Identification of Target Genes of the p16INK4A-pRB-E2F Pathway*

Richard Vernell‡, Kristian Helin‡§, and Heiko Müller‡§

From the ‡Department of Experimental Oncology, European Institute of Oncology, 20141 Milan, Italy and §The FIRC Institute of Molecular Oncology, 20139 Milan, Italy

Deregulation of the retinoblastoma protein (pRB) pathway is a hallmark of human cancer. The core members of this pathway include the tumor suppressor protein, pRB, which through binding to a number of cellular proteins, most notably members of the E2F transcription factor family, regulates progression through the cell division cycle. With the aim of identifying transcriptional changes provoked by deregulation of the pRB pathway, we have used cell lines that conditionally express a constitutively active phosphorylation site mutant of pRB (pRBACDK) or p16INK4A (p16). The expression of pRBACDK and p16 resulted in significant repression and activation of a large number of genes as measured by high density oligonucleotide array analysis. Transcriptional changes were found in genes that are essential for DNA replication and cell proliferation. In agreement with previous results, we found a high degree of overlap between genes regulated by p16 and pRB. Data we have obtained previously for E2F family members showed that 74 of the genes repressed by pRB and p16 were induced by the E2Fs and 23 genes that were induced by pRB and p16 were repressed by the E2Fs. Thus, we have identified 97 genes as physiological targets of the pRB pathway, and the further characterization of these genes should provide insights into how this pathway controls proliferation. We show that Gibbs sampling detects enrichment of several sequence motifs, including E2F consensus binding sites, in the upstream regions of these genes and use this enrichment in an in silico filtering process to refine microarray derived gene lists.

Data generated in the last decade have pointed to a central role for the retinoblastoma protein (pRB) pathway in regulating the progression through the G1 phase of the mammalian cell cycle (1). The core members of this pathway include, in addition to pRB (and its family members p107 and p130), the D-type cyclins that, in association with CDK4 and CDK6, promote proliferation of the cell cycle, in part through the phosphorylation of the pRB family members, and the INK4 family of cyclin-dependent kinase inhibitors that specifically bind and inhibit the activity of CDK4 and CDK6. The high frequency by which alterations have been identified in the pRB pathway in human cancer indicates an understanding of the central role of its core members in regulating cell proliferation. We have led several researchers to suggest that the deregulation of the pRB pathway is an obligatory event in human cancer (e.g., see Ref. 2).

The biochemical mechanism by which pRB is restricting cell proliferation is widely believed to involve protein-protein interactions. Several hundred proteins have been reported to bind to pRB (3), however, the relevance of most of these interactions is poorly understood. The most studied and the best understood targets for pRB are members of the E2F transcription factor family (4, 5). The E2F transcription factors are essential for the proper transcriptional regulation of a number of genes, whose gene products control the progression through the cell cycle. These genes include CCNE1 (cyclin E1), CCNA2 (cyclin A2), and CDC25A, which are all essential for the entry into the S phase of the cell cycle, and genes that are involved in the regulation of DNA replication, such as CCND1, DHFR, and TK1 (thymidine kinase) (6, 7). When E2Fs bind to the promoters of these genes in complexes with pRB or its family members, gene expression is repressed. The repression is effectively relieved by CDK1-mediated phosphorylation of the pRB family members, resulting in transcription of the E2F-regulated genes and subsequent progression through the cell cycle.

Because of the central role of the E2Fs in regulating the progression through the cell division cycle, we and others have used gene expression profiling and chromosomal immunoprecipitation assays to identify novel E2F target genes (5, 8–13). The data obtained from these expression profiles have shown that the E2Fs, in addition to having a role in regulating the expression of genes involved in cell cycle progression and DNA replication, also regulate the expression of genes controlling differentiation, DNA repair, and apoptosis. The aim of the current study was to identify transcriptional changes provoked by the reintroduction of a functional pRB pathway in human tumors. By performing such a study we expected to identify genes whose deregulation, as a result of a non-functional pRB pathway, could contribute to transformation. Moreover, we also expected that the gene expression would give valuable therapeutic biomarkers, which could be useful when compounds aimed at restoring a functional pRB pathway in tumor cells will be tested. To perform these studies, we chose the widely used human osteosarcoma cell line U-2 OS as an experimental model because these cells do not express p16. Moreover, because of the high level of CDK activity in U-2 OS cells, the majority of pRB is in its hyperphosphorylated state. Reintroduction of p16 and a non-phosphorylatable version of pRB
results in cell cycle arrest in these cells, which can be abrogated by re-expression of the E2Fs (14–18). Here, we report the identification of 74 genes, whose expression is repressed by p16 and pRB, and induced by E2F. Gibbs sampling performed in the upstream regions of these genes detects the enrichment of several sequence motifs, including E2F consensus binding sites. The enriched sequence motifs were used in a search of roughly 10,000 human upstream regulatory regions to identify novel E2F target gene candidates and to refine the microarray data.

