Development of Genetic Transformation and Gene Delivery Methods in the Marine Microalga *Micromonas commoda*

A thesis submitted to the School of Environmental Sciences at the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Yixuan Li

May 2026

Registration No.: 100365031/1

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

Nitrogen availability in the oceans is a key factor influencing the productivity and diversity of microalgae. To overcome nitrogen limitation and competition, marine microalgae have evolved diverse nitrogen metabolic strategies. The green alga *Micromonas commoda* and the diatom *Thalassiosira pseudonana* are globally distributed marine microalgae that exhibit distinct nitrogen utilization strategies. Cyanate serves as an alternative nitrogen source in marine environments, and the key enzyme involved in cyanate metabolism is cyanase (*CynS*), which is present in both *M. commoda* and *T. pseudonana*. However, they display different capacities for cyanate utilization, and the role of *CynS* in nitrogen metabolism is not clear.

In *M. commoda*, functional genomic studies have been constrained by the lack of efficient and stable genome editing tools. To investigate the function of the *CynS* gene in *M. commoda*, the development of efficient and cost-effective transformation methods is essential. In this study, a robust and highly efficient polyethylene glycol (PEG)-mediated transformation method was established for DNA delivery into *M. commoda*. Transformed cell lines were able to grow on 0.3% seawater agarose plates containing zeocin and expressed enhanced yellow fluorescent protein (*eYFP*). Furthermore, a CRISPR/Cas9 construct was assembled using Golden Gate cloning, resulting in a 76-nt deletion within the coding region of the *CynS* gene and generating a frameshift mutation in *M. commoda*.

The PEG-mediated transformation of *T. pseudonana* was optimized. Prior to the transformation, Alcalase was employed as the sole enzyme for efficient cell wall removal. Single-cell colonies were isolated on selective plates and the expression of enhanced green fluorescent protein (*eGFP*) was confirmed.

In addition, the abilities of *M. commoda* and *T. pseudonana* to utilize different nitrogen sources were further investigated. *M. commoda* was able to utilize nitrate, ammonium, and cyanate as sole nitrogen sources, but exhibited growth retardation when grown on urea. In contrast, *T. pseudonana* was capable of utilizing nitrate, nitrite, ammonium, and urea as sole nitrogen sources, but showed no growth on cyanate. Stable isotope probing (SIP) results revealed low nitrogen uptake rates under cyanate treatment, indicating negligible direct assimilation of cyanate.

This thesis presents an efficient transformation and genome editing system for *M. commoda*, together with an optimized gene delivery method for *T. pseudonana*. This provides a foundation for functional gene characterization in marine microalgae.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

Abstract **2**

List of Figures **8**

List of Tables **20**

List of Abbreviations **22**

Acknowledgements **25**

Preface **27**

1 Introduction **29**

1.1 Introduction to marine microalgae 29

1.1.1 Marine microalgae and their ecological roles 29

1.1.2 Chlorophyta and the model green alga *Micromonas* 32

1.1.3 Diatom and the model species *Thalassiosira pseudonana* . . . 34

1.2 Nitrogen cycle in oceans and strategies of microalgae to utilize different nitrogen sources 36

1.2.1 Nitrogen cycle in oceans 36

1.2.2 Cyanate utilization in marine microorganisms 39

1.2.3 Adaptive strategies of marine microalgae to utilize different nitrogen sources 42

1.3 Molecular toolkits for marine microalgae 47

1.3.1 Overview of microalage transformation 47

1.3.2 Genetic engineering in diatoms 49

1.3.3 Genetic engineering in prasinophytes 51

1.4 Motivation and aims of this thesis 55

2 PEG-based transformation methods for *Micromonas commoda* and *Thalassiosira pseudonana* 57

2.1 Introduction 58

2.2 Materials and Methods 63

2.2.1 Cell cultures for *Micromonas commoda* and *Thalassiosira pseudonana* 63

2.2.2 Growth Curves 63

2.2.3 Extraction of *M. commoda* genomic DNA 65

2.2.4 Construct for *eYFP* and *shble* expression in *M.commoda* . . 66

2.2.5 Construct for 2A self-cleaving peptides expression in *M.commoda* 70

2.2.6 PEG-mediated transformation 74

2.2.7 Screening colonies and cultures 76

2.2.8 Microscopy 76

2.2.9 Western blot for protein expression analysis 77

2.3 Results 78

2.3.1 *Micromonas commoda* 78

2.3.2 *Thalassiosira pseudonana* 92

2.4 Discussion 94

3 Development of CRISPR/Cas9 in *Micromonas commoda* 99

3.1 Introduction 99

3.2 Materials and Methods 103

3.2.1 Cell culture 103

3.2.2 Designing sgRNAs and *in vitro* assay 103

3.2.3 Identification of U6 small nuclear RNA (snRNA) 104

3.2.4 Extraction of *M. commoda* and *M. pusilla* genomic DNA . . 106

3.2.5 Golden gate assembly 106

3.2.6 Transformation 109

3.2.7 Screening of knockout colonies 109

3.3 Results 112

3.3.1 Identification of U6 snRNA 112

3.3.2 The design of sgRNA for *CynS* gene knockout 114

3.3.3 Testing of cleavage efficiencies 114

3.3.4 Construction of knockout plasmids 117

3.3.5 Screening and verification of knockout transformant 121

3.3.6 Identification of DNA damage repair genes in *M. commoda* . 121

3.4 Discussion 122

4 Nitrogen utilization of *Micromonas commoda* and *Thalassiosira pseudonana* 129

4.1 Introduction 129

4.2 Methods 132

4.2.1 Cell cultures for *Micromonas commoda* and *Thalassiosira pseudonana* 132

4.2.2 Measurements of growth curves 134

4.2.3 Stable isotope incubation of *T. pseudonana* 134

4.2.4 Transformation and screening for *T. pseudonana* mutants . . 136

4.3 Results 139

4.3.1 The nitrogen utilization strategies of *T. pseudonana* wild-type and mutant cells lines 139

4.3.2 The nitrogen strategies of *Micromonas commoda* 148

4.4 Discussion 151

4.4.1 Nitrate reductase in *T. pseudonana* 151

4.4.2 The putative role of cyanase in diatom *T. pseudonana* and green algae *M. commoda* 154

5 Conclusions and future work 157

5.1 Discussion and conclusions 157

5.1.1 Establishment of a molecular toolkit for the marine picoeukaryote *Micromonas commoda* 157

5.1.2 PEG-mediated transformation method on *T. pseudonana* . . 159

5.1.3 Differentiation of nitrogen utilization in *Micromonas commoda* and *Thalassiosira pseudonana* 160

5.1.4 Functional divergence among nitrogen assimilation pathways in the marine diatom *Thalassiosira pseudonana* 161

5.2 Future work 162

5.2.1 Optimization of the PEG-mediated transformation method on *M. commoda* 162

5.2.2 CRISPR/Cas9 editing efficiency in *M. commoda* 163

5.2.3 Cyanate assimilation by *M. commoda* 164

5.2.4 Future investigation of cyanate metabolism in *T. pseudonana* 165

Bibliography **166**

Appendix **207**

List of Figures

1.1 Primary, secondary and tertiary endosymbiotic events in the microalgae evolution (Otte et al., 2025). 30

1.2 *Micromonas* phylogeny and distribution. (**A**) A consensus neighbor-joining (NJ) distance 18*S* rRNA gene tree illustrating the distinct *Micromonas* clades. (Inset) *Micromonas* thin section showing the nucleus (n), chloroplast (c), flagellum (f), and mucronate extension. (**B**) Sea surface temperature (SST) and locations where *Micromonas* (solid lines and circles around the isolates used in this work) and *Ostreococcus* (dashed lines) 18*S* rDNA sequences have been recovered (Worden et al., 2009). 33

1.3 Global distribution of centric genus *Thalassiosira* and the structure of *Thalassiosira pseudonana*. (A) Distribution of the genus *Thalassiosira* in both V4 and V9 datasets of Tara Oceans (Pierella Karlusich et al., 2025). Bubble areas are scaled to the % reads for the genus among eukaryotic reads at each station location, whereas crosses indicate absence of detection. The size fraction where the genus was most prevalent in samples is shown alongside the map. (B) Differential interference contrast (DIC) images of *T. pseudonana*. *Fr*, frustule; *Ch*, chloroplast. (C) Scanning electron microscope (SEM) images of *T. pseudonana* frustules. *Fp*, fultoportula; *G*, girdle; *GBs*, girdle bands. Scale bars = 1 μm. (B,C) are based on Fattorini and Maier (2021). 35

1.4 The marine nitrogen cycle (Hutchins and Capone, 2022). Red arrows indicate transfers or exchanges and blue arrows indicate microbial transformations. AOA, ammonia-oxidizing archaea; DDA, diatom–diazotroph association; DNRA, dissimilatory nitrate reduction to ammonia; NCD, non-cyanobacterial diazotroph; NOB, nitrite-oxidizing bacteria; Ox, oxidation; UCYN-A, unicellular cyanobacteria group A. 37

1.5 Scheme of cyanase-catalyzed bicarbonate-dependent decomposition of cyanate to ammonia and bicarbonate (Johnson and Anderson, 1987). 40

1.6 Distribution and phylogenetic analysis of *cynS* (Mao et al., 2022). **(A)** Biogeographic distribution and taxonomic composition of *cynS* and *ureC* transcripts across the global ocean, including the surface layer (SRF), deep chlorophyll maximum (DCM), and mesopelagic zone (MES). The red rectangle in subpanel (c) highlights the high expression of *cynS* in *Micromonas* in polar area. **(B)** Unrooted maximum-likelihood phylogenetic tree of *CynS* sequences. *CynS* from prasinophytes, diatoms, and dinoflagellates are highlighted. . . 41

1.7 Schematic of the canonical urea pathway across different cellular compartments (Horák et al., 2020). Red arrows indicate key enzymes, blue arrows show downstream enzymes, and yellow ellipses denote key metabolites. 47

1.8 Maximum-likelihood phylogenetic tree and transformation status of sequenced diatom genomes based on the 18S sequences. The phylogenetic tree depicts the evolutionary relationships among diatom species with complete genome sequences. Species with established genetic transformation protocols are highlighted with dots at the leaves. Representative morphologies of key species are illustrated alongside their respective phylogenetic positions. The tree is divided into three major groups based on the phylogeny: multi-polar centrics (green), araphid pennates (blue), and raphid pennates (brown). 50

2.1 Standard curve of *M. commoda*. 64

2.2 Overview of Golden Gate cloning assembly for transformation vector used in transforming *M. commoda*. L0 modules are created by PCR and then cloned into pCR4-TOPO vectors. Key components include the promoter (RPS9-prom) and terminator (RPS9-term), ampicillin resistance gene (AMP), kanamycin resistance gene (KAN), and target gene fragment *eYFP* and *shBle*. Specific 4nt overhangs are shown in different colors. The enzyme cleavage sites were marked as pink arrows. 68

2.3 Plasmid maps of pMCshBle:*eYFP* (a) and L1-RPS9:2A:*eYFP* (b) . 72

2.4 Antibiotic sensitivity of *M. commoda* in nourseothricin (a), cycloheximide (b), chloramphenicol (c), zeocin (d), blasticidin S (e) and postive control (f). Error bar represents mean ± SD (n = 3). . . 79

2.5 Comparison of codon usage. Codon usage as a fraction of triplets used per amino acid. Gene in *M.commoda* (red), *shBle* gene (dark gray), and enhanced yellow fluorescent protein gene (*eYFP*, light gray). 81

2.6 Restriction enzyme digestion and PCR analyses of L1 and L2 constructs of pMC*shBle*:*eYFP*. (a) Restriction digestion of L1-RPS9:*shBle* using EcoRI; (b) PCR screening of *E. coli* colonies carrying L1-RPS9:*eYFP*; (c) PCR screening of *E. coli* colonies carrying the L2 construct pMC*shBle*:*eYFP*. Orange triangles indicate the correct constructs as confirmed by Sanger sequencing. . 83

2.7 PCR analyses of four RPS9:*shBle*:2A:eYFP plasmids. Orange triangles indicate the correct constructs as confirmed by Sanger sequencing. 84

2.8 Screening of transformants using plasmid L1-RPS9:*shBle*. (a) Map of the transformation plasmid L1-RPS9:*shBle*; (b) Transformed cell cultures in K medium after 30 days of incubation; (c) Agarose gel electrophoresis showing that two transformants (PEG4000 and PEG8000 treatments) exhibit a 514 bp band corresponding to the RPS9 promoter and partial *shBle* gene sequence, whereas the wild type (WT) shows no corresponding band. The marker used is a 100 bp ladder. 86

2.9 Validation of *eYFP* and *shBle* genes in transformants. (a) PCR screening of the *eYFP* gene. Lanes 1–7 correspond to picked colonies, and the plasmid control is the transformed plasmid *pMCshBle:eYFP*. A band of 720 bp was detected in all colonies and the plasmid control, but no band was observed in the wild-type (WT). (b) PCR screening of the RPS9 promoter and partial *shBle* gene. Lanes 1–7 correspond to picked colonies, and the plasmid control is *pMCshBle:eYFP*. A band of 514 bp was detected in all colonies and the plasmid control, but no band was observed in the WT. The DNA marker used was a 100 bp ladder. 87

2.10 PEG-mediated transformation of *M. commoda*. (a) Transformants resistant to zeocin were selected in 0.3% agarose plate. (b) Fluorescence microscopy analysis of *eYFP* expression in *M. commoda* wild-type (WT) and transformatn L2SY. Shown are chlorophyll autofluorescence, *eYFP* fluorescence, bright-field images, and a merged image. Scale bars, 5 μm. (c) The growth curve of *M. commoda* WT and transformants in control medium with three replicates. (d) The maximum growth rate of *M. commoda* WT and transformants in control medium with three replicates. 88

2.11 Western blot analysis of wild-type and transformant strains was performed using an anti-eGFP antibody. (a) Colony L2SY exhibited a single band at 27 kDa, corresponding to *eYFP*. Four 2A transformants displayed a single band at 43.3 kDa, consistent with the fused protein *shBle*-2A-*eYFP*. No bands were detected for the wild type. (b) Protein loading control based on Ponceau S. 90

2.12 Confocal images of wild type and transformants.Shown are chlorophyll autofluorescence, *eYFP* fluorescence, bright-field images, and a merged image. Scale bars, 5 μm. 91

2.13 PEG-mediated transformation of *T. pseudonana* (Walker et al., 2026). (a) Harvested cells were subjected to varying concentrations of alcalase and incubation durations during the protoplasting step of the PEG transformation protocol. Following recovery, half of each reaction was plated on full-salt F/2 agar supplemented with 100 μg/mL nourseothricin. (b) Alcalase treatment using 10 μg/mL (30mAnson) for 40min. (c) Alcalase treatment using 1000 μg/mL (3 Anson) for 40min. (d) Agarose gel electrophoresis showing that transformants using different alcalase treatments (1.5 Anson,40min; 3 Anson, 20min; 3 Anson, 40min) exhibit a 212bp band corresponding to the *eGFP* gene sequence, whereas the wild type (WT) and water (negative control) shows no corresponding band. The marker used is a 100 bp ladder. (e) Fluorescence microscopy of wild-type (i.e., untransformed) and PEG-transformed *T. pseudonana* cells. Cells were transformed with pBIG1, which contains *eGFP*. 93

3.1 CRISPR/Cas9 gene editing system. Cas9 uses guide RNA (gRNA) to recognize target DNA sequences. When gRNA spacer pairs with a complementary DNA sequence adjacent to a PAM site, Cas9 cleaves the DNA, generating a blunt double-strand break via its RuvC and HNH nuclease domains (Pickar-Oliver and Gersbach, 2019). 100

3.2 Knockout plasmid maps of (a) MC56dYFP and (b) MC58dYFP. . . 110

3.3 Knockout plasmid maps of (a) Mp56dYFP and (b) Mp58dYFP. . . . 111

3.4 Alignment of U6 snRNA sequence in *M. pusilla* and *M. commoda* . 112

3.5 Alignment of U6 snRNA from representative species. 113

3.6 The *CynS* gene in *M. commoda*. (a) the map of the *M. commoda* *CynS* gene. Primers F2 and R2 were used to amplify the *CynS* gene template. sgRNAs 3–10 were designed for targeted genome editing. The yellow region represents the N-terminal segment of *cynS*. The green region denotes the catalytic domain of *cynS* and the red regions represent the ultra-conserved residues identified in Mao et al. (2022). (b) agarose gel results of plasmids F2-R2 pCR4-TOPO linearization checked by EcoRI digestion. 115

3.7 Results of in vitro test. (a) Agarose gel results of linearizated cyanase gene template plasmids digested by NcoI (4859bp). The marker used is a 1kb DNA ladder. (b) DNA template for *in vitro* T7 transcription of sgRNA. Bands 3–10 represent DNA template with designed sgRNA 3–10. The marker used is the ultra low range (ULR) DNA ladder.(c) sgRNA quality check by using 12% urea-TBE Polyacrylamide Gel Electrophoresis (PAGE). Bands 3–10 represent transcription of sgRNA 3–10. The marker used is the low range ssRNA ladder. (d) *In vitro* test results of sgRNAs and target sizes for successful cutting. Linearized plasmids of cyanase template without sgRNAs was set as a negative control. Yellow triangles indicate selected sgRNAs for following knockout plasmid constructs. 116

3.8 PCR results of L1 constructs for knockout plasmids. (a) Restriction digestion analysis of L1-RSP:cas9:*eYFP* using EcoRI. The estimated length is 9631bp; (b) Restriction digestion analysis of L1-RSP:cas9 using XmaI and MdeI. The estimated length is 4063bo and 5301bp; (c) Restriction digestion analysis of L1-MpU6:sgRNA5 using EcoRI. The estimated length is 4803bp; (d) Restriction digestion analysis of L1-MpU6:sgRNA6 and L1-MpU6:sgRNA8 using EcoRI. The estimated length of both plasmids are 4803bp; (e) Restriction digestion analysis of L1-McU6:sgRNA5, L1-McU6:sgRNA6 and L1-McU6:sgRNA8 using EcoRI. The estimated length of both plasmids are 5136bp. Red triangles indicate the correct constructs as confirmed by Sanger sequencing. DNA marker used was a 1kb ladder. 118

3.9 PCR results of L2 constructs for knockout plasmids. (a) Restriction digestion analysis of plasmid Mc5&6 and Mc5&8 using PstI and BamHI. The estimated DNA length are 4599bp and 7751bp. (b) PCR screening of *E. coli* colonies carrying plasmid Mc5&6dYFP, targeting DNA length of 1655bp. (c) PCR screening of *E. coli* colonies carrying Mc5&8dYFP, targeting DNA length of 1655bp; (d) PCR screening of *E. coli* colonies carrying Mp5&6dYFP, targeting band of 2798bp. (e) Restriction digestion analysis of Mp5&6dYFP using EcoRV and NcoI. The estimated length of both plasmids are 2053bp and 8911bp. Red triangles indicate the correct constructs confirmed by Sanger sequencing. DNA marker used was a 1kb ladder. 120

3.10 Screening and verification of mutant C34. (a) Screen of some colonies using three plasmid including Mp5&8dYFP, Mc5&8dYFP and Mc5&6dYFP. WT is positive control. The red triangles indicate C34 which potentially possess the edited target genes and further screen via Sanger sequencing. (b) Alignment of Sanger sequencing results with the cyanase gene from C34. (c) Alignment of amino acid sequence between WT and mutant C34. The green area marked the cyanase domain in WT sequence and the red font indicated the frameshift mutation in C34. (d) PCR results of cyanase gene in WT and C34. The DNA marker was a 1kb ladder. (e) PCR results of cas9 gene in C34. WT was a negative control. . 122

4.1 Schematic of compartmentation of major nitrogen assimilation, turnover, and links between the connected pathways in *P. tricornutum* (Smith et al., 2019). 132

4.2 Vector map of TPcynS 7&36. 137

4.3 PCR of cyanase gene in primary and secondary clones. (a) primary colonies selected on Aquil plates containing 100 µg/ml nourseothricin. Red arrows indicate mosaic colonies containing wild-type and potentially successfully edited cell lines; (b,c) secondary colonies selected from primary colonies on Aquil plates containing 100 µg/ml nourseothricin. Blue arrows indicate colonies which possess the edited target genes with no wild-type copies, and were isolated and confirmed via Sanger sequencing; (d) Sanger sequencing results of C3, C31 and C32. 140

4.4 Screening of bi-allelic knockouts and homologous recombination (HR) events by PCR and Sanger sequencing. (a) PCR-based screening of bi-allelic knockout of the *cynS* gene. (b) Transformant U1 carried a 26-bp deletion in one allele, whereas the corresponding region in the other allele was replaced by the NAT resistance cassette through homologous recombination. The HR-modified allele was not amplified by the primers used to analyse the target region. (c) HR at the *NitR* gene region, where transformants N1 and N2 showed no PCR amplification, indicating bi-allelic replacement. (d) Confirmation of HR-mediated integration of FCP:NAT resistance gene cassette at target loci in by inverse PCR in *NitR* mutants (Belshaw et al., 2023). (e) Sanger sequencing confirmation of *cyanase* knockout transformants C9M and C32, as well as the 26-bp deletion in one allele of transformant U1. (f) Overview of the CRISPR/Cas9-mediated HR used to generate the *URE* and *NitR* mutants. The Tp:Cas9_sgRNA vector expressed Cas9 and the sgRNA, while the 5′flank_TpFCP:NAT:3′flank donor vector provided the homologous repair template for NAT cassette integration (Belshaw et al., 2023). 141

4.5 Physiological responses including cell density, Chla fluorescence, PSII Fv/Fm and cell size of NitR mutants N1 (a) and N2 (b) under 100 μM nitrate and N limited condition, as well as *URE* mutant U1 (c) under 50 μM urea and N limited condition. Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean ± SD (n = 3). 142

4.6 Physiological responses of *T. pseudonana* to N-limitation and versatile nitrogen sources, including cell density (a) and PSII Fv/Fm (b). Error bar represents mean ± SD (n = 3). 144

4.7 Physiological responses of *T. pseudonana* to 100 μM cyanate. Cell density (a), Chla fluorescence (b), PSII Fv/Fm (c) and cell size (d) under 100 μM cyanate and N limited condition. (e) The isotopic composition δ values of *T. pseudonana* wide-type cells in 100 μM cyanate. (f) The specific uptake rate of 100 μM cyanate, 100 μM ammonium and 50 μM urea. Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean ± SD (n = 3). 145

4.8 After deep nitrogen limited treatment, physiological responses of *T. pseudonana* and its *cynS* mutants (a), *NitR* mutants (b) and *URE* mutant (c) to 100 μM nitrite, as well as growth rate in four days (d). Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean ± SD (n = 3). 147

4.9 After one-day nitrogen limited treatment, physiological responses of *T. pseudonana* and its *cynS* mutants (a), *NitR* mutants (b) and *URE* mutant (c) to 100 μM nitrite, as well as growth rate in four days (d). Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean ± SD (n = 3). 148

4.10 Physiological responses of *M. commoda* under different nitrogen sources. After nitrogen limited treatment (a), The physiological responses including Chla fluorescence (b) and Fv/Fm (c) with initial cell density at 1.6×10^6 cells/mL. After nitrogen limited treatment (d), The physiological responses including Chla fluorescence (e) and Fv/Fm (f) with initial cell density at 4×10^6 cells/mL. Error bar represents mean ± SD (n = 3). 149

4.11 Physiological responses of *M. commoda* under 100 μM cyanate and ammonium, as well as 5 μM ammounium added everyday. (a) Chla fluorescence; (b) Fv/Fm; (c) microscopy images of *M.commoda* culture. Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean ± SD (n = 3). 151

4.12 Physiological responses of *M. commoda* to cyanate gradient. (a) Chla fluorescence; (b) Fv/Fm; (c) maximum growth rate. Error bar represents mean ± SD (n = 3). 152

List of Tables

1.1 Summary of genome resources and genetic transformtion in Prasinophytes 54

2.1 PEG transformation on microalgae 62

2.2 Antibiotics and their concentrations 66

2.3 Primers used for amplification of fragments for L0 assembly. Uppercase letters indicate the original sequence, lowercase letters indicate the BsaI recognition site and defined overhangs, and the underlined sequences represent the defined overhangs. 69

2.4 Amino acid sequences of 2A peptides 73

2.5 Sequences of codon optimized 2A peptides 73

2.6 *Micromonas commoda* sensitivity to PEG concentration at different incubation times. "+" means cells can survive. "-" means cells can not survive. 85

3.1 Overview of sgRNAs designed and tested for cleavage efficiencies. Each sgRNA is followed by a specific PAM sequence. High out-of-frame scores ($>$ 66) are recommanded to minimize the occurrence of in-frame deletions. 105

3.2 PCR primers for creating Golden gate cloning L0 modules. BsaI sites are shown in bold. 4nt specific overhangs are shown in underlined. Lowercase letters represent extra added sequences, and uppercase letters indicate the part of primer combine with the initial template. 107

3.3 DNA damage repair genes in *M. commoda* 123

4.1 The amount of enriched reagent (^{15}N , 98% or 99%) and reagent at natural abundance needed for making 200ml of 0.1M enriched nitrogen stock (δ=500‰). 135

List of Abbreviations

AOA	Ammonia-oxidizing archaea
Chl a	Chlorophyll a
CPS	Carbamoyl phosphate synthase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CynS	Cyanase
DCR	endoribonuclease Dicer
DDAs	Diatom–diazotroph associations
DDR	DNA double-strand breaks
DHA	Docosahexaenoic acid
DIN	Dissolved inorganic nitrogen
DNA-PK	DNA–dependent protein kinase
DOC	Dissolved organic carbon
DON	Dissolved organic nitrogen
DSB	Double-strand DNA break
eGFP	Enhanced green fluorescent protein
EGT	Endosymbiotic gene transfer
EPA	Eicosapentaenoic acid
eYFP	Enhanced yellow fluorescent protein
FACS	Fluorescence-activated cell sorting
FCP	Fucoxanthin chlorophyll-a/c-binding protein
gDNA	Genomic DNA
GFP	Green fluorescent protein

GS/GOGAT	Gutamine synthetase/glutamine oxoglutarate amino transferase or glutamate synthase
HDR	Homology-directed repair
HGT	Horizontal gene transfer
HR	Homologous recombination
HRP	Horseradish peroxidase
HRM	High-resolution melt curve analysis
INDELs	Insertion–deletions
MNs	Meganucleases
NADH	Nicotinamide adenine dinucleotide hydride
NAT	Nourseothricin
nCas9	Cas9 nickase
NCBI	National Center for Biotechnology Information
NCDs	Non-cyanobacterial diazotrophs
NCMA	National Center for Marine Algae and Microbiota
NHEJ	Non-homologous end joining
NiR	Nitrite reductase
NOB	Nitrite-oxidizing bacteria
NOS	Nitric oxide synthase
NR	Nitrate reductase
ODZs	Oxygen-depleted zones
OUC	Ornithine–urea cycle
PAM	Protospacer adjacent motif
RCC	Roscoff Culture Collection
RI	Random insertion
RNAi	RNA interference
RNP	Ribonucleoprotein
RT-qPCR	Real-time quantitative PCR
RPS9	Ribosomal protein S9
SDM	Site-Directed Mutagenesis

sgRNAs	Single-guide RNAs
SIP	Stable isotope probing
snRNA	Small nuclear RNA
sRNAs	Small non-coding RNAs
SOW	Synthetic ocean water
T7EI	T7 endonuclease I assay
TAG	Triacylglycerol
TALEN	Transcriptional activator-like effector nuclease
ThiM	Thiazole kinase
UAL	Urea amidolyase
URE	Urease
WT	Wide-type

Acknowledgements

I am extremely grateful to my supervisors, Thomas Mock (University of East Anglia) and Yunyun Zhuang (Ocean University of China), for their invaluable advice, continuous support, and patience throughout my PhD studies. Thomas Mock has always been supportive and thoughtful in guiding my research. His extensive knowledge and broad experience provided clear scientific direction and insightful advice at every stage of this project. I was continually inspired by his creative ideas and encouraged to explore new research questions with confidence and enthusiasm. Yunyun Zhuang introduced me to this collaborative PhD project at UEA and has guided me with great patience and care, always offering steady and reliable support whenever needed. This PhD is financially supported by the China Scholarship Council (Certification No. 202106330018) and National Science Foundation of China (Cyanate utilization by marine eukaryotic phytoplankton and its ecological role; No.42276134), which are gratefully acknowledged.

I would like to thank all members of the Mock group. In particular, I sincerely thank Longji Deng, Nigel Belshaw, and Lulu Wang for their support with my lab work and experimental discussions. I am especially grateful to Antonia Otte for valuable discussions on experimental methods and for helping me with the bioinformatics analysis. I also thank all other members of the laboratory for the interesting suggestions, and insights into my research.

I would like to express my gratitude to my collaborators. Particularly thanks to Alexandra Z. Worden and Olga Matantseva for sharing their expertise and plasmids for CRISPR/Cas applications in *Micromonas commoda*. I am grateful to Emma Jane Lougheed Walker and Bogumil J. Karas for their important contributions to the

review paper and to the optimisation of the PEG-mediated transformation protocol for *Thalassiosira pseudonana*. I also thank Xuewei Mao for her support and valuable suggestions throughout this research. My thanks also go to Alina Marca for her guidance in the design and measurement of the stable isotope probing experiments. I would further like to thank Stephen Robinson and Heloise Vinette for providing training for nucleofection experiments in their laboratory.

I am deeply thankful to my friends: Shunan Fu, Mingliang Gu, Yunzhu He, Yang Li, Qiuxia Liao, Siyu Meng, Chuang Sun, Xueqiong Sun, Xiangpeng Wang, Yao Xiong, Ruoting Xiong, Yuhang Zheng and Xiaoyu Zhu. Your friendship, encouragement and the experiences we have shared have meant more to me than words can express. Your companionship has been a tremendous source of motivation throughout my PhD journey.

Finally, I owe my deepest gratitude to my parents and my grandfather. Your unconditional love and support have been the greatest source of strength, allowing me to move forward with confidence and perseverance.

Preface

This statement confirms that the work contained in this thesis was conceived, planned, conducted, interpreted and written by Yixuan Li. Prof. Thomas Mock, my primary supervisor, was involved throughout all stages of this PhD, including reviewing the five chapters contained within this thesis. Involvement of other members of the Mock lab and collaborators is outlined below.

Chapter 1 introduces the ecological importance of diatoms and green algae in marine environment. It explains the need for molecular tools for model species in diatoms and green algae. A publication, explained below, is included which gives a broad overview of molecular tools in diatoms. In this publication, I am the first author.

Genetic engineering in diatoms: advances and prospects. I conceived and designed the structure of this manuscript and wrote the first version. Dr. Longji Deng, Dr. Emma Jane Lougheed Walker, Dr. Bogumil J. Karas and Prof. Thomas Mock contributed to writing and editing the manuscript. Dr. Longji Deng designed Figures 1 and 2.

Chapter 2 describes the development of PEG-based transformation systems for *Micromonas commoda* and *Thalassiosira pseudonana*. Verification of *eYFP* and *eGFP* expression was conducted by Dr. Longji Deng. Western blot were performed with Xueqiong Sun and Dr. Longji Deng. Transformation in *T. pseudonana* was conducted in collaboration with Dr. Longji Deng, with the protocol development and discussion contributed by Dr. Emma Jane Lougheed Walker. The 20% PEG 8000 solution and alcalase used in the transformation protocol were provided by Dr. Emma Jane Lougheed Walker.

Chapter 3 details the development of a CRISPR/Cas9 system in *M. commoda*.

The plasmid containing the codon-optimized *Cas9* gene and the *M. pusilla* U6 promoter was provided by Dr. Olga Matantseva and Prof. Alexandra Z. Worden. The identification of the U6 snRNA sequence in *M. commoda* was conducted under the supervision of Dr. Amanda Hopes.

Chapter 4 examines nitrogen utilization in *M. commoda* and *T. pseudonana*. The plasmid used for the knockout of cyanase gene in *T. pseudonana* was constructed by Dr. Xuewei Mao during her postgraduate visit to Thomas Mock's laboratory. The nitrate reductase (*NitR*) and urease (*URE*) mutants were generated by Dr. Nigel Belshaw. Stable isotope probing (SIP) experiments were designed and conducted under the supervision of Dr. Alina Marca. Physiological experiments were designed and conducted under the supervision of Prof. Yunyun Zhuang.

1

Introduction

1.1 Introduction to marine microalgae

1.1.1 Marine microalgae and their ecological roles

Marine microalgae are unicellular eukaryotes ranging in size from ca. 0.95μm to $\sim$1000 μm. They are commonly classified into several major taxonomic phyla, including Bacillariophyceae (diatoms), Chlorophyta (green algae), Chlorarachniophyceae, Chrysophyta (golden algae), Cryptophyceae (cryptomonads), Dinophyta (dinoflagellates), Euglenophyta (euglenids), Haptophyta (haptophytes) and Rhodophyta (red algae).

Microalgal evolution has been shaped by endosymbiotic gene transfer (EGT) and horizontal gene transfer (HGT), particularly in lineages with complex plastids (Otte et al., 2025). Together with mutation, natural selection, genetic drift and gene flow, these processes have driven the reorganization of mosaic genomes from different microalgal species. Cyanobacteria are the prokaryotic group of oxygenic photoautotrophs, which began to oxygenate the atmosphere around 2.1 – 2.4 billion years ago (Schirrmeister et al., 2015). The earliest oxygenic photosynthetic eukaryotes have arisen from the wholesale engulfment of a cyanobacterium by a eukaryotic host cell (Figure 1.1). The engulfed cyanobacterium become a membrane-bounded organelle called the “plastid” (Falkowski et al., 2004). This is the primary endosymbiosis, which is thought to occur about 1.5 billion years ago

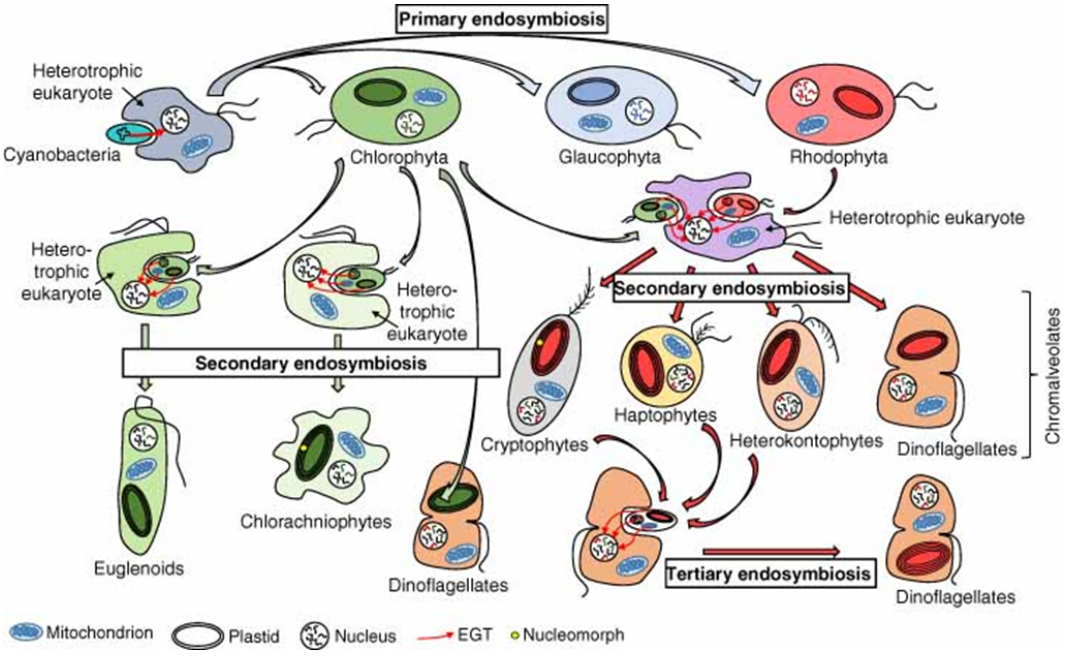


Figure 1.1: Primary, secondary and tertiary endosymbiotic events in the microalgae evolution (Otte et al., 2025).

(Yoon et al., 2004) and was split to chlorophyta, glaucophyta (freshwater) and rhodophyta. In the secondary endosymbiosis, chlorophyta were engulfed by another heterotrophic eukaryote, giving rise to euglenoids, chlorarachniophytes, as well as rhodophyta gave rise to the plastid in the cryptophyte, haptophyte, and heteroktophytes (ochrophytes) that occurred about 1.3 billion years ago (Yoon et al., 2004). It should be noted that dinoflagellates have undergone secondary or tertiary endosymbiosis.

The majority of microalgae are photosynthetic and are widely distributed throughout the oceanic photic zone (upper ca. 200m), where they contribute nearly half of global primary production and play essential roles in regulating global biogeochemical cycles (Field et al., 1998; Falkowski et al., 2008; Naselli-Flores and Padisák, 2023). In surface oceans, microalgae assimilate atmospheric CO_2 and synthesize organic carbon through photosynthesis, thereby contributing to the reduction of atmospheric CO_2 concentrations (Geider et al., 2001). Subsequently, fixed carbon is transferred through marine trophic networks or exported to deeper ocean layers via the sinking of organic matter, a key component of biological carbon pump (Tréguer et al., 2018). Some microalgae

exhibit mixotrophic capabilities, combining both phototrophic and heterotrophic modes (Cohen, 2022). This dual nutritional strategy improves their resilience in marine food webs and promotes a $\sim 35\%$ increase in carbon flux to higher trophic levels (Ward and Follows, 2016). Together with their rapid turnover on timescales of days, microalgae-fixed carbon is efficiently routed into marine food webs, supporting efficient energy fluxes and nutrient recycling.

Microalgae play a key role in marine biogeochemical cycles by coupling primary production with microbial processes, particularly through their interactions with heterotrophic bacteria (Ward and Follows, 2016; Hamilton et al., 2024). They form intricate associations with both symbiotic and free-living bacteria that involve resource exchange, chemical signaling, and community regulation (Amin et al., 2015). Bacteria provide microalgae with nutrients such as nitrogen and vitamins, and microalgae supply bacteria with organic matter and other essential metabolites (Segev et al., 2016; Paerl et al., 2017; Cirri and Pohnert, 2019; Kumar et al., 2024; Li et al., 2026). For instance, in the co-culture of coccolithophore *Emiliania huxleyi* and bacterium *Phaeobacter inhibens*, the maximum cell abundance of *E. huxleyi* reach approximately 25% higher than that in the absence of bacteria (Segev et al., 2016). In addition, a substantial fraction of carbon fixed by microalgal photosynthesis is released as dissolved organic carbon (DOC) and rapidly assimilated by heterotrophic marine bacteria (Azam et al., 1983; Amin et al., 2012). Organic compounds such as indole-3-acetic acid and tryptophan further function as signaling molecules in microalgal–bacterial interactions, mediating nutrient exchange and contributing to algal bloom (Amin et al., 2015).

1.1.2 Chlorophyta and the model green alga *Micromonas*

Green algae (Chlorophyta) is a diverse and ecologically important group of microalgae in both marine and freshwater ecosystems. Among them, Mamiellales represent a dominant lineage globally distributed from tropical to polar waters (Monier et al., 2016; Yung et al., 2022; Rey Redondo et al., 2024), and they comprise the genera *Bathycoccus*, *Ostreococcus*, *Micromonas*, *Mamiella*, and *Mantoniella* (Figure 1.2). Each genus exhibits a distinct biogeographic distribution pattern, reflecting specific adaptations to different ecological niches (Yung et al., 2022). Mamiellales genomes are compact (13–21 Mb), and eight high-quality whole-genome resources with curated gene annotations are currently available (https://phycocosm.jgi.doe.gov/mycocosm/species-tree/tree;XOhQH3?organism=mamiellophyceae). All described species to date are haploid. In addition, large-scale metagenomic sequencing efforts have generated a substantial number of Mamiellales genome sequences, further expanding available genomic resources (Sunagawa et al., 2015; Delmont et al., 2022; Boulton et al., 2025).

Micromonas are tiny ($\leq$2 μm), motile unicellular green algae possessing a single chloroplast and a mitochondrion. Two isolated *Micromonas* species have been sequenced and annotated, providing valuable model systems for experimental validation of gene function and genomic variation. The genome of *Micromonas commoda* RCC299 is fully assembled, whereas the *Micromonas pusilla* CCMP1545 genome was published as a high-quality draft comprising 21 scaffolds corresponding to 19 chromosomes (van Baren et al., 2016). The nuclear genome of *Micromonas* is highly reduced, with sizes of 20.9 Mb in *M. commoda* RCC299 and 21.9 Mb in *M. pusilla* CCMP1545. Originating from primary endosymbiosis, *Micromonas* is considered less derived than *Ostreococcus*, providing valuable insights into early green algal evolution and the genomic features of the proto-prasinophyte, the common ancestor of land plants and prasinophytes. Notably, *Micromonas* also exhibits exceptionally high genetic diversity, comprising

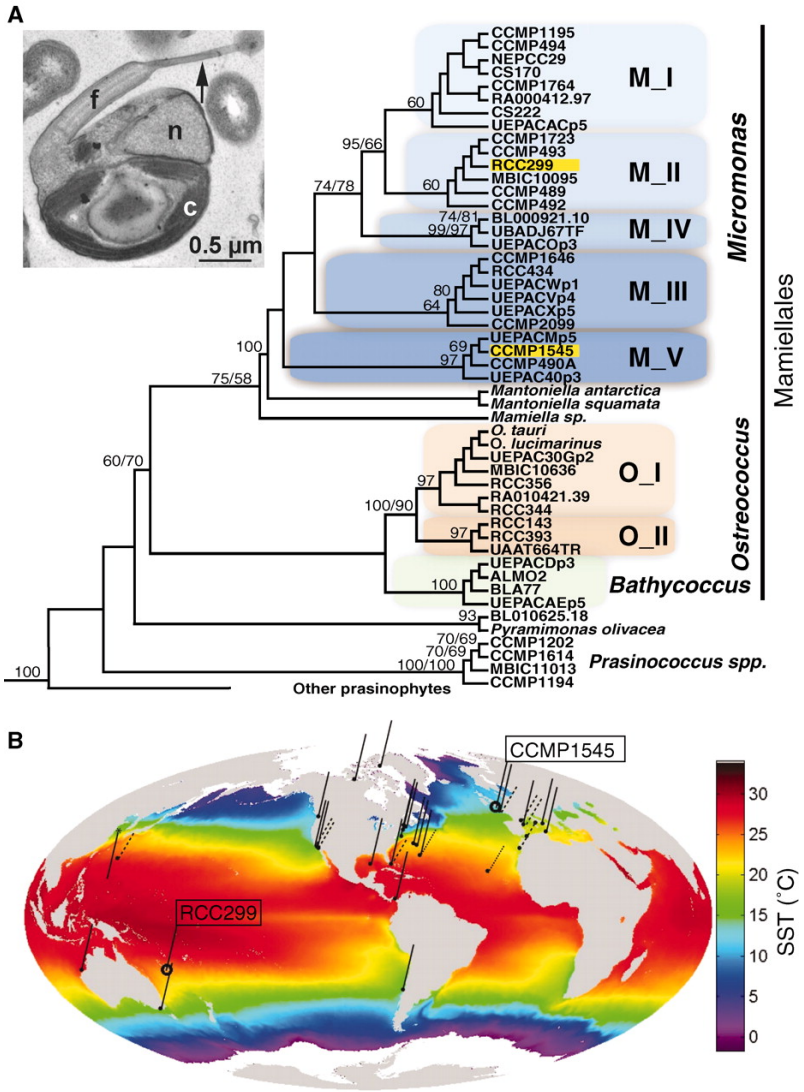


Figure 1.2: *Micromonas* phylogeny and distribution. (**A**) A consensus neighbor-joining (NJ) distance 18*S* rRNA gene tree illustrating the distinct *Micromonas* clades. (Inset) *Micromonas* thin section showing the nucleus (n), chloroplast (c), flagellum (f), and mucronate extension. (**B**) Sea surface temperature (SST) and locations where *Micromonas* (solid lines and circles around the isolates used in this work) and *Ostreococcus* (dashed lines) 18*S* rDNA sequences have been recovered (Worden et al., 2009).

numerous deeply divergent clades, with temperature acting as a key driver shaping both its diversification and global distribution. (Demory et al., 2019).

Since the publication of the first genome assembly, *Micromonas* has been widely used to study nutrients usage including nitrogen, phosphate, iron and vitamin B1 (McDonald et al., 2010; McRose et al., 2014; Guo et al., 2018; Hehenberger et al., 2023; Meyer et al., 2025), photosynthesis (Cuvelier et al., 2017; Dalessandri et al.,

2025) and symbiotic relationship between algae and bacteria (Ferrer-González et al., 2023; Hamilton et al., 2024). *M. commoda* RCC299 exhibits the most divergent repertoire of nitrogen transporters among available Mamiellales genomes (McDonald et al., 2010). The presence of distinct nitrogen-related gene clusters, together with the likely duplication of *NRT2*, suggests an enhanced capacity for nitrogen uptake and potentially higher expression levels. *M. commoda* RCC299 was isolated from the South Pacific Ocean, an oligotrophic environment characterized by low nutrient availability. These genomic features are consistent with adaptive evolution to nutrient-limited conditions. *Micromonas* is also a potential sentinel for ocean change (Demory et al., 2019). Taken together, *Micromonas* is a valuable model organism for the study of ecological and evolutionary processes in eukaryotic phytoplankton.

1.1.3 Diatom and the model species *Thalassiosira pseudonana*

Diatoms represent a key group of marine microalgae that dominate global ocean microbiomes (Pierella Karlusich et al., 2025). They are unicellular eukaryotes with plastids acquired through secondary endosymbiosis, derived from a red algal ancestor (Figure1.1). Diatoms contribute approximately 20% of global photosynthesis and account for nearly 40% of marine primary productivity, playing a critical role in regulating marine carbon cycle (Nelson et al., 1995). Characterized by silicified cell walls with intricate morphologies, diatoms are the largest contributors to biosilicification in the world and have been widely applied in bionanotechnology, including drug delivery, biosensing, and gene delivery (Dolatabadi and de la Guardia, 2011; Delalat et al., 2015; Tréguer et al., 2021; Wang et al., 2023a).

Cosmopolitan *Thalassiosira* is among the most abundant genera of diatoms, exhibiting higher abundance at high latitudes (Figure 1.3A). *Thalassiosira pseudonana* is one of the most prevalent species within this genus and is widely distributed in coastal waters worldwide, tolerating a broad range of salinities

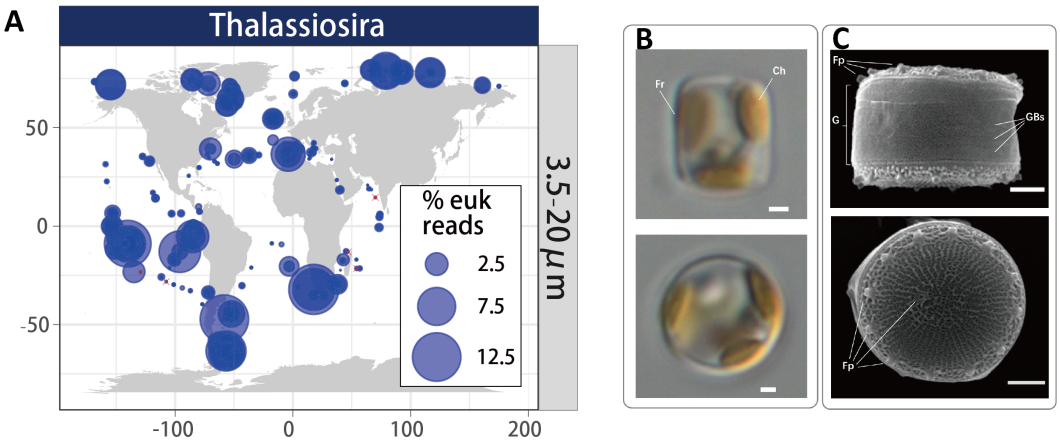


Figure 1.3: Global distribution of centric genus *Thalassiosira* and the structure of *Thalassiosira pseudonana*. (A) Distribution of the genus *Thalassiosira* in both V4 and V9 datasets of Tara Oceans (Pierella Karlusich et al., 2025). Bubble areas are scaled to the % reads for the genus among eukaryotic reads at each station location, whereas crosses indicate absence of detection. The size fraction where the genus was most prevalent in samples is shown alongside the map. (B) Differential interference contrast (DIC) images of *T. pseudonana*. *Fr*, frustule; *Ch*, chloroplast. (C) Scanning electron microscope (SEM) images of *T. pseudonana* frustules. *Fp*, fultoportula; *G*, girdle; *GBs*, girdle bands. Scale bars = 1 μm. (B,C) are based on Fattorini and Maier (2021).

(0.5–37‰) and temperatures (4–25°C; (Poulsen and Kröger, 2023)). With a high growth rate (> 2 doublings per day), *T. pseudonana* is now widely used as a model organism for investigating the molecular mechanisms underlying diatom biology, particularly the bio-morphogenesis of the silica cell wall (Figures 1.3B and C). It was the first eukaryotic marine alga with a sequenced genome in 2004 (Armbrust et al., 2004), revealing a relatively small nuclear genome of 34 Mb encoding 11,242 protein-coding genes. Subsequent studies have further improved genome and proteome annotations, enhancing its utility as a genetic and functional model system (Chen et al., 2023b). The first genetic transformation for *T. pseudonana* was successfully performed in 2006 (Poulsen et al., 2006) and the genetic modification based on CRISPR/Cas was reported later (Hopes et al., 2016; Belshaw et al., 2023; Messemer et al., 2026). The increasing availability of advanced molecular tools has promoted *T. pseudonana* as a model organism for investigating the ecology, evolution and gene function of diatoms (Mock et al., 2008; Martin et al., 2011; Ashworth et al., 2013; Kamalanathan et al., 2019).

1.2 Nitrogen cycle in oceans and strategies of microalgae to utilize different nitrogen sources

1.2.1 Nitrogen cycle in oceans

Nitrogen is an essential element for all living organisms and a fundamental component of cellular macromolecules, including amino acids, nucleic acids, and chlorophyll. Nitrogen in the marine environment exists in two major forms: inorganic nitrogen and organic nitrogen. Inorganic nitrogen mainly includes compounds such as nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+), nitric oxide (NO), and dinitrogen (N_2) (Figure 1.4). These forms are generally considered the main nitrogen sources for most marine microorganisms due to their relative abundance and bioavailability. Organic nitrogen comprises a diverse range of compounds, including amino acids, urea, cyanate, and other dissolved organic nitrogen (DON) molecules. Many marine microorganisms are capable of utilizing organic nitrogen, but this ability varies between species, reflecting differences in metabolic capacity and environmental adaptation.

In marine ecosystems, nitrogen availability limits primary productivity in approximately half of the global ocean, thus exerting a further influence on marine food webs and the biogeochemical cycle (Moore et al., 2013). Anthropogenic environmental change in marine ecosystems is intensifying. Reactive nitrogen produced through the Haber–Bosch process and released via agricultural runoff, together with nitrogen oxides emitted from fossil-fuel combustion, is increasingly entering marine environments and has substantially increased nitrogen input to the ocean and is estimated to have roughly doubled the turnover rate of the global nitrogen cycle (Gruber and Galloway, 2008). Ocean warming, acidification, enhanced seawater stratification, and ocean deoxygenation exert multiple effects on marine nitrogen biogeochemistry, especially influencing microbial nitrogen transformation pathways (Hutchins and Capone, 2022).

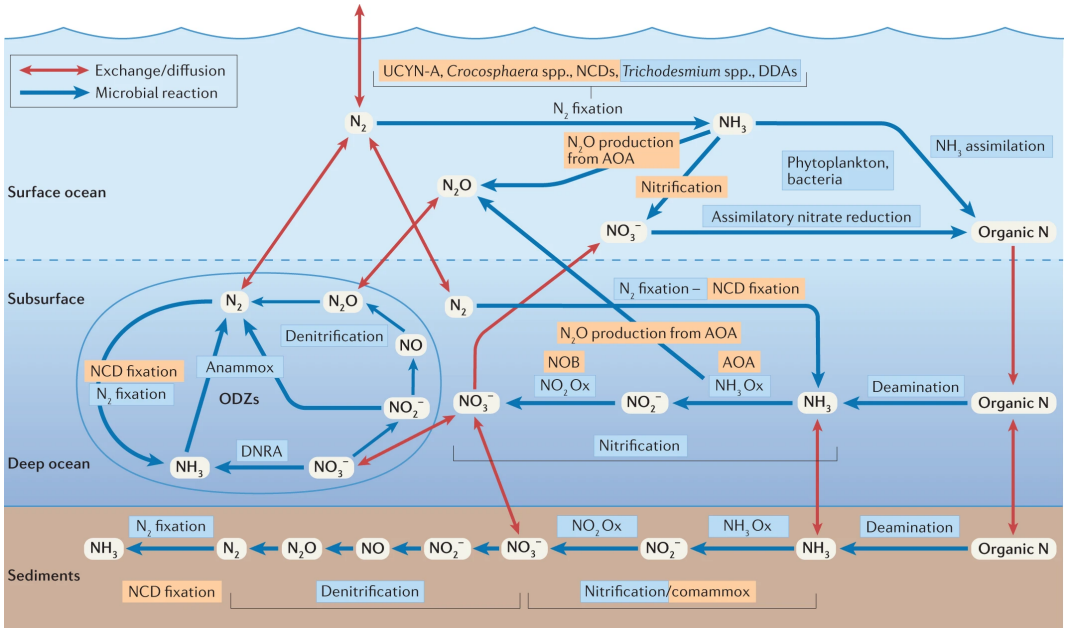


Figure 1.4: The marine nitrogen cycle (Hutchins and Capone, 2022). Red arrows indicate transfers or exchanges and blue arrows indicate microbial transformations. AOA, ammonia-oxidizing archaea; DDA, diatom–diazotroph association; DNRA, dissimilatory nitrate reduction to ammonia; NCD, non-cyanobacterial diazotroph; NOB, nitrite-oxidizing bacteria; Ox, oxidation; UCYN-A, unicellular cyanobacteria group A.

Traditionally, the nitrogen cycle comprises several major processes that transform nitrogen between different chemical forms, including nitrogen fixation, nitrification, denitrification, anammox, assimilation, and ammonification (Figure 1.4). Assimilation, nitrogen fixation, nitrification, and ammonification represent nitrogenous nutrient uptake and recycling pathways that generate and retain bioavailable nitrogen within the marine system. In contrast, nitrogen loss pathways include denitrification and anammox, both of which convert fixed nitrogen into gaseous forms and remove it from the system. The nitrogen cycle does not function in a strictly sequential manner but instead consists of distinct yet interconnected processes mediated by diverse microorganisms, forming a complex microbial network (Kuypers et al., 2018).

The largest freely accessible global nitrogen inventory in the ocean is in the form of N_2, but it is only accessible to microorganisms that carry the enzyme nitrogenase. Marine nitrogen fixers (diazotrophs) are the group that can convert N_2 to

bioavailable nitrogen predominantly in the form of NH_3. They are phylogenetically and physiologically diverse, including free-living and symbiotic cyanobacteria, non-cyanobacterial diazotrophs (NCDs) and diatom–diazotroph associations (DDAs) (Zehr et al., 2001; Schvarcz et al., 2022; Tschitschko et al., 2024). Marine diazotrophs are widely distributed throughout the global ocean, from oligotrophic tropical waters to polar regions. They contribute up to 50% of the new (exogenous) nitrogen input in oligotrophic environments, helping to offset losses from processes such as denitrification and anammox on global scales (Karl et al., 1997; Zehr and Capone, 2020). In addition to their role in nitrogen cycling, diazotrophs also contribute significantly to carbon and nitrogen export to the deep ocean, influencing long-term ocean carbon sequestration (Bonnet et al., 2023). However, there are several challenges facing marine diazotrophs. Diazotrophy is energetically expensive and highly sensitive to O_2 , and also being constrained by the availability of key nutrients such as nitrogen, iron and phosphorus. In the presence of sufficient bioavailable nitrogen, fast-growing non-diazotrophic phytoplankton can outcompete diazotrophs for other limiting nutrients (Ward et al., 2013). Additionally, the availability of fixed nitrogen suppresses N_2 fixation, as many diazotrophs preferentially assimilate dissolved inorganic nitrogen rather than incur the energetic cost of nitrogen fixation (Knapp, 2012; Turk-Kubo et al., 2023). These physiological and environmental constraints regulate diazotroph activity, ecological niche differentiation, and global biogeographic distribution (Hutchins and Boyd, 2016; Zehr and Capone, 2020; Hutchins and Sañudo-Wilhelmy, 2022).

Nitrification is a microbial process that oxidizes fully reduced nitrogen (NH_3 or NH_4^+) first to NO_2^- and then to NO_3^-, with nitrate being the predominant form of bioavailable nitrogen in the marine environment. This process is carried out primarily by ammonia-oxidizing archaea (AOA), which are diverse and widely distributed throughout the open ocean (Francis et al., 2005). Nitrite-oxidizing bacteria (NOB) subsequently oxidize nitrite to nitrate in oxygenated waters, but they constitute only a minor fraction of the bacterial community (0.1%) with the

main members belonging to the phylum *Nitrospinae* (Ward, 2011; Damashek et al., 2019). Ammonia oxidizers are the most abundant microbial functional group in subsurface waters, but their contribution to global oceanic CO_2 fixation and carbon sequestration remains poorly understood.

Bacteria and some eukaryotes can convert oxidized nitrogen (NO_3^-, NO_2^- , NO and N_2O) to N_2 under O_2-depleted conditions via denitrification, resulting in the loss of fixed nitrogen from marine systems (Devol, 2015). Anammox is another key nitrogen loss process in which NH_3 is oxidized with NO_2^- to produce N_2, making it a substantial contributor to nitrogen removal in oxygen-depleted zones (ODZs) (Zhang et al., 2020). Recent studies also show that sinking and suspended particles in the water column can create localized low-oxygen microenvironments that support the activity of denitrifying and anammox microorganisms (Fuchsman et al., 2017; Karthäuser et al., 2021). In addition, these particles likely supply ammonium through the remineralization of organic matter, providing a substrate for the anammox process.

1.2.2 Cyanate utilization in marine microorganisms

Cyanate (CNO^-) is a low-molecular-weight organic nitrogen compound widely distributed in marine environments, typically at nanomolar concentrations (Widner et al., 2013, 2016; Widner and Mulholland, 2017; Widner et al., 2018). There are several potential sources of cyanate in oceans, including terrestrial inputs, phytoplankton release, photochemical or spontaneous degradation of dissolved organic matter (Widner et al., 2016; Wang et al., 2024c). Increased cyanate concentrations in regions with high chlorophyll *a* suggest that cyanate production is closely related to phytoplankton metabolic activity in marine systems (Widner and Mulholland, 2017). Intracellular cyanate originates from the abiotic decomposition of nitrogen-containing compounds such as urea and carbamoyl phosphate. The observed linear increase in cyanate concentrations during the late stationary phase of diatom cultures, together with the absence of

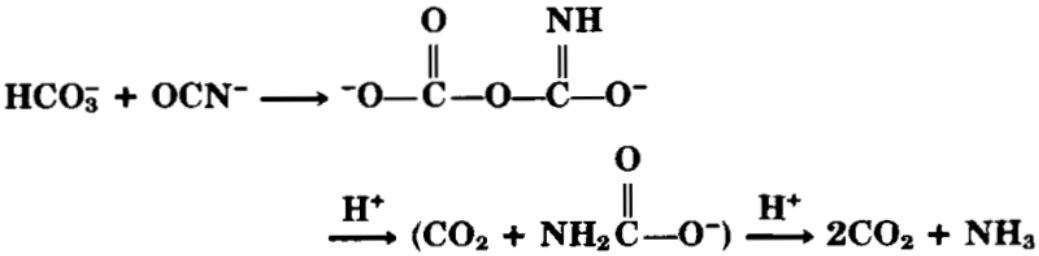


Figure 1.5: Scheme of cyanase-catalyzed bicarbonate-dependent decomposition of cyanate to ammonia and bicarbonate (Johnson and Anderson, 1987).

cyanate production in the dark, suggests that cyanate formation is closely related to light-dependent photochemical processes involving diatom-derived organic matter (Widner et al., 2016; Zhu et al., 2023). Cyanate metabolism was first investigated in *Escherichia coli* and then was identified across a broad range of organisms, including cyanobacteria, bacteria, fungi, and microalgae (Espie et al., 2007; Elleuche and Pöggeler, 2008; Yu et al., 2019; Díez et al., 2023). Central to this process is the enzyme cyanase (*CynS*), which catalyses the conversion of cyanate to carbon dioxide and ammonia through a bicarbonate-dependent reaction (Figure 1.5). Phylogenetic analysis shows that horizontal gene transfer of *cynS* is widespread across diverse lineages (Figure 1.6B). Marine metagenomic analyses show that *cynS* is as prevalent as *ureC*, indicating that cyanate is an important nitrogen source in marine microbial communities (Figure 1.6A).

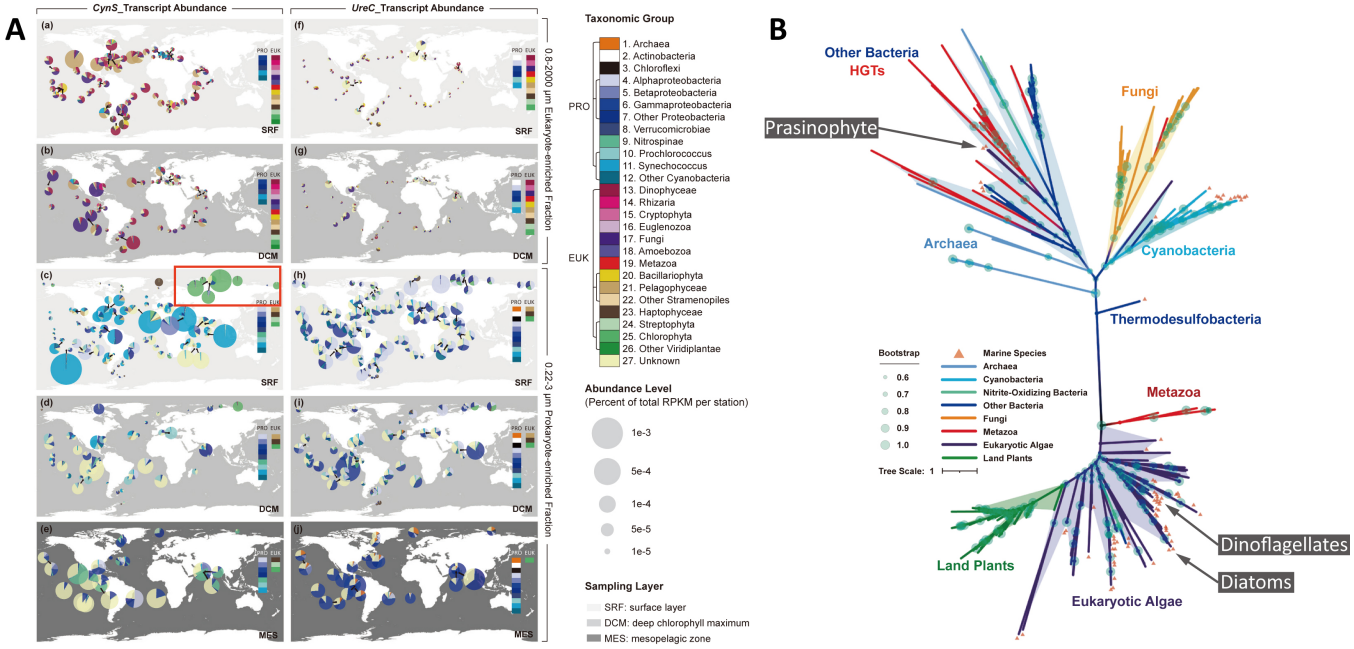


Figure 1.6: Distribution and phylogenetic analysis of *cynS* (Mao et al., 2022). **(A)** Biogeographic distribution and taxonomic composition of *cynS* and *ureC* transcripts across the global ocean, including the surface layer (SRF), deep chlorophyll maximum (DCM), and mesopelagic zone (MES). The red rectangle in subpanel (c) highlights the high expression of *cynS* in *Micromonas* in polar area. **(B)** Unrooted maximum-likelihood phylogenetic tree of *CynS* sequences. *CynS* from prasinophytes, diatoms, and dinoflagellates are highlighted.

CynS functions in three main roles: extracellular cyanate assimilation, intracellular cyanate detoxification, and metabolic regulation. Growth using cyanate as sole nitrogen source has been well documented in bacteria (Johnson and Anderson, 1987; Kunz and Nagappan, 1989), cyanobacteria (Espie et al., 2007; Kamennaya and Post, 2011; Sato et al., 2023), Thaumarchaeota (Palatinszky et al., 2015; Kitzinger et al., 2019) and fungi (Elleuche and Pöggeler, 2008; Linder, 2019), but microalgae appear largely unable to utilize extracellular cyanate (Guérin et al., 2025). Dinoflagellate *Prorocentrum donghaiense* is capable of growth on cyanate as the sole nitrogen source based on preliminary growth and nitrogen uptake results, but it occurs at lower rates relative to nitrate and ammonium (Hu et al., 2012, 2014). Environmental transcriptomic analyses reveal that the *CynS* gene in *Pelagomonas calceolata* is upregulated in nitrate-poor oceanic regions, supporting its persistence in oligotrophic environments (Guérin et al., 2025). However, laboratory experiments show that *CynS* expression in *P. calceolata* is downregulated in the presence of cyanate, indicating that this gene is unlikely to be involved in the catabolism of extracellular cyanate to ammonia. In an enrichment experiment, cyanate addition had negligible effects on both cellular chlorophyll fluorescence and the abundance of natural phytoplankton communities collected at the north Pacific subtropical gyre (Sato et al., 2023). The function of *CynS* in microalgae remains poorly understood. Nevertheless, *CynS* genes are widespread and diverse among marine microalgae across global oceans (Mao et al., 2022). Elucidating the role of *CynS* in microalgae is essential for improving our understanding of microalgal nitrogen metabolism and its contribution to the marine nitrogen cycle.

1.2.3 Adaptive strategies of marine microalgae to utilize different nitrogen sources

Nitrogen (N) is an essential element for life, but the proportion of biologically available inorganic nitrogen such as nitrate, nitrite, and ammonium constitutes only a small fraction of the total nitrogen pool on Earth (Schlesinger, 1997). As a

result, nitrogen availability frequently limits primary production in many ecosystems. Microalgae contribute the primary production in marine ecosystems and also play a critical role in the marine nitrogen cycle. The types and concentrations of available nitrogen sources strongly influence the distribution and abundance of microalgae. However, the dissolved nitrogen sources on which phytoplankton mainly rely are often limited across vast areas of the low-latitude surface ocean (Moore et al., 2013). Nitrogen limitation can induce multiple physiological responses in microalgae, including growth retardation and reduced capacity to adapt to environmental changes (Aranguren-Gassis et al., 2019; Smith et al., 2019). Furthermore, microalgae experience intense competition for available nitrogen sources and have therefore evolved diverse assimilation strategies to utilize different forms of nitrogen, enabling them to maintain their growth and ecological competitiveness. However, abnormal increases in nitrogen concentrations in the marine environment can promote the proliferation of harmful algal bloom which may be formed by organisms including diatoms, chrysophytes, and dinoflagellates. (Borkman et al., 2012; Gobler and Sunda, 2012; Sun et al., 2020). In recent years, anthropogenic activities have accelerated the eutrophication of coastal waters, resulting in more frequent red tide events and significant impacts on coastal ecosystems (Anderson et al., 2008; Sun et al., 2020).

1.2.3.1 Strategies on using inorganic nitrogen

Microalgae possess highly efficient carbon and nitrogen metabolic networks that enable rapid assimilation and transformation of nutrients (Yew et al., 2019). Most microalgae utilize dissolved inorganic nitrogen such as nitrate, nitrite, and ammonium as their primary nitrogen sources (Zehr and Kudela, 2011). There are significant differences in the expression of inorganic nitrogen metabolism–related pathways among microalgae species (Zhang et al., 2025). For example, diatoms possess highly efficient nitrogen assimilation systems, whereas dinoflagellates have relatively less complete pathways, which may contribute to their reduced capacity for inorganic nitrogen utilization compared to diatoms (Gao et al., 2018;

Pechkovskaya et al., 2020; Abassi and Ki, 2022). Nitrate and nitrite are transported into cells via specific transporters and are subsequently reduced to ammonium by nitrate reductase (*NR*) and nitrite reductase (*NiR*), separately. Nitrogen metabolism is closely coordinated with photosynthesis and carbon flux in microalgae (Falkowski and Stone, 1975). In the diatom *Phaeodactylum tricornutum*, *NR* knockout mutants exhibit more rapid and abrupt carbon repartitioning compared to wild type (McCarthy et al., 2017). Furthermore, genes associated with ribosome biogenesis, photosystems, chlorophyll biosynthesis, and carbon fixation and metabolism are downregulated following nitrate resupply, highlighting the key role of nitrate assimilation in coordinating carbon and nitrogen metabolism. Under nitrate-limited conditions, diatoms and some green algae redirect carbon flux from storage carbohydrates, such as chrysolaminarin and starch, toward the accumulation of neutral lipids, particularly triacylglycerols (TAGs) (Yu et al., 2009; Rismani-Yazdi et al., 2012). Therefore, nitrate depletion is an effective strategy to enhance TAG accumulation in diatoms such as *T. pseudonana* and *P. tricornutum* (Yu et al., 2009). Ammonium is the nitrogen form incorporated directly into organic nitrogen compounds, primarily through its assimilation into glutamate via the glutamine synthetase/glutamine oxoglutarate amino transferase or glutamate synthase (GS/GOGAT) cycle (Miflin and Lea, 1975). Ammonium is the preferred inorganic nitrogen source because its assimilation requires less energy than nitrate or nitrite. In many microalgal species, the presence of ammonium can suppress the uptake and assimilation of other nitrogen sources (Dortch, 1990; Huang et al., 2020; Fichtbauer et al., 2023).

Nitric oxide (NO) is a key signaling molecule produced by the enzyme nitric oxide synthase (NOS), playing an important regulatory role in diverse physiological processes in microalgae. NO functions as a redox-active messenger that modulates cellular responses to environmental and metabolic stress. For example, elevated intracellular NO levels induce the expression of death-specific proteins when photosynthetic electron transport is disrupted in the marine diatom *Skeletonema costatum* (Chung et al., 2008). In addition, NO-mediated stress signaling

influences the perception of substratum properties and results in reduced growth and impaired photosynthetic performance in the marine diatom *P. tricornutum* (Vardi et al., 2008). Furthermore, NO signaling is related to nitrogen assimilation processes, as it can be generated as a by-product of nitrate reduction and may serve as an indicator of intracellular nitrogen status. This comes out that NO is not only a stress signal but also an important regulatory molecule controlling fundamental cellular processes in marine microalgae (Astier et al., 2021).

1.2.3.2 Strategies on using organic nitrogen

Dissolved organic nitrogen (DON) can be broadly divided into labile, semilabile, and refractory fractions according to its bioavailability. The labile fraction (LDON), including amino acids, nucleic acids, urea, and proteinaceous compounds, is absorbed and assimilated by many microalgal species due to its low molecular complexity and high bioavailability (Wheeler et al., 1974; Sutton and Sposito, 2005; Huang et al., 2020; Li et al., 2026). ABC transporters are involved in the uptake of these organic nitrogen compounds into cells (Smith et al., 2019; Ji et al., 2020). Some microalgae even exhibit higher uptake and utilization rates for organic nitrogen than for inorganic nitrogen sources (Hu et al., 2014; Huang et al., 2020). For example, dinoflagellates generally do not show a significant advantage over diatoms in utilizing inorganic nitrogen, but they are often more competitive in exploiting organic nitrogen sources (Smayda, 1997). Urea is a major component of dissolved organic nitrogen in marine environments. Many dinoflagellate blooms in coastal waters have been associated with elevated urea concentrations, suggesting that urea may promote the competitive dominance of dinoflagellates over other phytoplankton under urea-enriched conditions (Li et al., 2009; Ou et al., 2014; Glibert and Burford, 2017). Furthermore, microalgal species exhibit selective utilization of DON components. During DON aging in marine environments, labile compounds are progressively depleted, leaving more refractory and hydrophobic components (Zheng et al., 2022; Li et al., 2025b). Diatoms preferentially utilize labile and hydrophilic DON, whereas dinoflagellates exhibit greater capacity to

exploit refractory and hydrophobic DON (Chen et al., 2023a; Li et al., 2025a). These differences in DON utilization strategies may contribute to the ecological succession from diatoms to dinoflagellates.

In addition to utilizing extracellular urea, marine photosynthetic stramenopiles have been reported to possess the ornithine–urea cycle (OUC, Figure 1.7), an anabolic pathway that enables nitrogen recycling and contributes to enhanced the efficiency of nitrogen utilization (Allen et al., 2011; Price et al., 2016; Smith et al., 2019). Carbamoyl phosphate synthase (CPS) acts as a key enzyme, channeling both carbon and nitrogen into the OUC in the form of carbamoyl phosphate, thereby linking primary metabolism to nitrogen assimilation (Allen et al., 2011; Horák et al., 2020). The distinct evolutionary origins of CPS suggest that different copies may perform different metabolic functions across lineages. In stramenopiles, the original CPS participates in OUC, and a duplicated copy is used for pyrimidine biosynthesis. In other photosynthetic eukaryotes, such as green algae and plants, the cyanobacterium-derived CPS primarily functions in pyrimidine synthesis, whereas rhodophytes and glaucophytes retain eukaryotic pyrimidine-associated CPS without involvement in OUC (Horák et al., 2020). Through the coordinated activity of downstream OUC enzymes, this cycle supports the recycling of nitrogenous compounds, helps maintaining intracellular carbon–nitrogen balance, and plays a critical role in sustaining cellular productivity under fluctuating nitrogen conditions in marine environments. Furthermore, the OUC may interact with other nitrogen metabolism pathways, such as amino acid biosynthesis and polyamine synthesis, enhancing overall nitrogen balance within the cell (Allen et al., 2011).

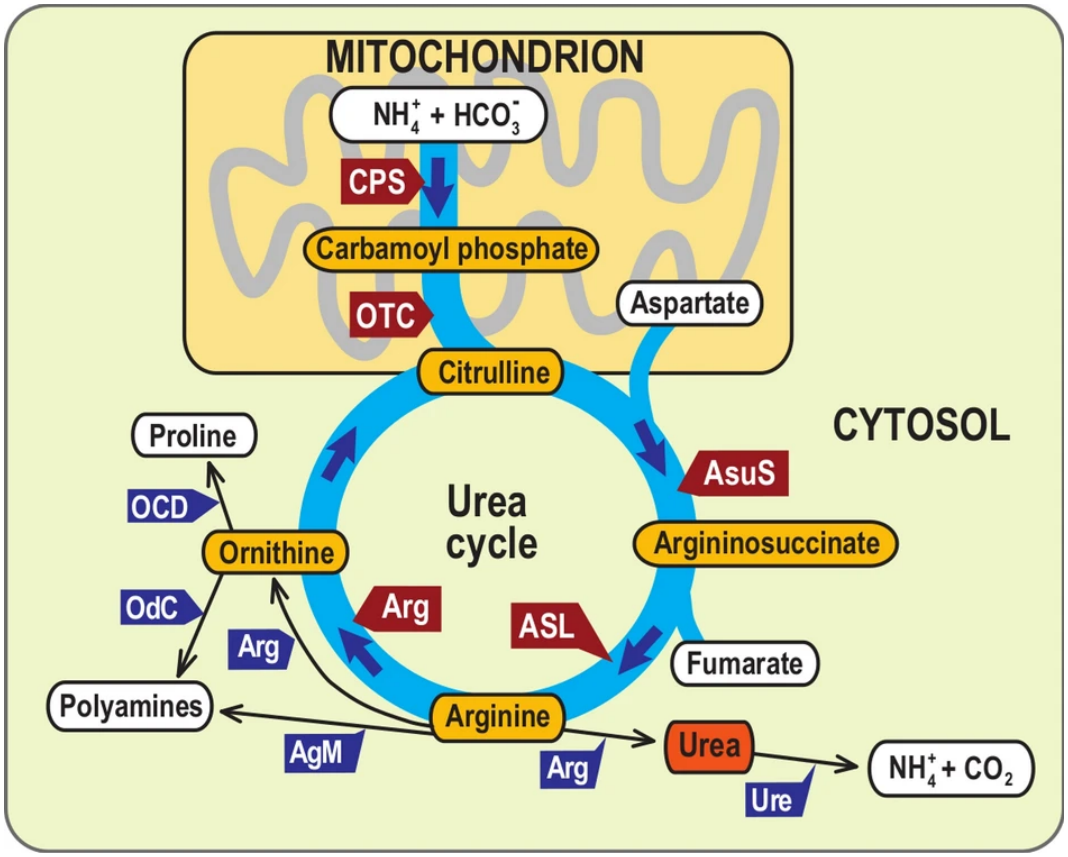


Figure 1.7: Schematic of the canonical urea pathway across different cellular compartments (Horák et al., 2020). Red arrows indicate key enzymes, blue arrows show downstream enzymes, and yellow ellipses denote key metabolites.

1.3 Molecular toolkits for marine microalgae

1.3.1 Overview of microalage transformation

Reverse genetics aims to alter gene expression in order to investigate gene function and to generate lines with predefined phenotypes. Given their importance in ecology, evolution, and bioengineering, molecular toolkits for microalgae have undergone substantial development. Together with advances in metabolomics, these tools pave the way for deciphering metabolic pathways and enable targeted metabolic engineering linked to metabolite production (Allen et al., 2011; Blaby et al., 2013; Dolatabadi and de la Guardia, 2011; Abbriano et al., 2018). This helps us improve our understanding of the biology and evolution of microalgae.

Microalgae have emerged as powerful photosynthetic cell factories in biotechnology. They have complex metabolic pathways that generate diverse value-added compounds with applications in commercial production (Sreenikethanam et al., 2022). Traditionally, different microalgal species have been selected for specific purposes to produce distinct bioproducts (Pulz and Gross, 2004). However, wild-type strains have inherent limitations: their yields can only be increased by optimizing culture conditions such as light, nutrients, and temperature, and even then, the maximum productivity remains relatively low. Genetic engineering is able to overcome this constraint in metabolic capacity and then enhance productivity. This is typically achieved through overexpression or targeted knockdown of relevant genes, enabling the production of commercially valuable compounds such as biofuels, antibodies, and carotenoids (Trentacoste et al., 2013; Hempel et al., 2017; Bule et al., 2018). For example, overexpression of a Δ5-elongase and phytase gene from other species in *P. tricornutum* enables the accumulation of high-value docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA) (Hamilton et al., 2014; Pudney et al., 2019).

In summary, the development of genetic modification tools in microalgae is critical not only for uncovering fundamental genetic and regulatory mechanisms, but also for expanding their potential as platforms for biotechnological applications. However, targeted genome editing remains feasible for only a limited number of species, largely due to technological constraints such as low transformation efficiency, lack of efficient selectable markers, and limited availability of species-specific molecular toolkits. Nowadays, many important microalgae species remain challenging on stable transformation and targeted genome editing (Faktorova et al., 2020). This limitation significantly hinders the functional characterization of genes and the exploration of their metabolic and regulatory networks. Developing robust genetic tools is essential to unlock the potential of these species for both basic research and biotechnological applications.

1.3.2 Genetic engineering in diatoms

Before any genetic loci can be modified using targeted approaches, there need to be methods for delivering and establishing genetic material in diatom cells. Usually, these transformation protocols are developed using the expression of marker genes such as green fluorescent protein (GFP) and yellow fluorescent protein (YFP), which, based on their fluorescence, help to assess the stability of the genetic modification. The latter is an important criterion for many applications in reverse genetics with diatoms (Falciatore et al., 1999; Faktorova et al., 2020; Sabatino et al., 2015; Yin and Hu, 2021). Once a transformation protocol has been developed, many endogenous genetic loci can be manipulated. The most used genetic modifications in diatoms include gene overexpression, knockdown, and knockout. Furthermore, some diatoms such as *T. pseudonana* can be used for efficient gene replacement facilitated by homology-directed repair (Belshaw et al., 2023). Until now, the genetic transformation systems have been established in 20 diatom species (Figure 1.8).

Gene overexpression was first successfully employed in diatoms in 1995 (Dunahay et al., 1995). This approach typically involves the knock-in of an additional copy of the target gene driven by a strong constitutive promoter, such as the fucoxanthin chlorophyll-a/c-binding protein (*FCP*) promoter. Overexpression can confer new phenotypic traits, allowing researchers to infer potential functions of the target gene.

Gene knockdown is the process that targets mRNA instead of DNA, thereby resulting in the decrease of the encoded protein. The first report of gene knockdown was based on inhibiting the expression of the GUS reporter and the endogenous cryptochrome gene in *P. tricornutum* using antisense RNA technology (De Riso et al., 2009). Since then, gene silencing has been widely applied in diatoms (Trentacoste et al., 2013; Shrestha and Hildebrand, 2017; Sabatino et al., 2022; Wang et al., 2023a). However, the molecular mechanism underlying gene silencing remains to be completely resolved because it does not appear to be following the canonical mechanism of RNA interference (RNAi) mediated by small

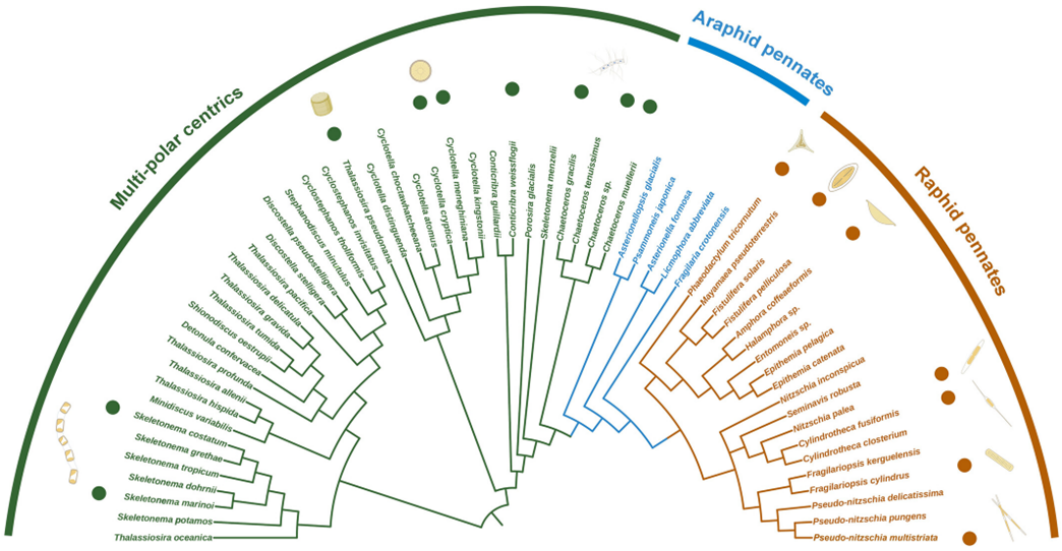


Figure 1.8: Maximum-likelihood phylogenetic tree and transformation status of sequenced diatom genomes based on the 18S sequences. The phylogenetic tree depicts the evolutionary relationships among diatom species with complete genome sequences. Species with established genetic transformation protocols are highlighted with dots at the leaves. Representative morphologies of key species are illustrated alongside their respective phylogenetic positions. The tree is divided into three major groups based on the phylogeny: multi-polar centrics (green), araphid pennates (blue), and raphid pennates (brown).

non-coding RNAs (sRNAs) and the activity of the endoribonuclease Dicer (DCR) because the mRNA of the silenced genes is still present, as observed in *T. pseudonana*, for instance, Kirkham et al. (2017). However, new data based on comparative genome and transcriptome analyses and reverse genetics with *P. tricornutum* have in fact revealed that key RNAi effectors have diversified in diatoms. Nevertheless, this new study suggests the presence of distinct RNAi pathways in diatoms. For instance, the *P. tricornutum* DCR was found to process 26–31-nt-long double-stranded sRNAs originating mostly from transposons covered by repression-associated epigenetic marks (Grypioti et al., 2024).

