Diel gene regulation in *P. tricornutum*

**Running title:** Diel gene regulation in *P. tricornutum*

**Corresponding authors:**

Atle M. Bones
Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway
Tel: +47-74598692; Fax: +47-73596100; email: atle.bones@bio.ntnu.no

Olav Vadstein
Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway
Tel: +47-73594058; Fax: +47-73591283; email: olav.vadstein@biotech.ntnu.no

**Research category:**
System biology, molecular biology and gene regulation
Gene regulation of carbon fixation, storage and utilization in the diatom *Phaeodactylum tricornutum* acclimated to light/dark cycles

Matilde Skogen Chauton\(^1\), Per Winge\(^2\), Tore Brembu\(^2\), Olav Vadstein\(^1\), Atle M. Bones\(^2\)

\(^1\)Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

\(^2\)Cell-, Molecular biology and Genomics Group, Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

\(^\dagger\) Equal contribution
Abstract

Regulation of carbon metabolism in the diatom *Phaeodactylum tricornutum* at cell, metabolite and gene expression level in exponential fed-batch cultures is reported. Transcriptional profiles and cell chemistry sampled simultaneously at all time-points provide a comprehensive data set on carbon incorporation, fate and regulation. An increase in Nile Red fluorescence (a proxy for cellular neutral lipids) was observed throughout the light period, and water-soluble glucans increased rapidly in the light period. A near-linear decline in both glucans and lipids was observed during the dark period, and transcription profile data indicated that this decline was associated with onset of the mitosis. More than 4,500 transcripts that were differentially regulated during the light/dark cycle are identified, many of which were associated with carbohydrate and lipid metabolism. Genes not previously described in algae and their regulation in response to light was integrated in this analysis together with proposed roles in metabolic processes. Some very fast light responding genes in e.g. fatty acid biosynthesis were identified and allocated to biosynthetic process. Transcripts and cell chemistry data reflect the link between light energy availability and light energy consuming metabolic processes. Our data confirm spatial localization of processes in carbon metabolism to e.g. either plastids or mitochondria, or glycolysis/gluconeogenesis which are localized to the cytosol, chloroplast and to mitochondria. Localization and diel expression pattern may be of help to determine the roles of different isoenzymes, and mining of genes involved in light responses and circadian rhythms.

Keywords: Microarray, Calvin cycle, glycolysis, TCA cycle, fatty acid metabolism, condensin complex
Diel gene regulation in *P. tricornutum*

**Introduction**

The ecological success of diatoms has drawn attention from phycologists for a long time, and their usefulness in many areas (e.g. as feed organisms in aquaculture or raw material producers in bioenergy) has made this group of microalgae known to a broader audience. The recent completion of the genome sequence of *Phaeodactylum tricornutum* (Bowler et al., 2008) has provided a useful model organism for this group, and knowledge of its physiology can now be extended to the genetic and metabolic level, resulting in a deeper understanding of some of the factors that regulate major cell processes in diatoms (Parker et al., 2008; Fernie et al., 2012). The transcriptome of *P. tricornutum* has been studied in a few contexts, e.g. silicon metabolism (Sapriel et al., 2009) and light acclimation (Nymark et al., 2009).

Carbon metabolism in autotrophic algae starts with photosynthesis and carbon acquisition, followed by incorporation of carbon into different organic metabolites. In diatoms, chrysolaminaran (1,3-β-D-glucan) is one of the main sinks for carbon fixed during light periods, and it is also incorporated into glucans through gluconeogenesis (Myklestad and Granum, 2009). In darkness glucans provide energy for nutrient assimilation and carbon skeletons for synthesis of other biomolecules through glycolysis and the tricarboxylic acid (TCA) cycle. Triosephosphate exported from the chloroplast might serve as a precursor for hexose sugars (glucose and fructose) or might be transformed into pyruvate and enter the TCA cycle in the mitochondrion. Metabolism of storage carbohydrates is a dynamic process, and water-soluble glucans are located in vacuoles (Granum and Myklestad, 2002; Kroth et al., 2008). Cells use monomeric glucose in various metabolic pathways, and polymeric 1,3-β-D-glucans therefore have to be broken down into monomers. Synthesis of polymers (that must be reduced to monomers again) is considered a way to reduce the large oscillations in
Diel gene regulation in *P. tricornutum*

cellular glucose levels that would otherwise occur over the light/dark cycles (Vårum et al., 1986).

Another fraction of the fixed carbon is incorporated via acetyl-Coenzyme A (acetyl-CoA) into fatty acids (FAs). Short FAs with few double bonds are combined with glycerol on the endoplasmic reticulum (ER) and stored as neutral triacylglycerols (TAGs) in lipid droplets in the cytosol or in chloroplasts (Fan et al., 2011). Other fatty acids are elongated and desaturated, and combined with phosphorus or carbohydrates (glycans) into more complex phospho- or glycolipids to provide polar membrane material. Fatty acids store more energy in smaller molecules, but the processes of lipid synthesis and catabolism are slower compared to carbohydrate metabolism (Lancelot and Mathot, 1985). It has been shown that other factors such as nutrient depletion (Reitan et al., 1994; Breteler et al., 2005) or excess light energy (Norici et al., 2011) may increase cellular lipid content in diatoms. Along with mapping and identification of gene coding regions of microalgae genomes, genetic engineering is a rapidly developing field that may provide a tool to enhance e.g. lipid synthesis and storage in algae for use in biofuel production (Radakovits et al., 2010).

Cell metabolism occurs in cyclic patterns, and some cycles are cued by external geophysical zeitgebers such as light/dark periods (e.g. photosynthetic carbon reduction) while other cycles (e.g. the TCA cycle) are cued by internal biological clocks. The internal molecular clocks are related to gene transcription and translation regulated in feedback loops (Tessmar-Raible et al., 2011), and circadian control of many cell processes has been studied in the green alga *Chlamydomonas reinhardtii* (Schulze et al., 2010). Control points can be at translational level (Mittag, 2003) or post-translational level (Zhang et al., 2011), and sometimes regulated as a feedback signal system. As an example, oscillations of sugar content contributed to sugar-
Diel gene regulation in *P. tricornutum*

responsive transcripts in plant cells (Haydon et al., 2011). In general, major cellular processes in microalgae, such as photosynthesis, nutrient uptake and incorporation, cell growth and division, show diel regulation that follow the light/dark periods (Sanchez et al., 2009, Radakovits et al. 2010). However, this general pattern may be altered in species with high growth rate, where some cells undergo more than one cell division per day (Granum et al., 2009). Several studies have investigated diel regulation of the transcriptome in Cyanobacteria, and 25-90% of the expressed genes were found to exhibit oscillating expression profiles (Stöckel et al., 2008; Zinser et al., 2009; Shi et al., 2010; Straub et al., 2011). Biochemical pathways involved in carbon metabolism appeared to be tightly regulated, with processes such as photosynthesis and carbon assimilation dominating during the day, and catabolic utilization of energy reserves occurring mainly during the night (Dron et al., 2012). Similar analyses of eukaryotic plankton are rare, but about 94% of the expressed genes in the prasinophyte *Ostreococcus tauri* were differentially expressed during a light/dark cycle (Monnier et al., 2010).

In this work, we describe the transcriptional regulation of glucan and lipid metabolism in *P. tricornutum* acclimated to light/dark cycles and surplus nutrient supply, based on a time series of cell chemistry measurements and simultaneous data from microarray analysis. The fate of the incorporated carbon was followed through gluconeogenesis or glycolysis, and from acetyl-CoA into the TCA cycle or to fatty acid synthesis and elongation. Analysis of data from the transcriptional level was integrated with cell chemistry measurements, focusing on the uptake and metabolism of major carbon pools in the cells. Our results show that genes encoding enzymes involved in processes such as the TCA cycle and fatty acid biosynthesis are highly coordinated over the light/dark cycle, and the transcriptional data indicate activity that may
Diel gene regulation in *P. tricornutum*

ultimately result in the synthesis of carbon skeletons in the chloroplast during
daytime, and breakdown of carbon-rich compounds in the cytosol and mitochondria
and cell division during night-time.
Results and discussion

Two samples of 50 ml were removed from each culture at selected time points over a course of 26.5 hours; one sample was used for cell chemistry measurements and the other sample was subjected to RNA extraction. The sampling points were (hours after light on): T1=0.5, T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), T7=23.5 (darkness), and (the next day) T8=27. The chronologically last sample (T8) corresponds to a time point between T1 and T2 when considering the light/dark cycle (discussed below). In the analyses of gene expression the latest point in darkness (T7) was used as reference.

