PNAS
www.pnas.org

Supplementary Information for

Multi-omics resolution of molecular events of a day during the life of Chlamydomonas

Daniela Strenkert, Stefan Schmollinger, Sean D. Gallaher, Patrice A. Salomé, Samuel O. Purvine, Carrie D. Nicora, Tabea Mettler-Altmann, Eric Soubeyrand, Andreas Weber, Mary Lipton, Gilles J. Basset and Sabeeha S. Merchant

Corresponding authors: Sabeeha Merchant
Email: merchant@chem.ucla.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1815238116

Other supplementary materials for this manuscript include the following:

Additional data (separate file)

Dataset S1. Average expression estimates for all Chlamydomonas genes over the diurnal cycle.
Dataset S2. Untargeted proteomics detection of soluble Chlamydomonas proteins over the diurnal cycle.
Dataset S3. Co-expression networks in Chlamydomonas based on our study.
Dataset S4. Expression estimates and list of genes involved in DNA replication. (excel file).
Dataset S5. Expression estimates and list of genes involved in photosynthesis. (excel file).
Dataset S6. Expression estimates and list of genes involved in respiration. (excel file).
Dataset S7. Expression estimates and list of genes involved in fermentation. (excel file).
Dataset S8. Expression estimates and list of genes involved in the CCM complex. (excel file).
Dataset S9. Expression estimates and list of genes involved in the CCM complex in our study and Zones *et al*., 2015. (excel file).
Dataset S10. Expression estimates and list of truncated hemoglobin genes. (excel file).

Supplementary Methods

Library Preparation and Reads Analysis

Non-protein-coding RNA was filtered by alignment to a pseudo-assembly consisting of *C. reinhardtii* rRNA and snoRNA sequences. The remaining reads were aligned to the latest assembly of the Chlamydomonas nuclear genome (v5 from the Phytozome portal) combined with v4 of the organellar genomes (1). Alignment was performed with STAR_2.5.2b (2) using default parameters except --alignIntronMax 5000. Relative expression estimates in terms of FPKMs were generated from the BAM files using cuffdiff v2.0.2 (3) with default parameters except -max-bundle-frags 1000000000 --multi-read-correct --library-type fr-firststrand (3).

Quantitative real time PCR

Oligonucleotide sequences used in this study:
EIF1A-F GACATCGTGCTGGTGGGTCTT;
EIF1A-R TACGAATGTGCTCCGGCAG;
LHCSR3.1-F CACAACACCTTGATGCGAGATG;
LHCSR3.1-R CCGTGTCTTGTCAGTCCCTG;
LHCSR3.2-F TGTGAGGCACTCTGGTGAAG;
LHCSR3.2-R CGCCTGTTGTCACCATCTTA;
PSBS1-F TTCGTGTGTACGCCGCTTCG;
PSBS2-F TTCGTGTGTACGCTGCCACT;
PSBS-R AGCTTGAGGGGGAACTGCTT.

Proteomics

MS/MS data search

The MS/MS spectra from all LC-MS/MS datasets were converted to ASCII text (.dta format) using MSConvert (http://proteowizard.sourceforge.net/tools/msconvert.html) which precisely assigns the charge and parent mass values to an MS/MS spectrum as well as converting them to centroid. The data files were then interrogated via target-decoy approach (4) using MSGFPlus (5) using a +/- 20 ppm parent mass tolerance, partial tryptic enzyme settings, and a variable posttranslational modification of oxidized Methionine. All MS/MS search results for each dataset were collated into tab separated ASCII text files listing the best scoring identification for each spectrum.

Data analysis

MS/MS search results for each biological replicate were collated to a tab delimited text file using in-house program MAGE Extractor. These results were then imported into a SQL Server (Microsoft, Redmond, WA) database and filtered to 1% FDR by adjusting the Q-Value provided by MSGF+. The in-house program MASIC (https://omics.pnl.gov/software/masic, 36) was run on each dataset, which provides ion statistics and general information extracted from the instrument binary rawfile. Specifically, both the maximum observed MS ion count for each MS/MS spectrum's precursor signal (termed PeakMaxIntensity) and a selected ion chromatogram (SIC, termed StatMomentsArea) are provided for subsequent peptide quantitation. The results from all datasets' MASIC output were collated using the in-house program MAGE File Processor, with a tab delimited text file listing all relevant data as the result. These results were also imported into SQL Server and connected to the filtered MSGF+ results via Dataset_ID (internal to PNNL's data management system), and relevant MS/MS scan number. Unique peptide sequences (with post-translational modifications counted as separate sequences) were grouped per dataset with maximum PeakMaxIntensity values provided (available in Dataset S2). This data was then pivoted to provide a crosstab of peptides, with protein information carried through the query. The crosstab was then imported into InfernoRDN (an implementation of the R statistical package, https://omics.pnl.gov/software/infernordn), Log_2 transformed and mean central tendency normalized (boxplot alignment). The normalized peptide abundance values were then exported into Excel (Microsoft), anti-logged, and proteins grouped with values summed using the Pivot Table function.

Supplementary Results

Survey of synchronization techniques across systems

The study of any rhythmic event during an organism development requires efficient synchronization of all individuals within the population, as well as synchronization of all cells within each organism. Two distinct types of rhythms have been the focus of research for the past decades: cell-cycle (or how one cell turns into two), and diurnal or circadian (or how one organism arranges gene expression over a 24-h day). *Chlamydomonas reinhardtii* is one of the few organisms where the same synchronization (or entrainment) regime can work for both types of rhythms.
For comparison purposes with other systems, we have compiled a list of synchronization protocols, and discuss their relevance for Chlamydomonas rhythmic behaviors.

Bacteria (*E. coli*)

Cell cycle length: ~ 40 minutes.
Circadian period: none.
Synchronization method: selection by filtration and size (6, 7).
Basis of synchronization: cells adhere to cellulose filters, and release daughter cells into the flow-through. Collected daughter cells are smaller in diameter and near synchronous relative to the stage of the cell-cycle.

Synchronization method: temperature-sensitive DnaC mutant (8).
Basis of synchronization: DnaC is an essential DNA replication protein that interacts with the replicative DNA helicase DnaB. DnaC is required to initiate DNA replication; shifting the mutant to the restrictive temperature causes cell-cycle arrest after the completion of replication forks.

Synchronization method: evocation of stringent response (9).
Basis of synchronization: treating cells with serine hydroxamate, an analog of tRNA charging, causes a cell-cycle arrest after DNA replication and before chromosome segregation. Synchronization is achieved by relieving the chemical inhibition by providing fresh medium.

mammalian cell cultures

Cell cycle length: ~12-13 hours.
Circadian period: ~ 22-23 hours.
Synchronization method: chemically-induced cell-cycle arrest.
Basis of synchronization: early G1 arrest can be achieved by treating cells with mimosine or lovastatin; hydroxyurea can arrest cells in S phase; cells can also be selectively blocked during M phase with treatment with microtubule-disrupting agents like nocodazole.
Note: these methods are only effective in synchronizing cell cultures for cell-cycle studies, but they do not synchronize their circadian clocks.

Synchronization method: centrifugal elutriation.
Basis of synchronization: many cell types will grow larger as they progress towards M phase before cytokinesis. This method is similar in principle with bacterial size selection, but uses a constant flow of medium in the direction opposite of the centrifugal force to push cells out of the chamber for collection.
Note: this method will allow the isolation of multiple cell types at any stage of the cell-cycle, without the addition of drugs or changing culture medium. Its potential for synchronization of the circadian clock is not known.

Synchronization method: release from serum starvation.

Basis of synchronization: removal of serum from the culture medium causes cells to arrest at the G_0 stage of the cell-cycle. Adding serum back to the culture medium stimulates cell-cycle entry.
Note: this method synchronizes both the cell-cycle and circadian rhythms in cell cultures. Most studies agree that the circadian clock gates the timing of cell divisions.

cyanobacteria (*Synechococcus elongatus PCC7942*)

Cell-cycle length: 10-24 hours.
Circadian period: ~24 hours.
Synchronization method: dark incubation.
Basis of synchronization: a culture in stationary stage is diluted in fresh BG11 growth medium and transferred in the dark for 1-2 days. Cells initiate cell-cycle progression after being transferred into constant light.
Note: this method will also synchronize the circadian clock by resetting all cells in the culture by the dark-to light transition.

Cell-cycle length: 10-24 hours.
Circadian period: ~24 hours.
Synchronization method: alternating light-dark cycles.
Basis of synchronization: in contrast to plants, light is perceived indirectly via the redox status of the plastoquinone pool by the circadian input kinase CikA (10). Resetting of the clock is therefore coupled to the metabolic state of the cell rather than based on photoreceptor-driven changes in transcript or protein abundance.
Note: will not synchronize cell division. With appropriate microscopes and reporters, one can however track cell division and the circadian clock in individual cells to study gating between the two rhythms (11).

Arabidopsis thaliana (and other plants)

Cell-cycle length: 10-24 hours.
Circadian period: ~24 hours.
Synchronization method: alternating light-dark cycles, total daylength of 24 hours.
Basis of synchronization: the clock genes *CCA1* and *LHY* are induced by light perceived by red/far-red light photoreceptors (phytochromes) and blue light photoreceptors (cryptochromes). The duration of the light part of the cycle may be as short at 4-6 hours. Skeleton photoperiods (two light pulses of less than one hour each) may also substitute for shorter daylengths by mimicking the dark-to light and light-to dark transitions.
Note: this method is not generally applicable to synchronize cell divisions.

Synchronization method: alternating warm-cool cycles, total day length of 24 hours.
Basis of synchronization: warmer temperatures mimic the day, while cooler temperatures mimic night-time. Temperature differentials as small as 1ºC may be sufficient for synchronization.
Note: this method is not generally applicable to synchronize cell divisions. Congruent light-dark and warm-cool cycles tend to reinforce circadian synchronization; providing cool-warm cycles antiphase to light-dark cycles may alter the phase relationship of light-responsive and temperature-responsive clock-controlled genes.

Chlamydomonas reinhardtii

Cell-cycle length: 8-72 hours depending on growth conditions.
Circadian period: 24 hours.
Synchronization method: size selection by Rastgeldi Threshold Centrifuge.

Basis of synchronization: a cell suspension is continuously centrifuged on a special rotor with interconnected openings (or thresholds) that allow cells of a given size to go through. This method works well to separate zoospores from haploid cells in a mixed culture.
Note: this technique was not applied to the separation of early- and late-G_1 phase cells.

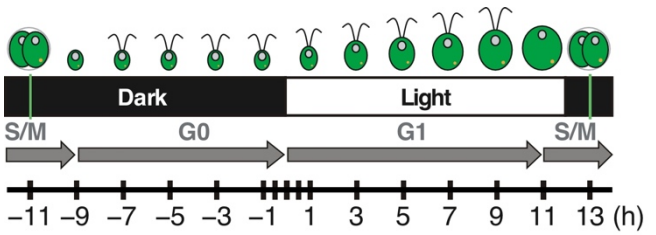
Synchronization method: alternating light-dark cycles, total daylength of 24 hours.
Basis of synchronization: cell growth occurs during the light part of the day, fueled by photosynthesis. Cell division takes place shortly after dusk, either once or multiple times depending on cell volume at the time of cell-cycle commitment. Therefore, one Chlamydomonas mother cell may give birth to two, four, eight, 16 or even 32 daughter cells after dusk. Photoperiod may vary from 12L: 12D, 14L: 10D or 12L: 4D, with temperatures ranging from 24ºC to 35ºC.
Note: Chlamydomonas cells will be synchronized both for cell-cycle and circadian rhythms when the total day length totals 24 hours. It is not known whether the algal circadian clock can be entrained to 12L: 4D cycles.

