A transcriptomic perspective of cell cycle regulation in dinoflagellates

David Morse*, Philip Daoust[1] and Siham Benribague

Institut de Recherche en Biologie Végétale,

Département de Sciences Biologiques,

Université de Montréal, Montréal, Québec, Canada

H1X 2B2

* Author to whom correspondence should be addressed

[1] Present address: Genome Quebec Innovation Center, Montreal, Canada

Key Words: Dinoflagellate, cell cycle, cyclin, cyclin-dependent kinase, S phase, M phase

Abstract

Dinoflagellates are a group of unicellular and generally marine protists, of interest to many because of their ability to form the large algal blooms commonly called “red tides”. The large algal concentrations in these blooms require sustained cell replication, yet to date little is known about cell cycle regulation in these organisms. To address this issue, we have screened the transcriptomes of two dinoflagellate species with budding yeast cell cycle pathway components. We find most yeast cell cycle regulators have homologs in the dinoflagellate, suggesting that the yeast model is appropriate for understanding regulation of the dinoflagellate cell cycle. The dinoflagellates are lacking several components essential in yeast, but a comparison with a broader phylogenetic range of protists reveals these components are usually also missing in other organisms. Lastly, phylogenetic analyses show that the dinoflagellates contain at least three cyclin-dependent kinase (CDK) homologs (belonging to the CDK1, CDK5 and CDK8 families), and that the dinoflagellate cyclins belong exclusively to the A/B type. This suggests that dinoflagellate CDKs likely play a limited role outside regulation of the cell cycle.

Introduction

Dinoflagellates are one of four phyla of protists within the Alveolata, a superphylum also containing the apicomplexans, the chromerids and the ciliates (Bachvaroff et al., 2011). These different groups all have flattened cortical vesicles termed alveolae and express the protein alveolin that specific to these organelles (Gould et al., 2008). However, dinoflagellates have several features that distinguish them from the other members of the alveolates, most notably with respect to their nuclear organization. Dinoflagellates can have large genomes, some reaching as high as 250 Gbp (Hou and Lin, 2009), which represents an amount roughly 80-fold greater than the haploid human genome. Furthermore, this extensive genetic baggage is organized within the nucleus differently from what is typically found. In most eukaryotes, nuclear DNA is wrapped around an octamer of histone basic proteins to form nucleosomes. Furthermore, nucleosomes can be stacked together with the degree of compaction depending on post-translational modifications of the

histone N terminal regions. The degree of compaction also varies over the cell cycle, with interphase chromosomes more spread out in order to be more readily accessible for replication and transcription and mitotic chromosomes more compact to facilitate separation of the sister chromatids. Dinoflagellate chromosomes differ from this in that they appear permanently condensed and neither histone proteins nor nucleosomes have ever been observed (Bodansky et al., 1979; Rizzo et al., 1982). Despite this, it is likely that histones serve a functional role, even if present at very low levels, since dinoflagellate transcriptomes contain conserved sequences encoding the four core histone genes as well as an array of histone modifying proteins (Bayer et al., 2012; Roy and Morse, 2012). Instead of histones, dinoflagellates employ other basic proteins to counteract the high level of negative charges on the DNA, including a small histone-like protein similar to bacterial HU (Wong et al., 2003) and a dinoflagellate/viral nuclear protein (DVNP) (Gornik et al., 2012). Interestingly, the basal lineage *Perkinsus marina* has neither DVNP nor permanently condensed chromosomes, while the early branching *Hematodinium* displays both characteristics, thus suggesting acquisition of DVNP accompanied formation of the unusual dinoflagellate nuclear phenotype. The DNA in the chromosomes has been proposed to be a liquid crystal (Rill et al., 1989), a structure possibly facilitated by the ten-fold reduction in protein/DNA ratio compared to other eukaryotes (Rizzo and Nooden, 1972).

In addition to this unusual chromosome structure, which might be expected to pose mechanical difficulties to replication and transcription, mitosis is also atypical. In almost all dinoflagellates, the nuclear envelope remains intact throughout mitosis and the spindle is cytoplasmic. In consequence, contact between the chromosomes and the cytoplasmic microtubule array is indirect and must involve proteins that span the nuclear envelope. Microtubules were observed to pass through cytoplasmic channels spanning the nucleus and to make contact with densely staining structures on the cytoplasmic face of the nuclear envelope (Bhaud et al., 2000; Fritz and Treimer, 1983; Oakley and Dodge, 1974, 1976). Typical spindle poles are not readily apparent in dividing cells.

In light of these unusual physical features, the question of how closely the regulatory machinery, which allows passage through S-phase and M-phase, mirrors that in typical eukaryotes is pertinent to an understanding of the dinoflagellate cell cycle. This is expected to be of particular value as dinoflagellates are notorious for their ability to form red tides, also called harmful algal blooms (Glibert et al., 2005). We have tested for the presence of proteins similar to the cell cycle actors involved in the yeast cell in transcriptomes from two dinoflagellate species. We find that in general the dinoflagellate cell cycle pathway is well described by the yeast model.

Methods

Sequences corresponding to all proteins involved in cell cycle progression in the budding yeast KEGG (Okuda et al., 2008) pathway (map04111) were used to query transcriptomes of *Lingulodinium polyedrum* (Beauchemin et al., 2012; Roy et al., 2014) and *Symbiodinium* spp (CassKB8 and Mf1.05b) (Bayer 2012) with tBLASTn at a threshold value of e^{-05}. All dinoflagellate sequences identified were then mapped onto the yeast KEGG pathway and the presence or absence of dinoflagellate examples in the yeast pathway shown by differential coloring of the KEGG map entries. A full list of the genes tested, along with the best hit E-values in a dinoflagellate transcriptome, is provided in Table S1.

In cases where a given protein was not found in the dinoflagellate transcriptomes, it was also used to query a number of other species using tBLASTn at a threshold of e^{-05}, in order to assess if its absence unique to dinoflagellates. Queries were restricted to one species at a time among the following: the fungi *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, the animals *Homo sapiens* and *Drosophila melanogaster*, the higher plants *Arabidopsis thaliana* and *Oryza sativa*, the trichomonad *Trichomonas vaginalis*, the trypanosome *Trypanosoma cruzi*, the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*, the ciliates *Paramecium tetraurelia* and *Tetrahymena thermophila*, and finally the apicomplexans *Toxoplasma gondii* and *Cryptosporidium parvum*.

Phylogenetic analyses were performed using an online version of RAxML (Stamatakis et al., 2005) available at the CIPRES science gateway (Miller et al., 2010)

using the default parameters. All sequences were aligned using ClustalW, checked visually, and imported to the CIPRES gateway in a PHYLIP format. Trees were visualized using Dendroscope (Huson and Scornavacca, 2012).

Results and Discussion

S phase transcriptional activation

The transcription of yeast genes involved in S-phase entry is mediated by two transcription factors called SBF (for Swi4/6 cell cycle box (SCB) Binding Factor) and MBF (for Mlu I cell cycle box (MCB) Binding Factor). Each transcription factor is a heterodimer containing the transcriptional coactivator Swi6p (Switching deficient-6) and either Swi4p (in the case of SBP) (Nasmyth and Dirick, 1991) or Mbp1p (Mlu I-box binding protein) (Koch et al., 1993) as the DNA binding component. These transcription factors are normally held inactive by the regulatory factor Whi5p (Whiskey, anecdotally named because of a bet involving Irish whiskey), an analog of the metazoan retinoblastoma tumor suppressor Rb (Costanzo et al., 2004). Interestingly, while the genetic circuit is conserved between yeast and metazoans, the proteins involved are all different (Cross et al., 2011). It is thus not surprising that the dinoflagellate transcriptomes lack homologs to the yeast (or the human) transcription factors and to the proteins that regulate their activity.