Cell chemistry

Four cell specific chemical variables are show in Figure 1. Our predictions were that when cultures are in steady state sampling is independent of chronology and only dependent on the point of time with respect to the light/dark cycle. As seen from our data, the T8 measurements fit well between T1 and T2, and thus support steady state. Steady state was also confirmed by stable biomass during a period of at least 5 days before sampling, and this phenomenon is well known (Droop 1974). As a consequence T8 is plotted in accordance with its position in the light/dark cycle and not chronology, in the remaining figures.

Total water-soluble carbohydrates in the cells increased from 5 pg cell\(^{-1}\) to 11.5 pg cell\(^{-1}\) (average of two cultures) during the light period, and the rate of increase was very high towards the end of the light period (Figure 1A). The same pattern has been observed in other diatoms (Vårum et al., 1986; Granum et al., 2009). A few hours into the light period, cellular carbon increased rapidly from 8 to 12.4 pg cell\(^{-1}\) and thus increase approximately 50% during the light period (Figure 1B). Cellular nitrogen
Diel gene regulation in *P. tricornutum*

increased from 1.5 to 2 pg cell\(^{-1}\) during the first hours of the light period, and was stable until it decreased during the dark period (Figure 1B). The observed increase in both cellular carbon and nitrogen (per cell) during the light was followed by a decrease in the dark, and there was no accumulation of N or C in the cells over a light/dark cycle. Hockin and co-workers (2012) studied the proteome and transcriptome of *T. pseudonana* and found that the response of carbon metabolism to nitrogen starvation in this diatom is more similar to cyanobacteria than other photosynthetic eukaryotes. Cell numbers were stable during the light period, on average 7.41x10\(^5\) ml\(^{-1}\) (±2.22%, average of 4 measurements), and increased steadily during the dark period. On the last sampling (T8, 27 hours after the start) the cell numbers were 1.02x10\(^6\), which is an increase of 38%. These numbers indicate that cell division took place during the dark and a few hours into the next light period, and bio-dilution may partially explain the observed decrease in N per cell in the dark.

Optical density, on the other hand, increased 35% during the light period and decreased rapidly in the dark (data not shown). The increase in optical density during the day may reflect cell volume increment, rather than an increase in cell numbers.

Nile red fluorescence (NRF) was used as a proxy for neutral lipids (Greenspan and Fowler, 1985; Chen et al., 2009). The fluorescence was stable and low from the middle of the dark period to the first three hours into the light period. Over the next hours the signal intensity nearly doubled and remained so until the beginning of the dark period when it decreased to a level close to level at the beginning of the light period (Figure 1A). This observation is in accordance with previous studies in the centric diatom *Cyclotella meneghiniana* (Sicko-Goad et al., 1988) and the eustigmatophyte *Nannochloropsis sp.* (Sukenik and Carmeli, 1990). Nile red is also used to stain proteins (Demeule et al., 2007) and although excitation/emission
Diel gene regulation in *P. tricornutum*

Wavelengths differ from those for lipid analyses, it may be a source of interference with the lipid measurements. Larsen et al. (2011) quantified lipids and proteins in a biofilm using Nile red and other fluorochromes, and found that oleic acid fluoresced almost twice as much as palmitic and stearic acid when the lipids where precipitated in water. *P. tricornutum* is rich in palmitic and palmitoleic acids (Domergue et al., 2003) and Nile red as a proxy for neutral lipids may be biased due to the specific fatty acid composition. Gravimetric measurements of total neutral lipids (Bligh and Dyer, 1959) sampled 27 h after start of the experimental period showed that *P. tricornutum* contained 10.1(±3.5)% (of dry weight) neutral lipids.

**Gene expression**

Global transcriptome analyses using whole-genome oligonucleotide microarrays for *P. tricornutum* showed that 4,567 genes were significantly regulated (p>0.05) during at least one of the eight time points in the study, equivalent to about 44% of the whole transcriptome. We generated a coexpression network based on Pearson correlation values between the gene pairs of the 1,936 most responsive genes in the experiment (Figure 2). Six clusters of coregulated genes were identified; together, the clusters formed a circle similar to the coexpression network observed in a diurnal experiment performed on the cyanobacterium *Cyanothece sp.* (Stöckel et al., 2008).

Cluster 1 mainly contains genes involved in ribosome biogenesis and processing of pre-rRNAs. A hallmark of this cluster is that the genes are down-regulated around midday. Cluster 2 contains genes with highest expression late in the day and during night, and includes genes associated with many cellular processes, among them cell division and the TCA-cycle. Cluster 3 is composed mainly of genes that are down-regulated after transition from dark to light and includes genes involved in nitrogen
Diel gene regulation in *P. tricornutum*

metabolism (UreG, urease and cyanate hydratase), as well as genes central in the glyoxylate cycle (malate synthase and isocitrate lyase). Cluster 3 also includes genes coupled to mitochondrial β-oxidation such as beta-ketoacyl-CoA thiolase (KCT3) and multifunctional fatty acid oxidation complex subunit alpha (Phatr2_35240). Cluster 4 includes genes coupled to various processes with maximum transcriptional activity during midday. Notably, many of them encode heat shock proteins, chaperones, and proteins involved in protein modification/degradation. Cluster 5 includes genes that are down-regulated during night-time, and encode proteins involved in many cellular processes such as carotenoid, chlorophyll and fatty acid biosynthesis. Cluster 6 is mainly composed of genes transcribed from the chloroplast as well as nuclear-encoded ribosomal proteins.

Carbon fixation

Glycerate 3-phosphate is one of the primary products of CO₂ assimilation and is phosphorylated to 1,3-bisphosphoglycerate by phosphoglycerate kinase. Four central enzymes in carbon fixation were highly co-regulated and showed highest expression early in the light period and reduced activity in the dark period: phosphoribulokinase (Prk), phosphoglycerate kinase (Phatr2_29157), glyceraldehyde-3-phosphate dehydrogenase (GapC4) and the recently described Stramenopile protein triosephosphate isomerase/glyceraldehyde-3-phosphate dehydrogenase (TPI/GapC3) (Figure 3). The Prk converts D-ribulose 5-phosphate to D-ribulose 1,5-bisphosphate, which acts as the acceptor for CO₂ during the production of glycerate 3-phosphate by ribulose bisphosphate carboxylase/oxygenase (RubisCo). 3-phospho-D-glycerate in turn is phosphorylated to 1,3-diphosphoglycerate by phosphoglycerate kinase (Phatr2_29157). The Stramenopile-specific fusion protein TPI/GapC3 might be one
Diel gene regulation in *P. tricornutum*

of the proteins responsible for the reduction of 1,3-diphosphoglycerate to glyceraldehyde 3-phosphate (catalysed by the GAPDH domain) as well as the isomerization of glyceraldehyde 3-phosphate to dihydroxyacetone phosphate (catalysed by the N-terminal triosephosphate isomerase domain) (Liaud et al., 2000). TPI/GapC3 is predicted to encode a chloroplast-localized protein, WoLF PSORT (Horton et al., 2007); however, previous work by Liaud and co-workers suggested that this protein was localized to the mitochondria. Triosephosphate is transported out of the chloroplast by triosephosphate transporters. There are 3 triosephosphate transporters in *P. tricornutum*; of these the plastidic Tpt3, which is related to *Arabidopsis thaliana* TPT/APE2, is most light responsive. The triosephosphate exported from the chloroplast may serve as a precursor for hexose sugars (glucose and fructose) or may be transformed into pyruvate and enter the TCA cycle in the mitochondrion (glycolysis). D-ribulose 5-phosphate is regenerated through the sequential action of several enzymes of the Calvin cycle, including fructose bisphosphatase (FBP, FBPC1), transketolase (TkI) and the plastidic ribulose-phosphate 3-epimerase (Rpe). The expression pattern of these genes was closely related to the other carbon fixation enzymes, although not induced to the same degree by light (Figure 3). The sedoheptulose bisphosphatase (SBPase) was not significantly regulated at any of the time points.

