Gradual evolution of cell cycle regulation by cyclin-dependent kinases during the transition to animal multicellularity

Alberto Perez-Posada[1], Omaya Dudin[1], Eduard Ocaña-Pallarès[1], Iñaki Ruiz-Trillo[1,2,3]*, Andrej Ondracka[1]*

1 Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Catalonia, Spain

2 Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Av. Diagonal, 645, 08028 Barcelona, Catalonia, Spain

3 ICREA, Passeig Lluís Companys 23, 08010, Barcelona, Catalonia, Spain

*Corresponding authors

Email: inaki.ruiz@ibe.upf-csic.es

Email: andrej.ondracka@gmail.com

Abstract

Progression through the cell cycle in eukaryotes is regulated on multiple levels. The main driver of the cell cycle progression is the periodic activity of cyclin-dependent kinase (CDK) complexes. In parallel, transcription during the cell cycle is regulated by a transcriptional program that ensures the just-in-time gene expression. Many core cell cycle regulators are present in all eukaryotes, among them cyclins and CDKs; however, periodic transcriptional programs are divergent between distantly related species. In addition, many otherwise conserved cell cycle regulators have been lost and independently evolved in yeast, a widely used model organism for cell cycle research. To gain insight into the cell cycle regulation in a more representative opisthokont, we investigated the cell cycle regulation at the transcriptional level of *Capsaspora owczarzaki*, a species closely related to animals. We developed a protocol for cell cycle synchronization in *Capsaspora* cultures and assessed gene expression over time across the entire cell cycle. We identified a set of 801 periodic genes that grouped into five clusters of expression over time. Comparison with datasets from other eukaryotes revealed that the periodic transcriptional program of *Capsaspora* is most similar to that of animal cells. We found that orthologues of cyclin A, B and E are expressed at the same cell cycle stages as in human cells and in the same temporal order. However, in contrast to human cells where these cyclins interact with multiple CDKs, *Capsaspora* cyclins likely interact with a single ancestral CDK1-3. Thus, the *Capsaspora* cyclin-CDK system could represent an intermediate state in the evolution of animal-like cyclin-CDK regulation. Overall, our results demonstrate that *Capsaspora* could be a useful unicellular model system for animal cell cycle regulation.

Keywords

cell cycle, evolution of cell cycle, periodic gene, transcriptional program, cell division, synchronization of cell cultures, opisthokonta, holozoan

Author's summary

When cells reproduce, proper duplication and splitting of the genetic material is ensured by cell cycle control systems. Many of the regulators in these systems are present across all eukaryotes, such as cyclin and cyclin-dependent kinases (CDK), or the E2F-Rb transcriptional network. Opisthokonts, the group comprising animals, yeasts and their unicellular relatives, represent a puzzling scenario: in contrast to animals, where the cell cycle core machinery seems to be conserved, studies in yeasts have shown that some of these regulators have been lost and independently evolved. For a better understanding of the evolution of the cell cycle regulation in opisthokonts, and ultimately in the lineage leading to animals, we have studied cell cycle regulation in Capsaspora owczarzaki, a unicellular amoeba more closely related to animals than fungi that retains the ancestral cell cycle toolkit. Our findings suggest that, in the ancestor of Capsaspora and animals, cyclins oscillate in the same temporal order as in animals, and that expansion of CDKs occurred later in the lineage that led to animals.

Introduction

The cell cycle is an essential and fundamental biological process that underpins the cell division and proliferation of all cells. Progression through the cell cycle involves multiple layers of regulation [1]. The main regulatory networks that govern the transitions between cell cycle stages are broadly conserved in eukaryotes, both on the level of individual regulators [2], as well as on the level of network topology [3]. However, it is still not well understood how this conserved regulatory network is deployed in cells with different cellular lifestyles and how it changes across evolution.

Among the main regulators of the progression through the cell cycle are cyclins and cyclin-dependent kinases (CDKs), two gene families broadly conserved across eukaryotes [1,2]. Cyclins and CDKs have undergone independent expansions and subfunctionalization in every major lineage of eukaryotes, including opisthokonts [4–6]. In animals, there are multiple cyclins and CDKs that form discrete complexes, activating specific downstream effectors in different phases of the cell cycle [7,8]. Cyclin D-CDK4,6 complexes control entry into the cell cycle in response to mitogenic factors [9–11]. The G1/S transition is driven by the Cyclin E/CDK2 complex [12,13], and progression through S phase is controlled by the Cyclin A/CDK2 complex [13]. Lastly, cyclin B/CDK1 drive completion of mitosis [14,15]. In contrast, in the budding yeast *Saccharomyces cerevisiae* one single CDK sequentially binds to nine cyclins in three temporal waves [16,17]: Cln1-2 are expressed in G1 and mark the commitment to a new cycle [18–21], Clb5,6 promote DNA replication at S phase [22,23], and Clb1-4 drive progression through mitosis [24,25]. The fission yeast *Schizosaccharomyces pombe* has a single CDK that also binds different cyclins at G1,S, and M: Cig1,2 drive progression through G1 and S phase [26,27], and Cdc13 drives progression through mitosis [28,29]. However, a single CDK-Cyclin complex can drive progression through the entire cell cycle in this species [30].

In addition to the cyclin-CDK activity, the cell cycle is also regulated at the transcriptional level by timing the expression of genes required in its different phases. For instance, the E2F-Rb network of transcription factors controls initiation of DNA replication in animals at the G1/S transition [31–33]. In yeasts,

transcriptional regulation of the G1/S transition is driven by SBF and MBF, two transcription factor complexes that bear no homology to E2F [34]. Recent findings show that these transcription factors were acquired through lateral gene transfer during fungal evolution [2]. In addition, other transcription factors are the mitotic Fkh1 and Fkh2 [35,36] in *S. cerevisiae* or the Hcm1 transcription factor, controlling progression through G2 and mitosis [37]. In human cells, a protein of the same family, FoxM1, also regulates gene expression in mitosis [32,38]. Although oscillatory transcriptional activity during the cell cycle is present in numerous species and cell types [32,33,46–50,37,39–45], the genes affected by cell cycle-regulated transcription are divergent between distantly related species [51]. Likewise, even among different human cell types, periodic expression of only a fraction of genes is common to all of them [32,52].

Yeasts have historically been a powerful model system to understand the control of the cell cycle in animals. However, it has become clear that many otherwise conserved cell cycle regulators have been lost and independently evolved in the fungal lineage [2,3,53–55]. Thus, we sought to investigate the cell cycle control in another organism within opisthokonts that has retained the ancestral cell cycle regulation. We focused on *Capsaspora owczarzaki* (hereafter *Capsaspora*), a species more closely related to animals than yeasts, easy to culture, and for which good genomic resources are available [56–59]. This amoeba has a life cycle that includes three distinct stages that differ both in their morphology and transcriptional and proteomic profiles: amoebas with filopodia that proliferate in adherent cultures, an aggregative multicellular stage in which cells produce an extracellular matrix, and a cystic form that lacks filopodia [60–62]. Moreover, *Capsaspora* has a compact, well-annotated genome, with many homologs to animal genes [59]. With recent advances that allow transfection in the laboratory [63], *Capsaspora* is becoming a tractable model organism.

In this work, we have established a protocol to synchronize cell cycle progression in *Capsaspora* and have characterized its cell division and transcriptional profile across the entire cell cycle. We found that globally, the periodic transcriptional program of *Capsaspora* is enriched in genes that date back to eukaryotic origin, and it resembles human cells more than the periodic transcriptomes of yeasts. Out of four human cyclin types, *Capsaspora* contains

131 homologs of cyclins A, B, and E. We found that these three cyclins are
132 transcriptionally regulated during the cell cycle and have a conserved temporal
133 order and cell cycle stage with human cells. In contrast, *Capsaspora* only
134 contains one ancestral copy of the CDK, which likely form complexes with all of
135 the *Capsaspora* cyclins. We also found that orthologs of many other cell cycle
136 regulatory genes have a conserved timing of expression compared with animal
137 cells. Our findings suggest that the cyclin-CDK system of animals evolved
138 gradually, through an intermediate stage where one single CDK was able to
139 interact with several cyclins at distinct stages of the cell cycle. Thus, while
140 expansion and subfunctionalization of animal cyclins occurred earlier, expansion
141 of CDKs occurred concomitantly with the emergence of animals.

Results

Synchronization of cell cultures in *Capsaspora*

Synchronization of cell cultures is a basic experimental tool required to study the cell cycle [64]. Cell synchronization methods include temperature-sensitive strains, elutriation, cell sorting and commercially available inhibitors that arrest the cell cycle. Arrest and release approaches have previously been used to assess cell cycle progression in several organisms [39,42,65,66]. Hydroxyurea, a widely used S-phase inhibitor in yeast cells [67], was already shown to inhibit cell proliferation in *Capsaspora* cultures during the adherent cell stage [62]. Therefore, to check if cell cycle inhibition occurred before entering S phase and was reversible, we treated *Capsaspora* adherent cultures with hydroxyurea and assessed DNA content by flow cytometry. Upon hydroxyurea treatment, cells exhibited 1C DNA content, indicating arrest in G1 phase (Fig. 1A). Upon wash and release into fresh media, we observed synchronous progression through the cell cycle as assessed by DNA content (Fig. 1C). This indicates that hydroxyurea inhibits the cell cycle in S phase in *Capsaspora* and that its effect is reversible.

To measure the timing of cell cycle stages in *Capsaspora*, we treated two biological replicates of adherent proliferative *Capsaspora* cultures with hydroxyurea, and we later released them into fresh medium. Samples were taken at 45-minute intervals, starting from 2 hours after release, taking a total of 16 time points that were analyzed using flow cytometry (Fig. 1B). Following release, cells spent approximately 1.5 hours duplicating their DNA content, and after 8 hours from release a G1 peak appeared again, indicating completion of the cell cycle (Fig. 1C). These observations were reproduced in the two replicates. At later time points, we noticed co-occurrence of 1C and 2C peaks. This may be due to some cells progressing through the cell cycle more rapidly than others [68], causing loss of synchrony, or due to an irreversible arrest in a fraction of the cells upon HU treatment [69]. Nevertheless, our time course encompasses a complete round of the cell cycle, as we observed an increase in 1C at later time points.

Dynamics and morphology of cell division in *Capsaspora*

To characterize the dynamics of cell division in *Capsaspora*, we used time-lapse microscopy in synchronized cells. In parallel, we analyzed synchronized cells for DNA content by flow cytometry and characterized the cell morphology during cell division using fluorescence microscopy. Out of a total of 100 mitotic cells observed by live imaging (Fig. 2A), the highest fraction underwent cytokinesis at approximately 10 hours (Fig. 2C). Cells seemed to round up and slightly detach from the plate surface while retaining their filopodia (Fig. 2A, Video 1). This phenomenon has been previously characterized in other eukaryotic cells lacking a rigid cell wall, such as animal cells or *Dictyostelium* amoeboid cells [70–77]. Cells took an average of 3 minutes to completely undergo cytokinesis (Fig. 2A, Video 1), measured as the time from rounding up to the splitting of two daughter cells. This value is at least twice as fast as in animal cell types, where the average time for cytokinesis is 6 minutes [78–80], and it is also considerably faster than in yeasts [81–84] when considered in relation to the total generation time of *Capsaspora* (estimated to be around 12 hours in our culture conditions). As shown in Fig. 2D, the measured area of daughter cells is roughly the same, suggesting that cell division is symmetric and yields two equally sized cells.

To investigate the morphology of mitotic cells, we stained tubulin and DNA. On each cell, we observed one dot of dense concentrations of tubulin, from which microtubules emanate (Fig. 2B, white arrows). These dots duplicated and remained close, first associated with the nucleus and then surrounding densely packed DNA. A central, thicker spindle emerged and grew as DNA separated and moved to opposite poles of the dividing cell. A similar phenomenon has been described for *Dictyostelium* [85]. Previous studies have reported the absence of proteins from the gamma-tubulin ring complex and the yeast spindle pole body in *Capsaspora* [86], which together with our observations suggests that *Capsaspora* mitotic spindle is organized without a microtubule organizing center (MTOC), or that it possesses an independently evolved MTOC, such as in yeast [87].

Detection of periodically expressed genes during the *Capsaspora* cell cycle

In many cell types, cell division cycles are accompanied by a transcriptional program of periodic gene expression over time. To understand the transcriptome dynamics during the cell cycle of *Capsaspora,* we performed time-series RNA-seq experiments. We used RNA extracts from the same two biological replicates as in the flow cytometry assay and sequenced them using Illumina HiSeq v4. We processed the sequencing reads using Kallisto [88] (Supplementary File 1). Spearman correlations by gene expression profiles showed that time points are grouped according to the temporal order of sampling (Supplementary Fig. S2). This indicates that gene expression is not shifted over time and was reproducible between the two replicates. To detect periodicity patterns in gene expression, we applied two algorithms, JTK_CYCLE [89] and RAIN [90], on an average dataset of the two replicates where non-expressed genes were filtered out (Supplementary Fig. S1A). We assigned two ranks to every gene according to the p-values calculated by JTK_CYCLE and RAIN (Fig. 3B), and assigned the final periodicity rank as the sum of JTK and RAIN ranks. We applied a conservative cutoff to identify genes that are undoubtedly periodically transcribed by taking the top 800 genes ranked within the top 2000 ranking for each independent dataset (Fig. 3B) (Supplementary Fig. S1B). This cutoff corresponds to 10% of the total number of genes in *Capsaspora*, a fraction similar to those observed in other species [39,42,50,91]. Although false negatives with higher ranks might have been discarded due to our conservative approach, we confirmed that top-ranked genes showed oscillatory behavior. We manually included *Capsaspora* Cyclin A (CAOG_04719T0) [4,59] (see below) to the list of periodic genes despite ranking in position 1916, as it has a periodic expression profile (Fig. 5A).

For a gene expression dataset containing only genes identified as periodic, we observed stronger correlations by time points (Fig. 3C) or pairwise genes than for the entire dataset (Supplementary Fig. S3A). We also observed a strong correlation between initial and late time points (Fig. 3C), suggesting that the cultures indeed completed the entire cell cycle despite the loss of synchrony. A principal coordinate analysis using data for the top-ranked periodic genes

retrieved a grouping of time points that resembles a circle with two components explaining around 70% of the total distance (Fig. 3D, Supplementary Fig. S3B), and a similar layout is obtained on a t-SNE plot of periodically expressed genes (Supplementary Fig. S4). The results from these analyses indicate that there is a set of periodically transcribed genes during the cell cycle of synchronous *Capsaspora* cells. This set of genes represents the periodic transcriptional program of *Capsaspora* (Supplementary File 2).

Gene expression is clustered in periodic waves during the cell cycle

The cell-cycle-associated transcriptional program responds to the requirements of the cell at a given moment. For example, in many cell types, genes necessary for DNA replication or mitosis are transcribed only at the time of their biochemical activity. The detection of periodic genes in *Capsaspora* prompted us to classify them into temporal clusters, by centering their expression profiles to the mean and grouping them using hierarchical clustering based on a dissimilarity matrix by Euclidean distance (Fig. 4A). Five clusters were detected according to the similarity in expression profiles over time (Fig. 4A, Supplementary Fig. S4A, Supplementary Fig. S4B), and we obtained very similar results by using k-means clustering (Supplementary Fig. S5). We then associated these gene clusters to the cell cycle stage during the peak of expression: the G1/S cluster, S cluster, G2/M cluster, M cluster, and G1 cluster. Next, we calculated gene ontology (GO) enrichment for every cluster against the whole periodic transcriptional program using Ontologizer [92] Parental-Child-Union calculation (Fig. 4B, Supplementary File 3).

The G1/S cluster contains 194 genes peaking in the initial time points, from 2 to 3.5 hours after release. Genes found here exhibit the largest differences in gene expression between time points and are enriched in GO terms related to DNA replication, deoxyribonucleotide biosynthesis, and chromosome organization. There is also enrichment in the GO term “response to stress”, suggesting an effect of the treatment with hydroxyurea [68,69]. The small cluster assigned to S/early G2 contains 36 genes peaking between 3.5 to 5.75 hours enriched in the GO term “nucleosome binding” and includes several histone genes.

In the G2/M cluster, we found 176 genes with the peak of expression at 6.5-7.25 hours post-release. This cluster is enriched in the GO term “non-coding RNA metabolic process”, and contains genes related to tRNA maturation such as RTCB, DDX1 [93], and tRNA ligases. It has been previously reported that tRNA synthesis can increase during the cell cycle in several systems [94–96]. To our knowledge, however, there is still no link reported between tRNA modification and progression through the cell cycle.

The M cluster is the largest one in the dataset, with 241 genes peaking between 9.5 and 11 hours. Genes reach the highest expression during time points when the cells enter mitosis, reaching a plateau which reflects the partial asynchrony of the cells at the time. This cluster is enriched for genes annotated with chromosome segregation, organelle fission, and diverse cytoskeletal components like spindle proteins and myosins. All these GO terms can be linked to mitotic cell division.

The G1 cluster has 157 genes. Genes in this cluster show higher expression levels in both the late and initial time points of the experiment. Many GO terms enriched in this cluster are related to the mitochondrion and diverse metabolic processes that indicate an increase in cell metabolism as the cell progresses through the cell cycle [97].

