Genome-Wide Analysis of Core Cell Cycle Genes in the Unicellular Green Alga *Ostreococcus tauri*

Steven Robbens,† Basheer Khadaroo,*[1] Alain Camasses,* Evelyne Derelle,* Conchita Ferraz,‡ Dirk Inzé,† Yves Van de Peer,† and Hervé Moreau**

*Université Paris VI, Laboratoire Arago, Modèles en Biologie Cellulaire et Evolutive, Banyuls sur Mer, France; †Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology (VIB), Ghent University, Ghent, Belgium; and ‡Institut de Génétique Humaine, Montpellier, France

The cell cycle has been extensively studied in various organisms, and the recent access to an overwhelming amount of genomic data has given birth to a new integrated approach called comparative genomics. Comparing the cell cycle across species shows that its regulation is evolutionarily conserved; the best-known example is the pivotal role of cyclin-dependent kinases in all the eukaryotic lineages hitherto investigated. Interestingly, the molecular network associated with the activity of the CDK-cyclin complexes is also evolutionarily conserved, thus, defining a core cell cycle set of genes together with lineage-specific adaptations. In this paper, we describe the core cell cycle genes of *Ostreococcus tauri*, the smallest free-living eukaryotic cell having a minimal cellular organization with a nucleus, a single chloroplast, and only one mitochondrion. This unicellular marine green alga, which has diverged at the base of the green lineage, shows the minimal yet complete set of core cell cycle genes described to date. It has only one homolog of CDKA, CDKB, CDKD, cyclin A, cyclin B, cyclin D, cyclin H, Cks, Rb, E2F, DP, DEL, Cdc25, and Wee1. We have also added the APC and SCF E3 ligases to the core cell cycle gene set. We discuss the potential of genome-wide analysis in the identification of divergent orthologs of cell cycle genes in different lineages by mining the genomes of evolutionarily important and strategic organisms.

Introduction

All living organisms undergo cell division, of which the regulation is highly conserved throughout evolution (Stals and Inzé 2001). The eukaryotic cell cycle is regulated at multiple points, and cell division is ensured by cyclin-dependent kinase–cyclin (CDK-cyclin) complexes, heterodimers composed of a CDK subunit that binds a regulatory cyclin subunit. CDK-cyclin complexes are present in all eukaryotic lineages hitherto studied (Joubès et al. 2000). Their activity is, furthermore, controlled by evolutionarily conserved regulatory mechanisms: phosphorylation/dephosphorylation of the CDK subunit, binding of CDK inhibitors (CKI), cytoplasmic sequestration of the cyclin subunit, and specific ubiquitylation targeting of the cyclin subunit and CKI to proteasome-mediated proteolysis (Deshaies and Ferrell 2001; Obaya and Sedivy 2002).

Cell cycle control genes have been found in the different lineages investigated, including the animal, yeast, and plant lineages. Even though specific regulatory mechanisms are present in all lineages, some have evolved differently, such as the retinoblastoma (Rb/E2F/DP) pathway, which is present in animals and plants but absent in yeast (Rubin et al. 2000). Comparative studies of the cell cycle among model organisms belonging to different eukaryotic lineages can, thus, provide crucial information in distinguishing between the core cell cycle common to all phyla and lineage-specific adaptations. Most of the comparative analysis on the cell cycle regulation in ophistokonts (metazoans and fungi) has already yielded the identification of several evolutionarily conserved cell cycle control genes. Although many of these genes are also known in higher plants, their precise role is hard to grasp because of the high complexity of the plant model genomes; namely, the presence of multiple copies of key genes such as CDKs and cyclins. For example, genome-wide analysis shows that cell division control might involve nine CDKs (one of CDKA, four of CDKB, three of CDKD, and one of CDKF) and 30 cyclins (10 of cyclin A, nine of cyclin B, 10 of cyclin D, and one of cyclin H) in *Arabidopsis thaliana* (Vandepoele et al. 2002). The function of each copy is very difficult to investigate because their independent roles are blurred: silencing one copy does not necessarily yield the complete phenotype associated with the gene, as part or all of the function of the silenced copy can be rescued by the other copies. Thus, there is a need for a simpler green lineage–specific model organism that can be used to unravel the cell cycle specificities of this phylum. Furthermore, studies in the major "classical" model organisms are not sufficient to account for the common features and the particularities of each model. It is, for instance, difficult to determine whether the presence of only one CDK in yeast is a primeval feature inherited from the ancestral eukaryotic cell or a more recent simplification after the separation between the ophistokonts and the green lineage. These questions can only be answered by the study of new model organisms that occupy key phylogenetic positions. Undoubtedly the genome-wide analysis of their cellular functions, such as the core cell cycle genes, will help in the understanding of the complex green lineage–specific adaptations.

Ostreococcus tauri is a marine unicellular green alga of the Prasinophyceae clade that belongs to the Chlorophyta group of the Planta kingdom (Courties et al. 1998). Because Prasinophyceae have diverged early at the base of the Chlorophyta and consequently of the green lineage (Bhattacharya and Medlin 1998), *O. tauri* holds a key phylogenetic position in the eukaryotic tree of life. It is,

Steven Robbens and Basheer Khadaroo have participated equally to this work.

[1]Present address: Molecular Biology Group, Trieste, Italy.

Key words: cell division cycle, cyclin-dependant kinase, cyclin, green alga, *Ostreococcus tauri*

E-mail: h.moreau@obs-banyuls.fr.

Mol. Biol. Evol. 22(3):589–597. 2004
doi:10.1093/molbev/msi044
Advance Access publication November 10, 2004

therefore, a potentially powerful model to differentiate between the processes that are common to all eukaryotes (i.e., inherited from the "ancestral eukaryotic cell") and specific adaptations that have occurred after the separation of the different lineages. *O. tauri* is the smallest free-living eukaryotic cell described to date (Courties et al. 1994), with a diameter of no more than 1 μm. Furthermore, *O. tauri* has a minimal cellular organization, with a nucleus, a single chloroplast, and only one mitochondrion (Chrétiennot-Dinet et al. 1995). It has a nude plasma membrane without scales or flagella, a reduced cytoplasm (Chrétiennot-Dinet et al. 1995), and a small 12.5-MB to 13-Mb genome, which is currently being sequenced (Derelle et al. 2002). The high-throughput sequencing step of its complete genome is now finished (data not shown), and first analyses indicate that most of the genes have high similarities with genes belonging to the higher plant lineage and can, thus, be easily annotated by sequence similarity. Here, we compare the core cell cycle genes of *O. tauri* with those of *A. thaliana* and discuss new features of their evolution.

