the plant journal

SEB Society for Experimental Biology

The Plant Journal (2016) **85,** 161–176 doi: 10.1111/tpj.13081

RESOURCE

Flux balance analysis of primary metabolism in the diatom *Phaeodactylum tricornutum*

Joomi Kim[1,*], Michele Fabris[2,3,4,5], Gino Baart[3,4,5,6], Min K. Kim[7], Alain Goossens[3,4], Wim Vyverman[5], Paul G. Falkowski[1] and Desmond S. Lun[7,8,9]

[1]*Environmental Biophysics and Molecular Ecology Program, Department of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901, USA,*

[2]*Plant Functional Biology and Climate Change Cluster (C3), Faculty of Science University of Technology, Sydney, New South Wales, Australia,*

[3]*Department of Plant Systems Biology, Vlaams Instituut voor Biotechnologie, B–9052, Gent, Belgium,*

[4]*Department of Plant Biotechnology and Bioinformatics, Ghent University, B–9052, Gent, Belgium,*

[5]*Department of Biology, Laboratory of Protistology and Aquatic Ecology, Ghent University, B–9000, Gent, Belgium,*

[6]*Centre of Microbial and Plant Genetics Lab for Genetics and Genomics and Leuven Institute for Beer Research, Leuven University, Gaston Geenslaan 1, B–3001, Leuven, Belgium,*

[7]*Center for Computational and Integrative Biology and Department of Computer Science, Rutgers University, Camden, NJ 08102, USA,*

[8]*Department of Plant Biology and Pathology, Rutgers University, New Brunswick NJ 08901, USA, and*

[9]*School of Information Technology and Mathematical Sciences, University of South Australia, Mawson Lakes, South Australia, Australia*

Received 29 June 2014; revised 4 November 2015; accepted 9 November 2015; published online 21 November 2015.
*For correspondence (e-mail jkim@marine.rutgers.edu).

SUMMARY

Diatoms (Bacillarophyceae) are photosynthetic unicellular microalgae that have risen to ecological prominence in oceans over the past 30 million years. They are of interest as potential feedstocks for sustainable biofuels. Maximizing production of these feedstocks will require genetic modifications and an understanding of algal metabolism. These processes may benefit from genome-scale models, which predict intracellular fluxes and theoretical yields, as well as the viability of knockout and knock-in transformants. Here we present a genome-scale metabolic model of a fully sequenced and transformable diatom: *Phaeodactylum tricornutum*. The metabolic network was constructed using the *P. tricornutum* genome, biochemical literature, and online bioinformatic databases. Intracellular fluxes in *P. tricornutum* were calculated for autotrophic, mixotrophic and heterotrophic growth conditions, as well as knockout conditions that explore the *in silico* role of glycolytic enzymes in the mitochondrion. The flux distribution for lower glycolysis in the mitochondrion depended on which transporters for TCA cycle metabolites were included in the model. The growth rate predictions were validated against experimental data obtained using chemostats. Two published studies on this organism were used to validate model predictions for cyclic electron flow under autotrophic conditions, and fluxes through the phosphoketolase, glycine and serine synthesis pathways under mixotrophic conditions. Several gaps in annotation were also identified. The model also explored unusual features of diatom metabolism, such as the presence of lower glycolysis pathways in the mitochondrion, as well as differences between *P. tricornutum* and other photosynthetic organisms.

Keywords: *Phaeodactylum tricornutum*, computational model, intermediate metabolism, glycolysis, ancient eukaryotic metabolism, biofuels.

INTRODUCTION

Diatoms (Bacillarophyceae) are photosynthetic eukaryotic microalgae that constitute the most abundant and diverse group of marine eukaryotic phytoplankton. They account for approximately half of the ocean's primary productivity, or approximately 20% of the world's productivity (Nelson *et al.*, 1995; Field *et al.*, 1998). They are excellent candidates for sustainable production of biofuel feedstock, not only because many strains are hardy, versatile and efficient in carbon fixation, but also because their major storage products are lipids (Hildebrand *et al.*, 2012; Levitan *et al.*, 2014). Interest in harnessing them to obtain lipids as feedstocks for biodiesel or nutritional supplements has grown over the past decade, because they potentially synthesize orders of magnitude more oil per hectare compared to terrestrial plants (Lebeau *et al.*, 2003, Chisti, 2007; Hu *et al.*, 2008, Graham *et al.*, 2011; Levitan *et al.*, 2014). They also do not have to compete for agricultural land, and may be particularly environmentally beneficial when biofuel feedstock production can be combined to treat wastewater (Hoffman, 1998; Mehta and Gaur, 2005; Clarens *et al.*, 2010).

Sequencing has made it possible to reconstruct genome-scale metabolic networks of model organisms, or organisms of interest for biotechnological purposes. Flux balance analysis (FBA) is a widely used approach for studying metabolic networks. This approach considers all the known metabolic reactions in an organism, based on genomic information or biochemical data, and calculates the fluxes of metabolites through this network using linear programming, knowledge of reaction stoichiometries and biomass composition, and additional constraints, such as limits on nutrient uptake or metabolite excretion rates. It may be used to predict growth rates, theoretical yields of metabolites under various conditions, and the viability of knockout or knock-in transformants. Furthermore, FBA may be used to guide the rational engineering of microorganisms, i.e. it may identify pathways in microbes that lead to efficient re-distribution of metabolites towards a desired end product or goal. This method has been successfully applied to the design of *Escherichia coli* strains for over-production of a diverse range of metabolites, including ethanol (Pharkya and Maranas, 2006), various amino acids (Pharkya *et al.*, 2003; Lee *et al.*, 2007), hydrogen and vanillin (Pharkya *et al.*, 2004), and in *Saccharomyces cerevisiae* strains to over-produce TCA metabolites (Zelle *et al.*, 2008). It has also been used to determine optimal gene knockouts to maximally increase respiration rates in *Geobacter sulfurreducens* and hence increase its bioremediation capabilities (Izallalen *et al.*, 2008).

Flux balance analysis (FBA) has been used to model photosynthetic metabolism in the cyanobacteria *Synechocystis* sp. and *Arthrospira platensis*, and the green algae *Chlamydomonas reinhardtii* (Boyle and Morgan, 2009; Klanchui *et al.*, 2012; Knoop *et al.*, 2013) and *Chlorella* sp. FC2 IITG (Muthuraj *et al.*, 2013). However, to date, only one FBA model of a stramenopile exists (Prigent *et al.*, 2014). Stramenopiles, which include diatoms, chrysophytes (golden algae) and brown algae, as well as some non-photosynthetic micro-organisms, differ from green algae and cyanobacteria in terms of cell structure and metabolism, pigmentation, nutrient uptake and recycling, mixotrophy, the presence of a urea cycle, and mechanisms for carbon concentration (Kooistra *et al.*, 2007). Diatoms in particular have some highly unusual features, including the presence of the Entner–Doudoroff pathway, connected to the lower half of glycolysis in the mitochondrion (Kroth *et al.*, 2008, Fabris *et al.*, 2012). Major questions remain concerning the origin and function of mitochondrial glycolysis. Traditionally, cytosolic glycolysis has been viewed as the original or ancestral form of eukaryotic energy metabolism (Whatley *et al.*, 1979; Blackstone, 1995). Alternatively, it has been hypothesized that cytosolic glycolysis originated from the eubacteria that eventually became mitochondria, and that mitochondrial glycolysis is an ancient eukaryotic trait (Martin and Muller 1998). Evidence for the latter hypothesis is provided by the observations that (i) diatoms and non-photosynthetic oomycetes have a mitochondrial glycolytic pathway, and (ii) all glycolytic enzymes of the eukaryotic cytosol analyzed to date are more similar to eubacterial homologs than to homologs in archaebacteria, the lineage that contains ancestral eukaryotes (Martin *et al.*, 1993; Keeling and Doolittle, 1997; Liaud *et al.*, 2000).

Given that diatoms (i) possess highly unusual features that may be relics of ancient eukaryotic metabolism, which has a bearing on questions regarding the origins of eukaryotic glycolysis, and (ii) are excellent candidates for biodiesel feedstocks, the absence of a genome-scale model of a diatom represents a large gap in our knowledge base. This paper presents a genome-scale metabolic model of the pennate diatom *P. tricornutum*. This organism was chosen because it is sequenced, and data on biomass composition are available (Bowler *et al.*, 2008; Frada *et al.* 2013; Guerra *et al.*, 2013). The metabolic network is based on the *P. tricornutum* genome, biochemical literature, online bioinformatic databases (Karp *et al.*, 2000; Fabris *et al.*, 2012), and extensive manual curation. The model simulates global growth behaviors under autotrophic, mixotrophic and heterotrophic conditions, as well as knockout conditions that explore the role of lower glycolysis in the mitochondrion. Although *P. tricornutum* is not normally capable of growing solely on glucose in the dark, this condition was simulated as a means to understand the organism either under genetic modification (Zaslavskaia *et al.*, 2001) or during night-time metabolism. Mixotrophic growth conditions are

of special interest for biofuel production (Wang *et al.*, 2012). The results of *in silico* predictions for the autotrophic case were validated by experimental evidence from previous studies, and results for the mixotrophic case were compared with data from a recent isotope labeling study (Zheng *et al.*, 2013). Our goal is to provide further knowledge of this system, as well as an interactive platform for integrating high-throughput data in the future, which may provide insights into the optimal method to engineer diatoms for bioproducts.

RESULTS AND DISCUSSION

A stoichiometric model of metabolism was constructed for *P. tricornutum* from genomic databases and literature data. The network includes glycolysis, the TCA cycle, oxidative and reductive pentose phosphate pathways, the phosphoketolase pathway, and amino acid, nucleotide, chlorophyll, chrysolaminarin and lipid synthesis pathways.

Network construction

The metabolic network of *P. tricornutum* consists of 587 metabolites and 849 metabolic reactions, not including the biomass equation (Table 1). For many of the pathways in *P. tricornutum*, most of the enzymes of the pathway are annotated but with gaps (Table S1). In these cases, the missing enzymes were included in the model on the assumption that the alga possesses the complete pathway, but the enzymes were missed in annotation. In the course of constructing the model, several missing enzymes were identified and given a confidence score of 1, based on the convention described by Thiele and Palsson (2010) (Table S2). A confidence score of 1 indicates that there is no genomic evidence for the enzyme, but the enzyme is included in the model so that known pathways may be completely included.

