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[1,24,30], 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 Swart[17,&], Anastasios 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,*]

[#]Equal contribution

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

[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, 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, Banyul sur Mer, France*

[14]*Institute de Biologie de l'ENS, Département de biologie, École Normale Supérieure, CNRS, INSERM, 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, MA, 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 de Rosario, 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, Jeddah, 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, Catalonia, Spain*

[45]*Catalan Institution for Research and Advanced Studies, Barcelona, Catalonia, 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: *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: *Department of Microbiome Science, Max Planck Institute for Developmental Biology, Tübingen, Germany*

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

ABSTRACT

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

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

INTRODUCTION

The ocean represents the largest continuous planetary ecosystem, hosting an enormous variety of organisms. These range from some of the largest creatures on Earth to a vast microscopic biota including unicellular eukaryotes (protists). Despite their small size, protists play key roles in marine biogeochemical cycles and harbor tremendous evolutionary diversity[1-3]. Notwithstanding their significance 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[4]. Most of the biological information available is based on comparison of proteins from cultured genome-sequenced species to homologs in genetically tractable model taxa, such as yeast[5-9]. 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 that represent neither the most ecologically relevant protists nor the breadth of eukaryotic diversity.

The development of genetic tools requires reliable information about gene organization and regulation of the emergent model species. Over the last decade, some of this information has become available through genome[5,6,8,10] and transcriptome sequencing initiatives[7,9,11,12] resulting in nearly 120 million unigenes from protists[13]. Insights from these studies have enabled the phylogenetically-informed approach[7] used for selecting and developing key marine protists into model systems in the Environmental Model Systems (EMS) Project presented herein. Forty-one scientific 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, which can be manipulated, and which encompass multiple ecologically important marine protists.

Here, we summarize detailed methodological achievements by this collaborative effort 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 14 emerging model species presented herein, and the collective set of genetic tools from the overall collaborative project, will not only extend our knowledge of cell biology, evolution and functional biodiversity, but also serve as platforms to advance microbial biotechnology.

RESULTS

Overview of taxa in the EMS Initiative

Taxa were selected from multiple eukaryotic supergroups[1,7] to maximize the potential to compare fundamental aspects 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 major 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 some protists have been developed in the past[17,18], the challenge for this initiative was to develop novel genetic tools for species which not only require different cultivation conditions but are also phenotypically extremely diverse. It should be noted that not all major lineages could be pursued. For example, amoebozoans did not feature in the 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, in particular *Dictyostelium discoideum*[14]. Below we summarize the taxa interrogated and overall EMS initiative outcomes (**Fig. 1**), provide evidence and detailed protocols for 14 taxa, for which no transformation systems have been previously reported (Category A), and 7 taxa for which existing protocols[15,17,18,25,40,44,45,51] were advanced (Category B) (**Figs. 2 and 3; Table 1**). We also review an already published transformation protocol[56] in 1 species (Category C), and we discuss the unsuccessful

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

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[1,8] (**Fig. 1**). Genome sequences are available for the picoprasinophytes (<3 µm cell diameter) tested herein, specifically, *Micromonas commoda*, *Micromonas pusilla*, *Ostreococcus lucimarinus* and *Bathycoccus prasinos*. As part of the EMS initiative, we report on the first genetic systems for *Bathycoccus*, a scaled, non-motile genus, and *Micromonas*, a motile, naked genus with larger genomes than *Bathycoccus* and *Ostreococcus*[8]. We also report on the first genetic systems for *Tetraselmis striata* and *Ostreococcus lucimarinus*. The latter was transformed based on an adapted homologous recombination system for *Ostreococcus tauri*[19].

O. lucimarinus (RCC802) and *B. prasinos* (RCC4222) were transformed using protocols adapted from *O. tauri*[20]. 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 KanMX 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 the site of integration and/or the number of integrated genes (**Fig. 3a; Suppl. Fig. 1; Suppl. sheet 1**).

The *O. tauri* construct did not work in *B. prasinos*, while the use of the *B. prasinos* histone H4 and high affinity phosphate transporter 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 clones that were luminescent, and one non-luminescent clone (**Fig. 3b; Suppl. sheet 1)**, 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 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 *M. polaris*[8], and the other encoding an eGFP gene flanked by endogenous promoter and terminator sequences from the ribosomal protein S9. Acclimated mid-exponential *M. commoda* cells grown in L1 medium at 21 ^{o}C were spun at 5000 x *g* for 10 min, the pellet was resuspended in Buffer SF (Lonza) premixed with carrier DNA (pUC19) and the plasmid, and 3 x 10^7 cells were used per reaction. After applying the EW-113 pulse, 100 µl of ice-cold recovery buffer (10 mM HEPES-KOH, pH 7.5; 530 mM sorbitol; 4.7% [w/v] PEG 8000) was added to each well and incubated for 5 min at room temperature. Each reaction was then transferred into 2 ml L1, placed at 21 °C and light was increased stepwise over 72 h. 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; Suppl. sheet 1**). Additionally, the RLU was 1500-times 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) has been successfully transformed using microprojectile bombardment (**Suppl. Fig. 2**). 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. 2)**. The terminator sequence was obtained from the *T. striata g*lyceraldehyde-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 (seawater-distilled water in a 1:1 volume ratio) 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*[21], and there is one report of nuclear transformation of a calcifying coccolithophore species[22]. 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[23]. However, transformation of *Emiliania huxleyi*, which is the most prominent and ecologically significant coccolithophore, has not been achieved yet[24].

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. 3d, Suppl. Fig. 3)** and the sequence 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. Uniquely, they represent an intermediate stage of the endosymbiotic process, since their plastids still harbor a relict nucleus (nucleomorph)[5]. 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[25].

A. amoebiformis (CCMP2058) cells (1 x 10^7) were resuspended in 100 µl of Gene Pulse Electroporation Buffer (BioRad) with 20 to 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 1,000-fold higher than the previous study[25] (**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. 3e**).