Taken together, the GO enrichment analyses show that gene expression clusters contain conserved genes involved in the cell cycle in *Capsaspora.*

Conserved temporal order of cyclin and CDK expression in *Capsaspora*

In eukaryotes, the cell cycle events are regulated by cyclins in complex with CDKs. While the cyclin and CDK gene families are broadly conserved across eukaryotes [2,3], some of the subfamilies are lineage specific and have radiated differently. In budding yeast, two types of cell cycle cyclins can be found: cyclin B and Cln-type cyclins. Both types bind to one single CDK, Cdk1 [16]. In animals, this ancestral CDK expanded and specialized [4] resulting in multiple cyclin-CDK partners involved in different phases of the cell cycle: CDK2 binds to cyclins E and A at the onset and later stages of S phase [13], cyclin B binds to Cdc2 in

mitosis [14], and CDK4 and CDK6 bind to cyclin D types during G1 [98]. To our knowledge, how cyclin-CDK binding partnership grew in complexity remains unclear.

To provide some insights into the evolution of cyclin-CDK binding partnership, we used our temporal gene expression dataset to infer the timing of cyclin-CDK activity during the cell cycle in *Capsaspora*. Due to inconsistencies in published work [4,59], we first revised the classification of cyclin and CDK genes found in *Capsaspora* by phylogenetic profiling using a complete taxon sampling for Holozoa (the clade comprising animals and their closest unicellular relatives) and validated the gene previously reported as *Capsaspora* CDK1 [59] by Sanger sequencing and transcriptomics (see Methods) (Supplementary File 4, Supplementary File 5, Supplementary Figs 6-8). Despite the limited phylogenetic resolution, *Capsaspora* and other unicellular holozoan sequences appeared in earlier branching positions to the metazoan Cyclin A, B and E clades, this being compatible with *Capsaspora* having orthologs of these cyclins (Supplementary Fig. S6). The phylogeny does not suggest the presence of a Cyclin D ortholog in *Capsaspora*. In the CDK phylogeny (Fig. 5-Fig. supplement 2), sequences from *Capsaspora* and other filastereans branch as a sister-group to the metazoan CDK1 clade (100% of UFBoot), whereas the branching of sequences from other unicellular holozoans is more uncertain concerning the CDK1 and CDK2-3 metazoan clades. From this, we envision two possible evolutionary scenarios. In a first scenario, an ancestral duplication of CDK1-3 into CDK1 and CDK2-3 occurred in a common ancestor of Holozoa. As most unicellular holozoans have only one sequence within the CDK1-3 clade, this would imply differential losses of either CDK1 or CDK2-3 in Ichthyosporea, Filasterea and Choanoflagellatea, and Metazoa conserving both paralogs. Despite *Salpingoeca rosetta* has two sequences within the CDK1-3 clade, both paralogs are likely to descend from a duplication event occurred in the Choanoflagellatea lineage, with *Monosiga brevicollis* losing one of the two copies. In a second scenario, which we find more parsimonious, the duplication would have occurred in the lineage leading to Metazoa, but the limiting phylogenetic signal would not have allowed reconstruction of the real phylogenetic pattern of the CDK1-3 clade. Thus, we propose that the subfunctionalization of CDK1-3 is a specific feature of Metazoa,

with Capsaspora retaining the ancestral CDK1-3 gene instead of having a CDK1 ortholog as previously reported [59].

We report a clear temporal ordering of expression of the putative *Capsaspora* Cyclins A, B, and E (Fig. 5A). Cyclin E belongs to the G1/S cluster, cyclin A clusters together with S-phase genes, and cyclin B is in the M cluster, peaking at mitosis (Fig.4) together with *Capsaspora* CDK1-3 (Fig. 5B), although we found the CDK1/2/3 transcript expressed at high levels also during the rest of the cell cycle. In conclusion, our results suggest that cyclins A, B and E follow the same temporal order and cell cycle phases as cyclins in human cells.

The *Capsaspora* periodic transcriptional program includes ancient eukaryotic genes and is similar to that of animal cells

Besides the cyclin-CDK system, other regulators are periodically expressed during the cell cycle [32,39,40]. To characterize the periodic expression program in *Capsaspora* in comparison to other species, we identified orthologs of cell cycle regulators from other species in *Capsaspora* using OrthoFinder [99] (Supplementary File 5) and using a list of one-to-one *Capsaspora*-human orthologs from a set of phylogenies of *Capsaspora* [61]. From these sources, we identified which periodic human genes with known functions in the cell cycle (as described in [32,39,40,100]) have also a periodic ortholog in *Capsaspora*. We found that numerous DNA replication genes are upregulated in the G1/S cluster in *Capsaspora*, including DNA polymerase subunits, replication factors, and proteins CDC45 and PCNA (Fig. 5C). Among human genes that peak in mitosis, we found *Capsaspora* orthologs of Aurora Kinase A (AURKA), protein regulator of cytokinesis 1 (PRC1) and the anaphase-promoting complex (APC) subunit UBE2S expressed in the G2/M cluster (Fig. 5D). In human cells, AURKA regulates the assembly of the centrosome and the mitotic spindle [101], mitotic cyclins and cohesins are degraded by the APC [1], and PRC1 regulates cytokinesis by cross-linking spindle midzone microtubules [102]. Although we did not find regulatory APC subunits CDC20 and CDH1 [103,104] among periodically expressed genes in *Capsaspora*, the UBE2S peak in mitosis suggests that APC activity might also be transcriptionally regulated during the cell cycle in

363 *Capsaspora*. During M phase, we also observed upregulation of different
364 kinesins, microtubule motors with conserved function in the mitotic spindle, and
365 centromere proteins; a more detailed overview is provided in Supplementary File
366 7. We also identified *Capsaspora* periodic genes belonging to the same
367 orthogroups as other cell cycle regulators in human cells, but without a one-to-
368 one ortholog relationship (Supplementary File 8).

369 In addition to the examples above, we were interested in the similarities of
370 the periodic transcriptional program of *Capsaspora* with those of other eukaryote
371 species. First, to understand the evolutionary origin of the periodic genes found
372 in *Capsaspora*, we calculated gene age enrichment for every cell cycle cluster.
373 We assigned a gene age to the orthogroups by Dollo parsimony [105] and
374 compared the enrichment ratios for non-periodic genes and the five clusters
375 separately with the gene age of the whole transcriptome of *Capsaspora*. Three
376 of the clusters of periodic genes presented significant enrichment in pan-
377 eukaryotic genes (Fig. 6A, Fig. 6—source data 1). Our data thus shows that a
378 large fraction of genes in the periodic transcriptional program of *Capsaspora*
379 belong to gene families originating early in eukaryotic evolution.

380 Next, we compared our dataset of *Capsaspora* periodic genes with datasets
381 of cell cycle synchronized cells of different organisms, namely three different cell
382 types of *Homo sapiens* (Hela Cells, U2OS cells, and foreskin primary fibroblasts)
383 [32,41,50], *Saccharomyces cerevisiae* [43], *Schizosaccharomyces pombe* [91],
384 and *Arabidopsis thaliana* [44]. For each dataset, we took the published lists of
385 periodic genes and corrected for the number of genes in each species (Fig. 6—
386 source data 2). We set a threshold of less than 10% of genes to be periodic for
387 human cells and the yeasts, and less than 5% for *A. thaliana*. Thus 1790 periodic
388 genes were identified in HeLa cells, 1245 in U2OS cells, 461 in fibroblasts, 592
389 in *S. cerevisiae*, 499 in *S. pombe*, and 1060 in *A. thaliana* datasets. We found
390 1925 orthogroups that contained at least one periodic gene from either of the
391 datasets (Fig. 6B), and named these “periodic orthogroups”. Of these, one third
392 had orthologs in all five species.

393 We first computed the number of periodic genes shared between pairs of
394 species (defined by their presence in the same orthogroup). Overall, all species

share a small number of periodic genes (Fig. 6—source data 3, Fig. 6—source data 4), with the number of genes being highest for *Capsaspora* with *H. sapiens* cell lines (167, 81 and 99 periodic orthogroups with HeLa cells, U2OS cells, and fibroblasts, respectively) rather than yeasts or *A. thaliana* (Fig. 6D, Supplementary File 11, Supplementary File 12). Still, the periodic expression of the majority of the genes is not shared between species, consistent with former findings of 2% to 5% of periodic genes shared between different organisms [51]. This indicates that the periodic transcriptional program is divergent both between species and even between different cell types within an organism.

Despite the low numbers of periodically expressed genes in common, we wondered whether the periodic transcriptional program in each species indeed evolved independently, meaning that the pairs of genes that share the periodic expression between species are observed only by chance. To that end, we calculated the expected number of shared periodic orthogroups by chance as the product of ratios of periodic orthogroups from each species within their orthogroups in common (Fig. 6C). We detected that, for most species, the number of shared periodic genes is higher than by chance, especially for *Capsaspora* and the core cell cycle gene set of *H. sapiens* (defined in [32]) (Fig. 6D, Supplementary File 11, Supplementary File 12). We found the same when comparing periodic one-to-one orthologs [61] (Fig. 6D) or when defining periodic genes on every species using the same method that we applied to *Capsaspora* (Supplementary Fig. S9, Supplementary Files 10-12). Therefore, these findings are robust with respect to the methods used to identify periodically expressed genes and to assign orthology relationships. Thus, although the cell cycle periodic expression program largely evolves fast and independently, our data suggest there is a core set of genes of conserved oscillatory expression during the cell cycle (Fig. 6B).

Overall, our cross-species comparison of the periodic gene expression programs revealed that the *Capsaspora* periodic gene expression program is more similar to human cells that to current unicellular model systems for the cell cycle. Furthermore, including a new species in the global analysis, we discovered a previously unappreciated core set of genes for which periodic expression is deeply conserved.

Discussion

In this study, we have used synchronized *Capsaspora* adherent cells to gain insight into key aspects of the cell cycle, such as cell division and periodic gene expression, in a unicellular relative of animals. Previously, the cell cycle has been studied in only a handful of species due to the inability to obtain synchronous cell cultures. With this synchronization protocol, the cell cycle of a closer relative of animals can be studied in cultures that can be synchronized from DNA replication to cell division.

Our experimental setup made possible to characterize mitotic cell division in *Capsaspora*, which we found relies on microtubule-based structures, as previously described in other eukaryote species (Forth and Kapoor, 2017). Our observations suggest the presence of a putative non-centrosomal microtubule organizing center (MTOC) in *Capsaspora*, which raises new questions about the mechanisms of cell division in this species. As non-centrosomal MTOCs have independently evolved in many different species [107], it may well be that *Capsaspora* has a non-centrosomal, independently evolved MTOC, or that their microtubules are able to self-arrange, as previously shown in other systems [108].

Synchronization of *Capsaspora* cell cultures allowed us to study transcription during the cell cycle using RNA sequencing. We identified five waves of gene expression across time, with most genes being expressed in the G1/S transition and in mitosis. As in previously studied organisms, these waves of transcription can be grouped in clusters containing genes related to the main events of the phases of the cell cycle. The periodic transcriptional program of *Capsaspora* is enriched in genes that emerged at the onset of eukaryotes, showing that the cell cycle relies in numerous genes, such as DNA replication proteins and cytoskeleton components, which are common to all eukaryotes due to their roles in fundamental cellular processes. Although transcriptional activity during the cell cycle is present in numerous species and cell types [51], the genes affected by cell-cycle-regulated transcription are divergent between distantly related species, likely due to the fact that transcriptional regulation adapts to the environment and lifestyle of each particular cell type. Our observations suggest

that this occurs largely by old genes gaining and losing periodic regulation, rather than new species-specific genes evolving to be regulated during the cell cycle. This is consistent with Jensen et al.'s observations that periodicity in complex activity can evolve rapidly in different lineages by recruiting different partners of the same complex, which preserves the periodic regulation of the entire complex [51]. Still, in contrast to previous analysis with a more limited set of species, our analysis clearly revealed a core set of genes of which the periodic regulation is deeply conserved among eukaryotes.

From the temporal order of gene expression of *Capsaspora* cyclins, we conclude that cyclins A, E, and B follow the same order and are associated with the same cell cycle stage as in *H. sapiens* cells. In contrast to human cells, where these cyclins bind their respective partner CDKs, *Capsaspora* only possesses one ancestral CDK1-3, which, although periodically expressed with a peak in M-phase, exhibits high transcript levels throughout the cell cycle. Due to this, we propose that CDK1-3 might be the binding partner of cyclin A, B, and E in *Capsaspora* and that it might be involved in all phases of the cell cycle (Fig. 7). Nevertheless, in the absence of biochemical data, we cannot exclude the possibility that *Capsaspora* cyclins A and E bind other non-canonical CDKs. Interestingly, in knock-out mice where all CDKs except for CDK1 were deleted, all cyclins were also found to bind CDK1[109], which suggests that animal cyclins A, B and E are also able to bind CDK1. Given these data, we reconstruct a likely evolutionary scenario of the evolution of the cyclin-CDK system in opisthokonts. The ancestral opisthokont likely possessed a single CDK, with a role in multiple cell cycle phases, and B-type cyclins. Cyclins underwent duplication and subfunctionalization first, acquiring roles in regulating distinct phases of the cell cycle while binding to the single CDK. This evolutionary intermediate state is present in *Capsaspora*. During the emergence of animals, CDKs also underwent expansion and subfunctionalized to bind specific cyclins, forming discrete complexes active in a particular stage. This suggests that there was a gradual evolution of the cyclin-CDK control of the cell cycle during the emergence of animals.

Author contributions

A.P., O.D. and A.O. set up the methodology, and performed the time-series experiment. A.P. performed the flow cytometry, analyzed the gene expression data, sequenced the *Capsaspora* CDK1 gene, performed the comparative analysis, and wrote the original draft. O.D. performed the microscopy. E.O.P. performed the phylogenetic analysis. I.R.T. provided supervision. A.O. conceived the project, performed the *Capsaspora* RNA extraction, and provided supervision. All authors reviewed and edited the manuscript.

Acknowledgements

We thank all the members of Iñaki Ruiz-Trillo's lab for discussion and comments on the manuscript, Daniel Richter for discussion on the statistical tests, Xavier Grau-Bové for helpful discussions on the data analysis, Arnau Sebé-Pedrós for providing resources of the *Capsaspora* phylome, and Meritxell Antó-Subirats for technical support. We also acknowledge the UPF Flow Cytometry Core Facility for assistance with flow cytometry, and the CRG Genomics Unit for mRNA library preparation and Illumina sequencing.

This work was funded by a European Research Council Consolidator Grant (ERC-2012-Co-616960) to I.R.T; and a grant from the Spanish Ministry for Economy and Competitiveness (MINECO; BFU2014-57779-P, with European Regional Development Fund support) to I.R.T. A.P. was supported by a "la Caixa" Foundation (ID 100010434) fellowship, whose code is LCF/BQ/ES16/11570008. O.D. was supported by a Swiss National Science Foundation Early PostDoc Mobility fellowship (P2LAP3_171815) and a Marie Sklodowska-Curie individual fellowship (MSCA-IF 746044). E.O.P. was supported by a pre-doctoral FPI grant from MINECO. A.O. was supported by a Marie Sklodowska-Curie individual fellowship (MSCA-IF 747086). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Materials and methods

Cell cultures and culture synchronization

Capsaspora cells were incubated at 23ºC in ATCC medium 1034 (modified PYNFH medium). Two independent cultures of *Capsaspora* at 30-50% confluency were treated using 10mM Hydroxyurea (Sigma Aldrich, Saint Louis, MO, USA, #H8627) in culture medium, and left incubating for approximately 14 hours.

Cells were released from Hydroxyurea by washing prior to elution in fresh medium; samples were collected by scraping and washing two hours after release, and from there on every 45 minutes until thirteen hours. A total of sixteen time points were taken, constituting a time window comprising one event of genome duplication and one mitotic division.

We also tested different concentrations and incubation times of hydroxyurea, nocodazole (Sigma-Aldrich, #M1404), and aphidicolin (Sigma-Aldrich, #A0781). Only 10mM hydroxyurea for longer than thirteen hours showed arrest of the cell cycle, while Nocodazole had no observable effect by DNA content measurement, and the rest ruined the samples due to insolubility of the compound.

DNA content measurement

Cells were washed in PBS and fixed in 70% ethanol in PBS, then incubated in RNAse A (Sigma-Aldrich, #R6148) (one volume in three volumes of 1xPBS) for 24 hours at 37 ºC. Cells were then incubated in a final concentration of 20µg/ml propidium iodide (Sigma-Aldrich, #P4170-25MG) for 72 hours at 4 ºC.

Samples were analyzed by flow cytometry using a BD LSR Fortessa analyser (Becton Dickinson, Franklin Lakes, NJ, USA). SSC-A and FSC-A were used to detect populations of stained cells. Single cells were gated by FSC-H and FSC-A. An average of 10,000 events per sample were recorded. PI-positive cells were detected using a 561 nm laser with a 610/20 band pass filter (red

550 fluorescence). To estimate the cell count, Texas Red-A was plotted as
551 histograms using FlowJo 9.9.3 (FlowJo LLC, Ashland, OR, USA).