In addition to gene overexpression and silencing in diatoms, the last decade has revolutionized our ability to modify the diatom genome by enabling targeted modifications using sequence-specific nucleases. Modifications are achieved through the application of meganucleases (MNs), transcriptional activator-like effector nucleases (TALENs), and especially CRISPR/Cas-assisted genome editing. These genetic tools introduce double-strand DNA breaks (DSB) at specific loci, with

subsequent repair either via homology-directed repair (HDR) or non-homologous end joining (NHEJ). The latter is the most frequent repair mechanism in diatoms, acting on single-nucleotide mutations and small insertion–deletions (INDELs). MNs, TALENs, and CRISPR/Cas have been used successfully to induce targeted mutations in diatom genomes (Daboussi et al., 2014; Fortunato et al., 2016). TALENs have the advantage of high specificity in genome editing and have been applied in both *P. tricornutum* and *T. pseudonana* (Daboussi et al., 2014; Fortunato et al., 2016; Weyman et al., 2015). However, a typical TALEN system requires pairing two TALENs to target DNA sequences of 30–36 base pairs, with each side consisting of 18 repeats (Malzahn et al., 2017). Thus, the design is complex and not as versatile as the CRISPR/Cas technology. Therefore, MNs and TALENs were gradually replaced in diatom research by the effective CRISPR/Cas technology such as Hopes et al. (2016). The first reports for successful CRISPR/Cas-based genome editing leading to stable gene knockouts were based on work with *P. tricornutum* (Nymark et al., 2016) and *T. pseudonana* (Hopes et al., 2016). Following these landmark papers, CRISPR/Cas-mediated homologous recombination (HR) was developed in *P. tricornutum*, *T. pseudonana*, and *Chaetoceros muelleri*, providing targeted genome editing at endogenous loci (Moosburner et al., 2020; Belshaw et al., 2023; Yin and Hu, 2023). The first application of DNA-free gene knockout by direct delivery of CRISPR/Cas9 ribonucleoproteins (RNPs) in diatom was published in 2018 (Serif et al., 2018). Furthermore, the Cas9 nickase (nCas9) and dead (d)Cas9 have also been developed for diatom research (Nawaly et al., 2020; Matsui et al., 2024; Guo et al., 2024).

1.3.3 Genetic engineering in prasinophytes

Prasinophytes form a basal algal group in the green lineage and are widespread marine green algae from polar to tropical regions. Although many species have sequenced genomes, few are available to genetic transformation (Table 1.1).

Ostreococcus was the first prasinophytes species in which genetic transformation

methods were established, including electroporation and PEG-basd transformation (van Ooijen et al., 2012; Sanchez et al., 2019). The Km gene (conferring eukaryotic G418 resistance) was used as a selectable marker, and luciferase served as a reporter in transformations. Moreover, efficient gene targeting via homologous recombination (HR) enables precise knockouts and reporter insertions in *Ostreococcus* (Lozano et al., 2014). Random insertion (RI) and targeted insertion via HR rely on distinct DNA repair mechanisms. RI is mediated by NHEJ of double-strand breaks, whereas HR is recombinase-dependent and involves strand exchange with homologous DNA sequences (Shinohara et al., 1992; Heyer et al., 2010). *O. tauri* RCC745 Δ*ThiM* deletion mutants were generated using HR to investigate the function of thiazole kinase (ThiM) in thiamine (vitamin B_1) metabolism (Paerl et al., 2017). Specifically, this approach enabled the assessment of the ability of *O. tauri* to grow on thiamine precursors in the absence of *ThiM*, thereby confirming its role in the salvage pathway of vitamin B_1 biosynthesis.

The protocols adapted from *O. tauri* using electroporation were transferable to *O. lucimarinus* (RCC802) and *B. prasinos* (RCC4222) (Faktorova et al., 2020). *O. lucimarinus* (RCC802) was successfully transformed by electroporation using an *O. tauri*-derived construct, resulting in numerous G418-resistant colonies with luciferase expression. In contrast, transformation of *B. prasinos* (RCC4222) required species-specific regulatory sequences, and although fewer transformants were obtained, successful genomic integration and luciferase expression were achieved in a subset of clones, indicating the importance of genomic context for transgene expression.

In general, electroporation is the most commonly used method for the stable introduction of exogenous DNA into cells. However, transformation protocols successful in *Bathycoccus* and *Ostreococcus* failed in *Micromonas*. Instead, Lonza nucleofection enabled successful transformation of *M. commoda* (CCMP2709) using codon-optimized constructs expressing either luciferase or *eGFP* under exogenous or endogenous regulatory elements (Faktorova et al., 2020). Nevertheless, due to the limited availability of effective selectable markers,

transformants were overgrown by wild-type cells during prolonged incubation.

T. striata (KAS-836) was successfully transformed via microprojectile bombardment using gold particle–coated plasmids carrying the *bar* gene, with stable transformants obtained under zeocin or glufosinate selection.

Table 1.1: Summary of genome resources and genetic transformtion in Prasinophytes

Speices	Delivery system	Selective marker	Reporter	Approach	Genetic Transformation
Ostreococcus tauri RCC1115	PEG	*kanMX*	LUC	OE	(Sanchez et al., 2019)
Ostreococcus tauri 0TTH0595	Electroporation	*kanMX*	–	OE, HR	(van Ooijen et al., 2012) (Lozano et al., 2014)
Ostreococcus lucimarinus RCC802	Electroporation	*kanMX*	LUC	OE	(Faktorova et al., 2020)
Bathycoccus prasinos RCC4222	Electroporation	*kanMX*	LUC	OE	(Faktorova et al., 2020)
Micromonas commoda RCC299	Lonza Nucleofector	–	eGFP, LUC	OE	(Faktorova et al., 2020)
Tetraselmis striata KAS-836	Biolistics	*Ble*	–	OE	(Faktorova et al., 2020)

Ble: Bleomycin resistance gene; *eGFP*: enhanced green fluorescent protein gene; HR:homologous recombination; LUC: luciferase; *kanMX*; Kanamycin/ G418 resistance gene; LUC: luciferase; OE: overexpression; PEG: PEG–mediated transformation.

1.4 Motivation and aims of this thesis

The motivation of this thesis was to investigate the strategies of extracellular nitrogen assimilation in marine microalgae within the marine environment, as well as their intracellular nitrogen metabolism. Marine microalgae have evolved diverse nitrogen assimilation strategies to adapt to fluctuating ocean conditions and play a critical role in the marine nitrogen cycle. The green alga *M. commoda* and the diatom *T. pseudonana* are among the most abundant marine microalgal species. The development of genetic transformation systems is essential for enabling deeper investigation of their nitrogen metabolism. *T. pseudonana* has been extensively studied and is readily transformable, but transformation methods for *M. commoda* remain limited. To date, functional genomics studies in *M. commoda* have been constrained by the lack of efficient and stable genome editing tools.

Chapter 2 describes the development of stable and efficient polyethylene glycol (PEG)–mediated transformation protocols for DNA delivery into *M. commoda* and *T. pseudonana*. For *M. commoda*, the transformation vector contains the *shble* gene and enhanced yellow fluorescent protein (*eYFP*), both flanked by endogenous promoter and terminator sequences from the ribosomal protein S9 (RPS9) gene. Following transformation, single-cell colonies were isolated from 0.3% agarose plates supplemented with 25 μg/mL zeocin. The expression of *eYFP* was confirmed by fluorescence detection using a widefield microscope. For *T. pseudonana*, transformation vector contains selective marker *NAT* gene and *eGFP*. *T. pseudonana* cell wall were removed by alcalase and then the vector were delivered via PEG. Single-cell colonies were isolated on 0.8% agarose platecontaining 100μg/mL Nourseothricin. The expression of *eGFP* was confirmed byfluorescence detection using a widefield microscope.

Chapter 3 presents the results of genome editing in *M. commoda* using CRISPR/Cas9 to knockout the *CynS* gene. Eight sgRNAs were designed and validated through *in vitro* test. Four CRISPR/Cas9 plasmids were then assembled

using Golden Gate cloning and delivered into cells. Following screening of the transformed cell lines, cyanase knockout strains were successfully identified. This is the first report of CRISPR/Cas9–mediated gene knockout in *M. commoda*.

Chapter 4 investigates the ability of the marine microalgae *M. commoda* and *T. pseudonana* to utilize cyanate as a nitrogen source. *M. commoda* is capable of using cyanate as the sole nitrogen source, although its growth is limited. In contrast, *T. pseudonana* cannot grow on cyanate as the sole nitrogen source despite possessing the *CynS* gene. To elucidate the function of *CynS* in *T. pseudonana*, *CynS* knockout strains were generated. When cultured with nitrite as the nitrogen source, the knockout strains exhibited physiological differences compared with the wild type. Specifically, *CynS* knockout strains displayed increased cell size and a reduced maximum growth rate compared with that of wild type. These findings suggest that *CynS* functions in cyanate detoxification and may also influence metabolic regulation.

Lastly, chapter 5 concludes with a summary of this research of the main results, followed by suggestions for future research.

2

PEG-based transformation methods for *Micromonas commoda* and *Thalassiosira pseudonana*

The results of PEG-based transformation on *Thalassiosira pseudonana* have been published in *Nature communications* (https://doi.org/10.1038/s41467-026-68562-6). The main text and supporting information from the published article are included here, with figure and an associated short discussion added for this thesis.

2.1 Introduction

PEG-mediated genetic transformation was first reported in tobacco protoplasts in 1982 and later applied to many other organisms such as bacteria, algae, fungi (Krens et al., 1982; Kindle, 1990; Hanahan et al., 1991; Zhang et al., 2016). In this method, cells or protoplasts are induced to take up foreign DNA in the presence of PEG. To date, the underlying mechanism of PEG remains unclear and may vary between organisms. Some studies suggest that PEG destabilizes cell membrane, facilitating DNA entry into cells, whereas others propose that PEG increases the effective concentration of solutes, promoting DNA adsorption to the cell membrane followed by endocytosis (Lentz and Lee, 1999; Chen et al., 2008). Compared with other transformation methods, such as biolistic bombardment and electroporation, PEG-mediated transformation is generally inexpensive, technically simple, and time-efficient. It can yield relatively high numbers of transformants under optimized conditions, although its efficiency varies considerably among species and experimental systems. A major limitation is that PEG-mediated DNA uptake is most effective in cells lacking a rigid cell wall or in protoplasts, which restricts its application to species in which efficient cell wall removal and subsequent protoplast regeneration can be achieved (Davey et al., 2005). In addition, transformation efficiency is strongly influenced by the physicochemical properties of the PEG solution and treatment, including PEG molecular weight and concentration, pH, temperature, and incubation time. These parameters therefore need to be systematically optimized for each species.

PEG transformation has been used on many microalgae species (Table 2.1). The cell wall is the primary physical barrier to the entry of foreign DNA into cells. Its removal is critical for efficient PEG-mediated transformation. Protoplast release is influenced by multiple factors, including cell wall thickness, enzyme incubation conditions (temperature, duration, and pH), mechanical conditions such as gentle agitation, and the composition of the osmoticum (Sinha et al., 2003). As cell wall composition varies among species, it is necessary to optimize and evaluate cell wall

removal strategies for each species. For example, alcalase has been reported to be more effective for spheroplasting *P. tricornutum* cells than the cocktail of zymolyase, hemicellulase and lysozyme, resulting in a significant improvement in transformation efficiency (Karas et al., 2015; Walker et al., 2026). Once the cell wall has been removed, protoplasts become highly susceptible to rupture, shrinkage, and aggregation (Lai et al., 2020). Therefore, maintaining a stable osmotic pressure during and after enzymatic digestion is essential. Common osmoticum include sugar derivatives, such as sucrose, sorbitol, and mannitol, inorganic salts, such as KCl, NaCl, and $MgSO_4$, or combinations of these compounds (Chen et al., 2015; Yang et al., 2015). The suitable molecular weight of PEG for transformation varies between species (Table 2.1). It is necessary that testing the PEG conditions including molecular weight, concentration, incubation time prior transformation.

The nucleofection method was established in *M.commoda* (Faktorova et al., 2020). Nucleofection is an optimized electroporation. Electroporation relies on transient permeabilization of the cell membrane upon exposure to an electric field, allowing charged molecules such as DNA, RNA and protein to enter the cell (Weaver and Chizmadzhev, 1996; Kotnik et al., 2015). However, this method has some potential disadvantages such as relatively high cell mortality and low transfection efficiency. Nucleofection improves upon electroporation by employing optimized electrical parameters, including cell type-specific buffers and precisely controlled pulse programs, to facilitate the direct delivery of nucleic acids into the nucleus (Gresch et al., 2004). This approach has been shown to improve transfection efficiency and maintain higher cell viability across a wide range of cell types including microalgae (Chicaybam et al., 2013; Booth and King, 2020; Faktorova et al., 2020). In the transformation of *M. commoda*, two codon-optimized plasmids encoding luciferase gene and enhanced green fluorescent protein (*eGFP*) gene, separately, were delivered into cells by nucleofection (Faktorova et al., 2020). The expression of luciferase and *eGFP* was detected in transformants. However, because selectable marker was not introduced, transformants were subsequently overgrown by

wide-type (WT) cells. In addition, the Lonza nucleofector system is not commonly used for algal transformation, and the limited accessibility of required equipment further constrains its broader application.

Methods for delivering DNA to *T. pseudonana* include biolistic bombardment, electroporation, and bacterial conjugation (Poulsen et al., 2006; Karas et al., 2015; Messemer et al., 2026). These methods can achieve stable and efficient transformation, but have limitations. Biolistic bombardment and electroporation require specialized equipments. Bacterial contamination is the major concern in bacterial conjugation. Moreover, recent studies have reported unstable and unreliable reporter gene expression following conjugation, likely due to episomal rearrangements (Diamond et al., 2023). Therefore, the development of a simple, cost-effective, and robust transformation method remains an important objective for advancing genetic studies in *T. pseudonana*.

2A peptides are short self-cleaving peptides, typically 18–21 amino acids in length, that mediate the separation of polypeptides during translation in eukaryotic cells (Ahier and Jarriault, 2014). Commonly used 2A peptides include F2A from foot-and-mouth disease virus, E2A from equine rhinitis A virus, P2A from porcine teschovirus-1, and T2A from *Thosea asigna* virus (Ryan et al., 1991; Szymczak and Vignali, 2005). These peptides act through a unique translational mechanism known as "ribosomal skipping", in which peptide bond formation fails between the terminal glycine and proline residues of the 2A sequence, resulting in the production of two discrete proteins (Donnelly et al., 2001). However, incomplete ribosomal skipping can result in continued translation and the production of an uncleaved fusion protein. High cleavage efficiencies have been reported in systems such as silkworms, zebrafish, mice, and human cell lines, whereas substantially lower and variable efficiencies have been observed in microalgae (Kim et al., 2011; Wang et al., 2015; Plucinak et al., 2015; Koh et al., 2018; Wang et al., 2024b). By enabling the co-expression of multiple proteins at comparable levels from a single promoter, 2A peptides provide a potentially valuable strategy for reducing vector complexity and coordinating transgene expression (Minskaia et al., 2013; Liu et al.,

2017). Their performance in *M. commoda* has not been evaluated. Investigating the functionality of 2A peptides in *M. commoda* could facilitate the development of future multigene expression systems and genetic engineering applications in this species.

Table 2.1: PEG transformation on microalgae

	Species	Enzyme for digesting cell wall	PEG	Reference
Diatom	*Thalassiosira pseudonana*	Alcalase	20% PEG8000	(Walker et al., 2026)
	Phaeodactylum tricornutum	Alcalase	20% PEG8000	(Walker et al., 2026)
Coccolithophore	*Pleurochrysis carterae*	Proteinase K	20% PEG6000	(Endo et al., 2016)
Green algae	*Chlamydomonas reinhardtii*	–	20% PEG8000	(Díaz-Santos et al., 2016)
	Haematococcus pluvialis	Cocktail of cellulase and macerozyme R-10	20% PEG4000	(Guo et al., 2022)
	Chlorella ellipsoidea	Cocktail of cellulase, macerase and pectinase	20% PEG4000	(Kim et al., 2002)
	Chlorella sp. DT	Cocktail of macrease and cellulase	20% PEG6000	(Huang et al., 2006)
	Chlorella vulgaris	Cocktail of cellulase R-10 , Macerozyme R-10 and Pectinase	20% PEG4000	(Yang et al., 2015)
	Nannochloropsis sp.	Cocktail of cellulase R-10 , Macerozyme R-10 and Pectinase	20% PEG4000	(Zainal Abidin et al., 2020)
	Dunaliella salina	–	LiAc/35% PEG	(Chai et al., 2013)
	Ostreococcus tauri	–	30% PEG4000	(Sanchez et al., 2019)

2.2 Materials and Methods

2.2.1 Cell cultures for *Micromonas commoda* and *Thalassiosira pseudonana*

The marine green algae *Micromonas commoda* (RCC827) was obtained from the Roscoff Culture Collection (RCC). Cultures were maintained in modified K medium (1mM NO_3^-) prepared with artificial seawater at 35 PSU (practical salinity units) using Tropic Marin® Classic Sea Salt (Germany). The K nutrients were prepared according to the protocol of the National Center for Marine Algae and Microbiota (NCMA). Cultures were incubated at 20 °C under a 14:10 h light:dark cycle at an irradiance of 150 μmol photons m^{-2} s^{-1}. For genetic transformation experiments, K medium prepared with 22 PSU seawater was used.

Thalassiosira pseudonana (CCMP1335) was obtained from NCMA. Cells were cultured in synthetic ocean water (SOW) supplemented with f/2 nutrients. SOW was prepared according to the NCMA Aquil medium protocol, and f/2 nutrients were prepared following the NCMA protocol. Cultures were maintained at 20 °C under continuous light at an irradiance of 50 μmol photons m^{-2} s^{-1}.

2.2.2 Growth Curves

M. commoda cell density was counted using a hemacytometer under a light microscope at 20× magnification. Cell growth was also monitored by measuring chlorophyll a (Chl a) fluorescence with 460 nm excitation and 680 nm emission using a SpectraMax® iD3 Multi-Mode Microplate Reader. A standard curve was generated relating Chl a fluorescence to cell density (Figure 2.1). Each sample was measured in triplicate to ensure accuracy and reproducibility. The resulting standard curve provides an equation that describes the relationship between fluorescence intensity and cell density, enabling direct calculation of cell density from fluorescence measurements obtained using a plate reader. As shown in

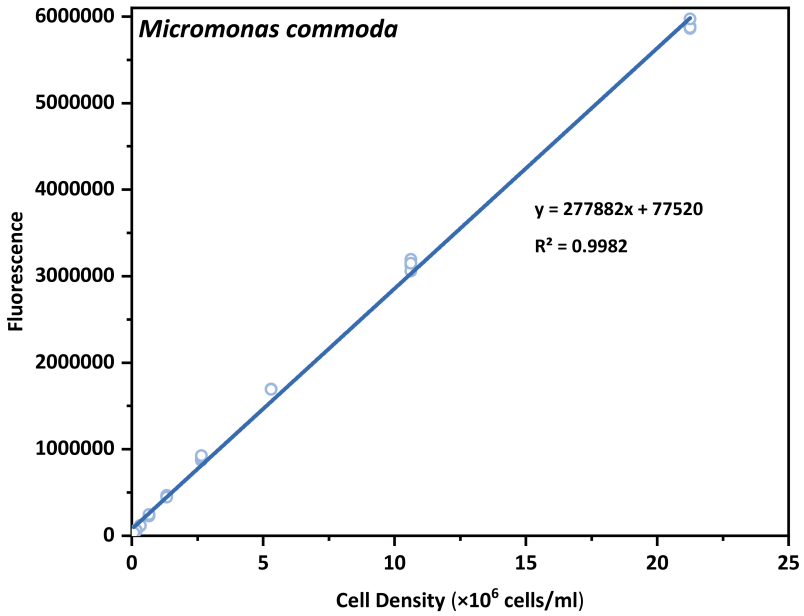


Figure 2.1: Standard curve of *M. commoda*.

Figure 2.1, the fitted equation is: y = 277882x + 77520.

T. pseudonana cell density was monitored using a Beckman Coulter Counter M4.

Growth rates (μ, d^{-1}) were calculated using the endpoints of the fluorescence data or cell density, and the following equation:

$$\mu = \frac{\ln(N_1) - \ln(N_0)}{t_1 - t_0}, \tag{2.1}$$

where N_1 and N_0 are the fluorescence values/cell density at timepoints t_1 and t_0, respectively. The difference in time $(t_1 - t_0)$ is expressed in days.

One millilitre (ml) sample of each biological replicate was used to measure the maximum quantum yield of photosystem II (Fv/Fm) with a Walz Phyto-Pam Phytoplankton Analyser fluorometer (Walz Gmbh, Effeltrich, Germany) using the Phyto-ED component. Fv/Fm data provides a quantification of photosynthetic efficiency of *in vivo* algae cultures by excitation of chlorophyll at different wavelengths and using the resulting data in the following equation:

$$F_V/F_m = \frac{F_m - F_0}{F_m} \tag{2.2}$$

Where F_m is the maximum fluorescence yield, and F_0 is the minimum fluorescence yield (Maxwell and Johnson, 2000).

2.2.2.1 Antibiotics test

To identify the suitable types and concentrations of antibiotics used in selective medium for *M. commoda*, five antibiotics including nourseothricin (Jena Bioscience, AB-102L), zeocin (Thermo Fisher, R25001), chloramphenicol (Sigma-Aldrich, C0378-5G), blasticidin S (Sigma-Aldrich, 203350-25MG), and cycloheximide (Sigma-Aldrich, 239763-1GM) were tested. *M. commoda* cultures in the exponential growth phase were collected and resuspended in K medium (35 PSU) containing different antibiotics. Antibiotics were prepared as stock solutions prior to use. The stock and working concentrations used in medium are shown in Table 2.2. The initial cell density for zeocin, chloramphenicol, blasticidin S, and cycloheximide treatments was set at $\sim$1.3 $\times$ 10^6 cells/mL (Chl a fluorescence was 443,224). The initial cell density for nourseothricin treatment was set at $\sim$3 $\times$ 10^6 cells/mL (Chl a fluorescence was 1,093,951). Cultures were incubated at 20 °C under a 14:10 h light:dark cycle at an irradiance of 150 μmol photons m^{-2} s^{-1} for 13 days.

2.2.3 Extraction of *M. commoda* genomic DNA

M. commoda was grown to exponential phase ($\sim$1x10^7 cells/ml) and its genomic DNA (gDNA) extracted using the Qiagen MagAttract HMW DNA Kit (Qiagen, Catalogue # 67563) according to the manufacturer's protocol.

Table 2.2: Antibiotics and their concentrations

Antibiotics	Stock Concentration (mg/ml)	Working concentration (μg/ml)
Zeocin	100	0, 25, 50, 100
Chloramphenicol	100	0. 100. 200, 300
Nourseothricin	100	0. 50, 100, 150
Blasticidin S	10	0, 5, 10, 20
Cycloheximide	100	0. 20. 50

2.2.4 Construct for *eYFP* and *shble* expression in *M.commoda*

2.2.4.1 Determining codon bias compared to *eYFP* and *shble* gene

M. commoda reference transcripts (ASM9098v2) were downloaded from National Center for Biotechnology Information (NCBI) datasets. Subsequently, emboss cusp (https://www.bioinformatics.nl/cgi-bin/emboss/cusp) was used to calculate codon usage for *M. commoda*, enhanced yellow fluorescent protein (*eYFP*) and *shBle* genes. Usage was expressed as a fraction (frequency of one codon/frequency of all codons for a specific amino acid).

2.2.4.2 Plasmid construction using golden gate cloning

Golden Gate cloning was carried out according to Weber et al. (2011) and Hopes et al. (2016). DNA fragments were flanked by specific type IIS restriction enzyme recognition sites and defined overhangs, enabling their directional and seamless assembly into a defined final construct. Golden gate cloning mainly have three assemblies. In Level 0 (L0) assembly, specific DNA elements, including promoters, coding sequences, and terminators, are amplified using primers containing BsaI restriction sites and defined overhangs, and then cloned into a TOPO vector. Level 1 (L1) assembly is assembling L0 modules into an L1 backbone with BbsI restriction sites to form a promoter–gene–terminator cassette, enabling expression of a single gene. Level 2 (L2) assembly is the final step that multiple L1 modules

are combined into an L2 backbone, generating an integrated plasmid capable of expressing all target genes according to the designed construct. The plasmid for transforming *M. commoda* was constructed using Golden Gate Cloning (Figure 2.2).

L0 assembly

The endogenous RPS9 promoter and terminator were selected according to Faktorova et al. (2020). Nested PCR was used to amplify the RPS9 promoter and terminator. First, the full RPS9 gene, comprising the promoter, coding sequence, and terminator, was amplified from gDNA using Phusion DNA polymerase (NEB) with the following primers: 5′-AGGTAGCTTGGTCAACATTG-3′ and 5′-AGGGATGTCAAGTCTTTGC-3′. This amplified fragment used as the template for subsequent amplification of the RPS9 promoter and terminator. The *shble* gene was amplified using PCR from pICH47732:FCP:*ShBle*, and the *eYFP* gene was PCR-amplified from plasmid pICH47742:FCP:Cas9YFP. Primers for amplification are listed in Table 2.3. All sequences were cloned into the pCR4-TOPO vector (Thermo Fisher). Five µl of the reaction was transformed into 50 µl of NEB 5-alpha chemically competent *Escherichia coli* for further cloning and extraction. Constructs were screened by digestion with EcoRI and then sequenced using two M13 primers 5′-GTAAAACGACGGCCAGT-3′ and 5′-CAGGAAACAGCTATGAC-3′.

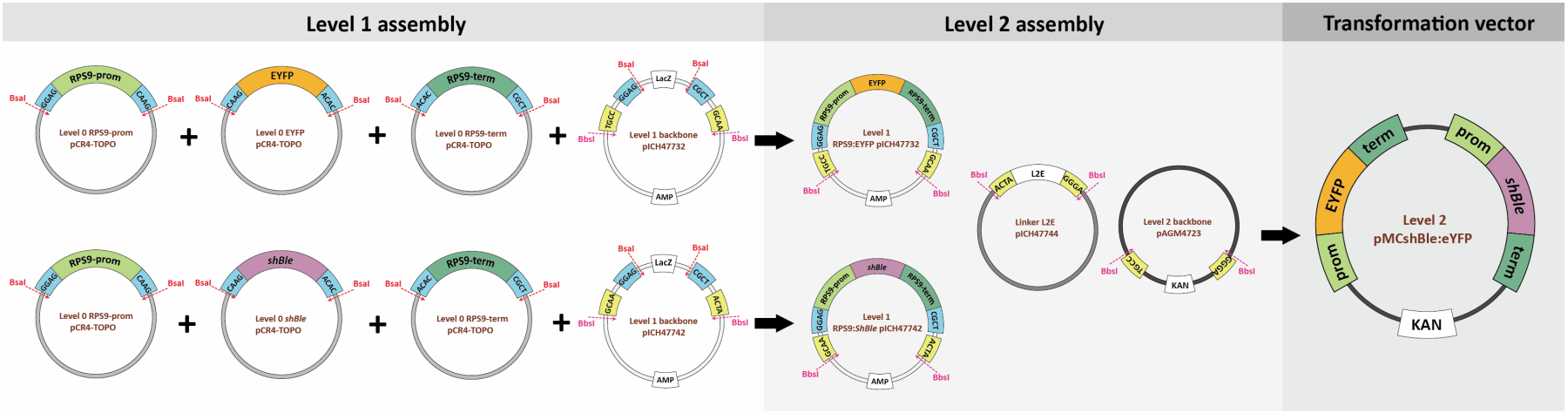


Figure 2.2: Overview of Golden Gate cloning assembly for transformation vector used in transforming *M. commoda*. L0 modules are created by PCR and then cloned into pCR4-TOPO vectors. Key components include the promoter (RPS9-prom) and terminator (RPS9-term), ampicillin resistance gene (AMP), kanamycin resistance gene (KAN), and target gene fragment *eYFP* and *shBle*. Specific 4nt overhangs are shown in different colors. The enzyme cleavage sites were marked as pink arrows.

Table 2.3: Primers used for amplification of fragments for L0 assembly. Uppercase letters indicate the original sequence, lowercase letters indicate the BsaI recognition site and defined overhangs, and the underlined sequences represent the defined overhangs.

Primer	Sequence	Annealing temperature (°C)
RPS9prom F	ggtctcaggagTCGTGGATTTTTGACAAAA	62
RPS9prom R	ggtctcaCTTGGCTGTCCCTTCGCCT	67
RPS9term F	ggtctcaACACCCGAGCTGAGGGC	66
RPS9term R	ggtctcaagcgCTCGCGCGGGTTCGTTT	71
shBle F	ggtctcacaagATGGCCAAGTTGACCAGT	65
shBle R	ggtctcagtgtTCAGTCCTGCTCCTCGGC	68
eYFP F	ggtctcaggtcATGGTGAGCAAGGGCGA	68
eYFP R	ggtctcagtgtTCACTTGTACAGCTCGTCCAT	65

L1 assembly

The Level 1 pICH47732 backbone was used to assemble the *eYFP* cassettes with the RPS9 promoter and RPS9 terminator, resulting in the plasmid RPS9:*eYFP* pICH47732. The antibiotic resistance cassette was assembled with three L0 modules—RPS9 promoter, *shBle*, and RPS9 terminator—into a L1-pICH47742 backbone, generating RPS9:*shBle*-pICH47742.

Golden Gate reactions for L1 and L2 assembly were adapted from the protocol by Hopes et al. (2016). Forty fmol of each component was mixed with 100 units of BsaI–HF v2 or BbsI–HF (NEB), 100 units of T4 DNA ligase (NEB) and 1× ligation buffer in 20μL reaction. The reaction was incubated at 37 °C for 5h, following by 50 °C for 5min and 80 °C for 10min. Five µl of the reaction was transformed into 50 μl of 10-beta competent *E. coli* for further cloning and extraction. Constructs were screened by digestion with EcoRI and sequenced using forward primer 5′-CCCACTCTGTGAAGACAA-3′ and reverse primer 5′-GCCAATATATCCTGTCAAACAC-3′.

L2 assembly

For L2 assembly, the L1 modules RPS9:*eYFP* (pICH47732) and RPS9:*shBle* (pICH47742) were assembled with the L2E linker (pICH47744) and the L2 destination vector pAGM4723, resulting in transformation plasmid pMC*shBle*:*eYFP* (Figure 2.3a). Constructs were screened by forward primer 5′-ACTACCTGAGCTACCAGTCC-3′ and reverse primer 5′-GCCAATATATCCTGTCAAACAC-3′, and then verified using whole-plasmid sequencing by Eurofins Genomics.

2.2.5 Construct for 2A self-cleaving peptides expression in *M.commoda*

To verify the cleavage efficiency of 2A self-cleaving peptides in *M. commoda*, four commonly used 2A peptides were selecte, including P2A, T2A, E2A and F2A. All four 2A sequences were codon-optimized, and a glycine–serine– glycine (GSG) spacer was fused to the N-terminus of each 2A peptide for better cleavage efficiency (Table 2.4 and 2.5). The 2A oligonucleotides were synthesized by Merck (United Kingdom).

L0 assembly

Touchdown PCR was used to generate double-stranded 2A DNA fragments using the forward primer 5′-ggtctcactgaGGCTCGGGC-3′ and the reverse primer 5′-ggtctcaccatCGGGCCCG-3′. PCR products were purified using the Monarch® Spin PCR & DNA Cleanup Kit (NEB).

The *shBle*(2A) fragment was PCR-amplified from L0-*shBle* using forward primer 5′-ggtctcacaagATGGCCAAGTTGACCAGT-3′ and reverse primer 5′-ggtctcaTCAGTCCTGCTCCTCGGC-3′. The *eYFP*(2A) fragment was amplified by PCR from L0-*eYFP* using forward primer 5′-ggtctcaATGGTGAGCAAGGGCGA-3′ and reverse primer 5′-ggtctcaTCAGTCCTGCTCCTCGGC-3′. The PCR products were then cloned

in the pCR4-TOPO vector and designated as L0-*shBle*2A and L0-*eYFP*2A.

L1 assembly

L0-RPS9-p, L0-RPS9-t, L0-*shBle*2A, L0-*eYFP*2A, and the 2A DNA fragments were assembled into the L1 backbone pICH47742, generating four plasmids containing different 2A linkers. The stop codon of *shBle* was subsequently removed using the Q5 site-directed mutagenesis (SDM) kit (NEB) with the forward primer 5′-GGCTCGGGCCAGTGCACGAACTA-3′ and the reverse primer 5′-GTCCTGCTCCTCGGCCACGAAGT-3′. The plasmid for verifying functions of 2A peptide is shown in Figure (2.3b).

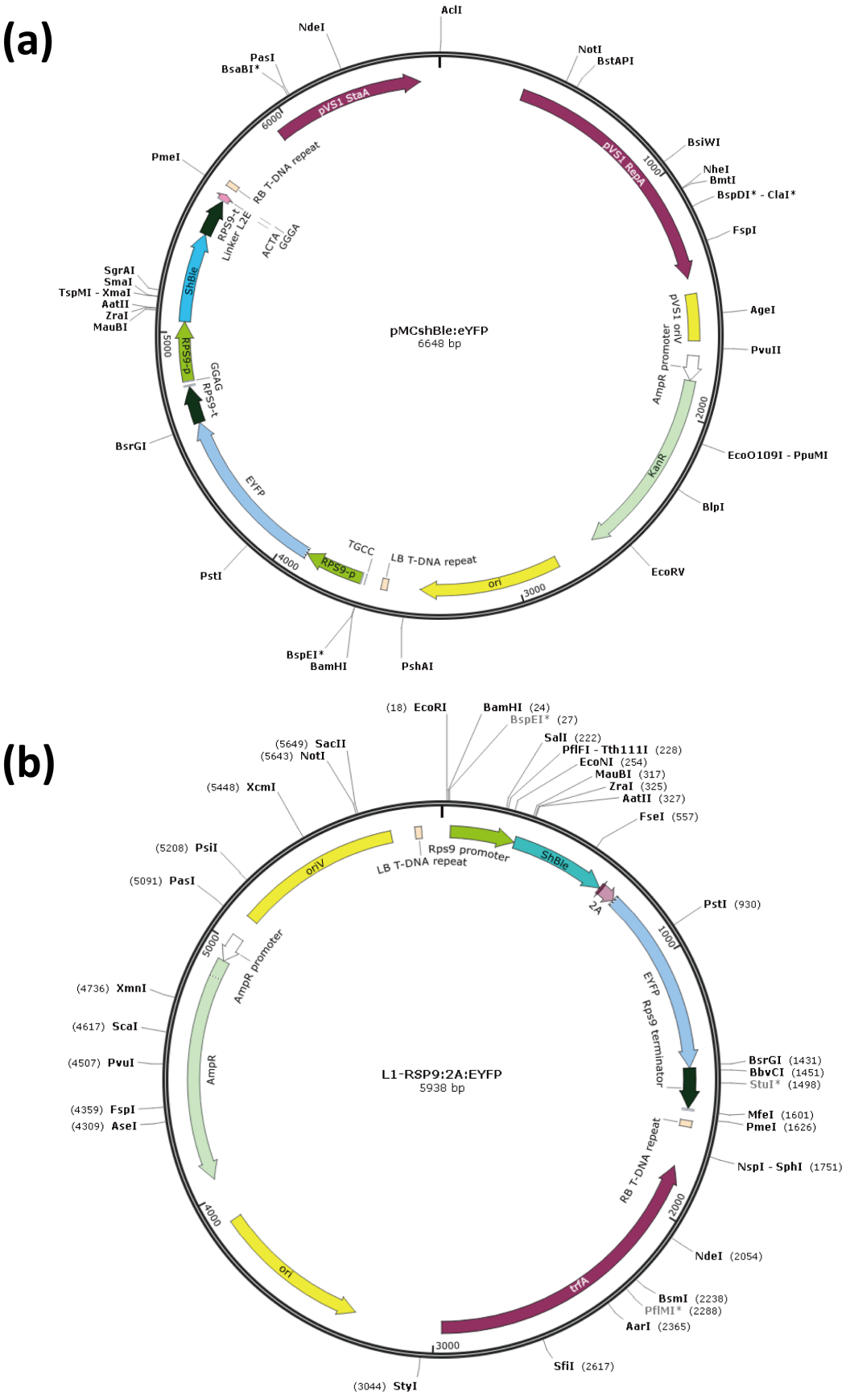


Figure 2.3: Plasmid maps of pMCshBle:*eYFP* (a) and L1-RPS9:2A:*eYFP* (b)

Table 2.4: Amino acid sequences of 2A peptides

2A	Amino acid sequence
T2A	GSGEGRGSLLTCGDVEENPGP
E2A	GSGQCTNYALLKLAGDVESNPGP
F2A	GSGVKQTLNFDLLKLAGDVESNPGP
P2A	GSGATNFSLLKQAGDVEENPGP

Table 2.5: Sequences of codon optimized 2A peptides

2A	Sequence
T2A	5′-GGCTCGGGCGAGGGCCGCGGCTCGCTCCTCACGTGCGGCGACGTCGAGGAGAACCCGGGCCCG-3′
E2A	5′-GGCTCGGGCCAGTGCACGAACTACGCGCTCCTCAAGCTCGCGGGCGACGTGGAGTCGAACCCGGGCCCG-3′
F2A	5′-GGCTCGGGCGTGAAGCAGACGCTCAACTTCGACCTCCTCAAGCTCGCGGGCGACGTGGAGTCGAACCCGGGCCCG-3′
P2A	5′-GGCTCGGGCGCGACGAACTTCTCGCTCCTCAAGCAGGCGGGCGACGTGGAGGAGAACCCGGGCCCG-3′

2.2.6 PEG-mediated transformation

2.2.6.1 *Micromonas commoda*

The plasmid was extracted using the Monarch® Plasmid Miniprep Kit. Plasmids were linearized with AclI at 37 °C for 4 h and purified using the QIAquick PCR Purification Kit (QIAGEN).

M. commoda cells at mid-exponential phase (approximately 1.5×10^7 cells/mL) were collected by centrifugation (Sigma 4K15) at 3000 rpm for 8 min. The supernatant was carefully removed, and the cell pellet was resuspended to a final volume of 500 μL. Next, 5 μL of single-stranded salmon sperm DNA (ssssDNA, Sigma-Aldrich D7656) was added. The ssssDNA was denatured at 95 °C for 5 min and cooled on ice to maintain single-strandedness previously.

Then, 5 μg of plasmid DNA (circular or linear) and 500 μL of 0.22 μm-filtered 40% PEG-4000 solution (w/v; containing 50 mM $CaCl_2$ and 0.8 mM NaCl, equilibrated to 30 °C) were added. The cells:DNA:PEG mixture was gently mixed by inversion and incubated at room temperature for 2–10 min. Following incubation, cells were centrifuged at 4000 rpm for 3 min using Eppendorf Centrifuge 5425R. The supernatant was removed, and cells were resuspended in 15 mL of K medium (22 PSU seawater) in a Corning® 25 cm^2 cell culture flask. Cells were recovered under low light intensity (5–10 μmol photons m^{-2} s^{-1}) at 20 °C for 36–48 h.

After recovery, cell density was measured, and $\sim 7 \times 10^7$ cells (1.5–2 mL) were embedded into 20 mL of 0.3% low-gelling agarose plates (Merck, A4018) supplemented with K medium and zeocin (25 μg/mL). The final agarose concentration was 0.27–0.28%. Plates were incubated at 20 °C under a 14:10 h light:dark cycle with an irradiance of 150 μmol photons m^{-2} s^{-1} for at least two weeks until colonies appeared.

2.2.6.2 *Thalassiosira pseudonana*

Approximately 1×10^5 cells/mL of *T. pseudonana* were inoculated into f/2 medium. After 4 days of growth (late exponential phase), cells were collected by centrifugation at 4000 rpm for 10 min (Sigma 4K15) and resuspended at a final concentration of 3×10^8 cells/mL.

For protoplasting, 1 mL of cells was transferred to a 50 mL Falcon tube containing 9 mL of L1 medium and 1000 μL of alcalase (Sigma-Aldrich, 126741). The mixture was placed on a slow rocking shaker ($\sim$15 rpm) and incubated at room temperature for varying incubation times.

After incubation, 40 mL of f/2 medium was gently added and the cells were pelleted by centrifugation at 3665 rpm for 5 min at 18 °C. The supernatant was carefully decanted, and the cells were gently resuspended in 500 μL of f/2 medium followed by the addition of 40 μg salmon sperm DNA (Invitrogen, 15632011). For all subsequent centrifugation steps, pellets were gently resuspended to avoid mechanical shearing of the protoplasts. Next, 250 μL of protoplasts was combined with 500 ng of the 8.2 kb episome pBIG1 and 1 mL of 20% PEG-8000 (w/v; containing 10 mM Tris pH 8, 10 mM $CaCl_2$, and 2.5 mM $MgCl_2$) equilibrated to 30 °C. The pBIG1 plasmid is provided by Dr Krisztina Sarkozi. The protoplast:DNA:PEG suspension was gently mixed by inversion and incubated at room temperature for 30 min.

Following PEG-mediated transformation, cells were transferred to 10 mL of f/2 medium and allowed to recover for 2 days prior to plating. After recovery, cells were centrifuged for 5 min at 4000 rpm (Sigma 4K15), resuspended in 500 μL of f/2 medium, and 250 μL of suspension was plated onto full-salt f/2 agarose plates supplemented with nourseothricin (100 μg/mL).

2.2.7 Screening colonies and cultures

2.2.7.1 *Micromonas commoda* transformants

Two mL *M. commoda* culture was collected by centrifugation, and the supernatant was removed. Cell pellets were resuspended in 20 μL of lysis buffer (10% Triton X-100, 20 mM Tris–HCl, pH 8, 10 mM EDTA), stored at −80°C for 15 min, and then incubated at 95 °C for 10 min.

The presence of *eYFP* was screened by PCR using primers 5′-CGAGGAGCTGTTCACCGGG-3′ and 5′-AACTCCAGCAGGACCATGTG-3′, and *shBle* was screened using Phusion High-Fidelity DNA Polymerase (NEB) with forward primer 5′-AAGTTGACCAGTGCCGTTC-3′ and reverse primer 5′-TCCCGGAAGTTCGTGGACA-3′.

2.2.7.2 *Thalassiosira pseudonana* transformants

Four mL *T. pseudonana* culture was collected by centrifugation, and the supernatant was removed. Cell pellets were resuspended in 20 μL of lysis buffer (10% Triton X-100, 20 mM Tris–HCl, pH 8, 10 mM EDTA), stored at −80°C for 15 min, and then incubated at 95 °C for 10 min. The presence of *eGFP* was screened by PCR using MyTaq Redmix (Meridian Bioscience) with forward primer 5′-ATGGCCGACAAGCAGAAGAA-3′ and reverse primer 5′-AACTCCAGCAGGACCATGTG-3′.

2.2.8 Microscopy

Cell fluorescence images were acquired using an inverted Zeiss Observer 7 microscope equipped with a 63× oil immersion objective (Carl Zeiss AG, Jena). The microscope was configured for multitrack, line-by-line acquisition to simultaneously capture brightfield, chlorophyll, and *eYFP* fluorescence.

Specifically, *eYFP* fluorescence was detected using a 488 nm bandpass filter, *eGFP* fluorescence was detected using a 550 nm bandpass filter, and chlorophyll autofluorescence was detected using a 610 nm bandpass filter.

2.2.9 Western blot for protein expression analysis

M. commoda wild-type and transformant cultures (50 mL, mid-exponential phase) were harvested by centrifugation. Pellets were resuspended in 50μL lysis buffer (50 mM Tris-HCl, pH 6.8, 2% SDS) containing 1× protease inhibitor (Roche), briefly vortexed, and incubated at room temperature for 30 min. Samples were then centrifuged at 13,000 rpm for 30 min at 4 °C to remove cell debris. The supernatants were transferred to clean 1.5 mL Eppendorf tubes and kept on ice. Protein concentrations were determined using the Pierce BCA Protein Assay Kit (Thermo Scientific). Absorbance was measured at 562 nm using a SpectraMax iD3 plate reader (Molecular Devices).

Denaturing protein gel electrophoresis was performed using NuPAGE Bis-Tris Mini Gels (Thermo Scientific) following the manufacturer's protocol for reduced samples with NuPAGE MES running buffer. For each sample, 40μg of extracted protein was mixed with reducing agent and incubated at 70 °C for 10 min. Protein samples (40 μg per lane) were loaded onto the gel together with a dual-color protein standard (10–250 kDa, Bio-Rad) and separated using an XCell SureLock Mini-Cell system for 40 min at 200 V. Proteins were transferred onto a PVDF membrane (Amersham, 0.2 μm) at 41 V for 30 min using 1× transfer buffer (25 mM Tris, 192 mM glycine, 10% methanol) supplemented with 0.1% NuPAGE Antioxidant. Protein transfer was visualized with Ponceau S staining to confirm protein loading. Membranes were then blocked with 5% milk in TBST (1× TBS with 0.1% Tween-20) for 1 h at room temperature.

After blocking, membranes were incubated with *eGFP* antiserum diluted 1:1000 in 5% milk TBST for 1 h at room temperature. Membranes were then washed three times with 1× TBST for 10 min each with gentle agitation at room temperature.

A horseradish peroxidase (HRP)-conjugated goat anti-rabbit secondary antibody (Promega) diluted 1:2500 in fresh 5% milk TBST was applied for 40–60 min. Blots were subsequently incubated with TMB substrate solution (Thermo Scientific) for 3 min and imaged to visualize chemiluminescent signals.

2.3 Results

2.3.1 *Micromonas commoda*

2.3.1.1 Antibiotics test

The sensitivity of *M. commoda* to five antibiotics was tested, including nourseothricin, cycloheximide, chloramphenicol, zeocin and blasticidin S (Figure 2.4). Under nourseothricin treatment, the initial cell density was set at 3×10^6 cells/mL. Compared with the positive control group, cells did not survive at 150 μg/mL nourseothricin after 10 days of incubation. No growth was observed at 100 μg/mL, although cells remained viable, and only limited inhibition was observed at 50 μg/mL (Figure 2.4a).

The initial cell density was set at 1.4×10^6 cells/mL under treatments with cycloheximide, chloramphenicol, zeocin, and blasticidin S (Figure 2.4 b–f). All treatment groups showed significant differences compared with the control group ($p < 0.01$, t-test). Under cycloheximide, chloramphenicol, and zeocin treatments, cells did not survive at all tested concentrations after 4, 6, and 9 days of incubation, respectively. Blasticidin S significantly inhibited cell growth prior to 9 days of incubation, but the growth began to recover by day 13.

2.3.1.2 Codon bias

To ensure efficient expression of exogenous *shBle* and *eYFP* genes, codon usage in *M. commoda* was analysed and compared with that of the *shBle* and *eYFP* genes

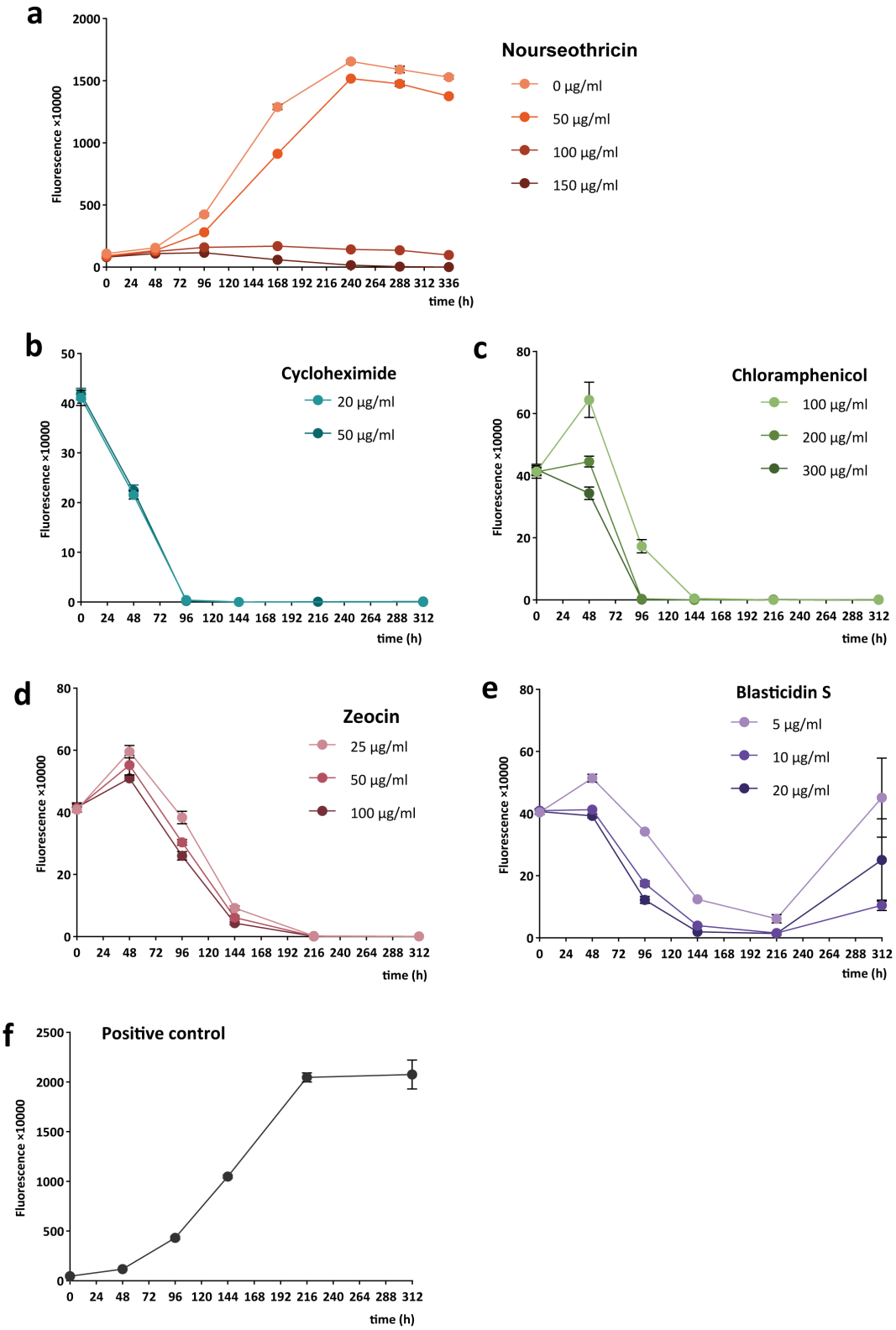


Figure 2.4: Antibiotic sensitivity of *M. commoda* in nourseothricin (a), cycloheximide (b), chloramphenicol (c), zeocin (d), blasticidin S (e) and postive control (f). Error bar represents mean ± SD (n = 3).

from pICH47732:FCP:*shBle* (for *Fragilariopsis cylindrus*) and pICH47742:FCP:Cas9YFP (for *T. pseudonana*), separately. Most codons used by *shBle* and *eYFP* are well represented in the *M. commoda* transcriptome across expressed genes (Figure 2.5). The expression levels under the RPS9 promoter are considered sufficient to confer resistance to zeocin and to produce a strong *eYFP* signal.

2.3.1.3 Golden Gate cloning

Golden Gate cloning was employed to construct plasmids to establish the *M. commoda* transformation system and verify functions of 2A self-cleaving peptides in *M. commoda*. By incorporating unique 4-base overhangs, multiple gene fragments with complementary overhangs can be precisely assembled and extraneous sequences can be eliminated, thereby enabling seamless ligation.

pMC*shBle*:*eYFP* construct

The plasmid co-expressing the *shBle* resistance gene and the *eYFP* reporter gene in *M. commoda* was successfully constructed and validated (Figure 2.6). The L1 assembly was designed to combine the promoter, gene of interest, and terminator into a single expression cassette, which was subsequently inserted into the L1 vector backbone. These vector backbones contain the *LacZ* gene to facilitate blue-white screening; successful assemblies disrupt *LacZ*, resulting in white colonies. The use of unique palindromic overhang sequences further minimized the likelihood of vector self-ligation and incorrect fragment assembly.

Specifically, the RPS9 promoter, *eYFP* gene, and RPS9 terminator were assembled into the vector backbone pICH47732 to generate L1-RPS9:*eYFP*. Similarly, the RPS9 promoter, *shBle* resistance gene, and RPS9 terminator were assembled into pICH47742 to produce L1-RPS9:*shBle*. Seven colonies containing L1-RPS9:*shBle* were selected and initially screened by restriction digestion using EcoRI, all of which showed the expected fragment size (5152 bp). Colonies 1–3 were

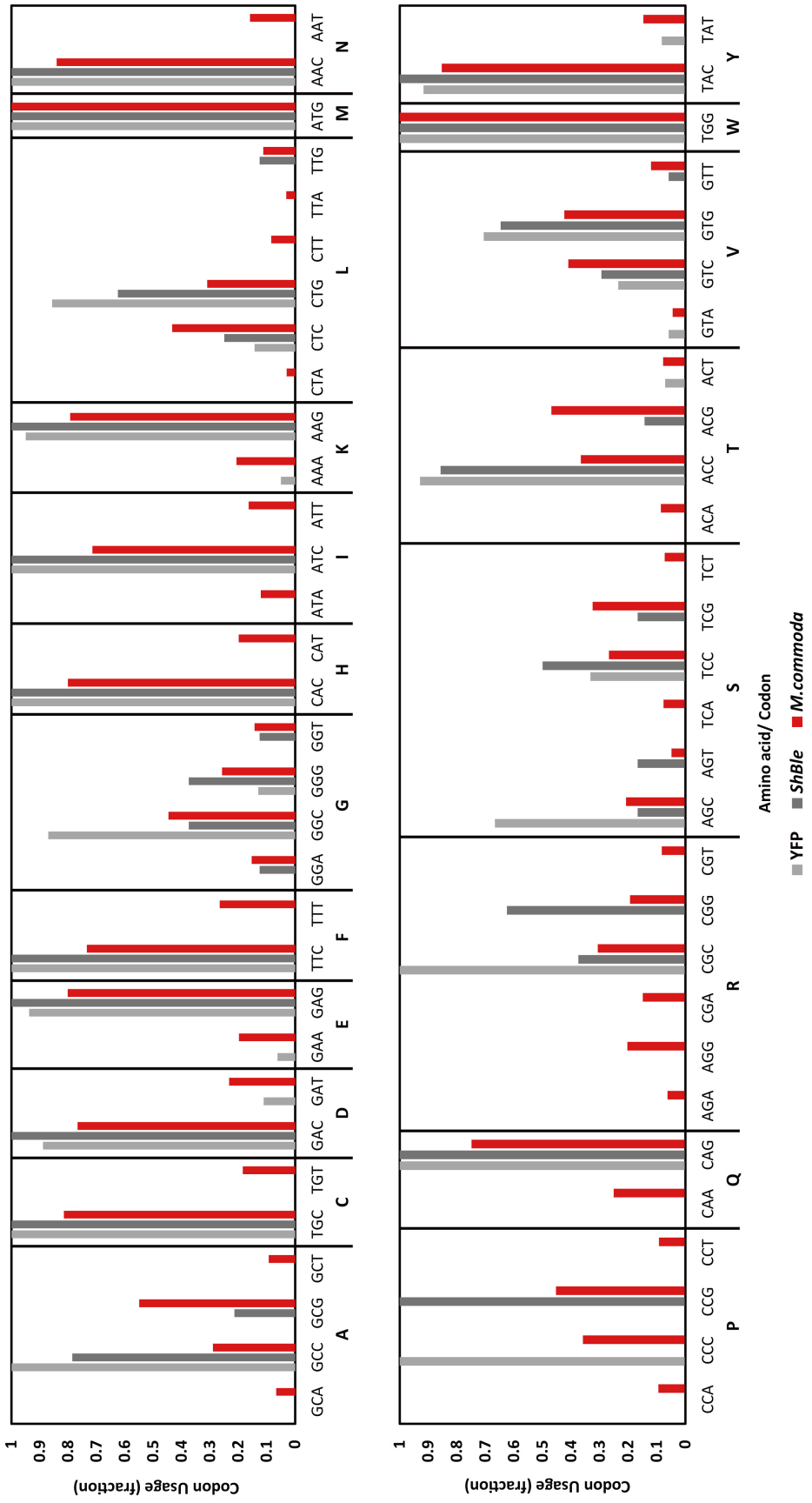


Figure 2.5: Comparison of codon usage. Codon usage as a fraction of triplets used per amino acid. Gene in *M.commoda* (red), *shBle* gene (dark gray), and enhanced yellow fluorescent protein gene (*eYFP*, light gray).

subsequently sequenced, and only L1-RPS9:*shBle*-1 was confirmed to have the correct sequence. In parallel, fourteen colonies containing L1-RPS9:*eYFP* were screened by PCR using primers targeting the vector backbone, designed to amplify the full insert. All samples yielded the expected product size (1237 bp). Colonies 1–3 were then sequenced, and only L1-RPS9:*eYFP*-2 was verified to be correct.

Subsequently, the validated L1 cassettes, together with the modular end linker L2E (pICH41744), were assembled into the L2 backbone pAGM4723, resulting in the final construct pMCshBle:*eYFP*. The L2 backbone pAGM4723 contains a red chromogenic marker (Cred), allowing visual selection. Colonies derived from uncut or unmodified templates retain the Cred gene and appear orange-red, whereas correctly assembled colonies appear white. Nine white colonies were selected and screened by PCR targeting part of the inserted cassette, yielding the expected fragment size (1156 bp). Colonies 1 and 2 were sequenced, and both were confirmed to contain the correct final construct.

RPS9:*shBle*:2A:*eYFP* constructs

To determine whether 2A self-cleaving peptides are functional in *M. commoda* and to identify the most efficient variant, a series of RPS9:*shBle*:2A:*eYFP* constructs were generated. In these constructs, the selective marker gene *shBle* and the fluorescent reporter gene *eYFP* were linked via four different 2A peptides (P2A, T2A, E2A, and F2A). Codon optimization of 2A peptide sequences and GSG spacer was performed based on the codon usage bias of *M. commoda* (Figure 2.5). Although the precise mechanism of the GSG spacer remains unclear, its inclusion at the N-terminus has been reported to enhance cleavage efficiency across multiple host systems (Szymczak et al., 2004; Kim et al., 2011; Wang et al., 2015).

The plasmids were constructed using L1 assembly, in which the RPS9 promoter, *shBle* gene, individual 2A sequences, *eYFP* gene, and the RPS9 terminator were sequentially assembled into the vector backbone pICH47732, generating L1-RPS9:*shBle*:2A:*eYFP* constructs. For each 2A construct, six colonies were selected and screened by PCR using primers targeting the vector backbone to

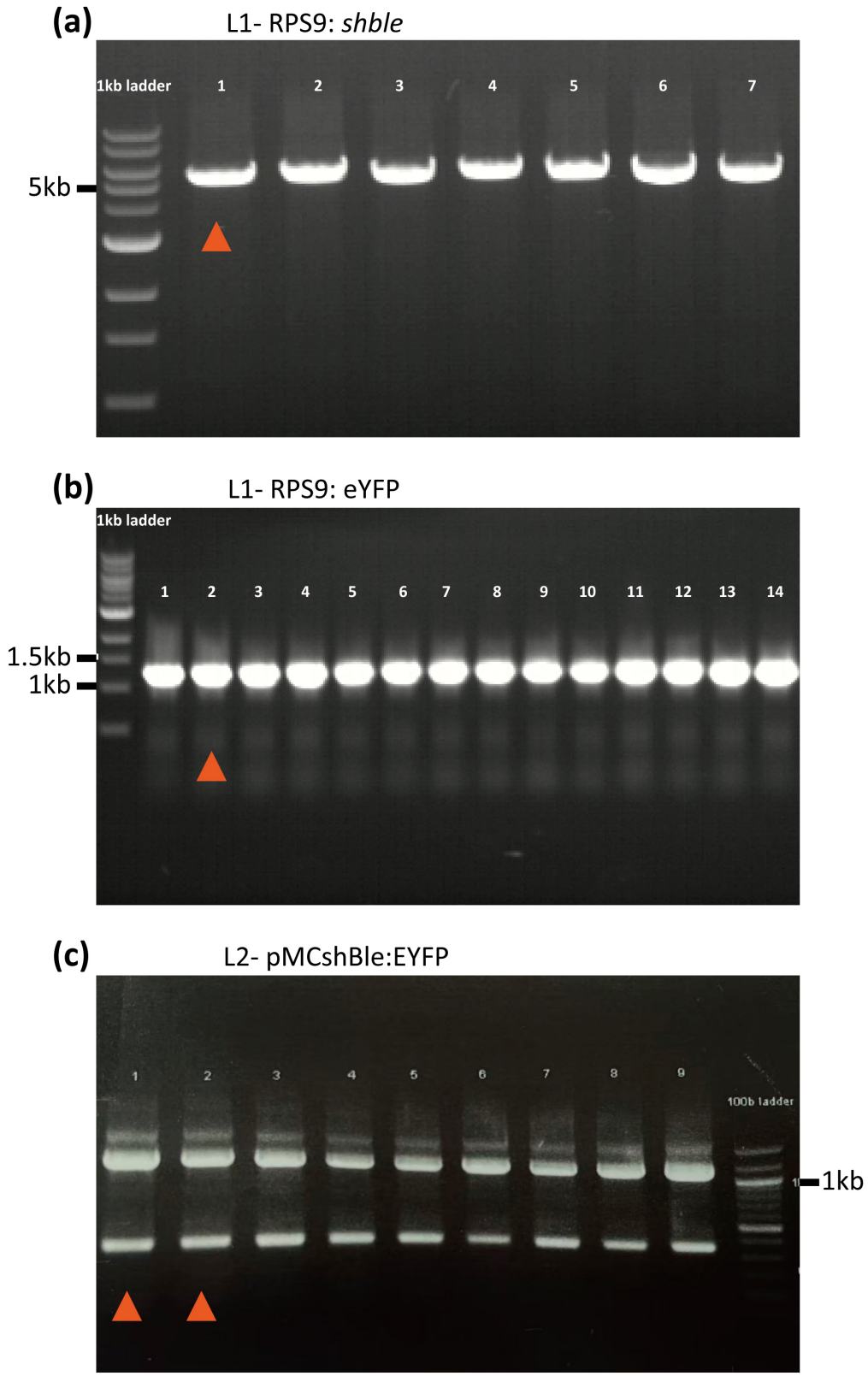


Figure 2.6: Restriction enzyme digestion and PCR analyses of L1 and L2 constructs of pMC*shBle*:*eYFP*. (a) Restriction digestion of L1-RPS9:*shBle* using EcoRI; (b) PCR screening of *E. coli* colonies carrying L1-RPS9:*eYFP*; (c) PCR screening of *E. coli* colonies carrying the L2 construct pMC*shBle*:*eYFP*. Orange triangles indicate the correct constructs as confirmed by Sanger sequencing.

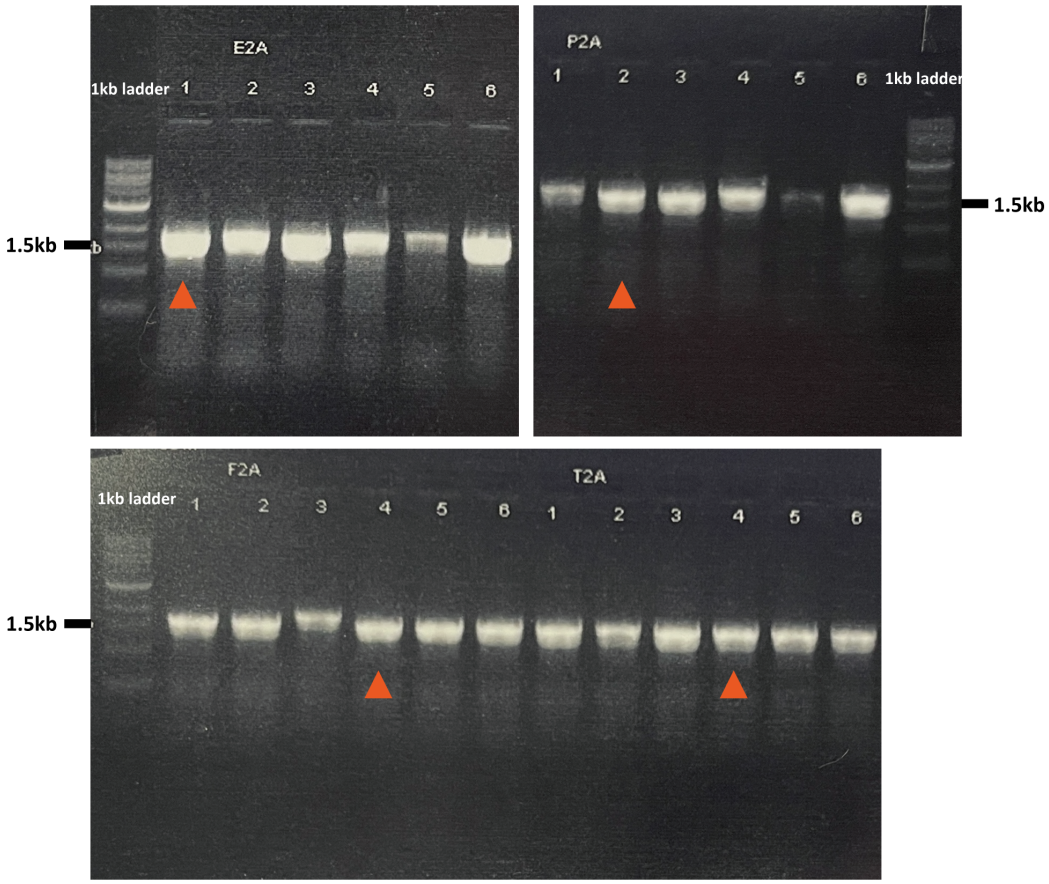


Figure 2.7: PCR analyses of four RPS9:*shBle*:2A:eYFP plasmids. Orange triangles indicate the correct constructs as confirmed by Sanger sequencing.

amplify the full insert. The majority of samples produced the expected fragment size (1656–1668 bp). Subsequently, two to three colonies per construct were sequenced, of which one clone for each 2A construct was confirmed to have the correct sequence (Figure 2.7).

2.3.1.4 PEG-mediated transformation and screening of transformants

Cell sensitivity to PEG concentration and incubation time was evaluated prior to transformation (Table 2.6). Two PEG molecular weights, PEG4000 and PEG8000, were selected for comparison. At a concentration of 20%, cells remained viable when the incubation time was shorter than 10 min, whereas complete loss of viability was observed after 15 min, with no recovery. Under identical conditions, 20% PEG 8000

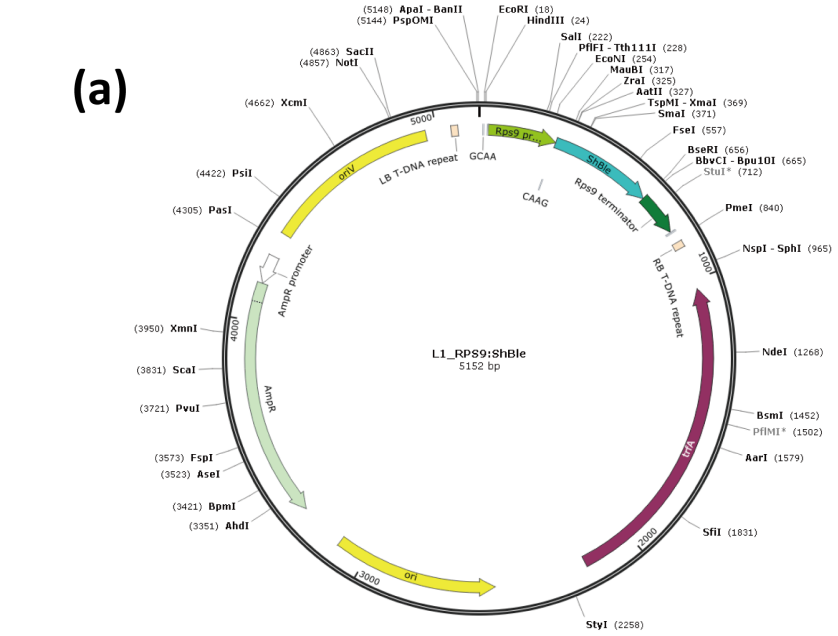
Table 2.6: *Micromonas commoda* sensitivity to PEG concentration at different incubation times. "+" means cells can survive. "-" means cells can not survive.

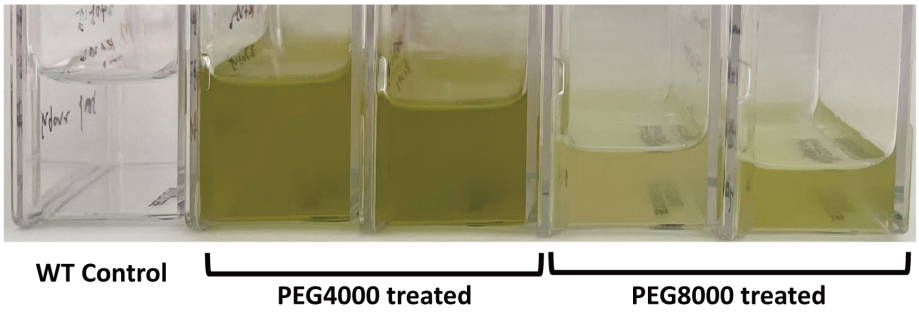
	2min	6min	10min	15min
20% PEG 4000 solution	+	+	+	–
30% PEG 4000 solution	– (recover after two weeks)	–	–	–
20% PEG 8000 solution	+	+	+	–

exhibited higher cytotoxicity than 20% PEG4000. Exposure to 30% PEG4000 for 2 min resulted in near-complete cell death. Following the PEG treatment, cells were transferred to K medium, and cultures began to turn green after approximately two weeks of incubation. These results demonstrate that 30% PEG4000 is not suitable for transformation of *M. commoda*.

The feasibility of transformation with 20% PEG4000 and PEG8000 was tested using the plasmid L1-RPS9:*shBle*, which confers resistance to zeocin (Figure 2.8a). Following transformation, cells were transferred to K medium at an initial density of 3×10^6cells/ml. The cells initially turned white, and then PEG4000-treated cells started to turn green after incubation for 20 days, and PEG8000-treated cells showed recovery after approximately 25 days. The transformed culture after 30 days of incubation are shown in Figure (2.8b). Cultures from each PEG treatment (PEG4000 and PEG8000) was selected for validation of the *shBle* gene by PCR. Primers were designed to amplify a 514 bp fragment spanning the RPS9 promoter and part of the *shBle* sequence. This fragment was detected in both PEG4000- and PEG8000-treated transformants but was absent in the wild-type control (Figure 2.8c).

PEG4000 was selected for subsequent transformation experiments due to its lower cytotoxicity. The plasmid pMC*shBle*:eYFP, conferring cells with zeocin resistance and enabling *eYFP* expression, was used for transformation. Transformed colonies appeared in 0.3% agarose selection plates after 14 days of incubation (Figure 2.10a). Transformation with 5 μg of linearised plasmid DNA yielded approximately 95 ± 45 colony-forming units (CFU) per plate, whereas transformation with 10 μg of





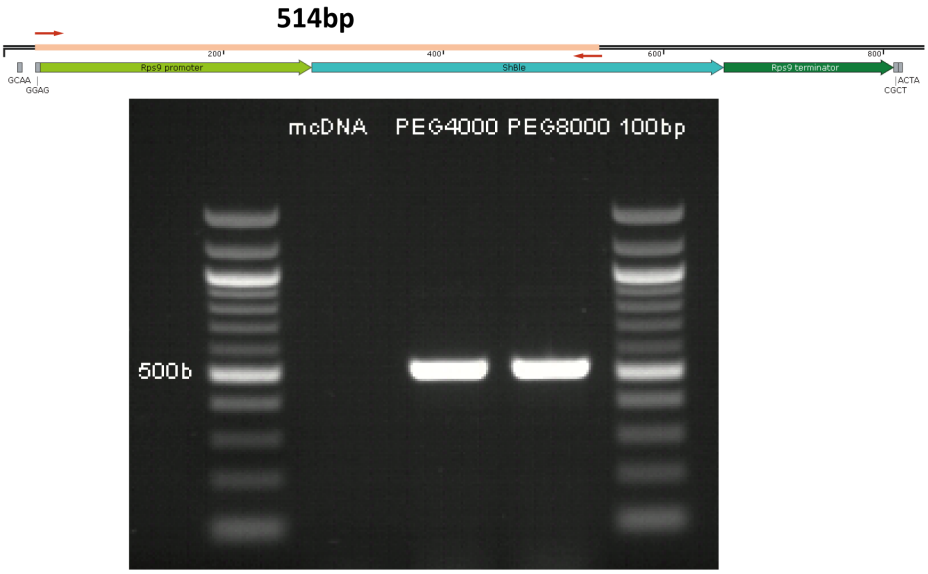


Figure 2.8: Screening of transformants using plasmid L1-RPS9:*shBle*. (a) Map of the transformation plasmid L1-RPS9:*shBle*; (b) Transformed cell cultures in K medium after 30 days of incubation; (c) Agarose gel electrophoresis showing that two transformants (PEG4000 and PEG8000 treatments) exhibit a 514 bp band corresponding to the RPS9 promoter and partial *shBle* gene sequence, whereas the wild type (WT) shows no corresponding band. The marker used is a 100 bp ladder.

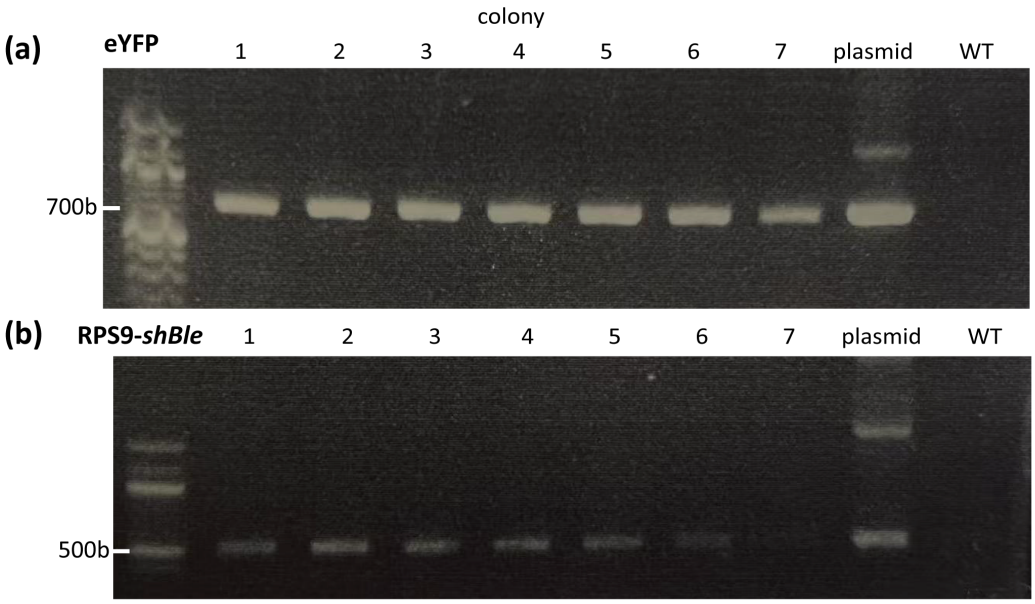


Figure 2.9: Validation of *eYFP* and *shBle* genes in transformants. (a) PCR screening of the *eYFP* gene. Lanes 1–7 correspond to picked colonies, and the plasmid control is the transformed plasmid *pMCshBle:eYFP*. A band of 720 bp was detected in all colonies and the plasmid control, but no band was observed in the wild-type (WT). (b) PCR screening of the RPS9 promoter and partial *shBle* gene. Lanes 1–7 correspond to picked colonies, and the plasmid control is *pMCshBle:eYFP*. A band of 514 bp was detected in all colonies and the plasmid control, but no band was observed in the WT. The DNA marker used was a 100 bp ladder.

circular plasmid DNA produced only a few colonies, which were not quantified. Both *eYFP* and *shBle* genes were confirmed in seven independent transformant colonies as well as in the plasmid pMC*shBle*:*eYFP*, but were absent in the wild-type strain (Figure 2.9). Expression of *eYFP* was further verified by fluorescence microscopy, showing strong signals in transformant cells, whereas the wild-type exhibited only very weak background fluorescence (Figure 2.10b).

Growth experiments were performed using transgenic strains to evaluate the impact of transformation on cellular phenotype (Figure 2.10c, d). The transformant L1-RPS9:*shBle*, conferring zeocin resistance, has been stably maintained in liquid culture for 10 months. The L2SY transformant, which exhibits both zeocin resistance and *eYFP* expression, has been maintained for 3 months. Growth analysis revealed that L1-RPS9:*shBle* displayed no significant difference compared to the wild-type (WT). In contrast, L2SY showed slower

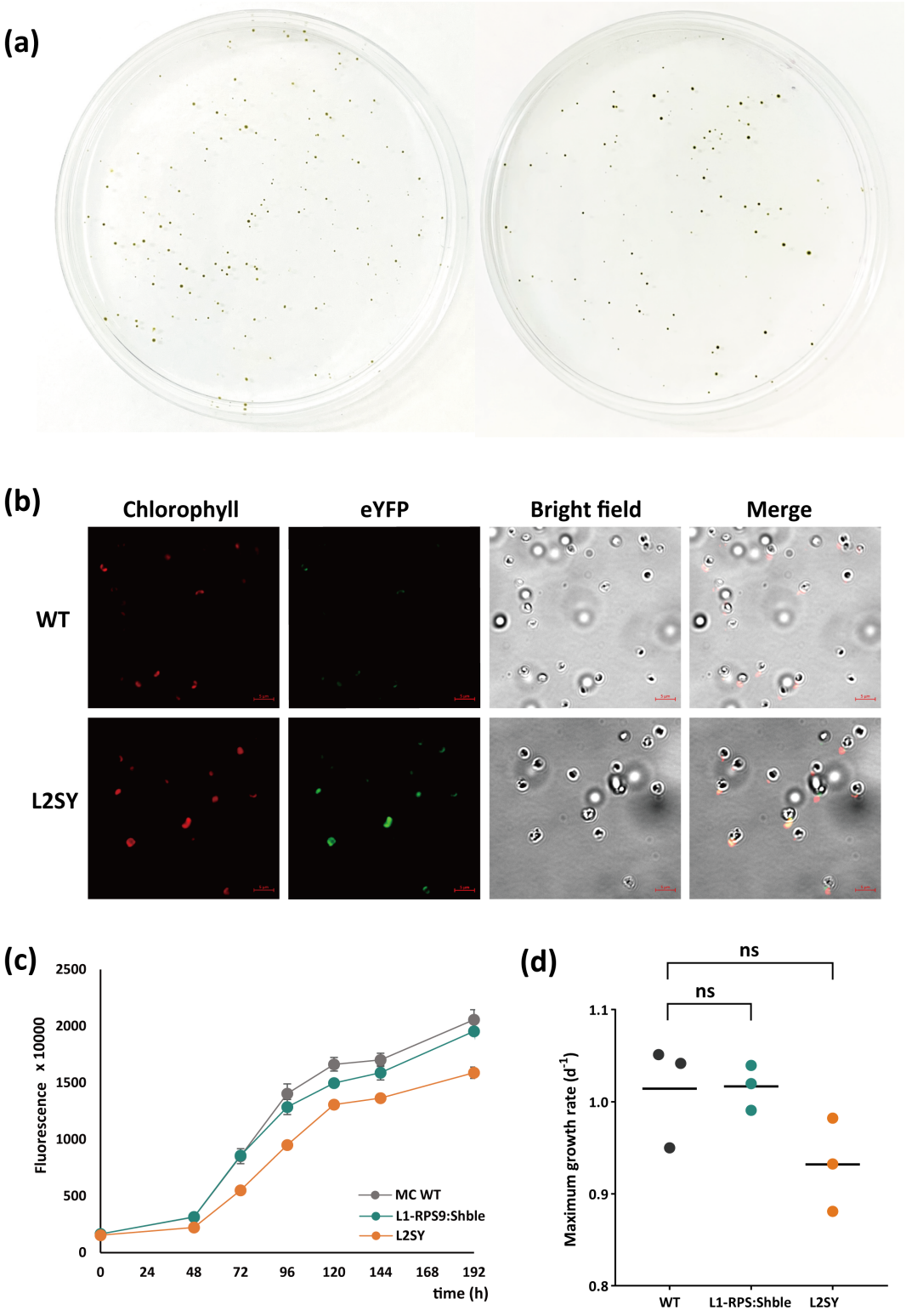


Figure 2.10: PEG-mediated transformation of *M. commoda*. (a) Transformants resistant to zeocin were selected in 0.3% agarose plate. (b) Fluorescence microscopy analysis of *eYFP* expression in *M. commoda* wild-type (WT) and transformatn L2SY. Shown are chlorophyll autofluorescence, *eYFP* fluorescence, bright-field images, and a merged image. Scale bars, 5 μm. (c) The growth curve of *M. commoda* WT and transformants in control medium with three replicates. (d) The maximum growth rate of *M. commoda* WT and transformants in control medium with three replicates.

growth relative to the WT after 48 h ($p < 0.01$). Importantly, the maximum growth rates of the wild-type and both transformants were not significantly different. These results suggest that the reduced growth observed in L2SY is likely attributable to an extended lag phase, potentially arising from transformation related effects or the metabolic burden of *eYFP* expression, which having minimal impact on exponential growth.

