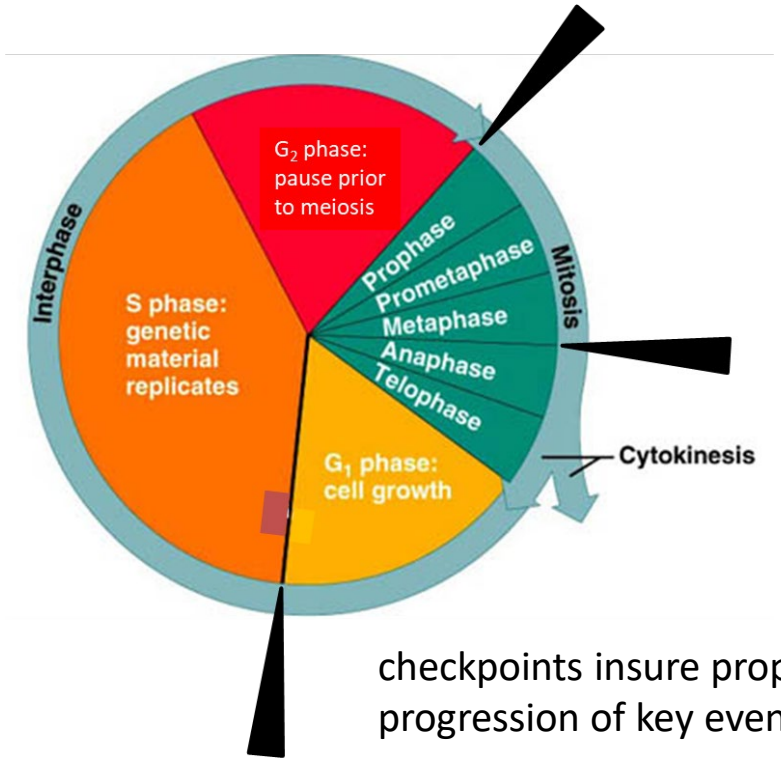
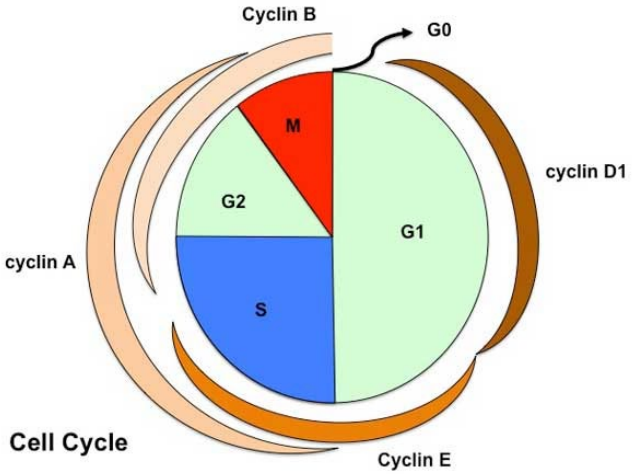
Key Events in the Life Histories of Cells

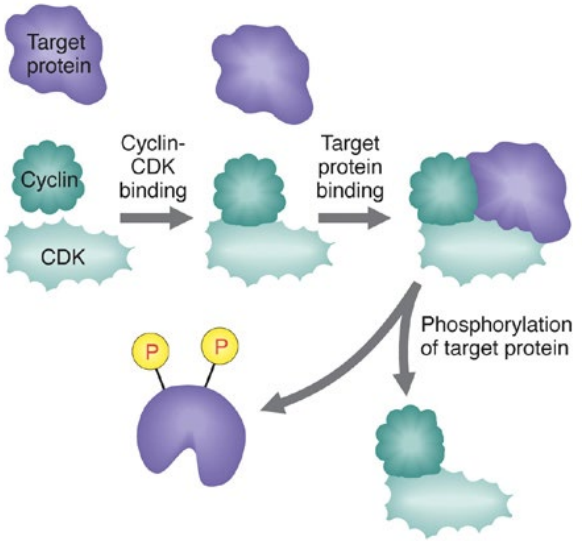
- The cell cycle and mitosis: deciding when to proceed through various steps of cell doubling.
- Alternative genome partitioning mechanisms in eukaryotes: mitosis and meiosis.
- Sex: mating-type recognition and cell fusion.
- Sexual reproductive systems: mating-type number, isogamy vs. anisogamy, and sex chromosomes.
- Senescence.

The Eukaryotic Cell Cycle is Driven by Oscillating Abundances of Cyclin Proteins and Cyclin-Dependent Protein Kinases (CDKs)



checkpoints insure proper progression of key events



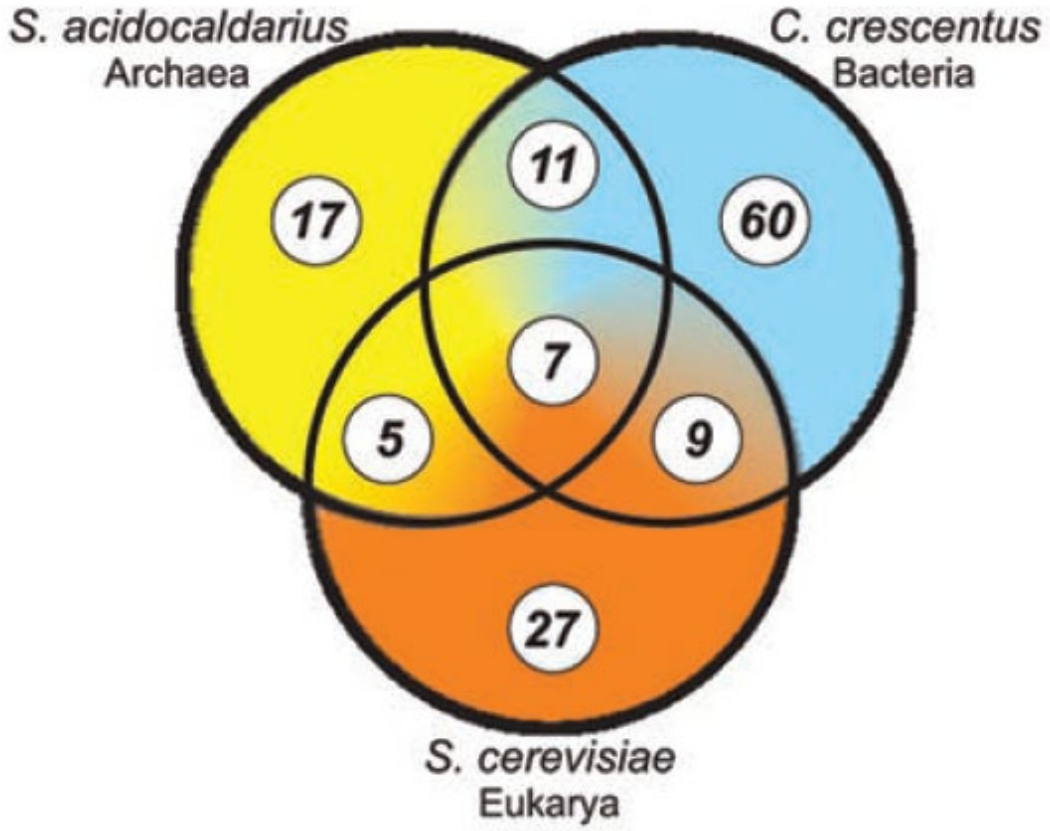


- The cell life cycle is divided into “stages” based on the state of cell growth and genome replication, although the relative durations of stages varies greatly among phylogenetic lineages.
- Oscillations in cyclin gene expression and protein decay drive the activation of specific CDKs, which in turn are involved in phosphorylation of other proteins associated with the cell cycle.
- Despite the centrality of the “cell cycle”, there is no “text-book”' molecular description of the processes.

Evolutionary Variation in the Cell Cycle: passive increases in network complexity in the absence of any known selective advantages of alterations.

- Development of molecular complexity often proceeds through gene duplication and subfunctionalization – a substantial amount of such duplication occurred in the FECA to LECA transition.
- Increased network complexity in multicellular species.
- Nonorthologous gene replacement – alterations in the underlying participants on a constant pathway.
- Redundancy of regulatory mechanisms.
- Multimeric protein complexes often become more complex relative to those in archaea – transitions from homomers to heteromers.