Localization of enzymes and metabolites

The reactions of the network were localized into four compartments: cytosol, mitochondria (matrix), chloroplast and peroxisome. The chloroplast was further divided into the stroma and lumen. Although the mitrochondrion comprises an inter-membrane space and matrix, the inter-membrane space is chemically equivalent to the cytosol; 'mitochondrion' therefore refers to the matrix. Subcellular localization was obtained using software from TargetP (Emanuelsson *et al.*, 2000) or Mitoprot (Claros and Vincens, 1996). When no signaling peptide was predicted, the recommendation in Thiele and Palsson (2010) was followed, and the enzyme was designated as cytosolic. In some cases, if these programs did not predict that an enzyme has a signal peptide but all the enzymes of its substrates and products were known to be located in a certain compartment, the enzyme was considered to be located in that compartment.

Biomass formation equations

The macromolecular composition of cells under each growth condition was based on experiments performed in chemostats (Methods S1), as well as data from the literature. The macromolecular composition consisted of amino acids, nucleotides, lipids, chlorophyll and carbohydrates (Table 1).

Model constraints

Flux maps for growth under autotrophic, mixotrophic and heterotrophic conditions were constructed using the network reconstruction and FBA (Figure 1a–d). During autotrophic growth, *P. tricornutum* fixes carbon dioxide and

Table 1 Biomass formation: standing stocks of cell constituents

	Amount (mol kg^{-1} biomass)
DNA	0.0009
RNA	0.001
Carbohydrate	0.15
Chlorophyll *a*	0.01
Growth-associated ATP	29.89
Non-growth-associated ATP	0.36-7.60
Monogalactosyl diacylglycerols	0.18
Triacylglycerols	0.14
Digalactosyl diacylglycerols	0.07
Sulphoquinovosyl diacylglycerols	0.10
Alanine	0.15
Arginine	0.07
Aspartic acid	0.12
Cysteine	0.01
Glutamic acid	0.14
Glycine	0.15
Histidine	0.02
Isoleucine	0.07
Leucine	0.11
Lysine	0.07
Methionine	0.02
Phenylalanine	0.08
Proline	0.10
Serine	0.11
Threonine	0.09
Tyrosine	0.04
Valine	0.09
Tryptophan	0.01
Ornithine	0.02
Asparagine	0.08
Glutamine	0.51

The coefficients for the equation describing the formation of biomass are given in units of moles per kg biomass (dry weight). Total amino acids, carbohydrate, lipid and chlorophyll were measured. DNA amount was calculated from knowledge of genome size (Bowler *et al.*, 2008). Total RNA amount was taken from Fernandez-Reiriz *et al.* (1989). The components of total lipids were taken from Yongmanitchai and Ward (1992). The overall amino acid content was based on our own measurements (Methods S1), but the amino acid composition was taken from Marsot *et al.* (1991) and Brown (1991). Growth-associated ATP maintenance was estimated using a *C. reinhardtii* model (Boyle and Morgan, 2009).

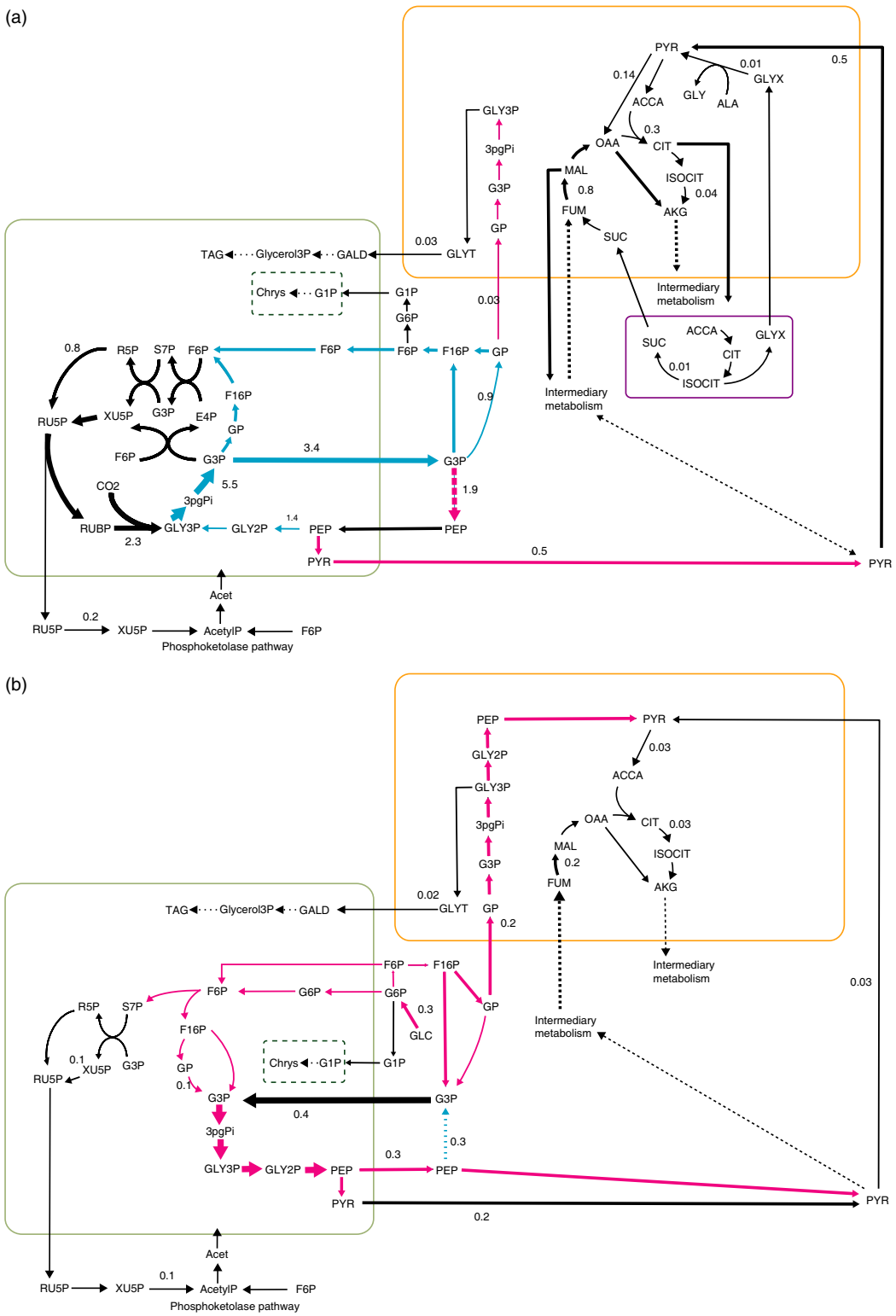

Figure 1. Metabolic flux maps for *P. tricornutum* under (a) autotrophic growth conditions, (b) heterotrophic (dark) conditions, (c) mixotrophic conditions comprising glucose and light, and (d) mixotrophic conditions comprising glucose, light and inorganic carbon.
The green rectangle represents the chloroplast, the orange rectangle represents the mitochondrion, the purple rectangle represents the peroxisome, and the dashed green line in the plastid represents a separate compartment for chrysolaminarin. Blue arrows are reactions that run in the direction of gluconeogenesis, and pink reactions run in the direction of glycolysis. The thickness of the arrows correlated with relative flux predictions. Abbreviations for metabolites are given in Table S2.

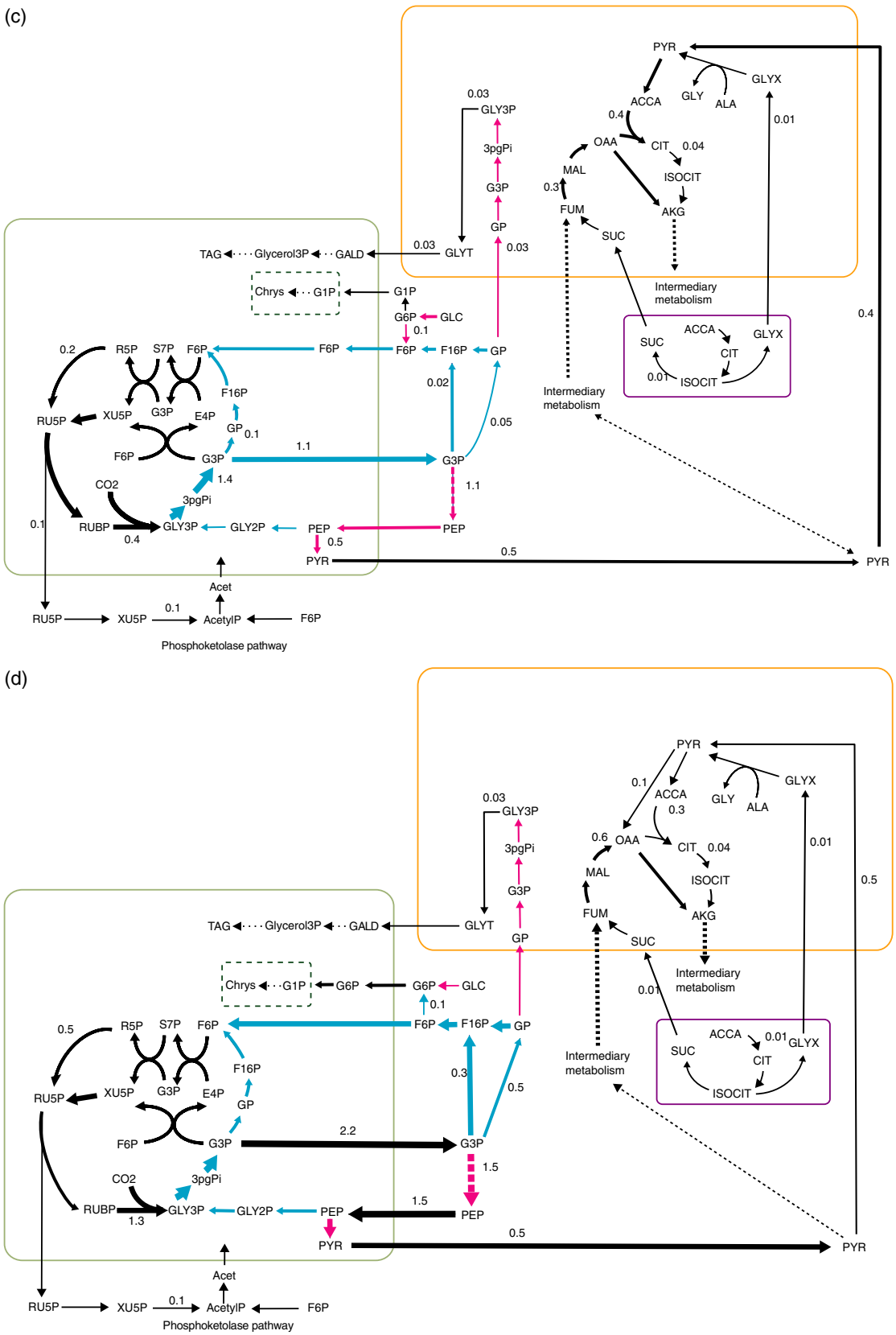

Figure 1. Continued.