2.3.1.5 Multigene expression using 2A self-cleaving peptides

To investigate the function of four self-cleaving peptides in *M. commoda*, four 2A-expressing plasmids (Figure 2.3b) were introduced into *M. commoda* via PEG-mediated transformation, separately. In each plasmid, the *shBle* gene was linked to *eYFP* through a 2A self-cleaving peptide. Following PEG transformation, colonies were selected on 0.3% agarose plates containing zeocin and then kept in K medium containing zeocin, confirming the expression of *shBle*. The expression of *eYFP* was then assessed by fluorescence microscopy. Strong *eYFP* fluorescence signals were observed in L2SY cells and all four 2A transformants expressing *eYFP* independently, whereas WT cells exhibited only weak fluorescence (Figure 2.12), indicating that *eYFP* is functional in the all four kind of 2A transformants.

Western blot was performed to evaluate the cleavage efficiency of the four self-cleaving peptides (Figure 2.11). L2SY cells showed a single band at 27 kDa, corresponding to *eYFP* alone. In contrast, all four 2A transformants displayed a single band at 43.3 kDa, corresponding to the fused protein *shBle-2A-eYFP*. These results suggest that 2A peptides have relatively low cleavage efficiency in *M. commoda*, but the function of *shBle* and *eYFP* remain largely unaffected.

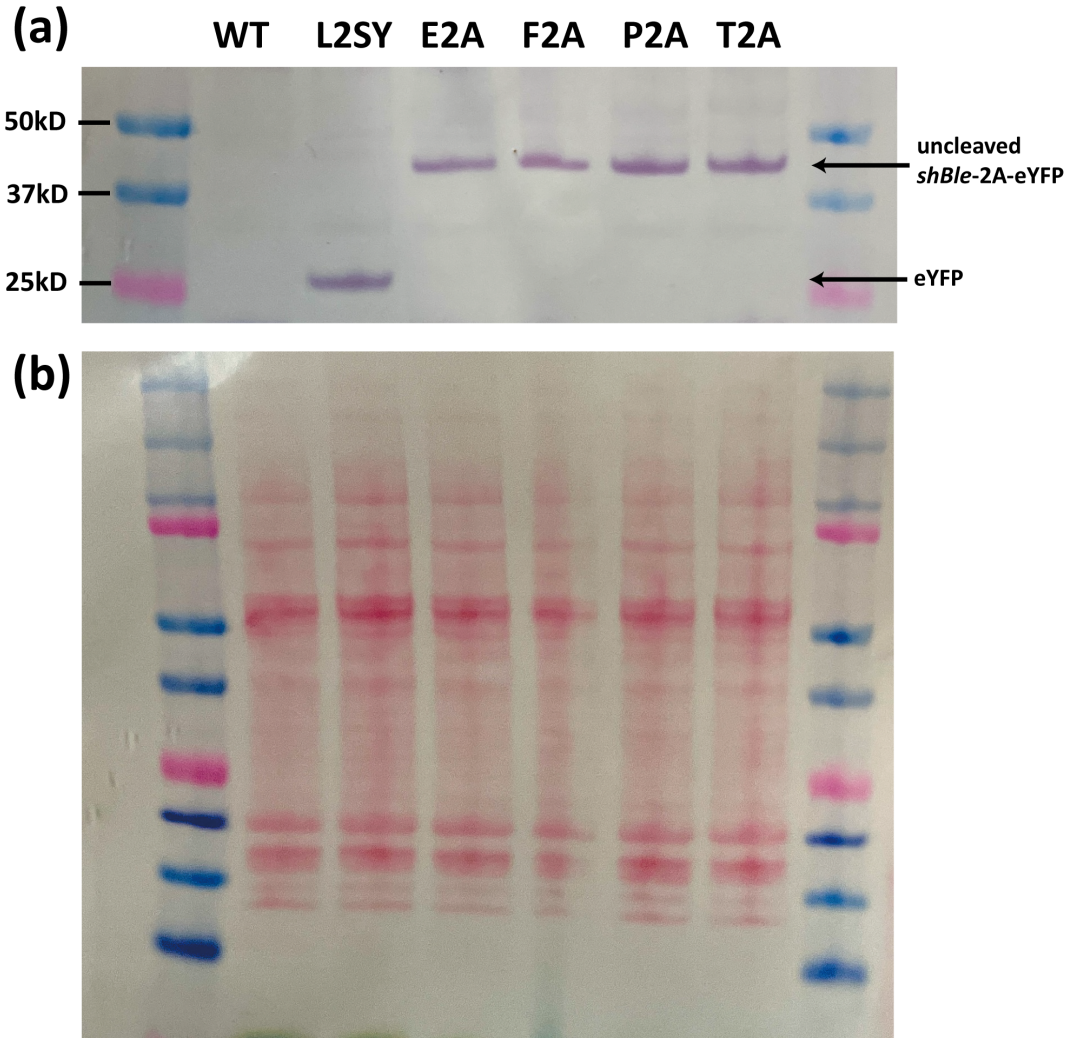


Figure 2.11: Western blot analysis of wild-type and transformant strains was performed using an anti-eGFP antibody. (a) Colony L2SY exhibited a single band at 27 kDa, corresponding to *eYFP*. Four 2A transformants displayed a single band at 43.3 kDa, consistent with the fused protein *shBle*-2A-*eYFP*. No bands were detected for the wild type. (b) Protein loading control based on Ponceau S.

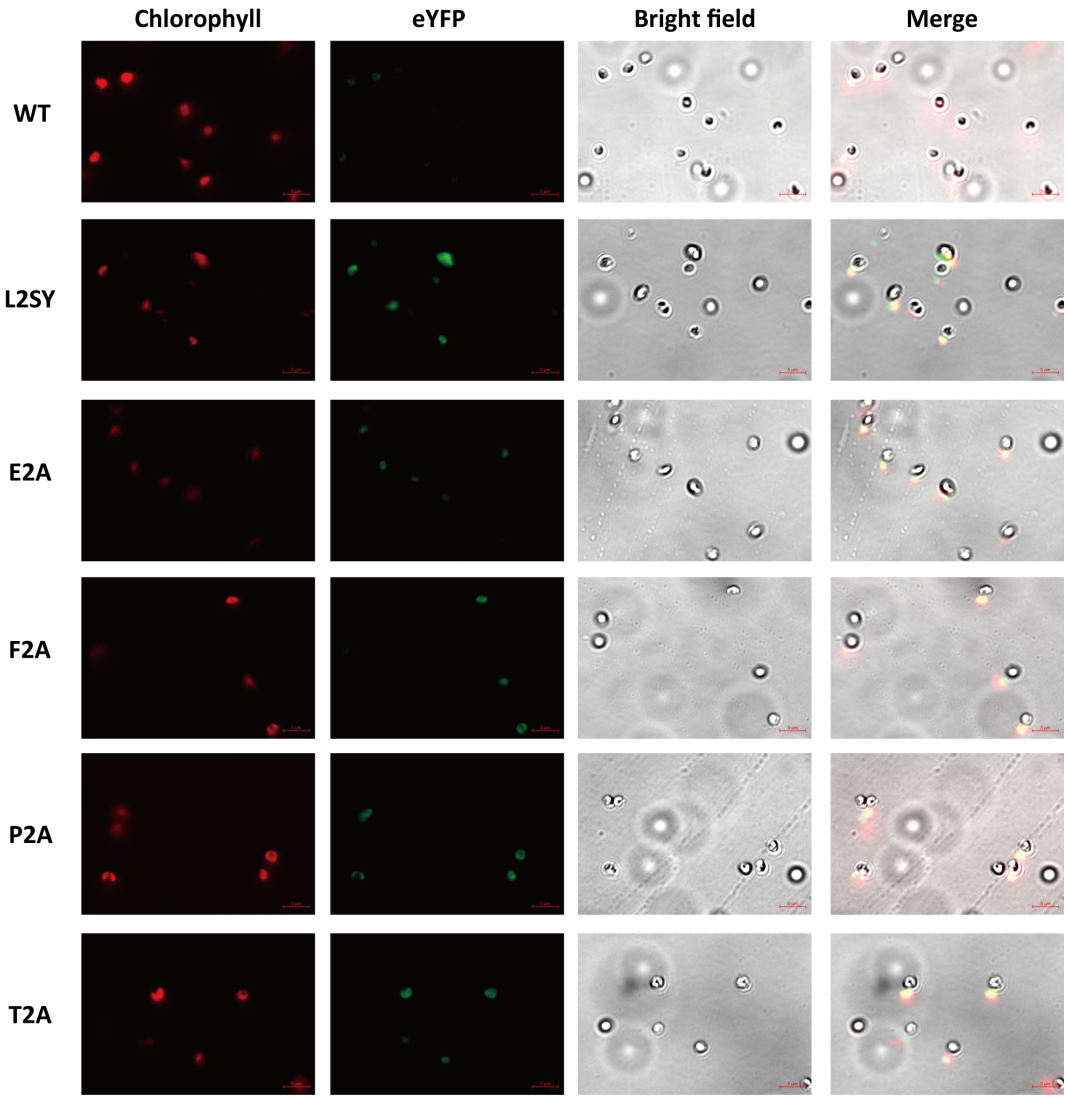


Figure 2.12: Confocal images of wild type and transformants.Shown are chlorophyll autofluorescence, *eYFP* fluorescence, bright-field images, and a merged image. Scale bars, 5 μm.

2.3.2 *Thalassiosira pseudonana*

To optimize the PEG transformation protocol for *T. pseudonana*, alcalase was used as the sole enzyme for the removal of cell-wall. Given the differences in cell-wall composition between species, cells were subjected to varying protoplasting conditions (Figure 2.13a) before PEG treatment. The circular episome pBIG1 (8.2 kb) conferring nourseothricin resistance and carrying *eGFP* was used in these experiments. All protoplasting conditions led to the emergence of colonies on the nourseothricin selection plates, with the harsher protoplasting conditions (i.e., $\geq$ 1.5 Anson alcalase) yielding more growth.

Transformants from the 500 μl alcalase treatment plate (i.e., 1.5 Anson, 40 min) were pooled and subjected to fluorescence microscopy to assess if *eGFP* was being expressed, as this would indicate successful episome transfer. Both *eGFP* signals and chla fluorescence were visible in transformants (Figure 2.13e), demonstrating that episome transfer had occurred. PCR results showed amplification of *eGFP* in the transformants, whereas no amplification was detected in the wild-type (Figure 2.13d).

To further characterize if protoplasting was truly occurring, *T. pseudonana* cells were incubated with 10 or 1000 μl of alcalase (i.e., 30 mAnson and 3 Anson, respectively) for 40 min (Figure 2.13b, c) and then imaged via microscopy. Based on these microscopy images, cells from the 10 μl treatment (i.e., 30 mAnson) showed a protoplasting rate of 4% (1/28 cells), whereas cells from the 1000 μl treatment (i.e., 3 Anson) demonstrated a protoplasting rate of 50% (13/20 cells). This finding underscores that alcalase can be used to degrade the cell wall of diatoms harboring silica frustules.

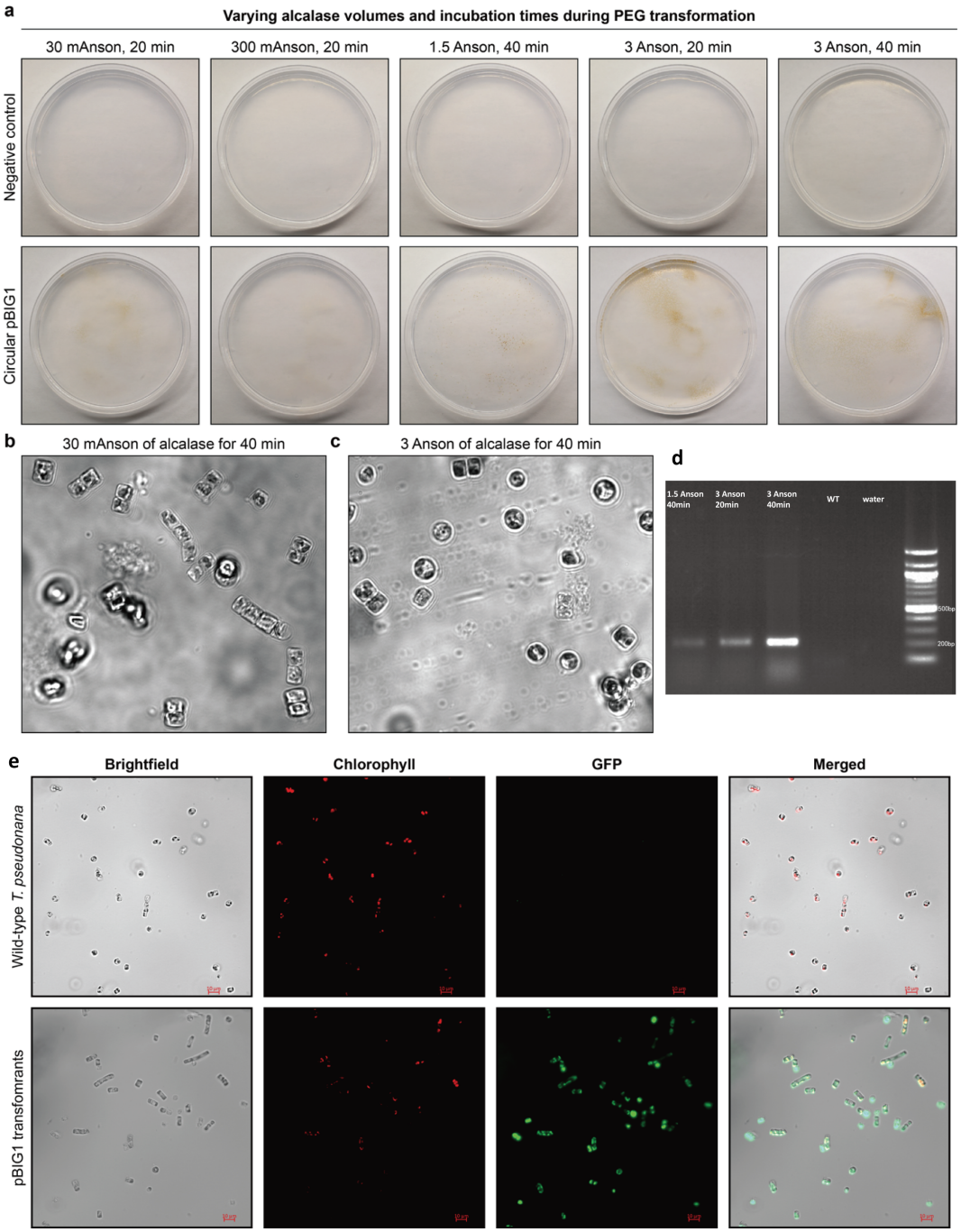


Figure 2.13: PEG-mediated transformation of *T. pseudonana* (Walker et al., 2026). (a) Harvested cells were subjected to varying concentrations of alcalase and incubation durations during the protoplasting step of the PEG transformation protocol. Following recovery, half of each reaction was plated on full-salt F/2 agar supplemented with 100 μg/mL nourseothricin. (b) Alcalase treatment using 10 μg/mL (30mAnson) for 40min. (c) Alcalase treatment using 1000 μg/mL (3 Anson) for 40min. (d) Agarose gel electrophoresis showing that transformants using different alcalase treatments (1.5 Anson,40min; 3 Anson, 20min; 3 Anson, 40min) exhibit a 212bp band corresponding to the *eGFP* gene sequence, whereas the wild type (WT) and water (negative control) shows no corresponding band. The marker used is a 100 bp ladder. (e) Fluorescence microscopy of wild-type (i.e., untransformed) and PEG-transformed *T. pseudonana* cells. Cells were transformed with pBIG1, which contains *eGFP*.

2.4 Discussion

In this study, I demonstrated the feasibility of PEG-mediated transformation in both the marine green alga *M. commoda* and the diatom *T. pseudonana*. To date, the only reported transformation method for *M. commoda* is Lonza nucleofection (Faktorova et al., 2020). PEG-mediated transformation is a straightforward approach that it does not require specialized equipment or buffers, making it more accessible. My results indicate that this method enables an efficient and stable transformation in *M. commoda*. One of the limitations of PEG-mediated transformation in diatoms is that it is effective only in cells lacking an intact cell wall. Here, we demonstrate that partial or complete removal of the cell wall significantly enhances PEG-mediated transformation efficiency in *T. pseudonana*.

The cell wall is the first physical barrier to entry when attempting to transform diatoms. Permeabilizing cell wall should increase the ability for biomolecules, like DNA and RNPs, to successfully enter the cell. My results show that protoplasting the cells also enabled efficient transformation through the PEG method (Figure 2.13a). Spheroplasting *P. tricornutum* cells with alcalase prior to conducting electroporation increased the transformation efficiency by 76- to 280-fold (Walker et al., 2026). These results suggest that the spheroplasting and protoplasting techniques can be applied to diatoms, though it will likely be necessary to optimize the intensity and duration of the enzyme treatment on a species-by-species basis along with other parameters in the respective protocols. This expectation aligns with the trajectory of bacterial conjugation in diatoms, which was originally established using plated *P. tricornutum* cells. Although initially assumed to have limited applicability, conjugation has since been successfully adapted to multiple species, including *C. meneghiniana*, *Nitzschia captiva*, and *Fistulifera solaris* (Sprecher et al., 2023; Maeda et al., 2023; Yin et al., 2024).

PEG4000 has also been widely used for transformation in various microalgal species (Ohnuma et al., 2014; Sanchez et al., 2019; Hirooka et al., 2022). The

underlying mechanisms of action remain unclear, and consequently the optimal PEG incubation time varies between species. For example, a short incubation of 2 min was sufficient for efficient transformation in *O. tauri*. However, incubation times of up to 40 min are recommended for *P. tricornutum* and *T. pseudonana* (Sanchez et al., 2019; Walker et al., 2026). In this study, we evaluated the sensitivity of *M. commoda* to PEG but did not systematically investigate the relationship between incubation time and transformation efficiency. Future work should include optimizating the incubation duration to determine the balance between sufficient membrane permeabilization and minimization of cytotoxic effects, which is likely to be critical for maximizing transformation efficiency in *M. commoda* and other microalgae.

The transformation efficiency is related to the physical state of DNA, which has been reported in other microalgae (Sanchez et al., 2019; Guo et al., 2022; Walker et al., 2026). Our results show that transformation using linearized plasmid has significantly higher efficiency than that of using plasmid. When using plasmid pMC*shBle*:*eYFP*, hundreds of colonies per plate were obtained with 5 μg of linearized DNA (Figure 2.10), but only a few colonies were observed with 10 μg of circular plasmid DNA. It should be noted that plasmid pMC*shBle*:*eYFP* contains a prokaryotic origin of replication and cannot replicate in *M. commoda*. This difference indicates that the physical state of the DNA plays a critical role in determining transformation efficiency. Linearized DNA contains free double-strand ends that can be recognized as DNA double-strand breaks (DSBs) by DNA damage-response (DDR) pathway (Sancar et al., 2004; Thomy et al., 2021; Lim et al., 2023). Homologous recombination (HR) and non-homologous end joining (NHEJ) are two main mechanisms for DSB repair in eukaryotic cells (Shrivastav et al., 2008). Random integration of exogenous DNA into the genome is primarily through NHEJ. In contrast, circular plasmids lack free DNA ends and are less efficiently processed by these repair mechanisms, resulting in a relatively low integration efficiency (Puchta, 2005). Therefore, linearized plasmids are more likely to yield more transformants in *M. commoda*.

Carrier DNA is a key factor in improving transformation efficiency in both PEG-mediated transformation and electroporation (Corellou et al., 2009; Ohnuma et al., 2014; Sanchez et al., 2019). Commonly used carrier DNA in microalgal transformation includes salmon sperm DNA and yeast tRNA, both of which enhance the stability and uptake of plasmid DNA by protecting it from nuclease-mediated degradation and reducing DNA loss during transformation. In this study, no colonies were obtained in the absence of carrier DNA during the transformation of *M. commoda*, indicating that carrier DNA is essential for the successful transformation in this system. Similar observations were reported in PEG-mediated transformation of *O. tauri* and electroporation in *P. tricornutum*, where the inclusion of carrier DNA resulted in five- to fifteen- fold increase in transformant yield compared to conditions without carrier DNA, separately (Sanchez et al., 2019; Walker et al., 2026). Therefore, addition of carrier DNA should be considered a critical factor when optimizing transformation protocols in microalgae.

In this study, the endogenous RPS9 promoter was selected to drive transgene expression in *M. commoda*, based on its role as a constitutively expressed ribosomal protein gene and existed transformation application (Faktorova et al., 2020). Ribosomal protein genes are typically highly expressed in cells, although their transcription can be modulated under environmental stress conditions (Bachy et al., 2018). Our results demonstrate that the RPS9 promoter efficiently drives the expression of heterologous genes, conferring zeocin resistance and strong *eYFP* fluorescence. The commonly used promoter varies across microalgal species. In diatoms, *FCP* promoters are widely used as constitutive promoters in many species such as *P. tricornutum*, *T. pseudonana*, *F. cylindrus*, and *Nitzschia captiva* (Hopes et al., 2016; Hopes, 2017; Belshaw et al., 2023; Sprecher et al., 2023). On the other hand, in non-photosynthetic species such as the diatom *N. putrida*, the promoter and terminator of nicotinamide adenine dinucleotide hydride (NADH)-ubiquinone reductase complex I were selected to drive transgene expression based on their high transcript abundances (Deng et al., 2025). However, it should be noted that not all

endogenous constitutive promoters are equally strong, and promoter selection should be tailored to experimental conditions. High-level constitutive expression of large proteins may impose a metabolic burden on picoeukaryotes. In this study, the *eYFP* expression may contribute to the observed delay in the lag phase growth (Figure 2.10c). Under this condition, inducible promoters provide an alternative strategy by enabling conditional gene expression in response to specific environmental conditions (Li et al., 2025c). Promoters responsive to nutrient availability, including nitrogen, iron, silicon, and phosphate, have been widely exploited as inducible systems. For example, the nitrate reductase (*NR*) promoter is commonly used in diatoms, where it is activated in the presence of nitrate as the sole nitrogen source and repressed when other nitrogen such as ammonium is available (Schellenberger Costa et al., 2013; Ifuku et al., 2015). This also gives ideas for optimizing foreign gene expression in *M. commoda*.

The efficiency of this cleavage process is influenced by factors such as ribosome structure, the specific 2A sequence, and the upstream coding context. High cleavage efficiency (approaching 100%) has been reported in systems such as silkworms, zebrafish, mice, and human cell lines, whereas much lower and more variable efficiencies have been observed in microalgae (Kim et al., 2011; Wang et al., 2015; Plucinak et al., 2015; Koh et al., 2018; Wang et al., 2024b). For example, four different 2A peptides enabled co-expression of *sfCherry* and *shBle* within a single expression cassette in the marine microalga *Nannochloropsis salina*, although cleavage efficiencies varied depending on the 2A sequence used (Koh et al., 2018). However, the performance of 2A peptides is not universally consistent across microalgal species. In the dinoflagellate *Breviolum minutum*, P2A-mediated cleavage was reported to be inefficient and inconsistent during co-expression of GFP and puromycin N-acetyltransferase (*pac*) (Gornik et al., 2022). Similarly, in the diatom *Phaeodactylum tricornutum*, 2A peptide-mediated cleavage has been reported to be variable and relatively inefficient (Diamond et al., 2023). Consistent with these observations, my western blot results (Figure 2.11) indicate that the cleavage efficiency of the four tested 2A peptides in *M. commoda* was low or

undetectable. Only the fused *shBle-eYFP* protein was detected, whereas no band corresponding to free *eYFP* was observed. However, the fusion configuration did not affected the expression of *eYFP* (Figure 2.12). This may be because the 2A peptides did not mediate efficient cleavage and instead functioned as a linker-like sequence between *shBle* and *eYFP*, with minimal impact on protein folding and expression.

Differences in ribosomal properties, particularly reduced sensitivity of ribosomal E site in microalgae, may contribute to the reduced efficiency of 2A-mediated ribosomal skipping (Koh et al., 2018). Cleavage efficiency may be improved by optimizing the 2A sequence through codon optimization or by using extended 2A peptide variants (Plucinak et al., 2015). In addition, the use of 2A peptides reduces the number of promoters required but increases the length of the polycistronic transcript. When expression is driven by the same promoter, the expression efficiency may be influenced by the length and composition of the linked coding sequences (Muñoz et al., 2019; Wang et al., 2024b). Furthermore, variation in transgene expression may arise from positional effects associated with random genomic integration (Wang et al., 2024b). To fully assess the applicability of 2A peptides as molecular tools in *M. commoda*, it will be important to systematically evaluate the relationships between 2A peptide sequence, cleavage efficiency, positional effects within the expression cassette, and the functional activity of the expressed proteins.

3

Development of CRISPR/Cas9 in *Micromonas commoda*

3.1 Introduction

The past decade has revolutionized our ability to modify microalgal genomes through the development of sequence-specific nucleases that enable targeted genetic modifications. Gene editing tools include meganucleases (MNs), zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeats-associated system (CRISPR/Cas) (Daboussi et al., 2014; Fortunato et al., 2016; Hopes et al., 2016; Greiner et al., 2017). These technologies introduce double-strand breaks (DSBs) at specific genomic loci, which are subsequently repaired via either homology-directed repair (HDR) or non-homologous end joining (NHEJ). NHEJ is the main repair pathway in microalgae and typically results in single-nucleotide mutations or small insertions and deletions (INDELs). Among these tools, the CRISPR/Cas9 system has emerged as the most versatile approach in microalgae, owing to the relative simplicity of its design compared to other approaches.

The native CRISPR/Cas9 system functions as part of the bacterial adaptive immune system by detecting and cleaving foreign DNA using sequence-specific guide RNAs (gRNAs) (Brouns et al., 2008). Short fragments of viral DNA are integrated into the host genome as spacers between clustered regularly interspaced palindromic

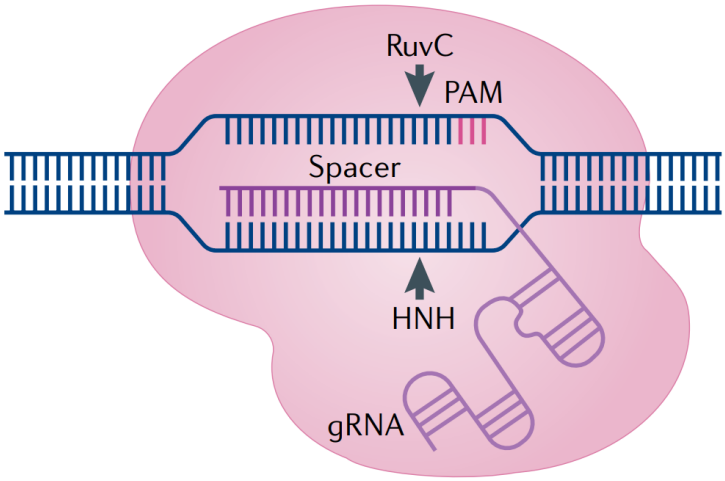


Figure 3.1: CRISPR/Cas9 gene editing system. Cas9 uses guide RNA (gRNA) to recognize target DNA sequences. When gRNA spacer pairs with a complementary DNA sequence adjacent to a PAM site, Cas9 cleaves the DNA, generating a blunt double-strand break via its RuvC and HNH nuclease domains (Pickar-Oliver and Gersbach, 2019).

repeats (CRISPR). In genome editing, the CRISPR RNA (crRNA) is fused with the trans-activating CRISPR RNA (tracrRNA) to generate a synthetic single guide RNA (sgRNA), which replace native gRNAs. The expression of sgRNA is often controlled by a promoter which recruits polymerase III for transcription of small non-coding RNAs. The U6 promoter is a popular choice for many organisms (Cong et al., 2013; Brooks et al., 2014; Hopes et al., 2016; Kim et al., 2021; Wang et al., 2024a), and can be easily identified for use in endogenous systems as the U6 gene contains regions of high conservation. The sgRNA directs the Cas9 endonuclease to a complementary DNA target site, enabling sequence-specific cleavage at the desired genomic locus (Bortesi and Fischer, 2015). Cas9 is a DNA endonuclease belonging to type II systems. It is able to recognize a protospacer adjacent motif (PAM) in the target DNA and then induce a DSB via its two nuclease domains (Figure 3.1). Each domain cleaves one DNA strand, that the HNH domain cleaves the DNA strand complementary to the guide RNA, and the RuvC domain cleaves the non-complementary strand (Brouns et al., 2008; Lander, 2016). Following cleavage, the resulting DNA break is repaired by the host cell through either NHEJ or HDR.

The first report of CRISPR/Cas-based genome editing in microalgae demonstrated

transient expression of Cas9 and successful gene editing in *Chlamydomonas reinhardtii* (Jiang et al., 2014). However, constitutive *in vivo* expression of Cas9 was found to be toxic. To overcome this limitation, ribonucleoprotein (RNP) complexes consisting of recombinant Cas9 protein and either synthetic or *in vitro* transcribed sgRNAs have been widely employed to achieve efficient genome editing in *C. reinhardtii* (Baek et al., 2016; Shin et al., 2016). Since then, CRISPR/Cas-mediated gene knockout has been successfully established in a broad range of microalgal species (Tran et al., 2025), including *P. tricornutum* (Nymark et al., 2016), *T. pseudonana* (Hopes et al., 2016; Belshaw et al., 2023), *Nannochloropsis spp.* (Wang et al., 2016), *Chlorella spp.* (Kim et al., 2024), *Tetraselmis spp.* (Chang et al., 2020), *Euglena gracilis* (Nomura et al., 2019, 2024), and *Porphyridium purpureum* (Jeon et al., 2021). In these applications, single sgRNA are typically used to induce targeted mutations via the error-prone NHEJ repair pathway. Due to low editing efficiencies and mosaic genotypes, an efficient and reliable screening of edited transformants is a crucial step in microalgal genome editing workflows.

To isolate and confirm the desired genetic modifications, steps include subcloning and screening were employed. Both are necessary because the initial cultures are often mosaic (Huang and Daboussi, 2017; Serif et al., 2018), necessitating the isolation of single-cell lines on solid plates or in liquid medium. Most transformed colonies can be isolated using antibiotics or chemical compounds on selective plates. However, the growth conditions on these plates should be tested when working with a new microalgal species. Clonal cell lines generally appear in the first round of subcloning. However, mosaic colonies containing a mixture of wild-type and mutant cells, as well as colonies carrying monoallelic mutations, are frequently obtained in diploid species. In such cases, an additional round of single-colony isolation or subcloning is required to obtain clones carrying biallelic mutations in the target gene. However, this process is time-consuming, taking more than one week to obtain clones, depending on the growth rates and the delivery system used (Yin and Hu, 2021; Belshaw et al., 2023). Transformed cells

expressing a fluorescence marker can be separated by fluorescence-activated cell sorting (FACS), which therefore can achieve high rates of cell recovery under near axenic conditions (Pereira et al., 2018; Deng et al., 2025). However, it should be noted that some species can not survive the FACS procedure due to mechanical pressure, which may cause disruptions of the plasma membrane and cell wall (Reckermann, 2000).

PCR amplification of the target gene followed by analyzing the electropherograms is a common method for analyzing the genetic mutation to be expected (Nymark et al., 2016; Hopes et al., 2016; Moosburner et al., 2022). Gene knockout cells often exhibit a distinct PCR-based genotyping pattern compared to WT cells (Hopes et al., 2016). When a small deletion or insertion occurs (less than 100 bp or a few hundred bp), gene-knockout cell-lines typically display a smaller band than WT due to the reduced fragment size (Li et al., 2025c). To further validate the modification, the expected sequence is cloned into a vector for transformation into *E. coli*, followed by Sanger sequencing and alignment with the reference genome. Additionally, alternative methods such as high-resolution melt curve analysis (HRM), T7 endonuclease I assay (T7EI), and software-based Tracking of INDELs by Deconvolution (TIDE) provide insight into the nature of genetic modifications (Nymark et al., 2016; Slattery et al., 2018; Moosburner et al., 2020). If the PCR reactions fail due to extensive genomic rearrangements, Southern blotting may be useful to detect fragment size shifts and confirm copy number variations compared to WT DNA (Zhang and Hu, 2014; Kira et al., 2016). To further corroborate the genetic modifications of coding genes, relative protein levels can be assessed by western blotting, whereas mRNA levels of the target gene can be analyzed using real-time quantitative PCR (RT-qPCR) (Görlich et al., 2019; Strauss et al., 2023).

3.2 Materials and Methods

3.2.1 Cell culture

The culture and growth conditions for the marine green alga *Micromonas commoda* (RCC827) are detailed in 2.2.1.

3.2.2 Designing sgRNAs and *in vitro* assay

For the *cyanase* (*cynS*) gene in *M. commoda*, two sgRNAs targeting conserved domains within the coding region were designed and tested following Hopes et al. (2016). Conserved domains were identified through protein sequence alignment.

The chimeric sgRNA sequence was as follows: NNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT, where the underlined region represents the 20 nt target sequence. A total of ten sgRNA targets were designed using RGEN Cas-Designer based on predicted off-target scores and GC content (Table 3.1). The sgRNA1 and sgRNA2 were excluded given that they positioned distal to the conserved domains, they were excluded, and only sgRNAs from sgRNA3 to sgRNA10 were selected as candidate targets.

An *in vitro* Cas9 cleavage assay was performed to evaluate the efficiency of all designed sgRNAs, following an established protocol (Benchling protocol). Eight sgRNAs were synthesized from oligonucleotides via PCR, verified on a 2% agarose gel, and purified using the Monarch PCR & DNA Cleanup Kit (NEB). The purified DNA templates were used for *in vitro* transcription using the HiScribe T7 High Yield RNA Synthesis Kit (NEB). The resulting RNAs were purified using the RNA Clean & Concentrator Kit (Zymo Research) to remove residual DNA and enzymes. RNA integrity and size were confirmed by electrophoresis on an 12% (7 M urea)

denaturing polyacrylamide gel (0.5× TBE) at 200 V for 40 min. Gel was stained with SYBR Gold nucleic acid stain (Invitrogen; CAT# S11494) and visualized using a transilluminator.

The *cynS* template was amplified by PCR using Phusion High-Fidelity DNA Polymerase (NEB) with the forward primer 5′-CCCGGGTTTGGAAAAAGTGC-3′ (F2) and reverse primer 5′-ATTGGAACGCAGAGTGAGGG-3′ (R2). The PCR product was verified on a 1% agarose gel and cloned into a pCR4-TOPO vector (Invitrogen). The resulting plasmid was then linearized using *NcoI*, ensuring no disruption to the *cynS* template sequence.

For the *in vitro* cleavage assay, 30 nM of linearized *cynS* template, 40 nM sgRNA, and 40 nM *Streptococcus pyogenes* Cas9 protein (NEB; M0386S) were incubated in reaction buffer at 37°C overnight. The reaction was terminated by adding stop buffer as described in Hopes (2017) and heating at 80°C for 10 min. Reaction products were analyzed on a 1× TAE agarose gel and visualized using a transilluminator to assess the sgRNA cleavage efficiency.

3.2.3 Identification of U6 small nuclear RNA (snRNA)

To identify the U6 small nuclear RNA (snRNA) in *M. commoda*, U6 sequences from multiple organisms were retrieved from GenBank and aligned using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) to determine conserved regions. Based on these alignments, the highly conserved central region of the U6 sequence was confirmed. An NCBI blastn search was then performed against the *M. commoda* genome using this conserved region to identify candidate U6 snRNA loci. The resulting hits were further examined for sequence similarity and the presence of characteristic U6 features.

Table 3.1: Overview of sgRNAs designed and tested for cleavage efficiencies. Each sgRNA is followed by a specific PAM sequence. High out-of-frame scores (> 66) are recommanded to minimize the occurrence of in-frame deletions.

sgRNA	Target	PAM	Out-of-frame Score	Direction	GC Contents
sgRNA1	TGTCGTCAAAAACGGACGCC	CGG	78.5	-	55
sgRNA2	GCGGACATGGATCGGCGGAA	CGG	72.8	-	65
sgRNA3	CATGAAGAAAGGGATCACGT	GGG	76.5	+	45
sgRNA4	CTGCAGCACGTGCCGTACAA	GGG	68.7	+	60
sgRNA5	ATGGACCCGGTGCCCTTGTA	CGG	68.7	-	60
sgRNA6	ATCGACTTCGACATGGAGTT	TGG	68	+	45
sgRNA7	AGGCGGATCCAAACGGGGAC	AGG	65.9	+	65
sgRNA8	CCGTTGCCTATTAAATTTGA	CGG	68.8	+	35
sgRNA9	AGCCCGTAGTCGAATGCGTG	GGG	67	+	60
sgRNA10	TGACAAAAATATCGCCCCGC	CGG	75.2	-	50

3.2.4 Extraction of *M. commoda* and *M. pusilla* genomic DNA

Both *M. commoda* and *M. pusilla* were grown to exponential phase and the genomic DNA (gDNA) were extracted using the Qiagen MagAttract HMW DNA Kit (Qiagen, Catalogue #67563) according to the manufacturer's protocol.

3.2.5 Golden gate assembly

3.2.5.1 L0 assembly

The L0 assembly of endogenous RPS9 promoter and terminator was detailed in 2.2.4.2. The MpCas9-U6Guide-RPL42 plasmid with both codon optimized Ccas9 gene and *M. pusilla* U6 promoter was received from Prof. Alexandra Z. Worden and Dr. Olga Matantseva. The Cas9 gene and *M. pusilla* U6 promoter were amplified by PCR using Phusion DNA polymerase (NEB) from MpCas9-U6-RPL42 plasmid. The *M. pusilla* U6 terminator was amplified by PCR from *M. pusilla* genome DNA. *M. commoda* U6 promoter and terminator were PCR-amplified from *M. commoda* genome DNA. The *eYFP* gene was amplified from pICH47742:FCP:Cas9YFP. The sgRNAs were amplified from DNA template for *in vitro* assay. Primers for amplification are listed in Table 3.2. All sequences were cloned into the pCR4-TOPO vector (Thermo Fisher). Plasmid DNA was extracted using Monarch plasmid miniprep kit.

The BsaI and BbsI restriction site in the L0–cas9 plasmid were removed using the Site-Directed Mutagenesis (SDM) kit (NEB, UK). The BsaI restriction site was removed firstly using forward primer 5′-ccagaaggcgcaAgtctcgggc-3′ and reverse primer 5′-atgtcctccttgaacgtgagcgagtcgt-3′. The mutation site is indicated by the uppercase letter. The resulting L0–cas9–SDM–Q plasmid was generated. The BbsI restriction site in the L0–cas9–SDM–Q plasmid was then removed using forward primer 5′-caacatcgtcaaAaagacggagg-3′ and reverse primer 5′-acctgcggcatcgaga-3′. The mutation site is also indicated by the uppercase letter. The resulting plasmid

Table 3.2: PCR primers for creating Golden gate cloning L0 modules. BsaI sites are shown in bold. 4nt specific overhangs are shown in underlined. Lowercase letters represent extra added sequences, and uppercase letters indicate the part of primer combine with the initial template.

Primer	Sequence	Annealing temp (°C)
cas9 F	**ggtctc**acaagGACAAGAAGTACTCGATC	66.7
cas9 R	**ggtctc**aGACCTTGCGCTTCTTCTTC	66.4
eYFP F	**ggtctc**aggtcATGGTGAGCAAGGGCGA	72.3
eYFP R	**ggtctc**agtgtTCACTTGTACAGCTCGTCCAT	69.5
McU6P F	**ggtctc**aggagAATTGGTTTATATGTGTGTCGA	67
McU6P R	**ggtctc**aAGGTATGGCATCCTGCTTTGGTA	68.1
McU6T F	**ggtctc**agtgcTTTTTTTAACGCGTGAAGAACA	67
McU6T R	**ggtctc**aagcgGCGTCCGAGTCTTACCAGA	72.2
MpU6P F	**ggtctc**aggagAAAGAGCGTGTTTTGAACG	68.1
MpU6P R	**ggtctc**aGACATGTCCTCATGCACAACA	66.6
MpU6T F	**ggtctc**aagcgCGTGTACTTTCGAAAGTA	66.9
MpU6T R	**ggtctc**agtgcTTTTTTACCGTGTCTGACTA	66.7
McsgRNA5 F	**ggtctc**aacctATGGACCCGGTGCCCTT	71
McsgRNA6 F	**ggtctc**aacctATCGACTTCGACATGGA	66.6
McsgRNA8 F	**ggtctc**aacctCCGTTGCCTATTAAATTTGA	65.5
MpsgRNA5 F	**ggtctc**atgtcATGGACCCGGTGCCCTT	71
MpsgRNA6 F	**ggtctc**atgtcATCGACTTCGACATGGA	66.6
MpsgRNA8 F	**ggtctc**atgtcCCGTTGCCTATTAAATTTGA	65.5
sgRNA R	**ggtctc**aGCACCGACTCGGTGCCA	69.6

L0–cas9–SDM–Q–K was used as L0 module for further assembly.

3.2.5.2 L1 assembly

The L1 assembly and screening methods were detailed in 2.2.4.2.

The Cas9 and *eYFP* cassettes were assembled with the RPS9 promoter and RPS9 terminator in L1 backbone, resulting in the plasmid RPS9:cas9:*eYFP*-pICH47732.

The plasmid was extracted using Monarch plasmid miniprep kit (NEB) and then screened using EcoRI digestion, with target length of 9631bp. The *eYFP* gene was removed using the Site-Directed Mutagenesis (SDM) kit (NEB, UK) with primer 5′-acacccgagctgagggccgg-3′ and 5′-gaccttgcgcttcttcttcgggtccg-3′, generating plasmid RPS9:cas9-pICH47732. The plasmid was extracted and then screened using XmaI and MdeI digestion, with target length of 4063bp and 5301bp. The correct constructs were confirmed by whole plasmid sequencing.

The sgRNA5, sgRNA6 and sgRNA8 were assembled with the McU6 promoter and terminator in L1 backbone, resulting in the plasmid McU6:sgRNA5-pICH47751, McU6:sgRNA6-pICH47761 and McU6:sgRNA8-pICH47761, separately. Plasmids were digested using EcoRI and the estimated length is 5136bp. Correct constructs were confirmed by Sanger sequencing using forward primer 5′-cccactctgtgaagacaa-3′ and reverse primer 5′-gccaatatatcctgtcaaacac-3′.

The sgRNA5, sgRNA6 and sgRNA8 were assembled with the MpU6 promoter and terminator in L1 backbone, resulting in the plasmid MpU6:sgRNA5-pICH47751, MpU6:sgRNA6-pICH47761 and MpU6:sgRNA8-pICH47761, separately. Plasmids were digested using EcoRI and the estimated length is 4803bp. Correct constructs were confirmed by Sanger sequencing using forward primer 5′-cccactctgtgaagacaa-3′ and reverse primer 5′-gccaatatatcctgtcaaacac-3′.

3.2.5.3 L2 assembly

For L2 assembly assembly of knockout plasmids with *M. commoda* U6 promoter, the L1 modules RPS9:cas9-*eYFP*-pICH47732, RPS9:*shBle*-pICH47742 and McU6:sgRNA5-pICH47751, L4E linker (pICH41780) and the L2 vector pAGM4723 were assembled with McU6:sgRNA6-pICH47761 and McU6:sgRNA8-pICH47761, separately, resulting in transformation plasmid L2-MC56 and L2-MC58. The *eYFP* were removed using the SDM kit with primers 5′-acacccgagctgagggccgg-3′ and 5′-gaccttgcgcttcttcttcgggtccg-3′, generating the plasmid L2-Mc56dYFP and L2-Mc58dYFP (Figure 3.2). Constructs were PCR-screened by forward primer

5′-gcgaacggcgagatccgcaa-3′ and reverse primer 5′-gacacgacctccgaccactc-3′, and the estimated length was 1654bp. The integrity of the constructs was then confirmed by whole-plasmid sequencing.

For assembly of knockout plasmids with *M. pusilla* U6 promoter, the L1 modules RPS9:cas9-pICH47732 and RPS9:*shBle*-pICH47742, MpU6:sgRNA5-pICH47751, L4E linker (pICH41780) and the L2 vector pAGM4723 were assembled with MpU6:sgRNA6-pICH47761 and MpU6:sgRNA8-pICH47761, separately, resulting in transformation plasmid L2-Mp56dYFP and L2-Mp58dYFP (Figure 3.3). The *E. coli* colonies carrying constructs were screened using forward primer 5′-caacatcgtcaaaaagacggagg-3′ and reverse primer 5′-gccaatatatcctgtcaaacac-3′, and the estimated length was 2798bp. Plasmids were then extracted and screened by EcoRV and NcoI disgestion, with estimated length of 2053bp and 8911bp. The construct Mp58dYFP was confirmed by whole-plasmid sequencing.

3.2.6 Transformation

The transformation method for *M. commoda* was detailed in 2.2.6.1.

3.2.7 Screening of knockout colonies

The lysis method for *M. commoda* colonies was detailed in 2.2.7.1.

Two μl of lysate were used as templates in Taq or Q5 PCR to amplify the CRISPR–Cas-targeted fragment of the cyanase gene. Two primer pairs were used for amplification. The first primer pair generated an expected amplicon of 903 bp, consisted of primers 5′-CCCGGGTTTGGAAAAAGTGC-3′ (forward) and 5′-ATTGGAACGCAGAGTGAGGG-3′ (reverse). The second pair consisted of a forward primer 5′-CTGCAGCACGTGCCGTACAA-3′ and the same reverse primer, yielding an expected product of 415 bp.

PCR products were analyzed by agarose gel electrophoresis to identify the lower

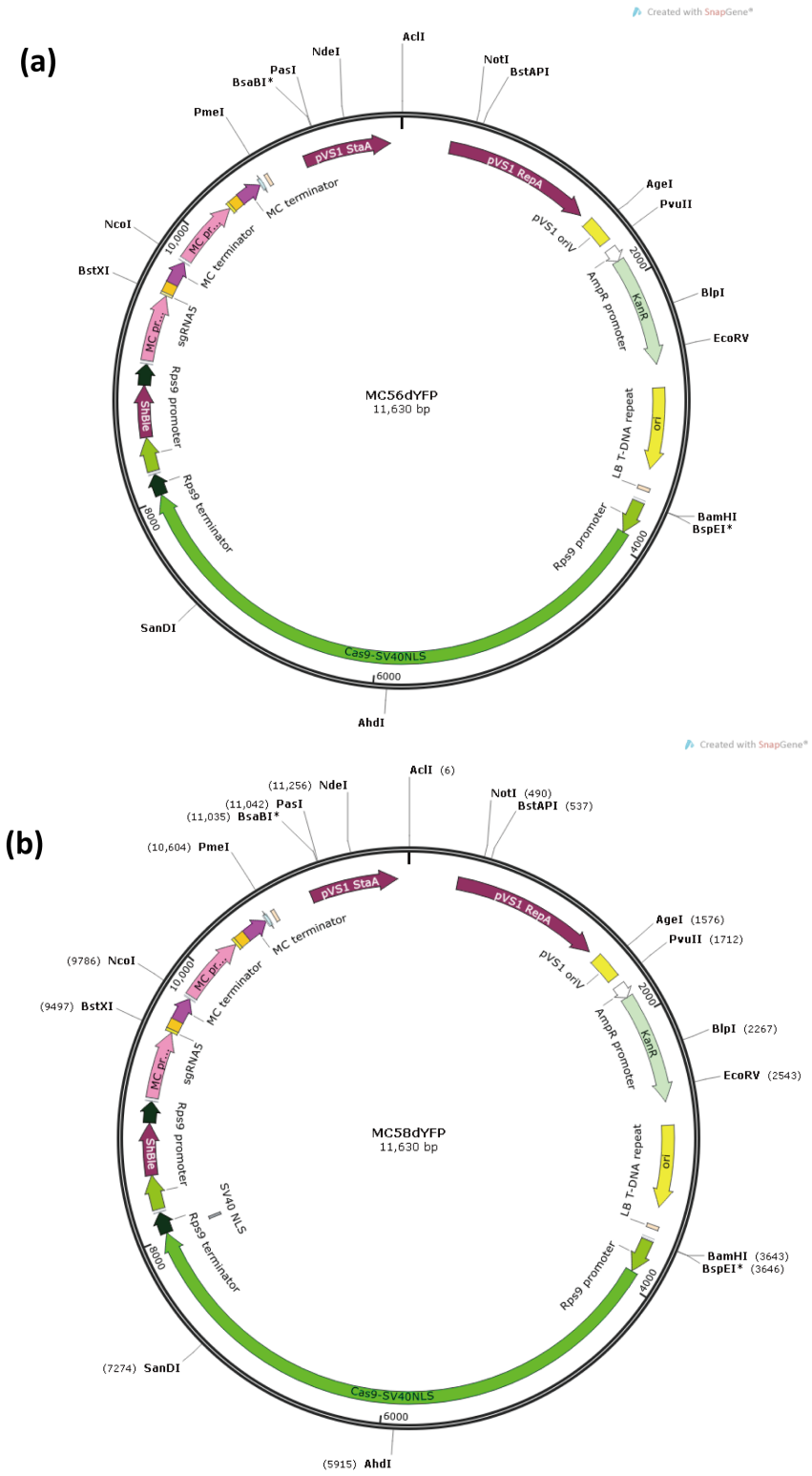


Figure 3.2: Knockout plasmid maps of (a) MC56dYFP and (b) MC58dYFP.

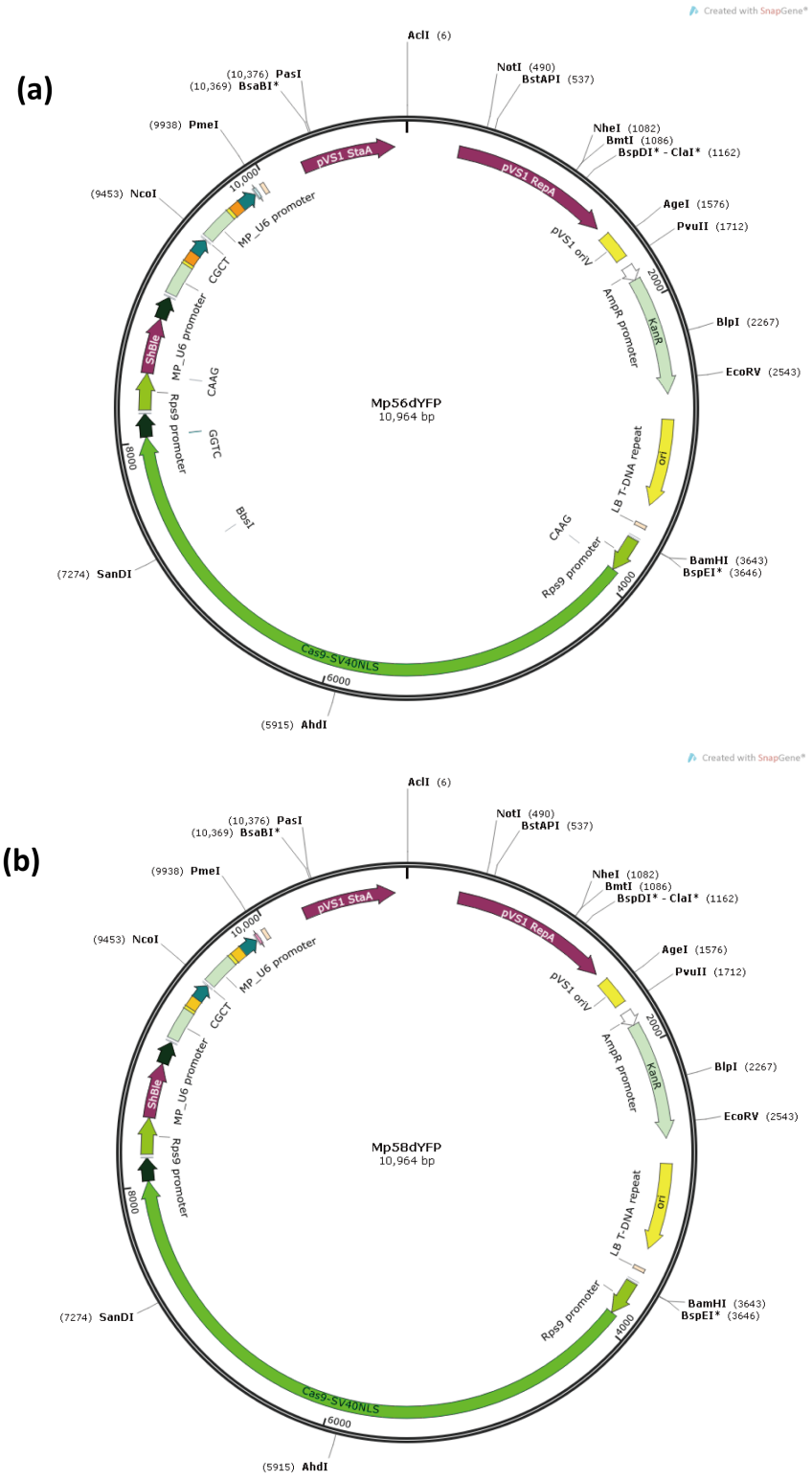


Figure 3.3: Knockout plasmid maps of (a) Mp56dYFP and (b) Mp58dYFP.

```
M.pusilla   GTTCATTCGTTGAACATCTGTTAAAATTGGAACGATACAGAGAAGATTAGCATG
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||
M.commoda   GTTCATTCGT-GAACATCTGTTAAAATTGGAACGATACAGAGAAGATTAGCATG

M.pusilla   GCCCCTGCGCAAGGATGACACTGCAACTCTGGGAAAGGAACCATATTTTTTT
            |||||||||||||||||||||||||||||||||||||||||||||||||||
M.commoda   GCCCCTGCGCAAGGATGACACTGCAACTCTGGGAAAGGAACCATATTTTTTT
```

Figure 3.4: Alignment of U6 snRNA sequence in *M. pusilla* and *M. commoda*

band corresponding to the cyanase gene fragment. The desired bands were excised from the gel and purified using a Zymoclean gel DNA recovery Kit. Purified DNA samples were subsequently sequenced by Eurofins.

3.3 Results

3.3.1 Identification of U6 snRNA

The U6 small nuclear RNA (snRNA) is conserved across diverse species (Figure 3.5) and typically initiates with a guanine (G) at the $5'$ end. Several conserved sequence motifs were identified, including `ATACAGAGAAGATTAGCA`, `CCCTC`, and `AAGGATGA`, which are characteristic of functional U6 snRNA.

The U6 snRNA sequence from *M. pusilla* was used as a query to identify the corresponding sequence in *M. commoda* (Figure 3.4). A candidate U6 snRNA sequence was identified with a highly significant E-value of 1.1×10^{-46}, indicating strong sequence conservation and homology.

Comparison of the two sequences reveals a high degree of conservation, with only minor nucleotide variation, supporting the identification of the *M. commoda* sequence as the U6 snRNA homolog.

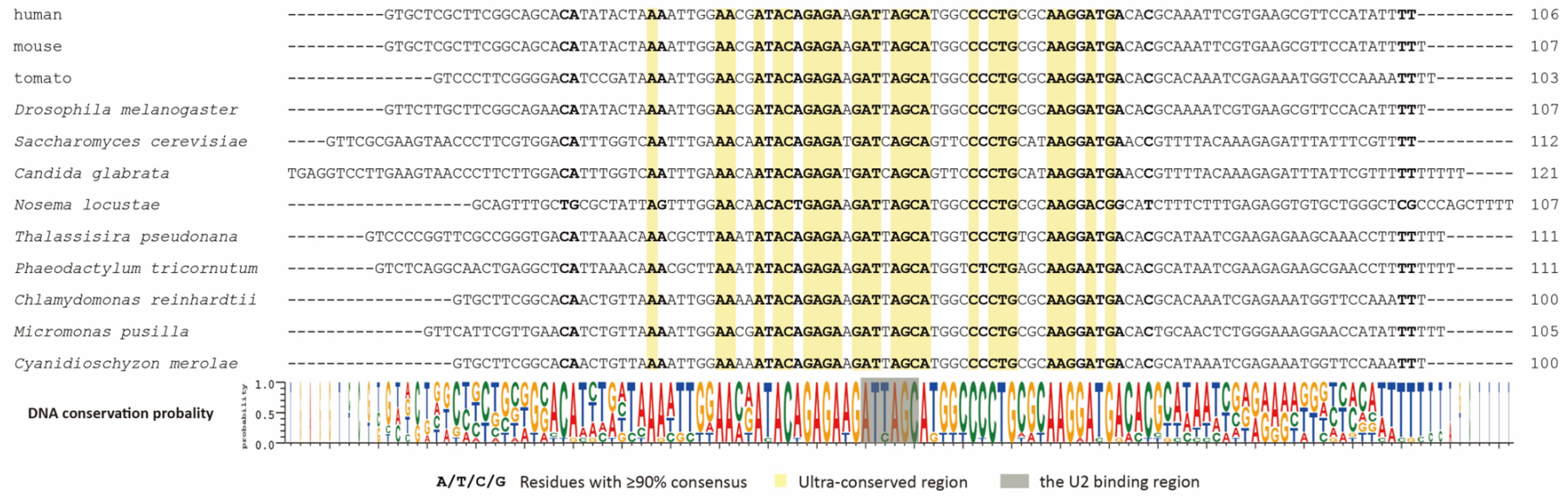


Figure 3.5: Alignment of U6 snRNA from representative species.

3.3.2 The design of sgRNA for *CynS* gene knockout

The *CynS* gene in *M. commoda* was searched and checked using NCBI and JGI databases, confirming that it is a single-copy gene lacking introns. Gene structure and predicted active sites are consistent with the conserved catalytic domain identified in *CynS* from representative species (Mao et al., 2022). Three key residues within the predicted active site including arginine (R), glutamic acid (E), and serine (S) are indicated in Figure 3.6a.

The consideration of designing Cas9 cleavage sites is effectively disrupting the three active sites within the C-terminal domain of the *CynS* gene. RGEN Cas-Designer was used to predict candidate sgRNA targets. The character of *Streptococcus pyogenes* Cas9 is specifically recognizing 20 nt target sequences adjacent to a PAM motif (NGG). Eight sgRNAs targeting regions both upstream and downstream of the domain were selected and aimed to knock out this functional region. No potential off-target sites were identified in the genome.

The *CynS* gene was amplified using primers F2 and R2 and subsequently cloned into the pCR4-TOPO vector to generate a cyanase template plasmid for the following *in vitro* test (Figure 3.6b). Restriction analysis using EcoRI, which has two recognition sites in the pCR4-TOPO vector, produced a fragment of the expected size (1069 bp), confirming successful cloning.

3.3.3 Testing of cleavage efficiencies

The plasmid of *CynS* gene template was linearized by *NcoI* digestion, yielding the expected fragment length of 4859 bp (Figure 3.7a). A DNA template encoding a T7 promoter, a 20-nt sgRNA target sequence, and an sgRNA scaffold was subsequently assembled (Figure 3.7b). Eight sgRNAs were synthesized and validated, each with the expected length of 99 nt (Figure 3.7c).

In vitro cleavage assays were performed by incubating the linearized *CynS* gene

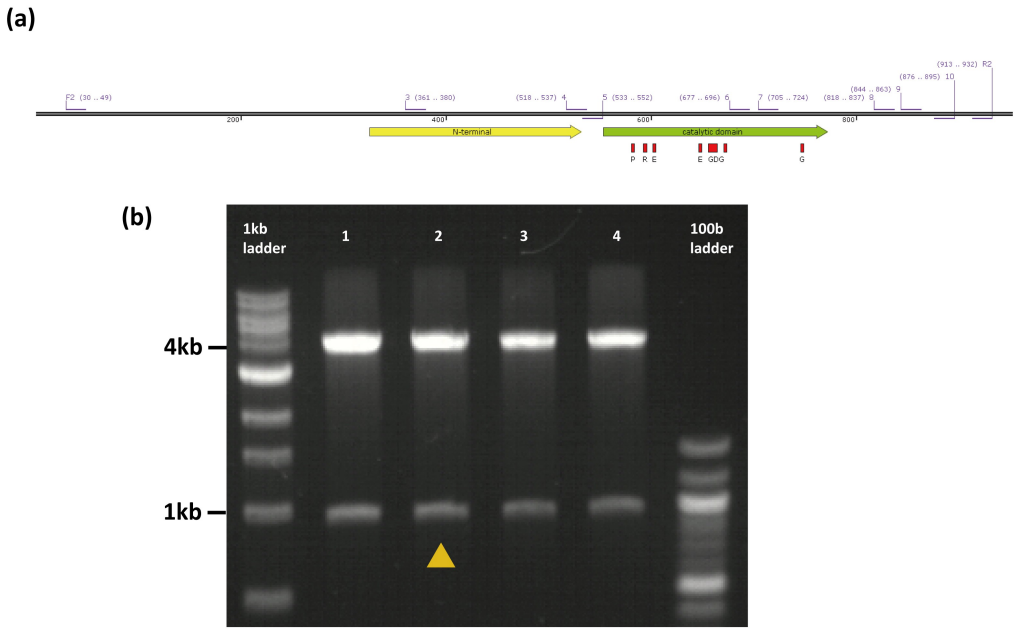


Figure 3.6: The *CynS* gene in *M. commoda*. (a) the map of the *M. commoda CynS* gene. Primers F2 and R2 were used to amplify the *CynS* gene template. sgRNAs 3–10 were designed for targeted genome editing. The yellow region represents the N-terminal segment of *cynS*. The green region denotes the catalytic domain of *cynS* and the red regions represent the ultra-conserved residues identified in Mao et al. (2022). (b) agarose gel results of plasmids F2-R2 pCR4-TOPO linearization checked by EcoRI digestion.

template with Cas9 enzyme and individual sgRNAs in sterile PCR tubes. The cleavage efficiency of each sgRNA was evaluated based on the intensity of the remaining undigested DNA bands and the sizes of the resulting cleavage fragments. The results indicated that sgRNA5, sgRNA6, sgRNA7, sgRNA8, and sgRNA9 exhibited relatively high cleavage efficiency (Figure 3.7d).

Since two sgRNAs are required to function as a pair for deletion of the target domain, the genomic positions of these five sgRNAs were further analyzed. Among them, only sgRNA5 was located upstream of the targeted domain, whereas sgRNA6, sgRNA7, sgRNA8, and sgRNA9 were located downstream (Figure 3.6a). Based on this positional arrangement, two sgRNA pairs were designed: sgRNA5&6 and sgRNA5&8, aiming to generate deletions of 158 bp and 299 bp, separately.

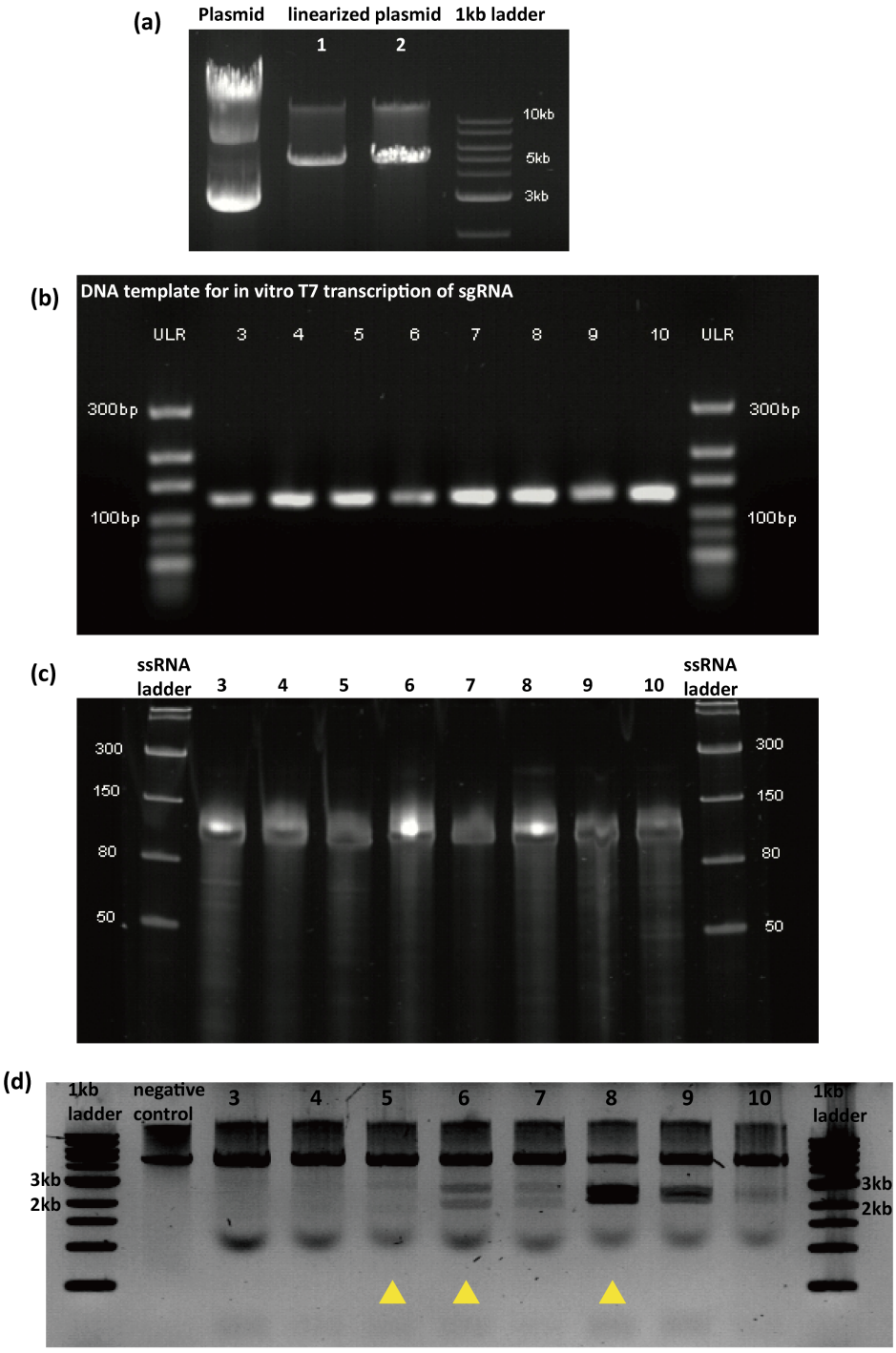


Figure 3.7: Results of in vitro test. (a) Agarose gel results of linearizated cyanase gene template plasmids digested by NcoI (4859bp). The marker used is a 1kb DNA ladder. (b) DNA template for *in vitro* T7 transcription of sgRNA. Bands 3–10 represent DNA template with designed sgRNA 3–10. The marker used is the ultra low range (ULR) DNA ladder.(c) sgRNA quality check by using 12% urea-TBE Polyacrylamide Gel Electrophoresis (PAGE). Bands 3–10 represent transcription of sgRNA 3–10. The marker used is the low range ssRNA ladder. (d) *In vitro* test results of sgRNAs and target sizes for successful cutting. Linearized plasmids of cyanase template without sgRNAs was set as a negative control. Yellow triangles indicate selected sgRNAs for following knockout plasmid constructs.

3.3.4 Construction of knockout plasmids

Plasmids for *CynS* gene knockout were constructed using Golden Gate assembly. In L1 assembly, essential expression cassettes were generated. Cassette containing the RPS9 promoter, *cas9*, *eYFP*, and the RPS9 terminator was assembled into the vector backbone pICH47742, resulting in the plasmid L1-RSP:cas9:*eYFP*. In this construct, *cas9* was directly fused to *eYFP*. Ten colonies were selected for plasmid extraction and screened by *EcoRI* digestion, of which five exhibited the expected fragment size of 9631 bp (Figure 3.8a). Colonies 1, 3, and 10 were further verified by sequencing, and only L1-RSP:cas9:*eYFP*–3 was confirmed to have the correct sequence. Subsequently, the *eYFP* was removed from plasmid L1-RSP:cas9:*eYFP*–3. The resulting constructs, L1-RSP:cas9–2 and L1-RSP:cas9–4, were both confirmed by sequencing to have successful *eYFP* removal and correct sequences (Figure 3.8b).

The sgRNA expression cassettes driven by U6 promoters from *M. commoda* and *M. pusilla* were assembled, separately. For constructs using the *M. pusilla* U6 promoter, MpU6:sgRNA5-4, MpU6:sgRNA6-2, and MpU6:sgRNA8-2 were sequence-verified as correct (Figure 3.8c, d). For constructs using the *M. commoda* U6 promoter, McU6:sgRNA5-1, McU6:sgRNA6-1, and McU6:sgRNA8-1 were confirmed to have correct sequences (Figure 3.8e).

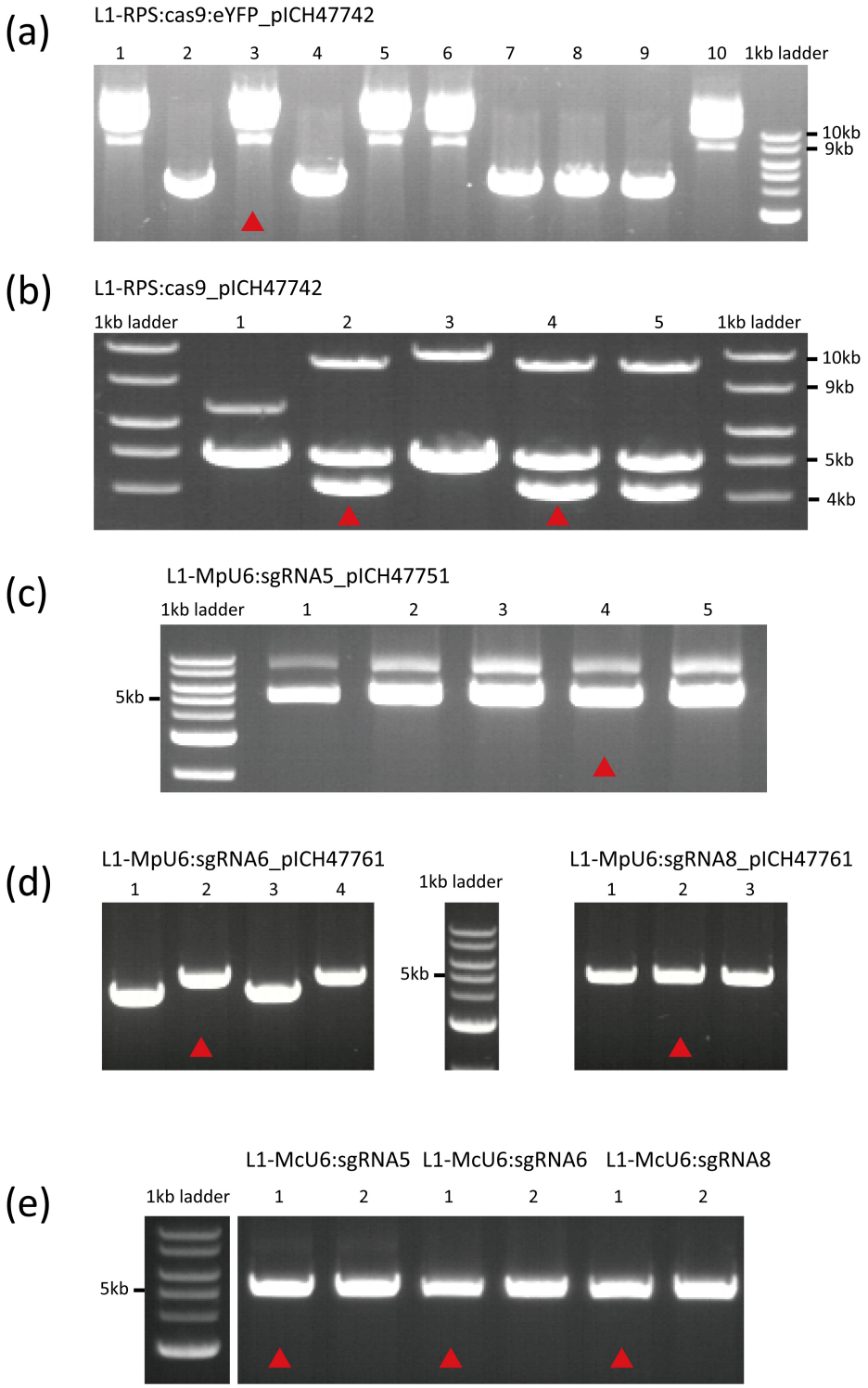


Figure 3.8: PCR results of L1 constructs for knockout plasmids. (a) Restriction digestion analysis of L1-RSP:cas9:*eYFP* using EcoRI. The estimated length is 9631bp; (b) Restriction digestion analysis of L1-RSP:cas9 using XmaI and MdeI. The estimated length is 4063bo and 5301bp; (c) Restriction digestion analysis of L1-MpU6:sgRNA5 using EcoRI. The estimated length is 4803bp; (d) Restriction digestion analysis of L1-MpU6:sgRNA6 and L1-MpU6:sgRNA8 using EcoRI. The estimated length of both plasmids are 4803bp; (e) Restriction digestion analysis of L1-McU6:sgRNA5, L1-McU6:sgRNA6 and L1-McU6:sgRNA8 using EcoRI. The estimated length of both plasmids are 5136bp. Red triangles indicate the correct constructs as confirmed by Sanger sequencing. DNA marker used was a 1kb ladder.

For L2 assembly, validated L1 cassettes, together with linker L4E, were assembled into the L2 backbone pAGM4723 to generate plasmids for *CynS* gene knockout in *M. commoda*.

For the construction of Mc5&6 and Mc5&8, the plasmid L1-RSP:cas9:*eYFP* was used. Nine colonies from each construct were selected for plasmid extraction and screened by *PstI* and *BamHI* digestion. Nearly all colonies exhibited expected DNA sizes of 4599 bp and 7751 bp (Figure 3.9a). Two colonies per construct were sequenced, confirming that Mc5&6-1, Mc5&6-2, and Mc5&8-1 had correct sequences, which could used for further transformation. Subsequently, the removal of *eYFP* was performed using Mc5&6-1 and Mc5&8-1 as templates, but the assembly efficiency was relatively low. Primers were designed to amplify the region from the latter part of *cas9* to the end of *shBle*. Successful removal of *eYFP* was expected to yield a fragment of 1655 bp. For construction of Mc5&6dYFP, only 4 out of 23 colonies showed the expected band size, and Mc5&6dYFP-16 was sequence-verified as correct (Figure 3.9b). For construct Mc5&8dYFP, only 2 out of 15 colonies showed the expected size, and Mc5&8dYFP-7 was confirmed by sequencing (Figure 3.9c).

For knockout constructs using the U6 promoter from *M. pusilla*, L1-RSP:cas9 without *eYFP* was used. The amplified fragment spanned from the latter part of *cas9* to the vector backbone, including part of cas9 cassette, *shBle* cassette and two sgRNA expression cassettes. For construction of Mp5&6dYFP, three colonies exhibited the expected size, and Mp5&6dYFP-10 was confirmed by sequencing (Figure 3.9d).

For the Mp5&8dYFP construct, assembly efficiency was low. A total of 96 colonies were screened by PCR, of which only four showed the expected band size. These candidates were further validated by digestion with *EcoRV* and *NcoI*. Mp5&8dYFP-53 and Mp5&8dYFP-62 were subsequently sequenced and confirmed to have correct sequences (Figure 3.9e).

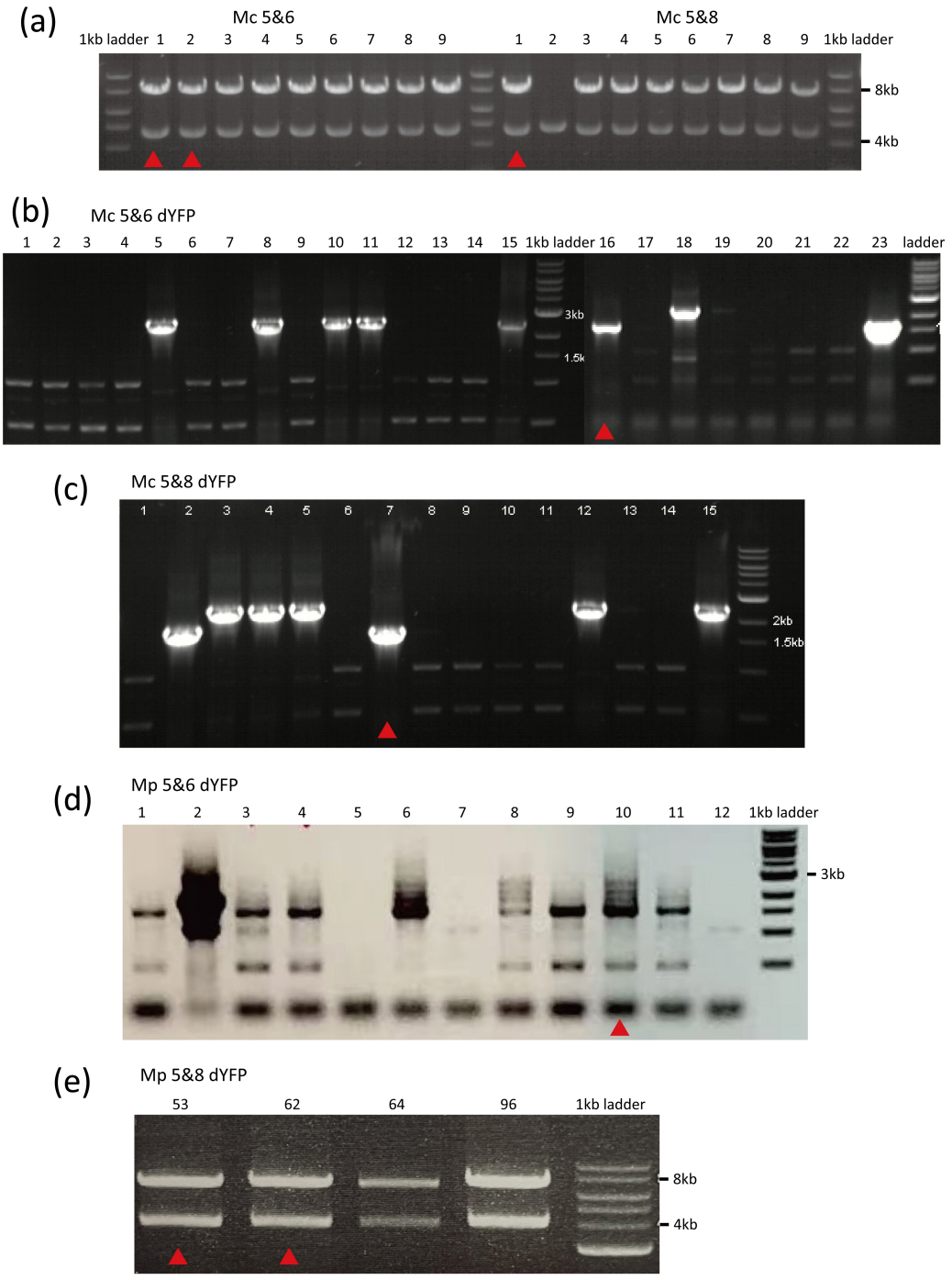


Figure 3.9: PCR results of L2 constructs for knockout plasmids. (a) Restriction digestion analysis of plasmid Mc5&6 and Mc5&8 using PstI and BamHI. The estimated DNA length are 4599bp and 7751bp. (b) PCR screening of *E. coli* colonies carrying plasmid Mc5&6dYFP, targeting DNA length of 1655bp. (c) PCR screening of *E. coli* colonies carrying Mc5&8dYFP, targeting DNA length of 1655bp; (d) PCR screening of *E. coli* colonies carrying Mp5&6dYFP, targeting band of 2798bp. (e) Restriction digestion analysis of Mp5&6dYFP using EcoRV and NcoI. The estimated length of both plasmids are 2053bp and 8911bp. Red triangles indicate the correct constructs confirmed by Sanger sequencing. DNA marker used was a 1kb ladder.

3.3.5 Screening and verification of knockout transformant

M. commoda is haploid. The knockout mutants can be directly isolated from primary clones. A total of 236 colonies were screened by colony PCR for CRISPR–Cas-mediated knockout, including 76 colonies from Mp5&8dYFP transformation, 76 from Mc5&8dYFP transformation, and 67 from Mc5&6dYFP transformation.

Two pairs of primers were designed to verify disruption of the cyanase gene. Initially, primers amplifying the full-length cyanase gene were used to screen colonies. However, the efficiency of deleting the designed fragment was low, and only a mutant C34 with a 76-nt deletion was identified (Figure 3.10a, d). Subsequently, a second pair of primers targeting a shorter region encompassing the designed knockout site was used for screening. However, no additional mutants were detected. Sanger sequencing confirmed that the mutation of C34 occurred at the sgRNA5 target site, resulting in a 76-nt deletion within the cyanase gene (Figure 3.10b). This deletion caused a frameshift mutation, leading to disruption of the cyanase domain (Figure 3.10c). Colony PCR also confirmed the presence of the *cas9* gene in nutant C34 (Figure 3.10e).

3.3.6 Identification of DNA damage repair genes in *M. commoda*

To investigate the DNA double-strand break repair (DDR) mechanism present in *M. commoda*, homologues of genes involved in NHEJ and HR were searched against the *M. commoda* RCC299 genome available in the JGI database (Table 3.3). Homologues of the major NHEJ-associated proteins, including the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and DNA ligase IV (LIG4). HR-associated proteins RAD51, XRCC2 and XRCC3 were detected.

The distribution of Ku70 and Ku80 homologues was further examined in the species included in the comparative analysis. Their presence was recorded separately for each species, as Ku70 and Ku80 were not assumed to be universally conserved across all examined taxa.



Figure 3.10: Screening and verification of mutant C34. (a) Screen of some colonies using three plasmid including Mp5&8dYFP, Mc5&8dYFP and Mc5&6dYFP. WT is positive control. The red triangles indicate C34 which potentially possess the edited target genes and further screen via Sanger sequencing. (b) Alignment of Sanger sequencing results with the cyanase gene from C34. (c) Alignment of amino acid sequence between WT and mutant C34. The green area marked the cyanase domain in WT sequence and the red font indicated the frameshift mutation in C34. (d) PCR results of cyanase gene in WT and C34. The DNA marker was a 1kb ladder. (e) PCR results of cas9 gene in C34. WT was a negative control.

3.4 Discussion

This study demonstrates the feasibility of applying CRISPR/Cas9 system in *M. commoda* to induce targeted genome modifications. The Cas9 nuclease from *Streptococcus pyogenes* (SpCas9) recognizes a 20-bp protospacer sequence adjacent to an PAM and introduces a blunt-ended DNA DSB approximately 3 bp upstream of the PAM site (Jinek et al., 2012). The break is predominantly repaired through

Table 3.3: DNA damage repair genes in *M. commoda*

DDR mechanism	Gene	JGI protein ID
NHEJ-associated	DNA-binding subunit of a DNA-dependent protein kinase (Ku70 autoantigen)	6114
	DNA-binding subunit of a DNA-dependent protein kinase (Ku80 autoantigen)	5294
	DNA ligase IV	9693
HR-associated	DNA repair protein RAD51/RHP55	(A) 3124, (B) 9261, (C) 264, (D) 4369
	XRCC2	2600
	XRCC3	6877

the errorprone NHEJ pathway following DSB in the absence of a repair template, resulting in INDELs at the target locus. These INDELs disrupt the coding sequence, frequently leading to frameshift mutations and loss of gene function. In this study, CRISPR/Cas9 mediated editing successfully generated mutations in the cyanase gene, confirming the functionality of this system in *M. commoda*.