Diatoms possess organic carbon concentrating mechanisms (CCMs) that provide a higher CO₂ concentration around Rubisco than that in the surrounding medium (Roberts et al., 2007); however, the details of this process are still unknown. Genes that may be coupled to CCMs in diatoms, such as the plastidic HCO₃⁻-transporters SLC4A_1 and SLC4A_2 and the carbonic anhydrases PtCa1 and PtCa2, were strongly down-regulated at dark (Figure 3). Interestingly, four genes previously linked
Diel gene regulation in *P. tricornutum*

to a proposed C4-pathway in diatoms (Reinfelder, 2000) appear to have higher transcriptional activity late in the day and during night. Phosphoenolpyruvate carboxylases PEPCase_1 and PEPCase_2, which convert phosphoenolpyruvate and HCO$_3^-$ to oxaloacetate, show a similar expression pattern (Supplementary Table S1). The same is also observed for phosphoenolpyruvate carboxykinase (PEPCK1) which decarboxylate oxaloacetate, and the NADP-dependent malic enzyme (Phatr2_51970) which decarboxylate malate. The mitochondrial pyruvate carboxylase (PYC1), which carboxylate pyruvate to oxaloacetate, is closely co-regulated with PEPCK1, while the plastid localized PYC2 shows lowest transcriptional activity late at night. The transcriptional activity of pyruvate-phosphate dikinase (PPdK) is also highest late at night. However, regulation of the alleged C4 enzymes may be an indirect effect of aerating the culture with 1-2% CO$_2$ (v/v) and could be connected to regulation of pH homeostasis at least during the night (Haimovich-Dayan et al., 2012).

Glucan biosynthesis and glycolysis

One of the key enzymes in gluconeogenesis is fructose-1,6-bisphosphate aldolase (Fba), which reversibly convert D-glyceraldehyde 3-phosphate into D-fructose 1,6-bisphosphate. Evolutionary origin and functional diversification of the various Fba found in diatoms is described by Allen et al. (2012). Fba3 is a class II gene encoding a cytosolic enzyme; minimum expression level was observed late at night and it showed a strong light dependency with about 16 times higher expression during the light period (Figure 4). This suggests that Fba3 has an important role in biosynthesis of hexose sugars and storage polysaccharides. Blue light induction of Fba3 has been reported previously (Tachibana et al., 2011). The other class II gene (FbaC2, with a putative plastid localization) showed less diel regulation.
Diel gene regulation in *P. tricornutum*

The first irreversible step of hexose sugar synthesis is the production of fructose 6-phosphate catalyzed by the cytosolic fructose 1,6-bisphosphatase (FBP). There are five fructose 1,6-bisphosphatase enzymes in *P. tricornutum*, but only two of the enzymes, FBP and FBPC1, are light-induced and show clear diel expression patterns (Figure 3). FBPC1 most likely encodes a plastidic form, while FBP is a cytosolic enzyme and the best candidate gene responsible for sugar biosynthesis (Kroth et al., 2008). Fructose 6-phosphate is converted to glucose 6-phosphate by glucose-6-phosphate isomerase (GPI), and the two GPIs observed here (GPI_1 and GPI_3) showed opposite regulation (Figure 4). GPI_3, which carries a chloroplast targeting signal, is closely co-regulated with enzymes of the Calvin cycle and is most likely coupled to gluconeogenesis. GPI_1 may be a key enzyme of glycolysis; it has a predicted cytosolic localization and showed highest transcriptional activity at night.

Conversion of glucose 6-phosphate to glucose 1-phosphate is catalyzed by phosphoglucomutase (PGM) (Figure 4). There are five PGM enzymes in *P. tricornutum*, but only two of these genes, *PGM_1* (eukaryotic type) and *Phatr2_48819* (cyanobacterial type), show clear diel expression patterns. Both proteins have chloroplast targeting signals. PGM_1 is related to *Arabidopsis thaliana* PGM, which is essential for starch biosynthesis, and display a gene expression pattern related to carbon fixation/Calvin cycle genes, suggesting that it has functions connected to gluconeogenesis. In contrast, the cyanobacterial-type PGM is strongly down-regulated by light, and has a low expression during the first 6 hours of the day with maximum transcriptional activity in the evening and during the night (Figure 4).

Recently the presence of a functional Entner-Doudoroff pathway (EDP), which converts glucose-6-phosphate to pyruvate and glyceraldehyde-3-phosphate, was reported in *P. tricornutum* (Fabris et al., 2012). The putative key enzyme of this
Diel gene regulation in *P. tricornutum*

pathway, 2-keto-3-deoxyphosphogluconate aldolase (EDA; Phatr2_34120), peaked at the beginning of the light period and reached the expression minimum at the beginning of the dark period (Supplementary Table S1). Both enzymes of the *P. tricornutum* EDP have a putative mitochondrial localization; however, the EDA expression pattern is opposite of most of the genes encoding mitochondria- and cytosol-localized glycolytic enzymes. The role of the EDP may be to feed pyruvate to the TCA cycle early in the light period, when pyruvate production through the “classic” glycolytic pathway is low.

Several of the genes encoding putatively plastid- or ER-localized enzymes in the gluconeogenesis/1,3-β-glucan pathway, such as *GPI*₃, *PGM1*, *Phatr2_23639*, *BGS1* and a putative 1,6-β-branching enzyme *Phatr2_50238* (Fabris et al., 2012) showed a coordinated diel expression pattern, with a peak at the beginning and a minimum at the end of the light period (Figure 4). In contrast, both endo-1,3-beta-glucosidases that were differentially expressed in the experiment (Phatr2_54681 and Phatr2_54973) reached maximum (yet low) expression during the dark period. Both the endo-1,3-beta-glucanases and the exo-1,3-beta-glucosidases identified on the *P. tricornutum* genome have a putative localization to ER/Golgi or vacuoles (Kroth et al., 2008). Based on the spatial and temporal separation of the different components, a possible pattern of glucan biosynthesis and metabolism emerges. During daytime, carbon fixed through photosynthesis is shuttled through the gluconeogenesis pathway in chloroplasts to produce the nucleotide sugar UDP-glucose. UDP-glucose is made available to BGS1 through ER connected to the plastid (CER), leading to synthesis of chrysolaminaran that may be transported through ER/Golgi to storage vacuoles. At night, expression of biosynthetic enzymes decreases and expression of endo-1,3-beta-glucanases and exo-1,3-beta-glucosidases increases, shifting the equilibrium towards
Diel gene regulation in *P. tricornutum*

degradation of chrysolaminaran to glucose. *P. tricornutum* encodes one glucokinase (Phatr2_15495) which may be cytosolic or localized to the vacuole (Figure 4). A glucose transporter (Phatr2_12520) with a predicted vacuolar localization (WoLF PSORT) showed highest expression during the dark period (Supplementary Table S1), and may act to shuttle glucose to the cytosol. A cytosolic glucokinase could convert glucose to glucose 6-phosphate, which may enter the glycolytic pathway via the cytosolic glucose-6-phosphate isomerase (GPI_1) that was, in fact, upregulated during the dark period.