552 **Cell microscopy and image analysis**

553 Microscopy pictures were taken using a Zeiss Axio Observer Z.1
554 Epifluorescence inverted microscope equipped with Colibri LED illumination
555 system and Axiocam 503 mono camera (Carl Zeiss microscopy, Oberkochen,
556 Germany). A Plan-Apochromat 100X/1.4 oil objective (Nikon Corporation, Tokyo,
557 Japan) was used for imaging fixed cells. For the live imaging, we used an EC
558 Plan-Neofluar 40x/0.75 air objective (Carl Zeiss microscopy).
559
560 Image analysis was done using ImageJ software [110]. For fixed cells, we
561 used the oval selection tool to draw the contour of each cell and measured cell
562 perimeter. As cells are spherical, we computed cell area using ImageJ. We
563 estimated the relative cell area of every pair of daughter cells by dividing each
564 measurement by the sum of the two daughter cells areas. All the calculation and
565 data plotting was done in R Software ver. 3.4.4[111].

566 **RNA isolation and sequencing**

567 Time point samples were washed in 1xPBS, poured in Trizol, and frozen at
568 -80ºC. Total RNA was purified using Zymo RNA miniprep kit (Zymo Research,
569 Irvine, CA, USA, #R2050). mRNA libraries were prepared using the TruSeq
570 Stranded mRNA Sample Prep kit (Illumina, San Diego, CA, USA, Cat. No. RS-
571 122-2101). Paired-end 50bp read length sequencing was carried out at the CRG
572 genomics core unit on an Illumina HiSeq v4 sequencer, with all samples from the
573 same replicate being pooled in the same lane.
574
575 *Capsaspora* adherent cultures cDNA was obtained by RT-PCR using
576 SuperScript® III Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA,
577 #18080044) following the manufacturer's instructions. PCR was performed using
578 Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA,
579 USA, #M0530L) following the manufacturer's instructions.

Transcriptomic analysis

RNA reads were mapped using Kallisto v0.43.1[88] using default parameters onto a set of the largest isoforms of the *Capsaspora* transcriptome[62]. The resulting time-series transcriptome in transcript-per-million (tpm) units is available in Supplementary File 1. We retrieved only the transcripts whose average expression level is above 1 tpm in the whole time series.

Gene expression level was normalized by subtracting the mean over time and dividing by the standard deviation. Normalized datasets were clustered according to their Spearman correlation values using hierarchical clustering (R gplots library[111,112]).

Identification of periodically expressed genes

Periodic genes were detected in *Capsaspora* by ranking using JTK_CYCLE [89] and RAIN [90] on the time-series transcriptomes and an average of the two replicates. We set JTK_CYCLE parameters to periods=14:16 and sampling interval=0.75, and ranked every gene by their BH.Q value. We set RAIN parameters to period=16 and delta=0.75, and ranked every gene by their Benjamini-Hochberg corrected p-value. We set a cutoff of the genes ranked below 2000 on each separate replicate and simultaneously ranked below 800 in the average dataset (see Supplementary Fig S1, Supplementary File 2).

Clustering analysis

Periodic genes were hierarchically clustered according to similarity of gene expression over time (averaged between two replicates), and clustered by k-means clustering[111] using standard parameters and k=5. Agreement between clustering methods was calculated as the number of genes belonging to the same pair of clusters divided by the size of the smallest cluster in the pair.

Calculation of Gene ontology enrichment

Gene ontology enrichment was calculated using Ontologizer[92] using the –c “Parent-Child-Union” –m “Bonferroni” options. Bonferroni-corrected p-values were taken as significant when below 0.05.

Phylogenetic classification of CDKs and Cyclins from early-holozoa taxa

Annotated sequences of cyclins and CDKs from *H. sapiens* and *S. cerevisiae* were retrieved from Cao et al.[4] and Swissprot[113], and *A. thaliana* sequences were taken from[113–115]. These sequences were used as queries to detect potential CDKs and cyclins orthologues in our dataset (Supplementary File 4) using BLAST+ v2.3.0[116,117]. Those sequences that aligned were BLASTed against a database including all proteins from *H. sapiens*, *S. cerevisiae* and *A. thaliana*. We only included in our phylogeny those sequences whose best hit against this database matched the original sequences used in the detection step. Proteins were aligned with MAFFT v7.123b.[118], using the -einsi option, and alignments were trimmed using trimAl v1.4.rev15[119] with the -gappyout option. Trimmed alignments were manually inspected and cleaned of poorly informative sequences except if that sequences corresponded to early-branching holozoa (ebH) taxa. Cleaned alignment were used as inputs for phylogenetic inference with IQ-TREE v1.6.7[120], using the -bb 1000 and -mset LG options. For CDKs, since the tree topology did not show any ebH sequences belonging to CDK families without orthologues in Metazoa, we performed a second phylogenetic inference using only the ebH and metazoan proteins to reduce the potential phylogenetic noise that may be introduced with the addition of non-informative divergence.

Sequences were classified into CDK/Cyclins families taking into account the tree topology and the UFBoot nodal support values[121]. Families were named according to the orthology relationship between the ebH protein and the *H. sapiens* sequence. For example, Sarc_g11690T was classified as CDK10 because its position in the tree suggests an orthology relationship to this *H. sapiens* protein (Supplementary Fig. S7), whereas Cowc_CAOG_08444T0 was classified as CDK11-CDK11B because it is orthologue to both *H. sapiens*

proteins (Supplementary Fig. S7). We found three ebH cyclin subfamilies without orthologues in Metazoa, which were named accordingly to the corresponding orthologue in *Saccharomyces cerevisiae* (PCL1, PCL2, PCL9; PCL5, CLG1; and PCL6-PCL7). Those ebH sequences showing ambiguous or poorly supported branching patterns were classified as uncertain.

We found that the current *Capsaspora* sequence reported by[59] as a CDK1 orthologue (CAOG_07905) is considerably shorter in length than other CDK orthologues. This could explain why it was not detected in the work by Cao et al.[4], if an e-value constraint was taken into consideration. By inspection of the genomic sequence from[122], we detected an assembly gap of 1kb neighboring the 3' end of the predicted annotation of CAOG_07905 (Supplementary Fig. S8A). In silico translation of the surroundings of this gap revealed protein domains conserved in H. *sapiens* and S. *cerevisiae* CDK1 proteins (Supplementary Fig. S8A-8C). Upon realization that the gene annotation of *Capsaspora* CDK1-3 was incomplete, we designed forward (5'-GCTCAAGGAGGTCATCCACC-3') and reverse (5'-CTCTCTGCCCGATTACAAGC-3') primers to PCR-amplify the unknown sequence from both genomic DNA and cDNA (Supplementary Fig. S8A-8B). Sanger sequencing of the amplified products revealed one more intron and one more exon, which were used to reconstruct the missing sequence in the assembly. We mapped the transcriptome of Replicate 2 – timepoint 9 against this reconstructed sequence using tophat2[123] with standard parameters. Identification of paired-end overlapping reads in the reconstructed exons using Tablet [124] verified the results of Sanger sequencing (Supplementary Fig. S8A). The updated cDNA and protein sequence of *Capsaspora* CDK1-3, which aligns much better with human and yeast sequences (Supplementary Fig. S8C), can be found in Supplementary File 6.

Identification of orthologous genes

Groups of orthologs were generated using OrthoFinder v2.1.2 [99] using default parameters (evalue 10e-3, MCL clustering) on a dataset of proteomes found in Supplementary File 5[113,125–127].

One-to-one orthologues were taken from a phylome of 6598 genes of *Capsaspora* reconstructed in [61] using the algorithm by [128] (http://phylomedb.org/phylome_100).

Determining cell cycle-regulated genes in H. sapiens, S. cerevisiae, S. pombe and A. thaliana

We downloaded the lists of cell cycle regulated genes which can be found at [32,41,43,44,50,91]. We translated the gene and probe names of their datasets using BioMart[129] and Uniprot[113] tools (see Supplementary File 10).

Gene Age enrichment analysis

We used Count software[105] to assign the emergence of every orthogroup –and therefore every *Capsaspora* gene- to a given ancestor in common between species by Dollo Parsimony. We defined six different ages for *Capsaspora*: 0 "*Capsaspora*-specific", 1 "Filozoa", 2 "Holozoa", 3 "Opisthokonta", 4 "Unikonta", and 5 "Paneukaryotic". All *Capsaspora* genes unassigned to any orthogroup were defined as "*Capsaspora*-specific". Gene age enrichment was calculated using contingency tables and significance by Fisher exact test using R software ver. 3.4.4.[111], and was corrected for multi-test hypothesis using Bonferroni correction.

Comparative analysis

Every periodic and non-periodic gene of each of the seven datasets were assigned to their respective orthogroup, if any. We obtained seven lists of orthogroups containing periodic genes of each species or cell type, and also defined lists of orthogroups containing genes (regardless of periodic) of each of the five species. A subset of periodic orthogroups (those containing at least one periodic gene from at least one of the seven datasets) was generated, and plotted

for periodicity, presence, or absence, in all the datasets using R gplots library[112].

Numbers and ratios of periodic shared orthogroups and one-to-one orthologues, as well as binomial test p-values (see Fig.6D, Supplementary Fig. S9, Supplementary File 11, Supplementary File 12), were calculated using R software ver. 3.4.4. [127]. For each pair-wise comparison of species, e.g. *Capsaspora* and *Homo sapiens* HeLa cells (Fig.6C), we took the number of orthogroups they have in common as a total population, *C*. Then we looked at the number of periodic orthogroups from each species that are within this total population, *p1* and *p2*, and calculated the null expectation (A_{exp}) as a product of the ratios of these two subpopulations within the population of orthogroups in common.

Reanalysis of cell cycle datasets using JTK_CYCLE and RAIN

We reanalyzed the datasets of gene expression of [44] (HU treatment dataset), [91] (three replicates of elutriation), [43] (two replicates of wildtype synchronous yeast cultures) , [41] (one replicate, thymidine block), [32] (four replicates of double thymidine block), and [50] (a dataset of ~8000 genes matching filtering criteria by the authors) using the same pipeline used in our *Capsaspora* datasets. For those with replicates, periodicity ranks were calculated for each replicate independently and summed at the end.

As every experiment comprised different numbers of cell cycles of different length, we set up JTK and RAIN parameters to look for periodicity in time lapses according to the author's reports (see Supplementary File 10). We corrected for the number of genes by setting a threshold of less than 10% of genes to be periodic. Overlap between datasets of periodic genes was calculated using R Software ver. 3.4.4. [111].

REFERENCES

1 Morgan DO. Cell Cycle: Principles of Control. *Yale J Biol Med* 2007. doi:10.1093/icb/icm066.

2 Medina EM, Turner JJ, Gordân R, Skotheim JM, Buchler NE. Punctuated evolution and transitional hybrid network in an ancestral cell cycle of fungi. *Elife* 2016; **5**: 1–23.

3 Cross FR, Buchler NE, Skotheim JM. Evolution of networks and sequences in eukaryotic cell cycle control. Philos. Trans. R. Soc. B Biol. Sci. 2011. doi:10.1098/rstb.2011.0078.

4 Cao L, Chen F, Yang X, Xu W, Xie J, Yu L. Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. *BMC Evol Biol* 2014. doi:10.1186/1471-2148-14-10.

5 Li Z. Regulation of the Cell Division Cycle in Trypanosoma brucei. *Eukaryot Cell* 2012; **11**: 1180–1190.

6 Gutierrez C. The Arabidopsis Cell Division Cycle. *Arab B* 2009. doi:10.1199/tab.0120.

7 Nieduszynski CA, Murray J, Carrington M. Whole-genome analysis of animal A- and B-type cyclins. *Genome Biol* 2002. doi:10.1186/gb-2002-3-12-research0070.

8 Satyanarayana A, Kaldis P. Mammalian cell-cycle regulation: Several cdks, numerous cyclins and diverse compensatory mechanisms. Oncogene. 2009. doi:10.1038/onc.2009.170.

9 Matsushime H, Roussel MF, Ashmun RA, Sherr CJ. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. *Cell* 1991; **65**: 701–713.

10 Matsushime H, Quelle DE, Shurtleff SA, Shibuya M, Sherr CJ, Kato JY. D-type cyclin-dependent kinase activity in mammalian cells. *Mol Cell Biol* 2015; **14**: 2066–2076.

754 11 Klein EA, Assoian RK. Transcriptional regulation of the cyclin D1 gene at
755 a glance. *J Cell Sci* 2008; **121**: 3853–3857.

756 12 Cook JG, Park C-H, Burke TW, Leone G, DeGregori J, Engel A *et al.*
757 Analysis of Cdc6 function in the assembly of mammalian prereplication
758 complexes. *Proc Natl Acad Sci* 2002. doi:10.1073/pnas.032677499.

759 13 Coverley D, Laman H, Laskey RA. Distinct roles for cyclins E and A
760 during DNA replication complex assembly and activation. *Nat Cell Biol*
761 2002. doi:10.1038/ncb813.

762 14 Pomerening JR, Sontag ED, Ferrell JE. Building a cell cycle oscillator:
763 Hysteresis and bistability in the activation of Cdc2. *Nat Cell Biol* 2003.
764 doi:10.1038/ncb954.

765 15 Pomerening JR, Sun YK, Ferrell JE. Systems-level dissection of the cell-
766 cycle oscillator: Bypassing positive feedback produces damped
767 oscillations. *Cell* 2005; **122**: 565–578.

768 16 Mendenhall MD, Hodge a E. Regulation of CDC28 cyclin-dependent
769 protein kinase activity during the cell cycle of the yeast *Saccharomyces*
770 *cerevisiae*. *Microbiol Mol Biol Rev* 1998.

771 17 Bloom J, Cross FR. Multiple levels of cyclin specificity in cell-cycle control.
772 Nat. Rev. Mol. Cell Biol. 2007. doi:10.1038/nrm2105.

773 18 Cross FR, Tinkelenberg AH. A potential positive feedback loop controlling
774 CLN1 and CLN2 gene expression at the start of the yeast cell cycle. *Cell*
775 1991. doi:10.1016/0092-8674(91)90394-E.

776 19 Nasmyth K, Dirick L. The role of SWI4 and SWI6 in the activity of G1
777 cyclins in yeast. *Cell* 1991. doi:10.1016/0092-8674(91)90444-4.

778 20 Richardson HE, Wittenberg C, Cross F, Reed SI. An essential G1 function
779 for cyclin-like proteins in yeast. *Cell* 1989. doi:10.1016/0092-
780 8674(89)90768-X.

781 21 Skotheim JM, Di Talia S, Siggia ED, Cross FR. Positive feedback of G1

cyclins ensures coherent cell cycle entry. *Nature* 2008. doi:10.1038/nature07118.

22 Epstein CB, Cross FR. CLB5: A novel B cyclin from budding yeast with a role in S phase. *Genes Dev* 1992. doi:10.1101/gad.6.9.1695.

23 Schwob E, Nasmyth K. CLB5 and CLB6, a new pair of B cyclins involved in DNA replication in Saccharomyces cerevisiae. *Genes Dev* 1993. doi:10.1101/gad.7.7a.1160.

24 Surana U, Robitsch H, Price C, Schuster T, Fitch I, Futcher AB *et al.* The role of CDC28 and cyclins during mitosis in the budding yeast S. cerevisiae. *Cell* 1991. doi:10.1016/0092-8674(91)90416-V.

25 Richardson H, Lew DJ, Henze M, Sugimoto K, Reed SI. Cyclin-B homologs in Saccharomyces cerevisiae function in S phase and in G2. *Genes Dev* 1992. doi:10.1101/gad.6.11.2021.

26 Martin-Castellanos C, Labib K, Moreno S. B-type cyclins regulate G1 progression in fission yeast in opposition to the p25rum1 cdk inhibitor. *EMBO J* 1996; **15**: 839–49.

27 Mondesert O, McGowan CH, Russell P. Cig2, a B-type cyclin, promotes the onset of S in Schizosaccharomyces pombe. *Mol Cell Biol* 2015; **16**: 1527–1533.

28 Booher R, Beach D. Involvement of cdc13+ in mitotic control in Schizosaccharomyces pombe: possible interaction of the gene product with microtubules. *EMBO J* 1988. doi:10.1002/j.1460-2075.1988.tb03075.x.

29 Booher RN, Alfa CE, Hyams JS, Beach DH. The fission yeast cdc2/cdc13/suc1 protein kinase: Regulation of catalytic activity and nuclear localization. *Cell* 1989. doi:10.1016/0092-8674(89)90429-7.

30 Coudreuse D, Nurse P. Driving the cell cycle with a minimal CDK control network. *Nature* 2010. doi:10.1038/nature09543.

31 Goodrich DW, Wang NP, Qian YW, Lee EYHP, Lee WH. The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle. *Cell* 1991. doi:10.1016/0092-8674(91)90181-W.

32 Grant GD, Brooks L, Zhang X, Mahoney JM, Martyanov V, Wood TA *et al.* Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors. *Mol Biol Cell* 2013; **24**: 3634–3650.

33 Ishida S, Huang E, Zuzan H, Spang R, Leone G, West M *et al.* Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. *Mol Cell Biol* 2001; **21**: 4684–4699.

34 Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M, Brown PO. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 2001. doi:10.1038/35054095.

35 Gefeng Z, Spellman PT, Volpe T, Brown PO, Botstein D, Davis TN *et al.* Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. *Nature* 2000; **406**: 90–94.

36 Reynolds D, Shi BJ, McLean C, Katsis F, Kemp B, Dalton S. Recruitment of Thr 319-phosphorylated Ndd1p to the FHA domain of Fkh2p requires C1b kinase activity: A mechanism for CLB cluster gene activation. *Genes Dev* 2003; **17**: 1789–1802.

