Genetic tool development in marine protists:

Emerging model organisms for experimental cell biology

Drahomíra Faktorová[1,#,*], R. Ellen R. Nisbet[2,#,%], José A. Fernández Robledo[3,#], Elena Casacuberta[4,#], Lisa Sudek[5,#], Andrew E. Allen[6,7], Manuel Ares Jr.[8], Cristina Aresté[4], Cecilia Balestreri[9], Adrian C. Barbrook[2], Patrick Beardslee[10], Sara Bender[11], David S. Booth[12], François-Yves Bouget[13], Chris Bowler[14], Susana A. Breglia[15], Colin Brownlee[9], Gertraud Burger[16], Heriberto Cerutti[10], Rachele Cesaroni[17], Miguel A. Chiurillo[18], Thomas Clemente[10], Duncan B. Coles[3], Jackie L. Collier[19], Elizabeth C. Cooney[20], Kathryn Coyne[21], Roberto Docampo[18], Christopher L. Dupont[7], Virginia Edgcomb[22], Elin Einarsson[2], Pía A. Elustondo[15,$], Fernan Federici[23], Veronica Freire-Beneitez[24,25], Nastasia J. Freyria[3], Kodai Fukuda[26], Paulo A. García[27], Peter R. Girguis[28], Fatma Gomaa[28], Sebastian G. Gornik[29], Jian Guo[5,8], Vladimír Hampl[30], Yutaka Hanawa[31], Esteban R. Haro-Contreras[15], Elisabeth Hehenberger[20], Andrea Highfield[9], Yoshihisa Hirakawa[31], Amanda Hopes[32], Christopher J. Howe[2], Ian Hu[2], Jorge Ibañez[23], Nicholas A.T. Irwin[20], Yuu Ishii[33], Natalia Ewa Janowicz[30], Adam C. Jones[11], Ambar Kachale[1], Konomi Fujimura-Kamada[34], Binnypreet Kaur[1], Jonathan Z. Kaye[11], Eleanna Kazana[24,25], Patrick J. Keeling[20], Nicole King[12], Lawrence A. Klobutcher[35], Noelia Lander[18], Imen Lassadi[2], Zhuhong Li[18], Senjie Lin[35], Jean-Claude Lozano[13], Fulei Luan[10], Shinichiro Maruyama[33], Tamara Matute[23], Cristina Miceli[36], Jun Minagawa[34,37], Mark Moosburner[6,7], Sebastián R. Najle[4,38], Deepak Nanjappa[21], Isabel C. Nimmo[2], Luke Noble[39,**], Anna M.G. Novák Vanclová[30], Mariusz Nowacki[17], Isaac Nuñez[23], Arnab Pain[40,41], Angela Piersanti[36], Sandra Pucciarelli[36], Jan Pyrih[24,30], Joshua S. Rest[42], Mariana Rius[19], Deborah Robertson[43], Albane Ruaud[23,***], Iñaki Ruiz-Trillo[4,44,45], Monika A. Sigg[12], Pamela A. Silver[46,47], Claudio H. Slamovits[15], G. Jason Smith[48], Brittany N. Sprecher[35], Rowena Stern[9], Estienne C. Swart[17,***], Anastasios D. Tsaousis[24,25], Lev

Tsypin[49,50], Aaron Turkewitz[49], Jernej Turnšek[6,7,46,47], Matus Valach[16], Valérie Vergé[13], Peter von Dassow[23,51], Tobias von der Haar[24], Ross F. Waller[2], Lu Wang[52], Xiaoxue Wen[10], Glen Wheeler[9], April Woods[48], Huan Zhang[35], Thomas Mock[32,*], Alexandra Z. Worden[5,53,*] & Julius Lukeš[1,*]

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

[2]*Department of Biochemistry, University of Cambridge, Cambridge, UK*

[3]*Bigelow Laboratory for Ocean Sciences, East Boothbay, USA*

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

[5]*Monterey Bay Aquarium Research Institute, Moss Landing, USA*

[6]*Integrative Oceanography Division, Scripps Institution of Oceanography, University of California, San Diego, USA*

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

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

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

[10]*School of Biological Sciences, University of Nebraska, Lincoln, USA*

[11]*Gordon and Betty Moore Foundation, Palo Alto, USA*

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

[13]*Sorbonne Université, CNRS UMR7621, Observatoire Océanologique, Banyuls sur Mer, France*

[14]*Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure, CNRS, INSERM, Université PSL, Paris, France*

[15]*Centre for Comparative Genomics and Evolutionary Bioinformatics, Dalhousie University, Halifax, Canada*

[16]*Department of Biochemistry and Robert-Cedergren Centre for Bioinformatics and Genomics, Université de Montréal, Montreal, Canada*

[17]*Institute of Cell Biology, University of Bern, Bern, Switzerland*

[18]*Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, USA*

[19]*School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, USA*

[20]*Department of Botany, University of British Columbia, Vancouver, Canada*

[21]*University of Delaware College of Earth, Ocean and Environment, Lewes, USA*

[22]*Woods Hole Oceanographic Institution, Woods Hole, USA*

[23]*Facultad Ciencias Biológicas, Pontificia Universidad Católica de Chile, Fondo de Desarrollo de Areas Prioritarias, Center for Genome Regulation and Millennium Institute for Integrative Biology (iBio), Santiago de Chile, Chile*

[24]*School of Biosciences, University of Kent, Canterbury, Kent, UK*

[25]*Laboratory of Molecular and Evolutionary Parasitology, University of Kent, UK*

[26]*Graduate School of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan*

[27]*Department of Mechanical Engineering, Massachusetts Institute of Technology, Boston, USA*

[28]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, USA*

[29]*Centre for Organismal Studies, University of Heidelberg, Heidelberg, Germany*

[30]*Department of Parasitology, Faculty of Science, Charles University, BIOCEV, Vestec, Czech Republic*

[31]*Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan*

[32]*School of Environmental Sciences, University of East Anglia, Norwich, UK*

[33]*Graduate School of Life Sciences, Tohoku University, Sendai, Miyagi, Japan*

[34]*Division of Environmental Photobiology, National Institute for Basic Biology, Okazaki, Aichi, Japan*

[35]*Department of Marine Sciences, University of Connecticut, Groton, USA*

[36]*School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy*

[37]*Department of Basic Biology, School of Life Science, Graduate University for Advanced Studies, Okazaki, Aichi, Japan*

[38]*Instituto de Biología Molecular y Celular, CONICET, and Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina*

[39]*Center for Genomics and Systems Biology, New York University, New York, USA*

[40]*Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

[41]*Center for Zoonosis Control, Global Institution for Collaborative Research and Education, Hokkaido University, Sapporo, Japan*

[42]*Department of Ecology and Evolution, Stony Brook University, Stony Brook, USA*

[43]*Lasry Center for Biosciences, Clark University, Worcester, USA*

[44]*Departament de Genètica Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Spain*

[45]*Catalan Institution for Research and Advanced Studies, Barcelona, Spain*

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

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

[48]*Department of Environmental Biotechnology, Moss Landing Marine Laboratoriess, Moss Landing, USA*

[49]*Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, USA*

[50]*Department of Biology, California Institute of Technology, Pasadena, USA*

[51]*Instituto Milenio de Oceanografia de Chile, Concepción, Chile*

[52]*Institute of Oceanography, Minjiang University, Fuzhou, China*

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

[%] Present address: *School of Biosciences, University of Nottingham, Sutton Bonington, UK*

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

[**]Present address: *Institute de Biologie de l'ENS, Département de biologie, École Normale Supérieure, CNRS, INSERM, Paris, France*

[***]Present address: *Max Planck Institute for Developmental Biology, Tübingen, Germany*

[#]Equal contribution

[*]Corresponding authors: dranov@paru.cas.cz (DF); t.mock@uea.ac.uk (TM); azworden@geomar.de (AZW); jula@paru.cas.cz (JL)

ABSTRACT

Diverse microbial ecosystems underpin life in the sea. Among these microbes are many unicellular eukaryotes that span the diversity of the eukaryotic tree of life. However, genetic tractability has been limited to a few species, which do not represent eukaryotic diversity or environmentally relevant taxa. Here, we report on the development of genetic tools in a range

of protists primarily from marine environments. We present evidence for foreign DNA delivery and expression in 13 species never before transformed and advancement of tools for 8 other species, as well as potential reasons for why transformation of yet another 17 species tested was not achieved. Our resource in genetic manipulation will provide insights into the ancestral eukaryotic lifeforms, general eukaryote cell biology, protein diversification and the evolution of cellular pathways.

INTRODUCTION

The ocean represents the largest continuous planetary ecosystem, hosting an enormous variety of organisms, which include microscopic biota such as unicellular eukaryotes (protists). Despite their small size, protists play key roles in marine biogeochemical cycles and harbor tremendous evolutionary diversity[1,2]. Notwithstanding their importance for understanding the evolution of life on Earth and their role in marine food webs, as well as driving biogeochemical cycles to maintain habitability, little is known about their cell biology including reproduction, metabolism, and signalling[3]. Most of the biological knowledge available is based on comparison of proteins from cultured species to homologs in genetically tractable model taxa[4-7]. A major impediment to understanding the cell biology of these diverse eukaryotes is that protocols for genetic modification are only available for a small number of species[8,9] that represent neither the most ecologically relevant protists nor the breadth of eukaryotic diversity.

The development of genetic tools requires reliable information about gene organisation and regulation of the emergent model species. Over the last decade, genome[4-6] and transcriptome sequencing initiatives[7] have resulted in nearly 120 million unigenes being identified in protists[10], which facilitates the developments of genetic tools used for model species Insights from these studies enabled the phylogenetically-informed approach[7] for

selecting and developing key marine protists into model systems in the Environmental Model Systems (EMS) Project presented herein. Forty-one research groups took part in the EMS Project, a collaborative effort resulting in the development of genetic tools that significantly expand the number of eukaryotic lineages that can be manipulated, and which encompass multiple ecologically important marine protists.

Here, we summarize detailed methodological achievements and analyse results to provide a synthetic 'Transformation Roadmap' for creating new microeukaryotic model systems. Although the organisms reported here are diverse, the paths to overcome difficulties share similarities, highlighting the importance of building a well-connected community to overcome technical challenges and accelerate the development of genetic tools. The 13 emerging model species presented herein, and the collective set of genetic tools from the overall collaborative project, will not only extend our knowledge of marine cell biology, evolution and functional biodiversity, but also serve as platforms to advance protistan biotechnology.

RESULTS

Overview of taxa in the EMS Initiative

Taxa were selected from multiple eukaryotic supergroups[1,7] to maximize the potential of cellular biology and to evaluate the numerous unigenes with unknown functions found in marine protists (**Fig. 1**). Prior to the EMS initiative, reproducible transformation of marine protists was limited to only a few species such as *Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, and *Ostreococcus tauri* (**Suppl. Table 1**). The EMS initiative included 39 species, specifically, 6 archaeplastids, 2 haptophytes, 2 rhizarians, 9 stramenopiles, 12 alveolates, 4 discobans, and 4 opisthokonts (**Fig. 1**). Most of these taxa were isolated from coastal habitats, the focus area of several culture collections[7]. More than 50% of the selected

species are considered photoautotrophs, with another 35% divided between heterotrophic osmotrophs and phagotrophs, the remainder being predatory mixotrophs. Almost 20% of the chosen species are symbionts and/or parasites of marine plants or animals, 5% are associated with detritus, and several are responsible for harmful algal blooms (**Suppl. Table 2**).

