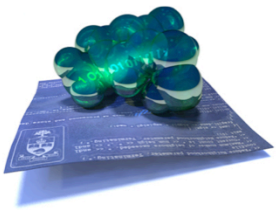
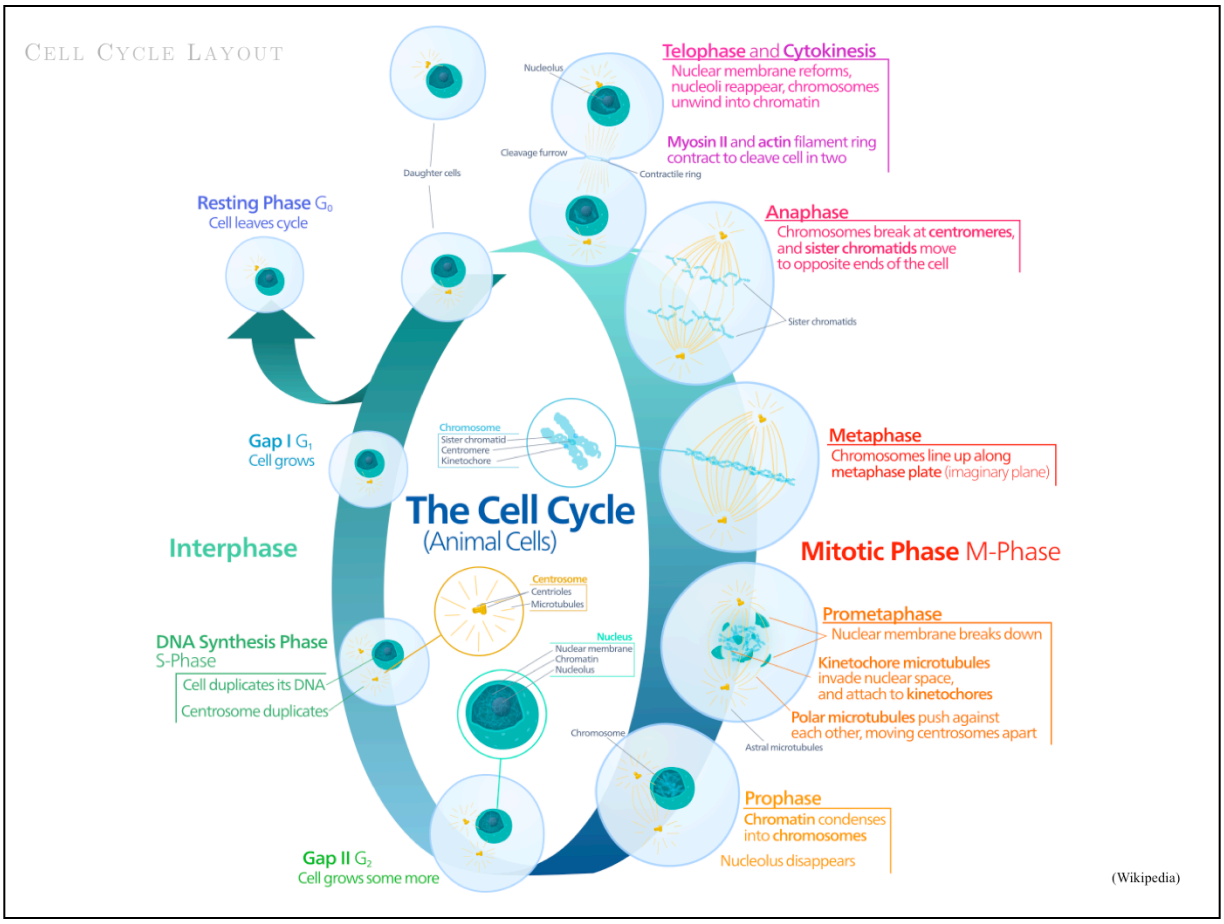
A
Bioinformatics
Course

Cell Cycle



Boris Steipe

department of biochemistry – department of molecular genetics
University of Toronto



The general layout of the cell cycle is universally conserved. If you have not come across its main aspects in previous courses, Wikipedia is a possible starting point: https://en.wikipedia.org/wiki/Cell_cycle

