1. INTRODUCTION

This introduction will first describe the general properties of the eukaryotic cell cycle and review what is known about this process in the two model organisms budding yeast *Saccharomyces cerevisiae*, fission yeast *Schizosaccharomyces pombe* and in multicellular eukaryotes. The second part will focus on the role of transcriptional regulation during the cell cycle and present the use of DNA microarray technology as a global approach to the study of this complex mechanism.

1.1 CELL CYCLE

1.1.1 General overview

Cells reproduce themselves during the cell cycle, which is defined as the series of events that occurs from the birth of the cell to its subsequent division into two new cells. During the cell cycle, all components necessary for survival, most notably the genome, must be doubled in amount and equally distributed to the two newly formed cells at division. The cell cycle has been the subject of investigation over the last 150 years since the "cell theory" was first formulated in 1858 by Virchow in his book *Cellularpathologie*. He was the first one to clearly state that cells form the fundamental structural and functional units of all living organisms and that a new cell can only arise from a pre-existing one.

A century of studies has since revealed that the basic processes and control mechanisms involved in this process are universal in eukaryotes and has led to the view of the cell cycle as a highly regulated developmental sequence that brings about the reproduction of the cell. Most cell cycle research has focused on identifying the components of the cell cycle machinery involved in its control and progression.

In all eukaryotic cells, key regulatory steps are responsible for deciding if the cell will keep dividing or will enter a quiescent state (called stationary phase or G_0) or undergo differentiation, depending on the environmental and nutritional conditions as well as external signals in mammalian cells. Once the cell commits to a new round of DNA synthesis and mitosis, the process is irreversible and the cell cycle machinery is responsible for checking that all ongoing processes have been completed before moving to the next stage.

If any abnormality (unreplicated DNA, damaged DNA, insufficient cell growth) is detected, the cycle will pause and progression will be restored only when the situation has reverted to normal. Many diseases, most notably cancer, have been linked with the aberrant behaviour of some of the protein complexes that drive the cell cycle. Molecules involved in G1/S transition appear to be deregulated in most human tumors (Malumbres M. and Carnero A., 2003) and cardiovascular diseases (Boehm M. and Nabel E.G., 2003) and unsuccessful mitosis can be detected in human neurodegenerative diseases (Alzheimer) (Vincent I. *et al.*, 2003), just to mention few examples. This shows the great medical importance of cell cycle research. A better understanding of this process is therefore needed especially to identify new useful drug targets and better strategies for the treatment of these diseases.

1.1.2 Eukaryotic cell cycle

The eukaryotic cell cycle is usually divided into four separate phases called G1, S, G2 and M, the most crucial events being nuclear and cell division (Fig. 1.1).

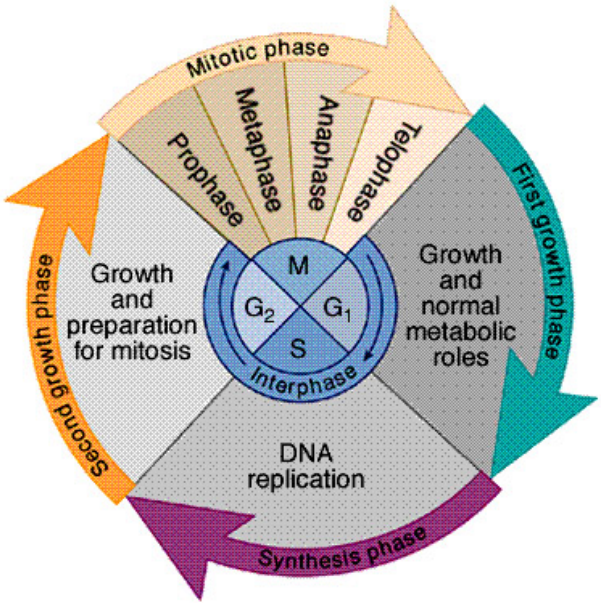


Fig. 1.1 Schematic representation of the eukaryotic cell cycle.
From http://www.bmb.psu.edu/courses/biotc489/notes/biointeract.htm

At the beginning of mitosis, in prophase, the DNA condenses into chromosomes, and the cell's microtubules are rearranged to form the mitotic spindle that will provide a

platform for chromosome separation. As mitosis progresses, the cell pauses in a state called metaphase, in which the duplicated chromosomes are aligned on the mitotic spindle, ready for segregation. The next stage called anaphase starts when the chromosomes move to the poles of the spindle, where they decondense and reform the intact nuclei. The cell then divides in two by a process called cytokinesis (Fig. 1.2).

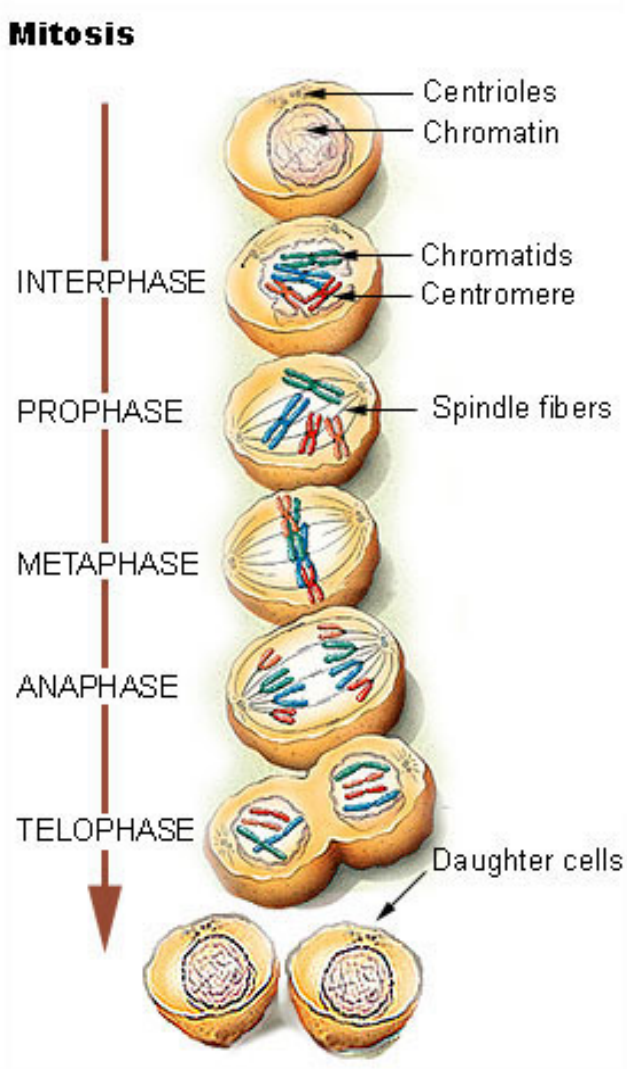


Fig. 1.2 Mitotic division and its phases.
From http://training.seer.cancer.gov/module_anatomy/unit2_1_cell_functions_2.htm

The other, much longer part of the cycle is known as interphase and includes G1, S and G2. DNA replication takes place during a relatively small portion of interphase, called the S (Synthesis) phase. The interval between M and S is called G1 (Gap1) phase, and the interval between S and M is called G2 (Gap2) phase.

G1 and G2 provide extra time for cell growth: the cell requires time to double its mass before dividing. If not, it would become smaller and smaller after each round of division. During G1, the cell "measures" its own size and only when a critical mass has been reached, commits to DNA replication and completion of a division cycle. Similarly, G2 phase provides a safe gap, allowing the cell to ensure that DNA replication is complete before entry into mitosis. In addition, if cells in G1 have not yet committed themselves to DNA replication, they can pause in a specialized resting state, called G0, where they can remain until the environmental conditions are favourable to resume proliferation.

The molecular machinery responsible of running the cell cycle is a large one and involves many components. For example, DNA replication on its own requires primases and DNA polymerases to synthesise the DNA strand complementary to a pre-existing template, topoisomerases and helicases to unwind the DNA strands and ligases to link strands together. Such complex machinery requires an accurate control. The strategy that has been evolved to achieve the tight regulation of the cell cycle events can be explained by two concepts. The first is the existence of key molecular complexes driving cell cycle progression, called cyclin-dependent kinases (CDKs). The second is that of "checkpoints", which can be defined as specific points when the cell makes sure that all the ongoing processes are completed before progressing to the next stage. I shall first consider the CDKs.

1.1.3 Cell cycle control mechanisms: cyclin-dependent kinases

The process that led to the identification of the key molecules involved in orchestrating cell cycle progression has benefited from studies carried out in different organisms. First came the discovery in yeast of cell division cycle (cdc) mutants whose characterisation allowed to identify genes which, when mutated, make the cell unable to divide and are therefore required for essential cell cycle events (Hartwell L.H. *et al.*, 1970; Nurse P. *et al.*, 1976). The product encoded by one of those genes was later identified in fission yeast as the protein kinase cdc2p (Nurse P. and Bissett Y., 1981).

At the same time, a protein-like substance named maturation-promoting factor (MPF) was isolated in *Xenopus* egg cytoplasm showing chromosome condensation activity (Wasserman W.J. and Masui Y., 1976). This was later on shown to be a complex

made of a protein kinase and a cyclin (Lohka M.J. *et al.*, 1988). In addition, a cyclin was discovered in sea urchin embryos, a protein that is synthesised and destroyed in correspondence with cell cycle division in blastomeres (Evans T. *et al.*, 1983).

These protein complexes are formed from two basic types of components, a protein kinase subunit (called CDK) with a catalytic function, and an activating protein called cyclin with a regulatory function (Cross F., 1995; Wuarin J. and Nurse P., 1996; Stern B. and Nurse P., 1996; Nurse, 1997).

CDK activity is tightly regulated through several mechanisms: (1) binding by specific cyclin cofactors, (2) binding by CDK inhibitors, and (3) inhibitory or activating phosphorylation (Tyers M. and Jorgensen P., 2000).

