Advanced genetic tools enable synthetic biology in the oleaginous microalgae *Nannochloropsis sp.*

Eric Poliner[1,2], Eva Farré[3], *Christoph Benning[2,3,4]

[1]Cell and Molecular Biology Program, Michigan State University, East Lansing, Michigan

[2]MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan

[3]Department of Plant Biology, Michigan State University, East Lansing, Michigan

[4]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan

*Corresponding author: Christoph Benning, benning@msu.edu, (517) 355-1609, orcid.org/0000-0001-8585-3667

Abstract

Nannochloropsis is a genus of fast-growing microalgae that are regularly used for biotechnology applications. *Nannochloropsis* species have a high triacylglycerol content and their polar lipids are rich in the omega-3 long-chain polyunsaturated fatty acid, eicosapentaenoic acid. Placed in the heterokont lineage, the *Nannochloropsis* genus has a complex evolutionary history. Genome sequences are available for several species, and a number of transcriptomic datasets have been produced, making this genus a facile model for comparative genomics. There is a growing interest in *Nannochloropsis* species as models for the study of microalga lipid metabolism and as a chassis for synthetic biology. Recently, techniques for gene stacking, and targeted gene disruption and repression in the *Nannochloropsis* genus have been developed. These tools enable gene specific, mechanistic studies, and have already allowed the engineering of improved *Nannochloropsis* strains with superior growth, or greater bioproduction.

Keywords

Nannochloropsis, algal biotechnology, marker-free engineering, gene stacking, synthetic biology, episomes

Key Message

Nannochloropsis is emerging as a facile chassis for synthetic biology of microalgae and beyond.

Introduction

Algae are highly efficient at turning solar energy into biomass, and are sources of unique bioproducts, such as omega-3 fatty acids (Mühlroth et al. 2013; Zou et al. 2000), carotenoids (Yaakob et al. 2014), and interesting polysaccharides, such as agarose, alginate, and β-1,3-glucans(Chew et al. 2017; Sheehan et al. 1998; Yaakob et al. 2014). Several groups have screened algae for productivity and production of valuable compounds (Rodolfi et al. 2009; Sheehan et al. 1998; Unkefer et al. 2017). *Nannochloropsis* was identified as a genus with rapid growth, and high lipid content, including triacylglycerol (TAG) (Rodolfi et al. 2009) and the omega-3 (*ω3*) long-chain polyunsaturated fatty acid, eicosapentaenoic acid (EPA) (Zou et al. 2000). Under nutrient-replete conditions *Nannochloropsis* species have a lipid content of approximately 25-30% of dry-weight (Jia et al. 2015; Meng et al. 2015; Rodolfi et al. 2009; Xiao et al. 2015). Abiotic stresses, such as high-light or nutrient deprivation, in particular nitrogen (N) deprivation, cause microalgae to pause growth and accumulate storage compounds. Under these conditions, *Nannochloropsis* species accumulate high quantities of TAG, up to 60% of biomass (Jia et al. 2015; Meng et al. 2015; Rodolfi et al. 2009; Simionato et al. 2013; Vieler et al. 2012; Xiao et al. 2015). In recent years, genomes for several *Nannochloropsis* species have become available (Table 1) (Corteggiani Carpinelli et al. 2013; Radakovits et al. 2012; Vieler et al. 2012; Wang et al. 2014) and molecular tools have been developed (Table 2-4) (Ajjawi et al. 2017; Kilian et al. 2011; Poliner et al. 2017; Wei et al. 2017c), making species in this genus excellent microalgal models for comparative genomics (Hu et al. 2014; Wang et al. 2014).

Synthetic biology is an emerging field based on rationally designing biological systems (Andrianantoandro et al. 2006). To develop systems that behave as desired, an approach described as design, build, test is used to iteratively test refinements and determine how elements of the system influence the outcome (Agapakis 2014). The design phase is often based on information drawn from genome-wide data and databases of related systems. Organisms with high-quality genome-wide data and advanced genetic engineering tools that can be redesigned are known as chassis organisms. Metabolic maps are built using genome assemblies, functional annotation, and databases of known enzymatic pathways. Regulatory networks are coming into focus through integrating RNA-seq, chromatin immunoprecipitation DNA sequencing (CHIP-Seq), and databases of transcription factors and their target DNA motifs. In order to build or refine biosynthetic pathways or develop chassis organisms, several molecular tools are needed to modify

the genome of an organism. While molecular tools, such as mutant libraries, transgenic overexpression, and reporter protein fusions are instrumental in gaining a molecular understanding of biological processes, they are by themselves insufficient to create optimized biological systems. Redesigned algae will require a new generation of tools that enable precise and marker-free knockout mutants, and high-capacity gene stacking systems that can robustly and predictably express multiple genes. Finally, in order to test the synthesized system, highly facile methods to select or screen for the desired modifications as quickly as possible post transformation are needed.

The *Nannochloropsis* genus is also an emerging algal model for genetic engineering of lipid accumulation (Ajjawi et al. 2017; Hu et al. 2014; Li et al. 2014b). Several *Nannochloropsis* species seem particularly amenable to transgene expression, with a moderate GC (guanine cytosine) content and simple gene structure facilitating genetic engineering (Jinkerson et al. 2013; Vieler et al. 2012; Wang et al. 2014). Several endogenous promoters and terminators are in use, including bidirectional promoters that are helpful in stacking transgenes, i.e. expressing multiple genes (Jinkerson et al. 2013; Kilian et al. 2011; Moog et al. 2015; Nobusawa et al. 2017; Poliner et al. 2017). Methods exist for targeted DNA insertion by homologous recombination into the genome (Dolch et al. 2017; Kilian et al. 2011; Nobusawa et al. 2017). CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) based methods have been developed, making targeted gene disruption and editing possible (Ajjawi et al. 2017; Wang et al. 2016). Genetic engineering toolkits are becoming publicly available and should accelerate development of *Nannochloropsis* species as chassis organisms.

Taxomony and Evolution of the *Nannochloropsis* genus

Nannochloropsis is a genus in the heterokont phylum. Photosynthetic heterokonts are secondary endosymbionts originating from an unicellular heterotrophic eukaryotic cell that engulfed a red alga, which, over time, became a plastid. The evolutionary relationships of organisms in this clade are complex and still not well defined. The red-algal type plastid shows several signatures of its origin distinct from green lineage plastids (Janouskovec et al. 2010; Keeling 2009; Wei et al. 2013). Brown algae and diatoms are also algae of the heterokont lineage that share 57% and 51% of genes with *N. gaditana* CCMP526 genome, respectively (Radakovits et al. 2012).

The *Nannochloropsis* genus was established by Hibberd based on morphological characteristics (Hibberd 1981; Lubián 1982) and constituent species identified (*oceanica*,

granulata, *limnetica*, *salina*, *gaditana*) based on the 18S ribosomal RNA (18S rDNA) and/or the plastid genome RuBisCO large subunit-encoding gene (*rbcL*) (Andersen et al. 1998; Fawley et al. 2015; Suda et al. 2002; Vieler et al. 2012). Whole genome phylogeny of six species of *Nannochloropsis* showed *N. gaditana* and *N. salina* to be closely related and separate from *N. oceanica*, *N. granulata*, and *N. oculata* (Wang et al. 2014). Fawley et al. proposed a separate genus (*Microchloropsis*) for *N. gaditana* and *N. salina* (Fawley et al. 2015); however in this review the *Nannochloropsis* genus will be inclusive to *N. salina* and *N. gaditana*.

The *Nannochloropsis* genus seems to possess a number of nuclear genes derived from endosymbiotic gene transfer, particularly in lipid biosynthetic and carbohydrate degradation pathways, e.g. glycosyl hydrolases (Wang et al. 2014). The diverse genetic background of the *Nannochloropsis* genus may have contributed to its oleaginousness, with a particularly large set of putative lipid biosynthetic genes. For example, *Nannochloropsis oceanica* possess 11 type-2 diacylglycerol acyltransferases (DGATs) referred to also as DGTTs (Vieler et al. 2012; Wang et al. 2014; Xin et al. 2017; Zienkiewicz et al. 2017), and a high copy number of other predicted lipid biosynthetic genes such as: enoyl-ACP reductase (ENR), ketoacyl-ACP synthase (KAS), ketoacyl-ACP reductase (KAR), acyl-ACP thioesterase (TE), long-chain fatty acyl-CoA synthetase (LC-FACS), phosphatidic acid phosphatase (PAP), and lysophosphatidyl acyltransferase (LPAT) (Wang et al. 2014). Of these, only DGTTs (Li et al. 2016a; Wei et al. 2017a; Xin et al. 2017; Zienkiewicz et al. 2017), and LPATs (Nobusawa et al. 2017) have been characterized. It has been proposed that the ancestral heterotroph, the endosymbiotic red alga, and additional horizontal gene transfer contributed to the present genome (Wang et al. 2014).

The secondary endosymbiosis event also led to interesting cellular structure characteristics, such as four membranes surrounding the plastid, complicating intracellular trafficking (Keeling 2009; Kroth and Strotmann 1999; Murakami and Hashimoto 2009). Trafficking of nuclear encoded proteins into the plastid has been extensively studied in diatoms (Bolte et al. 2009), which led to the specialized protein location prediction software HECTAR (Gschloessl et al. 2008). While in several *Nannochloropsis* species examples of protein localization by fluorescent protein (FP) fusions have been published, including to the plastid (Moog et al. 2015), endoplasmic reticulum (ER) (Gee and Niyogi 2017; Poliner et al. 2017), mitochondria (Ma et al. 2017) and lipid droplets (Nobusawa et al. 2017; Zienkiewicz et al. 2017), there have not been investigations into the signals and mechanisms of protein localization. The transport of metabolites across subcellular

compartments also has not been studied in the *Nannochloropsis* genus, although it represents a plausible target for optimizing metabolite production (Loira et al. 2017).

Genetic diversity has also arisen frequently by horizontal gene transfer during algae evolution (Bowler et al. 2008; Diner et al. 2017), and has enabled adaption to unique environments or metabolic niches (Schonknecht et al. 2013). Several lipid biosynthetic genes appear to be related to bacterial homologs and likely were acquired by horizontal gene transfer, including KAR, PAP, ENR, KAS, TE, and LC-FACS genes (Wang et al. 2014). An operon encoding proteins specialized in hydrogen generation possibly derived from bacteria has also been identified in *N. oceanica* (Vieler et al. 2012). Transkingdom gene transfer by bacterial conjugation to diatoms and the capacity of heterokonts to maintain episomal DNA indicates a possible route of gene acquisition (Diner et al. 2017; Karas et al. 2015). This evolutionary diversity of the *Nannochloropsis* genus and the genetic plasticity of the heterokonts provide an interesting model for symbiotic evolution and may lead to a chassis organism that can be robustly adapted for genetic engineering.

Genomes and transcriptomes across the *Nannochloropsis* genus

A complete genome sequence forms the foundation for gene-specific studies, and is a prerequisite for the drafting of metabolic and regulatory maps. Several genome assemblies have been generated for different species of *Nannochloropsis,* including multiple strains within a species, such as, *N. oceanica* CCMP1779 (Vieler et al. 2012) and IMET1 (Wang et al. 2014), *N. gaditana* CCMP526 (Radakovits et al. 2012; Wang et al. 2016) and B-31 (Corteggiani Carpinelli et al. 2013), *N. salina* CCMP537, *N. oculata* CCMP525 (Wang et al. 2014), and *N. granulata* CCMP529 (Table 1) (Wang et al. 2014). The genomes of the examined *Nannochloropsis* species are approximately 30 megabases, and contain 7-11,000 genes each (Radakovits et al. 2012; Vieler et al. 2012; Wang et al. 2014). The *N. gaditana* B-31 genome is estimated to be distributed over 30 chromosomes (Corteggiani Carpinelli et al. 2013), and the presumed *N. oceanica* IMET1 chromosomes separated by pulse-field gel electrophoresis as 22 individual genome fragments (Wang et al. 2014).