EXAMPLE: YEAST G1/S TRANSITION CONTROL

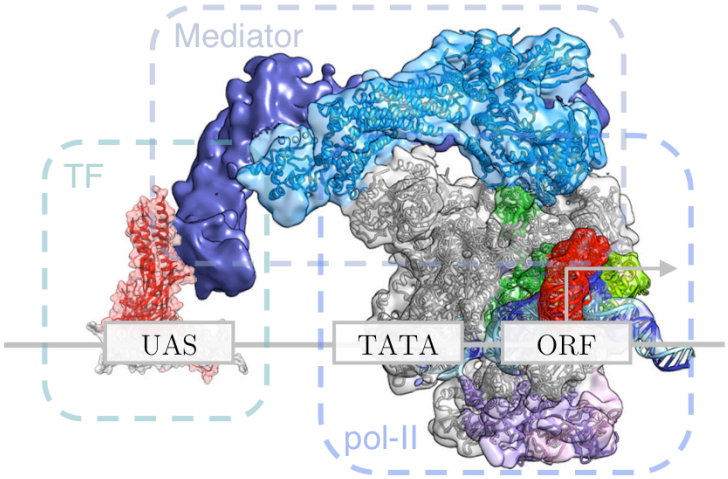
Definition of the G1/S transition switch **system** – "Bottom up"

A collection of facts (cf. McInerny 2011, Travesa 2013, Haase 2014)

Defined from carefully done,
individual wet-lab experiments:

Components

- Cdc14
- Swi6
- Mbp1
- Swi4
- Whi5
- CDK(**Cdc28**)/Cln1,2,3
- SCB and MCB UAS
  (Upstream Activating Sequences)
- Rad53
- CDK(**Cdc28**)/Clb1,2,6
- Nrm1



Note: among these genes, only **Cdc28** is "essential"!

We need to start somewhere on our quest to define a G1/S transition control *system* and literature annotated genes (or genes annotated to GO terms) is a good a starting point as any. Key roles are played by the MBF and the SBF complexes that bind to a large number of UAS elements of their target genes, enabling transcription of the downstream ORF through recruitment of polII via the mediator complex. This is generally how regulatory transcription factors work and this mechanism provides much room for refinements in the time-course of activation – via early or later binding of transcription factors to high- resp. low-affinity variations of the UAS sequence, and refinements in the logic of activation eg. by mixing UAS to provide for mutually exclusive occupation of sites, or producing a requirement to have multiple sites occupied for effective activation.

Like many regulatory systems, we have a protein (here: Swi4 and Mbp1) that serves to integrate many different regulatory signals and pathways, which then leads to a unified signal output. We often think of these proteins as a "hub" of the system. The protein transcription factor needs to be bound and brought into an active state. Much has been written about this system but here is a handful of the most salient facts:

McInerny CJ. (2011) Cell cycle regulated gene expression in yeasts. Adv Genet. 73:51-85.

Travesa A, *et al.* (2013) Repression of G1/S transcription is mediated via interaction of the GTB motifs of Nrm1 and Whi5 with Swi6. Mol Cell Biol. 33(8):1476-86.

Haase SB and Wittenberg C.(2014) Topology and control of the cell-cycle-regulated transcriptional circuitry. Genetics. 196(1):65-90.

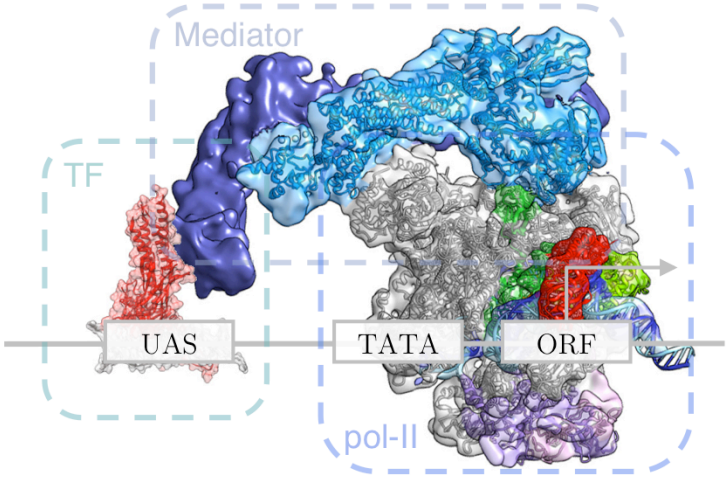
Definition of the G1/S transition switch system – "Bottom up"

A collection of facts (cf. McInerny 2011, Travesa 2013, Haase 2014)

Turning it on ...

