Genome **Biology**

RESEARCH

Open Access

Genome-wide analysis of the diatom cell cycle unveils a novel type of cyclins involved in environmental signaling

Marie JJ Huysman[1,2,3], Cindy Martens[2,3], Klaas Vandepoele[2,3], Jeroen Gillard[1], Edda Rayko[4], Marc Heijde[4], Chris Bowler[4], Dirk Inzé[2,3], Yves Van de Peer[2,3], Lieven De Veylder[2,3] and Wim Vyverman*[1]

Abstract

Background: Despite the enormous importance of diatoms in aquatic ecosystems and their broad industrial potential, little is known about their life cycle control. Diatoms typically inhabit rapidly changing and unstable environments, suggesting that cell cycle regulation in diatoms must have evolved to adequately integrate various environmental signals. The recent genome sequencing of *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* allows us to explore the molecular conservation of cell cycle regulation in diatoms.

Results: By profile-based annotation of cell cycle genes, counterparts of conserved as well as new regulators were identified in *T. pseudonana* and *P. tricornutum*. In particular, the cyclin gene family was found to be expanded extensively compared to that of other eukaryotes and a novel type of cyclins was discovered, the diatom-specific cyclins. We established a synchronization method for *P. tricornutum* that enabled assignment of the different annotated genes to specific cell cycle phase transitions. The diatom-specific cyclins are predominantly expressed at the G1-to-S transition and some respond to phosphate availability, hinting at a role in connecting cell division to environmental stimuli.

Conclusion: The discovery of highly conserved and new cell cycle regulators suggests the evolution of unique control mechanisms for diatom cell division, probably contributing to their ability to adapt and survive under highly fluctuating environmental conditions.

Background

Diatoms (Bacillariophyceae) are unicellular photosynthetic eukaryotes responsible for approximately 20% of the global carbon fixation [1,2]. They belong to the Stramenopile algae (chromists) that most probably arose from a secondary endosymbiotic process in which a red eukaryotic alga was engulfed by a heterotrophic eukaryotic host approximately 1.3 billion years ago [3,4]. This event led to an unusual combination of conserved features with novel metabolism and regulatory elements, as recently confirmed by whole-genome analysis of *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* [5-7], which are representatives of the two major architectural diatom types, the centrics and the pennates, respectively.

Besides their huge ecological importance, diatoms are interesting from a biotechnological perspective as producers of a variety of metabolites (including oils, fatty acids, and pigments) [8,9], and because of their highly structured mesoporous cell wall, made of amorphous silica [10]. Thus, understanding the basic mechanisms controlling the diatom life cycle will be important to comprehend their ecological success in aquatic ecosystems and to control and optimize diatom growth for commercial applications.

As predominant organisms in marine and freshwater ecosystems, diatoms often encounter rapid and intense environmental fluctuations (for example, light and nutrient supply) [11] that might have dramatic effects on cell physiology and viability. Therefore, cell cycle regulation in diatoms most probably involves efficient signalling of different environmental cues [12]. Recent studies illustrate how diatoms can acclimate rapidly to iron limitation

* Correspondence: Wim.Vyverman@UGent.be
[1] Protistology and Aquatic Ecology, Department of Biology, Ghent University, Krijgslaan 281-S8, 9000 Gent, Belgium

[13,14] and phosphorus scarcity [15] through biochemical reconfiguration or maintenance of internal reservoirs and how their cell fate can be determined by perception of diatom-derived reactive aldehydes [16,17]. Furthermore, in *P. tricornutum*, a new blue light sensor (cryptochrome/photolyase family member 1) has been discovered with dual activity as a 6-4 photolyase and a blue-light-dependent transcription regulator [18]. Thus, diatoms are expected to possess complex fine-tuned signalling networks that integrate diverse stimuli with the cell cycle. The recent availability of genome data of *T. pseudonana* [5] and *P. tricornutum* [6] now provides the basis to explore how the cell cycle machinery has evolved in diatoms.

Efficient molecular regulation of the cell cycle is crucial to ensure that structural rearrangements during cell division are coordinated and that the genetic material is replicated and distributed correctly. In eukaryotes, the mitotic cell cycle comprises successive rounds of DNA synthesis (S phase) and cell division (mitosis or M phase) separated from each other by two gap (G1 and G2) phases [19]. Passage through the different cell cycle phases is controlled at multiple checkpoints by an evolutionarily conserved set of proteins, the cyclin-dependent kinases (CDKs) and cyclins (reviewed in [19,20]). Together, these proteins can form functional complexes, in which the CDKs and cyclins act as catalytic and regulatory subunits, respectively. Various types of CDKs and cyclins exist and they generally regulate the cell cycle, but some can be involved in other processes, such as transcriptional control or splicing [21,22].

In eukaryotes, activity of CDK-cyclin complexes is mainly controlled by (de)phosphorylation of the CDK subunits and interaction with inhibitors or scaffolding proteins [23]. Regulators include CDK-activating kinases (CAKs) [24,25], members of the WEE1/MYT1/MIK1 kinase family and CDC25 phosphatases that carry out inhibitory phosphorylation and dephosphorylation [26], as well as CDK inhibitors (CKIs) [23] and the scaffolding protein CKS1/Suc1 [27,28].

The aim of this work was to reveal the molecular network of cell cycle regulators in *P. tricornutum*, a species used for decades as a model diatom for physiological studies. *P. tricornutum* is a coastal diatom, typically found in highly unstable environments, and its cells can easily acclimate to environmental changes [13,29]. Key cell cycle regulators (CDKs, CDK interactors, and cyclins) were annotated and their transcript expression profiled during synchronized growth in *P. tricornutum*. The results indicate that diatom cell division is controlled by a combination of conserved molecules found in yeast, animals and/or plants, and novel components, including diatom-specific cyclins that probably transduce the environmental status of the cells to the cell cycle machinery.

Results and discussion
Annotation of the cell cycle genes in diatoms
The following cell cycle gene families were selected for comprehensive analysis: CDKs, cyclins, CKS1/suc1, WEE1/MYT1/MIK1, CDC25, and CKIs. These gene families were annotated functionally on the basis of their homology with known cell cycle genes in other organisms (see Materials and methods). The results of this family-wise annotation are discussed below and summarized in Table 1 and Additional file 1. The nomenclature of all identified proteins is according to that used in other protists for which cell cycle gene annotation was available [30,31].

Cell cycle synchronization and expression analysis
To validate the predicted functions of the annotated genes, we examined their transcript expression during the cell cycle. To synchronize cell division in *P. tricornutum*, we subjected exponentially growing cells to a prolonged dark period, which arrests the cells in the G1 phase [32] (Figure 1; Additional file 2), and released the cells synchronously from this arrest point by illumination. A comparable method had been applied successfully to synchronize growth in a closely related diatom, *Seminavis robusta* [33]. Microscopic observations of the dark-arrested *P. tricornutum* cultures showed that all cells contained a single undivided chloroplast (Figure 1a, upper panel). Accordingly, in flow cytometric histograms, the dark-arrested cells showed only a 2C peak (Figure 1b and Additional file 2, t = 0), confirming the G1 phase identity of cells containing a single chloroplast. When cells were released from the dark arrest, the population of bi-chloroplastidic cells steadily increased and cells entered the S phase, as observed by flow cytometry (Additional file 2, upper panel). However, the level of synchrony decreased at later time points (from 10 h after the dark release onward), probably because cells entered the next cell division cycle at the moment other cells still had to pass through M phase (Additional file 2). To circumvent this problem and to obtain an enrichment of cells in M phase during the later time points (Additional file 2), the metaphase blocker nocodazole was added at the time of re-illumination [34], but without major effect on cell cycle progression (Additional file 2).

To monitor gene expression during the different cell cycle phases, exponentially growing cells were synchronized in the presence of nocodazole (Figure 1b, c). Automated analysis of the flow histograms indicated that G1-phase cells were dominant during the first 4 h of re-illumination; from 4 to 7 h, cells went through S phase, as seen by the broadening and lowering of the 2C peak, while cells went mainly through the G2 and M phases at 8 to 12 h (Figure 1b, c). In *S. robusta*, chloroplast division had been found to take place only after S-phase onset [33]. Chloroplast division in *P. tricornutum* was observed starting from 5 h after illumination, confirming the S-phase timing determined by flow

Table 1: Overview and evolutionary conservation of the different core cell cycle gene families

Cell cycle gene	Number of copies				
	Phatra	Aratha, b	Osttaa, c	Saccea, c	Homsaa, c
CDKA	2d	1	1	1	3
CDKB	-	4	1	-	-
CDKC	2	2	1	1	1
CDKD	1	3	1	-	1
CDKF	-	1	-	1	1
CYCA	1?	10	1	NA	NA
CYCB	2?	9	1	NA	NA
CYCD	1?	10	1	NA	NA
CYCH	1?	1	1	NA	NA
CDC25	-	-	1e	1	3
Wee1/Myt1/Mik1	1	1	2	2	2
CKS	1	2	1	1	2
CKI	-	7	1	1	8

[a]Abbreviations: Phatr, *Phaeodactylum tricornutum*; Arath, *Arabidopsis thaliana*; Ostta, *Ostreococcus tauri*; Sacce, *Saccharomyces cerevisiae*; Homsa, *Homo sapiens*. [b]Data taken from [67]. [c]Data taken from [30].[d]One of these genes shows some *CDKB* characteristics. [e]Classification uncertain because of weak phylogeny. NA, not available due to other classification nomenclature.

cytometry (Figure 1a, lower panel, and 1c). The duration of the cell cycle after the synchronization procedure was comparable with that of cultures grown under standard conditions (approximately one division per day; Additional file 3). For downstream analysis, at hourly intervals after illumination, samples were taken for expression analysis by real-time quantitative polymerase chain reaction (qPCR).
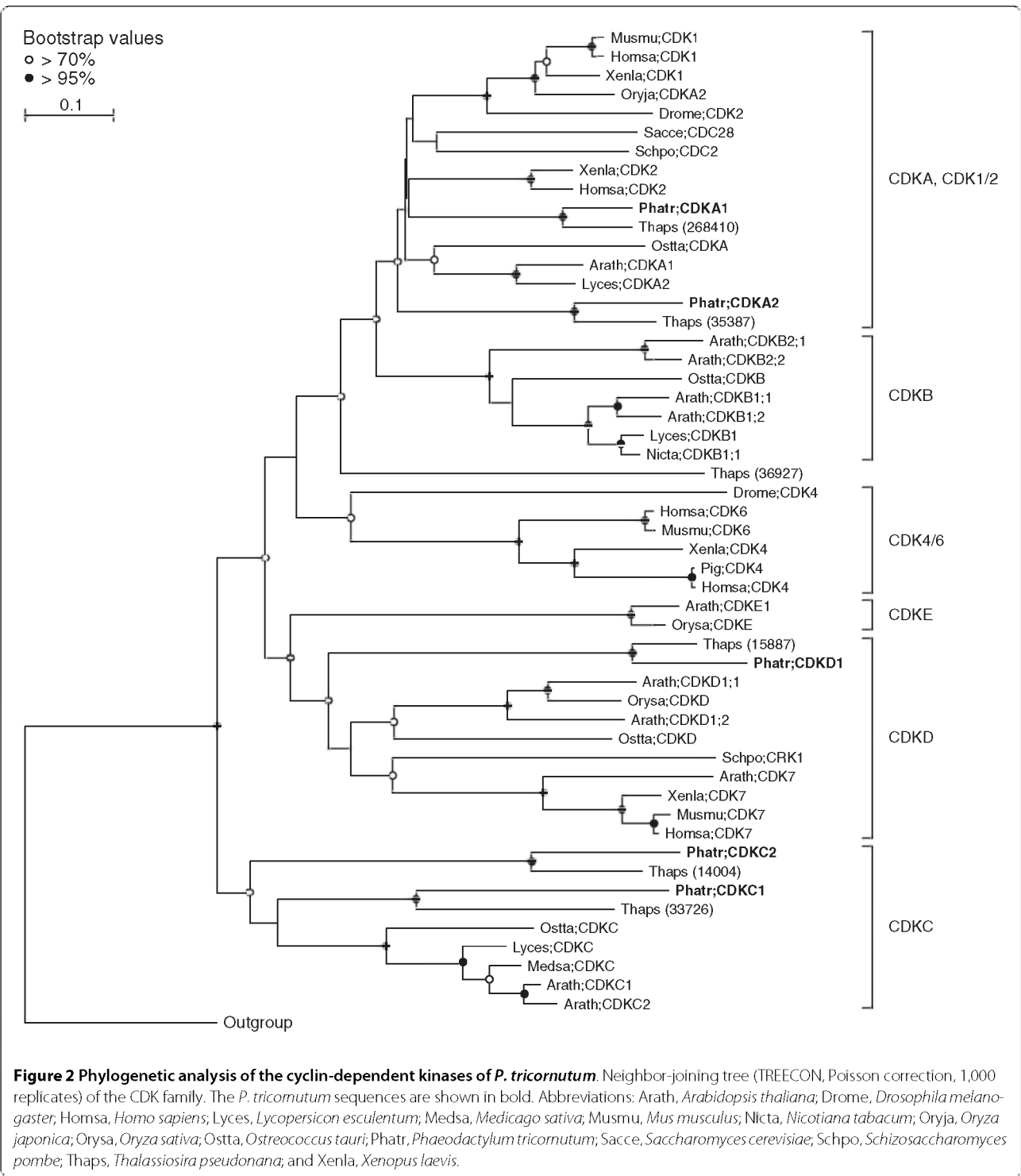
CDKs and CDK interactors
CDKs
CDKs are serine/threonine kinases that play a central role in cell cycle regulation and other processes, such as transcriptional control. Yeast uses only one single PSTAIRE-containing CDK for cell cycle progression [35,36], while higher organisms encode different CDKs implicated in cell division. The most conserved CDKs contain a PSTAIRE cyclin-binding motif [19,20]. In plants, the PSTAIRE-containing CDK had been designated CDKA and is active during both G1-to-S and G2-to-M transitions [19]. The plant-specific B-type CDKs contain a P [P/S]T [A/T]LRE motif and are active during the G2 and M phases [19]. In animals, three PSTAIRE (Cdk1, Cdk2, and Cdk3) and two P(I/L)ST(V/I)RE (Cdk4 and Cdk6) CDKs are involved in cell cycle control, although evidence has been found recently that only Cdk1 is really required to drive cell division [20,37].