Only a Small Fraction of Cell-Cycle-Dependent Genes are Conserved Across the Major Domains



Bernander et al. (2010, Cell Cycle)

Evolutionary Relationships Among the Cyclins: dramatic expansions in some, especially multicellular, lineages.

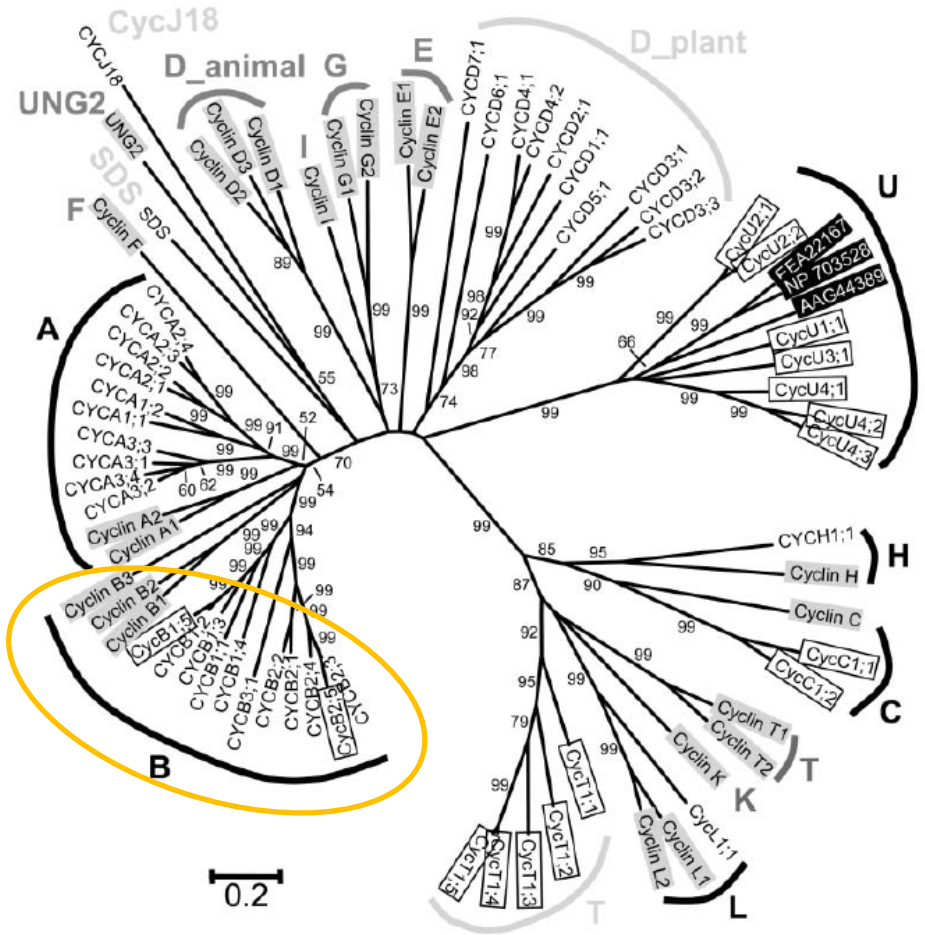


Figure 1. Unrooted NJ tree of the Arabidopsis and human cyclins, with bootstrap values higher than 50% shown for each clade. Ten and 13 families of cyclins are recognizable in Arabidopsis and human, respectively, five of which (A, B, C, H, and L) are shared by both species. The remaining five types of Arabidopsis cyclins are named as CycJ18-, D_plant-, T-, SDS-, and L-type, respectively. Twenty-two proteins from human are highlighted by gray boxes. Three proteins from protists are in inverted boxes: NP_703528, EAA22167, and AAG44389 from *Plasmodium falciparum*, *P. yoelii yoelii*, and *Trypanosoma cruzi*, respectively. Eighteen newly named Arabidopsis cyclins are in open boxes.

- Animals (grey boxes) use four major cyclin types; yeasts just one (B).

- Plants (no boxes and white boxes) share many cyclins with animals, but both groups also have unique families.

Wang et al. (2004, Plant Physiol.)

- 26 cyclin genes are known in the ciliate *Tetrahymena*, most of which exhibit just one peak during the complex meiotic cell cycle.

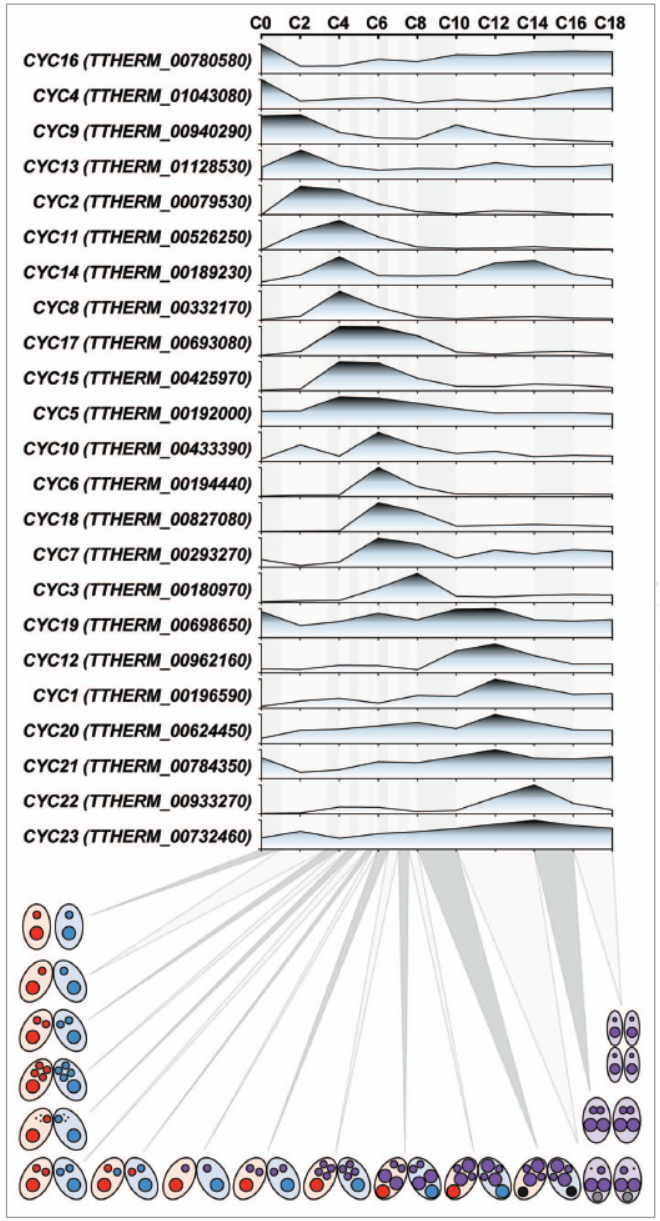


Figure 1. Normalized microarray expression values for *T. thermophila* cyclin genes during the 18 hour conjugation cycle (C0-C18). Raw data collected from Tetrahymena Gene Expression Database (tged.ihb.ac.cn/).[2] Values relative to the peak of expression during conjugation are shown. Shading increases linearly from 0% expression (white) to 100% expression (black). Three cyclin genes were omitted because their expression levels were extremely low during conjugation (*CYC24*/TTHERM_00842480, *CYC25*/TTHERM_00717540), or because they showed higher expression during vegetative growth or starvation (*CYC26*/TTHERM_00066840). Panel of conjugation events adapted from Miao et al.