bicarbonate, and uses water as the source of reductant, light as the source of energy, and micronutrients as essential components of the catalytic reactions. In order to constrain the model under this condition, a maximum uptake rate for CO_2 was imposed using the values 1.3 mmol CO_2 g^{-1} dry weight h^{-1}, a HCO_3 uptake rate of 0.65 mmol g^{-1} dry weight h^{-1}, or a maximum uptake of 1.95 mmol g^{-1} dry weight h^{-1} for total inorganic carbon (Ci); these values are based on measurements of uptake rates in *P. tricornutum* cells during steady-state photosynthesis (Burkhardt *et al*., 2001). The mixotrophic conditions, under which both photoautotrophy and glucose consumption occur, were as follows: (i) growth with glucose in the light, with no CO_2 or bicarbonate (Ci), and (ii) growth with glucose in the light, with Ci. Heterotrophic conditions, i.e. growth on glucose in the dark, were also modeled. In all cases, the input of carbon (number of carbons) was equal to the input of carbons in the autotrophic case to facilitate direct comparisons between conditions; the maximum glucose uptake rate for mixotrophic case 1 and the heterotrophic case were therefore set to 0.325 mmol g^{-1} dry weight h^{-1}, which represents the Ci uptake rate of the autotrophic case (1.95 mmol g^{-1} dry weight h^{-1}) divided by the number of carbons (six) in glucose. For mixotrophic case 2, half the carbon came from inorganic carbon and half from glucose.

Simulations of heterotrophic or mixotrophic growth may be seen as tools to assess the metabolic capabilities of this organism in different genetic backgrounds. Some wild-type *P. tricornutum* are able to utilize glucose when grown in the light (Zheng *et al*., 2013), although different strains have different efficiencies of glucose metabolism, while some cannot use glucose at all (Huang *et al*., 2015). Wild-type *P. tricornutum* cannot grow on the sugar heterotrophically. By inserting a human glucose transporter, Zaslavskaia *et al.* (2001) engineered a *P. tricornutum* strain that was able to grow in the dark as the sole carbon source. The model simulations for the heterotrophic condition may therefore be viewed as either a simulation of night-time metabolism, or growth in a different genetic background.

The biomass composition was set to be the same across all conditions (based on the autotrophic case). However, as biomass probably differs for mixotrophic and heterotrophic conditions in actual cultures, a sensitivity analysis was performed whereby different versions of the mixotrophic and heterotrophic conditions were tested, using different biomass values of protein, carbohydrate, lipids, chlorophyll *a* and RNA (Table S1). The DNA value was kept the same. In total, three versions of the mixotrophic biomass and three versions of the heterotrophic biomass were tested. Where available in the literature, data for *P. tricornutum* were used; if this was not available, data from other diatoms were used. For the mixotrophic cases, data for protein and carbohydrate were not available from diatom studies; for these cases, *C. reinhardtii* data were used. The three versions of heterotrophic biomass were based on diatom data.

Several mixotrophic and heterotrophic cases were tested to cover multiple possibilities for biomass and address various discrepancies in the literature; for example, the studies by Liu *et al.* (2009) and Cerón García *et al.* (2005) reported that mixotrophic *P. tricornutum* had approximately 14% and 64% less chlorophyll, respectively, compared with autotrophically grown cultures. A study using the diatom *Chaetoceros* sp. (Cheirsilp and Torpee, 2011) reported 4% less lipid, while a study using the diatom *Cyclotella cryptica* (Pahl *et al.*, 2010) reported 38% more lipid in heterotrophically grown cultures compared with autotrophic cultures. Version 3 for mixotrophy and version 3 for heterotrophy were the same as the corresponding version 1 except that RNA was doubled; this was tested because there are no available data for changes in RNA under mixotrophy or heterotrophy.

The sensitivity analysis revealed that changes in the relative amounts of lipid, carbohydrate, chlorophyll, protein or RNA did not change the overall flux pattern (the reactions that carried flux, or the directions of reversible reactions), but did change the amount of flux through those reactions slightly. The overall growth rate also varied, depending on which biomass equation was used, although the difference never varied by more than 16% compared with the original. As a result, we expect the model to reliably predict the overall flux distribution pattern, although we do not expect to predict the specific quantitative values of individual fluxes.

Autotrophic simulation results

Under autotrophic growth, the model predicts that the doubling time of the cell is 13.9 h, which is within the range of our experimental observations: 11.9–17.7 h (Table S1). The fluxes with the largest magnitudes were found for the reactions of the Calvin cycle and photophosphorylation (Figure 1a). Photophosphorylation provides the energy required to regenerate glyceraldehyde-3–phosphate (G3P) from glycerate-3–phosphate. As there are no known direct NAD(P)H transporters between the plastid and the cytosol, reducing equivalents in the form of G3P are transferred from the plastid to meet the demands for NAD(P)H in the cytosol. In contrast, ATP in the cytosol may be obtained through ATP/ADP translocators between the cytosol and mitochondrion, as well as through ATP-producing reactions in the cytosol.

Light drives linear electron flow from water to NADPH through the two photosystems (PSI and PSII), and pumps protons to generate an electrochemical proton gradient across the thylakoid membrane, which in turn generates ATP in the plastid. If linear electron flow were the only process to produce ATP and NADPH, this would generate an

ATP/NADPH ratio that is insufficient to fuel CO_2 assimilation by the Calvin cycle (Allen, 2002). In Viridiplantae (green algae and land plants), cyclic electron flow (CEF) around PSI plays a large role in making up the shortfall of ATP. In diatoms, however, the PSII to PSI ratio is two or more, and CEF around PSI is not easily accommodated.

Growth and maintenance ATP requirements were included to account for the unknown ATP requirements associated with transport, biosynthesis and polymerization, as well as cell maintenance, homeostasis and repair. In *C. reinhardtii*, growth-associated maintenance is estimated to be approximately 29.89 mmol ATP g^{-1} dry weight h^{-1} (Boyle and Morgan, 2009). Non-growth-associated ATP requirements (maintenance ATP) range from 0.36 to 7.60 mmol ATP g^{-1} dry weight h^{-1} for *Lactobacillus plantarum* and *E. coli*, respectively, and is reported to be approximately 1.50 mmol ATP g^{-1} dry weight h^{-1} in *C. reinhardtii* (Reed *et al.*, 2003; Boyle and Morgan, 2009). As these values are not known for *P. tricornutum*, the default growth and non-growth-associated maintenance requirements were set to be the same as that of *C. reinhardtii*, or 29.89 mmol ATP g^{-1} dry weight h^{-1} and 1.50 mmol ATP g^{-1} dry weight h^{-1}, respectively. Using this ATP maintenance requirement, our model predicts that, in *P. tricornutum*, ATP is synthesized in the plastid entirely through linear electron flow, i.e. there is no CEF around PSI. This is in agreement with the experimental results obtained by Bailleul *et al.* (2015), who reported that CEF played a negligible role in regulating ATP/NADPH levels in diatoms. More generally, they found that reducing equivalents generated in the plastid were exported to the mitochondria to generate additional ATP, i.e. a partial re-routing of photosynthate towards mitochondrial respiration to optimize the ATP/NADPH ratio in the plastid. This is in agreement with the model, which predicts that the flux through mitochondrion ATP synthase is approximately 1.2 times higher than through ATP synthase in the plastid.

In order to investigate how changes in ATP maintenance change the predicted flux distributions, as well as flux through CEF, the autotrophic condition was tested using two extremes for the non-growth ATP maintenance value, i.e. 0.36 and 7.60 mmol ATP g^{-1} dry weight h^{-1} (the ATP maintenance values for *L. plantarum* and *E. coli*, respectively), while keeping the growth-associated maintenance as 29.89 mmol ATP g^{-1} dry weight h^{-1}. The results show that changing the ATP maintenance value causes the amount of flux for individual reactions to change, although the reactions carrying flux, and the overall flux distribution, do not change. However, at the highest non-growth-associated ATP maintenance value of 7.6 mmol ATP g^{-1} dry weight h^{-1}, there is a small flux through CEF. More generally, flux through CEF appears when the non-growth ATP maintenance is set to approximately 7.4 mmol ATP g^{-1} dry weight h^{-1} or higher (Table 2). Therefore, *P. tricornutum* is predicted to have zero CEF only as long as the total ATP maintenance (growth-associated and non-growth-associated) is lower than 29.89 + 7.40 mmol ATP g^{-1} dry weight h^{-1}.

Table 2 Effect of changes in ATP maintenance on PSI, PSII and CEF flux (mmol g^{-1} dry weight h^{-1})

	Non-growth-associated ATP maintenance (mmol g^{-1} dry weight h^{-1})				
	0.36	1.5	7.4	7.6	1.5 (minimum PSII)
PSII	6.60	6.85	8.74	8.79	5.33
PSI	6.60	6.85	8.75	8.84	19.13
CEF	0.00	0.00	0.01	0.05	13.80

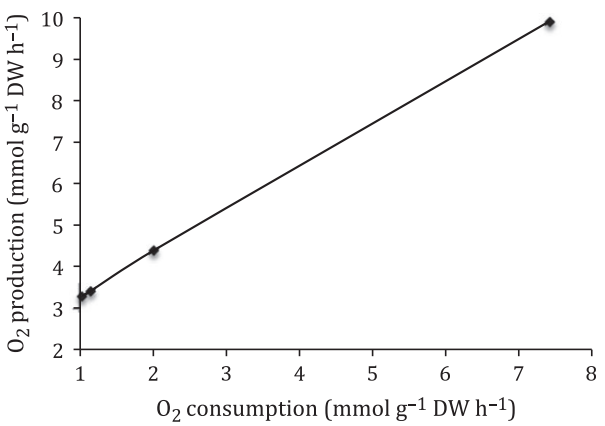

Figure 2. O_2 consumption for aerobic respiration increases linearly with O_2 production by PSII.
From left to right, the points represent ATP maintenance values of 0.36, 1.5, 7.6 and 50 mmol g^{-1} dry weight h^{-1}.