Five CDKs could be identified in *P. tricornutum* (Table 1), of which two clustered together with the CDKA (plant)/CDK1-2 (animal) family in the phylogenetic tree (Figure 2). CDKA1 contains the typical PSTAIRE cyclin-binding motif (Figure 3) and its mRNA levels were high during late G1 and S phase (Figure 4a), suggesting a role at the G1-to-S transition. CDKA2 shows a PSTALRE motif (Figure 3), which is a midway motif between the CDKA hallmark PSTAIRE and the plant-specific CDKB hallmark P [P/S]T [A/T]LRE. The mRNA levels of CDKA2 were elevated in G2/M cells (Figure 4a). No homologs of the metazoan CDK4/6 family were found in *P. tricornutum*.

CDKC, CDKD and CDKE (designated Cdk9, Cdk7 and Cdk8 in animals, respectively) are kinases related to CDKA [38]. C-type CDKs (CDKC and Cdk9) and Cdk8 have been shown to associate with transcription initiation complexes and, thus, to play a role in transcriptional control [39,40]. Additionally, CDKC2 is active in spliceosomal dynamics in plants [22] and CDKE controls floral cell differentiation [41]. We identified two C-type CDKs (Table 1), CDKC1 and CDKC2 (Figure 2a) with PITALRE and PLQFIRE cyclin-binding motifs, respectively (Figure 3). No CDKE homolog was found in *P tricornutum*. Both *CDKC* genes had relatively low mRNA levels throughout the cell cycle without any discernible cell cycle phase pattern (data not shown). Thus, like in other eukaryotes, *CDKC* expression probably does not depend on the cell cycle phase in *P. tri-*
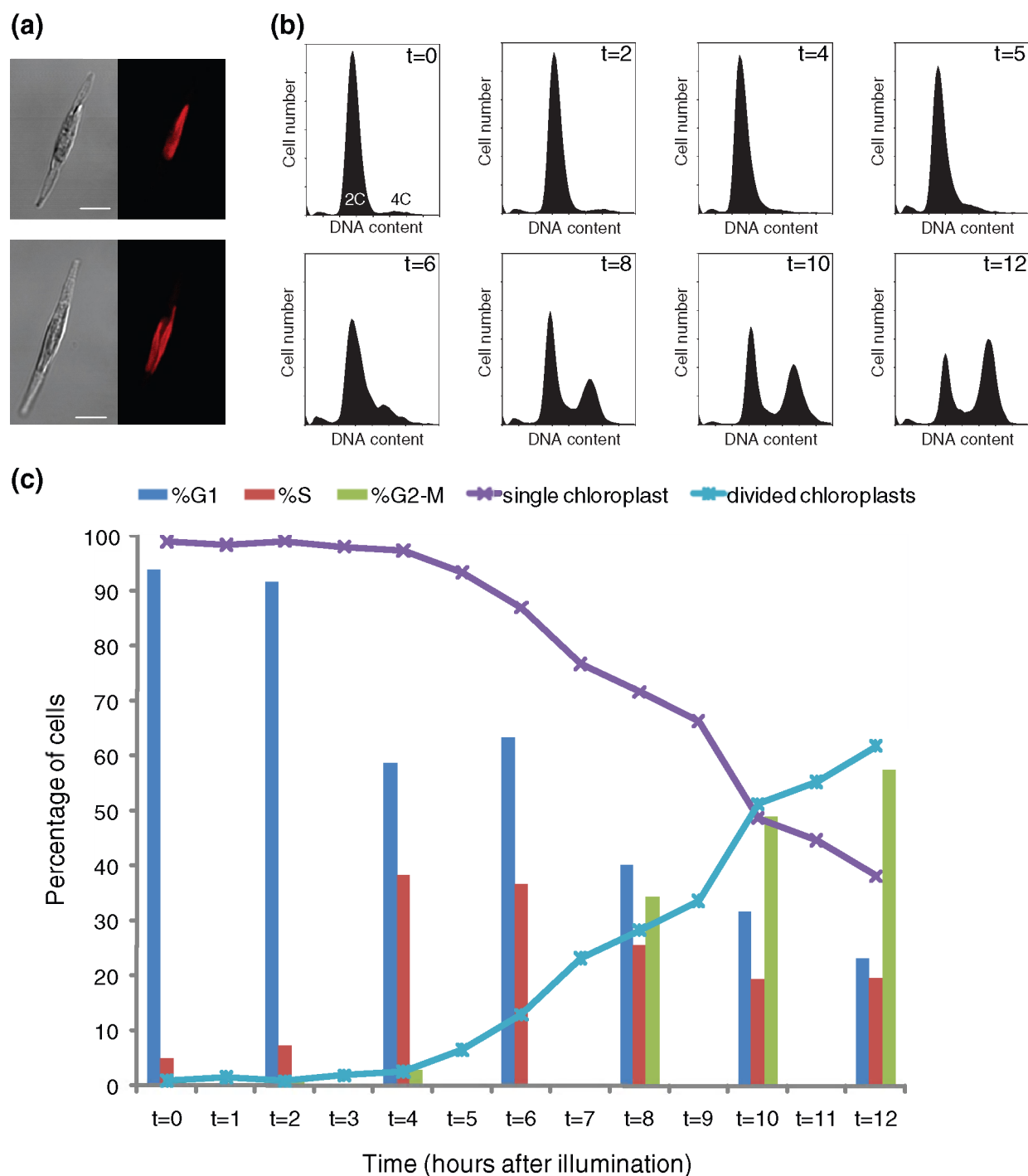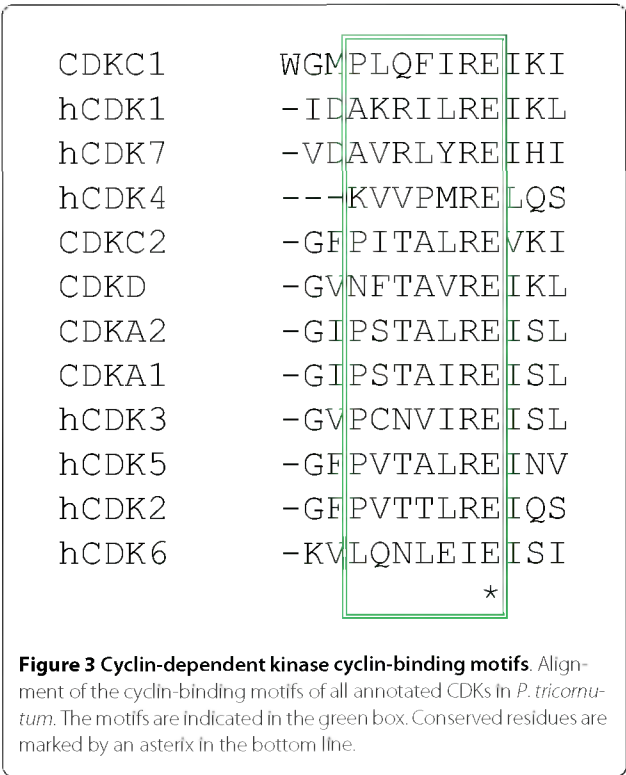
Figure 1 Synchronization of the cell cycle in *P. tricornutum*. (a) Confocal images of a dark-arrested cell (upper panel) showing a single parietal chloroplast and a cell after 12 h illumination (lower panel) showing divided and translocated daughter chloroplasts. Red, autofluorescence of the chloroplast. Scale bar: 5 μm. **(b)** Validation of synchronization of the cell cycle of *P. tricornutum* by flow cytometry. DNA content (abscissa) is plotted against cell number (ordinate). After a 20-h dark period, most of the cells are blocked in G1 phase (t = 0 to 4 h), indicated by the single 2C peak. After reillumination, cells proceed synchronously with their cell cycle, going through S phase (between t = 4 and 7 h), visible as the broadening and lowering of the 2C peak, and G2-M phase (t = 8 to 12 h), indicated by the accumulation of 4C cells. **(c)** Histogram indicating the proportion of cells in a certain cell cycle phase and chloroplast conformation during the cell cycle. Divided chloroplasts were observed starting from 5 h after illumination, after S-phase onset.

Figure 2 Phylogenetic analysis of the cyclin-dependent kinases of *P. tricornutum*. Neighbor-joining tree (TREECON, Poisson correction, 1,000 replicates) of the CDK family. The *P. tricornutum* sequences are shown in bold. Abbreviations: Arath, *Arabidopsis thaliana*; Drome, *Drosophila melanogaster*; Homsa, *Homo sapiens*; Lyces, *Lycopersicon esculentum*; Medsa, *Medicago sativa*; Musmu, *Mus musculus*; Nicta, *Nicotiana tabacum*; Oryja, *Oryza japonica*; Orysa, *Oryza sativa*; Ostta, *Ostreococcus tauri*; Phatr, *Phaeodactylum tricornutum*; Sacce, *Saccharomyces cerevisiae*; Schpo, *Schizosaccharomyces pombe*; Thaps, *Thalassiosira pseudonana*; and Xenla, *Xenopus laevis*.

cornutum, but it might be involved in other processes, such as transcription or splicing. One CDKD was identified (*CDKD1*) in *P. tricornutum* (Table 1 and Figure 2a). D-type CDKs are known to interact with H-type cyclins to form a CAK complex [24]. We found that *CDKD1* mRNA levels were high at the G1-to-S phase transition (Figure 4a). Another CDK variant, CDKF, has only been found in plants, where it functions as a CAK-activating kinase (CAKAK) [24]. No members of the CDKF family were identified in *P. tricornutum*, confirming that the CAKAK pathway is specific to plants and should have evolved within the green lineage (Table 1).

