5. CONSERVATION OF CELL CYCLE REGULATED GENE EXPRESSION

This chapter will focus on the conservation of periodic gene transcription through evolution. The fission yeast cell cycle program of gene expression compiled for this thesis has been compared with the one described for budding yeast (Cho R.J. *et al.*, 1998), (Spellman P.T. *et al.*, 1998) in order to identify a core set of genes whose function and regulation are conserved across yeast species. A similar comparison has been done, to a lesser extent, with the human cell cycle (Cho R.J. *et al.*, 2001; Whitfield M.L. *et al.*, 2002).

5.1 Fission yeast and budding yeast: what is conserved?

Fission yeast and budding yeast are only distantly related and they represent a good complementary system for the identification of conserved mechanisms among eukaryotes.

Two budding yeast microarray studies of cell cycle gene expression identified ~ 400 (Cho R.J. *et al.*, 1998) and 800 (Spellman P.T. *et al.*, 1998) periodic genes. For budding yeast two gene lists were used: the complete Spellman list of periodic genes and a second list of 301 genes, which were periodic in both the Spellman and Cho datasets. For fission yeast, the analysis was performed using both the complete list of 407 periodic genes as well as the list of 136 'high amplitude' periodic genes. Homologous *S. cerevisiae* genes were identified using a curated ortholog table including a total number of 2981 genes (Materials and methods, section 2.6.6). For each list, only genes with an ortholog were included in the analysis (Table 5.1 – the numbers in parentheses refer to the number of ortholog genes within each list).

The overlap of periodic genes was surprisingly small yet statistically highly significant. Table 5.1 shows the *P*-values associated with the different gene list comparisons, together with the sizes of the overlaps expected by chance for the same comparisons (calculated based on the size of the gene lists used and the total number of *S. pombe* genes with a *S. cerevisiae* ortholog as explained in section 2.6.6). The best results were obtained for the comparisons where the 'high amplitude' gene lists were used (Table 5.1- comparisons in bold), against the all periodic Spellman genes (P = 4.3e-32) and

against the Spellman/Cho overlap (*P* = 2.33e-35) respectively. The number of conserved cell cycle regulated genes increased when the lists of all periodic *S. pombe* genes and all periodic Spellman genes were compared but the significance was drastically reduced (*P* = 2.0e-22). Among the 'high amplitude' genes, 87 had a budding yeast ortholog and 42 of them were also periodic in both budding yeast studies (Table 5.2). This is the most conservative list of periodic/ortholog genes in the two yeasts. The additional 13 genes that were periodic in fission yeast ('high amplitude' *S. pombe* genes) and in the Spellman study are also listed in Table 5.2.

In addition, the five clusters of *S. pombe* cell cycle regulated genes were compared with the five clusters of *S. cerevisiae* periodic genes, as identified by Spellman *et al.* (1998); only the *S. pombe* 'high amplitude' genes were used for this comparison. Despite the many differences in the life cycles of the two yeasts, a significant number of orthologs was regulated at corresponding cell cycle phases in the two model organisms. Fission yeast clusters 1 to 4 overlapped significantly with the budding yeast G2/M (*P* = 1.0e-08), G1 (*P* = 8.3e-10), S (*P* = 5.4e-17) and S/G2 (*P* = 1.6e-03) phase genes, respectively. *S. pombe* cluster 4 and *S. cerevisiae* S/G2 genes showed a weaker overlap, probably as consequence of the fact that this cluster in fission yeast represents quite an heterogeneous group of genes, expressed over a relatively long G2 phase. *S. cerevisiae* M/G1 cluster showed a weak overlap with *S. pombe* cluster 2.

