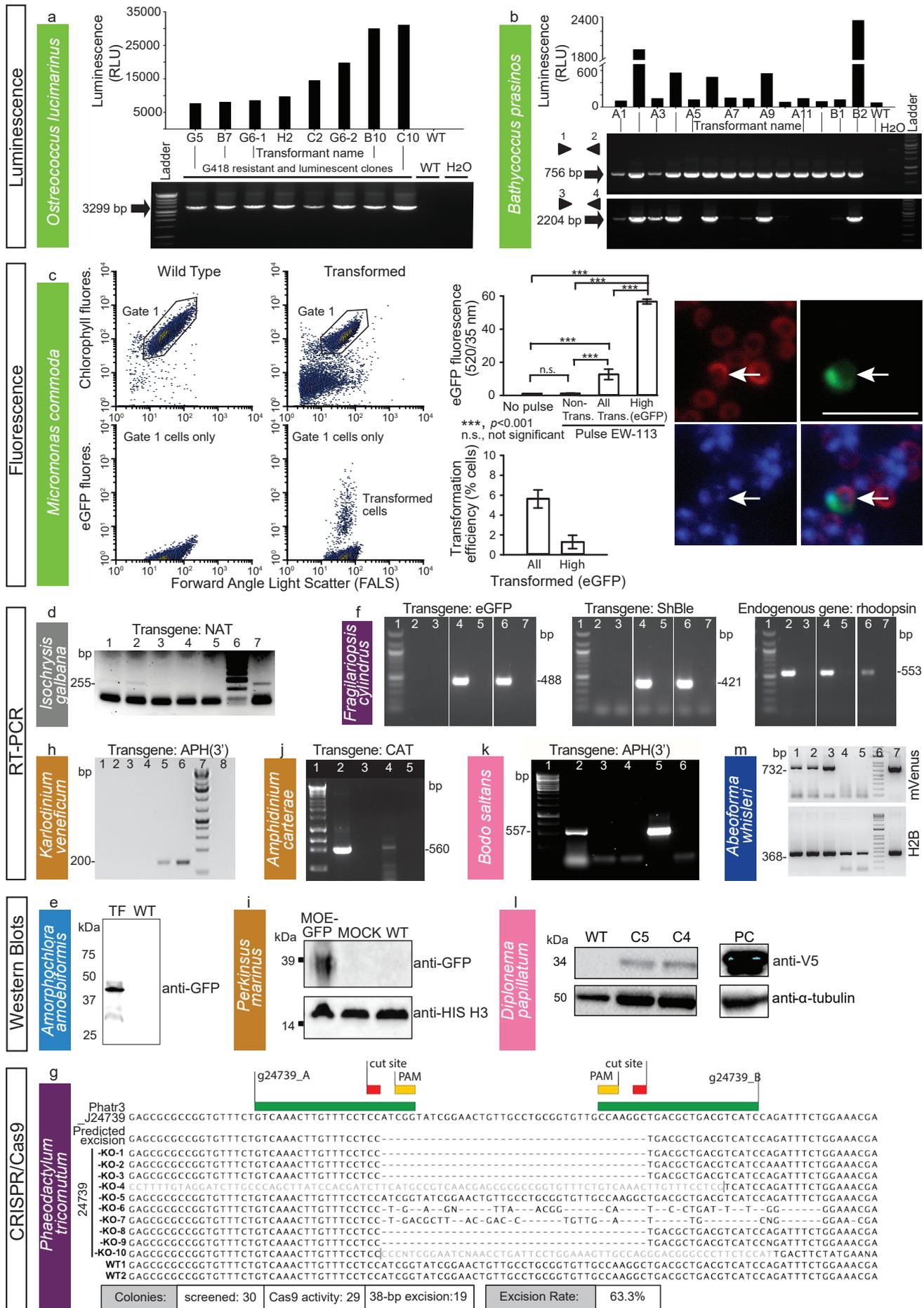