Progression through the cycle is achieved through alternation of a state of low CDK activity in G1 phase to a state of high CDK activity in S, G2 and M phases. S-phase is initiated when protein kinase activity increases from a very low to a moderate level, and a further increase of activity to a high level initiates mitosis (Fisher D.L. and Nurse P., 1996). Inactivation of the kinase activity at the end of mitosis resets the cell for a new cell cycle. The alternation between those two CDK states guarantees that DNA replication and chromosome segregation will occur only once per division cycle.

Protein kinases are present throughout the entire cycle, and they need to be activated at specific points and inactivated once their action is not required any longer. Such a tight regulation is achieved by collaboration with different cyclin partners. These proteins do not have enzymatic activity by themselves but by binding to their partner kinase they activate the kinase. The CDK subunit recognises and binds to a specific domain, called cyclin box, present in every cyclin (Morgan D.O., 1997). Cyclin concentration oscillates during the cycle and this is achieved by combining periodic gene transcription and protein degradation (Tyers M. and Jorgensen P., 2000). Cyclin degradation depends on a sequence located near the amino-terminus called the destruction box, which targets cyclins to the ubiquitination-dependent proteolytic pathway (Morgan D.O., 1997)**.** A multiprotein complex, called anaphase-promoting complex (APC), contains an ubiquitin protein ligase that catalyses the ligation of several ubiquitin molecules to a cyclin, targeting it to the proteasomes for degradation. Cyclin degradation results in CDK inactivation.

Even when the CDK-cyclin complex is formed it still requires further activation to be fully functional. Several kinases and phosphatases are responsible for this regulation

and some of them are also effectors of the checkpoint cascade, linking the cell cycle machine to such important signalling pathways. CDK activation is also achieved by phosphorylation of a threonine residue of the CDK subunit operated by a CDK-activating kinase (CAK) or by removing an inhibitory phosphorylation operated by a CDK inhibitor (CKI) at a tyrosine residue.

1.1.4 Cell cycle control mechanisms: checkpoints

The definition of checkpoints in yeast was proposed for the first time in 1989 by Hartwell and Weinert (Hartwell L.H. and Weinert T.A., 1989) as a surveillance system monitoring the status of the cell and able to arrest the cycle if abnormalities are detected. Since then, many of the components involved in this process have been identified and assigned to four different categories: signals, sensors, transducers and receivers. Signals, which activate the checkpoint, are detected by sensors, that are responsible for monitoring cell size, status of the chromosomes or of the mitotic spindle. If any anomaly is present, the sensor will detect it and in turn activate a signalling cascade involving first the transducers and then the receivers. Transducers transmit and amplify the checkpoint signal to the receivers, which are the downstream targets of the checkpoint. Receivers often are proteins involved in regulating cell cycle progression and therefore responsible for arresting the cell cycle at different stages as summarised in Table 1.1.

These checkpoints help to coordinate cell cycle events, because if the order of events is incorrect then an incomplete set of genetic information is transmitted to the newly formed cell, which may lead to cancer in higher eukaryotes (Hunter T. *et al.*, 1994; Hall M. and Peters G., 1996).

Table 1.1 Yeast checkpoints

Name of checkpoint	Defect	Arrest point(s)
S-M	Unreplicated DNA	S
DNA damage	Damaged DNA	G1/S, G2/M
Intra-S damage	Damaged DNA	Extended S
Re-replication	Uncompleted M	S onset
Spindle assembly (SAC)	Defective spindle	Metaphase/anaphase
Spindle orientation (SOC)	Spindle misorientation	M
Morphogenesis	Disorganised actin cytoskeleton	G2

Cell size	Smaller cell size	G1
Meiotic prophase	Incomplete recombination	Meiosis I

The S-M and DNA damage checkpoints (Murakami H. and Nurse P., 2000; Nurse P., 1997) ensure the dependence of mitosis on completion of DNA replication and on undamaged DNA, respectively. If the cell was forced to undergo mitosis before DNA replication was complete or DNA damage repaired, it would pass on broken or incomplete sets of chromosomes to its daughters. The DNA replication checkpoint is also active in meiosis together with another checkpoint responsible for detecting defects in recombination during meiotic prophase (Murakami H. and Nurse P., 2000).

The Re-replication checkpoint is responsible for DNA replication taking place only once per cycle (Stern B. and Nurse P., 1996), whereas the Intra-S checkpoint slows down DNA synthesis in response to DNA damage caused by Ionising Radiation (IR), UV or the alkylating agent methylmethane sulfonate (MMS) (Carr A. and Caspari T., 2004).

The spindle orientation checkpoint (SOC) delays anaphase onset when the actin cytoskeleton is not properly oriented (Gachet Y. *et al.*, 2001) and the spindle assembly checkpoint (SAC) delays mitosis when the mitotic spindle is not properly organised or chromosomes are detached from it (Amon A., 1999). If the cell progresses into anaphase and starts to divide before all the chromosomes are aligned on the mitotic spindle, the chromosomes will not be allocated equally between the daughter cells.

The G1 checkpoint, which senses cell size, is called START or Restriction point (R) in mammals. If the cell is too small, the cycle will pause until the cell reaches the right size needed to progress through S phase. In addition, START is probably one of the most crucial points in cell cycle progression and at least two distinct control elements are known to be required for passing it and initiate the cell cycle: cyclin-dependent kinases (CDKs) which drive onset and progression of S phase and transcription factors that activate the G1/S specific wave of gene transcription whose expression is essential to complete DNA synthesis. The rest of this introduction will describe the role of CDKs, checkpoints and transcriptional regulation in governing the cell cycle in yeast and higher eukaryotes.

1.2 YEAST CELL CYCLE

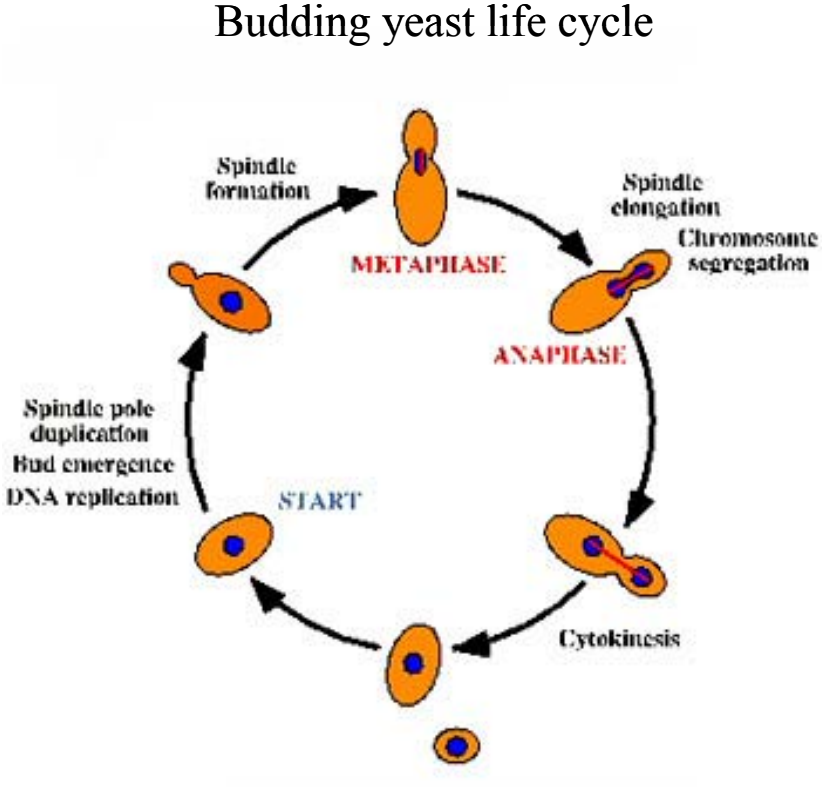
1.2.1 General overview

Yeasts are unicellular fungi – a large heterogeneous group of eukaryotic organisms. They are ideal for genetic studies of eukaryotic cell biology because they reproduce almost as rapidly as bacteria and have a genome size less than $1/100^{th}$ that of mammals. They are very well suited for identifying, cloning and characterizing the genes involved in controlling the cell cycle. The two major yeast model organisms are the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe*. Although the evolutionary lineages leading to budding and fission yeasts diverged many hundreds of millions of years ago (Heckman D.S. *et al.*, 2001; Sipiczki M., 2000), the two organisms have similar life cycles (Fig. 1.3).

S. cerevisiae has a long and distinguished experimental history; it was the first eukaryote to have its genome sequenced (Goffeau A. *et al.*, 1997) and, because of its well understood biology, it has also become a common model for molecular biology. *S. pombe* is evolutionary distant to budding yeast and has dozens of genes present in multicellular organisms (including disease genes) but not in budding yeast. A number of the key cell cycle regulatory molecules now being extensively studied in higher eukaryotes were first identified and characterized in fission yeast, and it is clear that many of the regulatory networks elucidated through genetic analysis in *S. pombe* are conserved in multicellular eukaryotes as well (Forsburg S.L. and Nurse P., 1991).

As in higher eukaryotes, the cell cycle processes of both yeasts can be divided into four discrete phases (G1, S, G2 and M) with the fission yeast cell cycle being characterised by relatively short G1 and S phases compared to budding yeast. Due to the differences in the length of cell cycle phases, the main control step takes place at the G1/S transition in budding yeast and at the G2/M transition in fission yeast (Fig. 1.3A and B).

Budding yeast life cycle



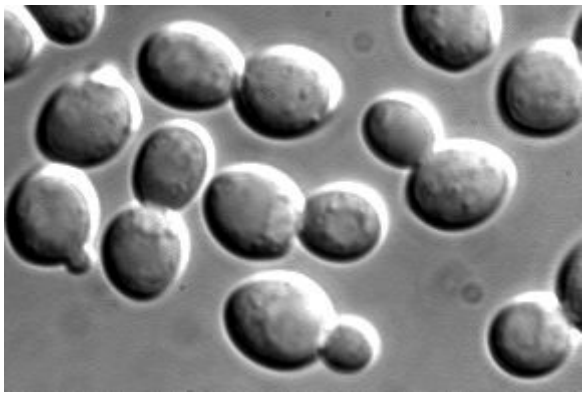


Fig. 1.3A Life cycles of both yeasts – Budding yeast.