On
- Cdc14
- Swi6
- Mbp1
- Swi4
- Whi5
- CDK(**Cdc28**)/Cln1,2,3
- SCB and MCB UAS
  (Upstream Activating Sequences)

- Rad53
- CDK(**Cdc28**)/Clb1,2,6
- Nrm1



Note: among these genes, only **Cdc28** is "essential"!

Turning it on ...

- Swi6 enters the nucleus after dephosphorylation by Cdc14.
- One wave of gene activation is associated with the MBF complex binding to MCB DNA elements. MBF consists of Mbp1 (with DNA binding activity) and Swi6 (with regulatory activity). The target genes predominantly function in DNA replication.
- Another wave of gene activation is associated with the SBF complex binding to SCB DNA elements. SBF consists of Swi4 (with DNA binding activity) and Swi6 (with regulatory activity). The target genes predominantly function in the preparation of budding, cell-wall biosynthesis etc.
- Swi4 expression itself is regulated through ECB motifs in its promoter: this causes expression during G1 and provides a basal amountof Swi4 transcription factor to prime its activity. Once Swi4 is switched on, it targets its own gene's promoter, thus providing a positive feedback loop.
- Binding of MBF/SBF to their respective UAS does in itself not lead to activation. Activation requires the action of a CDK (cyclin dependent kinase) complex of Cdc28 and Cln3.
- The activation of MBF is obscure, but depends on a Cdc28/Cln CDK.
- SBF is activated when phosphorylation by CDK removes the inhibitor Whi5 which is normally bound to Swi6 in the SBF complex. Phosphorylated Whi5 is excluded from the nucleus.

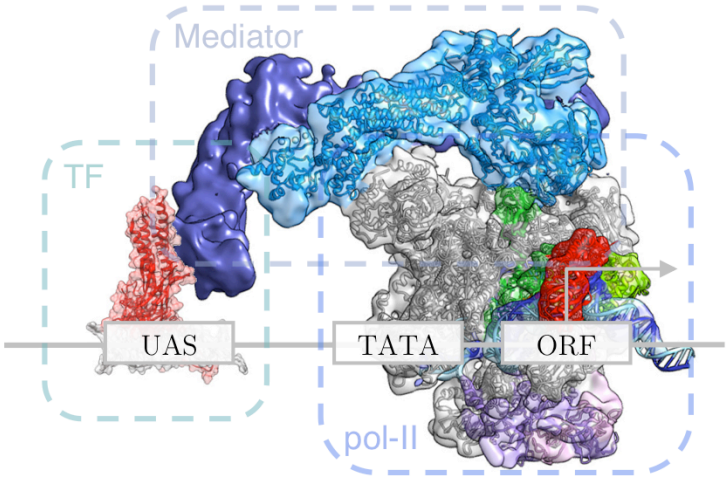
Example: Yeast G1/S transition control

Definition of the G1/S transition switch system – "Bottom up"

A collection of facts (cf. McInerny 2011, Travesa 2013, Haase 2014)

... and turning it off again.

- Cdc14
- Swi6
- Mbp1
- Swi4
- Whi5
- CDK(**Cdc28**)/Cln1,2,3
- SCB and MCB UAS
  (Upstream Activating Sequences)
- Rad53
- CDK(**Cdc28**)/Clb1,2,6
- Nrm1 Off



Note: among these genes, only **Cdc28** is "essential"!

... and turning it off again:

- MBF is repressed in a negative feedback loop, when one of its target genes, the inhibitor Nrm1, accumulates as the S-phase progresses; Nrm1 too binds to Swi6 but only in the MBF complex.
- Both MBF and SBF activity is turned off in G2 by a CDK complex that contains Cdc28 and the B-type cyclins Clb1/Clb2; these are activated by Cln1 and Cln2 which are induced by MBF and SBF – this provides a negative feedback loop that is responsible for timely conclusion of the S-phase expression wave.
- SBF and MBF complexes are disassembled towards the end of the S-phase when Swi6 is phosphorylated and excluded from the nucleus. This involves a CDK complex of Cdc28 and the B-type cyclin Clb6; the latter, again, is a MBF / SBF target gene and thus creates a negative feedback control.
- Rad53 can prevent the inactivation of MBF by phosphorylating Nrm1. This leads to sustained expression of the DNA replication and repair genes under the control of MBF. In this way, Rad53 performs its function as a checkpoint gene, prolonging S-phase gene expression if necessary. This is an interface between two systems.