In addition, we identified seven hypothetical CDKs (hCDKs; Additional file 1) with divergent cyclin-binding

```
CDKC1      WGMPLQFIREIKI
hCDK1      -IDAKRILREIKL
hCDK7      -VDAVRLYREIHI
hCDK4      ---KVVPMRELQS
CDKC2      -GFPITALREVKI
CDKD       -GVNFTAVREIKL
CDKA2      -GIPSTALREISL
CDKA1      -GIPSTAIREISL
hCDK3      -GVPCNVIREISL
hCDK5      -GFPVTALREINV
hCDK2      -GFPVTTLREIQS
hCDK6      -KVLQNLEIEISI
                      *
```

Figure 3 Cyclin-dependent kinase cyclin-binding motifs. Alignment of the cyclin-binding motifs of all annotated CDKs in *P. tricornutum*. The motifs are indicated in the green box. Conserved residues are marked by an asterix in the bottom line.

domains (Figure 3) that could not be integrated into the phylogenetic tree due to high sequence divergence. The expression levels of several of these hCDKs were modulated during the cell cycle (Figure 4a). The *hCDK1* mRNA levels were the highest during G2-M, whereas those of *hCDK6* were up-regulated during G1 phase and *hCDK2*, *hCDK3*, *hCDK4*, and *hCDK5* were predominantly

expressed at G1 and/or S phase. For *hCDK7*, no reproducible expression pattern was found (data not shown).
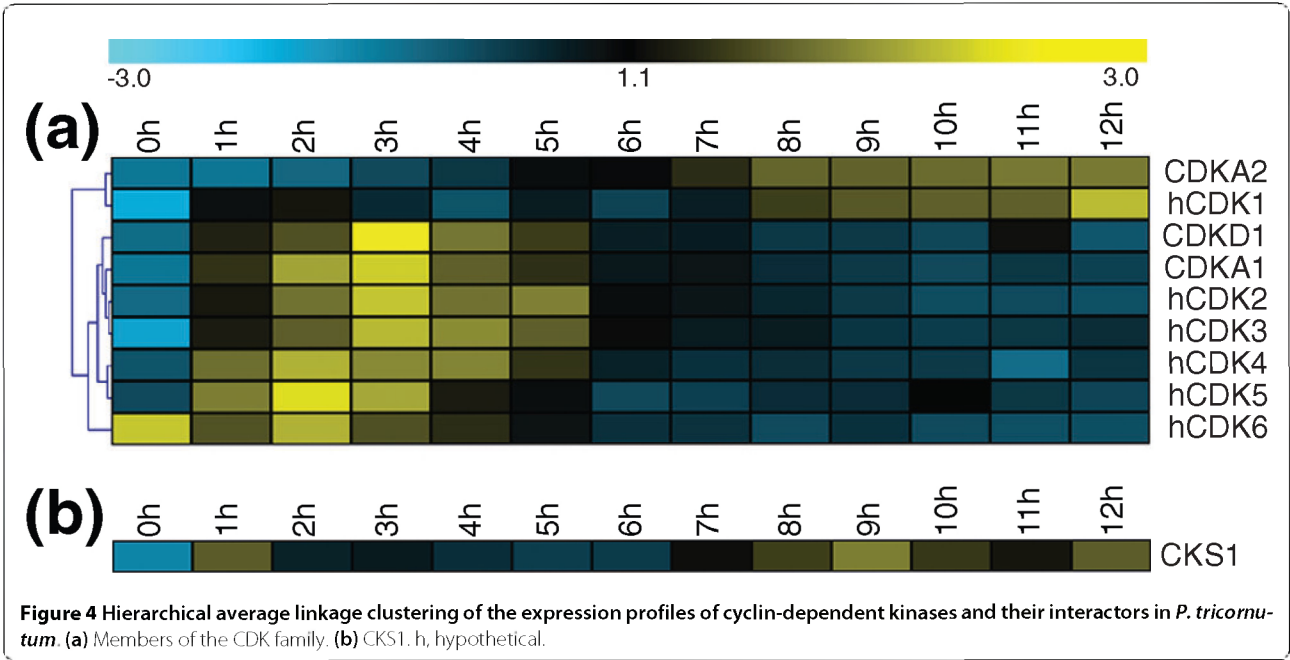
CDK subunit

CDK subunit (CKS) proteins act as docking factors that mediate the interaction of CDKs with putative substrates and regulatory proteins [27]. In *P. tricornutum*, one *CKS* gene was found (*CKS1*; Table 1) of which the mRNA levels were mainly high in G2/M cells (Figure 4b).

WEE1/MYT1/MIK1 kinases

WEE1/MYT1/MIK1 kinases inhibit cell cycle progression through phosphorylation of CDKs [26]. In yeast and animals, MYT1 is a membrane-associated kinase that phosphorylates Thr14 of Cdc2 proteins, as well as Tyr15, which is also a target of WEE1, a nucleus-localized kinase [42,43]. A single CKI could be identified in *P. tricornutum*, belonging to the MYT1 family (Table 1; Additional file 4) [42]. In *Arabidopsis thaliana*, the inhibitory kinase corresponds to WEE1 [44], while the green alga *Ostreococcus tauri* expresses both *WEE1* [30] and *MYT1* (unpublished data), like animals do [42] (Table 1). Expression of the *P. tricornutum MYT1* kinase was not associated with a specific cell cycle phase (data not shown). Because *MYT1* is probably implicated in stress responses during the cell cycle [45], it is possible that the imposed dark arrest or addition of nocodazole influenced the mRNA levels of *MYT1*, with too much variability in its expression profile as a consequence.

CDC25 phosphatase

As antagonists of the WEE1/MYT1/MIK1 kinases, CDC25 phosphatases activate CDKs [26]. In contrast to the presence of a counteracting kinase, no CDC25 phosphatase could be identified in *P. tricornutum* (Table 1) or in *T. pseudonana*. Both *Arabidopsis* and *Oryza sativa* also lack a

Figure 4 Hierarchical average linkage clustering of the expression profiles of cyclin-dependent kinases and their interactors in *P. tricornutum*. (a) Members of the CDK family. **(b)** CKS1. h, hypothetical.

functional CDC25 [46,47] and, in plants, CDC25-mediated regulatory mechanisms have been proposed to be replaced by a mechanism governed by the plant-specific B-type CDKs [48]. In *P. tricornutum*, no true B-type CDK homolog could be found, but CDKA2, classified by weak homology as A-type CDK class, possessed a PSTALRE cyclin-binding motif (Figure 3), which is halfway between the CDKA and CDKB hallmarks. This motif also occurred in the *Dictyostelium discoideum* CDC2 homolog [49] and in the *O. tauri* CDKB protein [30]. The PSTALRE motif is present as well in the CDKA2 homolog of *T. pseudonana* (Thaps3_35387; Figure 2a), confirming that this subtype could generally be found in diatoms. Moreover, *CDKA2* was expressed during G2-M (Figure 4a), the expected time of action of a B-type CDK. Although further in-depth biochemical research will be required to determine its true physiological function, the presence of this A/B-type CDK might explain the absence of a CDC25 phosphatase in diatoms. Alternatively, if the sequence of the CDC25 phosphatase had diverged to such an extent in diatoms, it might be not detectable by sequence homology, as already suggested for higher plants as well [50].

CDK inhibitors

CDK-cyclin complexes can be inactivated by CKIs, including the members of the INK4 family and the Cip/Kip family in animals [51], or Kip-related proteins and SIAMESE proteins in plants [52,53]. CKIs are mainly low-molecular-weight proteins that inhibit CDK activity by tight association in response to developmental or environmental stimuli [23,51,54]. Despite extensive sequence similarity searches for CKIs, no homologs could be identified in *P tricornutum*, which is not so surprising given the high sequence diversity of this cell cycle family [52]. These inhibitory proteins are most probably present in *P tricornutum*, but their identification will require more advanced molecular techniques.

Cyclins

The cyclin gene family is expanded in diatoms

We found a large number of highly diverged cyclin genes in diatoms, of which 24 are in *P. tricornutum* (Additional file 1). Due to their high divergence, indicated by the low bootstrap values in the phylogenetic tree, the classification into different subclasses was not clear (Figure 5), as it was for the 52 putative cyclins identified in *T. pseudonana* [55]. Moreover, many represent a novel class of cyclins, which we designated diatom-specific cyclins (dsCYCs).

To investigate whether the expansion of the cyclin gene family is specific to diatoms, we compared cyclin abundance among a representative set of Chromalveolates (Stramenopiles, Apicomplexa, and Ciliates; Table 2) for which genome data are available [56-64] and have been pre-processed in a previous study [65]. Because of the lack of cell cycle gene annotation in all investigated species, we

first screened for cyclin genes, which allowed us to create a reference dataset for analyzing cyclin evolution. We searched the different genomes for proteins that showed similarity to our cyclin HMMER profile and determined the number of proteins that contained an InterPro cyclin domain (Table 2). Generally, both detection methods yielded comparable results within all species (Table 2). An indication of the putative subclasses and function of the detected proteins is given by specific cyclin InterPro domains (Table 2). The proportion of the detected cyclin proteins relative to the predicted total gene number of each species revealed that, in the diatom genomes, cyclins are overrepresented compared to all investigated species, except for both *Cryptosporidium* species [57,58] and *Paramecium tetraurelia* [64] (Table 2). However, the total number of cyclins found in *Cryptosporidium* (12) is low compared to that in diatoms (28 in *P. tricornutum* and 57 in *T. pseudonana*). *Cryptosporidium* species are protozoan pathogens that depend on their hosts for nutrients. Moreover, Gene Ontology distribution for *Cryptosporidium* and *Plasmodium* is similar, indicating that no functional specialization of conserved gene families has occurred [58]. In *Paramecium tetrauleria*, the cyclin family is expanded as well. However, this species has a complex genome structure, possessing silent diploid micronuclei and polyploid macronuclei. Furthermore, *P. tetraurelia* underwent at least three whole-genome duplications, resulting in an apparent expansion of almost every gene family [64].

In conclusion, the large number of cyclin genes in both diatoms does not seem to be shared with its closest related species, indicating that diatom cyclins could have evolved separately to acquire new specific functions. Although the cyclin family has been found to be expanded in both diatoms, the size of the cyclin gene family in *T. pseudonana* is larger than that in *P. tricornutum*, which seems to result mainly from the presence of a larger number of diatom-specific cyclins in *T. pseudonana* (Figure 5). The biological cause of the changes in the cyclin family size remains unknown, although natural selection due to differential habitats might have played a role, or alternatively, random gene loss or gain might have occurred over long time stretches, as both species diverged at least 90 million years ago [6]. Genome sequence data of other diatom species are currently being generated (for example, for *Fragilariopsis cyclindrus* and *Pseudo-nitzschia multiseries*) and will help to shed light on cyclin gene family evolution in diatoms.

Conserved cyclins

Cyclins can be functionally classified into two major groups: the cell cycle regulators and the transcription regulators. Generally, during the cell cycle, specific cyclins are associated with G1 phase (cyclin D), S phase (cyclins A and E), and mitosis (cyclins A and B) [66]. In *P. tricornutum*, we identified a single A/B-type cyclin gene (*CYCA/B;1*; Figure 5), which gradually accumulated its mRNA
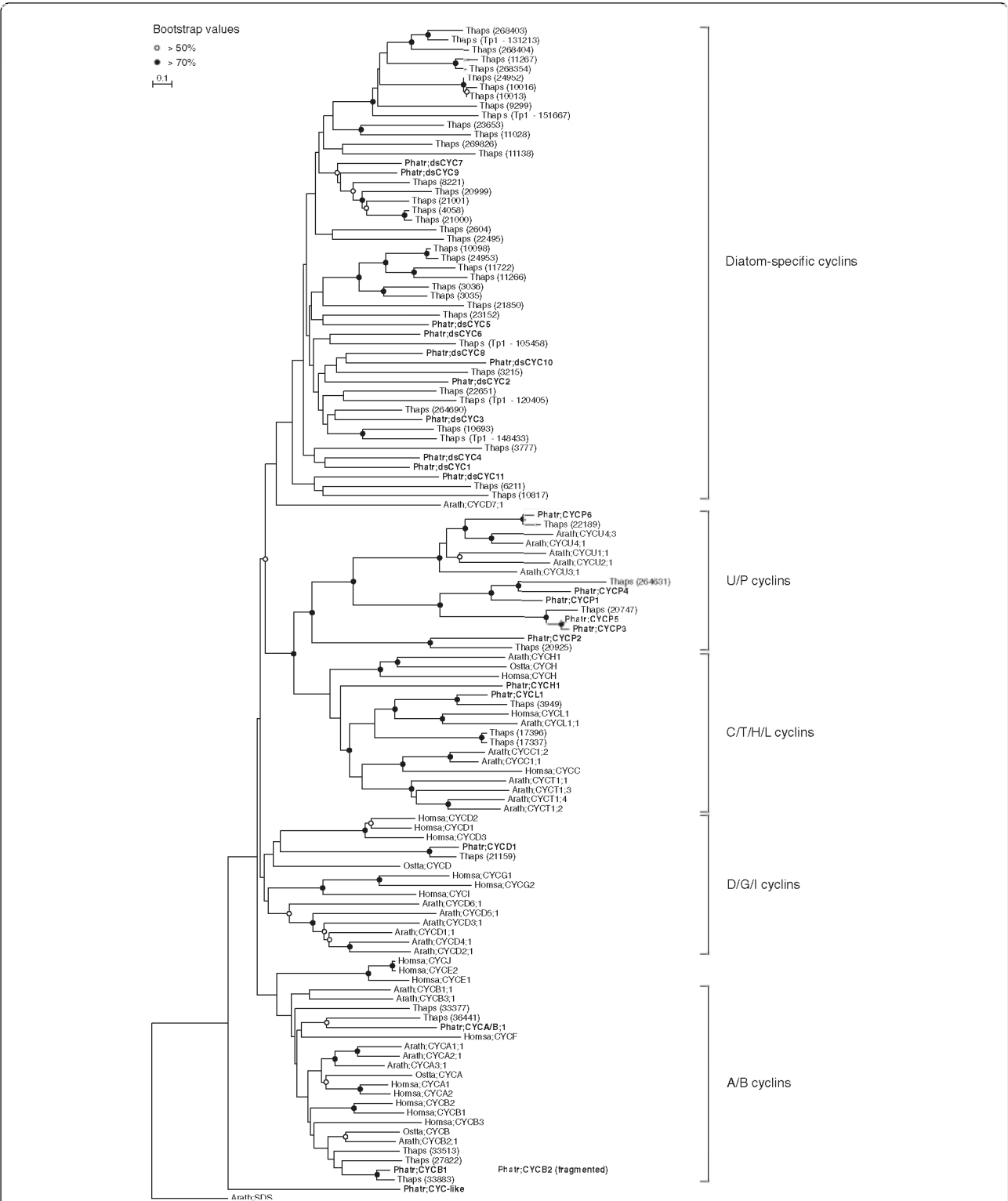
Figure 5 Phylogenetic analysis of the cyclins of *P. tricornutum*. Neighbor-joining tree (TREECON, Poisson correction, 500 replicates) of the cyclin family. The *P. tricornutum* sequences are shown in bold. Abbreviations: Arath, *Arabidopsis thaliana*; Homsa, *Homo sapiens*; Ostta, *Ostreococcus tauri*; Phatr, *Phaeodactylum tricornutum*; and Thaps, *Thalassiosira pseudonana*.