While some transformation systems for protists have been developed in the past[8,9,11], the challenge for this initiative was to develop genetic tools for species which not only require different cultivation conditions but are also phenotypically diverse. It should be noted that not all major lineages were explored. For example, amoebozoans did not feature in this aquatic-focused initiative, in part because they tend to be most important in soils, at least based on current knowledge, and manipulation systems exist for members of this eukaryotic supergroup, such as *Dictyostelium discoideum*[12]. The overall EMS initiative outcomes are summarized in Figure 1 and Table 1. We provide detailed protocols for 13 taxa, for which no transformation systems have been previously reported (Category A), and 8 taxa, for which existing protocols[9,11,13-19,20,21] were advanced (Category B) (**Figs. 2, 3 and 4; Table 1; Suppl. Tables 1-5 and Online Methods**). We also review an already published EMS transformation protocol[22] in 1 species (Category C), and we discuss unsuccessful transformation attempts for 17 additional taxa (**Fig. 1; Online Methods**). Finally, we synthesize our findings in a roadmap for the development of transformation systems in protists (**Fig. 5**).

***Archaeplastids*.** Prasinophytes are important marine green algae distributed from polar to tropical regions. They form a sister group to chlorophyte algae, and together, these two groups branch adjacent to land plants, collectively comprising the Viridiplantae, which are part of the Archaeplastida[1,23] (**Fig. 1**). Genome sequences are available for the picoprasinophytes (<3 µm cell diameter) tested herein, specifically, *Micromonas commoda*, *M. pusilla*, *Ostreococcus lucimarinus* and *Bathycoccus prasinos*. As part of the EMS

initiative, we report on the first genetic tools for *Bathycoccus*, a scaled, non-motile genus, and *Micromonas*, a motile, naked genus with larger genomes than *Bathycoccus* and *Ostreococcus*[22]. We also report on the first genetic tools for *Tetraselmis striata* and *Ostreococcus lucimarinus*. The latter was transformed based on an adapted homologous recombination system for *Ostreococcus tauri*[24,25].

O. lucimarinus (RCC802) and *B. prasinos* (RCC4222) were transformed using protocols adapted from *O. tauri*[24,25]. Briefly, using electroporation for transfer of exogenous genes, *O. lucimarinus* was transformed using a DNA fragment encoding the *O. tauri* high-affinity phosphate transporter (*HAPT*) gene fused to a luciferase gene and a kanamycin selection marker (**Table 1; Suppl. Table 3**), which resulted in transient luciferase expression 24 h after electroporation (**Table 1; Fig. 3a**). After 2 weeks of growth in low-melting agarose plates containing G418 (1 mg/ml), 480 colonies were obtained, picked, and grown in artificial seawater with the antibiotic neomycin. Of these, 76 displayed luminescence ≥ 2.5 fold above background (80 Relative Luminescence Units, RLU), with widely variable levels (200 to 31020 RLU), likely reflecting either variations in the site of integration and/or the number of integrated genes (**Fig. 3a; Suppl. Fig. 1; Online Methods**).

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

Although transformation methods successful for *Bathycoccus* and *Ostreococcus* failed in *Micromonas*, Lonza nucleofection was successful with *M. commoda* (CCMP2709) (**Table 1; Fig. 3c**) using 2 different codon-optimized plasmids, one encoding the luciferase gene (NanoLuc, Promega) flanked by an exogenous promoter and terminator sequence from the 5′- and 3′- untranslated regions (UTRs) of histone *H3* in *Micromonas polaris* (CCMP2099), and the other encoding an enhanced fluorescent protein *(eGFP)* gene flanked by endogenous promoter and terminator sequences from ribosomal protein S9 (**Suppl. Table 5**). Sensitivities to antibiotics were established (**Suppl. Table 3**). Constructs did not include a selectable marker, as we aimed to introduce and express foreign DNA while developing conditions suitable for transfection that supported robust growth in this cell wall-lacking protist (**Table 1**). Transformants revealed a significantly higher level of eGFP fluorescence than wild type cells, with 1.3% of the population showing fluorescence per cell 45-fold higher than both the non-transformed portion of the culture and the wild type cells (**Fig. 3c; Online Methods**). Additionally, the RLU was 1500-fold higher than controls when using the luciferase-bearing construct, such that multiple experiments with both plasmids confirmed expression of exogenous genes in *M. commoda*.

T. striata (KAS-836) was transformed using microprojectile bombardment (**Suppl. Fig. 2a**). Two selectable marker genes were tested, consisting of a putative promoter and 5’ UTR sequences from the *T. striata* actin gene and either the coding sequences of the *Streptoalloteichus hindustanus* bleomycin gene (conferring resistance to zeocin) or the *Streptomyces hygroscopicus bar* gene (conferring resistance to glufosinate) (**Table 1; Suppl. Fig. 2a; Online Methods**). The terminator sequence was obtained from the *T. striata* glyceraldehyde-3-phosphate dehydrogenase gene. Linearized plasmids were coated on gold particles and introduced into *T. striata* cells by using the PDS-1000/He Particle Delivery

System (BioRad). Transformants were successfully selected on half-strength f/2 at 50% salinity agar plates containing either 150 µg/ml zeocin or 150 µg/ml glufosinate.

***Haptophytes* (incertae sedis).** Haptophytes are a group of photosynthetic protists that are abundant in marine environments and include the major calcifying lineage, the coccolithophores. Genome sequences are available for *Emiliania huxleyi*[6] and *Chrysochromulina tobin*[26], and there is one report of nuclear transformation of a calcifying coccolithophore species[27] but transformation of *E. huxleyi*, the most prominent coccolithophore, has not been achieved yet[27]. Here, as part of the EMS initiative, a stable nuclear transformation system was developed for *Isochrysis galbana*, a species that lacks coccoliths, but represents an important feedstock for shellfish aquaculture[28].

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

***Rhizarians*.** Rhizarians include diverse non-photosynthetic protists, as well as the photosynthetic chlorarachniophytes that acquired a plastid *via* secondary endosymbiosis of a

green alga[4]. Uniquely, they represent an intermediate stage of the endosymbiotic process, since their plastids still harbor a relict nucleus (nucleomorph). Here, we report on an advanced transformation protocol for the chlorarachniophyte *Amorphochlora (Lotharella) amoebiformis* for which low-efficiency transient transformation has previously been achieved using particle bombardment[14].

A. amoebiformis (CCMP2058) cells were resuspended in 100 µl of Gene Pulse Electroporation Buffer (BioRad) with 20-50 µg of the reporter plasmid encoding eGFP-RubisCO fusion protein under the control of the native *rbcS1* promoter and subjected to electroporation (**Table 1**). Cells were immediately transferred to fresh ESM medium and incubated for 24 h. Transformation efficiency was estimated by the fraction of cells expressing eGFP, resulting in 0.03-0.1% efficiency, as enumerated by microscopy, showing an efficiency up to 1000-fold higher than in the previous study[14] (**Table 1**). Stable transformants were generated by manual isolation using a micropipette, and a transformed line has maintained eGFP fluorescence for at least 10 months without antibiotic selection (**Fig. 2; Fig. 4b; Online Methods**).

Stramenopiles. Stramenopiles are a diverse lineage harboring important photoautotrophic, mixotrophic (combining photosynthetic and phagotrophic nutrition) and heterotrophic taxa. As the most studied class in this lineage, diatoms (Bacillariophyceae) were early targets for the development of reverse genetics tool[11,29]. Diatoms are estimated to contribute approximately 20% of annual carbon fixation[30] and, like several other algal lineages, are used in bioengineering applications and biofuels[31]. Although other cold-adapted eukaryotes have yet to be transformed, here we present a protocol for the Antarctic diatom *Fragilariopsis cylindrus*[32]. A first transformation protocol has also been developed for *Pseudo-nitzschia multiseries*, a toxin-producing diatom[33]. Here we also present work for non-diatom

stramenopiles, including the first transformation protocol for the eustigmatophyte *Nannochloropsis oceanica*, and an alternative protocol for the labyrinthulomycete *Aurantiochytrium limacinum*[20], both of which are used for biotechnological applications. Furthermore, we report on advances for CRISPR/Cas-driven gene knockouts in *Phaeodactylum tricornutum*[8,13] and a more efficient bacterial conjugation system for *Thalassiosira pseudonana*[13].

Microparticle bombardment was used on *F. cylindrus* (CCMP1102) which was grown, processed and maintained at 4 °C in 24 h light. Exponential phase cells were harvested onto a 1.2 µm membrane filter which was then placed on an 1.5% agar Aquil plate for bombardment with beads coated with a plasmid containing zeocin resistance and *eGFP*, both controlled by an endogenous fucoxanthin chlorophyll *a/c* binding protein (FCP) promoter and terminator (**Table 1; Suppl. Table 3; Online Methods**)[34]. Transformation was performed using 0.7 µm tungsten particles and the biolistic particle delivery system PDS1000/He (BioRad). Rupture discs for 1350 and 1550 pounds per square inch (psi) gave the highest colony numbers with efficiencies of 20.7 colony forming units (cfu)/10^8 cells and 30 cfu/10^8 cells, respectively. Following bombardment, the filter was turned upside down and left to recover for 24 h on the plate, then cells were rinsed from the plate/filter and spread across five 0.8% agar Aquil plates with 100 µg/ml zeocin. Colonies appeared 3 to 5 weeks later. PCR on genomic DNA showed that 100% and 60% of colonies screened positive for the bleomycin gene (*ShBle*) for zeocin resistance and the gene encoding eGFP, respectively. Confirmed by fluorescence-activated cell sorting (FACS) and microscopy, eGFP was localized to the cytosol and was distinguishable from plastid autofluorescence (**Fig. 2**). Additional confirmation by PCR and RT-PCR (**Fig. 4c**) revealed that the *ShBle* and *eGFP* genes were present in the genomes of transformants after multiple transfers (> 10) 2 years later, indicating long-term stability.

Bacterial conjugation methods were improved in *T. pseudonana* (CCMP1335) using the silaffin precursor *TpSil3p* (**Table 1; Online Methods**) as the target gene. *TpSil3p* was fused to *eGFP* flanked by an FCP promoter and terminator, cloned into a pTpPuc3 episomal backbone, and transformed into mobilization plasmid-containing EPI300 *E. coli* cells (Lucigen). The donor cells were grown in SOC medium at 37 °C until OD_{600} of 0.3–0.4, centrifuged and resuspended in 267 µl SOC medium. Next, 200 µl of donor cells were mixed with *T. pseudonana* cells, co-cultured on pre-dried 1% agar plates, dark incubated at 30 °C for 90 min, then at 18 °C in constant light for 4 h, followed by selection in 0.25% agar plates containing 100 µg/ml NTC. Colonies were observed after 2 weeks, inoculated into 300 µl L1 medium and supplemented with 200 µg/ml NTC to reduce the number of false positives. Positive transformants were identified by colony PCR screening (**Suppl. Fig. 3**) and epifluorescence microscopy (**Fig. 2**).