EXAMPLE: YEAST G1/S TRANSITION CONTROL

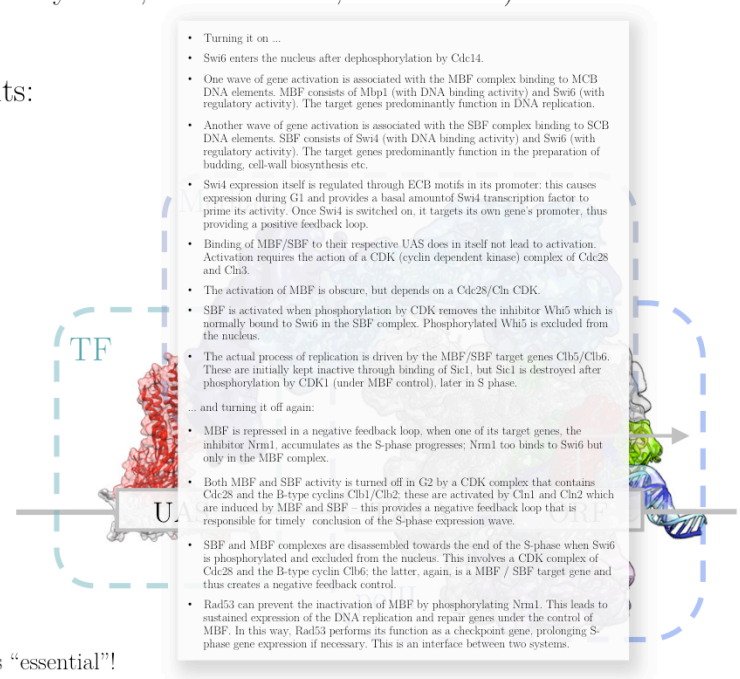
Definition of the G1/S transition switch system – "Bottom up"

A collection of facts (cf. McInerny 2011, Travesa 2013, Haase 2014)

Defined from carefully done, individual wet-lab experiments:

- Cdc14
- Swi6
- Mbp1
- Swi4
- Whi5
- CDK(**Cdc28**)/Cln1,2,3
- SCB and MCB UAS (Upstream Activating Sequences)
- Rad53
- CDK(**Cdc28**)/Clb1,2,6
- Nrm1

Note: among these genes, only **Cdc28** is "essential"!



- Turning it on ...
- Swi6 enters the nucleus after dephosphorylation by Cdc14.
- One wave of gene activation is associated with the MBF complex binding to MCB DNA elements. MBF consists of Mbp1 (with DNA binding activity) and Swi6 (with regulatory activity). The target genes predominantly function in DNA replication.
- Another wave of gene activation is associated with the SBF complex binding to SCB DNA elements. SBF consists of Swi4 (with DNA binding activity) and Swi6 (with regulatory activity). The target genes predominantly function in the preparation of budding, cell-wall biosynthesis etc.
- Swi4 expression itself is regulated through ECB motifs in its promoter; this causes expression during G1 and provides a basal amountof Swi4 transcription factor to prime its activity. Once Swi4 is switched on, it targets its own gene's promoter, thus providing a positive feedback loop.
- Binding of MBF/SBF to their respective UAS does in itself not lead to activation. Activation requires the action of a CDK (cyclin dependent kinase) complex of Cdc28 and Cln3.
- The activation of MBF is obscure, but depends on a Cdc28/Cln CDK.
- SBF is activated when phosphorylation by CDK removes the inhibitor Whi5 which is normally bound to Swi6 in the SBF complex. Phosphorylated Whi5 is excluded from the nucleus.
- The actual process of replication is driven by the MBF/SBF target genes Clb5/Clb6. These are initially kept inactive through binding of Sic1, but Sic1 is destroyed after phosphorylation by CDK1 (under MBF control), later in S phase.

... and turning it off again:

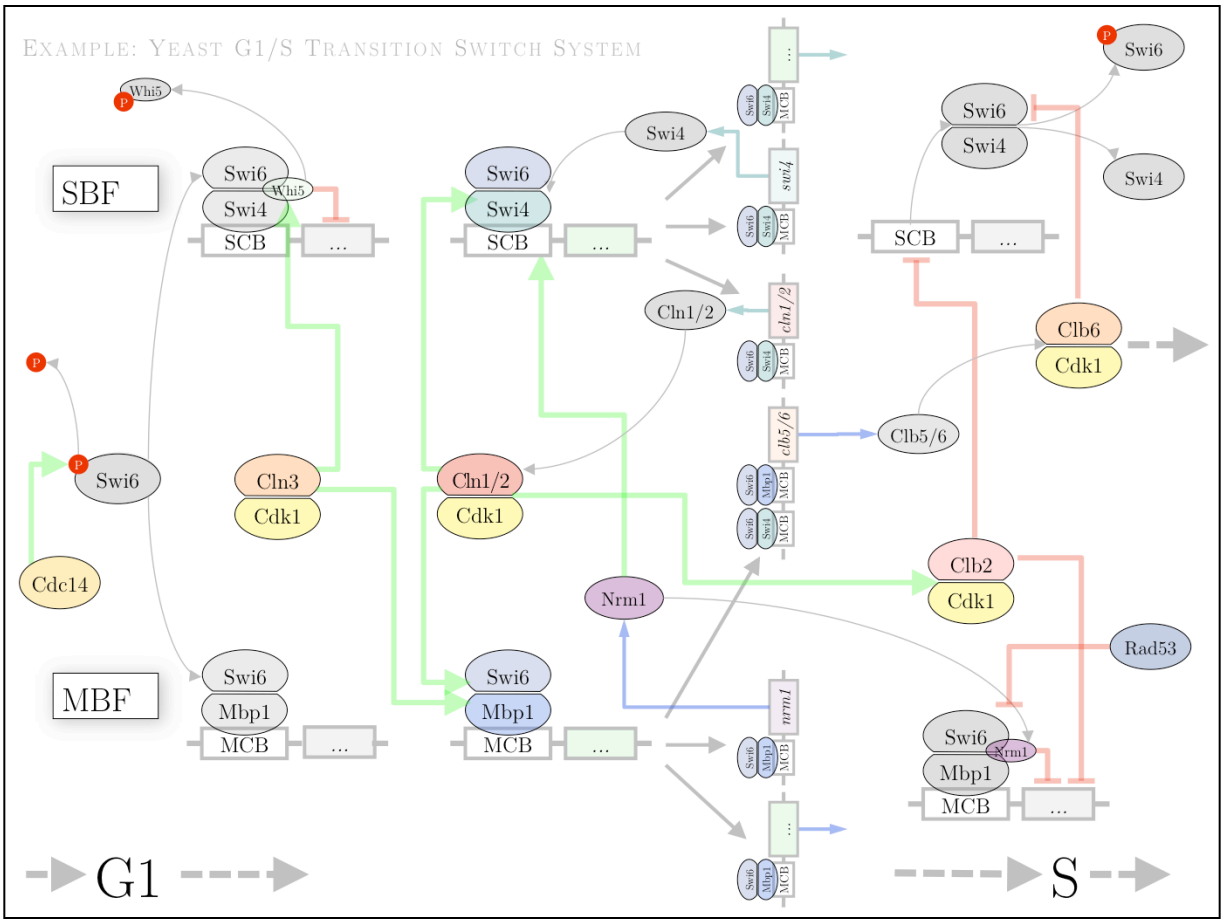
- MBF is repressed in a negative feedback loop, when one of its target genes, the inhibitor Nrm1, accumulates as the S-phase progresses; Nrm1 too binds to Swi6 but only in the MBF complex.
- Both MBF and SBF activity is turned off in G2 by a CDK complex that contains Cdc28 and the B-type cyclins Clb1/Clb2; these are activated by Cln1 and Cln2 which are induced by MBF and SBF – this provides a negative feedback loop that is responsible for timely conclusion of the S-phase expression wave.
- SBF and MBF complexes are disassembled towards the end of the S-phase when Swi6 is phosphorylated and excluded from the nucleus. This involves a CDK complex of Cdc28 and the B-type cyclin Clb6; the latter, again, is a MBF / SBF target gene and thus creates a negative feedback control.
- Rad53 can prevent the inactivation of MBF by phosphorylating Nrm1. This leads to sustained expression of the DNA replication and repair genes under the control of MBF. In this way, Rad53 performs its function as a checkpoint gene, prolonging S-phase gene expression if necessary. This is an interface between two systems.

Though factually accurate as far as I can tell, there are two things wrong with this description that make it only marginally useful:

(i) *Understanding* this collection of facts about symbols is not intuitive. It is remarkably difficult to organize this small body of knowledge. It is also remarkably difficult to commit it to memory.

(ii) The key concepts and ideas are missing from the description: where is the notion of a commited START (or restriction point); where do we mention that this network of interactions constitutes a "switch"?