Adapted from Vanderbilt Medical Center, Gould Lab webpage:
http://www.mc.vanderbilt.edu/vumcdept/cellbio/gould/html/news.html.
The image of budding yeast was taken from the University of Winnipeg, Simmons lab webpage:
http://io.uwinnipeg.ca/~simmons/2152web/2152/fungi2a.htm

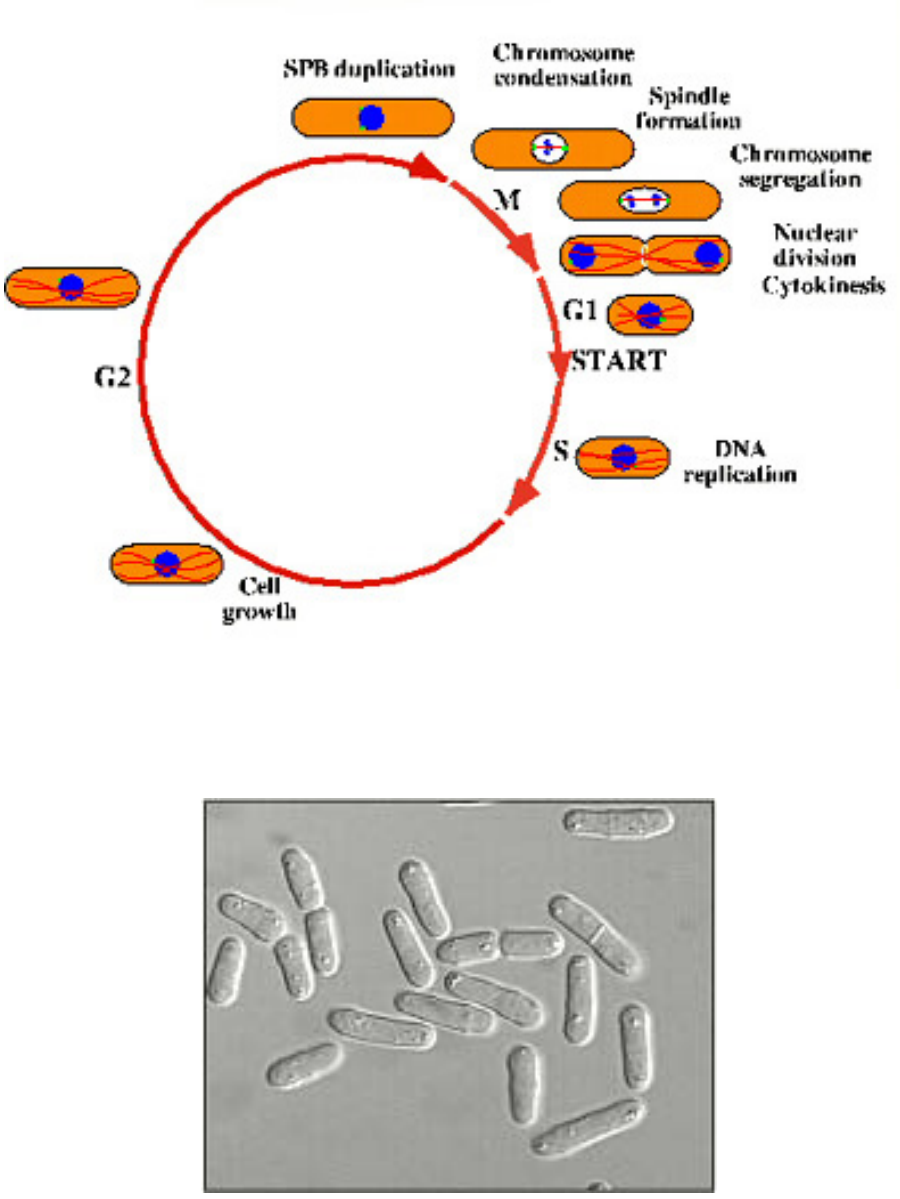


Fig. 1.3B Life cycles of both yeasts – Fission yeast.
Adapted from Vanderbilt Medical Center, Gould Lab webpage: http://www.mc.vanderbilt.edu/vumcdept/cellbio/gould/html/news.html.

Another important difference in their life cycle is that *S. cerevisiae* spends most of the time in a diploid state since the cells can mate under all growth conditions whereas *S. pombe* is haploid since mating is immediately followed by meiosis.

In both organisms cell cycle progression is driven by a single CDK and its cyclin partners. This section of the introduction will describe all the major players involved in the budding and fission yeasts' cell cycles and their counterparts in higher eukaryotes.

1.2.2 Budding yeast cell cycle

Five different CDKs are encoded by the *S. cerevisiae* genome. The most important one involved in cell cycle regulation is Cdc28p that acts in conjunction with nine different cyclins belonging to two families. These are three G1 cyclins, Cln1-3p, which function at START, and six B-type cyclins, Clb1-6p, which function in S, G2 and M phase (Nasmyth K., 1996).

The three G1 cyclins have different functions. Cln3p is needed to promote the activation of START-dependent transcription (regulated by the transcription factor complexes SBF and MBF as described later on), whereas Cln1p and Cln2p act downstream of Cln3p as effectors required to trigger START-regulated events such as budding and DNA replication. The genes coding for cyclins are transcribed in a periodic manner: the *CLN1*/*2* peak of expression coincides with START and is Cln3p-dependent. *CLN3* itself has a periodic behaviour with expression peaking at the M-G1 transition.

Similarly, of the six B-type cyclins, Clb5p/Clb6p governs entry into S phase whereas Clb1p-Clb4p regulates entry into mitosis. Clb5p and Clb2p play the major role and are supported in this by the other cyclin partners (Tyers M. and Jorgensen P., 2000). The peak of expression of their genes again coincides with the time their products are required: *CLB5*/*6* peak in late G1, *CLB3*/*4* in early G2 and *CLB1*/*2* in late G2.

Two proteins are responsible for inhibiting the kinase complexes, Sic1p and Far1p. Sic1p is specific for Clb-Cdc28p, while Far1p is specific for Cln-Cdc28p. In G1, the inhibitory effect of Sic1p (present in M and G1), together with proteolytic degradation of the cyclins operated by the APC complex, is responsible for switching off the kinase activity and allowing formation of the pre-replicative complexes (containing origin-recognition complex proteins or ORC and Cdc6p). Once these complexes are formed and bound to the DNA, Sic1p is inactivated by phosphorylation operated by Clb-Cdc28p itself, and kinase activity rises again allowing the MCM proteins to associate with the chromatin and DNA replication to start. Far1p acts as a repressor of *CLN1*/*2* expression, inhibiting their regulator Cln3p-Cdc28p and is inactivated by phosphorylation by Cln-

Cdc28p complexes similarly as Sic1p. If we imagine progressing through the budding yeast cell cycle then the series of events that we would witness is shown in Fig. 1.4. In early G1, Clb-Cdc28p activity is kept low because of the inhibitory role of Sic1p, low transcription of the Clb cyclin genes and degradation of the Clb proteins already synthesized. While cell size increases, Cln3p-Cdc28p activates *CLB5*/*6* transcription (mediated by the transcription factor MBF) causing S phase onset and *CLN1*/*2* transcription (mediated by the transcription factor SBF) leading to the accumulation of their products which results in Sic1p proteolysis, the factor responsible for Clb degradation. The cell is now replicating its DNA and the concentration of Clb cyclins starts to increase further once the Sic1p inhibitory effect has been removed (Nasmyth K., 1996).

Passage to G2 is characterised by reduction of Cln1p-Cdc28p and Cln2p-Cdc28p as a consequence of a drop in *CLN1*/*2* transcription caused by Clb-Cdc28p kinases. The CDK inhibitor Far1p also contributes to a reduction in *CLN1*/*2* transcription, probably by inhibition of Cln3p-Cdc28p activity. Sic1p is still degraded whereas the Clb cyclins are not targeted by the APC. The cell is now in G2 where the major players are the Clb-Cdc28p complexes and it can then enter M phase.

The end of mitosis is marked by an increase in *SIC1* transcription (mediated by the transcription factor Swi5p) and consequent inhibition of the Clb-Cdc28p complexes by Sic1p. At the same time the APC is activated (by Cdc20p and Cdh1p) and the Clb cyclins are degraded. At this stage the Cln-Cdc28p activity has the chance of increasing again thus driving the cell into a new cycle.