The results of the model simulations also show that as ATP maintenance is increased, the cell absorbs more light, and O_2 consumption for aerobic respiration increases linearly with O_2 production by PSII (Figure 2). This is in agreement with the study by Bailleul *et al.* (2015), which found that O_2 consumption increased linearly with O_2 production; a constant proportion of the electron flow from photosynthesis was re-routed for O_2 consumption, regardless of light intensity.

Additionally, mitochondrial respiration is directly involved in optimizing photosynthesis. Maintaining an optimal ATP/NADPH ratio should therefore be contingent on sufficient communication between the plastid and mitochondria. Our model results are in agreement with this: if the mitochondrial ADP/ATP translocator (which exports ATP out of the mitochondria in exchange for ADP) is removed, there is flux through CEF, and the flux through ATP synthase in the plastid increases approximately 1.7-fold (Table 3). Moreover, in agreement with the

Table 3 Knockouts in the mitochondria and plastid affect flux through PSI, PSII, CEF and ATP synthases

	Wild-type	PSII	PSI	ATPS_P	CEF	ATPS_M
No mitochondrial ATP/ADP translocator	6.9	5.5	5.3	5.3	6.4	6.2
No mitochondrial complex IV	6.9	18.3	19.1	19.1	14.3	16.0
No mitochondrial complex III	3.4	9.2	9.6	9.6	7.2	8.0
No plastid ATP/ADP translocator	0.0	12.8	13.8	13.8	7.9	9.8
No plastid transporters, or plastid ATP/ADP translocator	7.7	0.7	0.1	0.1	3.3	2.1

Predicted fluxes (mmol g^{-1} dry weight h^{-1}) for hypothetical knockouts in the mitochondria and plastids, on PSII, PSI and ATP synthase in the plastid (ATPS_P), cyclic electron flow (CEF), and ATP synthase in the mitochondria (ATPS_M).

experimental results of Bailleul *et al.* (2015), who found that mitochondrial dysfunction negatively affected photosynthesis, the model predicts that deletion of one of the enzymes of the electron transport chain in the mitochondria (such as cytochrome *c* oxidase or coenzyme Q/cytochrome *c* reductase), lowers flux through PSII by 22%. If the plastid ADP/ATP translocator (JGI ID 49533), which directly imports ATP into the plastid, is deleted, there is flux through CEF, and the flux through ATP synthase in the plastid increases by approximately 109%. If the majority of transporters that transfer metabolites (G3P, pyruvate, aspartate and oxaloacetate) in or out of the plastid are also deleted, the model predicts that flux through CEF increases further, and ATP synthesis in the mitochondrion decreases (Table 3).

In our model, transporters connect the plastid to the cytosol, which is connected to the mitochondrion by its own set of transporters. The plastidic transporters in this organism remain to be determined, but *P. tricornutum* is likely to at least have a triose phosphate translocator (Weber *et al.*, 2006). Our model contains a transporter for G3P, as well as transporters that are known in plants (Fischer, 2011). Using these transporters, pyruvate and G3P are removed from the plastid; the G3P is converted to dihydroxyacetone in the cytosol, and sent into the mitochondrion (Figure 1a). Pyruvate, fumarate and oxaloacetate enter the mitochondrion from the cytosol. The mitochondrion then releases 2–oxoglutarate and malate, as well as a small amount of glycerate, which is imported to the plastid. The flux distribution of metabolites is contingent on which transporters exist, but, even in a highly conservative approach, in which only transporters for G3P, glutamate and 2–oxoglutarate are included, the flux distribution shows partial re-routing of dihydroxyacetone flowing towards the mitochondria. Given that it is highly likely that the majority of transporters in this organism remain to be discovered, the actual flux pattern is probably closer to the flux distribution produced with the greater number of transporters.

The flux distribution solution obtained by FBA is not unique; there are multiple solutions or flux distributions that maximize biomass production. An optimization function, Euclidean norm minimization (discussed below), was therefore used in conjunction with FBA to choose the (unique) flux vector with the lowest overall magnitude that still produces maximum biomass flux. The rationale for this optimization function is the assumption that fluxes for metabolic reactions in organisms are not larger than necessary for biomass production. The predictions for CEF are partially a result of this optimization function. Indeed, flux variability analysis, which calculates the maximum and minimum allowable values of each flux while still achieving maximal biomass flux, shows that, with ATP maintenance of 1.5 mmol g^{-1} dry weight h^{-1}, CEF varies over a relatively large range: 0–489 mmol g^{-1} dry weight h^{-1}. To a degree, linear electron flow trades off with CEF. If an artificial limit is placed on the linear electron flow by limiting PSII to 5.33 mmol g^{-1} dry weight h^{-1}, which is the minimum allowable PSII flux (obtained via flux variability analysis), CEF flux increases from 0 to 13.8 mmol g^{-1} dry weight h^{-1}, for an ATP maintenance value of 1.5 mmol ATP g^{-1} dry weight h^{-1} (Table 2). More generally, when ATP maintenance is lower than 7.4 mmol ATP g^{-1} dry weight h^{-1}, increased demand for ATP is met by higher flux through linear electron flow. For values higher than 7.4 mmol ATP g^{-1} dry weight h^{-1}, increased ATP demand is met by increased flux through both linear electron flow and CEF (Table 2).

However, this is only part of the explanation for the CEF results. As the optimization function also results in solutions with the shortest pathways, the question remains why such a large proportion of the plastid's ATP is synthesized in mitochondria, which requires a longer pathway than simply using CEF to optimize NADPH/ATP levels in the plastid (which requires greater absorption of photons, which are not in limited supply in our model). The reason for this may lie in the fact that sending reductant into the mitochondrion is a way to provide carbon skeletons in the mitochondrion for use in anabolic pathways, in addition to facilitating mitochondrial ATP production. As such, it is more 'efficient' than simply using CEF to meet ATP demands, at least when ATP maintenance (growth- and non-growth associated) is lower than 29.89 + 7.4 mmol ATP g^{-1} dry weight h^{-1}. Additionally, photosynthetically

derived dihydroxyacetone may be sent into the mitochondria, where the lower half of glycolysis occurs in *P. tricornutum*; this produces NADH and ATP in the mitochondria independently of the TCA cycle, possibly allowing for more precision in optimizing fluxes for specific metabolite composition. CEF also plays a large role in adjusting the ATP/NAD(P)H ratio in plants and green algae; mitochondrial ATP synthesis is sometimes 'preferred' over CEF. For example, in the leaves of pea (*Pisum sativum*), mitochondrial ATP synthesis is 'chosen' over CEF, which mainly takes effect only at high irradiances combined with very low CO_2 concentrations (Harbinson and Foyer, 1991). More generally, the results are in agreement with numerous studies that demonstrate a highly inter-dependent relationship between plastids and mitochondria in plants and green algae (Hoefnagel *et al.*, 1998; Matsuo and Obokata, 2006). *P. tricornutum* is no exception; in fact, electron micrographs show that plastids and mitochondria in this organism are tightly associated physically (Bailleul *et al.*, 2015).

Heterotrophic and mixotrophic simulation results

Under heterotrophy, the doubling time of the cell is predicted to be approximately 21 h, approximately 54% longer than for the autotrophic case. In cultures, algae grown heterotrophically may sometimes have higher growth rates than under autotrophy, as light shading effects do not matter, and cells grow to greater density. However, when normalized against the carbon input, heterotrophic growth is less efficient, and models for several autotrophs predict this (Shastri and Morgan, 2005; Boyle and Morgan, 2009; Muthuraj *et al.*, 2013). Under heterotrophy, some carbon is lost due to respiration; there is net release of CO_2, so less carbon is available for biomass production than under autotrophy. Under autotrophy, on the other hand, CO_2 produced from respiration may be recycled.

The flux distribution for the heterotrophic case starts with the uptake of glucose, which is converted to glucose-6–phosphate in the cytosol (Figure 1b). In *P. tricornutum*, the complete set of glycolytic reactions is found in both the plastid and cytosol, and the lower half of these reactions is found in the mitochondrion. Under heterotrophic growth, these reactions run in the direction of ATP or NAD(P)H formation in the plastid (reactions indicated by pink arrows in Figure 1b); this provides the energy required for lipid synthesis in the plastid. In general, the highest fluxes under heterotrophy are found in the reactions of the electron transport chain and ATP synthase in the mitochondrion.

If a mixotrophic case is run with the same amount of Ci as the autotrophic case (1.95 mmol g^{-1} dry weight h^{-1}), plus extra carbon in the form of glucose, biomass production increases (or the doubling time decreases). If 0.325 mmol g^{-1} dry weight h^{-1} glucose is supplied in addition to 1.95 mmol g^{-1} dry weight h^{-1} Ci, the doubling time is 6.9 h, or approximately half the doubling time of the autotrophic case. When mixotrophic conditions are set so that the carbon input is the same as in the autotrophic case, whether it is (i) growth with glucose in the light without Ci (Figure 1c), or (ii) growth with glucose in the light with Ci (Figure 1d), the doubling times are the same as in the autotrophic condition. The flux distributions for these two versions of mixotrophy are similar, although the flux through the Calvin cycle in case 1 is lower, because more of the initial carbon is in the organic form. Additionally, the fluxes are opposite for glucose-6–phosphate isomerase (EC 5.3.1.9), since the two versions of mixotrophy have differing glucose inputs. In general, it is unsurprising that the flux distributions are similar for these two cases, as CO_2 produced in the mitochondrion may be recycled and moved into the plastid to fuel RuBisCO and carbon fixation.

P. tricornutum is predicted to have the components of the glycerol-3–phosphate shuttle. The shuttle adjusts the NADH/NAD+ ratio in the cytosol as it converts cytosolic NADH into NAD+, and contributes to oxidative phosphorylation in the mitochondrion. The model results predict that this shuttle is active under mixotrophic conditions, possibly due to the fact that a larger proportion of cytosolic G3P is turned into glycerate-1,3–diphosphate via G3P dehydrogenase (which produces NADH and consumes NAD+), compared to the autotrophic case. In many organisms, mechanisms such as the malate/aspartate shuttle also function to regenerate NAD+ in the cytosol, but this requires the presence of a malate dehydrogenase in the cytosol, which has not been found in *P. tricornutum*.