Table 5.1 Overlap of periodic genes between *S. pombe* and *S. cerevisiae*

***S. pombe* (ortholog number)**	***S. cerevisiae* (ortholog number)**	**Ortholog number periodic in both yeasts**	**Expected overlap in gene number**	***P*-value**
All periodic genes (252)	Spellman (322)	81	27.2	2.0e-22
All periodic genes (252)	Cho and Spellman (137)	55	11.6	7.7e-26
'High amplitude' periodic (87)	**Spellman (322)**	**54**	**9.4**	**4.3e-32**
'High amplitude' periodic (87)	**Cho and Spellman (137)**	**42**	**4.0**	**2.33-35**
'High amplitude' cluster 1 (28)	Spellman: G1 (127)	9	1.2	1.2e-06
'High amplitude' cluster 1 (28)	Spellman: G2/M (74)	9	0.7	1.0e-08

'High amplitude' cluster 2 (36)	Spellman: G1 (127)	13	1.5	8.3e-10
'High amplitude' cluster 2 (36)	Spellman: M/G1 (33)	3	0.4	6.9e-03
'High amplitude' cluster 3 (14)	Spellman: S (40)	10	0.2	5.4e-17
'High amplitude' cluster 4 (4)	Spellman: S/G2 (50)	2	0.1	1.6e-03

Table 5.2 Core set of periodically expressed genes in fission and budding yeasts

***S. pombe* ortholog**	***S. cerevisiae* ortholog**	**Function**
DNA replication		
pol1	*POL1*	DNA polymerase α
cdc20	*POL2*	DNA polymerase ε
ssb1	*RFA1*	Single-stranded DNA-binding protein
cdc18	*CDC6*	Regulator of DNA replication initiation
mrc1	*MRC1*	DNA replication checkpoint protein
cdc22	*RNR1*	Ribonucleotide reductase
psm3	*SMC3*	Cohesin
rad21	*MCD1*	Cohesin
pht1	*HTZ1*	Histone variant
hta1, hta2	*HTA1, HTA2*	Histone H2A
htb1	*HTB1, HTB2*	Histone H2B
hht1, hht2, hht3	*HHT1, HHT2*	Histone H3
hhf1, hhf2, hhf3	*HHF1, HHF2*	Histone H4
Mitosis and cell division		
plo1	*CDC5*	Polo kinase
ark1	*IPL1*	Aurora kinase
fin1	*KIN3*	NimA kinase
cut2	*PDS1*	Securin (sister chromatid separation)[a]
slp1	*CDC20*	Activator of anaphase promoting complex
wis3	*SPO12*	Putative cell-cycle regulator
klp5, klp6, klp8	*KAR3, KIP1*	Kinesin microtubule motor[b]
mob1	*MOB1*	Protein involved in mitotic exit/septation
sid2	*DBF2*	Kinase involved in mitotic exit/septation
myo3	*MYO1*	Myosin II heavy chain
mid2	*BUD4*	Protein involved in cytokinesis
ace2	*ACE2*	Transcription factor
imp2	*HOF1*	Protein involved in cell division
chs2	*CHS2*	Protein involved in septum formation
eng1	*DSE4*	Glucanase for cell separation
mac1	*TOS7*	Putative role in cell separation (*S. pombe*)
Others		
rum1	*SIC1*	Inhibitor of cyclin-dependent kinase[a]
mik1	*SWE1*	Kinase inhibiting cyclin-dependent kinase
cig2	*CLB1-CLB6*	B-type cyclin[b]
msh6	*MSH6*	Mismatch-repair protein

rhp51	*RAD51*	DNA repair protein
SPBC32F12.10	*PGM1*	Phosphoglucomutase, carbohydrate metabolism
SPAP14E8.02	*TOS4*	Unknown function
Additional conserved genes from the comparison between 'high amplitude' periodic and Spellman		
cfh1, cfh2, cfh3, cfh4	*SKT5*	Protoplast regeneration and killer toxin resistance protein
psu1	*SIM1*	Essential protein required for cell wall integrity, member of the SUN protein family
SPCC1322.04	*UGP1*	UDP-glucose pyrophosphorylase (UTP-glucose-1-P uridylyltransferase)
SPAP7G5.06	*GAP1*	Protein with high similarity to general amino acid permease
exg1	*EXG1*	Exo-beta-1,3-glucanase
SPBC1198.07	*DFG5*	Putative glycosylphosphatidylinositol (GPI)-anchored protein
SPBC19C7.04	YMR295C	Unknown function
SPAC19B12.02	*GAS1*	Protein with high similarity to 1,3-beta-glucanosyltransferase
SPCC338.12	*PBI2*	Member of the subtilisin N-terminal region containing family
psy1	*SSO1*	Syntaxin-like protein component of the plasma membrane docking/fusion complex