**TCA cycle**

The TCA cycle is located to the mitochondria, although some of the involved isoenzymes have a putative cytosolic localization (Kroth et al., 2008; Millar et al., 2011). Pyruvate produced through glycolysis enters the TCA cycle in two ways: either through conversion to acetyl-CoA by the mitochondrial pyruvate dehydrogenase complex or through conversion to oxaloacetate by pyruvate carboxylase (Figure 5). The genes *PDHA1, PDHB1, DHLTA_1, Phatr2_38009,* and *DLDH* encode components of the mitochondrial pyruvate dehydrogenase complex, and showed low and relatively stable expression throughout the diel cycle. In contrast, the majority of the genes encoding TCA cycle enzymes displayed a coordinated, cyclic expression pattern. The expression levels of these genes increased throughout the light period, reaching maximum levels around the light/dark transition, and decreased in the dark period. Thus, the TCA cycle is upregulated at the onset of night in order to provide the cells with ATP, carbon skeletons and other metabolites. A major sink for the carbon skeletons is in N assimilation, which occurs mainly in the
Diel gene regulation in *P. tricornutum*

light. Similar regulation of the TCA pathway genes has been observed in Cyanobacteria (Stöckel et al., 2008; Shi et al., 2010; Straub et al., 2011).

Fatty acid biosynthesis

*De novo* biosynthesis of fatty acids (FAs) occurs in the plastids of all plants and algae. Plastidial acetyl-CoA acts as a precursor for fatty acid biosynthesis, and is provided from different sources: a) from the conversion of pyruvate by the pyruvate dehydrogenase complex, which is present both in mitochondria and chloroplasts, and b) from the conversion of acetate by acetyl-CoA synthetase. *Nannochloropsis* sp. incorporated acetate into lipids during the light period; in the dark acetate was incorporated into non-lipid components at a lower rate (Sukenik and Carmeli, 1990).

Recently, a plastidial sodium-dependent pyruvate transporter was reported in plants (Furumoto et al., 2011). Both *P. tricornutum* (Phatr2_34802) and *T. pseudonana* (Thaps3_711) encode an orthologue with high similarity to the plant pyruvate transporters from *Flaveria trinervia* and *Arabidopsis thaliana*. Phatr2_34802 contains 35% identical and 48% conserved residues compared with the *A. thaliana* pyruvate transporter BASS2 (At2g26900). A phylogenetic analysis based on a peptide alignment of the plant pyruvate transporters and related transporters showed that Phatr2_34802 and Thaps3_711 clusters together with *A. thaliana* and *F. trinervia* BASS2 transporters with high significance (Supplemental Figure S1A). Phatr2_34802 contains a putative bipartite N-terminal presequence indicative of a plastid localization. *P. tricornutum* also encodes an acetyl-CoA transporter (Phatr2_54049) with unclear localization; however, the TargetP localization server (Emanuelsson et al., 2007) predicts a chloroplast transit peptide. All three genes encoding the plastidic pyruvate dehydrogenase complex in *P. tricornutum* showed a similar expression...
Diel gene regulation in *P. tricornutum*

profile throughout the diel cycle (Supplemental Figure S1B). The genes were unregulated or moderately upregulated at the beginning of the light period, but expression levels decreased throughout the day and reached a minimum around the light-dark transition.

A similar expression profile was observed for the majority of the genes encoding enzymes involved in FA biosynthesis (Figure 6). Of the nine FA biosynthesis genes that were regulated in the experiment, eight reached maximum expression levels after the onset of light and minimum expression levels towards the end of the light period. *FABFb*, encoding one of three 3-oxoacyl-[acyl-carrier-protein] synthases catalyzing the condensation reaction in fatty acid synthesis in *P. tricornutum*, showed a divergent expression pattern. It peaked twice during the diel cycle: after the onset and at the end of the light period. *FABFb* expression reached the lowest levels during the dark period. The two other 3-oxoacyl-[acyl-carrier-protein] synthase genes (*FABB* and *FABFa*) showed the “consensus” expression pattern of FA biosynthesis genes; furthermore, their maximum expression levels were 10 to 20 times higher than that of *FABFb* (Supplemental Table S1). Several genes encoding desaturases involved in biosynthesis of unsaturated fatty acids (*PtFAD2, PtFAD6, PTD5b* and *PTD6*) had a similar expression pattern, with maximum expression levels after the onset of light (Supplemental Table S1). These genes were among the most responsive to the dark-light transition. We were not able to identify any clear candidate gene product responsible for the final hydrolysis generating the fatty acid. Daylight appears to be anticipated, as expression increased throughout the dark period. The spike in expression at the beginning of the light period was still pronounced for most of the fatty acid biosynthesis genes. It could act as a response to a sudden flux/increase in carbon fixed by photosynthesis as the light is turned on. However, no increase in
neutral lipid levels was observed until three hours after onset of light. One explanation for this apparent lag could be that increased lipid production during this period is channeled into biosynthesis of polar lipids or free fatty acids, which are not detected by the Nile Red assay applied here.

Fatty acid β-oxidation

Fatty acid degradation in plants takes place in the peroxisomes via the β-oxidation cycle. In diatoms, however, there appear to be two fatty acid degradation pathways, similar to mammalian cells, where one pathway is located in the peroxisomes and the other is located in the mitochondria. Two key enzymes involved in mitochondrial β-oxidation, the multifunctional fatty acid oxidation complex subunit alpha (Phatr2_35240) and β-ketoacyl-CoA thiolase (KCT3, also known as the mitochondrial trifunctional protein β-subunit), were both down-regulated during late dark period/beginning of light period and showed highest expression late in the light period and in the dark period (Figure 6). Five other genes predicted to encode enzymes involved in mitochondrial beta-oxidation show a similar diel regulation pattern: Acyl-CoA dehydrogenase (Phatr2_11014), catalyzing the first step in beta-oxidation; enyol-CoA hydratase small chain (Phatr2_55192), catalyzing the second step in mitochondrial fatty acid beta-oxidation; 3-ketoacyl-CoA thiolase (Phatr2_45947); the short chain acyl-coenzyme A dehydrogenase (ACD1), and 3-hydroxyacyl-coenzyme A dehydrogenase (HAD2). The observed expression pattern suggests that the mitochondrial beta-oxidation pathway provide acetyl-CoA that can be used in the TCA cycle during the night. Metabolic energy and carbon can be stored as fat that is later used for synthesis of glucose and other carbohydrates via the glyoxylate cycle. Malate synthase (MS) and isocitrate lyase (Phatr2_14401) are two
Diel gene regulation in *P. tricornutum*

Central enzymes in the glyoxylate cycle, and enable the synthesis of hexoses using acetyl CoA as an intermediate. In addition to its role as an intermediate in the gluconeogenesis, Acetyl CoA is also a substrate in the glyoxylate cycle. Both genes were strongly down-regulated when the light was turned on (Supplemental Table S1). Phosphoenolpyruvate carboxykinase (PEPCK1), which catalyzes the final step in the production of phosphoenolpyruvate used for gluconeogenesis, showed a similar expression pattern.

The 3,5-Δ-2,4-dienoyl-CoA isomerase (*DEII*) is predicted to encode a peroxisome-localized enzyme involved in FA β-oxidation. *DEII* showed increased transcriptional activity during daytime and decreased activity late in the evening and during the night. None of the other enzymes predicted to be connected to the peroxisomal β-oxidation pathway, such as acyl-coenzyme A oxidase (AOX1), acyl-coenzyme A dehydrogenase (Phatr2_42907) and peroxisomal 2,4-dienoyl-CoA reductase (DER1), showed increased transcriptional activity late in the evening or at night.