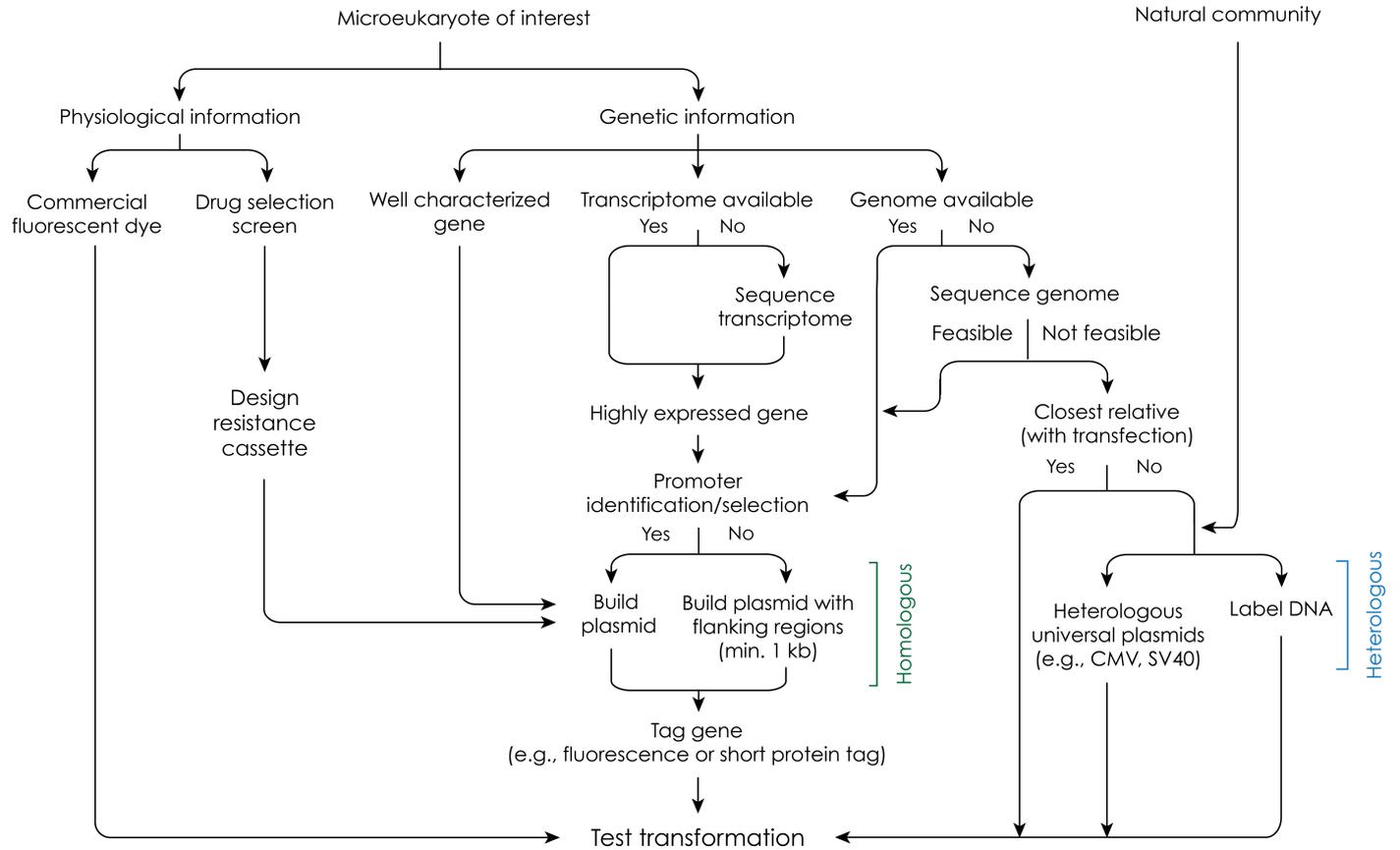