Table 2: Expansion of cyclin gene family in different representatives of the Chromalveolata

| | Stramenopiles | | | | Apicomplexa | | | | | | Ciliates | |
| | Diatoms | | Oomycetes | | | | | | | | | |
	Ph atr	Tha ps	Phy ra	Phy so	Cry ho	Cry pa	Pla fa	Pla yo	The an	The pa	Par te	Tet th
General												
Number of proteins matching the cyclin HMMER profile	28	57	19	19	12	12	5	5	8	8	144	29
Number of proteins with an InterPro cyclin domain	27	55	18	18	12	12	5	5	4	6	140	27
Specific InterPro domains												
IPR004367 Cyclin, C-terminal	7	9	4	5	-	-	-	-	-	-	19	6
IPR006670 Cyclin	6	18	7	7	2	2	2	1	2	1	94	20
IPR006671 Cyclin, N-terminal	18	45	9	10	1	1	-	-	-	-	96	20
IPR011028 Cyclin-like	27	55	18	18	11	11	5	5	4	6	140	27
IPR013763 Cyclin-related	21	47	13	13	6	6	3	2	2	2	72	22
IPR013922 Cyclin-related 2	1	1	1	1	3	3	1	1	1	1	21	1
IPR014400 Cyclin, A/B/D/E	2	4	2	3	1	1	-	-	-	-	42	3
IPR015429 Transcription regulator cyclin	4	3	6	5	2	2	2	2	2	2	6	3
IPR015432 Cyclin H	1	-	1	1	1	1	1	1	1	1	-	-
IPR015451 Cyclin D	-	4	-	-	1	-	-	-	-	-	-	-

Table 2: Expansion of cyclin gene family in different representatives of the Chromalveolata (Continued)

IPR015452 G2/mitotic-specific cyclin B3	-	-	-	-	-	-	-	-	-	-	3	1
IPR015453 G2/mitotic-specific cyclin A	1	1	3	3	-	-	-	-	-	-	2	-
IPR015454 G2/mitotic-specific cyclin B	1	-	-	-	-	-	-	-	-	-	9	1
IPR017060 Cyclin L	-	-	1	1	-	-	1	1	-	-	2	-
Total number of genes	10,402	11,776	15,743	19,027	3,994	3,952	5,268	5,268	3,792	4,035	39,642	27,000
Genome size (Mbp)	27.4	32.4	65	95	9.16	9.11	22.85	23.1	8.35	8.3		104
Cyclins/genes total (%)a	0.27	0.48	0.12	0.10	0.30	0.30	0.09	0.09	0.21	0.20	0.36	0.11
Cyclins/genes total (%)b	0.26	0.47	0.11	0.09	0.30	0.30	0.09	0.09	0.11	0.15	0.35	0.10

Abbreviations: Phatr, *Phaeodactylum tricornutum*; Thaps, *Thalassiosira pseudonana*; Phyra, *Phytophthora ramorum*; Physo, *Phytophthora sojae*; Cryho, *Cryptosporidium hominis*; Crypa, *Cryptosporidium parvum*; Plafa, *Plasmodium falciparum*; Playo, *Plasmodium yoelii yoelii*; Thean, *Theileria annulata*; Thepa, *Theileria parva*; Parte, *Paramecium tetraurelia*, Tetth, *Tetrahymena thermophila*. aNumber of cyclins versus total number of genes calculated with the number of proteins that match our cyclin HMMER profile. bNumber of cyclins versus total number of genes calculated with the number of proteins with a InterPro cyclin domain.

transcript during the G2 and M phases (Figure 6a). Both B-type cyclin genes (encoded by *CYCB1* and *CYCB2*) (Figure 5) were predominately expressed in G2/M cells, but mRNA levels of *CYCB2* accumulated earlier than those of *CYCB1* (Figure 6a). The single D-type cyclin (encoded by *CYCD1*; Figure 2b) was mainly expressed during S and G2/M phase progression (Figure 6a). As in plants, CYCE seems to be absent in diatoms [67].

Cyclins with a regulatory role during transcription include those belonging to the classes C, H, K, L, and T [39]. However, some cyclins involved in transcriptional control might also have a function in cell cycle regulation. For example, besides being a transcriptional regulator, the human C-type cyclin is also involved in the control of cell cycle transitions [68] and H-type cyclins can regulate the cell cycle through interaction with D-type CDKs, thereby forming a CAK complex [24,69,70]. The latter is probably also true for the P. tricornutum CYCH1 (Figure 5) because it was coexpressed with CDKD1 during the cell cycle (Figure 6a). The single L-type cyclin (encoded by CYCL1; Figure 5) showed elevated mRNA levels at G1 and during S phase (Figure 6a). In animals, cyclin L (also called Ania-6)

has previously been demonstrated to be an immediate early gene that could be involved in cell cycle re-entry [71,72].

Six cyclins in *P tricornutum* clustered together with P-type cyclins (PHO80-like proteins, also called U-type cyclins; Additional file 1 and Figure 5) that are believed to play a role in phosphate signalling [73,74]. The mRNA levels of all P-type cyclin genes (*CYCP1*, *CYCP2*, *CYCP3*, *CYCP4*, *CYCP5*, and *CYCP6*) were high early during the time series (Figure 6a). One cyclin gene did not cluster with any of the represented classes and was annotated as CYC-like (Figure 5). The mRNA levels of this gene peaked during the G1 and S phases (Figure 6a).

### Most diatom-specific cyclins are expressed early during the cell cycle

Eleven cyclin genes were identified that clustered only with cyclins of *T. pseudonana* (Figure 5). Therefore, we assigned these as dsCYC genes. *dsCYC3* and *dsCYC4* showed both high expression at the G2/M phases (Figure 6b). The mRNA levels of *dsCYC10* were slightly up-regulated at the G1-to-S transition and reached a peak late during the cell cycle (Figure 6b). As the other *dsCYC* genes displayed increased mRNA levels during the G1 and/or S phases (*dsCYC1*, *dsCYC2*, *dsCYC5*, *dsCYC6*, *dsCYC7*, *dsCYC8*,

dsCYC9, and *dsCYC11*; Figure 6b), some might function as immediate early genes controlled by light or mitogens.

Organisms living in aquatic environments, particularly in coastal regions, often have to cope with rapid and broad fluctuations in light intensity, temperature, nutrient availability, oxygen level, and salinity, all of which can have profound consequences on cell cycle progression. Comparative genome analyses of marine phytoplankton have revealed that coastal organisms contain genetic imprints indicative of adaptation to life under variable conditions [75,76], including distinct proteins coding for photosynthesis and light harvesting, additional two-component regulatory systems, novel carbon-concentrating mechanisms, transcription of transporters and assimilation proteins for the uptake of alternative nitrogen sources, and numerous metal transporter families and metal enzymes [75,76]. Similar adaptation imprints were also found in the diatom genomes [5,6]. Nevertheless, because diatoms generally dominate the microplankton in temperate waters and coastal upwelling regions under favorable conditions [77],

we expect diatoms to possess additional sophisticated fine-tuning systems enabling them to adjust the pace of the cell division rate in tune with the prevailing conditions.

Although in plants numerous copies of D-type cyclins integrate both external and internal signals into the cell cycle [19], in *P. tricornutum* only one *CYCD* was identified that was highly expressed late during the cell cycle (Figure 6a). Therefore, in diatoms CYCD probably does not play its classical role of G1-phase signal integrator, but might have acquired an alternative function in the G2-to-M transition as previously proposed for some D-type cyclins in plants [78]. On the other hand, the wide variety of *dsCYC* genes in diatoms expressed early during the cell cycle renders them plausible candidates to fulfil the task of signal integrators. Moreover, diatom-specific genes have been found to evolve faster than other genes in diatom genomes [6], indicating that these cyclin genes might have acquired novel and/or species-specific functions. Interestingly, other gene families expanded in diatoms include histidine kinases and heat shock factors, which are supposed to be involved in envi-
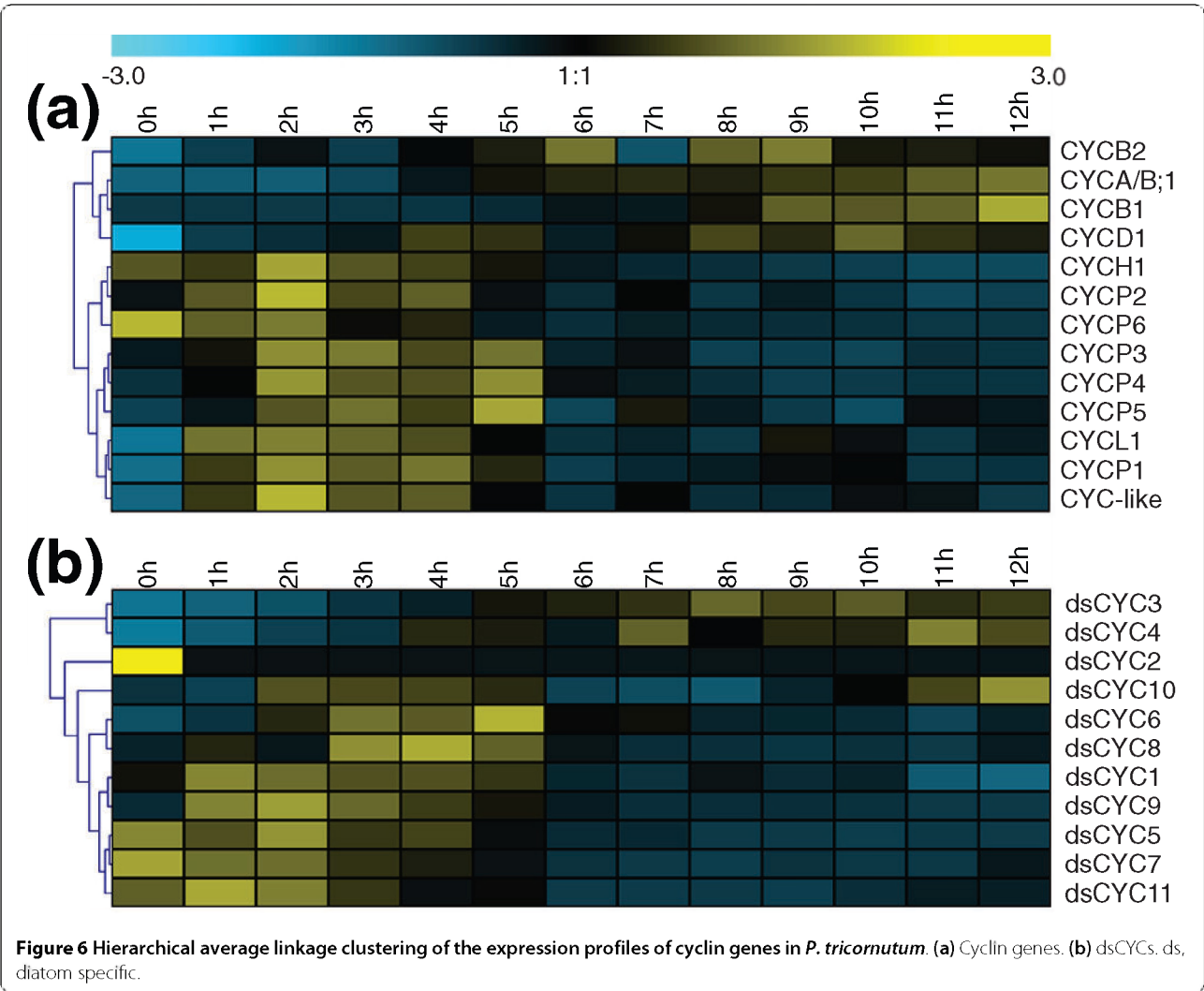
Figure 6 Hierarchical average linkage clustering of the expression profiles of cyclin genes in *P. tricornutum*. (a) Cyclin genes. (b) dsCYCs. ds, diatom specific.

ronmental sensing and expressed under certain growth conditions [6]. Thus, gene family expansion in diatoms could possibly be linked to the development of specific signal responses and adaptations to the environment.

dsCYCs respond to nutrient availability

To investigate the role of the *dsCYC* genes during the cell cycle, we analyzed them in more detail. More specifically, we examined whether their transcription is affected by nutrient deprivation. Analysis of recently published expressed sequence tag data [79,80] illustrates the differential expression of *dsCYC3*, *dsCYC7*, and *dsCYC10* across a range of environmental conditions (for example, nitrate-starved, nitrate-repleted, and iron-limited cultures). Moreover, a microarray analysis revealed that *dsCYC9* transcript levels were higher in cultures grown in the presence of silica than those without silica [81].