S-phase replication initiation

Eukaryotes initiate replication at many sites along the chromosomes, termed origins, in order to reduce the time required to copy large genomes. In yeast, origin sites are first recognized in early G1 phase by a complex formed by six ORC proteins (origin replication complex) Orc1 - 6 (Bell and Stillman, 1992). The ORC proteins are all ATPases, and only Orc1p is essential in yeast (Speck et al., 2005). The complex is stabilized by Cdc6p (Cell Division Cycle), an ATPase homologous to Orc1, and while essential in yeast (Cocker et al., 1996), archebacteria have only a single protein Orc1/Cdc6 (Barry and Bell, 2006). The dinoflagellates may employ a similar process, as both Orc1p and Cdc6p have hits with the same sequence in the dinoflagellate transcriptomes Fig. 1, Supp Table I). Cdc6p acts together with Cdt1p

(Cdc10-dependent transcript, also called Tah1 for Topo-A hypersensitive) to recruit a complex of six Mcm2-7 (Minichromosome maintenance) helicase proteins (Tanaka and Diffley, 2002) thereby forming what is termed a licensed pre-replication complex (pre-RC) (Bell and Dutta, 2002). Importantly, Cdc6p is a cyclin-dependent kinase (CDK) target, and when phosphorylated by the CDK that activates S-phase (S-CDK), becomes a target for ubiquitin mediated degradation. Since the S-CDK is activated in early S-phase and remains active until the onset of M-phase, Cdc6p accumulation occurs only in G1 phase. In consequence, the pre-RC cannot reform after the G1/S transition because Cdc6p remains absent, and this thus prevents re-replication of the genome. Given the possibility that the dinoflagellates could employ a single protein for the Orc1/Cdc6 function, then the machinery required for pre-RC formation is all present with the exception of Cdt1p. However, the yeast Cdt1p has little sequence homology with proteins of similar function in humans and plants (Sclafani and Holzen, 2007) so the lack of a homolog in the dinoflagellate transcriptomes is not unexpected.

Following pre-RC assembly, but prior to replication of the DNA itself, the helicase activity of the MCM complex is turned on. This activation requires a dimeric protein kinase called DDK (Dbf4-dependant kinase), formed from the Cdc7p protein kinase and its regulatory subunit Dbf4p (Dumbbell former) (Lei et al., 1997). Dinoflagellate transcriptomes contain a homolog for the Cdc7 kinase but not for the regulatory subunit. However, similar to what was noted above for Cdt1, BLAST searches do not recover a Dbt4 homolog in animals or plants suggesting that it function may be carried out by a non-homologous sequence. Phosphorylation of MCM leads to a conformational change that not only activates the enzyme but also allows binding of the replisome, the complex of proteins that will catalyze DNA replication (Yeeles et al., 2015). The conformational change in the Mcm brought about by the DDK allows binding of a single stranded DNA binding protein Cdc45p, which in turn aids in recruiting the replisome to the pre-RC (Petojevic et al., 2015).

The S-CDK also contributes to activation of the Cdc45 by a regulatory pathway involving a subunit of DNA polymerase ε called Dpb11 (DNA polymerase B

subunit 11) (Araki et al., 1995), the proteins Sld2 (Synthetic lethal with Dpb11) and Sld3 (Tanaka et al., 2007), as well as the GINS complex. GINS (named after the numbers Go, Ichi, Nii and San, which are five, one, two and three in Japanese) is a complex of four subunits, Sld5, Psf1 (Partner with Sld Five 1), Psf2 and Psf3, and without the GINS complex, Cdc45p and Dbf11p do not associate with the pre-RC (Takayama et al., 2003). Sld2 binds Dbp11 and is essential for replication in yeast (Kamimura et al., 1998), whereas Sld3 interacts with Cdc45 (Kamimura et al., 2001). The regulatory event catalyzed by the S-CDK is the phosphorylation of Sld2 and Sld3 (Tanaka et al., 2007; Zegerman and Diffley, 2007) and it is the complex of Dpb11 and the phosphorylated forms of Sld2 and Sld3 that bind the origin and allow binding of the replication initiation factor Cdc45. This elegant mechanism allows phosphorylation events catalyzed by the S-CDK (phosphorylation of both Sld2/3 and Cdc6) to not only initiation replication by polymerase recruitment but also to ensure there is a single and unidirectional passage from pre-RC formation to activation at this point in the cell cycle. The dinoflagellates, similar to their relatives the Apicomplexans, do not contain homologs of Sld2/3, Dbp11, or any of the GINS complex by BLAST searches (Fig. 1, Fig. 3), although it is again possible that functional homologs may be present.

M-phase regulation of cyclin-dependent kinase (CDK) activation

Induction of M phase CDK activity (M-CDK) by the mitotic cyclins Clb1p/Clb2p (Cyclin B) is essential for entry into mitosis, as cyclins not only activate the kinase but also direct it to specific substrates (Loog and Morgan, 2005). However, in addition to cyclins, a number of other factors are required for activity (Morgan, 1997). These include the presence of a Cdc28p kinase subunit Cks1p (Cdc28 kinase subunit, a regulatory subunit of the M-CDK (Hadwiger et al., 1989)), the absence of CDK inhibitors, the presence of an activating phosphate (on Thr161 of the yeast Cdc28p) and the absence of an inhibitory phosphate (on Tyr15) (Fig. 2). Dinoflagellate transcriptomes contain homologs for the cyclin dependent kinases, mitotic cyclins and Cks1. They also encode homologs of Cak1p (CDK-activating kinase, (Kaldis et al., 1996)) and the tyrosine kinase Swe1p (*Saccharomyces* Wee1)

that adds an inhibitory phosphate to the Cdc28 (Harvey et al., 2005). However, dinoflagellates do not encode the tyrosine phosphatase Mih1p (Mitotic inducer homolog, a member of the Cdc25 tyrosine phosphatase family) that removes the inhibitory phosphate added by Swe1p (Russell et al., 1989). This is unusual but not without precedent, as plants also lack a Cdc25-like phosphatase (Lipavska et al., 2011). The reason for the lack of a tyrosine phosphatase homolog in plants is not completely clear, but plants contain an unusual CDK homolog called CDKB in addition to the universally found CDKA. The CDKA form is used to initiate S-phase while the CDKB form is reserved for mitosis, and CDKA is thought to activate the CDKB thus eliminating a need for the tyrosine phosphatase in M-phase entry (Tulin and Cross, 2014). Interestingly, a phylogenetic reconstruction containing the dinoflagellate CDKs (Fig. 4) does place some dinoflagellate CDKs close to CDKB sequences, although this clade is not sufficiently well supported to exclude the possibility they are CDKAs.