All the light conditions (autotrophy and mixotrophy) also showed flux through either mitochondrial or plastidic pyruvate carboxylase (EC 6.4.1.1), which catalyzes the carboxylation of pyruvate to form oxaloacetate. This is in agreement with labeling experiments in *P. tricornutum* grown in the light on medium with inorganic carbon supplemented with U–^{13}C glucose (Zheng *et al.*, 2013). The labeling experiments revealed that carbon originating from U–^{13}C glucose was uniformly distributed across the metabolic network, and the majority of amino acids showed almost uniform ^{13}C enrichments, but some fragments derived from oxaloacetate and 2–oxoglutarate, which contain carbon fixed through anaplerotic reactions, were less enriched, implying that the carbon assimilated originated from dissolved inorganic carbon.

Other predictions

The TCA cycle is not predicted to run in a complete circle for all conditions, and the highest fluxes were between fumarate, malate and oxaloacetate. Under autotrophy, the majority of this oxaloacetate (approximately 85%) is then converted to 2–oxoglutarate and aspartate via aspartate aminotransferase (EC 2.6.1.1). The TCA cycle not only fuels ATP synthesis in the mitochondrion, but also provides

carbon skeletons for biosynthetic processes. Moreover, enzymes in the cytosol may bypass many of the steps in the TCA cycle. The non-cyclic flux pattern predicted by our model is consistent with findings from both labeling analyses and models in plants (Sweetlove *et al.*, 2010), in which non-cyclic flux modes were found to occur in both illuminated leaves and non-autotrophic tissues (seeds). The results for the heterotrophic condition were also in agreement with a flux balance model of heterotrophic Arabidopsis metabolism, which predicted that the reactions between 2–oxoglutarate and fumarate show no flux when the model is constrained solely by the need to synthesize biomass components in the correct proportions (Poolman *et al.*, 2009). A combination of a partial TCA cycle, which supports a low rate of mitochondrial oxidative phosphorylation, and substrate-level phosphorylation by glycolytic reactions, is sufficient to meet ATP demands.

The oxidative pentose phosphate pathway produces NADPH and pentose phosphates for biosynthesis of nucleotides, amino acids and fatty acids in the dark by decarboxylating glucose-6–phosphate. In contrast to higher plants, there is no complete oxidative pentose phosphate pathway in the plastid of *P. tricornutum*, although the complete oxidative pentose phosphate pathway appears to be present in the cytosol (Kroth *et al.*, 2008). In our results for the heterotrophic (dark) condition, there was no flux through the oxidative pentose phosphate pathway because regeneration of pentose phosphates occurs by other means, for example via phosphopentose epimerase (EC 5.1.3.1) in the cytosol or by reactions in the plastid. If future experimental work reveals that oxidative pentose phosphate pathway in the cytosol is in fact active, it suggests this pathway is active for reasons beyond maximizing biomass, i.e. the oxidative pentose phosphate pathway is part of a 'sub-optimal' flux solution and is not the metabolically favored pathway, unless other constraints are introduced.

The Entner–Doudoroff pathway and glycolysis in the mitochondrion

Diatoms are unusual in that they possess the complete second half of glycolysis in the mitochondria (Kroth *et al.*, 2008), as well as the Entner–Doudoroff pathway, an ancient glycolytic pathway that is predominantly restricted to prokaryotes (Fabris *et al.*, 2012). Mitochondrial glycolysis is currently believed to occur only in diatoms and non-photosynthetic oomycetes (Ginger *et al.*, 2010), and the purpose of this pathway in mitochondria is unknown. Smith *et al.* (2012) suggested that these enzymes may function in the direction of gluconeogenesis and that glycerate feeds into these reactions. Flux variability analysis suggests that this is a possibility, although, when using the optimization function, these enzymes are predicted to run in the direction of glycolysis, i.e. from dihydroxyacetone to glycerate-3–phosphate, for the growth conditions. These results depend on the presence of at least one transporter that allows glycolysis metabolites either into or out of the mitochondrion; a transporter for dihydroxyacetone phosphate is included in our model. Without it (or some other transporter), the glycolysis/gluconeogenesis pathway in the mitochondria becomes a metabolic 'dead end'. The inclusion of the transporter was therefore based on the assumption that the glycolysis/gluconeogenesis pathway in the mitochondrion is connected to the rest of the network.

In order to explore the role of glycolysis in the mitochondrion, various versions of our model were run, each with varying numbers of transporters between the mitochondrion and cytosol. Figure 1(a–d) is based on a 'liberal' version of the model, whereby multiple transporters were included to allow transfer of all the TCA metabolites between the mitochondrion and the cytosol. The assumption for this version was that *P. tricornutum* has all the transporters commonly found in plants or yeast. In a second version of the model, we limited the transporters to only those for which there is some degree of genomic evidence. This version (version 2) had only four putative transporters: an aspartate/glutamate shuttle, a malate/2–oxoglutarate shuttle, a citrate/malate shuttle, and a fumarate/succinate shuttle.

The flux results for version 2 were similar to those for the original version, but with a few differences: (i) CEF was active, (ii) the flux through the mitochondrial ATP synthase decreased approximately 56%, as there was less connection between the mitochondrion and the rest of the cell, (iii) pyruvate in the mitochondrion is converted into glycerate-3–phosphate via the action of hydroxypyruvate reductase (EC 1.1.1.29) and serine/pyruvate transaminase (EC 2.6.1.51), and (iv) glycerate-3–phosphate is converted into glycerone phosphate (GP) in the mitochondrion, and GP is transported out of the mitochondrion (Figure 3). The results underscore the importance of transporters for accurate flux predictions; the glycolysis/gluconeogenesis enzymes in the mitochondrion function in either direction, depending on the transporters included. Given that it is likely that we have under-represented the number of transporters in this organism, the actual flux distribution is probably closer to the results for the original version of the model.

The purpose of the Entner–Doudoroff pathway in diatoms remains unclear, but it may be a means to quickly adjust to environmental changes, enabling fast metabolic responses, despite the fact that it produces less energy per molecule of glucose than the Embden–Meyerhof–Parnas pathway (Fabris *et al.*, 2012). FBA predicts that there is no flux through the Entner–Doudoroff pathway, which is unsurprising given that the Entner–Doudoroff pathway is less efficient than the Embden–Meyerhof–Parnas pathway. When the entire glycolysis pathway in the mitochondrion is deleted, flux variability analysis shows that there is no allowable flux through the Entner–Doudoroff pathway; as

there is no outlet for the formation of G3P, it becomes a metabolic 'dead end'.

Comparing model results with isotopomer analysis of mixotrophic glucose metabolism

In the study by Zheng *et al.* (2013), *P. tricornutum* grown under mixotrophic glucose conditions was found to produce glycine and serine primarily from glyoxylate via the photorespiratory enzymes alanine/glyoxylate aminotransferase (AGAT, EC 2.6.1.44) and serine hydroxymethyltransferase (SHMT, EC 2.1.2.1), rather than via the more 'conventional' pathway through 3–phosphoglycerate dehydrogenase. Our results also showed that these pathways are active in forming glycine and serine, with no flux through the more conventional pathway (Figure 4). For the mixotrophic growth condition (with glucose and Ci), our model predicted that 99% of the serine was produced from SHMT in the mitochondrion, with the rest produced from cystathione β–synthase (CBS) in the cytosol. Approximately 98.8% of the total glycine was produced by SHMT in the cytosol and 1.3% was produced by AGAT in the mitochondrion. There was no flux through the more conventional pathway from 3–phosphoglycerate (which utilizes 3–phosphoglycerate dehydrogenase, phosphoserine aminotransferase and phosphoserine phosphatase).

When phosphoglycerate mutase, which competes for 3–phosphoglycerate in the cytosol, is deleted, this does not increase the flux through 3–phosphoglycerate, indicating that competition for 3–phosphoglycerate is not the reason for the alternative route. Deleting SHMT in the cytosol and mitochondrion, as well as CBS, is also not sufficient to send flux through the conventional pathway. However, when all of these enzymes are deleted, together with AGAT2 (EC 2.6.1.51), there is flux through the pathway from 3–phosphoglycerate. If all SHMT enzymes, CBS and AGAT are deleted, there is no biomass produced. The

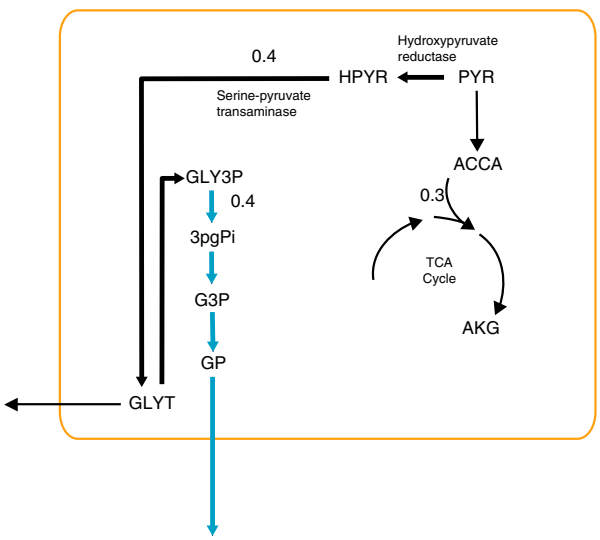

Figure 3. Metabolic flux map for *P. tricornutum* under autotrophic growth conditions with limited transporters for TCA metabolites.
The orange rectangle represents the mitochondrion. Abbreviations for metabolites are given in Table S2.