Stramenopiles. Stramenopiles are a diverse lineage with many important photoautotrophic, mixotrophic and heterotrophic taxa. As the most studied class in this lineage, diatoms (Bacillariophyceae) were early targets[26,18] for the development of reverse genetics tool[24]. Diatoms are estimated to contribute approximately 20% of annual carbon fixation[27] and, like several other algal lineages, are used in bioengineering applications and biofuels[28]. The transformation protocol for the Antarctic diatom *Fragilariopsis cylindrus*[29] represents the first transformation system for a cold-adapted eukaryote. A first transformation protocol has also been developed for *Pseudo-nitzschia multiseries*, a toxin-producing diatom[30]. 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*[31], both of which are used for biotechnological applications. Furthermore, we report on advances for CRISPR/Cas-driven gene knockouts in *P. tricornutum*[15,16] and a more efficient bacterial conjugation system for *T. pseudonana*[15].

Microparticle bombardment was used on *F. cylindrus* (CCMP1102) which was grown, processed and maintained at 4 °C in 24 hours light. Exponential phase cells (5 x 10^7 cells) were harvested onto a 1.2 µm membrane filter (Millipore) 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**)[32]. Transformation was performed using 0.7 µm tungsten particles and the biolistic particle delivery system PDS1000/He (BioRad). Rupture discs for 1,350 and 1,550 pounds per square inch (psi) gave the highest colony numbers (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 (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 zeocin resistance and eGFP genes, respectively. Confirmed by 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. 3f**) revealed that the *ShBle* gene for zeocin resistance and the gene encoding eGFP were present in the genomes of transformants after multiple transfers (> 10) two years later, indicating long-term stability.

Bacterial conjugation methods were improved in *T. pseudonana* (CCMP1335) using the silaffin precursor TpSil3p (**Table 1**) 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 donor cells were mixed with 4 x 10^7 *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 pour 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. 4**) 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[33]. To complement this approach and explore potentially higher efficiency methods for transformation with diatom episomal plasmids, we modified the existing conjugation-based method[15]. 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**). 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. 5**). 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 the bleomycin-resistance gene (*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 yeGFP:shble fusion (**Suppl. Fig. 6**). Approximately 1 x 10^8 cells were electroporated (**Suppl. Table 4**), adapting the electroporation protocol used for *Schizochytrium*[34]. 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**). 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. 6**).

Electroporation of *N. oceanica* (CCMP1779) was optimized based on observation of cells treated with fluorescein-conjugated 2000 kDa dextran and subsequent survival (**Table 1**). 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 control of the cytomegalovirus (CMV), the cauliflower Mosaic Virus 35S, or the VCP1 promoter previously described from *Nannochloropsis* sp.[35]. Transient expression of blue fluorescence (compared to cells electroporated simultaneously under the same conditions but 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. 7).** 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[16] for multiplexed targeted mutagenesis. Bacterial conjugation[15] 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 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**). 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. 8**). 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 14 days. Only colonies maintaining Cas9-2A-ShBle sequence on the delivered episome were able to grow on selection plates because Cas9 and the antibiotic-resistant gene (ShBle) were transcriptionally fused by the 2A peptide[36] (**Suppl. Fig. 8)**. Gel electrophoresis migration and sequencing of the genomic target loci confirmed the 38-bp excision and premature stop codon (**Fig. 3g**).

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. Techniques currently exist for transformation of only a small number of ciliates, perkinsids and apicomplexans[37,38]. Here, we present a first transformation protocol for *Karlodinium veneficum* (CCMP1975), a mixotroph (combining photosynthetic and phagotrophic nutrition) that produces fish-killing karlotoxins[39], and first evidence for

DNA delivery without gene expression in *Oxyrrhis marina* (CCMP 1788/CCMP 1795) and *Crypthecodinium cohnii* (CCMP 316), see **Suppl. Results, Table 1 and Supp. Fig. 17**) . Additionally, we report on improved transformation systems for *Perkinsus marinus* (PRA240) and *Amphidinium carterae* (CCMP1314) chloroplast, published recently as part of EMS initiative[40].

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 (*aph(3')*) 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 Nucleofector (Lonza). The preprogrammed Nucleofector optimization pulse codes, buffer SF/Solution I (Lonza), and 2 μg/μl of linearized DinoIII-neo were used. 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 three months, DNA and RNA were isolated from the resistant cultures as previously reported[42,43] and cDNA was synthesized using random hexamers. Out of 16 transformations, two 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 through nested PCRs and RT-PCRs, respectively (**Suppl. Fig. 9, Fig. 3h**).

We improved the transformation protocol[44,45] of *Perkinsus marinus*, a pathogen of marine mollusks, fish, and amphibians[46], based on adapting the *A. carterae* transformation system[47]. In brief, a newly formulated transformation 3R buffer (200 mM Na_2HPO_4; 70 mM NaH_2PO_4; 15 mM KCl; 1.5 mM $CaCl_2$; 150 mM HEPES-KOH, pH 7.3) was used to reduce

the cost of electroporation. We co-expressed 2 genes and efficiently selected transient and stable transformants using FACS (**Figs. 2 and 3i, Suppl. Fig. 10, Table 1**). 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. In brief, 5 x 10^7 cells were resuspended in 330 µl of fresh ATCC Medium 1886 and were mixed with 5.0 µg of linearized and circular [1:1] plasmid and 300 µl of glass beads (Sigma) in a 1.5 ml tube, vortexed for 30 s at maximum speed, and cells in 500 µl of culture medium were transferred to 6-well plates in a final volume of 3 ml. Two versions of the previously published Moe gene promoter were tested[44]. 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 bleomycin, blasticidin and puromycin have all been shown to confer resistence to transformed *P. marinus*; however, selection regimes are still relatively slow and inefficient, indicating further room for improvement[45].

We also report a new vector for the transformation of the *A. carterae* chloroplast, a core, photosynthetic dinoflagellate. *A. carterae,* like other dinoflagellates with a peridinin-containing chloroplast, contains a fragmented chloroplast genome made up of multiple plasmid-like minicircles. The previous transformation protocols made use of this, to introduce two vectors based on the *psbA* minicircle[40]. 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. This circular vector was introduced by biolistics to *A. carterae* (**Suppl. Fig. 11)**. Following selection with chloramphenicol, we were able to detect transcription of the chloramphenicol acetyl transferase gene via RT-PCR (**Fig. 3j)**. This result

suggests that all of the 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.