The plastid and mitochondrial genomes of representatives from five *Nannochloropsis* species were used to produce a pangenome of each organelle (Wei et al. 2013). The *N. oceanica* IMET1 plastid genome is 117,548 basepairs (bp) and contains 160 genes consisting of 126 protein-coding genes, and 34 RNA genes, and the mitochondrial genome is 38,057 bp and contains 63 genes consisting of 35 protein-coding genes and 28 RNA genes (Wei et al. 2013). The

158 *Nannochloropsis* genus plastid pangenome contains signatures of a red algal origin, including red
159 algal-type Rubisco and Rubisco activase genes (Starkenburg et al. 2014).
160 Extensive transcriptome data based on RNA-sequencing of cells grown under different
161 conditions reveal characteristic transcriptional changes, providing a whole-genome view of
162 possible adjustments to maintaining homeostasis. Examined conditions include phosphorus
163 (Mühlroth et al. 2017) and N deprivation (Corteggiani Carpinelli et al. 2013; Li et al. 2014b; Vieler
164 et al. 2012), alternating light:dark cycles (Poliner et al. 2015), varying light intensities (Alboresi
165 et al. 2016), and different growth phases of batch cultures (Radakovits et al. 2012) for various
166 *Nannochloropsis* species (Table 1). These datasets suggest that different aspects of metabolism
167 and other cellular processes, such as the cell cycle are coordinated on a transcriptional level in
168 response to environmental conditions. For example, N deprivation, which generally leads to a
169 transition from the normal cell division cycle to quiescence, also causes transcriptional
170 downregulation of photosynthesis and protein production, while lipid biosynthesis is upregulated
171 as was observed for *Nannochloropsis* species (Corteggiani Carpinelli et al. 2013; Li et al. 2014b;
172 Radakovits et al. 2012) and *Chlamydomonas* (Miller et al. 2010; Tsai et al. 2014). When *N.*
173 *oceanica* is grown in a light:dark cycle, there is phased expression of certain genes at different
174 times of day, including those involved in cell division at night, and anabolic processes during the
175 day (Poliner et al. 2015). A majority (64%) of the DNA-binding transcription factors and 56% of
176 other transcriptional regulators have phased expression during light:dark cycles. These genome-
177 wide datasets are an asset for further studies into metabolic adjustments occurring in response to
178 environmental changes and the underlying regulation. Links to currently published genome-wide
179 datasets are listed in Table 1.
180
181 **Light regulation and photosynthesis**
182 In heterokont algae, light, either by capture and conversion of solar energy during photosynthesis
183 or by perception through photosensory regulatory proteins, affects metabolite levels (pigments,
184 lipids, and carbohydrates) (Chauton et al. 2013; Poliner et al. 2015), coordinates the cell cycle
185 (Ashworth et al. 2013; Chauton et al. 2013; Huysman et al. 2013; Huysman et al. 2010; Poliner et
186 al. 2015), and may entrain a circadian clock (Braun et al. 2014). In *Nannochloropsis* species, high-
187 intensity light results in accumulation of TAG and a decrease in plastid size, thus maximizing
188 energy conversion while avoiding photodamage (Alboresi et al. 2016; Sukenik et al. 1989; Xiao

et al. 2015). The day:night cycle influences most organisms to coordinate behavior and/or metabolism with either phase (Ashworth et al. 2013; Chauton et al. 2013; Poliner et al. 2015). The transitory storage compounds used by *N. oceanica* during a light:dark cycle are under study with TAG and carbohydrates (measured in the form of hexoses) oscillating throughout a light:dark cycle (Fábregas et al. 2002; Poliner et al. 2015; Sukenik and Carmeli 1990). *Nannochloropsis* species accumulate TAG (Sukenik and Carmeli 1990) and carbohydrates during the day, which are both metabolized during the night, in accordance with transcriptional changes in genes encoding enzymes of the respective biosynthesis and utilization pathways (Poliner et al. 2015).

Nannochloropsis species are studied as a model for photosynthesis in secondary endosymbionts. The *Nannochloropsis* genus is notable for only possessing chlorophyll *a*, the unusual carotenoids violaxanthin and vaucheriaxanthin ester, and a xanthophyll cycle utilizing violaxanthin, antheraxanthin, and zeaxanthin (Alboresi et al. 2017; Cao et al. 2013; Chukhutsina et al. 2017). Red algae and their derived endosymbionts contain LHCr type antenna proteins that link core complex pigment protein components, and participate in energy transfer and photoprotection (Alboresi et al. 2017; Cao et al. 2013; Umetani et al. 2017). Characterization of the photosystem II (PSII) of *N. gaditana* identified the light harvesting complex proteins of the classes LHCx, LHCf, Red-CLH-like LHC, and LHCr that are characteristic of the red alga-type plastid (Umetani et al. 2017). Characterization of the *N. gaditana* photosystem I (PSI) discovered the absence of several subunits (PsaH, PsaK, PsaG) that are typically present in land plants, and identified the light harvesting complex proteins of the classes LHCr, LHCf, and LHCx associated with PSI (Alboresi et al. 2017).

Lipid and carbon metabolism

Lipid biosynthesis is the best characterized metabolic pathway in *Nannochloropsis* species, in particular the production of TAG and EPA. The *Nannochloropsis* genus is hypothesized to possess a cytosolic type-I fatty acid synthases (FAS) in addition to the plastid type-II FAS complex, but further studies to corroborate this hypothesis are needed (Alboresi et al. 2016; Poliner et al. 2015; Vieler et al. 2012). The TAG biosynthetic pathway involves the transfer of acyl chains to a glycerol backbone by the sequential action of glycerolphosphate acyltransferase (GPAT), LPAT, and DGAT, in addition phospholipidiacylglycerol acyltransferases (PDATs) have also been identified in *Nannochloropsis* species (Vieler et al. 2012). The four LPATs of *N. oceanica* have been

investigated for their roles in membrane lipid and TAG biosynthesis with LPAT1 and LPAT4 having primary roles in each process respectively, and LPAT2 and LPAT3 possibly playing roles in both processes (Nobusawa et al. 2017). Of the 13 DGAT-encoding genes, 6 are upregulated during N deprivation, a condition that also favors TAG accumulation (Zienkiewicz et al. 2017). A particularly robust DGAT of *Nannochloropsis* that functions in many different hosts has been identified (Zienkiewicz et al. 2017).

The *Nannochloropsis* genus contains the omega-3 fatty acid EPA in its membrane lipids (15-30% total fatty acids (TFA)) (Schneider et al. 1995; Schneider and Roessler 1994; Vieler et al. 2012; Zou et al. 2000). Reconstruction of the *N. oceanica* EPA biosynthetic pathway in *S. cerevisiae* by introducing four LC-PUFA fatty acid desaturases (FADs) and a fatty acid elongase (FAE), resulted in the production of EPA (0.1% TFA) (Poliner et al. 2017). FADs are named for the double bond introduced, a specific number of carbons from either the carboxyl (Δ, delta-) or methyl (*ω,* omega-) end of a fatty acid chain. Thus, omega-3 and delta-6 FADS act on the third carbon from the methyl end and the sixth carbon from the carboxyl end, respectively. The FADs of *N. oceanica* resemble those of other heterokont algae with their histidine box motifs for coordinating a diiron center, and contain in two cases (delta-5 and delta-6) a cytochrome b domain (Poliner et al. 2017). Eleven fatty acid elongases have been identified in *N. oceanica* (Vieler et al. 2012) and the delta-6 and palmitic fatty acid specific elongases from *N. oceanica* and *N. gaditana,* respectively, have been characterized in some detail. The palmitic acid elongase controls flux into the EPA pathway by conversion of 16:0 to 18:0 (Dolch et al. 2017), while the delta-6 elongase converts 18:3 to 20:3, two intermediates with low *in vivo* abundance (Poliner et al. 2017). EPA is likely produced in the ER but accumulates on diacylglyceryl-trimethylhomoserine (DGTS) and monogalactosyl diacylglycerol (MGDG), and to a lesser extent on digalactosyldiacylglycerol (DGDG) and phosphatidylglycerol (PG) (Simionato et al. 2013; Vieler et al. 2012). It has been proposed that EPA is imported into the plastid by a DGTS mediated transport (Schneider and Roessler 1994).

Carbohydrates play structural, storage, and osmoprotectant roles in *Nannochloropsis* species. In *N. oceanica* glucose is the predominant hexose in the total complex carbohydrate fraction, which contains smaller amounts of mannose, and trace amounts of rhamnose, fucose, arabinose, xylose, and galactose (Vieler et al. 2012). Marine *Nannochloropsis* species reduce their level of the sugar alcohol mannitol and the disaccharide trehalose content in response to low-salt

stress consistent with a role of these carbohydrates in osmoprotection (Pal et al. 2013). Heterokont algae lack starch but produce β-1,3 linked polysaccharides (chrysolamarinin), by the activity of a β-1,3-glucan synthase which is predicted to be encoded in the genome of *Nannochloropsis* species (Corteggiani Carpinelli et al. 2013; Vieler et al. 2012). Approximately 20% of alcohol insoluble polysaccharides are in this form in *N. oceanica* (Vieler et al. 2012). Chrysolamarinin is also a storage compound in diatoms (Chauton et al. 2013; Hildebrand et al. 2017), and has been suggested to have a similar role in *Nannochloropsis* species (Arnold et al. 2015; Li et al. 2014b; Xiao et al. 2015), but further studies are needed to confirm this hypothesis. Cellulose is a major polysaccharide in *Nannochloropsis* species, with approximately 80% of the of alcohol insoluble polysaccharides in this form in *N. oceanica* (Vieler et al. 2012). Cellulose serves as a major component of the cell wall (Jeong et al. 2017; Scholz et al. 2014; Vieler et al. 2012), but the cell wall is quite complex in *Nannochloropsis* species (Scholz et al. 2014). Four putative cellulose synthase-encoding genes have been identified (Jeong et al. 2017; Scholz et al. 2014). A large number of carbohydrate-degrading enzymes, 48-49 glycosyl hydrolases, with very diverse taxonomic relations, are found encoded across the pangenome (Wang et al. 2014). The complete repertoire of carbohydrate metabolism in the *Nannochloropsis* genus has yet to be fully established.

In order to understand the metabolic networks of *Nannochloropsis* species, a summary of possible chemical reactions in the form of a metabolic map has been generated. A mass-balanced metabolic map for *N. salina* CCMP537 has been produced by Loira and colleagues taking into account 9 organelles (as well as the plastid lumen) of the cell (Loira et al. 2017). The model was validated by modeling different growth conditions and comparing it to *in vivo* data. The conditions of N and phosphate deprivation were used to maximize lipid production and determine essential nutrients, respectively. The iNS934 map-based model indicated several genes whose disruption may result in increased TAG.

Transcriptional Regulation

The dynamic control of metabolism in response to environmental changes and intracellular cues is multilayered but inevitably involves transcription factors (TF, possess DNA binding activity) and transcriptional regulators (TR, regulators of TF activity) that modulate gene expression. Databases of known TFs and corresponding position-weighted matrices (PWM) (Stormo 2000)

can be used to systematically determine potential interacting TF-DNA sequences. Having a large number of closely related organisms is a valuable resource for *in silico* predictions of conserved regulatory DNA motifs (Hu et al. 2014). Genome sequencing and cataloging of TFs as a first step has been undertaken for the *Nannochloropsis* genus (Hu et al. 2014; Vieler et al. 2012). Comparative genomic studies of the heterokont lineage have identified TF signatures based upon organismal lifestyle (autotrophic, parasitic), multicellularity, or lifecycle stages (Buitrago-Flórez et al. 2014; Rayko et al. 2010; Thiriet-Rupert et al. 2016). TF prediction of *N. oceanica* implied the presence of 115 putative TFs and 109 putative TRs, which combined represent about 2% of the predicted proteins encoded in the genome (Vieler et al. 2012). The *Nannochloropsis* genus has a reduced number of TF and TR families (20-26) compared to land plants and green algae, possibly due to a simpler lifecycle and its unicellularity (Hu et al. 2014; Vieler et al. 2012). Putative TFs of the Myb family (29-35 members), a TF family known to regulate growth and metabolism in other organisms, are enriched in the *Nannochloropsis* genus (Hu et al. 2014; Vieler et al. 2012). Several TFs in *Nannochloropsis* species have been investigated but studies into their targets are only beginning. Despite the large amounts of predicted TFs and TRs, their roles need to be experimentally corroborated.