### Cell division

The goal of the previously described metabolic processes is to enable cell division, and more than 100 transcripts encoding proteins that are involved in DNA replication, cell cycle regulation and cell division showed a high degree of co-regulation during the time course of the experiment. Genes encoding proteins with key functions in mitosis had the lowest expression early in the day and increases to a maximum late in the evening and early night (Figure 7), showing that cell division is more frequent late in the evening and during night. These observations are in correspondence with previous results (Nelson and Brand, 1979; Vaulot et al., 1986; Bowler et al., 2010). Chromosomes undergo a major structural reorganization during mitosis, and the
Diel gene regulation in *P. tricornutum*

Compaction of interphase chromatin into mitotic chromosomes is regulated by a multisubunit ATPase complex, the condensin complex. The condensation of chromosomes is a continuous process that is initiated in early prophase and is required for accurate chromosome segregation in the anaphase (Swedlow and Hirano, 2003; Moser and Swedlow, 2011). Another protein with important functions during chromosome condensation is topoisomerase II (Uemura et al., 1987). The *P. tricornutum* orthologue of topoisomerase II (Phatr2_52174) and all components of the condensin complex I/II showed a highly co-regulated gene expression, with highest expression levels observed late in the evening and during early night. In addition, we observed that genes encoding proteins coupled to the spindle checkpoint are tightly co-regulated with the condensin complex genes. Aurora kinases are reported to phosphorylate condensin subunits during mitosis (Nakazawa et al., 2011). Both *P. tricornutum* Aurora-A, the INCENP orthologue (Phatr_45765), Borealin/CDCA8 related protein (Phatr2_46380) and CDC20 orthologue (Phatr2_12783) showed the same diel pattern of regulation as the condensin complex genes (Figure 7). These genes might serve as useful markers for cells undergoing mitosis.

**Conclusion**

This time series study of carbon metabolism, storage and regulation in the diatom *P. tricornutum* acclimated to a 16/8 h dark period highlight more than 4,500 transcripts that were differentially regulated, many of which are associated with carbohydrate and lipid metabolism and responsive to light. Orthologous genes not previously described in algae, such as a pyruvate transporter, have been identified. Examples of concerted transcriptional activities regulating steps of the same process/metabolic
Diel gene regulation in *P. tricornutum*

pathway are suggested, e.g. fast responding genes in fatty acid biosynthesis. More than 100 transcripts related to cell cycle regulation, DNA replication and cell division show a high degree of co-regulation.

Our data confirm the spatial regulation of different pathways in carbon metabolism, with some processes taking place in e.g. either plastids or mitochondria. Enzymes involved in glycolysis/gluconeogenesis, however, are localized to the cytosol, chloroplast and to mitochondria. Our data suggest a clear separation also in time of the processes that capture carbon (light period) and mitosis (dark period) in *P. tricornutum* acclimated to a regular light/dark cycle of 16/8 h. In general, pathways localized to chloroplasts, such as carbon fixation, glucan biosynthesis and lipid biosynthesis, are upregulated at the beginning of the light period. Pathways localized to the mitochondria, such as the TCA cycle and lipid beta-oxidation, peak towards the end of the light period. The localization combined with the diel expression pattern identified in this study may be of help in dissecting the roles of different isoenzymes. Furthermore, this dataset provides a valuable source to aid the work of mining genes involved in light responses and circadian rhythms.
Material and methods

Cultures and growth conditions

Phaeodactylum tricornutum Bohlin (originating from the CCMP 632 strain) was grown in axenic cultures at 20°C and illuminated with 150 µmol photons m⁻² s⁻¹ provided by fluorescent tubes (PHILIPS Master TL-D 36W/840) in cycles of 16/8 hour photoperiods. Growth medium was made from natural seawater filtered through 0.2 µm Tuffryn® membrane capsules (PALL GELMAN) before nutrient amendment and autoclaving. Nutrient amendment was according to Guillard’s f/2 recipe (Guillard, 1975), but P was increased to 1.1 of the f/2 recipe to prevent P limitation throughout the experimental period.

Three cultures of each treatment were grown in 2000 ml Nalgene optically clear flasks, and aerated with a mixture of air and 1-2% CO₂ (v/v) to prevent settling of cells and C limitation. Without addition of CO₂ its concentration and thus also pH, could vary dependent on presence/absence of light, with biased light/dark responses as a consequence. Cultivation was performed in exponential fed-batch modus: Fresh medium was added to the cultures every minute by micro pumps (Bio-Chem Valve™ Inc.) and medium doses increased exponentially based on information on actual culture volume and the desired dilution rate. Maximum specific growth rate was determined from measurements of optical density at 750 nm (OD₇₅₀) and cell numbers in batch cultures prior to the fed-batch study, and an average $u_{max}$ of 1.4 day⁻¹ was observed in different treatments (unlimited/N- or P-limitation). Both experimental cultures were filled and harvested several times until a stable optical density, i.e. biomass, for more than 5 consecutive days confirmed steady state growth. Steady state was used as a criterion for when sampling was initiated. An increase in cell numbers of 30% was observed during the dark phase of the experimental period,
indicating that the effective growth rate was slightly higher in the experimental fed-batch cultures, 1.55 day\(^{-1}\) (±0.12), based on a regression of cell numbers vs. time in the two replicate cultures. Both cultures were checked to confirm axenity by flow cytometric analyses of samples stained with SYBR® Green I during the cultivation period.

**Sampling and analyses**

At each sampling point during a 26.5 hours period two samples of 50 ml were removed from each culture. The sampling started 0.5 after the light was turned on in the morning on the first day, and was finalized after 27 hours. The sampling points were as follows (given as hours after “light on”): T1=0.5, T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), T7=23.5 (darkness), and T8=27 (next day). T8 represent the chronological midpoint between T1 and T2 in the figures, to emphasize the cyclicity of cell processes in acclimatized fed-batch cultures. In the analyses of gene expression sample T7 is used as reference. One 50 ml sample was used for measurements of OD\(_{750}\) and Nile Red-induced fluorescence (NRF), analyses of particulate carbon and nitrogen (CN) and water-soluble carbohydrates (data analyzed at all time points). Flow cytometric analyses of cell numbers were performed on samples from time points 1, 2, 4, 6 and 8. The other 50 ml sample was harvested by centrifugation for RNA extraction.

**Biomass and cell chemistry**

OD\(_{750}\) was measured in replicate aliquots of 2.5 mL from each culture in a Merck Spectroquant Pharo 100 spectrophotometer. The same samples were transferred to a Turner Designs Aquaflor\(^{TM}\) spectrofluorometer for measurement of autofluorescence.
Diel gene regulation in *P. tricornutum*

(AFL) before the samples were stained with Nile Red (0.25 mg mL\(^{-1}\) in acetone) and incubated for at least 15 minutes in darkness at room temperature prior to measurements of Nile Red fluorescence (Chen et al., 2009). Excitation/emission wavelengths are centered around 525/575 nm on this instrument, and the measured fluorescence is considered as representative of neutral lipids. The measured NRF was corrected for the AFL from the cells, and made biomass specific by normalization to cell counts.

Cells were counted in a Becton Dickinson FACScan flow cytometer, equipped with an argon-ion laser that provides excitation light of 488 nm. Subsamples of cells fixed with glutaraldehyde were analyzed by collecting signals of side scatter and autofluorescence (detector FL3 with a 650nm/LP filter), and Fluoresbrite™ Carboxylate YG 1 µm Microspheres (Polysciences, Inc.) were used as internal reference.

Samples for analysis of particulate C and N were collected on precombusted Whatman GF/F filters and stored at -23°C until analysis. Duplicate samples and blank filter pieces were bored out and treated with fuming HCl (37%) before they were packed into tin capsules and dried at 60°C. The samples were analyzed on an ECS 4010 Costech Instruments element analyzer.

Total water soluble carbohydrates was analyzed spectrophotometrically using a modified phenol-sulphuric method where absorbance readings at 485 nm are converted to carbohydrate content using glucose as standard (Granum and Myklestad, 2002).