The diatom *Pseudo-nitzschia multiseries* (15093C) and other members of this genus form buoyant linear chains with overlapping cell tips during active growth, and were unconducive to punctate colony formation on agar, where their growth is generally poor. To address this challenge, a low-gelation-temperature agarose seawater medium (LGTA) was developed to facilitate growth, antibiotic selection, and cell recovery. *P. multiseries* exhibited growth inhibition at relatively low concentrations under NTC, formaldehyde, and zeocin (**Suppl. Table 3**). Biolistic transformation of two other *Pseudo-nitzschia* species had been demonstrated at low efficiency[35]. To complement this approach and explore potentially higher efficiency methods for transformation with diatom episomal plasmids, we modified the existing conjugation-based method[13]. The published conjugation protocol was modified to enhance *P. multiseries* post-conjugation viability by reducing SOC content. An episomal version of the Pm_actP_egfp_actT expression cassette was transfected into *E. coli* EPI300+pTAMOB and used for conjugation (**Table 1; Online Methods**). After 48 h in L1

medium, cells were plated in LGTA and eGFP-positive cells were observed 7 days later (**Fig. 2**). PCR revealed the presence of plasmids in all eGFP positive colonies (**Suppl. Fig. 4**). Similarly, conjugation with the episome pPtPUC3 (bleomycin selection marker)-containing bacterial donors was followed under zeocin selection (200 μg/ml). After 7 days, only viable cells (based on bright chlorophyll fluorescence) contained the episome, as confirmed by PCR. Propagation of transformants after the first medium transfer (under selection) has so far been unsuccessful.

Stable transformation of *A. limacinum* (ATCC MYA-1381) was achieved by knock-in of a resistance cassette composed of *ShBle* driven by 1.3 kb promoter and 1.0 kb terminator regions of the endogenous glyceraldehyde-3-phosphate dehydrogenase gene carried in a pUC19-based plasmid (18GZG) along with the native 18S rRNA gene, and by knock-in of a similar construct containing a *eGFP:ShBle* fusion (**Suppl. Fig. 5**). Approximately 1 x 10^8 cells were electroporated, adapting the electroporation protocol used for *Schizochytrium*[36]. The highest transformation efficiency was achieved using 1 μg of linearized 18GZG plasmid with 2 pulses, resulting in a time constant of ~5 ms (**Table 1; Online Methods**). Expression of the fusion protein was confirmed by both the zeocin-resistance phenotype and the detection of eGFP (**Fig. 2**). Six 18GZG transformants derived from uncut and linearized plasmids were examined in detail. All maintained antibiotic resistance throughout 13 serial transfers, first in selective, and subsequently in non-selective media, and then again in selective medium. Integration of the plasmid into the genome was confirmed by PCR as well as by Southern blots using a digoxigenin-labeled *ShBle* gene probe, showing that 4 transformants had integrations by single homologous recombination, while in 2 transformants, additional copies of the antibiotic resistance cassette were integrated by non-homologous recombination elsewhere in the genome (**Suppl. Fig. 5**).

Electroporation of *N. oceanica* (CCMP1779) was optimized based on observation of cells treated with fluorescein-conjugated 2000 kDa dextran and subsequent survival (**Table 1; Online Methods**). A sorbitol concentration of 800 mM and electroporation at between 5 and 9 kV/cm resulted in highest cell recovery. These conditions were used during introduction of plasmids containing the gene for the blue fluorescent reporter *mTagBFP2* under the control of cytomegalovirus (CMV), the cauliflower Mosaic Virus 35S, or the VCP1 promoter previously described from *Nannochloropsis* sp.[37]. Transient expression of blue fluorescence (compared to cells electroporated simultaneously under the same conditions without plasmid) appeared within 2 h, lasted for at least 24 h, and disappeared by 48 h in subsets of cells electroporated with *mTagBFP2* under the control of CMV (**Suppl. Fig. 6**). The transient transformation was more effective when a linearized plasmid was used compared to a circular plasmid (**Table 1**). VCP1 did not induce blue fluorescence with a circular plasmid, while 35S gave inconsistent results with either circularized or linearized plasmids.

For *P. tricornutum* (CCAP1055/1), we adapted the CRISPR/Cas9 system[8] for multiplexed targeted mutagenesis. Bacterial conjugation[13] was used to deliver an episome that contained a Cas9 cassette and 2 single-guide RNA (sgRNA) expression cassettes designed to excise a 38 bp-long domain from the coding region of a nuclear-encoded, chloroplastic glutamate synthase (*Phatr3_J24739*) and introduce an in-frame stop codon after strand ligation (**Table 1; Online Methods**). The GoldenGate assembly was used to clone 2 expression cassettes carrying sgRNAs into a *P. tricornutum* episome that contained a *Cas9-2A-ShBle* expression cassette and the centromeric region CenArsHis (**Suppl. Fig. 7**). After their addition to a *P. tricornutum* culture, plates were incubated in a growth chamber under standard growth conditions for 2 days and transformed *P. tricornutum* colonies began to appear after 2 weeks. Only colonies maintaining *Cas9-2A-ShBle* sequence on the delivered

episome were able to grow on selection plates because *Cas9* and *ShBle* were transcriptionally fused by the 2A peptide[38] (**Suppl. Fig. 7**). Gel electrophoresis migration and sequencing of the genomic target loci confirmed the 38 bp-long excision and premature stop codon (**Fig. 4d**).

Alveolates. This species-rich and diverse group is comprised of ciliates, apicomplexans, and dinoflagellates (**Fig. 1**). As a link between apicomplexan parasites and dinoflagellate algae, perkinsids are key for understanding the evolution of parasitism, and also have potential biomedical applications[17]. Techniques currently exist for transformation of only a small number of ciliates, perkinsids and apicomplexans[39]. Here, we present a first transformation protocol for *Karlodinium veneficum* (CCMP1975), a phagotrophic mixotroph that produces fish-killing karlotoxins[40]. Experiments were also performed on *Oxyrrhis marina* (CCMP 1788/CCMP 1795), a basal-branching phagotroph that lacks photosynthetic plastids and *Crypthecodinium cohnii* (CCMP 316), a heterotroph used in food supplements. For both of these taxa first evidence of DNA delivery was achieved (**Table 1; Suppl. Results; Supp. Fig. 15; Online Methods**), a goal recently achieved for *C. cohnii* using electroporation[19]. Additionally, we report on improved transformation systems for *Perkinsus marinus* (PRA240) and *Amphidinium carterae* (CCMP1314) chloroplast, published recently as part of the EMS initiative[15].

K. veneficum (CCMP1975) was transformed based on electroporation and cloning the selectable marker gene aminoglycoside 3'-phosphotransferase (*nptII/neo*; note that *nptII/neo* is used synonymously with amino 3’-glycosyl phosphotransferase gene conferring resistance to kanamycin, neomycin, paromomycin, ribostamycin, butirosin and gentamicin B) into the backbone of the dinoflagellate-specific expression vector DinoIII-*neo*[41], which confers resistance to neomycin and kanamycin (**Table 1**). In brief, DinoIII-*neo* was linearized and

electroporated using the Nucleofector optimization pulse codes, buffer SF/Solution I (Lonza), and 2 μg/μl of linearized DinoIII-*neo*. Electroporated cells were selected under 150 μg/ml kanamycin 3 days post-electroporation. Fresh seawater with kanamycin was added every 2 weeks to the cultures and new subcultures were inoculated monthly. After 3 months, DNA and RNA were isolated from the resistant cultures as previously reported[42] and cDNA was synthesized using random hexamers. Out of 16 transformations, 2 cell lines (CA-137, DS-138) showed stable growth under kanamycin selection. CA-137 developed dense cultures after 3 months, and the resistance gene was detected in both DNA and RNA by nested PCR and RT-PCR, respectively (**Fig. 4e; Suppl. Fig. 8; Online Methods**).

We improved the transformation protocol[16,17] of *P. marinus*, a pathogen of marine mollusks, fish, and amphibians[43] (**Suppl. Table 5**). We co-expressed 2 genes and efficiently selected transient and stable transformants using FACS (**Table 1; Figs. 2** and **4f; Suppl. Fig. 9; Online Methods**). In addition, we established the integration profile of ectopic DNA once introduced into the *P. marinus* genome. We did not see evidence of integration through homologous recombination and observed a propensity for plasmid fragmentation and integration within transposable elements sites. An optimized alternative protocol for transformation using glass bead abrasion was also developed. Two versions of the previously published *Moe* gene promoter[16] were tested. Whereas the 1.0 kb promoter version induced expression after 2 or 3 days, the truncated version (0.5 kb) took 7 days for expression to be detected. Resistance genes to zeocin, blasticidin and puromycin have all been shown to confer resistance to transformed *P. marinus*; however, selection regimes are still relatively slow and inefficient, indicating further room for improvement[17].

We also report a new vector for the transformation of the *A. carterae* chloroplast, a photosynthetic dinoflagellate. *A. carterae*, like other dinoflagellates with a peridinin-containing chloroplast, contains a fragmented chloroplast genome made up of multiple

plasmid-like minicircles[40]. The previous transformation protocols made use of this to introduce two vectors based on the psbA minicircle[15]. Here, we show that other minicircles are also suitable for use as vectors. We created a new artificial minicircle, using the atpB minicircle as a backbone, but replacing the *atpB* gene with a codon-optimized chloramphenicol acetyl transferase (**Table 1; Online Methods**). This circular vector was introduced by biolistics to *A. carterae* (**Suppl. Fig. 10a**). Following selection with chloramphenicol, we were able to detect transcription of the chloramphenicol acetyl transferase gene via RT-PCR (**Fig. 4g**). This result suggests that all 20 or so minicircles in the dinoflagellate chloroplast genome would be suitable for use as artificial minicircles, thus providing a large pool of potential vectors.

Discobans. This diverse group, recently split into Discoba and Metamonada[44], includes heterotrophs, photoautotrophs, predatory mixotrophs, as well parasites. Discobans include parasitic kinetoplastids with clinical significance, such as *Trypanosoma brucei*, *T. cruzi* and *Leishmania* spp., for which efficient transformation protocols are available[45]. However, such protocols are missing for aquatic species. Here, we describe available transformation protocols for the kinetoplastid *Bodo saltans* and the heterolobosean *Naegleria gruberi*. The former was isolated from a lake, but identical 18S rRNA gene sequences have been reported from marine environments[46]. The latter is a freshwater protist that represents a model organism for closely related marine heterolobosean amoebas. Furthermore, we provide advanced methods that build on prior EMS results[18] for the diplonemid *Diplonema papillatum*.

B. saltans (ATCC 30904) was transformed with a plasmid containing a cassette designed to fuse an endogenous *EF-1*α gene with *eGFP* for C-terminal tagging. This cassette includes downstream of *eGFP*, a *B. saltans* tubulin intergenic region followed by the

selectable marker *nptII/neo* gene, conferring resistance to neomycin. *EF-1α* genes exist in tandem repeats. The homologous regions that flank the cassette were chosen as targets for inducing homology-directed repair, however, they target only one copy of the gene. As transcription in *B. saltans* is polycistronic[46], insertion of the tubulin intergenic region into the plasmid is essential for polyadenylation of the *EF1-α/GFP* fusion and *trans*-splicing of the *nptII/neo* gene (**Suppl. Table 5**). Selection of transfected cells began with 2 µg/ml of neomycin added 24 h after electroporation, and this concentration was gradually increased over 2 weeks to 5 µg/ml (**Table 1; Online Methods**). Cells were washed and subcultured into fresh selection medium every 4 days, and neomycin-resistant cells emerged 7-9 days post-electroporation. The eGFP signal was detected 2 days post-electroporation, albeit with low intensity. This may be due to the inefficient translation of *eGFP* since it has not been codon-optimized for *B. saltans* (**Fig. 2**). Genotyping analysis 9 months post-transfection confirmed the presence of the *nptII/neo* gene and at least partial plasmid sequence (**Fig. 4h; Suppl. Fig. 10b**). However, plasmid integration into the *B. saltans* genome through homologous recombination is still unconfirmed. This suggests either off-target plasmid integration or episomal maintenance.