Using CRISPR-Cas9 with a single sgRNA can introduce a variety of INDELs at a target locus (Nymark et al., 2016; Jiang et al., 2014). When two sgRNAs are employed, simultaneous cleavage at both target sites is required to achieve deletion of the intervening fragment. If the two sgRNA sites are not cleaved concurrently, independent mutations may occur at each site rather than excision of the fragment between them (Hopes et al., 2016). In this study, the design of two sgRNAs aimed to remove 299 bp fragment using sgRNA5 and sgRNA8, but only got one mutation with 76-nt deletion at the sgRNA5 site instead of the designed larger deletions (Figure 3.10b). Based on the *in vitro* assay, sgRNA8 exhibited significantly higher cleavage efficiency than sgRNA5 (Figure 3.7d), but the deletion observed in mutant C34 only occurred at the sgRNA5 target site (Figure 3.10b). Additionally, another eight randomly selected colonies were sequenced, none of which showed evidence of

INDELs. This discrepancy reflects the fundamental differences between *in vitro* and *in vivo* conditions. *In vitro* assays are conducted on naked DNA templates under optimal reaction conditions, but genome editing *in vivo* is affected by factors such as chromatin accessibility and sgRNA expression levels (Hsu et al., 2013). It is possible that the sgRNA8 target site is less accessible *in vivo* due to chromatin structure, leading to reduced cleavage efficiency.

Efficient deletion of a target fragment using paired sgRNAs requires the generation of simultaneous DSB at both target sites (Brooks et al., 2014; Hopes et al., 2016). In the absence of simultaneous cleavage, repair via the NHEJ pathway typically results in small INDELs at individual cut sites rather than large fragment deletions. The mutation observed in C34 suggests that DSB formation occurred only at the sgRNA5 site, indicating non-synchronous cleavage and a higher effective activity of sgRNA5 *in vivo* compared to sgRNA8. Although some studies have reported a positive correlation between *in vitro* and *in vivo* sgRNA efficiencies (Greiner et al., 2017; Mehravar et al., 2019), such relationships are not universally consistent. My results show the complexity of predicting sgRNA performance *in vivo* and highlight the importance of empirically testing multiple sgRNAs for each target locus. Future optimization could also include evaluating chromatin accessibility at target sites, refining sgRNA design using improved prediction algorithms, and increasing the number of sgRNA candidates to enhance the likelihood of achieving efficient and precise genome editing.

The CRISPR/Cas9 system requires both the Cas9 endonuclease and sgRNAs, which can be delivered into cells either via plasmid-based expression systems or as Cas9/single-guide RNA RNP. In this study, gene knockout was achieved by transforming plasmids encoding Cas9, two sgRNAs, and an antibiotic resistance cassette into cells, followed by random integration of these expression cassettes into the nuclear genome. However, this approach often results in relatively low gene editing efficiency due to variable transgene integration sites and inconsistent expression levels (Serif et al., 2018). Only one knockout mutant was obtained from 76 colonies following Mp5&8dYFP transformation in this study. Similarly low

editing efficiencies have been reported in early CRISPR/Cas9 applications in *C. reinhardtii* and *Nannochloropsis* spp., where approximately 300 *Nannochloropsis* colonies were screened but only two mutants were successfully generated (Jiang et al., 2014; Wang et al., 2016).

A disadvantage of vector-based transformation is the long-term expression of Cas9, which may increase the likelihood of off-target cleavage events. In contrast, RNP-based delivery represents a more efficient and precise genome editing strategy, as it enables transient Cas9 activity and reduces off-target effects (Serif et al., 2018; Yasui et al., 2026). In this approach, pre-assembled Cas9 protein and sgRNAs are directly introduced into cells, allowing genome editing to occur immediately without the need for intercellular transcription or translation (Zhang et al., 2021). As a result, RNP-mediated genome editing has been successfully applied across a wide range of microalgal species, including *C. reinhardtii* (Baek et al., 2016), *Nannochloropsis oceanica* (Naduthodi et al., 2019), *Chlorella vulgaris* (Kim et al., 2021), *Picochlorum celeri* (Krishnan et al., 2020), *Euglena gracilis* (Nomura et al., 2019), *Coccomyxa* sp. (Yoshimitsu et al., 2018), *Tetraselmis* sp. (Chang et al., 2020), *Phaeodactylum tricornutum* (Serif et al., 2018), and *Fistulifera solaris* (Yasui et al., 2026). It was reported that the RNP application in prasinophytes *Ostreococcus*, in which Cas9 accessed into cells and cleaved the target genomic DNA *in vivo* but no isolated knockout mutant cells (Kay, 2022). Given these advantages, an RNP-based delivery system may be a considerable method for improving genome editing efficiency in *M. commoda* and have significance of further applications.

DNA damage is inevitable and occurs continuously throughout the cell cycle. Endogenous factors, such as DNA replication errors, as well as exogenous mutagenic agents including ultraviolet (UV) radiation and ionizing radiation (IR), can induce substantial DNA damage (Kovács and Keresztes, 2002). Among the various forms of DNA lesions, DSBs are particularly lethal, which both strands of the DNA helix are severed, posing a significant challenge for accurate repair (Khanna and Jackson, 2001; Kastan and Bartek, 2004). To repair such damage, cells have evolved the DDR network that detects and repair DNA lesions (Sancar

et al., 2004).

The two principal pathways for the repair of DNA DSBs are NHEJ and HR (Onozawa et al., 2014). HR is generally regarded as an error-free mechanism because it utilizes a homologous DNA template to accurately restore the original DNA sequence (Sasaki et al., 2010). In contrast, NHEJ does not require a homologous template and instead directly ligates the broken DNA ends. It is inherently error-prone and results in INDELs at the repair junctions (Lieber, 2010).

The choice between HR and NHEJ is always regulated and determined by the phase of the cell cycle (Shrivastav et al., 2008; Nagaria et al., 2013), with additional species-specific differences. The eukaryotic cell cycle consists of four main stages: Gap 1 (G1), Synthesis (S), Gap 2 (G2), and Mitosis (M) (Matthews et al., 2022). During the G1 phase, cells grow and synthesize RNA and proteins required for DNA replication. In the S phase, DNA replication occurs, resulting in the duplication of chromosomes and the formation of sister chromatids. The G2 phase serves as a checkpoint stage in which cells verify the integrity of replicated DNA and prepare for mitosis. During the M phase, replicated chromosomes are segregated into two daughter nuclei, followed by cytokinesis. NHEJ does not require a homologous template and can function throughout the cell cycle, making it a major DSB repair pathway in many eukaryotic systems (Lieber, 2010; Lin et al., 2014). In contrast, HR is largely restricted to the late S and G2 phases, when DNA replication has been completed and sister chromatids are available to serve as templates for accurate repair (Heyer et al., 2010; Davis and Chen, 2013; Angstenberger et al., 2020).

In microalgae, the cell cycle is coupled to environmental cues such as light tensity, light–dark cycles and temperature (DuRand et al., 2002; Angstenberger et al., 2020; Kato and Nam, 2021; Pageault et al., 2024). For instance, in *M. pusilla*, cell division (M phase) is synchronized with the light-dark cycle, with the majority of mitotic events occurring during the first 8 hours of the dark period (DuRand et al., 2002).

Similarly, in *Tisochrysis lutea*, extension of the dark phase increases the proportion of cells entering the G2/M transition, indicating that cell cycle progression can be modulated by photoperiod length (Pageault et al., 2024). This temporal regulation of the cell cycle has important implications for genome editing efficiency, as the relative activities of HR and NHEJ are strongly influenced by the phase in which DNA damage occurs.

DDR mechanisms have not yet been directly characterized in *Micromonas*. The preference between the two major DSB repair pathways varies among organisms. For instance, yeast mainly rely on HR, but animal and plant cells prefer NHEJ (Manolis et al., 2001; Puchta, 2005; Stinson and Loparo, 2021). To gain insight into the DDR pathways potentially operating in *M. commoda*, genes involved in DNA repair were examined based on their conservation in related species and well-characterized model organisms. Key proteins involved in NHEJ include the DNA–dependent protein kinase (DNA-PK), the Ku70/80 heterodimer, DNA ligase IV (LIG4), and XRCC4 (Davis and Chen, 2013). DNA-PK is central to NHEJ in both programmed DNA rearrangements and the repair of accidental DNA breaks (Davis et al., 2014). It consists of a large catalytic subunit DNA-PKcs and the Ku70/80 heterodimer (Gottlieb and Jackson, 1993). The Ku70/80 is responsible for initial recognition and binding of DSBs, and homologues of these proteins have been identified in many plant species (Tamura et al., 2002; Gallego et al., 2003; Gu et al., 2014). LIG4 is a specialized DNA ligase that is considered essential for NHEJ and functions almost exclusively in this pathway (Grawunder et al., 1998). The DNA-PK, LIG4 and Ku70/80 homologues were identified in the *M. commoda* genome (JGI database), suggesting that core components of the NHEJ machinery are present in *M. commoda*. Genes associated with HR–mediated DNA repair including XRCC3 and RAD51 were also identified in the *M. commoda* genome (Bhattacharya et al., 2017), indicating that DSB repair in *M. commoda* may rely on a combination of the two mechanisms.

Applications of gene editing based on both NHEJ and HR have been reported in a wide range of microalgal species, including *T. pseudonana* (Hopes et al., 2016;

Belshaw et al., 2023), *C. reinhardtii* (Shin et al., 2016; Greiner et al., 2017; Angstenberger et al., 2020), *P. tricornutum* (Nymark et al., 2016; Moosburner et al., 2020), *N. oceanica* (Wang et al., 2016; Naduthodi et al., 2019), and *Cyanidioschyzon merolae* (Minoda et al., 2004; Tanaka et al., 2021). For example, CRISPR/Cas9-mediated genome editing was used to knock out the *urease* gene in *T. pseudonana* (Hopes et al., 2016), and CRISPR/Cas-mediated HR enabled efficient targeted gene replacement in the same organism (Belshaw et al., 2023). In eukaryotes, HR activity is often suppressed by the competing NHEJ pathway (Frit et al., 2014), which leads to a relatively low efficiency of precise gene targeting via HR in some microalgal systems (Lozano et al., 2014). As an alternative strategy for gene editing in microalgae, programmable endonucleases such as TALENs and CRISPR/Cas systems introduce DSBs at defined genomic loci and then facilitate precise and efficient gene editing via HDR in the presence of a homologous repair template (Gaj et al., 2013; Naduthodi et al., 2019; Belshaw et al., 2023). This provides a foundation for extending genome editing approaches in *M. commoda*, facilitating functional genomic studies.

4

Nitrogen utilization of *Micromonas commoda* and *Thalassiosira pseudonana*

4.1 Introduction

Nitrogen availability in the oceans is a key factor influencing the productivity and diversity of microalgae (Moore et al., 2013). Nitrogen limitation has been reported to constrain the growth of osmotrophic microalgae with cell sizes ranging from 2–15 μm in nutrient-depleted regions (Berthelot et al., 2025).

Nitrogen occurs in two major forms in the oceans, that is inorganic nitrogen and organic nitrogen. Dissolved inorganic nitrogen (DIN) including nitrate, nitrite and ammonium, is generally preferred by microalgae (Waser et al., 1998), but its availability is often limited in marine systems (McCarthy et al., 1997; Rees et al., 2006; Sharpe et al., 2023). In the euphotic zone, the mean concentration of nitrate is approximately 7 μM, whereas ammonium and nitrite concentrations are much lower, around 0.3 μM and 0.1 μM, respectively (Rees et al., 2006; Gruber, 2008; Sanz-Luque et al., 2015). Lower DIN concentrations are commonly observed in oligotrophic regions. For example, surface nitrate concentrations in the Atlantic Ocean are below detection limits (< 0.05 μM) at 70% of sampling stations, largely due to rapid uptake by phytoplankton (Rees et al., 2006). In contrast, dissolved

organic nitrogen (DON) is often present at higher concentration than DIN, even in nitrogen-limited regions (Bronk et al., 2007). Urea is a key component of the organic nitrogen pool and represents an important nitrogen source for some microalgae (McCarthy, 1972; Solomon et al., 2010). However, the bioavailability of other DON varies between microalgal species (Hu et al., 2014; Huang et al., 2020). Under conditions of limited nitrogen availability and competition for nitrogen resources, microalgae have evolved diverse strategies to utilize a wide range of nitrogen sources.

Cyanate has been recognized as an ecologically relevant nitrogen substrate in the ocean and may represent an alternative nitrogen source comparable in importance to urea in some marine environments (Widner et al., 2018; Mao et al., 2022). Cyanase, encoded by *cynS*, is the key enzyme involved in cyanate decomposition, but the organization and regulation of cyanate metabolism differ substantially among organisms. In prokaryotes, cyanate metabolism is commonly associated with *cyn* operons, although their composition and regulation vary between taxa. In *E. coli*, *cynT*, *cynS*, and *cynX* are arranged in an operon and are involved in cyanate transport and decomposition (Sung and Fuchs, 1988; Anderson et al., 1990). In cyanobacteria, a light-regulated *cynABDS* operon has been identified, in which *cynA*, *cynB*, and *cynD* together encode a putative ABC-type anion transport system, while *cynS* facilitates the metabolism of exogenous cyanate (Espie et al., 2007).

In eukaryotes, *cynS* homologues have been identified in a wide range of organisms, but the mechanisms underlying cyanate uptake, regulation, and utilization remain poorly understood (Elleuche and Pöggeler, 2008; Widner et al., 2018; Zarlenga et al., 2019). In marine microalgae, physiological evidence suggests that *P. donghaiense* can grow with cyanate as the sole nitrogen source (Hu et al., 2014). In addition, *cynS* expression responds significantly to nitrogen limitation in *Alexandrium fundyense* (Zhuang et al., 2015) and *Aureococcus anophagefferens* (Wurch et al., 2011). However, the presence or nitrogen-responsive expression of *cynS* does not necessarily demonstrate that cyanate can be assimilated as a

nitrogen source (Guérin et al., 2025). Therefore, physiological and functional characterization is required to determine whether *cynS* primarily contributes to cyanate assimilation and nitrogen metabolic regulation in different marine microalgae.

In microalgae, nitrate and nitrite are transported into cells via specific membrane transporters, among which members of the NRT2 family play a central role in high-affinity nitrate uptake (Forde and Cole, 2003; Smith et al., 2019). Once inside the cell, nitrate is reduced to nitrite by nitrate reductase (*NitR*) in the cytosol (Guerrero et al., 1981). Nitrite is subsequently transported into the chloroplast, where *NiR* catalyzes its reduction to ammonium (Sanz-Luque et al., 2015). In addition to inorganic nitrogen sources, urea can be imported via specific transport systems and metabolized to ammonium by either urease (*URE*) or urea amidolyase (UAL), depending on the species (Raven and Giordano, 2016; Liang et al., 2025). The resulting ammonium, whether derived from nitrate, nitrite, or urea, is then assimilated into carbon skeletons through the glutamine synthetase/glutamate synthase (GS/GOGAT) cycle, producing glutamate as a central nitrogen donor for cellular metabolism (Miflin and Lea, 1975). In addition to utilizing extracellular urea, marine photosynthetic stramenopiles possess the ornithine–urea cycle (OUC), which enables nitrogen recycling and contributes to enhanced nitrogen use efficiency (Allen et al., 2011). Beyond these core pathways, additional regulatory mechanisms coordinate nitrogen uptake and assimilation in response to environmental changes, including transcriptional control of transporter and enzyme genes as well as feedback regulation by intracellular nitrogen status. Figure 4.1 summarizes the complex nitrogen assimilation and metabolism in the diatom *P. tricornutum*.

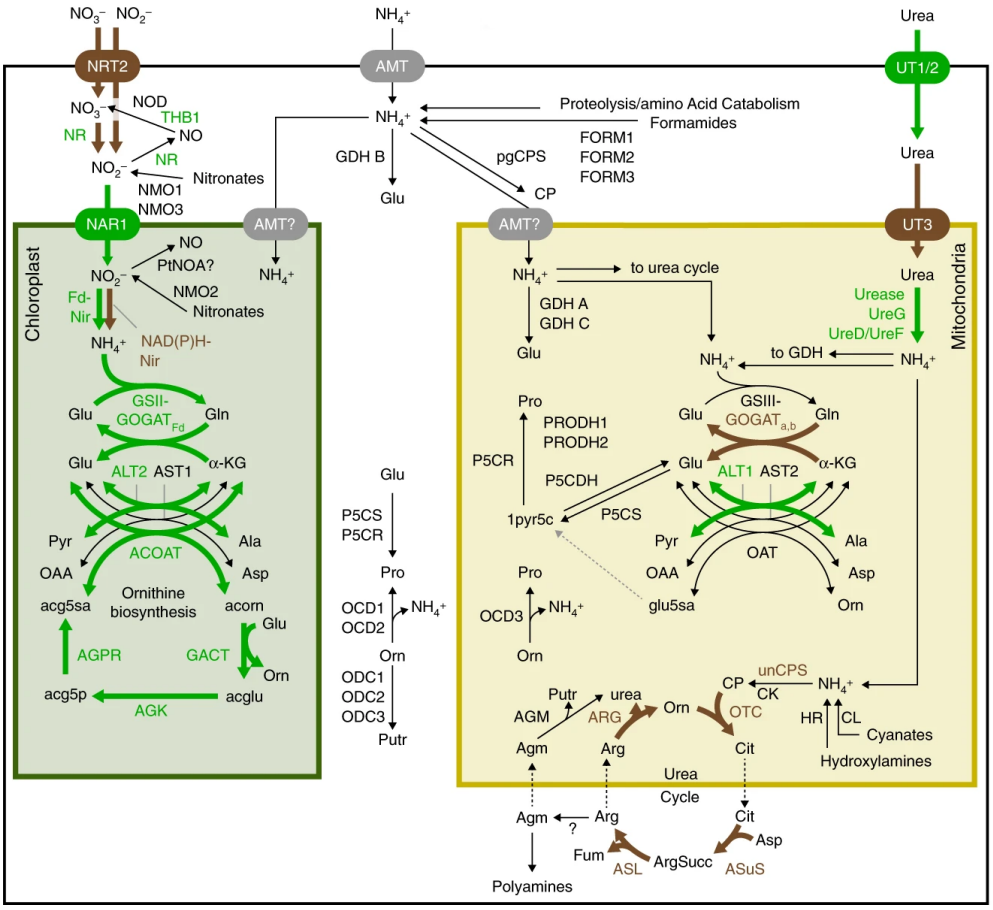


Figure 4.1: Schematic of compartmentation of major nitrogen assimilation, turnover, and links between the connected pathways in *P. tricornutum* (Smith et al., 2019).

4.2 Methods

4.2.1 Cell cultures for *Micromonas commoda* and *Thalassiosira pseudonana*

4.2.1.1 *Micromonas commoda*

The cell culture for *Micromonas commoda* is detail in 2.2.1.

M. commoda cells were grown at 20°C under a 14 h light / 10 h dark cycle with a photon flux density of 150 μmol photons m^{-2} s^{-1}. Cells at mid-exponential phase (Fv/Fm = 0.4) were collected by centrifugation (3000 $\times g$, 7 min), and the pellet was resuspended in 2×K medium without nitrogen (2K-N medium), in which all other

nutrients were supplied at double concentration except nitrogen. Cultures were monitored daily until the growth rate decreased to zero, indicating the completion of the nitrogen starvation experiment. For the add-back experiments, cultures were adjusted to initial concentrations of 1.5×10^6 cells/mL and 4×10^6 cells/mL. Starved cultures were diluted using filtered supernatant passed through a 0.22 µm filter and subsequently supplemented with nitrate, ammonium and cyanate to a final concentration of 100 μM, and urea to 50 μM.

4.2.1.2 *Thalassiosira pseudonana*

Thalassiosira pseudonana NitR mutants (N1 and N2) and the urease (*URE*) mutant were generated by Dr. Nigel Belshaw (Belshaw et al., 2023). The cyanase gene (*cynS*) knockout strain (C9M) was generated by Dr. Xuewei Mao. *T. pseudonana* wild-type (WT), U1, and C9 cells were cultured in synthetic ocean water (SOW) supplemented with Aquil nutrients, following the NCMA Aquil medium protocol. The *NitR* mutants (N1 and N2) were maintained in medium containing modified Aquil nutrients in which nitrate was replaced with nitrite.

For experiments of growth curve, *T. pseudonana* and mutant cultures were incubated at 20°C under continuous light with a photon flux density of 100 μmol photons $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$. Cells were grown in artificial seawater with modified Aquil nutrients containing 1 mM of the appropriate nitrogen source. Cells at mid-exponential phase were collected by centrifugation (5 min, 3000 $\times g$, 18°C), resuspended in nitrogen-free Aquil medium with all other nutrients at double concentration (2Aquil-N), and incubated for one day or until the Fv/Fm dropped below 0.1 depending on different experiments. For nitrogen add-in experiments, the N-starved cultures were adjusted to an initial concentration of approximately 2×10^5 cells/mL. Cultures were subsequently supplemented with nitrate, ammonium and cyanate to a final concentration of 100 μM, and with urea to 50 μM. A negative control (no-nitrogen) was included to verify that growth in the experimental conditions was due to the added nitrogen sources rather than

residual nitrogen in the medium. All treatments were established in triplicate.

4.2.2 Measurements of growth curves

The measurements of growth curve, growth rate and Fv/Fm for *M. commoda* and *T. pseudonana* are detailed in 2.2.2.

4.2.3 Stable isotope incubation of *T. pseudonana*

4.2.3.1 Preparation of ^{15}N enriched nitrogen stock

Isotope ratios are reported using delta (δ) notation in units of per mil (‰):

$$\delta^{15}N_{sample} = [\frac{(^{15}N/^{14}N)_{sample}}{(^{15}N/^{14}N)_{reference}} - 1] \times 1000 \quad (4.1)$$

where the $^{15}N/^{14}N$ reference is N_2 in air.

Three enriched nitrogen reagents including sodium nitrate (^{15}N , 98%), ammonium chloride (^{15}N , 99%) and urea ($^{15}N_2$, 98%) were obtained from Cambridge Isotope Laboratories, Inc. The nitrogen solution stocks of 0.1M enriched to 500‰ in ^{15}N were prepared using enriched compounds (^{15}N , 98%) and nitrogen compound at natural ^{15}N abundance and calculated based on the equation 4.1. To calculate the amount of enriched material (assumed as 100% enrichment instead of the 98 or 99% that is specified) that needs to be added to the reagent at natural abundance, the final composition of the solution is set as δ=500‰ and the ratio for the reference ($(^{15}N/^{14}N)_{reference}$) is 0.0037. Using these two values, the final ratio for the compound in solution ($(^{15}N/^{14}N)_{sample}$) is 0.00555. To make 200ml of 0.1M nitrogen solution stocks of enriched to 500‰ in ^{15}N, the amounts of compound at enriched abundance and that at natural abundance was listed in Table 4.1.

Table 4.1: The amount of enriched reagent (^{15}N , 98% or 99%) and reagent at natural abundance needed for making 200ml of 0.1M enriched nitrogen stock (δ=500‰).

Nitrogen source	Reagent at natural abundance	Reagent at enriched abundance
$NaNO_3$	1.7g	3.12mg
NH_4Cl	1.07g	1.88mg
urea	1.2g	2.14mg

4.2.3.2 Stable isotope incubation for *T. pseudonana*

Cultures of *T. pseudonana* wild-type (WT) were grown in SOW supplemented with Aquil medium in 5 L Erlenmeyer flasks. Incubation conditions and physiological measurements are detailed in Sections 4.2.1 and 4.2.2. Cells were pre-starved until Fv/Fm reached 0.1, after which nitrogen stock solutions were added to the medium.

For the urea and ammonium uptake experiments, cells were initially incubated in medium containing 1 mM nitrate at natural ^{15}N abundance. At mid-exponential phase, cells were collected, resuspended in 2×Aquil-N medium, and incubated until Fv/Fm = 0.1. The initial cell density was adjusted to 2×10^5 cells/mL. Enriched ammonium and urea (δ^{15}N = 500 ‰) stock solutions were added to achieve final concentrations of 100 µM ammonium and 50 µM urea, separately.

For the cyanate uptake experiment, enriched potassium cyanate (^{13}C, 99%; ^{15}N, 98%) could not be used due to limitations in the ^{13}C measurement range. Instead, the experiment was conducted in reverse: cells were first incubated with enriched nitrate (δ^{15}N = 500 ‰), resulting in a high intracellular ^{15}N concentration. After nitrogen limitation, cyanate at natural ^{15}N abundance was added to the medium.

6–7×10^7 cells were collected every 12h by centrifugation and rinsed with N-free Aquil medium for three times. Samples were freeze-dried at -106 ºC and 0.008 – 0.009 hPa for 24 hours in the geochemistry laboratory in University of East Anglia using a ScanVac CoolSafe 4-15 L freeze-dryer. Approximately 0.5mg of each

samples weweighted and nitrogen isotope abundance was then determined with Thermo Finnigan Delta plus XP isotope ratio mass spectrometer (IRMS) in Stable Isotope Platform in University of East Anglia. Specific uptake rates were calculated using equation 4.2 (Mulholland and Lee, 2009; Hu et al., 2014):

$$^{15}Nuptake(V) = \frac{(atom\%PN)_{final} - (atom\%PN)_{initial}}{(atom\%Nsource - atom\%PN)_{initial}} \times time \qquad (4.2)$$

where the source was the dissolved N pool that was enriched during uptake experiments.

4.2.4 Transformation and screening for *T. pseudonana* mutants

4.2.4.1 Biolistic transformation

The plasmid TPcynS 7&36 used for cyanase gene knockout was generated by Dr. Xuewei Mao (Figure 4.2). Plasmids were extracted using the Monarch Plasmid Miniprep Kit (NEB, UK).

Transformation was performed following the protocol described by Hopes et al. (2016). During transformation and subsequent screening, *T. pseudonana* WT cells were cultured in half salinity medium (1/2 SOW) containing 2× Aquil nutrients, and selected on 0.8% (w/v) 1/2 SOW agar plates supplemented with 2× Aquil nutrients. Prior to transformation, WT cells were maintained in 1/2 SOW Aquil liquid medium for at least five transfers in the exponential growth phase.

For transformation, WT cells were grown to a density of 1–1.5 × 10^6 cells/mL. Approximately 5 × 10^7 cells were collected by filtration onto a 0.2 μm membrane filter (Millipore) for each transformation and placed onto 1.5% agar shooting plates. Tungsten particles (0.7 mm) were coated with plasmid DNA (Figure 4.2) using 50 μL of 2.5 M $CaCl_2$ and 20 μL of 0.1 M spermidine. The coated particles were washed with 100% ethanol, resuspended in a final volume of 30 μL ethanol and

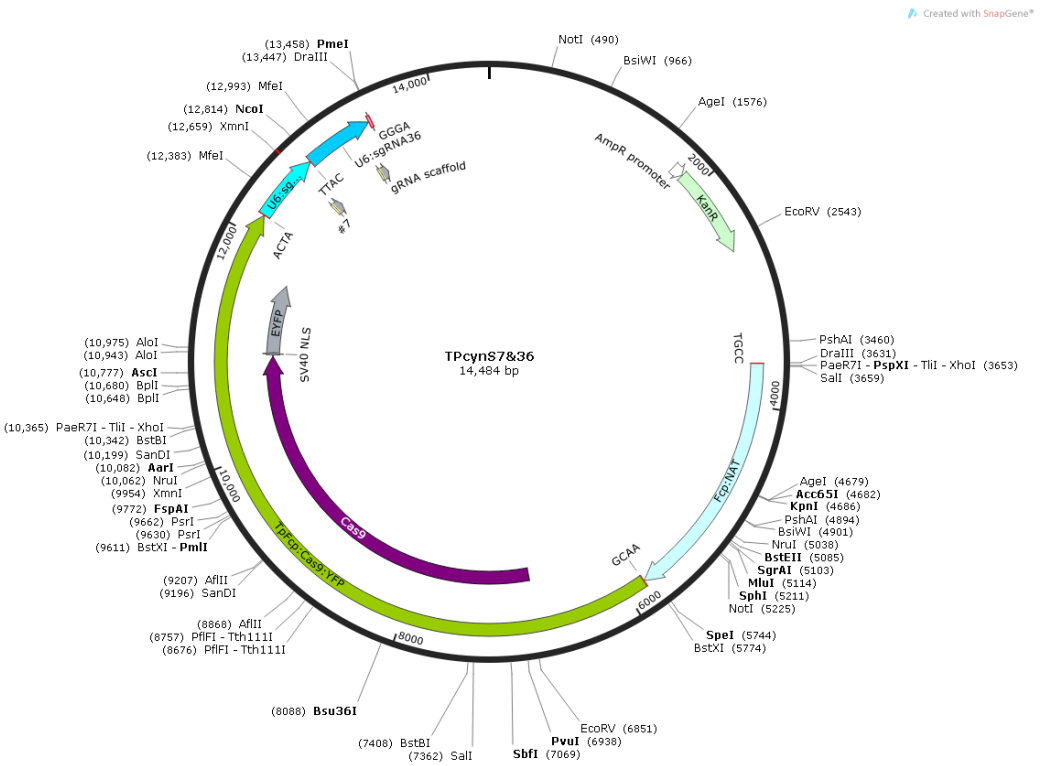


Figure 4.2: Vector map of TPcynS 7&36.

kept on ice until use. The PDS-1000/He Particle Delivery System (Bio-Rad) was sterilized and assembled, and the coated particles were loaded onto rupture discs (1335 or 1550 PSI). Cells were positioned 7 cm below the stopping screen in a vacuum chamber. Following transformation, filters were aseptically transferred into 25 mL of 1/2 SOW Aquil medium and incubated for 24 h under standard growth conditions. Cultures were then plated onto five individual 0.8% agar selection plates containing 100 μg mL^{-1} nourseothricin (NAT; Jena Bioscience, Catalogue #AB-102) and incubated at 20°C under continuous illumination until colonies appeared.

4.2.4.2 Selection and screen of transformed colonies

Single colonies were picked and resuspended in 2 mL of 1/2 SOW Aquil medium supplemented with 100 μg/mL NAT. Colony lysis was performed as described in Section 2.2.7.1. One millilitre of lysate was used as a template for PCR amplification of the CRISPR–Cas-targeted fragment of the *cynS* gene. Amplification was carried out using the forward primer

5′-GCCATCACTCAACACCAACAAG-3′ and the reverse primer 5′-CCCTCTCAGCAAACCATCCTAA-3′, generating a 708 bp fragment for the WT *cynS* gene and a 579 bp fragment for the knockout fragement. PCR products were analyzed on a 1% agarose gel to confirm fragment size.

Primary colonies were typically mosaic, consisting of a mixture of edited and wild-type cells, as previously reported (Hopes, 2017). Mosaic colonies showing evidence of editing were streaked onto selective agarose plates to obtain biallelic knockout subclones. Colonies from these subclones were PCR-confirmed and subsequently sent for Sanger sequencing, which verified biallelic editing. Confirmed knockout cell lines were maintained in 1/2 SOW Aquil medium supplemented with 100 μg/mL NAT.

4.2.4.3 Screening of *NitR* and *URE* mutants

The confirmation of *NitR* and *URE* mutants was performed as described in Belshaw et al. (2023). The *NitR* mutants N1 and N2 were screened using the forward primer 5′-CGGATATCAATGTGTTGGAGCACGAGA-3′ and reverse primer 5′-GGACCTTGATACGGAAGACTTGGTTGTTT-3′, which amplify a 510 bp fragment in WT cells. This fragment was absent in *NitR* mutants, as the target genomic region had been replaced by the FCP:NAT cassette.

The *URE* mutant U1 were screened using the forward primer 5′-CAAGGCGTTGATTGCCATGCAAA-3′ and reverse primer 5′-AACACACTTAAATCAGGAACAGGCAAA-3′, generating a 348 bp fragment in WT cells and a 322 bp fragment in the U1 mutant line.

4.3 Results

4.3.1 The nitrogen utilization strategies of *T. pseudonana* wild-type and mutant cells lines

4.3.1.1 Screening and verification of *cynS* knockout transformants

After 20 days of incubation on agar plates, fifteen colonies from primary clones were picked and screened for the *cynS* gene fragment by PCR (Figure 4.3a). The results showed that 46.7% of colonies (8 colonies) exhibited two bands, indicating mosaic populations with differing genotypes at a single locus. These eight mosaic colonies were then streaked onto selective agar plates for subcloning. A total of 32 colonies were subsequently picked and screened by colony PCR. Five colonies showed evidence of potential biallelic knockout, while 17 colonies exhibited mutations in only one allele. Biallelic knockout colonies C3, C31, and C32 were selected for Sanger sequencing, which confirmed them as homozygous knockouts of the edited *cynS* gene.

4.3.1.2 Screening and verification of *T. pseudonana* mutants

The generation of *NitR* mutants (N1 and N2) and the *URE* mutant (U1) by CRISPR/Cas-mediated homologous recombination (HR), as well as the initial confirmation were described in Belshaw et al. (2023). In both cases, the targeted loci were replaced by insertion of a resistance cassette (FCP:NAT), allowing mutant strains to grow in medium containing NAT.

To assess the long-term stability of these genome edits, PCR amplification of the targeted regions was performed on cultures maintained for five years post-transformation (Figure 4.4). In the *URE* mutant U1, one allele of the urease gene was replaced by the FCP:NAT cassette, while the second allele contained a 26 bp deletion, as confirmed by PCR and sequencing (Figure 4.4b, d). For the

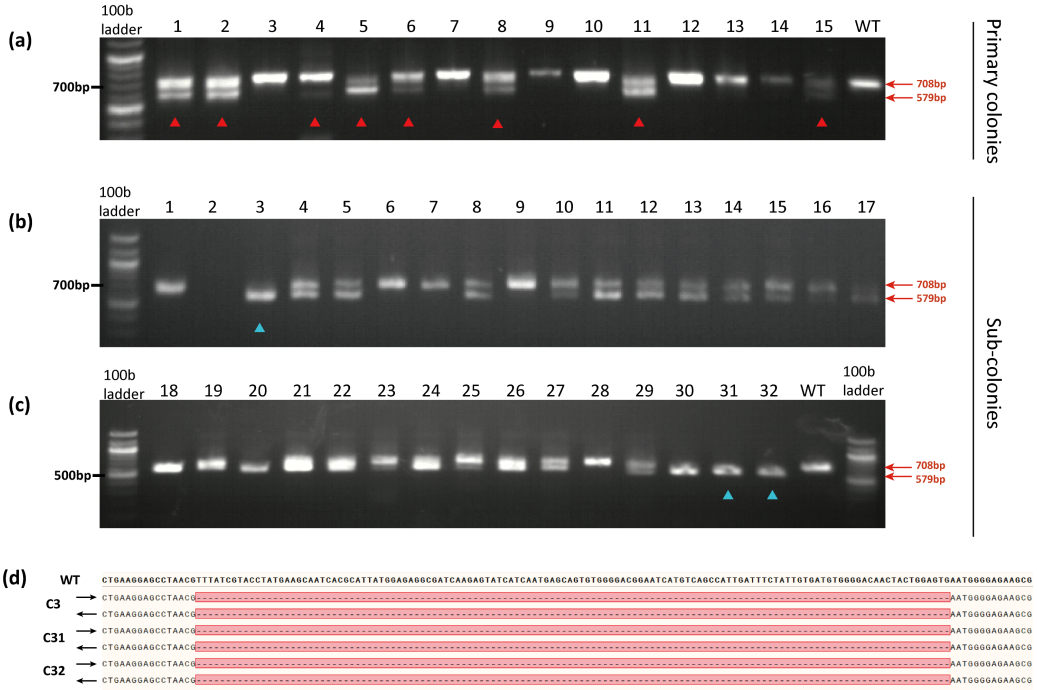


Figure 4.3: PCR of cyanase gene in primary and secondary clones. (a) primary colonies selected on Aquil plates containing 100 µg/ml nourseothricin. Red arrows indicate mosaic colonies containing wild-type and potentially successfully edited cell lines; (b,c) secondary colonies selected from primary colonies on Aquil plates containing 100 µg/ml nourseothricin. Blue arrows indicate colonies which possess the edited target genes with no wild-type copies, and were isolated and confirmed via Sanger sequencing; (d) Sanger sequencing results of C3, C31 and C32.

NitR mutants (N1 and N2), the targeted genomic region in both alleles was replaced by the FCP:NAT cassette, resulting in the absence of the expected PCR amplicon in mutant strains (Figure 4.4c). Similarly, the loss of the *cynS* gene in knockout strains C9M and C32 was verified by PCR, confirming a 128 bp deletion (Figure 4.4a, e).

The phenotyping of *NitR* and *URE* mutants were identified. *NitR* is essential for growth on nitrate as the sole nitrogen source. Both mutants N1 and N2 cannot grow in nitrate-containing medium and the cell size were significantly larger compared with negative control, indicating that nitrate reductase were inactivated in mutants (Figure 4.5a, b). Compared to negative control (N limited condition), the cell sizes of both *NitR* mutants were significantly larger and Fv/Fm were significantly lower in nitrate-containing medium after 2-day incubation, indicating the exist of nitrate also impacted the cell conditions. The loss of *URE* was significantly inhibited the

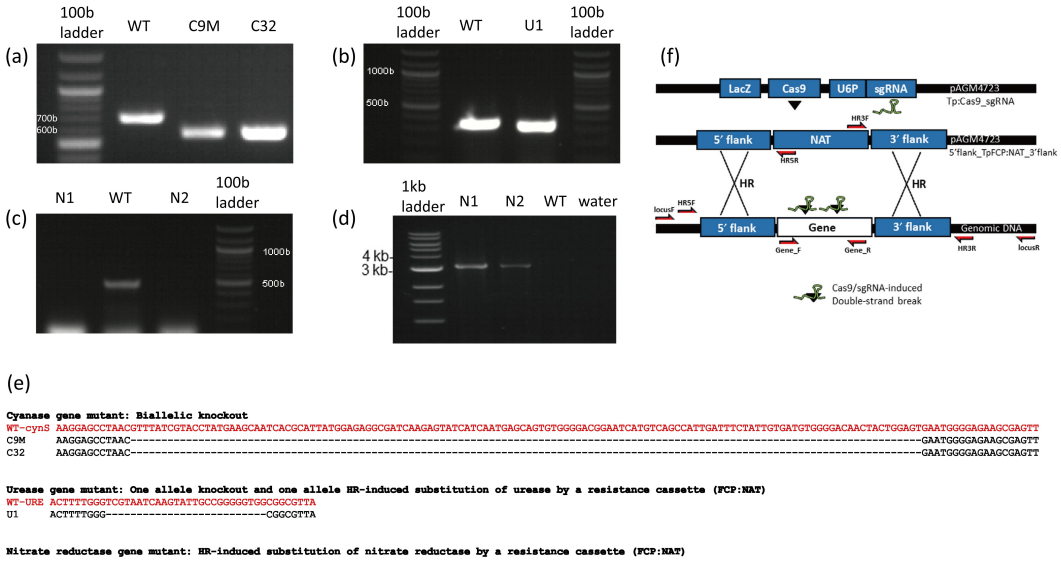


Figure 4.4: Screening of bi-allelic knockouts and homologous recombination (HR) events by PCR and Sanger sequencing. (a) PCR-based screening of bi-allelic knockout of the *cynS* gene. (b) Transformant U1 carried a 26-bp deletion in one allele, whereas the corresponding region in the other allele was replaced by the NAT resistance cassette through homologous recombination. The HR-modified allele was not amplified by the primers used to analyse the target region. (c) HR at the *NitR* gene region, where transformants N1 and N2 showed no PCR amplification, indicating bi-allelic replacement. (d) Confirmation of HR-mediated integration of FCP:NAT resistance gene cassette at target loci in by inverse PCR in *NitR* mutants (Belshaw et al., 2023). (e) Sanger sequencing confirmation of *cyanase* knockout transformants C9M and C32, as well as the 26-bp deletion in one allele of transformant U1. (f) Overview of the CRISPR/Cas9-mediated HR used to generate the *URE* and *NitR* mutants. The Tp:Cas9_sgRNA vector expressed Cas9 and the sgRNA, while the 5′flank_TpFCP:NAT:3′flank donor vector provided the homologous repair template for NAT cassette integration (Belshaw et al., 2023).

growth of U1 in urea, and its phenotyping was identical to N limited condition (Figure 4.5c).

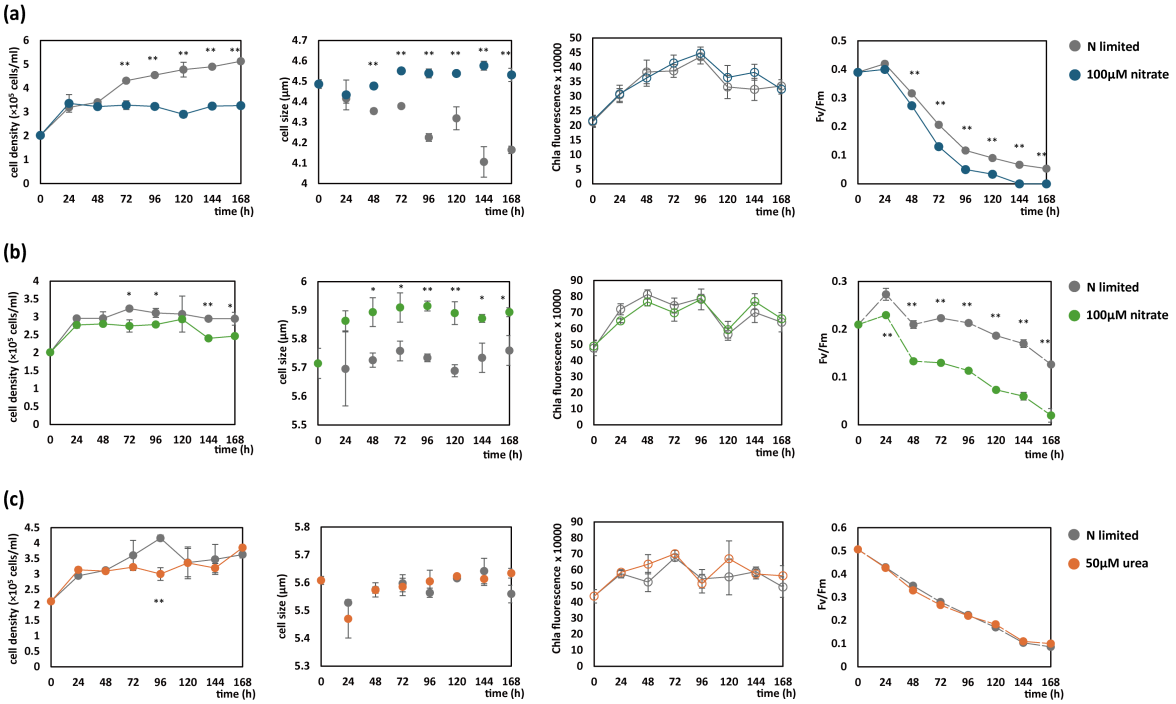


Figure 4.5: Physiological responses including cell density, Chla fluorescence, PSII Fv/Fm and cell size of NitR mutants N1 (a) and N2 (b) under 100 μM nitrate and N limited condition, as well as *URE* mutant U1 (c) under 50 μM urea and N limited condition. Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean $\pm$ SD (n = 3).

4.3.1.3 The nitrogen strategies of *T. pseudonana* wide-type cells

The growth of *T. pseudonana* WT cells in four nitrogen sources including nitrate, ammonium, urea and cyanate were measured (Figure 4.6). WT cells grew and reached peak cell densities between 60–96 h in all treatments except cyanate. The maximum cell densities in the nitrate and urea treatments ($1.88 \pm 0.15 \times 10^6$ cells/mL and $1.81 \pm 0.06 \times 10^6$ cells/mL, respectively) were significantly higher than those in ammonium ($1.51 \pm 0.10 \times 10^6$ cells/mL) and cyanate ($0.24 \pm 0.01 \times 10^6$ cells/mL) treatments (ANOVA, $p < 0.01$).

Fv/Fm values in nitrate, urea, and ammonium treatments recovered within 24 h, reached their maximum between 24–48 h, and subsequently decreased after 48 h, while corresponding cell densities continued to increase. The lag phase lasted 24 h for ammonium and urea, longer than the 12 h observed for nitrate. Cells in nitrate, urea, and ammonium entered the exponential growth phase between 36–60 h.

In the cyanate treatment, WT cell density was significantly lower than in other nitrogen treatments and began to diverge from the N-free control after 48 h (ANOVA, $p < 0.05$). Fv/Fm values remained below those in the N-free medium for the first 72 h and only began to increase after 108 h.

4.3.1.4 *T. pseudonana* growth and nitrogen uptake of cyanate

The growth of *T. pseudonana* WT was significantly inhibited in medium containing cyanate (Figure 4.7a,b). Fv/Fm values were even lower than those in the N-free treatment during the first 72 h. Cell size was significantly larger than in the N-limited treatment after 48 h of incubation.

To determine whether *T. pseudonana* could take up or assimilate cyanate, stable isotope analysis was performed. Cells were first incubated with enriched nitrate to the exponential phase and subsequently subjected to nitrogen starvation. The initial δ^{15}N of the cells was 339.7 ± 8.5‰. When incubated with cyanate ($\delta^{15}N_{\text{cyanate}} =$

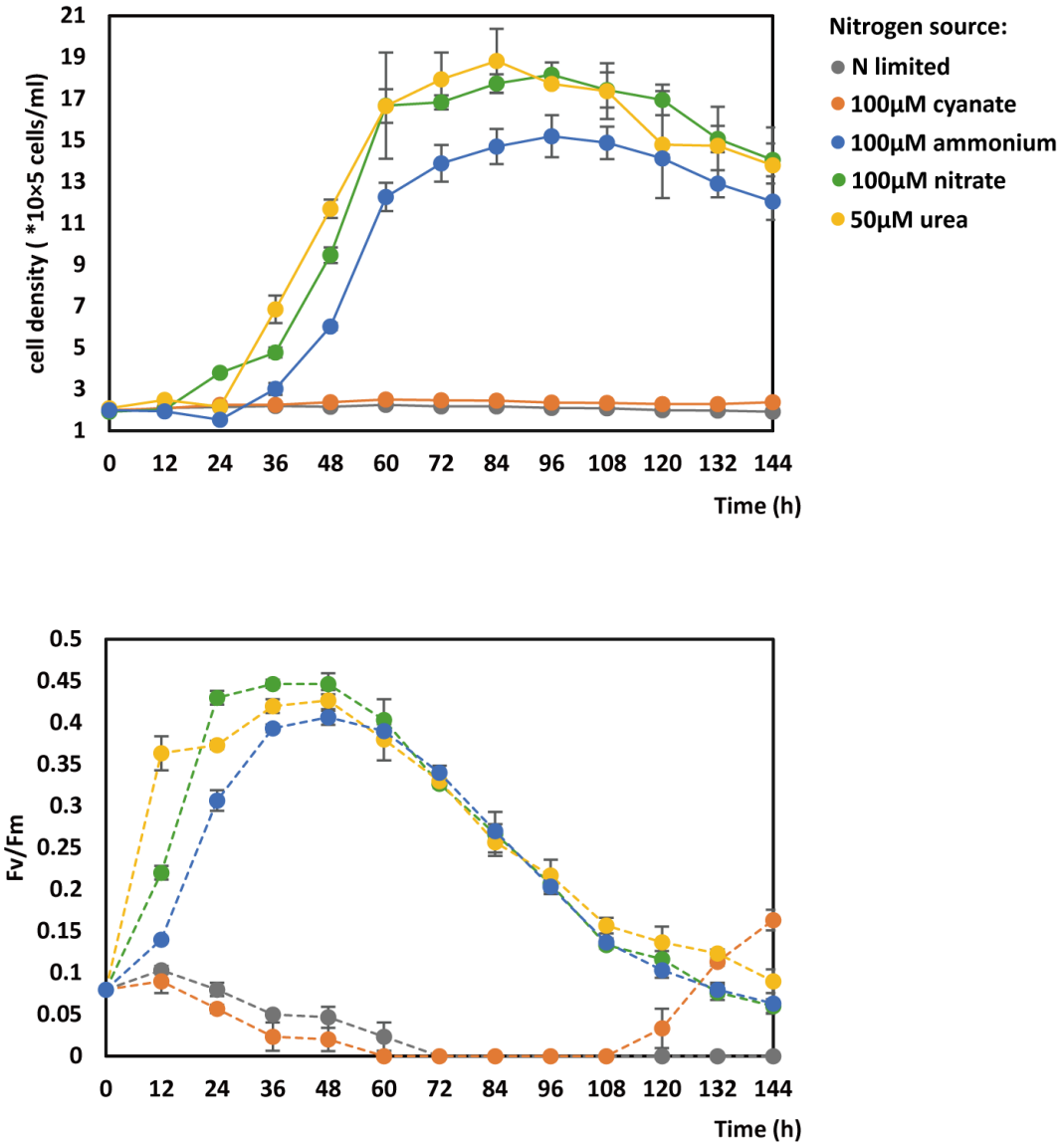


Figure 4.6: Physiological responses of *T. pseudonana* to N-limitation and versatile nitrogen sources, including cell density (a) and PSII Fv/Fm (b). Error bar represents mean ± SD (n = 3).

−4.1575‰) for 72 h, the δ^{15}N of cells decreased slowly, with no significant differences observed between sampling time points (Figure 4.7e). Specific uptake rates remained below 0.002 h^{-1} at all time points within 60 h, significantly lower than those observed for ammonium and urea treatments (Figure 4.7f).

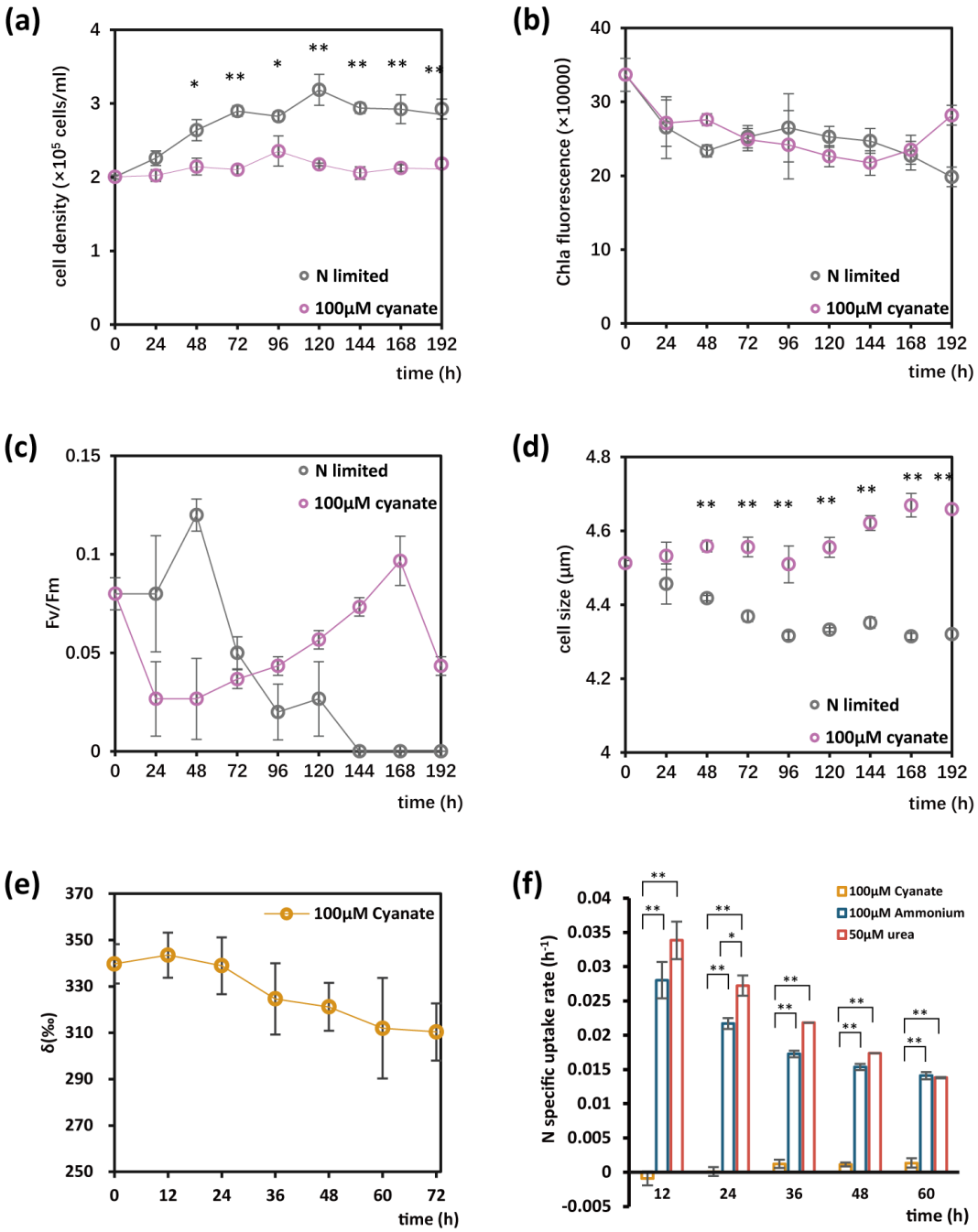


Figure 4.7: Physiological responses of *T. pseudonana* to 100 μM cyanate. Cell density (a), Chla fluorescence (b), PSII Fv/Fm (c) and cell size (d) under 100 μM cyanate and N limited condition. (e) The isotopic composition δ values of *T. pseudonana* wide-type cells in 100 μM cyanate. (f) The specific uptake rate of 100 μM cyanate, 100 μM ammonium and 50 μM urea. Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean $\pm$ SD (n = 3).

4.3.1.5 The growth conditions of *T. pseudonana* wild-type and mutant cell lines

To compare the effects of the loss of three nitrogen metabolism-related genes in *T. pseudonana*, WT, *cynS* knockout strains (C9M and C32), *NitR* mutants (N1 and N2), and the *URE* mutant (U1) were incubated in nitrite, a nitrogen source commonly used by these mutants.

Under deep nitrogen starvation (Fv/Fm <0.1), all strains exhibited their highest growth rate at 48 h (Figure 4.8d). The two *cynS* knockout strains (cell size > 5 μm) were significantly larger than WT (4.3–4.56 µm, $p < 0.01$) (Figure 4.8a). Their peak cell densities were significantly lower than WT (p <0.01), while Fv/Fm values showed no significant differences between *cynS* knockouts and WT. The two *NitR* mutants displayed distinct phenotypes (Figure 4.8b). N1 was similar to WT, whereas N2 exhibited a significantly larger cell size and a lower peak cell density compared to WT ($p < 0.01$). The *URE* mutant U1 (cell size > 5.7 μm) also showed significantly larger cell size and reduced growth rate compared to WT ($p < 0.01$) (Figure 4.8c).

Under one day of nitrogen starvation, Fv/Fm values for all strains were higher than 0.35, except for N2 (Fv/Fm = 0.2). All mutants reached their highest growth rate at 24 h, whereas WT peaked at 48 h (Figure 4.9d). Other phenotypic traits were consistent with those observed under deep nitrogen starvation.

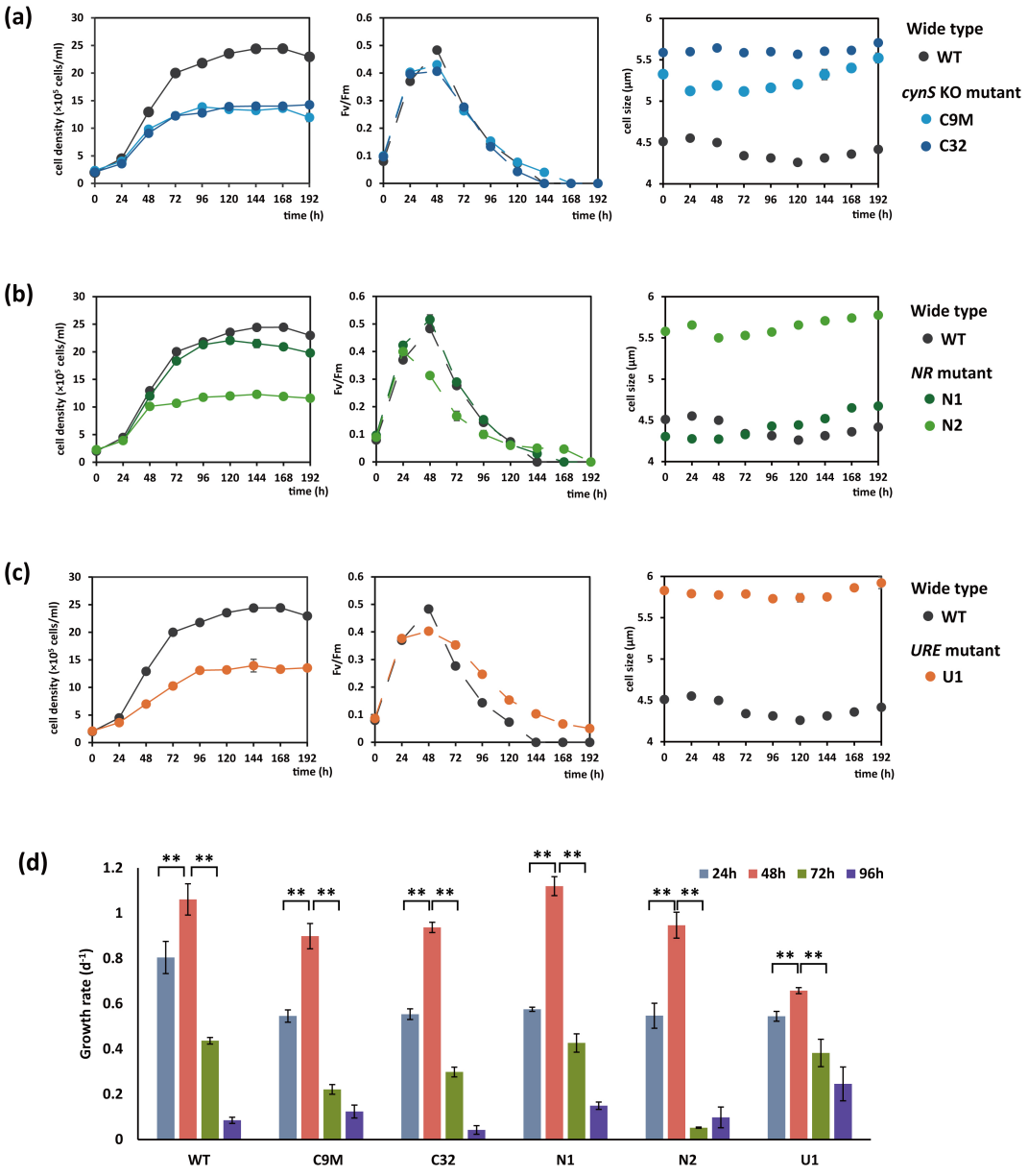


Figure 4.8: After deep nitrogen limited treatment, physiological responses of *T. pseudonana* and its *cynS* mutants (a), *NitR* mutants (b) and *URE* mutant (c) to 100 μM nitrite, as well as growth rate in four days (d). Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean $\pm$ SD (n = 3).

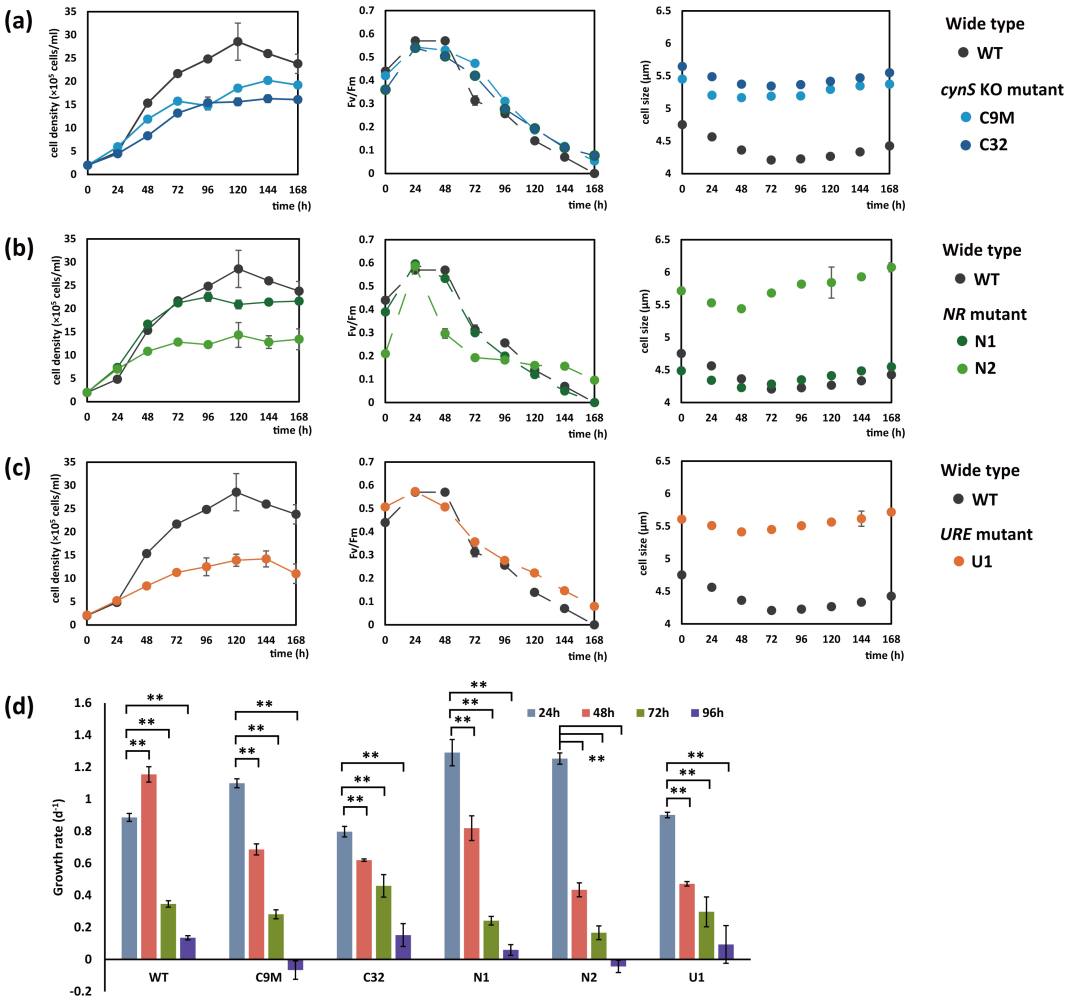


Figure 4.9: After one-day nitrogen limited treatment, physiological responses of *T. pseudonana* and its *cynS* mutants (a), *NitR* mutants (b) and *URE* mutant (c) to 100 μM nitrite, as well as growth rate in four days (d). Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean $\pm$ SD (n = 3).

4.3.2 The nitrogen strategies of *Micromonas commoda*

4.3.2.1 Growth conditions of *Micromonas commoda* in different nitrogen sources

When the initial cell density was set at both 1.6 $\times$ 10^6 cells/mL and 4 $\times$ 10^6 cells/mL, *M. commoda* grew well on inorganic nitrogen sources (nitrate and ammonium) (Figure 4.10). The ability of *M. commoda* to utilize urea was relatively limited. *M. commoda* was able to grow on cyanate, but the growth

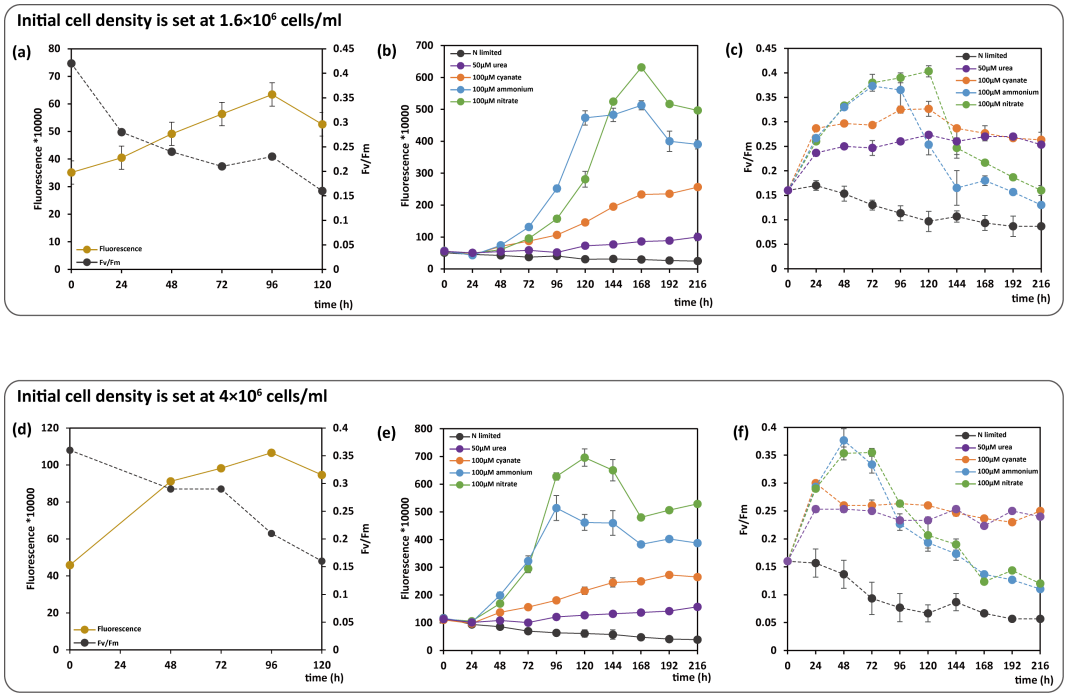


Figure 4.10: Physiological responses of *M. commoda* under different nitrogen sources. After nitrogen limited treatment (a), The physiological responses including Chla fluorescence (b) and Fv/Fm (c) with initial cell density at 1.6×10^6 cells/mL. After nitrogen limited treatment (d), The physiological responses including Chla fluorescence (e) and Fv/Fm (f) with initial cell density at 4×10^6 cells/mL. Error bar represents mean ± SD (n = 3).

pattern differed from that observed on nitrate and ammonium. The exponential growth phase was not clearly defined. Growth rates on cyanate remained low but relatively stable, and cell densities did not reach a maximum even after nine days of incubation. Based on the cyanate growth curve, *M. commoda* is capable of direct cyanate utilization. The observed slow growth rate (0.2–0.3 d^{-1}) and relatively stable Fv/Fm values (approximately 0.3 over eight days) may reflect limitations in cyanate uptake or the *CynS* activity.

4.3.2.2 Evidence of *M. commoda* using cyanate instead of ammonium

Cyanate can spontaneously decompose to ammonium when dissolved in water. It has been reported that the decomposition rate is 19.3 nM/h in 10 µM cyanate at 25 °C (Kamennaya et al., 2008). Based on this rate, approximately 4.8 µM/day of ammonium would be produced in a 100 µM cyanate medium. To simulate this

decomposition, a treatment was established in which 5 µM ammonium was added daily.

Results showed that growth in 100 µM cyanate was significantly higher than in the 5 µM ammonium daily treatment after 72 h, indicating that *M. commoda* was utilizing cyanate directly rather than the ammonium derived from its decomposition (Figure 4.11a). Fv/Fm values were similar between the 100 µM cyanate and 5 µM ammonium daily treatments, but both were lower than in the 100 µM ammonium treatment (Figure 4.11b). Microscopy of *M. commoda* cultures using DAPI staining indicated low bacterial contamination (Figure 4.11c). Rod-shaped bacteria (0.5-1μm) were only occasionally observed, with an estimated abundance of fewer than one bacterial cell per 100 algal cells. Bacteria represented less than 1% of the total cell population in *M.commoda* culture.

4.3.2.3 Growth kinetics for growth on cyanate

To further investigate the availability and potential toxicity of cyanate to *M. commoda*, cyanate was added to the medium at concentrations ranging from 0 to 300 µM. The results demonstrated that cyanate can serve as the sole nitrogen source for *M. commoda*, and its effect on growth is concentration-dependent (Figure 4.12).

Under cyanate concentrations up to 150 µM, the maximum growth rate increased with increasing cyanate concentration, indicating that cyanate effectively supports growth at moderate levels (Figure 4.12c). At higher concentrations (200 and 300 µM), although growth rates eventually exceeded those of the control, initial growth was significantly delayed, suggesting that the toxic effects of cyanate initially outweighed its benefits as a nitrogen source. These findings indicate that *M. commoda* can utilize cyanate directly, but higher concentrations impose physiological stress that slows early growth. Furthermore, Fv/Fm measurements revealed that photosynthetic efficiency was transiently reduced at 300 µM cyanate, consistent with inhibition of initial growth, but recovered during later stages of

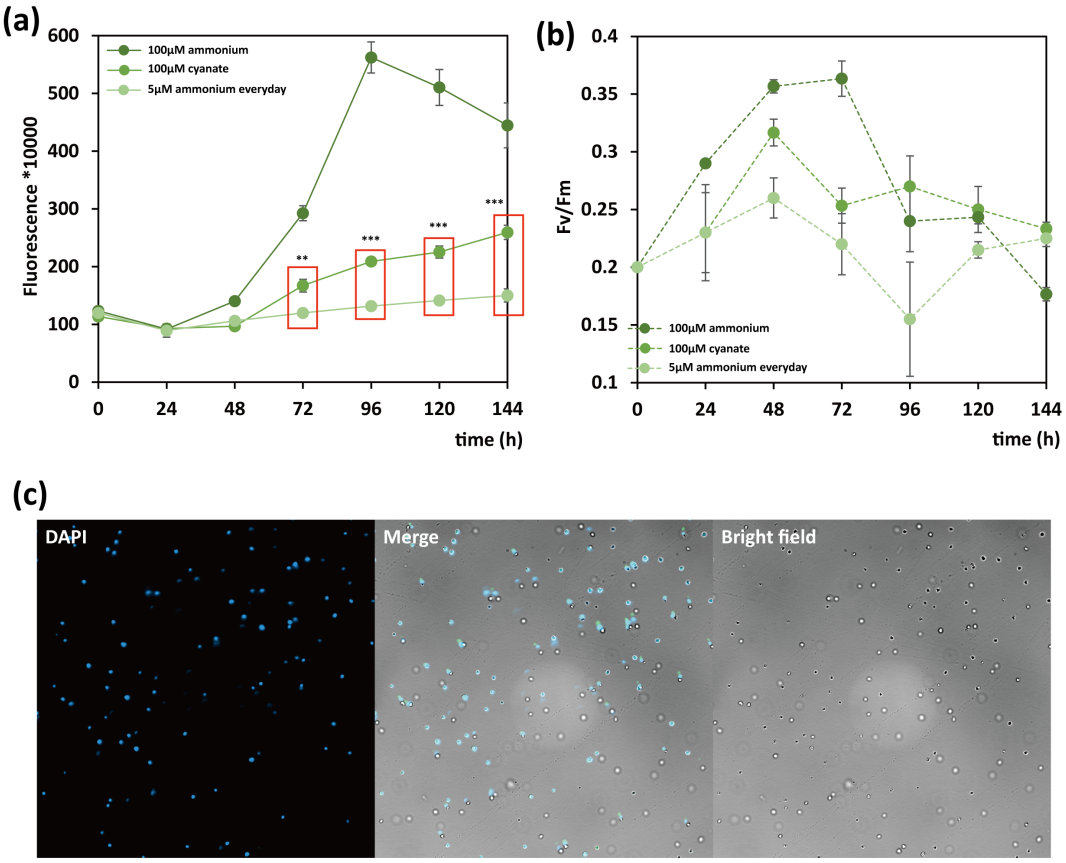


Figure 4.11: Physiological responses of *M. commoda* under 100 μM cyanate and ammonium, as well as 5 μM ammounium added everyday. (a) Chla fluorescence; (b) Fv/Fm; (c) microscopy images of *M.commoda* culture. Statistical significance is denoted as follows: $^{*}p < 0.05$ and $^{**}p < 0.01$. Error bar represents mean ± SD (n = 3).

growth (Figure 4.12b). This supports the hypothesis that cyanate toxicity is dose-dependent and that *M. commoda* can acclimate to moderate cyanate stress over time.

4.4 Discussion

4.4.1 Nitrate reductase in *T. pseudonana*

Nitrate is a major nitrogen source for marine microalgae, particularly diatoms (Lomas and Glibert, 2000). In diatoms, their capacity for rapid nitrate uptake and intracellular storage confers higher growth rates and a competitive advantage over

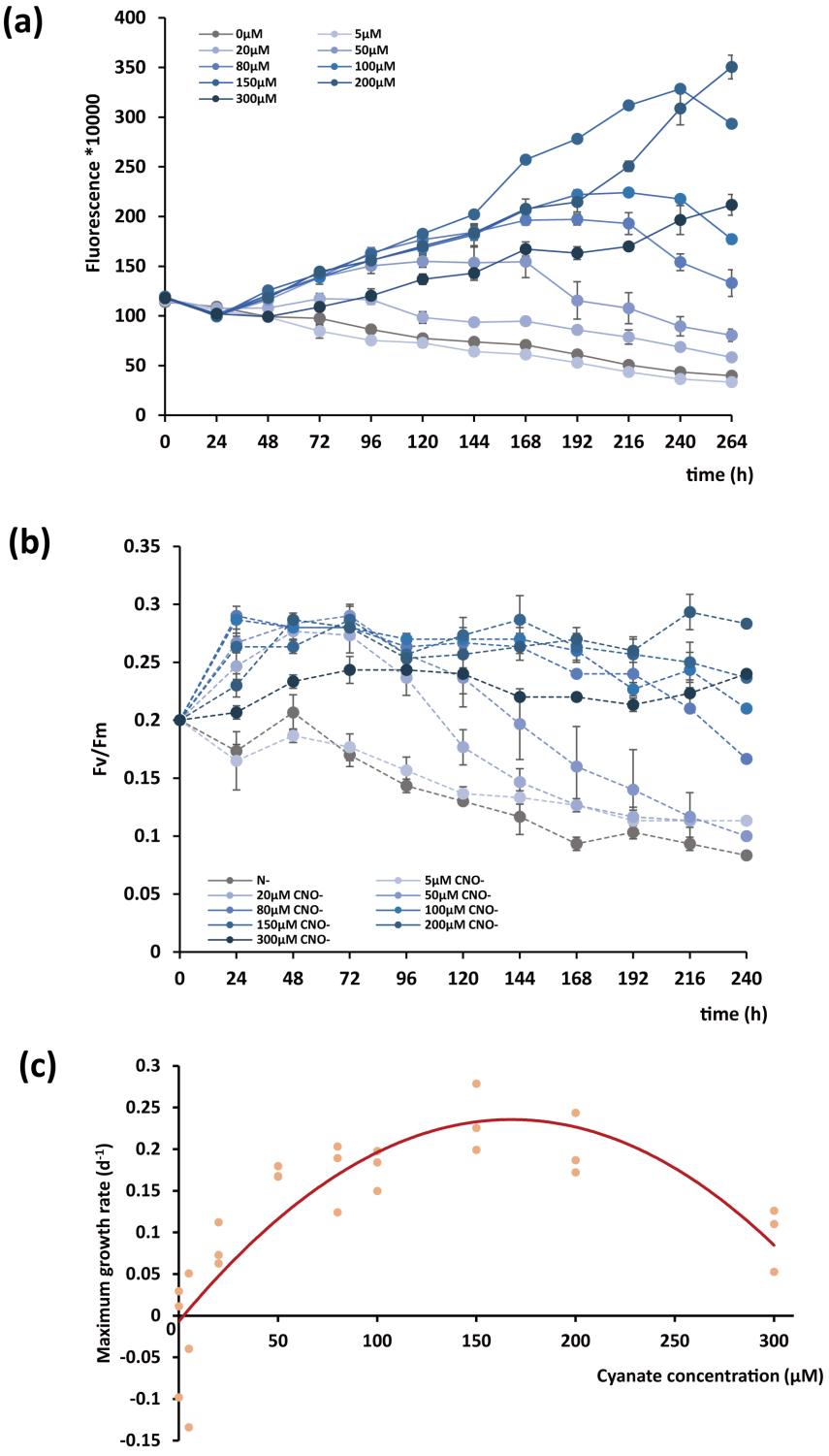


Figure 4.12: Physiological responses of *M. commoda* to cyanate gradient. (a) Chla fluorescence; (b) Fv/Fm; (c) maximum growth rate. Error bar represents mean ± SD (n = 3).

other microalgae and cyanobacteria during seasonal blooms (Cermeño et al., 2011). Under conditions of elevated nitrate availability, diatoms can take up nitrate in excess of their immediate growth requirements and store it in vacuoles (Raven, 1987; Lomas and Glibert, 2000; Clark et al., 2011; McCarthy et al., 2017). In diatoms, nitrate reduction mediated by *NitR* represents the first and rate-limiting step in nitrate assimilation. In addition, *NitR* is related to carbon metabolism, thereby influencing the partitioning of cellular energy between growth and nitrogen assimilation (McCarthy et al., 2017). We confirmed that *NitR* is essential for survival when nitrate is the sole nitrogen source (Figure 4.5), consistent with observations in other species (Camargo et al., 2007; McCarthy et al., 2017; Belshaw et al., 2023).

In this study, the physiological differences between nitrate-replete and nitrogen-limited treatments in *T. pseudonana NitR* mutants demonstrate that the loss of *NitR* has a profound impact on cellular physiology. Under the nitrate-replete conditions, *NitR* mutants exhibited significantly larger cell sizes, along with reduced growth rates and lower Fv/Fm, compared to nitrogen-limited treatments (Figure 4.5). These findings are consistent with previous studies. In diatom *P. tricornutum*, *NitR* knockout cells were shown to rapidly transport nitrate into the cell but were unable to assimilate it, resulting in intracellular nitrate accumulation (McCarthy et al., 2017). This accumulation leads to cell swelling and substantial alterations in biochemical composition, cellular physiology, and the global transcriptome. Such responses exert immediate and acute effects on central metabolic pathways and are mechanistically distinct from those observed under nitrogen limitation (Alipanah et al., 2015; Levitan et al., 2015; McCarthy et al., 2017). The impact of *NitR* loss was more severe under nitrate-replete conditions than under nitrogen limitation. In *P. tricornutum* NR-KO cells, nitrate addition triggers a rapid and systemic downregulation of central metabolic pathways, in contrast to the gradual acclimation observed under nitrogen-limited conditions, where cells adjust to a declining supply of nitrogen-derived products (McCarthy et al., 2017).

The regulation of *NitR* not only facilitates the nitrate assimilation but also contributes to the reduction of nitrite to NO *in vivo*, modulates multiple metabolic pathways, and helps prevent the formation of reactive oxygen and nitrogen species in microalgae (Meyer et al., 2005; Wei et al., 2014; Wang et al., 2023b). After mild nitrogen limitation, *NitR* mutants exhibited a faster growth response than WT when supplied with nitrite (Figure 4.9). This suggests that *NitR* does not function solely as a facilitator of nitrate assimilation in this context but may instead impose a regulatory constraint on early growth responses. However, the underlying mechanisms remain unclear and require further investigation.

4.4.2 The putative role of cyanase in diatom *T. pseudonana* and green algae *M. commoda*

Extracellular cyanate uptake and assimilation pathways have been identified in prokaryotes (Sung and Fuchs, 1992; Kamennaya and Post, 2011). In *E. coli*, cyanate degradation is mediated by the *cynTSX* operon (*cynT*, *cynS*, *cynX*), enabling efficient and tightly controlled cyanate utilization (Sung and Fuchs, 1988, 1992). In some cyanobacteria, such as *Synechococcus elongatus* PCC 7942 and *Prochlorococcus* sp. MED4, ABC-type cyanate transporters (*CynABD*) have been identified and function as active cyanate uptake systems. These transporters are often organized in an operon together with *cynS*, facilitating extracellular cyanate utilization (Espie et al., 2007; Kamennaya and Post, 2011). Although the *cynS* gene has been identified in several marine cyanobacterial genomes (Palenik et al., 2003; Rocap et al., 2003; Scanlan et al., 2009), in many cases they do not have encoding cyanate transport systems. In such species, cyanase only functions in the detoxification of intracellularly generated cyanate, generating from the intracellular urea and/or carbamoyl phosphate decomposition (Kamennaya and Post, 2011).

Although cyanate concentrations in marine environments are typically low (Widner et al., 2018; Kitzinger et al., 2019), environmental measurements generally reflect bulk steady-state concentrations rather than transient local

exposure levels experienced by individual cells. In this study, higher cyanate concentrations were applied to generate measurable physiological and isotopic responses, assess dose-dependent effects, and evaluate potential toxicity thresholds under controlled laboratory conditions. Similar experimental concentrations have also been used in previous studies investigating cyanate utilization in marine microorganisms (Hu et al., 2014; Palatinszky et al., 2015).

To date, no cyanate transporter has been identified in eukaryotic microalgae. Nevertheless, the role of *CynS* in nitrogen acquisition has been demonstrated in some species. For example, the dinoflagellate *Prorocentrum donghaiense* is capable of sustained growth on cyanate as the sole nitrogen source at relatively low growth rates (Hu et al., 2012, 2014). A similar growth pattern was observed in *M. commoda* (Figure 4.10), suggesting that cyanate can be utilized despite the apparent absence of dedicated transport systems. In *M. commoda*, the maximum growth rate was achieved at 150 μM cyanate, followed by a decline at higher concentrations (Figure 4.12). This indicates a concentration-dependent effect of cyanate, where low concentrations can support growth as an alternative nitrogen source, but higher concentrations exert inhibitory or toxic effects on cellular physiology compared to nitrogen-replete conditions. The ability to utilize cyanate in the absence of a characterized transporter suggests that uptake may occur via non-specific transport systems or passive diffusion, although the efficiency of such processes is likely limited. Elucidating the mechanisms of cyanate transport in eukaryotic microalgae is important for future research.

In many eukaryotic microalgae encoding *CynS*, cells are unable to utilize cyanate as a sole nitrogen source to support growth (data generated by collaborators and not published). It was reported that the *cynS* gene in *P. calceolata* was upregulated in oligotrophic environments, but was downregulated in the presence of cyanate and cells cannot grow when cyanate was the sole nitrogen source (Guérin et al., 2025). The function of *CynS* in *P. calceolata* is likely the recycling of intracellularly produced cyanate when nitrate supply is limited. It comes out that the removal of intracellular cyanate rather than external cyanate metabolism as an alternative N

source is the main role of *CynS* in some eukaryotic microalgae (Mao et al., 2022; Sato et al., 2023).

The diatom *T. pseudonana*, which also encodes a *CynS* gene, was unable to sustain growth on cyanate (Figure 4.7a–d). Stable isotope probing (SIP) further revealed extremely low nitrogen uptake rates under cyanate treatment, indicating negligible direct assimilation of cyanate (Figure 4.7e, f). Chlorophyll *a* fluorescence under cyanate treatment was comparable to that observed under nitrogen limitation. In addition, Fv/Fm was transiently suppressed during the first 72 h of exposure, and cells exhibited increased size relative to nitrogen-limited conditions. Together, these physiological responses suggest that cyanate can enter the cells but only at a low flux insufficient to meet cellular nitrogen demand for balanced growth. The *CynS* knockout mutants exhibited increased cell size, further supporting a toxic effect associated with intracellular cyanate accumulation (Figure 4.8). The loss of endogenous *CynS* was unable to detoxify or assimilate intracellular cyanate, which may lead to cyanate accumulation and subsequent cellular toxicity. My results indicate that *CynS* in *T. pseudonana* primarily functions in intracellular cyanate detoxification rather than serving a major role in extracellular cyanate assimilation. The deeper understanding of the *CynS* function in nitrogen metabolism deserves further investigation.