Fig. 4

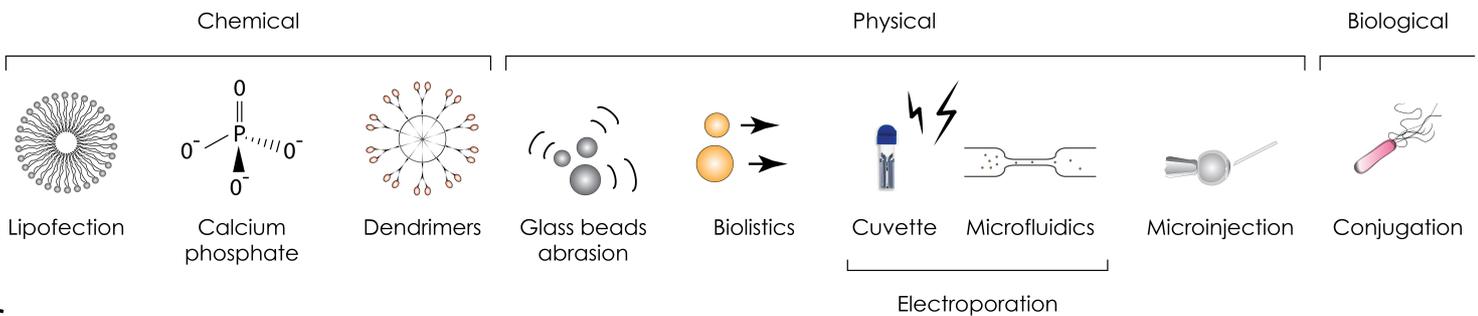
a

Vector construction



b

Transformation approaches



c

Protocol

Organism

- Description (life cycle stages)
- Strain
- Growth conditions and media
- Maximum cell density
- Autofluorescence
- Genome/transcriptome availability
- Cryopreservation

Resistance markers

- Antibiotics tested
- Assay used
- Resistance/sensitivity (IC₅₀)

Constructs

- Construct name
- Origin
- Annotation
 - Promoter
 - UTRs
 - Flanking regions
 - Resistance
- Nature
 - Autonomously replicating
 - Integrative
 - Transiently

Transformation

- Method
- Parameters
 - Buffer composition
 - Program
- Cell stage and density
- DNA concentration
- Drug selection
 - Concentration
 - Time
- Efficiency

Transformants

- Clones examined
- Proof of nucleic acid uptake
 - Isolation of total DNA
 - PCR verification
 - Southern blot
- Plasmid maintenance
 - Chromosomal integration
 - Episomal
- Transcription
 - RT-PCR
 - Northern blot
- Translation
 - Western blot
 - Epifluorescence microscopy
 - Protein localization
- Phenotype
 - Growth curve
 - Description

Table 1: Parameters used for successful transformation as shown in Figs. 2 and 3.

For additional information, see protocols.io, Suppl. table 5, and Suppl. doc.1. For contacting laboratories working with particular species, see details given in Suppl. table 6.