CDK activity decreases through M phase due to degradation of the cyclin by APC/C (the anaphase promoting complex/cyclosome), an E3 ubiquitin ligase (Pines, 2011). The APC/C requires one of two different adaptors to function in M phase exit, Cdc20p or Cdh1p (Cdc20 homolog) (Visintin et al., 1997). Cdc20p binds to APC/C when phosphorylated by CDK at the start of M phase, and allows the APC/C^{Cdc20} to target mitotic cyclins (Shirayama et al., 1998). Cdh1p is also phosphorylated by CDK, but only binds to APC/C when not phosphorylated (Zachariae et al., 1998). Since mitotic cyclins are targets of the APC/C, as cyclin levels begin to fall, CDK activity levels decrease. The reduction in kinase activity levels, in combination with activation of the protein phosphatase Cdc14 (Stegmeier and Amon, 2004), results in dephosphorylation of Cdh1p. APC/C^{Cdh1} activity is thus maintained at a high level when CDK activity is low.

DNA repair pathway

This pathway functions in yeast to arrest the cell cycle if DNA repair is ongoing. In dinoflagellates, the initial stages are present allowing unimpeded information flow from the clamp loader subunit Rad24p (Radiation sensitive 24) down to both the

checkpoint kinase Chk1p and to the general transcriptional co-repressor Tup1p (dTMP uptake 1) (Keleher et al., 1992). In yeast, Chk1p blocks passage through M phase (Rhind and Russell, 2000) by phosphorylating two main targets. The first of these is Pds1p (Precocious dissociation of sisters, also known as securin), which prevents anaphase entry by inhibiting Esp1p (Extra spindle pole bodies, also known as separin), the protease whose role is to degrade the cohesins that hold the sister chromatids together (Ciosk et al., 1998). The second is the cdc25 homolog Mih1p (Mitotic inducer homolog), a tyrosine phosphatase that activates the CDK by removal of the inhibitory phosphate previously added by Swe1p (Russell et al., 1989). A homolog of securin is indeed found in in dinoflagellate transcriptomes (Fig. 3), but the tyrosine phosphatase homolog Mih1 is absent as noted above.

Cohesin and chromatid separation

Cohesin is a complex of four subunits that holds together the two replicated DNA strands from S-phase until anaphase (Nasmyth, 2002). It is formed from Smc1p (Stability of minichromosomes), Smc3p, Scc1p (Sister Chromatid Cohesion) and Scc3p; in budding yeast these latter two proteins are also called Mcd1p (for Mitotic Chromosome Determinant) and Irr1p (for Irregular cell behavior). The two Smc subunits are rod-like proteins with a globular ATPase domain on one end and a dimerization domain at the other that allows them to form a V-shaped structure. Scc1 is thought to hold the two ATPase domains together, trapping the two sister chromatids in a ring shaped complex (Gruber et al., 2003), and it is Scc1 that is the target of the protease separin. Separin is thus essential for separation of the sister chromatids during anaphase. Interestingly, while a potential separase (Esp1) is found in dinoflagellates, the securin (Pds1) that regulates separase activity is not. Securin must be inactivated by APC/C mediated degradation, in combination with its specificity factor Cdc20, in order to liberate an active separin. However, while securin function is conserved across species, the sequence itself is not (Waizenegger et al., 2002). The dinoflagellate securin function may be encoded by a protein unique to dinoflagellates, as the transcriptomes also lack homologous sequences to securins from fission yeast (Cut2), drosophila (pimples) or vertebrates (PTTG).

Alternatively, these marine dinoflagellates may lack a securin- seawater temperatures are relatively cool, and while Pds1 is essential in yeast grown at 37 degrees, cells without Pds1 survive at lower temperatures (Yamamoto et al., 1996).

Spindle checkpoint

The spindle assembly checkpoint (SAC) blocks separation of the sister chromatids until the kinetochores on each DNA molecule are correctly attached to the spindle (Lara-Gonzalez et al., 2012). This pathway can be observed as a block at the metaphase to anaphase transition when the spindle is disassembled by nocodazole, and thus is presumably found in dinoflagellates since *Crypthecodinium* responds to nocodazole by a reversible prolongation of the G2/M phase (Yeung et al., 2000). The downstream target of the SAC is APC/C, and the mechanism linking the two is the sequestration of the APC/C activator Cdc20p by the protein Mad2p (Mitotic arrest-deficient) (Hwang et al., 1998). By a mechanism still not completely understood, unbound kinetochores catalyze Cdc20 sequestration, thus assuring that APC/C cannot target securin for degradation. A Mad2 homolog is found in the dinoflagellates, although upstream regulatory elements are missing (Fig. 3).

FEAR pathway

The FEAR pathway is named after cdc14 (Cdc fourteen early anaphase release), an essential protein in mitotic exit (Visintin et al., 1998). FEAR is one of two pathways used to regulate Cdc14p (the second, the MEN pathway, is described below). Cdc14p is a protein phosphatase that in yeast acts to dephosphorylate CDK targets and the APC/C regulatory subunit Cdh1p. Cdh1 is a CDK1 substrate, and cannot bind APC/C in its hyperphosphorylated form (at the onset of M phase). However, when dephosphorylated by Cdc14, Cdh1 binds to APC/C and maintains high levels of APC activity.

In yeast, Cdc14p activity toward its substrates is regulated by subcellular localization, specifically sequestration in the nucleus, and this sequestration is regulated by binding to the nucleolar protein Net1p (Nucleolar silencing establishing factor and telophase regulator) (Roccuzzo et al., 2015). Before

anaphase, Cdc14p is bound to Net1p that has been dephosphorylated by the PP2A protein phosphatase. The dinoflagellates have all three normally found PP2A subunits, the catalytic subunit Pph21p (Protein phosphatase) and its two regulatory subunits, Tpd3p (tRNA processing deficient) and Cdc55. The release of Cdc14p requires phosphorylation of Net1p, which occurs both by a direct phosphorylation catalyzed by the polo-like kinase Cdc5p (Yoshida and Toh-e, 2002) and by inactivation of PP2A, which allows phosphorylation of Cdc14p by the CDK1. In yeast, PP2A is inactivated by a pathway that passes from APC/C though the separase Esp1p and the kinetochore-associated protein Slk19p (Synthetic lethal Kar3p) (Queralt et al., 2006), but this pathway is unlikely in dinoflagellates as they lack Slk19 homologs. More importantly, dinoflagellates lack Net1p, making it unlikely that dinoflagellates can use the FEAR pathway at all to regulated Cdc14p (Fig. 3). This is curious, as dinoflagellates and yeast share the property a nuclear membrane that remains intact during mitosis (a closed mitosis), and thus nucleolar sequestration of Net1p would seem to be a particularly useful mechanism. However, in the context of the permanently condensed chromosomes of the dinoflagellates (Bhaud et al., 2000), the absence of Net1p is perhaps less surprising. The reason for this is that in addition to its role in sequestering Cdc14p, Net1p is involved in converting euchromatin to transcriptionally inactive heterochromatin in areas of the chromosome with deacetylated histones (Kasulke et al., 2002). Histones can not be involved in globally interchanging euchromatin and heterochromatin in dinoflagellates, since techniques as sensitive as Western blots and MS/MS analyses have failed to detect them (Roy and Morse, 2012).

MEN pathway

The mitotic exit network (MEN) pathway provides an alternative method to activate the Cdc14p phosphatase. In this pathway, the sequestered Cdc14p is released to the cytoplasm following phosphorylation by a dimer of Dbf2p (Dumbbell former) kinase and its regulatory subunit Mob1p (Mps one binder). This dimer can be activated either by phosphorylation of Mob1 by Mps1 (Monopolar spindle), a protein kinase also involved in the spindle-assembly checkpoint pathway, or by phosphorylation of

Dbf2p by the protein kinase Cdc15. Cdc15 is itself activated by the key regulator Tem1 (Termination of M phase), a monomeric GTPase present in the dinoflagellates. However, the usual activators of Tem1 (the GTP exchange factor Lte1, or the GTPase activating protein complex Bfa1/Bub2) are not found in the transcriptomes (Fig. 3).