Optimization of photo-bioreactors parameters for a single cell-doubling

Our photo-bioreactors (Photobioreactor FMT 150 from Photon System Instruments, Brno, Czech Republic) have a flat-panel design that allows uniform light exposure of cultures. All cultures parameters can be pre-programmed to specification and recorded during the course of the experiment. They include external parameters like temperature and light regime, as well as internal parameters like pH of the growth medium, aeration gas bubbling, optical density of the culture and medium flow-rate if in turbidostatic mode.
Sterile photobioreactors were filled with 350 ml High Salt Medium (HSM) and inoculated with Chlamydomonas cells at a starting optical density of 0.05. Cultures were aerated with 0.04% CO_2 (= air levels) and subjected to light-dark cycles (12 hours light, 12 hours dark provided by a mixture of red and blue LEDs for a total fluence rate of 200 µmol/m^2/s) and in-phase temperature cycles (12 hours at 28ºC followed by 12 hours at 18ºC). After reaching an optical density of 0.4 (~ 2-3 x 10^6 cells/mL), cultures were grown in turbidostatic mode, which maintains optical density constant, until the day of sample collection.
The combination of growth medium, light intensity, photoperiod and temperature cycles resulted in an exact doubling of cells shortly after lights-off, as measured by manual counting of cells under a microscope and by automated counting and cell diameter measurements with a Beckman Coulter Multisizer 3 (Beckman Coulter Inc., USA, 50 µm orifice).

Chlamydomonas cell division across the diurnal cycle

Chlamydomonas cells proliferate using a modified cell cycle called multiple fission, owing to the fact that one mother cell can give rise to 2-32 daughter cells; the number of daughter cells depends on the final volume of the mother cell at the time of division (12). Note that asymmetric division is sometimes used to describe cell division in Chlamydomonas, but this term is misleading, as daughter cells all have the same size. The word asymmetric rather applies to the distribution of materials between daughter cells.

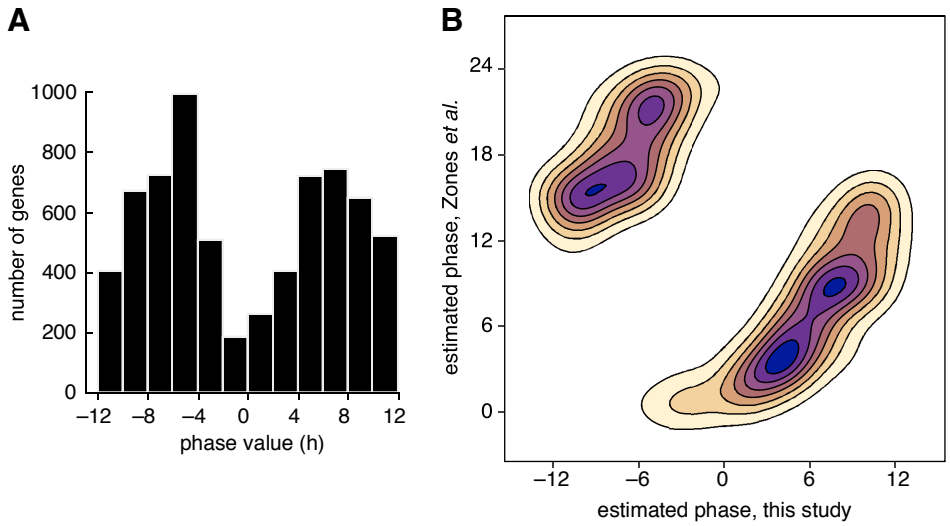


SOM Fig. 1. Diagram of Chlamydomonas cell division progression over the diurnal cycle.
Cells grow in volume and accumulate biomass during the day, using photosynthesis to fuel growth. Around dusk, cells will initiate the mitotic division and yield two daughter cells per mother cell. At night, daughter cells will regenerate their flagella, which were resorbed before cell division, and enter the quiescent phase of the cell cycle until the following dawn.
The green vertical line indicates the timing of cell division.

Under light-dark cycles, the Chlamydomonas cell cycle can be divided according to the stage of the diurnal cycle: G_0 or quiescent phase in the dark, G_1 or growth in the light, and DNA replication and cell division around dusk (SOM Fig. 1) (13, 14). Cells pass through a first checkpoint during the first six hours of G_1 called the "commitment point", similar to the yeast "start" cell-cycle checkpoint. As the name implies, cells that have gone through commitment will divide by the end of day. A second checkpoint, known as mitotic sizer, decides on how many divisions the mother cell will go through, and depends on its final volume before the first division (12). Once the number of divisions is established, the mother cell will undergo several rapid alternating DNA synthesis and cell division phases. All daughter cells will have the same size.

Comparison of rhythmicity across datasets

We have determined the expression pattern and peak expression window by collecting samples over one full diurnal cycle, centered on the dark-to light transition, with one sample every 1-2 hours. In another recent study from Jim Umen's laboratory, samples were collected over a full diurnal cycle, starting at the light-to dark transition, to highlight genes involved in cell division. Samples in this study were collected every 30 minutes to 1 hour. We processed both datasets with the algorithm JTK_CYCLE (15). By convention, lights on is taken as CT 0, so we adjusted our time series accordingly. Both datasets yield all possible phase values between 0 and 24 hours (SOM Fig. 2A). When plotting the phase of genes deemed to be rhythmic in both datasets, we observed good correlation, and provides a global validation of our results (SOM Fig. 2B).



SOM Fig. 2. Consistency of diurnal phases across independent datasets.
A, Distribution of phase values estimated with the JTK method of the algorithm MetaCycle. Genes were deemed expressed if at least one time point reached ≥ 1 FPKM, and rhythmic with a BH.Q ≤ $1x10^{-10}$. By convention, dawn is taken as reference point (phase "0").
B, Alignment of estimated phase values from this study and Zones *et al* (34). Values are plotted as a density plot. Note that phase values of +12 and –12 are equivalent.

Other rhythmic datasets are not as sample-dense (16, 17), so phase calling is not possible. However, a cursory exploration of the data does indicate that our dataset generally agrees with the gene expression patterns described (Table S2). We extracted expression profiles for the following sets of genes for comparison with our and other datasets: histone genes, MCM complex genes, PSI. Finally, we compiled a list of publications that described the rhythmic behavior of single genes or single molecules (18-24), and compared their results to our own. All results are given in Table S2.

Identification of Chlamydomonas orthologues

All Chlamydomonas gene identifiers are given in Supplementary Datasets S4-S11. We provide details about the identification of some of the most pertinent genes mentioned in the text below. For DNA replication-related genes, BLASTs were performed on the phytozome platform (https://phytozome.jgi.doe.gov/) with the yeast proteins as query, with the filtering option unchecked. Other gene lists were manually curated from the literature (25-34).

Ribonucleotide Reductase (RNR): the Chlamydomonas genome consists of one orthologues of the yeast RNR1 gene (the large subunit of the enzyme), and two orthologues for RNR2 (the small subunit): RNR1/RIR1: Cre12.g492950 (BLAST score: 1082, E value = 0)
RNR2/RIR2A: Cre12.g491050 (BLASTp score: 432, E value = 4.5E-149)
RNR2/RIR2B: Cre12.g509400 (BLASTp score: 397.5, E value = 4.3E-134).

Mini-Chromosome Maintenance (MCM) complex: the MCM complex is encoded by six genes in yeast (MCM2-MCM7), and each gene has a single orthologue in Chlamydomonas. Note that the name MCM1 is taken by a transcription factor that is not involved with the MCM complex.
MCM2: Cre07.g338000 (BLASTp score: 713, E value = 0).
MCM3: Cre06.g295700 (BLASTp score: 548.5, E value = 3.3E-179).
MCM4: Cre07.g316850 (BLASTp score: 569.3, E value = 0).
MCM5: Cre01.g023150 (BLASTp score: 572.8, E value = 0).
MCM6: Cre03.g178650 (BLASTp score: 591.7, E value = 0).
MCM7: Cre10.g455850 (BLASTp score: 614, E value = 0).

DNA polymerase delta: the catalytic subunit of the DNA-dependent DNA polymerase POLD1 is present as a single copy gene in Chlamydomonas and is the orthologue of yeast POL3. Of the two other subunits of DNA polymerase delta, the Chlamydomonas genome encodes a single orthologue for *POL31*, but appears to lack a sequence related to *POL32*).
POL3/POLD1: Cre01.g015250 (BLASTp score: 950.7, E value = 0).
POL31/POLD2: Cre08.g374050 (BLASTp score: 136.3, E value = 1.3E-34).
Histones: the yeast genome codes for one Histone H1 (HHO1), two Histone H2A (HTA1, HTA2), two Histone H2B (HTB1, HTB2), two Histone H3 (HHT1, HHT2), and two Histone H4 (HHF1, HHF2).

Histone H1: Cre13.g567450 (HON1, BLASTp score: 40.0, E value = 8.8E-4); Cre06.g275900 (HON2, BLASTp score: 36.2, E value = 1.1E-2).

Histone H2A: 27 genes with BLASTp scores ~ 180, E values ~ 1E-57 to 7E-59.
Cre06.g264750 (HTA14), Cre06.g264950 (HTA12/HTA13), Cre06.g265350 (HTA11/HTA42), Cre06.g266700 (HTA39), Cre06.g268050 (HTA10), Cre06.g268300 (HTA9), Cre06.g271350 (HTA40), Cre06.g273900 (HTA41), Cre06.g274200 (HTA42), Cre06.g274800 (HTA26), Cre06.g275850 (HTA6), Cre06.g276500 (HTA23), Cre06.g276950 (HTA24), Cre12.g504500 (HTA15/HTA16), Cre12.g504750 (HTA43), Cre12.g505550 (HTA17/HTA28), Cre12.g506250 (HTA18), Cre13.g570100 (HTA2), Cre13.g590800 (HTA37), Cre13.g591150 (HTA38), Cre17.g708550 (HTA20), Cre17.g709200 (HTA3/HTA30), Cre17.g710400 (HTA19), Cre17.g711700 (HTA5), Cre17.g713400 (HTA22/HTA25), Cre17.g714100 (HTA21), Cre17.g714500 (HTA1).

Histone H2A variants: three genes with BLASTp scores 115-126, E values 1.5E-29 to 6.7E-37. Cre13.g567700 (HAV), Cre06.g278088 (HAV1), Cre05.g241634 (HAV2).