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

Discobans. This diverse group, recently split into Discoba and Metamonada[48], 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[49]. However, such protocols are missing for aquatic species. Here, we describe the first available transformation protocols for the kinetoplastid *Bodo saltans* and the heterolobosean *Naegleria gruberi*. The former was isolated from a lake, but identical 18S rRNA sequence have been reported from the marine environment[50]. The latter is a fresh water protist that represents a model organism for closely related marine heterolobosean amoebas. Furthermore, we provide advanced methods that build on published EMS results[51] for the diplonemid *Diplonema papillatum*.

B. saltans (ATCC 30904) was transformed with a plasmid containing a cassette designed to fuse an endogenous EF-1α gene with eGFP for C-terminal tagging. This cassette includes downstream of the eGFP, a *B. saltans* tubulin intergenic region followed by the selectable marker gene *nptII / neo*, 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[52], 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. Square-wave electroporation (Nepa21) was used with a poring pulse of

250V (25 ms) and 5 transfer pulses of 60V (99 ms) in the presence of Cytomix buffer (120 mM KCl; 0.15 mM $CaCl_2$; 10 mM KH_2PO_4; 2 mM EGTA; 5 mM $MgCl_2$; 25 mM HEPES-KOH, pH 7.6). 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**). Cells were washed and subcultured into fresh selection medium every 4 days, and neomycin-resistant cells emerged 7 to 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. 3k; Suppl. Fig. 12**). However, plasmid integration into the *B. saltans* genome through homologous recombination is still unconfirmed. This suggests either off-target plasmid integration or that the plasmid is maintained episomally.

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**). 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, several resistant clones were recovered and expression of eGFP and/or hygromycin was confirmed by Western blotting (**Suppl. Fig. 13**). Expression of eGFP was observed by epifluorescence microscopy (**Fig. 2; Suppl. Fig. 13**) 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 of *SwaI*-linearised fragment (cut from p57-V5+NeoR plasmid) containing V5-tagged *nptII / neo* gene flanked by partial regulatory sequences derived from the hexokinase gene of the kinetoplastid *Blastocrithidia* (strain p57) (**Table 1**) using a published protocol[51]. 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. 14)** and the expression of the tagged *nptII / neo* protein was confirmed in 2 clones by Western blotting using α-V5 antibody (**Fig. 4l**).

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[9]. Genomic and transcriptomic data are available for multiple representatives characterized by diverse cell morphologies, some of which can even form multicellular structures[9,12]. Untill recently, only transient transformations have been achieved for some opistokonts such as the filasterean *Capsaspora owczarzaki*[53], the ichthyosporean *Creolimax fragrantissima*[54] and the choanoflagellate *Salpingoeca rosetta*[55]. 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 on recently published stable transformation protocol for *S. rosetta* achieved by using the selectable marker gene (puromycin N-acetyl-transferase - *PAC*) (**Fig. 2**)[56].

All *A. whisleri* life stages are highly sensitive to a variety of methods for transformation. However, we developed a Lonza 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**), 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 3m**).

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

Discussion

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

picture of the oceanic ecosystem. These studies highlight the diversity of ocean life, including the protistan component[2,3]. Remarkable strides have also been made in developing an overview of the genomes and predicted proteomes of these protists[13]. However, without genetic manipulation systems, these taxa have remained an untapped resource for providing deeper insights into their cell biology, with potentially valuable outcomes for evolutionary studies, nanotechnology, biotechnology, medicine, and pharmacology.

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

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 methods. Unlike established microbial model species, cultivation of marine protists can be challenging especially under axenic conditions and for predatory taxa that require co-cultivation with their prey. Nevertheless, 13 out of 35 species have been made 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 the 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; Suppl. Table 5; Suppl. Doc. 1**).

As exemplified in the new systems provided herein (**Figs. 2 and 3; Table 1**), 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, fluorescence-activated cell sorting (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 (e.g. **Fig. 2**, **Fig. 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 below according to their phylogenetic position (**Fig. 1**).

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

the cell. 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. The collaborative network facilitated many of the studies undertaken in this initiative. The names and contact details of all co-authors assigned to particular species are listed in **Suppl. Table 6**. Lessons from our integrated approach for instance include examples such as the adoption of the highly sensitive reporter *nanoluc* (*Nluc*) gene by several teams for initial rapid screening to find optimal conditions for transformation. If successful, teams usually continued with fluorescence-based or other methods, as well as a hypertonic recovery buffer. Furthermore, large amounts of carrier DNA usually facilitated successful initial transformations (e.g. *M. commoda*, *A. whisleri*) or improved existing protocols (*S. rosetta*[55]).

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

Symbiodinium microadriaticum were unsuccessful[57]. These and other lineages such as coccolithophores, warrant concerted future efforts.

The combination of new results presented herein and others arising from the EMS initiative[58] significantly expand 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. This high rate of success testifies to the benefits of taking on high-risk research in a large, open, and collaborative context[58]. Moreover, the new systems reported herein open a wealth of opportunities for exploring functional differences between members of relatively conserved protein families shared across eukaryotes, or even domains of life. Collectively, these novel and recently described protistan 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 for assistance, C. Poirier and M. Hamilton (Monterey Bay Aquarium Research Institute/GEOMAR) for FACS analysis; and V. K. Nagarajan, M. Accerbi, and P. J. Green (University of Delaware) who carried out *Agrobacterium* studies in *Heterosigma akashiwo*; and N. Kraeva, C. Bianchi and V. Yurchenko (University of Ostrava) for the help with designing the p57-V5+NeoR construct. We are also grateful to the protocols.io team 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 and other forms of grant support to the participating laboratories.

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.M., A.E.A., F.Y.B, C.B, Ch.B., H.C., T.C., J.L.C., K.C., 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., 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., 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., 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., 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., 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., 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., 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.M., A.Z.W. and J.L. with input from all authors.

COMPETING INTERESTS

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., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J. M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-

Lewis, S., Tara Oceans Coordinators., Acinas, S. G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M. E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P. & Karsenti, E. Eukaryotic plankton diversity in the sunlit global ocean. *Science* **348**, 1261605 (2015).