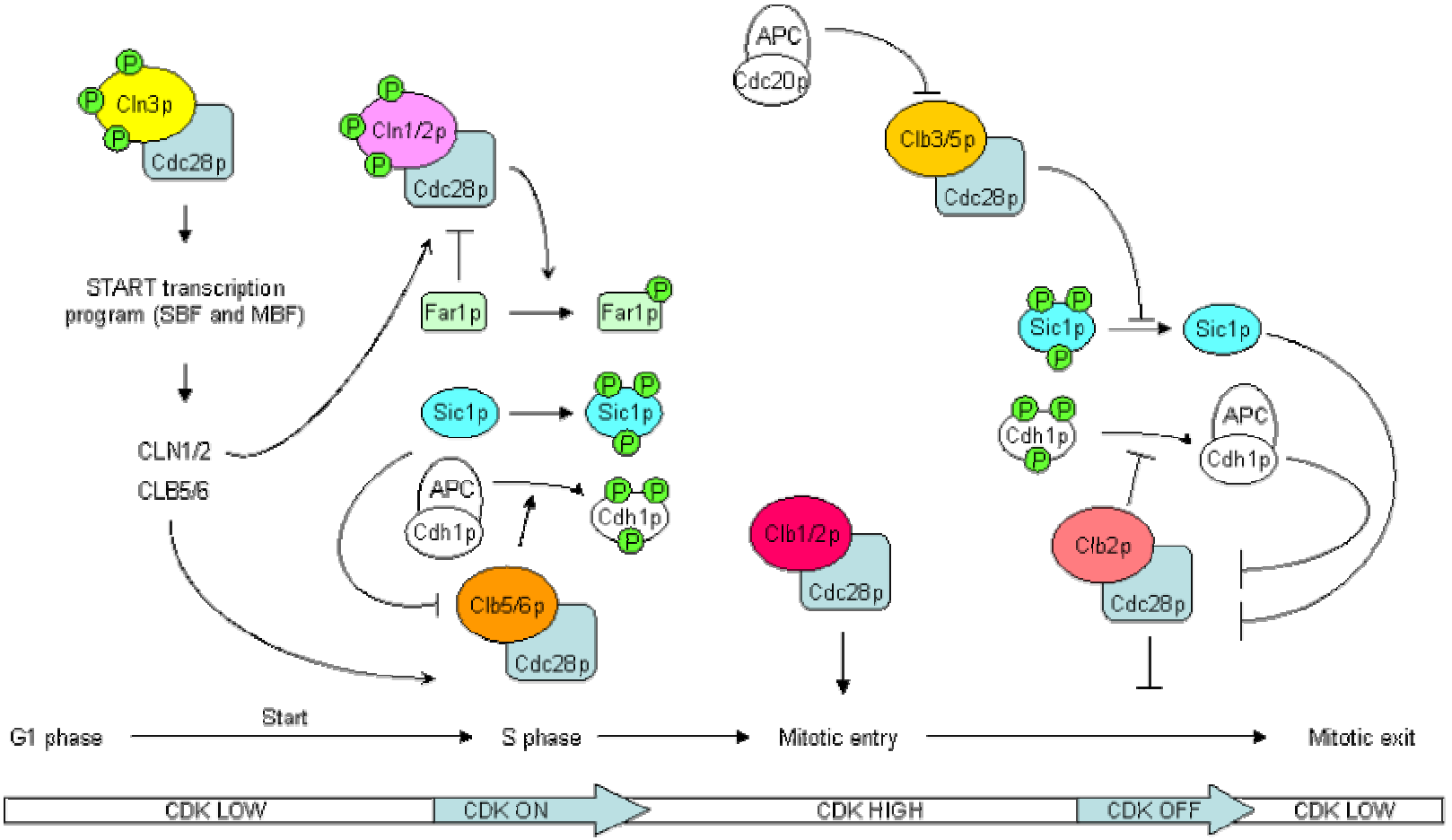


Fig. 1.4 **Central roles of CDK complexes in driving the budding yeast cell cycle.** Arrows indicate activation, blunt arrows inhibition and P stands for phosphorylation. Adapted from (Tyers M. and Jorgensen P., 2000).

1.2.3 Budding yeast checkpoints

The S-phase/DNA-damage checkpoint is a complex signalling pathway responsible for blocking or delaying cell cycle progression or DNA replication when the DNA is damaged. The proteins involved in this response are well conserved in eukaryotes and a unified model has been proposed for this checkpoint (Melo J. and Toczyski D., 2002). Several of the proteins involved in eukaryotic S-phase/DNA damage checkpoint can be found in the following table:

Table 1.2 Proteins involved in the S-phase/DNA damage checkpoint in yeast and mammals

Protein function	***S. cerevisiae***	***S. pombe***	**Mammals**
ATM/ATR kinases	Mec1 Tel1	Rad3 Tel1	ATR ATM
ATR-interacting proteins	Ddc2	Rad26	ATRIP
RFC-like proteins	Rad24	Rad17	Rad17
PCNA-like proteins	Ddc1 Rad17 Mec3	Rad9 Rad1 Hus1	Rad9 Rad1 Hus1
Mediators	Rad9 Mrc1	Crb2 Mrc1	BRCA1 Claspin
Effector kinases	Rad53 Chk1	Cds1 Chk1	Chk2 Chk1

Here I will briefly describe this model in budding yeast. The DNA damage can be detected by either Mec1p (related to human ATR) or Tel1p (related to human ATM). Mec1p responds to UV damage, double-strand breaks and replication blocks whereas Tel1p to double-strand breaks only. The sensing mechanism still remains unknown and hypotheses have been made that the cascade is activated by detection of abnormal DNA structures (Chrispell Forbes K. and Enoch T., 2000).

Before activating the signalling cascade that will result in a cell cycle arrest, these proteins need some additional factors to be loaded at the site of damage (Longhese M.P. *et al.*, 2003). These complexes in budding yeast are Rad24p and Ddc1p-Rad17p-Mec3p. The structural similarity of these proteins to others involved in DNA binding suggests a

model where Rad24p first binds to double-strand breaks then allows the loading of Ddc1p-Rad17p-Mec3p and then activates Mec1p kinase (bound to its partner Ddc2p). This kinase phosphorylates Rad9p (in response to DNA damage), Mrc1p (in response to DNA replication stress) or Chk1p. Phosphorylated Rad9p binds to Rad53p that inactivates Cdc5p, blocking cells in metaphase. Phosphorylated Chk1p in turn causes phosphorylation of Pds1p, an anaphase inhibitor, resulting again in a metaphase arrest (Murakami H. and Nurse P., 2000).

In budding yeast, the cell cycle arrests in metaphase-anaphase as a consequence of DNA damage whereas in fission yeast and higher eukaryotes the cycle is blocked at the G2-M transition. As described below (section 1.2.5) the G2/M arrest is achieved by an inhibitory phosphorylation of the CDK driving cell cycle progression. *CDC28* in *S. cerevisiae* is not involved in the DNA damage and DNA replication checkpoint control.

During mitosis the spindle checkpoint is responsible for halting the cycle when spindle defects or unattached chromosomes are detected. In budding yeast, the pathway is activated by the protein kinase Mps1p that phosphorylates Mad1p in a Bub1p, Bub3p and Mads2p dependent manner. Several of these proteins have been shown to interact with each other, and the most accepted model is that they form a complex that in turn interacts with Cdc20p (in a Mad2p-dependent manner), the activator of the APC. The result of this interaction is the inactivation of the APC ubiquitination activity. When APC is active, it is responsible for degradation of Pds1p, the inhibitor of sister-chromatid separation. Until Pds1p is present in the cell, sister-chromatid separation is not going to take place. Only when the spindle defects will be repaired, Mad2p will dissociate from Cdc20p leading to APC activation and Pds1p degradation, allowing the cycle to progress again (Amon A., 1999; Chrispell Forbes K. and Enoch T., 2000).

1.2.4 Fission yeast cell cycle

As in budding yeast, a single CDK encoded by *cdc2*, is directly involved in cell cycle regulation, controlling both onset of S-phase and M-phase (Fig. 1.5). Cdc2p activity is low in G1, moderate during S and G2 and high during M. Cdc2p is known to associate with four different cyclins: cdc13p, cig1p, cig2p and puc1p. Therefore, the cell cycle specific activity of cdc2p is determined by: (1) its associated cyclin, (2) inhibitory phosphorylation operated by the wee1p and mik1p kinases, which is removed by the

cdc25p phosphatase and (3) the CDK inhibitors rum1p (Sic1p analogue) and members of the Anaphase Promoting Complex (APC), ste9p and slp1p.

Cdc13p is an essential B-type cyclin in *S. pombe* (Fisher D.L. and Nurse P., 1996). During S phase, cdc13p begins to accumulate, due to a reduction in the concentration of its inhibitor rum1p (Benito J. *et al.*, 1998). Despite the formation of the cdc13p/cdc2p complex, its activity remains low in G2 phase because of the inhibition carried out by the tyrosine kinases wee1p and mik1p through phosphorylation of cdc2p. Activation during mitosis onset is due to an increase of the tyrosine phosphatase cdc25p (Russell P. and Nurse P., 1986), antagonist of protein kinases wee1p (Russell P. and Nurse P., 1987; Featherstone C. and Russell P., 1991) and mik1p (Lundgren K. *et al.*, 1991; Lee M.S. *et al.*, 1994). Cdc25p activity exceeds wee1p/mik1p activity causing dephosphorylation of cdc2p and therefore activation of the cdc13p/cdc2p complex. To exit mitosis, cdc13p is degraded by proteolysis and its removal results in very low level of CDK activity, preparing the cell for the next cycle.

The cig2p/cdc2p activity appears at the onset of S phase and falls after progression into S, due to cig2p proteolysis carried out by ubiquitination operated by the SCF (Skp1-Cullin-1-F-box) in G2 and M phases and by the APC in anaphase and G1 phase (Yamano H. *et al.*, 2004). Cig2p abundance is also controlled at the transcriptional level, with *cig2* gene expression peaking at G1/S.

Cig2p/cdc2p also appears to be involved in regulating the MBF transcription factor phosphorylating one of its components, res1p (Ayte J. *et al.*, 2001). This phosphorylation could either result in degradation of MBF via ubiquitination or in dissociation of res1p and res2p with loss of MBF-specific binding activity.

The functions of the other two complexes cig1p/cdc2p and puc1p/cdc2p are less understood. One of the roles of cig1p/cdc2p seems to be to relieve inhibitory effect of rum1p over cdc2p/cig2p and cdc2p/cdc13p, phosphorylating rum1p and therefore targeting it for degradation through the proteasome-dependent proteolytic pathway (Benito J. *et al.*, 1998). Puc1p/cdc2p has a specific G1 role and it is probably involved in the phosphorylation/inactivation of the CDK inhibitors, rum1p and ste9p (Martin-Castellanos C. *et al.*, 2000). Ste9p is responsible for degradation of the mitotic cyclins cdc13p and cig1p. CDK phosphorylation of ste9p promotes its degradation and prevents its interaction with the APC (Blanco M.A. *et al.*, 2000).

