Disruption of Ribosome Biogenesis Triggers a p21/p53-mediated Cell Cycle Checkpoint

Inauguraldissertation
zur Erlangung der Würde eines Doktors der Philosophie vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von:

Alessandro Di Cara

aus Palermo, Italien

Leiter der Arbeit: PD. Dr. George Thomas
Friedrich Miescher Institute for Biomedical Research

Basel, 2004

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Dr. G. Thomas, Dr. M.N. Hall, Dr B. Amati, Dr. N.E. Hynes, Dr. E.J. Oakeley

Basel, den
22/10/2004

Dekanin/Dekan
Hans-Jakob Wirz

Abstract

Cell cycle entry requires a dramatic increase in protein production. In order to cope with this demand, the cell must upregulate ribosome biogenesis. Given that ribosome biogenesis is the most energy-consuming anabolic process in a growing cell, and that changes in cellular ribosome content can alter the genetic program, we hypothesized that control mechanisms must exist to synchronize ribosome biogenesis and cell cycle progression. Here I report on a novel cell cycle checkpoint which is activated on the disruption of ribosome biogenesis and blocks cell cycle progression. Our studies, both in vitro and in vivo, show p21 and p53 as key mediators of this response.

To my parents,
Stefano Di Cara and Adriana Tumeo

Introduction

The aim of this project is the identification of a putative checkpoint mechanism that regulates cell cycle progression with respect to ribosome biogenesis. In the following sections, a brief description of the molecular mechanisms involved in the cell cycle, and its regulation, will be given. Particular emphasis will be placed on the mechanisms of regulation S phase entry. The subsequent sections will deal with the mechanisms of ribosome biogenesis and the regulation of this process during growth and proliferation in eukaryotic cells. Finally, a brief synopsis of the work already perfomed in this area of our lab will be provided, together with a description of the specific aims of my PhD project.

The cell cycle

The cell cycle represents the complete series of events from one cell division to the next. In eukaryotic cells this phenomenon can be divided in four major stages: G1, S, G2 and M. Each stage represents a different process undergone by the cell: The G1 part of the cycle is generally associated with cell growth and the gathering of resources for DNA replication; during S phase the cell duplicates its DNA; and the G2 phase represents an interphase in which cells grow further in size, duplicate their cellular organelles and gather resources for mitosis. Finally, mitosis (M phase) is the process of DNA partitioning and physical cell division. Some cell types also undergo a G0 phase, a quiescent state preceding the G1 phase.

A molecular overview

The task of describing all the molecular events that characterize all these phases is beyond the scope of this introduction. Instead, I will concentrate on the mechanisms regulating G1 to S phase progression, on which my work has focused. Cyclins and cyclin-dependant kinases are known to be major players in mediating the progression and passage through these two phases of the cycle.

It is generally thought that stimuli in the form of mitogens can kick-start this response by inducing the phosphorylation of a class of proteins known as pocket proteins or Rb family of proteins (Lukas, Lukas et al. 2004). The pocket proteins have a role in maintaining inhibitory complexes on promoters of key cell cycle regulatory genes. Inhibition is mediated through the direct interaction with members of the E2F family of transcription factors and the recruitment of hystone acetylases to the promoter. Upon phosphorylation, the pocket proteins are released from this complex permitting E2F-mediated transcription of the gene. A large number of genes that are regulated by E2F participate in cell cycle progression and the amplification of this process. Overall, these interactions constitute a positive feedback mechanism that allows the amplification of a mitogenic signal into a fully blown cellular response. The following sections describe the components of this pathway.

Cyclins and cyclin-dependant kinases

The phosphorylation of pocket proteins is mostly mediated by cyclin-dependent kinases (CDKs), a family of serine/threonine kinases. The activity and specificity of CDKs is modulated in a cell cycle-dependant manner by their association with cyclins. Cyclins are a group of structurally related proteins expressed during the different phases of the cell cycle. To date, at least 7 CDKs (1-7) and 8 major cyclin types (A-H) (Lew, Dulic et al. 1991) have been isolated from mammalian cells. Since each cyclin has a preferential CDK binding partner, the activity of each of the CDKs is controlled in a cell cycle-dependent manner. During late G1 phase, for example, cyclin E is actively expressed and binds CDK2, forming the cyclin E/CDK2 complex. Different cyclins can bind the same CDK albeit at different phases of the cell cycle: this phenomenon is best demonstrated by the recycling of CDK2 from cyclin E/CDK2 complexes into cyclin A/CDK2 complexes during S phase. Since cyclin A and cyclin E complexes share the same CDK but target different substrates (Nigg 1993), it implies that the substrate-specificity of the associated CDK is at least partly conferred by the cyclin.

In addition to their binding to cyclins, CDK activity is also regulated by the following mechanisms: phosphorylation of specific threonine residues by CDK-activating kinases (CAK) (Desai, Wessling et al. 1995); dephosphorylation of the Tyr 15- and Thr 14-

conserved residues; and the presence of cyclin-dependent kinase inhibitors (CKI) such as p21.

The composition of cyclin/CDK complexes with respect to the phases of the cell cycle is illustrated in (**Fig.1**). The major complex present in G1 is cyclin D/CDK(4,6), which is thought to be a molecular sensor that activates cell cycle progression in response to mitogenic stimuli (Sherr and Roberts 1999), but has also been associated with the regulation of cell growth (Sherr 1996). The initial phosphorylation of pocket proteins provided by this complex allows the transcription in the mid-to-late G1 of several genes involved in the cell cycle, of which cyclin E. Cyclin E complexes with CDK2 it promotes the phosphorylation of pocket proteins (Harbour, Luo et al. 1999), resulting in the release of inhibition of the rest of and the transcription of genes such as cyclin A. In a mechanism similar to the one described for cyclin E, cyclin A further participates in the phosphorylation of the Rb proteins, allowing the expression of yet another set of proteins which are required for cell cycle progression. The dogma on the ordered fashion of these events has recently been challenged by mouse knockout experiments. In cyclin D-deficient fibroblasts, for example, proliferation is nearly normal even if the cells show an increased requirement for mitogenic stimulation. Similarly, fibroblasts lacking both CDK4 and CDK6 show a normal cell cycle progression, albeit with a less efficient entry into S phase (Malumbres, Sotillo et al. 2004). These results show that both cyclin D and its associated CDKs (4 and 6) are dispensable for cell cycle progression. Another set of studies using knockouts of both cyclin E subtypes (Geng, Yu et al. 2003) or CDK2 (Ortega, Prieto et al. 2003) show that neither E type cyclins nor CDK2 are strictly required for either embryonic development or for continuous cell cycle progression. These effects might be explained by a certain degree of target overlap between the different cyclins and CDKs and/or by the capacity of some of the cyclins to bind other CDKs.

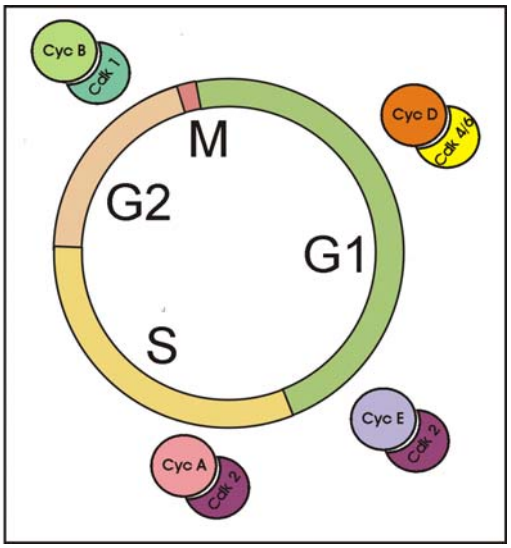


Fig.1 Diagram depicting the different phases of the cell cycle, together with the major cyclin/CDK complexes involved in each phase.

Pocket proteins

The Rb family of pocket proteins comprises three members: Rb; Rbl1/p107; and Rbl2/p130. All of these proteins are progressively phosphorylated during cell cycle progression, causing the release of inhibition on the E2F transcription factors.

The Rb protein is mostly involved in the passage of cells from the G1 to the S state, while Rbl2 is more involved in the control of the release from the G0 phase (Garriga, Limon et al. 1998; Classon, Salama et al. 2000). The other member of this family, Rbl1, is transcriptionally upregulated during the middle of the G1 phase (Beijersbergen, Carlee et al. 1995): its transcription depends partly on E2F and therefore on the successful phosphorylation of Rb and/or Rbl2 (Zhu, Xie et al. 1995). With regards to their levels of expression, Rb protein levels are nearly constant throughout the cell cycle while levels of Rbl1 and Rbl2 are modulated in opposite manners (Beijersbergen, Carlee et al. 1995). Upon its full phosphorylation in the late G1/early S phase, Rbl2 is degraded while Rbl1 levels gradually increase. The inverse tendency is observed when, at end of mitosis, cells re-enter G0.

The phosphorylation of pocket proteins is mostly mediated by CDKs and occurs at several sites, in a hierarchical manner. Recent data on Rbl2, for example, has shown the existence of 22 distinct phsophorylation sites, dependent on CDK(4,6) and/or CDK2

(Farkas, Hansen et al. 2002). Some of the sites require the previous phosphorylation of other sites, suggesting that phosphorylation proceeds in an orderly manner. Furthermore, while some sites can be phosphorylated by either of the CDKs mentioned, some of them are solely dependant on CDK2 or CDK(4,6) (Farkas, Hansen et al. 2002). Due to the different timing of CDK(4,6) and CDK2 activity, phosphorylation of these sites reflects a cell phase-specific behavior. Similar observations have also been made for Rbl1 (Farkas, Hansen et al. 2002) . Interestingly, two phosphorylation sites on Rbl2 are phosphorylated to as yet unidentified non-CDK kinases (Farkas, Hansen et al. 2002).

E2F transcription factors

The E2F family of transcription factors is composed of at least seven members (E2F1-7) that physically associate to form a heterodimer with a member of the DP family (usually DP1), which subsequently binds to the E2F site in target genes. E2F transcription regulates a number of genes involved in cell proliferation and differentiation. Identified E2F target genes include cell cycle regulators such as cyclin E, cyclin A, Rbl1, cyclin D1, Cdc2, and Cdc25A, and enzymes involved in DNA synthesis and DNA replication such as dihydrofolate reductase (DHFR), DNA polymerase α, thymidine kinase, PCNA, Cdc6 and the minichromosome maintenance (MCM) proteins (Dyson 1998; Ren, Cam et al. 2002). Interestingly, some E2F members, such as E2F1, stimulate their own production and that of their binding partners, such as DP1 (Ren, Cam et al. 2002). This mechanism is responsible of the upregulation of E2F1 protein levels during the mid-to-late S phase, which in turn stimulates the induction of further E2F-responsive genes and promotes S phase entry.

E2F transcriptional activity is modulated by multiple mechanisms, including negative regulation by the Rb family of pocket proteins. Different E2F/DP complexes bind to different Rb proteins: E2F1, E2F2, and E2F3 preferentially associate with pRb (Dyson 1998) and are potent transcriptional activators, whereas E2F4 and E2F5 predominantly interact with p107 and p130 and seem to be primarily involved in the active repression of E2F-responsive genes (Trimarchi and Lees 2002). E2F6 and E2F7 make up a subgroup that acts principally as a transcriptional repressor through a distinct pocket

protein-independent manner (Morkel, Wenkel et al. 1997; Di Stefano, Jensen et al. 2003).

The binding of pocket proteins to E2F members inhibits their transcriptional capacity, but also mediates the recruitment of acetylases (Brehm, Miska et al. 1998; Ferreira, Magnaghi-Jaulin et al. 1998) to the bound promoter, causing DNA conformational changes that further inhibit the transcription of the gene (O'Connor, Schaley et al. 2001). Since E2Fs constitute the docking sites for Rb proteins, their absence results in the disregulated activation of certain genes. Mice lacking E2F1, for example, have a higher propensity to develop tumors (Yamasaki, Bronson et al. 1998). Paradoxically, therefore, E2Fs can have a role both as oncogenes and tumor suppressors.

Cyclin-dependent kinase inhibitors

Cyclin-dependent kinase inhibitors are involved in the modulation of cyclin-dependant kinase activity during the cell cycle. This family of proteins can be divided in two broad categories:

(i) Ink4: $p16^{ink4a}$, $p15^{ink4b}$, $p18^{ink4c}$, $p19^{ink4d}$

(ii) CIP/Kip: $p21^{CIP1/Waf1}$, $p27^{Kip1}$, $p57^{Kip2}$

The members of the Ink4 group inhibit CDK4 and CDK6 by promoting their dissociation from cyclin D, while the CIP/Kip members inhibit all CDKs in a concentration-dependant manner (reviewed in (Sherr 1996)). Overexpression of any member of the CIP/Kip family causes a G1 block arrest in transfected cells (Pestell, Albanese et al. 1999), while overexpression of the Ink4 members such as p16 causes a reduction in the levels of the cyclin D/CDK4 complex (Hirai, Roussel et al. 1995; Quelle, Ashmun et al. 1995).

The role of cyclin D/CDK4 in promoting cyclin E production through the phosphorylation of pocket proteins has already been described; an additional participation of the cyclin D/CDK4 complex in this process is the sequestration of p21 and p27 CKI (Sherr and Roberts 1995). The free p21 and p27 proteins present in the cell participate in the formation and stabilization of cyclin D/CDK4 complexes (LaBaer, Garrett et al. 1997; Cheng, Olivier et al. 1999). The titration of these two CKIs relieves the inhibition of

cyclin E/CDK2 complexes. This, in turn, grants the progressive accumulation of cyclin E/CDK2 activity and the further hyperphosphorylation of Rb proteins. This system generates a hierarchical program of CDK activation, since the increase in CDK2 activity during G1 requires the inactivation of both the CIP/Kip proteins and is therefore dependent on the prior activation of the cyclin D pathway. Once CDK2 becomes active, it triggers the degradation of p27 by targeting it for phosphorylation and subsequent ubiquitination (Vlach, Hennecke et al. 1997). This event has two major effects: the destabilization of cyclin D/CDK(4,6) complexes; and the resulting release of p21, which is now able to inhibit cyclin E/CDK2. The temporary titration of p21 by cyclin E/CDK2 allows the formation of cyclinA/CDK2 complexes, which will then mediate S-phase progression. P21 and p27 differ mostly in their expression kinetics. Generally, p27 levels are high in quiescent cells and decrease in late G1: the protein levels are thought to be regulated by proteosome mediated degradation (Vlach, Hennecke et al. 1997). P21 levels are, instead, low in quiescent cells and increase during the late G1 phase (Macleod, Sherry et al. 1995). In addition to its normal regulatory mechanisms, high p21 levels can be induced by the p53 pathway (el-Deiry, Tokino et al. 1993). The Ink4 proteins are thought to mediate CDK activity modulation through the same pathway. Upon an anti-mitogenic stimulus, such as TGF-B, Ink4s are expressed and promote the dissociation of cyclin D/CDK4, which causes the releases of p21 and the inhibition of CDK2 activity (Reynisdottir, Polyak et al. 1995).

Cell cycle checkpoints and the role of p53

As described in the previous section, the progression through the cell cycle relies on an ordered series of events controlled by both activating and inhibitory interactions. The control of these mechanisms is of crucial importance for the development of multicellular organisms, where growth and division must be constrained. In this environment, cells proliferate only when subjected to stimuli such as growth factors and only when the integrity of the cell is not at stake. Cells have developed a number of “checkpoint” systems which are elicited to prevent cells from entering a new cell cycle phase before having completed the previous one, or when subject to a stress such as nutrient deprivation or UV damage. The most characterized of these checkpoints is the

so-called G1 “restriction point” (Pardee 1989). This checkpoint mechanism is placed the end of the G1 phase and constitutes a point of no return, after which cells are committed to enter S phase and duplicate their DNA. The molecular effectors of these checkpoint mechanisms are essentially the elements that participate in the normal cell cycle regulation: cyclins, CDKs, CDK inhibitors, and pocket proteins. Tumors constitute an example of a situation where these mechanisms have become defective, causing uncontrolled cell proliferation. Of particular interest in this manuscript is the role of the tumor suppressor p53, a frequent target of genetic alteration identified in human cancers (Hollstein, Sidransky et al. 1991).

Stress signals from various pathways, such as DNA strand breaks, stalled DNA replication forks, ribonucleotide deprivation, and hypoxia, all appear to converge on the activation of p53, reviewed in (Ashcroft, Kubbutat et al. 1999). The most studied example is probably the induction of p53 by DNA damage. In order to prevent entry into S phase with damaged DNA, a checkpoint mechanism is triggered in G1 through the ATM/ATR and Chk1/Chk2 kinases. These transducing kinases activate two major responses: the phosphorylation of the Cdc25 proteins; and the activation of the p53 transcription factor. The two mechanisms target, the same pathway but by two different means (**Fig.2**).

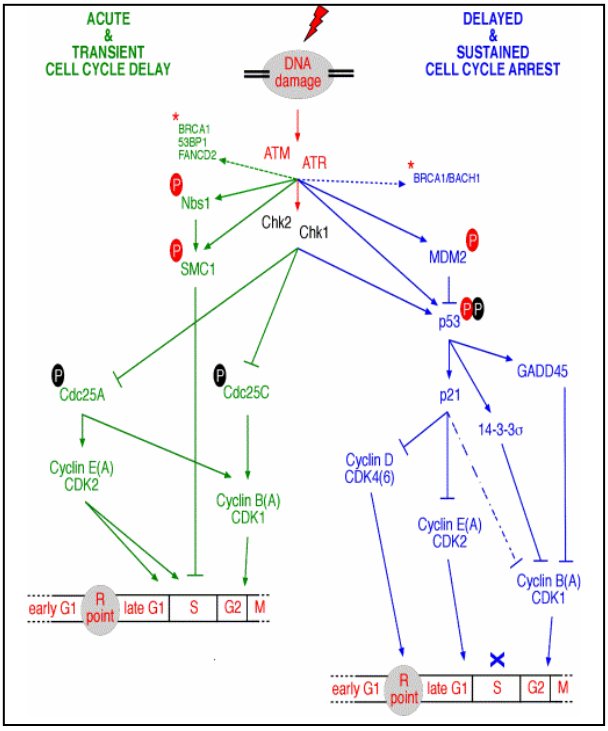


Fig.2 Pathways involved in the ATM/ATR-mediated response to DNA damage (Lukas, Lukas et al. 2004).

The phosphorylation of Cdc25a causes its degradation and decreases CDK2 activity which is dependant on the activatory dephosphorylation mediated by this protein (Zhao, Watkins et al. 2002; Sorensen, Syljuasen et al. 2003). The decrease in CDK2 activity affects the cyclin E/CDK2 and cyclin A/CDK2 complexes, which are not able to hyperphosphorylate the Rb pocket proteins. In the parallel pathway, p53 is activated by different mechanisms. Both the Chk1/2 kinases and ATM/ATR directly phosphorylate p53 and MDM2, a ubiquitin ligase that in its hypophosphorylated conformation targets p53 for degradation (Khosravi, Maya et al. 1999). The result is the activation and stabilization of p53 and the concomitant transcription of genes involved in cell cycle regulation and cell fate, such as $p21^{CIP/Kip}$ described earlier. p21 will also participate in the regulation of proliferation through the inhibition of CDK2 (Sherr and Roberts 1999),.and the sequestration of PCNA, an essential component of the DNA replication machinery (Li, Waga et al. 1994). Altogether, these mechanisms block the cell in G1, allowing the activation of the DNA repair mechanisms. The block in the cell cycle is maintained mostly through p21 (Bartek and Lukas 2001) until the DNA has been repaired. In case of severe DNA damage, a p53-dependent pro-apoptotic pathway is stimulated through mediators such as PUMA, NOXA, BAX, and PIG3 (Vousden and Lu 2002). p21 is actively degraded during the apoptotic response. This degradation is

considered to be one of the events pushing cells from stasis to apoptosis in response to DNA damage (Polyak, Waldman et al. 1996; Zhang, Fujita et al. 1999). P21 therefore has a role in cell stasis and protects from p53-mediated apoptosis.

To avoid the segregation of damaged DNA during mitosis the ATM/Chk(1,2) pathway also activates a block in G2. The mode of action is similar to the one just described for G1/S, but relies on the inhibition of CDK1, which normally associates with cyclin B and mediates progression from G2 to M. Cdc25c degradation by Chk1 phosphorylation (Brown, Lee et al. 1999), and p21 upregulation through p53, cause loss of CDK1 activity. In addition, p53 also induces a number of other factors, such as 14-3-3σ, which anchors CDK1 in the cytoplasm where it cannot induce mitosis, and Gadd45, which dissociates CDK1 from cyclin B1 (Zhan, Antinore et al. 1999). All the mechanisms just described constitute just an example of how a cellular stress such as DNA damage can cause, through p53, the arrest of normal cell cycle progression.

Ribosome biogenesis, the synthesis of new ribosomes

Ribosome biogenesis is the process responsible for the synthesis of new ribosomes: it has been mostly characterized in yeast but is conserved in mammalian cells (Eichler and Craig 1994). In eukaryotes, ribosome biogenesis takes place in the nucleolus and involves the synchronized production of both ribosomal RNA (rRNA) and ribosomal proteins. The eukaryotic ribosome is composed of two subunits distinguished by their different sedimentation rate: the 40S or small ribosomal subunit; and the 60S or large ribosomal subunit. Both subunits are composed of an intricate assembly of rRNA and ribosomal proteins.

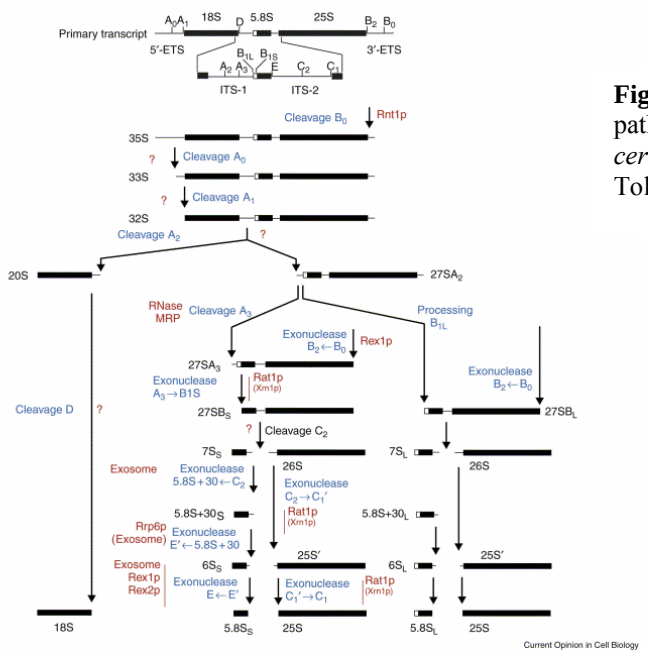


Fig.3. rRNA processing pathway in *Saccharomyces cerevisiae* (Fatica and Tollervey 2002).

One of the major steps in ribsosome biogenesis is the synthesis of the rRNA. The ribosome is composed of four mature rRNAs: the 18S rRNA in the 40S subunit; and the 5.8S, 28S, and 5S rRNAs in the 60S subunit. All, except the 5S, are derived from a same 45S rRNA precursor transcribed by RNA Pol I. The 45S rRNA precursor undergoes several processing steps to yield the mature rRNA (reviewed in (Fatica and Tollervey 2002)). The cleavage of the A2 site is of particular relevance, since it corresponds to the separation of the precursor of the 18S rRNA from that of the 28S/5.8S rRNA (**Fig.3**). Following this cleavage, the two rRNAs are extensively modified, a process requiring both endo- and exonucleases and the binding of ribosomal proteins. The 5S rRNA is transcribed separately by RNA polymerase III (Pol III) and does not undergo further processing. Ribosomal proteins, which bind to the rRNA, are transcribed by RNA polymerase II (Pol II). They are produced in the cytoplasm and then imported into the nucleus or nucleolus for assembly with the rRNA. Ribosome biogenesis therefore requires the coordination of events in different cellular compartments and the transcriptional activity of different polymerases. In mammalian cells, the means used for this coordination are still largely unknown.

Ribosome biogenesis during cell growth and cell proliferation

Cell growth and cell proliferation require that cells adjust their protein biosynthetic capacity in response to nutrient availability and external stimuli such as mitogens or growth factors. This involves coordinated changes both in the rate of translational initiation and in the abundance of the protein synthesis machinery itself, especially the number of ribosomes. Studies in both prokaryotic and eukaryotic cells have shown that the concentration of ribosomes within the cell can vary several-fold depending on growth rate (reviewed in (Gausing 1974; Woolford 1991)). Coordination between ribosome biogenesis and cell state is probably derived from the phase-specific requirements of ribosome numbers to meet demands for protein biosynthesis, but also from a need to preserve energy resources, since ribosome biogenesis requires a large amount of energy expenditure. It has been estimated that ribosome biogenesis constitutes 80% of the work of cell proliferation (Sollner-Webb and Tower 1986), and in *Saccharomyces cerevisiae* it consumes an extraordinary proportion of the cell's resources, accounting for >70% of all transcription (about 50% of all Pol II transcription) (Warner, Vilardell et al. 2001).

Ribosome biogenesis is a very complex process and can be regulated at different levels. It has been known for a long time, for example, that Pol I-mediated transcription of rRNA is tightly controlled. Amino acid starvation and nucleotide levels affects rRNA transcription rates (Grummt and Grummt 1976; Grummt, Smith et al. 1976), accomomodating rRNA production to the availability of nutrients and therefore to the rate of growth. From a cell cycle perspective, the rate of rRNA gene transcription decreases markedly in cells that are in S phase, whereas it becomes upregulated upon treatment with mitogenic signals. The primary transcription factor involved in the cell cycle-dependent regulation of rRNA transcription is TIF-IA which, together with UBF, controls the activity of Pol I (Schnapp, Pfleiderer et al. 1990; Jacob 1995; Datta, Budhiraja et al. 1997). Part of the effect of members of the pocket protein family on cell cycle progression is probably due to their effects on Pol I transcription. In fact, Rb and Rbl2 have been reported to inhibit the activity of Pol I through binding to its enhancer, UBF. Consistent with a role of pocket proteins in the inhibition of Pol I, rRNA synthesis is elevated in fibroblasts lacking both Rb and p130 (Ciarmatori, Scott et al. 2001).