Several approaches have been used to identify regulators of lipid biosynthesis in *Nannochloropsis* species. Hu and colleagues (Hu et al. 2014) took advantage of the extensive genome sequences available to predict conserved transcription factor binding sites (TFBS). They determined the enrichment of gene ontology (GO) terms associated with each motif, and enrichment of motifs associated with lipid biosynthetic genes. Using a TF catalogue and RNA-seq during N deprivation, Hu et al. identified TFs that showed positive or negative co-expression with lipid biosynthetic genes. Finally, they predicted putative connections between TFs and lipid biosynthetic gene promoters based on a TF-DNA motif database (Wingender 2008). One of these predicted lipid biosynthesis regulating TFs (bZIP1) was recently investigated (Kwon et al. 2017). In addition, Ajjawi and colleagues (Ajjawi et al. 2017) identified 20 TFs possibly involved in TAG accumulation based on changes in expression under N deprivation, and used CRISPR/Cas9 disruption to assess the predictions.

As a photosynthetic organism, light sensing is likely important for tuning metabolism in *Nannochloropsis* species. Aureochromes are heterokont specific photosensitive transcription activators, with a bZIP DNA binding domain and a photosensing dimerization LOV domain

(Vieler et al. 2012). In diatoms the aureochromes are implicated in regulating several processes, including cell cycle and light acclimation (Huysman et al. 2013; Mann et al. 2017; Schellenberger Costa et al. 2013). Cryptochromes are blue light photoreceptors derived from DNA repair enzymes, which have been characterized in diatoms (Coesel et al. 2009) and found to oscillate during light:dark cycles in diatoms (Ashworth et al. 2013) and *N. oceanica* (Poliner et al. 2015). Recently, the *Chlamydomonas reinhardtii* animal-like cryptochrome was reported to be involved in regulating the cell cycle and the circadian clock (Müller et al. 2017; Zou et al. 2017). Finally, although phytochromes were thought to be absent in most heterokonts, they have recently been identified in diatoms (Fortunato et al. 2016). *N. oceanica* possesses three aureochromes, an animal-like cryptochrome, and lacks phytochromes, but these proteins have not been functionally characterized (Vieler et al. 2012).

Transformation and gene expression platforms

The most widely adapted method for transformation of *Nannochloropsis* species is by electroporation (Kilian et al. 2011; Radakovits et al. 2012; Vieler et al. 2012), but other protocols based on biolistics (Kang et al. 2015a; Kang et al. 2015b) or agrobacterium have been developed (Beacham and Ali 2016; Cha et al. 2011) (Table 2). For insertion of a transgene into the genome by electroporation, a linear piece of DNA is required; a constructed DNA (construct) is therefore digested with restriction enzymes or PCR-amplified (Kilian et al. 2011; Li et al. 2014a; Poliner et al. 2017; Vieler et al. 2012). Each transformant is likely to have a distinct insertion site, and therefore may display different phenotypes due to genome context-specific regulation of the transgene's expression or disruption of endogenous genes by transgene insertion (Cha et al. 2011).

Introduction of circular DNA has so far had mixed success in producing transformants (Kilian et al. 2011; Li et al. 2014a; Vieler et al. 2012). Synthetic Genomics Inc. has described plasmid/episome maintenance in algae by use of autonomous replication sequences from a *Nannochloropsis* species (Dehoff et al. 2014), and the utilization of an *S. cerevisiae* centromere (autonomous replication sequence, CEN/ARS) region in pennate and centric diatoms (Diner et al. 2016a; Diner et al. 2017; Karas et al. 2015). In diatoms, episomes are maintained under antibiotic selection but are gradually lost without selection pressure (Diner et al. 2016a; Diner et al. 2017; Karas et al. 2015). The transgene expression levels from episomes are more uniform compared to genome integrated constructs between independent transformants, likely due to the absence of

insertion site-specific effects using this approach (Diner et al. 2016a; Karas et al. 2015). The capacity of episomes for maintenance of foreign DNA have been reported to be up to 94 kilobases (kb) in diatoms (Karas et al. 2015).

Antibiotic resistance marker genes

Several antibiotics are effective depending on the *Nannochloropsis* strain, including hygromycin, zeocin, and blastidicin, and genes conferring resistance are used to isolate *Nannochloropsis* transformants (Ajjawi et al. 2017; Kilian et al. 2011; Nobusawa et al. 2017; Radakovits et al. 2012; Vieler et al. 2012). Zeocin in combination with its respective marker gene is the most widely used selection in *Nannochloropsis* species due to its stringency at low concentrations (Table 2-4). However, it is mutagenic and can lead to secondary mutations (Lin et al. 2017). In diatoms, the selection agents nourseothricin and G418 are frequently utilized with their respective antibiotic resistance genes (Hildebrand et al. 2017; Karas et al. 2015; Zaslavskaia et al. 2001). Mutated endogenous proteins in conjunction with competitive inhibitors, such as phytoene desaturase and the inhibitor norflurazon can also be used as a selection marker in some algae (Huang et al. 2008). We have made a number of vectors containing *N. oceanica*-adapted antibiotic selection marker genes that will be available on Addgene (www.addgene.com). When several selection agents and resistance genes are available, multiple transgenic tools can be used in conjunction in one transgenic line (Fig. 1a). However, techniques for generation of transgenic algae without antibiotic resistance markers are necessary for deployment into open ponds. The removal of an antibiotic resistance marker gene could be achieved by the use of a recombinase or endogenous homologous recombination (Cheah et al. 2013). The cotransformation of an episome carrying an antibiotic resistance gene with an insertion construct without a selection marker may also enable generation of marker-free mutants after episome loss in the absence of selection pressure.

Transgenic expression in *Nannochloropsis* species

A fundamental technique for genetic engineering and synthetic biology is the overexpression of target genes by increasing transcriptional and/or translational efficiency. Most often, strong promoters that mediate high transcription rates are utilized to express heterologous or endogenous genes at elevated levels. Several endogenous promoters from a variety of *Nannochloropsis* species have been isolated and applied to transgenic expression (Table 2-4), including those driving the

genes encoding ubiquitin extension protein (UEP) (Dolch et al. 2017; Kang et al. 2015a; Radakovits et al. 2012; Wei et al. 2017a), β-tubulin (β-tub) (Kang et al. 2015a; Li et al. 2014a; Ma et al. 2017; Radakovits et al. 2012; Wang et al. 2016), lipid droplet surface protein (LDSP) (Kaye et al. 2015; Nobusawa et al. 2017; Poliner et al. 2017; Vieler et al. 2012), and elongation factor (EF) (Poliner et al. 2017; Zienkiewicz et al. 2017). Several bidirectional promoters are utilized for transgenic expression, including those driving the expression of the genes encoding violaxanthin chlorophyll binding proteins (VCP) (Kilian et al. 2011; Ma et al. 2017; Moog et al. 2015) or ribosomal subunits (Ribi) (Fig. 1b) (Poliner et al. 2017). To enhance translational efficiency of transgenes, a 5' UTR can include a consensus Kozak sequence (Dehoff and Soriaga 2014) or leader-enhancing sequence (Gallie et al. 1987; Li et al. 2016a; Xue et al. 2015).

Reporter genes are useful for evaluating transgenic strategies and understanding gene/promoter function. In several *Nannochloropsis* species, members of the classes of the fluorescent- (FP), luminescent- (lux), and chromoproteins (CP) are used (Table 2-4). Green fluorescent protein (GFP) is the most widely reported fluorescent protein and has been utilized for subcellular localization of fusion proteins throughout *Nannochloropsis* cells (Ma et al. 2017; Moog et al. 2015; Nobusawa et al. 2017). Other fluorescent proteins such as a red fluorescent protein (RFP, sfCherry) (Kang et al. 2015a), yellow FP (YFP, Venus variant) (Gee and Niyogi 2017; Nobusawa et al. 2017; Zienkiewicz et al. 2017), and cyan FP (CFP, Cerulean variant) (Poliner et al. 2017) are employed in different *Nannochloropsis* species. Luciferases have the advantage of a high signal to noise ratio, and specific substrates allow their use in combination. Codon-optimized firefly luciferase (Flux) and the ultra-bright NanoLuciferase (Nlux) for *in-vivo* assays in *N. oceanica* are in use (Poliner et al. 2017). The ultra-bright Nlux allows detection of very low protein quantities and is an effective photon donor for bioluminescence resonance energy transfer (BRET) (Hall et al. 2012; Suzuki et al. 2016). Chromoproteins are colored and do not need a substrate, while the β-Glucuronidase (GUS) reporter is an enzyme that produces a blue stain after conversion of 5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid (X-Gluc). A purple chromoprotein (shPCP) from the sea anemone *Stichodacyla haddoni* was successfully produced in *N. oculata* and used for screening of transformants (Shih et al. 2015), while GUS was utilized in *N. salina* (Li et al. 2014a).

Virus-derived 2A peptides allow production of two discrete proteins from a single transcript by preventing the formation of a peptide bond (peptide bond "skipping") (Fig. 1c). The

nascent protein interacts with the ribosome causing it to stall, then during the release of the ribosome a peptide bond is "skipped" between the final two amino acids (Sharma et al. 2012). 2A peptides are widely applied to different hosts and the efficiency of skipping depends on the peptide variant and host. After screening several variants and lengths in *N. oceanica*, the P2A peptide of 60 amino acids was found to be most efficient (Poliner et al. 2017). The F2A peptide has also been used in *N. salina* (Unkefer et al. 2017) and *Chlamydomonas* (Plucinak et al. 2015). In order to facilitate multigene expression, bidirectional promoters with P2A peptide sequences were assembled in the pNOC-stacked vector series and will be made available through Addgene (Poliner et al. 2017). Toolkits for assembly of multiple transgenic expression cassettes into a single vector have been developed for some algae (Hamilton et al. 2014; Jia et al. 2016), and would further facilitate gene stacking in *Nannochloropsis* species (Fig. 1d).

Generation of targeted gene disruption and transcriptional repression

RNA interference (RNAi) is a powerful technique that can suppress gene expression to varying degrees. RNAi by antisense or double-stranded RNA has been developed in a number of algae including *Chlamydomonas* (Moellering and Benning 2010; Rohr et al. 2004), diatoms (De Riso et al. 2009; Hildebrand et al. 2017; Schellenberger Costa et al. 2013), and several *Nannochloropsis* species (Ajjawi et al. 2017; Ma et al. 2017; Wei et al. 2017a; Wei et al. 2017c). RNAi using an inverted repeat of regions of a target gene has been found to be effective throughout the *Nannochloropsis* genus (Fig. 2a) (Table 4) (Ajjawi et al. 2017; Ma et al. 2017; Wei et al. 2017a; Wei et al. 2017c). In order to generate a strong and stable repression effect, several strategies are available including fusing the interfering RNA to an antibiotic resistance gene (De Riso et al. 2009; Wei et al. 2017c), co-silencing a gene that can be counter selected (Rohr et al. 2004), or expression of the interfering RNA from a bidirectional promoter that also drives expression of an antibiotic resistance gene. In case of essential genes, full disruption of the target gene can result in slow growth or lethality while transcriptional repression targeting the same gene may result in a moderate phenotype (Ajjawi et al. 2017).