Samples for gravimetric determination of total neutral (chloroform extractable) lipids (TNL) were collected by centrifugation of 1.25 L culture from each treatment when the experimental sampling was finished and the cultures refilled. The samples were
Diel gene regulation in *P. tricornutum*

*Sample preparation and RNA isolation*

From each time point a sample of 50 ml culture volume was collected from each culture flask. The samples were centrifuged at 4000g for 10 min at 15°C and resuspended in 1 ml f/2 medium. The resuspended cells were transferred to 2 ml tubes and centrifuged at 18000g for 1 min, the supernatant removed and the remaining cell pellet was flash frozen in liquid nitrogen and stored at 80°C. Samples harvested during the dark period were protected from stray light by covering tubes with aluminum foil, and as little light as possible was allowed in the lab during the sampling and centrifugation. Total RNA was isolated from the samples using the Spectrum™ Plant Total RNA kit (Sigma-Aldrich) as described in Nymark *et al.* (2009).

*DNA microarray experiments*

The RNA from each sample (83 - 200 ng) was reverse transcribed, amplified and labeled using the Low Input Quick Amp Labeling Kit, One-Color (Agilent p/n 5190-2305). 1650 ng cRNA from each sample was fragmented and hybridized on 4x44K *P. tricornutum* whole-genome 60-mer oligonucleotide microarrays (Agilent Technologies) in an Agilent G2545A Hybridization rotary oven (10 rpm, 65°C, 17.5 h). Hybridization was performed with the Gene Expression Hybridization Kit (Agilent p/n 5188-5242). The slides were washed with buffer 1 & 2 from Gene Expression Wash Buffer kit (Agilent p/n 5188-5327), and scanned twice at 5 μm resolution on a laser scanner (G2505 B from Agilent Technologies), using the “dynamic range
Diel gene regulation in *P. tricornutum*

expander” option in the scanner software. The resulting images were processed using Agilent Feature Extraction software v9.5. Microarray data are deposited in Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) with accession number GSE42514.

Statistical analysis

The single color scan data (feature extraction files) were analyzed, and genes with statistically significant differential expression were identified using the Limma package (version 3.2.3) (Smyth et al., 2005) and R (version 2.10.1). Spots identified as feature outliers were excluded from analysis, and weak or not detected spots were given reduced weight (0.5). The data was normalized using the quantile method and no background subtraction was performed. A design matrix was created and contrasts between time point T7 (23.5, last dark time point) against each of the time points (T1=0.5, T2=6, T3=10.5, T4=15.5, T5=16.5, T6=20, T7=23.5, and T8=27) were computed. The Benjamini and Hochberg’s method was used to estimate the false discovery rate (Benjamini and Hochberg, 1995). Genes with an adjusted P-value below 0.05 were considered to be statistically significant differential expressed. Genes selected for further analysis must have an average adjusted p value less than 0.05 in at least one time point. It should be noted that each gene could be represented with from 1-5 oligo probes.

Annotation of genes and pathways

The genes were annotated based on data from the *P. tricornutum* genome database at the DOE Joint Genome Institute (Bowler et al., 2008); http://genome.jgi-psf.org/Phatr2/Phatr2.home.html. Where no name was given, we used the according
Diel gene regulation in *P. tricornutum*

protein ID. The biochemical pathways were drawn based on data from the KEGG (Kanehisa et al., 2012) and DiatomCyc (Fabris et al., 2012) databases using Adobe Illustrator.

Network analysis

A table with log2 ratios for all genes with positive detection in at least one of the time points (10365 in total) was used to calculate Pearson correlation values for each gene-pair. To reduce the number of genes included in the analysis a filtered list of genes where only the most responsive genes was represented was made using the following filtering criteria: the gene must have an absolute log2 value >1.5 and adjusted p-value <0.05 for at least one of the time points. The resulting list of 1991 genes was used to extract gene-pairs with Pearson correlation values above 0.95, in total 44024 gene-pairs. The gene-pair list was imported into Cytoscape (Cline et al., 2007) and an unweighted network was made using the force-directed algorithm.
Diel gene regulation in *P. tricornutum*

Acknowledgements

This work is part of the LIPIDO project (Norwegian Research Council project no. 187161). Technical support from the national functional genomics lab for biological specimens (BTAC-FUGE) was made possible through a grant from the Norwegian Research Council (FUGE grant # 184146/S10). The authors would like to thank Leila Alipanah for assisting with harvesting of cultures and Torfinn Sparstad for help with the microarray hybridization.
References


Diel gene regulation in *P. tricornutum*

plastidial and microsomal Delta 12-fatty acid desaturases. Plant Physiology 131: 1648-1660


Hockin NL, Mock T, Mulholland F, Kopriva S, Malin G (2012) The response of diatom central carbon metabolism to nitrogen starvation is different from that of green algae and higher plants. Plant Physiology 158: 299-312


Diel gene regulation in P. tricornutum


Nakazawa N, Mehrotra R, Ebe M, Yanagida M (2011) Condensin phosphorylated by the Aurora-B-like kinase Ark1 is continuously required until telophase in a mode distinct from Top2. Journal of Cell Science 124: 1795-1807


Diel gene regulation in *P. tricornutum*


Sanchez S, Bouget FY, Sanchez F, Garnier L (2009) Functional genomics approaches to study the involvement of transcription factors in the microalgae Ostreococcus tauri circadian clock. Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology 153A: S207-S207


Diel gene regulation in *P. tricornutum*


Diel gene regulation in *P. tricornutum*

Figure legends

Figure 1. Variation in cell chemistry over a 16/8 light/dark cycle in *P. tricornutum.*

A, Water-soluble carbohydrates (pg cell\(^{-1}\)/squares) and cell specific Nile red-fluorescence (NRF, diamonds/right axis). B, Cellular content of C (pg cell\(^{-1}\)/triangles) and N (pg cell\(^{-1}\)/squares/right axis). Open/filled symbols represent replicate cultures and lines shows the average of the two replicate data series. The lower x axis shows hours from “light on” in the morning, and the upper x axis shows sampling times labelled T1 to T8. The last sampling point in the series (T8, sampled at 27 h after start) corresponds to a point between T1 and T2 (time 0.5 and 6 h after start) and has been indicated in the graphs (dashed lines and shaded symbols). Dark bar on top and lines indicate the period when the light is off.

Figure 2. Coexpression network of diel cycling genes in *P. tricornutum.* A network representation of the 1,936 most responsive genes in the experiment was made using Cytoscape and Pearson correlation values (r > 0.95) for each gene pair. In total 44003 gene pairs are included in the figure. Six clusters of tightly co-regulated genes are distinguished, and graphs showing the general expression pattern of a given cluster (normalized to time point T7=23.5) are presented for each cluster (labeled 1-6).

Figure 3. Coordinated regulation of genes coupled to carbon fixation in *P. tricornutum.* Expression levels at sampling points (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and T7=23.5 (darkness) are normalized to time point T7. Protein names and Genbank accession numbers: TPI/GapC3, triosephosphate isomerase/glyceraldehyde-3-phosphate dehydrogenase (EEC50801.1); Prk, phosphoribulokinase (ACI65926.1);
Diel gene regulation in *P. tricornutum*

Phosphoglycerate kinase (Phatr2_29157, EEC46310); GapC4, glyceraldehyde-3-phosphate dehydrogenase (EEC50484.1); FBP, fructose-1,6-bisphosphatase (EEC44418.1); FBPC1, fructose-1,6-bisphosphatase (EEC51753.1); Tkl, transketolase (EEC50321.1); Rpe, plastidic ribulose-phosphate 3-epimerase (EEC51500.1); Tpt3, plastidal triose-phosphate/phosphate translocator (EEC51811.1); PtCa1, plastidal beta carbonic anhydrase (EEC51057.1); PtCa2, plastidic beta carbonic anhydrase (EEC48791.1); SLC4A_1, plastidial HCO₃⁻ transporter (EEC48911.1); SLC4A_2, plastidial HCO₃⁻ transporter.