Histone H2B: 27 genes with BLASTp scores ~ 156, E values ~ 1E-48.
Cre06.g264800, (HTB14), Cre06.g264900, (HTB13), Cre06.g265400, (HTB11/HTB12/HTB42), Cre06.g266750, Cre06.g268100, (HTB10), Cre06.g268250, (HTB9), Cre06.g271376, Cre06.g273850, (HTB7), Cre06.g274250, Cre06.g274750, (HTB26/HTB6), Cre06.g275800, (HTB31/HTB4), Cre06.g276550, (HTB23), Cre06.g276900, (HTB24), Cre12.g504550, (HTB15), Cre12.g504700, (HTB16/HTB18/HTB28), Cre12.g505600, (HTB17), Cre12.g506200, Cre13.g570050, (HTB2), Cre13.g590750, (HTB37), Cre13.g591200, (HTB38), Cre17.g708600, (HTB20), Cre17.g709150, (HTB3/HTB30), Cre17.g710450, (HTB44), Cre17.g711750, (HTB22/HTB5), Cre17.g713450, Cre17.g714050, (HTB21), Cre17.g714550, (HTB1).

Histone H2B variants: one gene with BLASTp score 152, E value = 2.1E-47. Cre01.g062172 (HBV1).

Histone H3: 32 genes with BLASTp scores 233-236, E values ~ 1E-80.
Cre17.g714650 (HTR1/HTR47), Cre06.g267950 (HTR10), Cre06.g266650 (HTR11/HTR45), Cre06.g275750 (HTR12/HTR4), Cre06.g265000 (HTR13), Cre12.g504650 (HTR15/HTR28), Cre12.g504800 (HTR16), Cre12.g505500 (HTR17), Cre12.g506300 (HTR18), Cre17.g711850 (HTR19/HTR5), Cre13.g569950 (HTR2), Cre17.g708700 (HTR20), Cre17.g713950 (HTR21), Cre17.g713550 (HTR22), Cre06.g274850 (HTR23/HTR26/HTR6), Cre06.g276850 (HTR24), Cre06.g264650 (HTR27), Cre06.g274350 (HTR29), Cre17.g709050 (HTR3), Cre12.g506500 (HTR35), Cre17.g708150 (HTR36), Cre16.g650300 (HTR39), Cre16.g649900 (HTR40), Cre16.g648500 (HTR41), Cre06.g274000 (HTR7), Cre06.g271250 (HTR8), Cre06.g268350 (HTR9), Cre06.g265250, Cre06.g265500, Cre06.g274101, Cre06.g276600, Cre17.g710550.

Histone H3 variants: three genes with BLASTp scores 118-139, E values ~ 3E-33 to 1E-41. Cre02.g104800 (HTV1), Cre03.g197050 (HTV2), Cre16.g661450 (HTV3).

Histone H4: 32 genes with BLASTp scores 157-159, E values 4E-50 to 6E-51.
Cre06.g264600 (HFO14), Cre06.g265050 (HFO13), Cre06.g265200 (HFO34/HFO46), Cre06.g265450 (HFO12), Cre06.g266600 (HFO11/HFO45), Cre06.g268000 (HFO10), Cre06.g268400 (HFO9), Cre06.g271300 (HFO8), Cre06.g273950 (HFO7), Cre06.g274150 (HFO31/HFO32/HFO43), Cre06.g274300 (HFO29), Cre06.g274900 (HFO26/HFO6), Cre06.g275700 (HFO4), Cre06.g276650, Cre06.g276800 (HFO24), Cre12.g504600 (HFO15), Cre12.g504850 (HFO16), Cre12.g505450 (HFO17), Cre12.g506350 (HFO18), Cre12.g506450

(HFO35), Cre13.g570000 (HFO2), Cre16.g648550 (HFO41), Cre16.g649950 (HFO40), Cre16.g650250 (HFO39), Cre17.g708200 (HFO36), Cre17.g708650 (HFO20), Cre17.g709100, Cre17.g710500, Cre17.g711800, Cre17.g713500 (HFO22/HFO25), Cre17.g714000 (HFO19/HFO21/HFO3), Cre17.g714600.

Fermentation genes: all genes are based on published pathways for pyruvate metabolism (15, 18). Cre07.g324550 (LDH1), Cre01.g044800 (PFL1), Cre02.g095137 (PFR1), Cre03.g199800 (HYDA1), Cre09.g396600 (HYDA2), Cre06.g296750 (HYDEF), Cre06.g296700 (HYDG), Cre09.g391650 (HCP4), Cre17.g746997 (ADH1), Cre09.g396650 (PAT2), Cre09.g396700 (ACK1).

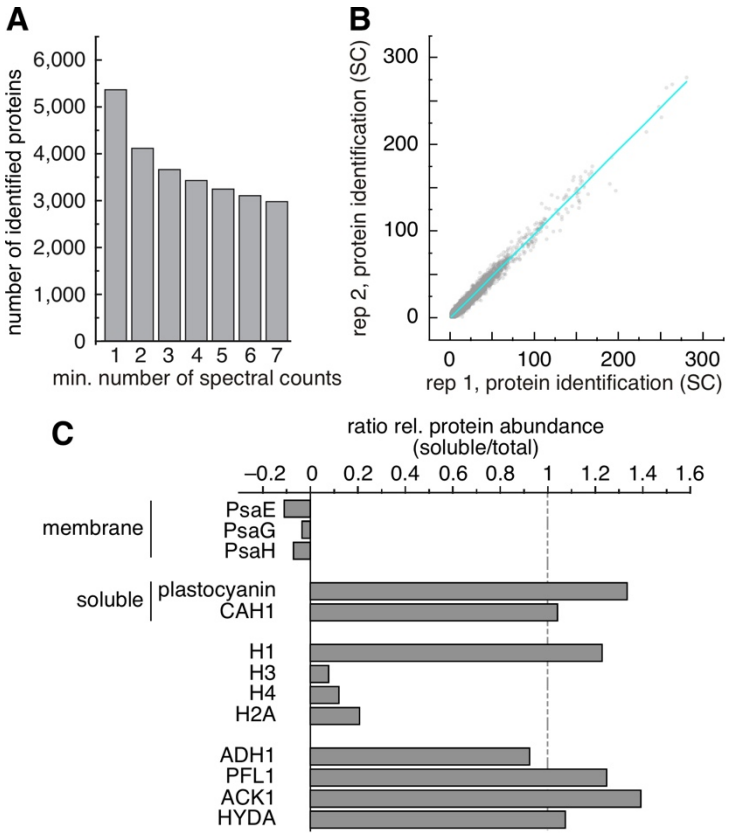
Ferredoxins: the Chlamydomonas encodes 13 ferredoxins (*PETF1/FDX1*, *FDX2-FDX12*, and *MFDX*) (24).

Cre14.g626700 (PETF/FDX1), Cre16.g658400 (FDX2), Cre06.g306350 (FDX3), Cre07.g334800 (FDX4), Cre17.g700950 (FDX5), Cre03.g183850 (FDX6), Cre01.g006100 (FDX7), Cre01.g005600 (FDX8), Cre12.g487900 (FDX9), Cre12.g559950 (FDX10), Cre06.g291650 (FDX11), Cre08.g374550 (FDX12), Cre12.g559950 (MFDX).

Quantitative extraction of proteins in soluble and total fractions

In order to complement our transcriptome analysis, we conducted an untargeted proteomics investigation of protein contents in diurnal cells. We collected samples in duplicate at six time-points over the diurnal cycle (CT–11, –5, –1, +1, +5, +11) from cultures grown in photo-bioreactors in the same settings as for RNAseq samples. We extracted soluble proteins by repeated freeze-thaw cycles, which we use routinely in the laboratory for the quantitative retrieval of plastocyanin, an abundant soluble protein (35).

Abundant soluble proteins were identified by LC-MS/MS and quantified using the software MASIC (36). To test for quantitative extraction of proteins, we also analyzed total proteins from one sample collected at CT+1. Efficiency of extraction by the freeze-thaw method was calculated as the ratio of Z-score normalized protein abundance for soluble and total protein samples. We identified a total of 3,163 soluble proteins with at least two peptides at a single time-point, and 1,934 soluble proteins with at least one peptide at each time-point (SOM Fig. 3A). Reproducibility between replicates is excellent (SOM Fig. 3B).



SOM Fig. 3. Quantitative recovery of soluble proteins by untargeted proteomics.
Data are based on biological duplicates collected from independent photo-bioreactors.
A, Summary of number of spectral counts and number of identified proteins across all time-points.
B, Assessment of reproducibility based on spectral counts detected for all proteins between replicates at each time point. Cyan line: linear regression fit.
C, Quantitative recovery of selected proteins in the soluble fraction, compared to total cell lysate. Extent of recovery is shown as the ratio between z-score normalized protein abundance (determined using the software MASIC) between the soluble proteome and total proteome at time +1.

Quantitative recovery of proteins: soluble vs total protein fraction.

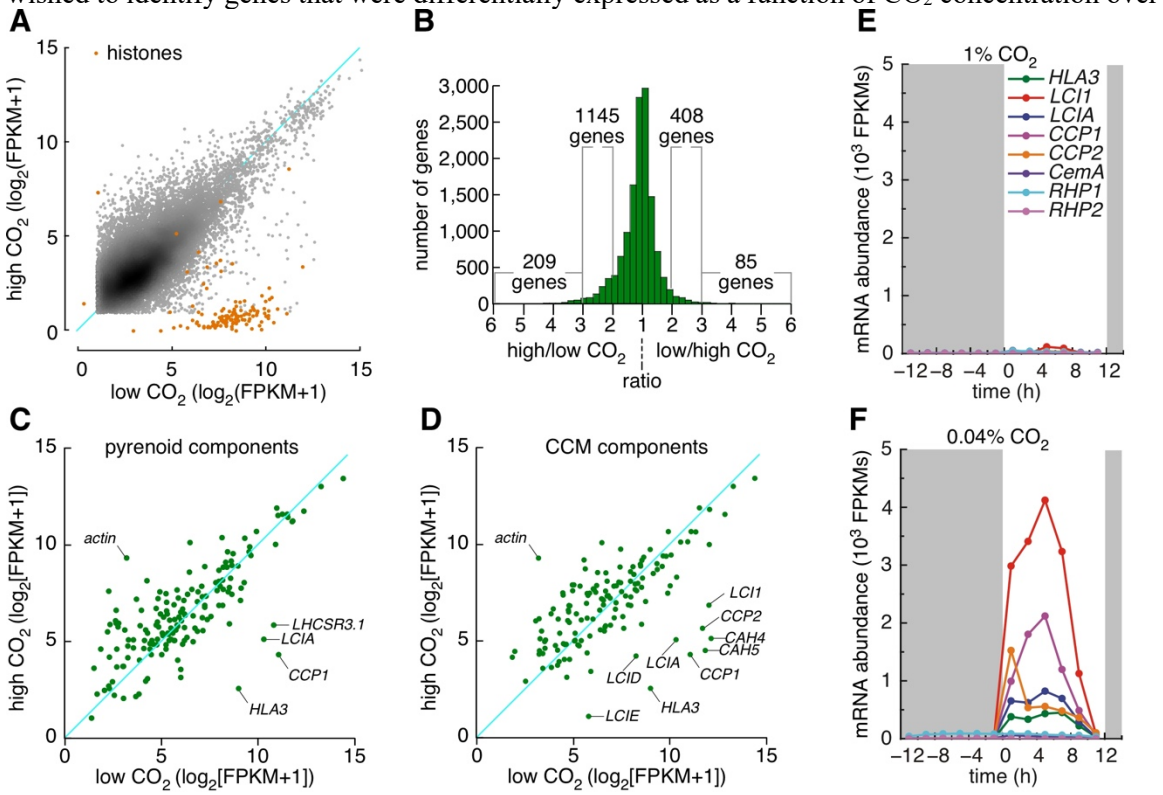
Under our conditions, known soluble proteins plastocyanin and carbonic anhydrase 1 (CAH1) are recovered quantitatively following freeze-thaw cycles (SOM Fig. 3C). As expected, plastid-localized, membrane proteins PsaE, PsaG and PsaH are poorly recovered in the soluble fraction.
Of all histones, only the H1 histone linker is quantitatively recovered, preventing us from determining the protein accumulation profile of histones H2A, H3 and H4 in our samples. Finally, enzymes involved in the fermentation of pyruvate ADH1, ACK1, PFL1 and HYDA2 are recovered in our soluble fractions. Since our total protein sample was collected at CT+1, when MCM and RNR proteins are not expressed, we cannot ascertain their quantitative retrieval in the soluble fraction. However, MCM proteins are known soluble proteins in other systems, and are expected to be released from the nucleus.