Nearly Constant Cell-Cycle Topology in the Face of Dramatic Repatterning of the Underlying Components

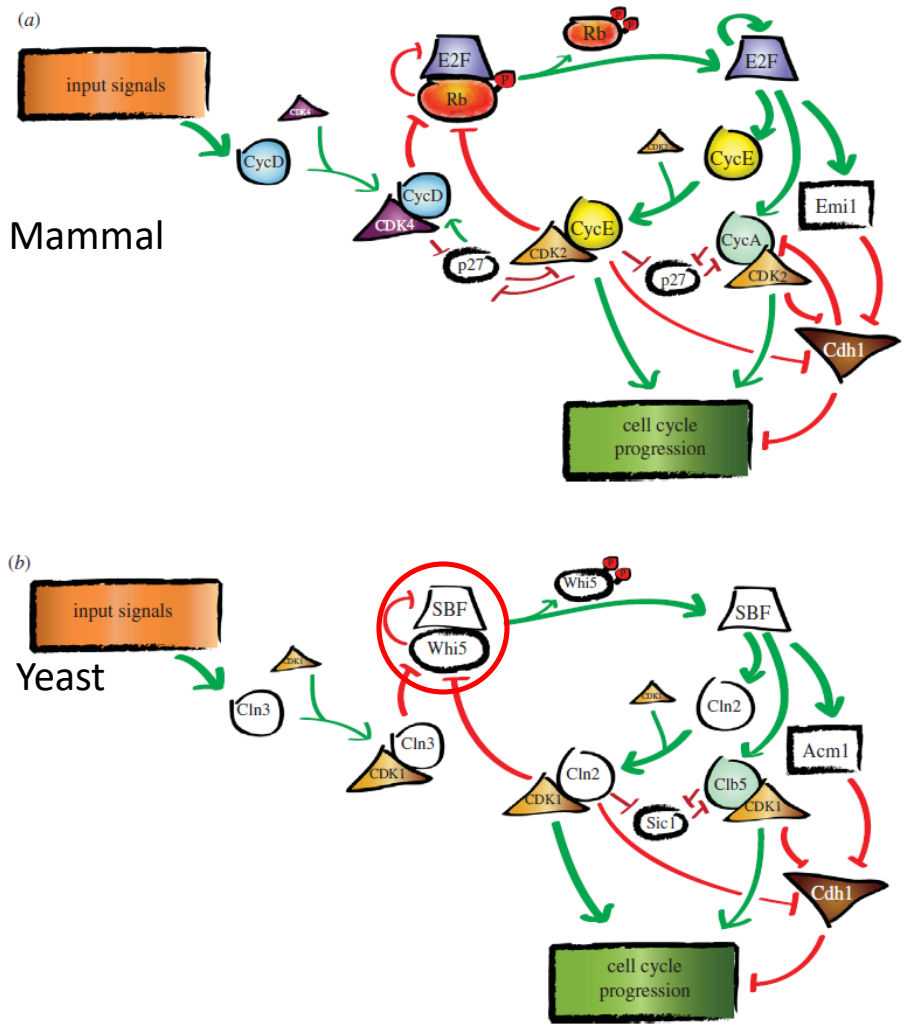


Figure 1. Schematic of (*a*) mammalian and (*b*) budding yeast G1–S control circuits indicates a common feedback-driven regulatory architecture. Shapes correspond to the type of protein (e.g. upward triangles denote cyclin-dependent kinases). Colour implies that the G1–S regulator has high sequence similarity (indicating homology, table 1) to the same regulator in another kingdom (animal, fungi, plant). The G1–S circuit in mammals is colourful (compared with budding yeast) because there are many identifiable sequence homologues between plants and animals.

- Triangles denote CDKs; circles denote cyclins.
- Green arrows denote activation steps; red lines with blunt ends denote inhibitory interactions.
- Proteins with colored enclosures have orthologous sequences in other eukaryotes, including plants.
- Those with the same color in mammal and yeast are orthologous to each other; those in white are unrelated.

Cross et al. (2001, Phil. Trans. Roy. Soc. B)

Three Potential Paths to Network Rewiring

Simple sequence divergence:

Parallel gain of nonorthologous participants:

Divergent resolution of redundancy:

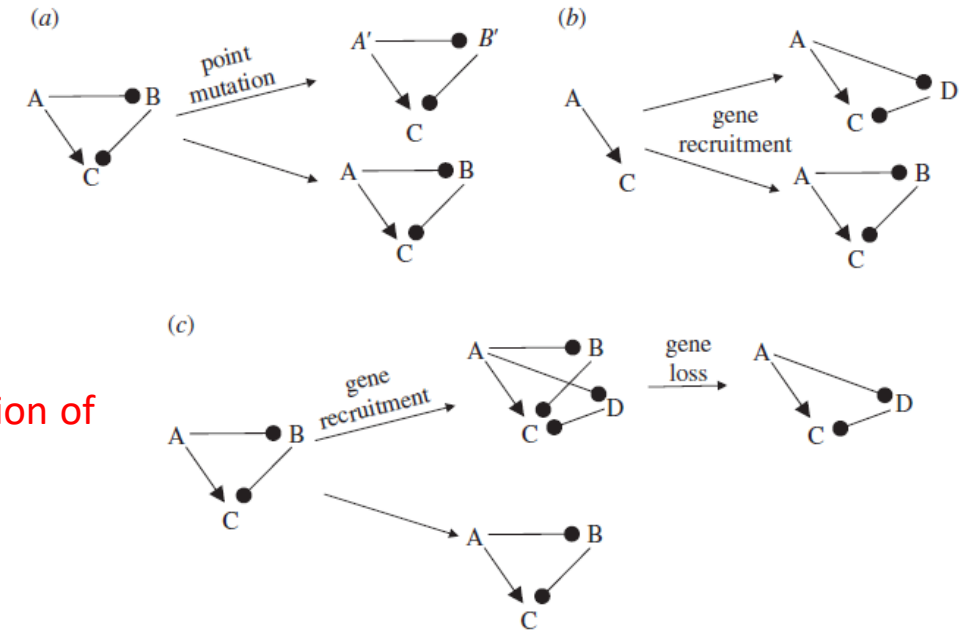


Figure 2. Schematic of possibilities for network evolution. (*a*) A simple network is conserved in two lineages, but along one or both lineages, highly accelerated sequence evolution results in loss of detectable homology in modern descendants (e.g. *A′* and A are direct sequence descendants of A in the ancestor, and have carried out the same network role throughout evolution, but A and *A′* are no longer sequence-alignable). In this case, both network and sequences are monophyletic: from a single origin, the network has retained the same topology, and all sequences have kept the same network role. (*b*) A simple network independently recruits new elements to elaborate the network (note that the 'sense' of the network remains the same, with A still activating the downstream C). In this case, the enhanced network is polyphyletic, as are the new sequences B and D. (*c*) Along one lineage, the network acquires an independent loop redundant with the B loop, allowing subsequent loss of B along this lineage, without losing network function at any step. In such a case, we consider the network monophyletic, even though the sequences are polyphyletic. Note that in the case of recruitment, B and D could be ancient relatives. Provided D did not carry out the indicated network role in precursor organisms, this still constitutes sequence polyphyly for this network.

Redundancy in the check-point mechanisms for limiting replication to one episode per cell cycle.

- The Cdc6 protein is responsible for “licensing” replication origins (part of the pre-replication complex).
- CDKs regulate Cdc6 stability in *S. cerevisiae* by at least three mechanisms: direct phosphorylation that leads to targeted proteolysis; nuclear exclusion; and direct binding.
- Does this redundancy reduce the error rate?