[a] Proteins encoded by these genes show little sequence homology but are functional homologs
[b] Protein families with various functions

5.2 Conserved genes across yeast species and their function

Among the genes whose periodic behaviour has been conserved in the two yeasts, two subsets of genes were identified: a group involved in DNA replication and another one involved in mitosis and/or cell division. The histone genes are a prominent group among those one whose function is linked to DNA replication. Histone genes show a periodic behaviour that is highly conserved across eukaryotes (Plumb M. *et al.*, 1983; Hereford L.M. *et al.*, 1981, Aves S.J. *et al.*, 1985).

The process of chromosomal DNA replication in yeast involves many components whose function appears conserved through evolution. The initiator is the origin recognition complex (ORC), which binds to replication origins and is required for their firing. The next step requires the binding of an ORC-interacting factor (*S. cerevisiae* Cdc6p/*S. pombe* cdc18p) and of the mini-chromosome maintenance (MCM) protein complex, responsible for the expansion of the unwound DNA at the replication origin. Once the DNA is unwound, short RNA-DNA primers are synthesised by the DNA polymerase α (Pol1p/pol1p) and then elongated by the DNA polymerase ε

(Pol2p/cdc20p). The process is then terminated by the ligation of all DNA fragments. The conservation of this fundamental process is reflected by the periodic behaviour of *CDC6*/*cdc18* and of the DNA polymerase α and ε coding genes, observed in both yeasts. Other genes whose function is directly involved in DNA replication and that were found to be cell-cycle regulated in both yeasts are: *RFA*/*ssb1*, coding for replication protein A whose function is essential for the formation of the replication fork (Wold M.S., 1997) and *RNR1*/*cdc22*, coding for the catalytic subunit of a ribonucleotide reductase which is responsible for catalyzing production of deoxyribonucleotides for DNA synthesis (Fernandez Sarabia M.J. *et al.*, 1993).

Another well conserved process between the two yeasts is the DNA-damage checkpoint, although some differences in the pathways can be identified (Melo J. and Toczyski D., 2002). The adaptor protein Mrc1p/mrc1p has a conserved function in both species within the checkpoint cascade and shows a periodic behaviour in both organisms.

Sister-chromatid cohesion is essential to ensure proper chromosome segregation in M phase and is regulated by two complexes, condensin and cohesin. When cohesins are cleaved, they dissociate from the chromosomes and lead to sister chromatid separation. Two cohesin-coding genes, *SMC3*/*psm3* and *MCD1*/*rad21* showed a conserved periodic behaviour in fission and budding yeast, although some differences have been detected in the dissociation of these proteins from the chromosomes in the two yeasts (Tomonaga T. *et al.*, 2000).

The mitotic cascade of events is activated by a cyclin-dependent kinase and regulated downstream by members of three separate kinase families: the Aurora, Polo and NIMA-related kinases. They are highly conserved in eukaryotes with a single family member in yeasts and multiple-gene families in humans. In yeasts these kinases control several mitotic processes such as spindle formation, CDK-phosphorylation, the anaphase promoting complex which in turns regulates M phase exit, and cytokinesis (Nigg E.A., 1998; O'Connell M.J. *et al.*, 2003; Andrews P.D. *et al.*, 2003). All three kinases in budding (*CDC5*, *IPL1* and *KIN3*) and in fission yeast (*plo1*, *ark1* and *fin1*) have a conserved periodic behaviour.

Microtubules contribute to the formation of the mitotic spindle whose function is to segregate the chromosomes; microtubule rearrangements are mainly due to kinesin motor proteins such as Kar3p and Kip1p in budding yeast and klp5p, klp6p and klp8p in fission yeast, a protein family involved in many different functions. Sister chromatids get

separated when Pds1p/cut2p is degraded by the anaphase promoting complex which is activated by Cdc20p/slp1p. All genes coding for these proteins were found to be periodic.