Furthermore, all the members of the Rb pocket protein family are able to interact with and inhibit Pol III activity (Sutcliffe, Cairns et al. 1999), which is responsible for the synthesis of tRNA and the 5S rRNA, which are necessary for translation and ribosome biogenesis respectively. Given the regulation of pocket proteins during the cell cycle, these interactions have probably evolved to modulate ribosome biogenesis as a function of cell cycle progression. Interestingly, p53 has also been shown to inhibit Pol I transcription, providing a mechanism to downregulate ribosome biogenesis during stress responses (Budde and Grummt 1999). In certain cell types, ribosome biogenesis is also regulated at the post-transcriptional level. During rRNA processing, as much as half of 18S rRNA can be degraded continuously in resting cells: this process is attenuated when cells receive growth stimuli (Cooper and Gibson 1971; Dudov and Dabeva 1983; Eichler and Craig 1994).

The regulatory mechanisms described so far rely on rRNA synthesis and availability. Importantly, efficient assembly of ribosomes requires coordination in the synthesis of their main components, rRNA and ribosomal proteins. In metazoa, the production of ribosomal proteins is regulated mainly at the translational level (Meyuhas, Nir et al. 1996; Thomas 2002). The transcripts encoding for ribosomal proteins are poorly translated in resting cells, whereas they are efficiently recruited to translating ribosomes when cells are exposed to growth and mitogenic signals. Translational regulation is conferred on these mRNAs by the presence of an unusual pyrimidine-rich sequence at their 5' end, referred to as a 5'-TOP (5'-Terminal OligoPyrimidine) tract. Regulation of 5'-TOP messages is mediated by the mTOR pathway, possibly through the activation of S6K (Meyuhas, Nir et al. 1996; Jefferies, Fumagalli et al. 1997). It is interesting that mTOR activity also stimulates the transcription of rRNA, possibly ensuring a certain degree of temporal coordination between production of rRNA and ribosomal proteins. mTOR appears to function as a hub, receiving input from mitogens, growth factors, nutrients, and energy levels (Zhang, Stallock et al. 2000; Dennis, Jaeschke et al. 2001; Shamji, Nghiem et al. 2003). It therefore constitutes an ideal gatekeeper for the regulation of ribosome biogenesis in response to the environment which the cell is exposed to. There is also evidence that the Rbl1 pocket proteins take advantage of this pathway to mediate a cell cycle-dependent inhibition of ribosomal protein synthesis, supported by the observation that overexpression of Rbl1 leads to the downregulation

of ribosomal protein translation, in correlation with a loss of S6 phosphorylation, a main substrate of S6K (Makris, Voisin et al. 2002).

Regulation of cell growth and cell proliferation by ribosome biogenesis

The previous section has highlighted the mechanisms that control ribosome biogenesis in response to cell growth and proliferation. Is there also a mechanism that controls cell growth and cell proliferation in response to ribosome availability or translation capacity? We hypothesize that such a mechanism exists for the following reasons:

(1) A feedback mechanism should exist in order, for example, to communicate to the cell cycle machinery to halt the upregulation of ribosome biogenesis when the required protein synthesis rate has been achieved, avoiding in this way unnecessary energy expenditures.
(2) The failure to upregulate ribosome biogenesis, in case of a defect, would create an imbalance in the number of ribosomes, potentially resulting in alterations to the translational program. It is thought that in mammalian cells mRNAs compete for association with ribosomes which are thought to be present in limiting amounts (Lodish 1974). The complexity of the 5' UTR of the mRNAs, their relative abundance, and their affinity for the translational machinery influences their rate of translation. Variations in ribosome amounts may lead to alterations in the pattern of the translated mRNA and have deleterious effects on cell homeostasis.

My project aimed at determining whether such control mechanisms exist. The publication by Volarevic et al. (Volarevic, Stewart et al. 2000) presents work previously performed in our laboratory on this subject. My PhD. project stemmed directly from this work.

The work by Volarevic et al. describes the generation of mice harboring a conditional deletion of the gene encoding the S6 ribosomal protein, a member of the 40S

ribosomal subunit. Deletion of S6 in adult mouse livers causes the abrogation of 40S synthesis due to a defect in the processing of the 18S rRNA precursor. The deletion of the S6 gene stops 40S ribosome biogenesis whithout affecting the pre-existing pool of ribosomes. Global translation was also shown to be unaltered by this defect.

To test the effects of the lack of ribosome biogenesis on cell growth and on cell proliferation, the livers were subjected to a starvation/refeeding experiment and to partial hepatectomy, respectively. The starvation/refeeding experimental paradigm relies on the fact that, when a mouse is starved of nutrients, the size of its liver decreases. Upon refeeding, however, the liver grows back to its original size, a process that relies solely on cell growth. Upon starvation and subsequent refeeding, ΔS6 livers grew back to their original size, showing that abrogation of ribosome biogenesis by S6 deletion does not affect liver growth under these conditions. The hepatectomy experiment, instead, is based on the property of the liver to regenerate to its original size after the removal of 70% of its mass. Regeneration relies principally on proliferation. When subjected to hepatectomy, ΔS6 livers failed to regenerate, showing that in the absence of ribosome biogenesis proliferation is compromised (**Fig.4A**).

A more detailed molecular characterization revealed that the G1 phase was unaltered in cells from ΔS6 livers, as shown by the normal kinetics of downregulation of p27 and accumulation of p21 and cyclin D protein levels. Furthermore, the activity of the cyclinD/CDK4 was normally induced, as in wild-type livers (**Fig.4B**). On the other hand, cells lacking S6 failed to progress through S phase and showed a lack of cyclin E, both at the protein and transcripts level, and as a consequence of CDK2 activity (**Fig.4C**). Altogether, these findings suggested the existence of a mechanismthat stops cell proliferation in response to defects in ribosome biogenesis (**Fig.5**).

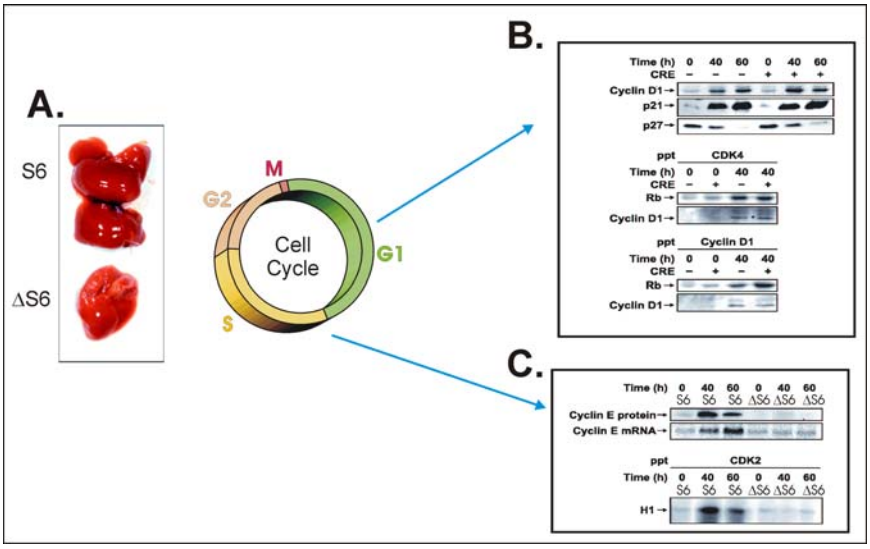


Fig.4 (**A**) Wildtype and ΔS6 livers 10 days after hepatectomy. (**B**) Characterization of G1-phase proteins at different timepoints after hepatectomy in wt and ΔS6 livers. (**C**) Characterization of S-phase proteins at different timepoints after hepatectomy in wt and ΔS6 livers.

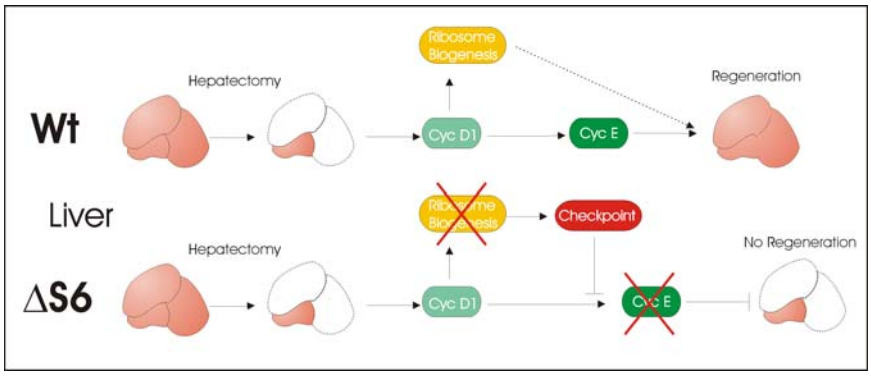


Fig.5 Model of the relationship between ribosome biogenesis and the cell cycle. In a wildtype situation, ribosome biogenesis is triggered by the cell cycle machinery and participates in cell cycle progression. In case of a defect in ribosome biogenesis, a checkpoint mechanism is activated upstream of cyclin E, and prevents cell cycle progression.

Aims of the PhD project

- **(1) To determine the origin of the mechanisms that cause the cell cycle block in cells defective in ribosome biogenesis**

The uncovering of such mechanism will provide clues on how ribosome biogenesis interacts with the cell cycle machinery.

- **(2) To establish an in vitro system that mimics the effetcs of deletion of the S6 gene in the mouse**

The aim is to gain experimental flexibility by using an in vitro model that can be more easily manipulated than the current mouse model and allows the testing of whether the phenotype observd in vivo is cell-autonomous.

Results

Characterization of a cell cycle checkpoint induced by the conditional deletion of the ribosomal protein S6 in the liver

o Characterization of the checkpoint by microarray analysis

Previous studies from our lab have shown that mouse livers harboring a conditional deletion of the S6 ribosomal protein (S6) gene are not able to regenerate in response to hepatectomy (Volarevic, Stewart et al. 2000), due to the failure of the hepatocytes to enter the first S phase. The absence of cyclin E induction, both at the protein and RNA levels, was thought to contribute to the defect in proliferation, although the involvement of additional lesions could not be excluded. Therefore, in order to gain a wider view of the genes and pathways possibly involved in the cell cycle block elicited by S6 depletion, I performed a microarray analysis using chips from Affymetrix. With this approach it was possible to measure the expression at the mRNA level of several thousand genes simultaneously. For this purpose I collected samples at different times following partial hepatectomy. Two types of mice were used: (1) mutant mice harboring a floxed version of the S6 gene and transgenic for a CRE recombinase under the control of the Mx promoter (Mx-Cre), where deletion was triggered through injection of polyI:polyC (I will refer to these mice as ΔS6/flox); and (2) wildtype mice harboring the floxed S6 gene, lacking the Mx-Cre transgene, and also injected with polyI:polyC (I will refer to these mice as S6/flox). Remnants of the livers were collected at: 0, 20, 30, and 40 hours after a two-thirds hepatectomy. The first S phase in this genetic background occurs about 40 hours after hepatectomy (Volarevic, Stewart et al. 2000). We used three mice per timepoint per genotype.

In order to determine the differences in gene expression between S6/flox and ΔS6/flox livers, I compared the two genotypes at each timepoint and derived a list of differentially expressed genes. I selected genes showing at least a 1.5-fold change between the two genotypes at any given phenotype and showing a change p-value of <0.01. These settings chosen for the fold change are slightly less stringent than the

arbitrary 2-fold change limit routinely used by other groups. The results, however, show that even slight changes in mRNA levels can cause substantial differences in protein levels. Decreasing the stringency in the fold change resulted in a substantial increase in candidate genes. Taking advantage of the fact that we had three replicates for each condition, I compensated this increase with a more stringent filter on the p-value change (less than 0.01 rather than 0.05). Because of the higher stringency in the analysis, this approach was aimed at obtaining a list of the best candidate genes rather than a comprehensive list of all the genes changing between the two genotypes (the results of the analysis are listed in Appendix A). As a proof of concept, we first measured the S6 mRNA levels in the samples (Fig.6). As expected, in the liver of ΔS6/flox mice the expression of the S6 transcript remained below detectable levels throughout the time course of the experiment, showing the efficiency of the conditional knockout.

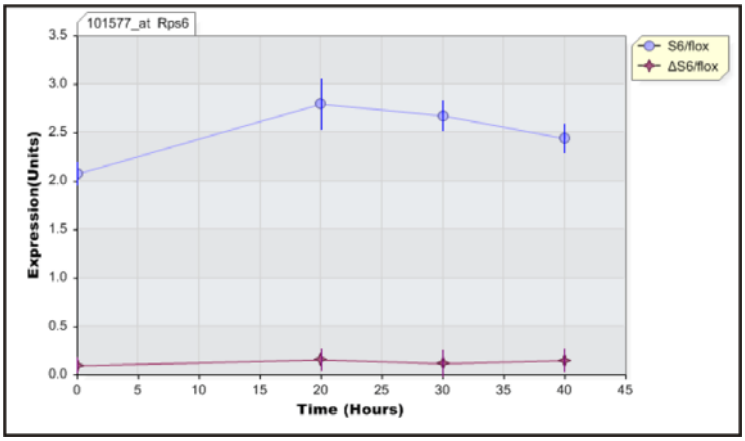


Fig.6 Normalized microarray data of the expression levels of S6 ribosomal protein in S6/flox (blue) versus ΔS6/flox (red) at different timepoints after hepatectomy. Each data point represents the average of three independent experiments. Error bars correspond to standard error.

As shown in (**Fig.7)**, the number of genes differentially expressed increased as a function of time, with a sharp increase at 30 hours after hepatectomy. A more accurate analysis showed that the majority of genes collected in the lists were more expressed in S6/flox livers. These differences could be caused either by the induction of these genes in S6/flox livers and/or by their downregulation in the ΔS6/flox condition.

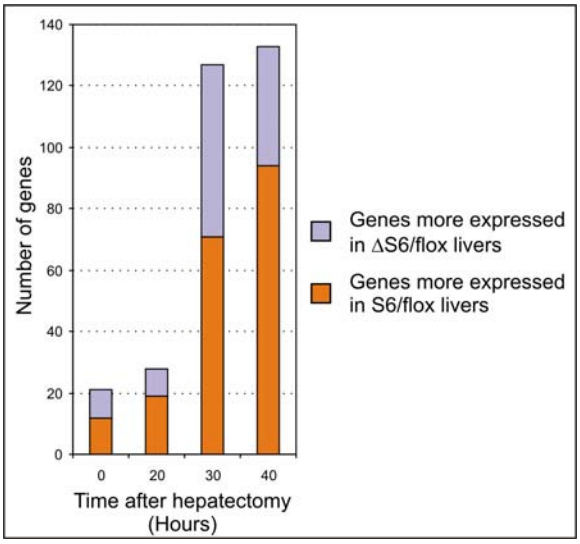


Fig.7 Histogram showing the number of genes found differentially expressed between S6/flox and ΔS6/flox at different timepoints after hepatectomy. Each bar is divided into genes expressed more in S6/flox or ΔS6/flox.

In order to clarify this point I performed an expression cluster analysis to identify genes displaying differences in expression between the two genotypes throughout the time course. I initially focused on the list of genes differentially expressed at 40 hours, for two reasons: first, at this timepoint I observed the highest number of differences between the two genotypes; and second, it is at 40 hours that the differences, with regards to entry into S phase, were originally detected between the two genotypes. The cluster analysis showed that in the liver of S6/flox the expression of a number of genes increased with time, reaching a peak at 40 hours after hepatectomy. Those same genes were not induced in the liver of ΔS6/flox (**Fig.8A**). On the other hand, the analysis of other timepoints led to the identification of genes expressed at higher levels in ΔS6/flox. In particular, expression of these genes was induced in both genotypes at 20 hours, although to a higher extent in ΔS6/flox. Furthermore, whereas in ΔS6/flox expression of these genes was maintained, it was instead gradually attenuated in S6/flox mice throughout the rest of the time course (**Fig.8B**). We decided to study these two major clusters in more detail by analyzing the genes that composed them. The following sections of this manuscript deal with this analysis.

o Lack of upregulation of cell cycle inducer genes in ΔS6/flox livers

I first analyzed the cluster of genes more expressed in S6/flox. As previously

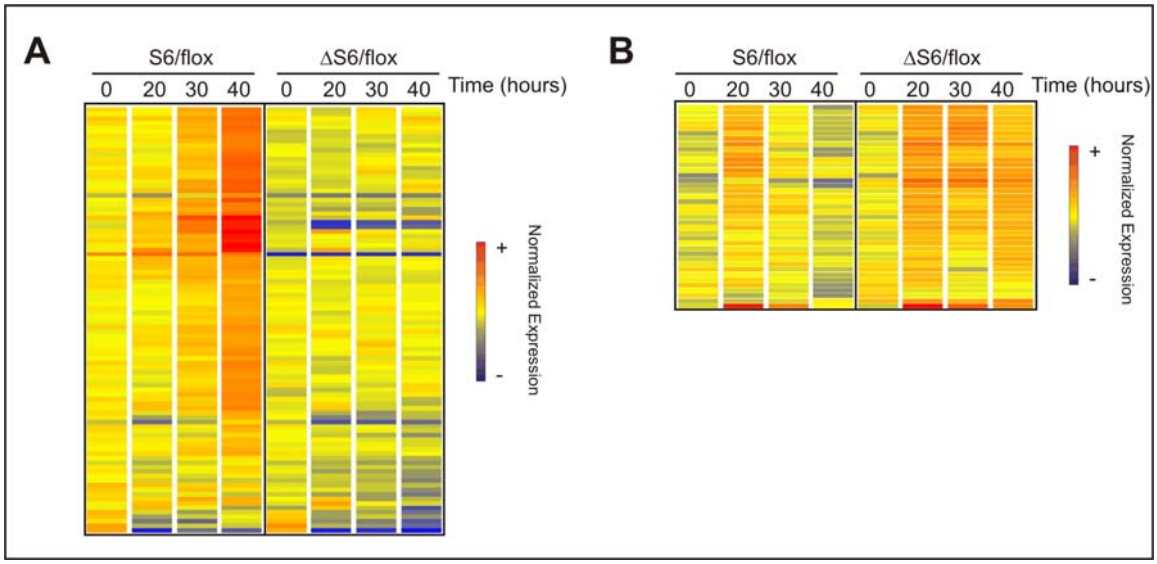
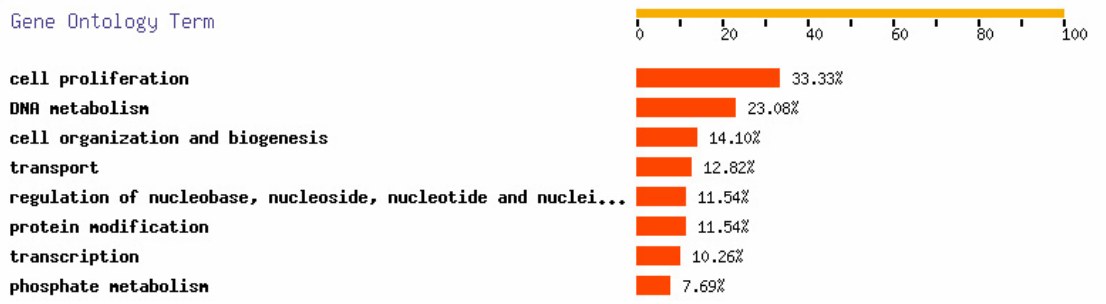


Fig.8 Expression cluster analysis of genes differentially expressed between S6/flox and ΔS6/flox at 40 hours after hepatectomy. Two distinct clusters are shown: (**A**) genes expressed more in S6/flox; and (**B**) genes expressed more in ΔS6/flox. Each horizontal line corresponds to a gene. Each column represents a timepoint. The colors reflect the levels of expression.

mentioned, these genes fail to be upregulated in ΔS6/flox. Since the over-representation of genes from a common functional class within an expression cluster can reveal a global pathway of activation, we grouped these genes by their biological function using the "Affymetrix GO analysis tool" (http://www.affymetrix.com) and "FatiGO" (Al-Shahrour, Diaz-Uriarte et al. 2004). The analysis resulted in the identification of several functional categories (Fig.9). Since I was principally interested in the mechanisms that control the cell cycle, I decided to focus the categories that could be directly linked to this process. Interestingly, I found that about 68% of the genes clustered in functional categories such as "cell proliferation", "DNA metabolism", and "regulation of nucleotide and nucleoside biosynthesis", which are directly involved in cell cycle progression. The genes found in these three categories were pulled together into a unique group referred to hereafter as CCIND (Cell Cycle Induced genes).



cell proliferation	Brrn	Anln	Aurkb	Brca1	Bub1	Ccna2	Ccnb2	
	Cdc20	Cdc6	Cdkn3	Cks1	Fen1	H2afx	Mad2l1	Mcm5
	Nek2	Plk1	Plk4	Rbl1	Ris2	Rrm1	Rrm2	Stk6
	Tacc3	Top2a	Uhrf1					
DNA metabolism	Aurkb	Brca1	Cbx5	Cenpa	Ezh2	Fen1	H2afx	H2afz
	H2av	Hat1	Hmgb1	Hmgb2	Mcm5	Rad51ap1		Rrm1
	Rrm2	Tk1	Top2a					
cell organization and biogenesis	Cbx5	Cenpa	Ezh2	H2afx	H2afz	H2av	Kif22	Rps6
	Tuba1	Tuba2	Tuba6					
transport	Fabp1	Hba-a1	Hbb-b1	Hmgb1	Sec61b	Tacc3	Tuba1	Tuba2
	Tuba6	Uhrf1						
protein modification	2700084L22Rik		Aurkb	Bub1	Ezh2	Hat1	Nek2	Plk1
	Plk4	Stk6						
regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism	Atpi	Ezh2	Hmgb2	Mcm5	Pbx3	Rbl1	Tcf19	Tmpo
	Uhrf1							
transcription	Ezh2	Hmgb2	Mcm5	Pbx3	Rbl1	Tcf19	Tmpo	Uhrf1
phosphate metabolism	Aurkb	Bub1	Nek2	Plk1	Plk4	Stk6		

Fig.9 Functional analysis, using gene otology, of genes expressed more in S6/flox at 40 hours after he patectomy. The top panel represents the percentage of genes present in each category. Categories representing less than 5% of the genes were not considered. The bottom panel lists the genes in each corresponding functional category. Each gene can have multiple functions and is therefore likely to be found in more than one category. The genes represented in blue constitute the list of selected CCIND genes.

similar expression profile and a related function. We set out to determine whether these genes shared common regulatory pathways. We found that about 49% of the CCIND gene promoters had peviously been reported to interact with E2F (Ren, Cam et al. 2002) (Fig.10).

CCIND genes E2F regulated	Aurkb	Brca1	Bub1	Ccna2	Cdc20	Cdc6	Cks1	Ezh2
	Fen1	Mad2l1	Mcm5	Rbl1	Rrm1	Rrm2	Top2a	H2afx
	Tk1							
Other CCIND genes	Anln	Atpi	Brrn	Cbx5	Ccnb2	Cdkn3	Hat1	Hmgb2
	Hmgb1	Nek2	Pbx3	Plk1	Plk4	Ris2	Stk6	Rad51ap1
	Tacc3	Tcf19	Tmpo	Uhrf1				

Fig.10 Table listing the cell cycle induced genes (CCIND) selected from the expression and functional clustering .The genes are divided according to whether they have been reported to depend on E2F transcription or not. Although other transcription factors are involved in the regulation of these genes, none were represented at to the same extent as E2F.

target genes are not induced in ΔS6/flox hepatocytes after hepatectomy

The E2F family of transcription factors comprises 7 members, all involved in both the positive and negative regulation of cell proliferation and differentiation. E2F1, in particular, has been characterized as a major inducer of cell proliferation (Chellappan, Hiebert et al. 1991) by mediating the transcription of several genes involved in DNA replication. E2F1 expression is regulated transcriptionally and peaks at the onset of S phase (Slansky, Li et al. 1993). I was therefore surprised by the absence of E2F1 in the list of differentially expressed genes. A closer look at the microarray raw data revealed that at 40 hours after hepatectomy E2F1 showed a p-value difference of 0.0057 between the two genotypes, a value well within the 0.01 threshold used. The gene, however, showed an overall fold change of expression between the two genotypes of 1.49, just below the 1.5 threshold set in our initial analysis, explaining why it was not found within our final gene lists (Fig.11A). It is important to note that a fold change threshold, as set in our experiment, is a purely arbitrary value, since even slight changes in the expression of a gene can greatly affect its protein levels. Using northern blot analysis we showed that E2F1 mRNA was undetectable in ΔS6/flox at all the tested timepoints, whereas it was indeed gradually induced in S6/flox, reaching its maximum at 40 hours (Fig.11B).

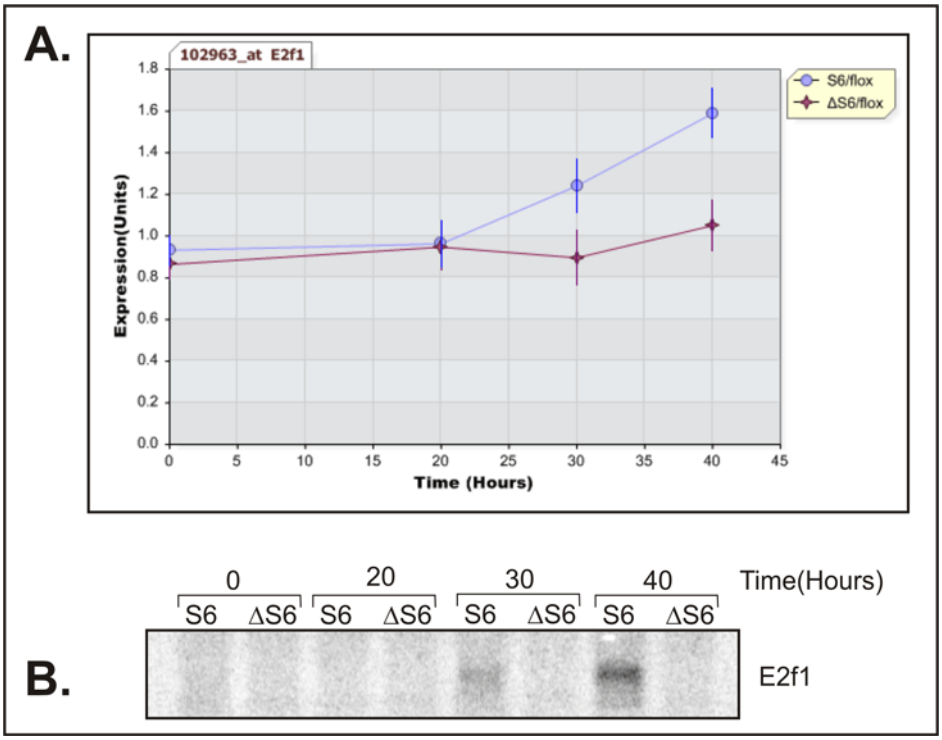


Fig.11 (**A**) Normalized microarray data of the expression levels of E2F1 in S6/flox (blue) versus ΔS6/flox (red) at different timepoints after hepatectomy. Each data point represents the average of three independent experiments. Error bars correspond to standard error. (**B**) Northern blot showing E2F1 mRNA levels in S6/flox and ΔS6/flox livers at different timepoints after hepatectomy.