3. Duarte, C. M., Regaudie-de-Gioux, A., Arrieta, J. M., Delgado-Huertas, A. & Agustí, S. The oligotrophic ocean is heterotrophic. *Annu. Rev. Mar. Sci.* **5**, 551-569 (2013).

4. 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).

5. Curtis, B. A., Tanifuji, G. & Burki, F. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**, 59–65 (2012).

6. Read, B. A., Kegel, J., Klute, M. J., Kuo, A., Lefebvre, S. C., Maumus, F., Mayer, C., Miller, J., Monier, A., Salamov, A., Young, J., Aguilar, M., Claverie, J. M., Frickenhaus, S., Gonzalez, K., Herman, E. K., Lin, Y. C., Napier, J., Ogata, H., Sarno, A. F., Shmutz, J., Schroeder, D., de Vargas, C., Verret, F., von Dassow, P., Valentin, K., Van de Peer, Y., Wheeler, G., *Emiliania huxleyi* Annotation Consortium, Dacks, J. B., Delwiche, C. F., Dyhrman, S. T., Glöckner, G., John, U., Richards, T., Worden, A. Z, Zhang, X. & Grigoriev, IV. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature* **499**, 209-213 (2013).

7. Keeling P. J., Burki, F., Wilcox, H. M., Allam, B. & Allen, E. E. 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. van Baren, M. J., Bachy, C., Reistetter, E. N., Purvine, S. O., Grimwood, J., Sudek, S., Yu, H., Poirier, C., Deerinck, T. J., Kuo, A., Grigoriev, IV., Wong, C. H., Smith, R. D., Callister, S. J., Wei, C. L., Schmutz, J. & Worden, A. Z. Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genomics* **17**, 1–22 (2016).

9. 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).

10. Strassert, J. F. H., Karnkowska, A., Hehenberger, E., Del Campo, J., Kolisko, M., Okamoto, N., Burki, F., Janouškovec, J., Poirier, C., Leonard, G., Hallam, S. J., Richards, T. A., Worden, A. Z., Santoro, A. E. & Keeling, P. J. Single cell genomics of uncultured marine alveolates shows paraphyly of basal dinoflagellates. *ISME J.* **12**, 304-308 (2018).

11. Joseph, S. J., Fernández-Robledo, J. A., Gardner, M. J., El-Sayed, N. M., Kuo, C. H., Schott, E. J.,Wang, H., Kissinger, J. C. & Vasta, G. R. The Alveolate *Perkinsus marinus*: Biological insights from EST gene discovery. *BMC Genomics* **11**, 228 (2010).

12. Grau-Bové, X., Torruella, G., Donachie, S., Suga, H., Leonard, G., Richards, T. A. & Ruiz-Trillo, I. Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife* **6**, e26036 (2017).

13. Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A. & Seeleuthner, Y. A global atlas of eukaryotic genes. *Nat. Commun.* **9**, 373 (2018).

14. Paschke, P., Knecht, D. A., Williams, T. D., Thomason, P. A., Insall, R. H., Chubb, J. R., Kay, R. R. & Veltman, D. M. Genetic engineering of *Dictyostelium discoideum* cells based on selection and growth on bacteria. *J. Vis. Exp.* **143**, e58981 (2019).

15. Karas, B. J., Diner, R. E., Lefebvre, S. C., McQuaid, J., Phillips, A. P., Noddings, C. M., Brunson, J. K., Valas, R. E., Deerinck, T. J., Jablanovic, J., Gillard, J. T., Beeri, K., Ellisman, M. H., Glass, J. I., Hutchison, C. A. 3rd., Smith, H. O., Venter, J. C., Allen, A. E., Dupont, C. L. & Weyman, P. D. Designer diatom episomes delivered by bacterial conjugation. *Nat. Commun.* **6**, 6925 (2015).

16. 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).

17. 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).

18. 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).

19. Lozano, J. C., Schatt, P., Botebol, H., Vergé, V., Lesuisse, E., Blain, S., Carré, I. A. & Bouget, F. Y. Efficient gene targeting and removal of foreign DNA by homologous recombination in the picoeukaryote *Ostreococcus*. *Plant J*. **78**, 1073-1083 (2014).

20. 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).

21. Hovde, B. T., Deodato, C. R., Hunsperger, H. M., Ryken, S. A., Yost, W., Jha, R. K., Patterson, J., Monnat, R. J. Jr., Barlow, S. B., Starkenburg, S. R. & Cattolico, R. A. 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).

22. 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).

23. 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).

24. Velmurugan, N. & Deka, D. Transformation techniques for metabolic engineering of diatoms and haptophytes: current state and prospects. *Appl Microbiol Biotechnol.* **102**, 4255-4267 (2018).

25. 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).

26. 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).

27. 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).

28. 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).

29. Mock, T., Otillar, R. P., Strauss, J., McMullan, M., Paajanen, P., Schmutz, J., Salamov, A., Sanges, R., Toseland, A., Ward, B. J., Allen, A. E., Dupont, C. L., Frickenhaus, S., Maumus, F., Veluchamy, A., Wu, T., Barry, K. W., Falciatore, A., Ferrante, M. I., Fortunato, A. E., Glöckner, G., Gruber, A., Hipkin, R., Janech, M. G., Kroth, P. G., Leese,

F., Lindquist, E. A., Lyon, B. R., Martin, J., Mayer, C., Parker, M., Quesneville, H., Raymond, J. A., Uhlig, C., Valas, R. E., Valentin, K. U., Worden, A. Z., Armbrust, E. V., Clark, M. D., Bowler, C., Green, B. R., Moulton, V., van Oosterhout, C. & Grigoriev, IV. Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**, 536–540 (2017).

30. Brunson, J. K., McKinnie, S. M. K., Chekan, J. R., McCrow, J. P., Miles, Z. D., Bertrand, E. M., Bielinski, V. A., Luhavaya, H., Oborník, M., Smith, G. J., Hutchins D. A., Allen, A. E. & Moore, B. S. Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom. *Science* **361**, 1356-1358 (2018).