Cytokinesis is conducted differently in different organisms but the major events appear to be universal. In *S. pombe* and animal cells, an actomyosin ring is formed at the cell equator, a membrane barrier is then synthesised to separate the two cytoplasms and then the two new cells will separate. In *S. pombe* an additional septum is synthesised behind the ring and then enzymatically digested to allow cell separation. In *S. cerevisiae* cells divide by budding. Positioning of the actomyosin ring differs significantly in the two yeasts but the downstream events are more conserved (Guertin D.A. *et al.*, 2002). The myosin II heavy chain protein responsible for contraction of the actomyosin ring (Myo1p/myo3p), the cdc15-like protein Hof1p/imp2p and Bud4p/mid2p, both components of the contractile ring, the Chs2p/csh2p septum component, the transcription factor Ace2p/ace2p, the glucanase Dse4p/eng1p involved in septum digestion, and the kinase Dbf2p/sid2p that regulates septum formation together with Mob1p/mob1p are all encoded by genes that show a conserved periodic behaviour.

As explained in the introduction to this thesis, the engine that drives the cell cycle is a complex formed by a kinase and its cyclin partners, highly conserved across eukaryotes. Three *S. pombe* cyclin coding genes (*cdc13*, *cig1* and *cig2*) appear to be cell cycle regulated and at least eight *S. cerevisiae* genes coding for cyclins are transcriptionally regulated. It should here be mentioned that in the case of cyclin genes it is difficult to establish a one to one relationship between orthologs, probably as a consequence of the duplication events that led in budding yeast to an increase in the number of cyclin genes. Also two of the CDK inhibitors have a conserved pattern of periodic gene expression, *SIC1*/*rum1* and *SWE1*/*mik1*. The gene *RAD51*/*rhp51,* coding for a protein regarded as a key player in homologous recombination and recombinational DNA repair and *MSH6/mhs6*, coding for a member of the eukaryotic DNA mismatch repair system, were found periodic in both model organisms.

5.3 Yeasts and humans: what is conserved?

A similar approach was adopted to compare cell-cycle regulated genes between yeasts and humans. Unfortunately, the lack of a reliable ortholog table for human genes made it impossible to run a systematic comparison as the one done between fission and budding

yeast. Therefore, only hand-curated human gene orthologs to the conserved/periodic fission/budding yeast genes were checked for periodicity. For this comparison, the datasets produced by Whitfield *et al*. (2002) and Cho *et al*. (2001) were considered. However, a statistical reanalysis of both studies revealed that cyclicity of human genes arises from experimental errors and that the reproducibility between experiments was very poor (Shedden K. and Cooper S., 2002; Cooper S. and Shedden K., 2003). Consequently, comparison was limited to those genes that were identified as periodic in human cells using traditional genetic approaches.

Human homologs to yeast genes coding for Aurora (STK6), Polo (SNK) and NIMA-related (NEK2) kinases showed a periodic behaviour underlying how conserved these protein families and their regulation are through evolution. Among the other human homologs that also appeared cell cycle regulated there were POLA (*POL1/pol1* homolog), CDC20 (*CDC20/slp1* homolog), RAD21 (*MCD1/rad21* homolog), RRM1 (*RNR1/cdc22* homolog), CDC6 (*CDC6/cdc18* homolog), which are all involved in DNA replication and have a conserved function across eukaryotes. Similarly, histones and several human cyclin genes (CCNE1 and 2, CCNG2, CCNA2, CCNF, CCNB1 and 2) were cell cycle regulated; periodic expression of these genes is a conserved mechanism of regulation from yeast to humans.

SNAI3 (*ACE2/ace2* homolog), which contains a zinc finger binding domain and acts as a transcriptional repressor, was also found periodic in the human studies. In this case homology is due only to the presence of the zinc finger domain and further studies are needed to investigate a potential functional homology between SNAI3 and ace2-type transcription factors.

Some genes were periodically expressed in fission yeast and humans only, including the protein kinase CDC2 and the protein phosphatase CDC25. Conversely, DNA replication MCM complex genes were periodic in budding yeast and humans only.

Despite the fact that absolute numbers of conserved genes among yeast species and between yeasts and humans are surprisingly small it is interesting that most of them have a well characterised regulatory function associated with basic cell cycle processes such as DNA replication, mitosis and cytokinesis. This differs from the larger list of cell cycle regulated genes and probably reflects the fact that periodic transcription of those genes plays a critical role in driving cell cycle progression.