On the base of this result, I hypothesized that a defect in E2F1 expression could explain the lack of induction of some of the CCIND genes in ΔS6/flox. To test this hypothesis further, I measured the levels of E2F1 protein. I was, however, unable to detect it through either western blotting or in vitro DNA pull-down assays, probably because of the relatively low abundance of this transcription factor in the liver. I therefore used the levels of Rbl1/p107 protein as readout for E2F1 activity. Rbl1 belongs to the pocket protein family of proteins that also includes Rb and Rbl2/p130. Rbl1 has been described as an E2F1 target and is upregulated in a cell cycle-dependent manner (Beijersbergen, Carlee et al. 1995; Hurford, Cobrinik et al. 1997). However, in contrast to other E2F targets present in our list, Rbl1 has been shown to be already upregulated in mid G1, rather than at the G1/S boundary, providing an earlier readout for E2F1 activity (Bartek, Bartkova et al. 1997). Consistent with these data, I observed the transcriptional upregulation of Rbl1 at 20 hours after hepatectomy in S6/flox. ΔS6/flox livers showed instead a complete lack of Rbl1 upregulation both at

the RNA and at a protein level (Fig.12), providing evidence for a lack of E2F1-regulated transcription in this genotype.

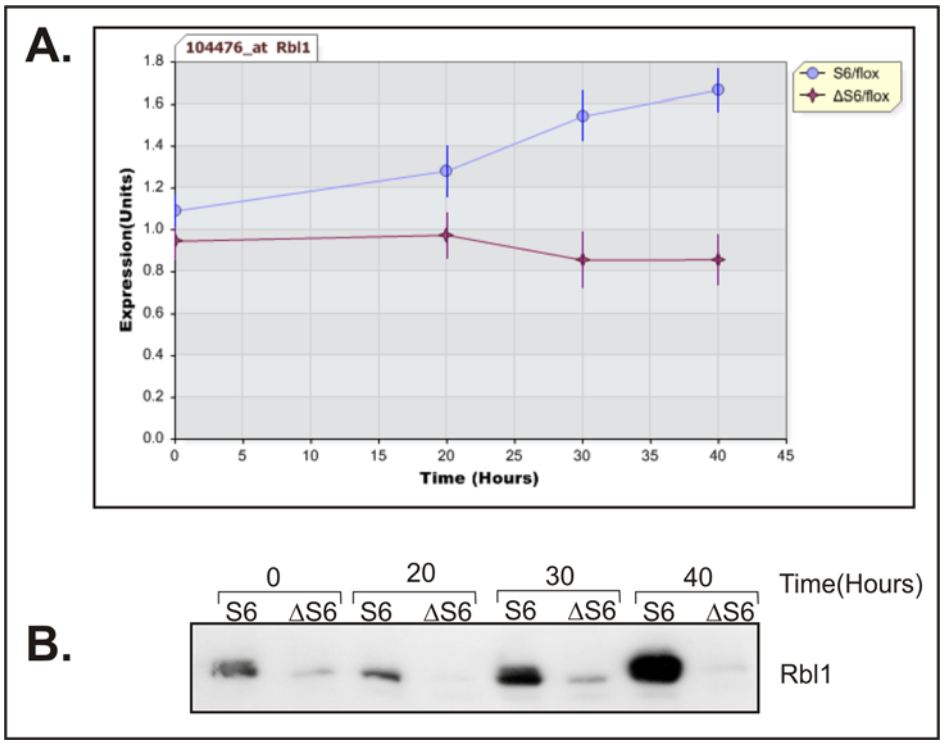


Fig.12 (**A**) Normalized microarray data of the expression levels of Rbl1/p107 in S6/flox (blue) versus ΔS6/flox (red) at different timepoints after hepatectomy. Each data point represents the average of three independent experiments. Error bars correspond to standard error. (**B**) Western blot showing Rbl1/p107 protein levels in S6/flox ΔS6/flox livers at different timepoints after hepatectomy.

o Rbl2 remains hypophosphorylated in ΔS6/flox livers

In cells entering the cell cycle, the upregulation of E2F1 relies on an auto-regulatory feed forward mechanism, where E2F1 induced its own expression through the binding of its own promoter (Johnson, Ohtani et al. 1994). This mechanism is based on the presence of an initial small pool of E2F1 protein and on the release of complexes containing pocket proteins that bind the promoter of E2F-responsive genes and actively repress their transcription (Johnson, Ohtani et al. 1994). The release of these inhibitory complexes is mediated by the phosphorylation of the pocket proteins from cyclin-dependent kinases (Cdks) and is a prerequisite for E2F1-mediated transcriptional

activation. Given the lack of E2F induction in ΔS6/flox, we tested the state of the pocket protein inhibitory complexes. In particular, we focused on Rbl2/p130 which, in a complex with E2F4, is known to mediate inhibition of transcription of E2F targets in resting livers (Garriga, Limon et al. 1998). Mitogenic stimuli trigger the release of the Rbl2/E2F4 inhibitory complex, through phosphorylation of Rbl1 and entry of the cells into G1 (Mayol, Garriga et al. 1995). We therefore decided to measure the levels of Rbl2 and its phosphorylation status during cell cycle progression. Using western blotting with an anti-p130-specific antibody we detected a slower migrating band, indicative of phosphorylation (Garriga, Limon et al. 1998), at 30 and 40 hours post hepatectomy, in extracts from livers of S6/flox but not in ΔS6/flox, suggesting a lack of Rbl2 phosphorylation in the livers of the latter.

Rbl2 has been reported to harbor 20 potential phosphorylation sites (Farkas, Hansen et al. 2002) and not all of them contribute to a detectable bandshift on a western blot. We therefore complemented this experiment with a more functional evaluation using electromobility-shift assays (EMSA). This assay measures the binding of protein complexes to a radioactively labeled oligonucleotide containing the E2F consensus-binding motif. It has been shown that, in resting livers, Rbl2/E2F4 complexes constitute the main binding activity to such an oligonucleotide (Garriga, Limon et al. 1998). These data were confirmed in our experiments, as shown in (Fig.13B). Incubation of the radiolabelled oligonucleotide with nuclear extracts from resting ΔS6/flox livers resulted in its upward shift, detectable through PAGE (Fig.13B, lane 2). The binding was competed by an excess of cold oligonucleotide, indicating the specificity of the interaction (Fig.13B, lane 3) and furthermore the presence of Rbl2 in the complex bound to the oligonucleotide was confirmed by the supershift obtained after pre-incubation with an anti-Rbl2 antibody (Fig.13B, lane 5). Using the same assay, we then analyzed the binding activity at different timepoints after hepatectomy for both S6/flox and ΔS6/flox. At 0 hours, binding of Rbl2/E2F4 complexes was detected for both ΔS6/flox and S6/flox. After hepatectomy, binding activity was gradually lost in S6/flox, as cells progressed into the cell cycle (Fig.13D) and in parallel with an increase in Rbl1 expression (Fig.13C). In ΔS6/flox, instead, at all timepoints analyzed the binding of Rbl2/E2F4 was indistinguishable from the one detected in resting liver. Altogether,

these experiments show that in ΔS6/flox livers, Rbl2 fails to be phosphorylated following hepatectomy: it remains in its DNA-bound conformation, therefore repressing the expression of the E2F target genes.

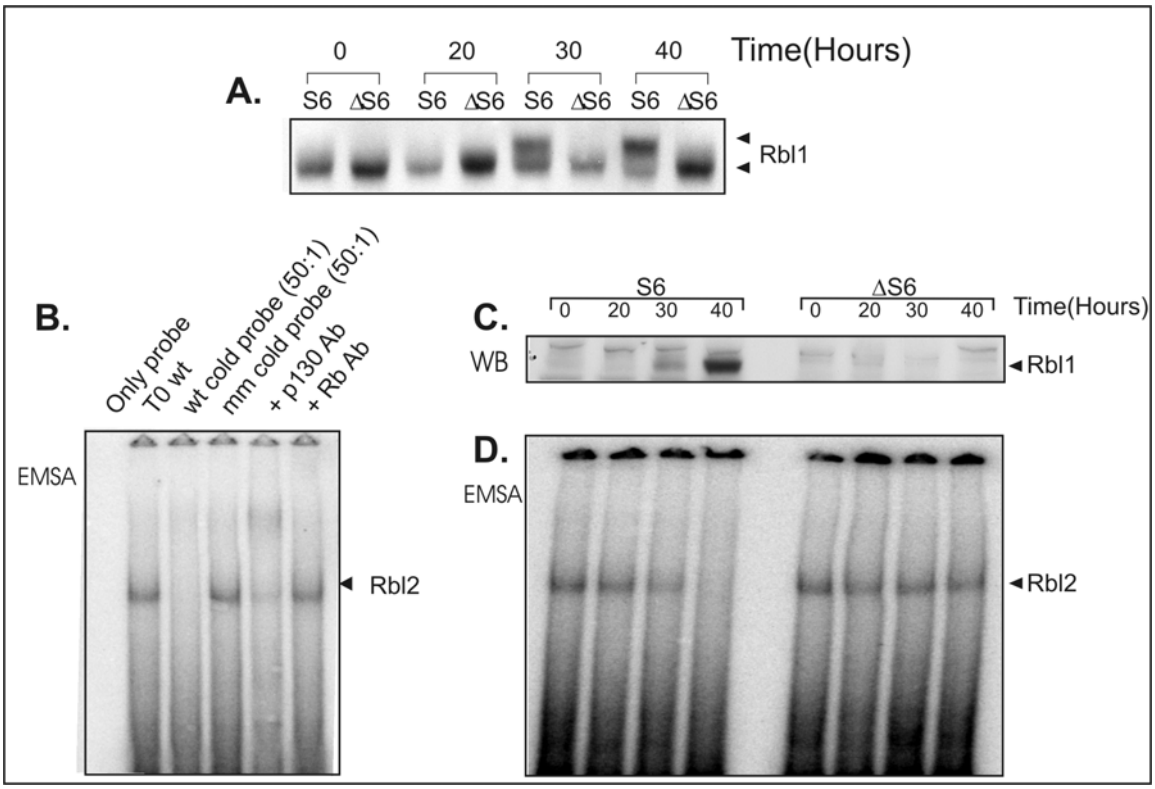


Fig.13 (**A**) Western blot showing Rbl2 levels in S6/floχ and ΔS6/flox at different timepoints after hepatectomy. (**B**) Electromobility shift assay to test specificity of binding and detection of Rbl2: *Lane1* E2 probe only (negative control); *Lane2* E2 probe + Time 0 S6/flox (positive control); *Lane3* competition with wt cold probe (50-fold excess; *Lane4* competition with mismatched cold probe (50-fold excess); *Lane5* supershift with anti-Rbl2 Ab; *Lane6* supershift with anti-Rb Ab. (**C**) Western blot showing Rbl1 levels in S6/floχ and ΔS6/flox at different timepoints after hepatectomy. (**D**) Electromobility shift assay to test the binding activity of the Rbl2 complex in S6/floχ and ΔS6/flox at different timepoints after hepatectomy

o **Cell cycle inhibitory genes are induced in ΔS6/flox livers**

As previously shown in (Fig.8), we also identified a second expression cluster, composed of genes which are expressed more in ΔS6/flox. In contrast to the CCIND, these genes are upregulated in both genotypes at 20 hours after hepatectomy, although at higher levels in ΔS6/flox livers. Furthermore, at subsequent timepoints, their expression is maintained at higher levels in ΔS6/flox livers, whereas it decreases in S6/flox livers. Because of these differences, we devised an alternative approach to

uncover candidate genes involved in the cell cycle block. We formed a list of genes that in our experiment were expressed more in ΔS6/flox at at least two different timepoints, since we reasoned that their differential expression over a longer time range could have an impact on the phenotype observed (Fig.14).

GeneSymbol	FoldChange	ChangePValue	Time (Hours)	Function
2610318G08Rik	1.70	5.30204E-05	30	---
	1.52	0.002868942	40	
Bax	1.86	0.007473629	0	induction of apoptosis by intracellular signals
	1.58	0.00135573	20	
Brix	1.93	0.002101344	30	ribosome biogenesis
	1.81	0.004325134	40	
Cdkn1a	3.87	0.00102304	0	regulation of cell cycle
	5.12	0.005218889	30	
E130105L11Rik	1.93	0.007452688	30	---
	2.00	0.007547999	40	
Eif5	1.50	0.00961459	0	regulation of translational initiation
	2.39	0.000403408	30	
	3.42	6.66179E-05	40	
Ephx1	1.81	0.00138372	0	proteolysis and peptidolysis
	1.62	0.006381256	30	
Ifrd1	4.54	0.000704678	30	differentiation
	4.69	0.001954447	40	
Phlda3	1.83	0.003219623	0	signal transduction
	2.29	0.009590066	20	
	2.01	0.001044018	40	

Fig.14 List of genes derived from microarray analysis that show an upregulation in ΔS6/flox at more than one of the timepoints collected. The genes selected show a fold change of >=1.5 and a change p-value of <0.01. The timepoints at which these genes show a significant differential expression are reported in the last column, together with the respective fold change and change p-value

Among the genes identified with such properties, Cdkn1a was of particular interest. This gene, also known as p21, codes for a cyclin-dependent kinase inhibitor known to affect cell cycle progression (Harper, Adami et al. 1993). P21 mRNA levels were already about four times higher in the liver of ΔS6/flox when compared to S6/flox at time 0, i.e. before the start of the regenerative response. The levels of p21 transcript increased steadily until 30 hours after hepatectomy in ΔS6/flox, whereas in S6/flox they peaked at 20 hours after hepatectomy and decreased at the subsequent timepoints. The p21 protein levels reflected the kinetics observed in the microarray data. In S6/flox, p21 protein levels increased after hepatectomy, as previously reported by other groups (Albrecht, Poon et al. 1998). In ΔS6/flox the levels of the p21 protein were higher than in S6/flox throughout the timecourse, except at 40 hours where, interestingly, p21 expression dropped to levels equivalent to S6/flox (Fig.15).

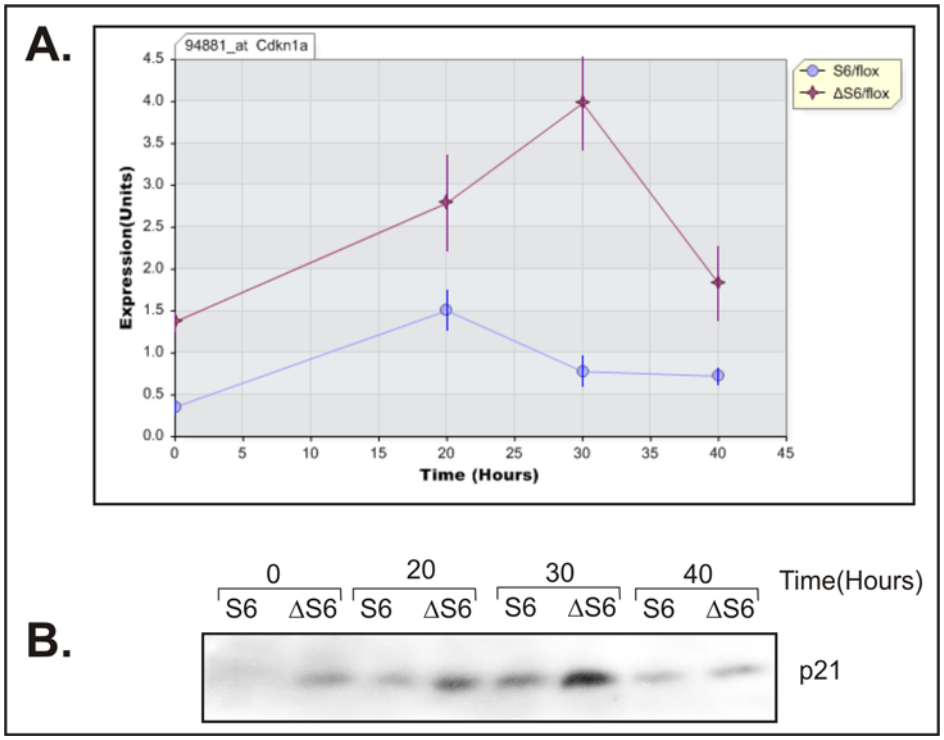


Fig.15 (**A**) Normalized microarray data of the expression levels of Cdkn1a/p21 in S6/flox (blue) versus ΔS6/flox (red) at different timepoints after hepatectomy. Each data point represents the average of three independent experiments. Error bars correspond to standard error. (**B**) Western blot showing Cdkn1a/p21 protein levels in S6/flox ΔS6/flox livers at different time points after hepatectomy

In many systems, transcription of p21 is known to be mediated by p53. We were therefore intrigued by the presence of Bax in the list, a well-known target of p53 (Miyashita and Reed 1995; Tang, Zhao et al. 1998). Given the involvement of p53 in a multitude of cellular checkpoints and its role in cell cycle regulation (Kuerbitz, Plunkett et al. 1992), I assessed whether other p53-targeted genes were differentially expressed. In addition to p21 and Bax, we found that the levels of Gadd45γ, cyclin G and Mdm2 mRNA were upregulated in ΔS6/flox at at least one of the timepoints, with a fold change of >1.5 and a change p-value of <0.05 (Fig.16). These results suggested a possible role of p21 and/or the p53 pathway in the ΔS6/flox S-phase block.

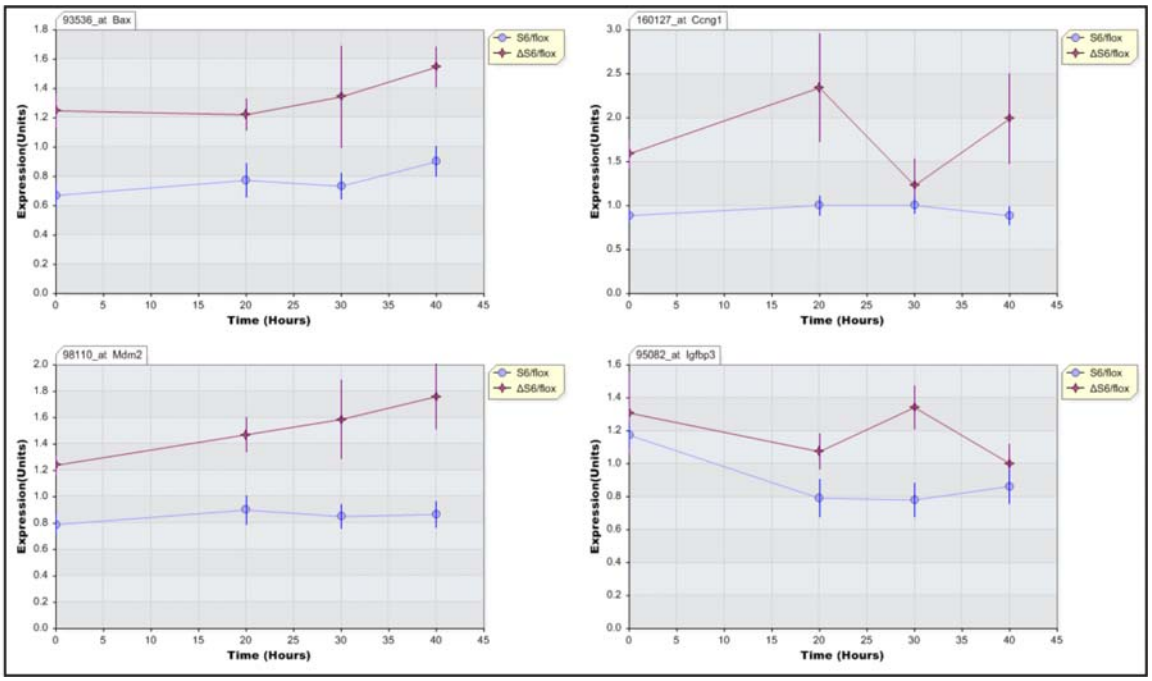


Fig.16 Normalized microarray data of the expression levels of p53 targets: Bax, cyclin G, Mdm2 and IGFBP3, at different timepoints after hepatectomy. Genotypes are expressed with different colors S6/flox (blue), ΔS6/flox (red). Each data point represents the average of three independent experiments. Error bars correspond to standard error

o Deletion of S6 is sensed before hepatectomy

Surprisingly even at time 0 the levels of p21 were higher in ΔS6/flox livers. The sample collected at time 0 corresponds to the piece of liver removed during the hepatectomy and is therefore subject to surgical stress. To exclude the possibility of p21 levels rising in response to the surgical procedure, I collected livers from ΔS6/flox mice in the absence of any manipulation (timepoint –1 hour) and measured the levels of p21. Even before any surgical procedure, p21 levels were higher in ΔS6/flox livers (Fig.17A).

In order discard the possibility that p21 is upregulated because of the different genetic background of these mice (ΔS6/flox mice contain an Mx-Cre transgene), I collected liver samples from ΔS6/flox Mx-Cre- and S6/flox Mx-Cre+ mice before the induction of the S6 excision (timepoint –72 hours). Using microarray analysis, I compared the samples derived from the –72 and –1 timepoints. At –72 hours the levels of p21 were identical in the two genotypes (Fig.17B). Upon deletion of S6, but before any surgical procedure, p21 levels rose in ΔS6/flox livers. In addition to a rise in p21 levels, livers

collected before surgery showed the presence of the 34S rRNA precursor (Fig.17C), which is indicative of a defective ribosome biogenesis caused by S6 deletion (Volarevic, Stewart et al. 2000). From this data I conclude that even before hepatectomy ΔS6/flox livers show a defect in ribosome biogenesis and an increased level of p21.

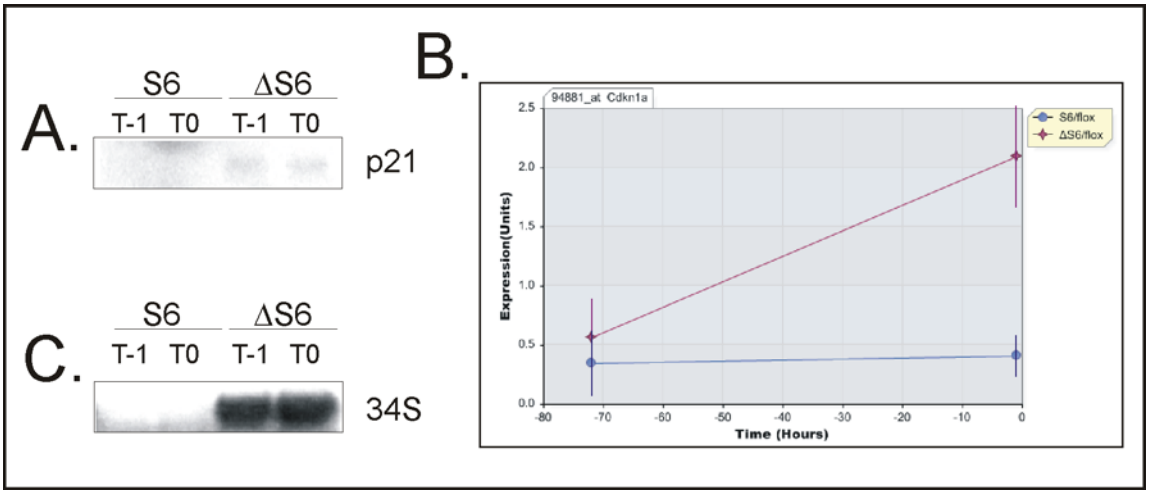


Fig.17 (**A**) Expression levels of p21 mRNA in S6/flox and ΔS6/flox liver before surgery (T-1) and beginning of surgery (T0). (**B**) Graph showing the microarray expression data of p21 in S6/flox (CRE+) and ΔS6/flox (CRE-) livers, before (T-72) and after (T-1) the induction of S6 deletion by pIpC injection. (**C**) Expression levels of the 34S rRNA precursor in S6/flox and ΔS6/flox liver before surgery (T-1) and at the beginning of surgery (T0).

Establishment of an in vitro model for studying how defects in ribosome biogenesis affect the cell cycle: results

The other aim of this part of the project was to establish an in vitro system that could mimic the disruption of ribosome biogenesis and the S-phase block observed in the ΔS6/flox mice. The establishment of such a system was prompted both by: (1) the inherent difficulty to dissect molecular pathways in vivo using a mouse model; and (2) the need for a model to test if the phenotype observed in vivo was cell autonomous. In order to establish such a system, I tested different experimental procedures to obtain S6-deficient cells in vitro and examine whether these systems reproduced the in vivo model. After successfully establishing an in vitro system, I characterized the mechanisms linking the loss of S6 to the block in S phase entry.

o Use of primary hepatocyte cultures as models for the ΔS6/flox proliferation phenotype

I first tested experimental procedures which relied on the use of cells derived from the S6/flox mice, since I thought they would be more likely to mimic the mouse model. I isolated mouse embryo fibroblasts (MEFs) from S6/flox animals, which harbored the Mx-Cre recombinase transgene (Cre+), and stimulated Cre expression by adding either polyinosinic-polycytidilic acid (pIpC), an inducer of IFN-α, or IFN-α itself to the culture medium. Unfortunately, using this approach I was unable to induce disruption of the S6 gene.

Therefore, I tested procedures in which I could induce the deletion of the gene in vivo and subsequently isolate and culture the S6 knockout cells. The injection of pIpC into S6/flox Cre+ mice causes full deletion of the gene in liver but not in other tissues (Volarevic, Stewart et al. 2000). Following this observation, I cultured hepatocytes isolated from newborn mice. Livers of newborn mice are not fully structured, and hepatocytes can be isolated through a simple collagenase digestion. Furthermore, in contrast to adult hepatocytes, hepatocytes from newborns have a higher proliferating potential in vitro (Hamamoto, Kamihira et al. 1999), undoubtedly an advantage when

studying cell cycle progression. This procedure required, however, the induction of the S6 deletion in newborn mice. Setting aside the difficulty of injecting mice of such a small size, both pIpC and IFN-α induced only a partial deletion of S6 in the livers of newborns.

I next devised a procedure to isolate hepatocytes from adult mice, knowing that a pIpC injection could induce full deletion of S6 in liver. I successfully isolated and cultured hepatocytes from both ΔS6/flox and S6/flox. Hepatocyte proliferation in response to insulin, in conjunction with either epidermal growth factor (EGF) or hepatocyte growth factor (HGF), was tested by measuring the incorporation of ^{3}H-labelled thymidine. Both treatments induced an increase in DNA synthesis in both ΔS6/flox and S6/flox hepatocytes. Consistent with the results in vivo, less thymidine was incorporated in ΔS6/flox hepatocytes (Fig.17), indicative of a defect in S-phase entry.