5

Conclusions and future work

5.1 Discussion and conclusions

This thesis was motivated by the need to investigate nitrogen-source utilization and the functions of genes involved in nitrogen metabolism in the marine microalgae *Micromonas commoda* and *Thalassiosira pseudonana*. Nitrogen-source utilization was examined in both species. Although both species possess a cyanase (*cynS*) gene, they exhibited contrasting capacities to utilize cyanate. To enable the functional characterization of *cynS* in *M. commoda*, a polyethylene glycol (PEG)-mediated transformation method was established and subsequently used to develop a CRISPR/Cas9 genome-editing system. A PEG-mediated transformation method was also established for *T. pseudonana*.

5.1.1 Establishment of a molecular toolkit for the marine picoeukaryote *Micromonas commoda*

Marine microalga *M. commoda* is one of the most abundant picoeukaryotic phytoplankton in the ocean and is distributed throughout the photic zone from tropical to polar regions (Worden et al., 2009; Bachy et al., 2022). It plays an important role in marine biogeochemical cycling (Massana, 2011). Since the publication of the *Micromonas* genome, this genus has been widely used to investigate evolution, nutrient utilization, and environmental adaptation (Worden

et al., 2009; McDonald et al., 2010; Guo et al., 2018; Simó et al., 2026). However, in the absence of well-developed gene editing tools, gene functions in *M. commoda* have been inferred indirectly rather than validated in their native cellular context. Prior to this study, nucleofection-mediated transient transformation had been demonstrated in *M. commoda* (Faktorova et al., 2020). However, stable transformant cell lines were not isolated, limiting its application for functional genetic studies. In addition, the requirement for a specialised Lonza Nucleofector device and proprietary buffers reduces the accessibility of this approach.

In this thesis, a stable PEG-mediated transformation method was developed for DNA delivery into *M. commoda*. Cell sensitivity to PEG concentration and incubation time was systematically evaluated, and suitable selectable markers were identified. Single-cell transformants were isolated and exhibited growth rates comparable to those of wild-type cells during the exponential phase. Expression of *shBle* conferred resistance to zeocin in both solid and liquid media, while *eYFP* expression was confirmed by fluorescence microscopy. Multigene expression using 2A self-cleaving peptides was also evaluated using a construct in which *shBle* was linked to *eYFP* through a 2A peptide. However, the 2A peptide exhibited relatively low cleavage efficiency in *M. commoda*. The functions of both *shBle* and *eYFP* were retained, suggesting that the uncleaved 2A sequence may function as a linker between *shBle* and *eYFP*.

Using the PEG-mediated DNA delivery method, a CRISPR/Cas9-based genome-editing system was established in *M. commoda*. The CRISPR/Cas9 plasmids were assembled using Golden Gate cloning and delivered into *M. commoda* cells. A conserved region of the *cynS* containing three key residues was targeted for deletion. Following screening of the transformed cell lines, a *cynS* knockout strain C34 with a 76-nt deletion was identified. This represents the first report of CRISPR/Cas9-mediated gene knockout in *M. commoda* and demonstrates the feasibility of targeted genome editing in this species.

Overall, the PEG-mediated transformation and gene editing approaches provide an

accessible and cost-effective platform for stable transgene expression and future functional genetic studies in *M. commoda*. Promoters, selectable markers, and fluorescent reporters suitable for genetic manipulation were identified and experimentally validated, providing a set of readily applicable genetic components for future studies. Importantly, these tools move functional studies in *M. commoda* beyond correlative genomic and physiological observations by enabling direct manipulation of candidate genes in their native cellular background. This paves ways for investigating downstream metabolic responses and examining interactions between cellular compartments. Furthermore, fluorescent protein fusions can also be used to determine the subcellular localization of candidate proteins, providing information on their cellular functions.

5.1.2 PEG-mediated transformation method on *T. pseudonana*

Existing methods for delivering DNA to *T. pseudonana* include biolistic bombardment (Poulsen et al., 2006), electroporation (Bugge and Robertson, 2019), and bacterial conjugation (Karas et al., 2015). Biolistics was the first method developed for transforming *T. pseudonana* and has been widely used to integrate non-replicative DNA into the nuclear genome (Hopes et al., 2016; Belshaw et al., 2023). This method requires expensive specialised equipment, (biolistic particle delivery system) which may not be accessible to many laboratories. Conjugation requires maintenance of the *E. coli* donor strain, in which plasmid incompatibility, high copy number, or toxic genetic elements can lead to plasmid rearrangement or loss (Slattery et al., 2018; Walker et al., 2026). An electroporation-mediated transformation protocol has also been reported for *T. pseudonana* (Bugge and Robertson, 2019), but this approach has seen limited application in subsequent genetic studies.

In this study, a PEG-mediated transformation protocol was optimized for *T. pseudonana*. Before transformation, alcalase was used as the sole enzyme for cell wall removal. Treatment with 3 Anson units of alcalase for 40 min resulted in a

protoplasting efficiency of approximately 50%. Transformation was achieved using 20% PEG 8000, and single-cell colonies were isolated on 1% agarose plates containing 100 μg/mL nourseothricin. Expression of *eGFP* was confirmed by fluorescence microscopy, demonstrating successful stable transformation.

This method provides an accessible and cost-effective transformation method of *T. pseudonana*. It may be particularly useful for laboratories with limited access to specialised transformation platforms and can reduce the technical complexity associated with routine genetic manipulation. Overall, the establishment of PEG-based method provides an alternative transformation option for *T. pseudonana* and expands the available toolkit for diatom genetic engineering.

5.1.3 Differentiation of nitrogen utilization in *Micromonas commoda* and *Thalassiosira pseudonana*

Nitrogen availability in the oceans is a key factor influencing the productivity and diversity of marine microalgae (Moore et al., 2013). To overcome nitrogen limitation and competition, marine microalgae have evolved diverse strategies for nitrogen acquisition and utilization. Their capacities to exploit available nitrogen sources play important roles in marine nitrogen cycling (Hutchins and Capone, 2022). Both *M. commoda* and *T. pseudonana* are widely distributed in coastal and oceanic waters (Worden et al., 2009; Poulsen and Kröger, 2023), and their abilities to utilize different nitrogen sources were investigated in this study. Both species can use nitrate and ammonium efficiently, but they show different utilization on organic nitrogen. *M. commoda* was able to utilize cyanate as a sole nitrogen source but showed limited growth on urea. Its response to cyanate was concentration dependent, with the highest growth rate observed at 150 μM cyanate and reduced growth at higher concentrations. By contrast, *T. pseudonana* was able to utilize urea as sole nitrogen sources but did not grow on cyanate. Stable isotope probing revealed very low nitrogen uptake under cyanate treatment, indicating negligible direct assimilation of cyanate-derived nitrogen. The reduced growth rates and Fv/Fm values relative to

nitrogen-limited controls further suggest that cyanate imposed physiological stress on *T. pseudonana*.

These responses indicate distinct organic nitrogen utilization strategies in the two species, particularly in their ability to exploit cyanate and urea as alternative nitrogen sources. Cyanate has been recognized as an ecologically relevant nitrogen substrate in the ocean and represent an alternative nitrogen source comparable in importance to urea in some marine environments (Widner et al., 2018; Mao et al., 2022). The ability of *M. commoda* to utilize cyanate may provide a competitive advantage under nitrogen-limited conditions and contribute to the transfer of cyanate-derived nitrogen into phytoplankton biomass and the marine food web. By contrast, *T. pseudonana* was unable to utilize cyanate but efficiently used urea as a sole nitrogen source. Although related nitrogen metabolism genes are present in both species, different microalgae may access distinct components of the dissolved nitrogen pool. These differences in nitrogen-use niches may influence the partitioning, recycling, and biological sources of nitrogen in marine ecosystems.

5.1.4 Functional divergence among nitrogen assimilation pathways in the marine diatom *Thalassiosira pseudonana*

Although *T. pseudonana* could not utilize cyanate as a nitrogen source, it possesses a *cynS* gene. To investigate its function, *cynS* knockout strains were generated. These mutants exhibited increased cell size and reduced maximum growth rates relative to the wild type, suggesting that *cynS* may contribute to cyanate detoxification and broader metabolic regulation. To compare the effects of disrupting different nitrogen metabolism-related genes, wild-type cells, *cynS* knockout strains, nitrate reductase (*NitR*) mutants, and a urease (*URE*) mutant were grown under nitrite-replete conditions that supported the growth of all strains. The mutants exhibited distinct physiological responses relative to the wild type, indicating that disruption of individual nitrogen metabolism genes can have broader consequences for cellular physiology beyond their immediate metabolic functions.

These findings highlight that nitrogen metabolism genes not only function in the direct utilization of their corresponding nitrogen sources, but can also influence cellular physiology when other nitrogen sources are supplied. These genes appear to form an interconnected metabolic network rather than acting independently. This network may expand to broader cellular processes to influence growth, cell size, photosynthetic performance, and metabolic allocation in response to changes in nitrogen availability. Further transcriptomic analysis will be used to identify shared and gene-specific metabolic and regulatory responses associated with these mutations and to clarify how nitrogen metabolism pathways are integrated with other cellular processes. This may improve our understanding of how coordinated nitrogen metabolic networks contribute to the ecological success of diatoms and their role in marine nitrogen cycling.

5.2 Future work

5.2.1 Optimization of the PEG-mediated transformation method on *M. commoda*

In this transformation system, the constitutive RPS9 promoter was used to drive stable gene expression. However, constitutive expression may lead to excessive production of recombinant proteins and impose a metabolic burden on cells (Li et al., 2025c). Future work could evaluate inducible or nutrient-responsive promoters, such as those regulated by nitrogen or iron availability, to enable more precise control of transgene expression under defined environmental or experimental conditions. The development of such regulatory elements would reduce the potential limitations of constitutive expression and further expand the molecular toolkit available for *M. commoda* transformation and functional genetic studies.

The efficiency of 2A peptides in *M. commoda* requires further validation. Previous studies have shown that 2A peptide-mediated cleavage efficiency can vary

substantially in microalgae, including the diatom *P. tricornutum* (Diamond et al., 2023). In this study, only one transformant was analysed for each 2A peptide construct. Therefore, additional independent transformants should be examined to determine the potential clone-specific and stochastic effects.

Alternative strategies for multigene expression should also be explored. Polycistronic expression of nuclear genes has been identified and validated in the green algae *Chlamydomonas reinhardtii* and *Chromochloris zofingiensis* (Gallaher et al., 2021). In *M. commoda*, full-length transcript isoform sequencing (Iso-Seq) could be used to identify naturally occurring polycistronic transcription units. Characterisation of such endogenous architectures may provide useful sequence features for the design of more efficient multigene expression systems and offer an alternative or complementary approach to 2A peptide-based strategies.

5.2.2 CRISPR/Cas9 editing efficiency in *M. commoda*

The low knockout efficiency observed in this study should be further investigated. One potential limitation is that successful editing depends on genomic integration and stable expression of the *Cas9* cassette. Editing efficiency could be improved by maintaining Cas9 expression independently of genomic integration. One possible strategy is to deliver the CRISPR/Cas cassette using a self-replicating episomal vector (Karas et al., 2015). Such episomes could provide sustained but temporary expression of the editing machinery and subsequently be eliminated by removing selection pressure, thereby reducing prolonged Cas9 exposure and the associated risk of off-target effects.

Alternatively, delivery of pre-assembled Cas9–single-guide RNA ribonucleoprotein complexes (RNPs) may further improve editing efficiency. Direct introduction of Cas9 protein and sgRNAs allows genome editing to occur immediately without the need for intracellular transcription or translation (Zhang et al., 2021). This strategy could bypass potential limitations associated with transgene integration and expression.

In addition to improving knockout efficiency, future work could extend the CRISPR/Cas9 system to homology-directed repair (HDR). CRISPR/Cas-mediated homologous recombination (HR) has been applied in several microalgal species (Naduthodi et al., 2019; Moosburner et al., 2020; Belshaw et al., 2023) and provide a route for precise genome modification in *M. commoda*. One potential strategy to enhance HDR efficiency is the use of replicating vectors to increase the intracellular copy number of donor repair templates. Geminiviral vectors have been used for in other systems (Yin et al., 2015), where increased donor-template availability can improve the probability of HR (Hopes, 2017).

The efficiency of sgRNA expression also requires further optimisation. In this study, the endogenous U6 snRNA sequence of *M. commoda* was identified by BLASTn, and two potential guanine (G) transcription start sites were identified downstream of a putative TATA box. However, the transcription start site and promoter boundaries have not been experimentally validated. Future work should map the endogenous U6 transcription start site using 5′ RACE and define the functional promoter region empirically. This information would enable more precise sgRNA expression and further improve CRISPR/Cas9 editing efficiency in *M. commoda*.

5.2.3 Cyanate assimilation by *M. commoda*

The growth of *M. commoda* in cyanate-containing media was observed. However, direct evidence of cyanate uptake into the cells is currently lacking. Therefore, cyanate uptake in *M. commoda* will be further investigated using stable isotope probing (SIP) approaches to directly quantify nitrogen assimilation from cyanate. Dual-labelled $^{13}C^{15}N$-cyanate should be used to quantify uptake at the single-cell level. To our knowledge, few marine eukaryotic phytoplankton species have been reported to directly utilize cyanate (Hu et al., 2014; Guérin et al., 2025). By directly measuring cyanate-derived nitrogen incorporation, this study will clarify the contribution of phytoplankton to cyanate cycling in marine ecosystems and improve our understanding of an overlooked pathway in the marine nitrogen cycle.

The cyanate utilization pathway in *M. commoda* should be further investigated at the transcriptional level. Time-resolved RNA-seq following cyanate addition could be used to identify candidate cyanate transporters, regulatory genes, and metabolic pathways associated with cyanate uptake and assimilation. Comparing transcriptional responses across a cyanate concentration gradient would also help distinguish genes involved in cyanate utilization from those responding to cyanate-induced stress, providing a more complete understanding of the molecular mechanisms underlying cyanate metabolism in *M. commoda*.

5.2.4 Future investigation of cyanate metabolism in *T. pseudonana*

Although *T. pseudonana* possesses a *cynS* gene, it cannot utilize cyanate as a sole nitrogen source, and the physiological phenotype of *cynS* mutants suggests that cyanase may have functions on other metabolisms. Future work should determine the function of *cynS* in cyanate detoxification and metabolic regulation.

Transcriptomic analysis of wild-type, *cynS*, *URE*, and *NitR* mutant strains under nitrogen-limited and nitrogen-replete conditions could be used to distinguish cyanase-specific responses from broader responses caused by disruption of nitrogen metabolism. Comparisons among these mutants would help identify shared transcriptional changes associated with nitrogen metabolic imbalance, as well as pathways specifically affected by loss of *cynS*. Particular attention should be given to nitrogen transport and assimilation, carbon metabolism, photosynthesis and cellular stress responses. Intracellular cyanate and related metabolites also should be quantified in wild-type and *cynS* mutant cells to test whether loss of cyanase leads to cyanate accumulation and other metabolic effects.

This could clarify why *T. pseudonana* retains *cynS* despite its inability to use cyanate as a nitrogen source and provide a broader understanding of how cyanase is integrated with other nitrogen metabolism pathways in diatoms.

Bibliography

Abassi, S. and Ki, J.-S. (2022). Increased nitrate concentration differentially affects cell growth and expression of nitrate transporter and other nitrogen-related genes in the harmful dinoflagellate *Prorocentrum minimum*. *Chemosphere*, 288:132526.

Abbriano, R., Vardar, N., Yee, D., and Hildebrand, M. (2018). Manipulation of a glycolytic regulator alters growth and carbon partitioning in the marine diatom *Thalassiosira pseudonana*. *Algal Research*, 32:250–258.

Ahier, A. and Jarriault, S. (2014). Simultaneous expression of multiple proteins under a single promoter in *Caenorhabditis elegans* via a versatile 2A-Based toolkit. *Genetics*, 196(3):605–613.

Alipanah, L., Rohloff, J., Winge, P., Bones, A. M., and Brembu, T. (2015). Whole-cell response to nitrogen deprivation in the diatom *Phaeodactylum tricornutum*. *Journal of Experimental Botany*, 66(20):6281–6296.

Allen, A. E., Dupont, C. L., Oborník, M., Horák, A., Nunes-Nesi, A., McCrow, J. P., Zheng, H., Johnson, D. A., Hu, H., Fernie, A. R., and Bowler, C. (2011). Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature*, 473:203–207.

Amin, S. A., Hmelo, L. R., van Tol, H. M., Durham, B. P., Carlson, L. T., Heal, K. R., Morales, R. L., Berthiaume, C. T., Parker, M. S., Djunaedi, B., Ingalls, A. E., Parsek, M. R., Moran, M. A., and Armbrust, E. V. (2015). Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature*, 522(7554):98–101.

Amin, S. A., Parker, M. S., and Armbrust, E. V. (2012). Interactions between

diatoms and bacteria. *Microbiology and Molecular Biology Reviews*, 76(3):667–684.

Anderson, D. M., Burkholder, J. M., Cochlan, W. P., Glibert, P. M., Gobler, C. J., Heil, C. A., Kudela, R. M., Parsons, M. L., Rensel, J. E. J., Townsend, D. W., Trainer, V. L., and Vargo, G. A. (2008). Harmful algal blooms and eutrophication: Examining linkages from selected coastal regions of the united states. *Harmful Algae*, 8(1):39–53.

Anderson, P. M., Sung, Y.-c., and Fuchs, J. A. (1990). The cyanase operon and cyanate metabolism. *FEMS Microbiology Reviews*, 7(3-4):247–252.

Angstenberger, M., de Signori, F., Vecchi, V., Dall'Osto, L., and Bassi, R. (2020). Cell synchronization enhances nuclear transformation and genome editing via Cas9 enabling homologous recombination in *Chlamydomonas reinhardtii*. *ACS Synthetic Biology*, 9(10):2840–2850.

Aranguren-Gassis, M., Kremer, C. T., Klausmeier, C. A., and Litchman, E. (2019). Nitrogen limitation inhibits marine diatom adaptation to high temperatures. *Ecology Letters*, 22(11):1860–1869.

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. A., Chaal, B. K., Chiovitti, A., Davis, A. K., Demarest, M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J., Kapitonov, V. V., Kröger, N., Lau, W. W. Y., Lane, T. W., Larimer, F. W., Lippmeier, J. C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. A., Saito, M. A., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P., and Rokhsar, D. S. (2004). The genome of the diatom *Thalassiosira Pseudonana*: Ecology, evolution, and metabolism. *Science*, 306(5693):79–86.

Ashworth, J., Coesel, S., Lee, A., Armbrust, E. V., Orellana, M. V., and Baliga, N. S. (2013). Genome-wide diel growth state transitions in the diatom *Thalassiosira*

pseudonana. *Proceedings of the National Academy of Sciences of the United States of America*, 110(18):7518–7523.

Astier, J., Rossi, J., Chatelain, P., Klinguer, A., Besson-Bard, A., Rosnoblet, C., Jeandroz, S., Nicolas-Francès, V., and Wendehenne, D. (2021). Nitric oxide production and signalling in algae. *Journal of Experimental Botany*, 72(3):781–792.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., and Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Marine Ecology Progress Series*, 10:257–263.

Bachy, C., Charlesworth, C. J., Chan, A. M., Finke, J. F., Wong, C.-H., Wei, C.-L., Sudek, S., Coleman, M. L., Suttle, C. A., and Worden, A. Z. (2018). Transcriptional responses of the marine green alga *Micromonas pusilla* and an infecting prasinovirus under different phosphate conditions. *Environmental Microbiology*, 20:2898–2912.

Bachy, C., Sudek, L., Choi, C. J., Eckmann, C. A., Nöthig, E.-M., Metfies, K., and Worden, A. Z. (2022). Phytoplankton surveys in the arctic fram strait demonstrate the tiny eukaryotic alga *Micromonas* and other picoprasinophytes contribute to deep sea export. *Microorganisms*, 10(5).

Baek, K., Kim, D., Jeong, J., Sim, S., Melis, A., Kim, J.-S., Jin, E., and Bae, S. (2016). DNA-free two-gene knockout in *Chlamydomonas reinhardtii* via CRISPR-Cas9 ribonucleoproteins. *Scientific Reports*, 6:30620.

Belshaw, N., Grouneva, I., Aram, L., Gal, A., Hopes, A., and Mock, T. (2023). Efficient gene replacement by CRISPR/Cas-mediated homologous recombination in the model diatom *Thalassiosira pseudonana*. *New Phytologist*, 238:438–452.

Berthelot, H., Zukowska, J., Henry, N., Noël, C., Thyssen, M., Leblanc, K., Planquette, H., Maguer, J.-F., Pepperkok, R., de Vargas, C., and Cassar, N. (2025). Nitrogen concentration shapes the size structure and the functional

diversity of phytoplankton communities in the southern Indian ocean. *ISME Communications*, 5(1):ycaf195.

Bhattacharya, S., Srinivasan, K., Abdisalaam, S., Su, F., Raj, P., Dozmorov, I., Mishra, R., Wakeland, E. K., Ghose, S., Mukherjee, S., and Asaithamby, A. (2017). RAD51 interconnects between DNA replication, DNA repair and immunity. *Nucleic Acids Research*, 45(8):4590–4605.

Blaby, I. K., Glaesener, A. G., Mettler, T., Fitz-Gibbon, S. T., Gallaher, S. D., Liu, B., Boyle, N. R., Kropat, J., Stitt, M., Johnson, S., Benning, C., Pellegrini, M., Casero, D., and Merchant, S. S. (2013). Systems-level analysis of nitrogen starvation–induced modifications of carbon metabolism in a *Chlamydomonas reinhardtii* starchless mutant. *The Plant Cell*, 25(11):4305–4323.

Bonnet, S., Benavides, M., Le Moigne, F. A. C., Wagener, T., Bidle, K. D., Picheral, M., Stemmann, L., Karl, D. M., and Zehr, J. P. (2023). Diazotrophs are overlooked contributors to carbon and nitrogen export to the deep ocean. *ISME Journal*, 17:47–58.

Booth, D. S. and King, N. (2020). Genome editing enables reverse genetics of multicellular development in the choanoflagellate *Salpingoeca rosetta*. *eLife*, 9:e56193.

Borkman, D. G., Smayda, T. J., Tomas, C. R., York, R., Strangman, W., and Wright, J. L. (2012). Toxic *Alexandrium peruvianum* (Balech and de Mendiola) balech and tangen in narragansett bay, rhode island (USA). *Harmful Algae*, 19:92–100.

Bortesi, L. and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, 33:41–52.

Boulton, W., Salamov, A., Grigoriev, I. V., Calhoun, S., LaButti, K., Riley, R., Barry, K., Fong, A. A., Hoppe, C. J. M., Metfies, K., Oetjen, K., Eggers, S. L., Müller, O., Gardner, J., Granskog, M. A., Torstensson, A., Oggier, M., Larsen, A., Bratbak, G., Toseland, A., Leggett, R. M., Moulton, V., and

Mock, T. (2025). Metagenome-assembled-genomes recovered from the Arctic drift expedition MOSAiC. *Scientific Data*, 12:204.

Bronk, D. A., See, J. H., Bradley, P., and Killberg, L. (2007). DON as a source of bioavailable nitrogen for phytoplankton. *Biogeosciences*, 4(3):283–296.

Brooks, C., Nekrasov, V., Lippman, Z. B., and Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 system. *Plant Physiology*, 166(3):1292–1297.

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, 321:960–964.

Bugge, J. and Robertson, D. (2019). Electroporation of *Thalassiosira pseudonana*. *protocols.io*.

Bule, M. H., Ahmed, I., Maqbool, F., Bilal, M., and Iqbal, H. M. N. (2018). Microalgae as a source of high-value bioactive compounds. *Frontiers in Bioscience (Scholar Edition)*, 10:197–216.

Camargo, A., Llamas, A., Schnell, R. A., Higuera, J. J., González-Ballester, D., Lefebvre, P. A., et al. (2007). Nitrate signaling by the regulatory gene NIT2 in *Chlamydomonas*. *The Plant Cell*, 19(11):3491–3503.

Cermeño, P., Lee, J. B., Wyman, K., Schofield, O., and Falkowski, P. G. (2011). Competitive dynamics in two species of marine phytoplankton under non-equilibrium conditions. *Marine Ecology Progress Series*, 429:19–28.

Chai, X. J., Chen, H. X., Xu, W. Q., and Xu, Y. W. (2013). Expression of soybean kunitz trypsin inhibitor gene skti in *Dunaliella salina*. *Journal of Applied Phycology*, 25:139–144.

Chang, K. S., Kim, J., Park, H., Hong, S.-J., Lee, C.-G., and Jin, E. (2020). Enhanced lipid productivity in AGP knockout marine microalga *Tetraselmis*

sp. using a DNA-free CRISPR-Cas9 RNP method. *Bioresource Technology*, 303:122932.

Chen, J., Yi, Q., Song, Q., et al. (2015). A highly efficient maize nucellus protoplast system for transient gene expression and studying programmed cell death-related processes. *Plant Cell Reports*, 34:1239–1251.

Chen, K., Achterberg, E. P., Li, K., Zhang, J., Xin, M., and Wang, X. (2023a). Governance pathway for coastal eutrophication based on regime shifts in diatom-dinoflagellate composition of the Bohai and Baltic Seas. *Water Research*, 240:121042.

Chen, P., Liu, H.-H., Cui, R., Zhang, Z.-L., Pang, D.-W., Xie, Z.-X., Zheng, H.-Z., Lu, Z.-X., and Tong, H. (2008). Visualized investigation of yeast transformation induced with Li+ and polyethylene glycol. *Talanta*, 77(1):262–268.

Chen, X. H., Yang, M. K., Li, Y. Y., and et al. (2023b). Improving the genome and proteome annotations of the marine model diatom *Thalassiosira pseudonana* using a proteogenomics strategy. *Marine Life Science & Technology*, 5:102–115.

Chicaybam, L., Sodré, A. L., Curzio, B. A., and Bonamino, M. H. (2013). An efficient low cost method for gene transfer to T lymphocytes. *PLOS ONE*, 8(3):e60298.

Chung, C. C., Hwang, S. L., and Chang, J. (2008). Nitric oxide as a signaling factor to upregulate the death-specific protein in a marine diatom, *Skeletonema costatum*, during blockage of electron flow in photosynthesis. *Applied and Environmental Microbiology*, 74(21):6521–6527.

Cirri, E. and Pohnert, G. (2019). Algae–bacteria interactions that balance the planktonic microbiome. *New Phytologist*, 223:100–106.

Clark, D. R., Miller, P. I., Woodward, M. S., and Rees, A. P. (2011). Inorganic nitrogen assimilation and regeneration in the coastal upwelling region of the Iberian Peninsula. *Limnology and Oceanography*, 56(5):1689–1702.

Cohen, N. R. (2022). Mixotrophic plankton foraging behaviour linked to carbon export. *Nature Communications*, 13:1302.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339:819–823.

Corellou, F., Schwartz, C., Motta, J.-P., Djouani-Tahri, E. B., Sanchez, F., and Bouget, F.-Y. (2009). Clocks in the green lineage: Comparative functional analysis of the circadian architecture of the picoeukaryote *Ostreococcus*. *The Plant Cell*, 21(11):3436–3449.

Cuvelier, M. L., Guo, J., Ortiz, A. C., van Baren, M. J., Tariq, M. A., Partensky, F., and Worden, A. Z. (2017). Responses of the picoprasinophyte *Micromonas commoda* to light and ultraviolet stress. *PLOS ONE*, 12(3):e0172135.

Daboussi, F., Leduc, S., Maréchal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D. F., Cavarec, L., and Duchateau, P. (2014). Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nature Communications*, 5:3831.

Dalessandri, S. M., Siersma, A., Krause, J. W., and Hoadley, K. D. (2025). Rapid vs. gradual changes in light significantly modulate diel photo-acclimation strategy and can lead to changes in comparative physiology in the green algae *Micromonas commoda*. *Limnology and Oceanography*, 70:3942–3957.

Damashek, J., Tolar, B. B., Liu, Q., Okotie-Oyekan, A. O., Wallsgrove, N. J., Popp, B. N., and Hollibaugh, J. T. (2019). Microbial oxidation of nitrogen supplied as selected organic nitrogen compounds in the south atlantic bight. *Limnology and Oceanography*, 64:982–995.

Davey, M. R., Anthony, P., Power, J. B., and Lowe, K. C. (2005). Plant protoplasts: status and biotechnological perspectives. *Biotechnology Advances*, 23(2):131–171.

Davis, A. J., Chen, B. P., and Chen, D. J. (2014). DNA-PK: A dynamic enzyme in a versatile DSB repair pathway. *DNA Repair*, 17:21–29.

Davis, A. J. and Chen, D. J. (2013). DNA double-strand break repair via non-homologous end-joining. *Translational Cancer Research*, 2(3):130–143.

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., and Falciatore, A. (2009). Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Research*, 37:e96.

Delalat, B., Sheppard, V., Rasi Ghaemi, S., and et al. (2015). Targeted drug delivery using genetically engineered diatom biosilica. *Nature Communications*, 6:8791.

Delmont, T. O., Gaia, M., Hinsinger, D. D., Frémont, P., Vanni, C., Fernandez-Guerra, A., Eren, A. M., Kourlaiev, A., d'Agata, L., Clayssen, Q., Villar, E., Labadie, K., Cruaud, C., Poulain, J., Da Silva, C., Wessner, M., Noel, B., Aury, J.-M., Sunagawa, S., Acinas, S. G., Bork, P., Karsenti, E., Bowler, C., Sardet, C., Stemmann, L., de Vargas, C., Wincker, P., Lescot, M., Babin, M., Gorsky, G., Grimsley, N., Guidi, L., Hingamp, P., Jaillon, O., Kandels, S., Iudicone, D., Ogata, H., Pesant, S., Sullivan, M. B., Not, F., Lee, K.-B., Boss, E., Cochrane, G., Follows, M., Poulton, N., Raes, J., Sieracki, M., Speich, S., de Vargas, C., Bowler, C., Karsenti, E., Pelletier, E., Wincker, P., and Jaillon, O. (2022). Functional repertoire convergence of distantly related eukaryotic plankton lineages abundant in the sunlit ocean. *Cell Genomics*, 2(5):100123.

Demory, D., Baudoux, A.-C., Monier, A., Simon, N., Six, C., Ge, P., Rigaut-Jalabert, F., Marie, D., Sciandra, A., Bernard, O., and Rabouille, S. (2019). Picoeukaryotes of the *Micromonas* genus: sentinels of a warming ocean. *The ISME Journal*, 13(1):132–146.

Deng, L., Li, Y., Tanizawa, Y., Nakamura, Y., Kamikawa, R., Hopes, A., and Mock, T. (2025). A genetic transformation system for the heterotrophic diatom *Nitzschia putrida* (Bacillariophyceae). *Journal of Phycology*, 61:1116–1123.

Devol, A. H. (2015). Denitrification, anammox, and N_2 production in marine sediments. *Annual Review of Marine Science*, 7:403–423.

Diamond, A., Diaz-Garza, A. M., Li, J., Slattery, S. S., Merindol, N., and Fantino, E., et al. (2023). Instability of extrachromosomal DNA transformed into the diatom *Phaeodactylum tricornutum*. *Algal Research*, 70:102998.

Díaz-Santos, E., Vila, M., Vigara, J., and León, R. (2016). A new approach to express transgenes in microalgae and its use to increase the flocculation ability of *Chlamydomonas reinhardtii*. *Journal of Applied Phycology*, 28:1611–1621.

Dolatabadi, J. E. N. and de la Guardia, M. (2011). Applications of diatoms and silica nanotechnology in biosensing, drug and gene delivery, and formation of complex metal nanostructures. *TrAC Trends in Analytical Chemistry*, 30:1538–1548.

Donnelly, M. L. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D., and Ryan, M. D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *Journal of General Virology*, 82(5):1013–1025.

Dortch, Q. (1990). The interaction between ammonium and nitrate uptake in phytoplankton. *Marine Ecology Progress Series*, 61(1-2):183–201.

Dunahay, T. G., Jarvis, E. E., and Roessler, P. G. (1995). Genetic transformation of the diatoms *Cyclotella cryptica* and *Navicula saprophila*. *Journal of Phycology*, 31:1004–1012.

DuRand, M. D., Green, R. E., Sosik, H. M., and Olson, R. J. (2002). Diel variations in optical properties of *micromonas pusilla* (Prasinophyceae). *Journal of Phycology*, 38(6):1132–1142.

Díez, J., López-Lozano, A., Domínguez-Martín, M. A., Gómez-Baena, G., Muñoz-Marín, M. C., Melero-Rubio, Y., and García-Fernández, J. M. (2023). Regulatory and metabolic adaptations in the nitrogen assimilation of marine picocyanobacteria. *FEMS Microbiology Reviews*, 47(1):fuac043.

Elleuche, S. and Pöggeler, S. (2008). A cyanase is transcriptionally regulated by arginine and involved in cyanate decomposition in *Sordaria macrospora*. *Fungal Genetics and Biology*, 45(11):1458–1469.

Endo, H., Yoshida, M., Uji, T., Saga, N., Inoue, K., and Nagasawa, H. (2016). Stable nuclear transformation system for the coccolithophorid alga *Pleurochrysis carterae*. *Scientific Reports*, 6:22252.

Espie, G. S., Jalali, F., Tong, T., Zacal, N. J., and So, A. K.-C. (2007). Involvement of the cynABDS operon and the CO_2-concentrating mechanism in the light-dependent transport and metabolism of cyanate by cyanobacteria. *Journal of Bacteriology*, 189:1013–1024.

Faktorova, D., N.-R. E. R. F. R. J. A. C. E., Sudek, L., Allen, A. E., and et al. (2020). Genetic tool development in marine protists: emerging model organisms for experimental cell biology. *Nature Methods*, 17(5):481–494.

Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C., and Bowler, C. (1999). Transformation of nonselectable reporter genes in marine diatoms. *Marine Biotechnology*, 1(3):239–251.

Falkowski, P. G., Fenchel, T., and DeLong, E. F. (2008). The microbial engines that drive earth's biogeochemical cycles. *Science*, 320(5879):1034–1039.

Falkowski, P. G., Katz, M. E., Knoll, A. H., Quigg, A., Raven, J. A., Schofield, O., and Taylor, F. J. R. (2004). The evolution of modern eukaryotic phytoplankton. *Science*, 305(5682):354–360.

Falkowski, P. G. and Stone, D. P. (1975). Nitrate uptake in marine phytoplankton: Energy sources and the interaction with carbon fixation. *Marine Biology*, 32(1):77–84.

Fattorini, N. and Maier, U. G. (2021). Targeting of proteins to the cell wall of the diatom *Thalassiosira pseudonana*. *Discover Materials*, 1:5.

Ferrer-González, F. X., Hamilton, M., Smith, C. B., Schreier, J. E., Olofsson, M., and Moran, M. A. (2023). Bacterial transcriptional response to labile exometabolites from photosynthetic picoeukaryote *Micromonas commoda*. *ISME Communications*, 3(1):5.

Fichtbauer, A., Temmink, R. J. M., and La Russa, M. (2023). Controlling the nitrogen environment for optimal *Rhodomonas salina* production. *Journal of Applied Phycology*, 35:1565–1574.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science*, 281(5374):237–240.

Forde, B. G. and Cole, J. A. (2003). Nitrate finds a place in the sun. *Plant Physiology*, 131:395–400.

Fortunato, A. E., Jaubert, M., Enomoto, G., Bouly, J.-P., Raniello, R., Thaler, M., Malviya, S., Bernardes, J. S., Rappaport, F., Gentili, B., Huysman, M. J. J., Carbone, A., Bowler, C., Ribera d'Alcalà, M., Ikeuchi, M., and Falciatore, A. (2016). Diatom phytochromes reveal the existence of far-red-light-based sensing in the ocean. *Plant Cell*, 28:616–628.

Francis, C. A., Roberts, K. J., Beman, J. M., Santoro, A. E., and Oakley, B. B. (2005). Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proceedings of the National Academy of Sciences of the United States of America*, 102(41):14683–14688.

Frit, P., Barboule, N., Yuan, Y., Gomez, D., and Calsou, P. (2014). Alternative end-joining pathway(s): Bricolage at DNA breaks. *DNA Repair*, 17:81–97.

Fuchsman, C. A., Devol, A. H., Saunders, J. K., McKay, C., and Rocap, G. (2017). Niche partitioning of the N cycling microbial community of an offshore oxygen deficient zone. *Frontiers in Microbiology*, 8:2384.

Gaj, T., Gersbach, C. A., and Barbas, Carlos F., I. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, 31(7):397–405.

Gallaher, S. D., Craig, R. J., Ganesan, I., Purvine, S. O., McCorkle, S. R., Grimwood, J., Strenkert, D., Davidi, L., Roth, M. S., Jeffers, T. L., Lipton, M. S., Niyogi, K. K., Schmutz, J., Theg, S. M., Blaby-Haas, C. E., and Merchant, S. S. (2021). Widespread polycistronic gene expression in green algae. *Proceedings of the National Academy of Sciences of the United States of America*, 118(7):e2017714118.

Gallego, M. E., Bleuyard, J.-Y., Daoudal-Cotterell, S., Jallut, N., and White, C. I. (2003). Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *The Plant Journal*, 35(5):557–565.

Gao, G., Xia, J., Yu, J., and Zeng, X. (2018). Physiological response of a red tide alga (*Skeletonema costatum*) to nitrate enrichment, with special reference to inorganic carbon acquisition. *Marine Environmental Research*, 133:15–23.

Geider, R. J., Delucia, E. H., Falkowski, P. G., Finzi, A. C., Grime, J. P., Grace, J., Kana, T. M., La Roche, J., Long, S. P., Osborne, B. A., Platt, T., Prentice, I. C., Raven, J. A., Schlesinger, W. H., Smetacek, V., Stuart, V., Sathyendranath, S., Thomas, R. B., Vogelmann, T. C., Williams, P. J. l. B., and Woodward, F. I. (2001). Primary productivity of planet earth: biological determinants and physical constraints in terrestrial and aquatic habitats. *Global Change Biology*, 7:849–882.

Glibert, P. M. and Burford, M. A. (2017). Globally changing nutrient loads and harmful algal blooms: Recent advances, new paradigms, and continuing challenges. *Oceanography*, 30(1):58–69.

Gobler, C. J. and Sunda, W. G. (2012). Ecosystem disruptive algal blooms of the brown tide species, *Aureococcus anophagefferens* and *Aureoumbra lagunensis*. *Harmful Algae*, 14:36–45.

Görlich, S., Pawolski, D., Zlotnikov, I., and Kröger, N. (2019). Control of biosilica morphology and mechanical performance by the conserved diatom gene silicanin-1. *Communications Biology*, 2:245.

Gornik, S. G., Maegele, I., Hambleton, E. A., Voss, P. A., Waller, R. F., and Guse, A. (2022). Nuclear transformation of a dinoflagellate symbiont of corals. *Frontiers in Marine Science*, 9:1035413.

Gottlieb, T. M. and Jackson, S. P. (1993). The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen. *Cell*, 72(1):131–142.

Grawunder, U., Zimmer, D., Fugmann, S., Schwarz, K., and Lieber, M. R. (1998).

DNA ligase IV is essential for V(D)J recombination and DNA double-strand breakrepair in human precursor lymphocytes. *Molecular Cell*, 2(4):477–484.

Greiner, A., Kelterborn, S., Evers, H., Kreimer, G., Sizova, I., and Hegemann, P. (2017). Targeting of photoreceptor genes in *Chlamydomonas reinhardtii* via Zinc-Finger Nucleases and CRISPR/Cas9. *The Plant Cell*, 29(10):2498–2518.

Gresch, O., Engel, F. B., Nesic, D., Tran, T. T., England, H. M., Hickman, E. S., Körner, I., Gan, L., Chen, S., Castro-Obregon, S., Hammermann, R., Wolf, J., Müller-Hartmann, H., Nix, M., Siebenkotten, G., Kraus, G., and Lun, K. (2004). New non-viral method for gene transfer into primary cells. *Methods*, 33(2):151–163.

Gruber, N. (2008). The marine nitrogen cycle: overview and challenges. In Capone, D. G., Bronk, D. A., Mulholland, M. R., and Carpenter, E. J., editors, *Nitrogen in the Marine Environment*, pages 1–50. Academic Press, San Diego.

Gruber, N. and Galloway, J. N. (2008). An earth-system perspective of the global nitrogen cycle. *Nature*, 451(7176):293–296.

Grypioti, E., Richard, H., Kryovrysanaki, N., Jaubert, M., Falciatore, A., Verret, F., et al. (2024). Dicer-dependent heterochromatic small RNAs in the model diatom species *Phaeodactylum tricornutum*. *New Phytologist*, 241:811–826.

Gu, J., Wang, Q., Cui, M., et al. (2014). Cloning and characterization of Ku70 and Ku80 homologues involved in DNA repair process in wheat (*Triticum aestivum* L.). *Plant Genetic Resources*, 12(S1):S99–S103.

Guerrero, M. G., Vega, J. M., and Losada, M. (1981). The assimilatory nitrate-reducing system and its regulation. *Annual Review of Plant Physiology*, 32:169–204.

Guo, C., Anwar, M., Mei, R., Li, X., Zhao, D., Jiang, Y., Zhuang, J., Liu, C., Wang, C., and Hu, Z. (2022). Establishment and optimization of PEG-mediated protoplast transformation in the microalga *Haematococcus pluvialis*. *Journal of Applied Phycology*, 34:1595–1605.

Guo, J., Wilken, S., Jimenez, V., Mock, T., Schıkora, M., Voss, B., Ghirardi, M. L., Green, B. R., Falkowski, P. G., Allen, A. E., Smith, J. A., and Coesel, S. (2018). Specialized proteomic responses and an ancient photoprotection mechanism sustain marine green algal growth during phosphate limitation. *Nature Microbiology*, 3(7):781–790.

Guo, W., Weng, Y., Ma, W., Chang, C., Gao, Y., Huang, X., and Zhang, F. (2024). Improving lipid content in the diatom *Phaeodactylum tricornutum* by the knockdown of the enoyl-CoA hydratase using CRISPR interference. *Current Issues in Molecular Biology*, 46:10923–10933.

Guérin, N., Seyman, C., Orvain, C., Bertrand, L., Gourvil, P., Probert, I., Vacherie, B., Brun, , Magdelenat, G., Labadie, K., Wincker, P., Thurotte, A., and Carradec, Q. (2025). Transcriptomic response of the picoalga *Pelagomonas calceolata* to nitrogen availability: new insights into cyanate lyase function. *Microbiology Spectrum*, 13(1):e02654–24.

Hamilton, M., Ferrer-Gonz'alez, F. X., and Moran, M. A. (2024). Heterotrophic bacteria trigger transcriptome remodelling in the photosynthetic picoeukaryote Micromonas commoda. *Environmental Microbiology Reports*, 16(3):e13285.

Hamilton, M. L., Haslam, R. P., Napier, J. A., and Sayanova, O. (2014). Metabolic engineering of *Phaeodactylum tricornutum* for the enhanced accumulation of omega-3 long-chain polyunsaturated fatty acids. *Metabolic Engineering*, 22:3–9.

Hanahan, D., Jessee, J., and Bloom, F. R. (1991). [4] Plasmid transformation of *escherichia coli* and other bacteria. In *Bacterial Genetic Systems*, volume 204 of *Methods in Enzymology*, pages 63–113. Academic Press.

Hehenberger, E., Guo, J., Wilken, S., Hoadley, K., Sudek, L., Poirier, C., Dannebaum, R., Susko, E., and Worden, A. Z. (2023). Phosphate limitation responses in marine green algae are linked to reprogramming of the tRNA epitranscriptome and codon usage bias. *Molecular Biology and Evolution*, 40(12):msad251.

Hempel, F., Maurer, M., Brockmann, B., Mayer, C., Biedenkopf, N., Kelterbaum, A., Becker, S., and Maier, U. G. (2017). From hybridomas to a robust microalgal-based production platform: molecular design of a diatom secreting monoclonal antibodies directed against the marburg virus nucleoprotein. *Microbial Cell Factories*, 16(1):131.

Heyer, W.-D., Ehmsen, K. T., and Liu, J. (2010). Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics*, 44:113–139.

Hirooka, S., Itabashi, T., Ichinose, T. M., Onuma, R., Fujiwara, T., Yamashita, S., Jong, L. W., Tomita, R., Iwane, A. H., and Miyagishima, S. (2022). Life cycle and functional genomics of the unicellular red alga *Galdieria* for elucidating algal and plant evolution and industrial use. *Proceedings of the National Academy of Sciences of the United States of America*, 119(41):e2210665119.

Hopes, A. (2017). *Expanding the molecular toolbox in diatoms: developing a transformation system, CRISPR-Cas and inverse yeast one-hybrid*. PhD thesis, University of East Anglia.

Hopes, A., Nekrasov, V., Kamoun, S., and et al. (2016). Editing of the urease gene by CRISPR-Cas in the diatom *Thalassiosira pseudonana*. *Plant Methods*, 12:49.

Horák, A., Allen, A., and Oborník, M. (2020). Common origin of ornithine–urea cycle in opisthokonts and stramenopiles. *Scientific Reports*, 10:16687.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., and Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, 31(9):827–832.

Hu, Z., Duan, S., Xu, N., and Mulholland, M. R. (2014). Growth and nitrogen uptake kinetics in cultured *Prorocentrum donghaiense*. *PLOS ONE*, 9(4):e94030.

Hu, Z., Mulholland, M. R., Duan, S., and Xu, N. (2012). Effects of nitrogen supply and its composition on the growth of *Prorocentrum donghaiense*. *Harmful Algae*, 13:72–82.

Huang, C. C., Chen, M. W., Hsieh, J. L., Lin, W. H., Chen, P. C., and Chien, L. F. (2006). Expression of mercuric reductase from *Bacillus megaterium* MB1 in eukaryotic microalga *Chlorella* sp. DT: an approach for mercury phytoremediation. *Applied Microbiology and Biotechnology*, 72:197–205.

Huang, K., Feng, Q., Zhang, Y., Ou, L., Cen, J., Lu, S., and Qi, Y. (2020). Comparative uptake and assimilation of nitrate, ammonium, and urea by dinoflagellate *Karenia mikimotoi* and diatom *Skeletonema costatum* s.l. in the coastal waters of the East China Sea. *Marine Pollution Bulletin*, 155:111200.

Huang, W. and Daboussi, F. (2017). Genetic and metabolic engineering in diatoms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372:20160411.

Hutchins, D. A. and Boyd, P. W. (2016). Marine phytoplankton and the changing ocean iron cycle. *Nature Climate Change*, 6:1072–1079.

Hutchins, D. A. and Capone, D. G. (2022). The marine nitrogen cycle: new developments and global change. *Nature Reviews Microbiology*, 20(7):401–414.

Hutchins, D. A. and Sañudo-Wilhelmy, S. A. (2022). The enzymology of ocean global change. *Annual Review of Marine Science*, 14:187–211.

Ifuku, K., Yan, D., Miyahara, M., Inoue-Kashino, N., Yamamoto, Y. Y., and Kashino, Y. (2015). A stable and efficient nuclear transformation system for the diatom *Chaetoceros gracilis*. *Photosynthesis Research*, 123:203–211.

Jeon, M. S., Han, S.-I., Jeon, M., and Choi, Y.-E. (2021). Enhancement of phycoerythrin productivity in *Porphyridium purpureum* using the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 ribonucleoprotein system. *Bioresource Technology*, 330:124974.

Ji, N., Zhang, Z., Huang, J., Zhou, L., Deng, S., Shen, X., and Lin, S. (2020). Utilization of various forms of nitrogen and expression regulation of transporters in the harmful alga *Heterosigma akashiwo* (raphidophyceae). *Harmful Algae*, 92:101770.

Jiang, W., Brueggeman, A. J., Horken, K. M., Plucinak, T. M., and Weeks, D. P. (2014). Successful transient expression of Cas9 and single guide RNA genes in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*, 13(11):1465–1469.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable Dual-RNA–Guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096):816–821.

Johnson, W. V. and Anderson, P. M. (1987). Bicarbonate is a recycling substrate for cyanase. *Journal of Biological Chemistry*, 262(19):9021–9025.

Kamalanathan, M., Chiu, M.-H., Bacosa, H., Schwehr, K., Tsai, S.-M., Doyle, S., Yard, A., Mapes, S., Vasequez, C., Bretherton, L., Sylvan, J. B., Santschi, P., Chin, W.-C., and Quigg, A. (2019). Role of polysaccharides in diatom *Thalassiosira pseudonana* and its associated bacteria in hydrocarbon presence. *Plant Physiology*, 180(4):1898–1911.

Kamennaya, N. A., Chernihovsky, M., and Post, A. F. (2008). The cyanate utilization capacity of marine unicellular cyanobacteria. *Limnology and Oceanography*, 53(6):2485–2494.

Kamennaya, N. A. and Post, A. F. (2011). Characterization of cyanate metabolism in marine *Synechococcus* and *Prochlorococcus* spp. *Applied and Environmental Microbiology*, 77(9):2912–2919.

Karas, B. J., Diner, R. E., Lefebvre, S. C., McQuaid, J., Phillips, A. P. R., Noddings, C. M., Brunson, J. K., Valas, R. E., Deerinck, T. J., Jablanovic, J., Gillard, J. T. F., Beeri, K., Ellisman, M. H., Glass, J. I., Hutchison, C. A., Smith, H. O., Venter, J. C., and Allen, A. E. (2015). Designer diatom episomes delivered by bacterial conjugation. *Nature Communications*, 6:6925.

Karl, D. M., Letelier, R., Tupas, L., Dore, J., Christian, J., and Hebel, D. (1997). The role of nitrogen fixation in biogeochemical cycling in the subtropical north Pacific Ocean. *Nature*, 388:533–538.

Karthäuser, C., Ahmerkamp, S., Marchant, H. K., Bristow, L. A., Pjevac, P., Schubert, C. J., and Kuypers, M. M. M. (2021). Small sinking particles control anammox rates in the Peruvian oxygen minimum zone. *Nature Communications*, 12:3235.

Kastan, M. B. and Bartek, J. (2004). Cell-cycle checkpoints and cancer. *Nature*, 432(7015):316–323.

Kato, S. and Nam, H. G. (2021). The cell division cycle of euglena gracilis indicates that the level of circadian plasticity to the external light regime changes in prolonged-stationary cultures. *Plants*, 10:1475.

Kay, H. M. M. (2022). *Post-Transcriptional Regulation of the Proteome in Ostreococcus tauri*. PhD thesis, The University of Edinburgh.

Khanna, K. K. and Jackson, S. (2001). DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genetics*, 27(3):247–254.

Kim, D. H., Kim, Y. T., Cho, J. J., Bae, J. H., Hur, S. B., Hwang, I., and Choi, T. J. (2002). Stable integration and functional expression of flounder growth hormone gene in transformed microalga, *Chlorella ellipsoidea*. *Marine Biotechnology*, 4:63–73.

Kim, J., Chang, K. S., Lee, S., and Jin, E. (2021). Establishment of a genome editing tool using CRISPR-Cas9 in *Chlorella vulgaris* utex395. *International Journal of Molecular Sciences*, 22:480.

Kim, J. H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K. Y., Kim, M.-K., Shin, B. A., and Choi, S.-Y. (2011). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLOS ONE*, 6(4):e18556.

Kim, J. S., Lee, S., Cho, S., and Jung, Y. (2024). Inducing heritable genomic deletions in APT gene of *Chlorella sorokiniana* using CRISPR/Cas9. *Algal Research*, 79:103435.

Kindle, K. L. (1990). High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America*, 87(3):1228–1232.

Kira, N., Ohnishi, K., Miyagawa-Yamaguchi, A., Kadono, T., and Adachi, M. (2016). Nuclear transformation of the diatom *Phaeodactylum tricornutum* using PCR-amplified DNA fragments by microparticle bombardment. *Marine Genomics*, 25:49–56.

Kirkham, A. R., Richthammer, P., Schmidt, K., Wustmann, M., Maeda, Y., Hedrich, R., Brunner, E., Tanaka, T., van Pée, K. H., Falciatore, A., and Mock, T. (2017). A role for the cell-wall protein silacidin in cell size of the diatom *Thalassiosira pseudonana*. *The ISME Journal*, 11:2452–2464.

Kitzinger, K., Padilla, C. C., Marchant, H. K., Hach, P. F., Herbold, C. W., Kidane, A. T., Littmann, S., and Daims, H. (2019). Cyanate and urea are substrates for nitrification by thaumarchaeota in the marine environment. *Nature Microbiology*, 4(2):234–243.

Knapp, A. N. (2012). The sensitivity of marine N_2 fixation to dissolved inorganic nitrogen. *Frontiers in Microbiology*, 3:1–14.

Koh, H. G., Kang, N. K., Kim, E. K., Jeon, S., Shin, S.-E., Lee, B., and Chang, Y. K. (2018). Advanced multigene expression system for *Nannochloropsis salina* using 2A self-cleaving peptides. *Journal of Biotechnology*, 278:39–47.

Kotnik, T., Frey, W., Sack, M., Haberl Meglič, S., Peterka, M., and Miklavčič, D. (2015). Electroporation-based applications in biotechnology. *Trends in Biotechnology*, 33(8):480–488.

Kovács, E. and Keresztes, (2002). Effect of gamma and UV-B/C radiation on plant cells. *Micron*, 33(2):199–210.

Krens, F. A., Molendijk, L., Wullems, G. J., and Schilperoort, R. A. (1982). *In vitro* transformation of plant protoplasts with Ti-plasmid DNA. *Nature*, 296:72–74.

Krishnan, A., Cano, M., Burch, T. A., Weissman, J. C., and Posewitz, M. C. (2020). Genome editing using Cas9-RNA ribonucleoprotein complexes in the high-productivity marine alga *Picochlorum celeri*. *Algal Research*, 49:101944.

Kumar, M., Tibocha-Bonilla, J. D., Füssy, Z., Lieng, C., Schwenck, S. M., Levesque, A. V., Al-Bassam, M. M., Passi, A., Neal, M., Zuniga, C., Kaiyom, F., Espinoza, J. L., Lim, H., Polson, S. W., Allen, L. Z., and Zengler, K. (2024). Mixotrophic growth of a ubiquitous marine diatom. *Science Advances*, 10(29):eado2623.

Kunz, D. A. and Nagappan, O. (1989). Cyanase-mediated utilization of cyanate in *Pseudomonas fluorescens* NCIB 11764. *Applied and Environmental Microbiology*, 55(1):256–258.

Kuypers, M. M. M., Marchant, H. K., and Kartal, B. (2018). The microbial nitrogen-cycling network. *Nature Reviews Microbiology*, 16(5):263–276.

Lai, Q., Wang, Y., Zhou, Q., and Zhao, Z. (2020). Isolation and purification of mesophyll protoplasts from *Ginkgo biloba* l. *Cytologia*, 85(1):27–32.

Lander, E. S. (2016). The Heroes of CRISPR. *Cell*, 164:18–28.

Lentz, B. R. and Lee, J. (1999). Poly(ethylene glycol) (PEG)-mediated fusion between pure lipid bilayers: a mechanism in common with viral fusion and secretory vesicle release? *Molecular Membrane Biology*, 16(4):279–296.

Levitan, O., Dinamarca, J., Zelzion, E., Lun, D. S., Guerra, L. T., Kim, M. K., Kim, J., Van Mooy, B. A. S., Bhattacharya, D., and Falkowski, P. G. (2015). Remodeling of intermediate metabolism in the diatom *Phaeodactylum tricornutum* under nitrogen stress. *Proceedings of the National Academy of Sciences of the United States of America*, 112(2):412–417.

Li, C., Yin, W., Pan, Y., and Hu, H. (2026). Interactions with bacteria shape diatom adaptation to carbon concentration changes. *Nature Communications*, 17:1289.

Li, J., Glibert, P. M., Zhou, M., Lu, S., and Lu, D. (2009). Relationships between nitrogen and phosphorus forms and ratios and the development of dinoflagellate blooms in the East China Sea. *Marine Ecology Progress Series*, 383:11–26.

Li, X., Zeng, C., Han, X., Fan, L., Zhao, J., Wang, Y., Wang, X., and Li, K. (2025a). The effect of the aging processes of Ulva prolifera-derived dissolved organic nitrogen associated with green tide on the diatoms-dinoflagellates succession in the Southern Yellow Sea, China. *Water Research*, 277:123370.

Li, X., Zeng, C., Han, X., Wang, X., and Li, K. (2025b). The bioavailability and component characteristics of the aging dissolved organic matter (DOM) from the macroalgae *Ulva prolifera* in seawater. *Marine Environmental Research*, 204:106946.

Li, Y., Deng, L., Walker, E. J. L., Karas, B. J., and Mock, T. (2025c). Genetic engineering in diatoms: advances and prospects. *The Plant Journal*, 121(6):e70102.

Liang, H., Li, T., Chen, Y., Wang, J., Aslam, M., Qin, H., Fan, W., Du, H., Kao, S.-j., and Lin, S. (2025). Urea amidolyase as an enzyme for urea utilisation in phytoplankton: Functional display in *Chlamydomonas reinhardtii*. *Molecular Ecology*, 34:e17734.

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry*, 79:181–211.

Lim, S., Yocum, R. R., Silver, P. A., and Way, J. C. (2023). High spontaneous integration rates of end-modified linear DNAs upon mammalian cell transfection. *Scientific Reports*, 13:6835.

Lin, S., Staahl, B. T., Alla, R. K., and Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife*, 3:e04766.

Linder, T. (2019). Cyanase-independent utilization of cyanate as a nitrogen source in ascomycete yeasts. *World Journal of Microbiology and Biotechnology*, 35(1):3.

Liu, Z., Chen, O., Wall, J. B. J., Zheng, M., Zhou, Y., Wang, L., Vaseghi, H. R.,

Qian, L., and Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Scientific Reports*, 7:2193.

Lomas, M. W. and Glibert, P. M. (2000). Comparisons of nitrate uptake, storage, and reduction in marine diatoms and flagellates. *Journal of Phycology*, 36(5):903–913.

Lozano, J.-C., Schatt, P., Botebol, H., Vergé, V., Lesuisse, E., Blain, S., Carré, I. A., and Bouget, F.-Y. (2014). Efficient gene targeting and removal of foreign DNA by homologous recombination in the picoeukaryote *Ostreococcus*. *The Plant Journal*, 78(6):1073–1083.

Maeda, Y., Nakamura, M., Watanabe, K., Okamoto, E., and Tanaka, T. (2023). Functional analysis of the putative centromere sequences of marine oleaginous diatom *Fistulifera solaris*. *Algal Research*, 74:103225.

Malzahn, A., Lowder, L., and Qi, Y. (2017). Plant genome editing with TALEN and CRISPR. *Cell & Bioscience*, 7:21.

Manolis, K. G., Nimmo, E. R., Hartsuiker, E., Carr, A. M., Jeggo, P. A., and Allshire, R. C. (2001). Novel functional requirements for non-homologous DNA end joining in *Schizosaccharomyces pombe*. *EMBO Journal*, 20(1-2):210–221.

Mao, X., Chen, J., van Oosterhout, C., Zhang, H., Liu, G., Zhuang, Y., and Mock, T. (2022). Diversity, prevalence, and expression of cyanase genes (*cynS*) in planktonic marine microorganisms. *The ISME Journal*, 16(2):602–605.

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011). Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom *Thalassiosira pseudonana*. *The ISME Journal*, 5(6):1057–1060.

Massana, R. (2011). Eukaryotic picoplankton in surface oceans. *Annual Review of Microbiology*, 65:91–110.

Matsui, H., Harada, H., Maeda, K., Sugiyama, T., Fukuchi, Y., Kimura, N., Nawaly, H., Tsuji, Y., and Matsuda, Y. (2024). Coordinated phosphate uptake

by extracellular alkaline phosphatase and solute carrier transporters in marine diatoms. *New Phytologist*, 241(3):1210–1221.

Matthews, H. K., Bertoli, C., and de Bruin, R. A. M. (2022). Cell cycle control in cancer. *Nature Reviews Molecular Cell Biology*, 23(2):74–88.

Maxwell, K. and Johnson, G. N. (2000). Chlorophyll fluorescence–a practical guide. *Journal of Experimental Botany*, 51(345):659–668.

McCarthy, J. J. (1972). The uptake of urea by natural populations of marine phytoplankton. *Limnology and Oceanography*, 17(5):738–748.

McCarthy, J. K., Smith, S. R., McCrow, J. P., Tan, M., Zheng, H., Beeri, K., Roth, R., Lichtle, C., Goodenough, U., Bowler, C., Dupont, C. L., and Allen, A. E. (2017). Nitrate reductase knockout uncouples nitrate transport from nitrate assimilation and drives repartitioning of carbon flux in a model pennate diatom. *The Plant Cell*, 29(8):2047–2070.

McCarthy, M., Pratum, T., Hedges, J., and Benner, R. (1997). Chemical composition of dissolved organic nitrogen in the ocean. *Nature*, 390:150–154.

McDonald, S. M., Plant, J. N., and Worden, A. Z. (2010). The mixed lineage nature of nitrogen transport and assimilation in marine eukaryotic phytoplankton: A case study of *Micromonas*. *Molecular Biology and Evolution*, 27(10):2268–2283.

McRose, D., Guo, J., Monier, A., Sudek, S., Wilken, S., Yan, S., Mock, T., Archibald, J. M., Begley, T. P., Reyes-Prieto, A., and Worden, A. Z. (2014). Alternatives to vitamin B1 uptake revealed with discovery of riboswitches in multiple marine eukaryotic lineages. *The ISME Journal*, 8(12):2517–2529.

Mehravar, M., Shirazi, A., Mehrazar, M. M., and Nazari, M. (2019). *In Vitro* prevalidation of gene editing by CRISPR/Cas9 ribonucleoprotein. *Avicenna Journal of Medical Biotechnology*, 11(3):259–263.

Messemer, A., Winge, P., and Brembu, T. (2026). Improvement of molecular tools for the model diatom *Thalassiosira pseudonana*. *Algal Research*, 94:104509.

Meyer, C., Lea, U. S., Provan, F., Kaiser, W. M., and Lillo, C. (2005). Is nitrate reductase a major player in the plant NO (nitric oxide) game? *Photosynthesis Research*, 83:181–189.

Meyer, M. G., White, V. J., Torano, O., Hannoush, H., Lankford, M., and Marchetti, A. (2025). The effects of iron limitation on the small chlorophyte *Micromonas* from the northeast pacific ocean. *Journal of Phycology*, 61:1251–1262.

Miflin, B. J. and Lea, P. J. (1975). Glutamine and asparagine as nitrogen donors for reductant-dependent glutamate synthesis in pea roots. *Biochemical Journal*, 149:403–409.

Minoda, A., Sakagami, R., Yagisawa, F., Kuroiwa, T., and Tanaka, K. (2004). Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. *Plant and Cell Physiology*, 45(6):667–671.

Minskaia, E., Nicholson, J., and Ryan, M. D. (2013). Optimisation of the foot-and-mouth disease virus 2A co-expression system for biomedical applications. *BMC Biotechnology*, 13:67.

Mock, T., Samanta, M. P., Iverson, V., Berthiaume, C., Robison, M., Holtermann, K., Durkin, C., BonDurant, S. S., Richmond, K., Rodesch, M., Kallas, T., Huttlin, E. L., Cerrina, F., Sussman, M. R., and Armbrust, E. V. (2008). Whole-genome expression profiling of the marine diatom *Thalassiosira pseudonana* identifies genes involved in silicon bioprocesses. *Proceedings of the National Academy of Sciences of the United States of America*, 105(5):1579–1584.

Monier, A., Worden, A. Z., and Richards, T. A. (2016). Phylogenetic diversity and biogeography of the mamiellophyceae lineage of eukaryotic phytoplankton across the oceans. *Environmental Microbiology Reports*, 8(4):461–469.

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., Galbraith, E. D., Geider, R. J., Guieu, C., Jaccard, S. L., Jickells, T. D., La Roche, J., Lenton, T. M., Mahowald, N. M., Marañón, E., Marinov, I., Moore,

J. K., Nakatsuka, T., Oschlies, A., Saito, M. A., Thingstad, T. F., Tsuda, A., and Ulloa, O. (2013). Processes and patterns of oceanic nutrient limitation. *Nature Geoscience*, 6(9):701–710.

Moosburner, M., Allen, A. E., and Daboussi, F. (2022). Genetic engineering in marine diatoms: current practices and emerging technologies. In Falciatore, A. and Mock, T., editors, *The Molecular Life of Diatoms*, pages 743–773. Springer, London.

Moosburner, M. A., Gholami, P., McCarthy, J. K., Tan, M., Bielinski, V. A., and Allen, A. E. (2020). Multiplexed knockouts in the model diatom *Phaeodactylum* by episomal delivery of a selectable Cas9. *Frontiers in Microbiology*, 11:5.

Mulholland, M. R. and Lee, C. (2009). Peptide hydrolysis and the uptake of dipeptides by phytoplankton. *Limnology and Oceanography*, 54(3):856–868.

Muñoz, C. F., Weusthuis, R. A., D'Adamo, S., and Wijffels, R. H. (2019). Effect of single and combined expression of lysophosphatidic acid acyltransferase, glycerol-3-phosphate acyltransferase, and diacylglycerol acyltransferase on lipid accumulation and composition in *Neochloris oleoabundans*. *Frontiers in Plant Science*, Volume 10 - 2019.

Naduthodi, M. I. S., Mohanraju, P., Südfeld, C., D'Adamo, S., Barbosa, M. J., and van der Oost, J. (2019). CRISPR–Cas ribonucleoprotein mediated homology-directed repair for efficient targeted genome editing in microalgae *Nannochloropsis oceanica* imet1. *Biotechnology for Biofuels*, 12:66.

Nagaria, P., Robert, C., and Rassool, F. V. (2013). DNA double-strand break response in stem cells: Mechanisms to maintain genomic integrity. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1830(2):2345–2353.

Naselli-Flores, L. and Padisák, J. (2023). Ecosystem services provided by marine and freshwater phytoplankton. *Hydrobiologia*, 850:2691–2706.

Nawaly, H., Tsuji, Y., and Matsuda, Y. (2020). Rapid and precise genome editing

in a marine diatom, *Thalassiosira pseudonana*, by Cas9 nickase (D10A). *Algal Research*, 47:101855.

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A., and Quéguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles*, 9(3):359–372.

Nomura, T., Inoue, K., Uehara-Yamaguchi, Y., Yamada, K., Iwata, O., Suzuki, K., and Mochida, K. (2019). Highly efficient transgene-free targeted mutagenesis and single-stranded oligodeoxynucleotide-mediated precise knock-in in the industrial microalga *Euglena gracilis* using Cas9 ribonucleoproteins. *Plant Biotechnology Journal*, 17:2032–2034.

Nomura, T., Kim, J.-S., Ishikawa, M., Suzuki, K., and Mochida, K. (2024). High-efficiency genome editing by Cas12a ribonucleoprotein complex in *Euglena gracilis*. *Microbial Biotechnology*, 17:e14393.

Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M., and Winge, P. (2016). A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports*, 6:24951.

Ohnuma, M., Yokoyama, T., Inouye, T., Sekine, Y., Kuroiwa, T., and Tanaka, K. (2014). Optimization of polyethylene glycol (PEG)-mediated DNA introduction conditions for transient gene expression in the unicellular red alga *Cyanidioschyzon merolae*. *The Journal of General and Applied Microbiology*, 60(4):156–159.

Onozawa, M., Zhang, Z., Kim, Y. J., Goldberg, L., Varga, T., Bergsagel, P. L., Kuehl, W. M., and Aplan, P. D. (2014). Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. *Proceedings of the National Academy of Sciences of the United States of America*, 111(21):7729–7734.

Otte, A., Hopes, A., and Mock, T. (2025). Evolution of microalgae and their

adaptations in different marine ecosystems. In *Encyclopedia of Life Sciences*, volume 4, pages 1–11. John Wiley & Sons, Ltd.

Ou, L., Lundgren, V., Lu, S., and Granéli, E. (2014). The effect of riverine dissolved organic matter and other nitrogen forms on the growth and physiology of the dinoflagellate *Prorocentrum minimum* (pavillard) schiller. *Journal of Sea Research*, 85:499–507.

Paerl, R. W., Bouget, F.-Y., Lozano, J. C., Vergé, V., Schatt, P., Allen, A. E., and Palenik, B. (2017). Use of plankton-derived vitamin B1 precursors, especially thiazole-related precursor, by key marine picoeukaryotic phytoplankton. *The ISME Journal*, 11(3):753–765.

Pageault, L., Charrier, A., Saint-Jean, B., Bougaran, G., Mairet, F., and Stachowski-Haberkorn, S. (2024). Cell cycle dynamics in the microalga *Tisochrysis lutea*: Influence of light duration and drugs. *Cells*, 13:1925.

Palatinszky, M., Herbold, C., Jehmlich, N., Pogoda, M., Han, P., von Bergen, M., Lagkouvardos, I., Karst, S. M., Kirkegaard, R. H., Albertsen, M., Nielsen, P. H., Wagner, M., and Daims, H. (2015). Cyanate as an energy source for nitrifiers. *Nature*, 524(7563):105–108.

Palenik, B., Brahamsha, B., Larimer, F. W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, W., Allen, E. E., McCarren, J., Paulsen, I., Dufresne, A., Partensky, F., Webb, E. A., and Waterbury, J. (2003). The genome of a motile marine *Synechococcus*. *Nature*, 424:1037–1042.

Pechkovskaya, S., Knyazev, N., Matantseva, O., Emelyanov, A., Telesh, I., Skarlato, S., and Filatova, N. (2020). Dur3 and Nrt2 genes in the bloom-forming dinoflagellate *Prorocentrum minimum*: Transcriptional responses to available nitrogen sources. *Chemosphere*, 241:125083.

Pereira, H., Schulze, P. S. C., Schüler, L. M., Santos, T., Barreira, L., and Varela, J. (2018). Fluorescence activated cell-sorting principles and applications in microalgal biotechnology. *Algal Research*, 30:113–120.

Pickar-Oliver, A. and Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology*, 20:490–507.

Pierella Karlusich, J. J., Cosnier, K., Zinger, L., Henry, N., Nef, C., Bernard, G., Scalco, E., Dvorak, E., Coordinators, T. O., Vieira, F. R. J., Delage, E., Chaffron, S., Ovchinnikov, S., Zingone, A., and Bowler, C. (2025). Patterns and drivers of diatom diversity and abundance in the global ocean. *Nature Communications*, 16:3452.

Plucinak, T. M., Horken, K. M., Jiang, W., Fostvedt, J., Nguyen, S. T., and Weeks, D. P. (2015). Improved and versatile viral 2A platforms for dependable and inducible high-level expression of dicistronic nuclear genes in *Chlamydomonas reinhardtii*. *The Plant Journal*, 82:717–729.

Poulsen, N., Chesley, P. M., and Kröger, N. (2006). Molecular genetic manipulation of the diatom *Thalassiosira pseudonana* (Bacillariophyceae). *Journal of Phycology*, 42:1059–1065.

Poulsen, N. and Kröger, N. (2023). *Thalassiosira pseudonana* (*Cyclotella nana*) (Hustedt) Hasle et Heimdal (Bacillariophyceae): A genetically tractable model organism for studying diatom biology, including biological silica formation. *Journal of Phycology*, 59:809–817.

Price, D. C., Farinholt, N., Gates, C., Shumaker, A., Wagner, N. E., Bienfang, P., and Bhattacharya, D. (2016). Analysis of gambierdiscus transcriptome data supports ancient origins of mixotrophic pathways in dinoflagellates. *Environmental Microbiology*, 18(11):4501–4510.

Puchta, H. (2005). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany*, 56(409):1–14.

Pudney, A., Gandini, C., Economou, C. K., Smith, R., Goddard, P., Napier, J. A., Spicer, A., and Sayanova, O. (2019). Multifunctionalizing the marine diatom *Phaeodactylum tricornutum* for sustainable co-production of omega-3 long

chain polyunsaturated fatty acids and recombinant phytase. *Scientific Reports*, 9(1):11444.

Pulz, O. and Gross, W. (2004). Valuable products from biotechnology of microalgae. *Applied Microbiology and Biotechnology*, 65(6):635–648.

Raven, J. A. (1987). The role of vacuoles. *New Phytologist*, 106(3):357–422.

Raven, J. A. and Giordano, M. (2016). *Combined Nitrogen*, pages 143–154. Springer International Publishing, Cham.

Reckermann, M. (2000). Flow sorting in aquatic ecology. *Scientia Marina*, 64:235–246.

Rees, A. P., Woodward, E. M. S., and Joint, I. (2006). Concentrations and uptake of nitrate and ammonium in the atlantic ocean between 60°N and 50°S. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(14):1649–1665. Atlantic Meridional Transect.

Rey Redondo, E., Xu, Y., and Yung, C. C. M. (2024). Genomic characterisation and ecological distribution of *Mantoniella tinhauana*: a novel mamiellophycean green alga from the western pacific. *Frontiers in Microbiology*, 15:1358574.

Rismani-Yazdi, H., Haznedaroglu, B. Z., Hsin, C., and Peccia, J. (2012). Transcriptomic analysis of the oleaginous microalga *Neochloris oleoabundans* reveals metabolic insights into triacylglyceride accumulation. *Biotechnology for Biofuels*, 5:74.

Rocap, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N. A., Arellano, A., Coleman, M., Hauser, L., Hess, W. R., Johnson, Z. I., Land, M., Lindell, D., Post, A. F., Regala, W., Shah, M., Shaw, S. L., Steglich, C., Sullivan, M. B., Ting, C. S., Tolonen, A., Webb, E. A., Zinser, E. R., and Chisholm, S. W. (2003). Genome divergence in two *Prochlorococcus ecotypes* reflects oceanic niche differentiation. *Nature*, 424:1042–1047.

Ryan, M. D., King, A. M. Q., and Thomas, G. P. (1991). Cleavage of foot-and-mouth

disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *Journal of General Virology*, 72(11):2727–2732.

Sabatino, V., Orefice, I., Marotta, P., Ambrosino, L., Chiusano, M. L., d'Ippolito, G., Romano, G., Fontana, A., and Ferrante, M. I. (2022). Silencing of a *Pseudo-nitzschia arenysensis* lipoxygenase transcript leads to reduced oxylipin production and impaired growth. *New Phytologist*, 233:809–822.

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

Sancar, A., Lindsey-Boltz, L. A., Ünsal-Kaçmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual Review of Biochemistry*, 73:39–85.

Sanchez, F., Geffroy, S., Norest, M., Yau, S., Moreau, H., and Grimsley, N. (2019). Simplified transformation of *Ostreococcus tauri* using polyethylene glycol. *Genes*, 10(5):399.

Sanz-Luque, E., Chamizo-Ampudia, A., Llamas, A., Galvan, A., and Fernandez, E. (2015). Understanding nitrate assimilation and its regulation in microalgae. *Frontiers in Plant Science*, 6:899.

Sasaki, M., Lange, J., and Keeney, S. (2010). Genome destabilization by homologous recombination in the germ line. *Nature Reviews Molecular Cell Biology*, 11(3):182–195.

Sato, M., Hashihama, F., and Takeda, S. (2023). Effects of cyanate enrichment on growth of natural phytoplankton populations in the subtropical Pacific. *Journal of Oceanography*, 79:1–12.

Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., Post, A. F., Hagemann, M., Paulsen, I., and Partensky, F. (2009).

Ecological genomics of marine picocyanobacteria. *Microbiology and Molecular Biology Reviews*, 73(2):249–299.

Schellenberger Costa, B., Sachse, M., Jungandreas, A., Bartulos, C. R., Gruber, A., Jakob, T., Kroth, P. G., and Wilhelm, C. (2013). Aureochrome 1a is involved in the photoacclimation of the diatom *Phaeodactylum tricornutum*. *PLoS ONE*, 8:e74451.

Schirrmeister, B. E., Gugger, M., and Donoghue, P. C. J. (2015). Cyanobacteria and the great oxidation event: evidence from genes and fossils. *Palaeontology*, 58(5):769–785.

Schlesinger, W. H. (1997). *Biogeochemistry: An Analysis of Global Change*. Academic Press, New York.

Schvarcz, C. R., Wilson, S. T., Caffin, M., Turk-Kubo, K. A., Orellana, M. V., Sañudo-Wilhelmy, S. A., Shoemaker, J. B., Varaljay, V. A., Lamb, J. J., Santoro, A. E., and Zehr, J. P. (2022). Overlooked and widespread pennate diatom–diazotroph symbioses in the sea. *Nature Communications*, 13:799.

Segev, E., Wyche, T. P., Kim, K. H., Petersen, J., Ellebrandt, C., Vlamakis, H., Barteneva, N., Paulson, J. N., Chai, L., Clardy, J., and Kolter, R. (2016). Dynamic metabolic exchange governs a marine algal-bacterial interaction. *eLife*, 5:e17473.

Serif, M., Dubois, G., Finoux, A. L., Teste, M.-A., Jallet, D., and Daboussi, F. (2018). One-step generation of multiple gene knock-outs in the diatom *Phaeodactylum tricornutum* by DNA-free genome editing. *Nature Communications*, 9:3924.

Sharpe, G., Zhao, L., Meyer, M. G., Gong, W., Burns, S. M., Tagliabue, A., Buck, K. N., Santoro, A. E., Graff, J. R., Marchetti, A., and Gifford, S. (2023). *Synechococcus* nitrogen gene loss in iron-limited ocean regions. *ISME Communications*, 3(1):107.

Shin, S. E., Lim, J. M., Koh, H., Kim, E. K., Kang, N. S., Jeon, S., Kwon, S. J., Shin, W. K., Lee, B. S., Hwangbo, K., Kim, J. S., and Jeong, W. J. (2016).

CRISPR/Cas9-induced knockout and knock-in mutations in *Chlamydomonas reinhardtii*. *Scientific Reports*, 6:27810.

Shinohara, A., Ogawa, H., and Ogawa, T. (1992). Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell*, 69(3):457–470.

Shrestha, R. P. and Hildebrand, M. (2017). Development of a silicon limitation inducible expression system for recombinant protein production in the centric diatoms *Thalassiosira pseudonana* and *Cyclotella cryptica*. *Microbial Cell Factories*, 16:145.

Shrivastav, M., De Haro, L. P., and Nickoloff, J. A. (2008). Regulation of DNA double-strand break repair pathway choice. *Cell Research*, 18:134–147.

Simó, R., Mangot, J.-F., Vergé, V., Lozano, J.-C., Saló, V., Ruiz-González, C., López-Sandoval, D., Castillo, Y. M., Kiene, R. P., Archer, S. D., Bouget, F.-Y., and González, J. M. (2026). Identification of a membrane transporter for dimethylsulfoniopropionate uptake in marine eukaryotic phytoplankton. *Nature Communications*.

Sinha, A., Wetten, A. C., and Caligari, P. D. S. (2003). Effect of biotic factors on the isolation of *Lupinus albus* protoplasts. *Australian Journal of Botany*, 51:103–109.

Slattery, S. S., Diamond, A., Wang, H., Therrien, J. A., Lant, J. T., Jazey, T., Lee, K., Klassen, Z., Desgagné-Penix, I., Karas, B. J., and Edgell, D. R. (2018). An expanded plasmid-based genetic toolbox enables Cas9 genome editing and stable maintenance of synthetic pathways in *Phaeodactylum tricornutum*. *ACS Synthetic Biology*, 7:328–338.

Smayda, T. J. (1997). Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography*, 42(5 part 2):1137–1153.

Smith, S. R., Dupont, C. L., McCarthy, J. K., Broddrick, J. T., Oborník, M., Horák, A., Füssy, Z., Cihlár, J., Kleessen, S., Zheng, H., McCrow, J. P., Hixson, K. K., Araújo, W. L., Nunes-Nesi, A., Fernie, A., Nikoloski, Z., Palsson, B. O., and Allen,

A. E. (2019). Evolution and regulation of nitrogen flux through compartmentalized metabolic networks in a marine diatom. *Nature Communications*, 10:4552.

Solomon, C. M., Collier, J. L., Berg, G. M., and Glibert, P. M. (2010). Role of urea in microbial metabolism in aquatic systems: a biochemical and molecular review. *Aquatic Microbial Ecology*, 59:67–88.

Sprecher, B. N., Buck, J. M., Ropella, L. L., Ramsperger, A., Kroth, P. G., and Yamada, N. (2023). Genetic transformation methods for diatom *Nitzschia captiva*: New tools to better understand dinotom endosymbiosis. *Algal Research*, 72:103136.

Sreenikethanam, A., Raj, S., Gugulothu, P., and Bajhaiya, A. K. (2022). Genetic engineering of microalgae for secondary metabolite production: Recent developments, challenges, and future prospects. *Frontiers in Bioengineering and Biotechnology*, 10:836056.

Stinson, B. M. and Loparo, J. J. (2021). Repair of DNA double-strand breaks by the nonhomologous end joining pathway. *Annual Review of Biochemistry*, 90:137–164.

Strauss, J., Deng, L., Gao, S., Toseland, A., Bachy, C., Zhang, C., Kirkham, A., Hopes, A., Utting, R., Joest, E. F., Tagliabue, A., Löw, C., Worden, A. Z., Nagel, G., and Mock, T. (2023). Plastid-localized xanthorhodopsin increases diatom biomass and ecosystem productivity in iron-limited surface oceans. *Nature Microbiology*, 8:2050–2066.

Sun, F., Wang, C., Wang, Y., Tu, K., Zheng, Z., and Lin, X. (2020). Diatom red tide significantly drive the changes of microbiome in mariculture ecosystem. *Aquaculture*, 520:734742.

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., d'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B. T., Royo-Llonch, M., Sarmento, H., Vieira-Silva, S., Dimier, C.,

Picheral, M., Searson, S., Kandels-Lewis, S., Bowler, C., de Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M. B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S. G., and Bork, P. (2015). Structure and function of the global ocean microbiome. *Science*, 348(6237):1261359.

Sung, Y. C. and Fuchs, J. A. (1988). Characterization of the cyn operon in *Escherichia coli* k12. *Journal of Biological Chemistry*, 263(29):14769–14775.

Sung, Y. C. and Fuchs, J. A. (1992). The *Escherichia coli* K-12 cyn operon is positively regulated by a member of the LysR family. *Journal of Bacteriology*, 174(11):3645–3650.

Sutton, R. and Sposito, G. (2005). Molecular structure in soil humic substances: The new view. *Environmental Science & Technology*, 39(23):9009–9015.

Szymczak, A. L. and Vignali, D. A. A. (2005). Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opinion on Biological Therapy*, 5(5):627–638.