CDK profiles in Lingulodinium

Cyclin-dependent kinases are the key players in regulating passage through cell cycle checkpoints, but can be used in cellular contexts that do not involve cell cycle control. Yeast contain six different CDKs (Malumbres, 2014), but only two of these (Cdc28p and Pho85p) are actually involved in regulation of the cell cycle. The other four CDKs (Sgv1p, Ctk1, Ssn3 and Kin28p) are involved in regulating transcription either by phosphorylating RNA polymerase II or as a subunit of TFIIH (Kin28p). Phylogenetic reconstructions were performed with *Lingulodinium* CDKs to assess their potential role. Five full length sequences were identified in the *Lingulodinium* transcriptome by a tBLASTn search using the yeast CDC28p. All contained the invariable amino acid residues found in most kinases and the typical motifs are conserved to a large extent (GEGTYG, DLKPQN, DFGLAR, and WYRAPE; Supp. Fig. 1).

An example of the phylogenetic reconstruction is shown in figure 4. This figure recovers the previously documented relationships between animal CDKs (Cao 2014), although the bootstrap support is significantly lower due to the inclusion of plant, dinoflagellate, ciliate and stramenopile sequences. The two yeast kinases (Pho85 and Cdc28) involved in cell cycle control are found in a well-supported clade together with their animal homologs, although support for the clade containing the transcription-related CDKs is much lower. Interestingly, four of the five *Lingulodinium* kinases are found in the cell-cycle clade, and only one loosely associated with the transcriptional regulation clade. Only one of the *Lingulodinium* CDKs (Lpo724212) is found in a clade where the phylogenetic relationships between the species are maintained.

Cyclin profiles in Lingulodinium

Cyclins are the obligate protein partners of CDKs, and are not only essential to activate the kinase but also to direct them to substrates that are appropriate for the particular cell cycle stage. Cyclin phylogeny in metazoans has identified 3 broad groups (Cao et al., 2014), the cyclin B group (containing cyclins A, B, D, E, J, F, G, I, O, CLB and CLN), the cyclin Y group (containing cyclins Y and PCL) and the cyclin C group (containing cyclins C, H, L, K, T, and Fam58). Again, while the phylogenetic resolution is insufficient to resolve these because of the wide range of species, it is clear that the dinoflagellate cyclins recovered from the transcriptome all cluster together, and appear to be more similar to cyclins A/B than to any of the other classes.

Conclusion

A transcriptome provides an invaluable catalog of genes that can be used to explore the biological and biochemical capabilities of an organism. Here we have interrogated the transcriptomes of two dinoflagellates species in order to assess the degree to which control of the cell cycle can be modeled by the yeast paradigm. We find that in general the yeast model fits remarkably well for a group of organisms so structurally different in nuclear structure from other eukaryotes. Interestingly, in cases where essential yeast genes are absent from the dinoflagellates, as a rule these are also absent from other phylogenetically related species.

Acknowledgements

The financial support of the National Science and Engineering Research Council of Canada (grant 171382-03 to DM) is gratefully acknowledged.

Competing interests

The authors have no competing interests.

Figure Legends

Figure 1 Summary of dinoflagellate genes homologous with cell cycle regulators of *Saccharomyces cerevisiae*. (A) A redrawn KEGG pathway of S-phase regulators. Red ovals represent genes found in dinoflagellates, while squares represent genes not found in dinoflagellates. Genes in green squares are essential in yeast. Red arrows represent activation, either due to phosphorylation (+p) or dephosphorylation (-p), while barred lines represent inhibition. Blue lines represent ubiquitinylation (+u).

Figure 2 Summary of dinoflagellate genes homologous with cell cycle regulators of *Saccharomyces cerevisiae*. (A) Redrawn KEGG pathway of M-phase regulators. Red ovals represent genes found in dinoflagellates, white squares are genes not found in dinoflagellates, and green squares are genes absent in dinoflagellates but that are essential in yeast. Red arrows represent activation, either due to phosphorylation (+p) or dephosphorylation (-p), while barred lines represent inhibition. Blue lines represent ubiquitinylation (+u).

Figure 3 Phylogenetic conservation of yeast cell cycle regulators

Presence (red) and absence (white) of homologs to different yeast genes across a wide phylogenetic range of species (BLASTp < e^{-05}). Species from top to bottom are *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Homo sapiens*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Oryza sativa*, *Trichomonas vaginalis*, *Trypanosoma cruzi*, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, *Paramecium tetraurelia*, *Tetrahymena thermophila*, *Toxoplasma gondii*, *Cryptosporidium parvum*, *Lingulodinium polyedrum*, and *Symbiodinium* spp (CassKB8 and Mf1.05b).

Figure 4 Dinoflagellate CDK phylogeny

RAxML phylogeny of five dinoflagellate cyclin dependent kinase sequences (red). Four fall into a well-supported clade of cell cycle regulators, including one found in a

well supported Cdk5 (yeastPho85) clade and the other three in the main Cdk1 (yeast Cdc28) clade. Only one is found in the poorly supported clade of transcriptional regulators.

Figure 5 Dinoflagellate cycle phylogeny

RAxML phylogeny of ten dinoflagellate cyclin N-terminal domain sequences (red and boxed). These all appear most closely related to the cyclin A/B family, and are thus likely to be involved in regulating the cell cycle.

Supplemental Figure 1 Multiple alignment of cyclin dependent kinase sequences

CDK sequences from *Arabidopsis* (CDKA, CDKB1, CDKB2), *Schizosaccharomyces* (cdc2), *Saccharomyces* (CDC28), *Homo* (CDK1), and *Lingulodinium* (all five Locus sequences) were aligned using ClustalW.

References

Araki H, Leem SH, Phongdara A, Sugino A (1995) Dpb11, which interacts with DNA polymerase II(epsilon) in Saccharomyces cerevisiae, has a dual role in S-phase progression and at a cell cycle checkpoint. Proc Natl Acad Sci U S A **92**: 11791-11795.

Bachvaroff TR, Handy SM, Place AR, Delwiche CF (2011) Alveolate phylogeny inferred using concatenated ribosomal proteins. The Journal of eukaryotic microbiology **58**: 223-233.

Barry ER, Bell SD (2006) DNA replication in the archaea. Microbiology and molecular biology reviews : MMBR **70**: 876-887.

Bayer T, Aranda M, Sunagawa S, Yum LK, Desalvo MK, Lindquist E, Coffroth MA, Voolstra CR, Medina M (2012) Symbiodinium transcriptomes: genome insights into the dinoflagellate symbionts of reef-building corals. PLoS One **7**: e35269.

Beauchemin M, Roy S, Daoust P, Dagenais-Bellefeuille S, Bertomeu T, Letourneau L, Lang BF, Morse D (2012) Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic. Proc Natl Acad Sci U S A **109**: 15793-15798.

Bell SP, Dutta A (2002) DNA replication in eukaryotic cells. Annual review of biochemistry **71**: 333-374.

Bell SP, Stillman B (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature **357**: 128-134.