EXPERIMENTAL PROCEDURES

Cell Culture and Cell Cycle Analysis—U-2 OS cells expressing tetra-cycline responsive p16 or constitutively active pRB (pRbΔCDK) alleles were described previously (17, 18). Cells were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% bovine calf serum (South American), 2 μg/ml puromycin, 200 μg/ml G418, and 2 μg/ml tetracycline. U-2 OS cells expressing HA-tagged ER-E2F1 (5) were cultured in Dulbecco’s modified Eagle’s medium containing 10% bovine calf serum and 2.5 μg/ml puromycin. Nearly confluent cultures of U-2 OS clones were trypsinized and plated at 5 × 10⁵ cells per 15-cm plate on the day before induction. Induction of E2F activity was accomplished by adding tetracycline to a final concentration of 30 μg/ml for 4 h. Exogenous p16 or pRbΔCDK proteins were induced by mixing cells twice with phosphate-buffered saline followed by growing in tetracycline-free media. After the first hour of culture the media was changed again. RNA was isolated at 12 h for pRbΔCDK, or 16 h for p16. For cell cycle analysis, cells were pulse labeled for 30 min in 20 μCi [3H]thymidine, harvested by trypsinization, fixed in 1% paraformaldehyde for 5 min, and permethanized in 70% methanol. DNA was denatured in 2 N HCl for 20 min and cells were neutralized in 0.1 M sodium borate, pH 8.5, for 2 min. Bromodeoxyuridine was detected using fluorescein isothiocyanate-labeled anti-bromodeoxyuridine monoclonal antibody (BD Biosciences), and DNA was stained in 0.5 ml of propidium iodide (10 μg/ml).

RNA Preparation—RNA for analysis by microarray and quantitative real-time reverse transcriptase-PCR was isolated using the RNeasy kit (Qiagen) and integrity was determined by formaldehyde-agarose gel electrophoresis. For use in quantitative PCR, the protocol included on-column treatment with DNase I according to the manufacturer’s protocol.

Immunostaining—Cells were cultured in the presence or absence of tetracycline for 16 h. Expression of HA-tagged pRbΔCDK protein was detected with anti-HA (12CA5) antibody. Exogenous p16 protein was detected with anti-p16 (Santa Cruz sc-467).

Quantitative Reverse Transcriptase-PCR—Changes in gene expression were confirmed using SYBR Green quantitative reverse transcriptase-PCR following the protocols from Applied Biosystems. Primers used can be found in Supplemental Materials.

High Density Oligonucleotide Microarrays—We used HG-U95 A, B, C, D, and E oligonucleotide microarrays (Affymetrix) containing 62,907 probe sets. Targets for hybridization to the microarrays were prepared from total RNA for analysis by microarray and quantitative real-time reverse transcriptase-PCR. Images were scanned with a Hewlett-Packard GeneArray scanner and the images were analyzed using Affymetrix’s Microarray suite 5.0 software. All targets were hybridized to two different sets of HG-U95 microarrays (chips Av2 through E). Chip replicas were analyzed as described in Ref. 5. The probe sets that were identified as regulated by the analysis of chip replicas were inserted into GeneSpring™ 5.1 software as gene lists for visualization, annotation, and clustering purposes. All data obtained in the array experiments will be made available.4

Generation and Use of the Promoter-Exon 1 Data Base—We used the University of California Santa Cruz genome browser (assembly released 28th of June 2002)5 to retrieve genomic DNA sequence for 16,031 human Reference Sequence accession numbers (NM_001234) available as of 02/11/02 including 1000 bp of upstream sequence as reported by the genome browser. Then, we blasted the exon 1 sequence as reported by the genome browser to 100 bp of the 5′-end of the cDNA as reported by the reference sequence entry. Only if this blast reported a hit with 99 or 100% identity in the correct orientation (plus/plus strand), did we consider the promoter-exon 1 sequence for further analysis. This procedure reduced the number of entries in the promoter-exon 1 database from 16,031 to 9,652 entries. All position weight matrix searches using this data base were performed exactly as described by Ref. 21.