Species	Transformation Method/ Device	Cell number	Vector;Amount (µg)	Promoter	Regulatory elements	Drug selection (µg/ml)	Time selection (Day)	Efficiency (%)	Transformation status (Stable/ Transient)	Reporter	Evidence of transformation	protocols.io link
Archaeplastids												
<i>Ostreococcus lucimarinus</i> (RCC802)	Electroporation Genepulser II	1-2x10 ⁹	Plasmid PotLuc; Linear; 5	HAPT, Histone H4 <i>O. tauri</i>	None	G418 (1000)	10-21	<0.0001	S	Luc	G418 resist, Luminescence, PCR	https://www.protocols.io/view/selection-of-stable-transformants-in-ostreococcus-zj2f4qce https://www.protocols.io/view/transient-luciferase-expression-in-ostreococcus-ot-hcib2ue https://www.protocols.io/view/transient-transformation-of-ostreococcus-species-o-g86bzze
<i>Bathycoccus prasinos</i> (RCC4222)	Electroporation Genepulser II	1-2x10 ⁹	Fusion PCR; pHAPT; pLucpH4;KanM; Linear; 5	HAPT, Histone H4 Endogenous	None	G418 (1000)	10-21	<0.0001	S	Luc	G418 resist, Luminescence PCR	http://dx.doi.org/10.17504/protocols.io.g86bzze http://dx.doi.org/10.17504/protocols.io.zj2f4qce http://dx.doi.org/10.17504/protocols.io.hcib2ue
<i>Micromonas commoda</i> (CCMP 2709)	Electroporation Lonza-Nucleofector	3 x 10 ⁷	RPS9pro ^{MS2} -eGFP-NLS-RPS9ter in pUC05-AMP; Circular; 10-20	Endogenous, ribosomal protein S9;	Endogenous, ribosomal protein S9;	n/a	2-6	5.6±1.3	T*	eGFP	Per cell eGFP Fluorescence, Fluorescence microscopy	http://dx.doi.org/10.17504/protocols.io.8p9hvr6
	Electroporation Lonza-Nucleofector	3 x 10 ⁷	H3pro ^{MS2} -LUC-H3ter in pUC05-AMP; Circular; 10-20	Histone H3 5' UTR from <i>M. polaris</i>	Histone H3 3' end formation -histone stem loop from <i>M. polaris</i>	n/a	3	n/a (Luc. assay is bulk, not per cell)	T*	NanoLuc®	Luminescence	http://dx.doi.org/10.17504/protocols.io.8p8hvrw
<i>Tetraselmis striata</i> (KAS-836)	Biorad Biolistics PDS-1000/He biolistics system	2.0x10 ⁷	pACTpro:BLE; Linear; 1.0	Actin, <i>T. striata</i>	Actin, <i>T. striata</i>	Zeocin (150)	21-28		S		Zeocin resist, PCR	http://dx.doi.org/10.17504/protocols.io.hjtb4nn
Haptophytes												