Bhaud Y, Guillebault D, Lennon J, Defacque H, Soyer-Gobillard MO, Moreau H (2000) Morphology and behaviour of dinoflagellate chromosomes during the cell cycle and mitosis. J Cell Sci **113**: 1231-1239.

Bodansky S, Mintz LB, Holmes DS (1979) The mesokaryote Gyrodinium cohnii lacks nucleosomes. Biochem Biophys Res Commun **88**: 1329-1336.

Cao L, Chen F, Yang X, Xu W, Xie J, Yu L (2014) Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. BMC Evol Biol **14**: 10.

Ciosk R, Zachariae W, Michaelis C, Shevchenko A, Mann M, Nasmyth K (1998) An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. Cell **93**: 1067-1076.

Cocker JH, Piatti S, Santocanale C, Nasmyth K, Diffley JF (1996) An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. Nature **379**: 180-182.

Costanzo M, Nishikawa JL, Tang X, Millman JS, Schub O, Breitkreuz K, Dewar D, Rupes I, Andrews B, Tyers M (2004) CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast. Cell **117**: 899-913.

Cross FR, Buchler NE, Skotheim JM (2011) Evolution of networks and sequences in eukaryotic cell cycle control. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **366**: 3532-3544.

Fritz L, Treimer RE (1983) An ultrastructural study of mitosis in a marine dinoflagellate Prorocentrum minimum. J. Protozool. **30**: 437-444.

Glibert PM, Anderson DM, Gentien P, Granéli E, Sellner K (2005) The global, complex phenomena of harmful algal blooms. Oceanography **18**: 132-141.

Gornik SG, Ford KL, Mulhern TD, Bacic A, McFadden GI, Waller RF (2012) Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates. Curr Biol **22**: 2303-2312.
Gould SB, Tham WH, Cowman AF, McFadden GI, Waller RF (2008) Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. Mol Biol Evol **25**: 1219-1230.
Gruber S, Haering CH, Nasmyth K (2003) Chromosomal cohesin forms a ring. Cell **112**: 765-777.
Hadwiger JA, Wittenberg C, Mendenhall MD, Reed SI (1989) The Saccharomyces cerevisiae CKS1 gene, a homolog of the Schizosaccharomyces pombe suc1+ gene, encodes a subunit of the Cdc28 protein kinase complex. Mol Cell Biol **9**: 2034-2041.
Harvey SL, Charlet A, Haas W, Gygi SP, Kellogg DR (2005) Cdk1-dependent regulation of the mitotic inhibitor Wee1. Cell **122**: 407-420.
Hou Y, Lin S (2009) Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes. PLoS One **4**: e6978.
Huson DH, Scornavacca C (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Systematic biology **61**: 1061-1067.
Hwang LH, Lau LF, Smith DL, Mistrot CA, Hardwick KG, Hwang ES, Amon A, Murray AW (1998) Budding yeast Cdc20: a target of the spindle checkpoint. Science **279**: 1041-1044.
Kaldis P, Sutton A, Solomon MJ (1996) The Cdk-activating kinase (CAK) from budding yeast. Cell **86**: 553-564.
Kamimura Y, Masumoto H, Sugino A, Araki H (1998) Sld2, which interacts with Dpb11 in Saccharomyces cerevisiae, is required for chromosomal DNA replication. Mol Cell Biol **18**: 6102-6109.
Kamimura Y, Tak YS, Sugino A, Araki H (2001) Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. EMBO J **20**: 2097-2107.
Kasulke D, Seitz S, Ehrenhofer-Murray AE (2002) A role for the Saccharomyces cerevisiae RENT complex protein Net1 in HMR silencing. Genetics **161**: 1411-1423.
Keleher CA, Redd MJ, Schultz J, Carlson M, Johnson AD (1992) Ssn6-Tup1 is a general repressor of transcription in yeast. Cell **68**: 709-719.
Koch C, Moll T, Neuberg M, Ahorn H, Nasmyth K (1993) A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase. Science **261**: 1551-1557.
Lara-Gonzalez P, Westhorpe FG, Taylor SS (2012) The spindle assembly checkpoint. Curr Biol **22**: R966-980.
Lei M, Kawasaki Y, Young MR, Kihara M, Sugino A, Tye BK (1997) Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. Genes Dev **11**: 3365-3374.
Lipavska H, Maskova P, Vojvodova P (2011) Regulatory dephosphorylation of CDK at G(2)/M in plants: yeast mitotic phosphatase cdc25 induces cytokinin-like effects in transgenic tobacco morphogenesis. Annals of botany **107**: 1071-1086.
Loog M, Morgan DO (2005) Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. Nature **434**: 104-108.

Malumbres M (2014) Cyclin-dependent kinases. Genome biology **15**: 122.
Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees, Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, pp. 1-8.
Morgan DO (1997) Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol **13**: 261-291.
Nasmyth K (2002) Segregating sister genomes: the molecular biology of chromosome separation. Science **297**: 559-565.
Nasmyth K, Dirick L (1991) The role of SWI4 and SWI6 in the activity of G1 cyclins in yeast. Cell **66**: 995-1013.
Oakley BR, Dodge JD (1974) Kinetochores associated with the nuclear envelope in the mitosis of a dinoflagellate. J Cell Biol **63**: 322-325.
Oakley BR, Dodge JD (1976) Mitosis and cytokinesis in the dinoflagellate Amphidinium carterae. Cytobios **17**: 35-46.
Okuda S, Yamada T, Hamajima M, Itoh M, Katayama T, Bork P, Goto S, Kanehisa M (2008) KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res **36**: W423-426.
Petojevic T, Pesavento JJ, Costa A, Liang J, Wang Z, Berger JM, Botchan MR (2015) Cdc45 (cell division cycle protein 45) guards the gate of the Eukaryote Replisome helicase stabilizing leading strand engagement. Proc Natl Acad Sci U S A **112**: E249-258.
Pines J (2011) Cubism and the cell cycle: the many faces of the APC/C. Nat Rev Mol Cell Biol **12**: 427-438.
Queralt E, Lehane C, Novak B, Uhlmann F (2006) Downregulation of PP2A(Cdc55) phosphatase by separase initiates mitotic exit in budding yeast. Cell **125**: 719-732.
Rhind N, Russell P (2000) Chk1 and Cds1: linchpins of the DNA damage and replication checkpoint pathways. J Cell Sci **113 (Pt 22)**: 3889-3896.
Rill RL, Livolant F, Aldrich HC, Davidson MW (1989) Electron microscopy of liquid crystalline DNA: direct evidence for cholesteric-like organization of DNA in dinoflagellate chromosomes. Chromosoma **98**: 280-286.
Rizzo PJ, Jones M, Ray SM (1982) Isolation and properties of isolated nuclei from the Florida red tide dinoflagellate Gymnodinium breve (Davis). The Journal of protozoology **29**: 217-222.
Rizzo PJ, Nooden LD (1972) Chromosomal proteins in the dinoflagellate alga Gyrodinium cohnii. Science **176**: 796-797.
Roccuzzo M, Visintin C, Tili F, Visintin R (2015) FEAR-mediated activation of Cdc14 is the limiting step for spindle elongation and anaphase progression. Nature cell biology **17**: 251-261.
Roy S, Letourneau L, Morse D (2014) Cold-Induced Cysts of the Photosynthetic Dinoflagellate Lingulodinium polyedrum Have an Arrested Circadian Bioluminescence Rhythm and Lower Levels of Protein Phosphorylation. Plant Physiol **164**: 966-977.
Roy S, Morse D (2012) A full suite of histone and histone modifying genes are transcribed in the dinoflagellate Lingulodinium, PLoS One.
Russell P, Moreno S, Reed SI (1989) Conservation of mitotic controls in fission and budding yeasts. Cell **57**: 295-303.