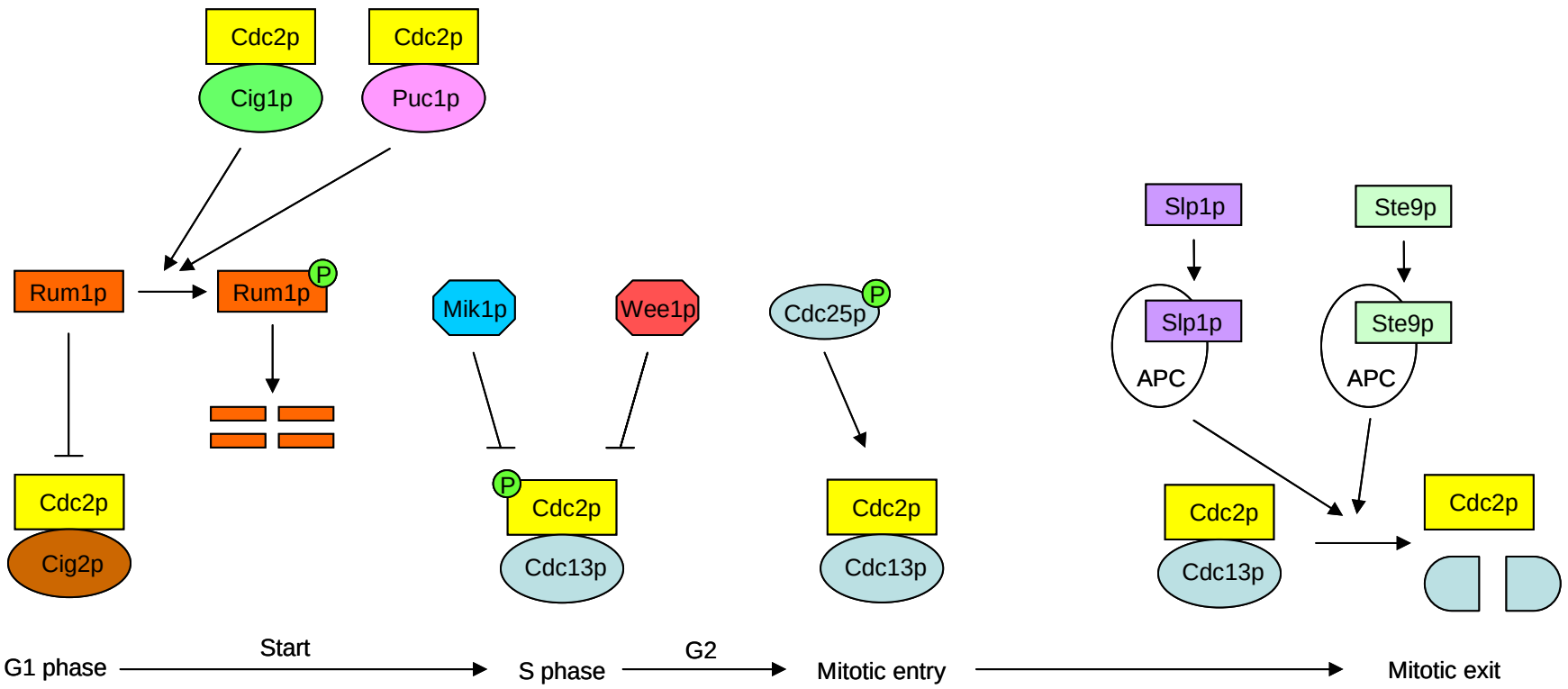


Fig. 1.5 Schematic representation of regulatory events during the fission yeast cell cycle.

Arrows indicate activation, blunt arrows indicate inhibition and P indicates phosphorylation.

1.2.5 Fission yeast checkpoints

If the cell cycle is perturbed by a block of DNA replication or by unrepaired DNA damage, then mitosis is prevented until the perturbation is corrected. This means that the replication and damage status of DNA is monitored and that this information is communicated by the regulatory pathway to mitotic cell cycle regulators leading to a block of mitosis. The components, identified so far, involved in this checkpoint in *S. pombe* are rad3p, rad26p, rad17p, rad1p, rad9p, hus1p, cut5p, crb2p, chk1p, mrc1p and cds1p.

First of all, the DNA damage needs to be sensed. Despite the fact that the sensing mechanism has not been fully elucidated, rad26p is known to be phosphorylated by rad3p in response to DNA damage as well as rad9p and hus1p. Rad9p, rad1p and hus1p physically bind to the damaged chromatin and rad17p appears to be required for the loading of this complex onto the DNA. Rad17p itself forms a complex with the Replication Factor C-like proteins (Green C.M. *et al.*, 2000).

Once DNA damage is detected, a signal needs to be sent to the cell cycle control machinery in order to stop cell cycle progression until the DNA replication can be restored. The link between the Rad checkpoint proteins and the CDK cdc2p is provided by two protein kinases, chk1p and cds1p. Chk1p is phosphorylated by crb2p (probably acting in a complex with cut5p) in response to DNA damage and acts by phosphorylating wee1p/mik1p (activating them) and cdc25p (inhibiting it). Mik1p appears to play a more important role than wee1p in the checkpoint response (Rhind N. and Russell P., 2001). This results in the inhibition of cdc2p/cdc13p complex and therefore in a temporary cell cycle G2-arrest that will last until the DNA is repaired or DNA replication restarted. Cds1p acts in the same way (via mrc1p activation) but seems to be specific for DNA replication blocks whereas chk1p is primarily involved in DNA damage checkpoints.

The spindle checkpoint operates in fission yeast in a very similar way to the one described for budding yeast (section 1.2.3). Interaction between mad2p and the APC slp1p causes stabilisation of cut2p and therefore prevents cut1p (bound to cut2p) from promoting chromatid separation (Chrispell Forbes K. and Enoch T., 2000).

1.3 CONSERVED CELL CYCLE MECHANISMS IN EUKARYOTES

CDKs are widely conserved among eukaryotes, from yeast to human. In both the budding and the fission yeast, a single CDK (Cdc28p and cdc2p, respectively) is responsible for catalysing all major cell cycle transitions. In higher eukaryotes, there has been an expansion in the number of CDKs that regulate the cell cycle, with up to five CDKs in humans and mice. This expansion allowed the specialisation of CDKs for particular cell cycle transitions in mammals: CDK4, CDK6 and CDK3 regulate G1 phase progression and entry into S phase; CDK2 is required for entry into S phase and DNA replication; and CDK1 is required for mitosis.

CDK4 and 6 are activated in response to growth factors and act in collaboration with cyclin D in promoting the transcription of proteins required for G1 and S phase. CDK1 and CDK2 appear to be the functional homologues to yeast cdc2p/Cdc28p and therefore are responsible for driving the cell cycle. CDK2 interacts with cyclin E at the beginning of S phase to trigger DNA synthesis, and then binds cyclin A throughout S phase. Both cyclin E/CDK2 and cyclin A/CDK2 activities are essential for initiation and completion of DNA replication and for ensuring that replication takes place only once in each cell cycle. Mitosis is then initiated by CDK1 in association with cyclin A and B. CDK3 is very closely related to CDK1 and 2 but its cyclin partner has not been identified yet. Similarly to what was described for yeasts (section 1.1.4), destruction of mitotic cyclins is essential for cell cycle progression and is again mediated by the APC.

CDK-cyclin complexes are negatively regulated by inhibitors belonging to two separate families: Ink4 family, which specifically inhibits CDK4 complexes and Cip/Kip family that targets CDK2-containing complexes.

Similarly to what has been described in yeast, CDKs are both positively and negatively regulated by phosphorylation. Inhibitory phosphorylation is catalysed by both a nuclear kinase, WEE1 and a cytoplasmic protein, MYT1. Inhibition is removed by members of the phosphatase family CDC25. In mammalian cells there are three forms of CDC25: CDC25A which specifically dephosphorylates cyclin E-CDK2 complexes early in the cycle and CDC25B/C which both function at the G2/M transition.

Cell cycle progression is tightly controlled in higher eukaryotes by a checkpoint system. If DNA is damaged, cell cycle will be delayed in G1 or G2 until the damage is repaired. ATM and ATR (Table 1.2) control the DNA damage response in mammalian

cells activating p53. Phosphorylation of p53 drives transcription of the CDK inhibitor p21, resulting in G1 arrest and transcription of 14-3-3 proteins (mammalian homologues to *S. pombe* Rad24 and Rad25), responsible for G2 arrest. ATM and ATR also phosphorylate the kinases Chk1 and Chk2. Chk2 activates p53 and both Chk1/Chk2 inhibit CDC25, causing a delay in mitosis onset. Both kinases also delay DNA replication targeting CDC25A for ubiquitin-degradation.

Such a high degree of conservation of cell cycle and checkpoint control in eukaryotic organisms underlies, once again, the fundamental importance of these pathways for the survival of the cell.

1.4 CELL CYCLE CONTROL OF GENE EXPRESSION

1.4.1 General overview

As discussed, progression through the cell cycle from yeast to human is the result of a series of events driven by CDK activity, including changes in gene expression. Periodic gene transcription of the cyclin and of many other genes involved in DNA replication is crucial for completion of the cycle. Changes in mRNA level of expression are normally due to transcription factors that specifically bind to promoter sequences of their target genes activating their transcription at certain stages of the cycle.

The first example of transcripts whose concentration was found to oscillate as a function of the cell cycle were the *Saccharomyces cerevisiae* histone mRNAs (Hereford L.M. *et al.*, 1981). Histone synthesis occurs specifically during S phase, which ensures that protein supply coincides with demand as chromosomal DNA is replicated. Other examples of cell cycle-regulated genes include those encoding most cyclins and many of the enzymes required for DNA synthesis. The expression of several genes fluctuates in a phase-dependent manner during the interphase period between mitoses.

Since their discovery, finding and studying genes whose transcription is cell cycle-regulated has been a challenging task for molecular biologists. By 1998, hundreds of cell cycle regulated genes were known in various organisms, but there were still many gaps: many genes and mechanisms important to the cell cycle regulation were unknown.

The development of the microarray technology (Duggan D.J. *et al.*, 1999; Brown P.O. and Botstein D., 1999; Ferea T.L. and Brown P.O., 1999; Lockhart D.J. and Winzeler E.A., 2000) which allows investigating the complete gene expression profile of an organism, made it possible to dissect the complex regulatory network that drives cell division and identify many of the proteins involved in its control (see section 1.5)**.**

1.4.2 Transcriptional cell cycle regulation in budding yeast

Three major types of cell cycle transcription factor complexes are known in *S. cerevisiae*: MBF / SBF, forkheads and Ace2p / Swi5p.