Quantification across the diurnal cycle.

The number of peptides detected for each time-point is derived from the analysis of a constant quantity of digested proteins. However, cells grow and accumulate biomass in the light, and this should include proteins. Indeed, protein content on a per cell basis does increase in the light, during which time is approximately doubles, as expected (Supplementary Fig. 3A). Protein content per cell remains constant at night, consistent with their quiescent stage. This needs to be kept in mind for normalization and comparison purposes across samples: if a protein is present as the same number of entities at CT+1 and CT+11, its relative protein abundance will in fact decrease by 50%. If a protein is present at half the number of entities at CT+11 relative to CT+1, then its relative protein abundance at Z+T11 will decrease by 75%. This only applies to light samples (CT+1, +5 and +11).

Impact of CO_2 concentration on CCM complex genes

Our growth conditions included bubbling with air levels of CO_2 (0.04%). In a previous study, Zones *et al*. (34) collected samples over one diurnal cycle while bubbling cultures with 1% CO_2. We wished to identify genes that were differentially expressed as a function of CO_2 concentration over



SOM Fig.4. Comparison of gene expression profiles over one diurnal cycle of Chlamydomonas grown with 1% CO_2 or 0.04% CO_2.

A, Scatterplot of maximum gene expression estimates (as $\log_2$[FPKM+1]) for all expressed genes (FPKM >1 in at least one time-point). Most histone genes (indicated as orange dots in the plot) are not poly-adenylated and appear strongly induced in our dataset.

B, Distribution of maximum expression ratio between datasets. The low/high CO_2 comparison is shown on the right side of the distribution, while the high/low CO_2 comparison is shown on the left side. Ratios are based on the $\log_2$[FPKM+1] referred to above.

Scatterplot of maximum gene expression estimates for genes encoding components of the pyrenoid (**C**), or to the CCM complex (**D**), between low and high CO_2 samples.

E, **F**, Relative expression levels, in FPKM, for selected genes in high (**E**) and low (**F**) CO_2.

the diurnal cycle. To this end, we modified the two datasets in two ways: first, we remapped all sequenced reads from our dataset to the Chlamydomonas nuclear genome only, to remove any reads derived from the organellar genomes (which represent ~40% of all reads, and affect FPKM values by changing the denominator for normalization). Next, we removed histones from both datasets, as they would appear highly induced in our dataset. Most histone transcripts are not poly-adenylylated, but there is sufficient bleed through while performing polyA selection during library preparation that many are deemed expressed with a cut-off of FPKM >1 (SOM Fig. 4a).
Of 17,617 non-histone genes, 13,702 were expressed in both datasets (at least one sample over the time-course has an FPKM value of 1 or higher). Using a cut-off of 2 ($\log_2$ ($FPKM_{low\ CO2}$ / $FPKM_{high\ CO2}$)), 493 genes are induced under low CO_2, while 1,354 genes were induced under high CO_2 (SOM Fig. 4B).
Most genes encoding proteins identified in the pyrenoid (36) or associated with the CCM complex maintain a relatively constant expression independent of CO_2 status (SOM Fig. 4C and 4D). Our analysis revealed known genes involved in the carbon concentrating mechanism such as *LCI* genes (*LCI1*, *LCIA*, *LCID* and *LCIE*), genes encoding the low CO_2-inducible chloroplast membrane proteins *CCP1* and *CCP2*, and carbonic anhydrases *CAH4* and *CAH5*. Our dataset now also allows the comparative analysis of the diurnal expression pattern of genes involved in the CCM, as exemplified in SOM Fig. 4E and 4F. Under our conditions of low CO_2 (0.04%), CCM genes clearly appear diurnally-rhythmic, with peak expression in the middle of the day.

Calculating respiration-derived and fermentation-derived ATP production in the dark

We estimate the amount of starch consumed at night to be ~ 15 µg per $1x10^6$ cells (Fig. 5B).
To respire this amount of starch over 12 hr during the night, cells would need to respire $0.5x10^6$ moles O_2 per $1x10^6$ cells:

15 µg starch = $15x10^{-6}$ gram hexoses / 180 (gram/mole hexose)
= $0.83x10^{-7}$ moles starch
6 O_2 molecules needed per hexose (one O_2 per C)
Total oxygen consumed at night:
$0.83x10^{-7}$ x 6 = $5x10^{-7}$ = $0.5x10^{-6}$ moles O_2 per $1x10^6$ cells

We estimate actual respiration rate to be ~ $0.2x10^{-9}$ mole O_2 per minute per $1x10^6$ cells (Fig. 5C).
Over 12 hr (= 720 minutes)
Total night respiration = 720 x $0.2x10^{-9}$ mole O_2 per $1x10^6$ cells
= $1.44x10^{-7}$ mole O_2 per $1x10^6$ cells
= $0.144x10^{-6}$ mole O_2 per $1x10^6$ cells
= 29% of total hypothetical oxygen consumed for respiration of 15 µg starch per $1x10^6$ cells.

Therefore:
29% of the starch consumed at night goes to respiration and generates 2x38 ATP per glucose molecule (38 ATP per pyruvate)
71% of the starch consumed at night goes to fermentation and generates 2x3 ATP per glucose molecule (3 ATP per pyruvate)
Respiration contribution for energy production: 2 x 38 x 29% = 22.04 (au)
Fermentation contribution for energy production: 2 x 3 x 71% = 4.26 (au)
Respiration contribution = 5.17 x Fermentation contribution

Gene discovery based on expression pattern

To facilitate the identification of genes that are co-expressed with genes of interest, we turned our expression dataset into a gene co-expression network, essentially following published methods (38, 39). We first normalized gene expression by calculating $\log_2$(FPKM+1) for all genes (nucleus-encoded and organellar). In a second step, we applied a quantile normalization with the *preprocessCore* package in R. Finally, we normalized expression of each gene by subtracting the quantile-normalized value at each time-point by the mean across the time course.

We used the mean-centered values to determine Pearson's correlation coefficient for all gene pairs with the function *cor* in R, and converted them into ranks. At this stage, each gene is characterized by its ranking relative to all other genes in the genome, and genes that are co-expressed will have small rank values. We then calculated the mutual rank (MR) for all gene pairs, where $MR_{a,b} = rank_{a,b} \times rank_{b,a}$. In the final step, MRs are converted to edge weights (between genes) with the formula $edge = e^{-(MR-1)/5}$. Edge values equal to or higher than 0.01 were considered significant. This approach is somewhat crude, as it does not directly take the temporal expression pattern of genes into account, but appears effective in detecting genes with distinctive expression patterns (Supplementary Fig. 6, 12).

Dataset S3 compiles all gene pairs that are significant (with edge ≥ 0.01).

From this table, we extracted co-expressed genes for *PSBS1/2*, *LHCSR1*, *LHCSR3.1/3.2*, plotted in Supplementary Fig. 12.

Supplementary Figures

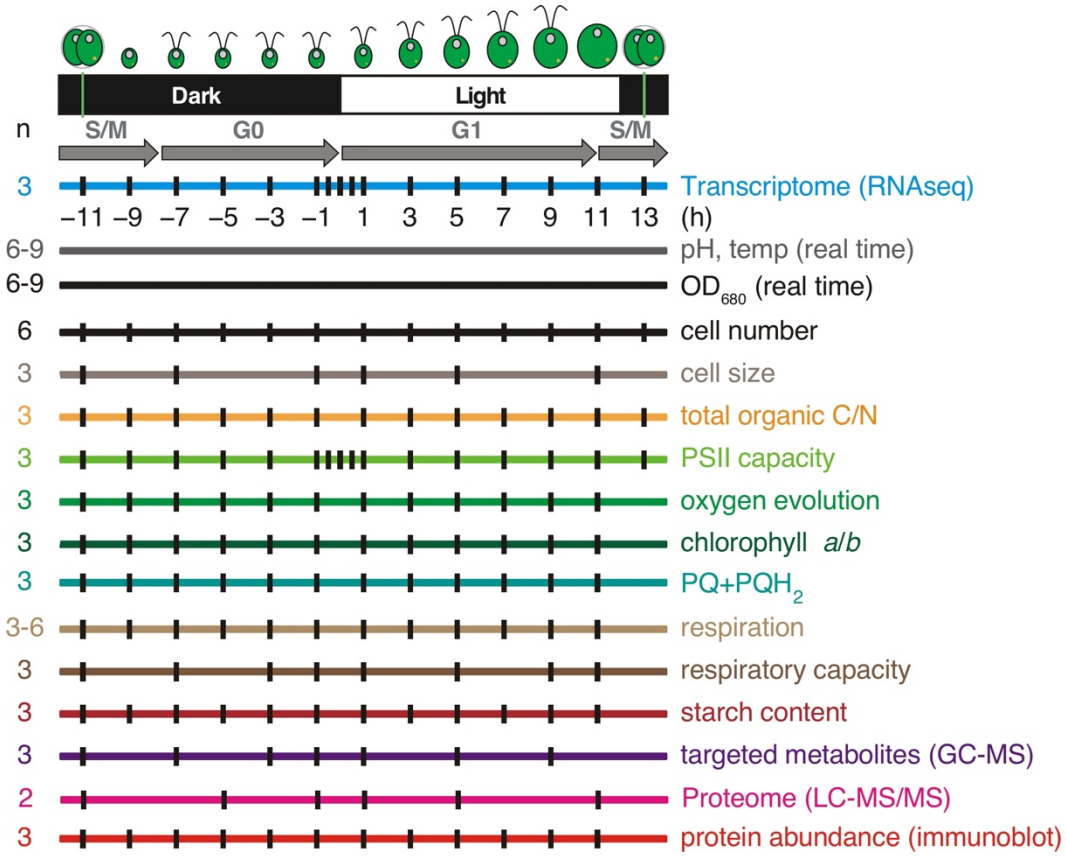


Fig. S1. Overview of the experimental design.
Schematic illustration of the Chlamydomonas cell cycle, adapted from Zones *et al* (34). Each horizontal line denotes an independent experiment, while each vertical line indicates times of sample collection during the diurnal cycle. By convention, dawn (dark-to light transition) is referred to as CT0. The number of biological replicates for each type of experiment is shown to the left of the corresponding horizontal line.
The green vertical lines indicate the timing of cell division.