To examine whether *dsCYC* expression could be responsive to nutrient status during the cell cycle, we monitored mRNA levels in parallel with cell growth during nutrient starvation-repletion experiments. Exponentially growing cultures were nutrient-starved for 24 h and re-supplied with only nitrate, phosphate, iron, trace metals, the combination of all nutrients (positive control), or no nutrients (negative control). Three hours after nutrient supply, samples were collected for expression analysis. After nitrate repletion, cells reinitiated cell division at almost comparable levels to the positive control cultures, whereas repletion with phosphate, iron, or trace elements did not differ from the negative control (Figure 7a), indicating that nitrate is a cell cycle rate-limiting nutrient in *P. tricornutum*, as reported for other diatom species [82,83]. Nitrogen starvation in diatoms generally leads to a G1-phase arrest [82,83]. Thus, increased mRNA levels of early cell cycle-regulated genes are to be expected at the time of cell cycle reinitiation after nitrate repletion. Accordingly, early cell cycle genes (*CYCP6*, *CYCH1*, and *hCDK5*) were induced in the nitrate replete and positive control cultures (Figure 7b). To exclude cell cycle effects during sampling, the starvation experiment was repeated for nitrate repletion, but after imposing a 24-h dark arrest after starvation and re-supply of nitrate in complete darkness. In these cultures, the expression of the early cell cycle genes did not differ from that of the negative control after nitrate supply (Figure 7c), confirming that expression of *CYCP6*, *CYCH1*, and *hCDK5* is linked to cell cycle re-entry rather than to the nitrate status of the cells.

In contrast to nitrate, cells resupplied with phosphate remained arrested (Figure 7a, b). Upon addition of phosphate, mRNA levels of *dsCYC5*, *dsCYC7* and *dsCYC10* were significantly higher than those of the negative control (Figure 7d), strongly suggesting that these genes might function as direct cell cycle signal integrators upon increase of phosphate levels. Upon replenishment with nitrate (in the

dark), iron or trace elements, no effects on *dsCYC* gene expression were observed (Figure 7d and data not shown).

Nitrogen, together with the micronutrient iron, is generally considered to be a major limiting factor of primary production in the oceans [84]. Phosphate limitation, on the other hand, is considered to be less common, although it has been reported in certain oceanic areas [85] and has been hypothized recently to have been more wide-spread during the glacial periods [86]. As an important constituent of adenosine triphosphate, nucleic acids, and phospholipids, phosphorus is an important molecule not only for growth, but for almost all metabolic activities. Recently, diatoms have been shown to reduce their phosphorus demand upon phosphorus limitation, and to maintain growth by substituting phospholipids with non-phosphorus membrane lipids, only when nitrogen is not limiting [15].

In summary, these results reveal that some dsCYCs might be involved in environmental cell cycle control, functioning as nutrient signal integrators. All phosphate-responding *dsCYC* genes were expressed early during the synchronized time series (Figure 6b), fitting with a function in linking nutritional status and cell division start.

Cell cycle biomarkers

The identification of the complete set of major cell cycle regulators in P. tricornutum, along with the determination of their temporal expression patterns, generates a basis for studying different cell cycle-related processes in diatoms. Diatom cell cycle biomarkers could be used to observe cell cycle effects in laboratory experiments, but they could also be highly valuable to monitor diatom life cycle events in the natural habitat, like bloom or rest periods.

To validate whether the expression data obtained through the synchronization experiment was applicable in cell cycle-associated studies, we selected diatom cell cycle genes with a defined expression pattern to test their value as cell cycle biomarkers. As a control experiment, we checked the expression of four early (*CYCH1*, *hCDK5*, *CDKA1*, and *CDKD1*) and two late (*CDKA2* and *CYCB1*) cell cycle genes during a 12-h light/12-h dark photoperiod (LD 12:12). Flow cytometry data during this 24-h time course of the grown cultures indicate that the cells show a low degree of 'natural' synchronization of cell division: in the morning, most cells are in the G1 phase, while in the evening, division takes place (Figure 8a). Thus, it was to be expected that genes determined as early and as late cell cycle genes would be induced in the morning and in the evening, respectively. Indeed, expression according to the different cell cycle distributions was found for all selected genes (Figure 8b, c), indicating that they would perform as good cell cycle markers in cell cycle-related studies and that the expression data obtained from the synchronization studies (Figures 4 and 6) could serve as a reliable basis to select appropriate marker genes.
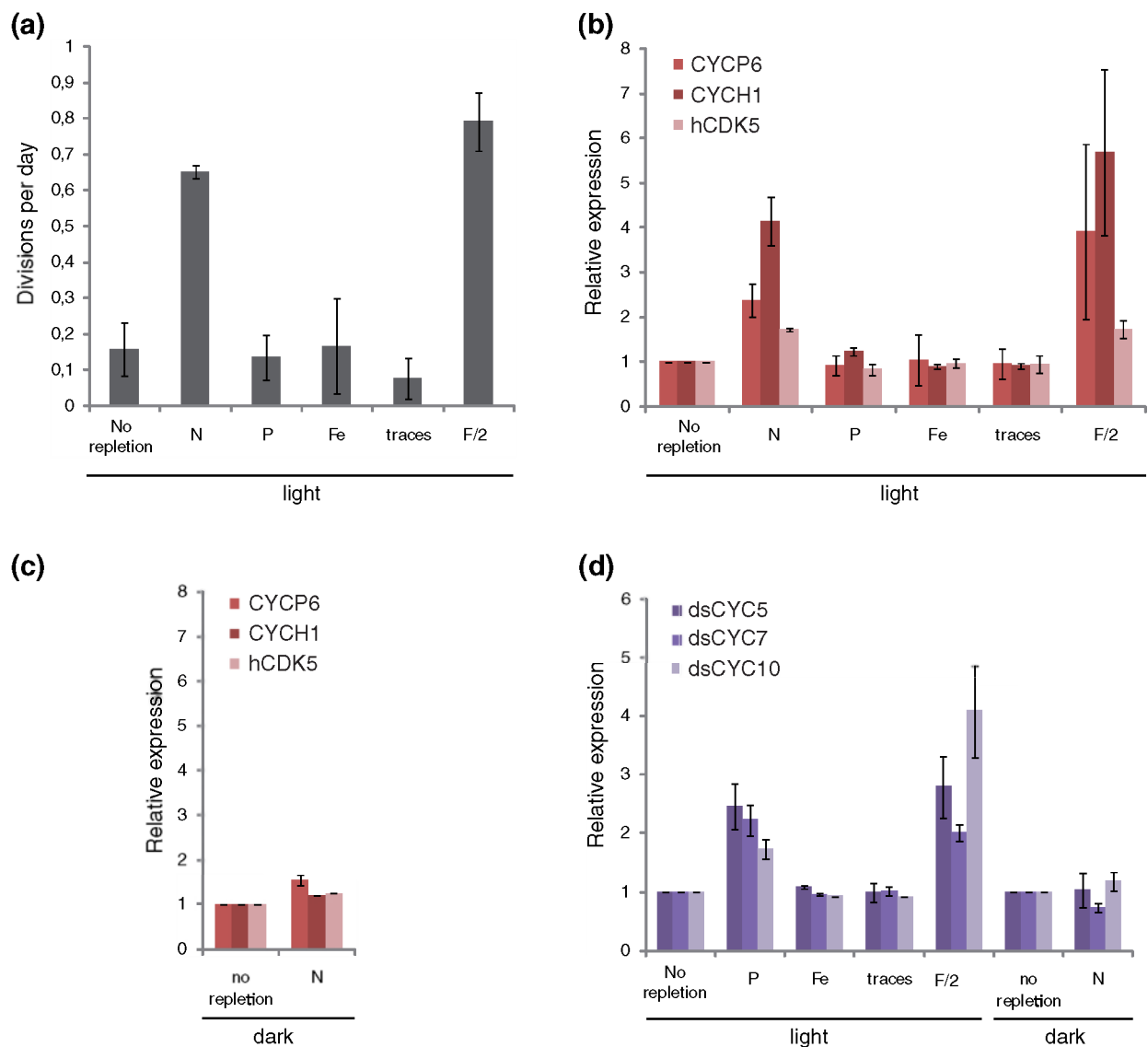
Figure 7 Nutrient response of diatom-specific cyclins. (a) Growth rate of different subcultures after repletion based on average cell density measurements at the time of and 3 days after repletion. These data indicate the ability of the cells to recover from starvation. **(b)** Expression profiles of early cell cycle genes at the time of sampling during the light experiment. **(c)** Expression profiles of early cell cycle genes at the time of sampling during the dark experiment. **(d)** dsCYCs responding to phosphate addition. Error bars represent standard errors of the mean of two biological replicates.

In a real case study, we used these cell cycle biomarkers to investigate whether the cell cycle in *P. tricornutum* would be regulated by an endogenous clock or a so-called circadian oscillator. Circadian regulation of cell division is well known to occur in eukaryotes and is particularly well-described for unicellular algae [87,88]. Although circadian regulation of light-harvesting protein-encoding genes and pigment synthesis has been reported in diatoms [89,90], we did not find any direct evidence that circadian regulation of the cell cycle exists in *P. tricornutum*. Comparison of cell cycle progression and cell cycle biomarker expression in cells under normal LD 12:12 or free-running LL 12:12 light

conditions indicate that neither the cell cycle itself nor mRNA accumulation of the main core cell cycle genes depends on a circadian oscillator (Additional files 5 and 6). These findings stress even more the importance of the development and use of efficient signalling networks that link environmental cues to cell growth in diatoms.

Conclusions

From the annotation and expression analyses, we conclude that the diatom cell cycle machinery shares common features with cell cycle regulatory systems present in other eukaryotes, including a PSTAIRE-containing CDK, con-
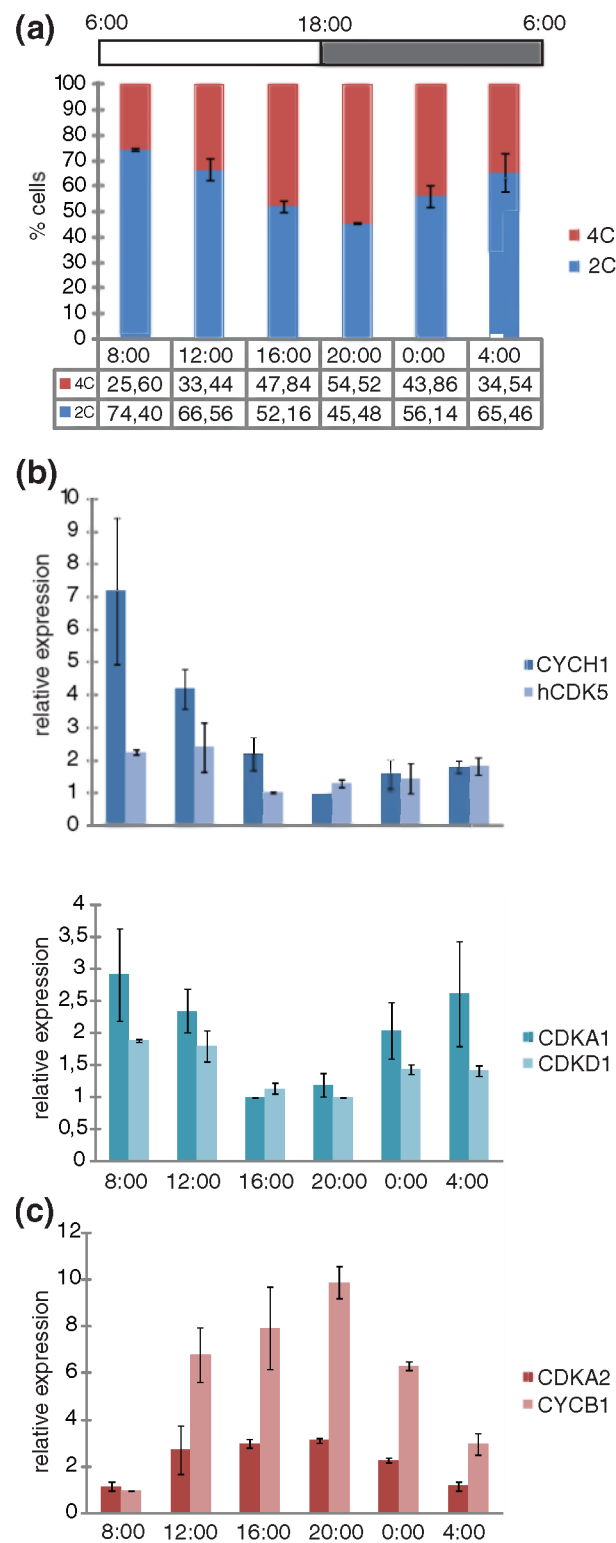
Figure 8 Validation of cell cycle marker genes. **(a)** DNA distributions (2C versus 4C) of exponentially growing cells entrained by a LD 12:12 photo-period during the time series **(b)** Expression profiles of early cell cycle genes (*CYCH1* and *hCDK5*; peak expression at t = 2 in the synchronization series (Figure 4 and 6)); and *CDKA1* and *CDKD1* (peak expression at t = 3 in the synchronization series (Figure 4)). **(c)** Expression profiles of late cell cycle genes (*CDKA2* and *CYCB1*). Error bars represent standard errors of the mean of two biological replicates.

served cyclin classes of types A, B, and D, and a MYT1 kinase. In addition, members of the retinoblastoma pathway for G1-S regulation involving the retinoblastoma protein and E2F/DP transcription factors [91-93] were also found in *P. tricornutum* (unpublished data). Components that were expected to be found in diatoms but could not be identified include a CDC25 phosphatase and CKIs. Possibly the function of the CDC25 phosphatase might be taken over by *CDKA2*, given its expression time and sequence similarity with B-type CDKs [48], whereas the lack of CKI identification by sequence similarity searches might be due to high sequence divergence [52].