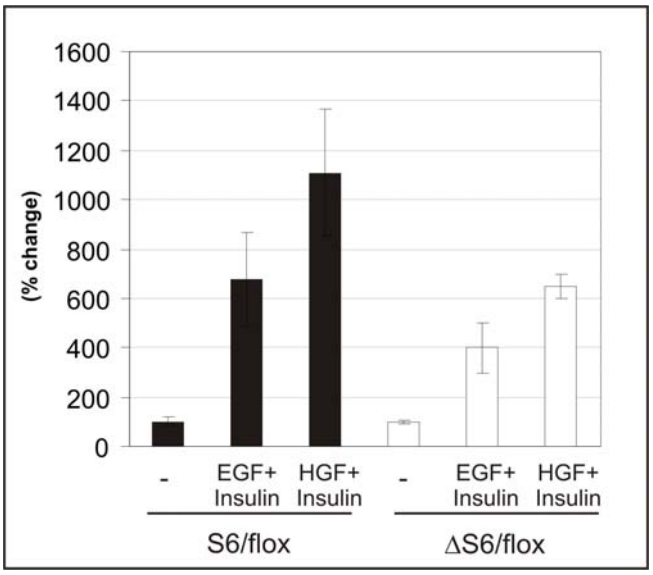


Fig.17 Levels of thymidine incorporation in S6/flox and ΔS6/flox primary adult hepatocytes. The cells were seeded for 4 hours after perfusion from the liver and starved for 18 hours, after which ^{3}H-thymidine was added. The cells were collected 72 hours after thymidine addition. Hepatocyte proliferation was induced by addition of Insulin (2 mg/ml) and either EGF (20 □g/ml) or HGF (10 □g/ml) after the period of starvation

- **SiRNA-mediated knockdown of S6 in A549 cells**

Primary adult Hepatocytes provided a useful tool to test if the effects caused by deletion of the S6 gene were cell autonomous, but they also presented some limitations because of their limited lifespan and proliferation rate. Therefore, I turned towards the use of siRNA as a method to knockdown expression of the S6 ribosomal protein. SiRNA-mediated knockdown involves the use of small RNA oligomers to trigger the

degradation of a target mRNA (Elbashir, Lendeckel et al. 2001). I transfected A549 cells with two different siRNAs directed against the human S6 ribosomal protein and compared them for the expression of S6 in cells which were either untreated, or treated with a non-silencing (NS) siRNA control, to correct for any non-specific effects caused by the siRNA treatment itself. Both of the siRNAs for S6, but not the NS siRNA, caused depletion of S6 mRNA at 24 hours after transfection (Fig.18A): these changes were also reflected at the protein level (Fig.18B). For all the subsequent experiments I used S6 siRNA (#2), which proved to be more efficient. Immunohistochemical analysis of S6 siRNA-treated cells using an antibody against S6 revealed a decrease in the staining of S6 in the nucleolus but not in the cytoplasm, where S6 is associated with ribosomes (Fig.18C).

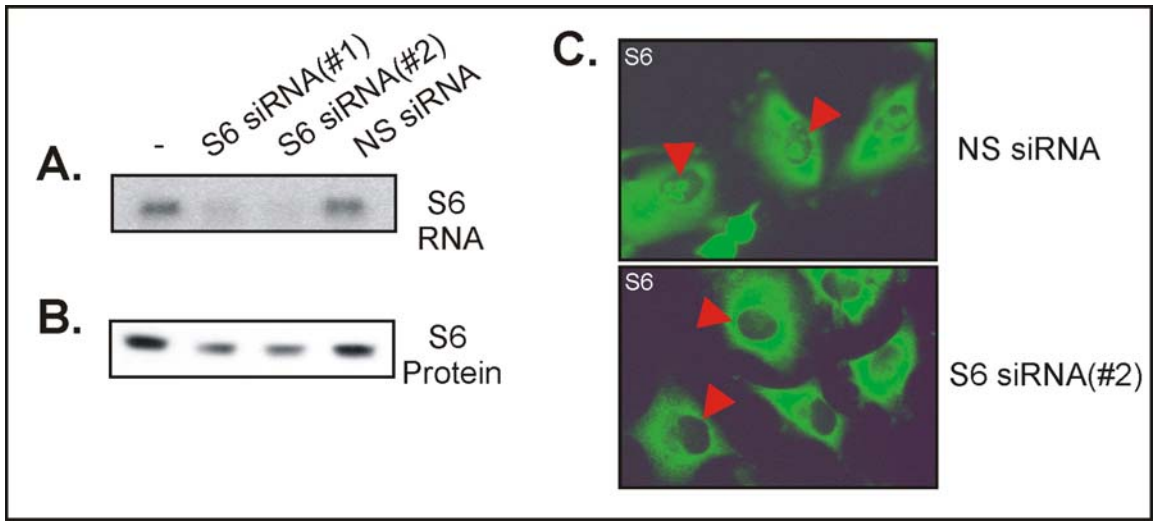


Fig.18 (**A**) RNA levels of ribosomal protein S6 24 hours after transfection with (i) nothing; (ii) 15nM S6 siRNA (#1); (iii) 15nM S6 siRNA (#2); (iv) 15nM non-silencing siRNA. (**B**) S6 protein levels using the same treatments as above. (**C**) Immunohistochemistry using S6 antibody: (top) cell treated for 24 hours with 15nM non-silencing siRNA; (bottom) cell treated for 24 hours with 15nM S6 siRNA(#2)

- **Knockdown of S6 by siRNA causes a reduction in the number of 40S ribosomal subunits and accumulation of a 34S precursor of 18S rRNA**

With the possibility of lowering S6 levels in vitro, I testedthe extent to which this system mimicked the phenotype of the ΔS6/flox mouse. In vivo, deletion of S6 affects the ability of hepatocytes to synthesize new 40S ribosomal subunits (Volarevic, Stewart et

al. 2000) due to a defect in rRNA processing that leads to the accumulation of a 34S precursor of the 18S rRNA. In contrast, no effect is observed on the production of 28S rRNA, which is associated with the 60S large ribosomal subunit (Volarevic, Stewart et al. 2000). To determine whether the knockdown of S6 in vitro caused a similar effect, I used polysome profiles of cytoplasmic extracts from S6 siRNA-treated cells. Like in ΔS6/flox animals, the depletion S6, as compared to NS siRNA-treated cells, caused a decrease in the numbers of both 40S subunits and polysomes, indicative of a possible effect on translation rates (Fig.19A). As a complement to this experiment, in order to determine the effects of S6 depletion on the production of 18S rRNA, I also used RNA gradients to measure mature rRNA levels. If we consider the ratio between 18S and 28S rRNA in cells treated with NS siRNAs to be 1, the value was clearly altered in cells treated with the S6 siRNA, due to a drop in production of 40S ribosomal subunits (Fig.19B). The defect in production of 18S was due to a failure in processing the 34S rRNA precursor as shown by northern blot using a specific probe for the 34S RNA precursor. The 34S RNA was observed in cells treated with the S6 siRNA but not in cells treated with NS siRNA.

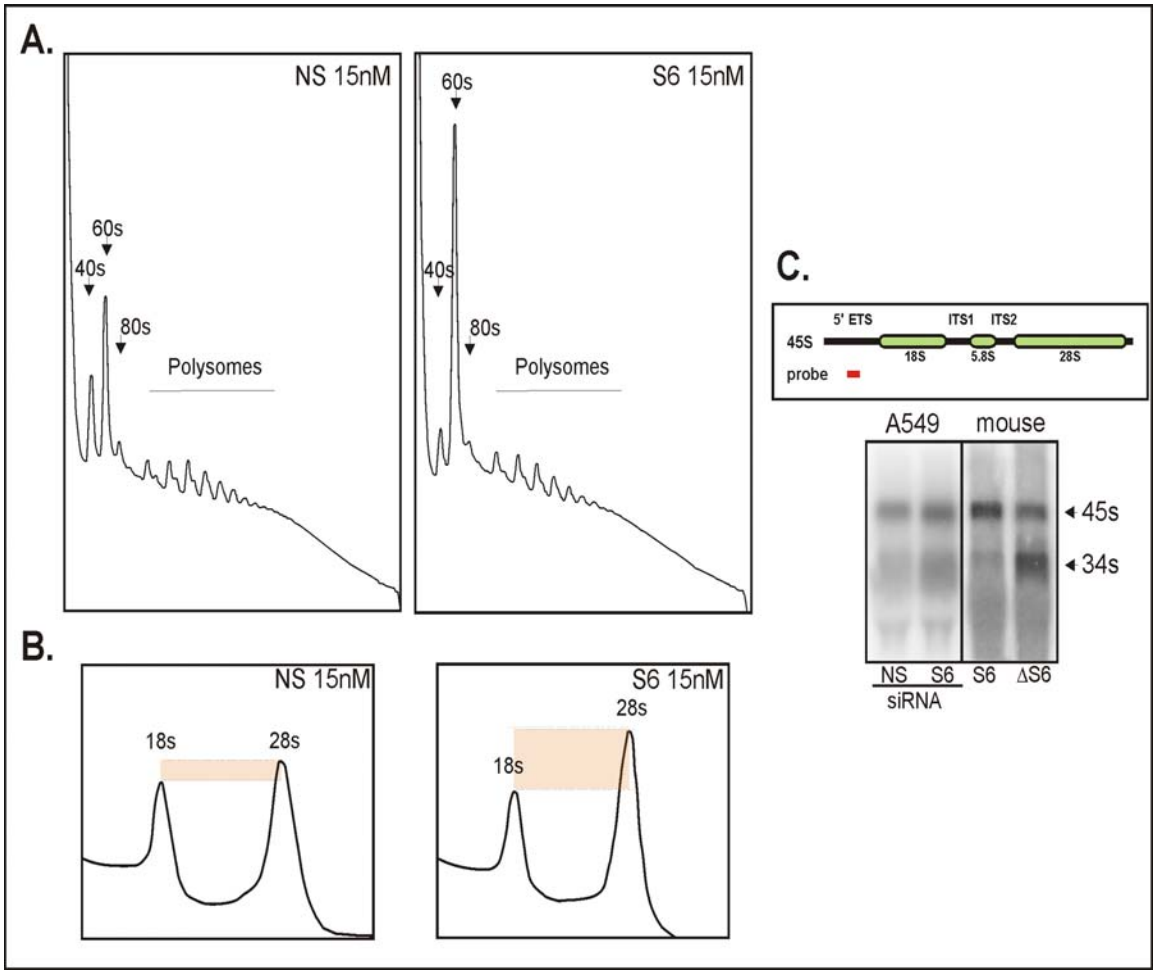


Fig.19 (**A**) Polysome profiles of cells treated with either non-silencing (NS) siRNA or 15nM S6 siRNA (#2) for 24 hours. (**B**) rRNA profile showing 18S and 28S rRNA of cells treated with either non-silencing (NS) siRNA or 15nM S6 siRNA (#2) for 24 hours. The red area represents the difference in height between the two rRNA peaks. (**C**) Northern blot against 34S rRNA precursor. In vitro experiments with siRNA and in vivo experiments with S6/flox mice after hepatectomy were compared: the diagram shows the location of the designed probe on the 45S rRNA precursor. Two different probes were used to detect the human and the mouse 34S. *Lane 1* NS siRNA 15nM, *Lane 2* S6 siRNA (#2) 15nM, *Lane 3* S6/flox liver 30 hours post-hepatectomy, *Lane 4* ΔS6/flox liver 30 hours post-hepatectomy

- **Knockdown of S6 causes a cell cycle block and induces a p21/p53 response**

After establishing that in vitro depletion of S6 causes a defect in ribosome biogenesis, I went on to test whether, as seen in ΔS6/flox, this defect could lead to a cell cycle block. To answer this question I perfomed a FACS analysis of cells using BrdU pulse-labelling. The cells were treated with BrdU for 10 minutes and then collected. BrdU is a modified nucleotide that integrates only in the DNA of cells which are undergoing DNA synthesis.

FACS analysis of the cells fixed and incubated with a fluorescent anti-BrdU antibody allows the quantification of BrdU-positive cells, providing a measure of the percentage of cells undergoing S phase. The cells were also treated with propidium iodide, to stain their DNA and allow the measurement of their distribution in the G1 and G2 phases of the cell cycle. In a first experiment, the cells were treated with either NS or S6 siRNA, and collected at 24 and 48 hours after treatment. At both timepoints, S6-treated cells displayed a decrease in S-phase and an increase in G1-phase population when compared to NS siRNA-treated cells (Fig.20). In agreement with these results, a biochemical characterization showed that the levels of Rbl1, a marker of cell proliferation, were lower in S6 siRNA-treated cells (Fig.21). The correlation between these data and the findings of the in vivo mouse model prompted us to test whether p21 and its upstream inducer, p53, were upregulated in cells treated with the S6 siRNA. Indeed, western blot analysis showed that both p21 and p53 were heavily induced in response to the S6 knockdown (Fig.21).

- **P21 and p53 can rescue the S6-induced cell cycle block in vitro**

To investigate whether p21 or p53 could be responsible for the block in S phase entry induced by downregulation of S6, I tested whether the simultaneous depletion of either of them could rescue the defect in cell proliferation caused by the lack of S6. For this purpose, A549 cells were treated with S6 siRNA in combination with either a p21 or p53 siRNA. At 24 and 48 hours after transfection, we measured the number of cells entering S phase through BrdU incorporation and FACS analysis. As a control, I used cells treated with S6 siRNA in combination with NS siRNA to compensate for any unspecific effect caused by the total amount of siRNA present (30nM total). Interestingly, depletion of p21 in S6 knockdown cells resulted in a partial rescue of S-phase entry. The effect, however, was even more pronounced in cells where p53 was downregulated, as shown by the number of cells in S phase, which nearly corresponds to the one of NS siRNA-treated cells (Fig.22A and B). I performed a biochemical analysis at 24 hours in order to determine, the efficacy of the depletion of the targeted proteins and to analyze the expression of p107. In Cells treated with p53 siRNA p21 protein levels were diminished, showing that, in this system, the upregulation of p21 is solely dependant on p53. With regards to proliferation, Rbl1 levels paralleled the results

obtained with the BrdU incorporation, showing rescue of the S6 knowckdown phenotype in cells treated with p53 siRNA (Fig.22C).

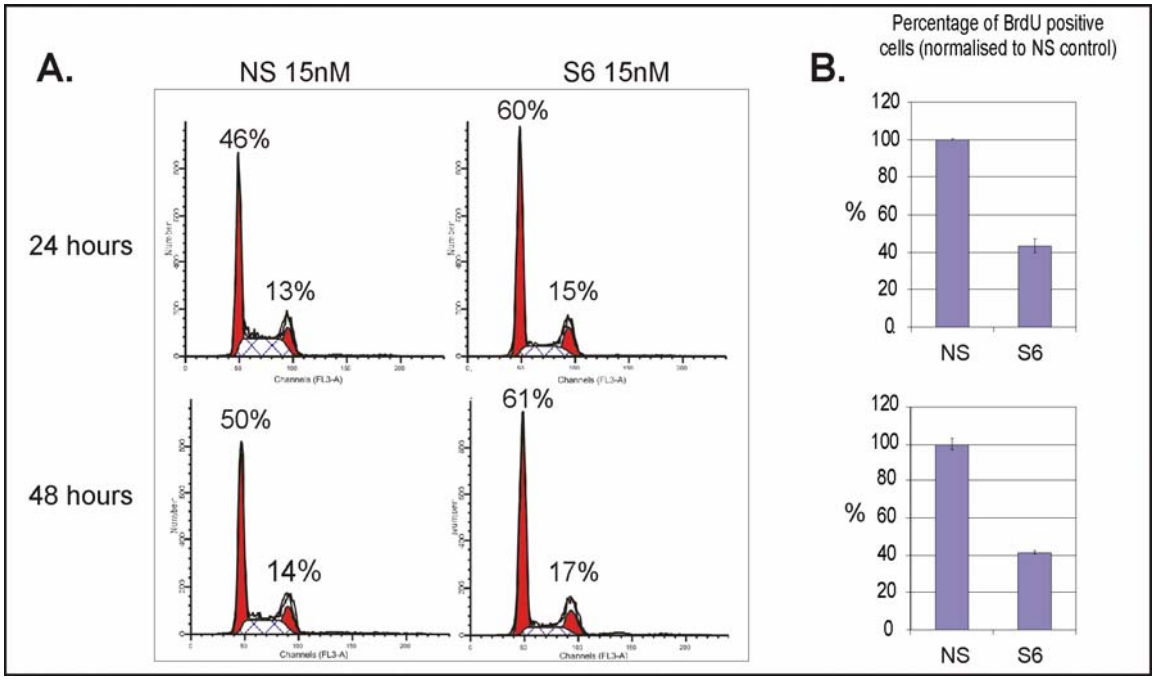


Fig.20 Cell cycle analysis of cells treated with either 15nM non-silencing (NS) siRNA or 15nM S6 siRNA. (**A**) Profile showing ploidy (x-axis) over number of cells (y-axis) at 24 or 48 hours after treatment. Values represent the percentage of cells in either G1 or G2 phase and are representative of one experiment. Replicate experiments gave similar results. (**B**) Percentage of BrdU-positive cells normalized to the NS control for each timepoint: values derive from average of three independent replicates

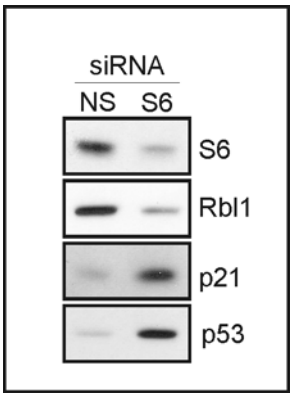


Fig.21 Biochemical analysis of cells treated with either 15nM non-silencing (NS) siRNA or 15nM S6 siRNA for 24 hours. The levels of S6, Rbl1, p21 and p53 are shown

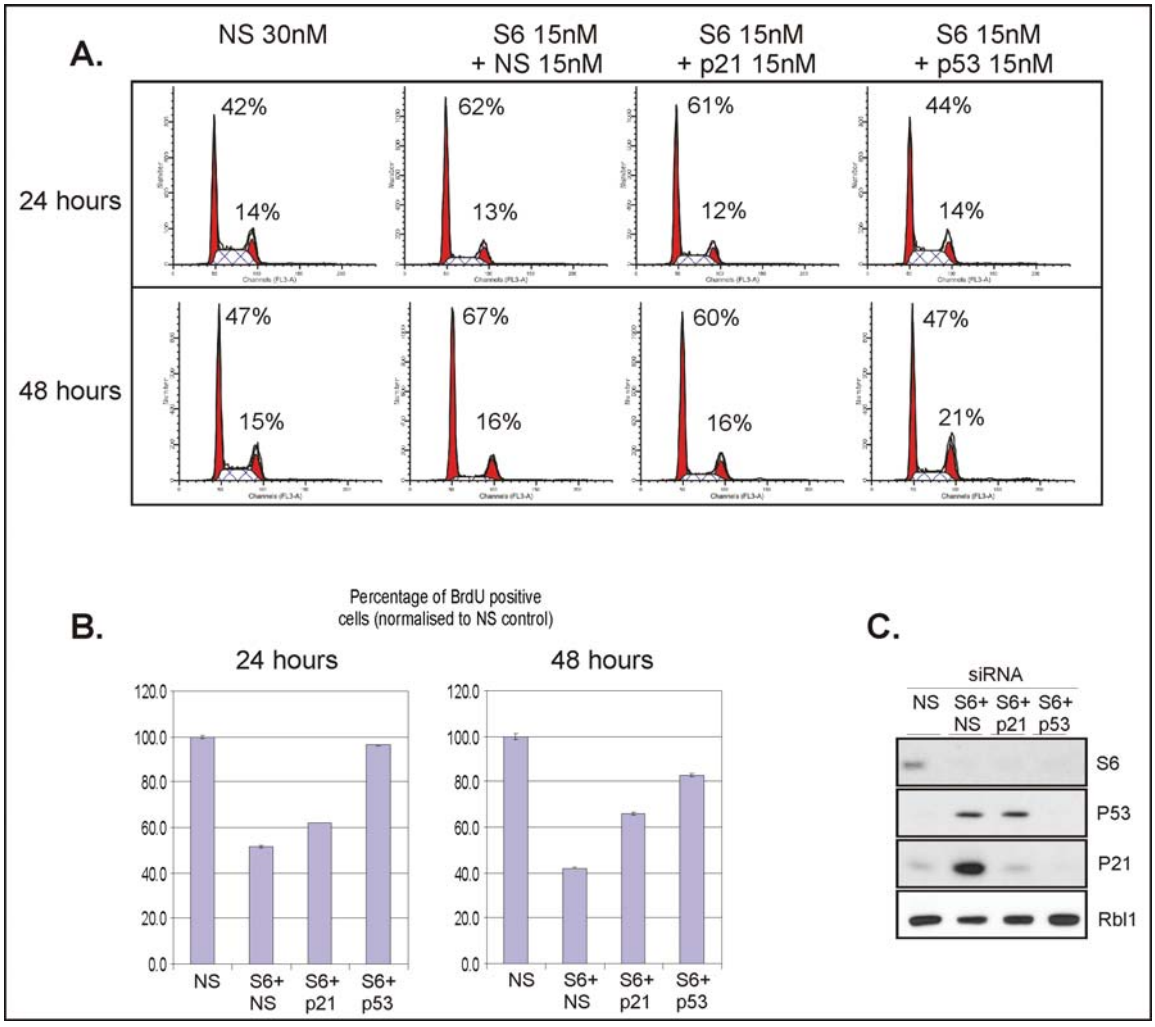


Fig.22 Cell cycle analysis of cells treated with either: (i) 30nM non-silencing (NS) siRNA; (ii) 15nM S6 + 15nM NS siRNA; (iii) 15nM S6 + 15nM p21 siRNA; (iv) 15nM S6 + 15nM p53 siRNA. (**A**) Profile showing ploidy (x-axis) over number of cells (y-axis) at 24 or 48 hours after treatment. Values represent the percentage of cells in either G1 or G2 phase and are representative of one experiment. Replicate experiments gave similar results. (**B**) Percentage of BrdU-positive cells normalized to the NS control for each timepoint: values derived from the average of three independent replicates. (**C**) Western blot showing levels of S6, p53, p21 and Rbl1 24 hours after siRNA transfection

Having established that depletion of both p21 and p53 can revert the block in S phase elicited by S6 downregulation, I assessed whether the same treatment also rescued cell division. For this purpose, I counted the number of cells at 24, 48 and 72 hours after transfection with different combinations of siRNAs. Surprisingly, in conditions that rescued S-phase entry the number of cells was much lower than the NS treated cells, (Fig.23). Frthermore at 48 hours after transfection, cells treated with S6 and p53

siRNAs showed an increase in G2 phase of about 6-10% as compared to NS siRNA-treated cells. Altogether evidence is provided showing that the effect of S6 depletion on the cell cycle is mediated by the activation of a p53-dependent pathway and, in part, upregulation of p21. However, although downregulation of p53 in S6-depleted cells rescues S-phase entry, it is not sufficient to rescue cell division.

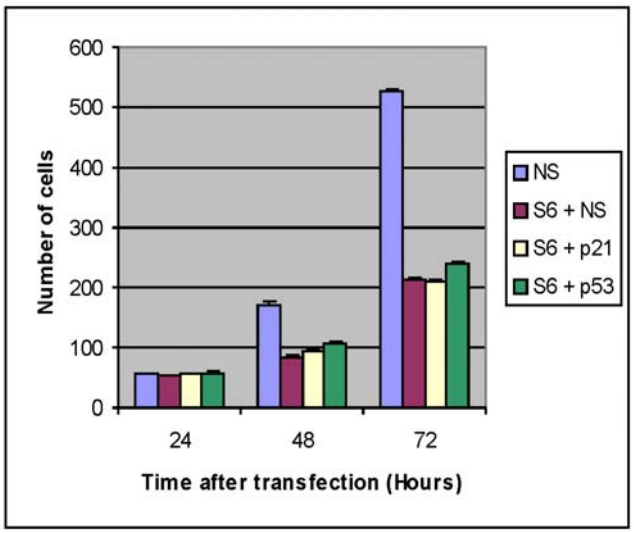


Fig.23 Histogram showing number of A549 cells present in a 10cm culture dish after treatment with either: (i) 30nM non-silencing (NS) siRNA; (ii) 15nM S6 + 15nM NS siRNA; (iii) 15nM S6 + 15nM p21 siRNA; (iv) 15nM S6 + 15nM p53 siRNA. Cells were collected at 24, 48 or 72 hours after transfection. The Values represent the average of three independent replicates

- **Knockout of p21 rescues S6-induced cell cycle block in ΔS6/flox livers**

In view of the results obtained in vitro, we were interested to see if the hepatocyte proliferation of ΔS6/flox animals could be rescued by the downregulation of p21. For this purpose, we generated S6/flox mice positive for the Mx-Cre transgene and carrying a homozygous p21 gene deletion. In so doing, upon induction of Cre expression by injection of pIpC, we obtained ΔS6/flox p21-/- mice. We performed partial hepatectomy on these mice and assessed the capacity of the liver cells to enter the first round of S phase. For this purpose, at we injected the mice intraperitoneally with BrdU following hepatectomy. Livers were then extracted at 40 hours after hepatectomy and conditioned for histoimmunochemistry with anti-BrdU antibodies. Interestingly, we found BrdU positive hepatocytes in the liver of the ΔS6/flox p21-/- mouse, indicating the presence, albeit limited, of proliferation. (Fig.24). These data suggest that, also in vivo,

p21 is a crucial player in the cell cycle checkpoint elicited by inhibition of ribosome biogenesis.