Position Weight Matrix Similarity (P WM Similarity)—The test for P WM similarity was performed in two different ways. 1) Using 50,000 random 25-mer oligonucleotide sequences, each sequence was presented to each P WM at the stringency indicated in Table II. A P WM score was calculated as described in Ref. 21. If the P WM score for that sequence was equal or higher than the stringency indicated in Table II, the oligonucleotide sequence was scored as a hit. In this test, similar matrices have the tendency to recognize similar sequences as a hit. The degree of similarity is quantified using the binomial distribution where the number of successes is the number of hits that two PWMs have in common, the number of trials is the number of hits of one PWM, and the probability of success is the number of hits of the second PWM divided by 50,000,000. The cumulative p value was then calculated. PWMs were judged similar if this value fell between 0.99999 and 1 (i.e. the number of successes lies far out in the tail of the distribution). The relatively high number of random oligonucleotide sequences was necessary because half of the PWMs would recognize only 1 hit in roughly 10,000 sequences. 2) To visualize the similarity of PWMs, the same random sequences were presented to the PWMs at a stringency that would recognize 5% of the sequences as a hit. By doing so, each sequence was assigned a pattern of 0s and 1s, where 0 stands for no hit with a given PWM and 1 stands for a hit with a given PWM. These patterns were analyzed by cluster analysis using the GeneSpring 5.1. software (Silicon Genetics).

Scoring Scheme for in Silico E2F Target Gene Filter—Based on the test for PWM similarity, each PWM was assigned to a PWM family: 1) E2F site (M00024, M00050, M00180, motif 5, motif “Kel et al.” (21), and motif “Farnham”); 2) GC-rich (motif 7, motif 20, motif 33, motif 63, and motif 71); 3) CCAT box (motif 4); and 4) CCAT-like (motif 45). If within a PWM family more than one PWM recognizes a binding site in a promoter, only the PWM with the highest enrichment factor is considered. The enrichment factors of the PWMs that recognize a binding site in a promoter are multiplied to give the final score. Multiplication of enrichment factors is justified because the probability to find two or more motifs in the same promoter is calculated by multiplying the probabilities of finding each motif in the promoter.

RESULTS

Previously, we have described the identification of several hundred genes whose expression changes significantly upon activation of E2F1, E2F2, or E2F3 (5). In total, between 7 and 10% of all genes tested in this screen showed changes in expression level. Computer-assisted motif searches in the promoters of these genes did not show a significant enrichment for the upstream regions of these genes detects the enrichment of several sequence motifs, including E2F consensus binding sites. The enriched sequence motifs were used in a search of roughly 10,000 human upstream regulatory regions to identify novel E2F target gene candidates and to refine the microarray data.

Although the more obvious explanation is that many of the genes are false positives, and several of them are not direct targets of the E2Fs, we do not believe that this is the case, because Northern blotting or quantitative PCR confirmed the microarray results for more than 95% of the genes tested (Ref. 5, and data not shown). Moreover, our current understanding of the E2F binding site consensus and more generally the mechanism of E2F-mediated regulation of transcription is limited. Several E2F-regulated genes have been described (e.g. CCNE1, CCNA2, and MYBL2) that do not contain perfect consensus E2F DNA binding sites (TTTSSCCG). There are several possibilities that could account for this observation. Although the more obvious explanation is that many of the genes are false positives, and several of them are not direct targets of the E2Fs, we do not believe that this is the case, because Northern blotting or quantitative PCR confirmed the microarray results for more than 95% of the genes tested (Ref. 5, and data not shown). Moreover, our current understanding of the E2F binding site consensus and more generally the mechanism of E2F-mediated regulation of transcription is limited. Several E2F-regulated genes have been described (e.g. CCNE1, CCNA2, and MYBL2) that do not contain perfect consensus E2F DNA binding sites (TTTSSCCG). There are several possibilities that could account for this observation. Although the more obvious explanation is that many of the genes are false positives, and several of them are not direct targets of the E2Fs, we do not believe that this is the case, because Northern blotting or quantitative PCR confirmed the microarray results for more than 95% of the genes tested (Ref. 5, and data not shown). Moreover, our current understanding of the E2F binding site consensus and more generally the mechanism of E2F-mediated regulation of transcription is limited. Several E2F-regulated genes have been described (e.g. CCNE1, CCNA2, and MYBL2) that do not contain perfect consensus E2F DNA binding sites (TTTSSCCG). There are several possibilities that could account for this observation.