For *N. gruberi* (ATCC 30224) two plasmids were designed. The first one carried the hygromycin B resistance gene (*hph*) with an actin promoter and terminator, along with an HA-tagged *eGFP* driven by the ubiquitin promoter and terminator. The second plasmid carried the *nptII/neo* gene instead. For each individual circular plasmid, 4 µg was electroporated (**Table 1; Online Methods**). About 48 h after electroporation, dead cells were removed from the suspension and viable cells were washed with PBS. Afterwards, 300 µg/ml of hygromycin B or 700 µg/ml of neomycin was added to the fresh media. One to 4 weeks later, resistant clones were recovered and expression of *eGFP* and/or hygromycin was confirmed by western blotting (**Suppl. Fig. 11**). Expression of *eGFP* was observed by

epifluorescence microscopy (**Fig. 2; Suppl. Fig. 11**) with ~80% of transformants maintaining hygromycin B or neomycin resistance in addition to expressing *eGFP*.

D. papillatum (ATCC 50162) was transformed by electroporation using 3 μg a *SwaI*-linearised fragment (cut from p57-V5+NeoR plasmid) containing the V5-tagged *nptII/neo* gene flanked by partial regulatory sequences derived from the hexokinase gene of the kinetoplastid *Blastocrithidia* (strain p57) (**Table 1; Online Methods**) using a published protocol[18]. About 18 h after electroporation, 75 μg/ml G418 was added to the medium and after 2 weeks 7 neomycin-resistant clones were recovered. Transcription of *nptII/neo* was verified in 4 clones by RT-PCR (**Suppl. Fig. 12**) and the expression of the tagged *nptII/neo* protein was confirmed in 2 clones by western blotting using the α-V5 antibody (**Fig. 4i**).

Opisthokonts

The opisthokont clade Holozoa includes animals and their closest unicellular relatives choanoflagellates, filastereans, ichthyosporeans, and corallochytreans. The establishment of genetic tools in non-metazoan holozoans promises to help illuminate the cellular and genetic foundations of animal multicellularity[47]. Genomic and transcriptomic data are available for multiple representatives characterized by diverse cell morphologies, some of which can even form multicellular structures[46]. Untill recently, only transient transformations had been achieved for some opistokonts such as the filasterean *Capsaspora owczarzaki*[48], the ichthyosporean *Creolimax fragrantissima*[49] and the choanoflagellate *Salpingoeca rosetta*[21]. Through the EMS initiative, we report on the first evidence for transient transformation of the ichthyosporean *Abeoforma whisleri*, isolated from the digestive tract of mussels, and review a recently published stable transformation protocol for *S. rosetta* achieved by using the selectable puromycin N-acetyl-transferase gene (**Fig. 2**)[22].

All *A. whisleri* life stages are highly sensitive to a variety of methods for transformation. However, we developed a 4D-nucleofection-based protocol using 16-well strips, wherein PBS-washed cells were resuspended in 20 µl of buffer P3 (Lonza) containing 40 µg of carrier plasmid (empty pUC19) and 1-5 µg of the reporter plasmid (*A. whisleri* H2B fused to mVenus fluorescent protein, mVFP) (**Table 1; Online Methods**), and subjected to code EN-138 (Lonza). Immediately after the pulse, cells were recovered by adding 80 µl of marine broth (Gibco) prior to plating in 12-well culture plates previously filled with 1 ml marine broth. After 24 h, ~1% of the culture was transformed based on the fraction of cells expressing *mVFP* in the nucleus (**Figs. 2** and **4j**).

Microbial eukaryotes in natural planktonic communities

Model organisms are typically selected based on criteria such as relative ease of isolation and asexual cultivation in the laboratory, however these attributes may not correlate with the capacity for uptake and expression of the exogenous DNA. We explored whether natural marine planktonic pico- and nanoeukaryote communities would take up DNA in a culture-independent setting. Microbial plankton from natural seawater was concentrated and electroporated with plasmids containing mTagBFP2 under the control of CMV or 35S promoters (**Suppl. Results; Online Methods**). In most trials, blue fluorescent cells were rare if detected at all (compared to control samples). However, in one natural community tested, a photosynthetic picoeukaryote population exhibited up to 50% of cells with transient expression of blue fluorescence when the CMV promoter was used (**Suppl. Fig. 13**). This suggests it might be possible to selectively culture eukaryotic microorganisms based on capacity to express exogenous DNA.

Discussion

The collaborative effort by the EMS initiative facilitated identification and optimization of the steps required to create new protist model systems, which culminated in the synthetic 'Transformation Roadmap' (**Fig. 5**). Our genetic manipulation systems for aquatic (largely marine) protists will enable deeper insights into their cell biology, with potentially valuable outcomes for aquatic sciences, evolutionary studies, nanotechnology, biotechnology, medicine, and pharmacology. Successes and failures with selectable markers, transformation conditions, and reporters were qualitatively compared across species (**Suppl. Tables 3** and **4**; **Table 1; Figs. 2, 3** and **4; Online Methods**).

For some of the selected species, the first step was to identify cultivation conditions for robust growth in the laboratory to either generate high cell densities or large culture volumes for obtaining sufficient biomass required for a variety of molecular biology experiments. Unlike established microbial model species, cultivation of marine protists can be challenging especially under axenic conditions or for predatory taxa that require co-cultivation with their prey. Nevertheless, 13 out of 35 species were rendered axenic prior to the development of transformation protocols. For the remaining species, we were unable to remove bacteria and therefore had to make sure that transformation signals were coming from the targeted protist rather than contaminants (**Suppl. Table 2**). Subsequent steps included the identification of suitable antibiotics and their corresponding selectable markers (**Table 1; Suppl. Table 3**), conditions for introducing exogenous DNA (**Table 1; Suppl. Table 4**), and selection of promoter and terminator sequences for designing transformation vectors (**Table 1; Online Methods – Suppl. Table 5 and Suppl. Notes 1**).

As exemplified in the new model systems provided herein (**Table 1; Figs. 2, 3** and **4**), a variety of methods were used to test whether exogenous DNA was integrated into the genome or maintained as a plasmid, and whether the introduced genes were expressed.

Approaches to show the former included inverse PCR, Southern blotting and whole genome sequencing, whereas approaches to demonstrate the latter included various combinations of PCR, RT-PCR, western blotting, epifluorescence microscopy, FACS, antibody-based methods, and/or growth assays in the presence of antibiotics to confirm transcription and translation of introduced selection and reporter genes (e.g. *eGFP*, *YFP*, *mCherry*). For fluorescent markers, it was first ensured that the wild type, or manipulated controls cells, had no signals conflicting with the marker (**Figs. 2** and **3c**), an important step because photosynthetic protists contain chlorophyll and other autofluorescent pigments. Overall transformation outcomes for each species were parsed into three groups according to the level of success or lack thereof (A = first transformation protocol for a given species; B = advanced protocol based on prior work; C = published protocol based on the EMS initiative) and are discussed according to their phylogenetic position (**Fig. 1**).

Our studies did not result in a universally applicable protocol because transformability and a range of other key conditions varied greatly across taxa and approaches, such as intrinsic features of the genome and differences in cellular structure and morphology. In general, electroporation proved to be the most common method for introducing exogenous DNA stably into cells. This approach was utilized for naked cells and protoplasts, yet frequently also worked, albeit with lower efficiency, on cells protected by cell walls. Linearized plasmids were most effective for delivery, and 5' and 3' UTRs-containing promotors of highly expressed endogenous genes provided the strongest expression of selective reporters and markers. If successful, teams usually continued with fluorescence-based methods. Furthermore, large amounts of carrier DNA usually facilitated successful initial transformations (e.g. *M. commoda*, *A. whisleri*) or improved existing protocols (*S. rosetta*[21]). We also provide the contact details of all co-authors who are assigned to particular species (**Suppl. Table 6)**.

Some lineages were difficult to transform, especially dinoflagellates and coccolithophores. Here, even if DNA appeared to be delivered (**Suppl. Table 5**), expression of the transformed genes could not be confirmed. Examples include the dinoflagellates *C. cohnii* and *Symbiodinium microadriaticum*, and the coccolithophore *E. huxleyi*. Thus, at least these 3 species need concerted future efforts.

The combination of results presented herein together with previously published protocols from the EMS initiative[50] significantly expands the segment of extant eukaryotic diversity amenable to reverse genetics approaches. Out of the 39 microbial eukaryotes selected for the initiative, exogenous DNA was delivered and expressed in more than 50% of them. The transformation systems enable us for the first time to shed light on the function of species-specific genes, which likely reflect key adaptations to specific niches in dynamic ocean habitats.

ACKNOWLEDGEMENTS

We thank M. Salisbury and D. Lacono, C. Poirier, M. Hamilton, C. Eckmann, H. Igel, C. Yung and K. Hoadley for assistance; and V. K. Nagarajan, M. Accerbi, and P. J. Green who carried out *Agrobacterium* studies in *Heterosigma akashiwo*; and N. Kraeva, C. Bianchi and V. Yurchenko for the help with designing the p57-V5+NeoR construct. We are also grateful to the protocols.io team (L. Teytelman and A. Broellochs) for their support and thank four anonymous reviewers for their constructive criticisms of our manuscript. This collaborative effort was supported by the Gordon and Betty Moore Foundation EMS Program of the Marine Microbiology Initiative (grant numbers: GBMF4972 and 4972.01 to F.Y.B.; GBMF4970 and 4970.01 to M.A. & A.Z.W.; GBMF3788 to A.Z.W.; GBMF 4968 and 4968.01 to H.C.; GBMF4984 to V.H.; GBMF4974 and 4974.01 to C. Brownlee; GBMF4964 to Y. Hirakawa; GBMF4961 to T. Mock; GBMF4958 to P.S.; GBMF4957 to A.T.;

GBMF4960 to G.J.S.; GBMF4979 to K.C.; GBMF4982 and 4982.01 to J.L.C.; GBMF4964 to P.J.K.; GBMF4981 to P.V.D.; GBMF5006 to A.E.A.; GBMF4986 to C.M.; GBMF4962 to J.A.F.R.; GBMF4980 and 4980.01 to S.L.; GBMF 4977 and 4977.01 to R.F.W.; GBMF4962.01 to C.H.S.; GBMF4985 to J.M.; GBMF4976 and 4976.01 to C.H.; GBMF4963 and 4963.01 to V.E.; GBMF5007 to C.L.D; GBMF4983 and 4983.01 to J.L; GBMF4975 and 4975.01 to A.T.; GBMF4973 and 4973.01 to I.R.T. and GBMF4965 to N.K.), by The Leverhulme Trust (RPG-2017-364) (to T. Mock and A. Hopes) and by ERD Funds (16_019/0000759) from the Czech Ministry of Education (to JL).