Sclafani RA, Holzen TM (2007) Cell cycle regulation of DNA replication. Annu Rev Genet **41**: 237-280.
Shirayama M, Zachariae W, Ciosk R, Nasmyth K (1998) The Polo-like kinase Cdc5p and the WD-repeat protein Cdc20p/fizzy are regulators and substrates of the anaphase promoting complex in Saccharomyces cerevisiae. EMBO J **17**: 1336-1349.
Speck C, Chen Z, Li H, Stillman B (2005) ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. Nat Struct Mol Biol **12**: 965-971.
Stamatakis A, Ludwig T, Meier H (2005) RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics **21**: 456-463.
Stegmeier F, Amon A (2004) Closing mitosis: the functions of the Cdc14 phosphatase and its regulation. Annu Rev Genet **38**: 203-232.
Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. Genes Dev **17**: 1153-1165.
Tanaka S, Diffley JF (2002) Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. Nature cell biology **4**: 198-207.
Tanaka S, Umemori T, Hirai K, Muramatsu S, Kamimura Y, Araki H (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature **445**: 328-332.
Tulin F, Cross FR (2014) A microbial avenue to cell cycle control in the plant superkingdom. Plant Cell **26**: 4019-4038.
Visintin R, Craig K, Hwang ES, Prinz S, Tyers M, Amon A (1998) The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. Mol Cell **2**: 709-718.
Visintin R, Prinz S, Amon A (1997) CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis. Science **278**: 460-463.
Waizenegger I, Gimenez-Abian JF, Wernic D, Peters JM (2002) Regulation of human separase by securin binding and autocleavage. Curr Biol **12**: 1368-1378.
Wong JT, New DC, Wong JC, Hung VK (2003) Histone-like proteins of the dinoflagellate Crypthecodinium cohnii have homologies to bacterial DNA-binding proteins. Eukaryot Cell **2**: 646-650.
Yamamoto A, Guacci V, Koshland D (1996) Pds1p is required for faithful execution of anaphase in the yeast, Saccharomyces cerevisiae. J Cell Biol **133**: 85-97.
Yeeles JT, Deegan TD, Janska A, Early A, Diffley JF (2015) Regulated eukaryotic DNA replication origin firing with purified proteins. Nature **519**: 431-435.
Yeung PK, New DC, Leveson A, Yam CH, Poon RY, Wong JT (2000) The spindle checkpoint in the dinoflagellate Crypthecodinium cohnii. Exp Cell Res **254**: 120-129.
Yoshida S, Toh-e A (2002) Budding yeast Cdc5 phosphorylates Net1 and assists Cdc14 release from the nucleolus. Biochem Biophys Res Commun **294**: 687-691.
Zachariae W, Schwab M, Nasmyth K, Seufert W (1998) Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. Science **282**: 1721-1724.
Zegerman P, Diffley JF (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature **445**: 281-285.