4 H. Muller, unpublished data.
changes induced by pRB and/or p16. Because endogenous E2F activity is repressed by the induction of pRB or p16 in U-2 OS cells, genes that in this setting show a pattern of regulation opposite to the one observed for induction of E2F activity are likely to be the relevant E2F target genes that are deregulated in cancer cells upon loss of p16 or pRB function.

We took advantage of two U-2 OS-derived cell lines that overexpress either pRbΔCDK (18) or p16 (17) in a tetracycline-regulated manner. Fig. 1 shows that both cell lines uniformly express high levels of p16 or pRbΔCDK as tested by immunostaining after removal of tetracycline (Fig. 1, A and B). This observation was confirmed by Western blot analysis showing that both pRbΔCDK and p16 reach high levels of expression 20 h after removal of tetracycline from the culture medium (Fig. 1, E and F). Both proteins were undetectable at time 0 h and reached near maximum levels of expression 16 h after induction. In the p16 cell line, endogenous pRB began to accumulate in its hypophosphorylated form (arrow) upon removal of tetracycline from the medium. F, Western immunoblot analysis demonstrating robust expression of HA-pRbΔCDK (indicated by arrow) upon removal of tetracycline.

G, expression of the E2F-repressed PAI1 gene was monitored upon induction of p16 or HA-pRbΔCDK by Northern blot analysis.
pRbΔCDK and p16 expression is capable of arresting the cell cycle (17, 18). Cell cycle arrest in turn can regulate the expression of hundreds of genes (e.g. see Ref. 29). We used the induction of PAI1 expression and the repression of cyclin E1 expression as parameters to identify the earliest time point where significant changes in the expression can be observed in the absence of detectable cell cycle arrest (Fig. 1, C and D). The time points chosen for further analysis were 12 h after the induction of pRbΔCDK and 16 h after the induction of p16.

The choice of early time points should ensure maximum specificity in terms of regulation of endogenous E2F activity in the absence of indirect interference because of cell cycle arrest.

We used oligonucleotide gene expression microarrays to identify genes whose expression changes significantly in the two cell lines after induction of pRbΔCDK and p16. Total RNA was isolated 0 (control) and 12 h after induction of pRbΔCDK expression and 0 (control) and 16 h after p16 expression. Each target cRNA was hybridized to two microarrays to facilitate statistical analysis of the results. The data were analyzed using Microarray suite version 5.0 (Affymetrix) and GeneSpring™ (Silicon Genetics) software. We also re-analyzed the microarray data gathered for the induction of E2F1, -2, and -3 activity derived from the Hu6800FL + Hu35k chipset (Affymetrix) that had been analyzed previously using Microarray suite version 4. During this procedure, around 5% of the genes considered previously were discarded as insignificant but the total number of genes found regulated was nearly doubled. These differences are entirely because of improvements in the statistical algorithm used by Microarray suite version 5.0, not to differences in the analysis settings.

To compare the data that were derived from two different chipsets, we restricted the following analyses to probe sets that were present on both chipsets. In total, we analyzed 42,089 probe sets (24,600 Unigenes) in the E2F induction experiments and 62,907 probe sets (40,915 Unigenes) in the pRbΔCDK and p16 induction experiments. 23,527 probe sets (16,897 Unigenes) were found to be comparable among each other and were used for further analysis.

Activation of the E2Fs (E2F1, E2F2, and E2F3) or induced expression of pRB or p16 resulted in significant changes in the expression of a number of genes. For each of the activated or induced proteins, we organized the up-regulated and the repressed genes into a total of 10 lists of regulated genes, 5 lists of up-regulated genes, and 5 lists of down-regulated genes. We asked how to combine the lists in a way that generates a high degree of specificity without excessive loss of potential candidate genes. We used published information about the identity of E2F target genes (Refs. 5 and 8–13, and references therein) and assembled a list of genes that had been studied in detail (bona fide target genes) or that had been found regulated consistently in different high throughput screens. This list contains 47 genes, 38 of which were present on both chip sets we analyzed (Fig. 2A). We reasoned that a meaningful combination of lists of regulated genes should maximize both the total number of known target genes as well as the proportion of known target genes relative to the total number of genes in the combined list. Because the activity of E2Fs is repressed by induction of pRB and/or p16, we focused our attention on combinations of lists with opposing effects on gene expression of the E2Fs on the one hand and pRB/p16 on the other hand. In total, we tested 62 combinations of lists (using the Boolean AND operator, see Appendix 1, Supplemental Materials).