Homologous recombination has been demonstrated in several *Nannochloropsis* species for insertion of an antibiotic resistance marker with flanking regions of 1 kb identical to the insertion site (Fig. 2b). Targeting efficiency has been reported to be high when transforming low-density cultures (Kilian et al. 2011). Several groups have used this technique for gene disruption (Table 4)

(Dolch et al. 2017; Gee and Niyogi 2017; Kilian et al. 2011; Nobusawa et al. 2017). Homologous recombination is adaptable to several purposes, such as, insertion of protein tags, insertion into neutral sites, or replacement of genes with altered functionality.

CRISPR/Cas9 is an RNA-guided nuclease-based approach that is dramatically expanding the capabilities of biologists to modify genomes, particularly through the ability to disrupt specific genes or perform precise gene editing (Fig. 2c) (Cong et al. 2013; Ran et al. 2013). In this system, two components have to be localized to the nucleus, a single-guide RNA (sgRNA) and the Cas9 nuclease, which together form a ribonucleoprotein complex. The sgRNA requires production without extraneous sequences or modifications on the termini; strategies for sgRNA production include the use of RNA polymerase III-driven promoters (most often the U6 promoter) (Nymark et al. 2016), direct introduction into the cell (Ajjawi et al. 2017), expression of modified tRNAs containing the sgRNA in a spliced region (Cermak et al. 2017), co-expression of a ribonuclease and sgRNA with cleavage sites (Cermak et al. 2017), and use of self-cleaving ribozymes (Cermak et al. 2017). The U6 promoter from diatoms appear to be active and suitable for sgRNA production (Nymark et al. 2016). However, there have been no reports of successful U6 promoter use in the *Nannochloropsis* genus, and two publications utilized alternative strategies (Table 4). Wang and colleagues (Wang et al. 2016) expressed the sgRNA from an V-ATPase promoter and had a low mutational efficiency. In a strategy developed by Ajjawi and coworkers (Ajjawi et al. 2017), the sgRNA is synthesized and introduced by transformation into a *Nannochloropsis* strain expressing Cas9. The generation of off-target mutations is an unresolved issue for CRISPR-based gene editing, which has not been examined in the *Nannochloropsis* genus. Several strategies exist to reduce the number of potential off-targets, including transient expression of one or both components (Ajjawi et al. 2017; Baek et al. 2016; Xie et al. 2017), paired nickase Cas9 (Ran et al. 2013), or high-fidelity Cas9 enzymes (Kleinstiver et al. 2016; Slaymaker et al. 2016).

Insertional mutatagenesis, whereby an antibiotic resistance gene is randomly integrated into a genome, results in gene disruptions and gene deletions (Figure 2d) (Li et al. 2016b; Tsai et al. 2014). Insertional mutagenesis conducted in *N. gaditana* produced mutants with a variety of growth and photosynthetic phenotypes, screening of which identified lines with enhanced light-use efficiency (Perin et al. 2015). While insertional mutagenesis is an efficient method for forward-genetic screens, ready-to-use lines with a disruption of a desired gene requires a mutant library that takes a significant investment to establish (Li et al. 2016b).

Altering metabolism in *Nannochloropsis* species by protein engineering

The usefulness of the aforementioned genetic engineering tools has been demonstrated by modifying different aspects of metabolism. The majority of studies have targeted lipid biosynthesis, either to enhance TAG or EPA production (Table 2-4). In several cases endogenous or heterologous (from *S. cerevisiae* or *C. reinhardtii*) DGATs have been overproduced in different *Nannochloropsis* species (Beacham and Ali 2016; Iwai et al. 2015; Li et al. 2016a; Wei et al. 2017a; Xin et al. 2017; Zienkiewicz et al. 2017). Overexpression of the endogenous DGAT1a-encoding gene in *N. oceanica* resulted in a 39% increase in TAG content per cell and RNAi repression resulted in a 20% decrease in TAG content per cell following N deprivation (Wei et al. 2017a). Overexpression of the endogenous DGGT5-encoding cDNA in *N. oceanica* resulted in a 3.5 fold increase in TAG (as %TFA) (Zienkiewicz et al. 2017). Furthermore, the DGTT7-encoding cDNA has also been overexpressed in *N. oceanica* IMET1 resulting in 69% and 129% increase in neutral lipid (% dry weight) content under N-replete and N-deprivation conditions (Li et al. 2016a). The malonyl-CoA transacylase of *N. oceanica* IMET1, which loads the malonyl group onto the acyl-carrier protein for fatty acid synthesis, has been characterized (Tian et al. 2013) and its overproduction resulted in a 36% (% dry weight) increase in lipids without compromised growth (Chen et al. 2017). In *N. salina* RNAi conducted against the pyruvate dehydrogenase complex (PDC) kinase (PDCK), in order to increase acetyl-CoA levels for fatty acid production *(21)*, resulted in enhanced TAG content at the expense of protein (Ma et al. 2017). Increased expression of the desaturase genes in the EPA pathway using gene stacking techniques achieved up to a 25% increase in EPA (as %TFA) (Poliner et al. 2017). Engineering efforts are only beginning and the enhanced productivity of modified *Nannochloropsis* strains will require further identification of targets to affect metabolite partitioning into the desired pathways, and/or the reduction of final product turnover.

Altering metabolism in *Nannochloropsis* species by regulatory engineering

Manipulation of entire metabolic pathways on a greater scale could be accomplished by TF engineering. Recently, TF overproduction, or inactivation (CRISPR) and repression (RNAi) of a TF encoding gene have been used in various *Nannochloropsis* species with the goal of increasing biomass and/or lipid production (Table 3-4) (Ajjawi et al. 2017; Kang et al. 2015b; Kang et al. 2017; Kwon et al. 2017). Overproduction of the bHLH2 TF, in *N. salina* resulted in an increased

growth rate and a greater biomass productivity, although the transcriptional reprogramming was not described (Kang et al. 2015b). Overproduction of the bZIP1 TF, a predicted lipid biosynthetic TF (Hu et al. 2014), in *N. salina* resulted in enhanced growth, and under stress conditions enhanced lipid content (Kwon et al. 2017). In the bZIP1 overexpressor lines, the expression of putative target genes involved in lipid biosynthesis (Hu et al. 2014) was increased under normal conditions and more dramatically under stress conditions (Kwon et al. 2017). Incredibly, the introduction of the *Arabidopsis thaliana* WRINKLED1 TF (WRI1), a regulator of seed oil production, into *N. salina* enhanced lipid accumulation, possibly by upregulating lipid biosynthetic genes containing the WRI1 motifs in the promoter (Kang et al. 2017). CRISPR inactivation of the ZnCys TF, likely involved in N assimilation, resulted in a large increase in TAG content in *N. gaditana* but a strong reduction of growth. An optimization of the balance between growth versus lipid accumulation was achieved through decreased expression of ZnCys by RNAi or by the CRISPR/Cas9 mediated insertion of an antibiotic resistance cassette into the 3' UTR of this gene (Ajjawi et al. 2017).

Additional challenges for the development of improved *Nannochloropsis* strains

To develop improved strains optimized for biosynthetic yield, either by increasing flux into the target pathway or by disruption of competing pathways multiple specific modifications are likely to be required. For example, stacking gene modifications will be necessary to introduce new pathways or for the optimization of existing pathways. Furthermore, the development of marker-free strategies or the utilization of auxotrophic markers for selection of genetically modified strains are necessary for deployment of engineered strains into open ponds exposed to the environment.

Disrupting competing metabolic pathways may yield enhanced productivity of certain bioproducts. Carbohydrates are a competing sink for lipid production and polysaccharide biosynthetic genes have been targeted for repression in green algae and diatoms to increase lipid accumulation (Daboussi et al. 2014; Hildebrand et al. 2017; Work et al. 2010). The tough cell wall of the *Nannochloropsis* genus takes considerable cellular resources to construct and is an impediment to efficient processing of cells to obtain bioproducts. Therefore, *Nannochloropsis* strains with weakened walls may be superior for bioproduction. Reducing the turnover of desired products may enhance their accumulation, and identification of the genes involved in product degradation will be an avenue towards enhancing the productivity of algae (Trentacoste et al. 2013; Xue et al. 2013).

The successful adoption of gene disruption technology will facilitate disruption of biosynthetic genes for essential metabolites, generating strains that require supplementation or gene complementation (auxotrophy). Nitrate reductase (NR)-deficient strains of *Chlamydomonas* strains can be complemented with a wild-type NR gene and selected on nitrate media (Kindle et al. 1989). The NR gene has been targeted in several studies in various *Nannochloropsis* species (Ajjawi et al. 2017; Kilian et al. 2011; Wang et al. 2016) and diatoms (Diner et al. 2016b; McCarthy et al. 2017). Auxotrophic selection based on the NR-disrupted strains may be particularly useful when paired with episomal artificial chromosomes, enabling a nutrient selection pressure on episome maintenance.

Call for an open alga

Establishing a model organism (or bioproduct chassis) requires dissemination of the skills and tools developed to a wide network of scientists. A niche is developing for third-party repositories (TPR) that facilitate the maintenance of accumulated biological materials and are making these materials more accessible than ever before. However, for a TPR to be sustainable, innovators must be willing to transfer their materials and utilize the TPR as part of their own workflow. A notable nonprofit TPR for plasmid and strain dissemination is Addgene that we are collaborating with by depositing the collection of *Nannochloropsis* engineering vectors developed in our studies. Some of the model *Nannochoropsis* species such as CCMP1779 (*N. oceanica*), CCMP526 (*N. gaditana*), and CCMP537 (*N. salina*) are publicly available from algae culture collections (NCMA, https://ncma.bigelow.org/). We are making an engineered NR knockout strain publicly available through NCMA, as well. As transgenic tools and knockout and chassis strains are produced, deposition with TPRs will accelerate innovation, allow researchers to do more with less, build genome-wide overexpression and knockout strain libraries, and establish quality controls and standardization in the field.

Author contributions

EP, EF, and CB wrote the manuscript. All authors read and approved the manuscript.

Acknowledgements

This work was supported by a National Science Foundation grant (IOS-1354721) to EF. In addition, parts of this work were supported by the Division of Chemical Sciences, Geosciences and Biosciences, Office of Basic Energy Sciences of the United States Department of Energy (DE-FG02-91ER20021) and MSU-AgBioResearch to CB.

Table 1. Publicly available whole-genome datasets produced in *Nannochloropsis* species.

Citation	Strains	Datasets	Public access
(Radakovits et al. 2012)	*N. gaditana* CCMP526	Genome assembly	http://nannochloropsis.genomeprojectsolutions-databases.com/
(Corteggiani Carpinelli et al. 2013)	*N. gaditana* B-31	Genome browser and BLAST server	www.nannochloropsis.org
(Wang et al. 2014)	*N. oceanica* IMET1 and CCMP531, *N. granulata* CCMP529, *N. oculata* CCMP525, *N. salina* CCMP537, *N. gaditana* CCMP526	Genome browser and BLAST server	http://www.bioenergychina.org:8989/
(Vieler et al. 2012)	*N. oceanica* CCMP1779	Genome browser and BLAST server	https://genome.jgi.doe.gov/Nanoce1779/Nanoce1779.home.html
(Poliner et al. 2015)	*N. oceanica* CCMP1779	RNA-Seq during light:dark cycles	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE69460
(Mühlroth et al. 2017)	*N. oceanica* CCMP1779	RNA-Seq under phosphate deprivation	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE95774
(Hu et al. 2014)	*N. oceanica* IMET1 and CCMP531, *N. granulata* CCMP529, *N. oculata* CCMP525, *N. salina* CCMP537, *N. gaditana* CCMP526	Predicted regulatory connections of lipid biosynthetic genes	http://www.singlecellcenter.org/en/NannoRegulationDatabase/Download.htm

Table 2. Genomic transformation protocols developed for the *Nannochloropsis* genus. Reporter or epitope indicates if a protein tag was included, reporter abbreviations are FP - fluorescent protein, CP - chromoprotein, and GUS - β-glucuronidase. Promoters used to drive transcription of transgenes *in vivo* are listed; the species of origin for non-*Nannochloropsis* promoters are indicated in italics. Endogenous promoter abbreviations are LDSP - lipid droplet surface protein, β-tub - β-tubulin, UEP - ubiquitin extension protein, HSP - heat shock protein, and VCP - violaxanthin–chlorophyll a binding protein. Selection agent refers to antibiotic selection used to isolate *Nannochloropsis* transformants.