Figure 1

Click here to download high resolution image

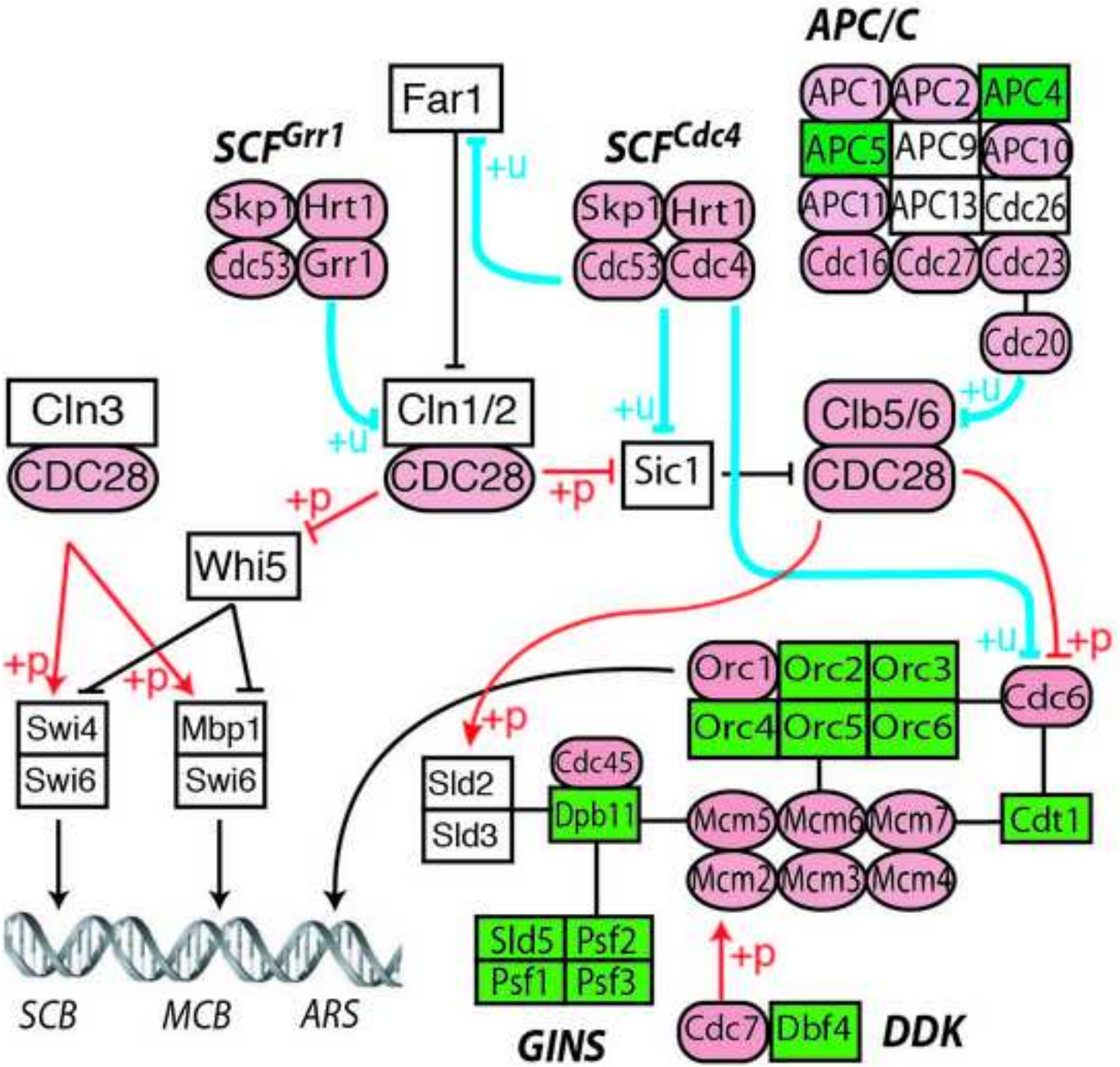


Figure 2

Click here to download high resolution image

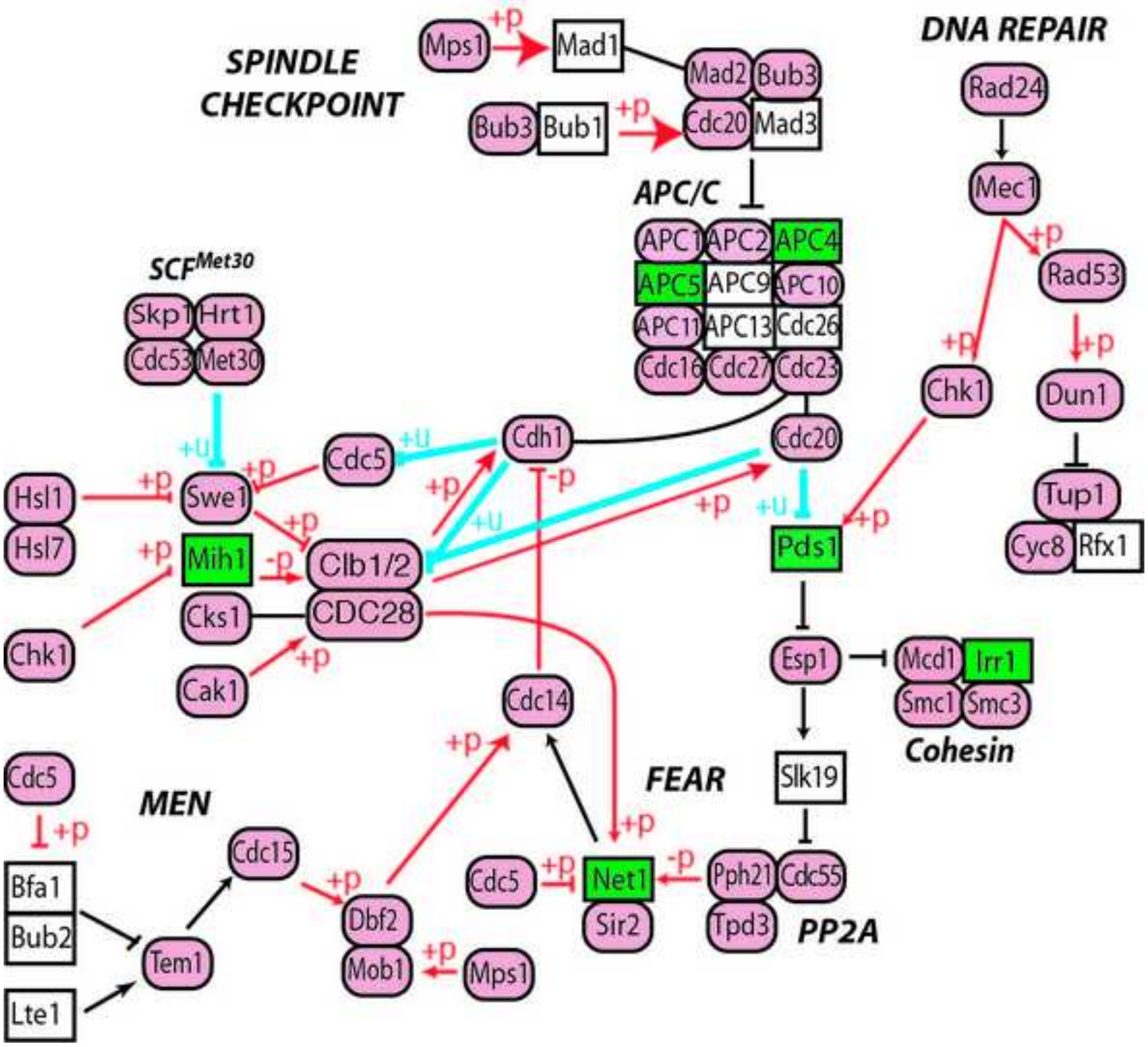


Figure 3

Click here to download high resolution image

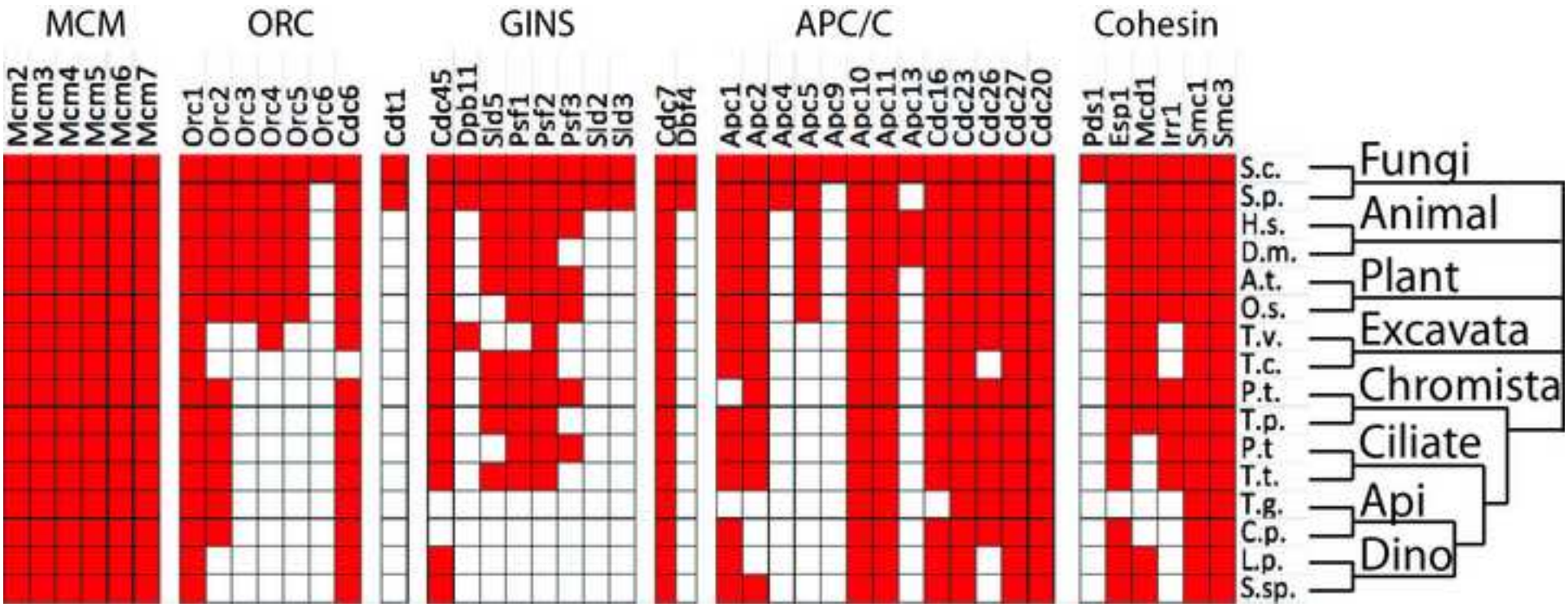


Figure 4

Click here to download high resolution image