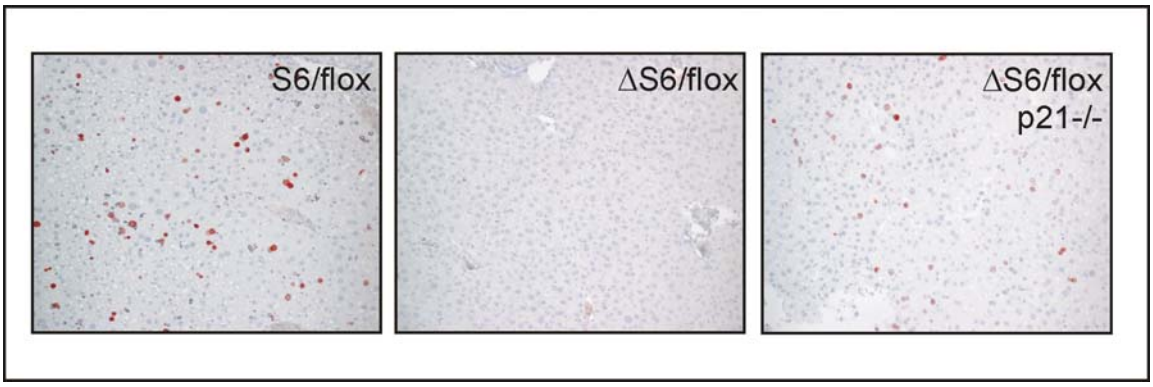


Fig.24 Immunohistochemical analysis of BrdU-positive cells in the livers of: (i) S6/flox; (ii) □S6/flox; and (iii) ΔS6/flox p21-/- 40 hours after partial hepatectomy. Proliferating cells are shown in red

o The ribosome biogenesis checkpoint is triggered by other ribosomal proteins

The data presented so far concern the effects on cell cycle progression of the depletion of S6, one of about 80 ribosomal proteins. To test whether the abrogation of other ribosomal proteins would result a cell cycle block, I performed experiments with siRNA directed against L7a and S23, two ribosomal proteins of the large and small ribosomal subunit respectively. The depletion of either of these two proteins caused an increase in the expression of p21 mRNA, as shown by northern blot analysis (Fig.25).

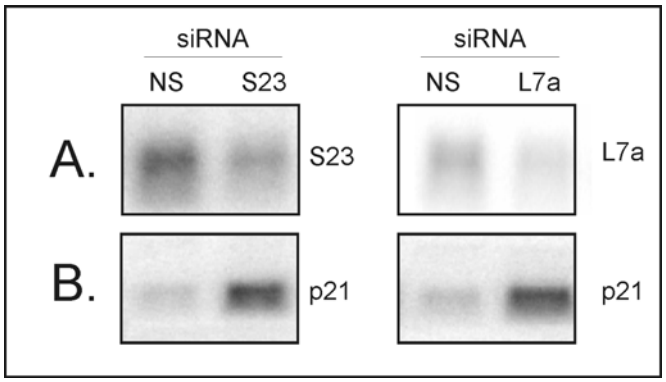


Fig.25 (**A**) mRNA levels of S23 and L7a after 24 hours' treatment of A549 cells with either: (i) 15nM non-silencing siRNA; or (ii) 15nM siRNA against S23 or L7a. (**B**) Levels of p21 mRNA in response to the same siRNA treatments

Discussion

S6 deletion in mouse livers inhibits the transcription of genes involved in S-phase progression

The microarray analysis of S6/flox versus ΔS6/flox livers following hepatectomy shows the differential expression of a number of genes involved in cell cycle progression. Therefore, the lack of cyclin E expression observed by (Volarevic, Stewart et al. 2000) is not the sole event responsible for the absence of cell proliferation in ΔS6/flox mice. Instead, the expression of a whole plethora of factors involved in cell cycle control and DNA metabolism fails to be induced in ΔS6/flox livers. The number of genes differentially expressed increases proportionally with time, in particular between 20 and 30 hours after hepatectomy. Since the 30-hours timepoint corresponds to the mid-to-late G1 phase, as shown by the rise in Rbl1 mRNA, the shift in the number of differentially expressed genes is most probably due to the upregulation of the cell cycle machinery in S6/flox but not in ΔS6/flox mice. It is not possible to rule out, however, a defect in other cellular processes that do not rely on transcriptional regulation. In ΔS6/flox livers, the correlation between the lack of expression cell cycle inducer genes, such as the cyclins, and the induction of cell cycle inhibitors, like p21, hints at the existence of a checkpoint mechanism that acts coordinately in response to S6 deletion and to the inhibition of ribosome biogenesis.

S6 inhibits E2F1-mediated transcription of genes involved in cell cycle progression

Transcription of most of the genes that fail to be upregulated in ΔS6/flox has been reported to be dependent, at least partly, on E2F transcription factors. During the G1 to S transition, there is an increase in E2F1-dependent transcription, as shown in S6/flox livers at 40 hours. In ΔS6/flox livers, a defect in E2F1 activity is suggested by the lack of E2F1 mRNA upregulation, as shown both by microarray and northern blot analysis and by the absence of induction of direct transcriptional targets of E2F1, such as Rbl1.

Hepatocytes in a resting liver are quiescent and confined in the G0 phase of the cell cycle. Partial hepatectomy triggers a proliferative response which requires the gradual exit from G0, an event largely dependent on the phosphorylation of the pocket protein Rbl2 (Mayol, Garriga et al. 1995; Garriga, Limon et al. 1998). Our experiments show that, after partial hepatectomy, Rbl2 remains hypophosphorylated in ΔS6/flox livers and maintains a DNA-bound conformation, as shown by electromobility shift assays. The presence of Rbl2/E2F4 complexes on the promoters of E2F-responsive genes would sterically hinder the binding of E2F1 and, therefore, its ability to activate transcription (Grana, Garriga et al. 1998). Since E2F1 activates its own transcription in a feed-forward mechanism, its expression will also be inhibited in these conditions.

Cyclin E is not sufficient to rescue the cell cycle arrest caused by S6 depletion

Together with Cdk4 and Cdk6, Cdk2 participates in the phosphorylation of pocket proteins such as Rbl2 and the consequent relief of their inhibition. Since Cdk(4,6) activity is not influenced by the deletion of S6 (Volarevic, Stewart et al. 2000), we can argue that a lack of Cdk2 activity could cause the defect observed in cell cycle progression. This hypothesis has been confirmed by a series of experiments performed in our lab (S. Fumagalli, unpublished results), which aimed at rescuing the S-phase entry in these livers by overexpressing cyclin E. The rationale behind this approach is that ectopic expression of cyclin E can induce S-phase entry even without the activation of the E2F pathway (Lukas, Herzinger et al. 1997). Two different approaches were used. The first relied on the use of adenovirus expressing human cyclin E, while in the other approach the ΔS6/flox mice were crossed with mice containing a knock-in of the cyclin E cDNA on the cyclin D locus (Geng, Whoriskey et al. 1999). Both systems successfully induced cyclin E expression. However, surprisingly the block in S-phase entry is maintained in these conditions. Co-immunoprecipitation experiments showed the presence of cyclin E/Cdk2 complexes, but a lack of any associated Cdk2 activity.

Upregulation of p21 inhibits S-phase entry in ΔS6/flox livers

Given the failure of cyclin E to rescue the phenotype, it is plausible that a lack of Cdk2 activity, rather than the absence of cyclin/Cdk2 complexes, would cause the inhibition of cell cycle progression. This was supported by the later observation that levels of p21, a Cdk2 inhibitor, are higher in ΔS6/flox livers, as shown by data from both microarray analysis and western blotting. It must be said that we initially excluded a role for p21 in this inhibition, since previous reports (Volarevic, Stewart et al. 2000) showed a normal regulation of p21 protein levels in ΔS6/flox livers. The reason for this discrepancy lies in the timing of p21 expression. Volarevic et al. measured the levels of p21 at 0 and 40 hours after hepatectomy and noticed the same extent of upregulation of the protein in both S6/flox and ΔS6/flox livers. This data confirms this observation but provides evidence that at timepoints preceding 40 hours the levels of p21 are markedly higher in ΔS6/flox livers.

To determine whether p21 is involved in the block of S-phase entry ΔS6/flox mice were crossed with p21 knockout mice. When subject to hepatectomy, ΔS6/flox p21-/- livers show proliferating cells 40 hours after surgery, albeit to a lower extent than the control S6/flox livers. p21 therefore participates in the G1 block caused by the deletion of S6. Although the experiments have yet to be performed, we expect that Cdk2 activity will be recovered in the absence of p21. The reason why the ΔS6/flox p21-/- liver did not provide a full rescue could be attributed to the participation of other proteins in this response or by the different timing of S-phase entry in this genetic background. Experiments with different kinetics should provide a clearer answer to this issue.

Defects in ribosome biogenesis are sensed early

p21 mRNA is already upregulated in resting ΔS6/flox livers. Similarly, in the same situation the accumulation of the 34S precursor of the 18S rRNA is detected. Both of these observations reflect that S6 depletion causes a defect in ribosome biogenesis and subsequent activation of the checkpoint, even in the absence of proliferative stimuli. In this respect, it has to be emphasized that ribosome biogenesis is a continuous event which takes place also in resting hepatocytes, and its perturbation, upon depletion of S6, would therefore be sensed in the period preceding partial hepatectomy. If this is the

case, however, why would there be an induction of cyclin D/Cdk4 activity in ΔS6/flox livers after hepatectomy? (Volarevic, Stewart et al. 2000). It is possible that the role of cyclin D/Cdk4 activity in this context is dissociated from its function in cell cycle progression, and that it participates instead in other cellular processes, such as cell growth or dedifferentiation. This hypothesis is supported by the fact that cyclin D is dispensable in the liver regenerative response (Geng, Whoriskey et al. 1999). Alternatively, induction of cyclin D and activation of Cdk4 could result from cellular events which are not sensitive to alterations in ribosome biogenesis.

Deletion of S6 by siRNA in A549 cells mimics the ΔS6/flox liver phenotype

In order to build an in vitro system that mimics the ΔS6/flox mouse model, siRNA technology was used to deplete S6 in the A549 cell line. The siRNA was effective in the downregulation of both S6 mRNA and protein, even at 24 hours after transfection. Interestingly, the immmunohistochemical analysis reveals that the siRNA treatment mainly affects the nucleolar pool of S6, whereas S6 cytoplasmic staining does not seem to decrease significantly. This observation probably reflects the depletion of nascent S6 proteins, which are synthesized in the cytoplasm and then transported in the nucleolus for 40S assembly. Instead, the cytoplasmic staining is due to the S6 protein associated with ribosomes, which are notoriously stable, although analysis of the polysome profiles would suggest that the content of the 40S ribosomal subunit decreases upon S6 depletion. Depletion of S6 by siRNA in A549 cells causes all the phenotypical characteristics associated with the deletion of the S6 gene in vivo, such as the impaired synthesis of the 40S ribosomal subunit and the concomitant accumulation of a 34S precursor of the 18S rRNA (Volarevic, Stewart et al. 2000). Furthermore, when compared to cells treated with non-silencing siRNA, cells where S6 was downregulated accumulated in G1, due to a defect in progression to S phase.

Deletion of S6 causes a p53-dependent cell cycle arrest

Similarly to what is observed in ΔS6/flox livers, downregulation of S6 by siRNA in A549 cells triggers the upregulation of p21. In the context of a cell cycle checkpoint, p21 upregulation is often associated with the transcriptional activity of p53 like, for example, in the response elicited by DNA damage (el-Deiry, Tokino et al. 1993). Although upregulation of p21 was observed in vivo, we lack direct evidence of an increase in p53 activity in that system. However, several p53 targets, such as Bax and Gadd45γ, are upregulated in ΔS6/flox livers, suggesting that p53 could play a role in the activation of the checkpoint. Using the in vitro system, I was able to take the model one step further and show that the p53 protein levels increase in response to the depletion of S6 by siRNA. Furthermore, I show that p21 induction is solely dependent on p53, since the depletion of p53 ablates the induction of p21. I also established that p21 and p53 are both involved in the cell cycle block induced by the lack of S6. In fact, if downregulation of p21 causes a partial rescue of S-phase entry, ablation of p53 fully rescues the ability of S6-deficient cells to enter S phase, to levels barely distinguishable from the ones of control cells. It is likely, as seen in vivo, that the upregulation of p53 in S6-deficient cells causes the induction of a number of genes, besides p21, that cooperate to cause a cell cycle arrest, explaining why the suppression of p21 alone is not sufficient to fully restore S-phase entry.

An important conclusion from the rescue experiments is that S phase can proceed even while the defect in ribosome biogenesis is still present. This supposes that the pool of ribosomes existing before S6 depletion provides sufficient translation capacity to drive cells through S phase. What is the reason, therefore, behind the upregulation of ribosome biogenesis during cell cycle progression? And why would a checkpoint mechanism exist at this stage? One possibility is that the upregulation of ribosome biogenesis in proliferating cells is mostly required to provide an adequate pool of ribosomes for the daughter cells, and that activation of the checkpoint would occur in circumstances where this task could not be met.

Is the upregulation of p53 only dependent on S6 and 40S disruption?

Using siRNA against two other ribosomal proteins, S23 and L7a, of the small and the large ribosomal subunit, respectively, I was able to detect the upregulation of p21 mRNA, suggesting that a defect of at least two other ribosomal proteins triggers a response similar to the one elicited by S6 downregulation. Furthermore, other groups have reported a relationship between defects in ribosome biogenesis, p53 upregulation and cell cycle inhibition. The manuscript from (Pestov, Strezoska et al. 2001), for example, describes studies on the protein Bop1 which they identified in a genetic screen for cDNAs that caused a reversible perturbation of cell cycle progression. Bop1 was shown to be involved in the processing of the precursor of the 28S rRNA, which assembles in the 60S large ribosomal subunit. Functional inactivation of Bop1 leads to the failure of 60S subunit synthesis, caused by defects in the maturation of the 28S rRNA, without affecting the production of 40S subunits. Cells defective in Bop1 function are blocked in G1 and show upregulation of p53. In the same background, inhibition of p53 by the E6 oncoprotein restores normal S-phase progression. Together with our data, these results show that defects in the biosynthesis of either the 40S or 60S ribosomal subunits block cell cycle progression through the induction of p53.

Upregulation of p53 in response to defects in ribosome biogenesis

The mechanisms that control upregulation of p53 in response to defects in ribosome biogenesis are still unclear, but two potential mechanisms have recently been proposed.The first one, described in (Sugimoto, Kuo et al. 2003), proposes a role for the $p19^{Arf}$ protein in the upregulation of p53. $P19^{Arf}$ mediates the stabilization of p53 by protecting it from Mdm2-mediated proteolisis (Tao and Levine 1999; Weber, Taylor et al. 1999; Zhang and Xiong 1999). The $p19^{Arf}$ mRNA is transcribed from an alternate reading frame of the $p16^{ink4a}$ locus and its expression is upregulated in response to oncogenic and hyperproliferative signals (Sherr 1998). Upregualtion of $p19^{Arf}$ induces a cell cycle block in both G1 and G2, which is mediated through p53 stabilization (Quelle, Zindy et al. 1995). Sugimoto et al. show that the overexpression of $p19^{Arf}$ inhibits ribosome biogenesis in a p53-independent manner, proposing a novel role for this

protein. This effect is elicited by $p19^{Arf}$-mediated degradation of B23, a protein involved in the maturation of 28S rRNA (Hadjiolova, Normann et al. 1994). The authors propose that the upregulation of $p19^{Arf}$ by oncogenic signals could trigger a p53-mediated cell cycle block and at the same time stop ribosome biogenesis through the disruption of 60S formation. Although we cannot exclude it, such a mechanism would be unlikely to work in our experimental system: in fact, as already mentioned, $p19^{Arf}$ expression is mainly induced by oncogenic stimuli, like overexpression of E2F1, which seem instead to be absent S6-depleted cells both in vitro and in vivo. Furthermore, Sugimoto et al. argue that defects in ribosome biogenesis parallel the inhibition of cell cycle progression, while our data suggest that it is indeed the defect in ribosome biogenesis that causes the cell cycle block. Last but not least, the production of 60S ribosomal subunits proceeds normally in S6-depleted cells. In conclusion, although appealing, the described model does not fit with the current data, but it could yet represent another coordination mechanism between the cell cycle and ribosome biogenesis.

Another model was proposed by three different groups, who all identified a subset of ribosomal proteins that directly promotes p53 stabilization. In particular, it was shown that L5, L11, and L23, all proteins from the 60S ribosomal subunit, were able to bind and inhibit the Mdm2 ubiquitin ligase, promoting the stabilization of p53 (Lohrum, Ludwig et al. 2003; Bhat, Itahana et al. 2004; Dai and Lu 2004; Jin, Itahana et al. 2004). The overexpression of L11, for example, led to the upregulation of p53 protein levels (Bhat, Itahana et al. 2004). According to the model of Bhat et al., non-ribosome-bound L11 interacts with Mdm2 and triggers the upregulation of p53 when in excess. Such a situation would be mimicked by the inhibition of rRNA transcription with low doses of actinomycin D, a treatment that is in fact shown by Bhat et al. to trigger L11-mediated upregulation of p53. Similarly, the authors argue, any physiological defect in 60S synthesis would lead to the accumulation of free L11, a consequent increase in p53 levels and, ultimately, to cell cycle block. Indeed, depletion of L11 was sufficient to eliminate the p53 upregulation caused by actinomycin D treatment. L5 and L23 show a similar behavior, although their depletion does not fully abolish the induction of p53 (Dai and Lu 2004; Jin, Itahana et al. 2004). All three of these proteins are dispensable for ribosome biogenesis and translation as shown by knockdown experiments using siRNA

(Bhat, Itahana et al. 2004; Dai and Lu 2004; Jin, Itahana et al. 2004). It is therefore likely that the proposed role for these proteins constitutes an extra ribosomal function. We cannot exclude the possibility that depletion of S6 causes an L11-, L5-, or L23-mediated induction of p53. If that were the case, however, our data would not agree with the fact that disruption of 60S synthesis is a prerequisite for this response. Interestingly, on top of the effect caused by the overexpression of L23, its downregulation also causes a weak induction of p53 and cell cycle arrest (Jin, Itahana et al. 2004), a situation reminiscent of the phenotype caused by S6 depletion. It is tempting to speculate that different mechanisms might sense either the lack or the excess of ribosomal proteins with respect to the amount of rRNA present. The effects of S6 overexpression could maybe provide some insight for such a mechanism.

Downregulation of p53 rescues S phase but not cell number

The experiments reorted show that although the downregulation of p53 can nearly fully rescue the S-phase arrest caused by S6 depletion in A549 cells, it does not result in an increase in cell number. In fact, the number of cells treated with a combination of p53 and S6 siRNAs remained roughly the same as the number of cells treated with the S6 siRNA alone, throughout the experiment. Further studies will be required to understand the mechanisms that participate in this block. We can exclude, however, the involvement of an apoptotic response, since no sub-G1 peak was detected by the FACS analysis. On the contrary, at 48 hours after transfection, the cells treated simultaneously with the S6 and p53 siRNAs showed a small but significant increase in the G2 population, as compared to the non-silencing controls. It is tempting to speculate that a second p53-independent checkpoint exists, that detects defects of ribosome biogenesis in the G2-M phase of the cell cycle. In the response to DNA damage, a mechanism operates, mediated by the Chk1 and Chk2 kinases, which inhibit Cdk1 activity (and therefore progression through mitosis) through the degradation of Cdc25c phosphatase. Interestingly, two publications from the same laboratory (Bhat, Itahana et al. 2004; Jin, Itahana et al. 2004) show FACS analysis profiles of cells treated with siRNA against L11. Although the authors claim that depletion of L11 does not induce a cell cycle block, they fail to comment on the

increase in G2 phase observed upon this treatment. Unfortunately, those papers do not provide data on cell numbers. This, and our observations, might be linked to the same mechanism.

Ribosome biogenesis as a therapeutic target

A recent publication by Rubi et al. (Rubbi and Milner 2003) provides an interesting report on the importance of the nucleolus and its function in ribosome biogenesis in cellular homeostasis. They first show that, surprisingly, the p53 response ttriggered by UV irradiation derives from the disruption of nucleolar integrity rather than from damage to the DNA itself, showing therefore that the nucleolus is under tight repair process monitoring. Furthermore, a good number of commercial drugs used for cancer treatment due to their pro-apoptotic properties are inhibitors of ribosome biogenesis. The author cites examples such as actinomycin D and cysplatin, suggesting that the disruption of ribosome biogenesis can therefore be lethal even for cancer cells. Indeed, our own studies go in the same direction. The A549 cells used in our studies are derived from a small cell lung carcinoma and proliferate fast in culture. Yet treatment with S6 siRNA and consequent disruption of ribosome biogenesis is sufficient to trigger a block in S phase through upregulation of p53. One problem, however, may arise from the fact that p53 is often mutated or repressed in malignant cancers. Nevertheless, our results suggest that downregulation of p53 allows progression through S phase but prevents cell division in cells where ribosome biogenesis is disrupted, reinforcing the idea that the process of ribosome biogenesis may be an ideal therapeutic target

Materials and methods

Mouse genotyping

Mice homozygous for RpS6-floxed genes were screened for the presence of the Mx-Cre transgene by PCR using a modification of the protocol from (Liu, Grinberg et al. 1998). The following nucleotides were used:

Cre 5' (5'-AATGCTTCTGTCCGTTTGCCGGT-3')
Cre 3' (5'-CCAGGCTAAGTGCCTTCTCTACA-3').

About 25mm of each mouse tail was cut using a sterile razor blade. The piece of tail was then incubated in 600µl 50mM NaOH solution for 3-4 hours until completely disaggregated. The NaOH was neutralized using 50µl 1M Tris-HCl pH 8.0. The samples were centrifuged for 6 minutes at 14,000 rpm on a bench centrifuge to pellet the debris. The supernatant was transferred into a fresh Eppendorf tube and then used for the PCR reaction. In order to increase the specificity of the analysis, a two-step PCR reaction was performed.

1µl DNA template	
4µl dNTP (20 mM)	Qiagen 272050
2.5µl PCR buffer (1.5mM $MgCl_2$)	Qiagen 1005476
0.3µl Taq polymerase	Qiagen 1005476
0.25µl primer 5'	Cre 5'
0.25µl primer 3'	Cre 3'
16.7µl H_20	

The following PCR reaction protocol was used:

95°C	5 min
95°C	1 min *
68°C	1 min *
72°C	1 min *
95°C	10 min

* Repeated for 15 cycles.

At the end of the reaction, 1µl of each PCR reaction was used as a template for the second PCR reaction. The second PCR reaction was performed in exactly the same manner, but the cycles were repeated for 40 cycles rather than 15.
The length of the amplified fragment is approximately 600 bp.

Hepatectomy

In all the experiments cited, the deletion of the gene in vivo was induced by a single intra-peritoneal injection of P(I)P(C) (P-1530). The pIpC solution was prepared by diluting the P(I)P(C) stock (5mg/ml) in PBS warmed at room temperature. 500µl solutions were then injected into each mouse using a BD Microlance 2006-04. The deletion was completed in the liver 72 hours post injection.

All the animals used for the experiments were males between 3 months and 6 months of age. Both Mx-Cre– and Mx-Cre+ animals were injected with 250µg P(I)P(C) 72 hours prior to surgery. The mice were anesthetized with a mixture made of 12% Rompun (2% solution, Bayer 00730551) and 8% Narketan 10 (10mg/ml solution, Vetoquinol AG, Belp, Bern, Switzerland) in PBS. 100µl anaesthetic was injected I.P. for each 10g body weight. To compensate for the anaesthetic-induced lowering of body temperature, the mice were placed on a heating plate at 38°C for the entire surgical procedure.

The mice were shaved and a skin incision was performed longitudinally from the plexus for 2cm. A small perpendicular incision was then made in the peritoneum just below the plexus. The incision was then extended by a small cut alongside the right side of the peritoneum until the thoracic cage was reached. The frontal and lateral lobes of the liver were then mechanically pushed out by digital massage. The lobes were ligated using a surgical silk thread around the extruded lobes with a double knot. The ligated lobes were excised using surgical scissors. The excised liver lobes were rinsed in PBS and immediately frozen in liquid nitrogen. The peritoneum was sutured using a surgical thread (Aesculap C0762113). The skin was closed using surgical clips (Becton Dickinson 427631). Animals were kept under a red heating lamp during the whole period of recovery from narcolepsy.

At given timepoints after partial hepatectomy the mice were sacrificed through cervical dislocation. The sutures were removed and the remaining liver lobes were washed in PBS to eliminate any residual blood and frozen immediately in liquid nitrogen.

RNA extraction

The RNA from the livers was extracted using the method of (Chomczynski P 1987). For microarray analysis, the RNA was further purified using the Rneasy Kits Purifications Protocol (Qiagen 74104).

Northern hybridization

Northern hybridization was performed using the method described in (Church and Gilbert 1984). The following oligonucleotide probes were used for the northern blot hybridization:

34S rRNA (human)	GGCGAGCGACCGGCAAGGCGGAGGTCGACCCACGC
E2F1	GATACACTGGACGGTCCTGGAAGCGTCGTAACGTCTGGGACGTCTTCTACCAGTATC
S6	GAGATGTTCAGCTTCATCTTGAAGCAGCTGAACGC

The oligonucleotides were labeled at the 5' terminus with γATP^{32}, using the T4 PNK enzyme. For 20µl:

18pmol oligonucleotide probe	
2µl 10x PNK buffer	New England Biolabs
5µl γATP^{32} 20µCi/µl	Amersham
1µl T4 PNK enzyme	New England Biolabs
Water to 20µl	

All the oligonucleotide probes were hybridized at 56°C overnight in a rotating oven. The excess was washed out by three rapid washes in a 0.1% SDS/2xSSC solution followed by a 1 hour wash at 56°C in the same solution. The membranes were then exposed to a phosphoimager screen. P21 northern blots were performed using a cDNA probe and labeled using the RandomPrime labeling kit (Boehringher Ingelheim).

Microarray analysis

Microarray hybridization was perfrormed following the protocols provided by Affymetrix (Affymetrix, Santa Clara, USA). The cRNA was hybridized to MGU-74A GeneChips™ (Affymetrix, Santa Clara, USA), washing and staining were performed in an Affymetrix Fluidics Station 450 and the samples were scanned in an Affymetrix GeneChip 3000 scanner as per manufacturer's instructions. Expression values were estimated using the robust multichip analysis algorithm (http://rmaexpress.bmbolstad.com). The expression analysis was performed using GeneSpring 6.2 (Silicon Genetics). Changes in gene expression were identified by requiring that a gene had a change in expression level at least 1.5-fold relative to the other genotype at one or more timepoints and pass a 1-way ANOVA (p<0.01) unless otherwise stated.

Electromobility shift assay

Electromobility gel shift assays (EMSA) used a modified protocol from (Garriga, Limon et al. 1998):

A small piece of frozen liver was left to defrost in 1ml hypotonic solution (25mM Tris-HCl pH7.5, 50mM KCl, 2mM $MgCl_2$, 1mM EDTA, 5mM DTT, 5µg/µl aprotinin, 5µg/µl leupeptin, 1mM PMSF) for a few seconds. The sample was then homogenized using a Dounce tissue homogenizer (Wheaton 357542) with a loose pestle, in order to avoid disruption of the nuclei. Following incubation in ice for 5 minutes to allow cell swelling, the samples were centrifuged at 15,000g for 15 minutes at 4°C. After removal of the supernatant (cytoplasmic extract), the pellet was resuspended in 500µl hypotonic solution and centrifuged again in the same conditions in order to wash away further cytoplasmic contaminants. At the end of these washes the nuclei were still intact. To disrupt the nuclei, the pellet was resuspended in 200µl 5x EMSA extraction buffer (100mM hepes pH 7.4, 500mM KCl, 5mM $MgCl_2$, 0.5mM NaF, 1mM DTT, 1mM PMSF, 10µg/µl aprotinin, 10µg/µl leupeptin). The nuclei were left to swell for 30 minutes in ice followed by a 10-minute centrifugation at 15,000g at 4°C. The supernatant was recovered and stored in aliquots of 20µl at –80°C.