31. Sakaguchi K., Matsuda T., Kobayashi T., Ohara J., Hamaguchi R., Abe E., Nagano N., Hayashi M., Ueda M., Honda D., Okita Y., Taoka Y., Sugimoto S., Okino N. & Ito M. Versatile transformation system that is applicable to both multiple transgene expression and gene targeting for thraustochytrids. *Applied and Environmental Microbiology* **78**: 3193-3202 (2012).

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

33. 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).

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

35. 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).

36. Duda, K., Lonowski, L. A., Kofoed-Nielsen, M., Ibarra, A., Delay, C. M., Kang, Q., Yang, Z., Pruett-Miller, S. M., Bennett, E. P., Wandall, H. H., Davis, G. D., Hansen, S. H. & Frödin, M. 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).

37. 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).

38. Rug, M. & Maier, A. G. Transfection of *Plasmodium falciparum. Methods Mol. Biol.* **923**, 75-98 (2013).

39. Barbrook, A. C., Howe, C. & Nisbet, R. Breaking up is hard to do: the complexity of the dinoflagellate chloroplast genome. *Perspect. Phycol*. https://doi.org/10.1127/pip/2018/0084 (2018).

40. Nimmo, I. C., Barbrook, A. C., Chen, J. E., Geisler, K., Smith, A. G., Aranda, A., Purton, P., Waller, R. F., Nisbet, R. E. R. & Howe, C. J. Genetic transformation of the dinoflagellate chloroplast. *eLife* **8**, e45292 (2019).

41. Sprecher, B. N., Zhang, H. & Lin, S. Nuclear gene transformation in a dinoflagellate. *bioRxiv* 602821 (2019).

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. Lin, S., Zhang, H. & Dubois, A. Low abundance distribution of Pfiesteria piscicida in Pacific and Western Atlantic as detected by mtDNA-18S rDNA real-time polymerase chain reaction. *J. Plankton Res*. **28**, 667-681 (2006).

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

45. 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).

46. Chambouvet, A., Gower, D. J., Jirků, M., Yabsley, M. J., Davis, A. K., Leonard, G., Maguire, F., Doherty-Bone, T. M., Bittencourt-Silva, G. B., Wilkinson, M. & Richards, T. A. Cryptic infection of a broad taxonomic and geographic diversity of tadpoles by Perkinsea protists. *Proc. Natl. Acad. Sci. USA* **112**, E4743-4751 (2015).

47. Ten Lohuis, M. R. & Miller, D. J. Genetic transformation of dinoflagellates (*Amphidinium* and *Symbiodinium*): expression of GUS in microalgae using heterologous promoter constructs. *Plant J.* **13**, 427-435, (1998).

48. Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., Agatha, S., Berney, C., Brown, M. W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., Del Campo, J., Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A. A., Hoppenrath, M., James, T. Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A., Lahr, D. J. G., Lara, E., LeGall, L., Lynn, D. H., Mann, D. G., Massana, R., Mitchell, E.

A. D., Morrow, C., Park, J. S., Pawlowski, J. W., Powell, M. J., Richter, D. J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F. W., Torruella, G., Youssef, N., Zlatogursky, V. & Zhang, Q. Revision to the classification, nomenclature and diversity of eukaryotes. *J. Euk. Microbiol*. **66**, 4-119 (2019).

49. 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).

50. Jackson, A. P., Quail, M. A. & Berriman, M. Insights into the genome sequence of a free-living Kinetoplastid: *Bodo saltans* (Kinetoplastida: Euglenozoa). *BMC Genomics* **9**, 594 (2008).

51. Kaur, B., Valach, M., Peña-Diaz, P., Moreira, S., Keeling, P. J., Burger, G., Lukeš, J. & Faktorová, D. Transformation of *Diplonema papillatum*, the type species of the highly diverse and abundant marine micro-eukaryotes Diplonemida (Euglenozoa). *Env. Microbiol*. **20**, 1030-1040 (2018).

52. 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).

53. Parra-Acero, H., Ros-Rocher, N., Perez-Posada, A., Kożyczkowska, A., Sánchez-Pons, N., Nakata, A., Suga, H., Najle, S. R. & Ruiz-Trillo, I. Transfection of *Capsaspora owczarzaki*, a close unicellular relative of animals. *Development* **145**, 162107 (2018).

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

55. 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).

56. Wetzel, L. A., Levin, T. C., Hulett, R. E., Chan, D., King, G. A., Aldayafleh, R., Booth, D. S., Sigg, M. A. & King, N. Predicted glycosyltransferases promote development and prevent spurious cell clumping in the choanoflagellate *S. rosetta*. *eLife* **7,** e41482 (2018).

57. Chen, J. E., Barbrook, A. C., Cui, G., Howe, C. J. & Aranda, M. The genetic intractability of *Symbiodinium microadriaticum* to standard algal transformation methods. *PLoS ONE* **14,** e0211936 (2019).

58. Waller, R. F., Cleves, P. A., Rubio-Brotons, M., Woods, A., Bender, S. J., Edgcomb, V., Gann, E. R., Jones, A. C., Teytelman, L., von Dassow, P., Wilhelm, S. W. & Collier, J. L. 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 phylogenetically-informed approach was used to select protists for concerted genetic manipulation efforts. 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 Supplementary Data. Delivery methods and construct types are shown pictorially. Overall, protocols and specific reagents are

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

Fig. 2. | Epifluorescence micrographs of transformed marine protists

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. For each of these fluorescence micrographs details of transformation protocol as well as a second form of evidence are described in the Results and summarized in Table 1.

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 species.