MBF (*Mlu*I binding factor) and SBF (SCB binding factor) operate at START (G1-S transition). The existence of such factors was first proposed when it was noticed that

most of the genes transiently transcribed during late G1 contained in their promoters one of two different UAS sequences, later named SCB element (Swi4/Swi6 cell cycle box - CACGAAA) and MCB element (*Mlu*I cell cycle box - ACGCGTNA) (Lowndes N.F. *et al.*, 1992). Further characterisation revealed that those sequences are specifically recognised by two different heterodimeric complexes that contain a common subunit, Swi6p, which is unable to bind DNA on its own and therefore requires an additional partner (Dirick L. *et al.*, 1992).

SBF is formed by Swi6p and Swi4p (Andrews B.J. and Herskowitz I., 1989) and preferentially binds to SCB elements (Primig M. *et al.*, 1992) whereas MBF contains Swi6p and Mbp1p (Koch C. *et al.*, 1993) and preferentially binds to MCB elements, although the distinction between SBF- and MBF-controlled genes is not possible based on the promoter sequences alone. Swi4p/Mbp1p DNA binding capacity is due to the presence of ankyrin repeats in the central part of those proteins, a characteristic feature of these transcription factor complexes conserved across yeast species (Koch C. *et al.*, 1993).

The precise mechanisms responsible for the regulation of MBF and SBF still remain unclear. Their activity is regulated at different levels: *SWI4* expression peaks at the M-G1 transition and Swi6p appears to be localised to the nucleus during S and to the cytoplasm during M but this is not enough to explain the G1 specific activation. Both transcription factors bind to the promoter region of their target genes during G1/S, before transcription is switched on (Koch C. *et al.*, 1996). A specific factor must therefore be responsible for their activation. Cln3p, in association with Cdc28p, has been proposed to be involved in such regulation (Koch C. *et al.*, 1996), and a recent study (Wijnen H. *et al.*, 2002) has proposed a similar model for Cln3p as for CDKs in mammals in activating E2F transcription factor. According to this view, Cln3 would interact with a Swi6p-specific inhibitor (still uncharacterised) therefore allowing MBF and SBF to activate transcription.

An alternative pathway for MBF regulation, functioning in parallel with the one just described, involves Stb1p; Cln-Cdc28p would in this case phosphorylate Stb1p causing its dissociation from MBF and therefore resulting in a down-regulation of MBF-dependent transcription (Costanzo M. *et al.*, 2003). In this case Stb1p would function as an activator of MBF-dependent transcription. Previous studies also demonstrated that the

SCB elements are not occupied by SBF in G2 and M and that the Clb1p-Clb4p kinase activity is responsible for this repression (Amon A. *et al.*, 1993).

Forkheads operate at the G2-M transition, and the characterisation of the components present in this complex was only completed very recently (Zhu G. *et al.*, 2000; Pic A. *et al.*, 2000; Kumar R. *et al.*, 2000; Koranda M. *et al.*, 2000; Jorgensen P. and Tyers M., 2000). It is formed by the forkhead-like transcription factors Fkh1p and Fkh2p, a MADS-box transcription factor Mcm1p and a transcriptional activator Ndd1p. It is the first example of an eukaryotic transcription factor complex containing both a MADS-box and a forkhead protein and the role of forkheads appears to be conserved from yeast to human (Alvarez B. *et al.*, 2001).

Fkh1p and Fkh2p seem to have overlapping roles in controlling G2-M transcription. Fkh2p cannot bind the DNA in the absence of Mcm1p (Kumar R. *et al.*, 2000). This two-element complex is bound to the promoter sequence of its targets throughout the cycle, and only the binding of an additional factor, Ndd1p, activates gene transcription (Koranda M. *et al.*, 2000; Jorgensen P. and Tyers M., 2000). Fkh1p is also able to bind various genes without any additional factor (Simon I. *et al.*, 2001). It is still unknown what activates members of the forkhead family of transcription factors; Fkh2p appears to be phosphorylated during the cell cycle (Pic A. *et al.*, 2000), and this might be an indication of a CDK regulation.

Ace2p and Swi5p are active in late M and early G1. Swi5p was first identified as specific activator of the HO gene, and Ace2p isolated afterwards because of its homology to Swi5p (Dohrmann P.R. *et al.*, 1992). There are many similarities between those two factors: their zinc-finger domains are 83% identical, *SWI5* and *ACE2* show a periodic behaviour peaking in M, their proteins are localised in the cytoplasm until M and then moved to the nucleus and their target genes peak in G1. Despite those similarities each factor activates a range of different genes, and this is due to a specific part of the proteins showing a promoter-specific activation (McBride H.J. *et al.*, 1999). Among the Swi5p targets is *SIC1* that is responsible for mitotic exit (Toyn J.H. *et al.*, 1997)

All the information collected in decades of research using traditional genetic approaches have been now integrated through the data collected using microarrays. The result has been the creation of a much more refined model for cell cycle control in *S. cerevisiae* (Futcher B., 2002). Two microarray studies of the budding yeast cell cycle have identified hundreds of periodic genes (Spellman P.T. *et al.*, 1998; Cho R.J. *et al.*,

1998) and assigned them to specific phases of the cycle (Spellman P.T. *et al.*, 1998). Cho *et al.* (1998) and Spellman *et al.* (1998) identified 416 and 800 periodic genes, respectively. Here I will discuss in more details the results of the Spellman study because they were analysed by the authors in more depth compared to the Cho dataset.

The 800 periodic genes identified by Spellman *et al.* were classified using a hierachical clustering algorithm (Eisen M.B. and Brown P.O., 1999) and their transcriptional regulation has been investigated by combining expression data together with transcription factor binding site search. In addition, the role of the cyclins Cln3p and Clb2p in controlling progression through the cycle was investigated looking at the expression profile of strains where Cln3p or Clb2p were induced. With this approach 363 genes were successfully assigned to a specific cell cycle stage, and the regulation understood to some degree for around 500 genes. The remaining 300 had smaller changes in expression and did not have good binding sites making it difficult to classify them. The results of this classification are summarised in the following table:

Table 1.3 Clusters of periodic genes in *S. cerevisiae* according to Spellman *et al.*, (1998)

Cluster	No of genes	Binding site	Regulator	Peak of expression	CLN3 effect	CLB2 effect
CLN2	119	ACGCGT	MBF, SBF	G1	I	R
Y	26	Unknown	Unknown	G1	-	-
FKS1	92	ACRMSAAA	SBF (MBF?)	G1	I	R
Histone	10	ATGCGAAR	Unknown	S	I	No effect
MET	20	AAACTGTGG	Met31p, Met32p	S	Some I	?
CLB2	35	MCM1 + SFF	Mcm1p + SFF	M	R	I
MCM	34	MCM1	Mcm1p	M/G1	R	I
SIC1	27	RRCCAGCR	Swi5p/Ace2p	M/G1	Maybe R	No effect

I: Induced, R: Repressed

The G1 clusters are:

1. “CLN2” which is regulated by MBF/SBF, includes mostly genes involved in DNA replication and the cyclin genes *CLN1*, *CLN2* and *CLB6* and appears to be induced by *CLN3* and repressed by *CLB2*.
2. “Y” which includes open reading frames located in Y elements at the chromosome ends. Nothing is known about their regulation or functional significance.
3. “FKS1” is represented by cell wall synthesis genes, the majority containing an SBF motif. They are induced by *CLN3* and repressed by *CLB2*.

The S and M clusters are:

1. “Histone” which appears to be induced by *CLN3*.
2. “MET” formed by genes involved in methionine biosynthesis, some of them induced by *CLN3* and most of them containing a binding site for Met31p/Met32p (Blaiseau P.L. *et al.*, 1997).
3. “CLB2” containing genes involved in mitosis and induced by *CLB2* and repressed by *CLN3*. Most of the components of this group (including *CLB1*, *CLB2*, *SWI5* and *BUD4*) have binding sites for Mcm1p/Fkh2p/Ndd1p.

The M/G1 clusters are:

1. “MCM” which includes genes involved in DNA replication (including the six MCM genes) that appear to be induced by *CLB2* and repressed by *CLN3*. The majority contain a binding site for Mcm1p and some an additional site for Mcm1p/Fkh2p/Ndd1p.
2. “SIC1” is regulated by Swi5p and Ace2p, some of the genes are repressed by *CLN3* whereas *CLB2* seems to have no effect on their regulation.

This was the first attempt made to clarify the transcriptional regulatory network in budding yeast. Several studies that followed have contributed to dissecting it in much more detail through the use of new experimental techniques such as crosslinking chromatin immunoprecipitation (ChIP) combined with DNA microarrays (ChIP-chip) (Simon I. *et al.*, 2001; Iyer V.R. *et al.*, 2001) and refined computational methods for the identification of transcription factor binding sites.

The model now proposed for the *S. cerevisiae* cell cycle is built around nine transcriptional regulators (Mbp1p, Swi4p, Swi6p, Mcm1p, Fkh2p, Fkh1p, Ndd1p, Swi5p and Ace2p) and two CDK complexes (Cln3p/Cdc28p and Clb2p/Cdc28p) as shown in Fig. 1.6.