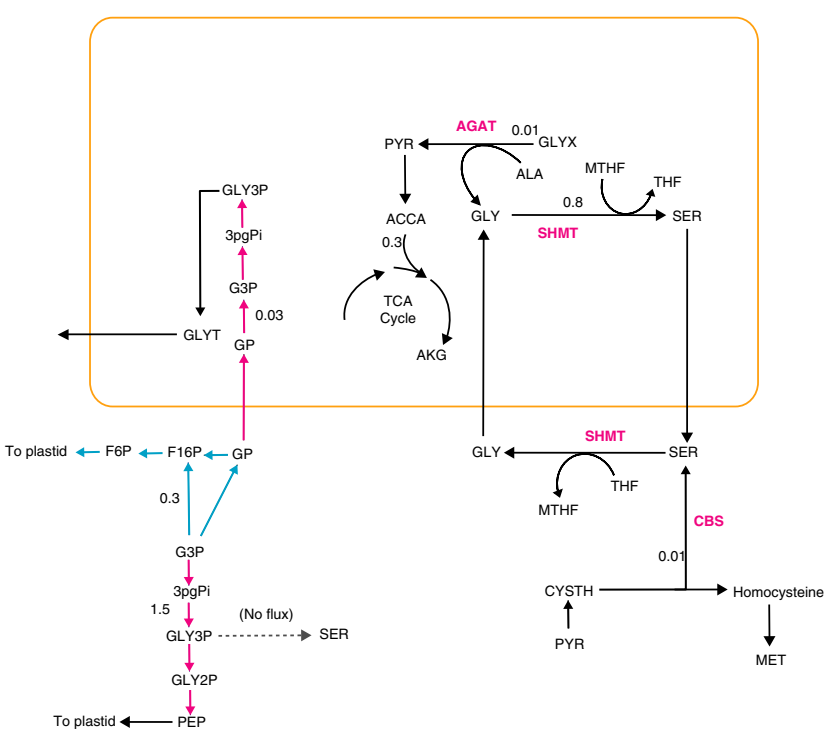

Figure 4. Overview of glycine and serine synthesis in *P. tricornutum* under mixotrophy (growth with glucose and inorganic carbon in the light), and related pathways.
The gray dashed arrow (no flux) represents the 'conventional' pathway for serine synthesis. The orange rectangle represents the mitochondrion. Abbreviations for metabolites and enzymes are given in Table S2.

optimization function may choose the SHMT, CBS and AGAT route for serine and glycine synthesis over the conventional pathway because these pathways lead to other amino acids that are synthesized concomitantly. Flux through CBS is also linked to methionine synthesis, for example, as the action of CBS leads to production of homocysteine, which may be used for methionine production via methionine synthase (EC 2.1.1.13).

The study by Zheng *et al.* (2013) also showed that there was flux through the Entner–Doudoroff and phosphoketolase pathways. Although our model did not predict flux through the Entner–Doudoroff pathway, the phosphoketolase pathway, which converts xylulose-5–phosphate to G3P and acetyl phosphate or fructose-6–phosphate to erythrose-4–phosphate and acetyl phosphate, was active for all growth conditions (Figure 1a–d). The acetyl phosphate that is formed from this pathway is converted to acetate and transferred into the plastid, where it is used to synthesize acetyl CoA, which feeds fatty acid synthesis.

Comparison to other model photosynthetic organisms

Flux distributions of *P. tricornutum* were compared to those for two other photosynthetic microbes for which FBA models exist: the green alga *Chlamydomonas reinhardtii*, and the cyanobacterium *Synechocystis* sp. PCC 6803 (Shastri and Morgan, 2005). As discussed above, *P. tricornutum* did not utilize cyclic electron flow to generate ATP in the plastid, and therefore had the highest calculated flux through the non-cyclic (i.e. linear) electron transport chain, normalized to CO_2 assimilated (Table 4). The output of O_2 was therefore slightly higher than for the other two organisms. However, as mentioned above, this was based on the assumption that *P. tricornutum* has ATP maintenance requirements that are not vastly different from those of *C. reinhardtii*.

In contrast to results from a *Synechocystis* model (Shastri and Morgan, 2005), malic enzyme was not found to be essential for heterotrophic growth; in fact, none of the conditions had flux through this enzyme. Additionally, phosphoenolpyruvate carboxylase, which converts bicarbonate and phosphoenolpyruvate to oxaloacetate, exists in *P. tricornutum*, but was not found to be necessary for maximizing biomass. This is in contrast to the chlorophyte alga, *Chlorella* sp. FC2 IITG, for which the flux through phosphoenolpyruvate carboxylase was high, especially under autotrophic conditions (Muthuraj *et al.*, 2013). This reaction has been suggested to play a role in adapting to low CO_2 conditions (Reinfelder *et al.*, 2000), but its flux was not found to increase as the CO_2 or bicarbonate uptake rate was set lower (to a third of the default case of 1.95 mmol g^{-1} dry weight h^{-1}). This is consistent with the view that this reaction may be part of a means to dissipate excess light energy in the diatom, or generate amino acid precursors, rather than play some role in maximizing biomass output under low CO_2. Further constraints on the model are needed in order to explore this effect.

Table 4 Comparison of selected fluxes with those in *Synechocystis* and *C. reinhardtii*

	Autotrophic growth fluxes (mol mol^{-1} CO_2)		
	P. tricornutum	*Synechocystis*	*C. reinhardtii*
Cyclic electron flow	0	0.5	1.9
Non-cyclic electron flow	5.5	4.8	3.5
Photons	5.5	5.3	5.4
O_2 output	0.7	0.6	0.4

In the plastids of land plants and green algae, some of the enzymes of the Calvin cycle, such as phosphoribulokinase, G3P dehydrogenase, fructose-1,6–bisphosphatase and seduheptulose-1,7–bisphosphatase, become inactive in the dark via a ferredoxin/thioredoxin system (Buchanan *et al.*, 1967; Holmgren *et al.*, 1977; Wolosiuk and Buchanan, 1977). In the current *P. tricornutum* model, deleting all of these enzymes for the dark condition is permissible (non-lethal), except in the case of G3P dehydrogenase; deletion of this enzyme in the plastid results in no biomass. Whether this is due to a key difference between diatoms and land plants/green algae, or the result of missing reactions or transporters in the model, remains to be elucidated. It is possible that light/dark regulation of plastidic enzymes simply differs in diatoms; in green algae and land plants, for example, the plastidic enzyme phosphoribulokinase (EC 2.7.1.19) is light-regulated via reversible reduction by reduced thioredoxin, while in the diatom *Odontella sinensis*, this enzyme is active under both light and dark conditions (Michels *et al.*, 2005).

Further work

The most recent *C. reinhardtii* model (Chang *et al.*, 2011) integrated biological and optical data to devise a model that predicts growth for various light sources, and experimentally verified transcripts for the enzymes in the network. This provides a higher predictive capacity and confidence in the model; it is a natural next step for progression of the *P. tricornutum* model. Additionally, knowledge of ATP requirements and transporters between compartments should increase the accuracy of predictions for the glycolysis/gluconeogenesis pathway in the mitochondria, the direction of flux for these reactions, the role that this pathway plays in carbon partitioning, and the percentage of ATP produced in the mitochondrion. Our model confirms recent findings showing that there is energetic coupling between plastids and mitochondria. Our model also predicts that glycerate produced in the mitochondria

is fed by the glycolysis pathway in the mitochondrion. Changes in these enzymes may therefore affect lipid synthesis, as glycerate may enter the plastid where it may potentially be shunted into glycerol, and ultimately into triacylglycerols. Glycerate also plays a role in photorespiration, although a plastidic glycerate kinase has yet to be identified in *P. tricornutum*. As our model also predicts flux through the phosphoketolase pathway, over-expression of this pathway, or of acetate kinase, which converts acetyl phosphate to acetate, may increase lipid synthesis. As genome annotation for this organism improves, and further constraints for the metabolic network are identified, the model will be an invaluable tool for determining the suitability of diatoms as a source of lipids for biodiesel. Gene expression studies will also be vital for modeling flux through pathways that are active outside conditions that biomass optimization, such as the Entner–Doudoroff pathway. Knowledge of transporters between organelles will also play an invaluable role in elucidating how diatoms have managed to be such successful contributors to primary productivity.

EXPERIMENTAL PROCEDURES

Network construction

A stoichiometric model of metabolism for *P. tricornutum* was constructed using genomic databases, DiatomCyc (http://www.diatomcyc.org/) (Fabris *et al.*, 2012) and journal articles. The reconstruction began using annotations from the KEGG database (http://www.genome.jp/kegg/) (Kanehisa and Goto, 2000) and DiatomCyc. The network includes the TCA cycle, oxidative and reductive pentose phosphate pathways, the Embden–Meyerhof–Parnas and Entner–Doudoroff pathways, the phosphoketolase pathway, and amino acid, nucleotide, chlorophyll, starch and lipid pathways. The reversibility of reactions was assessed using a combination of KEGG, MetaCyc (Caspi *et al.*, 2012) and BRENDA (Schomburg *et al.*, 2013). More specifically, the reversibility of reactions was first assessed using DiatomCyc, which is based on KEGG and MetaCyc. In most cases, there was agreement between all three databases regarding the reversibility of reactions. However, in some cases, there were discrepancies between the databases; in these cases, BRENDA was used to determine reaction reversibility.

Gaps in the network were addressed by first searching for the missing enzymes in other organisms, then using BLAST (Altschul *et al.*, 1990) to search the *P. tricornutum* database for similar amino acid sequences. Transporters were also searched for in this way. This led to inclusion of a malate/2–oxoglutarate antiporter (JGI ID 8990). The sequence was checked using the National Center for Biotechnology Information BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) program, to check hits in the Swissprot database to ensure that hits gave back malate/2-oxoglutarate antiporters; the first hit had an E value of 6E-62.

In some cases, 'gap' enzymes that were not found in databases were included; this was done in cases where the majority of the enzymes in a network were present, but the network was incomplete without the inclusion of these missing enzymes. One example is plastidic sedoheptulose bisphosphatase (EC 3.1.3.37). Both sedoheptulose 1,7–bisphosphate D–glyceraldehyde-3–phosphate lyase (which generates sedoheptulose 1,7–bisphosphate) and transketolase (which uses sedoheptulose-7–phosphate) are found in the plastid. Sedoheptulose bisphosphatase connects these two enzymes as it converts sedoheptulose 1,7–bisphosphate to sedoheptulose-7–phosphate. Another example in which a gap was filled was the case of NAD biosynthesis from aspartate. A reaction that produces nicotinate ribonucleotide and a separate reaction that consumes nicotinate adenine dinucleotide exist. Nicotinate mononucleotide adenyltransferase (EC 2.7.7.18) converts nicotinate ribonucleotide to nicotinate adenine dinucleotide, and fills the gap in this pathway. A list of all the gap-filling reactions is given in Table S1. These have a confidence score of 1. In total, 34 such reactions were included to fill gaps; this does not include transporter reactions.

Transporters between compartments

Although little is known about the shuttles that transport metabolites out of the plastid in this organism, several were assumed to exist as demands for both ATP and NAD(P)H for biosynthetic reactions outside the chloroplast must be met (Kramer and Evans, 2011). Triose phosphate translocators were included based on genomic information (Weber *et al.*, 2006). Isolated plastids from two marine diatoms were able to reduce added 3–phosphoglycerate, and inhibited by extraplastidic phosphate, which suggests that they have phosphate translocators (Wittpoth *et al.*, 1998). Other translocators were included as a means to bridge reactions that would otherwise be disconnected. For example, a UDP glucose pyrophosphorylase (EC 2.7.7.9) was encoded in the genome (JGI ID 23639); this catalyzes the conversion of UTP and glucose-1–phosphate to UDP-glucose and pyrophosphate, and is predicted to be located in the plastid, but 1,3-β–glucan synthase, which generates β–glucan from UDP-glucose, is located outside of the plastid. In this case, a transporter reaction was included to export UDP glucose out of the plastid. Other transporter reactions were included for similar reasons (export of amino acids, co-factors, sugars, etc.). *P. tricornutum* also has nucleotide transporters that shuttle nucleotides (synthesized in the cytosol) into the plastid (Ast *et al.*, 2009).