Szymczak, A. L., Workman, C. J., Wang, Y., Vignali, K. M., Dilioglou, S., Vanin, E. F., and Vignali, D. A. A. (2004). Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide–based retroviral vector. *Nature Biotechnology*, 22:589–594.

Tamura, K., Adachi, Y., Chiba, K., Oguchi, K., and Takahashi, H. (2002). Identification of Ku70 and Ku80 homologues in *Arabidopsis thaliana*: evidence for a role in the repair of DNA double-strand breaks. *The Plant Journal*, 29(6):771–781.

Tanaka, N., Mogi, Y., Fujiwara, T., Yabe, K., Toyama, Y., Higashiyama, T., and Yoshida, Y. (2021). CZON-cutter - a CRISPR-Cas9 system for multiplexed organelle imaging in a simple unicellular alga. *Journal of Cell Science*, 134(21):jcs258948.

Thomy, J., Sanchez, F., Gut, M., Cruz, F., Alioto, T., Piganeau, G., Grimsley, N., and Yau, S. (2021). Combining nanopore and illumina sequencing permits detailed analysis of insertion mutations and structural variations produced by PEG-mediated transformation in *Ostreococcus tauri*. *Cells*, 10(3):664.

Tran, Q. G., Le, T. T., Choi, D. Y., Cho, D. H., Yun, J. H., Choi, H. I., Kim, H. S., and Lee, Y. J. (2025). Progress and challenges in CRISPR/Cas applications in microalgae. *Journal of Microbiology*, 63(3):e2501028.

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., Bittner, L., Dugdale, R., Finkel, Z., Iudicone, D., Jahn, O., Guidi, L., Lasbleiz, M., Leblanc, K., Levy, M., and Pondaven, P. (2018). Influence of diatom diversity on the ocean biological carbon pump. *Nature Geoscience*, 11:27–37.

Tréguer, P. J., Sutton, J. N., Brzezinski, M., Charette, M. A., Devries, T., Dutkiewicz, S., Ehlert, C., Hawkings, J., Leynaert, A., Liu, S. M., Llopis Monferrer, N., López-Acosta, M., Maldonado, M., Rahman, S., Ran, L., and Rouxel, O. (2021). Reviews and syntheses: The biogeochemical cycle of silicon in the modern ocean. *Biogeosciences*, 18(4):1269–1289.

Trentacoste, M., Shrestha, R. P., Smith, S. R., Glé, C., Hartmann, A. C., Hildebrand, M., and Gerwick, W. H. (2013). Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth. *Proceedings of the National Academy of Sciences of the United States of America*, 110(49):19748–19753.

Tschitschko, B., Esti, M., Philippi, M., Landa, M., Wünderlich, D., Richter, K., Schwaderer, A. S., Huisman, J., Kuypers, M. M. M., and Lavik, G. (2024). Rhizobia–diatom symbiosis fixes missing nitrogen in the ocean. *Nature*, 630:899–904.

Turk-Kubo, K. A., Gradoville, M. R., Cheung, S., Cornejo-Castillo, F. M., Harding, K. J., Morando, M., Mills, M., and Zehr, J. P. (2023). Non-cyanobacterial diazotrophs: global diversity, distribution, ecophysiology, and activity in marine waters. *FEMS Microbiology Reviews*, 47(6):fuac046.

van Baren, M. J., Bachy, C., Reistetter, E. N., Purvine, S. O., Grimwood, J., Sudek, S., Yu, H., Poirier, C., Deerinck, T. J., Kuo, K.-C., Grigoriev, I. V., and Worden, A. Z. (2016). Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genomics*, 17(1):267.

van Ooijen, G., Knox, K., Kis, K., Bouget, F.-Y., and Millar, A. J. (2012). Genomic transformation of the picoeukaryote *Ostreococcus tauri*. *Journal of Visualized Experiments*, (65):e4074.

Vardi, A., Bidle, K. D., Kwityn, C., Hirsh, D. J., Thompson, S. M., Callow, J. A., Falkowski, P. G., and Bowler, C. (2008). A diatom gene regulating nitric-oxide signaling and susceptibility to diatom-derived aldehydes. *Current Biology*, 18(12):895–899.

Walker, E. J. L., Pampuch, M., Deng, L., Li, Y., Tran, G., Mock, T., and Karas, B. J. (2026). Breaking the cell wall for efficient DNA delivery to diatoms. *Nature Communications*, 17:1848.

Wang, H., Xie, X., Gu, W., Zheng, Z., Zhuo, J., Shao, Z., Huan, L., Zhang, B., Niu, J., Gao, S., Wang, X., and Wang, G. (2024a). Gene editing of economic macroalga *Neopyropia yezoensis* (Rhodophyta) will promote its development into a model species of marine algae. *New Phytologist*, 244(5):1687–1691.

Wang, J., Zhou, W., Chen, L., Liu, T., and Lu, X. (2024b). Genetic engineering tools for the filamentous microalga *Tribonema minus*. *Algal Research*, 83:103697.

Wang, L., Sun, Y., Zhang, R., Pan, K., Li, Y., Wang, R., Zhang, L., Zhou, C., Li, J., Li, Y., Zhu, B., and Han, J. (2023a). Enhancement of hemostatic properties of *Cyclotella cryptica* frustule through genetic manipulation. *Biotechnology for Biofuels*, 16:136.

Wang, Q., Lu, Y., Xin, Y., Wei, L., Huang, S., and Xu, J. (2016). Genome editing of model oleaginous microalgae *Nannochloropsis spp.* by CRISPR/Cas9. *The Plant Journal*, 88(6):1071–1081.

Wang, R., Liu, J., Xu, Y., Liu, L., and Mopper, K. (2024c). Unraveling sources of cyanate in the marine environment: Insights from cyanate distributions and production during the photochemical degradation of dissolved organic matter. *Frontiers in Marine Science*, 11:1373643.

Wang, Y., Johnson, G. I., Postles, A., and Coyne, K. J. (2023b). Nitrate reductase enzymes in the alga *Chattonella subsalsa* are regulated by environmental cues at the translational and post-translational levels. *Frontiers in Microbiology*, 14:1059074.

Wang, Y., Wang, F., Wang, R., Zhao, P., and Xia, Q. (2015). 2A self-cleaving peptide-based multi-gene expression system in the silkworm *Bombyx mori*. *Scientific Reports*, 5:16273.

Ward, B. A., Dutkiewicz, S., Moore, C. M., and Follows, M. J. (2013). Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation. *Limnology and Oceanography*, 58(6):2059–2075.

Ward, B. A. and Follows, M. J. (2016). Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux. *Proceedings of the National Academy of Sciences*, 113(11):2958–2963.

Ward, B. B. (2011). Nitrification. In Ward, B. B., Arp, D. J., and Klotz, M. G., editors, *Nitrification*, pages 326–346. ASM Press, Washington, DC, USA.

Waser, N. A. D., Harrison, P. J., Nielsen, B., Calvert, S. E., and Turpin, D. H. (1998). Nitrogen isotope fractionation during the uptake and assimilation of nitrate, nitrite, ammonium, and urea by a marine diatom. *Limnology and Oceanography*, 43(2):215–224.

Weaver, J. C. and Chizmadzhev, Y. (1996). Theory of electroporation: A review. *Bioelectrochemistry and Bioenergetics*, 41(2):135–160.

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLOS ONE*, 6(2):e16765.

Wei, L., Derrien, B., Gautier, A., Houille-Vernes, L., Boulouis, A., Saint-Marcoux, D., Malnoë, A., Rappaport, F., de Vitry, C., Vallon, O., Choquet, Y., and Wollman, F. A. (2014). Nitric oxide-triggered remodeling of chloroplast bioenergetics and thylakoid proteins upon nitrogen starvation in *Chlamydomonas reinhardtii*. *The Plant Cell*, 26(1):353–372.

Weyman, P. D., Beeri, K., Lefebvre, S. C., Rivera, J., McCarthy, J. K., Heuberger, A. L., Peers, G., Allen, A. E., and Dupont, C. L. (2015). Inactivation of *Phaeodactylum tricornutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant Biotechnology Journal*, 13:460–470.

Wheeler, P. A., North, B. B., and Stephens, G. C. (1974). Amino acid uptake by marine phytoplankters. *Limnology and Oceanography*, 19(2):249–259.

Widner, B., Mordy, C. W., and Mulholland, M. R. (2018). Cyanate distribution and uptake above and within the eastern tropical South Pacific oxygen deficient zone. *Limnology and Oceanography*, 63(1):177–192.

Widner, B. and Mulholland, M. R. (2017). Cyanate distribution and uptake in North Atlantic coastal waters. *Limnology and Oceanography*, 62(6):2538–2554.

Widner, B., Mulholland, M. R., and Mopper, K. (2013). Chromatographic determination of nanomolar cyanate concentrations in estuarine and sea waters by precolumn fluorescence derivatization. *Analytical Chemistry*, 85(14):6661–6666.

Widner, B., Mulholland, M. R., and Mopper, K. (2016). Distribution, sources, and sinks of cyanate in the coastal North Atlantic Ocean. *Environmental Science & Technology Letters*, 3(8):297–302.

Worden, A. Z., Lee, J.-H., Mock, T., Rouzé, P., Simmons, M. P., Aerts, J., Allen, A. E., Cuvelier, M. L., Derelle, E., Everett, M. V., Foulon, E., Grimwood, J., Gundlach, H., Henrissat, B., Napoli, C., McDonald, S. M., Parker, M. S., Rombauts, S., Salamov, A., Vandepoele, K., Xiong, Y., Zhu, Y., Henrissat, B., Van de Peer, Y., Moreau, H., and Grigoriev, I. V. (2009). Green evolution

and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*, 324(5924):268–272.

Wurch, L. L., Haley, S. T., Orchard, E. D., Gobler, C. J., and Dyhrman, S. T. (2011). Nutrient-regulated transcriptional responses in the brown tide-forming alga *Aureococcus anophagefferens*. *Environmental Microbiology*, 13:468–481.

Yang, B., Liu, J., Liu, B., Sun, P., Ma, X., Jiang, Y., Wei, D., and Chen, F. (2015). Development of a stable genetic system for *Chlorella vulgaris*—a promising green alga for CO_2 biomitigation. *Algal Research*, 12:134–141.

Yasui, R., Suzuki, S., Fujii, T., Yabu, Y., Maeda, Y., Murata, S., Kataoka, K., and Tanaka, T. (2026). Highly efficient genome editing using CRISPR/Cas9 ribonucleoprotein in the marine oleaginous diatom *fistulifera solaris*. *Scientific Reports*.

Yew, G. Y., Lee, S. Y., Show, P. L., Tao, Y., Law, C. L., Nguyen, T. T. C., and Chang, J. S. (2019). Recent advances in algae biodiesel production: from upstream cultivation to downstream processing. *Bioresource Technology Reports*, 7:100227.

Yin, K., Han, T., Liu, G., and et al. (2015). A geminivirus-based guide RNA delivery system for CRISPR/Cas9 mediated plant genome editing. *Scientific Reports*, 5:14926.

Yin, W. and Hu, H. (2021). High-efficiency transformation of a centric diatom *Chaetoceros muelleri* by electroporation with a variety of selectable markers. *Algal Research*, 55:102274.

Yin, W. and Hu, H. (2023). CRISPR/Cas9-mediated genome editing via homologous recombination in a centric diatom *Chaetoceros muelleri*. *ACS Synthetic Biology*, 12:1287–1296.

Yin, W., Zhang, Y., Huang, Q., Pan, Y., and Hu, H. (2024). Genetic transformation of the freshwater diatom *Cyclotella meneghiniana* via bacterial conjugation. *Algal Research*, 84:103794.

Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular Biology and Evolution*, 21(5):809–818.

Yoshimitsu, Y., Abe, J., and Harayama, S. (2018). Cas9-guide RNA ribonucleoprotein-induced genome editing in the industrial green alga *coccomyxa* sp. strain kj. *Biotechnology for Biofuels*, 11:326.

Yu, E. T., Zendejas, F. J., Lane, P. D., Gaucher, S., Simmons, B. A., and Lane, T. W. (2009). Triacylglycerol accumulation and profiling in the model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* (Bacillariophyceae) during starvation. *Journal of Applied Phycology*, 21(6):669–681.

Yu, X.-Z., Lei, S.-Y., Lin, Y.-J., Zhang, X.-H., and Zhao, F.-J. (2019). Interaction of cyanate uptake by rice seedlings with nitrate assimilation: gene expression analysis. *Environmental Science and Pollution Research*, 26:20208–20218.

Yung, C. C. M., Rey Redondo, E., Sanchez, F., Yau, S., and Piganeau, G. (2022). Diversity and evolution of mamiellophyceae: Early-diverging phytoplanktonic green algae containing many cosmopolitan species. *Journal of Marine Science and Engineering*, 10(2):240.

Zainal Abidin, A. A., Suntarajh, M., and Balia Yusof, Z. N. (2020). Transformation of a Malaysian species of *Nannochloropsis*: gateway to construction of transgenic microalgae as vaccine delivery system to aquatic organisms. *Bioengineered*, 11(1):1071–1079.

Zarlenga, D. S., Mitreva, M., Thompson, P., Tyagi, R., Tuo, W., and Hoberg, E. P. (2019). A tale of three kingdoms: Members of the phylum nematoda independently acquired the detoxifying enzyme cyanase through horizontal gene transfer from plants and bacteria. *Parasitology*, 146(4):445–452.

Zehr, J. P. and Capone, D. G. (2020). Changing perspectives in marine nitrogen fixation. *Science*, 368(6492):eaay9514.

Zehr, J. P. and Kudela, R. M. (2011). Nitrogen cycle of the open ocean: From genes to ecosystems. *Annual Review of Marine Science*, 3:197–225.

Zehr, J. P., Waterbury, J. B., Turner, P. J., Montoya, J. P., Omoregie, E., Steward, G. F., Hansen, A., and Karl, D. M. (2001). Unicellular cyanobacteria fix N_2 in the subtropical north Pacific Ocean. *Nature*, 412:635–638.

Zhang, C. and Hu, H. (2014). High-efficiency nuclear transformation of the diatom *Phaeodactylum tricornutum* by electroporation. *Marine Genomics*, 16:63–66.

Zhang, C., Zong, H., Zhuge, B., Lu, X., Fang, H., Zhu, J., and Zhuge, J. (2016). Protoplast preparation and polyethylene glycol (PEG)-mediated transformation of *Candida glycerinogenes*. *Biotechnology and Bioprocess Engineering*, 21:95–102.

Zhang, H., Zeng, C., Zhang, X., Wang, X., and Li, K. (2025). Differential responses and adaptive mechanisms of diatoms and dinoflagellates to changes in dissolved organic nitrogen components. *Harmful Algae*, 150:103005.

Zhang, X., Ward, B. B., and Sigman, D. M. (2020). Global nitrogen cycle: Critical enzymes, organisms, and processes for nitrogen budgets and dynamics. *Chemical Reviews*, 120(12):5308–5351.

Zhang, Y., Iaffaldano, B., and Qi, Y. (2021). CRISPR ribonucleoprotein-mediated genetic engineering in plants. *Plant Communications*, 2(2):100168.

Zheng, R., Cai, R., Yao, H., Zhuo, X., He, C., Zheng, Q., Shi, Q., and Jiao, N. (2022). Experimental insight into the enigmatic persistence of marine refractory dissolved organic matter. *Environmental Science & Technology*, 56(23):17420–17429.

Zhu, Y., Mulholland, M. R., Macías, T. A., Echevarría, M. A., Pérez, V. E., and Bernhardt, P. (2023). Cyanate dynamics under algal blooms and sediment resuspension events in a shallow micro-tidal estuary in the lower Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, 281:108188.

Zhuang, Y., Zhang, H., Hannick, L., and Lin, S. (2015). Metatranscriptome profiling reveals versatile n-nutrient utilization, co2 limitation, oxidative stress, and active toxin production in an alexandrium fundyense bloom. *Harmful Algae*, 42:60–70.

Appendix

This appendix includes publications arising from this PhD research. These publications provide additional reviews and methodological developments about the genetic engineering and DNA delivery strategies in marine diatoms.

- Genetic engineering in diatoms: advances and prospects.

- Breaking the cell wall for efficient DNA delivery to diatoms.

Check for updates

the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2025) **121,** e70102 doi: 10.1111/tpj.70102

FOCUSED REVIEW

Genetic engineering in diatoms: advances and prospects

Yixuan Li[1], Longji Deng[1], Emma Jane Lougheed Walker[2], Bogumil J. Karas[2] and Thomas Mock[1,*]

[1]School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK, and
[2]Department of Biochemistry, Schulich School of Medicine and Dentistry, Western University, London, Ontario N6A 5C1, Canada

Received 9 December 2024; revised 1 March 2025; accepted 4 March 2025.
*For correspondence (e-mail t.mock@uea.ac.uk).

SUMMARY

Diatoms are among the most diverse and ecologically significant groups of photosynthetic microalgae, contributing over 20% of global primary productivity. Their ecological significance, unique biology, and genetic tractability make them ideal targets for genetic and genomic engineering and metabolic reprogramming. Over the past few decades, numerous genetic methods have been developed and applied to these organisms to better understand the function of individual genes and how they underpin diatom metabolism. Additionally, the ability of diatoms to synthesize diverse high-value metabolites and elaborate mineral structures offers significant potential for applications in biotechnology, including the synthesis of novel pharmaceuticals, nutraceuticals, and biomaterials. This review discusses the latest developments in diatom genetic engineering and provides prospects not only to promote the use of diatoms in diverse fields of biotechnology but also to deepen our understanding of their role in natural ecosystems.

Keywords: diatoms, genetic engineering, metabolic engineering, synthetic biology.

INTRODUCTION

Diatoms are a diverse class of photosynthetic microalgae widely distributed across oceans, lakes, rivers, and even soils (Pinseel et al., 2020). Representing the most diverse phytoplankton class in today's oceans, diatoms are estimated to include approximately 100 000 species worldwide (Mann & Vanormelingen, 2013). With only about 0.1% of the total global plant biomass, diatoms contribute over 20% of the world's primary productivity (Bar-On et al., 2018; Leblanc et al., 2012; Malviya et al., 2016). This disproportionate impact underscores their critical role, especially in aquatic ecosystems, driving global biogeochemical cycles.

Given their ecological significance and diversity, there is an impetus to sequence and develop molecular tools for a variety of diatom species. The current availability of over 50 diatom genomes with >100 still to be sequenced as part of the 100 Diatom Genomes Project (100DGP) in addition to more than 500 diatom transcriptomes, provides a significant resource for studying diatom biology from genes to applications (Brodie et al., 2017). Based on these resources, researchers have gained important insights into diatom physiology, evolution, and their ecological functions. Since the first successful genetic transformation in *Cyclotella cryptica* and *Navicula saprophila* using biolistics, this method has rapidly become a standard and, therefore, is the most popular technique for the genetic modification of diatoms (Dunahay et al., 1995). Advances in transformation technologies, such as electroporation (Hu & Pan, 2020; Okada et al., 2023; Yin & Hu, 2021; Zhang & Hu, 2014) and conjugation (Karas et al., 2015), combined with the advent of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9) (Hopes et al., 2016; Nymark et al., 2016), have enabled precise and efficient editing of diatom genomes. These advancements have positioned diatoms as an attractive target for genetic engineering and metabolic reprogramming, facilitating applications in biotechnology, such as alternative fuels, pharmaceuticals, nutraceuticals, and the synthesis of novel biomaterials (Brodie et al., 2017). This review aims at providing an overview of the methods and applications involved in the genetic manipulation of diatoms, including the current progress in terms of further developing genomics resources as the foundation of genome engineering. Our review will also provide prospects for this fast-developing field of fundamental and applied research with diatoms.

© 2025 The Author(s).
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DIATOM GENOMIC RESOURCES

Thalassiosira pseudonana and *Phaeodactylum tricornutum* were the first diatoms to have their genome completely sequenced (Armbrust et al., 2004; Bowler et al., 2008). Characterized by small genome size (<35 Mb) and rapid growth (>one cell division per day under optimal conditions), both species quickly became models, providing key insights into diatom evolution, adaptation and ecological roles in their natural environments (Bowler et al., 2010; Mock et al., 2022). However, the availability of only two diatom genomes is insufficient to understand the diversity of diatoms (Nakov et al., 2018). To gain more comprehensive insights on diatom adaptation and evolution, additional diatom species have been subsequently sequenced. For example, the genome of *Fragilariopsis cylindrus* revealed how diatoms adapt to polar oceans (Mock et al., 2017) and the genome of *Nitzschia putrida* illuminated how diatoms evolve from a phototropic to a heterotrophic lifestyle (Kamikawa et al., 2022). Furthermore, the genome of oleaginous diatoms *Fistulifera solaris* and *C. cryptica* provide a foundation for improving lipid synthesis through genetic modifications (Tanaka et al., 2015; Traller et al., 2016), which is relevant for diverse biotechnological applications. However, a step-change in our understanding of diatom functional and evolutionary genomics will be provided by the '100 Diatom Genomes Project' (https://jgi.doe.gov/csp-2021-100-diatom-genomes/), which was initiated a few years ago by an international consortium of diatom researchers from different fields. The project is funded by the Joint Genome Institute (JGI), Department of Energy (USA).

Consequently, these fast-expanding diatom genomic resources will provide a wealth of novel biological information which can be accessed through public data banks such as the National Center for Biotechnology Information (NCBI), the Joint Genome Institute (JGI), the European Nucleotide Archive (ENA), the Ensembl Genome Browser, DiatOmicBase, and PLAZA (Table S1). However, to harness this wealth of information, it will be necessary to develop synthetic biology tools that allow for large-scale genome engineering, analogous to the methods that have been pioneered for prokaryotes (Fredens et al., 2019; Hutchison et al., 2016) and yeast (Schindler et al., 2024). In combination with multi-omics tools, the integration of these diverse resources and methods will provide holistic insights into the function of genes, genetic networks, and their role in diverse biological processes.

GENETIC ENGINEERING

There are two complementary approaches used so far for diatom genetic engineering: Forward and reverse genetics. The former is a phenotype-first approach, which starts with an observable phenotype as the outcome of random mutagenesis. The latter aims to purposefully modify a genetic locus of interest, resulting in novel phenotypes to be characterized to reveal the *in vivo* function of the genetic locus. Reverse genetics is an important component of synthetic biology to create new diatom biology, which in its most developed stage would generate a synthetic diatom based on artificially synthesized chromosomes that carry a novel combination of genes and genetic networks. Although there is no synthetic diatom yet, significant advancements in terms of large-scale chromosome-like assemblies and novel transformation protocols pave the way for achieving that goal.

Forward genetics

Forward genetics generates random mutations by either chemical and/or radiation mutagenesis. DNA mutations can also be caused by random integrations of exogenous DNA, including selective markers (Goold et al., 2024; Moosburner et al., 2022). High-throughput approaches are subsequently used to screen many of the treated cell lines to identify phenotypes that are significantly different from the wild-type (WT) controls. For instance, chemically mediated mutagenesis in diatoms using *N*-ethyl-*N*-nitrosourea (ENU) has identified a new type of uridine-5-monophosphate synthase (UMPS) (Sakaguchi et al., 2011) in *P. tricornutum*. UV-mediated mutagenesis has been used to introduce dominant genetic mutations in diatoms. For example, it has been successfully applied to increase the eicosapentaenoic acid (EPA) content in *P. tricornutum* by up to 44% (Alonso et al., 1996). Additionally, UV mutagenesis was applied to identify novel selectable markers, such as the genetic locus conferring resistance against the herbicide norflurazon, which was caused by a specific amino acid substitution in the phytoene desaturase (PDS) gene (Taparia et al., 2019). Thus, forward genetics is a valuable tool for introducing dominant mutations. However, the diploid nature of diatom genomes during most of the life cycle imposes challenges to obtaining biallelic (homozygous) gene inactivation by untargeted mutagenesis, especially if the mutagen primarily causes single-strand breaks in the DNA (Huang & Daboussi, 2017; Moosburner et al., 2022). Similar challenges exist for the application of reverse genetics approaches.

Reverse genetics

Before any genetic loci can be modified using targeted approaches, there need to be methods for delivering and establishing genetic material in the diatom cell. Usually, these transformation protocols are developed using the expression of marker genes such as green fluorescent protein (GFP) and yellow fluorescent protein (YFP), which, based on their fluorescence, help to assess the stability of the genetic modification. The latter is an important criterion for many applications in reverse genetics with

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

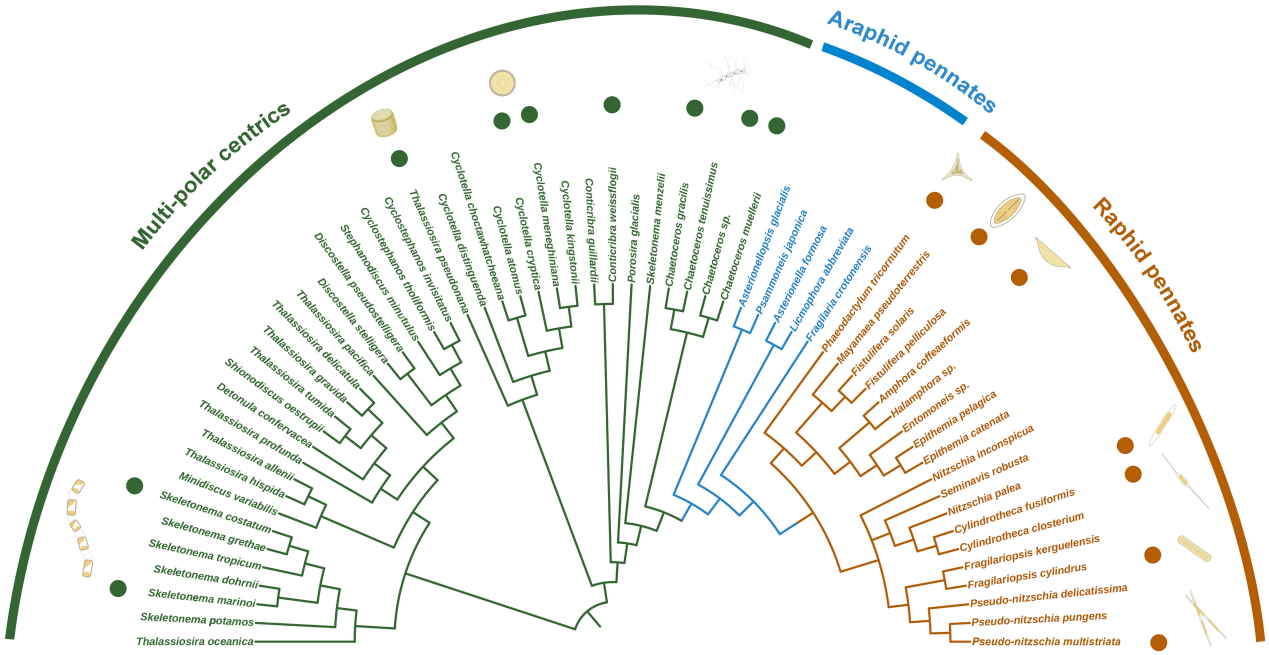


Figure 1. Maximum-likelihood phylogenetic tree and transformation status of sequenced diatom genomes based on the 18S sequences.
The phylogenetic tree depicts the evolutionary relationships among diatom species with complete genome sequences. Species with established genetic transformation protocols are highlighted with dots at the leaves. Representative morphologies of key species are illustrated alongside their respective phylogenetic positions. The tree is divided into three major groups based on the phylogeny: multi-polar centrics (green), araphid pennates (blue), and raphid pennates (brown).

diatoms (Faktorova et al., 2020; Falciatore et al., 1999; Sabatino et al., 2015; Yin & Hu, 2021). Once a transformation protocol has been developed, many endogenous genetic loci can be manipulated. The most used genetic modifications in diatoms include gene overexpression, knockdown, and knockout. Furthermore, some diatoms such as *T. pseudonana* can be used for efficient gene replacement facilitated by homology-directed repair (HDR). Until now, the genetic transformation systems have been established in 20 diatom species (Figure 1).

Gene overexpression, defined as an unnaturally high level of gene expression usually caused by the knock-in of an additional copy of the target gene driven by a strong constitutive promoter such as fucoxanthin chlorophyll-a/c-binding protein (*FCP*), may give cells new phenotypic characteristics, allowing researchers to draw conclusions about the potential function of the overexpressed gene. This method was successfully employed for the first time in diatoms in 1995 (Dunahay et al., 1995).

Gene knockdown is the process by which gene expression is at least reduced and sometimes completely abolished. It is reversible as the method targets mRNA instead of DNA, thereby resulting in the decrease of the encoded protein. The first report of gene knockdown was based on inhibiting the expression of the GUS reporter and the endogenous cryptochrome gene in *P. tricornutum* using antisense RNA technology (De Riso et al., 2009). Since then, gene silencing has been widely applied in diatoms (Sabatino et al., 2022; Shrestha & Hildebrand, 2017; Trentacoste et al., 2013; Wang et al., 2023). However, the molecular mechanism underlying gene silencing remains to be completely resolved because it does not appear to be following the canonical mechanism of RNA interference (RNAi) mediated by small non-coding RNAs (sRNAs) and the activity of the endoribonuclease Dicer (DCR) because the mRNA of the silenced genes is still present, as observed in *T. pseudonana*, for instance, (Kirkham et al., 2017). However, new data based on comparative genome and transcriptome analyses and reverse genetics with P. *tricornutum* have in fact revealed that key RNAi effectors have diversified in diatoms. Nevertheless, this new study suggests the presence of distinct RNAi pathways in diatoms. For instance, the *P. tricornutum* DCR was found to process 26–31-nt-long double-stranded sRNAs originating mostly from transposons covered by repression-associated epigenetic marks (Grypioti et al., 2024).

In addition to gene overexpression and silencing in diatoms, the last decade has revolutionized our ability to modify the diatom genome by enabling targeted modifications using sequence-specific nucleases. Modifications are achieved through the application of meganucleases (MNs), transcriptional activator-like effector nucleases (TALEN) and especially CRISPR/Cas-assisted genome editing. These genetic tools introduce double-strand DNA breaks (DSB) at specific loci, with subsequent repair either via homology-directed repair (HDR) or non-homologous end joining

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(NHEJ). The latter is the most frequent repair mechanism in diatoms acting on single-nucleotide mutations and small insertion–deletions (INDELS). MNs, TALENs, and CRISPR/Cas have been used successfully to induce targeted mutations in diatom genomes (Daboussi et al., 2014; Fortunato et al., 2016; Hopes et al., 2016). Among these three gene modification methods, meganucleases are the oldest method which recognizes and cleaves long DNA target sites (14–40 bp) and tolerates minor sequence changes in these sites (Chevalier & Stoddard, 2001). TALENs consist of a sequence-specific DNA-binding domain and a FokI cleavage domain (Christian et al., 2010). The customized TALE DNA-binding domain guides the nuclease to the target gene, while the FokI nuclease domain dimerizes to induce DNA cleavage (Bogdanove & Voytas, 2011; Voytas, 2013). The first proof of concept for using MNs and TALENs to edit diatom genomes was published by Daboussi et al. (2014). TALEN technology was applied to both *P. tricornutum* and *T. pseudonana* (Fortunato et al., 2016; Weyman et al., 2015). TALENs have the advantage of low off-target activity, ensuring high specificity in genome editing (Nemudryi et al., 2014). However, a typical TALEN system requires pairing two TALENs to target DNA sequences of 30–36 base pairs, with each side consisting of 18 repeats (Malzahn et al., 2017). Thus, the design is complex and not as versatile as the CRISPR/Cas technology. Therefore, MNs and TALENs were gradually replaced in diatom research by the effective CRISPR/Cas technology (e.g., Hopes et al., 2016). The native CRISPR-Cas9 system functions as part of the bacterial immune system by detecting and cutting foreign DNA using sequence-specific guide RNAs (gRNAs). In genome editing, the CRISPR RNA (crRNA) fused to the trans-activating CRISPR RNA (tracrRNA) generates single-guide RNAs (sgRNAs), which replace native gRNAs. The sgRNAs guide the Cas9 protein to cleave the DNA at the desired loci (Bortesi & Fischer, 2015). The first reports for successful CRISPR/Cas-based genome editing leading to stable gene knockouts were based on work with *P. tricornutum* (Nymark et al., 2016) and *T. pseudonana* (Hopes et al., 2016). Both papers came out in the same year, separated only by some weeks. Following these landmark papers, CRISPR/Cas-mediated homologous recombination (HR) was developed in *P. tricornutum*, *T. pseudonana*, and *Chaetoceros muelleri*, providing targeted genome editing at endogenous loci (Belshaw et al., 2022; Moosburner et al., 2020; Yin & Hu, 2023). In 2018, the first paper on DNA-free gene knockout by direct delivery of CRISPR-Cas9 ribonucleoproteins (RNPs) was published. This DNA-free method is still successfully used in the diatom community (Serif et al., 2018). Furthermore, a Cas9 nickase (nCas9) has recently been developed for diatom research. The native Cas9 nuclease has two active domains, that is, RuvC (D10A) and HNH (H840A). The RuvC domain is inactive in the Cas9 nickase (nCas9). Thus, nCas9 only introduces single-strand DNA breaks (nicks). The nCas9 system has recently been used in diatoms to suppress the off-target effects of the native Cas9 nuclease (Matsui et al., 2024; Nawaly et al., 2020; Nigishi et al., 2024).

Very recently, a dead (d)Cas9 has been used in diatom research for the first time (Guo et al., 2024). The dCas9 is characterized by its inability to cut DNA. However, it still retains the ability to bind to DNA (Pickar-Oliver & Gersbach, 2019). dCas9 in diatoms has been used for CRISPR interference (CRISPRi) to perform CRISPR-mediated knockdown in *P. tricornutum* (Guo et al., 2024). dCas9 can also be used for CRISPR activation, in which the dCas9 is fused with transcriptional activators, allowing the upregulation of target genes. However, this method (CRISPRa) has yet to be established for diatom research. Base editors (e.g., cytosine, adenosine) have not yet found their way into diatom research either (Lee et al., 2023). Base editors enable precise editing of single bases, which can be used to introduce single amino acid changes to confer, for instance, resistance to compounds such as cycloheximide and norflurazon (Stevens et al., 2001; Taparia et al., 2019).

MOLECULAR GENETIC TOOLKITS

Promoters

Promoters can be endogenous or heterologous in origin and can drive constitutive or inducible expression of the respective downstream open reading frame. Commonly used promoters for diatom research are listed in Table 1. Constitutive promoters facilitate stable gene expression and are not influenced by growth stage or environmental conditions. Examples of endogenous constitutive promoters commonly used in diatoms include those from genes encoding fucoxanthin chlorophyll-a/c-binding protein (*Fcp*, now called *Lhcf*), histone 4 (*H4*), elongation factor II (*ef2*), acetyl-CoA carboxylase (ACCase) and U6 small nuclear RNA (U6) (Apt et al., 1996; De Riso et al., 2009; Dunahay et al., 1995; Falciatore et al., 1999; Hopes et al., 2016; Nymark et al., 2016; Sabatino et al., 2015; Seo et al., 2015; Siaut et al., 2007). Recently, *in vivo* transcriptional activity of four new endogenous promoters from the NADH:ubiquinone oxidoreductase (*Nub*), Synaptobrevin/VAMP-like protein (*SVP*), predicted protein 45,582, nd Prohobitin (*Pbt*) was characterized in *P. tricornutum* along with additional alternative promoters (e.g., calmodulin-dependent protein kinase II, oxygen-evolving enhancer protein 3, glucose-6-phosphate isomerase, fructose bisphosphate aldolase), providing comparative data for advancing genetic engineering strategies (Garza et al., 2023; Windhagauer et al., 2021). However, it should be mentioned that not all endogenous constitutive promoters are equally strong. For example, *H4* appears to be

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Toolbox for diatom transformation

	Species	Delivery system	Selective marker	Reporter	Promoter	Approach	Cloning and assembly strategy	Reference
Raphic pennate	*Phaeodactylum tricornutum*	Biolistics Electroporation Conjugation PEG	*Sh ble, nat, bsr, sat, nptII, cat*	eGFP, LUC, GFP, GUS, YFP, CFP, VENUS, mCherry	*FCP, H4, EF2, U6, Nub, SVP, 45582, Pbt, Act2, GLNA, calm, oee3, flav, NR, lsi1, FBP1, ca1, AP1, CMV, CIP, CaMV, RSV-LTR, DIG/pUAS, XVE/OlexA*	Overexpression, TALENs, MNs Gene silencing, CRISPR	Restriction enzyme Golden Gate Gibson uLoop *In vivo*	Cochrane, Brumwell, Soltysiak, et al. (2020); Diamond et al. (2023); Falciatore et al. (2020); Garza et al. (2023); Kadono et al. (2022); Karas et al. (2015); Kassaw et al. (2022); Walker et al. (2024); Windhagauer et al. (2021)
	Fragilariopsis cylindrus	Biolistics	*Sh ble*	eGFP	*FCP*	Overexpression	Gibson	Faktorova et al. (2020)
	Pseudo-nitzschia multistriata	Biolistics Conjugation	*Sh ble, nat*	GFP	*H4*	Overexpression	Restriction enzyme Gibson	Russo et al. (2018); Sabatino et al. (2015)
	Pseudo-nitzschia arenysensis	Biolistics	*Sh ble*	GFP, GUS	*H4*	Overexpression Gene silencing	Restriction enzyme	Sabatino et al. (2015, 2022)
	Nitzschia captiva	Biolistics	*bsr, Sh ble*	eGFP	*FCP, CfNR*	Overexpression CRISPR (failed)	Restriction enzyme	Sprecher et al. (2023)
	Nitzschia palea	Electroporation	*nat*	GFP	*FCP*	Overexpression	Restriction enzyme	Okada et al. (2023)
	Nitzschia putrida	Biolistics	*nat*	eGFP	*NADH*	Overexpression	Golden Gate	Unpublished
	Fistulifera solaris (*Navicula* sp.)	Biolistics Electroporation	*nptII, sh Ble*	GFP, eGFP	*FCP, H4, PtFCP, RSV, CaMV*	Overexpression Gene silencing	Restriction enzyme	Maeda et al. (2021); Muto et al. (2013); Naser et al. (2022)
	Fistulifera saprophila (*Navicula saprophila*)	Biolistics	*nptII*		*AccI*	Overexpression	Restriction enzyme	Dunahay et al. (1995)
	Cylindrotheca fusiformis	Biolistics	*Sh ble*	eGFP	*FCP, Pδ, CfNR*	Overexpression	Restriction enzyme	Fischer et al. (1999); Poulsen and Kroger (2005)
	Amphora coffeaeformis	Biolistics	*nat*	eYFP	*pPhat1*	Overexpression	Restriction enzyme	Buhmann et al. (2014)

(continued)

© 2025 The Author(s). *The Plant Journal* published by Society for Experimental Biology and John Wiley & Sons Ltd.,
The Plant Journal, (2025), **121**, e70102

13653313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1. (continued)

	Species	Delivery system	Selective marker	Reporter	Promoter	Approach	Cloning and assembly strategy	Reference
Multipolar centrics	*Thalassiosira pseudonana*	Biolistics Electroporation Conjugation	*Sh ble*, *nat*	GFP, mEGFP YFP, mNeonGreen, mScarlet-i, mTurquoise2	*FCP*, *U6*, *NR*, *SIT*, *Thaps3_9619*, *BST2*	Overexpression CRISPR Gene silencing	Restriction enzyme Golden Gate Gibson uLoop *In vivo*	Cochrane, Brumwell, Shrestha, et al. (2020); Falciatore et al. (2020); Hopes et al. (2016); Karas et al. (2015); Nam et al. (2022)
	Thalassiosira weissflogii	Biolistics		GUS	*FCP*	Overexpression	Restriction enzyme	Falciatore et al. (1999)
	Skeletonema marinoi	Electroporation	*bleo*		*Isu4e*, *PtFCP*	Overexpression	Restriction enzyme	Johansson et al. (2019)
	Skeletonema costatum	Electroporation	*bleo*		*FCP*	Gene silencing	Restriction enzyme	Zhen et al. (2024)
	Cyclotella cryptica	Biolistics	*nptII*, *bleo*		*AccI*, *SIT*	Overexpression Gene silencing	Restriction enzyme	Dunahay et al. (1995); Wang et al. (2023)
	Cyclotella meneghiniana	Conjugation	*bsr*	GUS, eGFP	*FCP*, *RL14*	Overexpression	Restriction enzyme	Yin et al. (2024)
	Chaetoceros sp. *CCK09*	Biolistics	*nat*	GFP	*FCP*	Overexpression	Restriction enzyme	Miyagawa-Yamaguchi et al. (2011)
	Chaetoceros muelleri	Electroporation	*ble*, *nat*, *bsr*	GFP, GUS	*FCP*, *ACAT*, *U6*	Overexpression CRISPR	Restriction enzyme	Yin and Hu (2021, 2023)
	Chaetoceros gracilis	Electroporation	*nat*	LUC, mAG	*FCP*, *NR*	Overexpression	Restriction enzyme	Ifuku et al. (2015)

ACAT, acetyl-CoA acetyltransferase; *AccI*, acetyl-CoA carboxylase; *Act2*, purine permease, actin /actin-like protein; *AP1*, alkaline phosphatase; *bsr*, blasticidin-S resistance gene; BST2, bestrophin-like protein (THAPSDRAFT_4820); *Ca1*, carbonic anhydrase; *calm*, calmodulin-dependent protein kinase II; *CaMV*, *Cauliflower mosaic* virus 35S; *cat*, chloramphenicol acetyl transferase conferring resistance to chloramphenicol; *CfNR*, *Cylindrotheca fusiformis* nitrate reductase; *CIP*, *Chaetoceros lorenzianus*-infecting DNA virus; *CMV*, cytomegalovirus; CRISPR, clustered regularly interspaced short palindromic repeats; DIG/pUAS, digoxin-inducible system; *Ef2*, elongation factor 2; eGFP, enhanced green fluorescent protein gene; *Fbp1*, ferrichrome-binding protein 1; *FCP*, fucoxanthin-chlorophyll protein; *Flav*, flavodoxin; GFP, green fluorescent protein gene; *GLNA*, glutamine synthetase; GUS, beta-glucuronidase; *H4*, histone H4; *Isi1*, iron-starvation-induced gene 1; *Isu4e*, 80S ribosomal subunit 4e; LUC, Luciferase; mAG, green fluorescence of the monomeric Azami–Green protein; mEGFP, monomeric enhanced GFP; MN, meganucleases; *NADH*, NADH–ubiquinone reductase complex 1 MLRQ subunit; *nat*, nourseothricin acetyl transferase; *nptII*, neomycin phosphotransferase II (resistance to G418/geneticin); *NR*, nitrate reductase; *Nub*, NADH:ubiquinone oxidoreductase; *oee3*, oxygen-evolving enhancer protein 3; *Pbt*, prohibitin; PEG, polyethylene glycol (PEG)-mediated transformation; *Phat1/ PtFCP*, *P. tricornutum* FCP promoter; *Pδ*, promoter of frua3 gene; *RL14*, ribosomal protein L14; *RSV*, *Rous sarcoma* virus long terminal repeat promoter; *RSV-LTR*, *Rous sarcoma virus* long terminal repeat; *sat*, streptothricin acetyl transferase; *sh ble*, *Streptoalloteichus hindustanus* bleomycin resistance gene; *SIT1*, silicon transporter; *SVP*, Synaptobrevin/VAMP-like protein; TALEN, transcription activator-like effector nucleases; uLoop, universal Loop; *U6*, small nuclear RNA of the U6 complex; VENUS, improved version of yellow fluorescent protein; XVE/OlexA, β-estradiol-inducible system; YFP, yellow fluorescent protein gene; *45582*, *predicted*rotein (Phatr3_J45582).

© 2025 The Author(s).

13653313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a strong promoter, whereas *FCP* is considered weaker in *P. tricornutum* (Garza et al., 2023).

Inducible promoters can switch gene expression on and off in response to specific environmental or experimental conditions. This allows for the controlled expression of a target gene and minimizes the potential negative effects from the overproduction of recombinant proteins that could otherwise occur during constitutive expression. The availability of nutrients can lead to significant changes in the expression of certain diatom genes. Thus, promoters that respond to nutrients, such as those involved in nitrogen, iron, silicon, and phosphate acquisition and metabolism have been used as inducible promoters. For instance, the nitrate reductase (*NR*) promoter is commonly used for inducible expression in diatoms, whereby *NR* is activated by the presence of nitrate as the sole nitrogen source and repressed if replaced by ammonium (Ifuku et al., 2015; Poulsen et al., 2006; Poulsen & Kroger, 2005; Schellenberger Costa et al., 2013). However, even when *NR* is induced, the overall transgene expression levels remain relatively low (Chu et al., 2016; Poulsen & Kroger, 2005). Three iron-responsive promoters were functionally characterized in *P. tricornutum*, including promoters of the iron-starvation-induced protein1 (*Isi1*), ferrichrome-binding protein1 (*FBP1*) and flavodoxin (*Fld*) genes (Yoshinaga et al., 2014). Promoters of a putative silicon-related protein (Thaps3_9619) and silicon transporters (SIT) in *T. pseudonana* (*TpSIT1*, *TpSIT2*) and *C. cryptica* (*CcSIT1*) are potential alternatives because they likely have less detrimental effects on core metabolic functions compared with the *NR* promoter (Davis et al., 2017; Shrestha & Hildebrand, 2017). Likewise, the alkaline phosphatase (AP) gene is strongly induced in response to inorganic phosphate depletion (Lin et al., 2013). Testing its promoter in *P. tricornutum* (*pPhAP1*) revealed a much higher expression of transgenes compared with using either the *Fcp* or any tested *NR* promoters (Lin et al., 2017). The promoter regulating the expression of the secreted protein 1 (HASP1) gene is also responsive to phosphate depletion but repressed if phosphate is available (Slattery et al., 2022). Furthermore, there are CO_2-sensitive promoters available for diatom research, such as the promoter regulating the expression of the β-carbonic anhydrase 1 gene (ptca1) in *P. tricornutum*, which is significantly decreased when cells are grown in a high CO_2 environment (i.e., 5% CO_2, as opposed to air-levels of 0.04%) (Harada et al., 2005). However, it is worth mentioning that the induction and/or inactivation of certain promoters, such as the *pPhAP1*, relies on the availability of essential nutrients, which might cause transcriptional reprogramming as a response (Garza et al., 2023). Hence, the altered transcriptome might interfere with subsequent phenotyping studies to characterize the *in vivo* function of the modified target gene. To reveal if and how those associated effects imposed by the regulation of these promoters impact the physiology of the genetically modified cell lines, we suggest performing multi-omics studies to capture potential reprogramming on diverse levels of organization from transcriptomes to metabolomes.

Heterologous promoters include non-endogenous promoters obtained from other diatom species, non-diatom species, viruses, and even synthetic chemical promoters. For example, the *Lhcf2* promoter from *Cylindrotheca fusiformis* was used to drive transgene expression in *P. tricornutum* and *Fistulifera* sp. (Miyagawa et al., 2009; Muto et al., 2013). With the molecular characterization of diatom-infecting viruses (DIVs), their promoters have been studied in diverse diatoms, and some of them have shown high transgene expression (Kadono et al., 2015; Kadono et al., 2020; Kadono et al., 2022). Furthermore, promoters from the mammalian cytomegalovirus (CMV), the avian rous sarcoma virus (RSV-LTR) and the plant cauliflower mosaic virus 35S (CaMV 35S) have been successfully used for transgene expression in *P. tricornutum* (Sakaue et al., 2008). The CaMV 35S and RSV-LTR promoters can drive high levels of gene expression comparable to the endogenous *FcpA* promoter in the log phase of *P. tricornutum* (Kadono et al., 2020).

Furthermore, chemically inducible gene expression systems have been developed to achieve dynamic gene expression as a consequence of applying exogenous chemicals to diatom cultures. Six of these systems were tested in *P. tricornutum* (Kassaw et al., 2022). Among them, β-estradiol and digoxin demonstrated high levels of reversibility and tunability, making them effective for the chemical induction of transgene expression. The β-estradiol system is highly sensitive, with β-estradiol concentrations as low as 1 nM sufficient to trigger gene expression. The digoxin-inducible promoter enables tight control of gene expression, with activation levels directly correlating with digoxin concentrations ranging from 0 to 100 μM. Chemically inducible promoters have been used in several model organisms (e.g., *Escherichia coli* and yeast). In some cases, the use of endogenous inducible promoters carries the risk of causing unintended effects on metabolism, while constitutive promoters drive continuous expression of the target genes even when their activity is not required anymore, which can be toxic to the cell (e.g., overexpression of the Cas9 enzyme). Chemically inducible promoters overcome these limitations; however, the impact of the chemicals on the metabolism of diatoms largely is unknown and therefore needs to be empirically tested.

Selective markers

Selective markers confer resistance to specific antibiotics/chemicals or restore auxotrophies and are often used when genetically engineering diatoms. A transgenic vector typically contains the gene(s) of interest and a selective or auxotrophic marker, the latter enabling the identification of

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

transformants when cells are exposed to the selective agent. The selective marker genes most commonly used in diatoms confer resistance to various antibiotics, including *Shble* (phleomycin/zeocin resistance), *Nat* (nourseothricin resistance), *Bsr* (blasticidin-S resistance), *Cat* (chloramphenicol resistance), *Sat* (streptothricin resistance) and *nptII* (G418/geneticin resistance) (Apt et al., 1996; Buck et al., 2018; Dunahay et al., 1995; Falciatore et al., 1999; Karas et al., 2015; Muto et al., 2013; Poulsen et al., 2006; Xie et al., 2014; Zaslavskaia et al., 2000). In addition, the mutation of endogenous genes can confer resistance to specific chemical compounds, which have been developed as selectable markers in diatoms. For instance, the inactivation of uridine-5′-monophosphate synthase (PtUMPS) and adenine phosphoribosyl transferase (PtAPT) genes in *P. tricornutum*, leading to the resistance to 5-fluoroorotic acid (5-FOA) and 2-fluoroadenine (2-FA), respectively, provides a selection strategy without relying on exogenous antibiotic resistance markers (Serif et al., 2018). Similarly, the point mutation of the endogenous phytoene desaturase gene (PDS, PHATRDRAFT_45735) conferred resistance to the herbicide norflurazon in *P. tricornutum* (Taparia et al., 2019). However, the potential effects on metabolism, following the removal of selecting agents due to the loss of endogenous genes, should be taken into consideration. Moreover, auxotrophic complementation markers provide an alternative approach. Cas9-directed genome engineering has also been used to generate strains of *P. tricornutum* that are auxotrophic for histidine, uracil, or tryptophan; these auxotrophies can be restored by providing the respective biosynthetic gene in the transgenic vector (Slattery et al., 2020).

When considering microbial interactions involving diatoms, applying reverse genetics tools to study them is still in its infancy. Most work is still focused on managing the contamination of diatom cultures – often, the selective markers introduced during genetic engineering efforts can also be used to mitigate contamination, so long as the contaminant is susceptible to the respective antibiotic or chemical selective agent. For instance, resistance to antifungal drugs such as amphotericin, echinocandins, and azoles has been widely studied to identify key resistance genes (Jensen et al., 2015; Moirangthem et al., 2021; Morio et al., 2017; Rybak et al., 2022; Spettel et al., 2019). This work, if done with diatoms, will provide basic information for selecting antifungal markers, for instance. Furthermore, studying interactions between diatoms, bacteria, and viruses involves the depletion and rescue of bacteria and their associated viruses (Zhang et al., 2024). Thus, introducing resistance genes to diatom genomes represents an approach for controlling these interactions and therefore provides an opportunity to selectively modulate microbial communities for advancing our understanding of microbial interactions with diatoms as the host organisms.

Reporter systems

Reporters are widely used to monitor and track engineered diatom cells by visualizing the expression of transgenes or the regulation of promoters. The integration of reporter genes allows researchers to quickly identify transformed cells and employ techniques such as fluorescence-activated cell sorting (FACS) to isolate cells emitting fluorescence. Enzyme-based reporters offer high sensitivity but require the addition of a substrate to generate a detectable signal (e.g., bioluminescence). For instance, luciferase (LUC) and β-glucuronidase (GUS) genes have been expressed in several diatom species (Falciatore et al., 1999; Ifuku et al., 2015; Sabatino et al., 2015; Zaslavskaia et al., 2000). However, some reporter systems require the application of complex assays for visualization, which makes these systems more technically challenging to work with (Huttly, 2009).

For real-time studies and live-cell imaging, fluorescent reporters may be better alternatives. For instance, the breakthrough of tagging fluorescent proteins revolutionized the study of protein–protein interactions in living cells, leading to the discovery of interaction networks and previously unknown protein functions (Nam et al., 2024; Turnsek et al., 2021). These proteins emit stable fluorescence upon excitation by the corresponding wavelengths of light. Because of their broad applicability (e.g., operational in different cell types, easy detectability), fluorescent proteins are perhaps the most frequently used reporter systems in diatoms (Figure 2). Examples are the green fluorescent protein (GFP), enhanced green fluorescent protein (eGFP), yellow fluorescent protein (YFP), optimized YFP (VENUS), the red fluorescent protein mCherry, and the cyan fluorescent protein gene (CFP) (Table 1). New fluorophores, including mNeonGreen, mScarlet-i, and mTurquoise2, have been tested and validated in *T. pseudonana*, expanding the toolkit for protein localization studies in diatoms (Nam et al., 2022). Additionally, fluorescent proteins that have been established in other microalgae and diverse protists include the brightest blue fluorescent protein mTagBFP. It is believed that these fluorescent proteins should also work in diatoms (Faktorova et al., 2020). When selecting reporter genes, it is critical to consider the background conditions of the cells (e.g., presence of chlorophyll autofluorescence) to ensure the accurate detection of the reporters and to minimize interference with background fluorescence.

Various reporters offer a wide spectrum of colors, enabling researchers to select the most suitable reporter based on their objectives, including multicolor imaging to study multiple reporters simultaneously. If reporters are fused to target proteins or promoter sequences, it allows for the identification of the subcellular localization of the target genes and the regulation of the promoters under

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

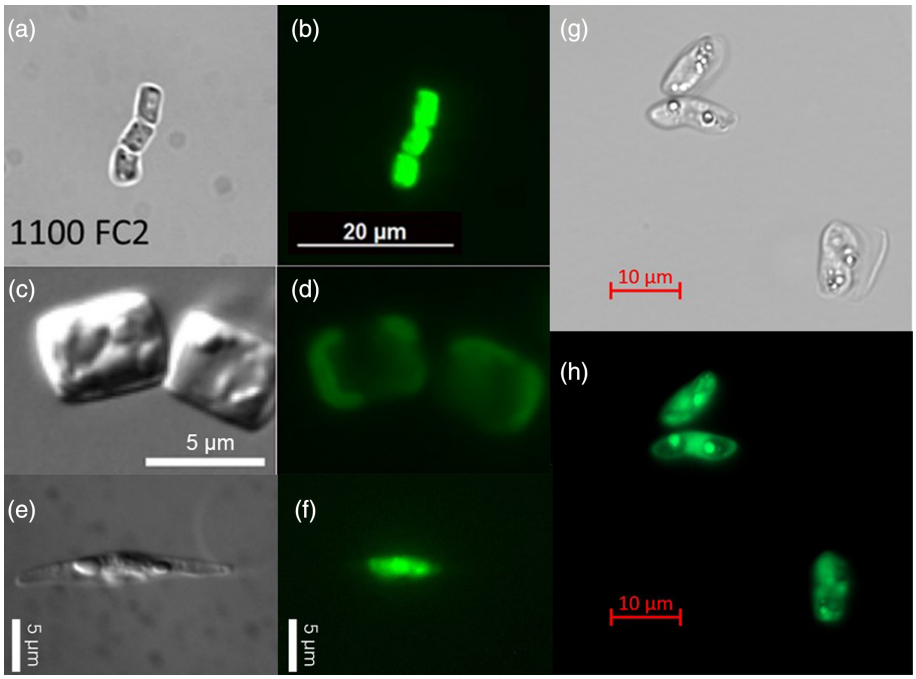


Figure 2. Expression of green fluorescent protein (GFP) or enhanced green fluorescent protein (eGFP) in four diatom species. Bright-field and corresponding GFP/eGFP fluorescence images are shown for each species.
(a, b) *Fragilariopsis cylindrus*, (c, d) *Thalassiosira pseudonana*, (e, f) *Phaeodactylum tricornutum*, and (g, h) *Nitzschia putrida*. Scale bars are 20 µm in (a, b), 10 µm in (g, h), and 5 µm in (c–f). Images have been taken by Jan Strauss, Amanda Hopes, and Longji Deng.

native conditions, respectively (Hao et al., 2022; Liu et al., 2016; Nojima et al., 2013; Pan et al., 2024; Shao et al., 2019). It should be noted that the folding and maturation of fluorescent reporter proteins can be a slow process that is influenced by the cellular environment (e.g., redox state). The delay between gene expression and protein maturation might introduce a bias between the detection of the fluorescence signals and the actual expression levels of the gene (Jullien & Gautier, 2015). Consequently, under such circumstances, it might be advantageous to combine the measurement of fluorescent signals with expression analyses of the target gene(s).

Cloning and assembly strategies

Restriction enzyme cloning likely is the most common cloning method in diatom research. Both the DNA fragment of interest and the vector are cut with the same restriction enzymes at specific recognition sites. The resulting complementary sticky or blunt ends allow the DNA fragment to be ligated into the vector using a DNA ligase, creating a recombinant DNA molecule. The first application for diatoms involved inserting the *nptII* gene into a plasmid, which conferred G418 resistance to the transformed *Cyclotella cryptica* (Dunahay et al., 1995). Type II restriction enzymes have long been the workhorses in diatom reverse genetics (Poulsen et al., 2006; Sabatino et al., 2022; Wang et al., 2023).

The Gibson assembly method is based on a seamless cloning strategy that joins multiple DNA fragments together in a specific order at a constant temperature using a T5 exonuclease, a DNA polymerase, and a DNA ligase (Gibson et al., 2009). This method is based on the assembly of overlapping fragments and is not constrained by restriction enzyme sites; hence, it offers flexibility in the design of vectors. Several studies in diatom research have successfully used this assembly method (Daboussi et al., 2014; Nigishi et al., 2024; Stukenberg et al., 2018). However, it should be noted that the efficiency decreases when the number of fragments increases beyond five due to challenges with ensuring correct annealing and maintaining overlap homology between fragments (Gibson et al., 2009). Thus, nicks or any kinds of base mutations in the overhangs will significantly impact the assembly success.

The golden gate assembly uses Type IIS restriction enzymes to enable precise and scar-free DNA assembly (Engler et al., 2008). The specified sticky ends on cuts are made at a defined distance from the recognition sites in plasmids and DNA fragments, achieving a hierarchical assembly of DNA parts (Bird et al., 2022). DNA parts are stored stably within plasmid vectors, allowing them to be efficiently reused in subsequent assembly steps. In diatoms, only *P. tricornutum* and *T. pseudonana* have been subjected to golden gate cloning (Belshaw et al., 2022; Hopes

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2016; Llavero-Pasquina et al., 2022; Mooshammer et al., 2020). Some of the golden gate components for both species are available from Addgene (Watertown, MA, USA) (https://www.addgene.org/). Furthermore, the modular cloning (MoClo) system is a hierarchical and modular assembly method based on golden gate cloning, combining standardized parts, such as promoters, coding sequences, and terminators in a predefined order (https://www.addgene.org/kits/marillonnet-moclo/). Several MoClo toolkits have been developed, providing reusable parts and vectors for mammalian (Weber et al., 2011), yeast (Lee et al., 2015), plants (Engler et al., 2014), bacteria (Iverson et al., 2016; Moore et al., 2016; Stukenberg et al., 2021), *Chlamydomonas reinhardtii* (Crozet et al., 2018), and cyanobacteria (Vasudevan et al., 2019). The use of a standardized syntax facilitates the sharing of non-species-specific modules (Patron et al., 2015). The MoClo toolkits have been used for diatoms (Nam et al., 2022; Russo et al., 2023) but not as commonly as other cloning methods.

The Loop assembly is another emerging and versatile DNA assembly system based on recursive DNA cloning. This method achieves high precision and reliability for assembling complex and larger constructs, with >80% average assembly efficiencies on over 200 different DNA constructs (Pollak et al., 2019). The universal Loop (uLoop) assembly, which is derived from the traditional Loop, has been successfully adapted for use in diatoms (Pollak et al., 2020). The diatom uLoop assembly kit is available from Addgene (https://www.addgene.org/kits/dupont-diatom-uloop/) and the uLoop library is still expanding to include new parts, such as newly characterized promoters and terminators (e.g. calm, oee3 and flav) in *P. tricornutum* (Garza et al., 2023). This open-access diatom uLoop library provides a modular and standardized approach for constructing diverse genetic components. Thus, it fosters collaboration between research groups and ensures data reproducibility through standardization.

There is an increasing use of *in vivo* assembly, particularly for constructing larger plasmids or those requiring the integration of multiple fragments. This approach leverages *Saccharomyces cerevisiae* (i.e., yeast) protoplast transformation methods (Kouprina & Larionov, 2008). A yeast artificial chromosome (YAC), often containing a CEN-ARS-HIS backbone, is included on a fragment that also incorporates essential elements for replication and selection in *E. coli*. The DNA fragments used for assembly can be chemically synthesized, PCR-amplified, or even directly obtained from isolated DNA, provided they are excised at desired positions using unique restriction enzymes. A minimum overlap of 40 base pairs (bp) between fragments is required for efficient recombination between complementary sequences; however, for assemblies involving a larger number of fragments, longer overlaps of 50–200 bp are recommended to enhance efficiency. The main advantage of this method is its capacity to assemble large plasmids, though it is more time-consuming and demands technical expertise. Notable applications of this technique include the assembly of entire mitochondrial genomes from *P. tricornutum* and *T. pseudonana* (Cochrane, Brumwell, Shrestha, et al., 2020; Cochrane, Brumwell, Soltysiak, et al., 2020) as well as the chloroplast genome of *P. tricornutum* (Walker et al., 2023). Furthermore, recent findings have demonstrated that plasmids can be directly assembled in *P. tricornutum* through non-homologous end joining, potentially simplifying the process of constructing genetic constructs, at least for this species (Walker et al., 2024).

DELIVERY SYSTEMS

Microparticle bombardment has been most widely used to deliver DNA, RNA, and proteins. It has been successfully employed for both nuclear and chloroplast transformation in many microalgal species (Hopes et al., 2016; Li & Bock, 2018; Schiedlmeier et al., 1994; Sodeinde & Kindle, 1993). The DNA vectors, RNAs, or proteins are coated onto the surface of nanoparticles either made of tungsten or gold. A particle delivery system is used to deliver them into the cells under high pressure. To minimize cell damage, particles between 0.7 and 1.1 μm are usually used for diatoms (Dunahay et al., 1995). However, considering that *T. pseudonana* cells are only about three times the size of these particles, it remains to be seen what the true impact is on the integrity of the diatom cells. At least they appear to have a high regenerative potential, evidenced by restoring the original phenotypes including the elaborate silica cell walls within 2–4 weeks after bombardment (Belshaw et al., 2022; Harada et al., 2005; Moosburner et al., 2020). Despite these shortcomings (e.g., low efficiency, destructive impact on cellular integrity), microparticle bombardment has been shown to be suitable for many diatom species. DNA plasmids in the size range of 5–10 kilobases (kb) are considered suitable for biolistic manipulation (Stewart et al., 2018). Biolistics can also be used to directly deliver RNA and proteins into diatom cells (Serif et al., 2018). This approach has been effectively applied to introduce Cas9/single-guide RNA ribonucleoprotein complexes (RNPs) for DNA-free genome editing (Serif et al., 2018). Meanwhile, triple gene knockouts were achieved in one step by delivering six RNP complexes simultaneously, although with relatively low efficiency (15%) compared with double gene knockout (23–52%).

Electroporation introduces foreign DNA into cells by temporarily creating pores in the cell membrane through the application of a brief electric pulse (Somiari et al., 2000; Tsong, 1991). This method transfers exogenous DNA independently of the cell's abilities to take it up and has been successfully used in several diatom species (Qin et al., 2012). It shows more than 10 times higher

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

transformation efficiency compared with conventional biolistics using small amounts of DNA (4–7 μg) (Naser et al., 2022; Yin & Hu, 2021; Zhang & Hu, 2014). Electroporation can also be used for co-transformation using more than a single vector (Zhang & Hu, 2014). Linear plasmids can be electroporated more efficiently compared with circular plasmids, achieving up to three times greater efficiency in diatom transformation (Yin & Hu, 2021). The spheroplasting electroporation method was recently developed for *P. tricornutum* (Walker et al., 2024). This method is characterized by high efficiency in delivering episomes, with amounts of as little as 1 ng and for plasmids as large as 55.6 kb. Additionally, this method has the advantage of high survival rates of the transformed diatom cells, with growth resuming within only 2 weeks (Ifuku et al., 2015; Yin & Hu, 2021; Zhang & Hu, 2014). Electroporation is a promising strategy, but the challenge lies in optimizing electroporation conditions, including the electric field strength and pulse duration to suit different diatom species. It is assumed that the cell wall significantly impedes DNA delivery during electroporation (Azencott et al., 2007). Therefore, it is at least necessary to adjust the protocol for different diatom species, and it is likely that this method, therefore, is not as widely applicable as biolistics.

The polyethylene glycol (PEG)-mediated transformation method is thought to facilitate the introduction of DNA into cells by promoting its passage across the cell membrane, though the precise mechanisms remain poorly understood. This method was evaluated in *P. tricornutum* during the development of the p0251s replicating plasmid (Karas et al., 2015); however, the results were highly inconsistent, yielding only a few colonies across several experiments. A recent breakthrough has occurred with the use of alcalase to protoplast *P. tricornutum* cells (Walker et al., in preparation), which dramatically increases the efficiency of PEG transformation (hundreds to thousands of transformants per reaction). This method shows great promise as it is highly efficient and does not require any specialized equipment, and it may become the preferred method for *P. tricornutum* and potentially other diatom species.

Using *E. coli* to directly introduce DNA vectors into diatom cells via bacteria-mediated conjugation is another delivery system that was developed in the recent past (Karas et al., 2015). The donor *E. coli* strain harbors a conjugative plasmid, which encodes for the DNA transfer machinery, and an episome containing the gene(s) of interest. The episome must also contain an origin of transfer (oriT) to be mobilized during conjugation and a selective marker. Furthermore, to be stably maintained extrachromosomally, the episome must contain a suitable autonomously replicating sequence (ARS). This sequence can vary between different diatom species; without an ARS, the episome must integrate into the diatom genome for continued propagation across successive generations. This method was first developed in *P. tricornutum* and *T. pseudonana* (Karas et al., 2015) and has since been explored in other diatom species.

High-copy number replication of the episome can be detrimental to the *E. coli* donor strain, particularly when CRISPR/Cas9 systems or other deleterious genes are expressed. To address the unstable characteristics of high-copy number plasmids, low-copy number and medium-copy number episomes have been developed. The pCC1BAC backbone (present in p0251s, Karas et al., 2015) contains two origins of replication: one that is constitutively expressed and facilitates single-copy number replication, and another that can be induced to high-copy number replication in the presence of L-arabinose when maintained in the EPI300 *E. coli* strain. Inducing high-copy number expression may be deleterious to the cell, but it can be useful when isolating large quantities of the episome for sequencing or other downstream applications. Another diatom episome, pPtPBR1, allows for constitutive medium-copy number expression in *E. coli* (Diner et al., 2016).

When using bacterial conjugation to deliver CRISPR/Cas9-containing episomes, this system achieves a comparable percentage of biallelic mutations when compared with microparticle bombardment (Moosburner et al., 2020; Sharma et al., 2018; Slattery et al., 2018). Compared with genome-integrated DNA vectors containing the Cas9 gene, extrachromosomal episomes carrying CRISPR/Cas9 systems can be easily removed by growing the diatom cells under non-selective conditions. This causes the loss of the episome over time (i.e, episome curing), which therefore avoids unwanted mutations by the constitutive expression of Cas9 if this gene has become part of the diatom genome. Thus, bacterial conjugation has the advantage of delivering genetic constructs without altering the host genome. However, the assembly of plasmids targeting multiple genes can lead to the instability of the constructs in *E. coli*, which needs to be considered when designing more complex multi-target constructs (Taparia et al., 2021).

CELL-LINE SCREENING POST-TRANSFORMATION

To isolate and confirm the desired genetic modifications, steps include subcloning and screening. Both are necessary because the initial cultures are often mosaic (Huang & Daboussi, 2017; Serif et al., 2018; Weyman et al., 2015), necessitating the isolation of single-cell lines on solid plates or in liquid medium. Most transformed diatom colonies can be isolated using antibiotics or chemical compounds on selective plates. However, the growth conditions on these plates should be tested when working with a novel diatom species. Clonal cell lines generally appear after the first round of subcloning. However, monoallelic (heterozygous) mutations frequently occur likely due to incomplete editing. In such cases, a second round of

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

subcloning is required to obtain bi-allelic (homozygous) mutations in the targeted gene. However, this process is time-consuming, taking more than 1 week to obtain clones, depending on the growth rates and the delivery system used (Belshaw et al., 2022; Yin & Hu, 2021). Transformed cells expressing a fluorescence marker can be separated by fluorescence-activated cell sorting (FACS), which therefore can achieve high rates of cell recovery under near axenic conditions (Nunez, 2001; Pereira et al., 2018). However, it should be noted that some diatom species may not survive the FACS procedure due to mechanical pressure, which may cause disruptions of the plasma membrane and cell wall (Reckermann, 2000).

PCR amplification of the target gene followed by analyzing the electropherograms is a common method for analyzing the genetic mutation to be expected (Hopes et al., 2016; Moosburner et al., 2022; Nymark et al., 2016). For gene overexpression and knockdown, selective marker or reporter genes are typically amplified for initial screening (De Riso et al., 2009; Haslam et al., 2020). Direct amplification of the target gene confirms its presence if the overexpressed gene is exogenous (Strauss et al., 2023). Gene-knockout cells often exhibit a distinct PCR-based genotyping pattern compared to WT cells (Hopes et al., 2016). When a small deletion occurs (less than 100 bp or a few hundred bp), gene-knockout cell lines typically display a smaller band than WT due to the reduced fragment size. Conversely, if the deletion spans several kilobases, no amplification is observed, necessitating additional confirmation using primers targeting the 3′ and 5′ flanking regions outside the deleted area (Belshaw et al., 2022). To further validate the modification, the expected sequence is cloned into a vector for transformation into *E. coli*, followed by Sanger sequencing and alignment with the reference genome. Additionally, alternative methods such as high-resolution melt curve analysis (HRM), T7 endonuclease I assay (T7EI), and software-based Tracking of Indels by Deconvolution (TIDE) provide insight into the nature of genetic modifications (Moosburner et al., 2020; Moosburner et al., 2022; Nymark et al., 2016; Slattery et al., 2018). If PCR reactions fail due to extensive genomic rearrangements, Southern blotting may be useful to detect fragment size shifts and confirm copy number variations compared to WT DNA (Kira et al., 2016; Zhang & Hu, 2014). To further corroborate the genetic modifications of coding genes, relative protein levels can be assessed by Western blotting, whereas mRNA levels of the target gene can be analyzed using real-time quantitative PCR (RT-qPCR) (Gorlich et al., 2019; Strauss et al., 2023).

Box 1. Summary

- Diatoms are important primary producers, and they are being used for diverse biotechnological applications.
- Multiomics resources available for many diatom species facilitate genetic engineering.
- Molecular tool kits include diverse promoters, selectable markers, and reporter systems.
- Cloning and assembling strategies have been developed even for large plasmids.
- The latest reverse genetics tools aim at building the first synthetic diatom.

Box 2. Open questions

- Considering that diatoms are the most species-rich group of algae, how many diatom species do we need to develop into model systems to address the most significant fundamental and applied questions?
- What kind of genetic engineering is required to identify the mechanisms underpinning the interactions of diatoms with other species?
- How can genetic engineering help to make use of diatoms in fast-emerging fields such as energy technology?
- Will it be possible to generate the first synthetic diatom for the development of novel carbon-capture mechanisms and the sustainable synthesis of high-value products?

PROSPECTS

Genetic engineering in diatoms has come a long way, and there are significant prospects because of three reasons: (i) Currently, there is a step change in the development of diatom genomics and multi-omics resources, which provide the foundation for discovering novel biology and the application of genetic tools to new species. (ii) Diatoms are as genetically tractable as any established model organisms in biology, that is, *Arabidopsis thaliana*, *yeast*, *Caenorhabditis elegans*. However, diatoms, unlike those established biological models, represent globally relevant organisms underpinning the largest food webs on Earth, and they drive global biogeochemical cycles responsible for the habitability of our planet. (iii) Their fast growth, high content of lipids, essential fatty acids, and their elaborate nanopatterned silica cell walls make them a target of the biotechnology sector, from nutraceuticals to material science. Consequently, for human societies to benefit from diatoms, we will need to continue to develop their genetic tractability. One current frontier is to design the first

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

synthetic diatom. To achieve that goal, synthetic chromosomes need to be assembled, transformed, and expressed in a diatom host, which remains challenging, but current work in this exciting field of research paves the way for building the first synthetic diatom soon. With respect to the latter, entire ~500 kbp chromosomes of *P. tricornutum* were successfully assembled in yeast and subsequently transferred to *E. coli*, establishing a critical proof of concept that demonstrates the ability of these organisms to maintain large diatom DNA fragments (Karas et al., 2013). Building on these results, it is proposed that all 25 chromosomes of *P. tricornutum* could be redesigned and resynthesized as 50 chromosomes that are ~400–500 kbp in size (Pampuch et al., 2022). This approach would allow for assembly in yeast and efficient propagation in *E. coli*. Furthermore, a recently developed rapid method for chloroplast genome assembly (Walker & Karas, 2025) could accelerate the creation of synthetic genomes with enhanced features, such as the removal of non-essential elements (e.g., repetitive sequences, transposons), genome recoding, and architectural reorganization. The delivery of whole chromosomes is anticipated to be possible through bacterial conjugation. Additionally, smaller fragments of approximately 50 kbp can now be delivered using an optimized electroporation protocol, enabling parallel testing of synthetic constructs (Walker et al., 2024). These recent advancements in the field of diatom genetic engineering, combined with many high-resolution reference genomes that will become available soon, position diatoms as robust and promising organisms for synthetic genomics to advance fundamental biological research and biotechnology (Boxes 1 and 2).

AUTHOR CONTRIBUTIONS

YL conceived and designed the structure of this manuscript, and she wrote the first version. All authors contributed to writing and editing the manuscript. LD designed Figures 1 and 2.

ACKNOWLEDGEMENTS

YL acknowledges the China Scholarship Council (Certification No. 202106330018) for funding through the International Cooperative Training Program to pursue her PhD thesis. TM acknowledges the School of Environmental Sciences at the University of East Anglia (UK) and the Natural Environment Research Council (NERC) (Grant NE/W005654/1) for partial support. Research in Karas lab is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), grant number RGPIN-2018-06172.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data of this study are available as part of the main article and the supplementary material.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table S1. Summary of sequenced diatom genomes. Resource from National Center for Biotechnology Information (NCBI), Joint Genome Institute (JGI), European Nucleotide Archive (ENA), Ensembl genome browser, DiatOmicBase, and PLAZA.

REFERENCES

Alonso, D.L., **Segura del Castillo, C.I.**, **Grima, E.M.** & **Cohen, Z.** (1996) First insights into improvement of eicosapentaenoic acid content in *Phaeodactylum tricornutum* (Bacillariophyceae) by induced mutagenesis. *Journal of Phycology*, **32**, 339–345. Available from: https://doi.org/10.1111/j.0022-3646.1996.00339.x

Apt, K.E., **Grossman, A.R.** & **Kroth-Pancic, P.G.** (1996) Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Molecular and General Genetics MGG*, **252**, 572–579. Available from: https://doi.org/10.1007/BF02172403

Armbrust, E.V., **Berges, J.A.**, **Bowler, C.**, **Green, B.R.**, **Martinez, D.**, **Putnam, N.H.** ***et al.*** (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*, **306**, 79–86.

Azencott, H.R., **Peter, G.F.** & **Prausnitz, M.R.** (2007) Influence of the cell wall on intracellular delivery to algal cells by electroporation and sonication. *Ultrasound in Medicine & Biology*, **33**, 1805–1817. Available from: https://doi.org/10.1016/j.ultrasmedbio.2007.05.008

Bar-On, Y.M., **Phillips, R.** & **Milo, R.** (2018) The biomass distribution on earth. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, 6506–6511. Available from: https://doi.org/10.1073/pnas.1711842115

Belshaw, N., **Grouneva, I.**, **Aram, L.**, **Gal, A.**, **Hopes, A.** & **Mock, T.** (2022) Efficient gene replacement by CRISPR/Cas-mediated homologous recombination in the model diatom *Thalassiosira pseudonana*. *The New Phytologist*, **238**, 438–452. Available from: https://doi.org/10.1111/nph.18587

Bird, J.E., **Marles-Wright, J.** & **Giachino, A.** (2022) A user's guide to Golden Gate cloning methods and standards. *ACS Synthetic Biology*, **11**, 3551–3563. Available from: https://doi.org/10.1021/acssynbio.2c00355

Bogdanove, A.J. & **Voytas, D.F.** (2011) TAL effectors: customizable proteins for DNA targeting. *Science*, **333**, 1843–1846. Available from: https://doi.org/10.1126/science.1204094

Bortesi, L. & **Fischer, R.** (2015) The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, **33**, 41–52. Available from: https://doi.org/10.1016/j.biotechadv.2014.12.006

Bowler, C., **Allen, A.E.**, **Badger, J.H.**, **Grimwood, J.**, **Jabbari, K.**, **Kuo, A.** ***et al.*** (2008) *The Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, **456**, 239–244. Available from: https://doi.org/10.1038/nature07410

Bowler, C., **Vardi, A.** & **Allen, A.E.** (2010) Oceanographic and biogeochemical insights from diatom genomes. *Annual Review of Marine Science*, **2**, 333–365. Available from: https://doi.org/10.1146/annurev-marine-120308-081051

Brodie, J., **Chan, C.X.**, **De Clerck, O.**, **Cock, J.M.**, **Coelho, S.M.**, **Gachon, C.** ***et al.*** (2017) The algal revolution. *Trends in Plant Science*, **22**, 726–738. Available from: https://doi.org/10.1016/j.tplants.2017.05.005

Buck, J.M., **Rio Bartulos, C.**, **Gruber, A.** & **Kroth, P.G.** (2018) Blasticidin-S deaminase, a new selection marker for genetic transformation of the diatom *Phaeodactylum tricornutum*. *PeerJ*, **6**, e5884. Available from: https://doi.org/10.7717/peerj.5884

Buhmann, M.T., **Poulsen, N.**, **Klemm, J.**, **Kennedy, M.R.**, **Sherrill, C.D.** & **Kroger, N.** (2014) A tyrosine-rich cell surface protein in the diatom *Amphora coffeaeformis* identified through transcriptome analysis and genetic transformation. *PLoS One*, **9**, e110369. Available from: https://doi.org/10.1371/journal.pone.0110369

Chevalier, B.S. & **Stoddard, B.L.** (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Research*, **29**, 3757–3774. Available from: https://doi.org/10.1093/nar/29.18.3757

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Christian, M., **Cermak, T.**, **Doyle, E.L.**, **Schmidt, C.**, **Zhang, F.**, **Hummel, A.** ***et al.*** (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, **186**, 757–761. Available from: https://doi.org/10.1534/genetics.110.120717

Chu, L., **Ewe, D.**, **Rio Bartulos, C.**, **Kroth, P.G.** & **Gruber, A.** (2016) Rapid induction of GFP expression by the nitrate reductase promoter in the diatom *Phaeodactylum tricornutum*. *PeerJ*, **4**, e2344. Available from: https://doi.org/10.7717/peerj.2344

Cochrane, R.R., **Brumwell, S.L.**, **Shrestha, A.**, **Giguere, D.J.**, **Hamadache, S.**, **Gloor, G.B.** ***et al.*** (2020) Cloning of *Thalassiosira pseudonana*'s mitochondrial genome in *Saccharomyces cerevisiae* and *Escherichia coli*. *Biology*, **9**, 358.