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

whole *pH4:KanMX pHAPT:luc* transgene. A band of the expected size was amplified in all transformants and absent in WT (control cells electroporated without constructs) and the PCR negative control (H_2O). (**b**) *In vivo* luminescence of 14 *B. prasinos* transformants resistant to G418 depicted as RLU per 5 s and a corresponding gel showing PCR amplification from DNA of transformants of *pH4:KanMx* (primers 1 and 2) and *pHAPT:luc* sequences (primers 3 and 4). (**c**) FACS analysis of *M. commoda* cells in two treatments: controls (wild type [WT] or "no-pulse", in which *eGFP* constructs were added but no electroporation pulse applied), and in the treatment to which the pulse was applied ("Transformed"). Note that lower panels include only the population of healthy cells selected using the depicted gates in upper panels. No compensation was used and PMT voltages were identical. Bar graphs show the mean and standard deviation of eGFP fluorescence from biological triplicates in the same experiment, analyzed as control (WT, i.e., "no-pulse" on cultures with constructs present) and pulsed (EW-113 treatment, with cells in the latter treatment analyzed as non-transformed, all eGFP and high eGFP (cells for which eGFP fluorescence was an order of magnitude higher than controls). Transformation efficiencies also reflect mean and standard deviations of biological triplicates. Images show natural chlorophyll fluorescence (red, top left), eGFP (green, top right) from a transformed cell (indicated by white arrow in all panels), with DAPI DNA staining (blue, bottom left), and the overlay of all three (bottom right) demonstrating the eGFP fluorescence is localized to the nucleus as designed. The scale bar represents 5 μm. (**d**) Expression of nourseothricin N-acetyl transferase (*nat*) in pIgNAT transformed *I. galbana*, from cells exhibiting resistance to nourseothricin. Expression of the *nat* transgene in WT (lane 3) and transformants (lanes 1 and 2) of *I. galbana* verified by RT-PCR. RT- negative and no template controls are shown in lanes 4 and 5, respectively, and 50 bp ladder (Bioline, lane 6), positive control PCR of genomic DNA from transformed cells (lane 7). (**e**) Western blot of expressed eGFP in *A. amoebiformis*, with total proteins extracted from transformants

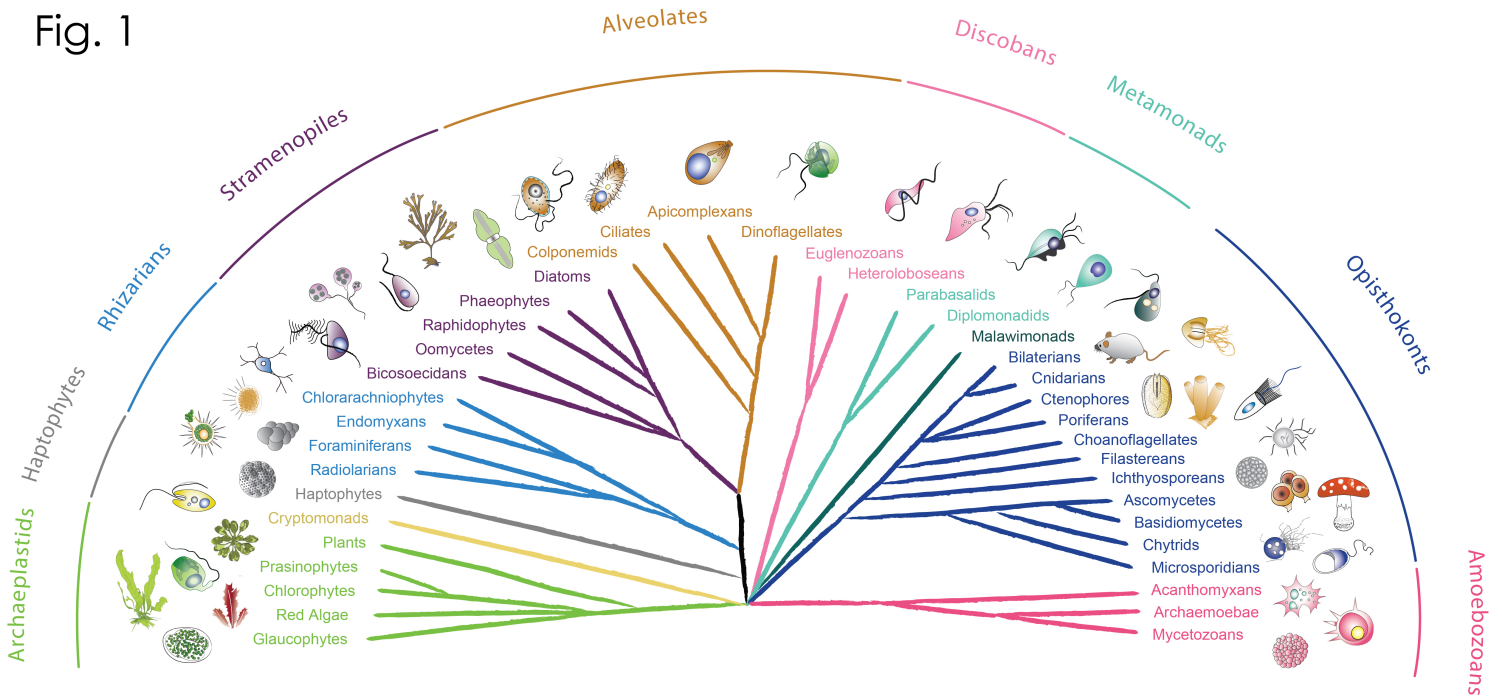
(TF) and WT cells using α-GFP monoclonal antibody (Takara; 1:1,000), and an α-mouse horseradish peroxidase (HRP)-coupled antibody (GE Healthcare; 1:10,000). (**f**) Expression of *eGFP*, *ShBle* and *Rhodopsin* (endogenous gene used as control) in WT (lane 2) and transformants (1350 FC2: lane 4, 1100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

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

Fig. 4. | ‘Transformation Roadmap’ for the creation of genetically tractable protists. (**a**) **Vector design and construction**. The three main goals for transformation of protists or other taxa are: 1) to determine an appropriate selection marker, 2) to design a suitable construct and 3) to find an appropriate transformation method for delivering DNA. Examples from this research exhibited different levels of success many with DNA delivered and expressed but ranging to DNA delivered without confirmation of gene expression or DNA

delivery not yet achieved (see **Fig. 1**). (**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).