The mitochondrial inner membrane is impermeable to NADH, so reducing equivalents must cross the membrane. An ATP/ADP translocase (which translocates ATP generated by oxidative phosphorylation with ADP) is also needed, as ATP and ADP do not diffuse freely across the inner mitochondrial membrane. Although little is known about the specifics of the transporters across the inner membrane of the mitochondria of *P. tricornutum*, certain shuttles exist across a wide range of organisms, such as fungi, plants, animals and algae; these were included in the model on the assumption that similar mechanisms exist in the diatom. Some mitochondrial transporters were also included based on genomic evidence.

Calculation of flux distributions

The calculation of flux distributions was performed using FBA coupled with Euclidean norm minimization (Bonarius *et al.*, 1996), or flux variability analysis (Mahadevan and Schilling, 2003). Briefly, using the network reconstruction, we generated a stoichiometric matrix S with m rows and n columns, where m is the number of metabolites and n is the number of reactions. Metabolites in different compartments of the cell are represented as distinct metabolites, i.e. separate metabolic entities. The ijth element S_{ij} of S is the stoichiometric coefficient of the ith metabolite in the jth reaction. To find an FBA solution, we solve

$$Z^* = \max \quad v_{\text{biomass}} \tag{1}$$

subject to $Sv = 0$ and $a \leq v \leq b$, where $v_{biomass}$ indicates the biomass flux component of the flux distribution vector v, a and b are lower and upper bounds, respectively, on the fluxes, and $\leq$ indicates component-wise inequality. The optimal solution z^* calculated by Eqn (1) is the predicted growth rate.

The optimal solution for Eqn (1) is generally not unique, allowing an infinite number of possibilities for the intracellular flux distribution. We obtained a unique flux distribution through Euclidean norm minimization, which has been shown to give good predictions of intracellular fluxes in a variety of systems from microbial to mammalian (Bonarius *et al.*, 1996; Blank *et al.*, 2005; Schuetz *et al.*, 2007). In particular, we found that FBA coupled with Euclidean norm minimization gives predicted fluxes that show a mean uncentered Pearson correlation of 0.88 with fluxes determined by ^{13}C metabolic flux analysis in a dataset of nine experiments performed on *Saccharomyces cerevisiae* (Table S3). Thus, after solving Eqn (1), we solve

$$\min \sum_i v_i^2 \quad (2)$$

subject to $Sv = 0$, $a \leq v \leq b$, and $v_{biomass} = z^*$. The optimal solution v^* of expression (2) is the predicted flux distribution. This optimal solution vector is unique, meaning that, if there are two vectors v_1 and v_2 that satisfy the constraints of problem (2) and achieve its optimum, then $v_1 = v_2$. To prove this, we note that the objective function of problem (2) is strictly convex and its feasible set is convex. Because the optimal solution vector to the minimization of a strictly convex function over a convex set is unique (Bertsekas, 2003), the optimal solution vector to problem (2) is unique.

All model calculations were performed using MOST (Metabolic Optimization and Simulation) software (Kelley *et al.*, 2014) with the Gurobi solver (http://www.gurobi.com/). The flux balance analysis model of *P. tricornutum* can be found in SBML format in Table S3.

Biomass composition

The biomass composition was formulated using a variety of measurements and sources in the literature. Small differences in biomass composition that were the result of different experimental conditions were not found to significantly affect flux distributions (http://www.gurobi.com/). Table S1 gives details of the biomass composition and Table 3 describes the experimental conditions.

ACKNOWLEDGMENTS

J.K. would like to thank Kuhn Ip, James Kelley and Tony Lane (Department of Computer Science, Rutgers University, New Brunswick, NJ) and Ilse M. Remmers (Department of Biotechnology Wageningen University, The Netherlands). J.K. acknowledges financial support from the Rutgers/National Science Foundation Integrative Graduate Education and Research Traineeship Project 'Solutions for Renewable and Sustainable Fuels for the 21st Century' and National Science Foundation Division of Graduate Education O903675 to Eric Lam (Department of Plant Biology and Pathology, Rutgers University, New Brunswick, NJ).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table S1. Data and calculations for biomass components, growth rate data, abbreviations used for compartments, reactions that were included to fill gaps in the model, calculations for Table 2, comparison of results showing the correlation between measured fluxes and results from FBA versus from FAB with Euclidean norm minimization, and flux variability analysis results.

Table S2. Reactions and metabolites that make up the FBA model, together with gene names, confidence scores, EC numbers, notes on compartments, and reversibility.

Table S3. Flux balance analysis model of *P. tricornutum* in SBML format.

Methods S1. Descriptions of experimental procedures used to obtain data for the model.

REFERENCES

Allen, J.F. (2002) Photosynthesis of ATP – electrons, proton pumps, rotors, and poise. *Cell*, **110**, 273–276.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Ast, M., Gruber, A., Schmitz-Esser, S., Neuhaus, H.E., Kroth, P.G., Horn, M. and Haferkamp, I. (2009) Diatom plastids depend on nucleotide import from the cytosol. *Proc. Natl Acad. Sci. USA*, **106**, 3621–3626.

Bailleul, B., Berne, N., Murik, O. *et al.* (2015) Energetic coupling between plastids and mitochondria drives CO_2 assimilation in diatoms. *Nature*, **524**, 366–369.

Bertsekas, D.P. (2003) *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific.

Blackstone, N.W. (1995) A units-of-evolution perspective on the endosymbiotic theory of the origin of the mitochondrion. *Evolution*, **49**, 785–796.

Blank, L., Kuepfer, L. and Sauer, U. (2005) Large-scale ^{13}C-flux analysis reveals mechanistic principles network robustness to null mutations in yeast. *Genome Biol.* **6**, R49.

Bonarius, H.P.J., Vassily, H., Meesters, K.P.H., de Gooijer, C.D., Schmid, G. and Tramper, J. (1996) Metabolic flux analysis of hybridoma cells in different culture media using mass balances. *Biotechnol. Bioeng.* **50**, 299–318.

Bowler, C., Allen, A.E., Badger, J.H. *et al.* (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, **456**, 239–244.

Boyle, N.R. and Morgan, J.A. (2009) Flux balance analysis of primary metabolism in *Chlamydomonas reinhardtii*. *BMC Syst. Biol.* **3**, 4.

Brown, M.R. (1991) The amino acid and sugar composition of 16 species of microalgae used in mariculture. *J. Exp. Mar. Biol.* **145**, 79–99.

Buchanan, B.B., Kalberer, P.P. and Arnon, D.I. (1967) Ferredoxin-activated fructose diphosphatase in isolated chloroplasts. *Biochem. Biophys. Res. Commun.* **29**, 74–79.

Burkhardt, S., Amoroso, G., Riebesell, U. and Sultemeyer, D. (2001) CO_2 and HCO_3^- uptake in marine diatoms acclimated to different CO_2 concentrations. *Limnol. Oceanogr.* **46**, 1378–1391.

Caspi, R., Altman, T., Dreher, K. *et al.* (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **40**, D742–D753.

Cerón García, M.C., Sánchez Mirón, A., Fernández Sevilla, J.M., Molina Grima, E. and García Camacho, F. (2005) Mixotrophic growth of the microalga *Phaeodactylum tricornutum*: Influence of different nitrogen and organic carbon sources on productivity and biomass composition. *Proc. Biochem.*, **40**, 297–305.

Chang, R.L., Ghamsari, L., Manichaikul, A. *et al.* (2011) Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism. *Mol. Syst. Biol.* **7**, 518.

Cheirsilp, B. and Torpee, S. (2011) Enhanced growth and lipid production of microalgae under mixotrophic culture condition: effect of light intensity, glucose concentration and fed-batch cultivation. *Bioresour. Technol.* **110**, 510–516.

Chisti, Y. (2007) Biodiesel from microalgae. *Biotechnol. Adv.* **25**, 294–306.

Clarens, A.F., Resurreccion, E.P., White, M.A. and Colosi, L.M. (2010) Environmental life cycle comparison of algae to other bioenergy feedstocks. *Environ. Sci. Technol.* **44**, 1813–1819.

Claros, M.G. and Vincens, P. (1996) Computation method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* **241**, 779–786.

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005–1016.

Fabris, M., Matthijs, M., Rombauts, S., Vyverman, V., Goossens, A. and Baart, G.J.E. (2012) The metabolic blueprint of Phaeodactylum tricornutum reveals a eukaryotic Entner-Doudoroff glycolytic pathway. *Plant J.* **70**, 1004–1014.

Fernandez-Reiriz, M.J., Perez-Camacho, A., Ferreiro, M.J., Blanco, J., Planas, M., Campos, M.J. and Labarta, U. (1989) Biomass production and variation in the biochemical profile (total protein, carbohydrates, RNA, lipids, and fatty acids) of seven species of marine microalgae. *Aquaculture*, **83**, 17–37.

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

Fischer, K. (2011) The import and export business in plastids: transport processes across the inner envelope membrane. *Plant Physiol.* **155**, 1511–1519.

Frada, M.J., Burrows, E.H., Wyman, K.D. and Falkowski, P.G. (2013) Quantum requirements for growth and fatty acid biosynthesis in the marine diatom Phaeodactylum tricornutum (Bacillariophyceae) in nitrogen replete and limited conditions. *J. Phycol.*, **49**, 381–388.

Ginger, M.L., McFadden, G.I. and Michels, P.A.M. (2010) Rewiring and regulation of cross-compartmentalized metabolism in protists. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 831–845.

Graham, J.M., Graham, L.E., Zulkifly, S.B., Pfleger, B.F., Hoover, S.W. and Yoshitani, J. (2011) Freshwater diatoms as a source of lipids for biofuels. *J. Ind. Microbiol. Biotechnol.* **39**, 419–428.

Guerra, L.T., Levitan, O., Frada, M.J., Sun, J.S., Falkowski, P.G. and Dismukes, G.C. (2013) Regulatory branch points affecting protein and lipid biosynthesis in the diatom Phaeodactylum tricornutum. *Biomass Bioenerg.* **59**, 306–315.