Fig. 2B shows the results of combining lists of E2F up-regulated genes with pRB and p16 down-regulated genes. Two distinct classes of lists are immediately evident. Single lists and combined lists composed only of E2F induction data contained many genes and a low percentage of known E2F targets (type I). Combined lists containing at least one list of E2F up-regulated genes and one list of pRB or p16 down-regulated genes contain significantly fewer genes but a much larger percentage of known E2F target genes (type II). The combination of the E2F1, E2F2, and E2F3 derived lists (or any duplicate combination thereof) does not lead to a significant increase in the percentage of known E2F target genes in the resulting combined list (see Fig. 2B). On the other hand, the p16 derived single list of down-regulated genes contains the highest percentage of known E2F target genes. Interestingly, both the percentage and the total number of known E2F target genes are higher for the single list of p16 down-regulated genes (7.8%, 28 genes) than in the triplicate combination of E2F up-regulated genes (1.4%, 5 genes), despite the similar total number of genes in those lists. This observation strongly suggests that induction of E2F activity alone is not a sufficient stimulus for the induced expression of some bona fide E2F target genes like CCNA2, DHFR, TS, and others in our system. The reason may be that we are using proliferating U-2 OS cells rather than quiescent cells employed in other studies (8, 10). Furthermore, U-2 OS cells seem to express low levels of BRG1, a necessary component for pRB-mediated repression of E2F target genes (30). At least some of the E2F target genes may therefore already be expressed at levels that cannot be increased further by induction of E2F activity, whereas they may be repressed by induction of pRB or p16, depending on the availability of co-repressors. Nevertheless, the combination of lists with E2F up-regulated genes and p16/pRB down-regulated genes leads to a significant increase in specificity of the resulting lists. Note that the number of genes identified in the combined lists is on average nearly 100 times larger than the number of genes that would be expected by chance, whereas the number of genes expected by chance is around 0.5 or less (Fig. 2C). We chose lists with more than 20% of known E2F target genes for further analysis. The lists satisfying this criterion are all composed of both the p16 and the pRB data and one or two E2F derived lists (see Fig. 2B, encircled).

We also compared p16 and pRB up-regulated genes to E2F down-regulated genes. In this case, an apparent increase in specificity as seen in Fig. 2B was not observed (data not shown). Nevertheless, the number of genes found up-regulated with p16 and pRB and down-regulated with at least one E2F was significantly higher than would be expected by chance (Fig. 2C). It is worth noting that none of the genes defined by this filter was a known E2F target gene as defined in Fig. 2A. We will call genes that are up-regulated by E2Fs and down-regulated by p16 and pRB cluster 1 genes, whereas the genes with opposite behavior will be called cluster 2 genes in the following sections.

Fig. 3A shows a gene tree (generated with GeneSpring software) that contains the cluster 1 and cluster 2 genes. Cluster 1 genes are more abundant than cluster 2 genes (55 Unigenes versus 22 Unigenes, see also Table 1). None of the cluster 2 genes is part of our list of known E2F target genes. Cluster 1 genes display a slightly more homogeneous behavior than cluster 2 genes. Fig. 3B shows the -fold change values for cluster 1 and cluster 2 genes in the five experiments. The -fold changes are somewhat more pronounced in the E2F experiments but no striking differences can be observed between cluster 1 and cluster 2 genes in their overall behavior (except for the opposite direction of change). We used quantitative PCR to confirm the microarray-derived gene expression profiles for 25 genes that are evenly distributed over cluster 1 and cluster 2. In all cases, the profiles observed in the microarray experiments were reproduced by this approach (Fig. 3, C and D). All cluster 1 and...
cluster 2 genes are listed in Table I. Table I contains three genes whose expression levels did not change significantly in the pRB and p16 microarray experiments. These genes are \textit{EGR1}, \textit{PAI1}, and \textit{CTGF}. For these genes, we have previously verified by Northern blotting that they are cluster 1 (\textit{EGR1}) or cluster 2 (\textit{CTGF}, \textit{PAI1}) genes (5). \textit{EGR1}, \textit{PAI1}, and \textit{CTGF} are therefore examples of false negatives. Less stringent analysis of the microarray data would eliminate these false negatives at the cost of including potential false positives. We observed that less stringent gene lists exhibit poorer performance in promoter motif searches and therefore preferred to work with high stringency gene lists even though we are aware that we will be missing a number of true pRB pathway target genes.