Most interestingly, we found a major expansion of the cyclin gene family in diatoms and discovered a new cyclin class, the diatom-specific cyclins. The latter are most probably involved in signal integration to the cell cycle because transcript levels of *dsCYC5*, *dsCYC7*, and *dsCYC10* depended on phosphate (this study), and *dsCYC9* was reported to be induced upon silica availability [81]. Besides their role in nutrient sensing, we hypothesize that transcription of some *dsCYC* genes might also be light-modulated, as illustrated by the high *dsCYC2* mRNA levels in dark-acclimated cells that drastically dropped after 1 h of light exposure (Figure 6b). In addition, this gene was recently found to be modulated upon blue light treatment [18]. The responsiveness of other *dsCYC* genes to different light conditions is currently under investigation.

The complete set of major diatom key cell cycle regulators identified in this study could serve as a set of marker genes for monitoring diatom growth both in the laboratory and in the field. As cell cycle-regulated transcription cannot be assumed to depict a cell cycle-regulatory role for a gene, the predicted functions of the individual diatom cell cycle genes await further experimental confirmation by molecular and biochemical studies, although they already provide first insights into the manner in which diatoms control their cell division. Therefore, this dataset will form a starting point for future experiments aimed at exploring and manipulating the diatom cell cycle.

Materials and methods
Culture conditions
P. tricornutum (Pt1 8.6; accession numbers CCAP 1055/1 and CCMP2561) [29] was grown in F/2 medium without silica (F/2-Si) [94], made with filtered and autoclaved sea water collected from the North Sea (Belgium). Cultures were cultivated at 18 to 20°C in a 12-h light/12-h dark regime (50 to 100 $\mu mol \cdot photons \cdot m^{-2} \cdot s^{-1}$) and shaken at 100 rpm. Under these conditions, the average generation time of *P. tricornutum* was calculated to be 0.93 ± 0.07 days (Additional file 3).

Family-wise annotation of the diatom cell cycle genes
In a first step, known plant and animal cell cycle genes were selected to construct a reference cell cycle dataset. The members of every cell cycle family were used to build family-specific HMMER profiles [95]. With these profiles, the predicted *P. tricornutum* and *T. pseudonana* proteomes were screened for the presence of core cell cycle families. Missing gene families were also screened against the raw genome sequence (using tBLASTN) to account for annotation errors (that is, missing genes). For each family, the putative *P. tricornutum* homologs found were validated by comparing them with the reference family members in a multiple alignment.

Phylogenetic analysis
Multiple alignments generated with MUSCLE [96] were manually improved with BioEdit [97]. To define subclasses within the gene families, phylogenetic trees were built that included the reference cell cycle genes from plants and animals. Both TREECON [98] and PHYLIP [99] were used to construct the neighbor-joining trees based on Poisson-corrected distances. To test the significance of the nodes, bootstrap analysis was applied using 1,000 replicates for all trees, except for the cyclin tree (500 replicates).

Synchronization of the cell cycle in P. tricornutum
P. tricornutum cells were arrested in the G1 phase by prolonged darkness (20 h). After release of the cells from this G1 checkpoint by reillumination, samples for cell cycle analysis and real-time qPCR were collected during 12 h at hourly intervals, starting at reillumination (t = 0). To prevent cells from entering a second cell cycle, nocodazole (2.5 mg/l; Sigma-Aldrich, St. Louis, Missouri, USA) was added to the cultures at t = 0. Synchronization was validated by flow cytometric analysis on a Partec CyFlow ML platform (with data acquisition software Flomax; Partec GmbH, Münster, Germany) on cells fixed with 70% ethanol, washed three times with 1× phosphate buffered saline and stained with 4',6-diamidino-2-phenylindole (final concentration of 1 ng/ml). For each sample, 10,000 cells were processed. Flow cytograms were analyzed with Multicycle AV for Windows (Phoenix Flow Systems, San Diego, California, USA) software to determine relative representations of the different cell cycle stages in the samples.

Nutrient starvation/repletion experiment
Exponentially growing cells (under constant light, 50 $\mu mol \cdot photons \cdot m^{-2} \cdot s^{-1}$) were collected by centrifugation 3 days after medium replenishment, and washed twice with natural seawater (North Sea, Belgium) to starve the cells. After 24 h starvation, the culture was subdivided into six subcultures and supplied with only nitrate (8.82×10^{-4} M $NaNO_3$; N), phosphate (3.62×10^{-5} M $NaH_2PO_4 H_2O$; P), iron (1.17×10^{-5} M $FeCl_3 6H_2O$; Fe), trace metals ($3.93 \times$

10^{-8} M $CuSO_4$ $5H_2O$, 2.60×10^{-8} M Na_2MoO_4 $2H_2O$, 7.65×10^{-8} M $ZnSO_4$ $7H_2O$, 4.20×10^{-8} $CoCl_2$ $6H_2O$ and 9.10×10^{-7} M $MnCl_2$ $4H_2O$; trace), the combination of all nutrients (concentrations as mentioned above; F/2), or no nutrients (no repletion). Samples were taken for real-time qPCR after 3 hours of incubation. Cell density and growth rate were monitored during 3 days after repletion using a Bürker counting chamber to assess the degree of starvation in the different subcultures. For each sample, the average cell density was counted from nine large squares (0.1 mm³) and growth rate was calculated from semi-log linear regression of the cell numbers plotted against time.

To exclude cell cycle effects upon nitrate repletion, the experiment was repeated with cells grown in a LD 12:12 photoperiod. Three days after medium replenishment, the cells were washed twice with natural seawater (North Sea, Belgium) to starve the cells and illuminated for 12 h. The cells were then incubated in the dark for 24 h and no nutrients and nitrate were supplied in the dark as mentioned above. Samples were taken for real-time qPCR after 3 hours of incubation in the dark.

Real-time qPCR
For RNA extraction, 5×10^7 cells were collected at each time point, fast frozen in liquid nitrogen, and stored at -70°C. To lyse the cells and extract RNA, TriReagent (Molecular Research Center, Inc., Cincinnati, Ohio, USA) was used initially. After addition of chloroform, RNA was purified from the aqueous phase by RNeasy purification, according to the manufacturer's instructions (RNeasy Min-Elute Cleanup kit; Qiagen, Hilden, Germany). Contaminating genomic DNA was removed by DNaseI (GE Healthcare, Little Chalfont, United Kingdom) treatment. RNA concentration and purity were assessed by spectrophotometry (NanoDrop ND-1000, Wilmington, Delaware, USA). Total RNA was reverse transcribed with Superscript II reverse transcriptase (Invitrogen, Carlsbad, California, USA) in a total volume of 40 μl with oligo(dT) primers. Finally, 1.25 ng (synchronization experiment and control experiment) or 10 ng (nutrient starvation/repletion experiment and circadian experiment) of cDNA was used as template for each qPCR reaction.

Samples in triplicate were amplified on the Lightcycler 480 platform with the Lightcycler 480 SYBR Green I Master mix (Roche Diagnostics, Brussels, Belgium) in the presence of 0.5 μM gene-specific primers (Additional file 1). The cycling conditions were 10 minutes polymerase activation at 95°C and 45 cycles at 95°C for 10 s, 58°C for 15 s, and 72°C for 15 s. Amplicon dissociation curves were recorded after cycle 45 by heating from 65°C to 95°C. In qBase [100], data were analyzed using the ΔC_t relative quantification method with the stably expressed histone H4 as a normalization gene (Additional file 7) [101]. Expres-

sion profiles of the synchronized cell cycle series were mean relative expression from three independent sample series. After normalization, the mean profiles were clustered using hierarchical average linkage clustering (analysis software TIGR MultiExperiment Viewer 3D (TMEV3D)).

Image acquisition
Confocal images were obtained with a scanning confocal microscope 100 M (Zeiss, Jena, Germany) equipped with the software package LSM510 version 3.2 (Zeiss, Jena, Germany) and a C-Apochromat 63× (1.2 NA) water-corrected objective. Chlorophyll autofluorescence was excited with HeNe illumination (543 nm).

Accession numbers
Sequence data from this article can be accessed through the Joint Genome Institute (JGI) portal [102]. Accession numbers of the cell cycle genes are listed in Additional file 1.

Additional material

Additional file 1 Cell cycle genes in *P. tricornutum.*. An Excel spreadsheet providing an overview of the annotated cell cycle genes in *P. tricornutum*.

Additional file 2 Cell cycle progression in nocodazole-treated versus untreated cells. A PDF figure file showing cell cycle progression in nocodazole-treated versus untreated cells. (a) Flow histograms plotting DNA content against cell number (left) and histograms indicating the ploidy distribution (2C versus 4C; right) during a 12-h time course of synchronized cells in the absence of nocodazole. At the later time points (t = 10 to 12), the level of synchrony decreased, indicated by the ploidy level equilibrium reached at these time points, probably resulting from cells entering the next cell cycle round, while other cells still have to pass through M phase. (b) Flow histograms plotting DNA content against cell number (left) and histograms indicating the ploidy distribution (2C versus 4C; right) during a 12-h time course of synchronized cells in the presence of nocodazole. At the later time points, an increasing enrichment of 4C cells can be observed because of a blockage of the cells at metaphase. Asterisk marks the apparently lower proportion of 2C cells after a 20-h dark treatment in the control series than in the nocodazole series, resulting from an acquisition artefact during flow cytometry, indicated by the increased peak broadness in the respective flow histogram.

Additional file 3 Growth curves of *P. tricornutum* cells under standard conditions. A PDF figure file showing growth curves of *P. tricornutum* cells under standard conditions (18°C, LD 12:12, 50 to 100 μmol·photons·m⁻²·s⁻¹). Error bars represent standard deviations.

Additional file 4 Phylogenetic tree of WEE1/MYT1/MIK1 family. A PDF figure file showing a Phylogenetic tree of WEE1/MYT1/MIK1 family. Neighbor-joining tree (PHYLIP, 1,000 replicates) of WEE1/MYT1/MIK1 family. The *P. tricornutum* sequence is shown in bold. Abbreviations: Arath, *Arabidopsis thaliana*; Drome, *Drosophila melanogaster*; Homsa, *Homo sapiens*; Musmu, *Mus musculus*; Orysa, *Oryza sativa*; Phatr, *Phaeodactylum tricornutum*; Schpo, *Schizosaccharomyces pombe*; Thaps, *Thalassiosira pseudonana*.

Additional file 5 Cell cycle versus circadian control. A PDF figure file showing cell cycle versus circadian control. Exponentially growing cultures entrained by a LD 12:12 photoperiod were subdivided in two cultures at the end of the light period 3 days after medium replenishment. Left and right: cells experiencing a normal (darkness; grey bar) and subjective (light; white bar) night, respectively. (a) Histograms plotting DNA distributions (2C versus 4C) of the cells during the 24-h time series. (b) Expression profiles of early cell cycle genes. (c) Expression profiles of late cell cycle genes. Error bars represent standard errors of the mean of two biological replicates.