Figure 4. Light/dark cycle regulation of genes encoding glycolytic and glucan biosynthesis enzymes in *P. tricornutum*. Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log₂ transformed gene expression ratios. +: GAPC3 is a protein fusion with triosephosphate isomerase (TPI). Abbreviations: C, cytosolic localization; ER, endoplasmatic reticulum localization; M, mitochondrial localization; P, plastid localization; V, vacuolar localization; PGM, phosphoglucomutase; GPI, glucose-6-phosphate isomerase; PFK, phosphofructokinase; FBP, fructose-1,5-bisphosphate phosphatase; Fba, fructose-1,6-bisphosphate aldolase; TIM, triosephosphate isomerase; GapC/GAPDH, glyceraldehyde-phosphate dehydrogenase; PGK, phospho-glycerate kinase; PGAM, phospho-glycerate mutase; PK, pyruvate kinase; G6Pase, glucose-6-phosphate isomerase; PGM, phosphoglucomutase; BGS, 1,3-beta-glucan synthase; BGER, beta-glucan elicitor receptor.
Figure 5. Light/dark cycle regulation of genes encoding TCA cycle enzymes in *P. tricornutum*. Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log2 transformed gene expression ratios. Asterisk indicates that the gene did not show significant differential (p<0.05) expression at any timepoint. Abbreviations: PDH, pyruvate dehydrogenase; DHLTA, dihydrolipoamide acetyltransferase; DLDH, dihydrolipoyl dehydrogenase; PYC, pyruvate carboxylase; CISY, citrate synthase; IDH, isocitrate dehydrogenase; OGD, 2-oxoglutarate dehydrogenase E1 component; SCSα, succinyl-CoA ligase alpha subunit; SDH, succinate dehydrogenase flavoprotein; FUM, fumarase; MDH, malate dehydrogenase.

Figure 6. Light/dark cycle regulation of genes encoding fatty acid metabolism enzymes in *P. tricornutum*. A, Fatty acid biosynthesis; B, Fatty acid beta-oxidation. Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log2 transformed gene expression ratios. Question mark indicates that no candidate gene for the enzyme activity has been identified. Asterisk indicates that the gene did not show significant differential expression (p<0.05) at any time point. Abbreviations: ACC, acetyl-CoA carboxylase; FABD, malonyl-CoA transacylase; FABB/FABF, 3-oxoacyl-[acyl-carrier-protein] synthase; FABG, 3-oxoacyl-(acyl-carrier-protein) reductase; FABZ,
Diel gene regulation in *P. tricornutum*

3R-hydroxyacyl-[acyl carrier protein] dehydrase; FABI, enoyl-ACP reductase; HAD, 3-hydroxyacyl-coenzyme A dehydrogenase; KCT, beta-ketoacyl-CoA thiolase.

Figure 7. Coordinated regulation of genes involved in cell division and mitosis. Expression levels at sampling points (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and T7=23.5 (darkness) are normalized to time point T7. Protein names and Genbank accession numbers: Aurora-A, Aurora kinase A (EEC48936.1); CYCB1, Cyclin B1 (EEC47805.1); CDC20_1, cell division cycle 20 (EEC47990.1); Phatr2_45765, INCENP protein (EEC48380.1); Pt-KIF7, kinesin family-like protein (EEC45857.1); dsCYC3, cyclin protein (EEC51859.1); Pt-KIF10, kinesin family-like protein (EEC45010.1); Phatr2_42614, abnormal spindle-like protein (EEC51580.1); CAPG_48922, non-SMC condensin I complex, subunit G (EEC44841.1); Phatr2_52174, DNA topoisomerase II (EEC47151.1); Phatr2_48134, kinesin motor domain protein (EEC45795.1); hCDK1, MAP kinase (EEC47173.1); pT-KIF12, kinesin family-like protein (EEC51889.1); CAPD2_43956, non-SMC condensin I complex subunit D2 (ACI65548.1); CAPH_36284, non-SMC condensin I complex, subunit H (EEC47398.1); CAPD3_49507, condensin-II complex subunit D3 (EEC44317.1); SMC4, condensin complex subunit SMC4 (ACI65666.1); CAPH2_50280, condensin-II complex subunit H2 (EEC43358.1); SMC2, condensin complex subunit SMC2 (EEC44274.1); CAPG2_44765, condensin-II complex subunit G2 (EEC49642.1); Borealin_46380, borealin-related protein (EEC47609.1). Genes which previously have not been assigned gene symbols or gene functions are marked with an underscore and corresponding Phatr2 ID (Phatr2: version 2.0 of the *P. tricornutum* genome).
Diel gene regulation in *P. tricornutum*

Supplementary Figure 1. Light/dark cycle regulation of genes encoding a pyruvate transporter and the plastidic pyruvate dehydrogenase complex in *P. tricornutum*. A, Phylogenetic analysis of the pyruvate transporter in plants and heterokonts. A Neighbour Joining (NJ) tree was constructed based on a protein alignment of the published pyruvate transporter from *A. thaliana* and *F. trinerva* and related proteins from plants, heterokonts and bacteria. A transporter from *Leptospira interrogans* (Acc. No. AAN51825) was chosen as outgroup. Different lineages are indicated as follows: Heterokonts, brown; Viridiplantae, green; Bacteria, black. B, Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log2 transformed gene expression ratios. Abbreviations: PDH, pyruvate dehydrogenase; DHLTA, dihydrolipoamide acetyltransferase; DLDH, dihydrolipoyl dehydrogenase.
Figure 1

A

B

pg carbohydrate per cell
pg N per cell
pg C per cell
pg N per cell

Hours

pg carbohydrate per cell
pg C per cell
pg N per cell

Hours

Downloaded from on January 22, 2018 - Published by Downloaded from on January 22, 2018 - Published by
Copyright © 2012 American Society of Plant Biologists. All rights reserved.
Figure 1. Variation in cell chemistry over a 16/8 light/dark cycle in *P. tricornutum*. A, Water-soluble carbohydrates (pg cell\(^{-1}\)/squares) and cell specific Nile red-fluorescence (NRF, diamonds/right axis). B, Cellular content of C (pg cell\(^{-1}\)/triangles) and N (pg cell\(^{-1}\)/squares/right axis). Open/filled symbols represent replicate cultures and lines shows the average of the two replicate data series. The lower x axis shows hours from “light on” in the morning, and the upper x axis shows sampling times labelled T1 to T8. The last sampling point in the series (T8, sampled at 27 h after start) corresponds to a point between T1 and T2 (time 0.5 and 6 h after start) and has been indicated in the graphs (dashed lines and shaded symbols). Dark bar on top and lines indicate the period when the light is off.
Figure 2

Hours

Relative gene expression

0 10 20 5 15 25

Hours

Relative gene expression

0 10 20 5 15 25

Hours

Relative gene expression

0 10 20 5 15 25

Hours

Relative gene expression

0 10 20 5 15 25

Downloaded from on January 22, 2018 - Published by www.plantphysiol.org
Copyright © 2012 American Society of Plant Biologists. All rights reserved.
Figure 2. Coexpression network of diel cycling genes in *P. tricornutum*. A network representation of the 1,936 most responsive genes in the experiment was made using Cytoscape and Pearson correlation values (*r* > 0.95) for each gene pair. In total 44003 gene pairs are included in the figure. Six clusters of tightly co-regulated genes are distinguished, and graphs showing the general expression pattern of a given cluster (normalized to time point T7=23.5) are presented for each cluster (labeled 1-6).
Figure 3