Citation	Strain(s)	Reporter or epitope	Promoter(s)	Selection agent(s)
(Vieler et al. 2012)	*N. oceanica* CCMP1779		LDSP	Hygromycin
(Radakovits et al. 2012)	*N. gaditana* CCMP526		β-tub, UEP, HSP	Zeocin
(Li et al. 2014a)	*N. oceanica* PP983 and MBIC10090, *N. granulata* MBIC10054, *N. salina* MBIC10063, *N. gaditana* CCAP849/5, *N.* oculata CCAP 849/1, *N. limnetica* KR1998/3	GUS	β-tub	Zeocin
(Cha et al. 2011)	*N. sp.* UMT-M3		*CMV* 35S	Hygromycin
(Ma et al. 2016)	*N. oculata* CS-179		*C. reinhardtii* HSP70A::RBCS2	Zeocin
(Moog et al. 2015)	*N. oceanica* CCMP1779	Green FP	VCP	Zeocin
(Shih et al. 2015)	*N. oculata* NIES-2146	shCP	*C. reinhardtii* HSP70A::RBCS2	
(Kang et al. 2015a)	*N. salina* CCMP1776	sfCherry	β-tub and UEP	Zeocin
(Chen et al. 2008)	*N. oculata*		*C. reinhardtii* HSP70A::RBCS2	
(Kilian et al. 2011)	*N. oceanica* W2J3B		VCP	Zeocin, Hygromycin, Blasticidin

Table 3. Overexpression studies conducted in the *Nannochloropsis* genus.
Methods describe the genetic techniques used in the study. Strains refers to the *Nannochloropsis* subspecies. The target gene refers to the function of the primary gene manipulated. Reporter or epitope indicates if a protein tag was included. Promoters used to drive transcription of transgenes *in vivo* are listed; the species of origin for non-*Nannochloropsis* promoters are indicated in italics. Endogenous promoter abbreviations are LDSP - lipid droplet surface protein, EF - elongation factor, Ribi - ribosomal subunit bidirectional, β-tub - β-tubulin, UEP - ubiquitin extension protein, HSP - heat shock protein, and VCP - violaxanthin–chlorophyll α binding protein. Selection agent refers to antibiotic selection used to isolate *Nannochloropsis* transformants.

Citation	Method(s)	Strain	Target gene(s)	Reporter(s) or epitope(s)	Promoter(s)	Selection agent(s)
(Poliner et al. 2017)	Multi-gene overexpression	*N. oceanica* CCMP1779	delta-9, delta-12, delta-5 fatty acid desaturases	Firefly luciferase, NanoLuciferase, Cyan FP (Cerulean variant)	EF, LDSP, Ribi	Hygromycin, Zeocin
(Zienkiewicz et al. 2017)	Overexpression	*N. oceanica* CCMP1779	diacylglycerol acyltransferase type 2-5	Yellow FP (Venus variant)	EF	Hygromycin
(Kaye et al. 2015)	Overexpression	*N. oceanica* CCMP1779	delta-12 fatty acid desaturase		LDSP	Hygromycin
(Li et al. 2016a)	Overexpression	*N. oceanica* CCMP1779	diacylglycerol acyltransferase type 2-7	Flag tag	Hsp20	Zeocin
(Kang et al. 2015b)	Overexpression	*N. salina* CCMP1776	basic helix loop helix 2	Flag tag	β-tub, UEP	Zeocin
(Beacham and Ali 2016)	Overexpression	*N. salina* CCAP 849/3	*S. cerevisiae* DGA1		*CMV* Tef, 35S	Hygromycin
(Chen et al. 2017)	Overexpression	*N. oceanica* IMET1	malonyl CoA-acyl carrier protein transacylase	Flag tag	HSP20	Zeocin
(Wei et al. 2017b)	Overexpression	*N. oceanica* IMET1	RuBisCO activase		β-tub, HSP70	Zeocin
(Iwai et al. 2015)	Overexpression, inducible expression	*N. oceanica* NIES-2145	*C. reinhardtii* diacylglycerol acyltransferase type 2-4		*C. reinhardtii* SQD, VCP	Zeocin
(Kang et al. 2017)	Overexpression	*N. salina* CCMP1776	*A. thaliana* WRINKLED1	Flag tag	β-tub, UEP	Zeocin
(Kwon et al. 2017)	Overexpression	*N. salina* CCMP1776	basic leucine zipper domain 1	Flag tag	β-tub, UEP	Zeocin

Table 4. Gene inactivation or repression studies conducted in the *Nannochloropsis* genus. Strain refers to the *Nannochloropsis* subspecies. Methods describe the genetic techniques used in the study. The target gene refers to the function of the primary gene manipulated. Reporter or epitope indicates if a protein tag was included, FP indicates fluorescent protein. Promoters used to drive transcription of transgenes *in vivo* are listed; the origin of transgenic promoters are indicated in italics. Endogenous promoter abbreviations are LDSP - lipid droplet surface protein, β-tub - β-tubulin, UEP - ubiquitin extension protein, HSP - heat shock protein, and VCP - violaxanthin–chlorophyll a binding protein. Selection agent refers to antibiotic selection used to isolate *Nannochloropsis* transformants.

Citation	Method(s)	Strain(s)	Target gene(s)	Reporter(s) or epitope(s)	Promoter(s)	Selection agent(s)
(Kilian et al. 2011)	Disruption by HR	*N. oceanica* W2J3B	nitrate reductase		VCP	Zeocin, Hygromycin, Blasticidin
(Wang et al. 2016)	Disruption by CRISPR/Cas9	*N. oceanica* IMET1	nitrate reductase		VCP, β-tub, V-ATPase	Hygromycin
(Ajjawi et al. 2017)	Disruption by CRISPR/Cas9, Repression by CRISPR/Cas9, RNAi	*N. gaditana* CCMP1894	ZnCys	Green FP, Flag tag	initiation factor 4AIII, 60S ribosomal protein L24, initiation factor 3, TCT	Hygromycin, Blasticidin
(Wei et al. 2017c)	Repression by RNAi	*N. oceanica* IMET1 and CCMP1779	carbonic anhydrase		β-tub	Zeocin
(Dolch et al. 2017)	Disruption by HR	*N. gaditana* CCMP526	palmitate elongase		UEP	Zeocin
(Perin et al. 2015)	Insertional mutagenesis	*N. gaditana* CCAP 849/5			UEP	Zeocin
(Wei et al. 2017a)	Overexpression and RNAi	*N. oceanica* IMET1	diacylglycerol acyltransferase type 1-1A	Green FP	VCP, UEP, β-tub	Zeocin
(Gee and Niyogi 2017)	Disruption by HR, complementation	*N. oceanica* CCMP1779	carbonic anhydrase	Flag tag, Yellow FP (Venus variant)	UEP	Zeocin
(Nobusawa et al. 2017)	Overexpression and disruption by HR	*N. oceanica* NIES-2145	lysophosphatidic acid acyltransferase 1-4	Green FP, Yellow FP (Venus variant)	LDSP	Zeocin, Hygromycin
(Ma et al. 2017)	Overexpression and repression by RNAI	*N. salina* CCMP537	pyruvate dehydrogenase kinase	Green FP	β-tub, VCP	Zeocin

Figure Legends

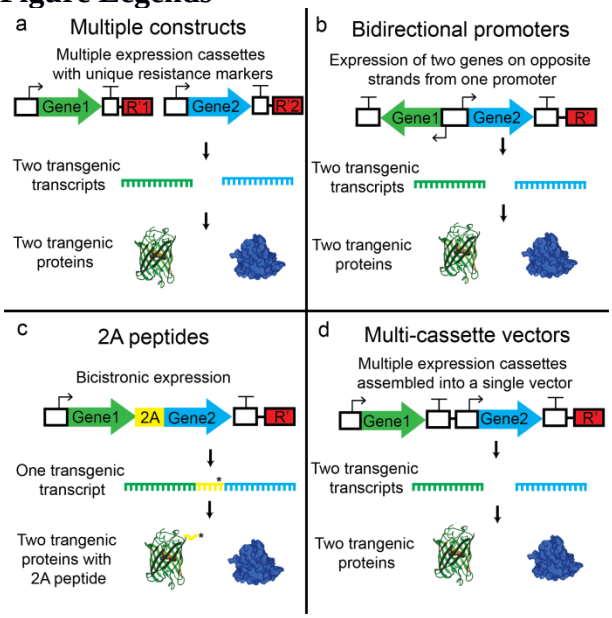

Fig. 1 Gene stacking strategies in *N. oceanica.*

Diagrams represent configuration for two protein coding transgenes. Gene1 represented by GFP is shown in green. Gene2 represented by NanoLuciferase is shown in blue. Promoters are indicated by arrows and terminators by a T (Synthetic Biology Open Language standard) (Galdzicki et al. 2014). Resistance markers are indicated as R' (red). *Nannochloropsis* expression cassettes without a plasmid backbone are shown. a. Constructs with unique selection markers (indicated as R'1 and R'2, red) can be introduced into one line. b. Bidirectional promoters regulate transcription on both DNA strands, and express two transcripts. c. Sequences encoding 2A peptides are placed between two protein-encoding genes. A peptide bond is not formed between the two final amino acids (*) during translation. Two discrete proteins are produced, the N' terminal protein (green) contains the majority of the 2A peptide (yellow) and the C terminal protein (blue) contains the last amino acid of the 2A peptide (not shown). d. Assembly of multiple expression cassettes into a single construct to produce multiple transcripts from different promoters. The image of the GFP protein was obtained from the NIH Image Gallery (https://www.flickr.com/photos/nihgov/) and is adapted under the terms of CC BY-NC 2.0. The NanoLuciferase image is adapted from Hall et al. with permission (Hall et al. 2012).

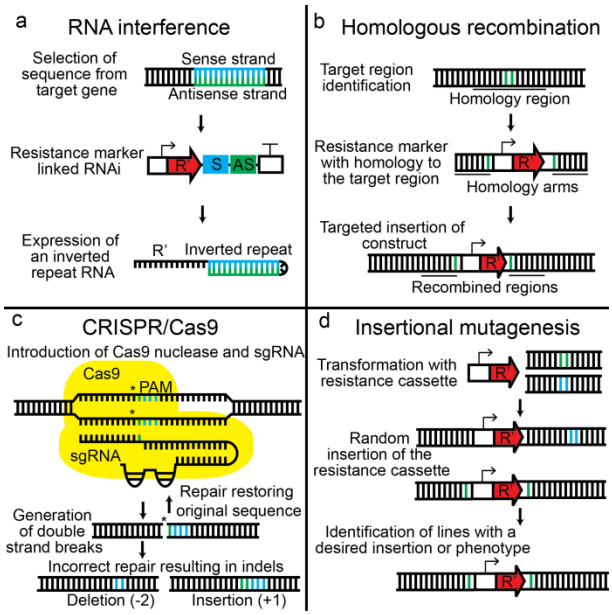

Fig. 2 Gene repression and inactivation techniques in the *Nannochloropsis* genus.