37 Pramila T, Wu W, Miles S, Noble WS, Breeden LL. The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle. *Genes Dev* 2006. doi:10.1101/gad.1450606.

38 Laoukili J, Kooistra MRH, Brás A, Kauw J, Kerkhoven RM, Morrison A *et al.* FoxM1 is required for execution of the mitotic programme and chromosome stability. *Nat Cell Biol* 2005. doi:10.1038/ncb1217.

39 Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE *et al.* Identification of Genes Periodically Expressed in the Human Cell

Cycle and Their Expression in Tumors. *Mol Biol Cell* 2002; **13**: 1977–2000.

40 Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, Sapinoso L *et al.* Transcriptional regulation and function during the human cell cycle. *Nat Genet* 2001. doi:10.1038/83751.

41 Bar-Joseph Z, Siegfried Z, Brandeis M, Brors B, Lu Y, Eils R *et al.* Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc Natl Acad Sci* 2008. doi:10.1073/pnas.0704723105.

42 Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB *et al.* Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. *Mol Biol Cell* 1998. doi:10.1091/mbc.9.12.3273.

43 Orlando DA, Lin CY, Bernard A, Wang JY, Socolar JES, Iversen ES *et al.* Global control of cell-cycle transcription by coupled CDK and network oscillators. *Nature* 2008; **453**: 944–947.

44 Menges M, Hennig L, Gruissem W, Murray JAH. Genome-wide gene expression in an Arabidopsis cell suspension. *Plant Mol Biol* 2003; **53**: 423–442.

45 Breyne P, Dreesen R, Vandepoele K, De Veylder L, Van Breusegem F, Callewaert L *et al.* Transcriptome analysis during cell division in plants. *Proc Natl Acad Sci* 2002; **99**: 14825–14830.

46 Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JCF *et al.* The transcriptional program in the response of human fibroblasts to serum. *Science (80-)* 1999. doi:10.1126/science.283.5398.83.

47 Kelliher CM, Leman AR, Sierra CS, Haase SB. Investigating Conservation of the Cell-Cycle-Regulated Transcriptional Program in the Fungal Pathogen, Cryptococcus neoformans. *PLoS Genet* 2016; **12**: 1–23.

867 48 Zones JM, Blaby IK, Merchant SS, Umen JG. High-Resolution Profiling of
868 a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii
869 Reveals Continuous Cell and Metabolic Differentiation. *Plant Cell* 2015;
870 **27**: 2743–2769.

871 49 Peña-Diaz J, Hegre SA, Anderssen E, Aas PA, Mjelle R, Gilfillan GD *et al.*
872 Transcription profiling during the cell cycle shows that a subset of
873 Polycomb-targeted genes is upregulated during DNA replication. *Nucleic*
874 *Acids Res* 2013; **41**: 2846–2856.

875 50 Dominguez D, Tsai YH, Gomez N, Jha DK, Davis I, Wang Z. A high-
876 resolution transcriptome map of cell cycle reveals novel connections
877 between periodic genes and cancer. *Cell Res* 2016; **26**: 946–962.

878 51 Jensen LJ, Jensen TS, De Lichtenberg U, Brunak S, Bork P. Co-evolution
879 of transcriptional and post-translational cell-cycle regulation. *Nature* 2006.
880 doi:10.1038/nature05186.

881 52 Giotti B, Joshi A, Freeman TC. Meta-analysis reveals conserved cell cycle
882 transcriptional network across multiple human cell types. *BMC Genomics*
883 2017. doi:10.1186/s12864-016-3435-2.

884 53 Rhind N, Russell P. Chk1 and Cds1: linchpins of the DNA damage and
885 replication checkpoint pathways. *J Cell Sci* 2000; **113 (Pt 2**: 3889–96.

886 54 Cooper K. Rb, whi it's not just for metazoans anymore. *Oncogene* 2006;
887 **25**: 5228–5232.

888 55 Kearsey SE, Cotterill S. Enigmatic variations: Divergent modes of
889 regulating eukaryotic DNA replication. Mol. Cell. 2003.
890 doi:10.1016/S1097-2765(03)00441-6.

891 56 Ferrer-Bonet M, Ruiz-Trillo I. Capsaspora owczarzaki. Curr. Biol. 2017.
892 doi:10.1016/j.cub.2017.05.074.

893 57 Hertel LA, Bayne CJ, Loker ES. The symbiont Capsaspora owczarzaki,
894 nov. gen. nov. sp., isolated from three strains of the pulmonate snail

Biomphalaria glabrata is related to members of the Mesomycetozoea. *Int J Parasitol* 2002; **32**: 1183–1191.

58 Ruiz-Trillo I, Inagaki Y, Davis LA, Sperstad S, Landfald B, Roger AJ. Capsaspora owczarzaki is an independent opisthokont lineage [2]. Curr. Biol. 2004. doi:10.1016/j.cub.2004.10.037.

59 Suga H, Chen Z, de Mendoza A, Sebé-Pedrós A, Brown MW, Kramer E *et al.* The Capsaspora genome reveals a complex unicellular prehistory of animals. *Nat Commun* 2013; **4**: 1–9.

60 Sebé-Pedrós A, Ballaré C, Parra-Acero H, Chiva C, Tena JJ, Sabidó E *et al.* The Dynamic Regulatory Genome of Capsaspora and the Origin of Animal Multicellularity. *Cell* 2016; **165**: 1224–1237.

61 Sebé-Pedrós A, Peña MI, Capella-Gutiérrez S, Antó M, Gabaldón T, Ruiz-Trillo II *et al.* High-Throughput Proteomics Reveals the Unicellular Roots of Animal Phosphosignaling and Cell Differentiation. *Dev Cell* 2016; **39**: 186–197.

62 Sebé-Pedrós A, Irimia M, del Campo J, Parra-Acero H, Russ C, Nusbaum C *et al.* Regulated aggregative multicellularity in a close unicellular relative of metazoa. *Elife* 2013; **2013**: 1–20.

63 Parra-Acero H, Ros-Rocher N, Perez-Posada A, Kożyczkowska A, Sánchez-Pons N, Nakata A *et al.* Transfection of Capsaspora owczarzaki, a close unicellular relative of animals. *Dev* 2018; **145**. doi:10.1242/dev.162107.

64 Banfalvi G. *Cell Cycle Synchronization Methods and Protocols*. 2011 doi:10.1007/978-1-61779-182-6_18.

65 Blajeski AL, Phan VA, Kottke TJ, Kaufmann SH. G1 and G2 cell-cycle arrest following microtubule depolymerization in human breast cancer cells. *J Clin Invest* 2002. doi:10.1172/JCI13275.

66 Jackman J, O'Connor PM. Methods for Synchronizing Cells at Specific

Stages of the Cell Cycle. In: *Current Protocols in Cell Biology*. 2001 doi:10.1038/nm1207-1391.

67 Slater ML. Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle. *J Bacteriol* 1973.

68 Fox MH, Read RA, Bedford JS. Comparison of synchronized Chinese hamster ovary cells obtained by mitotic shake-off, hydroxyurea, aphidicolin, or methotrexate. *Cytometry* 1987. doi:10.1002/cyto.990080312.

69 Singh A, Agarwal A, Xu Y jie. Novel Cell-killing mechanisms of hydroxyurea and the implication toward combination therapy for the treatment of fungal infections. *Antimicrob Agents Chemother* 2017. doi:10.1128/AAC.00734-17.

70 Aubin JE, Weber K, Osborn M. Analysis of actin and microfilament-associated proteins in the mitotic spindle and cleavage furrow of PtK2 cells by immunofluorescence microscopy. A critical note. *Exp Cell Res* 1979. doi:10.1016/0014-4827(79)90260-X.

71 Cadart C, Zlotek-Zlotkiewicz E, Le Berre M, Piel M, Matthews HK. Exploring the function of cell shape and size during mitosis. Dev. Cell. 2014. doi:10.1016/j.devcel.2014.04.009.

72 Cramer LP, Mitchison TJ. Investigation of the mechanism of retraction of the cell margin and rearward flow of nodules during mitotic cell rounding. *Mol Biol Cell* 1997. doi:10.1091/mbc.8.1.109.

73 Harris A. Location of cellular adhesions to solid substrata. *Dev Biol* 1973. doi:10.1016/0012-1606(73)90009-2.

74 Kunda P, Pelling AE, Liu T, Baum B. Moesin Controls Cortical Rigidity, Cell Rounding, and Spindle Morphogenesis during Mitosis. *Curr Biol* 2008. doi:10.1016/j.cub.2007.12.051.

75 Luxenburg C, Amalia Pasolli H, Williams SE, Fuchs E. Developmental

roles for Srf, cortical cytoskeleton and cell shape in epidermal spindle orientation. *Nat Cell Biol* 2011. doi:10.1038/ncb2163.

76 Mcconnell CH. The mitosis found in hydra. *Science (80-)* 1930. doi:10.1126/science.72.1859.170-b.

77 Zang JH, Cavet G, Sabry JH, Wagner P, Moores SL, Spudich JA. On the role of myosin-II in cytokinesis: division of *Dictyostelium* cells under adhesive and nonadhesive conditions. *Mol Biol Cell* 1997. doi:10.1091/mbc.8.12.2617.

78 Carvalho A, Desai A, Oegema K. Structural Memory in the Contractile Ring Makes the Duration of Cytokinesis Independent of Cell Size. *Cell* 2009. doi:10.1016/j.cell.2009.03.021.

79 Chakraborty P, Wang Y, Wei JH, van Deursen J, Yu H, Malureanu L *et al.* Nucleoporin Levels Regulate Cell Cycle Progression and Phase-Specific Gene Expression. *Dev Cell* 2008. doi:10.1016/j.devcel.2008.08.020.

80 Eppinga RD, Li Y, Lin JJCLC, Lin JJC. Tropomyosin and caldesmon regulate cytokinesis speed and membrane stability during cell division. *Arch Biochem Biophys* 2006. doi:10.1016/j.abb.2006.06.015.

81 Bi E, Maddox P, Lew DJ, Salmon ED, McMillan JN, Yeh E *et al.* Involvement of an actomyosin contractile ring in Saccharomyces cerevisiae cytokinesis. *J Cell Biol* 1998. doi:10.1083/jcb.142.5.1301.

82 Pollard TD, Wu JQ. Understanding cytokinesis: Lessons from fission yeast. Nat. Rev. Mol. Cell Biol. 2010. doi:10.1038/nrm2834.

83 Pu K-M, Akamatsu M, Pollard TD. The septation initiation network controls the assembly of nodes containing Cdr2p for cytokinesis in fission yeast. *J Cell Sci* 2015. doi:10.1242/jcs.160077.

84 Vallen EA, Caviston J, Bi E. Roles of Hof1p, Bni1p, Bnr1p, and Myo1p in Cytokinesis in Saccharomyces cerevisiae. *Mol Biol Cell* 2000. doi:10.1091/mbc.11.2.593.

85 Fukui Y, Inoue S. Cell division in Dictyostelium with special emphasis on actomyosin organization in cytokinesis. *Cell Motil Cytoskeleton* 1991; **18**: 41–54.

86 Azimzadeh J. Exploring the evolutionary history of centrosomes. Philos. Trans. R. Soc. B Biol. Sci. 2014. doi:10.1098/rstb.2013.0453.

87 Jaspersen SL, Winey M. THE BUDDING YEAST SPINDLE POLE BODY: Structure, Duplication, and Function. *Annu Rev Cell Dev Biol* 2004. doi:10.1146/annurev.cellbio.20.022003.114106.

88 Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016. doi:10.1038/nbt.3519.

89 Hughes ME, Hogenesch JB, Kornacker K. JTK-CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* 2010; **25**: 372–380.

90 Thaben PF, Westermark PO. Detecting rhythms in time series with rain. *J Biol Rhythms* 2014; **29**: 391–400.

91 Rustici G, Mata J, Kivinen K, Lió P, Penkett CJ, Burns G *et al.* Periodic gene expression program of the fission yeast cell cycle. *Nat Genet* 2004; **36**: 809–817.

92 Bauer S, Grossmann S, Vingron M, Robinson PN. Ontologizer 2.0 - A multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics* 2008. doi:10.1093/bioinformatics/btn250.

93 Popow J, Jurkin J, Schleiffer A, Martinez J. Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors. *Nature* 2014. doi:10.1038/nature13284.

94 Chen M, Gartenberg MR. Coordination of tRNA transcription with export at nuclear pore complexes in budding yeast. *Genes Dev* 2014. doi:10.1101/gad.236729.113.

95 Herrera MC, Chymkowitch P, Robertson JM, Eriksson J, Bøe SO, Alseth I

et al. Cdk1 gates cell cycle-dependent tRNA synthesis by regulating RNA polymerase III activity. *Nucleic Acids Res* 2018. doi:10.1093/nar/gky846.

96 White RJ, Gottlieb TM, Downes CS, Jackson SP. Cell Cycle Regulation of RNA Polymerase III Transcription. 1995 doi:10.1128/MCB.15.12.6653.

97 Schieke SM, McCoy JP, Finkel T. Coordination of mitochondrial bioenergetics with G1phase cell cycle progression. *Cell Cycle* 2008. doi:10.4161/cc.7.12.6067.

98 Obaya AJ, Sedivy JM. Regulation of cyclin-Cdk activity in mammalian cells. Cell. Mol. Life Sci. 2002. doi:10.1007/s00018-002-8410-1.

99 Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 2015. doi:10.1186/s13059-015-0721-2.

100 Santos A, Wernersson R, Jensen LJ. Cyclebase 3.0: A multi-organism database on cell-cycle regulation and phenotypes. *Nucleic Acids Res* 2015. doi:10.1093/nar/gku1092.

101 Fu J, Bian M, Jiang Q, Zhang C. Roles of Aurora Kinases in Mitosis and Tumorigenesis. *Mol Cancer Res* 2007. doi:10.1158/1541-7786.MCR-06-0208.

102 Li J, Dallmayer M, Kirchner T, Musa J, Grünewald TGP. PRC1: Linking Cytokinesis, Chromosomal Instability, and Cancer Evolution. Trends in Cancer. 2018. doi:10.1016/j.trecan.2017.11.002.

103 Li M, Zhang P. The function of APC/CCdh1 in cell cycle and beyond. *Cell Div* 2009. doi:10.1186/1747-1028-4-2.

104 Wang R, Burton JL, Solomon MJ. Transcriptional and post-transcriptional regulation of Cdc20 during the spindle assembly checkpoint in S. cerevisiae. *Cell Signal* 2017. doi:10.1016/j.cellsig.2017.02.003.

105 Csurös M. Count: Evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* 2010.

doi:10.1093/bioinformatics/btq315.

106 Forth S, Kapoor TM. The mechanics of microtubule networks in cell division. J. Cell Biol. 2017. doi:10.1083/jcb.201612064.

107 Wu J, Akhmanova A. Microtubule-Organizing Centers. *Annu Rev Cell Dev Biol* 2017. doi:10.1146/annurev-cellbio-100616-060615.

108 Mahoney NM, Goshima G, Douglass AD, Vale RD. Making microtubules and mitotic spindles in cells without functional centrosomes. *Curr Biol* 2006; **16**: 564–569.

109 Santamaría D, Barrière C, Cerqueira A, Hunt S, Tardy C, Newton K *et al.* Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* 2007. doi:10.1038/nature06046.

110 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012. doi:10.1038/nmeth.2019.

111 The R Core Team. *R: A language and environment for statistical computing.* 2018 doi:10.1038/sj.hdy.6800737.

112 Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T *et al.* Package 'gplots': Various R programming tools for plotting data. *R Packag version 2170* 2016. doi:10.1038/nnano.2010.132.

113 Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R *et al.* UniProt: The universal protein knowledgebase. *Nucleic Acids Res* 2017. doi:10.1093/nar/gkw1099.

114 Vandepoele K. Genome-Wide Analysis of Core Cell Cycle Genes in Arabidopsis. *PLANT CELL ONLINE* 2002. doi:10.1105/tpc.010445.

115 Wang G. Genome-Wide Analysis of the Cyclin Family in Arabidopsis and Comparative Phylogenetic Analysis of Plant Cyclin-Like Proteins. *PLANT Physiol* 2004. doi:10.1104/pp.104.040436.

116 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990. doi:10.1016/S0022-2836(05)80360-2.

117 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K *et al.* BLAST+: Architecture and applications. *BMC Bioinformatics* 2009. doi:10.1186/1471-2105-10-421.

118 Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002. doi:10.1093/nar/gkf436.

119 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009. doi:10.1093/bioinformatics/btp348.

120 Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015. doi:10.1093/molbev/msu300.

121 Minh BQ, Nguyen MAT, Von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 2013. doi:10.1093/molbev/mst024.

122 Denbo S, Aono K, Kai T, Yagasaki R, Ruiz-Trillo I, Suga H. Revision of the Capsaspora genome using read mating information adjusts the view on premetazoan genome. *Dev Growth Differ* 2019. doi:10.1111/dgd.12587.

123 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013. doi:10.1186/gb-2013-14-4-r36.

124 Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L *et al.* Using tablet for visual exploration of second-generation sequencing data. *Brief Bioinform* 2013. doi:10.1093/bib/bbs012.