Materials and Methods

Ostreococcus tauri Cultures

The *O. tauri* culture used in this study is the strain OTTH0595 (Courties et al. 1998). Cultures were grown in Keller medium (Sigma), diluted in 0.2 μm filtered Banyuls bay–sampled sea water (NaCl 38 g/L), at a temperature of 18°C, with a permanent irradiance of 60 μmol quanta/m^2/s, and under mild agitation. Growth was followed by flow cytometer analysis.

Annotation of the *Ostreococcus tauri* Cell Cycle Genes

All the genes were annotated based on their similarity with other cell cycle genes available in the public databases. These sequences were aligned using Blast (Altschul et al. 1990) against the *O. tauri* database, and the best hits were further manually annotated using Artemis (Rutherford et al. 2000). The mRNA expression of all the genes described in this study has been confirmed either by their presence among the ESTs sequenced from a cDNA library or from Northern blots or RT-PCR. All the sequences reported in this paper have been submitted to GenBank under the following accession numbers: AY675093 (CDKA), AY675094 (CDKB), AY675095 (CDKC), AY675096 (CDKD/CAK), AY675097 (CycA), AY675098 (CycB), AY675099 (CycD), AY675100 (CycH), AY330645 (Cdc25), AY675101 (Wee1), AY675102 (Rb), AY675103 (E2F), AY675104 (Del), AY675105 (Dp), AY675106 (Cks), AY675107 (Cdc20), AY675108 (CDH1/CCS52), AY675109 (Skp1), AY675110 (Apc1), AY675111 (Apc2), AY675112 (Apc5), AY675113 (Apc6/Cdc16), AY675114 (Apc10), AY675115 (Cdc26p), AY675116 (Apc7), AY675 117 (Apc8/Cdc23), AY675118 (Apc11), AY675119 (Apc4), and AY675120 (Apc3).

Phylogenetic Analysis

Sequences were aligned with ClustalW (Thompson, Higgins, and Gibson 1994). The sequence alignments were manually improved using BioEdit (Hall 1999). TreeCon (Van de Peer and De Wachter 1997) was used for constructing the neighbor-joining (Saitou and Nei 1987) trees based on Poisson-corrected distances, only taking into account unambiguously aligned positions. Bootstrap analysis with 500 replicates was performed to test the significance of the nodes.

Results and Discussion

Ostreococcus tauri Genome Status

The genome of the *O. tauri* strain OTH95 has been sequenced using the random sequencing method completed by an oriented walking strategy (data not shown). Approximately 120,000 reads corresponding to the extremities of 60,000 shotgun clones and 5,500 reads of the extremities of a BAC library have been assembled using the Phred-Phrap package. A total of 1,989 contigs longer than 2 kbp were obtained for an overall sevenfold depth of coverage. At this stage, specific oligonucleotides were designed at the extremities of the biggest contigs and used to specifically sequence the shotgun clones flanking these contigs. All the contigs obtained were physically located on chromosomes by using a direct hybridization approach on pulse field electrophoresis gels. A total of 5,441 nuclear protein-coding genes were identified using the EuGene gene prediction software version 1.64 (Schiex et al. 2001), which includes both intrinsic and extrinsic approaches for better performance. The mitochondria and chloroplast genomes have also been determined.

The nuclear genome size of *O. tauri* has been estimated at approximately 12.6 Mbp by pulsed-field electrophoresis (PFGE), whereas the total size obtained from the contigs is 12.4 Mbp. A total of 1,850 unique *O. tauri* expressed sequence tags (ESTs) have been sequenced, and around 99% of these sequences mapped with identity greater than 95% onto the genome by using Blast. This is further evidence of the completeness of the genome sequence. and this genome draft has then been used for the complete analysis of the *O. tauri* core cell cycle genes.

CDK-Cyclin Complexes

CDKs are universally conserved cell cycle regulators. Six classes of CDKs have been described in the plant model *A. thaliana* (Vandepoele et al. 2002). CDKA has a PSTAIRE cyclin-binding motif and is the plant ortholog of the universal eukaryotic cell cycle regulator CDK1 (Dorée and Hunt 2002). CDKB, whose expression is cell cycle regulated, belongs to a plant-specific CDK clade and plays a role at the G_2/M-phase transition. CDKD and CDKF are CDK activating kinases (CAK), which activate the CDK by phosphorylating the threonine residue in the T-loop (Jeffrey et al. 1995). CDKC and CDKE are not directly involved in the cell cycle control. According to this plant nomenclature (Joubès et al. 2000), four CDKs belonging to the A to D classes have been found in the genome of *O. tauri* (fig.1 and table 1), but only three (CDKA, CDKB, and CDKD) are involved in cell division control.