We were wondering whether E2F binding sites and/or other sequence motifs are present in the promoter regions of cluster 1 and cluster 2 genes. Sequence motif search algorithms can be used to identify motifs that are significantly overrepresented in a set of co-regulated genes. Some of these motifs may be overrepresented because they are specifically required for coordinated expression of this set of genes, whereas others are general constituents of a promoter. To distinguish between these possibilities, we generated a data base of 9652 human promoter-exon 1 sequences (see “Experimental Procedures” for details). We reasoned that specific motifs would be enriched in a set of co-regulated genes, whereas general motifs will not (Fig. 4A).

Transcription factor binding sites are generally described as PWMs. Kel et al. (21) recently described a set of experimentally proven E2F binding sites that in a modified position weight matrix search is capable of selecting E2F target genes \textit{in silico}. The PWM is calculated from a multiple alignment of binding sites of equal length. Given the PWM and applying the calculation as described (21), any sequence motif of the same length as the position weight matrix can be assigned a score. The higher the score the higher the homology of the sequence motif under study to the sequence motifs used to build the PWM. By fixing a certain score value cutoff (stringency), this approach yields yes/no answers. This way a given sequence motif is predicted to be homologous or not homologous to the sequences used to build the PWM. The method is associated with overprediction (false positives) and under prediction (false negatives) errors whose contribution to the total error is dependent on the stringency of the search (Fig. 4, B–D, upper panels). For example, if we used an E2F binding site matrix at a low stringency of the search, we would identify promoters as containing E2F sites even though...
they do not contain them, whereas at a high stringency of the search we would miss promoters that contain E2F binding sites. Because the optimal stringency (the one that minimizes the error sum) depends on the nature of the multiple alignment used to generate the PWM and not on the calculation itself, we tested multiple stringencies and show the associated error as a separate curve for every PWM tested (Fig. 4, B–D, upper panels). The under prediction error was estimated by generating the position weight matrix using all sites of the multiple alignment except for one and then calculating the stringency at which the missing site does not get predicted as being homologous anymore. This procedure is repeated for all sequences that are part of the multiple alignment. Then, the fraction of non-predicted sequences is plotted for every stringency. For example, we used the E2F binding sites identified in promoters that had been shown to bind E2F via chromatin IP as reported (11), generated a position weight matrix using all sites except for one, and calculated the score for the E2F site that had been left out. This procedure is repeated for all the sites. By the end of the procedure, every site has been assigned a score that varies slightly from one motif to the next. At low stringencies, the scores of all sites will be above that stringency and therefore all sites will be recognized as being E2F sites, whereas at higher stringencies some scores will be below that stringency and therefore these motifs will not be recognized as E2F sites. The fraction of these false negatives relative to all the sites constitutes the under prediction error at that stringency.

The data base we used to do motif searches contains 400 bp of sequence per gene around the transcriptional start site (see below). Therefore, the sequence motifs to be recognized are embedded in 400 bp of the promoter-exon 1 sequence. Accordingly, the overprediction error was estimated using the PWM generated from all sequences of the multiple alignment and calculating the highest stringency at which each of 400 random sequence motifs of the same length as the PWM gets predicted as being homologous. Then, for every stringency, the fraction of predicted sequences is plotted. The total prediction error is represented by the sum of the under prediction and the over-

Fig. 3. Gene expression profiles of pRB-p16-repressed and E2F-induced genes (cluster 1) and pRB-p16-induced and E2F-repressed genes (cluster 2 genes). A, gene tree showing behavior of cluster 1 and cluster 2 genes as measured by microarrays. Red, up-regulation; green, down-regulation; yellow, control state of expression. B, -fold expression changes of cluster 1 and cluster 2 genes as measured by microarrays. C, gene tree showing behavior of selected cluster 1 and cluster 2 genes as measured by quantitative PCR. Red, up-regulation; green, down-regulation; yellow, control state of expression. The names of the tested genes are shown in Supplemental Materials. D, -fold expression changes of cluster 1 and cluster 2 genes as measured by quantitative PCR.
### Gene Expression Profiling of the pRB Pathway