Promoters are indicated by arrows and terminators by a T (Synthetic Biology Open Language standard) (Galdzicki et al. 2014). Resistance markers are indicated as R' (red). Extended DNA is indicated by an overhanging backbone, and free ends of DNA are indicated by blunt ends. a. RNA interference by expression of sense (blue) and antisense (green) sequences of a target gene at the 3' end of a resistance gene. The sense and antisense sequences form an inverted repeat on the resistance marker transcript. b. Homology arms are the sequences flanking a desired insertion site (green). A resistance cassette with flanking homology arms to a target region is introduced by transformation. Homologous recombination between the homology arms of the resistance cassette and the target region results in a disrupted gene. c. CRISPR/Cas9 techniques utilize the Cas9 nuclease and an sgRNA, forming ribonucleoprotein complex (yellow). The PAM site (blue) and the final 3' nucleotide of the guide sequence (green) are indicated. Double stranded cuts are produced in the 3' end of the target region. Incorrect repair of the strand break prevents further Cas9 action. Insertion or deletion mutations are most likely to occur. d. Insertional mutagenesis by transformation with a resistance cassette. The resistance cassette is randomly inserted throughout the genome, resulting in mutant lines with unique insertion sites. Desired mutants (green) are identified by a phenotype screen, target gene screen, or from a mutant library.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Agapakis CM (2014) Designing Synthetic Biology. ACS Synthetic Biology 3:121-128

Ajjawi I, Verruto J, Aqui M, Soriaga LB, Coppersmith J, Kwok K, Peach L, Orchard E, Kalb R, Xu W, Carlson TJ, Francis K, Konigsfeld K, Bartalis J, Schultz A, Lambert W, Schwartz AS, Brown R, Moellering ER (2017) Lipid production in *Nannochloropsis gaditana* is doubled by decreasing expression of a single transcriptional regulator. Nature Biotechnology 35:647-652

Alboresi A, Le Quiniou C, Yadav SKN, Scholz M, Meneghesso A, Gerotto C, Simionato D, Hippler M, Boekema EJ, Croce R, Morosinotto T (2017) Conservation of core complex subunits shaped the structure and function of photosystem I in the secondary endosymbiont alga *Nannochloropsis gaditana*. New Phytologist 213:714-726

Alboresi A, Perin G, Vitulo N, Diretto G, Block MA, Jouhet J, Meneghesso A, Valle G, Giuliano G, Maréchal E, Morosinotto T (2016) Light Remodels Lipid Biosynthesis in *Nannochloropsis gaditana* by Modulating Carbon Partitioning Between Organelles. Plant Physiology:pp.00599.02016

Andersen RA, Brett RW, Potter D, Sexton JP (1998) Phylogeny of the Eustigmatophyceae Based upon 18S rDNA, with Emphasis on Nannochloropsis. Protist 149:61-74

Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. Molecular Systems Biology 2

Arnold AA, Genard B, Zito F, Tremblay R, Warschawski DE, Marcotte I (2015) Identification of lipid and saccharide constituents of whole microalgal cells by 13C solid-state NMR. Biochimica et Biophysica Acta (BBA) - Biomembranes 1848:369-377

Ashworth J, Coesel S, Lee A, Armbrust EV, Orellana MV, Baliga NS (2013) Genome-wide diel growth state transitions in the diatom *Thalassiosira pseudonana*. Proceedings of the National Academy of Sciences 110:7518-7523

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

Beacham TA, Ali ST (2016) Growth dependent silencing and resetting of DGA1 transgene in *Nannochloropsis salina*. Algal Research 14:65-71

Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier U-G (2009) Protein Targeting into Secondary Plastids. Journal of Eukaryotic Microbiology 56:9-15

Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kroger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jezequel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siaut M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van De Peer Y, Grigoriev IV (2008) The Phaeodactylum genome reveals the evolutionary history of diatom genomes. Nature 456:239-244

Braun R, Farre EM, Schurr U, Matsubara S (2014) Effects of light and circadian clock on growth and chlorophyll accumulation of *Nannochloropsis gaditana*. Journal of phycology 50:515-525

Buitrago-Flórez FJ, Restrepo S, Riaño-Pachón DM (2014) Identification of Transcription Factor Genes and Their Correlation with the High Diversity of Stramenopiles. PLoS ONE 9:e111841

Cao S, Zhang X, Xu D, Fan X, Mou S, Wang Y, Ye N, Wang W (2013) A transthylakoid proton gradient and inhibitors induce a non-photochemical fluorescence quenching in unicellular algae *Nannochloropsis* sp. FEBS Letters 587:1310-1315

Cermak T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, Belanto JJ, Starker CG, Mathre JW, Greenstein RL, Voytas DF (2017) A multi-purpose toolkit to enable advanced genome engineering in plants. Plant Cell:tpc.00922.02016

Cha T-S, Chen C-F, Yee W, Aziz A, Loh S-H (2011) Cinnamic acid, coumarin and vanillin: Alternative phenolic compounds for efficient Agrobacterium-mediated transformation of the unicellular green alga, *Nannochloropsis* sp. Journal of Microbiological Methods 84:430-434

Chauton MS, Winge P, Brembu T, Vadstein O, Bones AM (2013) Gene Regulation of Carbon Fixation, Storage, and Utilization in the Diatom *Phaeodactylum tricornutum* Acclimated to Light/Dark Cycles. Plant Physiology 161:1034-1048

Cheah YE, Albers SC, Peebles CAM (2013) A novel counter-selection method for markerless genetic modification in *Synechocystis* sp. PCC 6803. Biotechnology Progress 29:23-30

Chen HL, Li SS, Huang R, Tsai HJ (2008) Conditional Production of a Functional Fish Growth Hormone in the Transgenic Line of *Nannochloropsis Oculata* (Eustigmatophyceae). Journal of phycology 44:768-776

Chen J-W, Liu W-J, Hu D-X, Wang X, Balamurugan S, Alimujiang A, Yang W-D, Liu J-S, Li H-Y (2017) Identification of a malonyl CoA-acyl carrier protein transacylase and its regulatory role in fatty acid biosynthesis in oleaginous microalga *Nannochloropsis oceanica*: *Nannochloropsis* MCAT. Biotechnology and Applied Biochemistry

Chew KW, Yap JY, Show PL, Suan NH, Juan JC, Ling TC, Lee D-J, Chang J-S (2017) Microalgae biorefinery: High value products perspectives. Bioresource Technology 229:53-62

Chukhutsina VU, Fristedt R, Morosinotto T, Croce R (2017) Photoprotection strategies of the alga *Nannochloropsis gaditana*. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1858:544-552

Coesel S, Mangogna M, Ishikawa T, Heijde M, Rogato A, Finazzi G, Todo T, Bowler C, Falciatore A (2009) Diatom PtCPF1 is a new cryptochrome/photolyase family member with DNA repair and transcription regulation activity. EMBO reports 10:655-661

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339:819-823

Corteggiani Carpinelli E, Telatin A, Vitulo N, Forcato C, D'Angelo M, Schiavon R, Vezzi A, Giacometti GM, Morosinotto T, Valle G (2013) Chromosome Scale Genome Assembly and Transcriptome Profiling of *Nannochloropsis gaditana* in Nitrogen Depletion. Molecular Plant

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

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-e96

Dehoff P, Akella S, Soriaga L (2014) Autonomous replication sequences and episomal DNA molecules. US9447422 B2

Dehoff P, Soriaga L (2014) *Nannochloropsis* kozak consensus sequence. US9309523 B2

Diner RE, Bielinski VA, Dupont CL, Allen AE, Weyman PD (2016a) Refinement of the Diatom Episome Maintenance Sequence and Improvement of Conjugation-Based DNA Delivery Methods. Frontiers in bioengineering and biotechnology 4:65

Diner RE, Noddings CM, Lian NC, Kang AK, McQuaid JB, Jablanovic J, Espinoza JL, Nguyen NA, Anzelmatti MA, Jansson J, Bielinski VA, Karas BJ, Dupont CL, Allen AE, Weyman PD (2017) Diatom centromeres suggest a mechanism for nuclear DNA acquisition. Proceedings of the National Academy of Sciences 114:E6015-E6024

Diner RE, Schwenck SM, McCrow JP, Zheng H, Allen AE (2016b) Genetic Manipulation of Competition for Nitrate between Heterotrophic Bacteria and Diatoms. Frontiers in Microbiology 7

Dolch L-J, Rak C, Perin G, Tourcier G, Broughton R, Leterrier M, Morosinotto T, Tellier F, Faure J-D, Falconet D, Jouhet J, Sayanova O, Beaudoin F, Maréchal E (2017) A Palmitic Acid Elongase Affects Eicosapentaenoic Acid and Plastidial Monogalactosyldiacylglycerol Levels in *Nannochloropsis*. Plant Physiology 173:742-759

Fábregas J, Maseda A, Domínguez A, Ferreira M, Otero A (2002) Changes in the cell composition of the marine microalga, *Nannochloropsis gaditana*, during a light:dark cycle. Biotechnology letters 24:1699-1703

Fawley MW, Jameson I, Fawley KP (2015) The phylogeny of the genusNannochloropsis(Monodopsidaceae, Eustigmatophyceae), with descriptions ofN. australis sp. nov. andMicrochloropsis gen. nov. Phycologia 54:545-552

Fortunato AE, Jaubert M, Enomoto G, Bouly J-P, Raniello R, Thaler M, Malviya S, Bernardes JS, Rappaport F, Gentili B, Huysman MJJ, Carbone A, Bowler C, d'Alcalà MR, Ikeuchi M, Falciatore A (2016) Diatom Phytochromes Reveal the Existence of Far-Red-Light-Based Sensing in the Ocean. Plant Cell 28:616-628

Galdzicki M, Clancy KP, Oberortner E, Pocock M, Quinn JY, Rodriguez CA, Roehner N, Wilson ML, Adam L, Anderson JC, Bartley BA, Beal J, Chandran D, Chen J, Densmore D, Endy D, Grunberg R, Hallinan J, Hillson NJ, Johnson JD, Kuchinsky A, Lux M, Misirli G, Peccoud J, Plahar HA, Sirin E, Stan GB, Villalobos A, Wipat A, Gennari JH, Myers CJ, Sauro HM (2014) The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. Nat Biotechnol 32:545-550

Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TMA (1987) The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucleic Acids Research 15:3257-3273

Gee CW, Niyogi KK (2017) The carbonic anhydrase CAH1 is an essential component of the carbon-concentrating mechanism in *Nannochloropsis oceanica*. Proceedings of the National Academy of Sciences 114:4537-4542
Gschloessl B, Guermeur Y, Cock JM (2008) HECTAR: a method to predict subcellular targeting in heterokonts. BMC Bioinformatics 9:393
Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG, Otto P, Zimmerman K, Vidugiris G, Machleidt T, Robers MB, Benink HA, Eggers CT, Slater MR, Meisenheimer PL, Klaubert DH, Fan F, Encell LP, Wood KV (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS chemical biology 7:1848-1857
Hamilton ML, Haslam RP, Napier JA, 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
Hibberd DJ (1981) Notes on the taxonomy and nomenclature of the algal classes Eustigmatophyceae and Tribophyceae (synonym Xanthophyceae). Botanical Journal of the Linnean Society 82:93-119
Hildebrand M, Manandhar-Shrestha K, Abbriano R (2017) Effects of chrysolaminarin synthase knockdown in the diatom *Thalassiosira pseudonana*: Implications of reduced carbohydrate storage relative to green algae. Algal Research 23:66-77
Hu J, Wang D, Li J, Jing G, Ning K, Xu J (2014) Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae *Nannochloropsis*. Scientific Reports 4
Huang J, Liu J, Li Y, Chen F (2008) Isolation and Characterization of the Phytoene Desaturase Gene as a Potential Selective Marker for Genetic Engineering of the Astaxanthin-Producing Green Alga *Chlorella Zofingiensis* (Chlorophyta). Journal of phycology 44:684-690
Huysman MJ, Fortunato AE, Matthijs M, Costa BS, Vanderhaeghen R, Van den Daele H, Sachse M, Inze D, Bowler C, Kroth PG, Wilhelm C, Falciatore A, Vyverman W, De Veylder L (2013) AUREOCHROME1a-mediated induction of the diatom-specific cyclin dsCYC2 controls the onset of cell division in diatoms (*Phaeodactylum tricornutum*). Plant Cell 25:215-228
Huysman MJ, Martens C, Vandepoele K, Gillard J, Rayko E, Heijde M, Bowler C, Inze D, Van de Peer Y, De Veylder L, Vyverman W (2010) Genome-wide analysis of the diatom cell cycle unveils a novel type of cyclins involved in environmental signaling. Genome biology 11:R17
Iwai M, Hori K, Sasaki-Sekimoto Y, Shimojima M, Ohta H (2015) Manipulation of oil synthesis in *Nannochloropsis* strain NIES-2145 with a phosphorus starvation–inducible promoter from *Chlamydomonas reinhardtii*. Frontiers in Microbiology 6
Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proceedings of the National Academy of Sciences 107:10949-10954
Jeong SW, Nam SW, HwangBo K, Jeong WJ, Jeong B-r, Chang YK, Park Y-I (2017) Transcriptional Regulation of Cellulose Biosynthesis during the Early Phase of Nitrogen Deprivation in *Nannochloropsis salina*. Scientific Reports 7