125 Bhattacharya D, Price DC, Xin Chan C, Qiu H, Rose N, Ball S *et al.* Genome of the red alga Porphyridium purpureum. *Nat Commun* 2013. doi:10.1038/ncomms2931.

126 Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I *et al.* The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res* 2014. doi:10.1093/nar/gkt1069.

127 Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J *et al.* Ensembl 2018. *Nucleic Acids Res* 2018. doi:10.1093/nar/gkx1098.

128 Huerta-Cepas J, Capella-Gutiérrez S, Pryszcz LP, Marcet-Houben M, Gabaldón T. PhylomeDB v4: Zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Res* 2014. doi:10.1093/nar/gkt1177.

129 Smedley D, Haider S, Durinck S, Pandini L, Provero P, Allen J *et al.* The BioMart community portal: An innovative alternative to large, centralized data repositories. *Nucleic Acids Res* 2015. doi:10.1093/nar/gkv350.

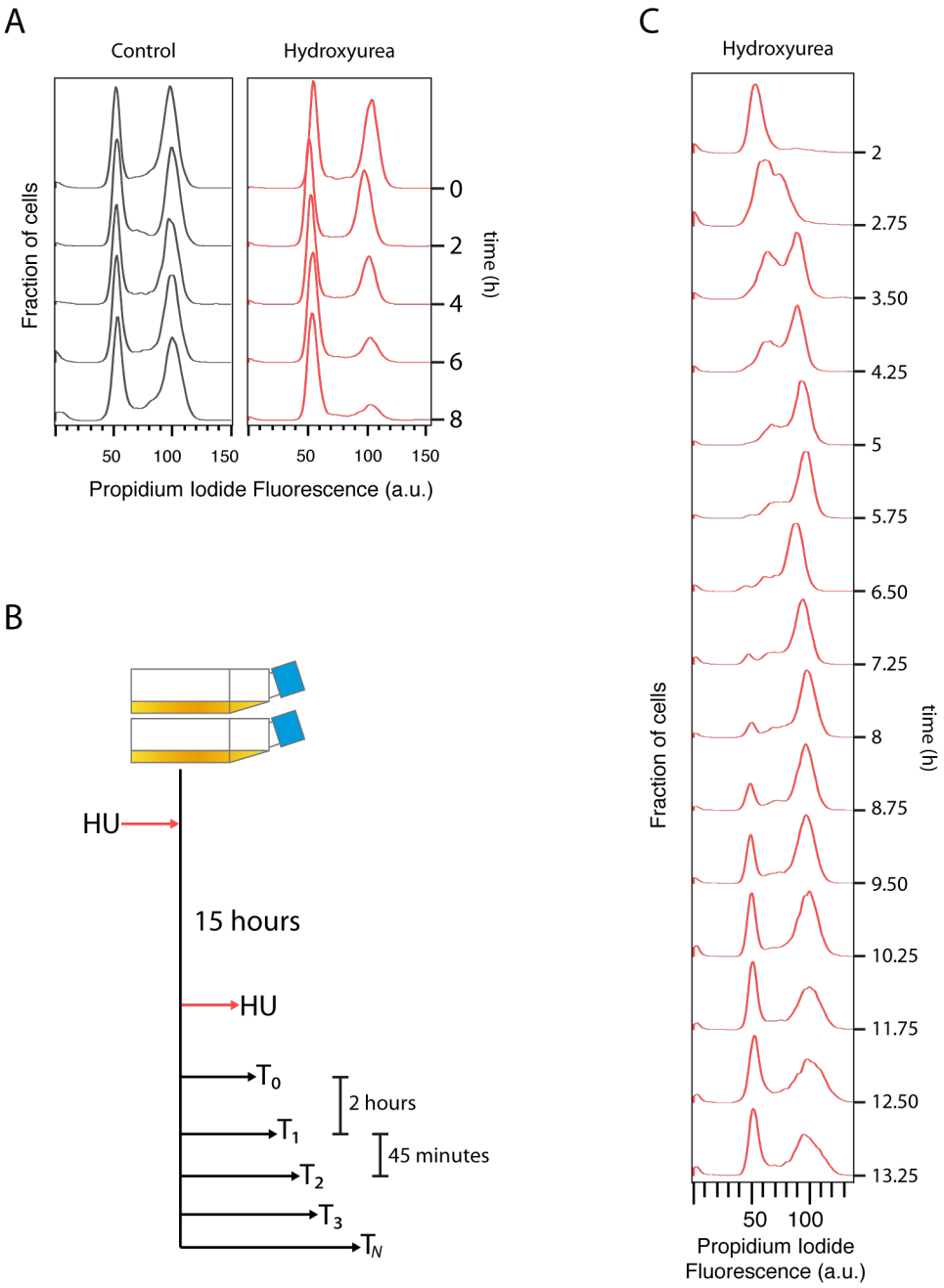


Fig. 1: Experimental setup for cell cycle synchronization in *Capsaspora*. A: Effect of Hydroxyurea in *Capsaspora* cells. Control cells were treated with an equal volume of distilled water. HU treated cells were sampled every 2 hours, and a major decrease in 2C cells is achieved after a minimum of 8 hours. **B:** Experimental layout used in this study. Two cell cultures growing independently for several generations were grown in fresh medium and kept in HU for a lapse of fifteen hours. After that, cells were released from HU and harvested every 45 minutes for 11 hours. Three different samples were taken from each culture at each time point. **C:** DNA content profile of synchronized *Capsaspora* cultures at representative time points of the whole experiment.

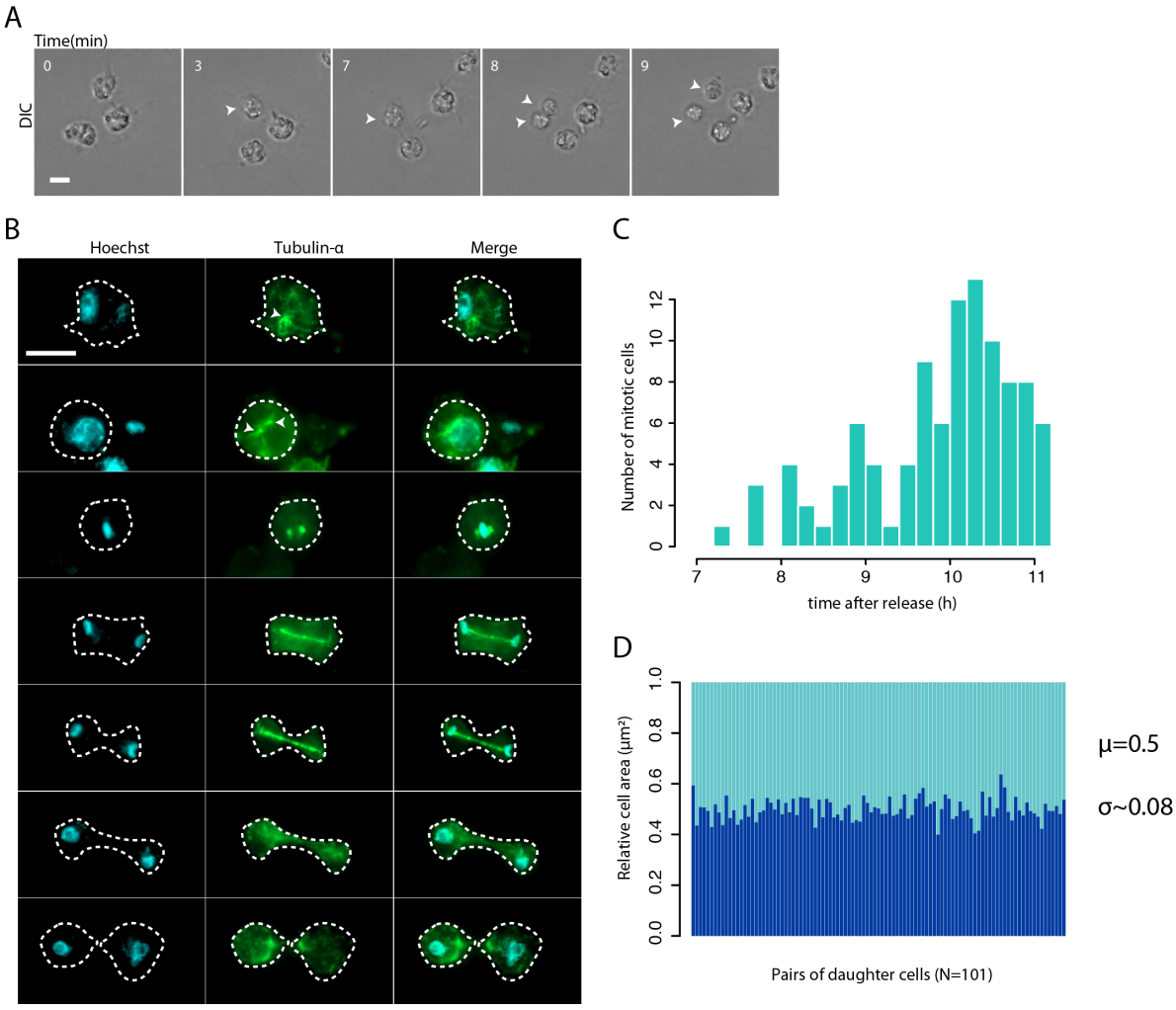


Fig. 2: Cell division in *Capsaspora*. A: Time lapse of live imaging of a synchronous culture. Numbers indicate minutes since round up. White arrows indicate a cell dividing during the time lapse. Scale bar: 5µm. **B:** Fluorescence immunostaining of DNA (cyan) and Tubulin alpha (green) in *Capsaspora* synchronous cultures at different stages of cell division. White arrows indicate structures with a high concentration of tubulin. White dashed outline indicates cell perimeter. Scale bar: 5µm. **C:** Histogram depicting the number of cells at the moment of division in different times of the time lapse. Range goes from 7h to 11h. **D:** Stacked barplot showing the normalized, relative cell area for each daughter cell in a total of 101 events of cell division.

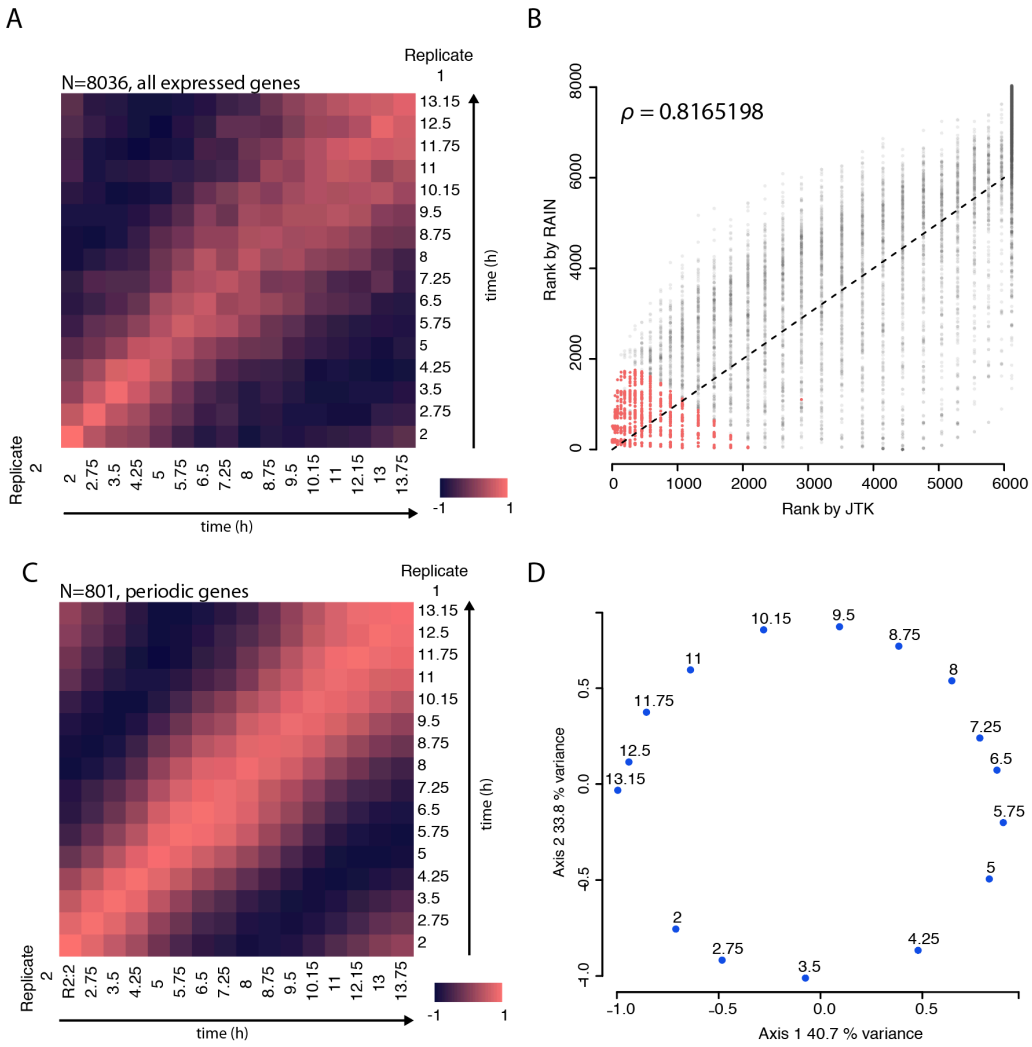


Fig. 3: Detection of periodic genes in *Capsaspora*. A: Pearson correlation for the normalized expression level of every gene between replicates. Bright red indicates positive correlation, and dark purple indicates negative correlation. **B:** Spearman correlation between the two methods used to detect cell cycle regulated genes in *Capsaspora*. Scatter plots depicting the rank assigned for every gene by the JTK_CYCLE and RAIN on an average dataset of the two time-series replicates. These two algorithms rely on different approaches to finally assign, for every gene, a p-value interpreted as the probability that it can be considered periodic. For its proper functioning, we set the two algorithms to look for periodic behavior in a lapse of 11 to 13 hours, with time lapses of 0.75 hours. We assigned two ranks to every gene according to the p-values calculated by JTK_CYCLE and RAIN. Each dot represents a gene expressed in the time series. Colored dots represent the 801 genes that were finally taken as periodic according to our criteria. **C:** Pearson correlation for the top 801 cell cycle regulated genes, also based on normalized gene expression level. See Fig. 4 for more details. **D:**

1140 Principal coordinate analysis on the set of 801 cell cycle regulated genes based on
1141 normalized gene expression level. Every dot represents a time point in the experiment.

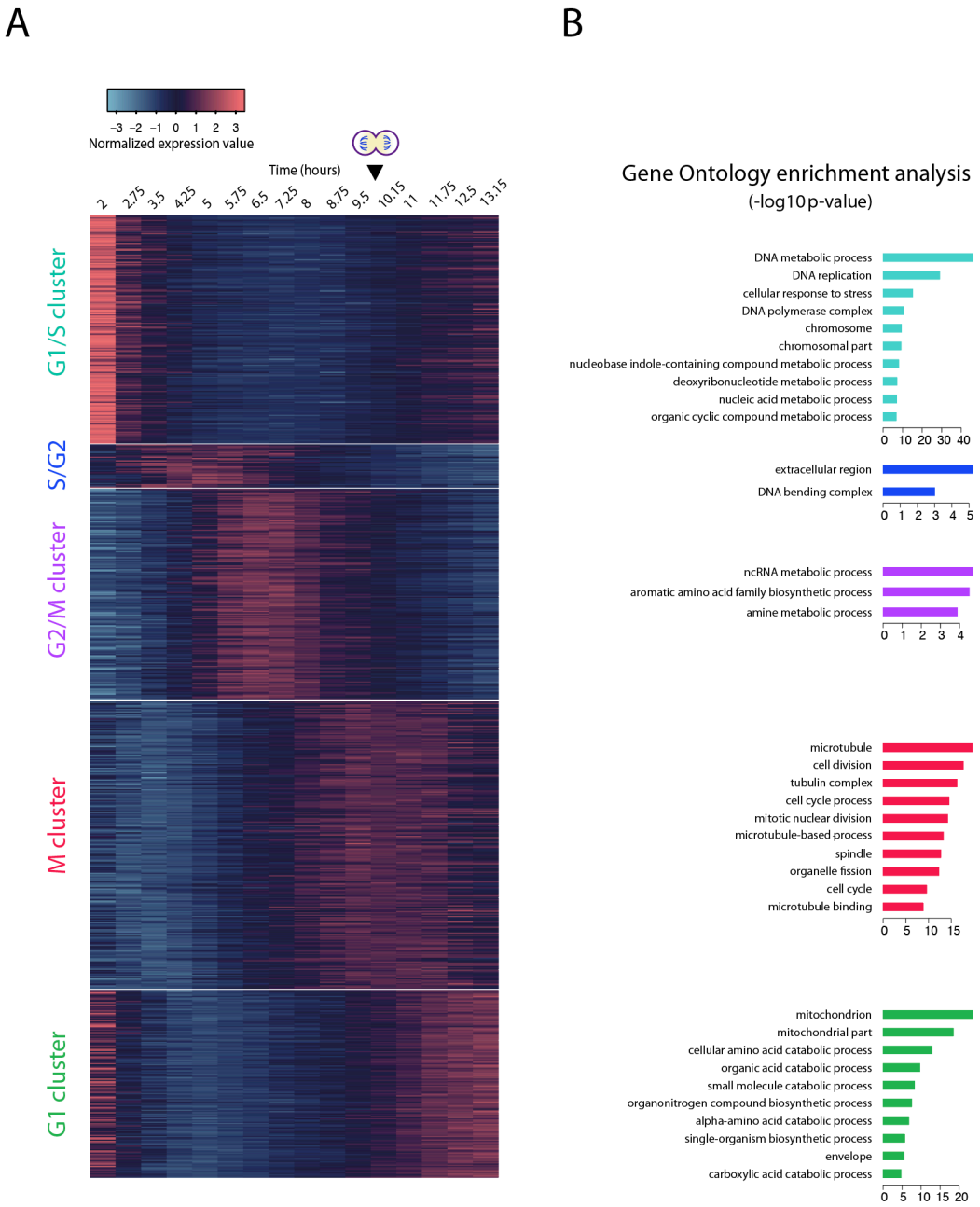


Fig. 4: The periodic transcriptional program of *Capsaspora*. A: heatmap of gene expression level depicting seven main clusters detected by Euclidean distance hierarchical clustering. Clusters were rearranged to visually represent their expression peaks over time. Black arrow and dividing cell indicate time of cell division (see Fig. 2). **B:** Top ten enriched gene ontology terms for every cluster shown in Fig. 4A. We considered an enrichment as significant when Bonferroni corrected p-value was lower than 0.05. For the full list of enriched GO terms, see Supplementary Table 4.1