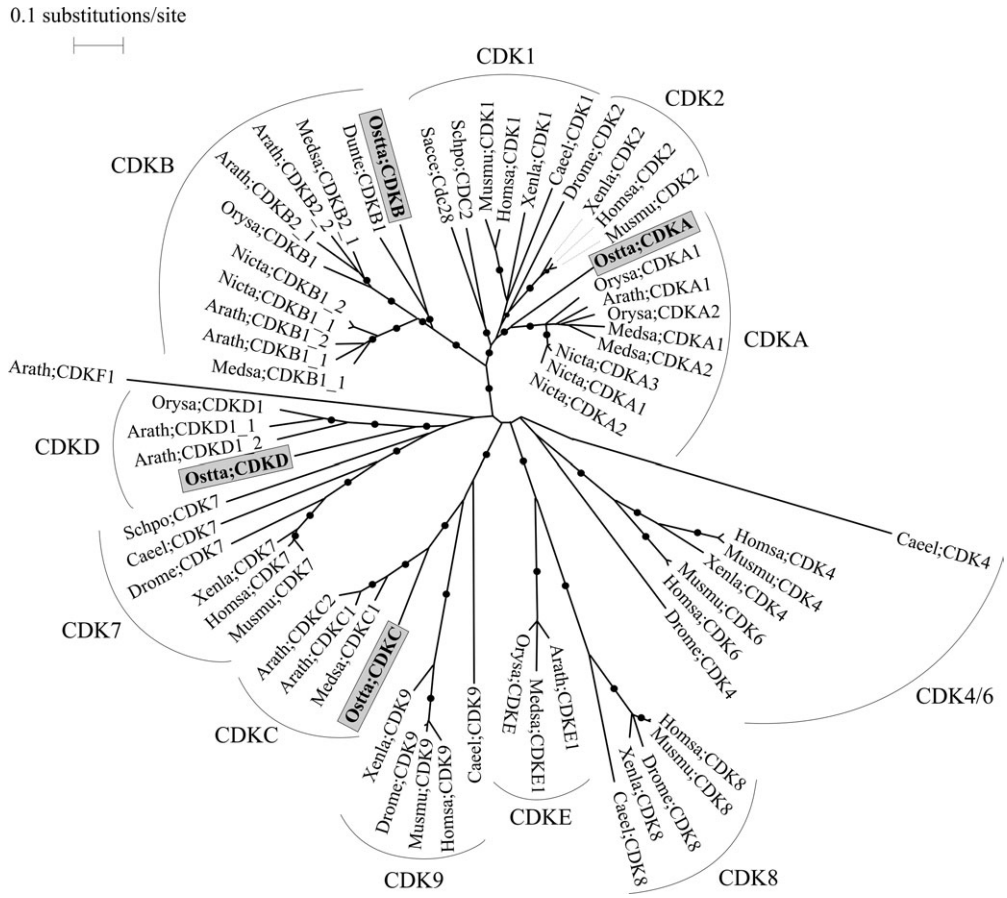

FIG. 1.—The CDK gene family. Unrooted neighbor-joining tree inferred from Poisson-corrected evolutionary distances for the CDK gene family involved in the cell cycle. The black dots indicate bootstrap values above 70 out of 500 samples. Arath: *Arabidopsis thaliana*; Medsa: *Medicago sativa*; Nicta: *Nicotiana tabacum*; Orysa: *Oryza sativa*; Dunte: *Dunaliella tertiolecta*; Sacce: *Saccharomyces cerevisiae*; Schpo: *Schizosaccharomyces pombe*; Musmu: *Mus musculus*; Homsa: *Homo sapiens*; Xenla: *Xenopus laevis*; Caeel: *Caenorhabditis elegans*; Drome: *Drosophila melanogaster*; Ostta: *Ostreococcus tauri*.

O. tauri CDKA contains the canonical PSTAIRE motif that is the hallmark of the central cell cycle regulator whose orthologs are Cdc2 in *S. pombe*, Cdc28 in *S. cerevisiae*, CDK1 in vertebrates, and CDKA in plants. This intronless gene is well conserved when compared with the PSTAIRE CDKs of other organisms and shows 66% sequence identity with the CDKA of *A. thaliana*. Moreover, the latter has four copies of the plant-specific B-class CDKs, whereas *O. tauri* has only one copy of a B-class–like CDK. *O. tauri* CDKB is intronless and contains a novel PSTALRE motif that is midway between the PSTAIRE CDKA and the P[S/P]T[A/T]LRE CDKB motifs. However, its overall sequence similarity and phylogenetic position confirms that this *O. tauri* gene is orthologous to higher plant CDKBs and is clearly not a divergent CDKA (fig. 1). Both *O. tauri* CDKA and CDKB have diverged early in their respective clade, confirming the phylogenetic position of *O. tauri* at the base of the green lineage (Courties et al. 1998). Furthermore, *O. tauri* has only one copy of CDKD bearing an NFTAIRE motif, homologous to the CDK-activating kinase (CAK), which positively regulates the activity of CDKA by phosphorylation of the threonine-161 residue. It is orthologous to the three CDKDs of *A. thaliana* (fig.1 and table 1).

Finally, only one PITAIRE motif *O. tauri* CDKC has been identified, as compared with the two CDKCs described in *A. thaliana* (fig.1 and table 1). PITAIRE CDKs have been reported to phosphorylate the carboxyl terminal domain (CTD) of the RNA polymerase II, and

Table 1
Comparison of CDK Genes of Metazoans, Yeasts, Plants, and *Ostreococcus tauri*

Genes	*O. tauri*	*A. thaliana*	*S. cerevisiae*	*H. sapiens*	*O. tauri*	*A. thaliana*	*S. cerevisiae*	*H. sapiens*
CDKA	1	1	1	3	PSTAIRE	PSTAIRE	PSTAIRE	PSTAIRE
CDKB	1	4	0	0	PSTALRE	P[S/P]T[A/T]LRE	—	—
CDKC	1	2	1	1	PITAIRE	PITAIRE	PITAIRE	PITALRE
CDKD	1	3	0	1	NFTAIRE	N[I/F/V]TALRE	—	NRTALRE
CDKF	—	1	1	1	—	YQSAFRE	PHNAKFE	PNQALRE

Table 2
Comparison of Green Lineage Cyclin Genes

	Number of Copies		Protein Motifs	
Genes	*O. tauri*	*A. thaliana*	*O. tauri*	*A. thaliana*
			LxCxE	
CycA	1	10	MIEVxEEY	No LxCxE
			MRNILVDW	
			HxKf	
CycB	1	9	MRAILIDW	Hx[R/K]F
CycD	1	10	No LxCxE	LxCxE except D4 and D6
CycH	1	1	IVRHEAK	MRAFYEAK

Table 3
Comparison of Several Core Cell Cycle Genes

	Number of Copies			
Genes	*O. tauri*	*A. thaliana*	*S. cerevisiae*	*H. sapiens*
Cdc25	1	1[a]	1	3
Wee1/Myt1/Mik1	1	1	2	2
Rb/p107/p130	1	1	0	3
E2F	1	3	0	6
DEL	1	3	0	1
DP	1	2	0	2
Cks	1	2	1	2
CKI	1?	7	1	8

[a] Very poorly conserved.

they do not participate directly in the cell cycle control (Barroco et al. 2003). No E-class or F-class CDKs have been found in the genome of *O. tauri*.

Cyclins are the regulatory binding partners of the CDKs, which confer the timing and substrate specificity to the activity of the CDK-cyclin complexes (Futcher 1996). *O. tauri* has the minimum set of cyclins described to date in any organism. More importantly, it has only one copy of each of the A-class, B-class, D-class, and H-class cyclins (Renaudin et al. 1996), thus, presenting even fewer cyclin gene copies than yeasts because *S. cerevisiae* has three G_1 cyclins (CLN) and six $S/G_2/M$ cyclins (CLB) (table 2 and fig. 2). The activities of the different cyclins in *S. cerevisiae* are redundant. In G_1-phase, threshold Cln3 kinase activity is necessary for going through the START point, at which time it switches on the other two G_1 cyclins, Cln1 and Cln2. The peak activity of these two cyclins induces the activity of the S-phase specific CDK-cyclin complexes by releasing the Clb5-associated and Clb6-associated kinases from the inhibitory CKI Sic1 (Schwob et al. 1994). The Clb5 and Clb6 kinase activities are necessary for progression from the G_1 to the end of the S-phases. Finally, the cyclins Clb1-4 are turned on at the G_2/M-phase transition, thus, leading the cell into M-phase (Mendenhall and Hodge 1998).