Jia B, Zheng Y, Xiao K, Wu M, Lei Y, Huang Y, Hu Z (2016) A vector for multiple gene co-expression in *Chlamydomonas reinhardtii*. Algal Research 20:53-56
Jia J, Han D, Gerken HG, Li Y, Sommerfeld M, Hu Q, Xu J (2015) Molecular mechanisms for photosynthetic carbon partitioning into storage neutral lipids in *Nannochloropsis oceanica* under nitrogen-depletion conditions. Algal Research 7:66-77
Jinkerson RE, Radakovits R, Posewitz MC (2013) Genomic insights from the oleaginous model alga *Nannochloropsis gaditana*. Bioengineered 4:37-43
Kang NK, Choi G-G, Kim EK, Shin S-E, Jeon S, Park MS, Jeong KJ, Jeong B-r, Chang YK, Yang J-W, Lee B (2015a) Heterologous overexpression of sfCherry fluorescent protein in *Nannochloropsis salina*. Biotechnology Reports 8:10-15
Kang NK, Jeon S, Kwon S, Koh HG, Shin S-E, Lee B, Choi G-G, Yang J-W, Jeong B-r, Chang YK (2015b) Effects of overexpression of a bHLH transcription factor on biomass and lipid production in *Nannochloropsis salina*. Biotechnology for Biofuels 8
Kang NK, Kim EK, Kim YU, Lee B, Jeong W-J, Jeong B-r, Chang YK (2017) Increased lipid production by heterologous expression of AtWRI1 transcription factor in *Nannochloropsis salina*. Biotechnology for Biofuels 10
Karas BJ, Diner RE, Lefebvre SC, McQuaid J, Phillips AP, Noddings CM, Brunson JK, Valas RE, Deerinck TJ, Jablanovic J, Gillard JT, Beeri K, Ellisman MH, Glass JI, Hutchison CA, 3rd, Smith HO, Venter JC, Allen AE, Dupont CL, Weyman PD (2015) Designer diatom episomes delivered by bacterial conjugation. Nat Commun 6:6925
Kaye Y, Grundman O, Leu S, Zarka A, Zorin B, Didi-Cohen S, Khozin-Goldberg I, Boussiba S (2015) Metabolic engineering toward enhanced LC-PUFA biosynthesis in *Nannochloropsis oceanica*: Overexpression of endogenous Δ12 desaturase driven by stress-inducible promoter leads to enhanced deposition of polyunsaturated fatty acids in TAG. Algal Research 11:387-398
Keeling PJ (2009) Chromalveolates and the Evolution of Plastids by Secondary Endosymbiosis. Journal of Eukaryotic Microbiology 56:1-8
Kilian O, Benemann CSE, Niyogi KK, Vick B (2011) High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp. Proceedings of the National Academy of Sciences of the United States of America 108:21265-21269
Kindle KL, Schnell RA, Fernández E, Lefebvre PA (1989) Stable nuclear transformation of Chlamydomonas using the Chlamydomonas gene for nitrate reductase. J Cell Biol 109:2589-2601
Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, Joung JK (2016) High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529:490-495
Kroth P, Strotmann H (1999) Diatom plastids: Secondary endocytobiosis, plastid genome and protein import. Physiologia Plantarum 107:136-141
Kwon S, Kang NK, Koh HG, Shin S-E, Lee B, Jeong B-r, Chang YK (2017) Enhancement of biomass and lipid productivity by overexpression of a bZIP transcription factor in *Nannochloropsis salina*: Engineering of *Nannochloropsis* with bZIP TF. Biotechnology and Bioengineering
Li D-W, Cen S-Y, Liu Y-H, Balamurugan S, Zheng X-Y, Alimujiang A, Yang W-D, Liu J-S, Li H-Y (2016a) A type 2 diacylglycerol acyltransferase accelerates the triacylglycerol

biosynthesis in heterokont oleaginous microalga *Nannochloropsis oceanica*. Journal of Biotechnology 229:65-71
Li F, Gao D, Hu H (2014a) High-efficiency nuclear transformation of the oleaginous marine *Nannochloropsis* species using PCR product. Bioscience, Biotechnology, and Biochemistry 78:812-817
Li J, Han D, Wang D, Ning K, Jia J, Wei L, Jing X, Huang S, Chen J, Li Y, Hu Q, Xu J (2014b) Choreography of Transcriptomes and Lipidomes of *Nannochloropsis* Reveals the Mechanisms of Oil Synthesis in Microalgae. Plant Cell 26:1645-1665
Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre PA, Fitz-Gibbon ST, Grossman AR, Jonikas MC (2016b) An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in *Chlamydomonas reinhardtii*. Plant Cell 28:367-387
Lin G, Wang Y, Guo L, Ding H, Hu Y, Liang S, Zhang Z, Yang G (2017) Verification of mutagen function of Zeocin in *Nannochloropsis oceanica* through transcriptome analysis. Journal of Ocean University of China 16:501-508
Loira N, Mendoza S, Paz Cortés M, Rojas N, Travisany D, Genova AD, Gajardo N, Ehrenfeld N, Maass A (2017) Reconstruction of the microalga *Nannochloropsis salina* genome-scale metabolic model with applications to lipid production. BMC Systems Biology 11
Lubián L (1982) Nannochloropsis gaditana sp. nov., una nueva Eustigmatophyceae marina. Lazaroa 4
Ma X, Pan K, Zhang L, Zhu B, Yang G, Zhang X (2016) Genetic transformation of *Nannochloropsis oculata* with a bacterial phleomycin resistance gene as dominant selective marker. Journal of Ocean University of China 15:351-356
Ma X, Yao L, Yang B, Lee YK, Chen F, Liu J (2017) RNAi-mediated silencing of a pyruvate dehydrogenase kinase enhances triacylglycerol biosynthesis in the oleaginous marine alga *Nannochloropsis salina*. Scientific Reports 7
Mann M, Serif M, Jakob T, Kroth PG, Wilhelm C (2017) PtAUREO1a and PtAUREO1b knockout mutants of the diatom *Phaeodactylum tricornutum* are blocked in photoacclimation to blue light. Journal of Plant Physiology 217:44-48
McCarthy JK, Smith SR, McCrow JP, Tan M, Zheng H, Beeri K, Roth R, Lichtle C, Goodenough U, Bowler CP, Dupont CL, Allen AE (2017) Nitrate Reductase Knockout Uncouples Nitrate Transport from Nitrate Assimilation and Drives Repartitioning of Carbon Flux in a Model Pennate Diatom. Plant Cell 29:2047-2070
Meng Y, Jiang J, Wang H, Cao X, Xue S, Yang Q, Wang W (2015) The characteristics of TAG and EPA accumulation in *Nannochloropsis oceanica* IMET1 under different nitrogen supply regimes. Bioresource Technology 179:483-489
Miller R, Wu G, Deshpande RR, Vieler A, Gartner K, Li X, Moellering ER, Zauner S, Cornish AJ, Liu B, Bullard B, Sears BB, Kuo MH, Hegg EL, Shachar-Hill Y, Shiu SH, Benning C (2010) Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiology 154:1737-1752
Moellering ER, Benning C (2010) RNA Interference Silencing of a Major Lipid Droplet Protein Affects Lipid Droplet Size in *Chlamydomonas reinhardtii*. Eukaryotic Cell 9:97-106
Moog D, Stork S, Reislöhner S, Grosche C, Maier U-G (2015) In vivo Localization Studies in the Stramenopile Alga *Nannochloropsis oceanica*. Protist 166:161-171

Mühlroth A, Li K, Røkke G, Winge P, Olsen Y, Hohmann-Marriott M, Vadstein O, Bones A (2013) Pathways of Lipid Metabolism in Marine Algae, Co-Expression Network, Bottlenecks and Candidate Genes for Enhanced Production of EPA and DHA in Species of Chromista. Marine Drugs 11:4662-4697

Mühlroth A, Winge P, Assimi AE, Jouhet J, Marechal E, Hohmann-Marriott MF, Vadstein O, Bones AM (2017) Mechanisms of phosphorus acquisition and lipid class remodelling under P limitation in a marine microalga. Plant Physiology

Müller N, Wenzel S, Zou Y, Künzel S, Sasso S, Weiß D, Prager K, Grossman A, Kottke T, Mittag M (2017) A Plant Cryptochrome Controls Key Features of the *Chlamydomonas* Circadian Clock and Its Life Cycle. Plant Physiology 174:185-201

Murakami R, Hashimoto H (2009) Unusual Nuclear Division in *Nannochloropsis oculata* (Eustigmatophyceae, Heterokonta) which May Ensure Faithful Transmission of Secondary Plastids. Protist 160:41-49

Nobusawa T, Hori K, Mori H, Kurokawa K, Ohta H (2017) Differently localized lysophosphatidic acid acyltransferases crucial for triacylglycerol biosynthesis in the oleaginous alga *Nannochloropsis*. The Plant Journal 90:547-559

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

Pal D, Khozin-Goldberg I, Didi-Cohen S, Solovchenko A, Batushansky A, Kaye Y, Sikron N, Samani T, Fait A, Boussiba S (2013) Growth, lipid production and metabolic adjustments in the euryhaline eustigmatophyte *Nannochloropsis oceanica* CCALA 804 in response to osmotic downshift. Applied Microbiology and Biotechnology 97:8291-8306

Perin G, Bellan A, Segalla A, Meneghesso A, Alboresi A, Morosinotto T (2015) Generation of random mutants to improve light-use efficiency of *Nannochloropsis gaditana* cultures for biofuel production. Biotechnol Biofuels 8:161

Plucinak TM, Horken KM, Jiang W, Fostvedt J, Nguyen ST, Weeks DP (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

Poliner E, Panchy N, Newton L, Wu G, Lapinsky A, Bullard B, Zienkiewicz A, Benning C, Shiu S-H, Farré EM (2015) Transcriptional coordination of physiological responses in *Nannochloropsis oceanica* CCMP1779 under light/dark cycles. Plant J 83:1097-1113

Poliner E, Pulman JA, Zienkiewicz K, Childs K, Benning C, Farre EM (2017) A toolkit for *Nannochloropsis oceanica* CCMP1779 enables gene stacking and genetic engineering of the eicosapentaenoic acid pathway for enhanced long-chain polyunsaturated fatty acid production. Plant biotechnology journal

Radakovits R, Jinkerson RE, Fuerstenberg SI, Tae H, Settlage RE, Boore JL, Posewitz MC (2012) Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. Nat Commun 3:686

Ran FA, Hsu Patrick D, Lin C-Y, Gootenberg Jonathan S, Konermann S, Trevino AE, Scott David A, Inoue A, Matoba S, Zhang Y, Zhang F (2013) Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell 154:1380-1389

Rayko E, Maumus F, Maheswari U, Jabbari K, Bowler C (2010) Transcription factor families inferred from genome sequences of photosynthetic stramenopiles. New Phytologist 188:52-66

Rodolfi L, Chini Zittelli G, Bassi N, Padovani G, Biondi N, Bonini G, Tredici MR (2009) Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. Biotechnology and Bioengineering 102:100-112