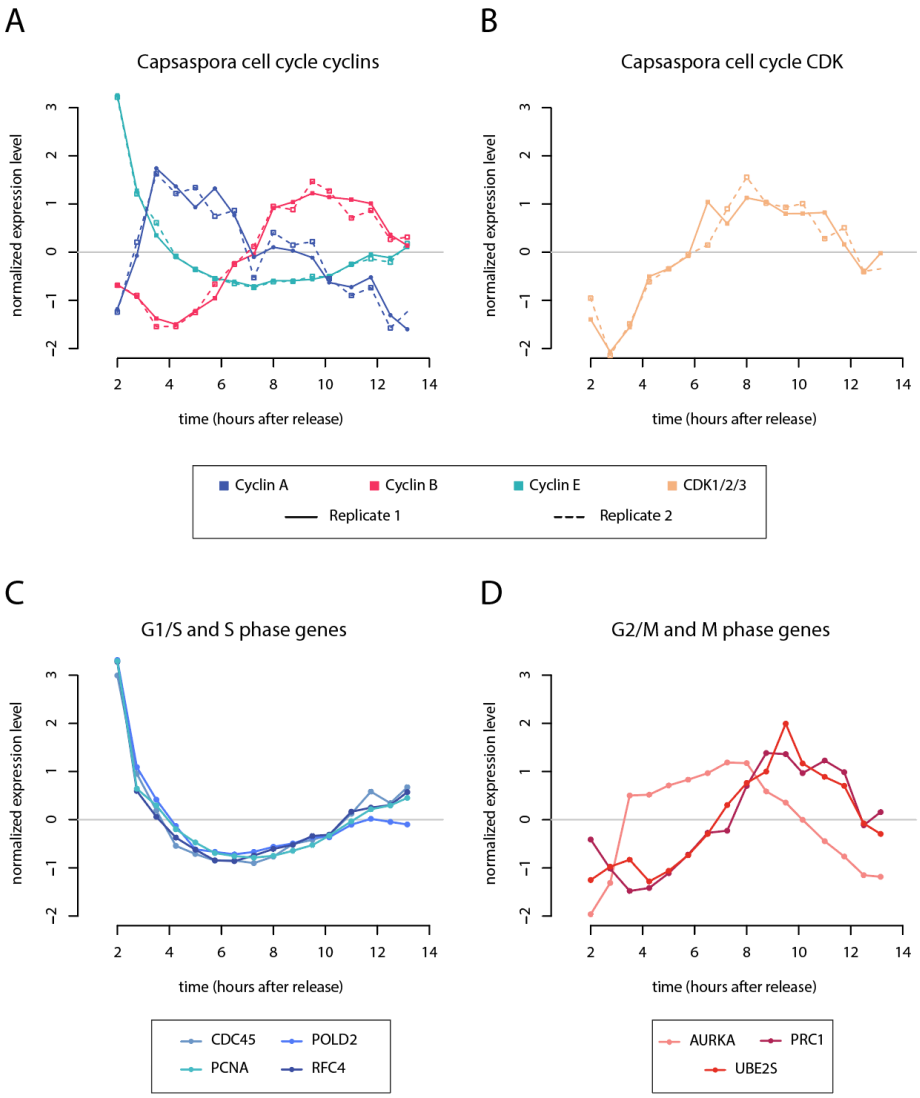


Fig. 5: Dynamics of the cyclin-CDK system and other cell cycle regulators in in *Capsaspora*. A: Normalized gene expression profile of the Cyclin A, B, and E genes found in *Capsaspora*. Normal and dashed lines indicate replicates 1 and 2, respectively. **B:** Normalized gene expression profile of the *Capsaspora* orthologue of CDK1/2/3 found by phylogenetic analyses (see Supplementary Fig. S7). Normal and dashed lines indicate replicates 1 and 2, respectively. **C:** Normalized gene expression level of *Capsaspora* orthologues of several G1/S regulators in animals. **D**: Normalized gene expression level of *Capsaspora* orthologues of several G2/M regulators in animals. Genes in C and D as described by [32,39,40,100]. A full list is depicted in Supplementary File 7.

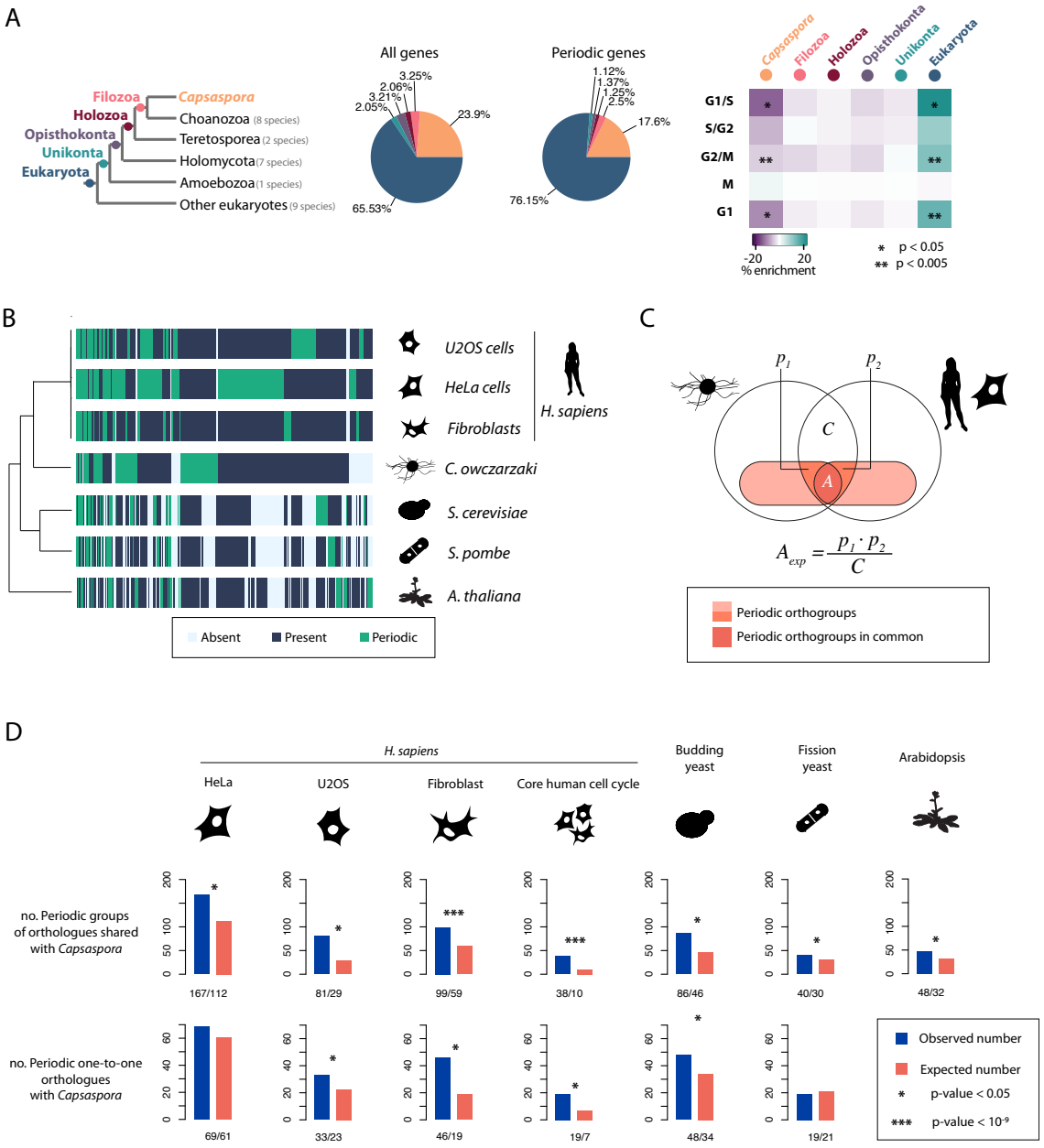


Fig. 6: The *Capsaspora* periodic transcriptional program has more resemblances to that of animal cells. (Left) Tree depicting gene ages considered in this study. (Center) Gene age stratification profiles of the *Capsaspora* genome and the periodic transcriptional program. (Right) Gene age enrichment/depletion on each cluster of periodic genes in *Capsaspora* (see Fig. 4) compared to non-periodic genes. **B:** Heatmap showing periodic orthogroups in common in the tree of species used in the comparative. **C:** Venn diagrams indicating how the binomial tests were calculated for each pair-wise comparison of species, e.g. *Capsaspora* and *Homo sapiens* HeLa cells. White circles indicate orthogroups from each species. The intersection between these represents the orthogroups in common, *C*. *p1* and *p2* areas represent periodic genes of each species within *C*. Null expectation probability (A_{exp}) is calculated as the product of *p1* and *p2*

1173 divided by *C*. P-values of all the binomial tests are provided in Supplementary File 9 and
1174 Supplementary File 11. **D**: Bar plots indicating the amount of shared periodic
1175 orthogroups and/or periodic one-to-one orthologues between pairs of cell types or
1176 species. P-values of all the binomial tests are provided in Supplementary File 9. **E:** Gene
1177 age enrichment/depletion of *Capsaspora* periodic genes with a periodic co-orthologue in
1178 *H. sapiens* HeLa cells. Comparison was done against the rest of the *Capsaspora* periodic
1179 transcriptional program. Color code for bar plots as in Fig.6A.

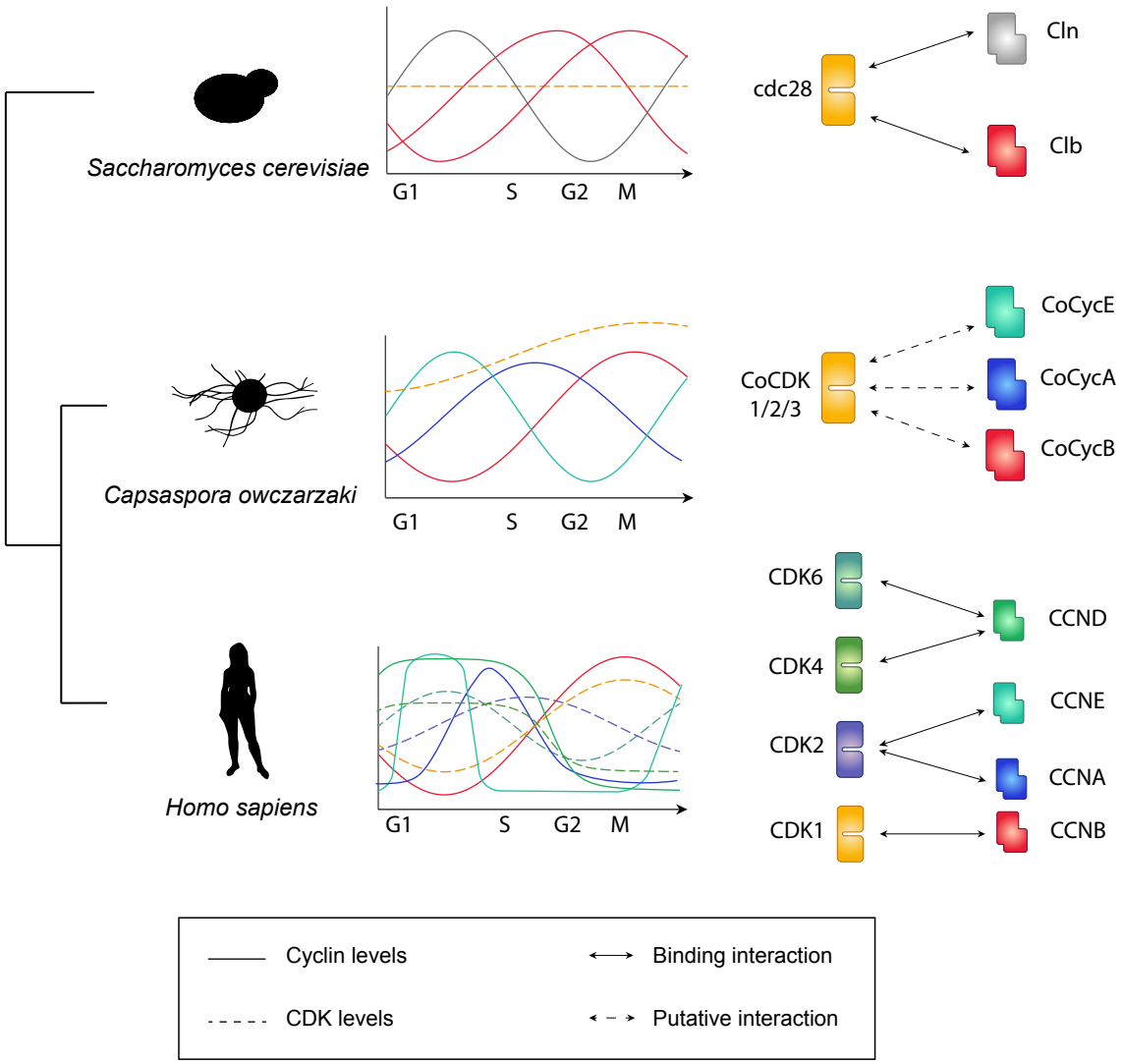
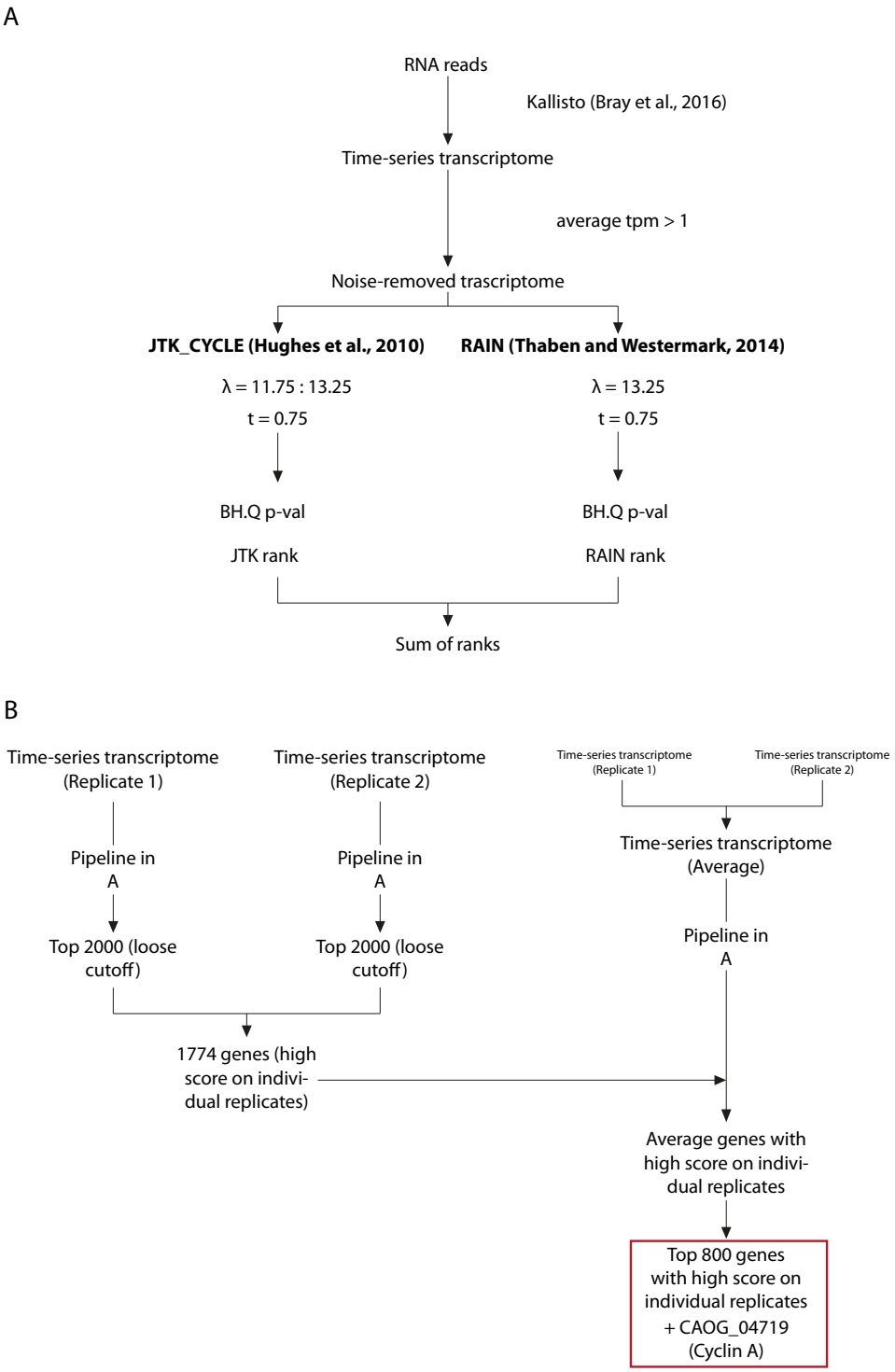
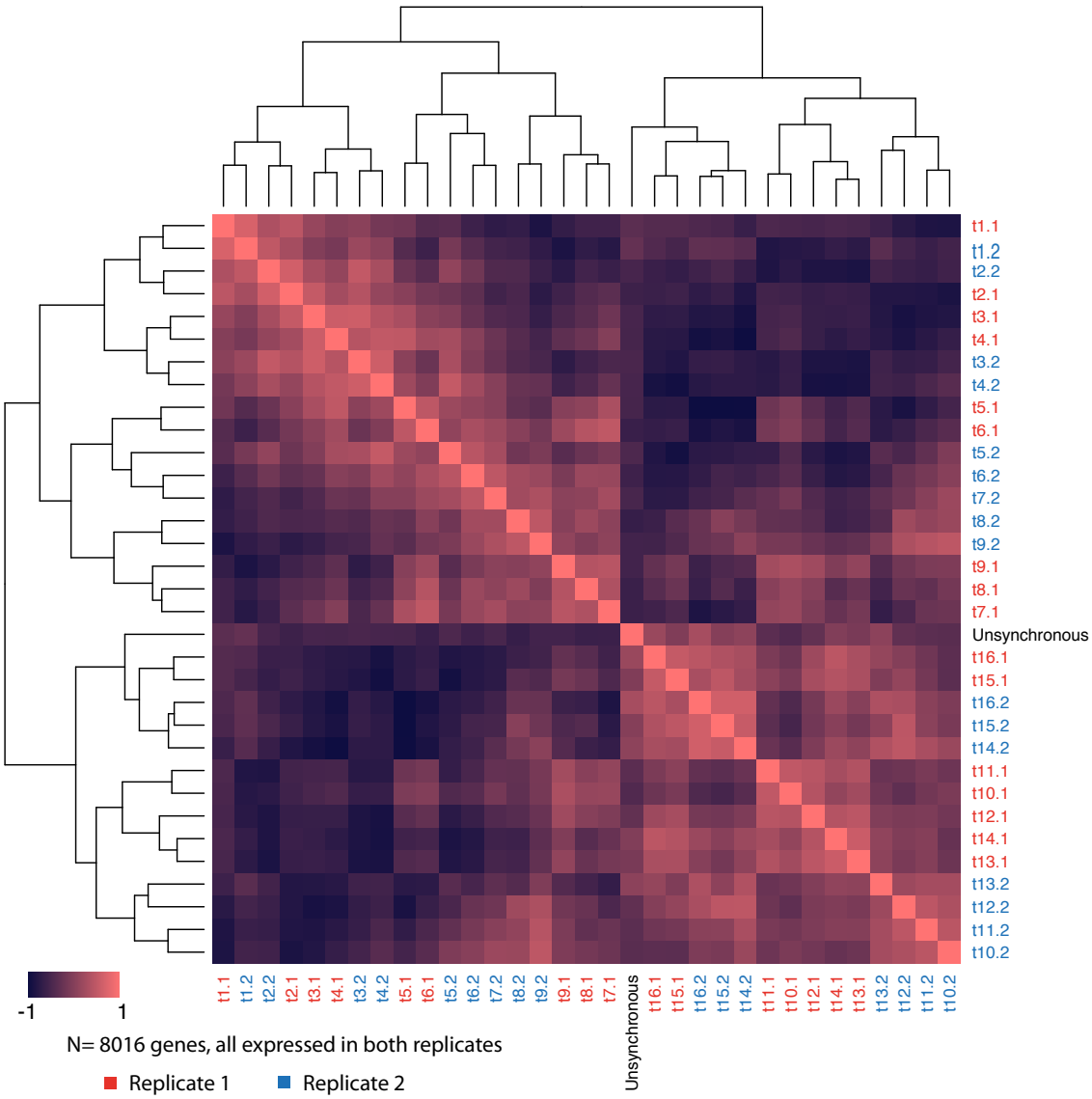