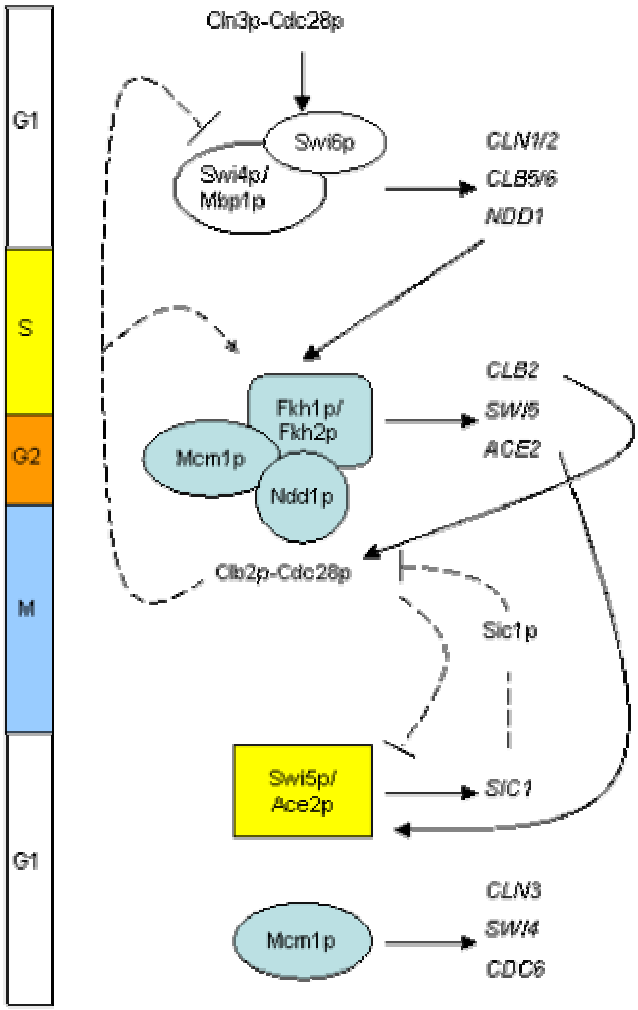


Fig 1.6. Serial regulation of gene transcription in *S. cerevisiae*.
Arrows indicate activation and blunt arrows inhibition. Adapted from (Tyers M. and Jorgensen P., 2000).

Simon *et al.* (2001) showed how a transcriptional regulator that is active in one specific cell cycle stage is responsible for the regulation of the transcriptional regulator that is active in the next stage of the cycle. Therefore, Cln3p-Cdc28p activates MBF

(Mbp1p/Swi6p) and SBF (Swi4p/Swi6p) in G1, which in turn switch on the transcription of the "CLN2" genes including the cyclin genes *CLN1/2* and *CLB5*/*6* (that drive the cycle through S phase) and the transcriptional activator *NDD1* (which is the limiting component of the complex activating G2/M transcription). SBF alone also participates in *CLB2* regulation at a later stage inducing its expression that as a result inhibits *CLN1/2* expression committing the cell to mitosis.

Mcm1p/Fkh2p/Ndd1p regulates the "CLB2" cluster in G2/M including *CLB2* that is responsible for the mitosis onset and *SWI5*/*ACE2* that regulate M/G1 genes. It is also responsible for the regulation of *CDC20* that encodes an APC activator and is therefore involved in mitosis exit. Mcm1p alone has been identified by McInerny *et al*. (1997) as a regulator of a subset of genes involved in pre-replication complex formation and in mating as well as *SWI4*, linking it to the regulation of SBF. This regulation does not involve forkheads. Similarly, Fkh1p alone was found to be able to bind several genes involved in chromatin structure and its regulation as well as the *CLB4* cyclin gene. Ace2p and Swi5p are regulators of M/G1 genes including *SIC1*, a specific Clb-Cdc28p inhibitor responsible for mitosis exit. Ace2p, Swi5p and Mcm1p are all able to bind *CLN3* and activate it. Exit from mitosis and increase of Cln3p concentration will prepare the cell for a new cycle.

It is now clear that partially redundant transcription factors have an overlapping set of genes that they control as well as specific ones for each factor (Simon I. *et al.*, 2001). This has been definitely shown by Iyer *et al.* (2001) for SBF and MBF, where SBF controls genes involved mainly in membrane/cell-wall formation and bud growth, and MBF controls genes for DNA replication.

In budding yeast, the transcriptional regulatory network that governs the cell cycle can be considered as a cycle itself, where all the transcription factors involved are progressively activated in a cascade. Some of those factors also appear to be able to inhibit their activators with a negative feedback resulting in a transition to the next cell cycle stage.

1.4.3 Transcriptional cell cycle regulation in fission yeast

More than thirty periodically expressed genes have been identified in fission yeast but far less transcription factors are known. In *S. pombe* the G1-S transition is characterised by a

transient peak of expression of a subset of genes: *cig2* (Connolly T. and Beach D., 1994), as well as *cdc18* (Kelly T.J. *et al.*, 1993), *cdc22* (Gordon C. and Fantes P., 1986), *cdt1* (Hofmann J.F. and Beach D., 1994), *cdt2* (Hofmann J.F. and Beach D., 1994; Obara-Ishihara T. and Okayama H., 1994), *mik1* (Christensen P.U. *et al.*, 2000; Ng S.S. *et al.*, 2001; Baber-Furnari B.A. *et al.*, 2000), *rad21* (Birkenbihl R.P. and Subramani S., 1995), *ste6* (Maqbool Z. *et al.*, 2003), *ssb1* (Parker A.E. *et al.*, 1997), *ste9* (Tournier S. and Millar J.B., 2000) and *suc22* (Harris P. *et al.*, 1996; Fernandez Sarabia M.J. *et al.*, 1993). Promoter sequences of those genes all contain MCB (*Mlu*I cell cycle box - ACGCGTNA) elements that are the binding sites for a transcription factor complex named DSC1 (DNA synthesis control) or MBF (*Mlu*I binding factor).

This complex contains the products of *cdc10*, *res1* and *res2* genes and is activated by rep2p (Lowndes N.F. *et al.*, 1992; Caligiuri M. and Beach D., 1993; Tanaka K. *et al.*, 1992; Zhu Y. *et al.*, 1994; Miyamoto M. *et al.*, 1994; Nakashima N. *et al.*, 1995; Zhu Y. *et al.*, 1997). Cdc10p, res1p and res2p are homologous to *S. cerevisiae* Swi6p, Swi4p and Mbp1. Their functional domains appear to be highly conserved: the carboxy-terminal region in cdc10p/Swi6p, the amino-terminal region in res1p/res2p/Swi4p/Mbp1p (representing the DNA binding domain) and several copies of an ankyrin motif present in the central part of all of these proteins (Ayte J. *et al.*, 1995; Sturm S. and Okayama H., 1996).

Cdc10p does not bind DNA directly and relies on res1p and res2p for this activity (Whitehall S. *et al.*, 1999). Cdc10p can form a heteromeric complex with res1p (Caligiuri M. and Beach D., 1993) or res2p (Zhu Y. *et al.*, 1994). Since both cdc10p/res1p and cdc10p/res2p recognise MCB elements, an overlapping role has been originally suggested for them, with cdc10p/res1p acting mainly at START and cdc10p/res2p controlling pre-meiotic S phase (Miyamoto M. *et al.*, 1994).

Different Res-protein domains appear to be required for cell cycle-dependent transcription: in the case of res2p, the extreme carboxy-terminal (C) and a domain including the amino-terminal ankyrin repeat are important for transcription activation whereas the C-terminus alone is responsible for rep2p binding (Whitehall S. *et al.*, 1999). The most important domain in res1p corresponds to a region overlapping the amino-terminal ankyrin repeat, and no specific function seems to be associated with the C-terminus (Whitehall S. *et al.*, 1999).

Whitehall *et al*. (1999) have modified the previous model proving that cdc10p, res1p and res2p are present in the MBF complex throughout the entire cell cycle in an inactive state (Whitehall S. *et al.*, 1999) and that transcriptional activation occurs when rep2p, a zinc-finger protein, specifically binds res2p (Nakashima N. *et al.*, 1995; Tahara S. *et al.*, 1998). Another zinc-finger protein, rep1p, was found to be involved in controlling initiation of pre-meiotic DNA synthesis interacting with cdc10p/res1p and cdc10p/res2p (Sugiyama A. *et al.*, 1994) and its role has always been thought to be meiosis specific until recently when White *et al.* (White S. *et al.*, 2001) suggested that rep1p might have an equally important role to rep2 in mitosis. This hypothesis requires further investigation.

As explained in section 1.3.2, in budding yeast the CDK Cln3p/Cdc28p is responsible for regulating the MBF/SBF complexes. Its fission yeast counterpart, cdc2p, and its cyclin partners surprisingly appear not to be involved in such regulation (Baum B. *et al.*, 1997). Tanaka *et al*. (Tanaka K. and Okayama H., 2000) showed that cdc10p/res2p is instead activated by the Pcl-like cyclin pas1p and its kinase partner pef1p, proposing a link between transcriptional regulation and progression through the cell cycle in *S. pombe*. Such a link between a CDK activity and MBF regulation has never been found in fission yeast and this hypothesis requires more evidence.

Another transcription factor named sep1p, belonging to the forkhead family along with *S. cerevisiae* Fkh1p and Fkh2p, has been identified during the characterisation of a mutant showing severe septation defects (Ribar B. *et al.*, 1997; Ribar B. *et al.*, 1999)**.** It is responsible for periodic transcription of *cdc15* whose product is involved in the formation of the actin ring during mitosis (Zilahi E. *et al.*, 2000) and whose peak of expression can be detected just before septation (Fankhauser C. *et al.*, 1995). Deletion of *sep1* results in loss of *cdc15* periodicity. No other sep1p targets have been identified so far, nor has the mechanism of its activity or regulation been clarified. The only other fission yeast forkhead transcription factor reported in the literature is mei4p, which has a meiosis specific role (Horie S. *et al.*, 1998).

The promoter sequence GNAACg/a that confers periodicity to cdc15 has been identified and shown to be present in a subset of genes all showing a peak of expression at the M/G1 transition (Anderson M. *et al.*, 2002). Those genes are: *cdc19*, *dmf1*, *fin1*, *plo1*, *ppb1*, *spo12* and *sid2*. The plo1p kinase is involved in controlling the expression of those genes and probably acts indirectly through a still uncharacterised transcription

factor named PBF by Anderson *et al*. (2002). The homologous *S. cerevisiae* genes *CDC5* (*plo1*), *DBF2* (*sid2*), *SPO12* (*spo12*) and *MCM2* (*cdc19*) that also show a periodic behaviour in budding yeast are under the control of Mcm1p/Fkh2p/Ndd1p forkhead-type transcription factor complex. The *S. pombe* forkhead sep1p protein might not be part of the putative PBF transcription factor (Anderson M. *et al.*, 2002).