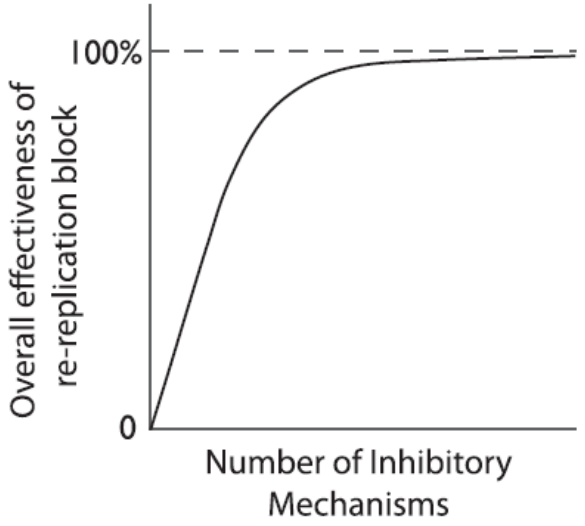
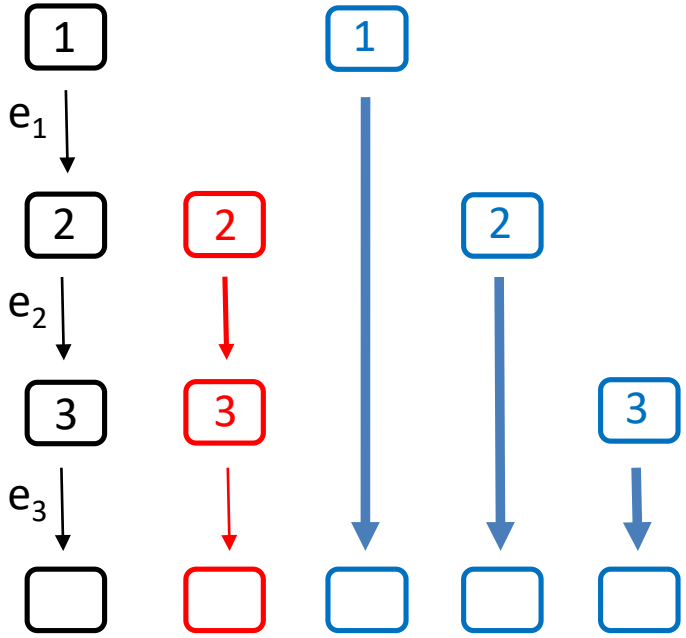
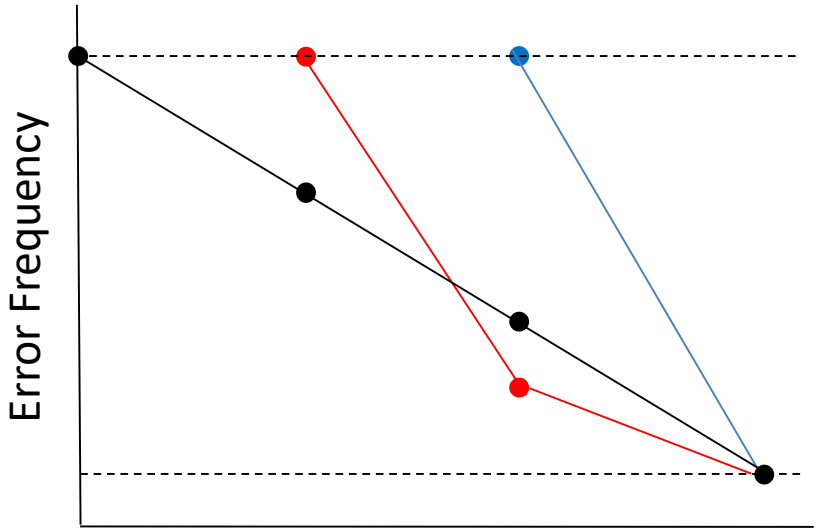


Figure 4. Model: The Evolution of Multiple Pre-RC-Regulatory Pathways
Details of the model are described in the main text.

Drury and Diffley (2009)

The Paradox of Robustness

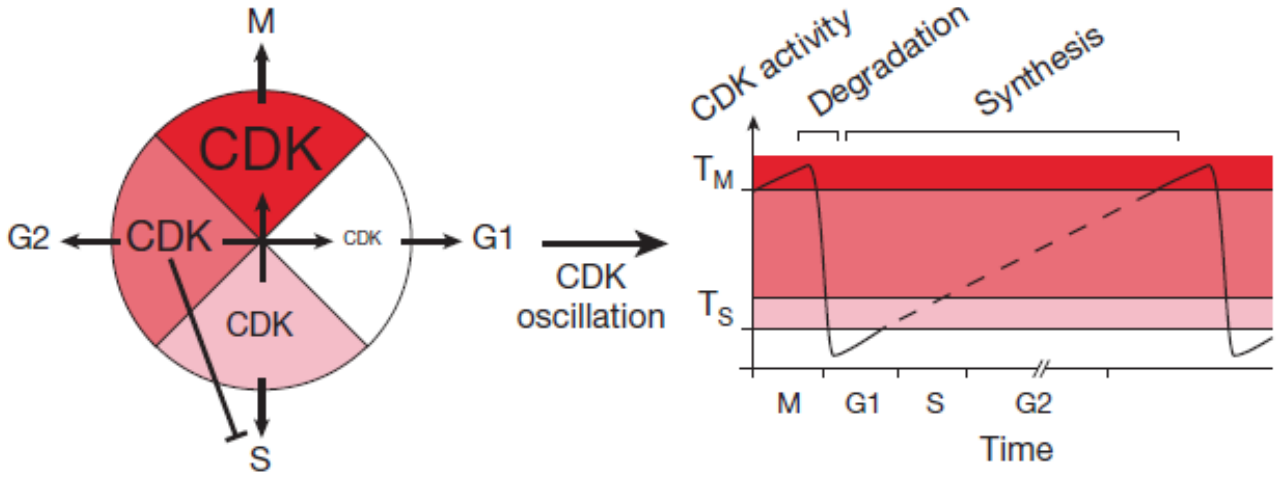




Why is the Cell-Cycle Pathway so Complex?

The *Schizosaccharomyces pombe* cell cycle can be engineered to run with an extremely simplified control mechanism that relies on just one major CDK fused to a single cyclin (Coudreuse and Nurse 2010).

This demonstrates the feasibility of a simple ancestral system for cell-cycle progression driven autonomously by a simple self-oscillating module, requiring no differential expression, interaction, and degradation of multiple participants.

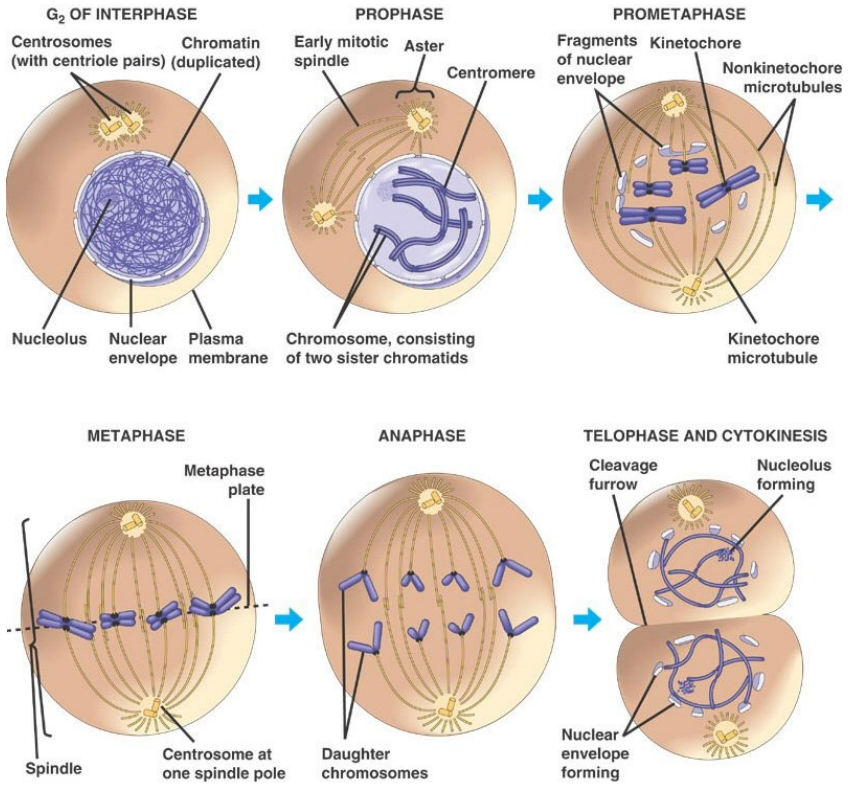


Major Transitions in the Evolution of Eukaryotic Mitosis

- Enormous expansion in organizational / pathway complexity prior to LECA.
- Enormous expansion in the complexity of molecular components prior to LECA.
- Enormous phylogenetic diversification in cell ultrastructural features following the establishment of eukaryotes.
- Coevolution of subcomponents of molecular complexes leads to isolation between phylogenetic lineages.

Eukaryotic mitosis: all key molecules involved in mitosis have homologs in archaea.

- Establishment of the eukaryotic mitotic mechanism involved the introduction of at least eight modifications not found in prokaryotes.



1) mitotic checkpoints;

2) switch from membrane- to spindle-based mechanism for chromosome segregation;

3) the evolution of centromeres or spindle-attachment sites;

4) molecules for chromosome sister-chromatid cohesion and condensation;

5) expansion in the numbers of origins of chromosome replication;

6) expansion of nucleosome size and number;

7) evolution of telomeres to prevent the erosion of the ends of linear chromosomes.

Mitotic Checkpoints: prevent division from initiating before all chromosomes have established connections with kinetochores.