Additional file 6 Cell cycle versus circadian control. A PDF figure file showing cell cycle versus circadian control. Flow histograms (DNA content plotted against cell number) of the different sampling points depicted in Additional file 5.

Additional file 7 Normalization gene evaluation. A PDF figure file showing normalization gene evaluation. **(a)** Real-time qPCR cycle threshold (Ct) values of candidate housekeeping genes during a 12-h (sampling every hour) synchronization time series. **(b)** Variation of Ct values of the candidate housekeeping genes during a 12-h (sampling every hour) synchronization time series. Error bars represent standard deviations.

Abbreviations

CAK: CDK-activating kinase; CDK: cyclin-dependent kinase; CKI: CDK inhibitor; CKS: CDK subunit; CYC: cyclin; D: dark; dsCYC: diatom-specific cyclin; hCDK: hypothetical CDK; L: light; qPCR: quantitative polymerase chain reaction.

Authors' contributions

MJJH performed the synchronization and expression experiments, analyzed the data and wrote the manuscript; CM was involved in the genome-wide annotation of cell cycle genes in *P. tricornutum* and *T. pseudonana* and helped write the manuscript; KV and ER were involved in the genome-wide annotation of diatom cell cycle genes. MJJH, CM, KV, JG, ER, MH, CB, DI, YVDP, LDV and WV helped to conceive and design the study, and read and approved the manuscript.

Author Details

[1]Protistology and Aquatic Ecology, Department of Biology, Ghent University, Krijgslaan 281-S8, 9000 Gent, Belgium, [2]Department of Plant Systems Biology, Flanders Institute for Biotechnology (VIB), Technologiepark 927, 9052 Gent, Belgium, [3]Department of Plant Biotechnology and Genetics, Ghent University, Technologiepark 927, 9052 Gent, Belgium and [4]Département de Biologie, Ecole Normale Supérieure, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 8186, rue d'Ulm 45, 75230 Paris Cedex 05, France

References

1. Hoek C Van den, Mann DG, Jahns HM: *Algae: An Introduction to Phycology* Cambridge: Cambridge University Press; 1995.
2. Field CB, Behrenfeld MJ, Randerson JT, Falkowski P: **Primary production of the biosphere: integrating terrestrial and oceanic components.** *Science* 1998, **281**:237-240.
3. Kooistra WH, De Stefano M, Mann DG, Medlin LK: **The phylogeny of the diatoms.** *Prog Mol Subcell Biol* 2003, **33**:59-97.
4. Li S, Nosenko T, Hackett JD, Bhattacharya D: **Phylogenomic analysis identifies red algal genes of endosymbiotic origin in the chromalveolates.** *Mol Biol Evol* 2006, **23**:663-674.
5. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou SG, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, et al.: **The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism.** *Science* 2004, **306**:79-86.
6. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret J-P, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, et al.: **The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes.** *Nature* 2008, **456**:239-244.
7. Montsant A, Jabbari K, Maheswari U, Bowler C: **Comparative genomics of the pennate diatom *Phaeodactylum tricornutum*.** *Plant Physiol* 2005, **137**:500-513.
8. Lebeau T, Robert J-M: **Diatom cultivation and biotechnologically relevant products. Part I: cultivation at various scales.** *Appl Microbiol Biotechnol* 2003, **60**:612-623.
9. Bozarth A, Maier U-G, Zauner S: **Diatoms in biotechnology: modern tools and applications.** *Appl Microbiol Biotechnol* 2009, **82**:195-201.
10. Kröger N: **Prescribing diatom morphology: toward genetic engineering of biological nanomaterials.** *Curr Opin Chem Biol* 2007, **11**:662-669.
11. Round FE, Crawford RM, Mann DG: *The Diatoms: Biology and Morphology of the Genera* Cambridge: Cambridge University Press; 1990.
12. Falciatore A, Ribera d'Alcalà MR, Croot P, Bowler C: **Perception of environmental signals by a marine diatom.** *Science* 2000, **288**:2363-2366.
13. Allen AE, Laroche J, Maheswari U, Lommer M, Schauer N, Lopez PJ, Finazzi G, Fernie AR, Bowler C: **Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation.** *Proc Natl Acad Sci USA* 2008, **105**:10438-10443.
14. Marchetti A, Parker MS, Moccia LP, Lin EO, Arrieta AL, Ribalet F, Murphy MEP, Maldonado MT, Armbrust EV: **Ferritin is used for iron storage in bloom-forming marine pennate diatoms.** *Nature* 2009, **457**:467-470.
15. Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblížek M, Lomas MW, Mincer TJ, Moore LR, Moutin T, Rappé MS, Webb EA: **Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity.** *Nature* 2009, **458**:69-72.
16. Vardi A, Formiggini F, Casotti R, De Martino A, Ribalet F, Miralto A, Bowler C: **A stress surveillance system based on calcium and nitric oxide in marine diatoms.** *PLoS Biol* 2006, **4**:e60.
17. Vardi A, Bidle KD, Kwityn C, Hirsh DJ, Thompson SM, Callow JA, Falkowski P, Bowler C: **A diatom gene regulating nitric-oxide signaling and susceptibility to diatom-derived aldehydes.** *Curr Biol* 2008, **18**:895-899.
18. Coesel S, Mangogna M, Ishikawa T, Heijde M, Rogato A, Finazzi G, Todo T, Bowler C, Falciatore A: **Diatom PtCPF1 is a new cryptochrome/ photolyase family member with DNA repair and transcription regulation activity.** *EMBO Rep* 2009, **10**:655-661.
19. Inzé D, De Veylder L: **Cell cycle regulation in plant development.** *Annu Rev Genet* 2006, **40**:77-105.
20. Morgan DO: **Cyclin-dependent kinases: engines, clocks, and microprocessors.** *Annu Rev Cell Dev Biol* 1997, **13**:261-291.
21. Coqueret O: **Linking cyclins to transcriptional control.** *Gene* 2002, **299**:35-55.
22. Kitsios G, Alexiou KG, Bush M, Shaw P, Doonan JH: **A cyclin-dependent protein kinase, CDKC2, colocalizes with and modulates the distribution of spliceosomal components in Arabidopsis.** *Plant J* 2008, **54**:220-235.
23. De Clercq A, Inzé D: **Cyclin-dependent kinase inhibitors in yeast, animals, and plants: a functional comparison.** *Crit Rev Biochem Mol Biol* 2006, **41**:293-313.
24. Umeda M, Shimotohno A, Yamaguchi M: **Control of cell division and transcription by cyclin-dependent kinase-activating kinases in plants.** *Plant Cell Physiol* 2005, **46**:1437-1442.
25. Kaldis P: **The cdk-activating kinase (CAK): from yeast to mammals.** *Cell Mol Life Sci* 1999, **55**:284-296.
26. Perry JA, Kornbluth S: **Cdc25 and Wee1: analogous opposites?** *Cell Div* 2007, **2**:12.
27. Pines J: **Cell cycle: reaching for a role for the Cks proteins.** *Curr Biol* 1996, **6**:1399-1402.
28. Harper JW: **Protein destruction: adapting roles for Cks proteins.** *Curr Biol* 2001, **11**:R431-R435.
29. De Martino A, Meichenin A, Shi J, Pan K, Bowler C: **Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions.** *J Phycol* 2007, **43**:992-1009.
30. Robbens S, Khadaroo B, Camasses A, Derelle E, Ferraz C, Inzé D, Peer Y Van de, Moreau H: **Genome-wide analysis of core cell cycle genes in the**

unicellular green alga *Ostreococcus tauri. Mol Biol Evol* 2005, **22**:589-597. A published erratum appears in *Mol Biol Evol* **22**:1158

31. Bisova K, Krylov DM, Umen JG: **Genome-wide annotation and expression profiling of cell cycle regulatory genes in** *Chlamydomonas reinhardtii. Plant Physiol* 2005, **137**:475-491.

32. Brzezinski MA, Olson RJ, Chisholm SW: **Silicon availability and cell-cycle progression in marine diatoms.** *Marine Ecol Progress Series* 1990, **67**:83-96.

33. Gillard J, Devos V, Huysman MJJ, De Veylder L, D'Hondt S, Martens C, Vanormelingen P, Vannerum K, Sabbe K, Chepurnov VA, Inzé D, Vuylsteke M, Vyverman W: **Physiological and transcriptomic evidence for a close coupling between chloroplast ontogeny and cell cycle progression in the pennate diatom** *Seminavis robusta. Plant Physiol* 2008, **148**:1394-1411.

34. Ng CKF, Lam CMC, Yeung PKK, Wong JTY: **Flow cytometric analysis of nocodazole-induced cell-cycle arrest in the pennate diatom** *Phaeodactylum tricornutum* Bohlin. *J Appl Phycol* 1998, **10**:569-572.

35. Mendenhall MD, Hodge AE: **Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast** *Saccharomyces cerevisiae. Microbiol Mol Biol Rev* 1998, **62**:1191-1243.

36. Moser BA, Russell P: **Cell cycle regulation in** *Schizosaccharomyces pombe. Curr Opin Microbiol* 2000, **3**:631-636.

37. Santamaría D, Barrière C, Cerqueira A, Hunt S, Tardy C, Newton K, Cáceres JF, Dubus P, Malumbres M, Barbacid M: **Cdk1 is sufficient to drive the mammalian cell cycle.** *Nature* 2007, **448**:811-815.

38. Joubès J, Chevalier C, Dudits D, Heberle-Bors E, Inzé D, Umeda M, Renaudin J-P: **CDK-related protein kinases in plants.** *Plant Mol Biol* 2000, **43**:607-620.

39. Oelgeschläger T: **Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control.** *J Cell Physiol* 2002, **190**:160-169.

40. Barrôco RM, De Veylder L, Magyar Z, Engler G, Inzé D, Mironov V: **Novel complexes of cyclin-dependent kinases and a cyclin-like protein from** *Arabidopsis thaliana* **with a function unrelated to cell division.** *Cell Mol Life Sci* 2003, **60**:401-412.

41. Wang W, Chen X: **HUA ENHANCER3 reveals a role for a cyclin-dependent protein kinase in the specification of floral organ identity in** *Arabidopsis. Development* 2004, **131**:3147-3156.

42. Mueller PR, Coleman TR, Kumagai A, Dunphy WG: **Myt1: a membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15.** *Science* 1995, **270**:86-90.

43. Liu F, Stanton JJ, Wu Z, Piwnica-Worms H: **The human Myt1 kinase preferentially phosphorylates Cdc2 on threonine 14 and localizes to the endoplasmic reticulum and Golgi complex.** *Mol Cell Biol* 1997, **17**:571-583.

44. De Schutter K, Joubès J, Cools T, Verkest A, Corellou F, Babiychuk E, Schueren E Van Der, Beeckman T, Kushnir S, Inzé D, De Veylder L: *Arabidopsis* **WEE1 kinase controls cell cycle arrest in response to activation of the DNA integrity checkpoint.** *Plant Cell* 2007, **19**:211-225.

45. Zhou BB, Elledge SJ: **The DNA damage response: putting checkpoints in perspective.** *Nature* 2000, **408**:433-439.

46. Bleeker PM, Hakvoort HWJ, Bliek M, Souer E, Schat H: **Enhanced arsenate reduction by a CDC25-like tyrosine phosphatase explains increased phytochelatin accumulation in arsenate-tolerant** *Holcus lanatus. Plant J* 2006, **45**:917-929.

47. Duan G-L, Zhou Y, Tong Y-P, Mukhopadhyay R, Rosen BP, Zhu Y-G: **A CDC25 homologue from rice functions as an arsenate reductase.** *New Phytol* 2007, **174**:311-321.

48. Boudolf V, Inzé D, De Veylder L: **What if higher plants lack a CDC25 phosphatase?** *Trends Plant Sci* 2006, **11**:474-479.

49. Michaelis C, Weeks G: **Isolation and characterization of a** *cdc2* **cDNA from** *Dictyostelium discoideum. Biochim Biophys Acta* 1992, **1132**:35-42.

50. Khadaroo B, Robbens S, Ferraz C, Derelle E, Eychenié S, Cooke R, Peaucellier G, Delseny M, Demaille J, Peer Y Van de, Picard A, Moreau H: **The first green lineage cdc25 dual-specificity phosphatase.** *Cell Cycle* 2004, **3**:513-518.

51. Sherr CJ, Roberts JM: **CDK inhibitors: positive and negative regulators of** G_1**-phase progression.** *Genes Dev* 1999, **13**:1501-1512.

52. Verkest A, Weinl C, Inzé D, De Veylder L, Schnittger A: **Switching the cell cycle. Kip-related proteins in plant cell cycle control.** *Plant Physiol* 2005, **139**:1099-1106.