In *O. tauri*, there is only one putative G_1 cyclin, the cyclin D, which contains one intron (table 2 and fig. 2). Surprisingly, the LxCxE retinoblastoma (Rb) binding motif that is normally present on cyclin D of animals and higher plants is not found on the putative *O. tauri* cyclin D but has been identified in the C-terminal part of the *O. tauri* cyclin A. *O. tauri* cyclin A also contains the well-conserved cyclin box motif MRNILVDW and a MIEVAEEY cyclin A–specific motif similar to the *A. thaliana* MRx[I/V]L[I/V]DW and LVEVxEEY cyclin A motifs. This gene has one intron, and it shares the highest homology of 26% sequence identity with the *A. thaliana* cyclin A2 (accession number PIR: D96505). The putative M-phase cyclin B has two introns, and it contains the well-conserved cyclin box motif MRAILVDW and the cyclin B–specific HxKF motif. Hence, this in silico analysis suggests that only one cyclin would be sufficient for each specific phase of the cell cycle, whereas two or more cyclins are present in *S. cerevisiae*, and even more genes are present in the multicellular organisms. Last, the cyclin H, which is the regulatory subunit of the CDKD, does not have an intron, and its sequence analysis yields a cyclin domain similar to homologs of cyclin H from *A. thaliana* with no particular feature (table 2 and fig. 2).

Therefore, *O. tauri* presents a minimal, yet complete, set of cell division control genes necessary to drive a eukaryotic cell through the complete division cycle: one of CDKA, one of CDKB, one of cyclin A, one of cyclin B, and one of cyclin D. Furthermore, the presence in this organism of two CDKs (the universal regulator CDKA and the green lineage–specific CDKB) supports the hypothesis that having one CDK would be a primeval feature that has been conserved in yeast but not a more recent simplification specifically acquired in this lineage. Furthermore, the simplification to only one copy for each cyclin type, in contrast to the usual high copy number of these genes in the other organisms, makes *O. tauri* a potentially powerful model for functional plant core cell cycle studies.

CDK subunits (CKS) are proteins that bind to the CDK protein, and their function is important in the transcriptional activation of Cdc20, the activating protein of the APC complex (see below) (Morris et al. 2003). Only one putative Cks has been found in the genome of *O. tauri* (table 3). This gene has four introns, and the encoded protein has a well-conserved N-terminus sharing an overall 78% and 82% similarity with *A. thaliana* Cks1 and Cks2, respectively.

CDK inhibitor (CKI) Kip-related proteins (KRPs) are inhibitors of CDK activities, and seven KRP genes have been found in *A. thaliana* (De Veylder et al. 2001). Despite many efforts, no such genes and no other related CDK inhibitors could be found in *O. tauri* by sequence similarity searches. Only a highly divergent sequence, sharing low sequence similarity to the KRP3 of *A. thaliana* has been found as a putative candidate (table 3). However, KRP genes are usually very divergent; for example, only few key amino acids are conserved between *A. thaliana* and animal inhibitors (De Veylder et al. 2001), and likewise for the CKI between the *S. cerevisiae* Sic1 and *S. pombe* Rum1 inhibitors (Sanchez-Diaz et al. 1998). Furthermore, no sequence conservation was found between the CKI from yeast and other lineages. This absence or very low sequence similarity means that the only possibility of identifying the *O. tauri* cell cycle inhibitors will be by using functional genetic and/or biochemical approaches.

Retinoblastoma (Rb/E2F/DP) Pathway

At the G_1/S transition, the Rb pathway is conserved among the animal and plant kingdoms but is absent in

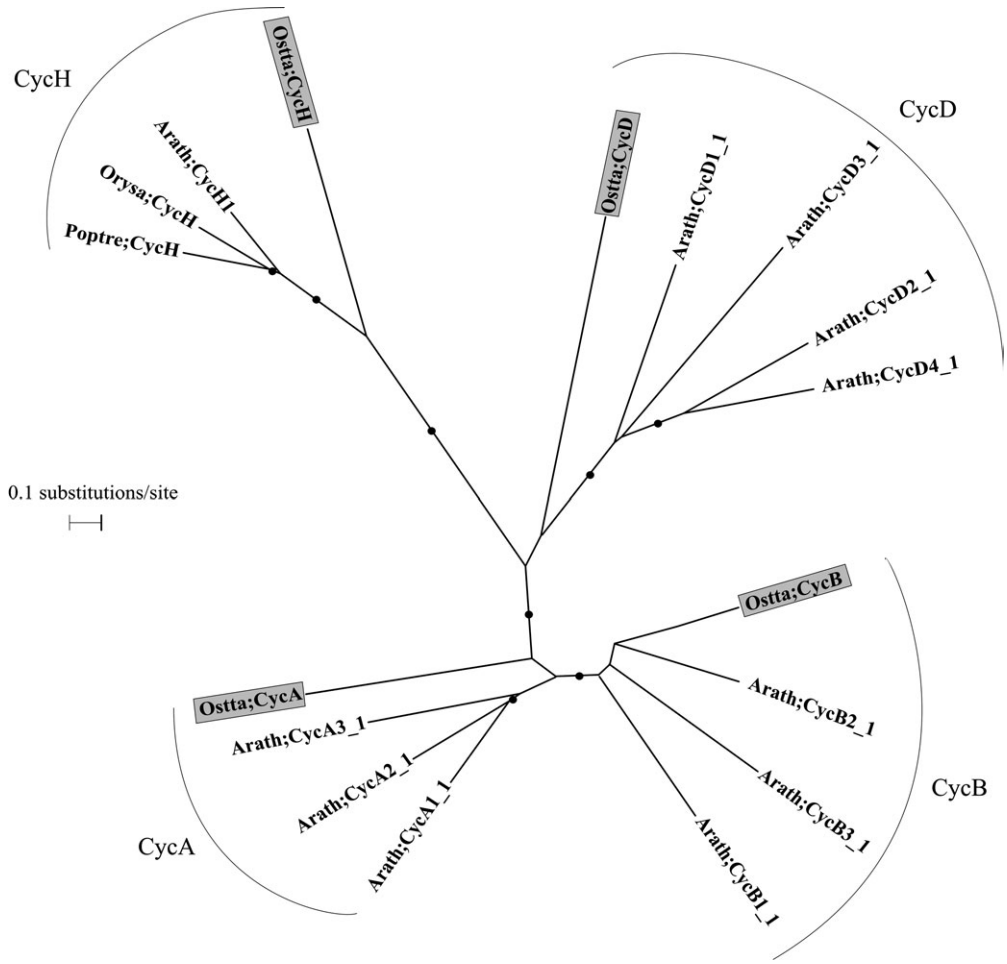

FIG. 2.—The cyclin gene family. Unrooted neighbor-joining tree inferred from Poisson-corrected evolutionary distances for the cyclin gene family involved in the cell cycle. The black dots indicate bootstrap values above 70 out of 500 samples. Arath: *Arabidopsis thaliana*; Orysa: *Oryza sativa*; Poptre: *Populus tremula*; Ostta: *Ostreococcus tauri*.