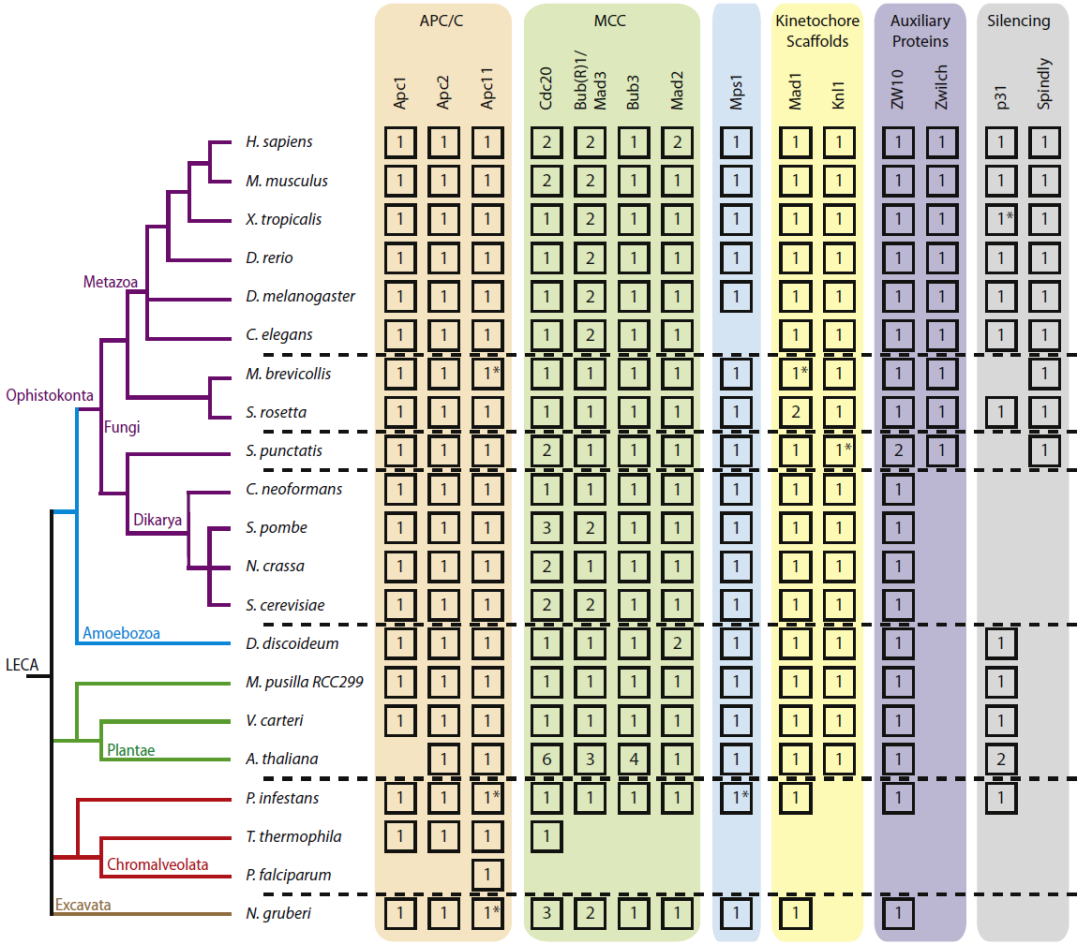


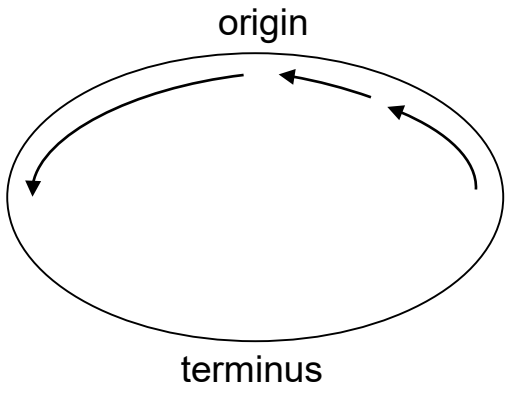
Figure 2. Homologs of the Core and Auxiliary MC Proteins
Schematic representation of eukaryotic tree of life in which a selection of eukaryotic species from the five different supergroups is indicated on the left. Checkpoint proteins are grouped in different functional groups (MCC, Mps1, kinetochore scaffolds, auxiliary proteins, silencing), and, whenever present, the number of homologs is indicated in black boxes (for gene IDs, see Table S1; for protein sequences, see Supplemental Sequence Files). Data on APC/C subunit homologs are adapted from (Eme et al., 2011); asterisks indicate potential homologs of MC subunits in genomic DNA from nonannotated genes (see Supplemental Experimental Procedures and Figure S3).

- Virtually all components appear to date back to LECA.
- Numerous core components appear to be missing from ciliates and *Plasmodium*.

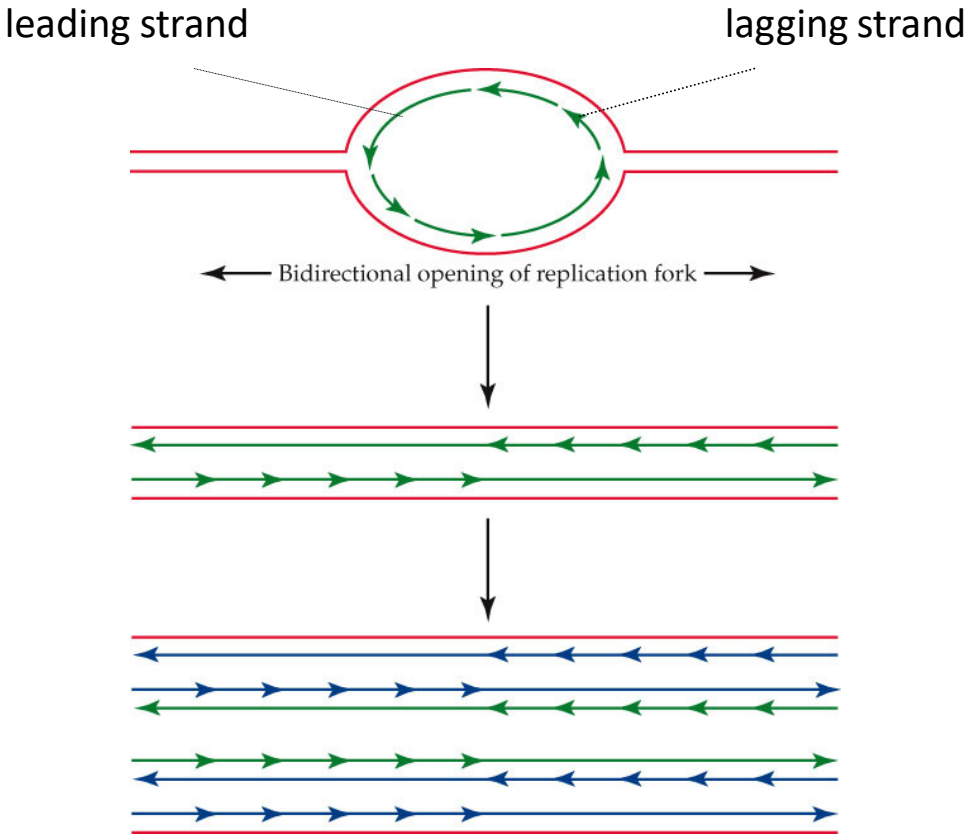
Vieugel et al. (2012, Cell)

Origins of Replication (ORIs): a Replication Feature Shared by All Organisms

- Bidirectional.
- Discontinuously produced Okazaki fragments are stitched together.



Prokaryotes



Eukaryotes

How are ORIs recognized?