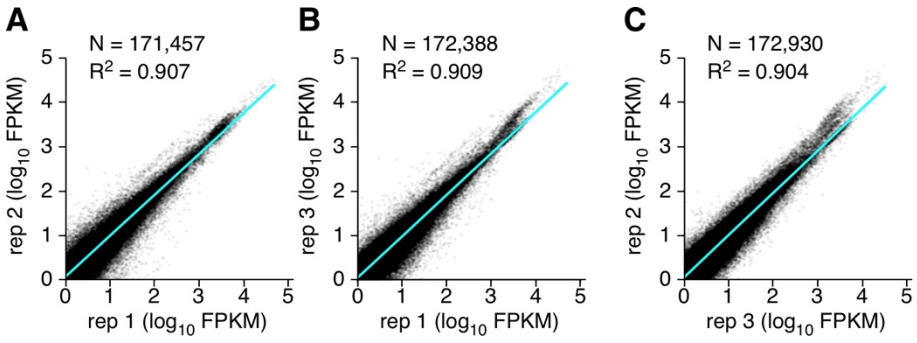


Fig. S2. Excellent reproducibility across replicates.
For each of our sixteen time-points, we collected cells from three independent photo-bioreactors and processed their RNA individually for RNAseq library preparation. To evaluate reproducibility, we compared transcript abundance (expressed as $\log_{10}$(FPKM)) between replicates for all time-points, resulting in over 170,000 pair-wise comparisons. Genes with expression levels < 1 FPKM (across all time-points) were removed from the analysis. The resulting scatterplots are plotted as 2D density plots; cyan line: linear regression fit to the data (with associated number of pair-wise comparisons and correlation coefficient R^2 given in each panel).
A. Comparison replicate 1 and replicate 2.
B. Comparison replicate 1 and replicate 3.
C. Comparison replicate 3 and replicate 2.

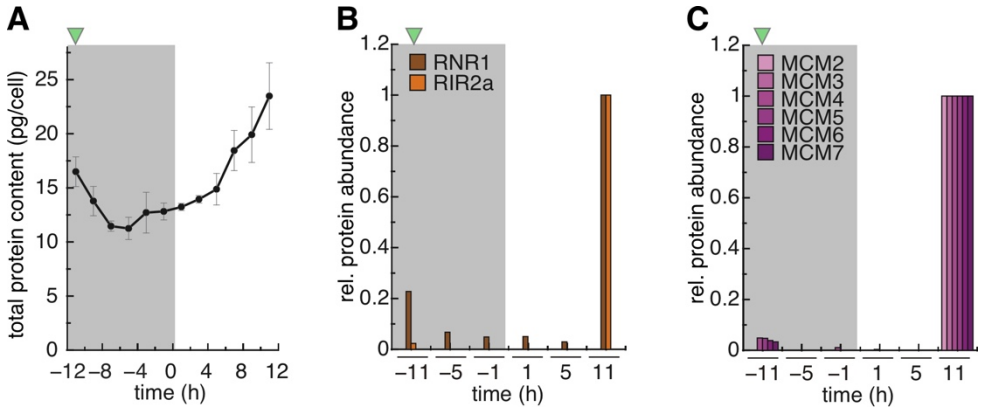


Fig. S3. Accumulation of proteins involved in DNA replication is restricted to the end of the day.

A, Total protein content in cells over the diurnal cycle.

Relative protein abundance for RNR1 and RIR2a **(B)** and MCM complex subunits **(C)**. Relative protein abundance is normalized to total protein content shown in **A**. The sample with highest relative protein abundance was set to 1. Data are from two independent experiments; quantification of all proteins can be found in Dataset S2.

The green triangles indicate the timing of cell division.

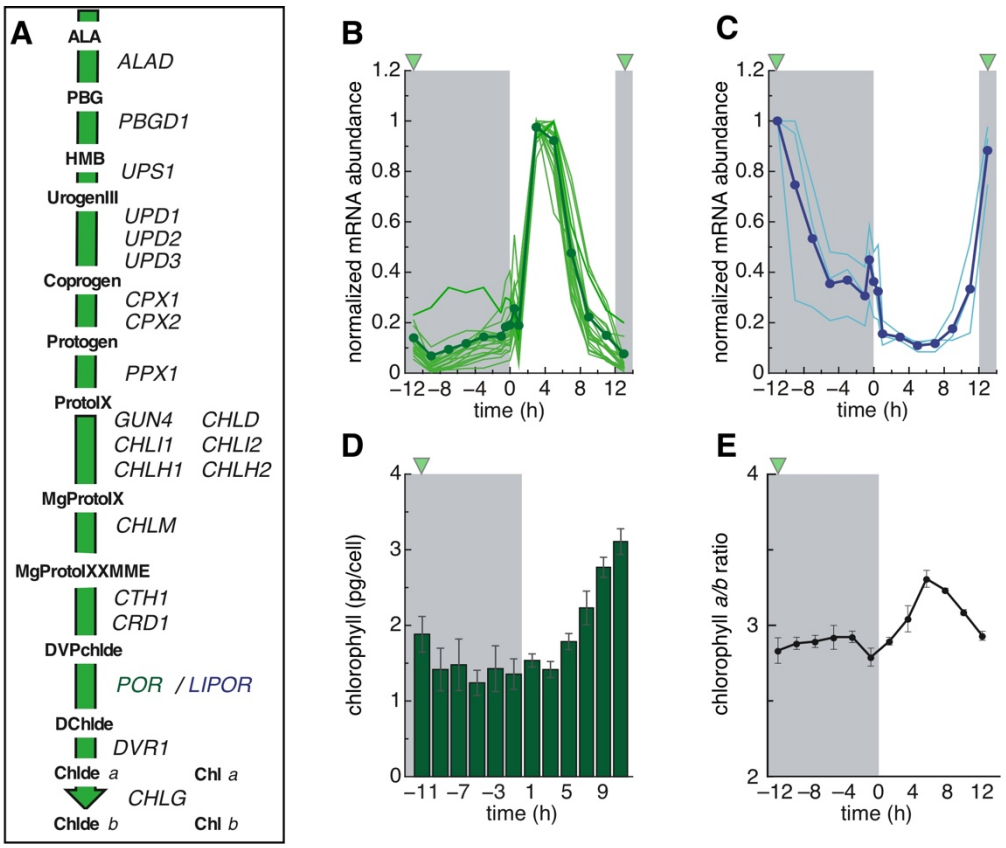
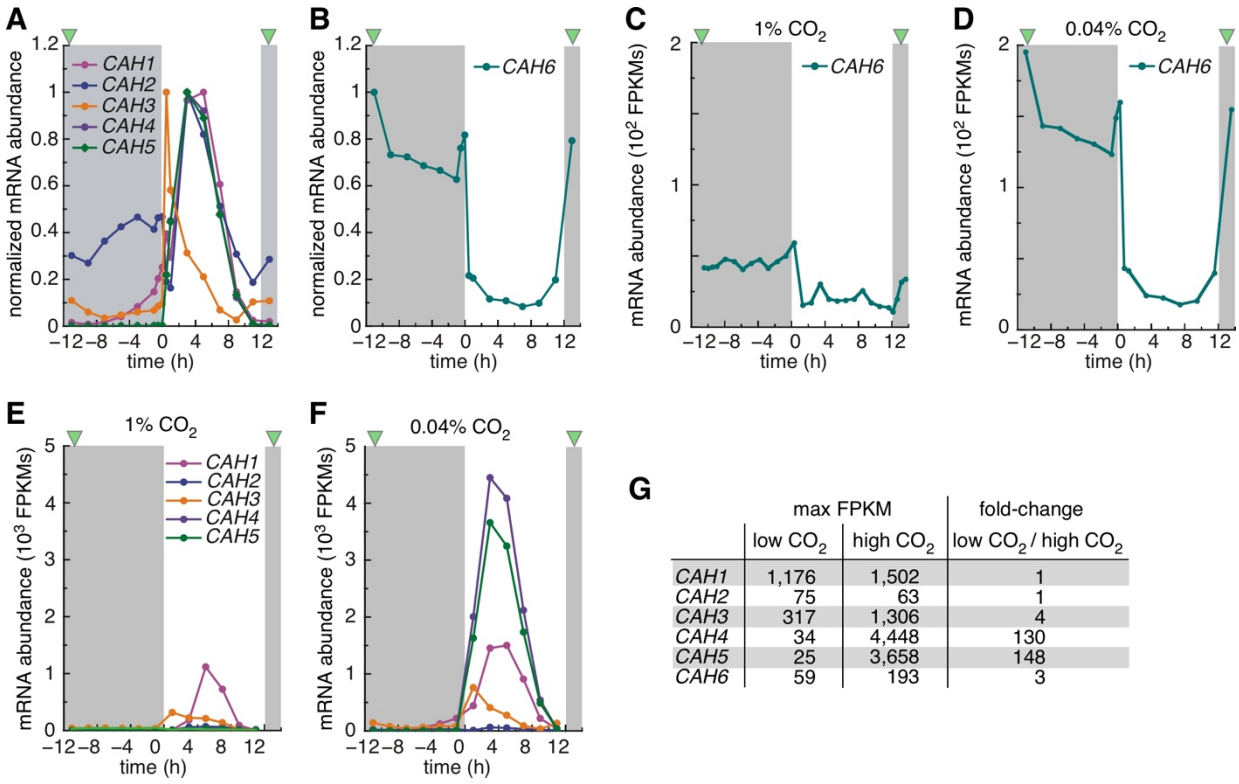


Fig. S4. Most chlorophyll biosynthetic genes are co-expressed.
A, Overview of the tetrapyrrole biosynthetic pathway.
B, mRNA abundance (normalized to 1 for the highest value) for all genes encoding enzymes listed in **A**, with the exception of LIPOR (light-independent protochlorophyllide oxidoreductase), shown in **C**. Thin lines: individual genes; thick lines with symbols: average of all genes for each panel.
D, Chlorophyll content in Chlamydomonas cells over the diurnal cycle. Chlorophyll *a* and *b* were measured spectrophotometrically after acetone/methanol extraction.
E, Chlorophyll *a/b* ratio.
The green triangles indicate the timing of cell division.



	max FPKM		fold-change
	low CO_2	high CO_2	low CO_2 / high CO_2
CAH1	1,176	1,502	1
CAH2	75	63	1
CAH3	317	1,306	4
CAH4	34	4,448	130
CAH5	25	3,658	148
CAH6	59	193	3

Fig. S5. Distinct expression patterns of the carbonic anhydrase gene family as a function of CO_2 concentration.

mRNA abundance (normalized to 1 for the highest value) for *CAH1-CAH5* (**A**) and *CAH6* (**B**). Carbonic anhydrase transcript levels, in FPKMs, over the diurnal cycle when grown in 1% CO_2 ((34), **C, E**) or 0.04% CO_2 (this study, **D, F**). *CAH6* expression (**C**, **D**), *CAH1-CAH5* expression (**E**, **F**).