Cochrane, R.R., **Brumwell, S.L.**, **Soltysiak, M.P.M.**, **Hamadache, S.**, **Davis, J.G.**, **Wang, J.** ***et al.*** (2020) Rapid method for generating designer algal mitochondrial genomes. *Algal Research*, **50**, 102014. Available from: https://doi.org/10.1016/j.algal.2020.102014

Crozet, P., **Navarro, F.J.**, **Willmund, F.**, **Mehrshahi, P.**, **Bakowski, K.**, **Lauersen, K.J.** ***et al.*** (2018) Birth of a photosynthetic chassis: a MoClo toolkit enabling synthetic biology in the microalga *Chlamydomonas reinhardtii*. *ACS Synthetic Biology*, **7**, 2074–2086. Available from: https://doi.org/10.1021/acssynbio.8b00251

Daboussi, F., **Leduc, S.**, **Marechal, A.**, **Dubois, G.**, **Guyot, V.**, **Perez-Michaut, C.** ***et al.*** (2014) Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nature Communications*, **5**, 3831. Available from: https://doi.org/10.1038/ncomms4831

Davis, A., **Crum, L.T.**, **Corbeil, L.B.** & **Hildebrand, M.** (2017) Expression of *Histophilus somni* IbpA DR2 protective antigen in the diatom *Thalassiosira pseudonana*. *Applied Microbiology and Biotechnology*, **101**, 5313–5324. Available from: https://doi.org/10.1007/s00253-017-8267-8

De Riso, V., **Raniello, R.**, **Maumus, F.**, **Rogato, A.**, **Bowler, C.** & **Falciatore, A.** (2009) Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Research*, **37**, e96. Available from: https://doi.org/10.1093/nar/gkp448

Diamond, A., **Diaz-Garza, A.M.**, **Li, J.**, **Slattery, S.S.**, **Merindol, N.**, **Fantino, E.** ***et al.*** (2023) Instability of extrachromosomal DNA transformed into the diatom *Phaeodactylum tricornutum*. *Algal Research*, **70**, 102998. Available from: https://doi.org/10.1016/j.algal.2023.102998

Diner, R.E., **Bielinski, V.A.**, **Dupont, C.L.**, **Allen, A.E.** & **Weyman, P.D.** (2016) Refinement of the diatom episome maintenance sequence and improvement of conjugation-based DNA delivery methods. *Frontiers in Bioengineering and Biotechnology*, **4**, 65. Available from: https://doi.org/10.3389/fbioe.2016.00065

Dunahay, T.G., **Jarvis, E.E.** & **Roessler, P.G.** (1995) Genetic transformation of the diatoms *Cyclotella cryptica* and *Navicula saprophila*. *Journal of Phycology*, **31**, 1004–1012. Available from: https://doi.org/10.1111/j.0022-3646.1995.01004.x

Engler, C., **Kandzia, R.** & **Marillonnet, S.** (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, **3**, e3647. Available from: https://doi.org/10.1371/journal.pone.0003647

Engler, C., **Youles, M.**, **Gruetzner, R.**, **Ehnert, T.M.**, **Werner, S.**, **Jones, J.D.** ***et al.*** (2014) A golden gate modular cloning toolbox for plants. *ACS Synthetic Biology*, **3**, 839–843. Available from: https://doi.org/10.1021/sb4001504

Faktorova, D., **Nisbet, R.E.R.**, **Fernandez Robledo, J.A.**, **Casacuberta, E.**, **Sudek, L.**, **Allen, A.E.** ***et al.*** (2020) Genetic tool development in marine protists: emerging model organisms for experimental cell biology. *Nature Methods*, **17**, 481–494. Available from: https://doi.org/10.1038/s41592-020-0796-x

Falciatore, A., **Casotti, R.**, **Leblanc, C.**, **Abrescia, C.** & **Bowler, C.** (1999) Transformation of nonselectable reporter genes in marine diatoms. *Marine Biotechnology (New York, N.Y.)*, **1**, 239–251. Available from: https://doi.org/10.1007/pl00011773

Falciatore, A., **Jaubert, M.**, **Bouly, J.P.**, **Bailleul, B.** & **Mock, T.** (2020) Diatom molecular research comes of age: model species for studying phytoplankton biology and diversity. *Plant Cell*, **32**, 547–572. Available from: https://doi.org/10.1105/tpc.19.00158

Fischer, H., **Robl, I.**, **Sumper, M.** & **Kröger, N.** (1999) Targeting and covalent modification of cell wall and membrane proteins heterologously expressed in the diatom *Cylindrotheca fusiformis* (Bacillariophyceae). *Journal of Phycology*, **35**, 113–120. Available from: https://doi.org/10.1046/j.1529-8817.1999.3510113.x

Fortunato, A.E., **Jaubert, M.**, **Enomoto, G.**, **Bouly, J.P.**, **Raniello, R.**, **Thaler, M.** ***et al.*** (2016) Diatom phytochromes reveal the existence of far-red-light-based sensing in the ocean. *Plant Cell*, **28**, 616–628. Available from: https://doi.org/10.1105/tpc.15.00928

Fredens, J., **Wang, K.**, **de la Torre, D.**, **Funke, L.F.H.**, **Robertson, W.E.**, **Christova, Y.** ***et al.*** (2019) Total synthesis of *Escherichia coli* with a recoded genome. *Nature*, **569**, 514–518. Available from: https://doi.org/10.1038/s41586-019-1192-5

Garza, E.A., **Bielinski, V.A.**, **Espinoza, J.L.**, **Orlandi, K.**, **Alfaro, J.R.**, **Bolt, T.M.** ***et al.*** (2023) Validating a promoter library for application in plasmid-based diatom genetic engineering. *ACS Synthetic Biology*, **12**, 3215–3228. Available from: https://doi.org/10.1021/acssynbio.3c00163

Gibson, D.G., **Young, L.**, **Chuang, R.Y.**, **Venter, J.C.**, **Hutchison, C.A., 3rd** & **Smith, H.O.** (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, **6**, 343–345. Available from: https://doi.org/10.1038/nmeth.1318

Goold, H.D., **Moseley, J.L.** & **Lauersen, K.J.** (2024) The synthetic future of algal genomes. *Cell Genomics*, **4**, 100505. Available from: https://doi.org/10.1016/j.xgen.2024.100505

Gorlich, S., **Pawolski, D.**, **Zlotnikov, I.** & **Kroger, N.** (2019) Control of biosilica morphology and mechanical performance by the conserved diatom gene silicanin-1. *Communications Biology*, **2**, 245. Available from: https://doi.org/10.1038/s42003-019-0436-0

Grypioti, E., **Richard, H.**, **Kryovrysanaki, N.**, **Jaubert, M.**, **Falciatore, A.**, **Verret, F.** ***et al.*** (2024) Dicer-dependent heterochromatic small RNAs in the model diatom species *Phaeodactylum tricornutum*. *The New Phytologist*, **241**, 811–826. Available from: https://doi.org/10.1111/nph.19429

Guo, W., **Weng, Y.**, **Ma, W.**, **Chang, C.**, **Gao, Y.**, **Huang, X.** ***et al.*** (2024) Improving lipid content in the diatom *Phaeodactylum tricornutum* by the knockdown of the enoyl-CoA hydratase using CRISPR interference. *Current Issues in Molecular Biology*, **46**, 10923–10933.

Hao, X., **Chen, W.**, **Amato, A.**, **Jouhet, J.**, **Marechal, E.**, **Moog, D.** ***et al.*** (2022) Multiplexed CRISPR/Cas9 editing of the long-chain acyl-CoA synthetase family in the diatom *Phaeodactylum tricornutum* reveals that mitochondrial ptACSL3 is involved in the synthesis of storage lipids. *The New Phytologist*, **233**, 1797–1812. Available from: https://doi.org/10.1111/nph.17911

Harada, H., **Nakatsuma, D.**, **Ishida, M.** & **Matsuda, Y.** (2005) Regulation of the expression of intracellular beta-carbonic anhydrase in response to CO_2 and light in the marine diatom *Phaeodactylum tricornutum*. *Plant Physiology*, **139**, 1041–1050. Available from: https://doi.org/10.1104/pp.105.065185

Haslam, R.P., **Hamilton, M.L.**, **Economou, C.K.**, **Smith, R.**, **Hassall, K.L.**, **Napier, J.A.** ***et al.*** (2020) Overexpression of an endogenous type 2 diacylglycerol acyltransferase in the marine diatom *Phaeodactylum tricornutum* enhances lipid production and omega-3 long-chain polyunsaturated fatty acid content. *Biotechnology for Biofuels*, **13**, 87. Available from: https://doi.org/10.1186/s13068-020-01726-8

Hopes, A., **Nekrasov, V.**, **Kamoun, S.** & **Mock, T.** (2016) Editing of the urease gene by CRISPR-Cas in the diatom *Thalassiosira pseudonana*. *Plant Methods*, **12**, 49. Available from: https://doi.org/10.1186/s13007-016-0148-0

Hu, H. & **Pan, Y.** (2020) Electroporation transformation protocol for *Phaeodactylum tricornutum*. In: Li, S., Chang, L. & Teissie, J. (Eds.) *Electroporation protocols: microorganism, mammalian system, and nanodevice*. New York, NY: Springer US, pp. 163–167.

Huang, W. & **Daboussi, F.** (2017) Genetic and metabolic engineering in diatoms. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **372**, 411. Available from: https://doi.org/10.1098/rstb.2016.0411

Hutchison, C.A., **Chuang, R.-Y.**, **Noskov, V.N.**, **Assad-Garcia, N.**, **Deerinck, T.J.**, **Ellisman, M.H.** ***et al.*** (2016) Design and synthesis of a minimal bacterial genome. *Science*, **351**, aad6253. Available from: https://doi.org/10.1126/science.aad6253

Huttly, A. (2009) Reporter genes. *Methods in Molecular Biology*, **478**, 39–69. Available from: https://doi.org/10.1007/978-1-59745-379-0_3

Ifuku, K., **Yan, D.**, **Miyahara, M.**, **Inoue-Kashino, N.**, **Yamamoto, Y.Y.** & **Kashino, Y.** (2015) A stable and efficient nuclear transformation system for the diatom *Chaetoceros gracilis*. *Photosynthesis Research*, **123**, 203–211. Available from: https://doi.org/10.1007/s11120-014-0048-y

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Iverson, S.V., **Haddock, T.L.**, **Beal, J.** & **Densmore, D.M.** (2016) CIDAR MoClo: improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. *ACS Synthetic Biology*, **5**, 99–103. Available from: https://doi.org/10.1021/acssynbio.5b00124

Jensen, R.H., **Astvad, K.M.**, **Silva, L.V.**, **Sanglard, D.**, **Jorgensen, R.**, **Nielsen, K.F.** ***et al.*** (2015) Stepwise emergence of azole, echinocandin and amphotericin B multidrug resistance in vivo in *Candida albicans* orchestrated by multiple genetic alterations. *The Journal of Antimicrobial Chemotherapy*, **70**, 2551–2555. Available from: https://doi.org/10.1093/jac/dkv140

Johansson, O.N., **Topel, M.**, **Pinder, M.I.M.**, **Kourtchenko, O.**, **Blomberg, A.**, **Godhe, A.** ***et al.*** (2019) Skeletonema marinoi as a new genetic model for marine chain-forming diatoms. *Scientific Reports*, **9**, 5391. Available from: https://doi.org/10.1038/s41598-019-41085-5

Jullien, L. & **Gautier, A.** (2015) Fluorogen-based reporters for fluorescence imaging: a review. *Methods and Applications in Fluorescence*, **3**, 042007. Available from: https://doi.org/10.1088/2050-6120/3/4/042007

Kadono, T., **Miyagawa-Yamaguchi, A.**, **Kira, N.**, **Tomaru, Y.**, **Okami, T.**, **Yoshimatsu, T.** ***et al.*** (2015) Characterization of marine diatom-infecting virus promoters in the model diatom *Phaeodactylum tricornutum*. *Scientific Reports*, **5**, 18708. Available from: https://doi.org/10.1038/srep18708

Kadono, T., **Tomaru, Y.**, **Sato, N.**, **Watanabe, Y.**, **Suzuki, K.**, **Yamada, K.** ***et al.*** (2022) Characterization of *Chaetoceros lorenzianus*-infecting DNA virus-derived promoters of genes from open reading frames of unknown function in *Phaeodactylum tricornutum*. *Marine Genomics*, **61**, 100921. Available from: https://doi.org/10.1016/j.margen.2021.100921

Kadono, T., **Tomaru, Y.**, **Suzuki, K.**, **Yamada, K.** & **Adachi, M.** (2020) The possibility of using marine diatom-infecting viral promoters for the engineering of marine diatoms. *Plant Science*, **296**, 110475. Available from: https://doi.org/10.1016/j.plantsci.2020.110475

Kamikawa, R., **Mochizuki, T.**, **Sakamoto, M.**, **Tanizawa, Y.**, **Nakayama, T.**, **Onuma, R.** ***et al.*** (2022) Genome evolution of a nonparasitic secondary heterotroph, the diatom *Nitzschia putrida*. *Science Advances*, **8**, eabi5075. Available from: https://doi.org/10.1126/sciadv.abi5075

Karas, B.J., **Diner, R.E.**, **Lefebvre, S.C.**, **McQuaid, J.**, **Phillips, A.P.**, **Noddings, C.M.** ***et al.*** (2015) Designer diatom episomes delivered by bacterial conjugation. *Nature Communications*, **6**, 6925. Available from: https://doi.org/10.1038/ncomms7925

Karas, B.J., **Molparia, B.**, **Jablanovic, J.**, **Hermann, W.J.**, **Lin, Y.-C.**, **Dupont, C.L.** ***et al.*** (2013) Assembly of eukaryotic algal chromosomes in yeast. *Journal of Biological Engineering*, **7**, 30. Available from: https://doi.org/10.1186/1754-1611-7-30

Kassaw, T.K., **Paton, A.J.** & **Peers, G.** (2022) Episome-based gene expression modulation platform in the model diatom *Phaeodactylum tricornutum*. *ACS Synthetic Biology*, **11**, 191–204. Available from: https://doi.org/10.1021/acssynbio.1c00367

Kira, N., **Ohnishi, K.**, **Miyagawa-Yamaguchi, A.**, **Kadono, T.** & **Adachi, M.** (2016) Nuclear transformation of the diatom *Phaeodactylum tricornutum* using PCR-amplified DNA fragments by microparticle bombardment. *Marine Genomics*, **25**, 49–56. Available from: https://doi.org/10.1016/j.margen.2015.12.004

Kirkham, A.R., **Richthammer, P.**, **Schmidt, K.**, **Wustmann, M.**, **Maeda, Y.**, **Hedrich, R.** ***et al.*** (2017) A role for the cell-wall protein silacidin in cell size of the diatom *Thalassiosira pseudonana*. *The ISME Journal*, **11**, 2452–2464. Available from: https://doi.org/10.1038/ismej.2017.100

Kouprina, N. & **Larionov, V.** (2008) Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nature Protocols*, **3**, 371–377. Available from: https://doi.org/10.1038/nprot.2008.5

Leblanc, K., **Arístegui Ruiz, J.**, **Armand, L.K.**, **Assmy, P.**, **Beker, B.**, **Bode, A.** ***et al.*** (2012) Global distributions of diatoms abundance, biovolume and biomass – gridded data product (NetCDF) – contribution to the MARE-DAT World Ocean atlas of plankton functional types. *PANGAEA*. Available from: https://doi.org/10.1594/PANGAEA.777384

Lee, J., **Lim, K.**, **Kim, A.**, **Mok, Y.G.**, **Chung, E.**, **Cho, S.I.** ***et al.*** (2023) Prime editing with genuine Cas9 nickases minimizes unwanted indels. *Nature Communications*, **14**, 1786. Available from: https://doi.org/10.1038/s41467-023-37507-8

Lee, M.E., **DeLoache, W.C.**, **Cervantes, B.** & **Dueber, J.E.** (2015) A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synthetic Biology*, **4**, 975–986. Available from: https://doi.org/10.1021/sb500366v

Li, Z. & **Bock, R.** (2018) Replication of bacterial plasmids in the nucleus of the red alga *Porphyridium purpureum*. *Nature Communications*, **9**, 3451. Available from: https://doi.org/10.1038/s41467-018-05651-1

Lin, H.-Y., **Shih, C.-Y.**, **Liu, H.-C.**, **Chang, J.**, **Chen, Y.-L.**, **Chen, Y.-R.** ***et al.*** (2013) Identification and characterization of an extracellular alkaline phosphatase in the marine diatom *Phaeodactylum tricornutum*. *Marine Biotechnology*, **15**, 425–436. Available from: https://doi.org/10.1007/s10126-013-9494-3

Lin, H.-Y., **Yen, S.-C.**, **Kuo, P.-C.**, **Chung, C.-Y.**, **Yeh, K.-L.**, **Huang, C.-H.** ***et al.*** (2017) Alkaline phosphatase promoter as an efficient driving element for exogenic recombinant in the marine diatom *Phaeodactylum tricornutum*. *Algal Research*, **23**, 58–65. Available from: https://doi.org/10.1016/j.algal.2017.01.007

Liu, X., **Hempel, F.**, **Stork, S.**, **Bolte, K.**, **Moog, D.**, **Heimerl, T.** ***et al.*** (2016) Addressing various compartments of the diatom model organism *Phaeodactylum tricornutum* via sub-cellular marker proteins. *Algal Research*, **20**, 249–257. Available from: https://doi.org/10.1016/j.algal.2016.10.018

Llavero-Pasquina, M., **Geisler, K.**, **Holzer, A.**, **Mehrshahi, P.**, **Mendoza-Ochoa, G.I.**, **Newsad, S.A.** ***et al.*** (2022) Thiamine metabolism genes in diatoms are not regulated by thiamine despite the presence of predicted riboswitches. *The New Phytologist*, **235**, 1853–1867. Available from: https://doi.org/10.1111/nph.18296

Maeda, Y., **Watanabe, K.**, **Kaha, M.**, **Yabu, Y.**, **Yoshino, T.**, **Matsumoto, M.** ***et al.*** (2021) Assessment on the oil accumulation by knockdown of triacylglycerol lipase in the oleaginous diatom *Fistulifera solaris*. *Scientific Reports*, **11**, 20905. Available from: https://doi.org/10.1038/s41598-021-00453-w

Malviya, S., **Scalco, E.**, **Audic, S.**, **Vincent, F.**, **Veluchamy, A.**, **Poulain, J.** ***et al.*** (2016) Insights into global diatom distribution and diversity in the world's ocean. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E1516–E1525. Available from: https://doi.org/10.1073/pnas.1509523113

Malzahn, A., **Lowder, L.** & **Qi, Y.** (2017) Plant genome editing with TALEN and CRISPR. *Cell & Bioscience*, **7**, 21. Available from: https://doi.org/10.1186/s13578-017-0148-4

Mann, D.G. & **Vanormelingen, P.** (2013) An inordinate fondness? The number, distributions, and origins of diatom species. *Journal of Eukaryotic Microbiology*, **60**, 414–420. Available from: https://doi.org/10.1111/jeu.12047

Matsui, H., **Harada, H.**, **Maeda, K.**, **Sugiyama, T.**, **Fukuchi, Y.**, **Kimura, N.** ***et al.*** (2024) Coordinated phosphate uptake by extracellular alkaline phosphatase and solute carrier transporters in marine diatoms. *The New Phytologist*, **241**, 1210–1221. Available from: https://doi.org/10.1111/nph.19410

Miyagawa, A., **Okami, T.**, **Kira, N.**, **Yamaguchi, H.**, **Ohnishi, K.** & **Adachi, M.** (2009) Research note: high efficiency transformation of the diatom *Phaeodactylum tricornutum* with a promoter from the diatom *Cylindrotheca fusiformis*. *Phycological Research*, **57**, 142–146. Available from: https://doi.org/10.1111/j.1440-1835.2009.00531.x

Miyagawa-Yamaguchi, A., **Okami, T.**, **Kira, N.**, **Yamaguchi, H.**, **Ohnishi, K.** & **Adachi, M.** (2011) Stable nuclear transformation of the diatom *Chaetoceros* sp. *Phycological Research*, **59**, 113–119. Available from: https://doi.org/10.1111/j.1440-1835.2011.00607.x

Mock, T., **Hodgkinson, K.**, **Wu, T.**, **Moulton, V.**, **Duncan, A.**, **van Oosterhout, C.** ***et al.*** (2022) Structure and evolution of diatom nuclear genes and genomes. In: Falciatore, A. & Mock, T. (Eds.) *The molecular life of diatoms*. New York: Springer, pp. 111–145.

Mock, T., **Otillar, R.P.**, **Strauss, J.**, **McMullan, M.**, **Paajanen, P.**, **Schmutz, J.** ***et al.*** (2017) Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature*, **541**, 536–540. Available from: https://doi.org/10.1038/nature20803

Moirangthem, R., **Kumar, K.** & **Kaur, R.** (2021) Two functionally redundant FK506-binding proteins regulate multidrug resistance gene expression and govern azole antifungal resistance. *Antimicrobial Agents and Chemotherapy*, **65**, e02415-20. Available from: https://doi.org/10.1128/aac.02415-20

Moore, S.J., **Lai, H.-E.**, **Kelwick, R.J.R.**, **Chee, S.M.**, **Bell, D.J.**, **Polizzi, K.M.** ***et al.*** (2016) EcoFlex: a multifunctional MoClo kit for *E. coli* synthetic

© 2025 The Author(s).
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.,

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

biology. *ACS Synthetic Biology*, **5**, 1059–1069. Available from: https://doi.org/10.1021/acssynbio.6b00031

Moosburner, M., **Allen, A.E.** & **Daboussi, F.** (2022) Genetic engineering in marine diatoms: current practices and emerging technologies. In: Falciatore, A. & Mock, T. (Eds.) *The molecular life of diatoms*. London: Springer Press, pp. 743–773.

Moosburner, M.A., **Gholami, P.**, **McCarthy, J.K.**, **Tan, M.**, **Bielinski, V.A.** & **Allen, A.E.** (2020) Multiplexed knockouts in the model diatom *Phaeodactylum* by episomal delivery of a selectable Cas9. *Frontiers in Microbiology*, **11**, 5. Available from: https://doi.org/10.3389/fmicb.2020.00005

Mooshammer, M., **Alves, R.J.E.**, **Bayer, B.**, **Melcher, M.**, **Stieglmeier, M.**, **Jochum, L.** ***et al.*** (2020) Nitrogen isotope fractionation during archaeal ammonia oxidation: coupled estimates from measurements of residual ammonium and accumulated nitrite. *Frontiers in Microbiology*, **11**, 1710. Available from: https://doi.org/10.3389/fmicb.2020.01710

Morio, F., **Jensen, R.H.**, **Le Pape, P.** & **Arendrup, M.C.** (2017) Molecular basis of antifungal drug resistance in yeasts. *International Journal of Antimicrobial Agents*, **50**, 599–606. Available from: https://doi.org/10.1016/j.ijantimicag.2017.05.012

Muto, M., **Fukuda, Y.**, **Nemoto, M.**, **Yoshino, T.**, **Matsunaga, T.** & **Tanaka, T.** (2013) Establishment of a genetic transformation system for the marine pennate diatom *Fistulifera* sp. strain JPCC DA0580 – a high triglyceride producer. *Marine Biotechnology (New York, N.Y.)*, **15**, 48–55. Available from: https://doi.org/10.1007/s10126-012-9457-0

Nakov, T., **Beaulieu, J.M.** & **Alverson, A.J.** (2018) Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (diatoms, Bacillariophyta). *The New Phytologist*, **219**, 462–473. Available from: https://doi.org/10.1111/nph.15137

Nam, O., **Grouneva, I.** & **Mackinder, L.C.M.** (2022) Endogenous GFP tagging in the diatom *Thalassiosira pseudonana*. *bioRxiv*, 2022.09.30.510313. Available from: https://doi.org/10.1101/2022.09.30.510313

Nam, O., **Musial, S.**, **Demulder, M.**, **McKenzie, C.**, **Dowle, A.**, **Dowson, M.** ***et al.*** (2024) A protein blueprint of the diatom CO_2-fixing organelle. *Cell*, **187**, 5935–5950.e18. Available from: https://doi.org/10.1016/j.cell.2024.09.025

Naser, I., **Yabu, Y.**, **Maeda, Y.** & **Tanaka, T.** (2022) Highly efficient genetic transformation methods for the marine oleaginous diatom *Fistulifera solaris*. *Marine Biotechnology (New York, N.Y.)*, **25**, 657–665. Available from: https://doi.org/10.1007/s10126-022-10189-x

Nawaly, H., **Tsuji, Y.** & **Matsuda, Y.** (2020) Rapid and precise genome editing in a marine diatom, *Thalassiosira pseudonana* by Cas9 nickase (D10A). *Algal Research*, **47**, 1855. Available from: https://doi.org/10.1016/j.algal.2020.101855

Nemudryi, A.A., **Valetdinova, K.R.**, **Medvedev, S.P.** & **Zakian, S.M.** (2014) TALEN and CRISPR/Cas genome editing systems: tools of discovery. *Acta Naturae*, **6**, 19–40.

Nigishi, M., **Shimakawa, G.**, **Yamagishi, K.**, **Amano, R.**, **Ito, S.**, **Tsuji, Y.** ***et al.*** (2024) Low-CO_2-inducible bestrophins outside the pyrenoid sustain high photosynthetic efficacy in diatoms. *Plant Physiology*, **195**, 1432–1445. Available from: https://doi.org/10.1093/plphys/kiae137

Nojima, D., **Yoshino, T.**, **Maeda, Y.**, **Tanaka, M.**, **Nemoto, M.** & **Tanaka, T.** (2013) Proteomics analysis of oil body-associated proteins in the oleaginous diatom. *Journal of Proteome Research*, **12**, 5293–5301. Available from: https://doi.org/10.1021/pr4004085

Nunez, R. (2001) Flow cytometry: principles and instrumentation. *Current Issues in Molecular Biology*, **3**, 39–45.

Nymark, M., **Sharma, A.K.**, **Sparstad, T.**, **Bones, A.M.** & **Winge, P.** (2016) A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports*, **6**, 24951. Available from: https://doi.org/10.1038/srep24951

Okada, K., **Morimoto, Y.**, **Shiraishi, Y.**, **Tamura, T.**, **Mayama, S.**, **Kadono, T.** ***et al.*** (2023) Nuclear transformation of the marine pennate diatom *Nitzschia* sp. strain NIES-4635 by multi-pulse electroporation. *Marine Biotechnology (New York, N.Y.)*, **25**, 1208–1219. Available from: https://doi.org/10.1007/s10126-023-10273-w

Pampuch, M., **Walker, E.J.L.** & **Karas, B.J.** (2022) Towards synthetic diatoms: the *Phaeodactylum tricornutum* Pt-syn 1.0 project. *Current Opinion in Green and Sustainable Chemistry*, **35**, 100611. Available from: https://doi.org/10.1016/j.cogsc.2022.100611

Pan, Y., **Zhang, W.**, **Wang, X.**, **Jouhet, J.**, **Marechal, E.**, **Liu, J.** ***et al.*** (2024) Allele-dependent expression and functionality of lipid enzyme phospholipid:diacylglycerol acyltransferase affect diatom carbon storage and growth. *Plant Physiology*, **194**, 1024–1040. Available from: https://doi.org/10.1093/plphys/kiad581

Patron, N.J., **Orzaez, D.**, **Marillonnet, S.**, **Warzecha, H.**, **Matthewman, C.**, **Youles, M.** ***et al.*** (2015) Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *The New Phytologist*, **208**, 13–19. Available from: https://doi.org/10.1111/nph.13532

Pereira, H., **Schulze, P.S.C.**, **Schüler, L.M.**, **Santos, T.**, **Barreira, L.** & **Varela, J.** (2018) Fluorescence activated cell-sorting principles and applications in microalgal biotechnology. *Algal Research*, **30**, 113–120. Available from: https://doi.org/10.1016/j.algal.2017.12.013

Pickar-Oliver, A. & **Gersbach, C.A.** (2019) The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology*, **20**, 490–507. Available from: https://doi.org/10.1038/s41580-019-0131-5

Pinseel, E., **Janssens, S.B.**, **Verleyen, E.**, **Vanormelingen, P.**, **Kohler, T.J.**, **Biersma, E.M.** ***et al.*** (2020) Global radiation in a rare biosphere soil diatom. *Nature Communications*, **11**, 2382. Available from: https://doi.org/10.1038/s41467-020-16181-0

Pollak, B., **Cerda, A.**, **Delmans, M.**, **Alamos, S.**, **Moyano, T.**, **West, A.** ***et al.*** (2019) Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *The New Phytologist*, **222**, 628–640. Available from: https://doi.org/10.1111/nph.15625

Pollak, B., **Matute, T.**, **Nunez, I.**, **Cerda, A.**, **Lopez, C.**, **Vargas, V.** ***et al.*** (2020) Universal loop assembly: open, efficient and cross-kingdom DNA fabrication. *Synthetic Biology (Oxford, England)*, **5**, ysaa001. Available from: https://doi.org/10.1093/synbio/ysaa001

Poulsen, N., **Chesley, P.M.** & **Kröger, N.** (2006) Molecular genetic manipulation of the diatom *Thalassiosira pseudonana* (Bacillariophyceae). *Journal of Phycology*, **42**, 1059–1065. Available from: https://doi.org/10.1111/j.1529-8817.2006.00269.x

Poulsen, N. & **Kroger, N.** (2005) A new molecular tool for transgenic diatoms: control of mRNA and protein biosynthesis by an inducible promoter-terminator cassette. *The FEBS Journal*, **272**, 3413–3423. Available from: https://doi.org/10.1111/j.1742-4658.2005.04760.x

Qin, S., **Lin, H.** & **Jiang, P.** (2012) Advances in genetic engineering of marine algae. *Biotechnology Advances*, **30**, 1602–1613. Available from: https://doi.org/10.1016/j.biotechadv.2012.05.004

Reckermann, M. (2000) Flow sorting in aquatic ecology. *Scientia Marina*, **64**, 235–246. Available from: https://doi.org/10.3989/scimar.2000.64n2235

Russo, M.T., **Rogato, A.**, **Jaubert, M.**, **Karas, B.J.** & **Falciatore, A.** (2023) *Phaeodactylum tricornutum*: an established model species for diatom molecular research and an emerging chassis for algal synthetic biology. *Journal of Phycology*, **59**, 1114–1122. Available from: https://doi.org/10.1111/jpy.13400

Russo, M.T., **Vitale, L.**, **Entrambasaguas, L.**, **Anestis, K.**, **Fattorini, N.**, **Romano, F.** ***et al.*** (2018) MRP3 is a sex determining gene in the diatom *Pseudo-nitzschia multistriata*. *Nature Communications*, **9**, 5050. Available from: https://doi.org/10.1038/s41467-018-07496-0

Rybak, J.M., **Barker, K.S.**, **Munoz, J.F.**, **Parker, J.E.**, **Ahmad, S.**, **Mokaddas, E.** ***et al.*** (2022) In vivo emergence of high-level resistance during treatment reveals the first identified mechanism of amphotericin B resistance in *Candida auris*. *Clinical Microbiology and Infection*, **28**, 838–843. Available from: https://doi.org/10.1016/j.cmi.2021.11.024

Sabatino, V., **Orefice, I.**, **Marotta, P.**, **Ambrosino, L.**, **Chiusano, M.L.**, **d'Ippolito, G.** ***et al.*** (2022) Silencing of a *Pseudo-nitzschia arenysensis* lipoxygenase transcript leads to reduced oxylipin production and impaired growth. *The New Phytologist*, **233**, 809–822. Available from: https://doi.org/10.1111/nph.17739

Sabatino, V., **Russo, M.T.**, **Patil, S.**, **d'Ippolito, G.**, **Fontana, A.** & **Ferrante, M.I.** (2015) Establishment of genetic transformation in the sexually reproducing diatoms *Pseudo-nitzschia multistriata* and *Pseudo-nitzschia arenysensis* and inheritance of the transgene. *Marine Biotechnology (New York, N.Y.)*, **17**, 452–462. Available from: https://doi.org/10.1007/s10126-015-9633-0

Sakaguchi, T., **Nakajima, K.** & **Matsuda, Y.** (2011) Identification of the UMP synthase gene by establishment of uracil auxotrophic mutants and the phenotypic complementation system in the marine diatom *Phaeodactylum tricornutum*. *Plant Physiology*, **156**, 78–89. Available from: https://doi.org/10.1104/pp.110.169631

Sakaue, K., **Harada, H.** & **Matsuda, Y.** (2008) Development of gene expression system in a marine diatom using viral promoters of a wide variety

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of origin. *Physiologia Plantarum*, **133**, 59–67. Available from: https://doi.org/10.1111/j.1399-3054.2008.01089.x

Schellenberger Costa, B., **Sachse, M.**, **Jungandreas, A.**, **Bartulos, C.R.**, **Gruber, A.**, **Jakob, T.** ***et al.*** (2013) Aureochrome 1a is involved in the photoacclimation of the diatom *Phaeodactylum tricornutum*. *PLoS One*, **8**, e74451. Available from: https://doi.org/10.1371/journal.pone.0074451

Schiedlmeier, B., **Schmitt, R.**, **Müller, W.**, **Kirk, M.M.**, **Gruber, H.**, **Mages, W.** ***et al.*** (1994) Nuclear transformation of *Volvox carteri*. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 5080–5084. Available from: https://doi.org/10.1073/pnas.91.11.5080

Schindler, D., **Walker, R.S.K.** & **Cai, Y.** (2024) Methodological advances enabled by the construction of a synthetic yeast genome. *Cell Reports Methods*, **4**, 100761. Available from: https://doi.org/10.1016/j.crmeth.2024.100761

Seo, S., **Jeon, H.**, **Hwang, S.**, **Jin, E.** & **Chang, K.S.** (2015) Development of a new constitutive expression system for the transformation of the diatom *Phaeodactylum tricornutum*. *Algal Research*, **11**, 50–54. Available from: https://doi.org/10.1016/j.algal.2015.05.012

Serif, M., **Dubois, G.**, **Finoux, A.L.**, **Teste, M.A.**, **Jallet, D.** & **Daboussi, F.** (2018) One-step generation of multiple gene knock-outs in the diatom *Phaeodactylum tricornutum* by DNA-free genome editing. *Nature Communications*, **9**, 3924. Available from: https://doi.org/10.1038/s41467-018-06378-9

Shao, Z., **Thomas, Y.**, **Hembach, L.**, **Xing, X.**, **Duan, D.**, **Moerschbacher, B.M.** ***et al.*** (2019) Comparative characterization of putative chitin deacetylases from *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* highlights the potential for distinct chitin-based metabolic processes in diatoms. *The New Phytologist*, **221**, 1890–1905. Available from: https://doi.org/10.1111/nph.15510

Sharma, A.K., **Nymark, M.**, **Sparstad, T.**, **Bones, A.M.** & **Winge, P.** (2018) Transgene-free genome editing in marine algae by bacterial conjugation – comparison with biolistic CRISPR/Cas9 transformation. *Scientific Reports*, **8**, 14401. Available from: https://doi.org/10.1038/s41598-018-32342-0

Shrestha, R.P. & **Hildebrand, M.** (2017) Development of a silicon limitation inducible expression system for recombinant protein production in the centric diatoms *Thalassiosira pseudonana* and *Cyclotella cryptica*. *Microbial Cell Factories*, **16**, 145. Available from: https://doi.org/10.1186/s12934-017-0760-3

Siaut, M., **Heijde, M.**, **Mangogna, M.**, **Montsant, A.**, **Coesel, S.**, **Allen, A.** ***et al.*** (2007) Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*. *Gene*, **406**, 23–35. Available from: https://doi.org/10.1016/j.gene.2007.05.022

Slattery, S.S., **Diamond, A.**, **Wang, H.**, **Therrien, J.A.**, **Lant, J.T.**, **Jazey, T.** ***et al.*** (2018) An expanded plasmid-based genetic toolbox enables Cas9 genome editing and stable maintenance of synthetic pathways *in Phaeodactylum tricornutum*. *ACS Synthetic Biology*, **7**, 328–338. Available from: https://doi.org/10.1021/acssynbio.7b00191

Slattery, S.S., **Giguere, D.J.**, **Stuckless, E.E.**, **Shrestha, A.**, **Briere, L.-A.K.**, **Galbraith, A.** ***et al.*** (2022) Phosphate-regulated expression of the SARS-CoV-2 receptor-binding domain in the diatom *Phaeodactylum tricornutum* for pandemic diagnostics. *Scientific Reports*, **12**, 7010. Available from: https://doi.org/10.1038/s41598-022-11053-7

Slattery, S.S., **Wang, H.**, **Giguere, D.J.**, **Kocsis, C.**, **Urquhart, B.L.**, **Karas, B.J.** ***et al.*** (2020) Plasmid-based complementation of large deletions in *Phaeodactylum tricornutum* biosynthetic genes generated by Cas9 editing. *Scientific Reports*, **10**, 13879. Available from: https://doi.org/10.1038/s41598-020-70769-6

Sodeinde, O.A. & **Kindle, K.L.** (1993) Homologous recombination in the nuclear genome of *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 9199–9203. Available from: https://doi.org/10.1073/pnas.90.19.9199

Somiari, S., **Glasspool-Malone, J.**, **Drabick, J.J.**, **Gilbert, R.A.**, **Heller, R.**, **Jaroszeski, M.J.** ***et al.*** (2000) Theory and in vivo application of electroporative gene delivery. *Molecular Therapy*, **2**, 178–187. Available from: https://doi.org/10.1006/mthe.2000.0124

Spettel, K., **Barousch, W.**, **Makristathis, A.**, **Zeller, I.**, **Nehr, M.**, **Selitsch, B.** ***et al.*** (2019) Analysis of antifungal resistance genes in *Candida albicans* and *Candida glabrata* using next generation sequencing. *PLoS One*, **14**, e0210397. Available from: https://doi.org/10.1371/journal.pone.0210397

Sprecher, B.N., **Buck, J.M.**, **Ropella, L.L.**, **Ramsperger, A.**, **Kroth, P.G.** & **Yamada, N.** (2023) Genetic transformation methods for diatom *Nitzschia captiva*: new tools to better understand dinotom endosymbiosis. *Algal Research*, **72**, 103136. Available from: https://doi.org/10.1016/j.algal.2023.103136

Stevens, D., **Atteia, A.**, **Franzen, L.-G.** & **Purton, S.** (2001) Cycloheximide resistance conferred by novel mutations in ribosomal protein L41 of *Chlamydomonas reinhardtii*. *Molecular Genetics and Genomics*, **264**, 790–795.

Stewart, M.P., **Langer, R.** & **Jensen, K.F.** (2018) Intracellular delivery by membrane disruption: mechanisms, strategies, and concepts. *Chemical Reviews*, **118**, 7409–7531. Available from: https://doi.org/10.1021/acs.chemrev.7b00678

Strauss, J., **Deng, L.**, **Gao, S.**, **Toseland, A.**, **Bachy, C.**, **Zhang, C.** ***et al.*** (2023) Plastid-localized xanthorhodopsin increases diatom biomass and ecosystem productivity in iron-limited surface oceans. *Nature Microbiology*, **8**, 2050–2066. Available from: https://doi.org/10.1038/s41564-023-01498-5

Stukenberg, D., **Hensel, T.**, **Hoff, J.**, **Daniel, B.**, **Inckemann, R.**, **Tedeschi, J.N.** ***et al.*** (2021) The Marburg collection: a Golden Gate DNA assembly framework for synthetic biology applications in *Vibrio natriegens*. *ACS Synthetic Biology*, **10**, 1904–1919. Available from: https://doi.org/10.1021/acssynbio.1c00126

Stukenberg, D., **Zauner, S.**, **Dell'Aquila, G.** & **Maier, U.G.** (2018) Optimizing CRISPR/Cas9 for the diatom *Phaeodactylum tricornutum*. *Frontiers in Plant Science*, **9**, 740. Available from: https://doi.org/10.3389/fpls.2018.00740

Tanaka, T., **Maeda, Y.**, **Veluchamy, A.**, **Tanaka, M.**, **Abida, H.**, **Marechal, E.** ***et al.*** (2015) Oil accumulation by the oleaginous diatom *Fistulifera solaris* as revealed by the genome and transcriptome. *Plant Cell*, **27**, 162–176. Available from: https://doi.org/10.1105/tpc.114.135194

Taparia, Y., **Dolui, A.K.**, **Boussiba, S.** & **Khozin-Goldberg, I.** (2021) Multiplexed genome editing via an RNA polymerase II promoter-driven sgRNA array in the diatom *Phaeodactylum tricornutum*: insights into the role of StLDP. *Frontiers in Plant Science*, **12**, 784780. Available from: https://doi.org/10.3389/fpls.2021.784780

Taparia, Y., **Zarka, A.**, **Leu, S.**, **Zarivach, R.**, **Boussiba, S.** & **Khozin-Goldberg, I.** (2019) A novel endogenous selection marker for the diatom *Phaeodactylum tricornutum* based on a unique mutation in phytoene desaturase 1. *Scientific Reports*, **9**, 8217. Available from: https://doi.org/10.1038/s41598-019-44710-5

Traller, J.C., **Cokus, S.J.**, **Lopez, D.A.**, **Gaidarenko, O.**, **Smith, S.R.**, **McCrow, J.P.** ***et al.*** (2016) Genome and methylome of the oleaginous diatom *Cyclotella cryptica* reveal genetic flexibility toward a high lipid phenotype. *Biotechnology for Biofuels*, **9**, 258. Available from: https://doi.org/10.1186/s13068-016-0670-3

Trentacoste, E.M., **Shrestha, R.P.**, **Smith, S.R.**, **Gle, C.**, **Hartmann, A.C.**, **Hildebrand, M.** ***et al.*** (2013) Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 19748–19753. Available from: https://doi.org/10.1073/pnas.1309299110

Tsong, T.Y. (1991) Electroporation of cell membranes. *Biophysical Journal*, **60**, 297–306. Available from: https://doi.org/10.1016/S0006-3495(91)82054-9

Turnsek, J., **Brunson, J.K.**, **Viedma, M.**, **Deerinck, T.J.**, **Horak, A.**, **Obornik, M.** ***et al.*** (2021) Proximity proteomics in a marine diatom reveals a putative cell surface-to-chloroplast iron trafficking pathway. *eLife*, **10**, 52770. Available from: https://doi.org/10.7554/eLife.52770

Vasudevan, R., **Gale, G.A.R.**, **Schiavon, A.A.**, **Puzorjov, A.**, **Malin, J.**, **Gillespie, M.D.** ***et al.*** (2019) CyanoGate: a modular cloning suite for engineering cyanobacteria based on the plant MoClo syntax. *Plant Physiology*, **180**, 39–55. Available from: https://doi.org/10.1104/pp.18.01401

Voytas, D.F. (2013) Plant genome engineering with sequence-specific nucleases. *Annual Review of Plant Biology*, **64**, 327–350. Available from: https://doi.org/10.1146/annurev-arplant-042811-105552

Walker, E.J.L. & **Karas, B.J.** (2025) Cloning a chloroplast genome in *Saccharomyces cerevisiae* and *Escherichia coli*. *Bio-Protocol*, **15**, e5162. Available from: https://doi.org/10.21769/BioProtoc.5162

Walker, E.J.L., **Pampuch, M.**, **Chang, N.**, **Cochrane, R.R.** & **Karas, B.J.** (2023) Design and assembly of the 117-kb P*haeodactylum tricornutum* chloroplast genome. *Plant Physiology*, **194**, 2217–2228. Available from: https://doi.org/10.1093/plphys/kiad670

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Walker, E.J.L., **Pampuch, M.**, **Tran, G.** & **Karas, B.J.K.** (2024) Spheroplasted cells: a game changer for DNA delivery to diatoms. *bioRxiv*, 2024.10.10.617634. Available from: https://doi.org/10.1101/2024.10.10.617634

Wang, L., **Sun, Y.**, **Zhang, R.**, **Pan, K.**, **Li, Y.**, **Wang, R.** ***et al.*** (2023) Enhancement of hemostatic properties of *Cyclotella cryptica* frustule through genetic manipulation. *Biotechnology for Biofuels and Bioproducts*, **16**, 136. Available from: https://doi.org/10.1186/s13068-023-02389-x

Weber, E., **Engler, C.**, **Gruetzner, R.**, **Werner, S.** & **Marillonnet, S.** (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One*, **6**, e16765. Available from: https://doi.org/10.1371/journal.pone.0016765

Weyman, P.D., **Beeri, K.**, **Lefebvre, S.C.**, **Rivera, J.**, **McCarthy, J.K.**, **Heuberger, A.L.** ***et al.*** (2015) Inactivation of *Phaeodactylum tricornutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant Biotechnology Journal*, **13**, 460–470. Available from: https://doi.org/10.1111/pbi.12254

Windhagauer, M., **Abbriano, R.M.**, **Ashworth, J.**, **Barolo, L.**, **Jaramillo-Madrid, A.C.**, **Pernice, M.** ***et al.*** (2021) Characterisation of novel regulatory sequences compatible with modular assembly in the diatom *Phaeodactylum tricornutum*. *Algal Research*, **53**, 102159. Available from: https://doi.org/10.1016/j.algal.2020.102159

Xie, W.H., **Zhu, C.C.**, **Zhang, N.S.**, **Li, D.W.**, **Yang, W.D.**, **Liu, J.S.** ***et al.*** (2014) Construction of novel chloroplast expression vector and development of an efficient transformation system for the diatom *Phaeodactylum tricornutum*. *Marine Biotechnology (New York, N.Y.)*, **16**, 538–546. Available from: https://doi.org/10.1007/s10126-014-9570-3

Yin, W. & **Hu, H.** (2021) High-efficiency transformation of a centric diatom *Chaetoceros muelleri* by electroporation with a variety of selectable markers. *Algal Research*, **55**, 102274. Available from: https://doi.org/10.1016/j.algal.2021.102274

Yin, W. & **Hu, H.** (2023) CRISPR/Cas9-mediated genome editing via homologous recombination in a centric diatom *Chaetoceros muelleri*. *ACS Synthetic Biology*, **12**, 1287–1296. Available from: https://doi.org/10.1021/acssynbio.3c00051

Yin, W., **Zhang, Y.**, **Huang, Q.**, **Pan, Y.** & **Hu, H.** (2024) Genetic transformation of the freshwater diatom *Cyclotella meneghiniana* via bacterial conjugation. *Algal Research*, **84**, 103794. Available from: https://doi.org/10.1016/j.algal.2024.103794

Yoshinaga, R., **Niwa-Kubota, M.**, **Matsui, H.** & **Matsuda, Y.** (2014) Characterization of iron-responsive promoters in the marine diatom *Phaeodactylum tricornutum*. *Marine Genomics*, **16**, 55–62. Available from: https://doi.org/10.1016/j.margen.2014.01.005

Zaslavskaia, L.A., **Lippmeier, J.C.**, **Kroth, P.G.**, **Grossman, A.R.** & **Apt, K.E.** (2000) Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *Journal of Phycology*, **36**, 379–386. Available from: https://doi.org/10.1046/j.1529-8817.2000.99164.x

Zhang, C. & **Hu, H.** (2014) High-efficiency nuclear transformation of the diatom *Phaeodactylum tricornutum* by electroporation. *Marine Genomics*, **16**, 63–66. Available from: https://doi.org/10.1016/j.margen.2013.10.003

Zhang, J., **Li, C.**, **Xiong, X.**, **Guo, K.**, **Feng, Y.**, **Jiang, Y.** ***et al.*** (2024) A removable and cosmopolitan dsRNA Toti-like virus causes asymptomatic but productive infection in a model diatom strain. *Algal Research*, **79**, 103491. Available from: https://doi.org/10.1016/j.algal.2024.103491

Zhen, Y., **Zhu, J.**, **Yue, M.** & **Mi, T.** (2024) Impacts of phosphoenolpyruvate carboxylase gene silencing on photosynthetic efficiency and carbon fixation in *Skeletonema costatum*. *Gene*, **933**, 148915. Available from: https://doi.org/10.1016/j.gene.2024.148915

© 2025 The Author(s).

1365313x, 2025, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70102 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [21/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

nature communications

Article

https://doi.org/10.1038/s41467-026-68562-6

Breaking the cell wall for efficient DNA delivery to diatoms

Received: 14 October 2024

Accepted: 12 January 2026

Published online: 22 January 2026

Check for updates

E. J. L. Walker[1], M. Pampuch[1], L. Deng[2], Y. Li[2], G. Tran[1], T. Mock[2] & B. J. Karas[1]

Diatoms produce 20% of the world's fixed organic carbon yet remain underutilized as cell factories due to limited genetic engineering tools. Here, we present optimized electroporation and polyethylene glycol (PEG) transformation methods for the model diatom *Phaeodactylum tricornutum*, enabling delivery of DNA and Cas9 ribonucleoprotein complexes with high efficiency. Transformants are recovered with as little as 1 ng of DNA, and linear or circular episomes as large as 55.6 kb are successfully introduced. The optimized electroporation protocol also reveals an unexpected capability: episomes can be assembled directly in the algal cell through non-homologous or homology-driven repair mechanisms, a process we term *diatom* in vivo *assembly* (DIVA). In addition, the PEG approach is adapted to successfully transform *Thalassiosira pseudonana*, demonstrating the applicability of our methods for engineering other diatom species. These tools could be used to accelerate diatom synthetic biology projects and, therefore, the development of sustainable technologies.

Climate change, population expansion, and soil depletion are among several challenges that existentially threaten current agricultural practices. Due to their comparatively rapid growth and higher photosynthetic efficiency, unicellular algae have been proposed as a sustainable agricultural alternative to traditional crops since the 1950s[1]. Diatoms (Bacillariophyta) are of particular interest for this purpose as they are widely distributed across the world's oceans[2], where they are one of the most photosynthetically productive taxa – accounting for an estimated 20% of total fixed organic carbon annually[3].

Phaeodactylum tricornutum is considered to be one of the most characterized model diatom species, having a fully sequenced telomere-to-telomere reference genome[4] in addition to molecular tools for DNA transformation[5–8], CRISPR-based genome engineering[9], protein expression[10], cloning of the whole organelle genomes[11,12], and more. These characteristics have made *P. tricornutum* an emerging chassis for diatom biotechnology[13] and the subject of a whole-genome synthesis effort[14]. The latter, termed the Pt-syn 1.0 project, has brought together an international consortium of phycologists with the aim of replacing all the native chromosomes in the nuclear, chloroplast, and mitochondrial genomes with synthetic counterparts. To actualize this, and to discover the full potential of diatoms as cell factories, a pipeline for efficiently and rapidly testing different DNA constructs in *P. tricornutum* must be established.

Existing methods for delivering DNA to *P. tricornutum* include biolistic bombardment[8], electroporation[5,7], polyethylene glycol (PEG) transformation[6], and bacterial conjugation[6]. Biolistics was the first method developed for transforming *P. tricornutum* and has been used to integrate non-replicative DNA into the nuclear genome[8,15], with reported efficiencies of 1 to 100 transformants per 10^8 cells. Though reasonably efficient, this method requires the use of expensive equipment (i.e., a biolistic particle delivery system) that many labs do not have access to; comparatively, conjugation, electroporation, and PEG transformation can all be conducted with more standard equipment.

Among these, conjugation is the most efficient transformation method for delivering episomes (4 transformants per 10^4 cells)[6], but also the most mutation prone and laborious to execute[6,9]. Conjugation requires the creation and maintenance of a donor strain of *Escherichia*

[1]Department of Biochemistry, Schulich School of Dentistry and Medicine, London, ON, Canada. [2]School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, UK. e-mail: bkaras@uwo.ca

coli, which must contain a suitable broad-host-range conjugative plasmid in addition to the episome of interest; complications can arise if these two constructs have incompatible replication machinery[16], if the episome is maintained at medium to high copy number[6], or if it contains elements that are toxic when expressed in *E. coli* (e.g., endonucleases)[9]. These issues can lead to episomal rearrangements or loss over time. As a result, donor strains often require repeated reisolation and plasmid sequencing to confirm construct stability – steps that can add days to weeks, and in some cases months, to a workflow.

The PEG and electroporation transformation methods offer a more direct path to episomal delivery, but each has limitations. The delivery of episomes through PEG transformation remains unutilized due to its low efficiency (≤1 transformant per 10^8 cells) and high variability between experiments[6]. Although PEG-mediated transformation has not yet been demonstrated in other diatom species, it is routinely applied in the microalga *Chlamydomonas reinhardtii*, where cell-wall removal followed by PEG-mediated agitation with DNA and glass beads enables efficient nuclear transformation[17]. Electroporation has a higher transformation efficiency (2.6 to 4.5 transformants per 10^5 cells)[5,7], but has not received wide adoption due to variability in efficiency and issues with reproducibility. Thus, we sought to optimize the PEG and electroporation methods for episome delivery to *P. tricornutum* with the hopes of developing a more rapid alternative to bacterial conjugation.

In this paper we describe optimized electroporation and PEG transformation protocols for engineering *P. tricornutum*. These methods are both faster and simpler to execute than bacterial conjugation, as well as more efficient than previously described protocols. We also demonstrate that it is possible to assemble episomes directly in the algal cell through non-homologous repair (NHR) and homology-directed repair (HDR) pathways, opening up avenues for exploring alternative assembly strategies that do not rely on traditional *E. coli* or *S. cerevisiae* cloning. Both protocols are suitable for rapid high throughput testing of several DNA constructs simultaneously and, due to their efficiency with low DNA amounts, could be used for the delivery of commercially synthesized episomes directly into the algal cell. Electroporation can also be used for the efficient delivery of Cas9 ribonucleoprotein (RNP) complexes, enabling DNA-free genetic engineering. The protoplasting techniques we establish in *P. tricornutum* are also used to successfully engineer *Thalassiosira pseudonana* cells via PEG transformation, demonstrating that these methods can be applied to other diatom species. These tools will be valuable as the Pt-syn 1.0 project begins to build and test different synthetic chromosome designs.

Results

Electroporation efficiency increases with enzymatic treatment of cells

Our initial attempts at electroporation using a previously described protocol[18] that was adapted from the methods of Zhang and Hu[5] yielded no transformants. After several failed attempts, we discovered that adjusting the electroporator capacitance from 25 to 50 μF yielded transformants for both episomes and marker cassettes, though efficiency was highly variable between experiments (Supplementary Data 1, Exp. 1 to 81; Supplementary Fig. 1). We used this method to deliver a linear 11 kb PCR-amplified episome, pPtGE31_ΔPtR (Addgene ID: 236260), to assess whether the construct would integrate into the genome, persist as a linear episome, or be converted into a circular episome within the *P. tricornutum* nucleus. We were able to recover DNA from these algal transformants in EPI300 *E. coli*, indicating that the linear episome had been circularized in the algal nucleus following electroporation (Supplementary Fig. 2). Whole-plasmid sequencing of four *E. coli* colonies confirmed that circularization had occurred via fusion of the construct's termini.

When this electroporation protocol was repeated using cells that had reached the late stationary phase of growth, we obtained up to 272 transformants per 10^8 cells with the same 11 kb PCR-amplified episome (Supplementary Data 1, Exp. 82 to 89). This represented a nearly 20-fold increase in efficiency when compared to transformations with early-log phase cultures (Exp. 80 and 81). Stationary phase cultures demonstrated noticeable differences in cell phenotype when compared to early-log phase cultures, with there being a higher abundance protoplasts (i.e., diatoms lacking a cell wall) in the older cultures. *P. tricornutum* cells can appear in three different morphotypes – fusiform, oval, and triradiate – and can switch morphotypes in response to environmental stimuli[19,20]. All morphotypes can generate protoplasts if the cell wall is sufficiently weakened. Ultimately, this observation led us to hypothesize that cell morphotype and cell wall integrity may influence electroporation transformation efficiency. Notably, a spheroplasted cell has only partial cell wall degradation and may not appear phenotypically different from a wild-type cell, whereas a fully protoplasted cell lacks a cell wall entirely and adopts a characteristically spherical morphology – provided it does not burst from osmotic or mechanical stressors.

Cell morphotype cannot be readily controlled in *P. tricornutum* cultures as this species displays a high degree of phenotypic plasticity; however, agar-plated cultures of CCAP 1055/1 tend to have a higher proportion of oval-type cells (~20–30% oval, ~70–80% fusiform) compared to liquid cultures (<10% oval, >90% fusiform). Thus, to test this hypothesis, we spheroplasted early-log phase *P. tricornutum* cells harvested from agar plates. Plate-derived cells were spheroplasted using the enzyme alcalase, a serine protease that has been shown to degrade the proteinaceous components of the *P. tricornutum* cell wall[21]. We observed that 100 μl of alcalase at an activity of approximately 3 Anson units per ml (i.e., 300 mAnson units) was enough to fully protoplast 3×10^8 plate-derived cells when resuspended in 375 mM D-sorbitol (Fig. 1a, b). The protoplasted cells were very fragile and prone to rupture from osmotic pressure or physical damage; treatment of 3×10^8 cells with 10 μl of alcalase (i.e., 30 mAnson units) was enough to protoplast some cells, whilst leaving others with partially degraded cell walls (i.e., spheroplasts; Supplementary Fig. 3).

Plate-derived cells (3×10^8) were treated with 0.3 mAnson units, 3 mAnson units, and 30 mAnson units of alcalase prior to the washing steps of the electroporation protocol. Transformation efficiency increased in all cases (Fig. 1c) when electroporating 500 ng of the same 11 kb PCR-amplified episome as before (Fig. 1d), with the 3 mAnson treatment demonstrating the highest increase in efficiency. Based on this result, we went forward with using 1 mAnson unit of alcalase per 10^8 cells when spheroplasting *P. tricornutum* for electroporation.

Electroporation efficiency of cells grown in liquid and solid media

We tested the electroporation efficiency of *P. tricornutum* cells grown in liquid media and on agar plates, with and without alcalase treatment across five biological replicates (Fig. 1e) using 500 ng of the same PCR-amplified 11-kb episome as before. Though transformation efficiency varied between experiments, plate-derived alcalase-treated cells were significantly more efficient (~1.2 transformants per 10^4 cells) than untreated plate-derived cells (~6.0 transformants per 10^7 cells), untreated liquid-derived cells (~4.2 transformants per 10^7 cells), and treated liquid-derived cells (~1.6 transformants per 10^6 cells, Supplementary Table 1). Statistical significance was determined using a Kruskal-Wallis test ($p = 0.0094$) followed by pairwise Wilcoxon tests with Bonferroni correction. Plate-derived, alcalase-treated cells yielded significantly more colonies than all other groups ($p < 0.05$), whereas no significant differences were observed among the remaining comparisons.

Based on this, all proceeding electroporations were conducted with cells derived from plates and spheroplasted with 1 mAnson unit

https://doi.org/10.1038/s41467-026-68562-6

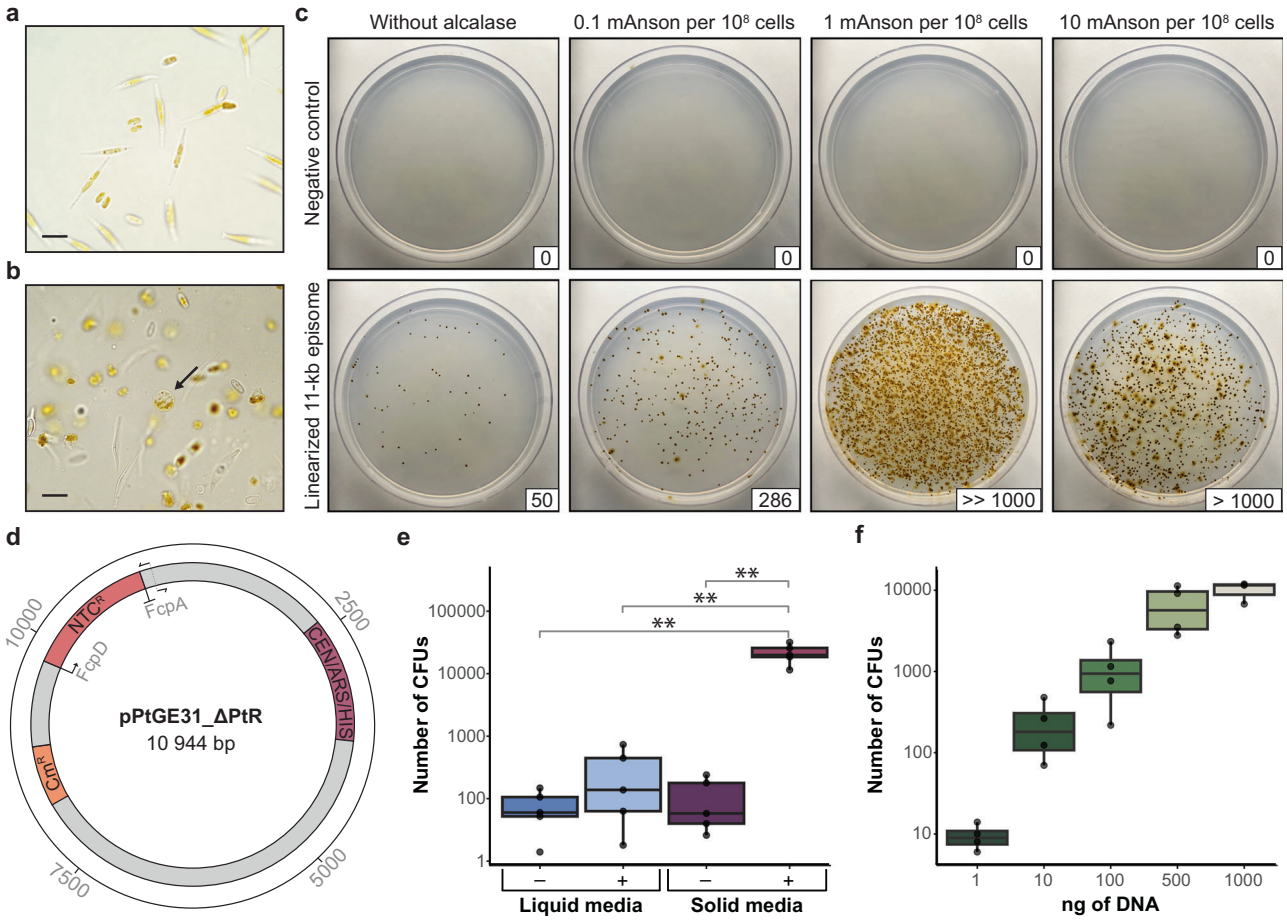


Fig. 1 | Spheroplasting *P. tricornutum* cells increases transformation efficiency. **a** Plate-derived *P. tricornutum* cells pre- and **b** post-treatment with 100 mAnson units of alcalase per 10^8 cells. The arrow indicates a completely protoplasted cell, and the scale bars represent 10 μm. **c** Comparison of agar plate-derived *P. tricornutum* cells treated with differing amounts of alcalase prior to electroporation. Cells were either electroporated with no DNA (negative control) or 500 ng of the 11 kb episome pPtGE31_ΔPtR, which was obtained through PCR amplification. The number of colony-forming units (CFUs) is depicted in the bottom right corner of each transformation plate. Following electroporation, half of the total reaction was plated on ¼-salt L1 plates supplemented with 100 μg/ml nourseothricin. **d** A plasmid map of pPtGE31_ΔPtR, the episome used for transformations in panels **c**, **e**, and **f** Primers depict where the plasmid was linearized during PCR-amplification. This plasmid contains a yeast centromere (CEN) and origin of replication (ARS) sequence, which enables episome replication and stability in *P. tricornutum*[6], as well as a bacterial chloramphenicol resistance marker (Cm^R) and an algal nourseothricin resistance marker (Ntc^R) driven by the FcpD/FcpA promoter-terminator pair[9]. **e** A boxplot demonstrating transformation efficiency for cells cultured in liquid or solid media with (+) or without (-) alcalase treatment (1 mAnson per 10^8 cells) prior to electroporation with 500 ng of PCR amplified pPtGE31_ΔPtR. Five biological replicates were performed. Asterisks denote pairs that are significantly different in terms of efficiency following a two-sided statistical analysis (Kruskal-Wallis test, $p = 0.0094$; Bonferroni-corrected Wilcoxon rank-sum tests: $p = 0.048$ for all comparisons to solid media-treated cells). **f** A boxplot demonstrating the number of *P. tricornutum* CFUs obtained with different nanograms of DNA during electroporation. The cells used in these experiments were derived from agar plates and treated with 1 mAnson of alcalase per 10^8 cells prior to electroporation; four biological replicates were performed. Box plots show the median (centre line), interquartile range (box, 25th–75th percentile), and whiskers extending to 1.5X the interquartile range. Points represent individual biological replicates.

of alcalase per 10^8 cells. This optimized method was compared with the original and adjusted electroporation protocols to demonstrate the impact that changing the capacitance parameter and spheroplasting have on electroporation efficiency (Supplementary Fig. 1).

Transformation efficiency with varying amounts of DNA

Our prior transformations were all conducted with ~500 ng of a PCR-amplified 11-kb plasmid, pPtGE31_ ΔPtR (Fig. 1d). We were interested to see how transformation efficiency varied when using differing amounts of this construct, and most importantly, what the minimum amount of DNA for transformation is. We electroporated as little as 1 ng and up to 1 μg of PCR-amplified pPtGE31_ ΔPtR across four biological replicates (Fig. 1f, data summarized in Supplementary Table 2). In all but one experiment, we were able to recover transformants when using just 1 ng of DNA during electroporation. Increasing the amount of DNA to 10 ng routinely gave rise to hundreds of transformants, with there being an average of 372 CFUs per reaction. Interestingly, there was only a small change in transformation efficiency when the amount of DNA was doubled from 500 ng to 1 μg, suggesting that 500 ng of linear pPtGE31_ ΔPtR was enough to nearly saturate the electroporation reaction.

Transforming linear and circular episomes to the algal nucleus

Our past explorations had relied upon the use of linear PCR-amplified DNA with blunt, non-phosphorylated termini. To test the efficiency of different structural forms of DNA, we prepared other variations of pPtGE31_ ΔPtR (Fig. 2a). The termini of the linear constructs varied in position along the episome and/or the properties of the termini (i.e., blunt or sticky ends, with or without 5′ phosphorylation). These constructs were adjusted to be equimolar prior to conducting electroporation in *P. tricornutum*. We found that there were nearly double the amount of CFUs when transforming any structural form of linear DNA compared to the circular episome (Fig. 2b).

https://doi.org/10.1038/s41467-026-68562-6

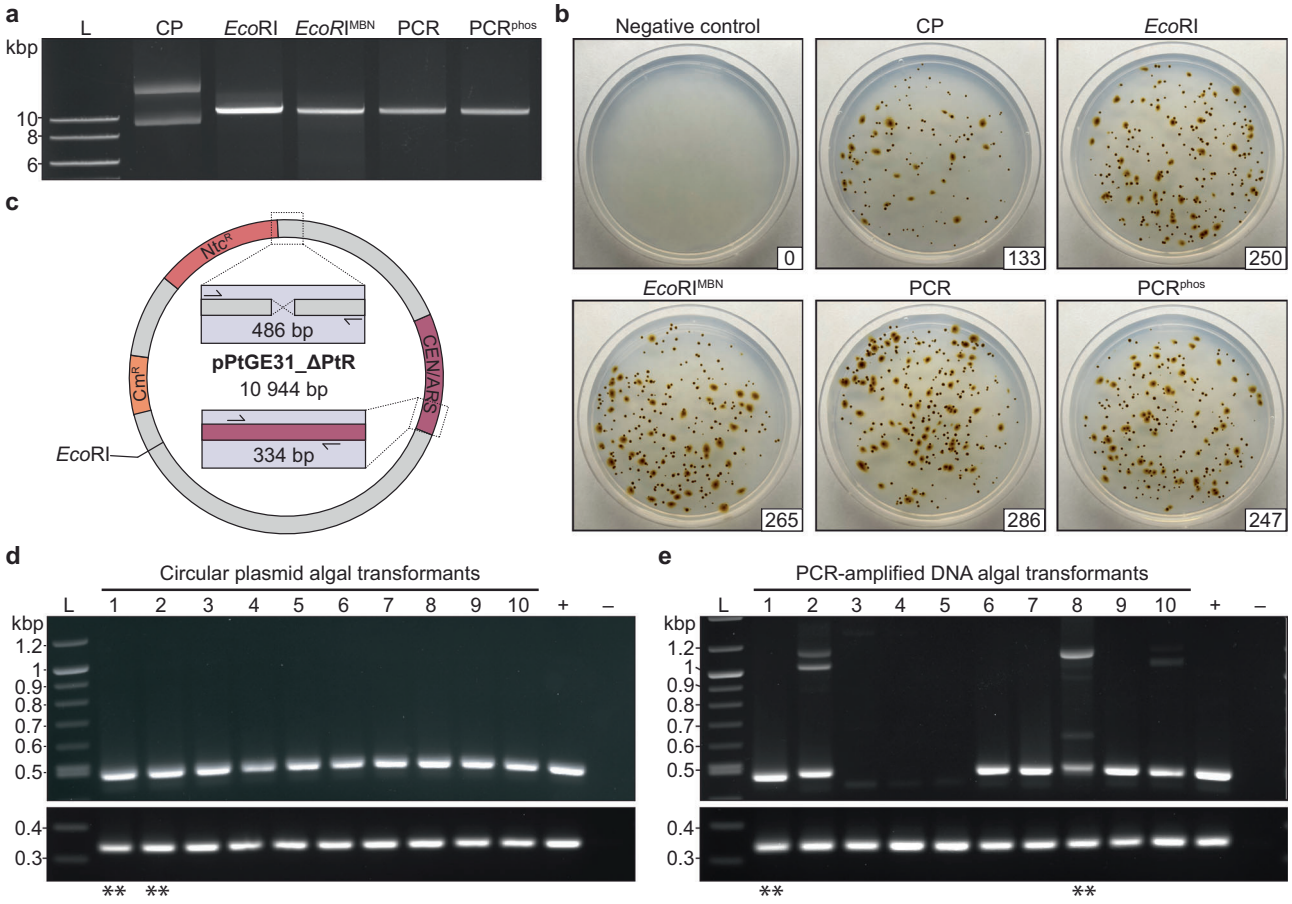


Fig. 2 | Comparing the transformation efficiency of linear, circularized, and phosphorylated DNA. a An agarose gel depicting different preparations of the episome pPtGE31_ ΔPtR. CP: circular plasmid, *Eco*RI: restriction enzyme digested plasmid, *Eco*RI^{MBN}: restriction enzyme digested plasmid with additional mung bean nuclease treatment to remove sticky ends, PCR: PCR-amplified linear plasmid DNA, and PCR^{phos}: PCR-amplified linear plasmid DNA generated with phosphorylated primers. **b** Comparison of *P. tricornutum* transformants electroporated with different preparations of pPtGE31_ ΔPtR. Following electroporation, one-tenth of the total reaction was plated on ¼-salt L1 plates supplemented with 100 μg/ml nourseothricin. Plates are labeled according to the type of DNA that was transformed, as described above. The number of CFUs is shown in the bottom right corner of each plate. **c** A plasmid map depicting the regions that were screened when assessing passaged algal transformants. The 486 bp junction spans the region where pPtGE31_ ΔPtR was split for PCR amplification, whereas the 334 bp junction spans an intact region of the plasmid backbone. **d** Screening 10 algal colonies that had been transformed with circular pPtGE31_ ΔPtR or **e** PCR-amplified pPtGE31_ ΔPtR. The primer pairs span the termini region (expected size of 463 bp, top panel) and the CEN/ARS region of the episome backbone (expected size of 334 bp, bottom panel). Primer pairs are listed in Supplementary Data 2. Dilute pPtGE31_ ΔPtR and genomic *P. tricornutum* DNA were used for the positive and negative controls, respectively. Double asterisks (**) indicate the algal episomes that would be recovered in *E. coli* and sent for whole plasmid sequencing.

We had also demonstrated that electroporated linear pPtGE31_ ΔPtR gets circularized in the *P. tricornutum* nucleus, but we had not yet performed comprehensive screening to determine if any indels were accruing during this repair process. To explore what occurs during termini fusion, ten colonies were repatched from the transformations with circular and PCR-amplified pPtGE31_ ΔPtR for additional screening. Colonies were screened using a set of primers that spans where the termini are expected to recombine for the PCR-amplified construct, as well as a region of the episome backbone (Fig. 2c). Colonies transformed with circular DNA did not demonstrate any unexpected insertions or deletions in the screened regions (Fig. 2d), as expected. Whole plasmid sequencing of circular colonies 1 and 2 confirmed that the episomes had not accrued any mutations or indels during transformation.

Colonies transformed with PCR-amplified DNA appeared to have indels in the region where termini fusion occurs but not in the intact backbone region (Fig. 2e). We conducted Sanger sequencing of the termini region for colonies 1, 2, and 6 to 10, which confirmed that indels of varying sizes had accrued in this region for every colony screened (Supplementary Fig. 4). The absence of the 486 bp band in lanes 3, 4, and 5 is most likely due to DNA degradation or deletions occurring at the episomal termini, either before (e.g., exonuclease activity) or during the circularization process. Whole plasmid sequencing of linear colonies 1 and 8 further confirmed the presence of indels in the termini region where the episome was being fused together, whereas all other regions of the episome had remained unchanged. Interestingly, the ~500-bp insertions in colonies 2, 8, and 10 were not complementary to any regions of the episome or algal genome. When these insertions were queried using BLASTn, they aligned to various regions of the salmon and/or trout genome accessions.

Single-stranded salmon sperm (ssss) DNA is used as a carrier during electroporation, and to our surprise, we had discovered that it could be captured during NHR of the episome termini. We attempted electroporation without ssssDNA to see if this would impact transformation efficiency, and it resulted in a 15-fold reduction in the number of CFUs (Supplementary Fig. 5), highlighting its importance as an additive for efficient transformation.

https://doi.org/10.1038/s41467-026-68562-6

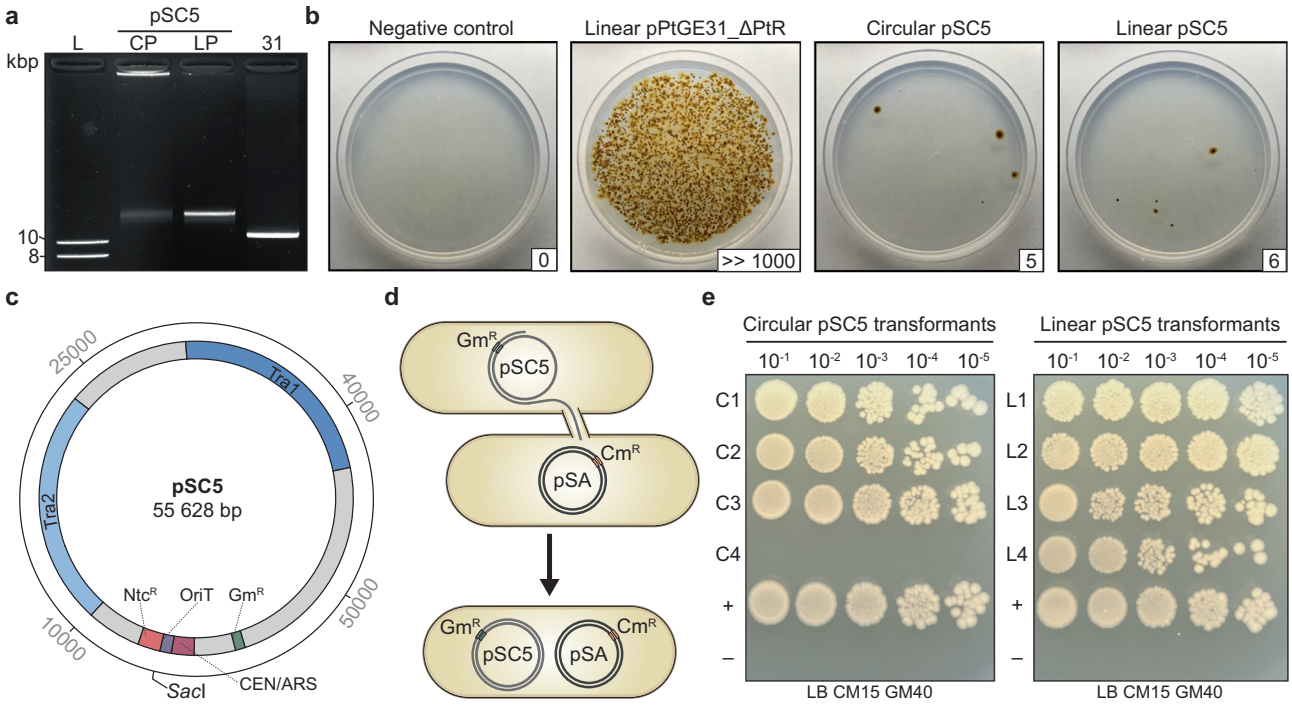


Fig. 3 | Electroporation and recovery of a 55.6 kb conjugative plasmid. a An agarose gel depicting the circular pSC5 (CP), linearized pSC5 (LP), and PCR-amplified pPtGE31_ΔPtR (31). **b** Electroporation of circular or linear variants of pSC5 and PCR-amplified pPtGE31_ ΔPtR into *P. tricornutum* cells. Following electroporation, half of the total reaction was plated on ¼-salt L1 plates supplemented with 100 μg/ml nourseothricin. Plates are labeled according to the type of DNA that was transformed, as described above. The number of CFUs is shown in the bottom right corner of each plate. **c** A plasmid map of pSC5. The transfer regions (Tra1 and Tra2) encode essential conjugative proteins. The origin of transfer (OriT) serves as a recognition sequence for the conjugative protein apparatus, facilitating the transfer of pSC5 during conjugation. The plasmid contains selection markers for bacteria (Gm^R, gentamycin resistance) and *P. tricornutum* (Ntc^R, nourseothricin resistance) as well as a yeast centromere (CEN) and replication origin (ARS), which enable episome stability and replication in *P. tricornutum*. **d** A donor strain of *E. coli* harboring pSC5, which confers gentamycin resistance, conjugates to a recipient strain of *E. coli* harboring pSAP, which confers chloramphenicol resistance. Successful conjugation leads to the creation of a transconjugant strain that is gentamycin and chloramphenicol-resistant. **e** Testing the conjugation efficiency of *E. coli* strains harboring episomes that had been recovered from *P. tricornutum* transformants electroporated with circular (C1 to C4) or linear (L1 to L4) pSC5. Strains harboring pSC5 or pSAP were used as the positive and negative controls, respectively. Dilutions of 10^{-1} to 10^{-5} were spot plated on LB plates supplemented with 15 μg/ml chloramphenicol and 40 μg/ml gentamicin.

Evaluation of genomic integration

We performed a plasmid/episome loss experiment by serially passaging four pPtGE31_ΔPtR transformants in non-selective media to check for genomic integration. The transformants chosen correspond to algal colonies 1 and 2 in Figs. 2d and 1 and 8 in Fig. 2e. After seven passages in liquid media with or without nourseothricin selection, serial dilutions of the cultures were then plated onto solid non-selective media to obtain single colonies. For each transformant, 100 single colonies were then re-patched onto plates supplemented with nourseothricin. If genomic integration of the episome had occurred, we would expect a high proportion of the single colonies to retain resistance. Conversely, if the episomes are maintained extra chromosomally, they should be lost over successive generations without selective pressure, and thus only a low proportion of colonies would retain resistance. As expected, only 2–7% of screened colonies retained resistance (Supplementary Table 3), demonstrating that genomic integration had not occurred in the assayed transformants.

Electroporation and recovery of a large plasmid

It has been demonstrated in other organisms that as episome size increases, electroporation efficiency decreases[22,23]. We sought to test the limits of electroporation in *P. tricornutum* by attempting transformation with pSC5, a 55.6 kb broad-host-range conjugative episome (Fig. 3c, Addgene ID: 188602)[24].

Equimolar amounts of circular and linear pSC5 were prepared along with PCR-amplified pPtGE31_ ΔPtR (Fig. 3a). Electroporation with either form of pSC5 gave rise to transformants, though the number of CFUs was orders of magnitude lower than electroporation with a linear 11 kb episome (Fig. 3b), as anticipated. To determine if the complete 55.6 kb episome had been successfully electroporated into *P. tricornutum*, we passaged four colonies from each of the transformations with circular and linear pSC5. DNA was isolated from the algal colonies and electroporated into *E. coli*. Here, if the *E. coli* transformants contain an intact version of pSC5, the episome will facilitate conjugation and self-mobilization into a recipient strain harboring pSAP[12] (Addgene ID: 206429), a 10 kb plasmid that contains a chloramphenicol resistance marker (Fig. 3d). Successful transconjugants will harbor pSC5 and pSAP, therefore carrying resistance to both gentamycin and chloramphenicol.

We performed conjugation using a pool of *E. coli* colonies for each transformation (Fig. 3e; C1–C4 are derived from circular pSC5 algal transformants, L1 to L4 are derived from linear pSC5 algal transformants). Conjugation was successful for three out of four (circular pSC5) and four out of four (linear pSC5) of the *E. coli* pools tested, demonstrating that intact pSC5 had been successfully electroporated and maintained in the majority of screened *P. tricornutum* transformants (Fig. 3e).

Electroporation and recombination of multiple DNA fragments

Assembly via a non-homologous repair pathway. The surprising discovery of salmon DNA inserted into the termini region of PCR-amplified pPtGE31_ΔPtR led us to hypothesize that *P. tricornutum* could assemble multiple DNA fragments into a single episome through NHR. To test this, we created an episome, pPtGE31_ShBle (Addgene ID:

https://doi.org/10.1038/s41467-026-68562-6

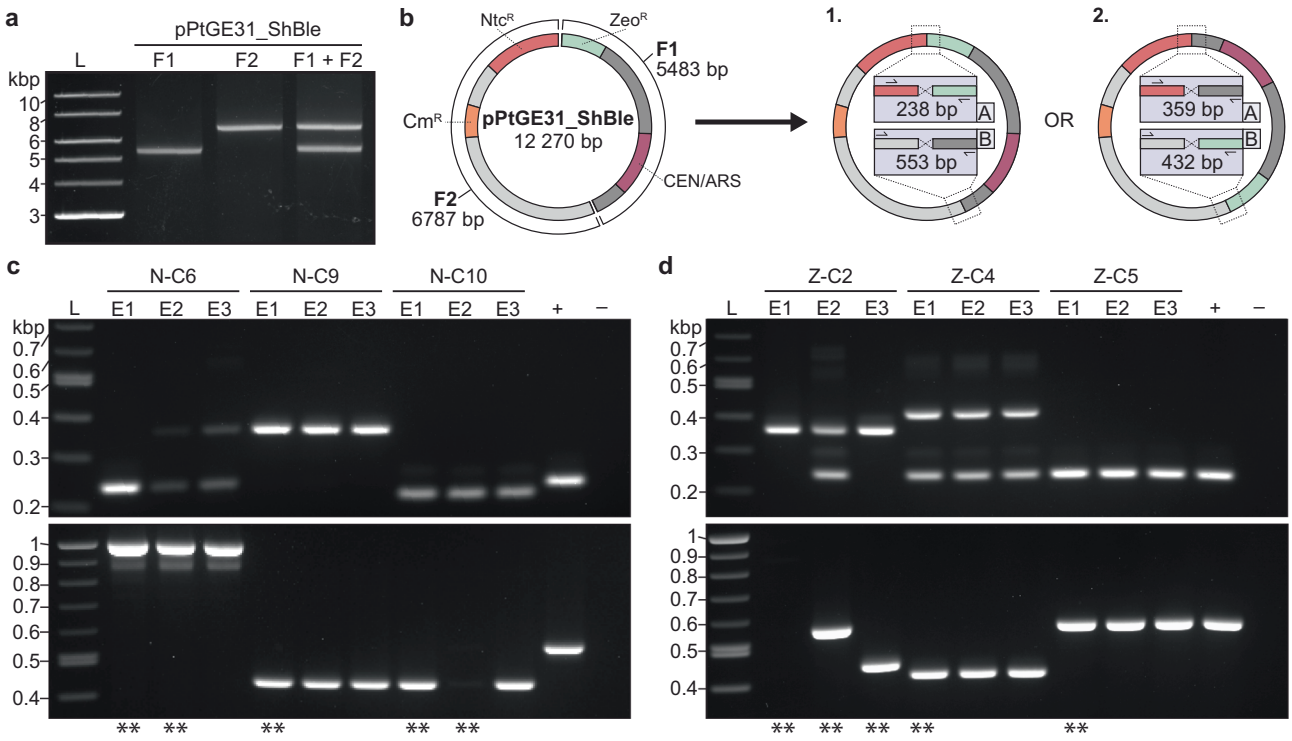


Fig. 4 | Electroporation and assembly of two non-overlapping fragments in *P. tricornutum*. **a** The episome pPtGE31_ShBle was amplified as two fragments (F1: fragment 1, F2: fragment 2). The fragments were adjusted to be equimolar before combining them into an assembly mixture (F1 + F2). **b** A plasmid map of pPtGE31_ShBle and a depiction of the different possible orientations post-assembly. The episome contains algal selection markers for Ntc^R and Zeo^R, a bacterial Cm^R selection marker, as well as a yeast CEN and ARS sequences. F1 and F2 do not share any overlapping sequences and thus can assemble in two different orientations. Primer sets were designed to span the two junctions, A and B, where recombination is expected to occur. The amplicons will vary in size depending on the orientation of the fragments. **c** Screening three *E. coli* transformants per algal colony using primers spanning junctions A (top panel, expected size of 238 or 359 bp) and B (bottom panel, expected size of 553 or 432 bp). DNA was isolated from algal transformants that were electroporated with F1 + F2 and selected initially with nourseothricin or **d** zeocin. Dilute pPtGE31_ ShBle and water were used for the positive and negative controls, respectively. Double asterisks (**) indicate the *E. coli* colonies that were sent for whole plasmid sequencing.

236261), that contains the NAT and ShBle markers conferring resistance to nourseothricin and zeocin, respectively. The episome pPtGE31_ShBle was then amplified as two non-overlapping fragments, each containing one of the algal selective markers. The fragments were adjusted to be equimolar prior to electroporation into *P. tricornutum* (Fig. 4a).