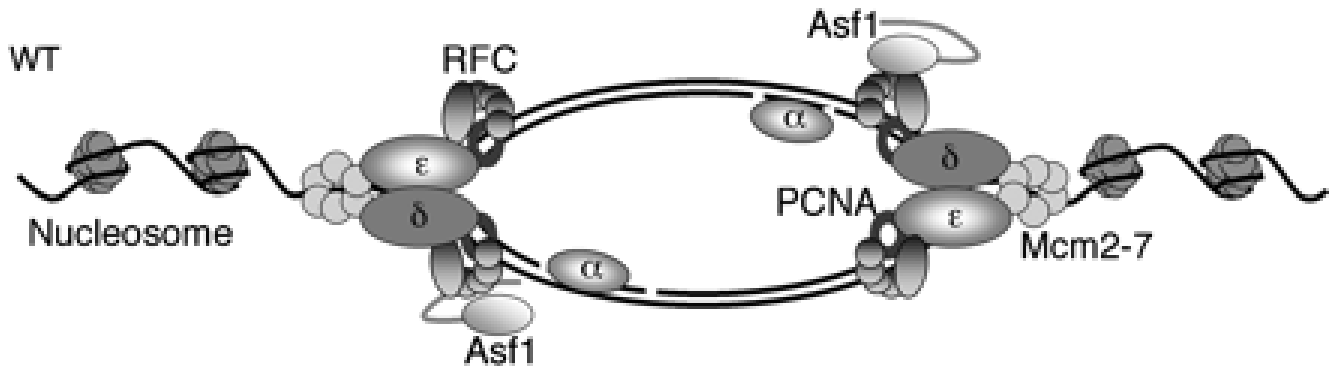
- **Budding yeast (*S. cerevisiae*)** – discrete, ~200 bp in length, and contain at least one copy of an 11-bp key binding-site sequence.
  - ORIs are autonomously replicating sequences (ARSs) that support replication when placed into artificial constructs. ~300 such elements per haploid genome.

- **Fission yeast (*S. pombe*)** – 0.5 to 1.0 kb in length, sometimes containing an 11-bp ARS-like sequence along with at least two 30 to 55-bp A/T-rich regions and other ill-defined stimulatory sequences.
  - ~50% of random intergenic sequences from *S. pombe* with lengths and A/T contents similar to verified ORIs exhibit potential origin activity.

- **Animals** – thousands per genome; 1.0 to 10.0 kb in length; ill-defined, but A/T rich; often with multiple initiation sites.

Major Protein Complexes Involved in Replication in Archaea and Eukaryotes

- DNA polymerases – replicate the DNA, using a single-strand template.
- Proliferating cell nuclear antigen (PCNA) – encircles the DNA and clamps the polymerase to the template.
- Minichromosome maintenance complex (MCM) – the helicase that unwinds replicating DNA.
- Replication factor C (RFC) – the clamp loader, which recognizes primer/template regions.



Architecture of the Replication Machinery: basic eukaryotic structures trace back to the archaea

- The participating proteins typically assemble into multimers.

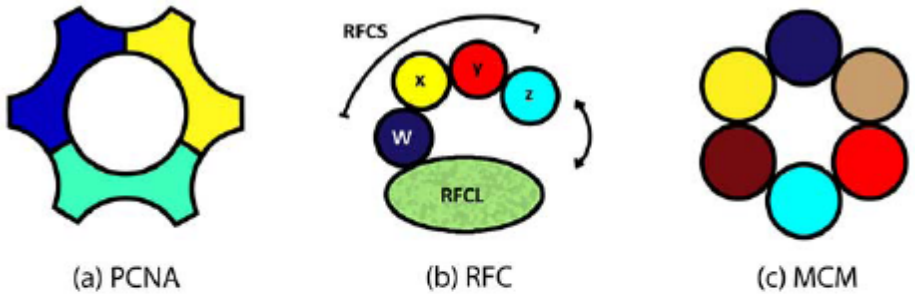


Figure 1. Structural schematic of the PCNA, RFC, and MCM complexes. (a) PCNA consists of 3 subunits forming a ring-like clamp that encloses the DNA polymerase and single stranded DNA. (b) RFC consists of a total of five subunits. Four small subunits (RFCS) form a chain, whose positions are labeled *w*, *x*, *y*, and *z*, that is anchored by *w* RFCS to one large subunit (RFCL). The complex opens between the terminal *z* RFCS and RFCL via an ATP driven conformation change. (c) The MCM complex consists of six MCM proteins in a hexameric ring.

Chia et al. (2010, PLoS ONE)

Table 1. Number of PCNA, RFCS, and MCM subunits found in Archaea and eukaryotes for literature [1,3,21–23,27,28,33,51,52,59,66,67] and this work.

	Number of distinct subunits		
Taxonomic Unit	**PCNA**	**RFCS**	**MCM**
Archaea			
Crenarchaeota	1,2,3	1,2	1
Euryarchaeota	1,2	1,2	1–4,8
Korarchaeota	1	1	1
Nanoarchaeota	1	1	1
eukaryotes	1	4	6
total number of subunits in structure	3	4	6

- Archaeal complexes contain the same numbers of subunits, but tend to be homomers.
  Eukaryotic complexes are heteromers, with subunits derived from duplications independent of those in archaea.

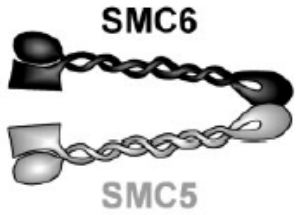
- No evidence that the emergence of increased molecular complexity was driven by adaptive processes.

Evolution of the SMC (Structural Maintenance of Chromosome) Proteins: condensins and cohesins.

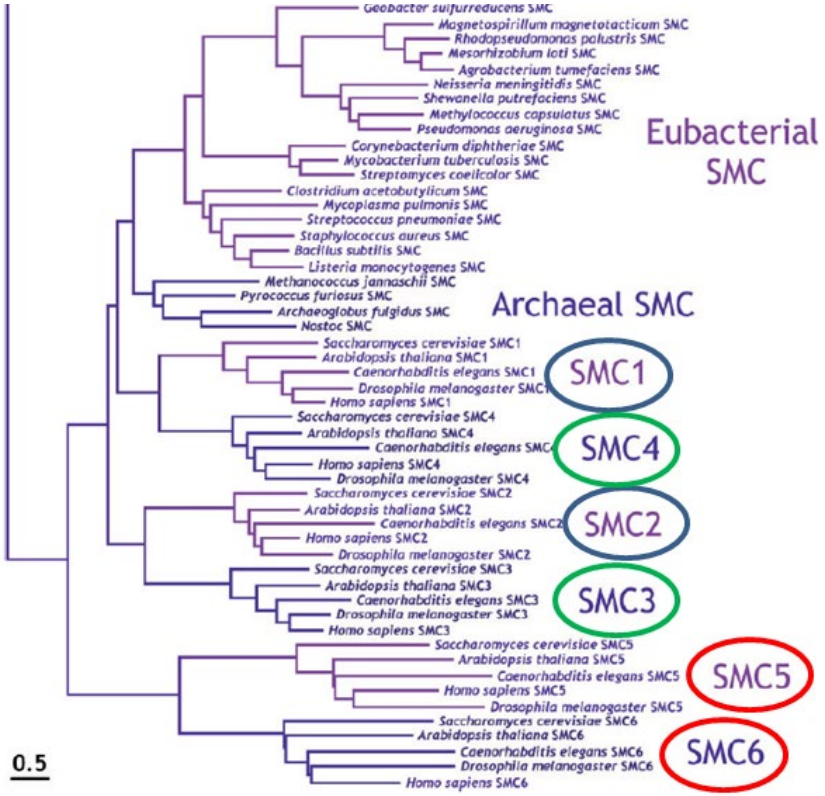
- Ancient coiled-coil proteins: all eukaryotes, and nearly all archaea and bacteria have them.
- Prokaryotes have just single homodimers.
- Eukaryotes generally have six paralogs, which form three heterodimers:

part of the condensin complex (metaphase)

part of the cohesin complex (prophase)



DNA repair

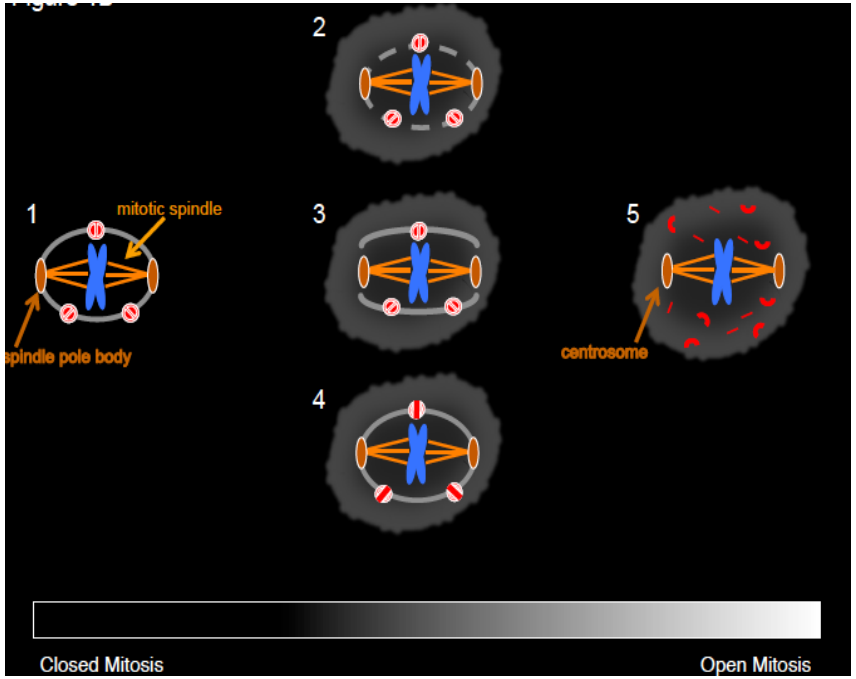


Gene duplications gave rise to the condensins and cohesins prior to LECA

Cobbe and Heck (2004, Mol. Biol. Evol.)