G, Comparison of *CAH* gene expression estimates, in FPKMs, across the two studies. Note that only time-points common to both experiments were considered.

The green triangles indicate the timing of cell division.

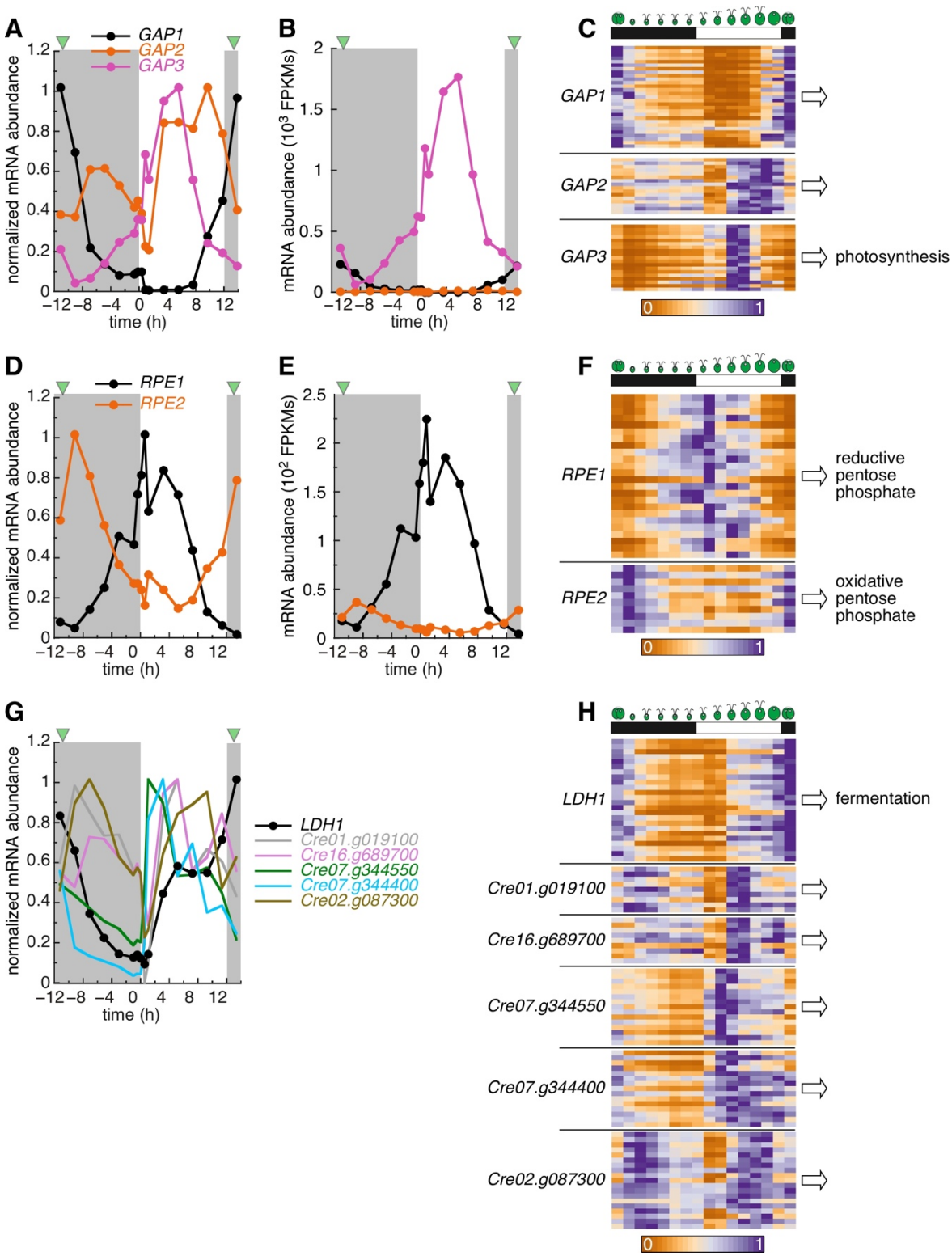


Fig. S6. Gene function predictions based on expression and co-expression patterns. Shown here are mRNA abundance (normalized to 1 for the highest value, or as FPKMs) for a number of gene families with distinct expression patterns. The green triangles indicate the timing of cell division.

Normalized mRNA abundance (**A**) and transcript abundance in FPKMs (**B**) for members of the glyceraldehyde dehydrogenase (*GAPDH*) family. **C**, Expression estimates for *GAP* co-expressed genes, shown as heatmaps, and putative cellular function of the respective *GAP* based on co-expressed genes.
Normalized mRNA abundance (**D**) and transcript abundance in FPKMs (**E**) for members of the ribulose phosphate-3-epimerase (*RPE*) family. **F**, Expression estimates for *RPE* co-expressed genes, shown as heatmaps, and putative cellular function of the respective *RPE* based on co-expressed genes.
Normalized mRNA abundance (**G**) for lactate dehydrogenase candidates. **H**, Expression estimates for lactate dehydrogenase candidates co-expressed genes, shown as heatmaps, and putative cellular function of the respective dehydrogenase based on co-expressed genes.

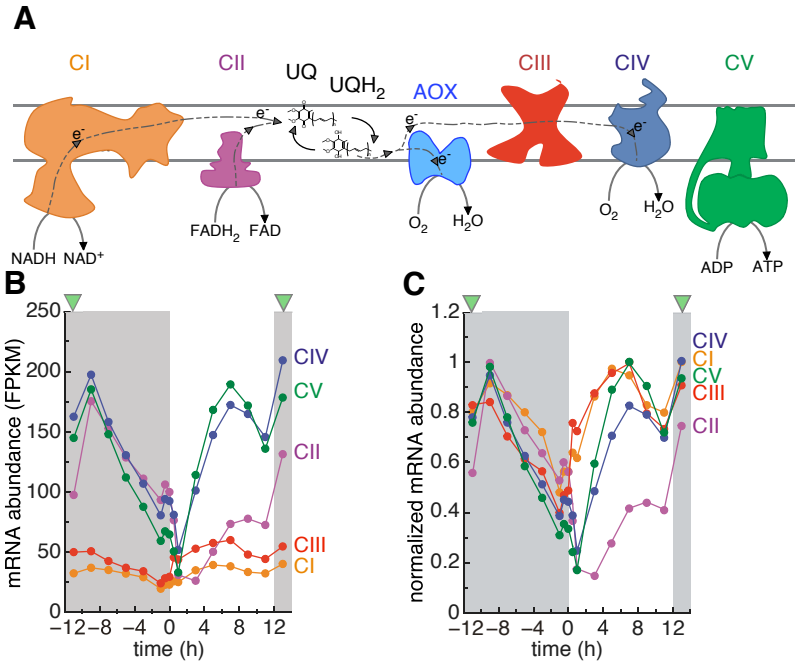


Fig. S7. Bimodal expression of complexes IV and V of the mitochondrial electron transport chain.

A, Schematic illustration of complexes (CI through CV) of the mitochondrial respiratory electron transport chain. UQ: ubiquinone; UQH2: ubiquinol.

B, Expression of nucleus-encoded subunits for CI to CV complexes over the diurnal cycle. Shown are average abundance (in FPKM) for all subunits within each complex. CI: orange; CII: purple; CIII: red; CIV: blue; CV: green.

The green triangles indicate the timing of cell division.

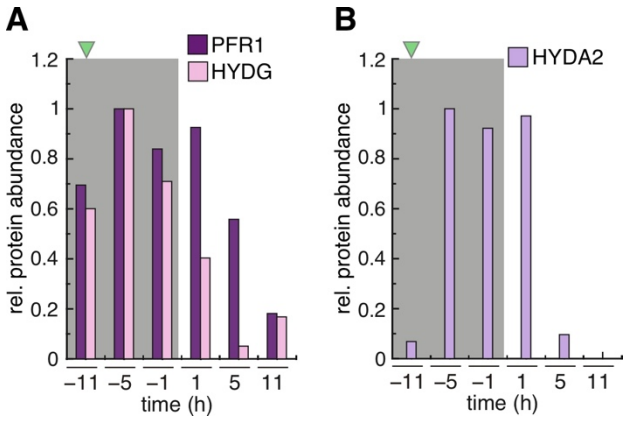


Fig. S8. Accumulation of proteins involved in pyruvate fermentation at night.
Relative protein abundance of fermentation enzymes PFR1, HYDG and HYDA2 in soluble fractions over the diurnal cycle. Relative protein abundance is normalized to total protein content shown in Supplementary Fig 3**A**. The sample with highest relative protein abundance was set to 1. Data are from two independent experiments; quantification of all proteins can be found in Dataset S2.
The green triangles indicate the timing of cell division.

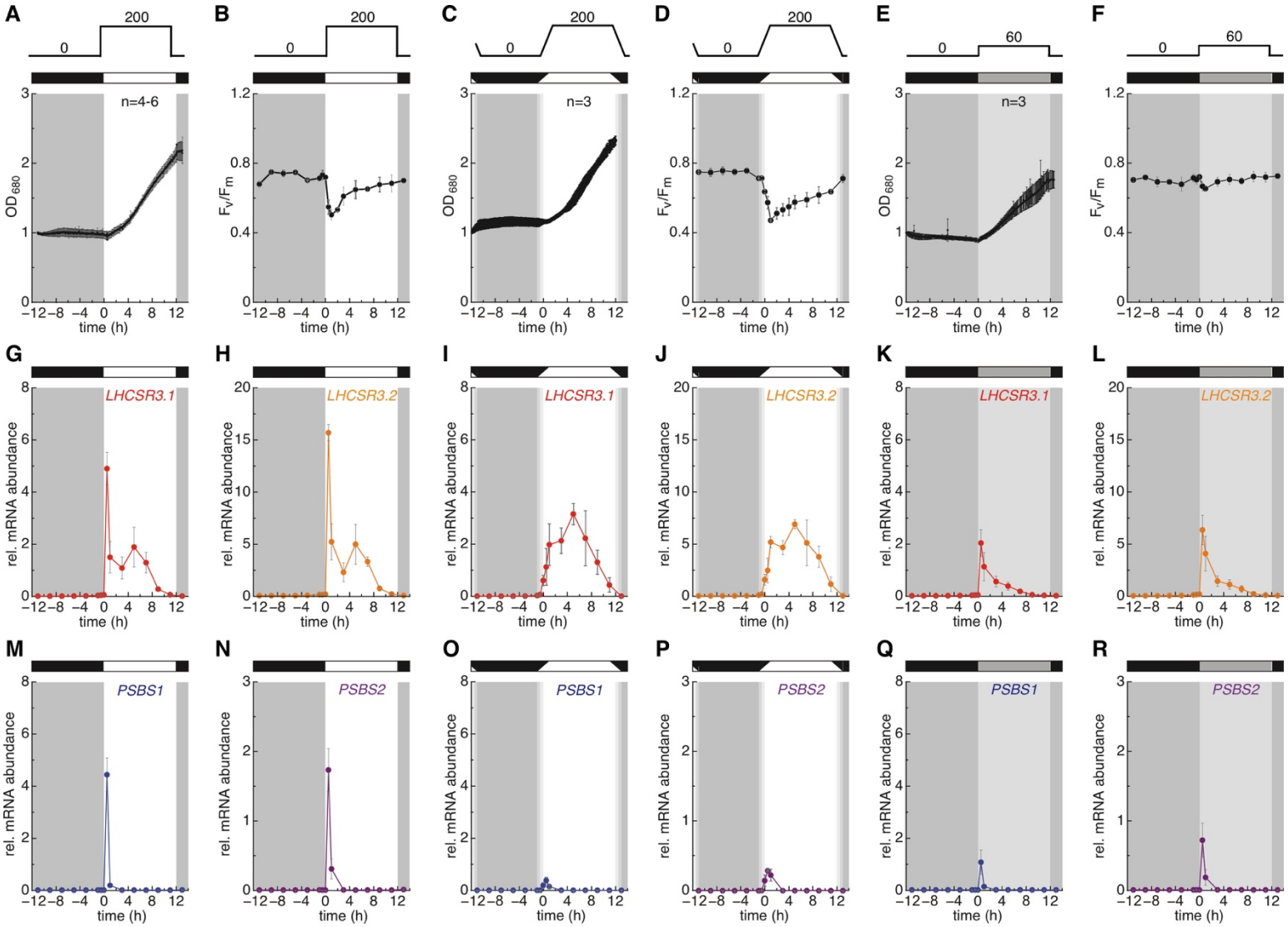

A
B
C
D
E
F
0
200
60
n=4-6
n=3
OD680
Fv/Fm
time (h)
G
H
I
J
K
L
LHCSR3.1
LHCSR3.2
rel. mRNA abundance
M
N
O
P
Q
R
PSBS1
PSBS2

Fig. S9. Evidence for two light signals driving expression of LHC-like stress response proteins.