The study of another mutant showing a cell separation defect has led to the identification of another transcription factor in *S. pombe*. Cells lacking the *eng1* gene are able to form a septum between mother and daughter cell at the end of mitosis but they cannot separate (Martin-Cuadrado A.B. *et al.*, 2003). Eng1p has a β-glucanase activity responsible for digesting the primary septum that allows cell separation as result of cytokinesis. In wild type cells, *eng1* mRNA shows a periodic behaviour peaking before septation. This periodicity is abolished in mutants lacking the *ace2* gene. *S. pombe* ace2p has been identified based on similarity with *S. cerevisiae* Ace2p, which in budding yeast acts as transcription factor responsible for the activation of the M/G1 wave of transcription, also including the β-glucanase *ENG1*.

1.4.4 Transcriptional cell cycle regulation in humans

Since the sequence and function of the MBF factor components are conserved between the two distantly related yeasts, homologues might also be expected in higher eukaryotes. However, only a functional homologue to MBF could be identified in metazoans, the E2F protein family, whose members show very little homology at the protein level with Swi4p/Swi6p. Functional homology relies on the fact that both complexes regulate gene transcription at the G1/S boundary, they bind a core consensus sequence CGCG, and they have a similar folding pattern (Breeden L.L., 2003).

E2F acts as a heterodimer binding to a member of the pocket protein family. In mammals at least six E2F (E2F1 to E2F6) proteins have been identified, together with three pocket proteins, pRB – product of the *retinoblastoma susceptibility* gene, p107 and p130. Different forms of E2F preferentially bind to certain pocket proteins and different heterodimers are formed at different stages of the cycle (Stevaux O. and Dyson N.J., 2002). E2F binds to p130 in G0, to p107 and pRB in early G1 and is found free in S phase. When bound to E2F, the pocket proteins have an inhibitory role. The inhibition needs to be removed for the cell to progress through the cycle.

Most changes in the E2F state take place at the G1/S transition, called 'Restriction point' (R) in mammalian cells. E2F is regulated by pRB which acts as a transcriptional inactivator. pRB is a negative regulator of cellular proliferation and it is functionally inactived in the majority of human tumors (Wang J.Y. *et al.*, 1994; Harbour J.W. and Dean D.C., 2000; Nevins J.R., 2001; Sherr C.J. and McCormick F., 2002). At START the CDK4/cyclin D and CDK6/cyclin E complexes phosphorylate the pocket protein bound to E2F causing its dissociation from the transcription factor that can therefore activate expression of its targets (McGowan C.H., 2003). One of its targets is the gene encoding cyclin E that forms complexes with CDK2, which are responsible for further phosphorylating the pocket proteins and keeping them in an inactive state. At the end of S phase, cyclin A substitutes for cyclin E in forming a complex with CDK2, resulting in E2F phosphorylation and downregulation of the E2F-dependent transcription.

E2F targets have recently been shown to be involved in several processes, beyond G1/S transition. They range from genes involved in triggering S-phase (Cyclin E and CDK2), in assembling the pre-replicative complex (ORC and MCM proteins) and in DNA synthesis (polymerase α) as well as genes involved in DNA repair (Stevaux O. and Dyson N.J., 2002). Surprisingly, E2F seems to also regulate genes with a mitotic role (Ishida S. *et al.*, 2001) and is itself a target of ATM (responsible for its activation), involved in response to DNA damage as described in session 1.2.5.

Similarly, another function that appears to be conserved from yeast to mammals is the one exerted by the forkhead family of transcription factors or forkhead box factors (FOX) in controlling mitosis. Within the large family of FOX factors in humans, the FOXO group is responsible for many fundamental cell processes, including cell cycle progression and DNA repair. This group includes FOXO1 (formerly known as FKHR), FOXO3a (FKHR-L1) and FOXO4 (AFX) (Burgering B.M. and Kops G.J., 2002). Their activity is regulated by the phosphatidyl-inositol 3OH kinase (PI(3)K)/protein kinase B (PKB) pathway. PKB phosphorylates FOXO members inhibiting their transcriptional activity at the G1 phase of the cycle. Beyond this stage PKB activity goes down allowing FOXO members activities to increase, in correspondence with G2/M, allowing mitosis execution (Alvarez B. *et al.*, 2001). The FOXO targets identified so far (based on homology with the budding yeast identified targets) are cyclin B and PLK, which are both involved in mitosis progression (Glover D.M. *et al.*, 1998).

Several microarray studies investigating gene expression in the human cell cycle have been published so far (Whitfield M.L. *et al.*, 2002; Cho R.J. *et al.*, 2001; Crawford D.F. and Piwnica-Worms H., 2001; Iyer V.R. *et al.*, 1999), as well as in mouse cells (Ishida S. *et al.*, 2001).

Whitfield *et al*. (2002) identified 874 periodic genes in the HeLa cancer cell line and have clustered them according to the cell cycle phase coincident with their peak of expression**.** They have also shown a good correlation between the timing of gene expression and the biological function of genes, as summarised in the following table.

Table 1.4 Periodic gene clusters in HeLa cells according to Whitfield *et al*. (2002)

Cell cycle phase	Cluster	Biological function	Genes
G1	Early DNA replication	DNA replication	CDC6, MCM2-6, DNA pol δ3, ORC1L
		DNA packaging	CHAF1A, CHAF1B, PCNA
		DNA repair	MSH2, FEN1, PCNA
		Cell cycle control	Cyclin E1, E2F1, CDC25A, Cyclin E2
S	Late DNA replication	Nucleotide metabolism	
		DNA replication	DNA polα and θ
		DNA repair and DNA recombination	RAD54, RAD 51
	Histone	Chromatin assembly/disassembly	H2A, H2B, H4, H1, SLBP, NPAT
G2	Tubulin		α-/β-tubulins (TUBA1-3, TUBB, TUBB2), BUB3
G2/M		Mitotic spindle organisation	Kinesins (KNSL1-6), TTK, Cyclin A2, CDC2, CKS1, Cylcin F, ESP1
G2/M		Mitotic spindle checkpoint	Cyclin B2, BUB1, BUB1B, CDC20, CENPE
		Centrosome duplication	STK15, PLK1, NEK2
M/G1		Actin cytoskeleton remodelling	KRAS2
		Cell adhesion	p120, vinculin
		Chromosome architecture	RAD21

Many of the periodic genes identified in this study have been previously reported as cyclic by Ishida *et al*. (2001) (Ishida S. *et al.*, 2001) and by Crawford and Piwnica-Worms (2001) (Crawford D.F. and Piwnica-Worms H., 2001). The Ishida study was performed on mouse embryo fibroblasts and aimed to identify most of the E2F-dependent genes, whereas Crawford and Piwnica-Worms studied the S-G2 specific transcription in HeLa cells.

Comparison with the other two studies by Iyer *et al*. (1999) and Cho *et al*. (2001) did not give equally good results. In the case of Iyer *et al*., the authors themselves recognised that most of the gene changes on the arrays are due to a wound-healing response of the fibroblasts, making it difficult to clearly identify the cell cycle response. Cho *et al.,* (2001) identified 731 periodic genes using human fibroblasts synchronised with a double-thymidine block protocol (Cho R.J. *et al.*, 2001). The overlap with the Whitfield study is only 96 genes, and this is probably due to the many differences in terms of cell lineage, microarray technology and analysis methods between the two studies and also the different degree of synchrony achieved in the different experiments.

1.5 CELL CYCLE RESEARCH USING DNA MICROARRAYS

With an increasing number of organisms now being sequenced, new methodologies are required to make sense of this huge amount of information and to complement more traditional 'gene by gene' approaches. DNA microarrays are increasingly popular tools for studying global gene expression, and the model organisms have been the first for which comprehensive genome-wide surveys of cell cycle regulated genes have been performed, including yeast (Cho R.J. *et al.*, 1998; Spellman P.T. *et al.*, 1998), bacteria (Laub M.T. *et al.*, 2000; Wei Y. *et al.*, 2001), protozoa (Bozdech Z. *et al.*, 2003), plants (Hennig L. *et al.*, 2003; Menges M. *et al.*, 2002) and humans (Cho R.J. *et al.*, 2001; Whitfield M.L. *et al.*, 2002).

In this study spotted arrays were used in two-color hybridisation experiments, as pioneered by Pat Brown and colleagues at Stanford University. Briefly, arrays of thousands of discrete DNA sequences are printed on a glass microscope slide using a robotic "arrayer". To compare the relative abundance of each of these gene sequences in two DNA or RNA samples (for example a 'test' cell state and a 'reference' cell state) the two samples are first labelled using different fluorescent dyes (say, a red dye and a green dye). They are then mixed and hybridised to the arrayed DNA spots. After hybridisation, fluorescence measurements are made with a confocal laser scanner that illuminates each DNA spot and measures fluorescence for each dye separately; these measurements are used to determine the ratio and, in turn, the relative abundance of the sequence of two mRNA or DNA samples for each gene. Computational methods are then applied to the microarray data to identify co-regulated clusters of genes, and motif-finding algorithms are used to find promoter elements characteristic of each cluster (Brown P.O. and Botstein D., 1999; Duggan D.J. *et al.*, 1999).

The research I have performed is inspired by that done with budding yeast (Cho R.J. *et al.*, 1998; Spellman P.T. *et al.*, 1998), as described in the Introduction, section 1.4.2. My work has provided a global picture of the expression profile during the fission yeast cell cycle. This will serve as a framework for the identification and further characterization of previously unknown genes that play a key role in the regulation of the cell cycle. In addition, since fission yeast and budding yeast are only distantly related, complementary studies in the two yeasts have proven to be very fruitful in understanding cell cycle regulation (Forsburg S.L., 1999). Both differences and similarities in the

regulation of gene expression between the two yeasts will be insightful to learn more about transcriptional control of the cell cycle and gene function in eukaryotes.