#### Table I

| PS_U95 | GenBank™ accession | UID | Common name | Cluster | Known target | Function |
|--------|-------------------|-----|-------------|---------|--------------|----------|
| 886_at | M60527            | 709 | DCK         | 1       | Yes          | Deoxycytidine kinase |
| 5418_at | A1344559          | 1634 | CDC25A      | 1       | Yes          | Inflammation |
| 1491_at | M31166            | 2050 | PTPX3       | 1       | Yes          | Signal transduction |
| 1467_at | U12535            | 2132 | EPS8        | 1       | Yes          | DNA replication/repair |
| 34314_at | X59543           | 2934 | RRM1        | 1       | Yes          | DNA replication/repair |
| 1516_g_at | A1476891         | 6877 | FLJ10483    | 1       |              |          |
| 5017_at | A912772           | 7988 |             |         |              |          |
| 41060_at | M74093            | 9700 | CCNE1       | 1       | Yes          | Cell cycle |
| 63032_at | A1653152          | 9914 | FST         | 1       |              | Signaling |
| 32662_at | X59543           | 2941 | PBX3        | 1       |              | Transcription |
| 789_at | X52541           | 326035 | EGK1   | 1       |              | Transcription |
| 44722_at | A1224533         | 351623 | DFKZp762L0311 | 1 |            |              |

The listed genes were found regulated by pRB and p16 and one of the E2Fs (E2F1, E2F2, or E2F3). Cluster 1 genes are up-regulated by E2F and down-regulated by pRB and p16. Cluster 2 genes are down-regulated by E2F and up-regulated by pRB and p16.
prediction errors (Fig. 4, B–D, upper panels).

Using the procedure described above, we characterized PWMs for the experimentally verified E2F binding sites described (21), as well as for a number of multiple alignments of motifs that have been identified by Gibbs sampling performed on the promoters of known E2F target genes (see below). The total error curve associated with each multiple alignment describes the predictive properties of the corresponding PWM. For the described motifs (21), the minimal total error is found at stringencies around 0.85 (Fig. 4B, upper panel). At this stringency, E2F sites that are present in the promoter of a gene of interest can be identified with the highest confidence. If E2F sites are specifically enriched in the promoters of cluster 1 and/or cluster 2 genes, this enrichment should be detectable at any given stringency. The 50 resulting numbers of genes containing an E2F site at any given stringency can be seen from the lower panel, shows the enrichment factor with increasing stringency of the search (Fig. 4, B–D, middle panels).

Because it is not obvious that an enrichment of E2F binding sites in our clusters 1 and 2, should it be detected, is biologically relevant rather than an artifact, we included a set of genes that had been selected via chromatin immunoprecipitation using E2F1- and E2F4-specific antibodies (E2F-ChIP genes) as a positive control (9, 11). In these genes, E2F binding sites are present in the promoters with very high probability. As a further control, we included the promoters of the known E2F target genes as depicted in Fig. 2A.

Next, we tested cluster 1, E2F-ChIP genes, and known E2F targets for their content of E2F binding sites that are predicted to be homologous to the experimentally proven E2F binding sites used in the study by Kel et al. (21) (see Table II). Fig. 4B, upper panel, shows the errors associated with that search. As can be seen from the middle panel in Fig. 4B, in the range of stringency values from 0.84 to 0.9 that have the best predictive properties (minimal total error sum), all three sets of genes contain significantly more genes predicted to contain an E2F site than could be expected by chance. Fig. 4B, lower panel, shows an increase in the enrichment factor with increasing stringency of the search. Fig. 4, C and D, shows the situation obtained for searches using a position weight matrix with similar predictive properties but describing a general motif. In this case, the curves observed for cluster 1 genes, E2F-ChIP genes, and known target genes fluctuate around the mean (Fig. 4, C and D, middle panel) and no enrichment can be seen with increasing stringency of the search (Fig. 4, C and D, lower panel).

These results suggest (i) that the enrichment for E2F binding sites present in the E2F-ChIP set of genes is detectable using our data base of promoters in conjunction with the PWM search described by Kel et al. (22) and (ii) that the cluster 1 genes show a similar if not even a more pronounced enrichment for E2F binding sites as compared with known E2F target genes. For cluster 2 genes, no such enrichment was detectable (data not shown). The reason for the lack of enrichment is unclear at the moment, but may be purely technical because our promoter data base contained only 10 promoter sequences for the 23 genes belonging to cluster 2. With this low number of promoters, the standard deviation associated with the expected num-
ber of hits is too large to permit reliable detection of enrichment. Therefore, the following analysis was restricted to cluster 1 genes only.