yeasts. An Rb homolog has also been described in the unicellular green alga *Chlamydomonas reinhardtii* (Umen and Goodenough 2001). However, this alga has a peculiar cell division, and the question of whether this Rb pathway is characteristic for multicellular organisms has remained unclear (Cross and Roberts 2001). In algae, plants, and metazoans, the retinoblastoma pathway induces the expression of S-phase–specific genes. The Rb protein sequesters and inactivates the DNA-bound heterodimer transcription factor comprising the E2F protein and its dimerization partner (DP) protein (Weinberg 1995). It also recruits the chromatin remodeling machinery for silencing the target genes (Shen 2002) and is responsible for maintenance of quiescence (Sage et al. 2003). At the G_1/S transition, the CDK-cyclin complex phosphorylates Rb. Once phosphorylated, the Rb protein frees the E2F/DP complex, which is able to induce the transcription of its target genes. In contrast to vertebrates, which have three copies of pocket proteins (Rb, p107, and p130), *O. tauri*, similar to *A. thaliana* and *C. reinhardtii*, has only one homolog of Rb gene (fig. 2 and table 3).

The E2F family of transcription factors comprise the subfamilies of activating E2Fs, inhibitory E2Fs, dimerization partner proteins (DPs), and DP-like and E2F-like proteins (DELs) (Shen 2002). The three E2F, two DP, and three DEL genes identified in *A. thaliana* have approximately 22% overall sequence similarity (Vandepoele et al. 2002). E2FA and E2FB have four binding domains, a DNA-binding, a DP-binding, an Rb-binding, and a transactivation-binging domain. When bound to DPA or DPB proteins, they are transcriptional activators that are repressed by Rb protein. E2FC lacks the transactivation-binding domain and is a homolog of animal inhibitory E2Fs E2F4-6. Also, the three *A. thaliana* DEL proteins, of which E2F7 is the recently described animal homolog (Di Stefano, Jensen, and Helin 2003), each have two DNA-binding domains but do not have either DP-binding, Rb-binding, or transactivation-binding domains. Hence, DEL proteins contribute to the class of E2F inhibitory subfamilies. Only one E2F, one DP, and one DEL gene have been identified in the genome of *O. tauri* by alignments of their sequences with their orthologs from *A. thaliana* (figs. 3 and 4). The phylogenetic analysis of the E2F family confirms the early divergence of *O. tauri* genes with respect to the higher plants (fig. 4). The *O. tauri* E2F is a homolog of activating E2Fs and has a conserved binding domain to the DP-binding and Rb-binding domains and DNA-binding and transactivation-binding

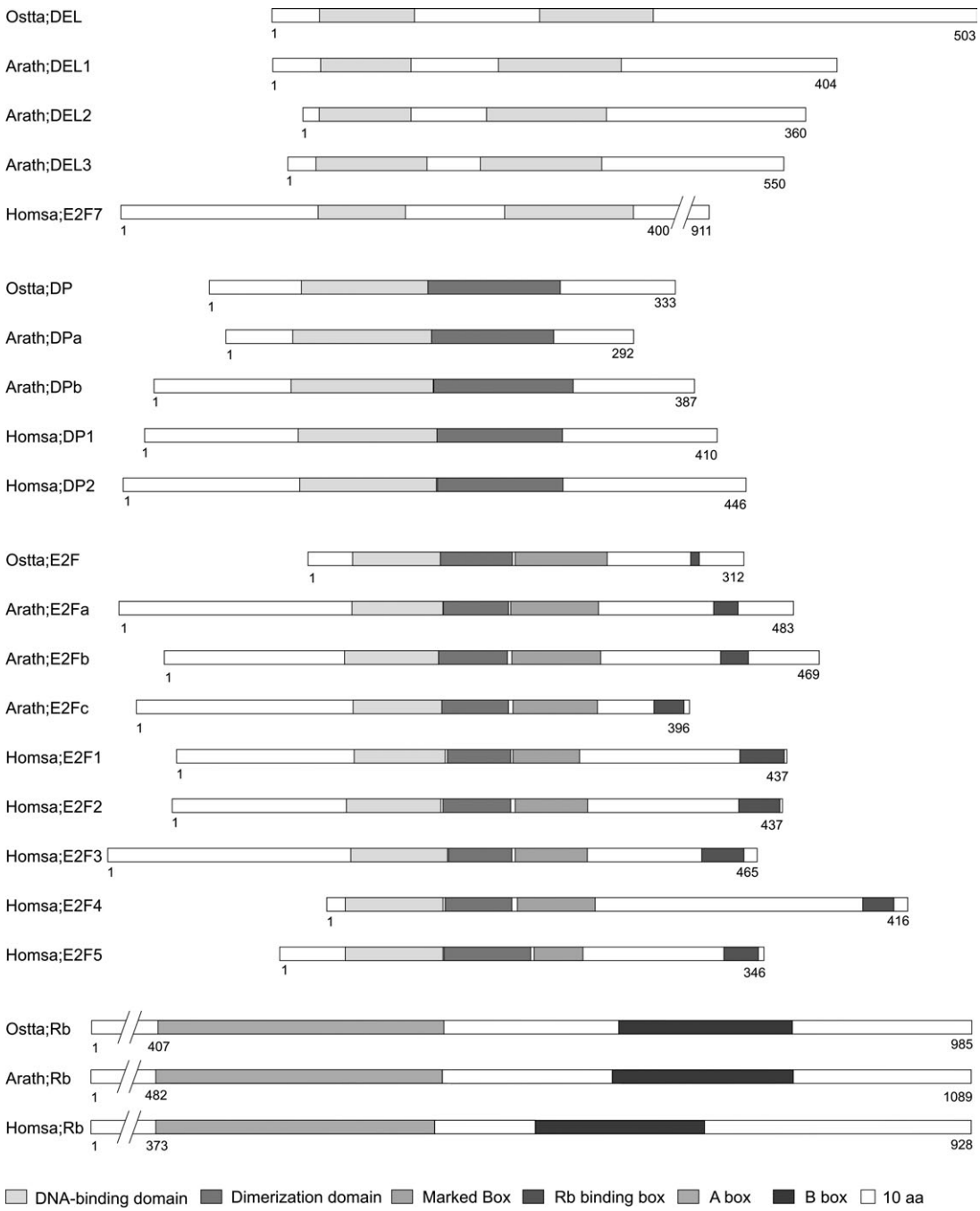

FIG. 3.—Schematic representation of the E2F family and Rb genes. Structural organization of the DEL, DP, E2F, and Rb proteins of *Ostreococcus tauri* (Ostta) compared with *Arabidopsis thaliana* (Arath) and *Homo sapiens* (Homsa). The DNA-binding, dimerization, Marked, and Rb- binding domains are indicated with colored boxes.

domains, whereas DEL has only two DNA-binding domains but no other domain (fig. 3).