The bait DNA probe used was based on the E2F consensus sequence from the E2 viral promoter: the mutated probe contains point mutations on the consensus sequence (Santa Cruz, sc-2507, sc-2508). For the formation of double-stranded DNA, equimolar amounts of the corresponding complementary strands were diluted in oligo-annealing buffer (20mM Tris HCl pH 7.5, 10mM $MgCl_2$, 50mM NaCl), incubated at 95°C for 5 minutes in a thermoblock and then left to cool gradually to RT. 25pmol of each dsDNA probe was then labelled using a T4 PNK enzyme (New England Biobabs). The mixture was incubated at 37°C for 1 hour.

Labelling mixture:

25pmol dsDNA	
γATP ^{32}P (6,000Ci/μl)	7μl
T4 PNK enzyme	2μl
PNK buffer	2.5μl
H_20	To 25μl

The reaction was performed by adding 10μg nuclear extracts to 80,000 cpm double-stranded probe and 2μg salmon sperm DNA. The reaction was kept in ice for 10 minutes for the binding to occur and then for 20 minutes at room temperature. The samples were run on a non-denaturating 4% PAGE for 2 hours at 280V. The gels were dried and exposed to film. Cold probe competition experiments and supershifting were performed in a similar manner, but the reaction was incubated for 1 hour on ice prior to the room temperature incubation. For cold competition, a 50-fold excess of cold wt or mismatched probe was used. For supershift, 1μl Rbl2 antibody was used.

FACS analysis and BrdU labeling

BrdU labeling was performed by adding 10μg/μl BrdU to the cell medium for 10 minutes before collection. The cells were then trypsinized and washed with PBS/BSA 1%. An anti-BrdU FITC-conjugated antibody (BD-347183) was used for BrdU detection. The washes and incubations were performed according to the manufacturer's instructions. FACS analysis was performed using a FacsCalibur machine in association with the Modfit program.

RNA interference

All the siRNAs were a kind gift from Francois Natt (Novartis, Basel, Switzerland). The sequences were developed using a proprietary algorithm developed by Novartis, which screens for target specificity.

Name	Target sequence
S6	GTA AGA AAG CCC TTA AAT Att
p21	TCA GCT GCT CGC TGT CCA Ctg
p53	TTC CAC TCG GAT AAG ATG Ctg
L7a	TTG TTC TCC ACC AAG GTG Gtg
S23	AAG GGT TGG CCT TTA GGG Ctg

The siRNAs were transfected in A549 cells using oligofectamine (Invitrogen, Cat. No. 12252-011) following the manufacturer's instructions. 8µl oligofectamine was mixed with 12µl optimem. After 10 minutes' incubation, both solutions were mixed and the transfection mix was incubated for an additional 20 minutes. Just prior to transfection, target plates were washed once with 5ml optimem and the cells were covered with 1ml optimem and 200µl transfection mix. After 5 hours' incubation at 37°C and 5% CO_2, the transfection mix was replaced with 5ml RPMI 1640 supplemented with 10% FCS. For 10-cm plates the values were all multiplied by three (8µl oligofectamine were used for a 6-cm dish).

The oligofectamine/siRNA mixture was added to cells in optimem medium and incubated for 5 hours. The medium was subsequently replaced with RPMI 1640 medium containing 10% FCS. The cells were collected at 24 or 48 hours after siRNA treatment. In most cases, 15nM (final) of each siRNA was added to the culture medium. A non-silencing (NS) siRNA (Quiagen, 1022076) was used as the negative control. In siRNA rescue experiments, the differences in total siRNA amounts were compensated for by the addition of the corresponding amount of NS siRNA.

Polysome profiles and RNA sucrose gradients

Polysome profile analysis was performed using protocols described in (Jefferies, Fumagalli et al. 1997). For RNA sucrose gradients, cells were extracted with solution D

(Chomczynski P 1987) and 0.2% SDS was added to the final extract. The samples were incubated for 1 minute at 60°C and spun in a microfuge for 5 minutes. The supernatant was layered on a 5.1-28.6% exponential sucrose gradient (100mM LiCl/10mM EDTA/10mM TrisCl pH7.4/SDS 0.2%) and ultracentrifuged at 60,000 rpm for 85 minutes at 20°C.

Western blot

Proteins were extracted using 50mM Tris HCl pH 7.4/5mM EDTA/250mM NaCl/50mM NaF/0.1% Triton X-100 + protease inhibitor tablets (Complete-Mini, Roche). For liver extracts, the samples were homogenized using a Dounce homogenizer. Protein samples were resolved by conventional SDS-PAGE and transferred onto PVDF membrane (Millipore), using a semidry transfer apparatus (LTF-Labortechnik). Blocking of the membrane was in Tris-Buffer Saline (TBS) supplemented with 0.02% Tween 20 and 5% milk. The following primary antibodies were used: mouse monoclonal anti-S6 (Novartis), rabbit polyclonal anti-L11 (a gift of Karen Vousden), rabbit polyclonal anti-p53 (Santa Cruz Biotechnology, FL-393), mouse monoclonal anti-p21 (Pharmingen, cat. 556431) rabbit polyclonal anti-p107 (Santa Cruz Biotechnolgy, C-18), rabbit plyclonal anti-p130 (Santa Cruz Biotechnology, C20)

Bibliography

Al-Shahrour, F., R. Diaz-Uriarte, et al. (2004). "FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes." Bioinformatics **20**(4): 578-80.

Albrecht, J. H., R. Y. Poon, et al. (1998). "Involvement of p21 and p27 in the regulation of CDK activity and cell cycle progression in the regenerating liver." Oncogene **16**(16): 2141-50.

Ashcroft, M., M. H. Kubbutat, et al. (1999). "Regulation of p53 function and stability by phosphorylation." Mol Cell Biol **19**(3): 1751-8.

Bartek, J., J. Bartkova, et al. (1997). "The retinoblastoma protein pathway in cell cycle control and cancer." Exp Cell Res **237**(1): 1-6.

Bartek, J. and J. Lukas (2001). "Mammalian G1- and S-phase checkpoints in response to DNA damage." Curr Opin Cell Biol **13**(6): 738-47.

Beijersbergen, R. L., L. Carlee, et al. (1995). "Regulation of the retinoblastoma protein-related p107 by G1 cyclin complexes." Genes Dev **9**(11): 1340-53.

Bhat, K. P., K. Itahana, et al. (2004). "Essential role of ribosomal protein L11 in mediating growth inhibition-induced p53 activation." Embo J **23**(12): 2402-12.

Brehm, A., E. A. Miska, et al. (1998). "Retinoblastoma protein recruits histone deacetylase to repress transcription." Nature **391**(6667): 597-601.

Brown, A. L., C. H. Lee, et al. (1999). "A human Cds1-related kinase that functions downstream of ATM protein in the cellular response to DNA damage." Proc Natl Acad Sci U S A **96**(7): 3745-50.

Budde, A. and I. Grummt (1999). "p53 represses ribosomal gene transcription." Oncogene **18**(4): 1119-24.

Chellappan, S. P., S. Hiebert, et al. (1991). "The E2F transcription factor is a cellular target for the RB protein." Cell **65**(6): 1053-61.

Cheng, M., P. Olivier, et al. (1999). "The p21(Cip1) and p27(Kip1) CDK 'inhibitors' are essential activators of cyclin D-dependent kinases in murine fibroblasts." Embo J **18**(6): 1571-83.

Chomczynski P, S. N. (1987). "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction." Anal Biochem **162**(1): 156-9.

Church, G. M. and W. Gilbert (1984). "Genomic sequencing." Proc Natl Acad Sci U S A **81**(7): 1991-5.

Ciarmatori, S., P. H. Scott, et al. (2001). "Overlapping functions of the pRb family in the regulation of rRNA synthesis." Mol Cell Biol **21**(17): 5806-14.

Classon, M., S. Salama, et al. (2000). "Combinatorial roles for pRB, p107, and p130 in E2F-mediated cell cycle control." Proc Natl Acad Sci U S A **97**(20): 10820-5.

Cooper, H. L. and E. M. Gibson (1971). "Control of synthesis and wastage of ribosomal ribonucleic acid in lymphocytes. II. The role of protein synthesis." J Biol Chem **246**(16): 5059-66.

Dai, M. S. and H. Lu (2004). "Inhibition of MDM2-mediated p53 ubiquitination and degradation by ribosomal protein L5." J Biol Chem.

Datta, P. K., S. Budhiraja, et al. (1997). "Regulation of ribosomal RNA gene transcription during retinoic acid-induced differentiation of mouse teratocarcinoma cells." Exp Cell Res **231**(1): 198-205.

Dennis, P. B., A. Jaeschke, et al. (2001). "Mammalian TOR: a homeostatic ATP sensor." Science **294**(5544): 1102-5.

Desai, D., H. C. Wessling, et al. (1995). "Effects of phosphorylation by CAK on cyclin binding by CDC2 and CDK2." Mol Cell Biol **15**(1): 345-50.

Di Stefano, L., M. R. Jensen, et al. (2003). "E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes." Embo J **22**(23): 6289-98.

Dudov, K. P. and M. D. Dabeva (1983). "Post-transcriptional regulation of ribosome formation in the nucleus of regenerating rat liver." Biochem J **210**(1): 183-92.

Dyson, N. (1998). "The regulation of E2F by pRB-family proteins." Genes Dev **12**(15): 2245-62.

Eichler, D. C. and N. Craig (1994). "Processing of eukaryotic ribosomal RNA." Prog Nucleic Acid Res Mol Biol **49**: 197-239.

el-Deiry, W. S., T. Tokino, et al. (1993). "WAF1, a potential mediator of p53 tumor suppression." Cell **75**(4): 817-25.

Elbashir, S. M., W. Lendeckel, et al. (2001). "RNA interference is mediated by 21- and 22-nucleotide RNAs." Genes Dev **15**(2): 188-200.

Farkas, T., K. Hansen, et al. (2002). "Distinct phosphorylation events regulate p130- and p107-mediated repression of E2F-4." J Biol Chem **277**(30): 26741-52.

Fatica, A. and D. Tollervey (2002). "Making ribosomes." Curr Opin Cell Biol **14**(3): 313-8.

Ferreira, R., L. Magnaghi-Jaulin, et al. (1998). "The three members of the pocket proteins family share the ability to repress E2F activity through recruitment of a histone deacetylase." Proc Natl Acad Sci U S A **95**(18): 10493-8.

Garriga, J., A. Limon, et al. (1998). "Differential regulation of the retinoblastoma family of proteins during cell proliferation and differentiation." Biochem J **333 (Pt 3)**: 645-54.

Gausing, K. (1974). "Ribosomal protein in E. coli: rate of synthesis and pool size at different growth rates." Mol Gen Genet **129**(1): 61-75.

Geng, Y., W. Whoriskey, et al. (1999). "Rescue of cyclin D1 deficiency by knockin cyclin E." Cell **97**(6): 767-77.

Geng, Y., Q. Yu, et al. (2003). "Cyclin E ablation in the mouse." Cell **114**(4): 431-43.

Grana, X., J. Garriga, et al. (1998). "Role of the retinoblastoma protein family, pRB, p107 and p130 in the negative control of cell growth." Oncogene **17**(25): 3365-83.

Grummt, I. and F. Grummt (1976). "Control of nucleolar RNA synthesis by the intracellular pool sizes of ATP and GTP." Cell **7**(3): 447-53.

Grummt, I., V. A. Smith, et al. (1976). "Amino acid starvation affects the initiation frequency of nucleolar RNA polymerase." Cell **7**(3): 439-45.

Hadjiolova, K. V., A. Normann, et al. (1994). "Processing of truncated mouse or human rRNA transcribed from ribosomal minigenes transfected into mouse cells." Mol Cell Biol **14**(6): 4044-56.

Hamamoto, R., M. Kamihira, et al. (1999). "Growth and differentiation of cultured fetal hepatocytes isolated various developmental stages." Biosci Biotechnol Biochem **63**(2): 395-401.

Harbour, J. W., R. X. Luo, et al. (1999). "Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1." Cell **98**(6): 859-69.

Harper, J. W., G. R. Adami, et al. (1993). "The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases." Cell **75**(4): 805-16.

Hirai, H., M. F. Roussel, et al. (1995). "Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6." Mol Cell Biol **15**(5): 2672-81.

Hollstein, M., D. Sidransky, et al. (1991). "p53 mutations in human cancers." Science **253**(5015): 49-53.

Hurford, R. K., Jr., D. Cobrinik, et al. (1997). "pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes." Genes Dev **11**(11): 1447-63.

Jacob, S. T. (1995). "Regulation of ribosomal gene transcription." Biochem J **306 (Pt 3)**: 617-26.

Jefferies, H. B., S. Fumagalli, et al. (1997). "Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70s6k." Embo J **16**(12): 3693-704.

Jin, A., K. Itahana, et al. (2004). "Inhibition of HDM2 and activation of p53 by ribosomal protein L23." Mol Cell Biol **24**(17): 7669-80.

Johnson, D. G., K. Ohtani, et al. (1994). "Autoregulatory control of E2F1 expression in response to positive and negative regulators of cell cycle progression." Genes Dev **8**(13): 1514-25.

Khosravi, R., R. Maya, et al. (1999). "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage." Proc Natl Acad Sci U S A **96**(26): 14973-7.

Kuerbitz, S. J., B. S. Plunkett, et al. (1992). "Wild-type p53 is a cell cycle checkpoint determinant following irradiation." Proc Natl Acad Sci U S A **89**(16): 7491-5.

LaBaer, J., M. D. Garrett, et al. (1997). "New functional activities for the p21 family of CDK inhibitors." Genes Dev **11**(7): 847-62.

Lew, D. J., V. Dulic, et al. (1991). "Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast." Cell **66**(6): 1197-206.

Li, R., S. Waga, et al. (1994). "Differential effects by the p21 CDK inhibitor on PCNA-dependent DNA replication and repair." Nature **371**(6497): 534-7.

Liu, J.-L., A. Grinberg, et al. (1998). "Insulin-Like Growth Factor-I Affects Perinatal Lethality and Postnatal Development in a Gene Dosage-Dependent Manner: Manipulation Using the Cre/loxP System in Transgenic Mice." Mol Endocrinol **12**(9): 1452-1462.

Lodish, H. F. (1974). "Model for the regulation of mRNA translation applied to haemoglobin synthesis." Nature **251**(5474): 385-8.

Lohrum, M. A., R. L. Ludwig, et al. (2003). "Regulation of HDM2 activity by the ribosomal protein L11." Cancer Cell **3**(6): 577-87.

Lukas, J., T. Herzinger, et al. (1997). "Cyclin E-induced S phase without activation of the pRb/E2F pathway." Genes Dev **11**(11): 1479-92.

Lukas, J., C. Lukas, et al. (2004). "Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time." DNA Repair (Amst) **3**(8-9): 997-1007.

Macleod, K. F., N. Sherry, et al. (1995). "p53-dependent and independent expression of p21 during cell growth, differentiation, and DNA damage." Genes Dev **9**(8): 935-44.

Makris, C., L. Voisin, et al. (2002). "The Rb-family protein p107 inhibits translation by a PDK1-dependent mechanism." Oncogene **21**(51): 7891-6.

Malumbres, M., R. Sotillo, et al. (2004). "Mammalian cells cycle without the D-type cyclin-dependent kinases Cdk4 and Cdk6." Cell **118**(4): 493-504.

Mayol, X., J. Garriga, et al. (1995). "Cell cycle-dependent phosphorylation of the retinoblastoma-related protein p130." Oncogene **11**(4): 801-8.

Meyuhas, D., S. Nir, et al. (1996). "Aggregation of phospholipid vesicles by water-soluble polymers." Biophys J **71**(5): 2602-12.

Miyashita, T. and J. C. Reed (1995). "Tumor suppressor p53 is a direct transcriptional activator of the human bax gene." Cell **80**(2): 293-9.

Morkel, M., J. Wenkel, et al. (1997). "An E2F-like repressor of transcription." Nature **390**(6660): 567-8.

Nigg, E. A. (1993). "Targets of cyclin-dependent protein kinases." Curr Opin Cell Biol **5**(2): 187-93.

O'Connor, R. J., J. E. Schaley, et al. (2001). "The p107 tumor suppressor induces stable E2F DNA binding to repress target promoters." Oncogene **20**(15): 1882-91.

Ortega, S., I. Prieto, et al. (2003). "Cyclin-dependent kinase 2 is essential for meiosis but not for mitotic cell division in mice." Nat Genet **35**(1): 25-31.

Pardee, A. B. (1989). "G1 events and regulation of cell proliferation." Science **246**(4930): 603-8.

Pestell, R. G., C. Albanese, et al. (1999). "The cyclins and cyclin-dependent kinase inhibitors in hormonal regulation of proliferation and differentiation." Endocr Rev **20**(4): 501-34.

Pestov, D. G., Z. Strezoska, et al. (2001). "Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: effects of nucleolar protein Bop1 on G(1)/S transition." Mol Cell Biol **21**(13): 4246-55.

Polyak, K., T. Waldman, et al. (1996). "Genetic determinants of p53-induced apoptosis and growth arrest." Genes Dev **10**(15): 1945-52.

Quelle, D. E., R. A. Ashmun, et al. (1995). "Cloning and characterization of murine p16INK4a and p15INK4b genes." Oncogene **11**(4): 635-45.

Quelle, D. E., F. Zindy, et al. (1995). "Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest." Cell **83**(6): 993-1000.

Ren, B., H. Cam, et al. (2002). "E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints." Genes Dev **16**(2): 245-56.

Reynisdottir, I., K. Polyak, et al. (1995). "Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta." Genes Dev **9**(15): 1831-45.

Rubbi, C. P. and J. Milner (2003). "Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses." Embo J **22**(22): 6068-77.

Schnapp, A., C. Pfleiderer, et al. (1990). "A growth-dependent transcription initiation factor (TIF-IA) interacting with RNA polymerase I regulates mouse ribosomal RNA synthesis." Embo J **9**(9): 2857-63.

Shamji, A. F., P. Nghiem, et al. (2003). "Integration of growth factor and nutrient signaling: implications for cancer biology." Mol Cell **12**(2): 271-80.

Sherr, C. J. (1996). "Cancer cell cycles." Science **274**(5293): 1672-7.

Sherr, C. J. (1998). "Tumor surveillance via the ARF-p53 pathway." Genes Dev **12**(19): 2984-91.

Sherr, C. J. and J. M. Roberts (1995). "Inhibitors of mammalian G1 cyclin-dependent kinases." Genes Dev **9**(10): 1149-63.

Sherr, C. J. and J. M. Roberts (1999). "CDK inhibitors: positive and negative regulators of G1-phase progression." Genes Dev **13**(12): 1501-12.

Slansky, J. E., Y. Li, et al. (1993). "A protein synthesis-dependent increase in E2F1 mRNA correlates with growth regulation of the dihydrofolate reductase promoter." Mol Cell Biol **13**(3): 1610-8.

Sollner-Webb, B. and J. Tower (1986). "Transcription of cloned eukaryotic ribosomal RNA genes." Annu Rev Biochem **55**: 801-30.

Sorensen, C. S., R. G. Syljuasen, et al. (2003). "Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A." Cancer Cell **3**(3): 247-58.

Sugimoto, M., M. L. Kuo, et al. (2003). "Nucleolar Arf tumor suppressor inhibits ribosomal RNA processing." Mol Cell **11**(2): 415-24.

Sutcliffe, J. E., C. A. Cairns, et al. (1999). "RNA polymerase III transcription factor IIIB is a target for repression by pocket proteins p107 and p130." Mol Cell Biol **19**(6): 4255-61.

Tang, H. Y., K. Zhao, et al. (1998). "Constitutive expression of the cyclin-dependent kinase inhibitor p21 is transcriptionally regulated by the tumor suppressor protein p53." J Biol Chem **273**(44): 29156-63.

Tao, W. and A. J. Levine (1999). "P19(ARF) stabilizes p53 by blocking nucleo-cytoplasmic shuttling of Mdm2." Proc Natl Acad Sci U S A **96**(12): 6937-41.

Thomas, G. (2002). "The S6 kinase signaling pathway in the control of development and growth." Biol Res **35**(2): 305-13.

Trimarchi, J. M. and J. A. Lees (2002). "Sibling rivalry in the E2F family." Nat Rev Mol Cell Biol **3**(1): 11-20.

Vlach, J., S. Hennecke, et al. (1997). "Phosphorylation-dependent degradation of the cyclin-dependent kinase inhibitor p27." Embo J **16**(17): 5334-44.

Volarevic, S., M. J. Stewart, et al. (2000). "Proliferation, but not growth, blocked by conditional deletion of 40S ribosomal protein S6." Science **288**(5473): 2045-7.

Vousden, K. H. and X. Lu (2002). "Live or let die: the cell's response to p53." Nat Rev Cancer **2**(8): 594-604.

Warner, J. R., J. Vilardell, et al. (2001). "Economics of ribosome biosynthesis." Cold Spring Harb Symp Quant Biol **66**: 567-74.

Weber, J. D., L. J. Taylor, et al. (1999). "Nucleolar Arf sequesters Mdm2 and activates p53." Nat Cell Biol **1**(1): 20-6.

Woolford, J. L., Jr. (1991). "The structure and biogenesis of yeast ribosomes." Adv Genet **29**: 63-118.

Yamasaki, L., R. Bronson, et al. (1998). "Loss of E2F-1 reduces tumorigenesis and extends the lifespan of Rb1(+/-)mice." Nat Genet **18**(4): 360-4.

Zhan, Q., M. J. Antinore, et al. (1999). "Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45." Oncogene **18**(18): 2892-900.

Zhang, H., J. P. Stallock, et al. (2000). "Regulation of cellular growth by the Drosophila target of rapamycin dTOR." Genes Dev **14**(21): 2712-24.

Zhang, Y., N. Fujita, et al. (1999). "Caspase-mediated cleavage of p21Waf1/Cip1 converts cancer cells from growth arrest to undergoing apoptosis." Oncogene **18**(5): 1131-8.

Zhang, Y. and Y. Xiong (1999). "Mutations in human ARF exon 2 disrupt its nucleolar localization and impair its ability to block nuclear export of MDM2 and p53." Mol Cell **3**(5): 579-91.

Zhao, H., J. L. Watkins, et al. (2002). "Disruption of the checkpoint kinase 1/cell division cycle 25A pathway abrogates ionizing radiation-induced S and G2 checkpoints." Proc Natl Acad Sci U S A **99**(23): 14795-800.

Zhu, L., E. Xie, et al. (1995). "Differential roles of two tandem E2F sites in repression of the human p107 promoter by retinoblastoma and p107 proteins." Mol Cell Biol **15**(7): 3552-62.

Bioinformatics

Preface

It is now well established that DNA microarray technology is an incredibly potent tool for biologists. It allows the simultaneous measurement of the expression levels of tens of thousands of transcripts, with full genome coverage arrays existing for many different organisms. The problem is, however, that while the result output from this technology is steadily increasing, the speed with which data can be analyzed is not. I became aware of this problem whilst analyzing my own microarray data, which consisted of a large number of conditions and replicates. In order to perform the analysis of my experiments, I automated several steps of this process to make it faster and more user-friendly. The software tools described in the following paragraphs were primarily developed for analysis of the microarray data generated in the course of my project. They found, however, a wider application amongst other users. Together with Dr. Edward Oakeley, at the FMI, I developed a set of software tools that could be used routinely for microarray analysis. Three programs were developed: ChangeMaker, which deals with the microarray analysis itself; MicroPlot which is used to visualize the data; and PromoterPlot which takes the next step in microarray analysis and deals with the identification of common patterns of transcription factors within promoters

ChangeMaker

o Introduction

The ChangeMaker program automates the comparison of microarray replicates.
The motivation behind this program came from the analysis of livers after hepatectomy in response to the deletion of the ribosomal S6 protein. The experimental setup for this analysis consisted of the comparison of triplicate repeats of two genotypes over four timepoints, thus a total of 24 samples. Given the particular variability of expression data in animal models, requiring that genes must show the same behavior in all the replicates is often overly sensitive to outliers, with the effect that we might miss some potentially interesting candidate genes. It was therefore worth providing a degree of flexibility to the analysis, by analyzing genes that changed in most, but not necessarily all, of the replicates. I needed, therefore, to generate a set of data that would show the genes which changed significantly between each genotype, for each timepoint and in at least two of my three replicates. The software tools provided at the time by Affymetrix were limited and time-consuming for this type of analysis. This application was developed jointly with Dr. Edward Oakeley and was programmed in the VB.Net language.

o Methods

We devised a program that would make these comparisons automatically, without using the Affymetrix software. The input of the program is the Affymetrix CEL files which correspond to the 75th percentile of pixel brightness in each oligo spot from the raw scan generated by the Affymetrix GCS 3000 scanner. The CEL files are loaded into the program as groups of replicates with a user-defined name (Fig.1)

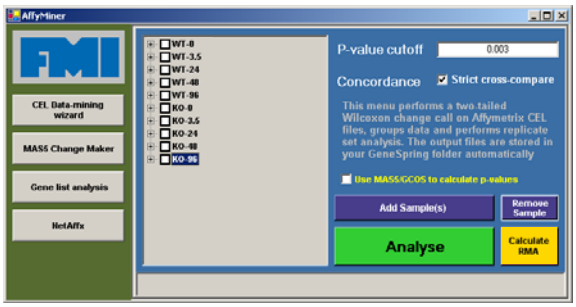


Fig.1 Screen shot of ChangeMaker: the user loads the CEL files and places them into groups of replicates

The user is then prompted to choose which sets of replicates are to be compared. These can represent either comparisons of different timepoints or comparisons between different treatments or both (Fig.2).

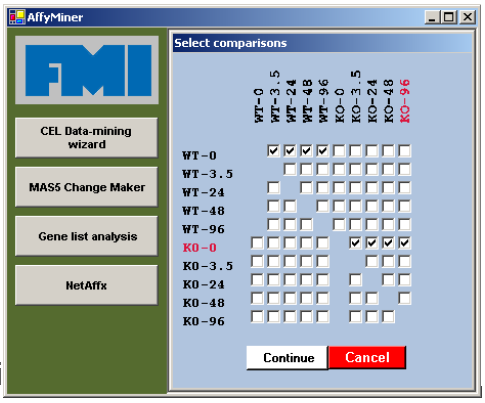


Fig.2 Screen shot of ChangeMaker. By clicking on the tickbox, the user selects the comparisons to be performed

Following thi

1) The original CEL file and corresponding library files were very simple text files which could be loaded and parsed directly by our program. However, with the advent of the GeneChip Operating System (GCOS) from Affymetrix these files became binary so we had to make use of Affymetrix's file-loader tool (file API) to handle newer files. Once the CEL files had been loaded, we had numerical values for the hybridization of each oligonucleotide on the array.