Rohr J, Sarkar N, Balenger S, Jeong B-r, Cerutti H (2004) Tandem inverted repeat system for selection of effective transgenic RNAi strains in *Chlamydomonas*: Tandem inverted repeat system for efficient RNAi. The Plant Journal 40:611-621

Schellenberger Costa B, Sachse M, Jungandreas A, Bartulos CR, Gruber A, Jakob T, Kroth PG, Wilhelm C (2013) Aureochrome 1a Is Involved in the Photoacclimation of the Diatom *Phaeodactylum tricornutum*. PLoS ONE 8:e74451

Schneider JC, Livne A, Sukenik A, Roessler PG (1995) A mutant of *Nannochloropsis* deficient in eicosapentaenoic acid production. Phytochemistry 40:807-814

Schneider JC, Roessler P (1994) Radiolabeling Studies Of Lipids And Fatty Acids In *Nannochloropsis* (Eustigmatophyceae), An Oleaginous Marine Alga1. Journal of phycology 30:594-598

Scholz MJ, Weiss TL, Jinkerson RE, Jing J, Roth R, Goodenough U, Posewitz MC, Gerken HG (2014) Ultrastructure and Composition of the *Nannochloropsis gaditana* Cell Wall. Eukaryotic Cell 13:1450-1464

Schonknecht G, Chen WH, Ternes CM, Barbier GG, Shrestha RP, Stanke M, Brautigam A, Baker BJ, Banfield JF, Garavito RM, Carr K, Wilkerson C, Rensing SA, Gagneul D, Dickenson NE, Oesterhelt C, Lercher MJ, Weber APM (2013) Gene Transfer from Bacteria and Archaea Facilitated Evolution of an Extremophilic Eukaryote. Science 339:1207-1210

Sharma P, Yan F, Doronina VA, Escuin-Ordinas H, Ryan MD, Brown JD (2012) 2A peptides provide distinct solutions to driving stop-carry on translational recoding. Nucleic Acids Research 40:3143-3151

Sheehan J, Dunahay T, Benemann J, Roessler P (1998) Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae; Close-Out Report.

Shih C-H, Chen H-Y, Lee H-C, Tsai H-J (2015) Purple Chromoprotein Gene Serves as a New Selection Marker for Transgenesis of the Microalga *Nannochloropsis oculata*. PLOS ONE 10:e0120780

Simionato D, Block MA, La Rocca N, Jouhet J, Marechal E, Finazzi G, Morosinotto T (2013) The response of Nannochloropsis gaditana to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. Eukaryotic Cell 12:665-676

Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. Science 351:84-88

Starkenburg SR, Kwon KJ, Jha RK, McKay C, Jacobs M, Chertkov O, Twary S, Rocap G, Cattolico R (2014) A pangenomic analysis of the *Nannochloropsis* organellar genomes reveals novel genetic variations in key metabolic genes. BMC Genomics 15:212

Stormo GD (2000) DNA binding sites: representation and discovery. Bioinformatics 16:16-23

Suda S, Atsumi M, Miyashita H (2002) Taxonomic characterization of a marine Nannochloropsis species, N. oceanica sp. nov. (Eustigmatophyceae). Phycologia 41:273-279

Sukenik A, Carmeli Y (1990) Lipid Synthesis and Fatty Acid Composition in *Nannochloropsis* Sp. (Eustigmatophyceae) Grown in a Light-Dark Cycle. Journal of phycology 26:463-469
Sukenik A, Carmeli Y, Berner T (1989) Regulation of Fatty Acid Composition by Irradiance Level in the Eustigmatophyte *Nannochloropsis* Sp. Journal of phycology 25:686-692
Suzuki K, Kimura T, Shinoda H, Bai G, Daniels MJ, Arai Y, Nakano M, Nagai T (2016) Five colour variants of bright luminescent protein for real-time multicolour bioimaging. Nat Commun 7:13718
Thiriet-Rupert S, Carrier G, Chénais B, Trottier C, Bougaran G, Cadoret J-P, Schoefs B, Saint-Jean B (2016) Transcription factors in microalgae: genome-wide prediction and comparative analysis. BMC Genomics 17
Tian J, Zheng M, Yang G, Zheng L, Chen J, Yang B (2013) Cloning and stress-responding expression analysis of malonyl CoA-acyl carrier protein transacylase gene of *Nannochloropsis gaditana*. Gene 530:33-38
Trentacoste EM, Shrestha RP, Smith SR, Gle C, Hartmann AC, Hildebrand M, Gerwick WH (2013) Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth. Proceedings of the National Academy of Sciences 110:19748-19753
Tsai CH, Warakanont J, Takeuchi T, Sears BB, Moellering ER, Benning C (2014) The protein Compromised Hydrolysis of Triacylglycerols 7 (CHT7) acts as a repressor of cellular quiescence in *Chlamydomonas*. Proc Natl Acad Sci USA 111:15833-15838
Umetani I, Kunugi M, Yokono M, Takabayashi A, Tanaka A (2017) Evidence of the supercomplex organization of photosystem II and light-harvesting complexes in *Nannochloropsis granulata*. Photosynthesis Research
Unkefer CJ, Sayre RT, Magnuson JK, Anderson DB, Baxter I, Blaby IK, Brown JK, Carleton M, Cattolico RA, Dale T, Devarenne TP, Downes CM, Dutcher SK, Fox DT, Goodenough U, Jaworski J, Holladay JE, Kramer DM, Koppisch AT, Lipton MS, Marrone BL, McCormick M, Molnár I, Mott JB, Ogden KL, Panisko EA, Pellegrini M, Polle J, Richardson JW, Sabarsky M, Starkenburg SR, Stormo GD, Teshima M, Twary SN, Unkefer PJ, Yuan JS, Olivares JA (2017) Review of the algal biology program within the National Alliance for Advanced Biofuels and Bioproducts. Algal Research 22:187-215
Vieler A, Wu G, Tsai C-H, Bullard B, Cornish AJ, Harvey C, Reca I-B, Thornburg C, Achawanantakun R, Buehl CJ, Campbell MS, Cavalier D, Childs KL, Clark TJ, Deshpande R, Erickson E, Armenia Ferguson A, Handee W, Kong Q, Li X, Liu B, Lundback S, Peng C, Roston RL, Sanjaya, Simpson JP, TerBush A, Warakanont J, Zäuner S, Farre EM, Hegg EL, Jiang N, Kuo M-H, Lu Y, Niyogi KK, Ohlrogge J, Osteryoung KW, Shachar-Hill Y, Sears BB, Sun Y, Takahashi H, Yandell M, Shiu S-H, Benning C (2012) Genome, Functional Gene Annotation, and Nuclear Transformation of the Heterokont Oleaginous Alga *Nannochloropsis oceanica* CCMP1779. PLoS Genet 8:e1003064
Wang D, Ning K, Li J, Hu J, Han D, Wang H, Zeng X, Jing X, Zhou Q, Su X, Chang X, Wang A, Wang W, Jia J, Wei L, Xin Y, Qiao Y, Huang R, Chen J, Han B, Yoon K, Hill RT, Zohar Y, Chen F, Hu Q, Xu J (2014) *Nannochloropsis* Genomes Reveal Evolution of Microalgal Oleaginous Traits. PLoS Genetics 10:e1004094
Wang Q, Lu Y, Xin Y, Wei L, Huang S, Xu J (2016) Genome editing of model oleaginous microalgae *Nannochloropsis* spp. by CRISPR/Cas9. The Plant Journal 88:1071-1081

Wei H, Shi Y, Ma X, Pan Y, Hu H, Li Y, Luo M, Gerken H, Liu J (2017a) A type-I diacylglycerol acyltransferase modulates triacylglycerol biosynthesis and fatty acid composition in the oleaginous microalga, *Nannochloropsis oceanica*. Biotechnology for Biofuels 10

Wei L, Wang Q, Xin Y, Lu Y, Xu J (2017b) Enhancing photosynthetic biomass productivity of industrial oleaginous microalgae by overexpression of RuBisCO activase. Algal Research 27:366-375

Wei L, Xin Y, Wang D, Jing X, Zhou Q, Su X, Jia J, Ning K, Chen F, Hu Q, Xu J (2013) *Nannochloropsis* plastid and mitochondrial phylogenomes reveal organelle diversification mechanism and intragenus phylotyping strategy in microalgae. BMC Genomics 14:534

Wei L, Xin Y, Wang Q, Yang J, Hu H, Xu J (2017c) RNAi-based targeted gene knockdown in the model oleaginous microalgae *Nannochloropsis oceanica*. The Plant Journal 89:1236-1250

Wingender E (2008) The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. Brief Bioinformatics 9:326-332

Work VH, Radakovits R, Jinkerson RE, Meuser JE, Elliott LG, Vinyard DJ, Laurens LML, Dismukes GC, Posewitz MC (2010) Increased Lipid Accumulation in the *Chlamydomonas reinhardtii* sta7-10 Starchless Isoamylase Mutant and Increased Carbohydrate Synthesis in Complemented Strains. Eukaryotic Cell 9:1251-1261

Xiao Y, Zhang J, Cui J, Yao X, Sun Z, Feng Y, Cui Q (2015) Simultaneous accumulation of neutral lipids and biomass in *Nannochloropsis oceanica* IMET1 under high light intensity and nitrogen replete conditions. Algal Research 11:55-62

Xie Y, Wang D, Lan F, Wei G, Ni T, Chai R, Liu D, Hu S, Li M, Li D, Wang H, Wang Y (2017) An episomal vector-based CRISPR/Cas9 system for highly efficient gene knockout in human pluripotent stem cells. Scientific Reports 7

Xin Y, Lu Y, Lee Y-Y, Wei L, Jia J, Wang Q, Wang D, Bai F, Hu H, Hu Q, Liu J, Li Y, Xu J (2017) Producing Designer Oils in Industrial Microalgae by Rational Modulation of Co-evolving Type-2 Diacylglycerol Acyltransferases. Molecular Plant

Xue J, Niu Y-F, Huang T, Yang W-D, Liu J-S, Li H-Y (2015) Genetic improvement of the microalga *Phaeodactylum tricornutum* for boosting neutral lipid accumulation. Metabolic Engineering 27:1-9

Xue Z, Sharpe PL, Hong S-P, Yadav NS, Xie D, Short DR, Damude HG, Rupert RA, Seip JE, Wang J, Pollak DW, Bostick MW, Bosak MD, Macool DJ, Hollerbach DH, Zhang H, Arcilla DM, Bledsoe SA, Croker K, McCord EF, Tyreus BD, Jackson EN, Zhu Q (2013) Production of omega-3 eicosapentaenoic acid by metabolic engineering of *Yarrowia lipolytica*. Nature Biotechnology 31:734-740

Yaakob Z, Ali E, Zainal A, Mohamad M, Takriff M (2014) An overview: biomolecules from microalgae for animal feed and aquaculture. Journal of Biological Research-Thessaloniki 21:6

Zaslavskaia LA, Lippmeier JC, Kroth PG, Grossman AR, Apt KE (2001) Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. Journal of phycology 36:379-386

Zienkiewicz K, Zienkiewicz A, Poliner E, Du Z-Y, Vollheyde K, Herrfurth C, Marmon S, Farré EM, Feussner I, Benning C (2017) *Nannochloropsis*, a rich source of diacylglycerol

acyltransferases for engineering of triacylglycerol content in different hosts. Biotechnology for Biofuels 10
Zou N, Zhang C, Cohen Z, Richmond A (2000) Production of cell mass and eicosapentaenoic acid (EPA) in ultrahigh cell density cultures of *Nannochloropsis* sp. (Eustigmatophyceae). European Journal of Phycology 35:127-133
Zou Y, Wenzel S, Müller N, Prager K, Jung E-M, Kothe E, Kottke T, Mittag M (2017) An Animal-Like Cryptochrome Controls the *Chlamydomonas* Sexual Cycle. Plant Physiology 174:1334-1347