AUTHOR CONTRIBUTIONS

The project was conceived and designed by A.C.J., J.Z.K., S.B., D.F., J.L., R.E.R.N., J.A.F.R., E.C., L.S., A.Z.W., T. Mock, A.E.A., F.Y.B, C. Brownlee, C. Bowler, H.C., T.C., J.L.C., K.C., C.L.D., V.E., V.H., Y. Hirakawa., C.J.H., P.J.K., N.K., S.L., C.M., J.M., I.R.T., P.A.S., C.H.S., G.J.S., A.D.T., P.V.D., A.T. and R.F.W. Data analysis was carried out by M.A.J., C.A., C. Balestreri, A.C.B., P.B., D.S.B., S.A.B., G.B., R.C., M.A.C., D.B.C., E.C.C., R.D., E.E., P.A.E., F.F., V.F.B., N.J.F., K.F., P.A.G., P.R.G., F.G., S.G.G., J.G., Y. Hanawa, E.R.H.C., E.H., A. Highfield, A. Hopes, 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., J.C.L., F.L., S.M., T. Matute, M.M., S.R.N., D.N., I.C.N., L.N., A.M.G.N.V., M.N., I.N., A. Pain, A. Piersanti, S.P., J.P., J.S.R., M.R., D.R., A.R., M.A.S., E.C.S., B.N.S., R.S.,T.V.H., L.T., J.T., M.V., V.V., L.W., X.W., G.W., A.W. and H.Z. The manuscript was written by D.F., R.E.R.N., J.A.F.R., E.C., L.S., T. Mock, A.Z.W. and J.L. with input from all authors.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E. & Keeling, P. J. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* **347**, 1257594 (2015).

2. de Vargas, C. et al. Eukaryotic plankton diversity in the sunlit global ocean. *Science* **348**, 1261605 (2015).

3. Collier, J. L. & Rest, J. S. Swimming, gliding, and rolling toward the mainstream: cell biology of marine protists. *Mol. Biol. Cell.* **30**, 1245-1248 (2019).

4. Curtis, B. A. et al. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**, 59–65 (2012).

5. Armbrust, E. V. et al. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306**: 79-86 (2004).

6. Read, B. A. et al. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature* **499**, 209-213 (2013).

7. Keeling P. J. et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* **12**, e1001889 (2014).

8. Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M. & Winge, P. A. CRISPR/Cas9 system adapted for gene editing in marine algae. *Sci. Rep.* **6**, 24951 (2016).

9. Hopes, A., Nekrasov, V., Kamoun, S. & Mock, T. Editing of the urease gene by CRISPR-Cas in the diatom *Thalassiosira pseudonana*. *Plant Methods* **12**, 49 (2016).

10. Carradec, Q. et al. A global atlas of eukaryotic genes. *Nat. Commun.* **9**, 373 (2018).

11. Apt, K. E., Kroth-Pancic, P. G. & Grossman, A. R. Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Mol. Gen. Genet*. **252**, 572-579 (1996).

12. Paschke, P. et al. Genetic engineering of *Dictyostelium discoideum* cells based on selection and growth on bacteria. *J. Vis. Exp*. **143**, e58981 (2019).

13. Karas, B. J. et al. Designer diatom episomes delivered by bacterial conjugation. *Nat. Commun.* **6**, 6925 (2015).

14. Hirakawa, Y., Kofuji, R. & Ishida, K. Transient transformation of a chlorarachniophyte alga, *Lotharella amoebiformis* (chlorarachniophyceae), with *uidA* and *egfp* reporter genes. *J. Phycol*. **44**, 814–20 (2008).

15. Nimmo, I. C. et al. Genetic transformation of the dinoflagellate chloroplast. *eLife* **8**, e45292 (2019).

16. Fernández Robledo, J. A., Lin, Z. & Vasta, G. R. Transfection of the protozoan parasite *Perkinsus marinus*. *Mol. Biochem. Parasitol*. **157**, 44-53 (2008).

17. Sakamoto, H., Hirakawa, Y., Ishida, K. I., Keeling, P. J., Kita, K. & Matsuzaki M. Puromycin selection for stable transfectants of the oyster-infecting parasite *Perkinsus marinus*. *Parasitol. Int*. **69**, 13-16 (2018).

18. Kaur, B. et al. Transformation of *Diplonema papillatum*, the type species of the highly diverse and abundant marine micro-eukaryotes Diplonemida (Euglenozoa). *Env. Microbiol*. **20**, 1030-1040 (2018).

19. Diao, J., Song, X., Zhang, X., Chen, L. & Zhang W. Genetic engineering of *Crypthecodinium cohnii* to increase growth and lipid accumulation. *Front Microbiol*. 9, 492 (2018).

20. Sakaguchi K. et al. Versatile transformation system that is applicable to both multiple transgene expression and gene targeting for thraustochytrids. *Appl. Environ. Microbiol.* **78**: 3193-3202 (2012).

21. Booth, D., Middleton, H. & King, N. Choanoflagellate transfection illuminates their cell biology and the ancestry of animal septins. *Mol. Biol. Cell*. **29**, 3026-38 (2018).

22. Wetzel, L. A. et al. Predicted glycosyltransferases promote development and prevent spurious cell clumping in the choanoflagellate *S. rosetta*. *eLife* **7,** e41482 (2018).

23. van Baren, M. J. et al. Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genomics* **17**, 1–22 (2016).

24. Lozano, J. C. et al. Efficient gene targeting and removal of foreign DNA by homologous recombination in the picoeukaryote *Ostreococcus*. *Plant J*. **78**, 1073-1083 (2014).

25. Van Ooijen, G., Knox, K., Kis, K., Bouget, F. Y. & Millar, A. J. Genomic transformation of the picoeukaryote *Ostreococcus tauri. J. Vis. Exp*. **65**, e4074 (2012).

26. Hovde, B. T. et al. Genome sequence and transcriptome analyses of *Chrysochromulina tobin*: Metabolic tools for enhanced algal fitness in the prominent order Prymnesiales (Haptophyceae). *PLoS Genet*. **11**, e1005469 (2015).

27. Endo, H., Yoshida, M., Uji, T., Saga, N., Inoue, K. & Nagasawa, H. Stable nuclear transformation system for the coccolithophorid alga *Pleurochrysis carterae*. *Sci. Rep*. **6**, 22252 (2016).

28. Dörner, J., Carbonell, P., Pino, S. & Farias, A. Variation of fatty acids in *Isochrysis galbana* (T-Iso) and *Tetraselmis suecica*, cultured under different nitrate availabilities. *J. Fish. Aquacult.* **5**, 1-3 (2014).

29. Dunahay, T. G., Jarvis, E. E. & Roessler, P. G. Genetic transformation of the diatoms *Cyclotella cryptica* and *Navicula saprophila*. *J. Phycol.* **31**, 1004–1012 (1995).

30. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**, 237-240 (1998).

31. Mishra, M., Arukha, A. P., Bashir, T., Yadav, D. & Prasad, G. B. K. S. All new faces of diatoms: Potential source of nanomaterials and beyond. *Front. Microbiol.* **8**, 1239 (2017).

32. Mock, T. et al. Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**, 536–540 (2017).

33. Brunson, J. K. et al. Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom. *Science* **361**, 1356-1358 (2018).

34. Kroth, P. G. Genetic transformation: A tool to study protein targeting in diatoms. *Methods Mol. Biol*. **390**, 257–267 (2007).

35. Sabatino, V., Russo, M. T., Patil, S., d'Ippolito, G., Fontana, A. & Ferrante, M. I. Establishment of genetic transformation in the sexually reproducing diatoms Pseudo-*nitzschia multistriata* and *Pseudo-nitzschia arenysensis* and inheritance of the transgene. *Marine Biotech.* **17**, 452-462 (2015).

36. Ono, K., Aki, T. & Kawamoto, S. Method for introducing a gene into labyrinthulomycota. US Patent 7,888,123 (2011).

37. Kilian, O., Benemann, C. S., Niyogi, K. K. & Vick, B. High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp. *Proc. Natl. Acad. Sci. USA* **108**, 21265-21269 (2011).

38. Duda, K. et al. High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases of CRISPR/Cas9 nicase pairs. *Nucl. Acids. Res.* **42**, e84 (2014).

39. Donald, R. G. & Roos, D. S. Stable molecular transformation of *Toxoplasma gondii*: a selectable dihydrofolate reductase-thymidylate synthase marker based on drug-resistance mutations in malaria. *Proc. Natl. Acad. Sci. USA* **90**, 11703-11707 (1993).

40. Barbrook, A. C., Howe, C.J. & Nisbet, R.E.R. Breaking up is hard to do: the complexity of the dinoflagellate chloroplast genome. *Perspect. Phycol.* **6**, 31-37 (2019).

41. Sprecher, B. N., Zhang, H. & Lin, S. Nuclear gene transformation in the dinoflagellate *Oxyrrhis marina*. *Microorganisms* **8**, 126 (2020).

42. Zhang, H., Campbell, D. A., Sturm, N. R., Rosenblad, M. A., Dungan, C. F. & Lin, S. Signal recognition particle RNA in dinoflagellates and the perkinsid *Perkinsus marinus*. *Protist* **164**, 748-761 (2013).

43. Chambouvet, A. et al. Cryptic infection of a broad taxonomic and geographic diversity of tadpoles by Perkinsea protists. *Proc. Natl. Acad. Sci. USA* **112**, E4743-4751 (2015).

44. Adl, S. M. et al. Revision to the classification, nomenclature and diversity of eukaryotes. *J. Euk. Microbiol.* **66**, 4-119 (2019).

45. Matthews, K. R. 25 years of African trypanosome research: From description to molecular dissection and new drug discovery. *Mol. Biochem. Parasitol.* **200**, 30-40 (2015).

46. Opperdoes, F. R., Butenko, A., Flegontov, P., Yurchenko, V. & Lukeš, J. Comparative metabolism of free-living *Bodo saltans* and parasitic trypanosomatids. *J. Eukaryot. Microbiol.* **63**, 657-678 (2016).

47. Richter, D. J., Fozouni, P., Eisen, M. & King, N. Gene family innovation, conservation and loss on the animal stem lineage. *eLife* **7**, 1–43 (2018).

48. Parra-Acero, H. et al. Transfection of *Capsaspora owczarzaki*, a close unicellular relative of animals. *Development* **145**, 162107 (2018).

49. Suga, H. & Ruiz-Trillo, I. Development of ichthyosporean sheds light on the origin of metazoan multicellularity. *Dev. Biol*. **377**, 284-292 (2013).

50. Waller, R. F. et al. Strength in numbers: collaborative science for new experimental model systems. *PLoS Biol.* **16**, e2006333 (2018).

FIGURE LEGENDS

Fig. 1. | Phylogenetic relationships and transformation status of marine protists. A schematic view of the eukaryotic tree of life with effigies of main representatives. Colour-coordinated species we have attempted to genetically modify are listed below. Current transformability status is schematized in circles indicating: DNA delivered and shown to be expressed (yellow; for details see text and Table 1); DNA delivered, but no expression seen

(grey); no successful transformation achieved despite efforts (blue). The details of transformation of species that belong to “DNA delivered” and “Not achieved yet” categories are described in the **Suppl. Table 5**.