Cultures were grown in photo-bioreactors under one of three light regimes, depicted on top of each panel for reference: 200 µmol photons /m^2/s during the light period with abrupt dawn and dusk transitions (**A**, **B**, **G**, **H**, **M**, **N**); 200 µmol photons /m^2/s during the light period with a 2 hour linear light gradient at each transition (**C**, **D**, **I**, **J**, **O**, **P**); 60 µmol /m^2/s during the light period with abrupt dawn and dusk transitions (**E**, **F**, **K**, **L**, **Q**, **R**).

A, **C**, **E**, Optical density of cell cultures at 680 nm.

B, **D**, **F**, Photosystem II capacity (Fv/Fm) values over the diurnal cycle.

Relative mRNA abundance of *LHCSR3.1* (**G**, **I**, **K**), *LHCSR3.2* (**H**, **J**, **L**), *PSBS1* (**M**, **O**, **Q**) and *PSBS2* (**N**, **P**, **R**) under all three light regimes, determined by quantitative RT-PCR. Results are shown as averages ± s.d. (n=3).

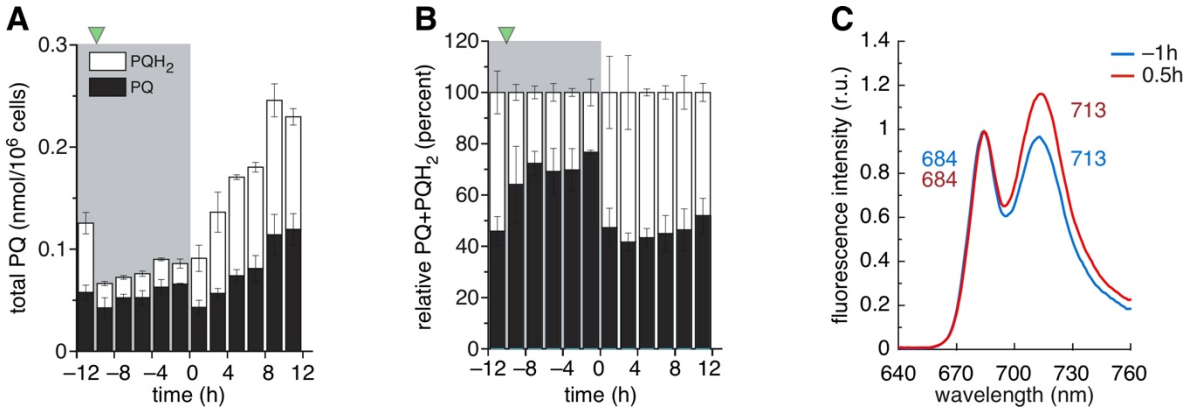


Fig. S10. Accumulation of reduced plastoquinone in the light and 77ºK chlorophyll fluorescence.

Plastoquinone (PQ, black bars) and reduced plastoquinone (plastoquinol, PQH_2, white bars) levels over the diurnal cycle, shown as absolute amounts per cell (**A**) or as percentage of total plastoquinone levels at each time point (**B**). Data shown as average ± s.d. (n=3).

C, Chlorophyll fluorescence at 77ºK before and after dawn. Chlorophyll fluorescence was recorded between 640 nm and 760 nm in samples collected 1 hour before (–1) and 30 min after (+0.5) dawn. Excitation wavelength was 435 nm. Fluorescence was normalized to the value at 685 nm. This experiment was performed twice with similar results.

The green triangles indicate the timing of cell division.

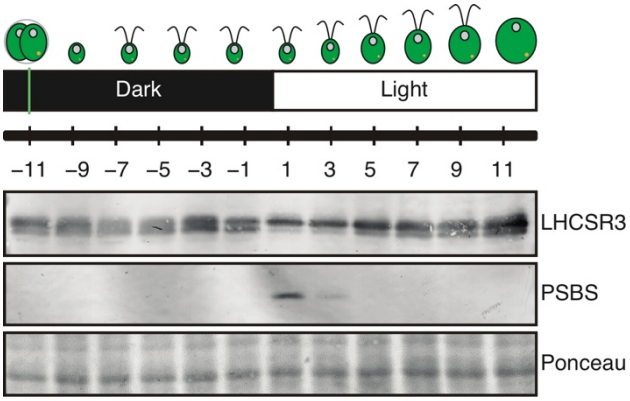


Fig. S11. LHCSR3 and PSBS accumulation is dependent on light.
Total protein from samples covering the diurnal cycle were separated by denaturing SDS-PAGE and interrogated with antibodies raised against LHCSR3 and PSBS. Equal protein amounts were loaded, and confirmed with Ponceau S stain. This experiment was performed three times with similar results.
The green vertical line indicates the timing of cell division.

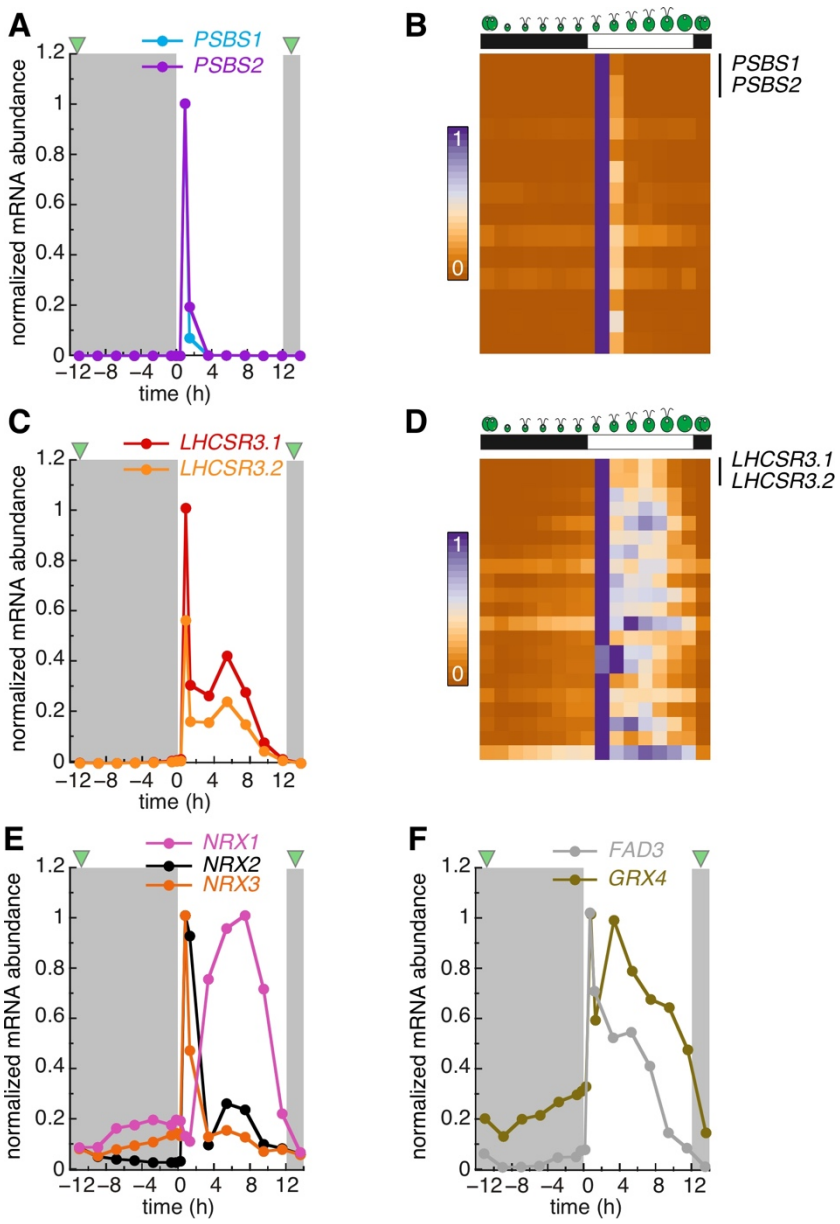


Fig. S12. *PSBS* and *LHCSR* co-expressed genes.
A, *PSBS1* and *PSBS2* expression (normalized to 1 for the highest value) over the diurnal cycle.
B, Expression profile of genes that belong to the *PSBS1* and *PSBS2* co-expression network, shown as a heatmap.
C, *LHCSR3.1* and *LHCSR3.2* normalized expression profile over the diurnal cycle.
D, Expression profile of genes that belong to the *LHCSR3.1* and *LHCSR3.2* co-expression network, shown as a heatmap.
E, Normalized expression profile for nucleoredoxin genes *NRX1-NRX3*. Note the sharp increase in mRNA levels at dawn for *NRX2* and *NRX3*.
F, Normalized expression profile for the fatty acid desaturase *FAD3* and the glutaredoxin *GRX4*. Note the two peaks in expression: one right after dawn (acute response) and a later peak, similar to *LHCSR3* genes.
The green triangles indicate the timing of cell division.

Table S1. Relative representation of selected mRNAs.

	nucleus-encoded				organellar-encoded				
functional categories[1]	number of genes	% transcripts [2]	% transcripts at peak [3]	peak (h)	number of genes	% transcripts [2]	% transcripts at peak [3]	peak (h)	~protein (units/cell)
chloroplast ETC	66	13.8	29.8	+7	33	81.5	89.5	+9	~ $3x10^6$
CBB cycle	21	3.6	5.4	+3	1	10.3	12.2	-1	
tetrapyrrole syn.	27	0.4	0.9	+3	3	0.3	0.7	-11	
Total ribosomes	152	34.2	45.6	+1					~ $2x10^6$
80S	76	32.8	42.1	+1					
70S cp	38	1.4	3.4	+1	20	5.1	8.8	+1	
70S mt	38	0.1	0.1	+1					
nucleosomes	100	1.2	6.1	+11					~ $0.5x10^6$
mitochondrial ETC	96	1.4	2.4	-9	7	97.6	98.9	-1	
TCA + glx cycle	26	0.7	1.9	-9					

[1] abbreviations are: Electron transfer chain (ETC), Calvin-Benson-Bassham (CBB), synthesis (syn), chloroplast (cp), mitochondria (mt), TriCarboxylic Acid (TCA), glyoxylate (glx).