<i>Isochrysis galbana</i> (CCMP 1323)	Biolistics PDS-1000/He	1-2x10 ⁶	pIgNAT; Circular; 1.0	Hsp70 <i>E. huxleyi</i>	Heterologous	Nourseothricin (80-150)	14	<0.0001	S	None	Nourseothricin resistance, PCR, RT-PCR	https://www.protocols.io/view/biolistic-transformation-of-isochrysis-galbana-2puqdnw https://www.protocols.io/view/method-for-electroporation-of-isochrysis-galbana-c-hmab42e
Rhizarians												
<i>Amorphochlora (Lotharella) amoebiformis</i> (CCMP 2058)	Electroporation Gene Pulser Xcell	0.5-1x10 ⁷	GFP-Rubisco; Circular; 30-50	rbcS1, Endogenous	rbcS1 Endogenous	*Manual selection of fluorescent cells	n/a	n/a	S/T	GFP	Fluorescence, Western blot	http://dx.doi.org/10.17504/protocols.io.35hqa36
Stramenopiles												
<i>Fragilariopsis cylindrus</i> (CCMP 1102)	Biorad Biolistics PDS-1000/He biolistics system	5x10 ⁷	pUC:FCP:Shble:FCP:eGFP; Circular; 1.0	FCP, Endogenous	None	Zeocin (100)	21 – 49	0.00003 (30 cfu/10 ⁸ cells)	S	eGFP	Zeocin resist, Fluorescence, PCR, RT-PCR	http://dx.doi.org/10.17504/protocols.io.z39f8r6 https://www.protocols.io/view/biolistic-transformation-of-polar-diatom-fragilari-z39f8r6
<i>Thalassiosira pseudonana</i> (CCMP 1335)	Bacterial conjugation	4x10 ⁷	TipSll3p-eGFP in pTpPuc3; Circular; n/a	Endogenous	Endogenous	Nourseothricin (100 in plates, 200 in liquid culture)	~14	~10	T	eGFP	Nourseothricin resistance, colony PCR, fluorescence	http://dx.doi.org/10.17504/protocols.io.nbzdp6 http://dx.doi.org/10.17504/protocols.io.7ghjt6
<i>Pseudo-nitzschia multiseriata</i>	Conjugation	1x10 ⁵	Pm_actP_egfp_actI; pPIPUC3	Pm actin; Pt fcpB,	None, other than contained in Promoter/term	Manual selection of fluorescent cells in LGTA; zeocin [200]	24h, 7	<0.1%	T	eGFP, shble	Fluorescence, vector targeted PCR on gDNA	http://dx.doi.org/10.17504/protocols.io.4pzgvp6
<i>Aurantiochytrium limacinum</i> (ATCC MYA-1381)	Bio – Rad Gene Pulser (165-2076) NEPA21	1x10 ⁸	18GZG 18GeZG plasmid; Linear; 1-10	Endogenous GAPDH	None	Zeocin (100)	5-7	44 per µg of DNA	S	eGFP, shble	Zeocin resist., PCR, Southern, Fluorescence	http://dx.doi.org/10.17504/protocols.io.8xyhxpw
<i>Nannochloropsis oceanica</i> (CCMP 1779)	Electroporation Genepulser II	1x10 ⁹	pMOD, Linear/Circular; 0.1-1	CMV	None	None	0.1-1	20 (linear) 1-2 (circular)	T	mTagBFP2	Fluorescence, PCR, RT-PCR	http://dx.doi.org/10.17504/protocols.io.h3nb8me