Cells transformed with both fragments simultaneously were plated on nourseothricin- or zeocin-containing plates initially (Supplementary Fig. 6). Out of the 116 colonies passaged post-electroporation, 94 demonstrated double resistance to both nourseothricin and zeocin (Supplementary Table 4). Twenty transformants, ten derived from each initial transformation plate, were passaged onto double-selection media, then DNA isolation and PCR screens were performed. These colonies are denoted with an N- or Z- to depict which initial selection plate they were derived from (N: nourseothricin, Z: zeocin).

Two orientations are expected when the fragments amplified from pPtGE31_ShBle are re-assembled (Fig. 4b). We designed a PCR screen that would not only assess if both fragments had been recombined, but could also determine the orientation of the fragments. When screening algal transformants, we only used the primer set spanning junction B as junction A contains primers that bind to the FcpD promoter and FcpA terminator, which are present in both the episome and algal genome. Of the twenty colonies screened, nine had amplification across junction B, with only five colonies demonstrating the expected amplicon sizes (Supplementary Fig. 7). The other four colonies either demonstrated multiple amplicons, some of which were the expected sizes, and/or amplicons that were larger than expected. Based on this initial screen, DNA from colonies N-C6, N-C9, N-C10, Z-C2, Z-C4, and Z-C5 was transformed into *E. coli* for further screening.

Three *E. coli* transformants were assessed per algal colony using the primer sets spanning junctions A and B (Fig. 4c, d). Interestingly, the *E. coli* transformants derived from algal colonies N-C6, N-C10, and Z-C2 demonstrated different banding patterns, suggesting that these algal colonies likely had a pool of assembled episomes. All *E. coli* transformants from algal colony N-C9 depicted the expected banding pattern for orientation 2, whereas all *E. coli* transformants from algal colony Z-C5 depicted the expected banding pattern for orientation 1. DNA from *E. coli* colonies N-C6-E1/E2, N-C9-E1, N-C10-E1/E2, Z-C2-E1/E2/E2, Z-C4-E1, and Z-C5-E1 were sent for whole-plasmid sequencing to validate if assembly had occurred.

The sequencing results demonstrated that recombination of fragments 1 and 2 had occurred in all assessed *P. tricornutum* transformants (Supplementary Fig. 8). In the majority of assembled episomes, multiple partial copies of fragments 1 and 2 had recombined. The most excessive recombination appeared *E. coli* colony Z-C2-E2, which harbored a 41.5 kb episome containing eight partial or whole copies of both fragments 1 and 2. Only *E. coli* colonies N-C9-E1, N-C10-E1, Z-C2-E3, and Z-C5-E1 contained the expected assembled episomes. Sequencing of *E. coli* transformants derived from the same algal colony revealed that algal colonies N-C6, N-C10, and Z-C2 possessed a pool of assembled episomes, as we had hypothesized.

Assembly via a homology-directed repair pathway. Though it is possible to assemble episomes via the NHR pathway, screening and sequencing demonstrated that the majority of *P. tricornutum*

https://doi.org/10.1038/s41467-026-68562-6

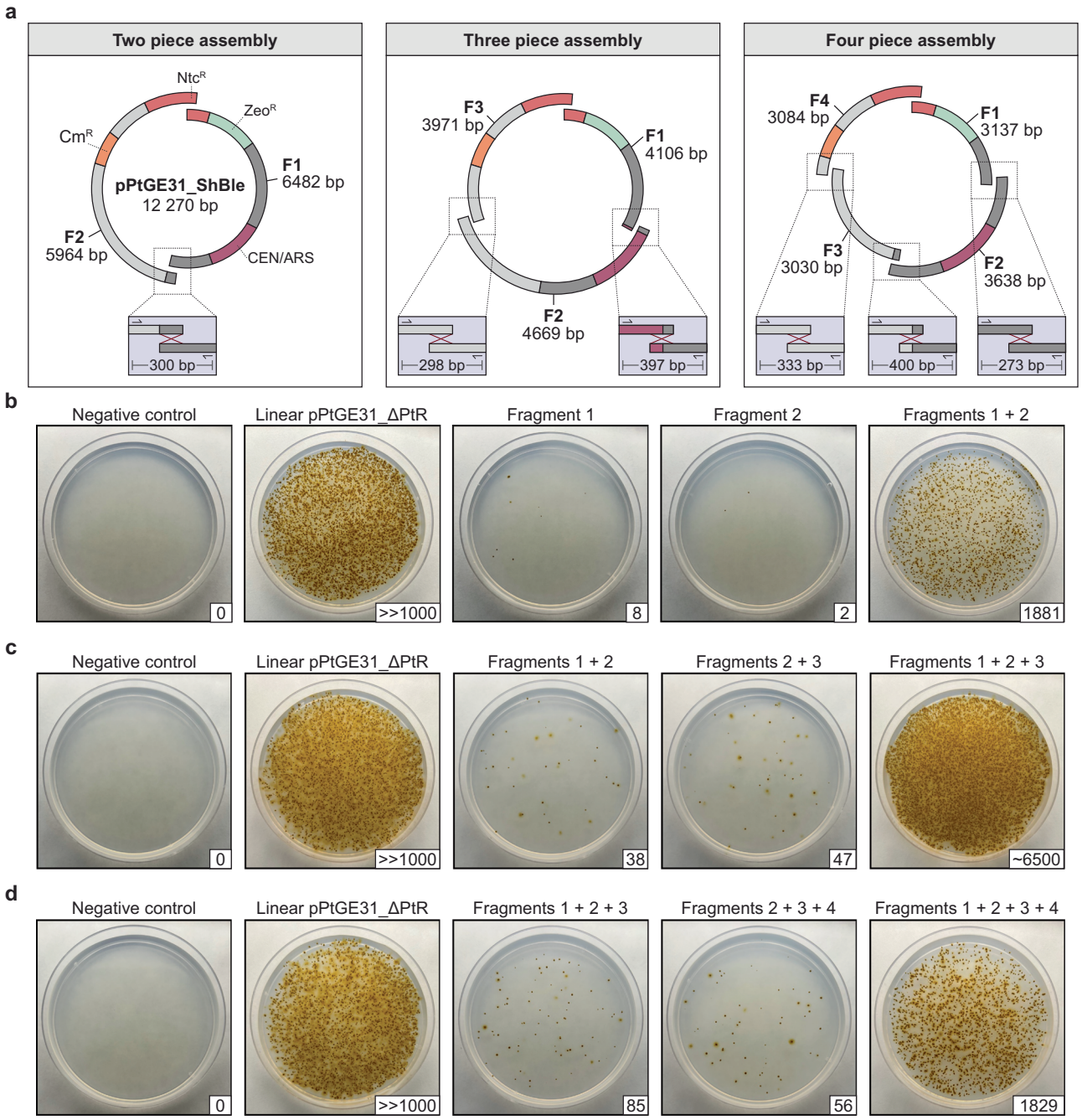


Fig. 5 | Electroporation and assembly of overlapping fragments in *P. tricornutum*. **a** The episome pPtGE31_ShBle was amplified as two, three, or four overlapping fragments (F1: fragment 1, F2: fragment 2, F3: fragment 3, F4: fragment 4). The fragments were adjusted to be equimolar before combining them into their respective assembly mixtures. Primers were designed to screen the other overlapping junction(s), as demonstrated in the bottom boxes. The episome contains a bacterial Cm^R marker, the algal Ntc^R/Zeo^R markers, and a yeast CEN/ARS. Electroporation of **b** two, **c** three, or **d** four overlapping fragments. For each condition, partial fragment combinations that cannot reconstitute the Ntc^R marker were also electroporated to assess template episome carry-over. Linear pPtGE31_ΔPtR (500 ng) was included as a positive control. Following electroporation, half of each reaction was plated on ¼-salt L1 medium supplemented with 100 μg/ml nourseothricin. Colony-forming units (CFUs) are indicated in the bottom right of each plate.

transformants contained undesirable assemblies of the non-overlapping fragments. To be able to reliably assemble episomes in the algal nucleus, an alternative strategy is needed. We sought to test whether overlapping fragments could be simultaneously transformed and assembled through the HDR repair pathway. It was hypothesized that the overlapping sequences between fragments would enable site-specific recombination at the termini, thereby creating episomes of an expected size and directionality.

To test this hypothesis, we split the episome pPtGE31_ShBle into two, three, or four fragments that overlapped with each other by ~160 bp (Fig. 5a). This length was chosen somewhat arbitrarily as a starting point, as we were uncertain about the minimum or maximum overlap size needed for successful HDR. Fragments were generated by PCR amplification using PtGE31_ΔPtR (Addgene ID: 236260) or pPtGE31_ShBle (Addgene ID: 236261) as template DNA (primers listed in Supplementary Data 2). In all assembly designs, one of the overlaps was positioned within the open reading frame (ORF) for

nourseothricin N-acetyl transferase (NAT), which confers nourseothricin resistance. Thus, when the assembly fragments are electroporated, nourseothricin-resistant transformants should only appear if homologous recombination has occurred within the NAT ORF or if there is carry-over of the PCR template DNA. Primers were designed to screen the other overlapping termini for successful recombination (Fig. 5a, bottom boxes).

We first tested the two-piece HDR assembly by electroporating fragments 1 and 2 individually or simultaneously (Supplementary Fig. 5b). When electroporated individually, fragments 1 and 2 gave rise to 8 and 2 CFUs, respectively, indicating that carry-over of the PCR template DNA had occurred. However, when the fragments were electroporated simultaneously, there were more than 1800 CFUs – far more than what would be anticipated by carry-over alone. Ten algal colonies were selected for screening of the other fragment junction (Supplementary Fig. 9a), all of which demonstrated the expected 300 bp band, suggesting that successful homologous recombination had occurred. Episomes from algal colonies C1, C2, and C3 were recovered in *E. coli*. Three *E. coli* colonies were screened per initial algal colony, all of which demonstrated the correctly sized band (Supplementary Fig. 9b). To confirm if the fragments had been correctly recombined, episomes from *E. coli* colonies C1-E1, C2-E1, and C3-E1 were sent for whole plasmid sequencing. This demonstrated that colonies C1-E1 and C3-E1 contained the correctly assembled episome (Supplementary Fig. 10), but the episome recovered from C2-E1 had two partial copies of fragment 1.

The same methodological approach was used to electroporate three overlapping fragments simultaneously (Fig. 5c). Electroporation of fragments 1 and 2 or 2 and 3 gave rise to 38 and 47 CFUs respectively, whereas electroporation of all three fragments gave rise to ~6500 CFUs. Ten algal colonies were screened for the junctions between F1/F2 or F2/F3. All colonies contained the expected bands except for C1 and C2, which lacked an amplicon for the F1/F2 junction and demonstrated a weak amplicon for the F2/F3 junction (Supplementary Fig. 9c). We anticipated that assembly may have failed in these colonies, but to further investigate this, we attempted to recover the algal episomes from C1 and C2 in *E. coli*, along with the positively screened episomes from C3 and C4. We were able to recover *E. coli* transformants for all the tested episomes and decided to screen two CFUs per initial algal colony, all of which demonstrated correctly sized amplicons for both junctions (Supplementary Fig. 9d). Episomes isolated from *E. coli* colonies C1-E1, C2-E1, C3-E1, and C4-E1 were sent for whole plasmid sequencing, with only C1-E1 deviating from the expected assembly (Supplementary Fig. 10).

To further test the limits of the HDR assembly pathway, we attempted electroporation of four overlapping fragments simultaneously. Electroporation of only three of the four fragments (i.e., F1 + F2 + F3 or F2 + F3 + F4) gave rise to 85 and 56 transformants, demonstrating the carry-over of template DNA once again (Fig. 5d). However, when all four fragments were combined and transformed simultaneously, there was a > 20-fold increase in the number of CFUs. Ten algal colonies were screened across the remaining three junctions, and all but C1 and C7 demonstrated the expected amplicons (Supplementary Fig. 9e). Colonies C1 and C7 specifically lacked the 273 bp amplicon for the junction between F1/F2. Algal episomes from C1, C2, C3, and C7 were transformed into *E. coli*; however, C7 did not give rise to any bacterial transformants, suggesting that incomplete assembly had occurred. Two *E. coli* colonies were screened for the episomes derived from C1, C2, and C3, all of which demonstrated the expected amplicons for the three junctions (Supplementary Fig. 9f). Whole plasmid sequencing of C1-E1, C2-E1, and C3-E1 revealed that correct assembly had occurred in algal colonies C1 and C3 (Supplementary Fig. 10), but C2 contained multiple partial copies of fragments 1 and 4.

In total, 26 of the 30 algal colonies screened across the various assemblies demonstrated that recombination had successfully occurred across the tested junctions (Supplementary Fig. 10). From the episomes recovered in *E. coli*, 7 out of 10 had the correctly assembled episomes when fully sequenced. This demonstrates that it is both feasible and reliable to assemble episomes through the HDR pathway in *P. tricornutum*.

Exploring the limits and applications of HDR assembly

The discovery of DIVA prompted us to further investigate the practical boundaries and potential applications of this assembly method. We first sought to determine the size of overlaps required to direct HDR. Our earlier assemblies relied on ~160 bp overlaps, but in *Saccharomyces cerevisiae*, as little as 20 bp is sufficient for recombination[25]. To test whether this was also true in *P. tricornutum*, we PCR-amplified NAT using primers that generated 0, 20, or 40 bp of overlap with the FcpD and FcpA promoter/terminator pair (Supplementary Fig. 11a). We then PCR-amplified pPtGE30 (Addgene ID: 236256) in two fragments, positioning the termini such that the NAT constructs would assemble into the *Sh Ble* locus, which uses the same promoter/terminator elements.

The NAT constructs were electroporated both individually and with the episomal fragments (Supplementary Fig. 11b). Electroporation of NAT-0, NAT-20, and NAT-40 alone confirmed that the marker with up to 40 bp of promoter/terminator sequences does not confer nourseothricin resistance. The NAT-0 assembly produced no transformants, whereas both the NAT-20 and NAT-40 assemblies yielded ~100 transformants. These results indicate that overlaps as short as 20 bp are sufficient for HDR-mediated assembly.

We next examined whether de novo synthesized constructs could be inserted directly into episomes in *P. tricornutum*. We commercially synthesized two constructs: (i) CytB-eGFP (833 bp), which carries the CytB mitochondrial targeting peptide identified by TargetP (2.0)[26], and (ii) eGFP-T2A-$NAT_{2\text{-}63}$ (1025 bp), in which eGFP is linked to residues 2 to 63 of the NAT protein via the T2A ribosomal skip sequence[27]. The CytB-eGFP construct was designed to insert into the *Sh Ble* locus of pPtGE31_ShBle (Supplementary Fig. S11c), whereas the eGFP-T2A-$NAT_{2\text{-}63}$ construct was designed to insert into a PCR-amplified version of pPtGE31_ΔPtR containing an incomplete NAT cassette (residues 47–188; Supplementary Fig. 11d). In both designs, the synthesized constructs included 50 bp of overlaps at each terminus to their respective insertion sites.

For the eGFP-T2A-NAT design, nourseothricin-resistant colonies should only arise if correct assembly restores the complete NAT gene or if there is unintended carry-over of the template episome. This reduces the likelihood of false positive transformants. Both de novo assembly designs yielded at least one hundred colonies (Supplementary Fig. 11e, f).

Ten nourseothricin-resistant colonies from each assembly were passaged and screened for GFP expression by confocal fluorescence microscopy. Of the CytB-eGFP transformants, 4/10 showed moderate mitochondrial eGFP localization, while all 10/10 of the eGFP-T2A-NAT transformants exhibited strong cytosolic eGFP expression (representative images in Supplementary Fig. 12). Three colonies per design were subsequently recovered in *E. coli* and sequenced (C3, C5, C7 for CytB-eGFP; C1, C2, C3 for eGFP-T2A-NAT). All recovered episomes contained the corresponding de novo construct; however, only one CytB-eGFP episome (C5-E1) and one eGFP-T2A-NAT episome (C1-E1) displayed the expected assembly configuration. The remaining episomes showed unexpected recombination events, including combinations of multiple episomal fragments and, in some cases, insertion of multiple copies of the de novo construct or unaligned DNA.

Delivery of RNPs via the optimized electroporation protocol

RNP complexes can be delivered to *P. tricornutum* through biolistic bombardment[28], providing a means for DNA-free genome engineering. We anticipated that it would also be possible to electroporate RNPs

https://doi.org/10.1038/s41467-026-68562-6

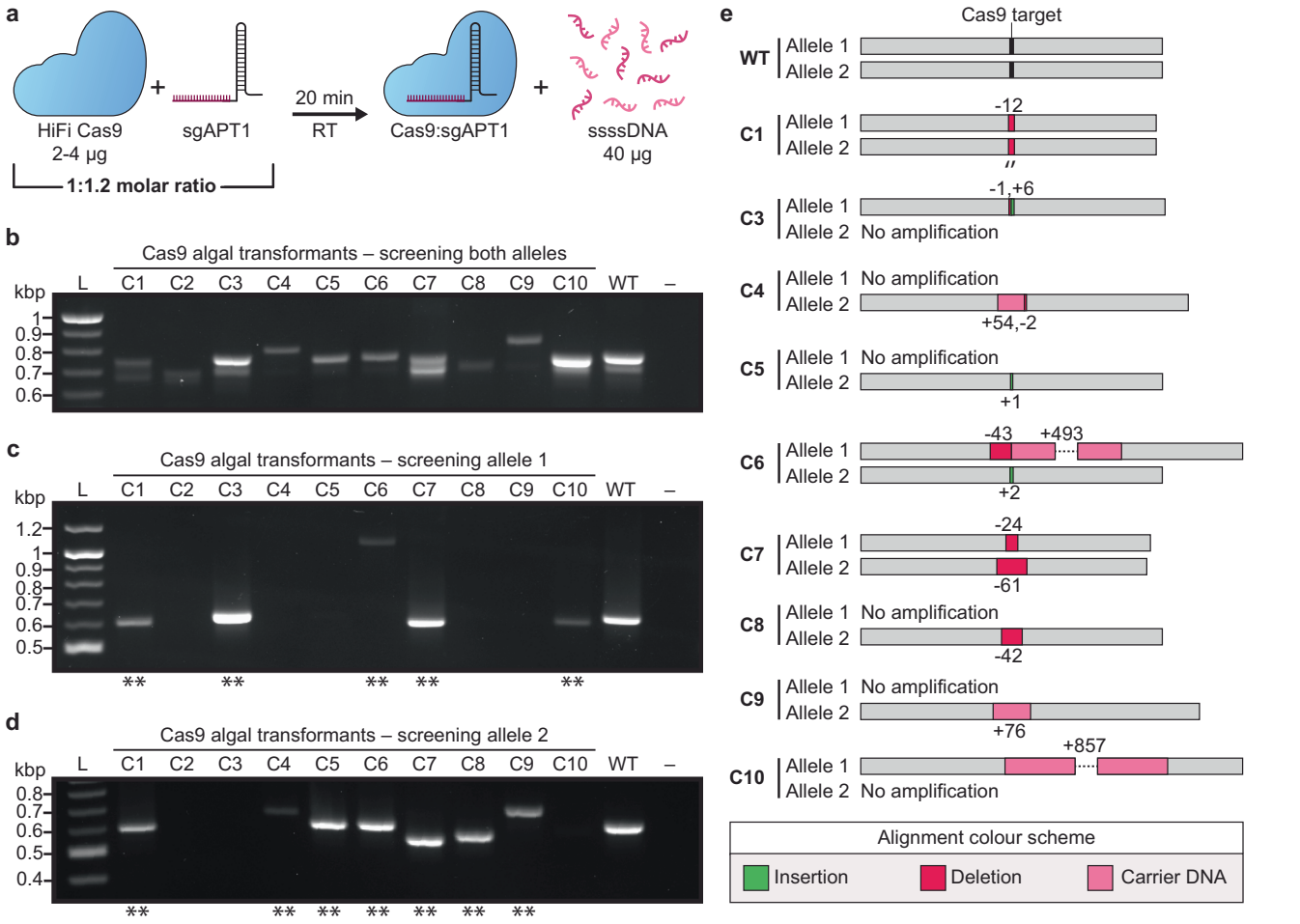


Fig. 6 | Electroporation of a Cas9:sgAPT1 complex into *P. tricornutum* for DNA-free engineering. a The HiFi Cas9 complex is combined with an sgRNA that targets the *PtAPT* locus (sgAPT1). After twenty minutes at room temperature, the complex is combined with 40 µg of carrier DNA (ssssDNA) prior to electroporation. **b** Screening ten 2-FA resistant algal transformants with primers that bind to both *PtAPT* alleles, **c** allele 1, and **d** allele 2. The positive and negative controls consist of wild-type *P. tricornutum* DNA and water, respectively, for all PCR screens. Double asterisks (**) indicate the amplicons that were sent for Sanger sequencing. **e** Alignment of the Sanger sequences captured for algal colonies C1 and C3-C10.

into this species, but this had yet to be reported in any diatoms. As a proof-of-concept, we replicated the methodology Serif et al. had used to deliver RNPs[28] using our optimized electroporation protocol in lieu of biolistics.

Serif et al. had previously shown that knock-out of the *P. tricornutum Adenine Phosphoribosyl Transferase* (*PtAPT*) gene confers resistance to the cytotoxin 2-fluoroadenine (2-FA) at a dose of 10 µM[28]. PtAPT functions in the adenine salvage pathway, catalysing the transfer of phosphoribosyl pyrophosphate (PRPP) to adenine to form adenosine monophosphate (AMP); in the presence of 2-FA, PtAPT instead converts 2-FA into 2-fluoro-AMP, which is subsequently phosphorylated to form toxic fluorinated nucleotides[28].

Chemicals can vary based on the manufacturer and/or batch; thus, when setting forth on this series of experiments, we first tested the resistance of wild-type *P. tricornutum* (CCAP 1055/1) cultured on ½-salt L1 plates supplemented with adenine (5 µg/ml) and our batch of 2-FA (0, 2.5, 5, 10, 15, 20 µM). To our surprise, there was an absence of growth on the 0 µM 2-FA plate, suggesting that adenine alone posed a toxicity to wild-type *P. tricornutum* cells. We tested this theory by plating wild-type cells onto ½-salt L1 plates supplemented with varying levels of adenine (0, 1, 2.5, 5 µg/ml; Figure S13). The growth of wild-type algae was negligible after one-week on the 5 µg/ml adenine plate, but relatively unimpacted on the 0.5 and 1 µg/ml plates. This interesting revelation led us to hypothesize that adenine supplementation could negatively impact Δ*PtAPT* transformants as well, so when conducting our RNP electroporation experiments, we plated transformants on 2-FA plates with and without adenine supplementation.

The *PtAPT* locus was targeted using two guide RNAs (gRNAs), each of which allowed for biallelic knock-out; we opted to use the first gRNA (gAPT1) for our proof-of-concept experiments as it had generated 11 of the 16 2-FA-resistant transformants in the singleplex experiments by Serif et al.[28]. HiFi Cas9 and a single guide RNA for APT1 (sgAPT1) were complexed according to the manufacturer's guidelines (Fig. 6a). Then, the RNP complex was combined with ssssDNA and electroporated into spheroplasted *P. tricornutum* cells. We opted to include carrier DNA when performing the RNP experiments as ssssDNA is thought to bind to and neutralize the positive charge of the cell membrane during electroporation, making it easier for other charged biomolecules to pass through[29].

Transformants were recovered for one day before plating one-fifth of the reaction across ½ salt L1 plates supplemented with 10 µM or 5 µM 2-FA, with or without 5 µg/ml adenine (Supplementary Fig. 14). This pilot experiment generated thousands of 2-FA-resistant transformants across the plates that had been supplemented with adenine, but only a handful across the plates without supplemental adenine. This was an unexpected discovery given our previous observations. Ten 2-FA-resistant colonies were passaged for further screening at the *PtAPT* loci to determine if Cas9-mediated engineering had occurred.

We used the primer sequences reported by Serif et al.[28] (Supplementary Data 2) to screen the *PtAPT* alleles using PCR. When using primers that bind to both alleles, amplicons of varying sizes were visible across the 2-FA resistant transformants, most of which differed from that of the wild-type strain (Fig. 6b). Likewise, amplicons of varying sizes were visible when using primers specific to allele 1 (Figs. 6c) and 2 (Fig. 6d). This suggested that Cas9-mediated engineering had occurred, but to corroborate our findings, amplicons generated with allele-specific primers were sent for Sanger sequencing.

Alignment of the sequenced amplicons to the wild-type sequence for alleles 1 and 2 confirmed the presence of indels in all 12 alleles that could be analysed (Fig. 6e). We were able to generate sequencing data of both alleles for three of the transformants (C1, C6, and C7), with one clone (C1) demonstrating a homozygous 12 bp deletion across both alleles. Interestingly, there were large (>50 bp) insertions in four of the sequenced alleles. Querying the insertion sequences in NCBI BLAST revealed that carrier DNA had integrated into these loci during the double-strand break repair process. The alignments for C6 and C10 contained two and four non-contiguous regions from the salmon and/or trout genomes, demonstrating that multiple shorn pieces of carrier DNA had been integrated as opposed to a single large fragment.

Cotransformation of an RNP complex and episome

With the discovery that RNP complexes can be electroporated into *P. tricornutum*, we next sought to see if both an RNP and episome could be co-transformed simultaneously, and if so, what the rate of co-transformation is. The aim of this experiment was to develop a protocol that could significantly improve screening for gene deletions that do not produce an obvious phenotype. To test this, the PCR-amplified episome pPtGE31_ΔPtR was combined with the Cas9:sgAPT1 complex and electroporated into spheroplasted *P. tricornutum* cells. Transformants were first plated on nourseothricin or 2-FA selective plates, and then later repatched onto the alternative selection type to determine if co-transformation had occurred.

Our earlier RNP transformation experiment (Supplementary Fig. 14) had demonstrated that adenine supplementation is important when selecting for *PtAPT* knock-out colonies post electroporation; however, adenine is also seemingly toxic to untransformed *P. tricornutum* cells (Supplementary Fig. 13). Thus, when co-transforming pPtGE31_ΔPtR and Cas9:sgAPT1, the electroporated cells were plated on two different selection plates: (1) ½-salt L1 plates supplemented with 5 μg/ml adenine and 10 μM 2-FA, and (2) ½-salt L1 plates supplemented with 1 μg/ml adenine and 100 μg/ml nourseothricin. The lower adenine concentration was chosen to allow for the growth of co-transformed colonies whilst not unintentionally selecting against colonies that had only uptaken the episome.

After two weeks of growth on nourseothricin or 2-FA selection, 50 co-transformed colonies were repatched onto plates with the same selection type and grown for 1 week. Then, the colonies were repatched onto the alternative selection type to determine if any had double-resistance. Of the 100 screened colonies, 13 demonstrated double-resistance to nourseothricin and 2-FA (Supplementary Table 5), indicating co-transformation had successfully occurred at a rate of ~10%.

Revisiting the PEG transformation method

A polyethylene glycol (PEG) transformation method for *P. tricornutum* with an efficiency of ≤ 1 transformant per 10^8 cells was previously described by Karas et al.[6]. For this method, algal cells were protoplasted using a combination of zymolyase, lysozyme, and hemicellulose ahead of PEG treatment. We revisited this method using alcalase as the sole enzyme for cell wall removal, with the intention of providing a simple and equipment-independent alternative to electroporation that is particularly useful for labs or species where this method remains a practical barrier.

We followed the method described by Karas et al.[6], substituting the aforementioned enzyme cocktail with 10, 100, or 1000 μl of alcalase (i.e., 30 mAnson, 300 mAnson, and 3 Anson, respectively) for protoplasting ~3 ×10^8 cells. The 3 Anson treatment was potent enough to protoplast the majority (i.e., >80%) of the cells. As well, the protoplasts were transformed with 500 ng of linear pPtGE31_ΔPtR as opposed to 25 μg of a circular episome, the amount originally used by Karas et al., to be able to compare the efficiency of transformation against our electroporation experiments. Treatment with 3 Anson of alcalase generated approximately 160 transformants per 3 ×10^8 cells, an efficiency of ~4 transformants per 10^7 cells. The other two treatments yielded 0 transformants (30 mAnson of alcalase) and approximately 6 transformants (300 mAnson alcalase), so all future PEG transformations were conducted with 3 Anson of alcalase per 3–5 ×10^8 cells. This suggests that complete cell wall removal is paramount to efficient PEG transformation.

We then sought to test other parameters of the PEG method to further optimize the transformation efficiency. First, we began to include 40 μg of carrier DNA per PEG reaction as this had led to a ~15-fold increase in transformation efficiency during electroporation. Then, we tested various concentrations of PEG-8000 and found that a range from 20 to 30% (w/v) can be used (Fig. S15a). Going forward with 20% PEG, we then tested if the duration of protoplasting, inclusion of carrier DNA, and length and/or temperature of incubation would alter transformation efficiency (Fig. S15b). As anticipated, the inclusion of carrier DNA increased transformation efficiency >10-fold (Supplementary Fig. 15c). Longer protoplasting and incubation times also appeared to increase transformation efficiency to ~6 transformants per 10^6 cells (Supplementary Fig. 15c).

The optimized PEG transformation protocol was then used to transform 1 μg of pSC5 into protoplasted *P. tricornutum* cells, generating ~60 transformants (Supplementary Fig. 15d). DNA were isolated from four algal transformants and recovered in EPI300 *E. coli*. Bacterial transformants were then assessed for conjugation in the same manner as before. All recovered episomes demonstrated the ability to conjugate (Supplementary Fig. 15e), verifying that pSC5 had been intactly transformed into *P. tricornutum* through the PEG method.

PEG transformation to *T. pseudonana*

Electroporation parameters can vary drastically between species and are often finnicky to troubleshoot or optimize. To expand the applications of our alcalase-based transformation methods, we sought to adapt the optimized PEG transformation protocol for *T. pseudonana*, a model diatom species that possesses a more typical silicified frustule. Given the differences in cell wall composition between these species, *T. pseudonana* cells were subjected to varying protoplasting conditions (Supplementary Fig. 16a) before PEG treatment. The circular episome pBIG1 (8.2 kb) conferring nourseothricin resistance and carrying eGFP was used in these experiments. All protoplasting conditions led to the emergence of colonies on the nourseothricin selection plates, with the harsher protoplasting conditions (i.e., ≥ 1.5 Anson alcalase) yielding more growth.

Transformants from the 500 μl alcalase treatment plate (i.e., 1.5 Anson, 40 min) were pooled and subjected to fluorescence microscopy to assess if eGFP was being expressed, as this would indicate successful episome transfer. Both autofluorescence and eGFP signals were visible in the transformed cells (Supplementary Fig. 17), demonstrating that episome transfer had occurred.

To further characterize if protoplasting was truly occurring, *T. pseudonana* cells were incubated with 10 or 1000 μl of alcalase (i.e., 30 mAnson and 3 Anson, respectively) for 40 min (Supplementary Fig. 16b and c, respectively) and then imaged via microscopy. Based on these

microscopy images, cells from the 10 μl treatment (i.e., 30 mAnson) showed a protoplasting rate of ~4% (1/28 cells), whereas cells from the 1000 μl treatment (i.e., 3 Anson) demonstrated a protoplasting rate of ~50% (13/20 cells). This finding underscores that alcalase can be used to degrade the cell wall of diatoms harboring silica frustules.

Discussion

In this paper, we demonstrate that partial or complete removal of the cell wall greatly increases the efficiency of electroporation and PEG transformation in *P. tricornutum*. These optimized protocols enabled us to push the limits and possibilities of transformation in diatoms. Episomes as large as 55.6 kb could be intactly delivered through either approach, and it was possible to electroporate RNP complexes directly to the algal nucleus for DNA-free genome engineering. Perhaps most interestingly, we also discovered that episomes can be directly assembled in the diatom cell with high efficiency – a feat that has been predominantly limited to *E. coli* and *Saccharomyces cerevisiae*, the workhorses of synthetic biology. These methods are not only more efficient than what has been reported in the past[5–7], but present as faster and simpler alternatives to bacterial conjugation for the reliable transformation of episomes.

The cell wall is the first physical barrier to entry when attempting to transform diatoms. It follows that permeabilizing this wall should increase the ability for biomolecules, like DNA and RNPs, to successfully enter the cell. Although spheroplasting and protoplasting techniques are commonly used for transforming other algal and plant species[30,31], these approaches had yet to be fully investigated in diatoms until this study. We found that spheroplasting *P. tricornutum* cells with alcalase prior to conducting electroporation increased the efficiency by ~76- to 280-fold (Supplementary Table 1), with cells derived from plated cultures demonstrating the highest transformation rates. Protoplasting the cells also enabled efficient transformation through the PEG method (~ 6 transformants per 10^6 cells).

Though *P. tricornutum* is unique in both its poorly silicified cell wall[32] and willingness to grow on solid media, we were able to adapt the alcalase-based protoplasting technique for *T. pseudonana*, a diatom species with a silicified cell wall that is typically cultured in liquid media. These results suggest that the spheroplasting and protoplasting techniques can be applied to other diatoms, though it will likely be necessary to optimize the intensity and duration of the enzyme treatment on a species-by-species basis along with other parameters in the respective protocols. This expectation aligns with the trajectory of bacterial conjugation in diatoms, which was originally established using plated *P. tricornutum* cells: although initially assumed to have limited applicability, conjugation has since been successfully adapted to multiple species, including *Cyclotella meneghiniana*, *Nitzschia captiva*, and *Fistulifera solaris*[33–35]. This precedent underscores the realistic potential for broader adaptability of the methods presented here.

The surprising discovery of carrier DNA integrated into the termini of a once linear episome led us to explore the potential for assembly directly in *P. tricornutum*. It was possible to simultaneously electroporate multiple episomal fragments that could recombine through NHR or HDR to form a single episome. The HDR pathway is much more robust for assembly purposes, with ~87% (26/30 colonies) screening positively for the recombined fragment junctions and 70% of recovered episomes having the expected sequence during the episomal assembly experiments (Supplementary Fig. 10). This capacity for in vivo recombination led us to coin the term diatom in vivo assembly (DIVA) to describe episome assembly occurring directly within the algal cell.

To further define the practical limits of DIVA, we assessed the minimal homology required for HDR-mediated assembly (Supplementary Fig. 11a, b). Using NAT constructs with 0, 20, or 40 bp of overlap to an episome, we found that 20 bp is sufficient to direct recombination – consistent with observations in *S. cerevisiae* and indicating that any fragment of interest can be prepared for HDR assembly by adding ≥20 bp of overlaps via PCR.

We next asked whether DIVA could support the direct insertion of de novo synthesized constructs. Two proof-of-concept assemblies – a CytB-eGFP fusion carrying a predicted mitochondrial targeting peptide, and an eGFP-T2A-NAT construct designed to minimize false positives – each yielded ≥100 transformants. Imaging revealed mitochondrial eGFP localization in 4/10 CytB-eGFP transformants and robust cytosolic expression in all 10/10 eGFP-T2A-NAT transformants screened. Sequencing of recovered episomes confirmed integration of the intended constructs, though several recombinants carried unexpected rearrangements, highlighting the need to further define the fidelity and constraints of DIVA.

Together, these results demonstrate that episomes can be assembled de novo in *P. tricornutum* via DIVA and establish a foundation for further exploration of this process. At present, however, the limits and most appropriate applications of DIVA remain unclear. Key parameters like the upper limit on fragment number, the size range of fragments, and the optimal overlap length, will require systematic investigation before DIVA can be deployed as a reliable or generalizable assembly strategy. We anticipate that the clearest advantages of DIVA may emerge in contexts where rapid or pooled assembly directly in *P. tricornutum* is useful – such as screening promoter libraries or constructing variant pools for phenotype-based enrichment.

The balance between homology-driven and non-homologous repair pathways is also likely to influence the accuracy of DIVA. Angstenberger et al. have demonstrated that knockdown of *ligIV*, an essential component of the non-homologous end joining (NHEJ) pathway, significantly increases the efficiency of homologous recombination when integrating a homology-bearing transgenic cassette into the *P. tricornutum* genome[36]. They also identified other potential targets of the NHEJ pathway (i.e., *Ku70/80*, *XRCC4*, and *BRCT*), none of which have yet to be reported in the literature for *P. tricornutum*. Knockdown of these genes may improve the fidelity of HDR-based DIVA, although the DNA repair mechanisms governing episomal assembly may differ from those involved in chromosomal integration.

Past methods for Cas9-based engineering in *P. tricornutum* included the delivery of RNPs through biolistics[28] or the bacterial conjugation of episomes expressing the Cas9 complex[9]. Our paper adds electroporation and PEG transformation as additional tools for Cas9 genome engineering, with it being possible to deliver episomes through either method or directly electroporate RNPs into *P. tricornutum*. We generated thousands of 2-FA resistant transformants by electroporating an RNP complex targeting *PtAPT*, >200-fold more efficient than what was previously reported with biolistics[28]. It was also possible to cotransform an episome and the *PtAPT* RNP complex at a rate of ~10%. This offers an alternative method for engineering regions of the genome that are not counter-selectable, as a proportion of transformants that have uptaken the episome, which can be selected for, will have also uptaken the RNP and incurred genome editing. The capacity to perform multigene engineering by electroporating multiple RNP complexes was not explored in this study, but we believe this will be possible as this has been previously accomplished via biolistics[28].

Through this series of proof-of-concept experiments, we also uncovered an unexpected paradox: while adenine supplementation (5 μg/ml) is visibly toxic to wild-type *P. tricornutum*, it substantially improves the recovery of Δ*PtAPT* transformants under 2-FA selection. Because *PtAPT* functions in the adenine salvage pathway, it is not immediately clear why supplemental adenine would have any beneficial effect on Δ*PtAPT* strains; if anything, one might expect the opposite. This result suggests that additional adenine salvage, recycling, or purine-metabolic pathways may be active in *P. tricornutum*, and that supplemental adenine may mitigate 2-FA toxicity through these alternative routes. Elucidating the underlying metabolic

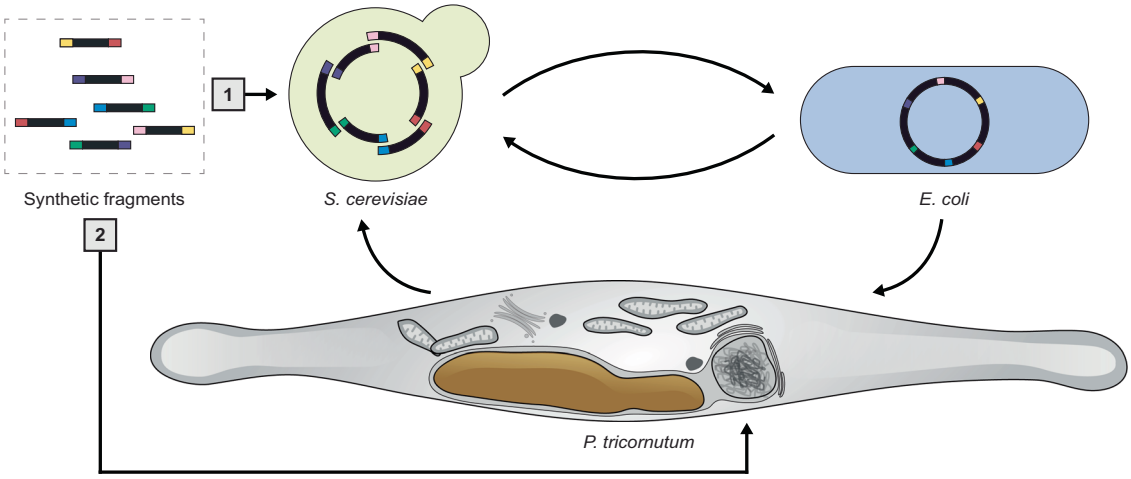


Fig. 7 | The canonical [1] and alternative [2] pathways for delivering and assembling episomes in *P. tricornutum*. Prior to this work, overlapping synthetic fragments would be assembled in *S. cerevisiae*. Assembled episomes would be transferred to *E. coli* and then conjugated to *P. tricornutum*. In this pathway, the episome is shuffled between three host organisms during the design-build-test cycle. The alternative pathway provides the opportunity to directly assemble episomes in the agal cell by electroporation of the synthetic fragments.

regulation will require further investigation, but is beyond the scope of the present study.

Complete sequencing data could not be captured for all the analyzed *PtAPT* mutants. This is likely due to the presence of large deletions that extend beyond the range of the primer binding locations. For reference, deletions as large as ~2.7 kb have been observed in Cas9-edited *P. tricornutum* loci before[37] and the primer sets for screening the *PtAPT* alleles were only positioned up to ~450 bp away from the cut site. The alleles that could be sequenced demonstrated that Cas9 engineering had occurred in all assessed clones and revealed the unexpected insertion of carrier DNA in some clones. Based on this finding, we anticipate that it will be possible to intentionally introduce heterologous sequences into the *P. tricornutum* genome by simultaneous electroporation of an RNP along with a selectable DNA fragment bearing homology arms to the targeted locus. Episome-based Cas9 genome engineering has been used to integrate heterologous sequences into the *P. tricornutum* genome before[38], so we anticipate that RNP-based integration of similarly sized heterologous inserts should be possible.

The methods presented in this study pave the way for accelerated engineering in *P. tricornutum*. When we first proposed the Pt-syn 1.0 project, we envisioned using bacterial conjugation to introduce synthetic chromosomes into the algal cell[14]. Now, we can envision the in vivo assembly of synthetic chromosomes directly in the algal nucleus or the piece-by-piece replacement of endogenous chromosomes with synthetic parts (Fig. 7). Challenges and questions remain about the limits of these approaches that future work will have to address before these feats can be accomplished, but in any case, this work has led us one step closer towards actualizing Pt-syn 1.0 – a diatom controlled by a wholly-synthetic genome.

To conclude, we establish improved electroporation and PEG-based approaches that increase the efficiency and reliability of DNA delivery into *P. tricornutum*. These methods allow transformation with very small amounts of DNA, support the uptake of large episomes, and enable both non-homologous and homology-directed assembly of episome fragments directly in the algal nucleus. We also show that Cas9 RNPs can be delivered by electroporation for DNA-free genome editing, and that episomes and RNPs can be co-transformed to facilitate screening of edits that lack obvious phenotypes. Finally, the adaptation of the PEG method to *T. pseudonana* demonstrates the broader applicability of this approach to other diatom species. Together, these results provide an optimized set of tools for diatom genetic engineering and will support future efforts that require rapid testing and construction of synthetic DNA in algal systems.

Methods

Microbial strains and growth conditions

P. tricornutum (Culture Collection of Algae and Protozoa, CCAP 1055/1) was cultured in liquid L1 media or on ½-salt L1 1% agar plates (w/v) in a growth chamber set to 18 °C and a light/dark cycle of 16/8 h. Cultures were irradiated with white LED lights (Mars Hydro, model: TSL 2000) set to photon flux densities of ~75 µE m − 2 s − 1 (liquid cultures) or ~55 µE m − 2 s − 1 (agar plate cultures). Agar plates were placed in clear polystyrene trays to reduce desiccation. L1 media was prepared as previously described and lacks silica[6].

In brief, liquid L1 medium was prepared by combining 1 L aquil salts (synthetic seawater), 2 ml NP stock, 1 ml L1 trace metals stock, and 0.5 ml f/2 vitamin solution. Once combined, 2 M NaOH was added until a pH of ~8.0 was achieved and the solution was sterilized through a 0.2-µm filter. For solid medium, sterile liquid L1 medium was mixed 1:1 with autoclaved 2% (w/v) Bacto-agar (all equilibrated to ~60 °C) and poured into petri dishes.

Aquil salts were prepared by mixing equal volumes of separate 2X anhydrous and hydrous salt solutions. The 2X anhydrous salt solution was prepared by dissolving the following in 500 ml ddH_2O: NaCl (24.5 g), Na_2SO_4 (4.09 g), KCl (0.7 g), $NaHCO_3$ (0.2 g), KBr (0.1 g), H_3BO_3 (0.03 g or 3 ml of a 10 mg ml^{-1} stock), and NaF (0.003 g or 300 µl of a 10 mg ml^{-1} stock). The 2X hydrous salt solution was prepared by dissolving $MgCl_2{\cdot}6H_2O$ (11.1 g) and $CaCl_2{\cdot}2H_2O$ (1.54 g) in 500 ml ddH_2O.

The NP stock solution was prepared at 100X by dissolving $NaNO_3$ (37.5 g per 100 ml) and $NaH_2PO_4{\cdot}H_2O$ (2.5 g per 100 ml) in ddH_2O. The trace metal stock was prepared at 1000X by dissolving the following in 1 L of ddH_2O: $FeCl_3{\cdot}6H_2O$ (3.15 g) and $Na_2EDTA{\cdot}2H_2O$ (4.36 g), followed by addition of the following components from concentrated stocks: $CuSO_4{\cdot}5H_2O$ (9.8 g L^{-1}), 0.25 ml; $Na_2MoO_4{\cdot}2H_2O$ (6.3 g L^{-1}), 3.0 ml; $ZnSO_4{\cdot}7H_2O$ (22.0 g L^{-1}), 1.0 ml; $CoCl_2{\cdot}6H_2O$ (10.0 g L^{-1}), 1.0 ml; $MnCl_2{\cdot}4H_2O$ (180.0 g L^{-1}), 1.0 ml; H_2SeO_3 (1.3 g L^{-1}), 1.0 ml; $NiSO_4{\cdot}6H_2O$ (2.7 g L^{-1}), 1.0 ml; Na_3VO_4 (1.84 g L^{-1}), 1.0 ml; and K_2CrO_4 (1.94 g L^{-1}), 1.0 ml. Lastly, the vitamin stock was prepared at 2000X by adding thiamine-HCl (200 mg L^{-1}), biotin (10 ml L^{-1} of a 0.1 g L^{-1} stock), and cyanocobalamin (1 ml L^{-1} of a 1 g L^{-1} stock).

T. pseudonana (Bigelow National Center for Marine Algae and Microbiota, catalog number: CCMP1335) was cultured in liquid F/2

media in a growth chamber set to 20 °C and 24 hour white light at a photon flux density of 50 μE m^{-2} s^{-1}.

S. cerevisiae strain VL6-48 (American Type Culture Collection [ATCC], catalog number: MYA-3666) was grown in 2X YPAD media or plated on complete minimal media lacking histidine with 1% agar (w/v) and 1 M D-sorbitol. Yeast cultures were incubated at 30 °C, with liquid cultures placed on a shaker set to 225 rpm. Agar plates were sealed with parafilm to reduce desiccation. Dropout media were prepared according to the manufacturer specifications; in brief, a minimal base media (Takara, catalog number: 630411) was combined with a -HIS dropout supplement (Takara, catalog number: 630415) in ddH_2O. Once combined, 2 M NaOH was added until a pH of ~5.6 was reached. YPAD media was sterilized through a 0.2-μm filter; dropout media was sterilized by autoclaving.

E. coli strain EPI300 (LGC Biosearch Technologies, Lucigen, catalog number: EC300110) was grown in lysogeny broth (Miller formula) with or without 1% agar (w/v) and supplemented with 15 μg/ml chloramphenicol and/or 40 μg/ml gentamycin and/or 100 μg/ml L-arabinose. Cultures were placed in an incubator set to 37 °C, with liquid cultures placed on a shaker set to 225 rotations per minute (rpm).

PCR amplification

Episomes and/or fragments for transformation and/or assembly were amplified with GXL polymerase (Takara, catalog number: R050A) using the rapid PCR protocol (all primers are listed in Supplementary Data 2). Of note, we used both phosphorylated and non-phosphorylated primers when amplifying pPtGE31_ΔPtR as a linear episome. PCR reactions were conducted with a total of 25 to 35 cycles depending on amplification efficiency. Fragments that were amplified from an episome template were treated with 10 units (0.5 μl) of *Dpn*I (New England Biolabs, catalog number: R0176) and incubated at 37 °C for 30 min before deactivation at 80 °C for 20 min. After deactivation, fragments were column-purified using the PureLink PCR Purification Kit (Invitrogen, catalog number: K310001) and eluted with ddH_2O. Where possible, primers were designed using Primer3 (version 4.1.0).

PCR screening of algal and *E. coli* transformants

DNA isolated from algal and *E. coli* transformants was screened for the presence of specific episomal regions using a PCR assay. Here, 1 μl of DNA isolated through alkaline lysis was used as template in a 10 μl SuperPlex PCR reaction (Takara, catalog number: 638543). For all algal transformants, episomes were analzyed with a single primer set at a time as there was rampant off-target binding when multiple primer pairs were used simultaneously (all primers are listed in Supplementary Data 2). We were able to successfully conduct multiplex PCR using *E. coli* template DNA; however, when screening the NHR assembled episomes, we chose to perform singleplex screens due to the unexpected recombination of the fragments. Reactions were carried out according to the manufacturer guidelines for a total of 30 cycles.

DNA isolation

Episomes were isolated from *P. tricornutum*, *S. cerevisiae*, and *E. coli* using an alkaline lysis protocol that was previously described in Karas et al.[6], with species-specific modifications detailed below. Briefly, cell pellets were resuspended in 250 μl of P1 (Qiagen, catalog number: 19051), with or without additional protoplasting enzymes as indicated below, and lysed with 250 μl P2 buffer (Qiagen, catalog number: 19052). Samples were mixed by 5–10 inversions and incubated for 1 min before neutralization with 250 μL P3 buffer (Qiagen, catalog number: 19053). Lysates were centrifuged at 16,000 x G for 10 min, and ~700 μl of supernatant was transferred to a fresh tube. DNA was precipitated from the lysate by adding 750 μl ice-cold 100% isopropanol, which was then centrifuged at 16,000 × g for 10 min. The DNA pellets were washed with 500 μl ice-cold 70% ethanol and centrifuged for 5 min. Pellets were air-dried and resuspended in 30–50 μl sterile ddH_2O.

Steps 1 to 2 for *E. coli*: (**1**) *E. coli* colonies were grown overnight in il LB supplemented with the appropriate antibiotic. (**2**) 1.5 to 3.0 ml of culture was pelleted by centrifugation at 10,000 x G.

Steps 1 to 3 for *S. cerevisiae*: (**1**) 3–5 ml of *S. cerevisiae* transformants (pooled) were harvested from dropout media plates ~5 days after assembly. (**2**) The cells were pelleted by centrifugation at 3000 x G for 10 min. (**3**) The P1 resuspension buffer consisted of 240 μl P1 with 5 μl of 1.4 M b-Mercaptoethanol and 5 μl Zymolyase solution (Zymolyase solution: 200 mg Zymolyase 20 T (BioShop, catalog number: ZYM001.100), 9 ml H_2O, 1 ml 1 M Tris pH 7.5, 10 ml 50% glycerol, stored at -20 °C). Following resuspension, the cells were incubated at 37 °C for 60 min ahead of the P2 step.

Steps 1 to 3 for *P. tricornutum*: (**1**) *P. tricornutum* transformants were repatched at least twice and grown to high density on a ¼-salt L1 plates supplemented with the appropriate antibiotic(s) (i.e., 100 μg/ml nourseothricin or zeocin for single selection, double selection plates contained both antibiotics at 50 μg/ml). On the second passage, colonies were patched as large streaks to ensure there were enough cells for DNA isolation – typically, a single plate would be used to passage 6 colonies that were struck out to each cover approximately a sixth of the plate. Plates were grown until a reasonable cell density was achieved, which typically took 4 days for nourseothricin plates and upwards of 6 days for zeocin, 2-FA, or double-selection plates. (**2**) Cells were scraped and resuspend in 250 μl of resuspension buffer, which consisted of 245 μl of P1 (Qiagen, catalog number: 19051) and 5 μl of alcalase (i.e., 15 mAnson, Sigma-Aldrich, catalog number: 126741). (**3**) The resuspended cell mixture was vortexed for 1 to 3 seconds, then placed in a heating block set to 56 °C for 10 min – this is within the optimal temperature range for alcalase activity, according to the manufacturer. After 10 min passed, the cells were vortexed again for 1 to 3 seconds to ensure there were no cell clumps.

For *E. coli*-derived episomes smaller than 11 kb, we used column purification kits to isolate and purify DNA (New England Biolabs, catalog number: T1110). DNA was eluted with ddH_2O equilibrated to 56 °C.

To isolate a high concentration of pSC5 (~56 kb) from *E. coli*, alkaline lysis was performed using 27 ml of culture (1.5 ml culture per lysis reaction, 18 reactions in total). The DNA pellets were resuspended with ~25 μl of $sddH_2O$ and combined into a single 1.5 ml tube (~ 460 μl total). Half of the isolated DNA was then digested using *Sac*I-HF, as described below. Then, a tenth volume of sodium acetate (3 M, pH 5.2) and two volumes of ice-cold 100% ethanol were added to the samples. The solutions were inverted to mix and placed into a -80 °C freezer for an hour before performing centrifugation at 16,000 x G for 10 min at 4 °C. The DNA pellets were then washed twice with 500 μl of ice-cold 70% ethanol before being decanted, dried, and resuspended with 50 μl of $sddH_2O$. This additional sodium precipitation step was performed to further concentrate the samples and reduce the carryover of contaminants, particularly from the digest reaction.

DNA concentrations were measured by fluorometry using the DeNovix dsDNA Broad Range Assay (FroggaBio, catalog number: DSDNA-BROAD-EVAL) and purity ratios were assessed using a Nanodrop spectrophotometer. For the experiments shown in Figs. 2 to 4, samples were adjusted to be equimolar prior to conducting electroporation.

Restriction digests of episomes

Circular pPtGE31_ ΔPtR was isolated a *P. tricornutum* transformant and recovered in EPI300 *E. coli*. After isolating the episome from *E. coli*, approximately 10 μg was digested using 3 μl of *Eco*RI-HF (New England Biolabs, catalog number: R3101) at 37 °C for 1 hour. An additional reaction was conducted where 10 μg was digested using 3 μl of *Eco*RI-HF and 3 μl of mung bean nuclease (New England Biolabs, catalog number: M0250) at 37 °C for 1 hour. Complete digestion was

confirmed by agarose gel analysis, then samples were purified using a DNA Cleanup Kit (New England Biolabs, catalog number: T1030).

Circular pSC5 was isolated from EPI300 *E. coli* as described above. Prior to conducting the sodium acetate precipitation step, approximately 40 μg of pSC5 was digested with 18 μl of *Sac*I-HF (New England Biolabs, R3156) at 37 °C for 1 hour. Complete digestion was confirmed by agarose gel analysis, then the sample was purified by sodium acetate precipitation, as described above.

Agarose gel electrophoresis

DNA from plasmid isolations or PCR screens were assayed using 1–2% (w/v) agarose gels stained with ethidium bromide. Uncropped images for all gels presented in this manuscript can be viewed in the Source Data file.

Preparation of Cas9:sgAPT1 complex

HiFi Cas9 nuclease (catalog number: 1081060) and synthetic sgRNA were purchased from Integrated DNA Technologies (IDT) and prepared according to the manufacturer's guidelines. Briefly, 2 to 4 μg of Cas9 (62 μM) was combined with sgRNA (100 μM) in a 1:1.2 molar ratio and then incubated at room temperature for 20 min to form the Cas9:sgRNA complex. After incubation, the complex was set on ice for up to 1 hour, during which the *P. tricornutum* cells were prepared for electroporation. Once the cells were ready, 40 μg of ssssDNA was added to a 50 μl preparation of the cells, followed by the Cas9:sgRNA complex. The mixture was pipetted up-and-down a few times before transferring to an electrocuvette and promptly pulsed. For the co-transformation experiment, we added 500 ng of PCR-amplified pPtGE31_ ΔPtR to the cells ahead of electroporation.

Yeast assembly of pPtGE31_ShBle

The episome pPtGE31_ShBle was created by yeast assembly of three overlapping fragments. First, the episome pPtGE31_ ΔPtR was amplified as two fragments using primers that would add 40 bp overlaps for the promoter and terminator region of the ShBle cassette and split the episome in the HIS3 region. Then, the ShBle cassette was amplified from the p0521s plasmid (Addgene ID: 62862)[6] using primers that would add 40 bp overlaps for the insertion site in pPtGE31_ ΔPtR (all primers are listed in Supplementary Data 2). Yeast assembly was conducted as previously described[39] using ≥ 100 ng of each fragment in place of a bacterial donor.

Electroporation to *P. tricornutum*

Prior to attempting this methodology, it is worth noting that there is a great degree of variance between electroporation equipment, and this can drastically impact the efficiency of transformation to *P. tricornutum*. If others attempt this protocol, it may be necessary to initially attempt electroporation with several different parameters to determine which settings work best with different set-ups. For our set-up (500 V, 50 μF, 400 Ω), an empty electrocuvette generates a time constant of ~32 to 35 milliseconds. A neighboring lab has successfully performed electroporation to *P. tricornutum* using an older electroporator set to the originally described parameters by Zhang and Hu (500 V, 25 μF, 400 Ω)[5].

P. tricornutum cells were grown in liquid media or on agar plates for the experiments conducted in Fig. 1 – all successive experiments used cells derived from agar plates. For liquid cultures, 1 ml of cells adjusted to 1–2 × 10^8 cells/ml were passaged into 50 ml of L1 media and placed in a growth chamber on a shaker set to ~100 rpm. For agar plate cultures, 250 μl of cells adjusted to 1 × 10^8 cells were spread onto a ½-salt L1 plate with 1% agar (w/v). Liquid and agar plate cultures were grown for 4 days. After 4 days, cells were scraped from the agar plates using 750 μl of L1 and transferred to a 1.5 ml tube. An additional 750 μl of L1 was used to remove any remaining cells from the plate. Then, 10^{-1} to 10^{-2} dilutions of the liquid and agar plate cultures were prepared and used to estimate cell density with a hemocytometer. Liquid and agar plate cultures were adjusted to have a final cell density between 2 and 4 × 10^8 cells/ml; for the liquid cultures, this requires concentrating the cells via centrifugation and resuspension in a smaller volume. Here, the cells are transferred from the flask to a 50 ml conical tube and pelleted by centrifugation at 2000 x *g* for 10 min at 18 °C. The pellet was resuspended with 375 mM D-sorbitol to a concentration between 2 and 4 × 10^8 cells/ml and transferred to a 1.5 ml tube. Cells derived from agar plates were pelleted by centrifugation at 750 x *g* for 4 min; here, we used a microcentrifuge, which could pellet the cells at a lower centrifugal speed. As before, the cells were resuspended using 375 mM D-sorbitol to a concentration between 2 and 4 × 10^8 cells/ml. The final cell concentration was equalized between the liquid and agar plate-derived cultures for every biological replicate.

To spheroplast, 1 mAnson of alcalase (Sigma-Aldrich, catalog number: 126741) was added for every 1 × 10^8 cells. At this stage, the cells will have been previously concentrated to a density between 2 – 4 × 10^8 cells/ml and already resuspended in 375 mM D-sorbitol. Alcalase was diluted 10-times in sddH_2O to reduce the pipetting error associated with small volumes. After the addition of alcalase, cultures were placed on a rocking shaker set to ~15 rpm at room temperature for 20 min, during which spheroplasting occurs. Once 20 minutes had passed, the cells were pelleted by centrifugation at 325 x G for 4 min. The supernatant is decanted, and the cells are gently resuspended with 1 ml of ice-cold 375 mM D-sorbitol using a P1000 pipette. Following this step, all materials (i.e., cells, DNA, D-sorbitol, electrocuvettes) should be kept on ice for the duration of this protocol.

The electroporation protocol we used is adapted from Zhang and Hu[5], with adjustments made by Kassaw et al.[18] and Pampuch (unpublished). First, the spheroplasted cells were pelleted by centrifugation at 325 x *g* for 4 min. The supernatant is decanted, and the pellet is gently resuspended with 1 ml of ice-cold 375 mM D-sorbitol. This process of centrifugation and resuspension is repeated another 2 times, such that in total, the cells will have been resuspended with 375 mM D-sorbitol five times (this includes centrifugation steps prior to and post-spheroplasting). After the fifth wash, the cells are spun at 750 x G for 3 min, and a P200 pipette is used to gently remove any remaining supernatant. The cells are gently resuspended with 100 to 200 μl of 375 mM D-sorbitol such that the total volume of resuspended cells will be between ~150 to 250 μl. Into sterile 1.5 ml tubes, 50 μl of cells are aliquoted along with 4 μl (10 μg/μl) of single-stranded salmon sperm DNA (ssssDNA, Sigma-Aldrich, catalog number: D7656) and the DNA to be transferred. Of note, the ssssDNA should be heated to 95 °C for 2 to 5 min and then cooled on ice prior to this to ensure it remains single-stranded. We used between 1 to 1000 ng of an 11 kb episome for most transformations in this paper, with 500 ng of this construct appearing sufficient to saturate the electroporation reaction. Negative controls consisted of only ssssDNA in the cell mixture.

Cells are gently pipetted up and down 2 times before transferring to an ice-cold sterile electrocuvette (Fisher Scientific, catalog number: FB102), ensuring that no air bubbles have been transferred and the cells lie evenly across the bottom. The electrocuvette is pulsed in an electroporator (BioRad, Gene Pulser Xcell System, catalog number: 1652660) set to the following parameters – voltage: 500 V, capacitance: 50 μF, and resistance: 400 Ω. The time constant varied from 30 to 32.5 milliseconds for the experiments conducted in Figs. 1–4. Immediately after pulsing, 1 ml of L1 media is gently added to the electrocuvette without disturbing the cells. The electrocuvettes are left to rest at room temperature for 10 min before fully resuspending the cells and transferring the mixture to 10 ml of L1 in 50 ml Falcon tubes. The lids of the Falcon tubes are left loose for air transfer, and the cultures are moved to a growth chamber to recover for 16 to 24 h.

After recovery, the cultures are pelleted by centrifugation at 2000 x G for 10 min at 18 °C. The pellet is resuspended with 500 μl of L1 media, then 50 μl (one-tenth), 125 μl (a quarter), and/or 250 μl (half) of

the cells are spread onto ¼-salt L1 plates supplemented with the appropriate antibiotic(s).

PEG transformation of *P. tricornutum*

The optimized PEG transformation protocol is similar to the methods described by Karas et al.[6], with some important modifications. First, 250 µl of wild-type cells adjusted to 1×10^8 cells were spread onto a ½-salt L1 plate with 1% agar (w/v). After 4 to 5 days of growth, cells were scraped from the agar plates using 1.5 µl of L1 and transferred to a 1.5 ml tube. The cells were counted via a hemocytometer and adjusted to a concentration of 2–4 $\times 10^8$ cells; in our experience, the density of the scraped cells typically falls within this range.

For protoplasting, 1 ml of cells was transferred to a 50 ml Falcon tube containing 9 ml of L1 and 1000 µl of alcalase (i.e., 3 Anson); here, alcalase is used in place of the zymolyase, hemicellulase, and lysozyme cocktail previously used by Karas et al.[6]. This mixture was then placed on a slow rocking shaker (~15 rpm) and incubated at room temperature for up to an hour. The protoplasting efficiency can be estimated by examining the cells at ≥200x magnification during incubation. The majority of cells (>80%) should have formed protoplasts before proceeding to the next step. In our hands, the protoplasts remain intact during this reaction, likely because the cells are incubated in full-salt L1 medium, which has a high osmolarity. A relatively high volume of alcalase is required compared to the electroporation method, which may reflect reduced enzyme activity in the presence of L1 salts.

After incubation, 40 ml of L1 was gently added to the cells, which were then pelleted by centrifugation at 1500 x *G* for 5 min at 18 °C. The supernatant was carefully decanted, and the cells were gently resuspended with 500 µl of L1 followed by 8 µl of ssssDNA (10 µg/µl). For all further centrifugation steps, the pellet should be resuspended gently to avoid mechanically shearing the protoplasts. Then, 250 µl of protoplasts were combined with episomal DNA (circular or linear, 500 ng to 1 µg) and 1 ml of 20% PEG-8000 (w/v; also contains 10 mM Tris pH 8, 10 mM $CaCl_2$, 2.5 mM $MgCl_2$) equilibrated to 30 °C.

The protoplast:DNA:PEG suspension was gently mixed through inversion and then incubated at RT for 30 min. After incubation, the mixture was centrifuged for 7 min at 1500 x *g*. The supernatant was removed and cells were resuspended with 1 ml of L1, and then transferred into a 50 ml Falcon tube containing an additional 9 ml of L1 (total volume of 10 ml). The cells were then recovered in the growth chamber for ~48 h (i.e., 2 days). After recovery, the cells were centrifuged for 5 min at 1500 x *g* and 18 °C. The supernatant was decanted, and the pellet was resuspended in 500 µl of L1, then 125–250 µl of the cells were plated on ¼ × L1 supplemented with nourseothricin (100 µg/ml). Plates were left in the growth chamber for at least two weeks before being pictured.

PEG transformation of *T. pseudonana*

Approximately 1×10^5 cells/ml of *T. pseudonana* was inoculated into f/2 medium. After 4 days of growth (end of exponential phase), cells were pelleted by centrifugation at 1500 x *g* for 10 min and resuspended to a concentration of 3 x 10^8 cells/ml. The same transformation protocol was followed as described for *P. tricornutum*; however, the protoplasting step was conducted in f/2 medium with differing amounts of alcalase and incubation lengths (Supplementary Fig. 12a). Approximately 500 ng of the 8.2 kb episome pBIG1 was combined with 40 µg of carrier DNA for transformation. Following PEG transformation, the cells were recovered in 10 ml of f/2 media and recovered for 2 days prior to plating. Post-recovery, the cells were centrifuged for 5 min at 1500 x *g*. The cells were then resuspended in 500 µl of f/2 media, and 250 µl was plated on full-salt f/2 agarose plates supplemented with nourseothricin (100 µg/ml).

Screening *P. tricornutum* colonies transformed with two fragments – NHR pathway

Non-overlapping fragments from pPtGE31_ShBle were electroporated both individually and simultaneously into *P. tricornutum* (Supplementary Fig. 3). Cells that were electroporated with both fragments were plated onto ¼-salt L1 plates supplemented with 100 µg/ml nourseothricin or 100 µg/ml zeocin. Transformants that appeared following electroporation with both fragments were passaged twice – first, onto media supplemented with the same antibiotic as the initial transformation, and then onto media supplemented with the alternative antibiotic (Supplementary Table 4). Twenty transformants, ten derived from each initial transformation plate, that demonstrated resistance to both nourseothricin and zeocin were passaged for a third time onto double-selection media (¼-salt L1 supplemented with 50 µg/ml nourseothricin and zeocin). DNA isolation was performed using cells from this third passage, followed by PCR screening.

Screening *P. tricornutum* colonies transformed with multiple fragments – HDR pathway

Overlapping fragments were adjusted to a final concentration of 56 nM, which is four times more concentrated than the linear episome that was used as a positive control in these experiments (14 nM). Different combinations of the fragments were electroporated to measure the amount of carry-over for the template plasmid, pPtGE31_ShBle. For the experiments exploring the limits of HDR assembly, the molarity of the insert was ~4-5x greater than the episomal fragments. The CytB-eGFP and eGFP-T2A-NAT constructs were commercially synthesized from IDT and PCR amplified to generate ample DNA ahead of assembly; the sequences of these constructs are available in the Source Data file, along with the primers in Supplementary Data 2.

In both series of experiments, cells were plated onto ¼-salt L1 plates supplemented with 100 µg/ml nourseothricin and grown for at least two weeks before passaging 10 colonies from the assembly plates onto the same type of selection plate. The repatched colonies were grown for one week before passaging again for DNA isolation and/or microscopy.

Recovery of *P. tricornutum* episomes in E. coli

Episomes were isolated from *P. tricornutum* transformants as described above. For each electroporation reaction, 2 to 4 µl of *P. tricornutum* DNA was combined with 25 to 50 µl of homemade electrocompetent EPI300 cells (derived from LGC Biosearch Technologies, catalog number: EC02T110) in a sterile 1.5 ml tube. The mixture was gently pipetted twice and then transferred to a sterile 2 mm electrocuvette, ensuring no air bubbles were transferred. The electrocuvette is pulsed in an electroporator set to the following parameters – voltage: 2.5 kV, capacitance: 25 µF, and resistance: 200 Ω. This generated a time constant of 4.9 to 5.2 milliseconds. Electroporated cells were resuspended by pipetting 1000 µl of SOC media up-and-down in the electrocuvette, and then transferred to sterile 1.5 ml tubes to recover at 37 °C and 225 rpm for approximately 1 hour. After recovery, half of the reaction volume (~400 µl) was spread across a 1.5% agar (w/v) LB plate supplemented with 15 µg/ml chloramphenicol (pPtGE31_ ΔPtR and pPtGE31_ShBle recovery) or 40 µg/ml gentamycin (pSC5 recovery), which were then transferred to a 37 °C incubator. Transformants appeared across the plates within 24 h, with there being a range from 4 to ~100 colonies per transformation.

For recovery of pPtGE31_ ΔPtR and pPtGE31_ShBle episomes. Single *E. coli* colonies were picked and inoculated into 5 ml of LB media supplemented with 15 µg/ml chloramphenicol. Cultures were grown overnight at 37 °C and 225 rpm. The following day, 500 µl of culture

was inoculated into 4.5 ml of LB media supplemented with 15 µg/ml chloramphenicol and 100 µg/ml L-arabinose. These episomes contain a pCC1BAC backbone, which can be induced to high-copy number replication in EPI300 by adding L-arabinose to the media. Induced cultures were grown for at least five hours prior to episome isolation, with 1.5 ml of culture being used for DNA isolation, as described above.

For recovery of the conjugative plasmid pSC5. DNA isolated from eight *P. tricornutum* pSC5 transformants were electroporated into EPI300 cells. Four colonies per transformation were inoculated into 5 ml of LB supplemented with 40 µg/ml gentamycin. Additionally, EPI300 strains carrying pSC5 (positive control) or pSAP (recipient cell line) were inoculated from glycerol stocks into 5 ml of LB supplemented with 40 µg/ml gentamycin (pSC5) or 15 µg/ml chloramphenicol (pSAP). The cultures were placed at 37 °C and 225 rpm. After the cultures reached saturation (~18 h), 125 µl of the pooled *E. coli* transformants and pSC5 cultures were inoculated into 5 ml of LB supplemented with 40 µg/ml gentamycin. For the pSAP culture, 5 ml of culture was diluted into 45 ml of LB supplemented with 15 µg/ml chloramphenicol. The cultures were grown at 37 °C and 225 rpm for 2 h. Then, the cultures were pelleted by centrifugation at 3000 x *g* for 12 min at 10 °C, decanting the supernatant post centrifugation. The pooled transformants and pSC5 cultures were each resuspended with 100 µl of 10% glycerol, whereas the 50 ml pSAP culture was resuspended with 1000 µl of glycerol. Strains were kept on ice during these steps to cease bacterial growth.

For conjugation, 100 µl of the donor cells (transformant pools [x8] and pSC5) were mixed with 100 µl of the recipient cells (pSAP). The cells were mixed by pipetting up and down 4 times before transferring the whole mixture to a 2% agar (w/v) LB plate. Additionally, 100 µl of the recipient was plated by itself to serve as a negative control. Conjugation plates were incubated at 37 °C for 1 hour. Then, plates were scraped using 2 ml of sddH_2O. The scraped cells were vortexed for 5 seconds to ensure any clumps were dispersed. Transconjugants were spot plated onto 1.5% agar (w/v) LB plates supplemented with 40 µg/ml gentamycin and 15 µg/ml chloramphenicol with dilutions ranging from 10^0 to 10^{-8}.

Sequencing and alignment

Amplicons generated by amplifying the episome termini region in *P. tricornutum* transformants were sent for Sanger sequencing (London Regional Genomics Centre). Episomes isolated from *E. coli* were sent for Oxford Nanopore whole-plasmid sequencing (Plasmidsaurus Inc. and Flow Genomics). The resulting sequences were manually aligned to the respective template episomes. Regions that could not be manually aligned were queried using a local alignment algorithm (EMBOSS water tool)[40] to search for homology in the episome. Sequences that could not be aligned manually or algorithmically were queried using BLASTn (NCBI)[41].

Analyses of *P. tricornutum* transformation efficiency

The total number of colonies per transformation reaction was estimated by manually counting the number of CFUs per plate and then multiplying this value by 2 or 10, depending on whether half or a tenth of the total reaction was plated, respectively. In instances where the transformation plate was too dense to reliably count by eye, five 1 cm^2 quadrants were sampled from the plate and counted using a dissecting microscope with 10x magnification. Here, the total number of CFUs was estimated by averaging the number of colonies per quadrant and multiplying this value (i.e., colonies per cm^2) by the total surface area of the petri dish.

Transformation efficiencies were estimated by dividing the total number of CFUs by the number of cells per reaction. When graphically comparing transformation reactions, all calculations and plotting were done in RStudio (version 2025.05.1 + 513); otherwise, we used an Excel spreadsheet (version 16.103.4) to keep track of data and tabulate transformation efficiencies (available in the Source Data file).

Statistics and reproducibility

For experiments in which transformation efficiency was quantitatively assessed (Fig. 1e, f), four to five independent biological replicates were performed to account for variability between electroporation experiments and to enable statistical analysis. For all other experiments, which were designed as qualitative proof-of-concept assays (e.g., assessment of NHR and HDR assembly), at least two independent biological replicates were performed. Where appropriate, representative plate images from a single biological replicate are shown.

Differences in transformation efficiency across the four treatment groups (plate-derived untreated, plate-derived alcalase-treated, liquid-derived untreated, and liquid-derived alcalase-treated) were statistically assessed using a two-sided Kruskal-Wallis test. Pairwise comparisons between selected groups (LU vs. PS, LS vs. PS, PU vs. PS) were performed using Wilcoxon rank-sum tests with Bonferroni correction for multiple comparisons.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Sequencing information is available in the Supplementary Datasets 3–5. The raw *P. tricornutum* colony counts used to estimate transformation efficiency for the graphs in Fig. 1e, f are available in the Source Data File. All constructs used in *P. tricornutum* are also deposited to Addgene, as referenced throughout the text. Source data are provided with this paper.

References

1. Spoehr, H. A. Chlorella as a Source of Food. *Proc. Am. Philos. Soc.* **95**, 62–67 (1951).
2. Malviya, S. et al. Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci.* **113**, E1516–E1525 (2016).
3. 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).
4. Giguere, D. J. et al. Telomere-to-telomere genome assembly of Phaeodactylum tricornutum. *PeerJ* **10**, e13607 (2022).
5. Zhang, C. & Hu, H. High-efficiency nuclear transformation of the diatom *Phaeodactylum tricornutum* by electroporation. *Mar. Genom.* **16**, 63–66 (2014).
6. Karas, B. J. et al. Designer diatom episomes delivered by bacterial conjugation. *Nat. Commun.* **6**, 6925 (2015).
7. Miyahara, M., Aoi, M., Inoue-Kashino, N., Kashino, Y. & Ifuku, K. Highly efficient transformation of the diatom Phaeodactylum tricornutum by multi-pulse electroporation. *Biosci. Biotechnol. Biochem.* **77**, 874–876 (2013).
8. Apt, K. E., Grossman, A. R. & Kroth-Pancic, P. G. Stable nuclear transformation of the diatomPhaeodactylum tricornutum. *Mol. Gen. Genet.* **252**, 572–579 (1996).
9. Slattery, S. S. et al. An expanded plasmid-based genetic Toolbox Enables Cas9 genome editing and stable maintenance of synthetic pathways in Phaeodactylum tricornutum. *ACS Synth. Biol.* **7**, 328–338 (2018).
10. Slattery, S. S. et al. Phosphate-regulated expression of the SARS-CoV-2 receptor-binding domain in the diatom Phaeodactylum tricornutum for pandemic diagnostics. *Sci. Rep.* **12**, 7010 (2022).

11. Cochrane, R. R. et al. Rapid method for generating designer algal mitochondrial genomes. *Algal Res.* **50**, 102014 (2020).
12. Walker, E. J. L., Pampuch, M., Chang, N., Cochrane, R. R. & Karas, B. J. Design and assembly of the 117-kb Phaeodactylum tricornutum chloroplast genome. *Plant Physiol.* **194**, 2217–2228 (2024).
13. Russo, M. T., Rogato, A., Jaubert, M., Karas, B. J. & Falciatore, A. Phaeodactylum tricornutum: An established model species for diatom molecular research and an emerging chassis for algal synthetic biology. *J. Phycol.* **59**, 1114–1122 (2023).
14. Pampuch, M., Walker, E. J. L. & Karas, B. J. Towards synthetic diatoms: The Phaeodactylum tricornutum Pt-syn 1.0 project. *Curr. Opin. Green. Sustain. Chem.* **35**, 100611 (2022).
15. Taparia, Y. et al. A novel endogenous selection marker for the diatom Phaeodactylum tricornutum based on a unique mutation in phytoene desaturase 1. *Sci. Rep.* **9**, 8217 (2019).
16. Walker, E. J. L., Jaafar, T., Kaneshan, A. & Karas, B. J. Resolving replication incompatibility between chloroplast and conjugative plasmids in E. coli. 2025.10.16.682956 Preprint at https://doi.org/10.1101/2025.10.16.682956 (2025).
17. Kindle, K. L. High-frequency nuclear transformation of Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. USA* (1990).
18. Kassaw, T. K., Paton, A. J. & Peers, G. Episome-based gene expression modulation platform in the model diatom Phaeodactylum tricornutum. *ACS Synth. Biol.* **11**, 191–204 (2022).
19. Martino, A. D., Meichenin, A., Shi, J., Pan, K. & Bowler, C. Genetic and phenotypic characterization of Phaeodactylum tricornutum (Bacillariophyceae) accessions. *J. Phycol.* **43**, 992–1009 (2007).
20. Song, Z., Lye, G. J. & Parker, B. M. Morphological and biochemical changes in *Phaeodactylum tricornutum* triggered by culture media: Implications for industrial exploitation. *Algal Res.* **47**, 101822 (2020).
21. Avasthi, P., MacQuarrie, C. D. & Radkov, A. Inducing protoplast formation in *Phaeodactylum tricornutum* by silica deprivation, enzymatic treatment, or cytoskeletal inhibition. The Stacks https://doi.org/10.57844/arcadia-fh8f-xz51 (2023)
22. Ohse, M., Takahashi, K., Kadowaki, Y. & Kusaoke, H. Effects of plasmid DNA sizes and several other factors on transformation of Bacillus subtilis ISW1214 with plasmid DNA by electroporation. *Biosci. Biotechnol. Biochem* **59**, 1433–1437 (1995).
23. Szostková, M. & Horáková, D. The effect of plasmid DNA sizes and other factors on electrotransformation of *Escherichia coli* JM109. *Bioelectrochem. Bioenerg.* **47**, 319–323 (1998).
24. Cochrane, R. R. et al. Superior conjugative plasmids delivered by bacteria to diverse fungi. *BioDesign Res.* **2022**, 9802168 (2022).
25. Gibson, D. G. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res* **37**, 6984–6990 (2009).
26. Almagro Armenteros, J. J. et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* **2**, e201900429 (2019).
27. Diamond, A. et al. Instability of extrachromosomal DNA transformed into the diatom *Phaeodactylum tricornutum*. *Algal Res.* **70**, 102998 (2023).
28. Serif, M. et al. One-step generation of multiple gene knock-outs in the diatom Phaeodactylum tricornutum by DNA-free genome editing. *Nat. Commun.* **9**, 3924 (2018).
29. Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355–360 (1995).
30. Hwang, H.-J., Kim, Y. T., Kang, N. S. & Han, J. W. A simple method for removal of the chlamydomonas reinhardtii cell wall using a commercially available Subtilisin (Alcalase). *J. Mol. Microbiol. Biotechnol.* **28**, 169–178 (2018).
31. Woo, J. W. et al. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* **33**, 1162–1164 (2015).
32. Borowitzka, M. A. & Volcani, B. E. The Polymorphic Diatom Phaeodactylum Tricornutum: Ultrastructure Of Its Morphotypes. *J. Phycol.* **14**, 10–21 (1978).
33. Yin, W., Zhang, Y., Huang, Q., Pan, Y. & Hu, H. Genetic transformation of the freshwater diatom *Cyclotella meneghiniana* via bacterial conjugation. *Algal Res.* **84**, 103794 (2024).
34. Sprecher, B. N. et al. Genetic transformation methods for diatom *Nitzschia captiva*: New tools to better understand dinotom endosymbiosis. *Algal Res.* **72**, 103136 (2023).
35. Maeda, Y., Nakamura, M., Watanabe, K., Okamoto, E. & Tanaka, T. Functional analysis of the putative centromere sequences of marine oleaginous diatom *Fistulifera solaris*. *Algal Res.* **74**, 103225 (2023).
36. Angstenberger, M., Krischer, J., Aktaş, O. & Büchel, C. Knock-down of a *ligIV* homologue enables DNA integration *via* homologous recombination in the marine diatom *Phaeodactylum tricornutum*. *ACS Synth. Biol.* **8**, 57–69 (2019).
37. Slattery, S. S. et al. Plasmid-based complementation of large deletions in Phaeodactylum tricornutum biosynthetic genes generated by Cas9 editing. *Sci. Rep.* **10**, 13879 (2020).
38. Moosburner, M. A. et al. Multiplexed knockouts in the model diatom Phaeodactylum by episomal delivery of a selectable Cas9. *Front Microbiol* **11**, 5 (2020).
39. Karas, B. J. et al. Transferring whole genomes from bacteria to yeast spheroplasts using entire bacterial cells to reduce DNA shearing. *Nat. Protoc.* **9**, 743–750 (2014).
40. Rice, P., Longden, I. & Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000).
41. Sayers, E. W. et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **50**, D20–D26 (2022).

Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-06172 and RGPIN-2025-05428) awarded to BJK., and funding from the Natural Environment Research Council (NERC) grant NE/Z504130/1 awarded to TM. We would like to acknowledge Drs Graham Peers and Tessema Kassaw for sharing their knowledge on electroporation to *P. tricornutum* with us. Our explorations into this method began with a protocol that was shared by Dr. Kassaw, who assisted us with initial troubleshooting. We would also like to acknowledge Karen Nygard, a microscopy specialist at Western University's Biotron Facilities, who assisted with capturing the microscopic images shown in Fig. 1a, b.

Author contributions

M.P. and G.T. established the parameters necessary for conducting electroporation in the Karas lab and performed experiments 1 through 79 in Supplementary Data 1. E.J.W. established the spheroplasting technique for electroporation and the protoplasting technique for PEG transformation in *P. tricornutum*. EJW also performed all other *P. tricornutum* experiments detailed in the manuscript. L.D. and Y.L. performed the PEG transformation experiments in *T. pseudonana* and generated the images used for Supplementary Figs. 16 and 17. E.J.W. wrote the manuscript and created all other figures in Adobe Illustrator; M.P., L.D., Y.L., T.M., and B.J.K. edited the manuscript and provided crucial insight. B.J.K. supervised all experimentation done by E.J.W., M.P., and G.T., and assisted with experimental design, data analysis, troubleshooting, and more. T.M. supervised all experimentation done by L.D. and Y.L.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-68562-6.

Correspondence and requests for materials should be addressed to B. J. Karas.

Peer review information *Nature Communications* thanks Kentaro Ifuku and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026