Morphological Diversity: Open vs. Closed Mitosis

- Driven by the expansion of “selfish” transposable elements?
- Scaling issues with increasing genome size?
- Essentially every variant with respect to degree of openness and spindle location exists across phylogenetic lineages.



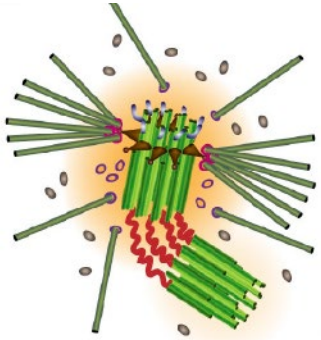
Some Fungi

Animals, Plants

Major comparative treatises by Raikoff and Kubai; Sazer et al. (2014, Curr. Biol.)

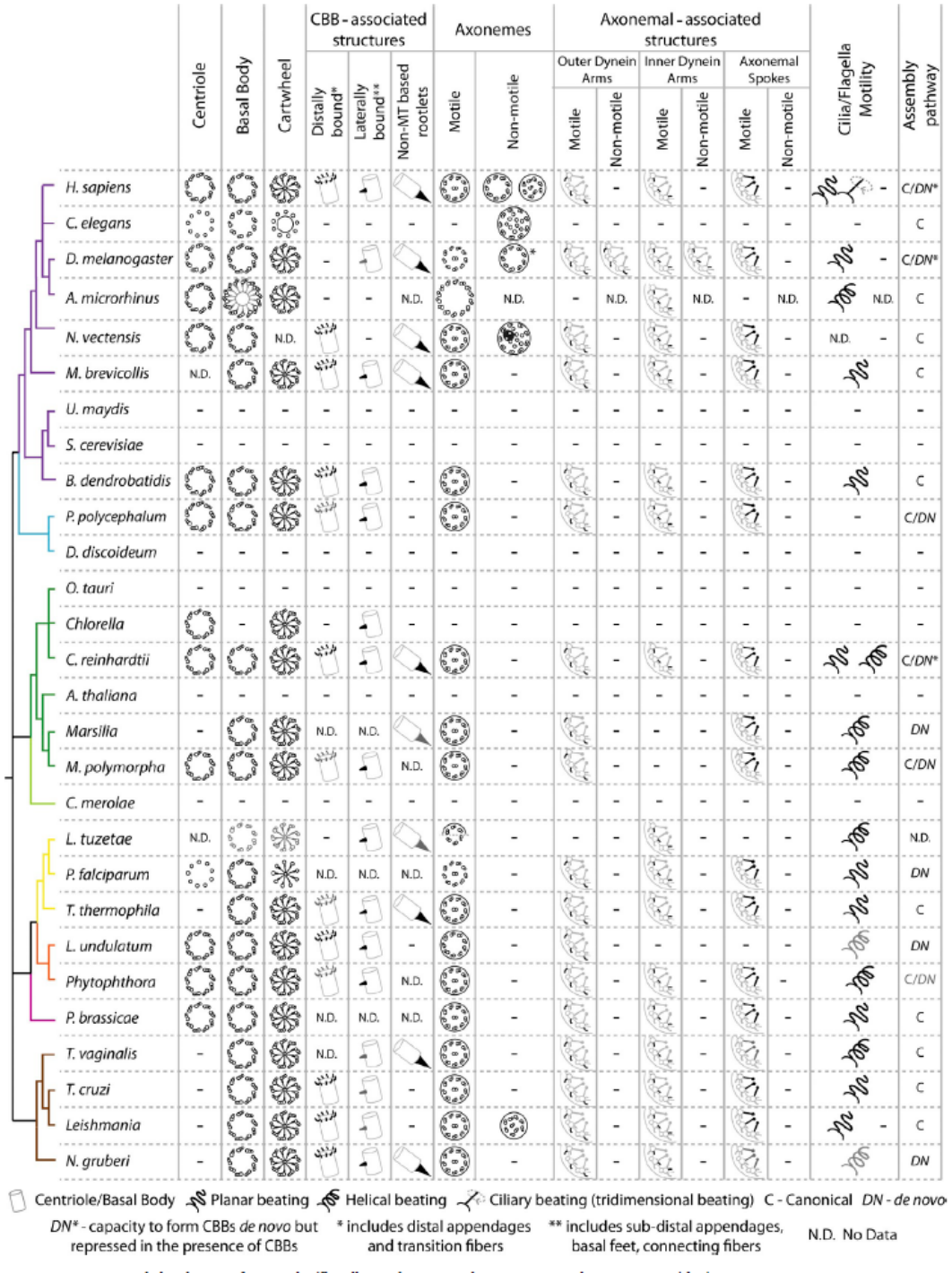
Phylogenetic Diversification of Centriole Components

- Present in LECA; absent from all prokaryotes.



- Involved in cytokinesis, generally serving as cores of the centrosome, which anchors the mitotic spindle.
- Also serve as “basal bodies,” initiating the assembly of cilia and flagella.
- The basic nine-element cylindrical structure is universal, but the units of assembly can consist of 1, 2, or 3 microtubules.
- Some lineages have no centrioles or basal bodies (e.g., vascular plants, most fungi, some red algae); some have only a centriole (some green algae); and some have only a basal body (tyrpanosomes).

Carvalho-Santos et al. (2011, J. Cell Biol.)



Centriole Biogenesis

- Centriolar proteins appear to evolve unusually rapidly, so that some may not be found by standard orthology searches.