Furthermore, the three *O. tauri* genes have diverged very early in each group, as observed for the CDK and cyclin genes. As for the CDKs and cyclins, *O. tauri* has the minimal but complete set of genes for the Rb pathway comprising one Rb gene, one E2F gene, one DP gene, and one DEL gene. Finding the Rb pathway in *O. tauri* confirms that its presence in *C. reinhardtii* is evolutionarily conserved, and the more parsimonious explanation is that Rb was present in the ancestral eukaryotic cell and has been subsequently lost in the yeast phylum. Recently, an unrelated gene called Whi5, which substitutes the role of Rb in yeast at the G_1/S transition by inhibiting the transactivation activity of the SBF and MBF transcription factors, reinforces the hypothesis of the loss of the Rb gene family (Costanzo et al. 2004).

Cdc25/Wee1 Control of CDK-Cyclin Activity

The kinase Wee1 and the phosphatase Cdc25 regulate the G_2/M transition in metazoans and yeasts by

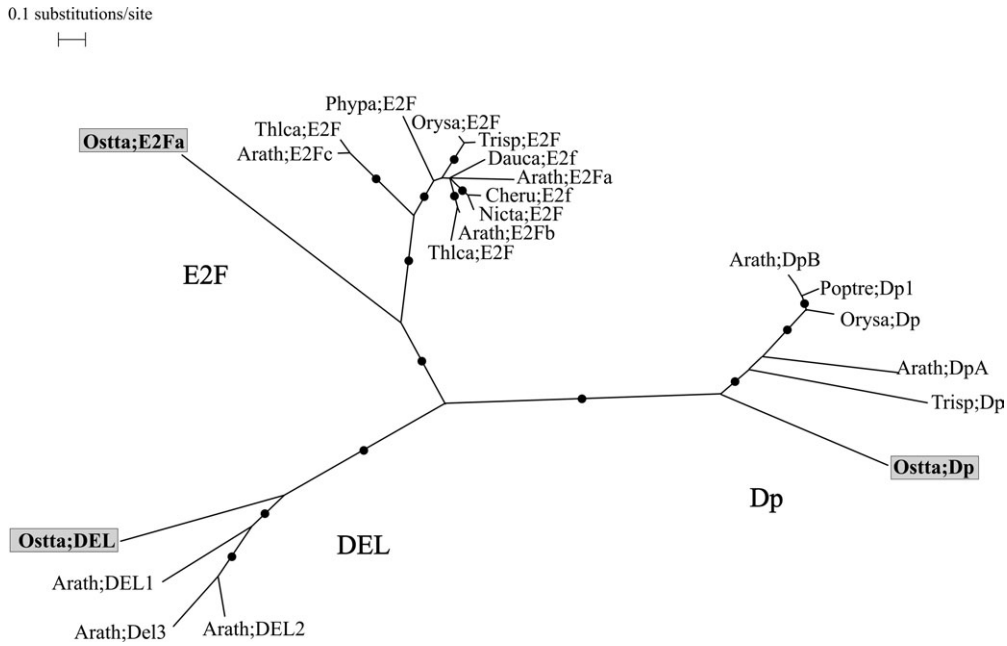

FIG. 4.—The E2F gene family. Unrooted neighbor-joining tree inferred from Poisson-corrected evolutionary distances for the E2F, Dp, and DEL families. The black dots indicate bootstrap values above 70. Arath: *Arabidopsis thaliana*; Nicta: *Nicotiana tabacum*; Orysa: *Oryza sativa*; Cheru: *Chenopodium rubrum*; Dauca: *Daucus carota*; Thlca: *Thlaspi caerulescens*; Trisp: *Triticum sp.*; Phypa: *Physcomitrella patens*; Poptre: *Populus tremula*; Ostta: *Ostreococcus tauri*.

posttranslational regulation of the CDK-cyclin complex. In plants, an ortholog of the wee1 gene has been found, but the cdc25 gene has never been identified either in the fully sequenced genomes of higher plants such as *A. thaliana* and rice or in that of the unicellular green alga *C. reinhardtii*. The absence of the M-phase inducer Cdc25 phosphatase in plants is puzzling because the other actors of this regulation pathway, namely the CDK-cyclin B complex and the Wee1 kinase, are evolutionarily conserved. Furthermore, the activating dephosphorylation at the M-phase entry is conserved in plants (Zhang, Letham, and John 1996; McKibbin, Halford, and Franci 1998).

An intronless putative wee1 gene has been identified in the genome of *O. tauri* (table 3). It is similar to the Wee1 kinase of *A. thaliana*, maize, and *C. reinhardtii* and potentially inhibits the activity of the CDK-cyclin complex by its inhibitory phosphorylation. Surprisingly, the ortholog of the activating phosphatase cdc25 gene has been identified in *O. tauri* (table 3). It is the first time that cdc25 is described in the green lineage, and it shows that the cdc25 gene is present at the base of this lineage. The *O. tauri* cdc25 gene codes for a protein that is able to rescue the *S. pombe* c*dc25-22* conditional mutant. Furthermore, microinjected *O. tauri* Cdc25 specifically activates starfish oocytes arrested in the prophase of the first meiotic division, thus, causing germinal vesicle breakdown. In vitro phosphatase assays, namely antiphosphotyrosine Western blotting and the histone H1 kinase assay confirmed the in vivo activity of *O. tauri* Cdc25 (Khadaroo et al. 2004).