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[25]		
Bigelowiella natans			

Stramenopiles

Fragilariopsis cylindrus	A		R
Thalassiosira pseudonana	B[15,17]		R
Seminavis robusta			
Pseudo-nitzschia multiseries	A		R
Heterosigma akashiwo			
Aurantiochytrium limacinum	B[31]		R
Caecitellus sp.			
Nannochloropsis oceanica	A		
Phaeodactylum tricornutum	B[15,18]		R

Alveolates

Euplotes crassus			
Euplotes focardii			
Chromera velia			
Perkinsus marinus	B[44,45]		R
Oxyrrhis marina	A		
Hematodinium sp.			
Fugacium (Symbiodinium) kawagutii			
Alexandrium catenella			
Breviolum (Symbiodinium) sp.			
Crypthecodinium cohnii	A		
Amphidinium carterae (chloroplast)	B[40]		R
Karlodinium veneficum	A		R

Discobans

Bodo saltans	A		R
Diplonema papillatum	B[51]		R
Eutreptiella gymnastica			
Naegleria gruberi	A		R

Opisthokonts

Pirum gemmata			
Sphaeroforma arctica			
Abeoforma whisleri	A		
Salpingoeca rosetta	C[55,56]		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

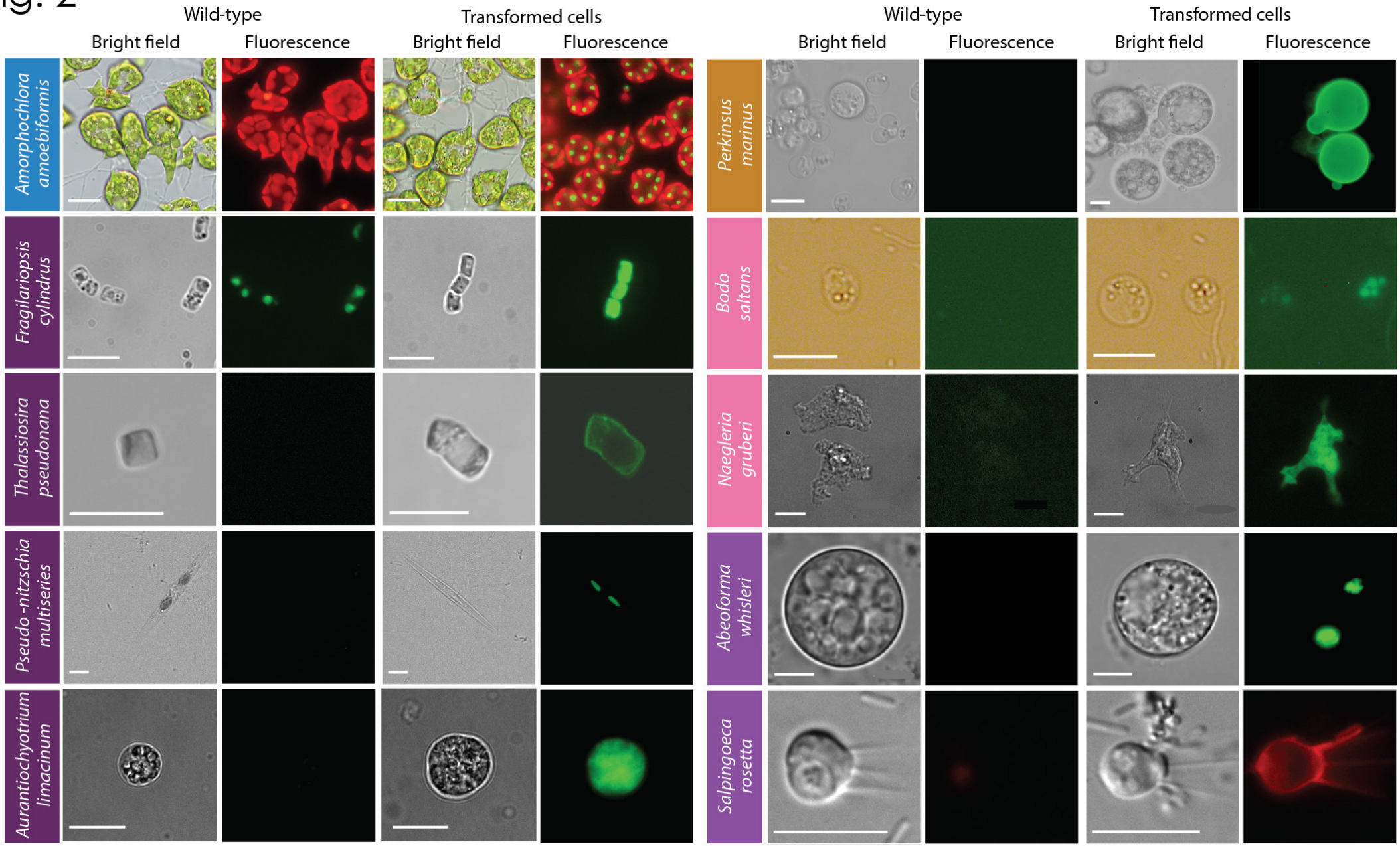

Wild-type
Transformed cells
Bright field
Fluorescence
Bright field
Fluorescence
Amorphochlora amoebiformis
Fragilariopsis cylindrus
Thalassiosira pseudonana
Pseudo-nitzschia multiseries
Aurantiochytrium limacinum
Wild-type
Transformed cells
Bright field
Fluorescence
Bright field
Fluorescence
Perkinsus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