[2] percentage of mean expression over the time-course.

[3] percentage at time of maximal gene expression.

Table S2. Comparison of this dataset with all other published data collected over a diurnal cycle.

observation	peak †	peak † (this study)	photon flux density (µmol /m^2/s)	% CO_2	temperature (ºC)	light-dark cycle	reference
MCM		13 (~0-58)	200	0.04	18: 28	12L: 12D	this study
PSI		5-8					16 time-points
histones		12-15 (4-2,118)					(3 replicates)
whole transcriptome	*see SOM Fig.* 2		250	1	28	12L: 12D	(34)
MCM	11-12 (~0-88)	13 (~0-58)					28 time-points
PSI	2-7	5-8					(2 replicates)
histones	11-15 (~0-3) [1]	12-15 (4-2,118)					
MCM	15	13 (~0-58)	195	NA	33 [2]	14L: 10D	(17)
PSI	1-5	5-8					11 time-points
histones	NA (0-61) [3]	12-15 (4-2,118)					(1-3 replicates)
MCM	nd	13 (~0-58)	NA	5	24	12L: 12D	(16)
PSI	NA	5-8					6 time-points
histones	nd	12-15 (4-2,118)					(6 replicates)
rpo	–2-2	–7					
psa	2-6	6					
psbB	6-10	6					
histones	14 [4]	12-15	~67 *	5	23	14L: 10D	(24)
starch	10-12	11	200	2	34	12L: 12D	(23)
starch	12	11	550-900	2	34	12L: 12D	(19)
chlorophyll	11 [5]	11	~100 **	NA	25	12L: 12D	(22)
rbcL	1	1	~117 ***	5	25	12L: 12D	(20)
psbA	6-9	6					
psbD	6-9	6					

chlorophyll	10-12 [5]	11	NA	NA	25	12L: 12D	(18)
chlorophyll	10 [5]	11	~49 ****	NA	25	12L: 12D	(21)

† peak expression time is given from 0-24 (0 referring to dawn, by convention), except when the peak expression window straddles the dark-to light transition, in which case it is reported from –12 to +12.
* estimated, assuming fluorescent light sources, based on a reported 5,000 lux light intensity.
** estimated, based on a reported 700 foot candle light intensity provided by cool white fluorescent lights.
*** estimated, based on a reported 800 foot candle light intensity provided by cool white fluorescent lights.
**** estimated, assuming fluorescent light sources, based on a reported 4,000 lux light intensity.
[1] most histones are not polyadenylylated, and are only barely detectable under classical polyA selection protocols. Peak expression timing is based on the most abundant histone RNAs.
[2] cells were first entrained to light-dark cycles at the constant temperature of 21ºC, then shifted to 31ºC before sample collection.
[3] expression estimates reported as read counts per gene, not FPKM. Peak time could not be reliably estimated from the dataset.
[4] The probe used in the article hybridizes to Cre17.g709050 (histone H3), Cre17.g709100 (histone H4), Cre17.g709150 (histone H2B), Cre17.g709200 (histone H2A).
[5] chlorophyll levels follow an S-shaped profile, with an initial lag from 0-3 hours after dawn, and a leveling off between 10-12, comparable to our own data (Fig. S4).
NA: not available.
nd: not determined.

Supplementary references:

1. Gallaher SD*, et al.* (2018) *High-throughput sequencing of the chloroplast and mitochondrion of* Chlamydomonas reinhardtii *to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates*. Plant J **93**(3):545-565.
2. Dobin A*, et al.* (2013) *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics **29**(1):15-21.
3. Trapnell C*, et al.* (2012) *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks*. Nat Protoc **7**(3):562-578.
4. Elias JE & Gygi SP (2010) *Target-decoy search strategy for mass spectrometry-based proteomics*. Methods Mol Biol **604**:55-71.
5. Kim S & Pevzner PA (2014) *MS-GF+ makes progress towards a universal database search tool for proteomics*. Nat Commun **5**:5277.
6. Helmstetter, C.E., *A ten-year search for synchronous cells: obstacles, solutions, and practical applications.* Front Microbiol, 2015. **6**: p. 238.
7. Helmstetter, C.E. and D.J. Cummings, *Bacterial Synchronization by Selection of Cells at Division.* Proceedings of the National Academy of Sciences of the United States of America, 1963. **50**(4): p. 767-774.
8. Carl, P.L., *Escherichia coli mutants with temperature-sensitive synthesis of DNA.* Mol Gen Genet, 1970. **109**(2): p. 107-22.
9. Ferullo, D.J. and S.T. Lovett, *The stringent response and cell cycle arrest in Escherichia coli.* PLoS Genet, 2008. **4**(12): p. e1000300.
10. Ivleva, N.B., et al., *Quinone sensing by the circadian input kinase of the cyanobacterial circadian clock.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17468-73.
11. Yang, Q., et al., *Circadian gating of the cell cycle revealed in single cyanobacterial cells.* Science, 2010. **327**(5972): p. 1522-6.
12. Craigie, R.A. and T. Cavalier-Smith, *cell volume and the control of the Chlamydomonas cell cycle.* J. Cell Sci., 1982. **54**: p. 173-191.
13. Cross, F.R. and J.G. Umen, *The Chlamydomonas cell cycle.* Plant J, 2015. **82**(3): p. 370-92.
14. Voigt, J. and P. Munzner, *The Chlamydomonas cell cycle is regulated by a light/dark-responsive cell-cycle switch.* Planta, 1987. **172**(4): p. 463-72.
15. Hughes, M.E., J.B. Hogenesch, and K. Kornacker, *JTK_CYCLE: An Efficient Nonparametric Algorithm for Detecting Rhythmic Components in Genome-Scale Data Sets.* Journal of Biological Rhythms, 2010. **25**(5): p. 372-380.
16. Idoine, A.D., et al., *The Diurnal Logic of the Expression of the Chloroplast Genome in Chlamydomonas reinhardtii.* PLoS One, 2014. **9**(10).
17. Tulin, F. and F.R. Cross, *A microbial avenue to cell cycle control in the plant superkingdom.* Plant Cell, 2014. **26**(10): p. 4019-38.
18. Bourguignon, L.Y.W. and G.E. Palade, *Incorporation of Polypeptides into Thylakoid Membranes of Chlamydomonas-Reinhardtii - Cyclic Variations.* Journal of Cell Biology, 1976. **69**(2): p. 327-344.
19. Garz, A., et al., *Cell-to-cell diversity in a synchronized Chlamydomonas culture as revealed by single-cell analyses.* Biophys J, 2012. **103**(5): p. 1078-86.
20. Herrin, D., A. Michaels, and E. Hickey, *Synthesis of a Chloroplast Membrane Polypeptide on Thylakoid-Bound Ribosomes during the Cell-Cycle of Chlamydomonas-Reinhardii 137.* Biochimica Et Biophysica Acta, 1981. **655**(2): p. 136-145.
21. Iwanij, V., N.H. Chua, and P. Siekevitz, *Synthesis and Turnover of Ribulose Biphosphate Carboxylase and of Its Subunits during Cell-Cycle of Chlamydomonas-Reinhardtii.* Journal of Cell Biology, 1975. **64**(3): p. 572-585.

22. Janero, D.R. and R. Barrnett, *Thylakoid membrane biogenesis in Chlamydomonas reinhardtii 137+. II. Cell-cycle variations in the synthesis and assembly of pigment.* J Cell Biol, 1982. **93**(2): p. 411-6.
23. Juppner, J., et al., *Dynamics of lipids and metabolites during the cell cycle of Chlamydomonas reinhardtii.* Plant J, 2017. **92**(2): p. 331-343.
24. Walther, Z. and J.L. Hall, *The uni chromosome of Chlamydomonas: histone genes and nucleosome structure.* Nucleic Acids Res, 1995. **23**(18): p. 3756-63.
25. Atteia, A., et al., *Bifunctional aldehyde/alcohol dehydrogenase (ADHE) in chlorophyte algal mitochondria.* Plant Mol Biol, 2003. **53**(1-2): p. 175-88.
26. Atteia, A., et al., *Anaerobic energy metabolism in unicellular photosynthetic eukaryotes.* Biochim Biophys Acta, 2013. **1827**(2): p. 210-23.
27. Burgess, S.J., et al., *Identification of the Elusive Pyruvate Reductase of Chlamydomonas reinhardtii Chloroplasts.* Plant Cell Physiol, 2016. **57**(1): p. 82-94.
28. Catalanotti, C., et al., *Fermentation metabolism and its evolution in algae.* Front Plant Sci, 2013. **4**.
29. Happe, T., B. Mosler, and J.D. Naber, *Induction, localization and metal content of hydrogenase in the green alga Chlamydomonas reinhardtii.* Eur J Biochem, 1994. **222**(3): p. 769-74.
30. Mus, F., et al., *Anaerobic acclimation in Chlamydomonas reinhardtii: anoxic gene expression, hydrogenase induction, and metabolic pathways.* J Biol Chem, 2007. **282**(35): p. 25475-86.
31. Sawyer, A. and M. Winkler, *Evolution of Chlamydomonas reinhardtii ferredoxins and their interactions with (FeFe)-hydrogenases.* Photosynth Res, 2017.
32. Schmollinger, S., et al., *Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism.* Plant Cell, 2014. **26**(4): p. 1410-1435.
33. Terauchi, A.M., et al., *Pattern of Expression and Substrate Specificity of Chloroplast Ferredoxins from Chlamydomonas reinhardtii.* Journal of Biological Chemistry, 2009. **284**(38): p. 25867-25878.
34. Zones, J.M., et al., *High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation.* Plant Cell, 2015. **27**(10): p. 2743-69.
35. Merchant, S. and L. Bogorad, *Rapid degradation of apoplastocyanin in Cu(II)-deficient cells of Chlamydomonas reinhardtii.* J Biol Chem, 1986. **261**(34): p. 15850-3.
36. Monroe, M.E., et al., *MASIC: a software program for fast quantitation and flexible visualization of chromatographic profiles from detected LC-MS(/MS) features.* Comput Biol Chem, 2008. **32**(3): p. 215-7.
37. Zhan, Y., et al., *Pyrenoid functions revealed by proteomics in Chlamydomonas reinhardtii.* PLoS One, 2018. **13**(2): p. e0185039.
38. Aoki, Y., et al., *ALCOdb: Gene Coexpression Database for Microalgae.* Plant Cell Physiol, 2016. **57**(1): p. e3.
39. Wisecaver, J.H., et al., *A Global Coexpression Network Approach for Connecting Genes to Specialized Metabolic Pathways in Plants.* Plant Cell, 2017. **29**(5): p. 944-959.