We have to ask: is there a better way to represent this knowledge? In particular, a way to focus on the conceptual aspects of the system, since these will help us construct a mental map in which we can integrate detailed information from our collection of facts.



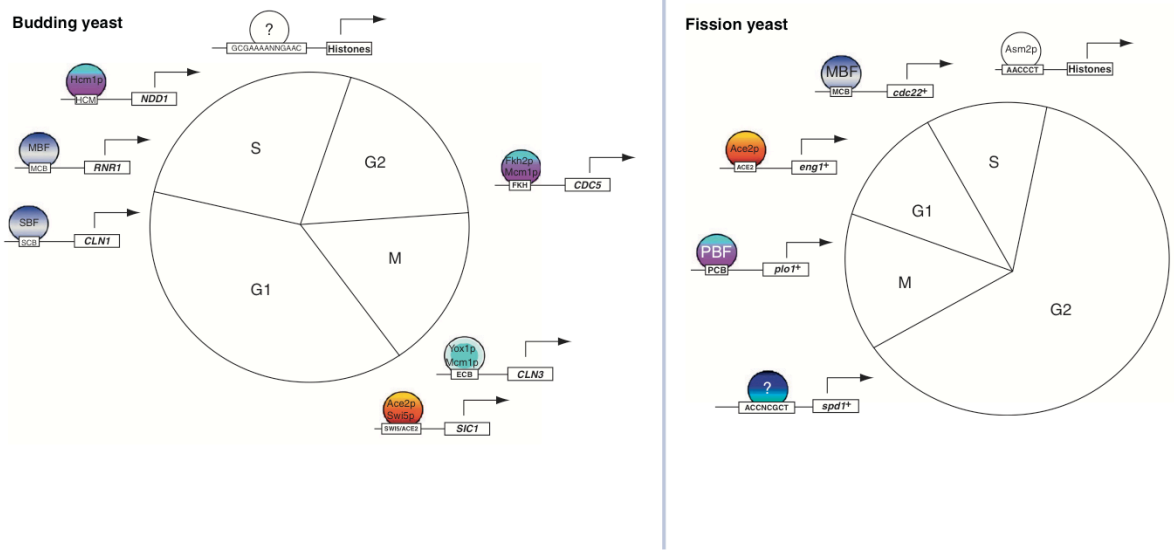
Here is a summary of the facts, modelled in a style that is popular among authors in molecular biology. The sketch will evoke others you have seen on the topic, in particular Haase (2014).

The sketch summarizes the facts by drawing out the activating and inhibiting relationships between the transcription units and their products with green and red lines respectively, and protein movement with grey curves. Active complexes are colored and inactive proteins are grey. In general, changes progress from G1 to S, left to right.

Although the relationships are visualized here, you may find this sketch to be just as impenetrable as the description of the facts, albeit for different reasons. Indeed, a defining element of the *structure* of the cell-cycle are specific points such as START in yeast (or the "restriction point" in higher eukaryotes) that represent an irreversible commitment to replication. Such commitment is not "a gene". Nor is it obvious which of the connections comprise feed-forward loops that switch the system into the "on" state, and feedback loops that switch it "off" again. Finally it is very hard to figure out which of the components provide "internal" control, and which ones constitute the input and output of the system. These points are the most severe criticism of sketches like these (and enumeration of components + activities): **the concept behind the system, the purpose, or meaning if you will is not expressed explicitly**.

We will address possible approaches in the "Systems Model" unit.

Cell cycle conservation



Cell cycle regulated gene expression in budding and fission yeasts.

For each yeast species, a single gene is shown to represent the group of genes that are transcribed at particular cell cycle times. Where known the name or DNA sequence of the upstream activating sequence and the transcription factor are indicated. Related transcription factors in the two yeasts are shown by the same color.

(From: McInerny 2011)

The estimated divergence time of budding yeat and fission yeast is 590 MYA. To put this into perspective, these two yeasts are about as distantly related to each other as you are to the river lamprey! General aspects, and some details of the cell cycle are conserved – but there are major difference too. Our learning units will focus on tranlating model-organism knowledge into other species, which we will call “YFO” to analyse such similarities and differences.

http://steipe.biochemistry.utoronto.ca/abc

B O R I S . S T E I P E @ U T O R O N T O . C A

DEPARTMENT OF BIOCHEMISTRY & DEPARTMENT OF MOLECULAR GENETICS
UNIVERSITY OF TORONTO, CANADA