Fig. 2. | Epifluorescence micrographs of transformed marine protists

Representative images of transformants and wild-type cell lines of 10 selected protist species. Colored boxes behind species names refer to phylogenetic supergroup assignments given in Fig. 1. Representative data of at least two independent experiments are shown. The fluorescent images show the expression of individual fluorescent marker genes introduced via transformation for all organisms shown, except in the case of *A. amoebiformis.* For the latter, red depicts the natural autofluorescence of photosynthetic pigments in the cell, while the additional green spheres in the transformant fluorescence panel shows introduced GFP fluorescence (see **Suppl. Fig. 15C** for a trace of these different regions in the cell). Scale bars are as follows: 10 µm for *A. amoebiformis*, *T. pseudonana*, *A. limacinum*, *B. saltans*, *N. gruberi*, *A. whisleri*, and *S. rosetta*; 15 µm for *P. marinus*; 20 µm for *F. cylindrus*; 100 µm for *P. multiseries.*

Fig. 3. | Various methods were used to demonstrate successful transformation in different archaeplastid species – Luminescence and Fluorescence.

Luminescence and Fluorescence (by FACS and epifluorescence microscopy) were used to verify expression of introduced constructs in three archaeplastids - *O. lucimarinus*, *B. prasinos*, and *M. commoda* (**a, b, c**). For the latter, red in the image depicts the natural autofluorescence of photosynthetic pigments in the cell, while green shows introduced eGFP fluorescence and blue shows the DAPI stained nucleus; the overlay shows colocalization of eGFP and nucleus signals. See Supplementary Fig. 15d for a trace of these different regions

in the cell. NS, not significant; trans., transformed. Representative data of at least two independent experiments are shown. For detailed figure description see **Suppl. Notes 2**.

Fig. 4. | Various methods were used to demonstrate successful transformation in different species - RT-PCR, western blot and sequencing.

Western blot, RT-PCR or sequencing (in case of Cas9-induced excision by CRISPR) were used to verify expression of introduced constructs in one haptophyte – *I. galbana* (**a**), one rhizarian – *A. amoebiformis* (**b**), two stramenopiles – *F. cylindrus* and *P. tricornutum* (**c, d**), three alveolates – *K. veneficum*, *P. marinus* and *A. carterae* (**e, f, g**), two discobans – *B. saltans* and *D. papillatum* (**h, i**) and one opisthokont – *A. whisleri* (**j**). Note that *nptII/neo* is used synonymously with amino 3'-glycosyl phosphotransferase gene (*aph(3')*) conferring resistance to kanamycin and neomycin. Representative data of at least two independent experiments are shown. For detailed figure description see **Suppl. Notes 2**.

Fig. 5. | 'Transformation Roadmap' for the creation of genetically tractable protists. (**a**) **Vector design and construction for microeukaryotes of interest and a natural community**. (**b**) **Transformation approaches**. Different symbols represent methods (e.g. chemical, physical or biological) for introducing DNA/RNA/protein into a living cell. (**c**) **Protocol**. Key methodological steps for successful transformation are listed in an abbreviated form (for particular examples, see Table 1 and text).

Table 1: Parameters used for successful transformation as shown in Figs. 2, 3 and 4. For additional information see **Suppl. Tables 5, 7 and 8** and **Suppl. Notes 1** and protocols.io. For contacting laboratories working with particular species, see details given in **Suppl. Table 6.**

ONLINE METHODS

Studied species and used transformation methods.

For the full list of vector sequences and maps see **Suppl. Notes 1** and for detailed description of **Figs. 3 and 4** see **Suppl. Notes 2.** Antibiotic concentrations effective for selection of transformats can be found in **Suppl. Table 3**, the details of the transformation methods applied to this study in **Suppl. Table 4** and contact details for individual laboratories in **Suppl. Table 6**. Full list of protists (including details of culture collection) and links to the complete step-by steps transformation protocols and published vector sequences are listed in **Suppl. Table 5**. The protocols.io links listed in **Table 1** and **Suppl. Table 5** are summarized in **Suppl. Tables 7** and **8.**

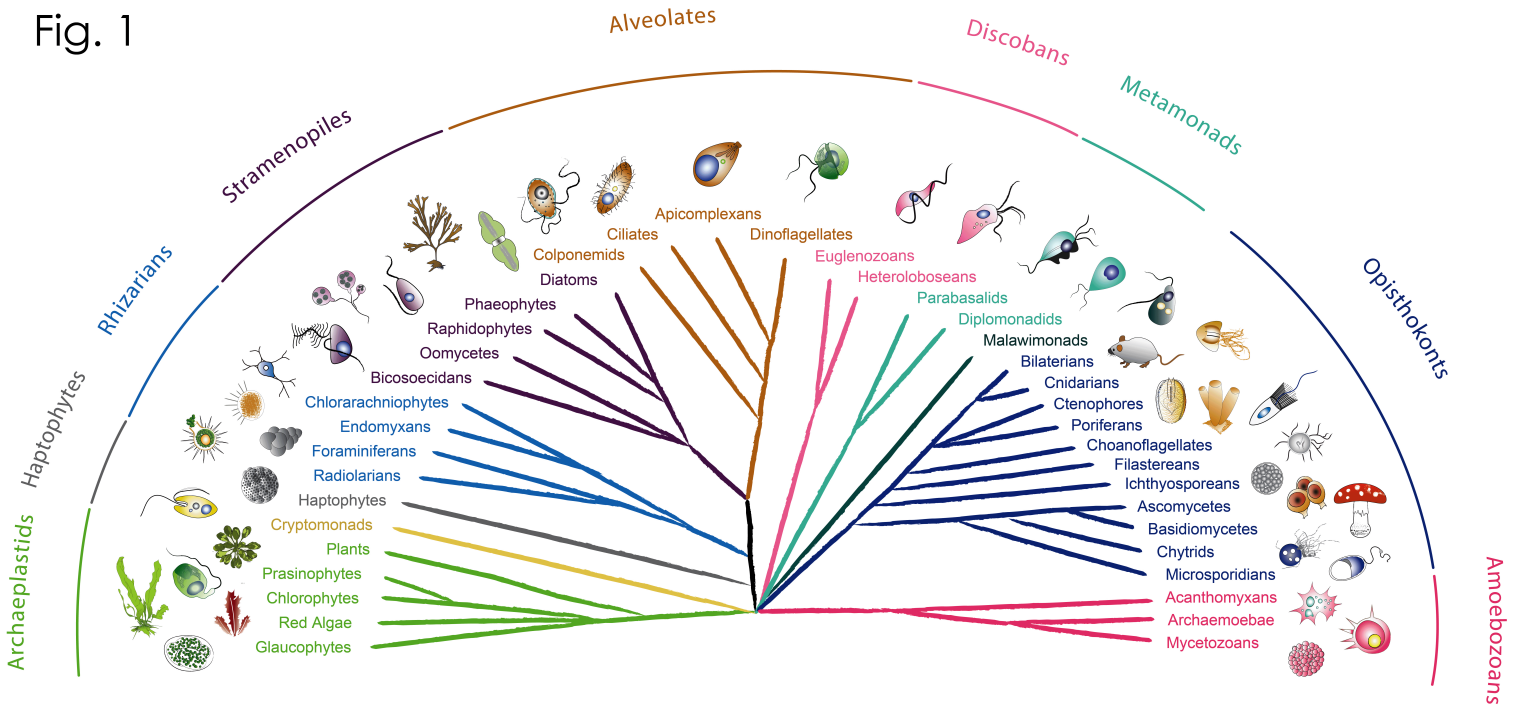
Life Sciences Reporting Summary

Further information is available in the Life Sciences Reporting Summary.

Data availability

The data that support the findings of this study are available from the corresponding authors as well as the other authors upon request (for the contacts see **Suppl. Table 6**). Source data for **Fig. 3, Fig. 4** and **Suppl. Figs. 9B, C; 11A** and **12B, C** are available online.

Fig. 1



Archaeplastids

Ostreococcus lucimarinus	A	R
Bathycoccus prasinos	A	R
Micromonas commoda	A	
Micromonas pusilla		
Tetraselmis striata	A	R
Pyramimonas parkeae		

Haptophytans

Isochrysis galbana	A	R
Emiliania huxleyi		

Rhizarians

Amorphochlora amoebiformis	B[14]	
Bigelowiella natans		

Stramenopiles

Fragilariopsis cylindrus	A	R
Thalassiosira pseudonana	B[9,13]	R
Seminavis robusta		
Pseudo-nitzschia multiseries	A	R
Heterosigma akashiwo		
Aurantiochytrium limacinum	B[20]	R
Caecitellus sp.		
Nannochloropsis oceanica	A	
Phaeodactylum tricornutum	B[9,11]	R

Alveolates

Euplotes crassus		
Euplotes focardii		
Chromera velia		
Perkinsus marinus	B[16,17]	R
Oxyrrhis marina	A	
Hematodinium sp.		
Fugacium (Symbiodinium) kawagutii		
Alexandrium catenella		

Breviolum (Symbiodinium) sp.		
Crypthecodinium cohnii	B[19]	
Amphidinium carterae (chloroplast)	B[15]	R
Karlodinium veneficum	A	R

Discobans

Bodo saltans	A	R
Diplonema papillatum	B[18]	R
Eutreptiella gymnastica		
Naegleria gruberi	A	R

Opisthokonts

Pirum gemmata		
Sphaeroforma arctica		
Abeoforma whisleri	A	
Salpingoeca rosetta	C[21,22]	R

Natural Communities

Pico-Eukaryotes	A	

TRANSFORMATION STATUS

- DNA delivered and expressed
- DNA delivered
- Not yet achieved

A First successful transformation of a particular species
B Alternative transformation / or improvement of the existing transformation method
C Review of an already published / existing protocol

CONSTRUCT

- Homologous
- Heterologous
- Labeled DNA
- mRNA
- FITC–Dextran

DELIVERY SYSTEM

Chemical · Electroporation · Conjugation · Biolistics · Glass beads abrasion · Microinjection