Time points

log2 ratios

TPI/GapC3
Phosphoribulokinase (Prk)
Phosphoglycerate kinase_29157
GapC4
FBP
FBPC1
TkI
Rpe
Tpt3
PtCa1
PtCa2
SLC4A_1
SLC4A_2
Figure 3. Coordinated regulation of genes coupled to carbon fixation in *P. tricornutum*. Expression levels at sampling points (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and T7=23.5 (darkness) are normalized to time point T7. Protein names and Genbank accession numbers: TPI/GapC3, triosephosphate isomerase/glyceraldehyde-3-phosphate dehydrogenase (EEC50801.1); Prk, phosphoribulokinase (ACI65926.1); Phosphoglycerate kinase (Phatr2_29157, EEC46310); GapC4, glyceraldehyde-3-phosphate dehydrogenase (EEC50484.1); FBP, fructose-1,6-bisphosphatase (EEC44418.1); FBPC1, fructose-1,6-bisphosphatase (EEC51753.1); TkI, transketolase (EEC50321.1); Rpe, plastidic ribulose-phosphate 3-epimerase (EEC51500.1); Tpt3, plastidal triose-phosphate/phosphate translocator (EEC51811.1); PtCa1, plastidal beta carbonic anhydrase (EEC51057.1); PtCa2, plastidic beta carbonic anhydrase (EEC48791.1); SLC4A_1, plastidial HCO₃⁻ transporter (EEC48911.1); SLC4A_2, plastidial HCO₃⁻ transporter.
Figure 4. Light/dark cycle regulation of genes encoding glycolytic and glucan biosynthesis enzymes in *P. tricornutum*. Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log2 transformed gene expression ratios. +: GAPC3 is a protein fusion with triosephosphate isomerase (TPI). Abbreviations: C, cytosolic localization; ER, endoplasmatic reticulum localization; M, mitochondrial localization; P, plastid localization; V, vacuolar localization; PGM, phosphoglcomutase; GPI, glucose-6-phosphate isomerase; PFK, phosphofructokinase; FBP, fructose-1,5-bisphosphate phosphatase; Fba, fructose-1,6-bisphosphate aldolase; TIM, triosephosphate isomerase; GapC/GAPDH, glyceraldehyde-phosphate dehydrogenase; PGK, phospho-glycerate kinase; PGAM, phospho-glycerate mutase; PK, pyruvate kinase; GPI, glucose-6-phosphate isomerase; PGM, phosphoglucomutase; BGS, 1,3-beta-glucan synthase; BER, beta-glucan elicitor receptor.
Figure 5

Glycolysis

Pyruvate

Dihydro-lipoamide-E

Lipoamide-E

S-acetyldihydro-lipoamide-E

PDHA1

PDHB1

Acetyl-CoA

DHLTA_1

38009

S-acetyldihydro-lipoamide-E

Dihydro-lipoamide-E

Lipoamide-E

CISY1

54834

MDH

OGD1

45017

IDH1

40430

FUM1

51720

SDH1

SDH2

Fatty acid metabolism

Fatty acid biosynthesis

PDHA1

PDHB1

DLDH

PYC1

DLDH

Glycolysis

2-oxoglutarate

S-succinyl-dihydrolipoamide-E

Dihydrolipoamide-E

Malate

Citrate

Isocitrate

Fumarate

Succinate

Succinyl-CoA

Oxaloacetate

MDH

IDH1

OGD1

2.0

1.5

1.0

0.5

0.0

-0.5

-1.0

-1.5

-2.0

-3.0

>3.0

www.plantphysiol.org on January 22, 2018. Published by www.plantphysiol.org. Copyright © 2012 American Society of Plant Biologists. All rights reserved.
Figure 5. Light/dark cycle regulation of genes encoding TCA cycle enzymes in *P. tricornutum*. Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log2 transformed gene expression ratios. Asterisk indicate that the gene did not significant differential (p<0.05) expression at any timepoint. Abbreviations: PDH, pyruvate dehydrogenase; DHLTA, dihydrolipoamide acetyltransferase; DLDH, dihydrolipoyl dehydrogenase; PYC, pyruvate carboxylase; CISY, citrate synthase; IDH, isocitrate dehydrogenase; OGD, 2-oxoglutarate dehydrogenase E1 component; SCSα, succinyl-CoA ligase alpha subunit; SDH, succinate dehydrogenase flavoprotein; FUM, fumarase; MDH, malate dehydrogenase.
Figure 6

A. Citrate cycle

- Acetyl-CoA
  - Malonyl-CoA
    - Malonyl-(acp)
    - Acetyl-(acp)
    - Acetoacetyl-(acp)
    - D3-hydroxybutryl-(acp)
    - Crotonyl-(acp)
    - Butyryl-(acp)

B. Fatty acid biosynthesis

- Hexadecanoyl-CoA
  - (S)-3-Hydroxy-hexadecanoyl-CoA
    - trans-Hexadec-2-enoyl-CoA
    - Hexadecanoyl-CoA
      - Acetyl-CoA
        - Acetoacetyl-CoA
          - 3-Oxohexadecanoyl-CoA
            - 3-Oxo-hexadecanoyl-CoA
              - Acetyl-CoA
                - Elongation
                  - Hexadecanoyl-(acp)
                    - Hexadecanoic acid
                      - Citrate cycle
Figure 6. Light/dark cycle regulation of genes encoding fatty acid metabolism enzymes in *P. tricornutum*. A, Fatty acid biosynthesis; B, Fatty acid beta-oxidation. Colored squares indicate expression levels at time point (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and the data are normalized to time point T7=23.5 (darkness). Light and dark periods are indicated in the bar above the squares. The color scale indicates log2 transformed gene expression ratios. Question mark indicates that no candidate gene for the enzyme activity has been identified. Asterisk indicates that the gene did not show significant differential expression (p<0.05) at any time point. Abbreviations: ACC, acetyl-CoA carboxylase; FABD, malonyl-CoA transacylase; FABB/FABF, 3-oxoacyl-[acyl-carrier-protein] synthase; FABG, 3-oxoacyl-(acyl-carrier-protein) reductase; FABZ, 3R-hydroxyacyl-[acyl carrier protein] dehydrase; FABI, enoyl-ACP reductase; HAD, 3-hydroxyacyl-coenzyme A dehydrogenase; KCT, beta-ketoacyl-CoA thiolase.
Figure 7. Coordinated regulation of genes involved in cell division and mitosis. Expression levels at sampling points (hours after light on) T1=0.5, T8=27 (day 2), T2=6, T3=10.5, T4=15.5, T5=16.5 (darkness), T6=20 (darkness), and T7=23.5 (darkness) are normalized to time point T7. Protein names and Genbank accession numbers: Aurora-A, Aurora kinase A (EEC48936.1); CYCB1, Cyclin B1 (EEC47805.1); CDC20_1, cell division cycle 20 (EEC47990.1); Phatr2_45765, INCENP protein (EEC48380.1); Pt-KIF7, kinesin family-like protein (EEC45857.1); dsCYC3, cyclin protein (EEC51859.1); Pt-KIF10, kinesin family-like protein (EEC45010.1); Phatr2_42614, abnormal spindle-like protein (EEC51580.1); CAPG_48922, non-SMC condensin I complex, subunit G (EEC44841.1); Phatr2_52174, DNA topoisomerase II (EEC47151.1); Phatr2_48134, kinesin motor domain protein (EEC45795.1); hCDK1, MAP kinase (EEC47173.1); pT-KIF12, kinesin family-like protein (EEC51889.1); CAPD2_43956, non-SMC condensin I complex subunit D2 (ACI65548.1); CAPH_36284, non-SMC condensin I complex, subunit H (EEC47398.1); CAPD3_49507, condensin-II complex subunit D3 (EEC44317.1); SMC4, condensin complex subunit SMC4 (ACI65666.1); CAPH2_50280, condensin-II complex subunit H2 (EEC43358.1); SMC2, condensin complex subunit SMC2 (EEC44274.1); CAPG2_44765, condensin-II complex subunit G2 (EEC49642.1); Borealin_46380, borealin-related protein (EEC47609.1). Genes which previously have not been assigned gene symbols or gene functions are marked with an underscore and corresponding Phatr2 ID (Phatr2: version 2.0 of the P. tricornutum genome).