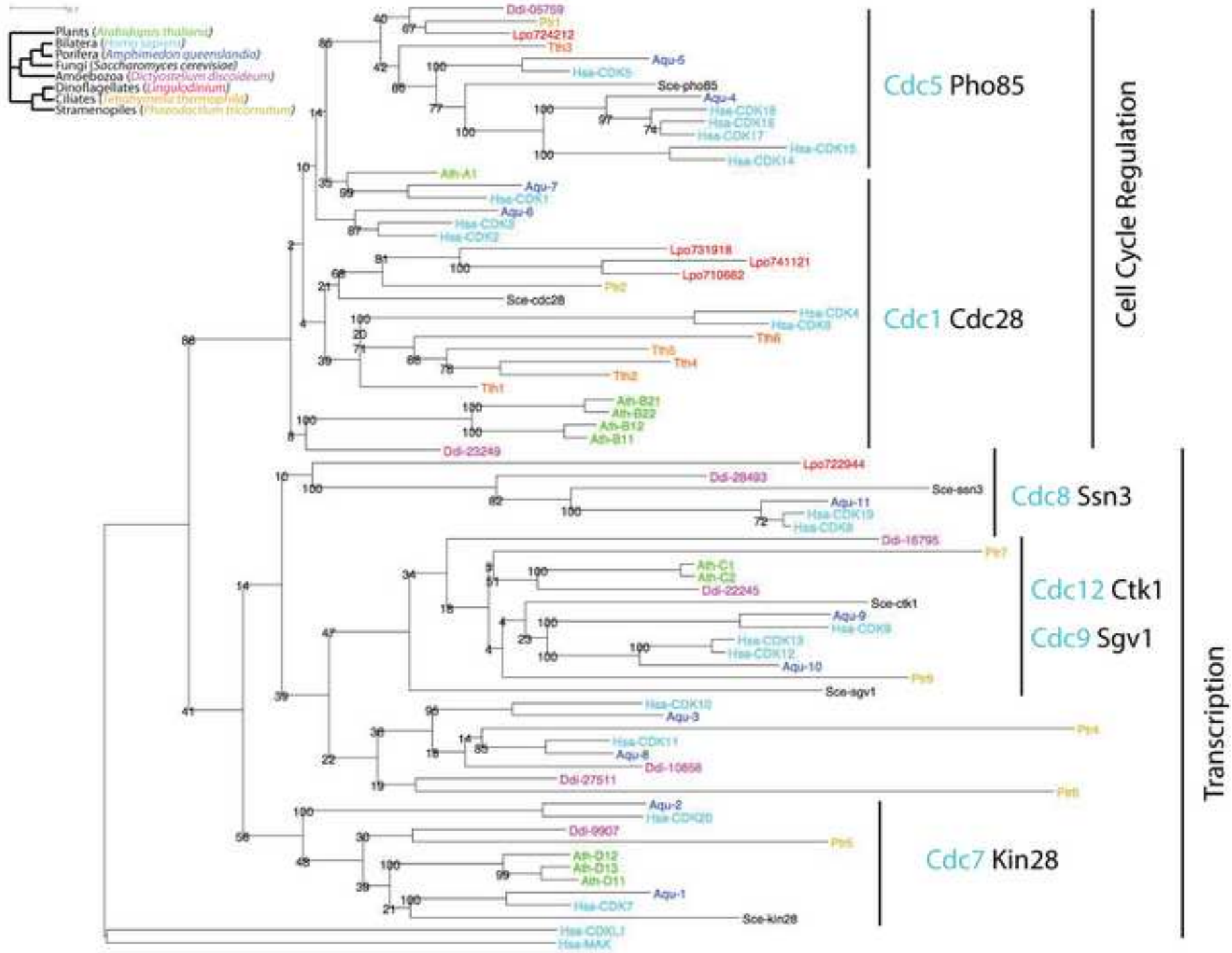
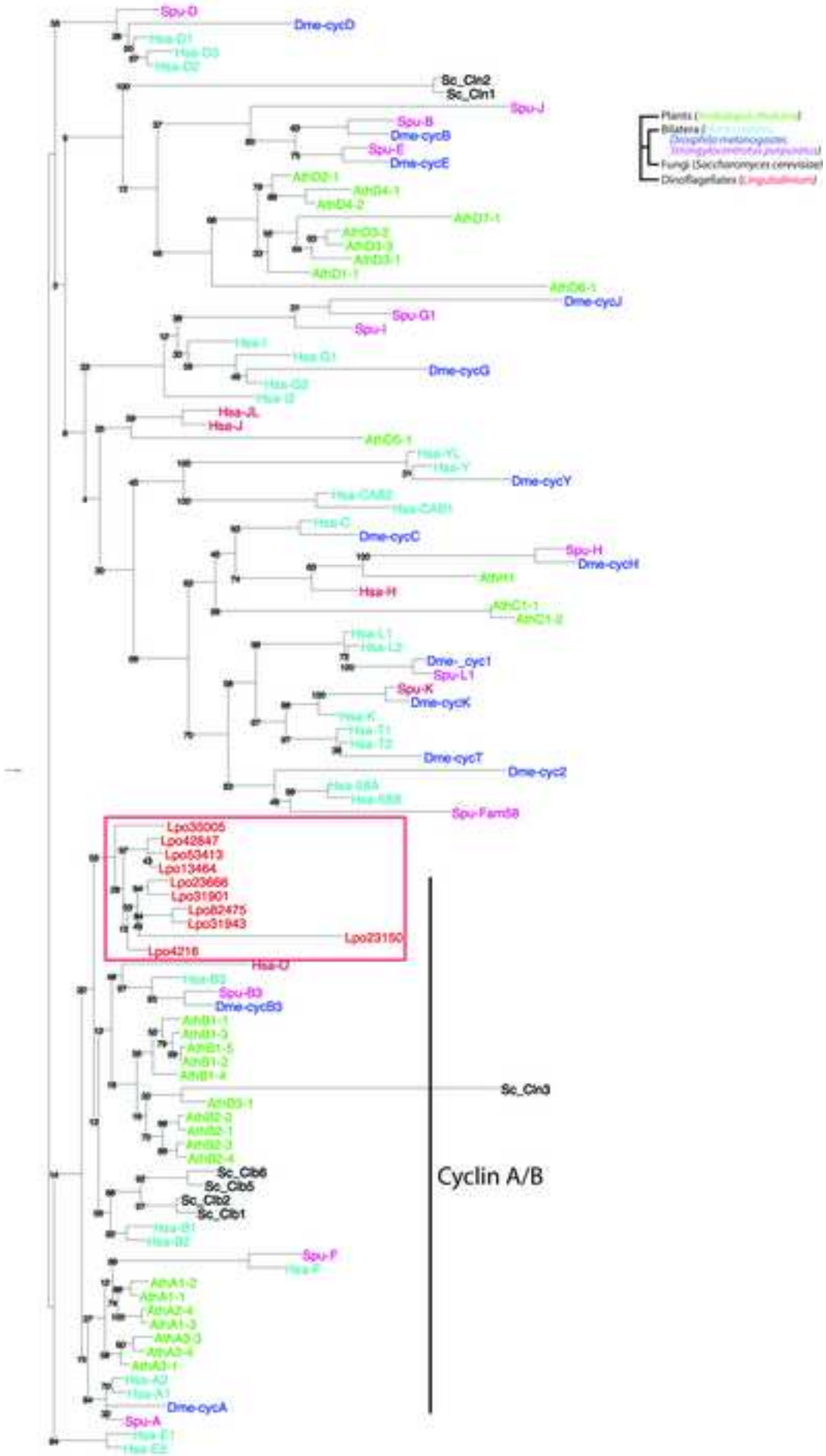


Figure 5

Click here to download high resolution image



Supplement Table 1

Click here to download Supplement Material: Supp Table I Dino cell cycle homologs FINAL.xlsx

Supplement Figure 1

Click here to download Supplement Material: Supp Fig 1CDK alignment.pdf