R Resistance selection · CRISPR/Cas9 · Did not work

Fig. 2

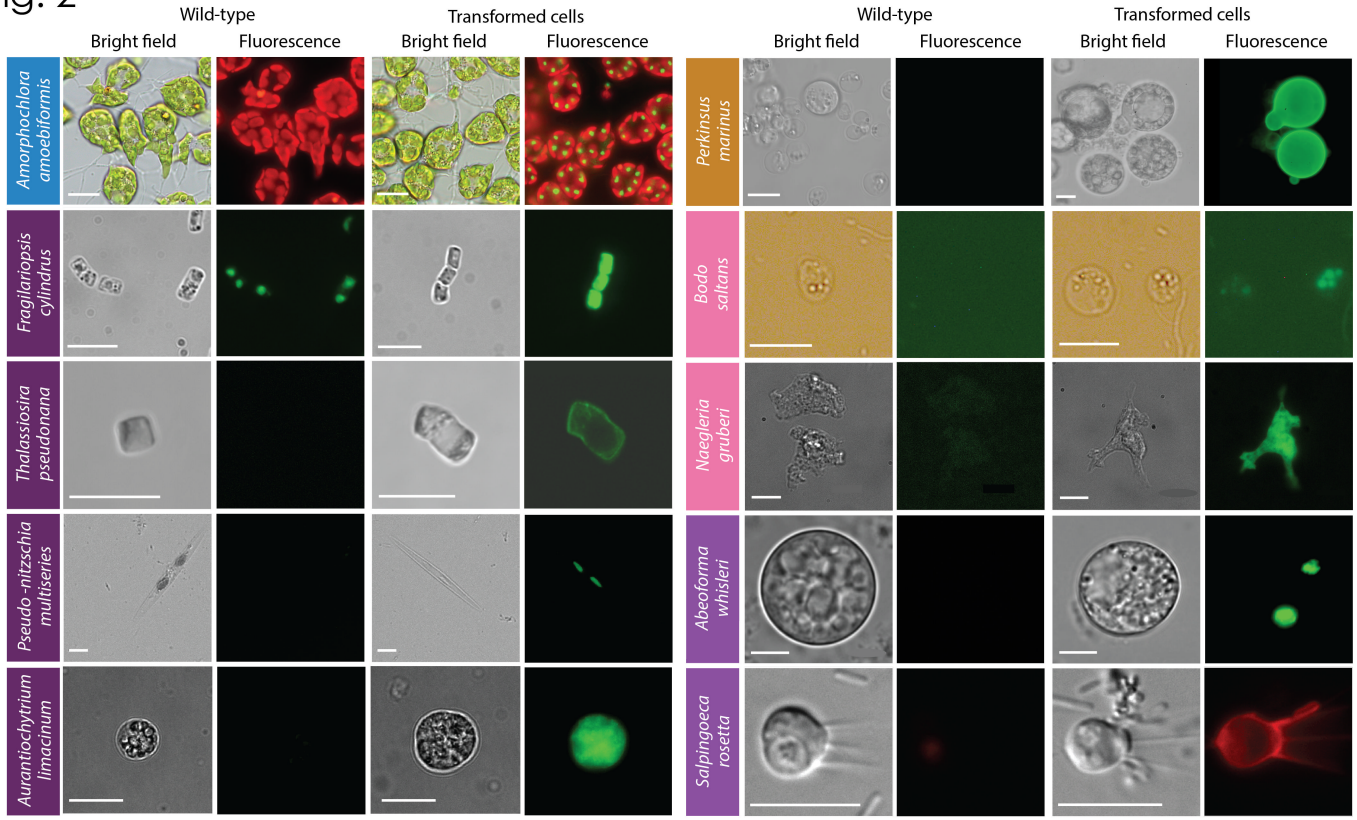

Wild-type
Transformed cells
Bright field
Fluorescence
Bright field
Fluorescence
Amorphochlora amoebiformis
Fragilariopsis cylindrus
Thalassiosira pseudonana
Pseudo-nitzschia multiseries
Aurantiochytrium limacinum
Wild-type
Transformed cells
Bright field
Fluorescence
Bright field
Fluorescence
Perkinsus marinus
Bodo saltans
Naegleria gruberi
Abeoforma whisleri
Salpingoeca rosetta

Fig. 3

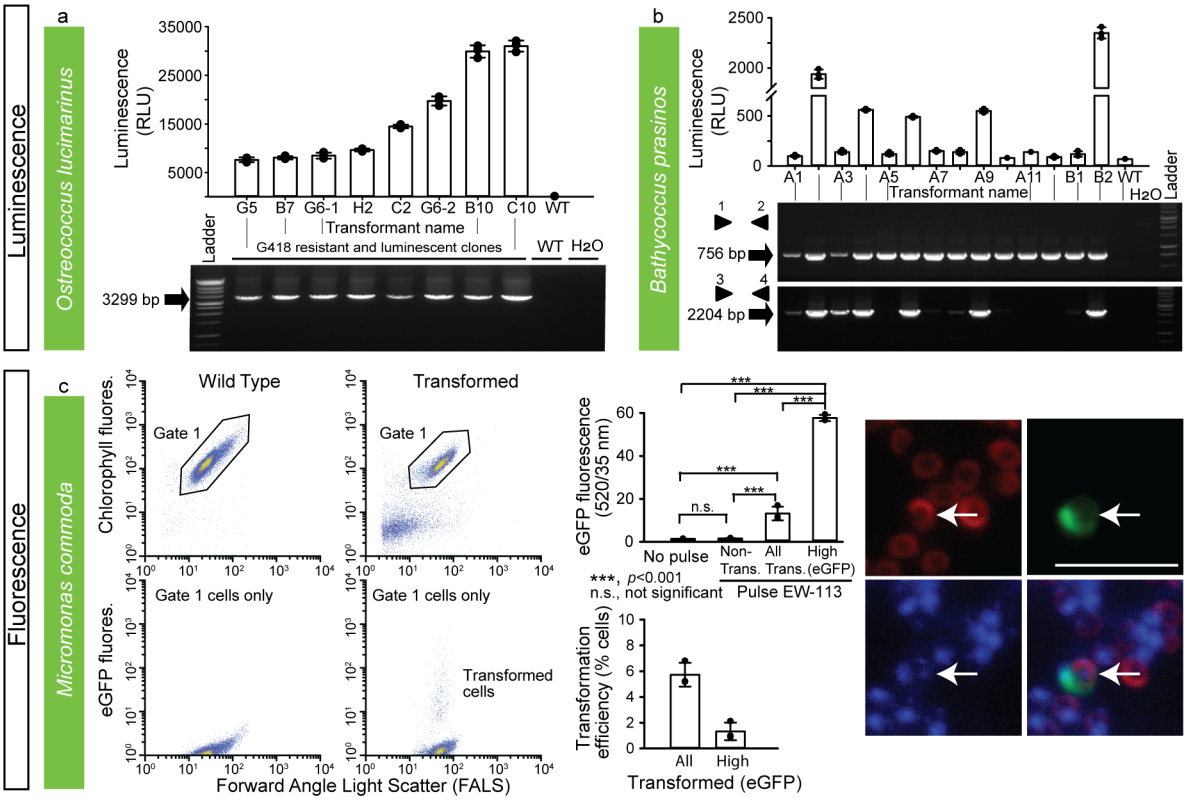

Luminescence
a
Ostreococcus lucimarinus
Luminescence (RLU)
35000
25000
15000
5000
G5 B7 G6-1 H2 C2 G6-2 B10 C10 WT
Transformant name
G418 resistant and luminescent clones
WT H2O
Ladder
3299 bp
b
Bathycoccus prasinos
Luminescence (RLU)
2500
2000
500
0
A1 A3 A5 A7 A9 A11 B1 B2 WT H2O
Transformant name
Ladder
1 2
756 bp
3 4
2204 bp
Fluorescence
c
Micromonas commoda
Wild Type
Transformed
Chlorophyll fluores.
Gate 1
Gate 1 cells only
eGFP fluores.
Transformed cells
Forward Angle Light Scatter (FALS)
eGFP fluorescence (520/35 nm)
***
n.s.
No pulse
Non-Trans.
All Trans.
High (eGFP)
Pulse EW-113
***, p<0.001
n.s., not significant
Transformation efficiency (% cells)
All
High
Transformed (eGFP)

Fig. 4

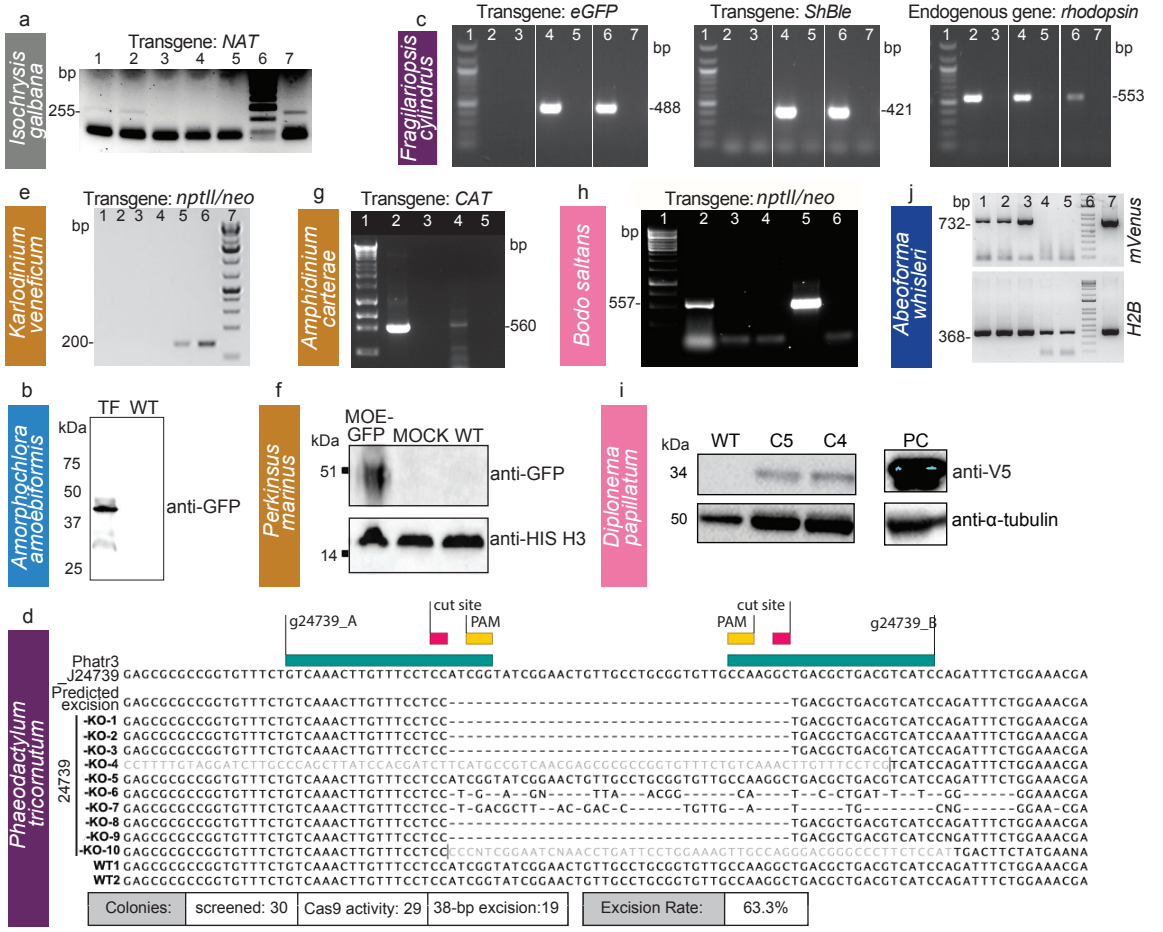
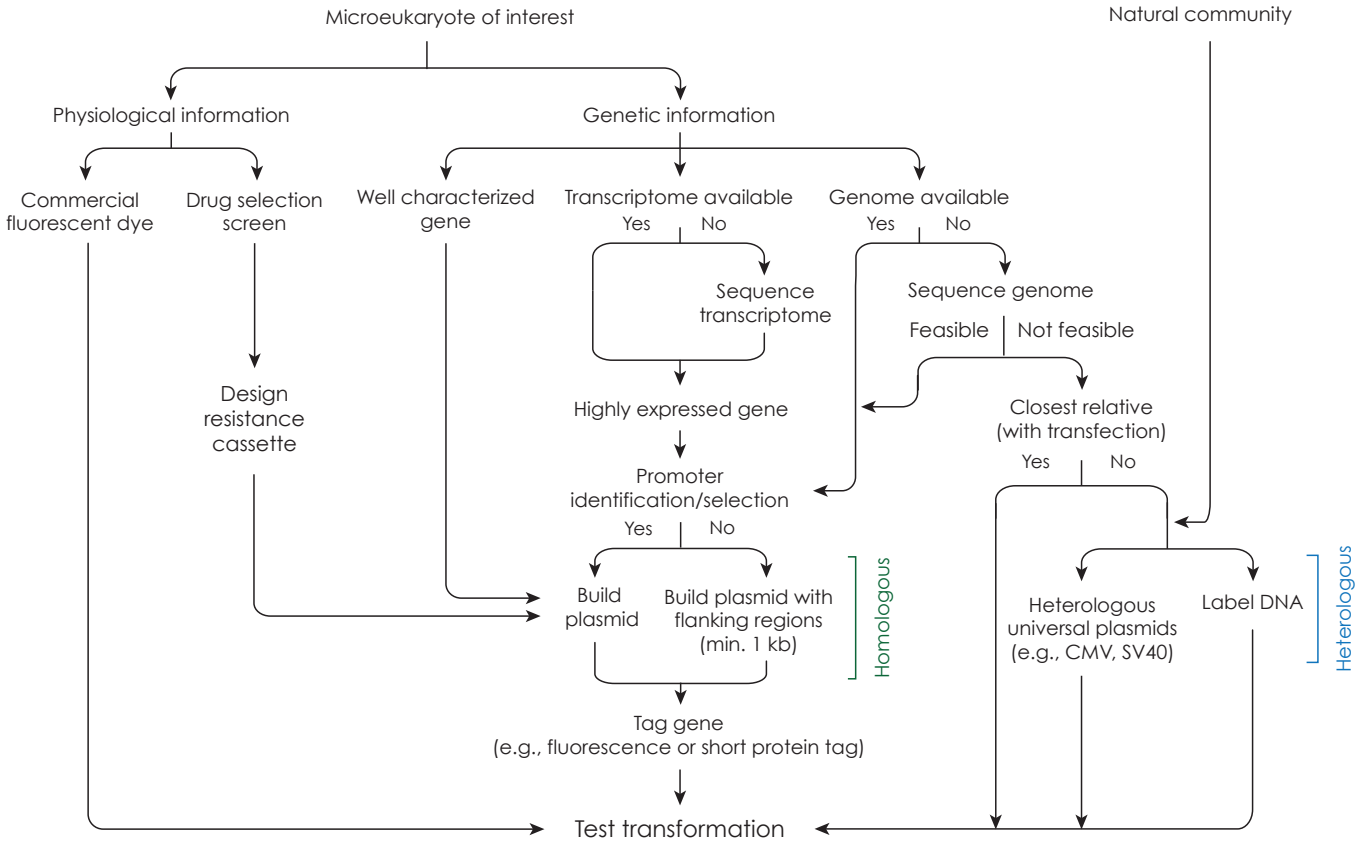