Fig. 7: a model of the dynamics of the Cyclin/CDK system in *Capsaspora*. Bold lines indicate cyclin levels and dashed lines indicate CDK levels across the cell cycle. In *Capsaspora*, several cell cycle cyclins may be binding to a single CDK that is not transcriptionally stable, an intermediate system that falls between yeast and human cyclin/CDKs. This suggests the expansions of cell cycle cyclins predate the expansion of cell cycle CDKs, which later acquired specific binding to cell cycle cyclins in animals.

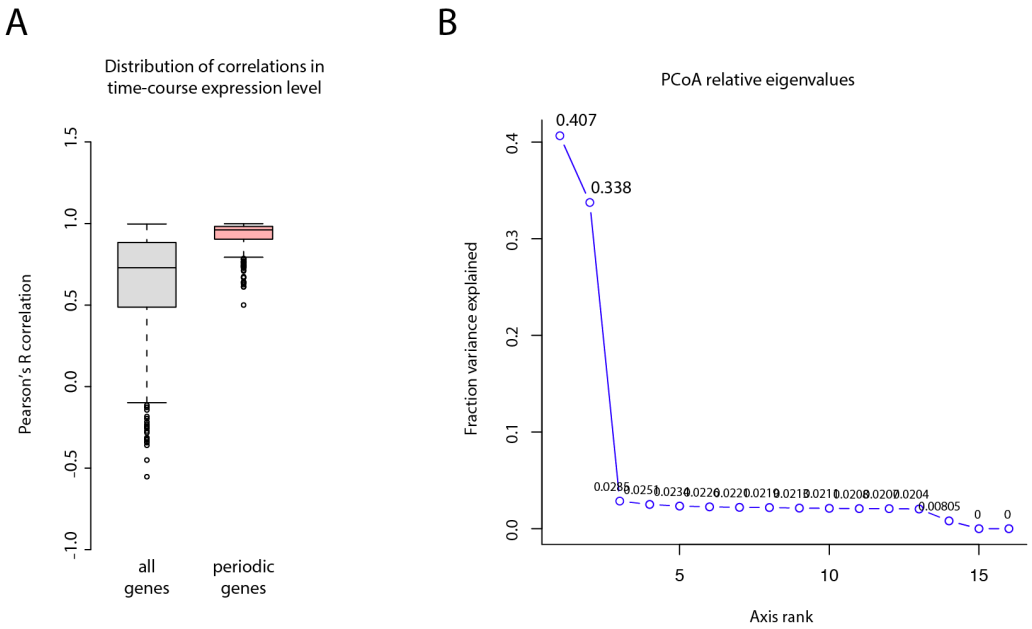


Supplementary Fig. S1: Computational pipeline used to detect periodic genes in *Capsaspora*. **A:** Pipeline for ranking the transcripts on each experiment. RNA reads were processed using Kallisto, noise transcripts were filtered out, and JTK and RAIN were run in parallel with the indicated setup. Bonferroni-corrected p-values were ranked, and the sum of ranks was used as a final rank. **B:** Pipeline used to detect periodic transcripts in *Capsaspora*. Samples were treated separately to detect periodic transcripts with a loose

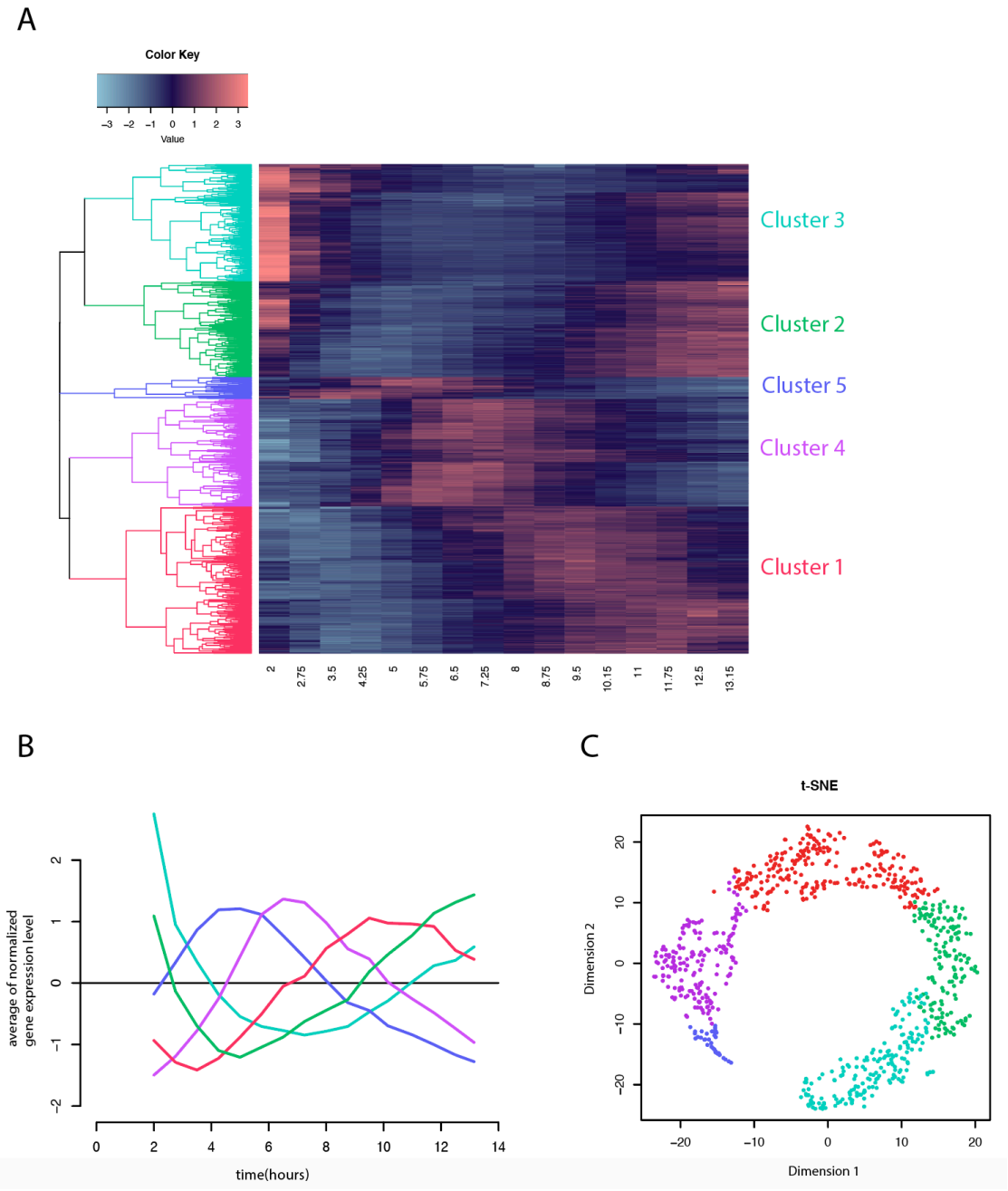
1196 cutoff. This gene set was used to filter periodic genes ranked from an average dataset
1197 of the two replicates, out of which the top 800 (10% of the number of genes in
1198 *Capsaspora*) were selected.



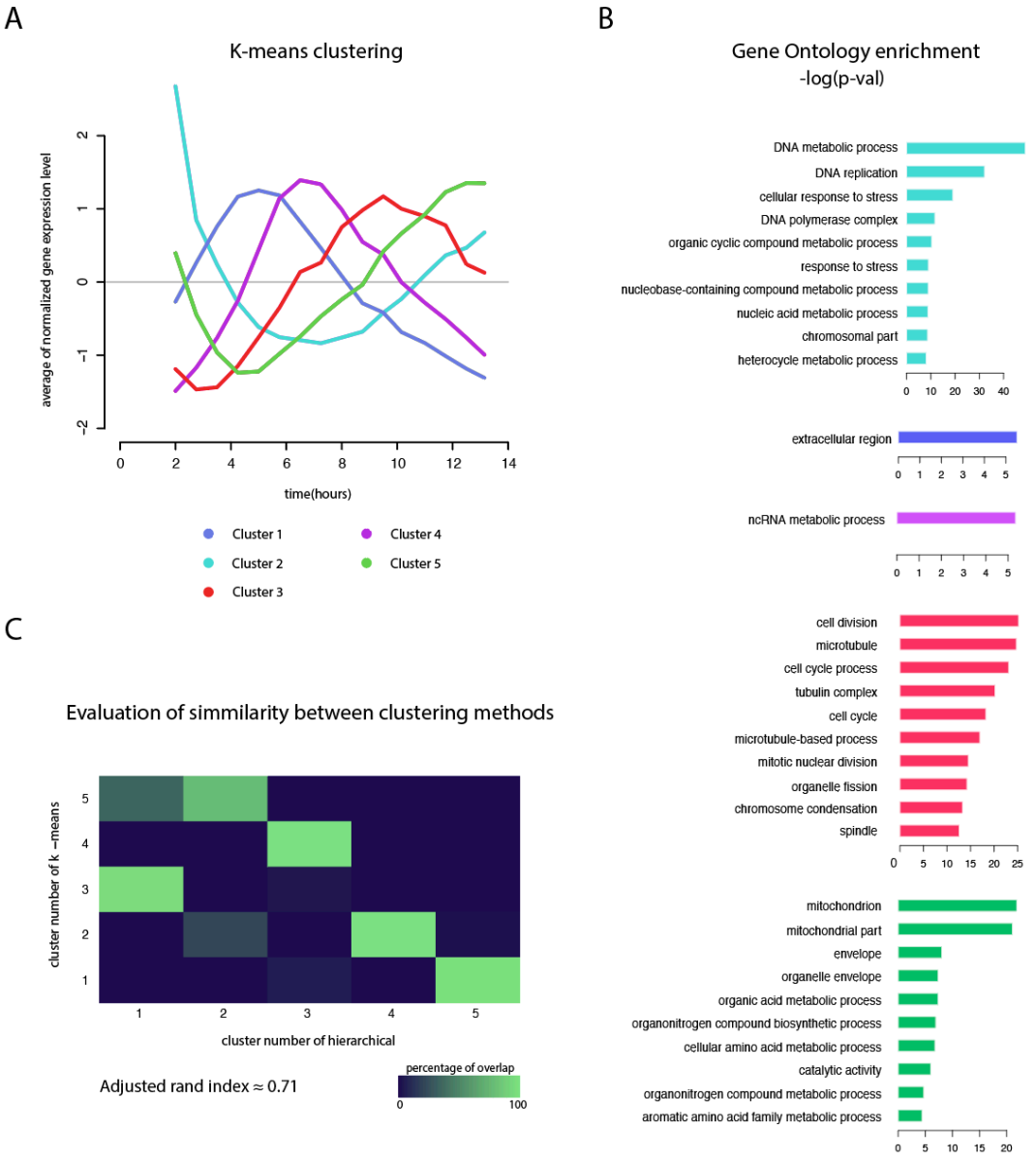
Supplementary Fig. S2: clustering of time points of each replicate and the unsynchronized culture sampled as a control, based on a dissimilarity matrix of Pearson correlation. All the genes with an average expression level above 1 tpm in both replicates were used.



Supplementary Fig. S3: **A:** Distributions of Pearson correlation between replicates of a set of randomly chosen 801 genes and the 801 genes defined as periodic. **B:** Fraction of variance explained by the different relative eigenvalues of the principal coordinate analysis (see Fig. 3D).

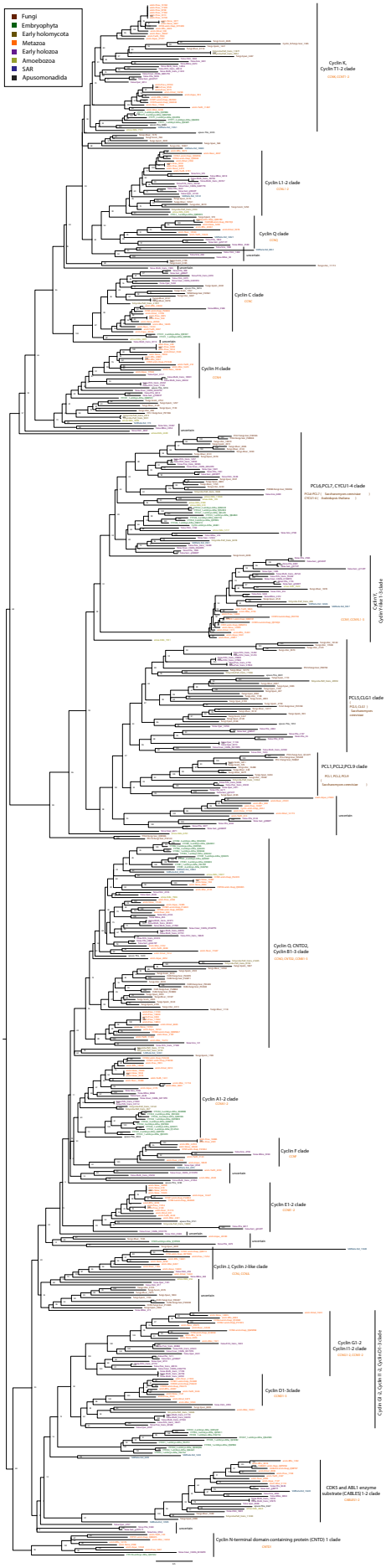


Supplementary Fig. S4: A: Hierarchical clustering of the expression profiles of the 801 periodic genes found in *Capsaspora*, and the color equivalences with the final clusters shown at Fig. 4A. **B:** Average expression level of *Capsaspora* periodic genes grouped by hierarchical clustering. **C:** t-SNE plot of all 801 genes in the periodic transcriptional program of *Capsaspora*, showing circle pattern as in Fig. 3D. Color code in both Fig.s follows the color code in Fig. 4.