The presence of the first functional green-lineage Cdc25 dual-specificity phosphatase discovered in *O. tauri* indicates that this gene was probably present in the ancestor of the eukaryotic cell. In this respect, it should be noted that a putative Cdc25-like gene has also been identified in the complete genome of *C. reinhardtii*. Furthermore, because its activity is necessary in higher plants, the most parsimonious hypothesis is that the sequence of Cdc25 in higher plants has diverged so much that it can no longer be recognized by sequence homology analysis (Khadaroo et al. 2004). This has been very recently confirmed by the identification in *A. thaliana* of a poorly conserved Cdc25-related protein having a tyrosine-phosphatase activity stimulating the kinase activity of *A. thaliana* CDKs (Landrieu et al. 2004).

Ubiquitin Ligases APC and SCF

Both the anaphase promoting complex (APC) and Skp1/Cdc53/F box protein (SCF) complex, which are the key enzymes for tagging proteins in the ubiquitin pathway, are the E3 ligases responsible for the specificity of protein degradation by the 26S proteasome (Irniger 2002; Cope and Deshaies 2003). The two evolutionarily conserved functions of the APC is the cell cycle–specific targeting of securin and the mitotic cyclin B for proteasome-mediated destruction. The APC is composed of at least 13 protein subunits initially identified in yeast and animals (Schwickart et al. 2004). All vertebrate APC subunits have their homologs in plants (Capron, Okresz, and Genschik 2003; Tarayre et al. 2004). Interestingly, although many genes (and notably cell cycle genes [see above]) are present as multiple copies in *A. thaliana*, its APC genes are present as single copies, except for APC3, which is represented by two slightly different genes (table 4, modified from Capron, Okresz, and Genschik [2003]). In *O. tauri*, putative orthologs of most of these genes have also been found as single copies, including only one copy of APC3. They show very high similarity scores with the *A. thaliana* APC genes. Furthermore,

Table 4
Evolutionary Conservation of the Subunits of the APC E3 ligase

APC subunits	Number of Copies				Protein Motifs
	O. tauri	*A. thaliana*	*S. cerevisiae*	*H. sapiens*	
APC1	1	1	1	1	Rpn 1 and 2 proteasome repeats
APC2	1	1	1	1	Cullin domain
APC3	1	2	1	1	TPR repeats
APC4	1	1	1	1	
APC5	1	1	1	1	
APC6/Cdc16	1	1	1	1	TPR repeats
APC7	1	1	Unclear	1	TPR repeats
APC8/Cdc23	1	1	1	1	TPR repeats
APC9	Unclear	Unclear	1	Unclear	
APC10/Doc1p	1	1	1	1	Doc domains
APC11p	1	1	1	1	Ring-H2 domain
Mnd2p	Unclear	1	1	1	
Swm1p	1	1	1	1	
Cdc26p	1	1	1	1	
Activators APC					
Cdc20	1	5	1	1	
CDH1/Ccs52	1	3	1	2	

putative orthologs of the two activators (Cdc20 and Cdh1) regulating the activity of the APC complex have also been found in the *O. tauri* genome. In contrast with the many putative orthologs of Cdc20 and Cdh1 in *A. thaliana* (Capron, Okresz, and Genschik 2003), there seems to be only one Cdc20 and one Cdh1 gene in *O. tauri* (table 4). Thus, *O. tauri* has a complete set of APC genes, with a minimal number of activators.

The Skp1/Cullin/F (SCF) box protein is the other evolutionarily conserved E3 ligase that is responsible for cell cycle control; namely, targeting the CKI for proteasome-mediated proteolysis, which is essential for the cell to progress in late G_1-phase (Deshaies and Ferrell 2001). The annotation of SCF genes of *O. tauri* has shown only one Skp1 gene and four cullin-domain proteins, of which two are putative Cdc53 genes and one is the Skp2 putative gene. Skp2 has been identified by using the *A. thaliana* Skp2 gene (Del Pozo, Boniotti, and Gutierez 2002), which contained both an F-box and a leucine-rich domain. These putative genes need to be functionally assayed to confirm this annotation.

Once more, *O. tauri* presents a conserved set of APC and SCF genes with a lower copy number of genes than in *A. thaliana*.

Conclusion

The data above reveal that the cell cycle control in the unicellular marine green alga *Ostreococcus tauri* is the simplest described to date (and one of the most complete across the different lineages). It has the minimum set of cyclins for driving the cell cycle and has indeed conserved the Rb pathway, which has been lost in the yeast phylum. It has also retained the plant-specific B-class CDK, and it presents the first green-lineage Cdc25 phosphatase, which has only been identified as a very poorly related gene in higher plants. *O. tauri* displays the minimum, yet complete, set of core cell cycle, and its functional analysis will undoubtedly yield valuable information providing a clear picture not blurred by the activity of other functionally redundant members of the gene family.

Acknowledgments

We thank Dr. W. Zacchariae for precious help in the annotation of E3 ligases and K. Vandepoele and Dr. P. Rouzé for their technical help. This work has been supported by the Génopole Languedoc-Roussillon, the CNRS, the Marie Curie predoctoral fellowship number HPMT-CT-2000-00211 for S.R. and the Ph.D. scholarship from the French Embassy in Mauritius for B.K.

Literature Cited

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

Barroco, R. M., L. De Veylder, Z. Magyar, G. Engler, D. Inze, and V. Mironov. 2003. 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. **60**:401–412.

Bhattacharya, D., and L. Medlin. 1998. Algal phylogeny and the origin of land plants. Plant Physiol. **116**:9–15.

Capron, A., L. Okresz, and P. Genschik. 2003. First glance at the plant APC/C, a highly conserved ubiquitin-protein ligase. Trends Plant Sci. **8**:83–89.

Chrétiennot-Dinet, M. J., C. Courties, A. Vaquer, J. Neveux, H. Claustre, J. Lautier, and M. C. Machado. 1995. A new marine picoeukaryote: *Ostreococcus tauri* gen. et sp. Nov. (Chlorophyta, Prasinophyceae). Phycologia **4**:285–292.

Cope, G. A., and R. J. Deshaies. 2003. COP9 signalosome: a multifunctional regulator of SCF and other cullin-based ubiquitin ligases. Cell **114**:663–671.

Costanzo M, J. L. Nishikawa, X. Tang, J. S. Millman, O. Schub, K. Breitkreuz, D. Dewar, I Rupes, B. Andrews, and M. Tyers. 2004. CDK Activity Antagonizes Whi5, an Inhibitor of G1/S Transcription in Yeast. Cell **117**:899–913.

Courties, C., R. Perasso, M. J. Chrétiennot-Dinet, M. Guoy, L. Guillou, and M. Troussellier. 1998. Phylogenetic analysis

and genome size of *Ostreococcus tauri* (Chlorophyta, Prasinophyceae). J. Phycol. **34**:844–849.