RT-PCR
a
Isochrysis galbana
Transgene: NAT
bp
255-
c
Fragilariopsis cylindrus
Transgene: eGFP
Transgene: ShBle
Endogenous gene: rhodopsin
-488
-421
-553
e
Karlodinium veneficum
Transgene: nptII/neo
200-
g
Amphidinium carterae
Transgene: CAT
-560
h
Bodo saltans
Transgene: nptII/neo
557-
j
Abeoforma whisleri
732-
368-
mVenus
H2B
Western Blots
b
Amorphochlora amoebiformis
TF WT
kDa
75
50
37
25
anti-GFP
f
Perkinsus marinus
MOE-GFP MOCK WT
51
14
anti-GFP
anti-HIS H3
i
Diplonema papillatum
WT C5 C4
PC
34
50
anti-V5
anti-α-tubulin
CRISPR/Cas9
d
Phaeodactylum tricornutum
g24739_A
cut site
PAM
g24739_B
Phatr3 J24739
Predicted excision
24739
-KO-1
-KO-2
-KO-3
-KO-4
-KO-5
-KO-6
-KO-7
-KO-8
-KO-9
-KO-10
WT1
WT2
Colonies: screened: 30
Cas9 activity: 29
38-bp excision:19
Excision Rate: 63.3%

Fig. 5

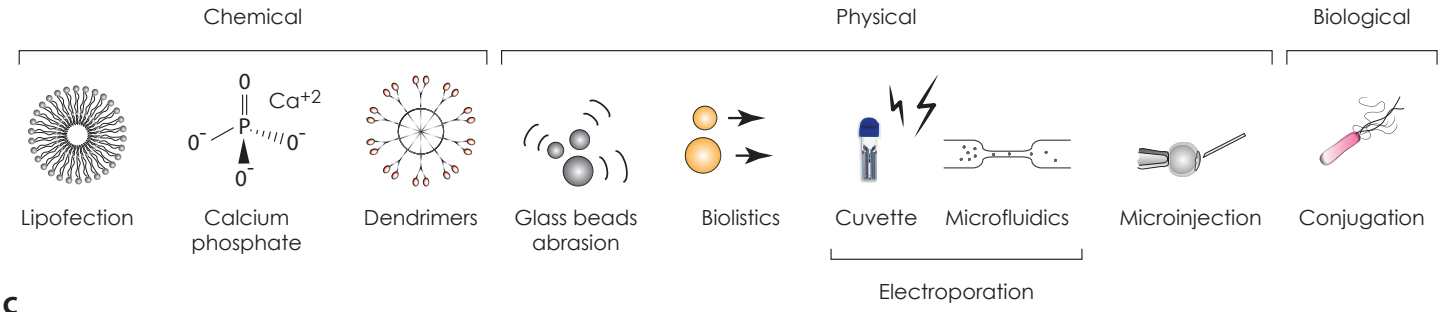
a

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 (IC_{50})

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, 3 and 4.
For additional information, see protocols.io links, Suppl. Table 5 and Supplementary Notes 1. For contacting laboratories working with particular species, see details given in Suppl. Table 6.

Species	Transformation Method/ Device	Cell number (input)	Vector;Amount (µg)	Promotor	Regulatory elements	Drug selection (µg/ml)	Time selection (Day)	Efficiency (%)	Transformation status (Stable/ Transient)	Reporter	Evidence of transformation	protocols.io link
Archaeoplastids												
Ostreococcus lucimarinus (RCC802)	Electroporation Genepulser II	1-2x10^{9}	Plasmid PotLuc; Linear; 5	HAPT, Histone H4 *O. tauri*	None	G418 (1000)	10-21	<0.0001	S	Luc	G418 resist, Luminescence, PCR	https://www.protocols.io/view/selection-of-stable-transformants-in-ostreococcus-zj2f4qe https://www.protocols.io/view/transient-luciferase-expression-in-ostreococcus-ot-hcib2ue https://www.protocols.io/view/transient-transformation-of-ostreococcus-species-o-g86bzze
Bathycoccus prasinos (RCC4222)	Electroporation Genepulser II	1-2x10^{9}	Fusion PCR; *p*HAPT: *p*LucpH4: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.zj2f4qe http://dx.doi.org/10.17504/protocols.io.hcib2ue
Micromonas commoda (CCMP 2709)	Electroporation Lonza-Nucleofector	3 x 10^{7}	RPS9pro^{Mco}-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^{Mpo}-LUC-H3ter in pUC05-AMP; Circular; 10-20	Histone H3 5' UTR from *M. polaris*	Histone H3 3' end formation -histone stem loop from *M. polaris*	n/a	3	n/a (Luc. assay is bulk, not per cell)	T*	NanoLuc®	Luminescence	http://dx.doi.org/10.17504/protocols.io.8p8hvrw
Tetraselmis striata (KAS-836)	Biorad Biolistics PDS-1000/He biolistics system	2.0x10^{7}	pACTpro:Ble; Linear; 1.0	Actin, *T. striata*	Actin, *T. striata*	Zeocin (150)	21-28		S		Zeocin resist, PCR	http://dx.doi.org/10.17504/protocols.io.hjtb4nn
Haptophytes												
Isochrysis galbana (CCMP 1323)	Biolistics PDS-1000/He	1-2x10^{6}	*p*IgNAT; Circular; 1.0	Hsp70 *E. huxleyi*	Heterologous	Nourseothricin (80-150)	14	<0.0001	S	None	Nourseothricin resistance, PCR, RT-PCR	https://www.protocols.io/view/biolistic-transformation-of-isochrysis-galbana-2pugdnw https://www.protocols.io/view/method-for-electroporation-of-isochrysis-galbana-c-hmab42e
Rhizarians												
Amorphochlora (Lotharella) amoebiformis (CCMP 2058)	Electroporation Gene Pulser Xcell	0.5-1×10^{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.35hgq36
Stramenopiles												
Fragilariopsis cylindrus (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
Thalassiosira pseudonana (CCMP 1335)	Bacterial conjugation	4x10^{7}	TpSli3p-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.nbzdap6 http://dx.doi.org/10.17504/protocols.io.7ghhjt6
Pseudo-nitzschia multiseries	Conjugation	1x10^{5}	Pm_actP_egfp_actT; pPtPUC3	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
Aurantiochytrium limacinum (ATCC MYA-1381)	Bio – Rad Gene Pulser (165-2076) NEPA21	1×10^{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
Nannochloropsis oceanica (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
Phaeodactylum tricornutum	Bacterial-conjugation	4x10^{7}	hCas9-2A-shble PtpBR episome	FcpF-hCas9 psRNA-sgRNA	Cen6-Arsh4-His3 centromere	Phleomycin (50)	10-16	1.25e^{-5} ~500 cfu	S	shble (Cas9) yfp VENUS	Phleoycin resistance, PCR maintained	http://dx.doi.org/10.17504/protocols.io.4bmgsk6 http://dx.doi.org/10.17504/protocols.io.7gihjue

(CCAP1055/1)			100uL E.coli OD_{600}=0.9			Zeocin (100)					episome, PCR Cas9 target site	http://dx.doi.org/10.17504/protocols.io.7gnhjve
Alveolates												
Perkinsus marinus (ATCC PRA240)	Electroporation LONZA-Nucleofector Glass beads abrasion (425-600 µm)	5-7x10^7	*p*PmMOE-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
Oxyrrhis marina (CCMP 1788, CCMP 1795)	Electroporation Gene Pulser Xcell; Chemical ($CaCl_2$)	1-5×10^6 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%	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-fitc-dextran-int-3cmgiu6 https://www.protocols.io/view/co-incubation-protocol-for-transforming-heterotrop-hmzb476
Crypthecodinium cohnii (CCMP 316)	Electroporation LONZA-Nucleofector		Stained DNA (739 bp); Linear; 1	None	None	n/a	n/a	<0.001	T	Fluorescence	Fluorescence	https://www.protocols.io/view/transfection-of-crypthecodinium-cohnii-using-label-z26f8he
Amphidinium carterae (chloroplast) (CCMP 1314)	Biorad Biolistics PDS-1000/He biolistics system	2.5x10^7	*p*AmpAtpBChl; 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
Karlodinium veneficum (CCMP 1975)	Electroporation	4x10^5	linear-DinoIII-*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)												
Bodo saltans (submitted to ATCC)	Electroporation Nepa21	1-1.5x10^7	Bs-EF1-α C- terminal tagging; Linear; 3-5	Endogenous	Endogenous	G418 (5)	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.7fchjiw
Diplonema papillatum (ATCC 50162)	Electroporation LONZA-Nucleofector	5x10^7	*p*57-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
Naegleria gruberi (ATCC 30224)	Electroporation BioRad Gene Pulser xCell	5x10^6	*p*NAEG-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												
Abeoforma whisleri (ATCC PRA-279)	Electroporation LONZA-Nucleofector	3x10^5	Awhis_H2Bvenus+ *p*UC19; 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.zexf3fn
Salpingoeca rosetta (ATCC PRA-390)	Electroporation LONZA-Nucleofector	4x10^5	SrActmCherry-CCTLL + *p*UC19; 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