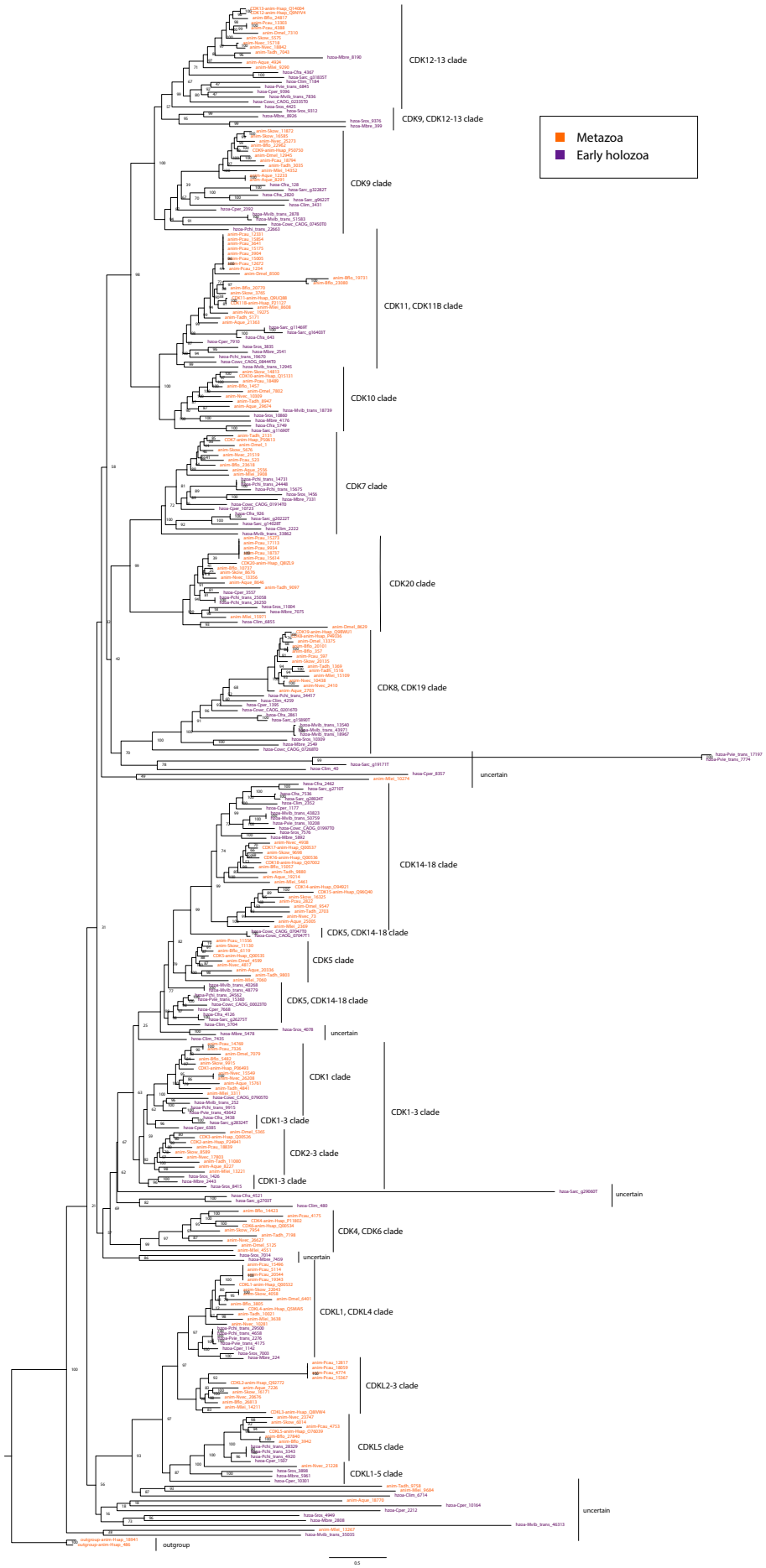
Supplementary Fig. S5: K-means clustering of the periodic transcriptional program of *Capsaspora*. **A:** Average expression level of *Capsaspora* periodic genes grouped by k-means clustering. **B:** Top ten enriched GO terms of each cluster of periodic genes generated by k-means clustering. GO terms were considered significant when Bonferroni-corrected p-value was lower than 0.05. Full list available at Supplementary Fig. S3. **C:** Agreement between clustering methods. Heatmap showing the percentage of overlap between clusters by two methods. Overlap is calculated as the number of genes belonging to the same pair of clusters divided by the size of the smallest cluster in the pair.

1228 **Fig. 4—source data 1**: List of significant (Bonferroni p-value < 0.05) GeneOntology
1229 enrichments of each hierarchical cluster of periodic genes in *Capsaspora*. Available on
1230 Figshare: https://figshare.com/s/4d642c9854efe6d879a7



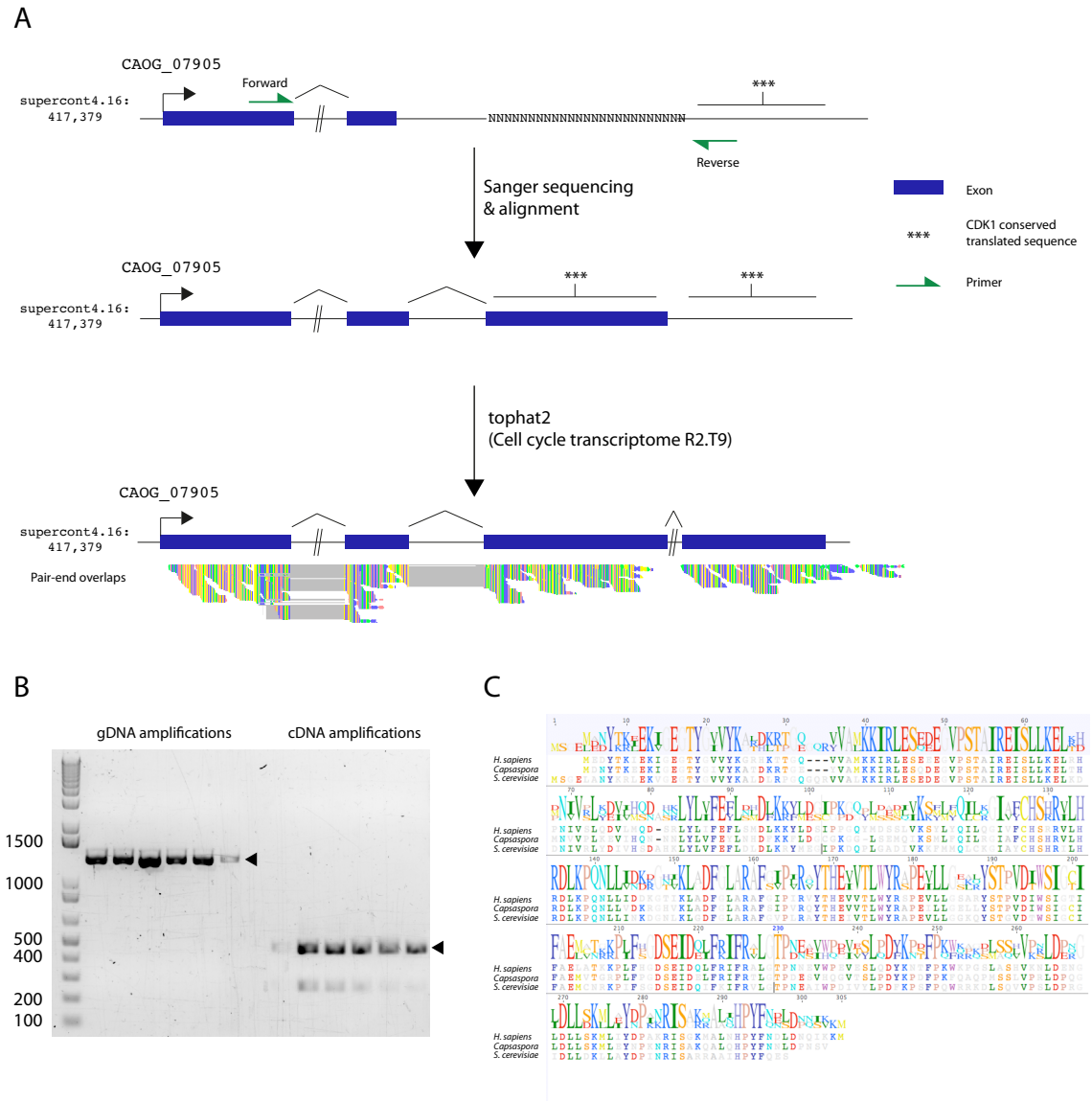
1231

Supplementary Fig. S6: Unrooted maximum likelihood phylogenetic tree (IQ-TREE) inferred from cyclins sequences of 30 eukaryotic species (see methods). Nodal support values (1000- bootstrap replicates by UFBoot) are shown in all nodes. Eukaryotic sequence names are abbreviated with the four-letter code (see methods) and colored according to their major taxonomic group (see panel).

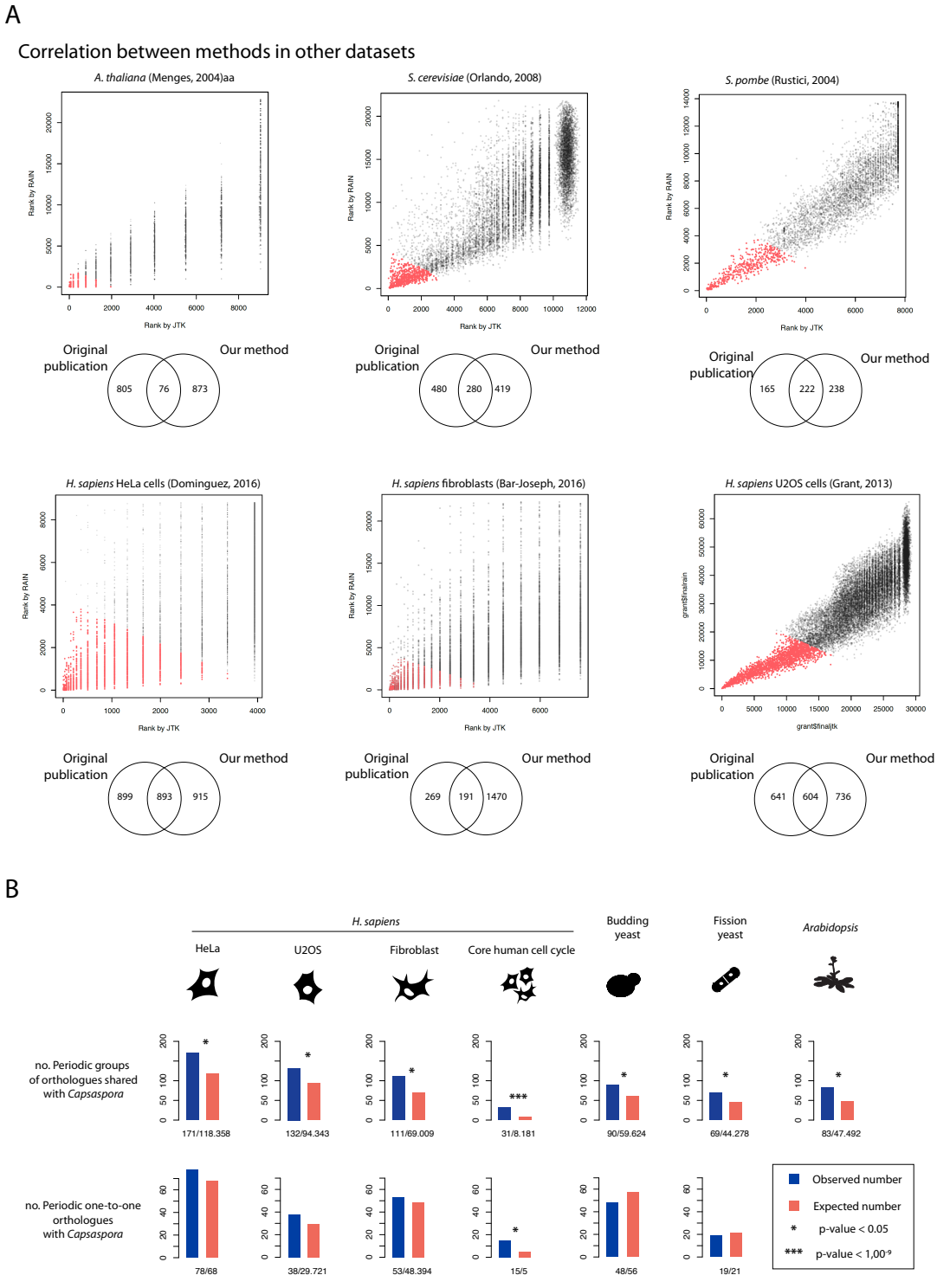


1237

Supplementary Fig. S7: Maximum likelihood phylogenetic tree (IQ-TREE) inferred from CDK sequences of early-branching holozoan species and animals (see methods). Statistical support values (1000-replicates UFBoot) are shown in all nodes. Eukaryotic sequence names are abbreviated with the four-letter code (see methods) and colored according to their major taxonomic group (see panel).



Supplementary Fig. S8. A: Schematic representation of the genomic locus of *Capsaspora* CDK1-3 gene showing exons, splicing sites, non-annotated regions of predicted sequence, and mapping of mRNA reads. B: PCR amplifications of *Capsaspora* CDK1 using primers detailed in Methods and A, using genomic DNA and cDNA as templates. Arrows indicate size of the products sent for sequencing. C: Alignment of H. *sapiens* and S. *cerevisiae* CDK1 genes, and the *Capsaspora* updated CDK1-3 sequence, using Geneious v8.1.9.



Supplementary Fig. S9: A Reanalysis of previous cell cycle datasets in model organisms. **A:** Scatter plots of ranks by JTK and RAIN for each dataset of each species used in the comparative analysis (see Fig. 6.D and Results section). Datasets were processed as indicated in Supplementary File 10, and Material and methods. Depending on the dataset, we could recover between one third and more than half of the originally described periodic genes, except *Arabidopsis* where the agreement was very low. **B:** Bar plots indicating the amount of shared periodic orthogroups and/or periodic one-to-

1259 one orthologues between pairs of cell types or species, using our own lists of periodic
1260 genes. P-values of all the binomial tests are provided in Supplementary File 9.

Video 1: Synchronized cells of Capsaspora undergoing cell division. Time interval between frames is 1 minute. The movie is played at 7fps. Scale bar = 5µm. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 1: Tables of transcript per million of each replicate. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 2: List of periodic genes in *Capsaspora*, containing information for each gene about cluster membership, gene age, and orthologs in other species. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 3: Gene Ontology enrichments for all the clusters in the periodic transcriptional program of *Capsaspora*. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 4: FASTA formatted sequences of cyclins and CDKs used in the phylogenetic analyses (see Supplementary Fig. S6 and Supplementary Fig. S7). Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 5: List of species used in the phylogenetic analyses and in the generation of groups of orthologues. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 6: FASTA formatted sequences of newly annotated *Capsaspora* CDK1-2-3 CDS and protein translation. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 7: List of cell cycle regulators in humans (described in [32,39,40,100]), and their respective orthologs in *Capsaspora* defined by OrthoFinder and/or phylome data (see Results and Methods). Bold indicates genes that have been plotted in Fig. 5C or 5D. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 8: List of cell cycle regulators in humans (described in [32,39,40,100]), that also have at least one periodic co-ortholog in *Capsaspora*, defined by OrthoFinder (see Results and Methods). Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 9: Metrics of gene age enrichment and depletion for the gene clusters of the periodic transcriptional program of *Capsaspora*, and their corresponding Fisher Test p-values. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 10: Procedure used to retrieve identifiers for the datasets of H. sapiens, S. cerevisiae, S. pombe, and A. thaliana, and parameters used to set up JTK_CYCLE and RAIN in the reanalysis. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 11: Metrics of shared periodic orthogroups (OG) and one-to-one orthologues between *Capsaspora* and the rest of cell types and species. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7

Supplementary File 12: Metrics of shared periodic orthogroups (OG) for all comparisons between pairs of species, using periodic genes from the literature and using our own sets of periodic genes. Available on Figshare: https://figshare.com/s/4d642c9854efe6d879a7