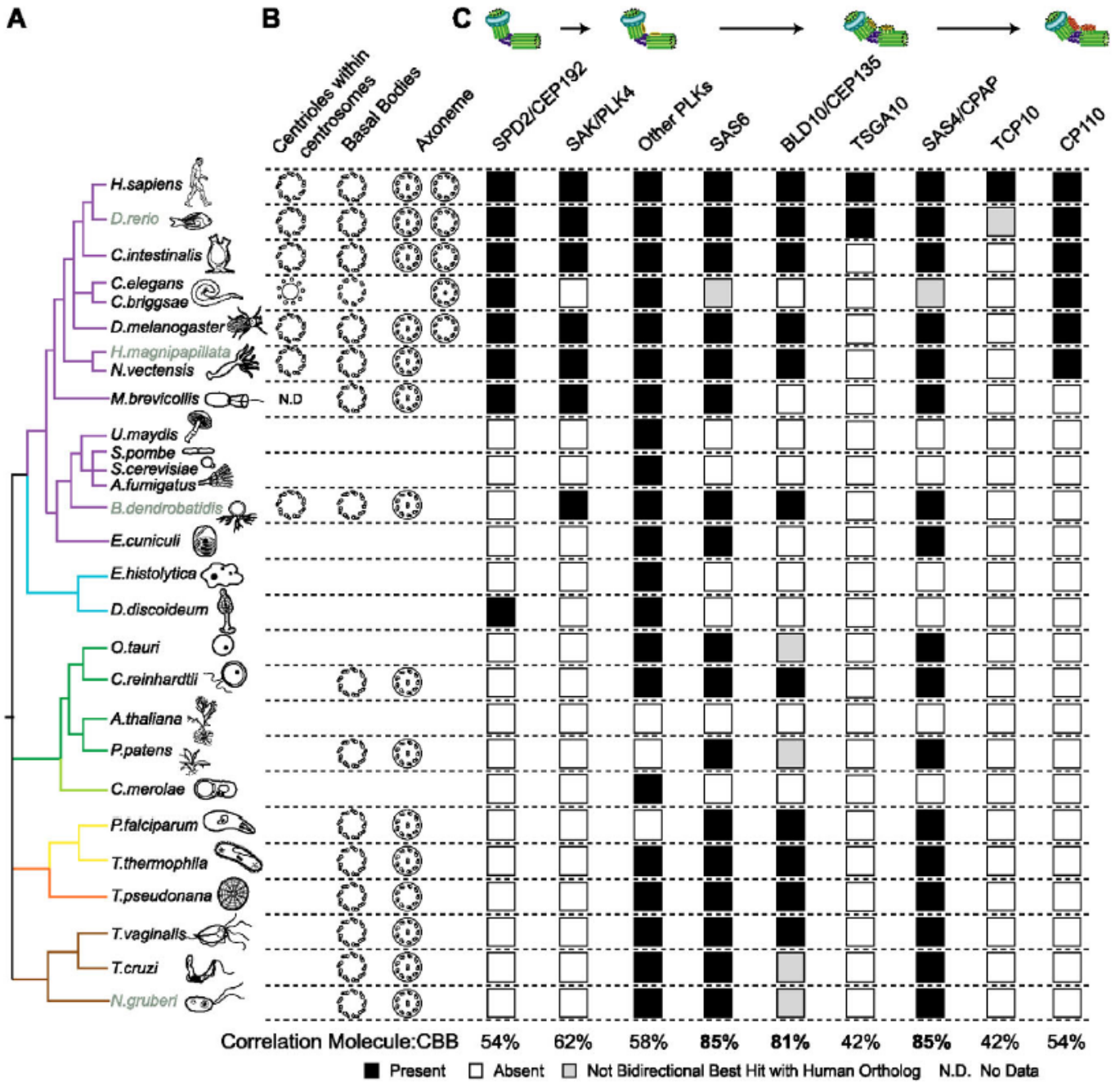


Fig. 2. Phylogenetic distribution of families of molecular players in CBB biogenesis. (A) Simplified taxonomic tree representing crown eukaryotic groups in

Carvalho-Santos et al. (2011, J. Cell Sci.)

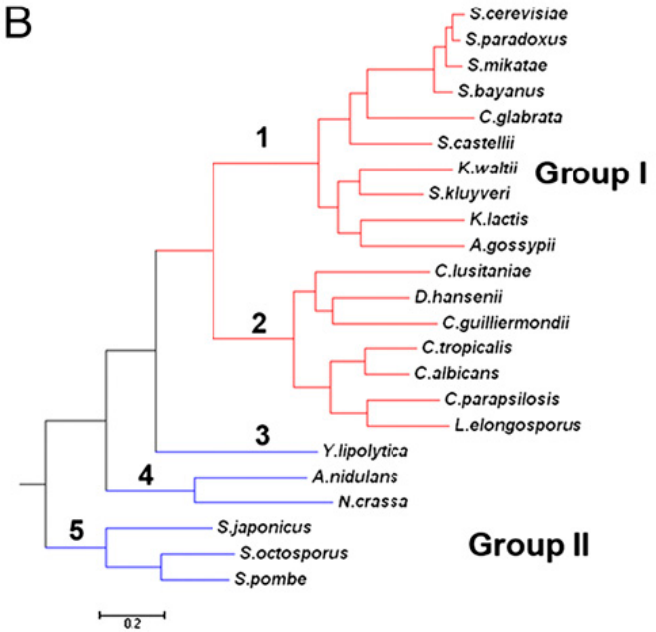
PCNA Coevolves with its Interacting Partners in Ways That Promote Interspecific Incompatibilities: an example of effectively neutral, passive emergence of species isolation.

Key changes in the Interdomain Connecting Loop (to which PCNA interactors bind):

A

		IDCL region											
Group I	*S. cerevisiae*	D	I	D	A	D	F	L	K	I	E	E	L
	S. castellii	D	I	D	A	D	F	L	K	I	E	D	L
	K. waltii	D	I	D	A	D	F	L	D	I	E	G	M
	K. lactis	E	I	E	T	D	F	L	E	I	S	D	M
	S. kluyveri	D	I	D	A	D	F	L	E	I	E	G	M
	A. gossypii	E	I	Y	A	D	F	L	E	I	D	Q	T
	C. lusitaniae	D	I	D	S	E	F	L	R	I	D	D	M
	D. hansenii	D	I	D	S	E	F	L	K	I	D	D	M
	C. tropicalis	D	F	D	S	E	F	L	Q	I	E	D	M
	C. albicans	D	I	D	S	E	F	L	Q	I	D	D	M
	C. parapsilosis	N	I	D	S	E	F	L	K	I	D	D	M
Group II	*Y. lipolytica*	T	I	D	Q	E	H	L	G	I	P	D	T
	A. nidulans	D	I	D	Q	E	H	L	A	I	P	E	T
	N. crassa	D	I	D	Q	E	H	L	G	I	P	D	T
	Sc. pombe	D	I	D	Q	E	H	L	G	I	P	D	I

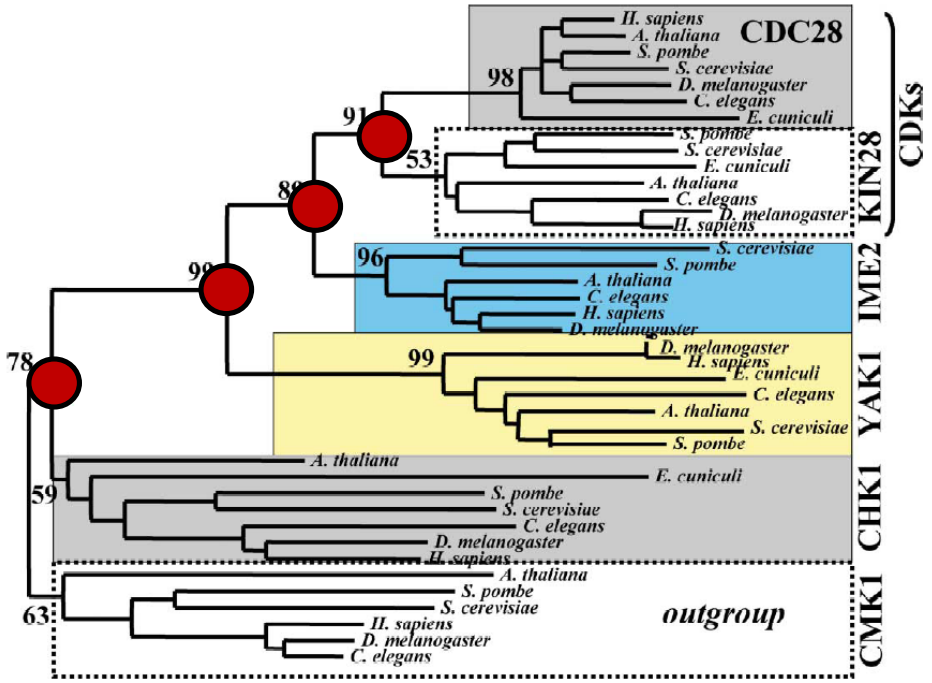
B



- In “yeast two-hybrid” analyses, the IDCL of Group I species is able to bind with interacting proteins of other Group I species, but not with those from Group II.

Zamir et al. (2012, PNAS)

Did the Cyclin-CDK Mechanism of Cell-Cycle Control Evolve Before or After the Origin of Mitosis?



Cell-cycle control kinases

Meiotic-specific regulation

Mitotic-specific regulation

DNA-damage checkpoint

Calmodulin kinase

Other kinases

Figure 3. A Maximum Likelihood Phylogenetic Tree of the Conserved CDC28 Family Kinases from Seven Eukaryotic Species
The color code and designations are as in Figure 2. Clades with kinases not involved in CCC are boxed with a dashed line. The CMK1 clade is used as the outgroup.

- Duplications preceded speciation events suggesting that the basic eukaryotic cell cycle dates to LECA.
- Branching order should be consistent with order of modifications, and suggests an order of events of mitosis → meiosis → cell-cycle control.

Krylov et al. (2003, Curr. Biol.)