53. Churchman ML, Brown ML, Kato N, Kirik V, Hülskamp M, Inzé D, De Veylder L, Walker JD, Zheng Z, Oppenheimer DG, Gwin T, Churchman J, Larkin JC:

SIAMESE, a plant-specific cell cycle regulator, controls endoreplication onset in *Arabidopsis thaliana. Plant Cell* 2006, **18**:3145-3157.

54. Sherr CJ, Roberts JM: **Inhibitors of mammalian G₁ cyclin-dependent kinases.** *Genes Dev* 1995, **9**:1149-1163.

55. Montsant A, Allen AE, Coesel S, De Martino A, Falciatore A, Mangogna M, Siaut M, Heijde M, Jabbari K, Maheswari U, Rayko E, Vardi A, Apt KE, Berges JA, Chiovitti A, Davis AK, Thamatrakoln K, Hadi MZ, Lane TW, Lippmeier JC, Martinez D, Parker MS, Pazour GJ, Saito MA, Rokhsar DS, Armbrust EV, Bowler C: **Identification and comparative genomic analysis of signaling and regulatory components in the diatom** *Thalassiosira pseudonana. J Phycol* 2007, **43**:585-604.

56. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RHY, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CMB, Dorrance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Krampis K, Lamour KH, Lee M-K, McDonald WH, Medina M, et al.: *Phytophthora* **genome sequences uncover evolutionary origins and mechanisms of pathogenesis.** *Science* 2006, **313**:1261-1266.

57. Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, Deng M, Liu C, Widmer G, Tzipori S, Buck GA, Xu P, Bankier AT, Dear PH, Konfortov BA, Spriggs HF, Iyer L, Anantharaman V, Aravind L, Kapur V: **Complete genome sequence of the apicomplexan,** *Cryptosporidium parvum. Science* 2004, **304**:441-445.

58. Xu P, Widmer G, Wang Y, Ozaki LS, Alves JM, Serrano MG, Puiu D, Manque P, Akiyoshi D, Mackey AJ, Pearson WR, Dear PH, Bankier AT, Peterson DL, Abrahamsen MS, Kapur V, Tzipori S, Buck GA: **The genome of** *Cryptosporidium hominis. Nature* 2004, **431**:1107-1112. A published erratum appears in *Nature* **432**:415

59. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M-S, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DMA, et al.: **Genome sequence of the human malaria parasite** *Plasmodium falciparum. Nature* 2002, **419**:498-511.

60. Carlton JM, Angiuoli SV, Suh BB, Kooij TW, Pertea M, Silva JC, Ermolaeva MD, Allen JE, Selengut JD, Koo HL, Peterson JD, Pop M, Kosack DS, Shumway MF, Bidwell SL, Shallom SJ, Van Aken SE, Riedmuller SB, Feldblyum TV, Cho JK, Quackenbush J, Sedegah M, Shoaibi A, Cummings LM, Florens L, Yates JR, Raine JD, Sinden RE, Harris MA, Cunningham DA, et al.: **Genome sequence and comparative analysis of the model rodent malaria parasite** *Plasmodium yoelii yoelii. Nature* 2002, **419**:512-519.

61. Gardner MJ, Bishop R, Shah T, de Villiers EP, Carlton JM, Hall N, Ren Q, Paulsen IT, Pain A, Berriman M, Wilson RJM, Sato S, Ralph SA, Mann DJ, Xiong Z, Shallom SJ, Weidman J, Jiang L, Lynn J, Weaver B, Shoaibi A, Domingo AR, Wasawo D, Crabtree J, Wortman JR, Haas B, Angiuoli SV, Creasy TH, Lu C, Suh B, et al.: **Genome sequence of** *Theileria parva,* **a bovine pathogen that transforms lymphocytes.** *Science* 2005, **309**:134-137.

62. Pain A, Renauld H, Berriman M, Murphy L, Yeats CA, Weir W, Kerhornou A, Aslett M, Bishop R, Bouchier C, Cochet M, Coulson RMR, Cronin A, de Villiers EP, Fraser A, Fosker N, Gardner M, Goble A, Griffiths-Jones S, Harris DE, Katzer F, Larke N, Lord A, Maser P, McKellar S, Mooney P, Morton F, Nene V, O'Neil S, Price C, et al.: **Genome of the host-cell transforming parasite** *Theileria annulata* **compared with** *T. parva. Science* 2005, **309**:131-133.

63. Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, Wortman JR, Badger JH, Ren Q, Amedeo P, Jones KM, Tallon LJ, Delcher AL, Salzberg SL, Silva JC, Haas BJ, Majoros WH, Farzad M, Carlton JM, Smith RK Jr, Garg J, Pearlman RE, Karrer KM, Sun L, Manning G, Elde NC, Turkewitz AP, Asai DJ, Wilkes DE, Wang Y, Cai H, et al.: **Macronuclear genome sequence of the ciliate** *Tetrahymena thermophila,* **a model eukaryote.** *PLoS Biol* 2006, **4**:e286.

64. Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, Segurens B, Daubin V, Anthouard V, Aiach N, Arnaiz O, Billaut A, Beisson J, Blanc I, Bouhouche K, Câmara F, Duharcourt S, Guigo R, Gogendeau D, Katinka M, Keller A-M, Kissmehl R, Klotz C, Koll F, Le Mouël A, Lepère G, Malinsky S, Nowacki M, Nowak JK, Plattner H, et al.: **Global trends of whole-genome duplications revealed by the ciliate** *Paramecium tetraurelia. Nature* 2006, **444**:171-178.

65. Martens C, Vandepoele K, Peer Y Van de: **Whole-genome analysis reveals molecular innovations and evolutionary transitions in chromalveolate species.** *Proc Natl Acad Sci USA* 2008, **105**:3427-3432.

66. Sánchez I, Dynlacht BD: **New insights into cyclins, CDKs, and cell cycle control.** *Semin Cell Dev Biol* 2005, **16**:311-321.
67. Vandepoele K, Raes J, De Veylder L, Rouzé P, Rombauts S, Inzé D: **Genome-wide analysis of core cell cycle genes in *Arabidopsis*.** *Plant Cell* 2002, **14**:903-916.
68. Liu Z-J, Ueda T, Miyazaki T, Tanaka N, Mine S, Tanaka Y, Taniguchi T, Yamamura H, Minami Y: **A critical role for cyclin C in promotion of the hematopoietic cell cycle by cooperation with c-Myc.** *Mol Cell Biol* 1998, **18**:3445-3454.
69. Fisher RP, Morgan DO: **A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase.** *Cell* 1994, **78**:713-724.
70. Yamaguchi M, Fabian T, Sauter M, Bhalerao RP, Schrader J, Sandberg G, Umeda M, Uchimiya H: **Activation of CDK-activating kinase is dependent on interaction with H-type cyclins in plants.** *Plant J* 2000, **24**:11-20.
71. Berke JD, Sgambato V, Zhu P-P, Lavoie B, Vincent M, Krause M, Hyman SE: **Dopamine and glutamate induce distinct striatal splice forms of Ania-6, an RNA polymerase II-associated cyclin.** *Neuron* 2001, **32**:277-287.
72. Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JCF, Trent JM, Staudt LM, Hudson J Jr, Boguski MS, Lashkari D, Shalon D, Botstein D, Brown PO: **The transcriptional program in the response of human fibroblasts to serum.** *Science* 1999, **283**:83-87.
73. Kaffman A, Herskowitz I, Tjian R, O'Shea EK: **Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85.** *Science* 1994, **263**:1153-1156.
74. Torres Acosta JA, de Almeida Engler J, Raes J, Magyar Z, De Groodt R, Inzé D, De Veylder L: **Molecular characterization of *Arabidopsis* PHO80-like proteins, a novel class of CDKA;1-interacting cyclins.** *Cell Mol Life Sci* 2004, **61**:1485-1497.
75. Palenik B, Ren Q, Dupont CL, Myers GS, Heidelberg JF, Badger JH, Madupu R, Nelson WC, Brinkac LM, Dodson RJ, Durkin AS, Daugherty SC, Sullivan SA, Khouri H, Mohamoud Y, Halpin R, Paulsen IT: **Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment.** *Proc Natl Acad Sci USA* 2006, **103**:13555-13559.
76. Peers G, Niyogi KK: **Pond scum genomics: the genomes of *Chlamydomonas* and *Ostreococcus*.** *Plant Cell* 2008, **20**:502-507.
77. Irigoien X, Harris RP, Verheye HM, Joly P, Runge J, Starr M, Pond D, Campbell R, Shreeve R, Ward P, Smith AN, Dam HG, Peterson W, Tirelli V, Koski M, Smith T, Harbour D, Davidson R: **Copepod hatching success in marine ecosystems with high diatom concentrations.** *Nature* 2002, **419**:387-389.
78. Schnittger A, Schöbinger U, Bouyer D, Weinl C, Stierhof Y-D, Hülskamp M: **Ectopic D-type cyclin expression induces not only DNA replication but also cell division in *Arabidopsis* trichomes.** *Proc Natl Acad Sci USA* 2002, **99**:6410-6415.
79. The Diatom EST Database   [http://www.biologie.ens.fr/diatomics/EST/]
80. Maheswari U, Mock T, Armbrust EV, Bowler C: **Update of the Diatom EST Database: a new tool for digital transcriptomics.** *Nucleic Acids Res* 2009, **37**:D1001-D1005.
81. Sapriel G, Quinet M, Heijde M, Jourdren L, Tanty V, Luo G, Le Crom S, Lopez PJ: **Genome-wide transcriptome analyses of silicon metabolism in *Phaeodactylum tricornutum* reveal the multilevel regulation of silicic acid transporters.** *PLoS One* 2009, **4**:e7458.
82. Olson RJ, Vaulot D, Chisholm SW: **Effects of environmental stresses on the cell cycle of two marine phytoplankton species.** *Plant Physiol* 1986, **80**:918-925.
83. Vaulot D, Olson RJ, Merkel S, Chisholm SW: **Cell-cycle response to nutrient starvation in 2 phytoplankton species, *Thalassiosira weissflogii* and *Hymenomonas carterae*.** *Mar Biol* 1987, **95**:625-630.
84. Falkowski PG, Barber RT, Smetacek VV: **Biogeochemical controls and feedbacks on ocean primary production.** *Science* 1998, **281**:200-207.
85. Wu J, Sunda W, Boyle EA, Karl DM: **Phosphate depletion in the western North Atlantic Ocean.** *Science* 2000, **289**:759-762.
86. Pichevin LE, Reynolds BC, Ganeshram RS, Cacho I, Pena L, Keefe K, Ellam RM: **Enhanced carbon pump inferred from relaxation of nutrient limitation in the glacial ocean.** *Nature* 2009, **459**:1114-1117.
87. Goto K, Johnson CH: **Is the cell division cycle gated by a circadian clock? The case of *Chlamydomonas reinhardtii*.** *J Cell Biol* 1995, **129**:1061-1069.
88. Moulager M, Monnier A, Jesson B, Bouvet R, Mosser J, Schwartz C, Garnier L, Corellou F, Bouget F-Y: **Light-dependent regulation of cell division in *Ostreococcus*: evidence for a major transcriptional input.** *Plant Physiol* 2007, **144**:1360-1369.

89. Oeltjen A, Marquardt J, Rhiel E: **Differential circadian expression of genes *fcp2* and *fcp6* in *Cyclotella cryptica*.** *Int Microbiol* 2004, **7**:127-131.
90. Ragni M, Ribera d'Alcalà M: **Circadian variability in the photobiology of *Phaeodactylum tricornutum*: pigment content.** *J Plankton Res* 2007, **29**:141-156.
91. Weinberg RA: **The retinoblastoma protein and cell cycle control.** *Cell* 1995, **81**:323-330.
92. Claudio PP, Tonini T, Giordano A: **The retinoblastoma family: twins or distant cousins?** *Genome Biol* 2002, **3**:reviews3012.
93. de Jager SM, Murray JA: **Retinoblastoma proteins in plants.** *Plant Mol Biol* 1999, **41**:295-299.
94. Guillard RRL: **Culture of phytoplankton for feeding marine invertebrates.** In *Culture of Marine Invertebrate Animals* Edited by: Smith WL, Canley MH. New York: Plenum Press; 1975:29-60.
95. Eddy SR: **Profile hidden Markov models.** *Bioinformatics* 1998, **14**:755-763.
96. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32**:1792-1797.
97. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp Ser* 1999, **41**:95-98.
98. Peer Y Van de, De Wachter R: **TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment.** *Comput Appl Biosci* 1994, **10**:569-570.
99. PHYLIP   [http://evolution.genetics.washington.edu/phylip/]
100. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J: **qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data.** *Genome Biol* 2007, **8**:R19.
101. Siaut M, Heijde M, Mangogna M, Montsant A, Coesel S, Allen A, Manfredonia A, Falciatore A, Bowler C: **Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*.** *Gene* 2007, **406**:23-35.
102. DOE Joint Genome Institute (JGI) Portal   [http://www.jgi.doe.gov/]