2) Calculate change *p*-values using a signed Wilcoxon rank test (as recommended by Affymetrix) on the raw hybridization data for every pair-wise replicate comparison (Fig.3). The algorithm for calculating the *p*-values was rewritten by E. Oakeley to reduce the time complexity of the calculations from 2^N to N^3 (N=number of oligonucleotides per transcript, typically 32). Any probe showing a change *p*-value of <0.003 was discarded.

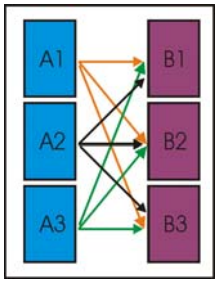


Fig.3 Schema representing an example of comparisons between two groups (A and B) each containing three replicates (1, 2, 3). The algorithm will perform all the pairwise comparisons and generate nine separate lists

3) Once the lists of genes significantly changing for each comparison has been established, the lists are crossed together to determine whether each transcript shows a consistent change in multiple pair-wise comparisons. For example, in an experiment with three overlapping lists, four regions of intersection can be found, indicated in yellow and white on (Fig.4). These correspond to regions where the probes have been shown to change in 2/3 of the comparisons performed (in yellow) or in all of them (white). The program pulls these lists together and creates three main lists. The first one contains genes that change in all the replicates compared, the second contains genes that change in two thirds of replicates, and the final one contains all the probes that have changed in at least one of the replicates. Taken together, these lists constitute three different stringencies of analysis. The exercise is not trivial, since the number of overlapping regions increases exponentially with the number of lists to be analyzed. The number of overlapping regions corresponds to $(2^n–1)–n$. Using the example described in (Fig.3), we obtain nine lists. This would correspond to 502 overlapping regions for each condition analyzed, which excludes any possibility of performing this task manually.

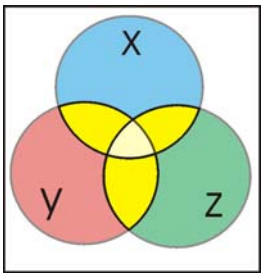


Fig.4 Venn diagram showing an example of intersection between three lists

4) The resulting main lists are exported to Genespring, a statistical program used in the institute for microarray analysis. The lists are further divided into genes increasing or decreasing in relation to the baseline sample provided in the comparisons (Fig.5).

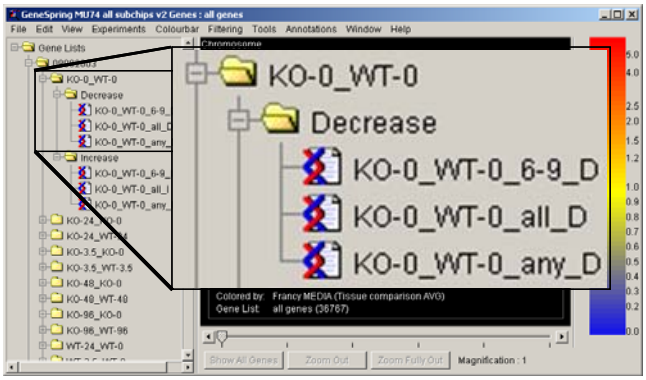


Fig.5 Screen shot of Genespring 6.1 showing the display of the lists resulting from the ChangeMaker program

o Conclusion

The ChangeMaker program provides a valuable tool for accelerating the analysis of microarray data using the raw oligonucleotide hybridizations, especially when dealing with replicates. The program does not consider fold change or expression level but simply patterns of variation based on the idea that if a single RNA class changes in concentration then all hybridizing targets should move in the same direction. More convertional analyses based on expression estimates may then be performed later using GeneSpring. The resulting gene lists are divided according to the direction and stability of the change observed, therefore allowing a certain degree of flexibility in the analysis of replicates.

MicroPlot

o Introduction

This program was designed as a web application to visualize microarray data.

The analysis of microarrays can provide direct results but can also constitute a library of information, which can be used at a later timepoint. This information must therefore be easily accessible and not confined to users experienced in microarray analysis. This application has been programmed in ASP.Net language and makes use of the .netcharting© module.

o Method

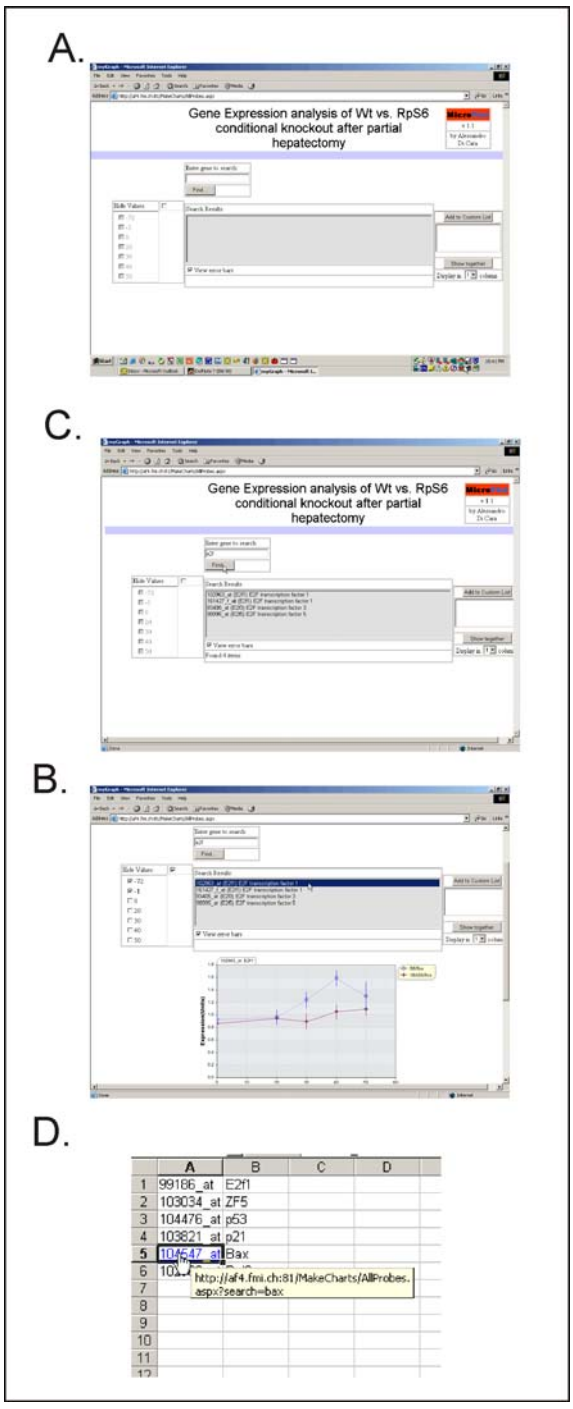
I developed a web application which displays, upon request, a graph of the expression data of a selected gene. The first part of the program, the loader, creates a database containing the expression data and the standard error (in case replicates are present) of each transcript in the experiment. The data can be inserted into the loader as an Excel worksheet. Once the database has been created, it is possible to query the data through a web page. The user inputs the name of the gene, its symbol or the Affymetrix probeset ID (Fig.6A), and a list of the transcripts matching the specified criteria appears (Fig.6B). Upon selection of one of the genes, a chart is displayed showing the expression data and the standard error of each timepoint in each genotype (Fig.6C).

The searches can also be performed using parameters on the web address. For example: http://<ServerName>/AllProbes.aspx?search=e2f1. This will direct the web page directly to the E2f1 expression data. This function becomes very useful for linking Excel spreadsheets to a graph output (Fig.6D).

o Conclusion

MicroPlot provides the possibility of viewing microarray expression data through a web page. It therefore allows the sharing of results between collaborators and other interested people, who may have no experience in microarray data analysis. Furthermore, it allows the generation of graphs which can be saved and pasted into other applications (the graphs of microarray data shown in this thesis were generated

using this program). Additional features could provide links to other databases to provide further data and insights in the gene being analyzed.



PromoterPlot

PromoterPlot was designed with the aim of finding common regulatory regions between genes. Using microarray analysis it is possible to identify clusters of genes which are co-expressed. However, to date it is very difficult to determine if the genes obtained are regulated by the same mechanisms without prior knowledge of the transcription factors involved. Originally, PromoterPlot was designed as a tool to view the positioning of transcription factors on promoter sequences. The use of this program was then applied to another project by E. Oakeley, which aimed at finding common transcriptional factor motifs amongst genes. This led to the development of a pattern-finding algorithm that would find common patterns of transcription factors on promoters. This program was written in close collaboration with E. Oakeley. I was mostly involved in the creation of the graphical interface and the programming of the pattern-finding algorithm. The following manuscript, embedded in this thesis, relates the functioning of this software and shows its application on the study of the c-jun promoter. The microarray data used in the manuscript derives from experiments from Karsten Schimdt, a former fellow from Brian Hemmings' laboratory at the FMI. The program was developed as a web application in ASP.Net and Javascript.

Nucleic Acids Research, 2005, Vol. 33, Web Server issue ***W423–W426***
doi:10.1093/nar/gki413

PromoterPlot: a graphical display of promoter similarities by pattern recognition

Alessandro Di Cara, Karsten Schmidt, Brian A. Hemmings and Edward J. Oakeley*

Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland

Received February 22, 2005; Revised and Accepted March 21, 2005

ABSTRACT

PromoterPlot (http://promoterplot.fmi.ch) is a web-based tool for simplifying the display and processing of transcription factor searches using either the commercial or free TransFac distributions. The input sequence is a TransFac search (public version) or FASTA/Affymetrix IDs (local install). It uses an intuitive pattern recognition algorithm for finding similarities between groups of promoters by dividing transcription factor predictions into conserved triplet models. To minimize the number of false-positive models, it can optionally exclude factors that are known to be unexpressed or inactive in the cells being studied based on microarray or proteomic expression data. The program will also estimate the likelihood of finding a pattern by chance based on the frequency observed in a control set of mammalian promoters we obtained from Genomatix. The results are stored as an interactive SVG web page on our server.

INTRODUCTION

The initial objective of this work is to develop a viewing tool to display the results of TransFac searches in a graphic form. In this paper, we will describe a software tool we have developed for combining expression data with promoter analysis.

Promoter analysis is a process that has historically been very difficult to perform in higher eukaryotes (1). The first question one must consider is how exactly to define a promoter? In the context of this article, we will be using the definition that the 'core promoter' should occupy a region 500 bp upstream of the start of transcription (2). The challenge with this definition is how should we obtain the starts of transcription? The variable and sometimes large 5′-untranslated regions of mammalian messages make any reference to the start of translation rather weak, and the poor sequence conservation between starts (3) makes computer prediction a complex and sometimes unreliable task. We purchased the promoter resources for human, mouse and rat from Genomatix (http://www.genomatix.de) which, as of spring 2004, contained ~156 000 starts of transcription. Many of these had been experimentally mapped using oligo capping technology (4). Non-commercial resources also exist, most notably the Eukaryotic Promoter Database (http://www.epd.isb-sib.ch/) (5) and the various genome sequence repositories.

There are two major schools of thought regarding promoter analysis at the present time: first, the sequence-based approach where short regions of sequence conservation between regulatory sequences are assembled in an attempt to predict regions of micro-conservation that might be important in the control of gene expression, e.g. the MEME motif discovery program (6) which is a tool for discovering motifs in a group of related sequences. MEME was one of the first such programs and it strives to develop position-dependent probability matrices for finding every possible letter at each position in a putative pattern. The motifs found do not contain gaps but can be rather short so that gaps are modeled by the occurrence of additional motifs with un-conserved relative spacings. The size of these motifs is automatically calculated by the program. One further program is MotifSampler, which uses Gibbs sampling to assign a probability distribution to the chance of finding apparently conserved regions of sequence (7). These models do not require any prior knowledge of the underlying biology and as such it can be difficult to assess the mechanistic significance of any pattern found (8), and even when sequence conservation does occur it does not necessarily imply a conserved regulatory function. The second major approach for promoter analysis is the knowledge-based search for known transcription factor binding sites [reviewed in (9)]. This process relies on the collection of information from the scientific literature about the known binding sites from which a consensus target site is estimated. This effort is largely the work of the German

*To whom correspondence should be addressed. Tel: +41 61 697 6986; Fax: +41 61 697 3976; Email: edward.oakeley@fmi.ch
Present address:
Karsten Schmidt, The Clayton Foundation Laboratories for Peptide Biology, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

company Biobase GmbH (http://www.biobase.de) through their Match program associated with the TransFac database. Match provides extremely detailed reports of potential binding sites in target sequences; however, the complexity of the answers returned can be daunting. The challenge here is to filter the data so that we can extract biologically useful models for hypothesis formulation.

The objective of this work is to create a simple web application which would use expression data or proteomics data to filter the list of potential transcription factors predicted by TransFac. We then developed an advanced pattern recognition algorithm to extract patterns of conserved factors in the promoters under investigation. Looking for single transcription factors does not provide any measure of the significance of findings above that given already in TransFac, instead modules of multiple transcription factors in a defined order have been shown to be critical for modulating the expression of genes (10,11). Higher-order complexes may be considered as pairs, triplets or greater numbers of factors in a module. Models composed of three factors have already been shown to be more selective than those of only two (12). Models with even higher complexity are overly stringent and appear automatically when searching for patterns of three. If multiple overlapping binding sites are predicted for the same transcription factor, these are concatenated in the display process into a single big binding site. This helps to simplify the display process without having a negative effect on the pattern searching process. The output is presented as an interactive web page using the Adobe SVG plug-in (Adobe Systems Inc., San Jose, CA) for Internet Explorer. Our tool is distributed in two forms: one for local installation, which can automate the TransFac queries and promoter sequence extraction using in-house information resources (source code available on request); and a second which is freely available on the Internet (http://promoterplot.fmi.ch). The input required for the Internet-accessible version is a TransFac search, saved as a text file, for your promoters. A free version of TransFac is available from Biobase (http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi) which may be used for this purpose.

DESIGN AND IMPLEMENTATION

PromoterPlot was developed as a web-based application. It was designed using Visual Studio 2003 (ASP.NET 1.1) and uses IIS as a web server (Microsoft Corporation, Redland, WA). We recommend that users access our program using Microsoft Internet Explorer 6 or later on the Windows operating system as some users have reported problems with other browsers. An Alchemi grid is used for processing (www.alchemi.net).

Modeling factor patterns

Users can either use our tool as a simple TransFac viewer to simply display the results of a TransFac search without further processing or they can look for conserved potentially regulatory modules within the promoters. To identify these conserved modules, we developed a pattern-searching algorithm, which works by scanning the promoter sequence for patterns composed of three transcription factors (ABC). So as not to collect patterns of very distant factors, which are less likely to interact with each other, the patterns are selected according to the maximum base pair distance between the first and the third factor (C–A), in a user definable manner (default = 100 bp). All of the patterns discovered in this way are collected in a single list. The patterns which passed the following restrictions are retained: (i) the same order of transcription factors; (ii) an identical strand distribution of the factors; (iii) a conserved spacing C–A and B–A with a user-defined 'wobble' for the spacing conservation (default ± 10 bp); and (iv) the pattern must occur in more than one of the promoters analyzed (default = 2). This process is summarized in Figure 1.

In our initial analysis, we found that the TransFac Match search, using the 'minimize false positives' (FPs) setting, predicts on average one binding site every 20 bp. This high stringency search is very tempting because it does not generate

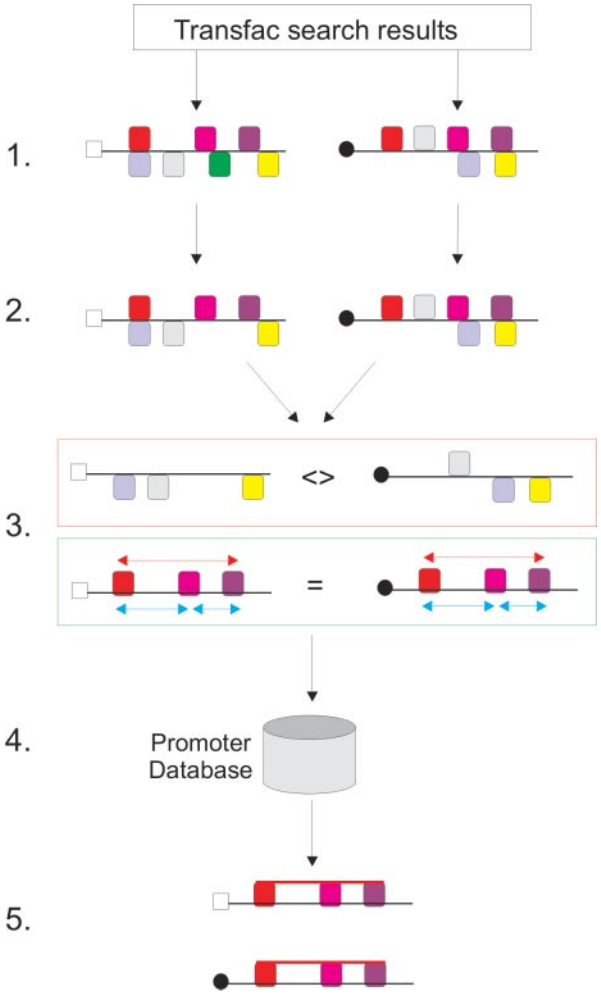


Figure 1. Pattern finding with PromoterPlot. 1: Transcription factor binding sites are predicted using the balance FP/FN option in TransFac. 2: Any factor for which there is evidence that it is not actually expressed or active in the tissue under investigation may be selectively removed from the analysis. 3: Patterns of three factors with conserved internal spacings and consistent binding strands which are found in two or more promoters are retained all others are discarded. 4: The frequency of the predicted patterns is compared with a database of mammalian promoters to estimate the probability of finding the observed results by chance. 5: The results are displayed as an interactive web page with matching genes from the database returned as Affymetrix IDs.

large numbers of binding sites; however, comparisons with published data are often not very good (13). Given that transcription factors bind cooperatively in nature it may be that sub-optimal sites are actually used *in vivo* which are stabilized by interaction with other proteins in a complex (13), but these sites often appear to be below the detection threshold for an FP search. The lower stringency search options of TransFac do a better job of finding the published interacting factors, with the minimize false negatives (FNs) option predicting as many as five binding sites per nucleotide and the balance FP/FN (SUM) option approximately one hit per nucleotide (if all vertebrate matrices are included in the search). However, when combined with our pattern discovery procedure, the SUM option appears to give a good balance between sensitivity and noise.

Clearly, the sequence of the promoter alone is unlikely to be sufficient for the effective modeling of transcription factor assembly because not every cell or developmental stage will necessarily use the same transcription factors to control expression in response to every possible stimulus (14). As not every transcription factor is in an active state in every cell, we can ask which transcription factors are expressed in the cell and what protein modifications might influence their activity? In our institute, we usually address this issue by looking directly at the expression of all transcription factors on Affymetrix microarrays. Any factor that is not expressed in any of the steps of the experiment is unlikely to be present to any significant quantity in regulatory complexes. Similarly, information may also be available from proteomics analysis of samples, which may reveal that certain factors are in inactive phosphorylation states or otherwise excluded from participation in the regulatory machinery. We decided that such information, when available, represents a useful resource for effective modeling and so users may optionally provide a file containing the names of transcription factors or TransFac matrix IDs which will be excluded from the analysis. Alternatively, during the course of expression data mining, we often find that transcription factors themselves are altered in their expression during an experiment. One might hypothesize that some of these changing factors have direct roles in the control of the other genes detected and a question that we are often presented with is 'can you show me what 'factor X' might be doing to my promoters?'. There are two solutions to this problem in our program: (i) if users have specific factors of interest in mind then they can provide a list of these and only patterns containing at least one of the named factors will be displayed; (ii) users can click on the factor names in the legend of the SVG display to see those promoters that contain patterns including specific factors.

Input data required

The potential inputs for PromoterPlot are TransFac result files (internet version), FASTA files (local version only, requires a local copy of TransFac) or Affymetrix IDs (local version only, requires a local copy of TransFac and a local promoter database). The FASTA headers should be kept as short as possible, but can contain start of transcription information in the following format: '>Some_Name' then '#transcription_start_position:start_color#'. Multiple starts can be supplied one after the other in this way. The final characters can be a short description of the gene, for example:

>12345_at#1985:gold#1951:silver#2001:bronze#MyGene

In this example, it would draw a promoter for the gene '>12345_at:MyGene' with a start colored 'gold' at position 1985, 'silver' at 1951 and 'bronze' at 2001. If you wish to use other colors, then you can also enter base-10 RGB values instead of the words in the format 'R, G, B'. We recommend that FASTA titles which do not use the above notation should take the format '>MyGeneName' and avoid using non-alphanumeric characters.

Because the results are active web pages containing server-side scripting, they are stored on our server for a maximum of 72 h for 'anonymous' searches (users get a session ID which they can use to access their data during this time). Users who would like to keep results for longer periods are encouraged to register (free) with a username and a password. Searches stored in this way are kept until the user deletes them or 3 months have passed without access.

Display of results

The primary objective of the display is to make each factor type visually distinct while retaining visual similarities between factors with similar names. It is clear that the same factor must always have the same appearance every time the program is run. The factors are represented by a filled box surrounded by a colored boarder (Figure 2). This two-color approach makes the process of discrimination much easier than with a single color. The fill colors are assigned automatically by taking the name of the transcription factor (e.g. STAT5) and converting the first three letters of the name into their corresponding ASCII values to give the red, green and blue color channels. If the name has fewer than three characters, then the missing characters are replaced by the ASCII code 00. The border color is generated in the same way but now the dominance is reversed so that the color is defined by the ASCII values of the final three characters of the name in the order blue, green and red. Thus, the border colors enable us to visually distinguish factors with very similar names.

The results are displayed in a web page that is composed of three frames: analysis, legend and output. Two viewing modes

Figure 2. The start(s) of transcription are represented by circles colored gold, silver or bronze. Each factor is colored according to its name with the fill color based on the name stem and the border color based on the name ending. The factors and promoters are clickable and clicking provides information about pattern groups, binding sites and other genes which share the same patterns.

exist: the first mode (pattern view) displays only the patterns discovered by the pattern-finding algorithm. Clicking on a factor or its corresponding entry in the legend will display all of the patterns containing that factor. Additionally, clicking on the box at the 5′ end of each promoter will display all of the patterns found in the selected sequence in any other promoter. The second mode (factor view) displays the individual transcription factors (optionally even those that are not members of patterns). Clicking on a factor will display the location of that factor in all of the promoters. In addition, it will provide the target binding site sequences and positions of these factors in the output window. Thanks to the SVG plug-in, it is possible to zoom and pan the analysis window, facilitating the display of large numbers of sequences. All of the results are stored in a password-protected user folder on our server. In the event that a pattern query takes a very long time to complete, it will continue to run on our server even if the web browser is closed and users may log back in at a later date to view their results.

Assessing the specificity of the results

We have purchased the sequences of 156 000 mammalian promoters from Genomatix. When patterns are predicted by the program, their frequency of occurrence in the test sequence is calculated and compared with their frequency in the control database. If we define a null hypothesis that there is no difference between the two frequencies, then we can test this using a Chi-square test with one degree of freedom (15). Each pattern has a vertical pin associated with it which is clickable. Moving the mouse over the pin brings that pattern to the foreground and makes the others translucent. Clicking on a pin hides all patterns except for selected one and displays the Affymetrix IDs for any of the mammalian promoters from the database which also contain this pattern. The color of the pin indicates the Chi-square result. Those patterns that fail the test have red pins, patterns that pass with a P-value <0.05 have green pins and those that pass with a P-value <0.01 have blue. To partially compensate for small numbers of test promoters, we perform a Yates' correction for discontinuity to reduce the risk of type I errors. Clicking on a gene shows the Affymetrix IDs for all database genes with multiple conserved patterns.

DISCUSSION

Here, we present a new bioinformatic tool (PromoterPlot) for automating the extraction of promoter patterns from microarray-based expression data. Binding sites are predicted using Biobase's 'Match' program from the TransFac suite (16). The pattern prediction process may be filtered to exclude factors that are not believed to be active in the experiment. The patterns identified are displayed in a simple interactive web-based graphical interface and stored on the server for future use in a password-protected user directory. We feel that this may be a useful application for visualizing promoter comparisons and to assist in the identification of regions of potential interest before engaging in time-consuming biochemical characterization. The database hits predicted are also useful for validation. The program may be accessed online at http://promoterplot.fmi.ch.

ACKNOWLEDGEMENTS

Funding to pay the Open Access publication charges for this article was provided by the Novartis Research Foundation.

Conflict of interest statement. None declared.

REFERENCES

1. Figeys,D. (2004) Combining different 'omics' technologies to map and validate protein–protein interactions in humans. *Brief. Funct. Genomic Proteomic*, **2**, 357–365.
2. Werner,T. (1999) Models for prediction and recognition of eukaryotic promoters. *Mamm. Genome*, **10**, 168–175.
3. Werner,T. (2002) Finding and decrypting of promoters contributes to the elucidation of gene function. *In Silico Biol.*, **2**, 23.
4. Maruyama,K. and Sugano,S. (1994) Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene*, **138**, 171–174.
5. Perier,R.C., Praz,V., Junier,T., Bonnard,C. and Bucher,P. (2000) The eukaryotic promoter database. *Nucleic Acids Res*, **28**, 302–303.
6. Bailey,T.L. and Elkan,C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, 14–17 August, Stanford, CA. AAAI Press, Menlo Park, CA, pp. 28–36.
7. Thijs,G., Lescot,M., Marchal,K., Rombauts,S., De Moor,B., Rouzé,P. and Moreau,Y. (2001) A higher order background model improves the detection of regulatory elements by Gibbs Sampling, *Bioinformatics*, **17**, 1113–1122.
8. Werner,T. (2000) Identification and functional modeling of DNA sequence elements of transcription. *Brief Bioinform.*, **1**, 372–380.
9. Frech,K., Quandt,K. and Werner,T. (1997) Finding protein-binding sites in DNA sequences: the next generation. *Trends Biochem. Sci.*, **22**, 103–104.
10. Christoffels,V.M., Grange,T., Kaestner,K.H., Cole,T.J., Darlington,G.J., Croniger,C.M. and Lamers,W.H. (1998) Glucocorticoid receptor, C/EBP, HNF3 and protein kinase A coordinately activate the glucocorticoid response unit of the carbamoylphosphate synthase I gene. *Mol. Cell. Biol.*, **18**, 6305–6315.
11. Klingenhoff,A., Frech,K., Quandt,K. and Werner,T. (1999) Functional promoter modules can be detected by formal models independent of overall nucleotide sequence similarity. *Bioinformatics*, **15**, 180–186.
12. Klingenhoff,A., Frech,K. and Werner,T. (2000) Regulatory modules shared within gene classes as well as across gene classes can be detected by the same *in silico* approach. *In Silico Biol.*, **2**, S17–S26.
13. Kel,A., Kel-Margoulis,O., Babenko,V. and Wingender,E. (1999) Recognition of NFATp/AP-1 composite elements within genes induced upon the activation of immune cells. *J. Mol. Biol.*, **288**, 353–376.
14. Kel,A., Reymann,S., Matys,V., Nettesheim,P., Wingender,E. and Borlak,J. (2004) A novel computational approach for the prediction of networked transcription factors of aryl hydrocarbon-receptor-regulated genes. *Mol. Pharmacol.*, **66**, 1557–1572.
15. Fisher,S. (1962) Use of chi-square in simple cross-over designs. *J. Ment. Sci*, **108**, 406–410.
16. Kel,A.E., Gossling,E., Reuter,I., Cheremushkin,E., Kel-Margoulis,O.V. and Wingender,E. (2003) MATCH: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.*, **31**, 3576–3579.