Courties, C., A. Vaquer, M. Troussellier, J. Lautier, M. J. Chrétiennot-Dinet, J. Neveux, M. C. Machado, and H. Claustre. 1994. Smallest eukaryotic organism. Nature **370**:255.

Cross, F. R., and J. M. Roberts. 2001. Retinoblastoma protein: combating algal bloom. Curr. Biol. **11**:R824–R827.

Del Pozo, J. C., M. B. Boniotti, and C. Gutierez. 2002 *Arabidopsis* E2Fc functions in cell division and is degraded by the ubiquitin-SCF (AtSKP2) pathway in response to light. Plant Cell **14**:3057–3071.

Derelle, E., C. Ferraz, P. Lagoda et al. (12 co-authors). 2002. DNA libraries for sequencing the genome of *Ostreococcus tauri* (chlorophyta, prasinophyceae): the smallest free-living eukaryotic cell. J. Phycol. **38**:1150–1156.

Deshaies, R. J., and J. E. Ferrell Jr. 2001. Multisite phosphorylation and the countdown to S phase. Cell **107**:819–822.

De Veylder, L., T. Beeckman, G. T. Beemster et al. (10 co-authors). 2001. Functional analysis of cyclin-dependent kinase inhibitors of *Arabidopsis*. Plant Cell **13**:1653–1668.

Di Stefano, L., M. R. Jensen, and K. Helin. 2003. E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. EMBO J. **22**:6289–6298.

Dorée, M., and T. Hunt. 2002. From Cdc2 to Cdk1: When did the cell cycle kinase join its cyclin partner? J. Cell Sci. **115**:2461–2464.

Futcher, B. 1996. Cyclins and the wiring of the yeast cell cycle. Yeast **12**:1635–1646.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41**:95–98.

Irniger, S. 2002. Cyclin destruction in mitosis: a crucial task of Cdc20. FEBS Lett. **532**:7–11.

Jeffrey, P. D., A. A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massague, and N. P. Pavletich. 1995. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature **376**:313–320.

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

Khadaroo, B., S. Robbens, C. Ferraz et al. (12 co-authors). 2004. The First Green Lineage cdc25 Dual-Specificity Phosphatase. Cell Cycle **3**:513–518.

Landrieu, I., M. Da Silva, L. De Vielder et al. (11 co-authors). 2004. A small CDC25 dual-specificity tyrosine-phosphatase isoform in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **101**:13380–13385.

McKibbin, R. S., N. G. Halford, and D. Francis. 1998. Expression of fission yeast cdc25 alters the frequency of lateral root formation in transgenic tobacco. Plant Mol. Biol. **36**:601–612.

Mendenhall, M. D., and A. E. Hodge. 1998. Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. **62**:1191–1243.

Morris, M. C., P. Kaiser, S. Rudyak, C. Baskerville, M. H. Watson, and S. I. Reed. 2003. Cks1-dependent proteasome recruitment and activation of CDC20 transcription in budding yeast. Nature **424**:1009–1013.

Obaya, A. J., and J.M. Sedivy. 2002. Regulation of cyclin-Cdk activity in mammalian cells. Cell. Mol. Life Sci. **59**:126–142.

Renaudin, J. P., J. H. Doonan, D. Freeman et al. (12 co-authors). 1996. Plant cyclins: a unified nomenclature for plant A-, B- and D-type cyclins based on sequence organization. Plant. Mol. Biol. **32**:1003–1018.

Rubin, G. M., M. D. Yandell, J. R. Wortman et al. (51 co-authors). 2000. Comparative genomics of the eukaryotes. Science **287**:2204–2215.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajadream, and B. Barrell. 2000. Artemis: sequence visualization and annotation. Bioinformatics **16**:944–945.

Sage, J., A. L. Miller, P. A. Perez-Mancera. 2003. Acute mutation of retinoblastoma gene function is sufficient for cell cycle re-entry. Nature **424**:223–228.

Saitou, N., and M. Nei. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

Sanchez-Diaz, A., I. Gonzalez, M. Arellano, and S. Moreno. 1998. The Cdk inhibitors p25rum1 and p40SIC1 are functional homologues that play similar roles in the regulation of the cell cycle in fission and budding yeast. J. Cell Sci. **111**:843–851.

Schiex, T., A. Moisan, and P. Rouzé. 2001. EuGène: an eucaryotic gene finder that combines several sources of evidence. Pp. 111-125 *in* O. Gascuel and M.-F. Sagot, eds. Computational Biology: Selected Papers (Lecture Notes in Computer Science), Vol. 2066. Springer-Verlag, Berlin.

Schwickart, M., J. Havlis, B. Haberman, A. Bogdanova, A. Camasses, T. Oelschlaegel, A. Shevenko, and W. Zachariae. 2004. Swm1/Apc13 is an evolutionarily conserved subunit of the anaphase-promoting complex stabilizing the association of Cdc16 and Cdc27. Mol. Cell. Biol. **24**:3562–3576.

Schwob, E., T. Bohm, M. D. Mendenhall, and K. Nasmyth. 1994. The B-type cyclin kinase inhibitor p40SIC1 controls the G1 to S transition in *S. cerevisiae*. Cell **79**:233–244.

Shen, W. H. 2002. The plant E2F-Rb pathway and epigenetic control. Trends Plant Sci. **7**:505–511.

Stals, H., and D. Inzé 2001. When plant cells decide to divide. Trends Plant Sci. **6**:359–364.

Tarayre, S., J. M. Vinardell, A. Cebolla, A. Kondorosi, and E. Kondorosi. 2004. Two classes of the Cdh1-type activators of the anaphase-promoting complex in plants: novel functional domains and distinct regulation. Plant Cell **16**:422–434.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Umen, J. G., and U. W. Goodenough. 2001. Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. Genes Dev. **15**:1652–1661.

Van de Peer, Y., and R. De Wachter. 1997. Construction of evolutionary distance trees with TREECON for Windows: accounting for variation in nucleotide substitution rate among sites. Comput. Appl. Biosci. **13**:227–230.

Vandepoele, K., J. Raes, L. De Veylder, P. Rouze, S. Rombauts, and D. Inze. 2002. Genome-wide analysis of core cell cycle genes in *Arabidopsis*. Plant Cell **14**:903–916.

Weinberg, R. A. 1995. The retinoblastoma protein and cell cycle control. Cell **81**:323–330.

Zhang, K., D. S. Letham, and P. C. John. 1996. Cytokinin controls the cell cycle at mitosis by stimulating the tyrosine dephosphorylation and activation of p34cdc2-like H1 histone kinase. Planta **200**:2–12.

Geoffrey McFadden, Associate Editor

Accepted November 2, 2004