<i>Phaeoactylum tricamutum</i> (CCAP1055/1)	Bacterial-conjugation	4x10 ⁷	hCas9-2A-shble PtpBR episome 100uL E.coli OD ₆₀₀ =0.9	FcpF-hCas9 psRNA-sgRNA	Cen6-Arsh4-His3 centromere	Phleomycin (50) Zeocin (100)	10-16	1.25e ⁻⁵ ~500 cfu	S	shble (Cas9) yfp VENUS	Phleoycin resistance, PCR maintained episome, PCR Cas9 target site	http://dx.doi.org/10.17504/protocols.io.4bmgsk6 http://dx.doi.org/10.17504/protocols.io.7gjhjue http://dx.doi.org/10.17504/protocols.io.7gnhjve
Alveolates												
<i>Perkinsus marinus</i> (ATCC PRA240)	Electroporation LONZA-Nucleofector Glass beads abrasion (425-600 µm)	5-7x10 ⁷	pPmMOE-GFP; Linear-Circular (1:1); 5	Endogenous	Endogenous	FACS Blasticidin(50-200) Puromycin (10-50) Bleo (50-200)	Drug: 20-60 FACS: 3	0.01-5	S	GFP, mCherry	Fluorescence Sequencing, PCR, Western blot	https://www.protocols.io/view/oyster-parasite-perkinsus-marinus-transformation-u-gv9bw96 https://www.protocols.io/view/glass-beads-based-transformation-protocol-for-perk-g36byre https://www.protocols.io/view/fluorescence-activated-cell-sorting-facs-of-perkin-hh2b38e
<i>Oxyrrhis marina</i> (CCMP 1788, CCMP 1795)	Electroporation Gene Pulser Xcell; Chemical (CaCl ₂)	1-5x10 ⁶ 1x10 ⁵	Fluorescently labelled DNA (5-25 µg) or FITC-dextran; mCherry	n/a Endogenous hsp90	n/a Endogenous hsp90	n/a	n/a	0.5-5%	S T	mCherry	Fluorescence	https://www.protocols.io/view/electroporation-of-oxyrrhis-marina-vcne2ve https://www.protocols.io/view/transfection-of-alexa488-labelled-dna-into-oxyrrhi-ha8b2hw https://www.protocols.io/view/electroporation-transformation-of-fic-dextran-int-3cmgiu6 https://www.protocols.io/view/co-incubation-protocol-for-transforming-heterotop-hmzb476 https://www.protocols.io/view/transfection-of-cryptocodium-cohnii-using-label-z26f8he
<i>Cryptocodium cohnii</i> (CCMP 316)	Electroporation LONZA-Nucleofector		Stained DNA (739 bp); Linear; 1	None	None	n/a	n/a	<0.001	T	Fluorescence	Fluorescence	http://dx.doi.org/10.17504/protocols.io.4r2gv8e
<i>Amphidinium carterae</i> (chloroplast) (CCMP 1314)	Biorad Biolistics PDS-1000/He biolistics system	2.5x10 ⁷	pAmpAtpBChl; Circular; 0.5	Endogenous	Endogenous	Chloramphenicol (20)	3 onwards	n/a	S	Ab res	RT-PCR, PCR Phenotype	http://dx.doi.org/10.17504/protocols.io.4r2gv8e
<i>Karlodinium veneficum</i> (CCMP 1975)	Electroporation	4x10 ⁵	linear-Dinoll-neo; Linear; 2	Endogenous	Endogenous	Kanamycin (150)	7	0.0005	S (3 mon)	n/a	RT-PCR, PCR	https://www.protocols.io/view/nucleofector-protocol-for-dinoflagellates-using-lo-qm8du9w
Discobans (Euglenozoans and Heteroloboseans)												
<i>Bodo saltans</i> (submitted to ATCC)	Electroporation Nepa21	1-1.5x10 ⁷	Bs-EF1- a C- terminal tagging; Linear; 3-5	Endogenous	Endogenous	G418 (3)	7-9		S	GFP	Fluorescence, PCR, RT-PCR	http://dx.doi.org/10.17504/protocols.io.s5jeg4n http://dx.doi.org/10.17504/protocols.io.7fchjw
<i>Diplonema papillatum</i> (ATCC 50162)	Electroporation LONZA-Nucleofector	5x10 ⁷	p57-V5+Neo ^R ; Linear; 3	Endogenous	Endogenous	G418 (75)	7-14	~5.5	S	n/a	Western blot (resistance marker), RT-PCR	http://dx.doi.org/10.17504/protocols.io.4digs4e
<i>Naegleria gruberi</i> (ATCC 30224)	Electroporation BioRad Gene Pulser xCell	5x10 ⁶	pNAEG-HYG; Circular; 4	Endogenous	Endogenous	Hygromycin (300) Neo (700)	15-28	80	T	GFP	Western blot (resistance marker), Fluorescence,	http://dx.doi.org/10.17504/protocols.io.hpub5nw http://dx.doi.org/10.17504/protocols.io.7w4hpgw
Opisthokonts												
<i>Abeoforma whisleri</i> (ATCC PRA-279)	Electroporation LONZA-Nucleofector	3x10 ⁵	Awhis_H2Bvenus+ pUC19; Circular; 1-5+ 40 carrier	Endogenous	Endogenous	n/a	10-15	1	T	Venus	Fluorescence, RT-PCR	http://dx.doi.org/10.17504/protocols.io.zex3fn
<i>Salpingoeca roseffa</i> (ATCC PRA-390)	Electroporation LONZA-Nucleofector	4x10 ⁵	SrActmCherry-CCILL + pUC19; Circular; 1-10 + 40 carrier	Endogenous	Endogenous	Puromycin (80)	10-12		S	mCherry	Gene Expression (Luc, Fluorescence)/resistance	http://dx.doi.org/10.17504/protocols.io.h68b9hw

*may be stable but overgrown by wild-type strain

Abbreviations: n/a, not applicable; NLS, Nuclear Localization Signal