Fig. 3

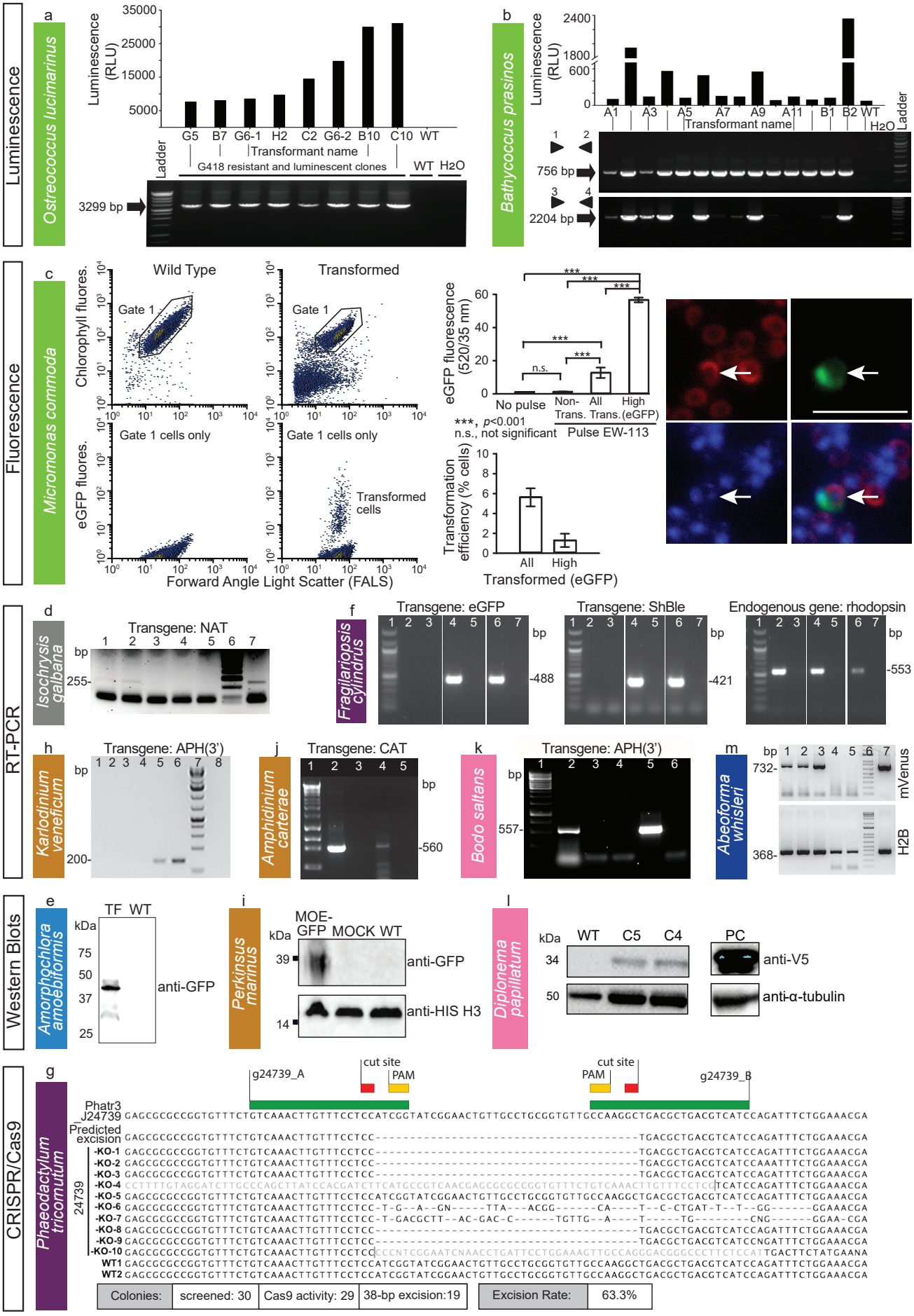

Luminescence
a
Ostreococcus lucimarinus
Luminescence (RLU)
35000
25000
15000
5000
G5 B7 G6-1 H2 C2 G6-2 B10 C10 WT
Transformant name
G418 resistant and luminescent clones
Ladder
WT H2O
3299 bp
b
Bathycoccus prasinos
Luminescence (RLU)
2400
1800
600
0
A1 A3 A5 A7 A9 A11 B1 B2 WT
Transformant name
H2O
Ladder
1 2
756 bp
3 4
2204 bp
Fluorescence
c
Micromonas commoda
Wild Type
Transformed
Gate 1
Chlorophyll fluores.
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)
***, p<0.001
n.s., not significant
Pulse EW-113
Transformation efficiency (% cells)
All
High
Transformed (eGFP)
RT-PCR
d
Isochrysis galbana
Transgene: NAT
bp
255-
f
Fragilariopsis cylindrus
Transgene: eGFP
-488
Transgene: ShBle
-421
Endogenous gene: rhodopsin
-553
h
Karlodinium veneficum
Transgene: APH(3')
200-
j
Amphidinium carterae
Transgene: CAT
-560
k
Bodo saltans
Transgene: APH(3')
557-
m
Abeoforma whisleri
732-
368-
mVenus
H2B
Western Blots
e
Amorphochlora amoebiformis
TF WT
kDa
75
50
37
25
anti-GFP
i
Perkinsus marinus
MOE-GFP MOCK WT
39
14
anti-GFP
anti-HIS H3
l
Diplonema papillatum
WT C5 C4
PC
34
50
anti-V5
anti-α-tubulin
CRISPR/Cas9
g
Phaeodactylum tricornutum
g24739_A
cut site
PAM
g24739_B
Phatr3 J24739
Predicted excision
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. 4

a

Vector construction

Microeukaryote of interest

Natural community

Physiological information

Genetic information

Commercial fluorescent dye

Drug selection screen

Well characterized gene

Transcriptome available

Yes

No

Genome available

Yes

No

Sequence transcriptome

Sequence genome

Feasible

Not feasible

Design resistance cassette

Highly expressed gene

Closest relative (with transfection)

Yes

No

Promoter identification/selection

Yes

No

Build plasmid

Build plasmid with flanking regions (min. 1 kb)

Homologous

Heterologous universal plasmids (e.g., CMV, SV40)

Label DNA

Heterologous

Tag gene (e.g., fluorescence or short protein tag)

Test transformation

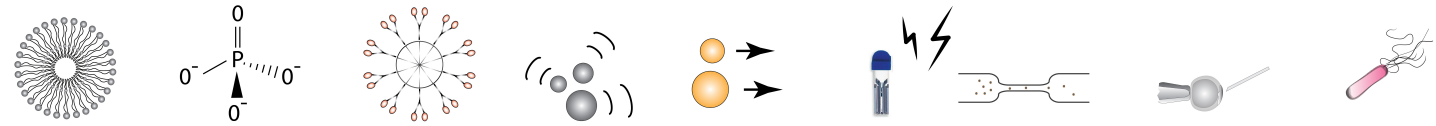
b

Transformation approaches

Chemical

Physical

Biological



Lipofection

Calcium phosphate

Dendrimers

Glass beads abrasion

Biolistics

Cuvette

Microfluidics

Microinjection

Conjugation

Electroporation

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 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)	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; 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.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}	pIgNAT; 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}	TpSil3p-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 (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.7gihjue 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- 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.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