Harbinson, J. and Foyer, C.H. (1991) Relationships between the efficiencies of photosystem I and II and stromal redox state in CO_2-free air. *Plant Physiol.* **97**, 41–49.

Hildebrand, M., Davis, A.K., Smith, S.R., Traller, J.C. and Abbriano, R. (2012) The place of diatoms in the biofuels industry. *Biofuels*, **3**, 221–240.

Hoefnagel, M.H.N., Atkin, O.K. and Wiskich, J.T. (1998) Interdependence between chloroplasts and mitochondria in the dark. *Biochim. Biophys. Acta*, **1366**, 235–255.

Hoffman, J.P. (1998) Wastewater treatment with suspended and nonsuspended algae. *J. Phycol.* **34**, 757–763.

Holmgren, A., Buchanan, B.B. and Wolosiuk, R.A. (1977) Photosynthetic regulatory protein from rabbit liver is identical with thioredoxin. *FEBS Lett.* **82**, 351–354.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M. and Darzins, A. (2008) Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant J.* **54**, 621–639.

Huang, A., Liu, L., Yang, C. and Wang, G. (2015) *Phaeodactylum tricornutum* photorespiration takes part in glycerol metabolism and is important for nitrogen-limited response. *Biotechnol. Biofuels*, **8**, 73.

Izallalen, M., Mahadevan, R., Burgard, A., Postier, B., Didonato, R., Sun, J., Schilling, C.H. and Lovley, D.R. (2008) *Geobacter sulfurreducens* strain engineered for increased rates of respiration. *Metab. Eng.* **10**, 267–275.

Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30.

Karp, P.D., Riley, M., Saier, M., Paulsen, I.T., Paley, S.M. and Pellegrini-Toole, A. (2000) The EcoCyc and MetaCyc databases. *Nucleic Acids Res.* **28**, 56–59.

Keeling, P.J. and Doolittle, W.F. (1997) Evidence that eukaryotic triosephosphate isomerase is of alpha-proteobacterial origin. *Proc. Natl Acad. Sci. USA*, **94**, 1270–1275.

Kelley, J.J., Lane, A., Li, X., Mutthoju, B., Maor, S., Egen, D. and Lun, D. (2014) MOST: a software environment for constraint-based metabolic modeling and strain design. *Bioinformatics*, **31**, 610–611.

Klanchui, A., Khannapho, C., Phodee, A., Cheevadhanarak, S. and Meechai, A. (2012) iAK692: a genome-scale metabolic model of *Spirulina platensis* C1. *BMC Syst. Biol.* **6**, 71.

Knoop, H., Grundel, M., Zilliges, Y., Lehmann, R., Hoffman, S., Lockau, W. and Steuer, R. (2013) Flux balance analysis of cyanobacterial metabolism: the metabolic network of *Synechocystis* sp. PCC 6803. *PLoS Comput. Biol.* **9**, e1003081.

Kooistra, W.H.C.F., Gersonde, R., Medlin, L.K. and Mann, D.G. (2007) The origin and evolution of the diatoms: their adaptation to a planktonic existence. In *Evolution of Primary Producers in the Sea* (Falkowski, P.G. and Knoll, A.H., eds). Waltham, MA: Academic Press Inc., pp. 210–239.

Kramer, D.M. and Evans, J.R. (2011) The importance of energy balance in improving photosynthetic productivity. *Plant Physiol.* **155**, 70–78.

Kroth, P.G., Chiovitti, A. and Gruber, A. *et al.* (2008) A model for carbohydrate metabolism in the diatom Phaeodactylum tricornutum deduced from comparative whole genome analysis. *PLoS ONE*, **3**, e1426. doi: 10.1371/journal.pone.0001426

Lebeau, T. and Robert, J.M. (2003) Diatom cultivation and biotechnologically relevant products. Part II: Current and putative products. *Appl. Microbiol. Biot.*, **60**, 624–632.

Lee, K.H., Park, J.H., Kim, T.Y., Kim, H.U. and Lee, S.Y. (2007) Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* **3**, 149.

Levitan, O., Dinamarca, J., Hochman, G. and Falkowski, P.G. (2014) Diatoms: a fossil fuel of the future. *Trends Biotechnol.* **32**, 117–124.

Liaud, M., Lichtle, C., Apt, K., Martin, W. and Cerff, R. (2000) Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. *Mol. Biol. Evol.* **17**, 213–223.

Liu, X., Duan, S., Li, A., Xu, N., Cai, Z. and Hu, Z. (2009) Effects of organic carbon sources on growth, photosynthesis, and respiration of *Phaeodactylum tricornutum*. *J. Appl. Phycol.* **21**, 239–246.

Mahadevan, R. and Schilling, C.H. (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **5**, 264–276.

Marsot, P., Cembella, A.D. and Colombo, J.C. (1991) Intracellular and extracellular amino acid pools of the marine diatom *Phaeodactylum tricornutum* (Bacillariophyceae) grown on unenriched seawater in high-cell-density dialysis culture. *J. Phycol.* **27**, 478–491.

Martin, W., Brinkmann, H., Savona, C. and Cerff, R. (1993) Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3–phosphate dehydrogenase genes. *Proc. Natl Acad. Sci. USA*, **90**, 8692–8696.

Martin, W. and Müller, M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature*, **392**, 37–41.

Matsuo, M. and Obokata, J. (2006) Remote control of photosynthetic genes by the mitochondrial respiratory chain. *Plant J.* **47**, 873–882.

Mehta, S.K. and Gaur, J.P. (2005) Use of algae for removing heavy metal ions from wastewater: progress and prospects. *Crit. Rev. Biotechnol.* **25**, 113–152.

Michels, A.K., Wedel, N. and Kroth, P.G. (2005) Diatom plastids possess a phosphoribulokinase with an altered regulation and no oxidative pentose phosphate pathway. *Plant Physiol.* **137**, 911–920.

Muthuraj, M., Palabhanvi, B., Misra, S., Kumar, V., Sivalingavasu, K. and Das, D. (2013) Flux balance analysis of *Chlorella* sp. FC2 IITG under photoautotrophic and heterotrophic growth conditions. *Photosynth. Res.* **118**, 167–179.

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

Pahl, S.L., Lewis, D.M., Chen, F. and King, K.D. (2010) Growth dynamics and the proximate biochemical composition and fatty acid profile of the heterotrophically grown diatom *Cyclotella cryptica*. *J. Appl. Phycol.* **22**, 165–171.

Pharkya, P. and Maranas, C.D. (2006) An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metab. Eng.* **8**, 1–13.

Pharkya, P., Burgard, A.P. and Maranas, C.D. (2003) Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol. Bioeng.* **84**, 887–899.

Pharkya, P., Burgard, A.P. and Maranas, C.D. (2004) Optstrain: a computational framework for redesign of microbial production systems. *Genome Res.* **14**, 2367–2376.

Poolman, M.G., Miguet, L., Sweetlove, L.J. and Fell, D.A. (2009) A genome-scale metabolic model of *Arabidopsis* and some of its properties. *Plant Physiol.* **151**, 1570–1581.

Prigent, S., Collet, G., Dittami, S.M. *et al.* (2014) The genome-scale metabolic network of *Ectocarpus siliculosus* (EctoGEM): a resource to study brown algal physiology and beyond. *Plant J.* **80**, 367–381.

Reed, J.L., Vo, T.D., Schilling, C.H. and Palsson, B.O. (2003) An expanded genome-scale model of *Escherichia coli* K–12 (*i*JR904 GSM/GPR). *Genome Biol.* **4**, R54.

Reinfelder, J.R., Kraepiel, A.M. and Morel, F.M.M. (2000) Unicellular C4 photosynthesis in a marine diatom. *Nature*, **407**, 996–999.

Schomburg, I., Chang, A., Placzek, S. *et al.* (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic Acids Res.* **41**, D764–D772.

Schuetz, R., Kuepfer, L. and Sauer, U. (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3**, 119.

Shastri, A.A. and Morgan, J.A. (2005) Flux balance analysis of photoautotrophic metabolism. *Biotechnol. Prog.* **21**, 1617–1626.

Smith, S.R., Abbriano, R.M. and Hildebrand, M. (2012) Comparative analysis of diatom genomes reveals substantial difference in the organization of carbon partitioning pathways. *Algal Res.* **1**, 2–16.

Sweetlove, L.J., Beard, K.F.M., Nunes-Nesi, A., Fernie, A.R. and Ratcliffe, R.G. (2010) Not just a circle: flux modes in the plant TCA cycle. *Trends Plant Sci.* **15**, 462–470.

Thiele, I. and Palsson, B. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121.

Wang, H., Fu, R. and Guofeng, P. (2012) A study on lipid production of the mixotrophic microalgae *Phaeodactylum tricornutum* on various carbon sources. *Afr J Microbiol Res.* **6**, 1041–1047.

Weber, A.P.M., Linka, M. and Bhattacharya, D. (2006) Ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot. Cell*, **5**, 609–612.

Whatley, J.M., John, P. and Whatley, F.R. (1979) From extracellular to intracellular: the establishment of mitochondria and chloroplasts. *Proc. R. Soc. Lond. B Biol. Sci.* **204**, 165–187.

Wittpoth, C., Kroth, P.G., Weyrauch, K., Kowallik, K.V. and Strotmann, H. (1998) Functional characterization of isolated plastids from two marine diatoms. *Planta*, **206**, 79–85.

Wolosiuk, R.A. and Buchanan, B.B. (1977) Thioredoxin and glutathione regulate photosynthesis in chloroplasts. *Nature*, **266**, 565–567.

Yongmanitchai, W. and Ward, O.P. (1992) Separation of lipid classes from *Phaeodactylum tricornutum* using silica cartridges. *Phytochemistry*, **31**, 3405–3408.

Zaslavskaia, L.A., Lippmeier, J.C., Shih, C., Ehrhardt, D., Grossman, A.R. and Apt, K.E. (2001) Trophic conversion of an obligate photoautotrophic organism through metabolic engineering. *Science*, **292**, 2073–2075.

Zelle, R.M., de Hulster, E., van Winden, W.A., de Waard, P., Dijkema, C., Winkler, A.A., Geertman, J.M., van Dijken, J.P., Pronk, J.T. and van Maris, A.J. (2008) Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. *Appl. Environ. Microbiol.* **74**, 2766–2777.

Zheng, Y., Quinn, A.H. and Sriram, G. (2013) Experimental evidence and isotopomer analysis of mixotrophic glucose metabolism in the marine diatom *Phaeodactylum tricornutum*. *Microb. Cell Fact.* **12**, 109.