Aknowledgements

I would like to thank George Thomas for providing me the opportunity to work in his lab and for exposing me to world-class science. A special thank you to Stefano Fumagalli, with whom I shared this project: I am grateful for all that I have learnt from you, I really enjoyed working as a team. Thanks also to Edward Oakeley for his collaboration and Dean Flanders for always providing me with the best material to work with. I am also grateful to Albert Tauler for helping me make the first steps in my PhD. I would also like to thank Francois Natt for providing the siRNAs and Xavier Grana and Angela Tyner for sharing their experimental protocols. Boheringher Ingelheim Foundation and the Novartis foundation provided financial support throughout my PhD.

I wouldn't have reached this point without the wonderful support that I received from my parents. Thank you for believing in me. Mum, thank you for being there for me and for spending all those long hours helping me out with my school homework. Dad, your Sunday trips to the natural history museum probably have something to do with all this. Celine, I know you had to wait a long time for this moment. Without you, these four years would have been very difficult and sad. Thank you for all your love, your patience and your constant encouragements, I love you. Maddalena, you can't read this from where you are, but thank you for everything you have given me, I still carry it in my heart; "volere e' potere" as you used to say.

Appendix A

List of candidate genes derived from microarray comparison of S6/flox and ΔS6/flox mouse livers at different time points after hepatectomy. The list is divided by timepoints, from 0 to 40.hours. Negative fold changes represent situations in which the ΔS6/flox showed a value superior to the S6/flox.

Time 0	ProbeSetID	GeneSymbol	FoldChange	ChangePValue
	93987_f_at	3110001N18Rik	3.23	5.74E-04
	93536_at	Bax	-1.86	7.47E-03
	96020_at	C1qb	1.73	1.12E-03
	161968_f_at	Ccr5	2.52	6.22E-03
	94939_at	Cd53	1.50	6.59E-03
	93445_at	Cd5l	2.45	9.98E-04
	94881_at	Cdkn1a	-3.87	1.02E-03
	98067_at	Cdkn1a	-1.91	3.26E-03
	103284_at	Cyp8b1	1.74	6.68E-03
	160265_at	Eif5	-1.50	9.61E-03
	101587_at	Ephx1	-1.81	1.38E-03
	100050_at	Idb1	-1.81	2.19E-03
	98110_at	Mdm2	-1.57	1.07E-03
	97497_at	Notch1	-1.69	1.88E-03
	98056_at	Phlda3	-1.83	3.22E-03
	96662_at	Ppap2b	1.83	8.48E-03
	101577_at	Rps6	21.43	6.79E-05
	104248_at	Ssr3	1.50	4.80E-03
	104249_g_at	Ssr3	1.69	1.05E-03
	96426_at	Tmsb4x	1.50	2.75E-03
	100397_at	Tyrobp	1.51	4.56E-03
Time 20	**ProbeSetID**	**GeneSymbol**	**FoldChange**	**ChangePValue**
	102807_at	9230112O05Rik	1.64	3.80E-03
	96058_s_at	Aldh2	-1.65	2.82E-03
	104072_at	Apcs	-1.92	3.06E-03
	93536_at	Bax	-1.58	1.36E-03
	98562_at	C1qa	2.28	2.32E-04
	96020_at	C1qb	2.33	1.39E-03
	92223_at	C1qg	2.28	1.46E-04
	103016_s_at	Cd68	1.79	5.66E-03
	103334_at	Crcp	-1.68	5.02E-04
	104354_at	Csf1r	1.57	3.69E-03
	100059_at	Cyba	1.54	6.67E-03
	97013_f_at	Cyba	1.54	8.20E-03
	102879_s_at	Fcgr1	1.50	1.54E-03
	99148_at	Fh1	-2.02	1.28E-04

	92217_s_at	Gp49b	1.52	5.08E-03
	92222_f_at	H2-Q1	2.67	2.12E-05
	98467_at	Itih4	-1.69	7.16E-04
	104093_at	Lsp1	1.53	6.24E-03
	94425_at	Ly86	1.51	9.34E-04
	102974_at	Marco	2.24	6.89E-03
	101468_at	Pfc	1.92	2.27E-04
	98056_at	Phlda3	-2.29	9.59E-03
	102013_at	Rdh6	-1.69	1.96E-03
	98915_at	Rnf149	-1.52	4.77E-03
	101577_at	Rps6	17.79	5.40E-04
	92858_at	Slpi	1.73	7.17E-03
	100397_at	Tyrobp	2.19	6.12E-03
	92558_at	Vcam1	1.56	1.37E-03
Time 30	**ProbeSetID**	**GeneSymbol**	**FoldChange**	**ChangePValue**
	102091_f_at	---	-1.59	2.45E-03
	99849_at	---	-1.79	3.81E-03
	96348_at	0610039C21Rik	-2.27	1.78E-03
	98033_at	1100001H23Rik	1.68	8.84E-03
	95409_at	1110019J04Rik	-1.93	2.10E-03
	96086_at	1110031B06Rik	1.96	9.04E-03
	160266_r_at	1110064N10Rik	-1.85	3.80E-03
	98594_at	1190002N15Rik	1.98	2.01E-03
	160676_at	1810012N18Rik	-1.80	5.76E-03
	95518_at	1810015C04Rik	-4.66	3.74E-03
	96122_at	2310016A09Rik	1.93	9.00E-03
	160283_at	2410005K20Rik	-1.85	3.24E-03
	97866_at	2510049I19Rik	1.88	3.90E-03
	103010_at	2610005L07Rik	-1.70	3.48E-03
	101962_at	2610007K22Rik	1.88	9.52E-03
	102385_at	2610318G08Rik	-1.70	5.30E-05
	98973_at	2610318G08Rik	-2.15	3.85E-03
	96925_at	2810024B22Rik	1.98	2.82E-04
	103071_at	2810429C13Rik	1.80	7.55E-05
	93987_f_at	3110001N18Rik	2.45	1.35E-03
	93437_f_at	4632419I22Rik	1.62	5.52E-03
	103389_at	Aass	-1.64	6.20E-03
	101515_at	Acox1	1.93	8.38E-03
	92796_at	Akp2	-3.77	3.47E-04
	102114_f_at	Angptl4	-1.77	2.55E-03
	96792_at	Apob	1.97	5.88E-03
	104328_at	Aqp9	2.89	7.95E-03
	160371_at	Arl6ip1	1.98	4.39E-03
	103697_at	AW061234	2.27	7.93E-03
	160384_at	Bat1a	-1.74	2.28E-03
	101582_at	BC003262	-1.97	1.06E-03
	103695_f_at	C330007P06Rik	1.74	8.23E-03
	93741_at	C87860	-1.87	1.07E-03
	93445_at	Cd5l	3.02	9.83E-03

100128_at	Cdc2a	2.25	4.86E-03
94881_at	Cdkn1a	-5.12	5.22E-03
101088_f_at	Cnbp1	-2.04	7.79E-03
93550_at	Csrp2	1.53	6.37E-03
102847_s_at	Cyp2a4	-3.42	5.66E-03
102701_at	Cyp2b20	-3.66	3.77E-03
160957_at	D12Ertd7e	1.58	1.19E-03
94502_at	D13Wsu50e	-1.64	2.93E-04
98071_f_at	Dck	1.97	4.56E-03
93843_at	Dhrs1	1.82	2.29E-03
102797_at	Dhrs3	1.73	3.30E-05
160203_at	Dnajc9	1.68	2.47E-03
100535_at	E130105L11Rik	-1.93	7.45E-03
104666_at	E430039K05Rik	-1.74	7.24E-03
97411_at	Ect2	2.29	1.42E-03
160265_at	Eif5	-2.39	4.03E-04
94393_r_at	Elovl2	1.88	1.14E-03
97317_at	Enpp2	-2.39	7.46E-03
101587_at	Ephx1	-1.62	6.38E-03
98076_at	Erp29	1.71	3.74E-03
160451_at	Etf1	-1.67	8.86E-04
98608_at	Etf1	-2.02	4.50E-03
100494_at	Fgf1	2.06	4.95E-03
92808_f_at	Fkbp4	2.51	3.83E-03
92697_at	Foxa1	1.77	6.51E-03
103036_at	G22p1	1.98	7.99E-03
96336_at	Gatm	1.76	3.82E-03
103498_at	Gcgr	1.51	1.95E-03
97819_at	Gsto1	-1.64	9.26E-03
99180_at	Gtpbp4	-1.52	4.10E-03
96710_at	H2av	1.93	9.34E-03
103534_at	Hbb-b1	5.80	7.23E-03
92590_at	Hmgcs2	1.96	3.29E-03
160104_at	Hsd3b7	2.04	1.39E-03
96594_at	Hspa4	3.58	2.00E-03
96172_at	Ian1	1.54	4.92E-03
160092_at	Ifrd1	-4.54	7.05E-04
97987_at	Igfals	2.65	2.83E-03
92739_at	Ivl	-1.94	2.31E-03
99632_at	Mad2l1	1.77	1.00E-03
100062_at	Mcm3	1.90	6.86E-03
160496_s_at	Mcm3	2.27	2.59E-03
93041_at	Mcm4	2.24	9.41E-04
102699_at	Mx2	-3.24	5.29E-04
98587_at	Nap1l1	2.10	6.64E-03
103907_at	Nedd4l	-1.53	7.49E-03
92569_f_at	Nol5	-1.91	9.97E-03
97824_at	Nola2	-1.73	9.54E-03
101002_at	Oazin	2.32	7.79E-03

	100737_at	Onecut1	1.77	3.46E-03
	99056_at	Pcbd	1.56	7.16E-03
	101065_at	Pcna	2.03	3.83E-03
	100554_at	Pdlim1	-1.82	1.45E-04
	93619_at	Per1	-1.52	2.57E-03
	99469_at	Pex6	1.58	4.31E-03
	97965_at	Pla2g6	-2.22	3.81E-03
	102663_at	Plaur	-1.50	1.90E-03
	103207_at	Pola1	1.86	5.96E-04
	99019_at	Por	-1.58	4.73E-03
	93485_at	Ptprd	1.52	1.65E-03
	160197_at	Pycrl	1.75	6.18E-03
	92659_at	Rapgef4	-1.59	2.96E-03
	104476_at	Rbl1	1.80	7.14E-03
	160759_at	Rfc2	1.78	2.97E-03
	103418_at	Rfc4	1.88	7.67E-04
	161787_f_at	Ris2	1.78	1.39E-03
	94024_at	Ris2	1.57	8.75E-03
	93782_at	Rnf4	-1.89	8.22E-03
	101889_s_at	Rora	-1.89	5.13E-03
	98081_at	Rpo1-3	-1.72	2.34E-03
	101577_at	Rps6	22.50	6.71E-05
	98950_at	Rragc	-1.53	8.69E-03
	100612_at	Rrm1	1.63	0.00665121
	100985_at	Siah1a	-1.57	3.43E-03
	98596_s_at	Siat9	-3.28	5.38E-03
	100916_at	Slc22a1	1.77	3.24E-04
	104560_at	Slc25a28	-1.56	3.32E-03
	94797_at	Slc26a1	3.30	8.24E-03
	101877_at	Slc31a1	2.09	5.37E-03
	99133_at	Slc3a2	-2.14	1.12E-03
	97309_at	St13	1.79	9.71E-03
	96326_at	Tat	-1.62	1.67E-03
	102354_at	Tcf19	1.78	7.05E-03
	102315_at	Tex292	-1.93	6.56E-03
	93728_at	Tgfb1i4	2.39	7.02E-03
	103794_i_at	Timd2	2.26	5.14E-03
	101964_at	Tkt	-1.55	2.98E-03
	93236_s_at	Tyms	1.63	4.49E-03
	93237_s_at	Tyms-ps	1.55	4.69E-04
	160605_s_at	Usp38	-1.52	6.15E-03
	95709_at	Vkorc1	1.54	8.89E-03
	99963_at	Zfp101	1.69	6.10E-03
	103753_at	Zzz3	-1.64	9.31E-03
Time 40	**ProbeSetID**	**GeneSymbol**	**FoldChange**	**ChangePValue**
	96156_at	1110008H02Rik	-2.17	5.70E-03
	95409_at	1110019J04Rik	-1.81	4.33E-03
	95690_at	1110030L07Rik	-1.61	1.82E-03
	104670_at	1700065A05Rik	1.59	4.67E-03

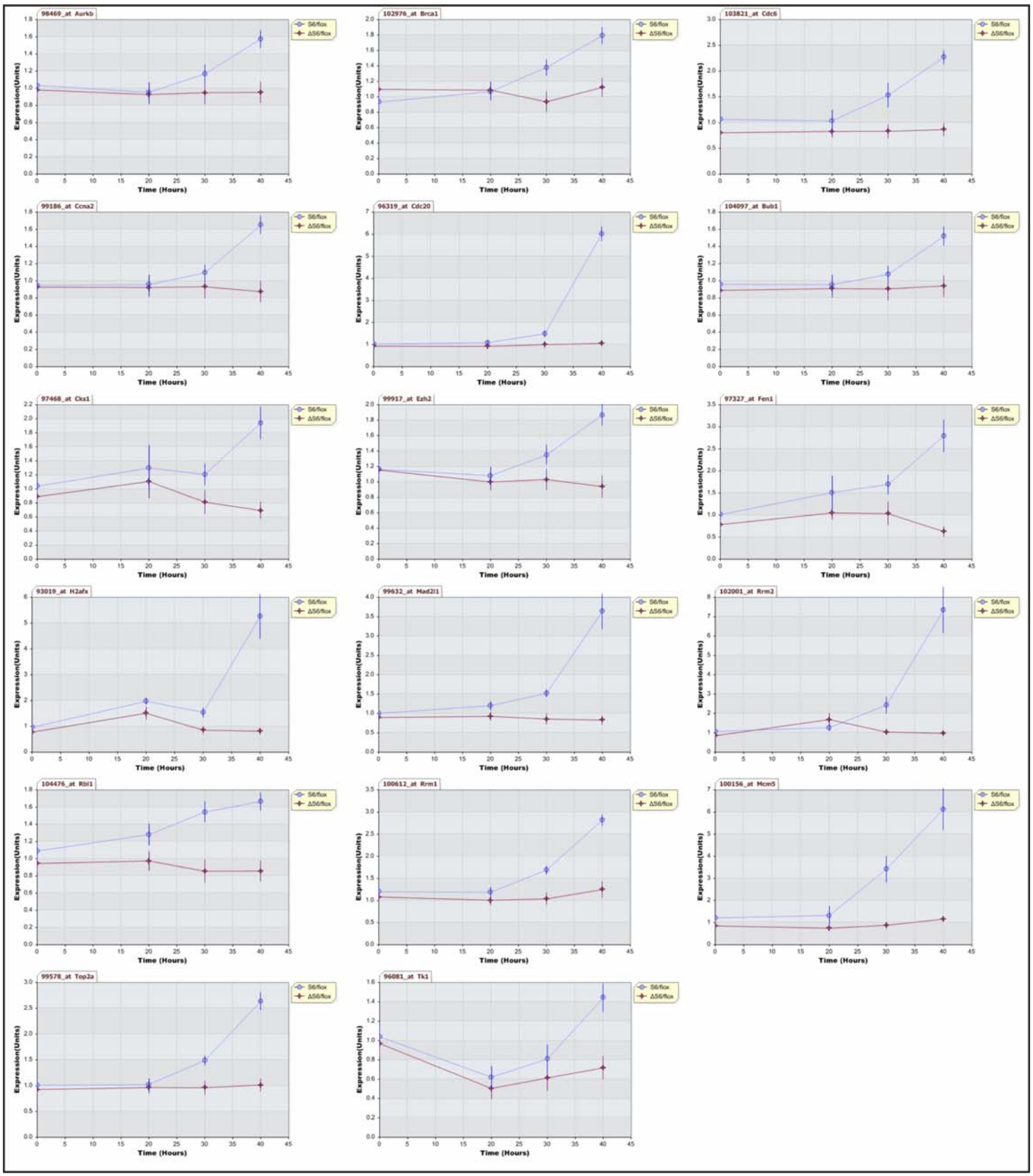
103027_at	1810030O07Rik	-1.54	6.26E-04
95406_at	1810037I17Rik	1.70	9.38E-03
104588_at	1810073K19Rik	-1.91	4.85E-06
160569_at	2310008M10Rik	1.98	1.70E-04
161332_f_at	2610036L13Rik	1.52	6.29E-03
102385_at	2610318G08Rik	-1.52	2.87E-03
100955_at	2700084L22Rik	2.76	2.85E-03
96016_at	2700094K13Rik	2.97	4.86E-03
93441_at	2700099C18Rik	1.81	5.13E-03
104423_at	2810047L02Rik	2.09	3.11E-04
100116_at	2810417H13Rik	2.57	1.48E-03
160279_at	4930588M11Rik	-1.68	9.90E-03
95753_at	A730011O11Rik	2.36	2.08E-03
104578_f_at	Actn1	1.69	2.38E-03
104579_r_at	Actn1	1.54	1.23E-03
98999_at	Adsl	-1.62	8.92E-03
95378_at	AI842396	-1.58	6.78E-04
160090_f_at	Aldo1	-1.83	1.64E-03
96784_at	Anln	2.82	3.59E-03
160371_at	Arl6ip1	1.97	4.03E-03
104745_at	Arl6ip2	-2.06	2.74E-04
93798_at	Atp1a1	-1.59	2.54E-04
93984_at	Atpi	1.64	6.72E-04
98469_at	Aurkb	1.65	3.60E-03
100944_at	AW112010	1.97	1.94E-03
98545_at	Bcap37	-1.86	4.87E-03
101521_at	Birc5	2.76	6.03E-04
102976_at	Brca1	1.60	8.94E-03
104097_at	Bub1	1.61	1.24E-03
160585_at	Bxdc1	-1.59	9.92E-03
102773_at	Car8	1.61	4.65E-03
104259_at	Cbx5	1.86	7.36E-03
99186_at	Ccna2	1.89	8.42E-05
160159_at	Ccnb1	5.84	1.53E-03
94294_at	Ccnb2	3.21	1.90E-04
96319_at	Cdc20	5.70	1.08E-04
94048_at	Cdc34	-1.58	4.36E-03
103821_at	Cdc6	2.63	5.08E-04
94217_f_at	Cdca3	2.03	7.25E-05
160638_at	Cdkn2c	1.97	8.78E-05
94971_at	Cdkn3	1.53	1.17E-03
96346_at	Cdo1	-1.68	3.76E-03
100616_at	Cenpa	2.01	5.56E-04
101538_i_at	Ces3	2.43	9.19E-03
104322_at	Ckap2	2.22	3.99E-04
97468_at	Cks1	2.79	6.28E-03
101093_at	Col4a1	-1.69	7.50E-03
94784_at	D030034H08	1.70	1.23E-03
160310_at	D19Bwg1357e	-1.82	7.85E-03

97295_at	D4Ertd421e	2.64	2.86E-03
93257_at	Ddx1	-1.53	1.15E-03
98489_at	Dlg7	1.60	5.69E-03
97880_at	Dlst	-1.66	1.89E-03
96680_at	Dnajb9	2.24	1.92E-03
100535_at	E130105L11Rik	-2.00	7.55E-03
97205_at	Eif3s1	-2.01	4.35E-03
160265_at	Eif5	-3.42	6.66E-05
99917_at	Ezh2	1.99	9.80E-03
94075_at	Fabp1	1.81	1.58E-03
97327_at	Fen1	4.46	4.39E-03
160648_at	Fignl1	2.86	2.89E-03
99546_at	Fkbp2	1.53	2.40E-03
160069_at	Gmnn	1.71	2.81E-03
103808_at	Golga5	-1.72	4.54E-03
93019_at	H2afx	6.48	6.55E-03
101954_at	H2afz	3.66	2.10E-03
96710_at	H2av	2.72	4.48E-04
97895_f_at	Hat1	1.72	1.28E-03
162457_f_at	Hba-a1	3.52	3.12E-03
94781_at	Hba-a1	11.60	5.00E-03
101869_s_at	Hbb-b1	13.31	6.06E-03
94805_f_at	Hist1h2ac	2.24	1.22E-03
93023_f_at	Hist2h3c2	1.72	7.43E-03
93095_at	Hmgb1	1.54	9.93E-03
93250_r_at	Hmgb2	3.34	3.31E-03
94795_at	Hsd3b4	4.89	7.94E-03
93752_at	Iars	-1.71	8.03E-03
160092_at	Ifrd1	-4.69	1.95E-03
95034_f_at	Ipo4	-1.78	9.18E-03
160501_at	Kif20a	3.33	1.49E-03
161856_f_at	Kif20a	1.60	1.26E-04
95118_r_at	Kif22	1.78	4.61E-03
160755_at	Kif2c	2.57	1.72E-04
97903_at	Leng5	1.54	9.22E-04
160517_at	Lmnb1	2.03	2.91E-03
99632_at	Mad2l1	4.39	3.90E-03
95721_at	Mapkapk2	-1.87	2.45E-03
160496_s_at	Mcm3	2.53	3.76E-03
100156_at	Mcm5	5.30	6.37E-03
104033_at	Mgea6	-1.60	8.82E-04
93342_at	Mki67ip	-1.63	4.00E-03
100885_at	Nek2	3.02	3.63E-03
161000_i_at	Nusap1	3.91	4.50E-03
100720_at	Pabpc1	-1.75	2.69E-04
93615_at	Pbx3	1.73	8.10E-03
93308_s_at	Pcx	-1.55	3.54E-03
98102_at	Pdha1	-1.80	5.28E-04
98056_at	Phlda3	-2.01	1.04E-03

99926_at	Pigr	2.00	7.03E-03
101350_g_at	Plk1	1.64	9.02E-03
93099_f_at	Plk1	2.15	1.70E-03
98996_at	Plk4	2.35	5.35E-03
99019_at	Por	-1.70	4.51E-03
94915_at	Ppib	1.60	6.34E-03
93495_at	Prdx4	1.93	1.17E-03
97560_at	Psap	1.62	4.45E-03
94953_at	Racgap1	1.60	2.70E-03
93676_at	Rad51ap1	1.52	4.81E-03
104476_at	Rbl1	1.94	1.79E-03
94024_at	Ris2	1.92	6.58E-04
101577_at	Rps6	16.29	1.10E-04
100612_at	Rrm1	2.26	2.31E-03
102001_at	Rrm2	7.54	5.83E-03
93548_at	Sec61b	1.84	5.31E-03
103345_at	Spna2	-1.58	4.03E-03
92639_at	Stk6	5.10	1.16E-03
97238_at	Tacc3	1.70	1.93E-03
102354_at	Tcf19	2.43	9.93E-03
104601_at	Thbd	-1.97	5.90E-03
96081_at	Tk1	2.01	9.78E-03
92782_at	Tmpo	1.55	3.23E-03
99578_at	Top2a	2.60	1.50E-03
100343_f_at	Tuba1	1.68	1.29E-03
98759_f_at	Tuba2	1.91	1.26E-03
101543_f_at	Tuba6	1.73	4.19E-04
93237_s_at	Tyms-ps	2.38	2.61E-03
99564_at	Uhrf1	2.19	3.02E-03
160321_at	Zfp216	-1.68	6.40E-04
101890_f_at	Zrf2	-1.86	5.29E-03

Appendix B

Graphics of CCIND genes expression levels



Curriculum Vitae

Born: 12/01/1978 in Palermo (Italy)
Nationality: Italian

Education

2000- 2004	**PhD in Cell Biology (obtained *cum laudae*)** Friedrich Miescher Institute**,** Basel (Switzerland) Supervisor: Dr. George Thomas
1996-2000	**BSc. (Honours) degree,** in **Molecular Genetics and Biotechnology** at the University of Sussex**,** Brighton (U.K.) Graduated with 2(i) honours.
1996	**Baccalaureate at the European School of Brussels,** Brussels (Belgium)

Research Experience

2000 – 2004	**Friedrich Miescher Institute (Switzerland)** PhD Thesis: "Disruption of Ribosome Biogenesis Triggers a p21/p53 Mediated Cell Cycle Checkpoint".
1999-2000	**University of Sussex, Brighton (U.K.)** Diploma Thesis: "Production and purification of the eIF4G translation factor using a baculovirus protein expression system".
1998-1999	**Pfizer Pharmaceuticals Ltd. (U.K).** "Studies of the localization of GABA-B receptors in humans".
1997 Aug. – Sep.	**Karolinska institute (Sweden)**. Research internship

Awards

Boehringer Ingelheim foundation PhD. scholarship

Publications

Disruption of ribosome biogenesis triggers a p53 mediated cell cycle checkpoint in vitro and in vivo
Di Cara A., Fumagalli S., Natt F. and Thomas G.
[Manuscript in review]

Hairpin RNA, a secondary structure of primary importance.
Svoboda P.**, Di Cara A.**
[Invited Review CMLS 2006, in press]

PromoterPlot a web service to find potential regulatory regions in mammalian promoters.
Di Cara A., Schmidt K., Hemmings B.A. and Oakeley E.J.
[Nucleic Acids Res. 2005 Jul 1;33]

Disruption of Ribosome Biogenesis Triggers a p21/p53 Mediated Cell Cycle Checkpoint
Di Cara A.
[PhD. Thesis (University of Basel, 2004)]

Ribosome biogenesis and novel cell cycle checkpoints
Di Cara A.
[Futura 3(16) 2001