Next, we asked whether similar enrichment could be detected for motifs other than E2F binding sites. To this end, we tested 280 publicly available PWMs (TRANSFAC) on cluster 1 genes, E2F-ChIP genes, and known E2F target genes. The results of some of these searches are shown in Table II. As expected, the matrices M00024, M00050, and M00180 that are describing the E2F binding site show significant enrichment. No significant enrichment was observed for Sp1 (M00008), YY1 (M00069), MYCMAX (M00123), p53 (M00272), the TATA element (M00216), and others (data not shown). In the case of the TATA element, the tendency was clearly toward exclusion from cluster 1, E2F-ChIP, and known E2F target genes, which is compatible with the lack of TATA elements in most known E2F target genes.

The fact that E2F binding sites are significantly overrepresented in three independently defined sets of E2F target genes (known targets, E2F-ChIP targets, and cluster 1 genes) suggests that it might be possible to identify sequence motifs that may cooperate with E2F binding sites in the regulation of E2F target genes. We turned to Gibbs sampling to identify such motifs. In particular, we used the AlignACE program (31, 32) to perform Gibbs sampling on cluster 1, E2F-ChIP, and known E2F target genes. We obtained more than 120 different motifs (data not shown) that are overrepresented in these sets of genes as measured by the background model implemented by AlignACE. To understand whether any of these motifs are biologically relevant for the regulation of E2F target genes, we needed to distinguish the general motifs (e.g. Sp1 site and TATA box) from the specific ones (e.g. E2F site). Therefore, we tested each of the motifs for enrichment in E2F target genes as described in Fig. 4. We call this test the motif specificity test.
Table II
Motifs enriched in E2F target gene promoters

| motif       | logo | stringency | average enrichment | # genes |
|-------------|------|------------|--------------------|---------|
| Kel et al.  | 0.9  | 4.4        | 284                |
| M00024      | 0.85 | 6.8        | 117                |
| M00050      | 0.92 | 3.8        | 462                |
| M00180      | 0.88 | 4.3        | 219                |
| Farnham     | 0.96 | 4.1        | 241                |
| Motif 4     | 0.96 | 4.5        | 215                |
| Motif 5     | 0.98 | 1.4        | 330                |
| Motif 7     | 0.96 | 3.3        | 557                |
| Motif 20    | 0.96 | 2.2        | 255                |
| Motif 33    | 0.96 | 4.2        | 584                |
| Motif 45    | 0.88 | 5.5        | 258                |
| Motif 63    | 0.92 | 4.3        | 630                |
| Motif 71    | 0.96 | 3.4        | 406                |
Whereas AlignACE measures overrepresentation of motifs based on statistical arguments considering only the set of genes that has been used as input, the motif specificity test measures the specificity of a motif to a set of genes of interest as compared with all the promoters in the data base. Most motifs failed the specificity test. However, a significant number of motifs did indeed show specific enrichment in the three independent sets of E2F target genes mentioned above (Table II, motifs 4, 5, 7, 20, 33, 45, 63, and 71). These observations suggest that the data base of promoters is of good quality and that the sets of target E2F target genes are sufficiently refined to allow for the identification of sequence motifs that are enriched together with E2F binding sites.

We were wondering whether the enrichment of sequence motifs other than E2F sites in E2F target genes could be used to generate an in silico filter for E2F target genes. Such a filter could be used to enhance the specificity of microarray-derived gene lists for known E2F target genes or to identify E2F target genes in other tissues, including genes that may not be expressed in U-2 OS cells. To build such a filter, we first characterized all the motifs that had shown significant enrichment in E2F target genes in more detail (Table II). In particular, we identified the stringency for each motif that represents a good compromise between maximum enrichment factor and minimum of total prediction error. As can be seen in Fig. 4B, lower panel, the maximum enrichment factor for E2F motifs is found in the region of high search stringency, a region that is associated with many false negatives but few false positive predictions, and not in the region of minimal total prediction error. Because for most genes in the promoter data base we do not know whether a binding site prediction is correct or not, determination of the true prediction error of a given PWM is impossible. Therefore, we preferred high stringency searches. For every motif, we determined the stringencies at which 5 to 25% of the known E2F target genes are recognized as containing that motif. Within that region we fixed the stringency that is associated with the best enrichment factor for further analysis (see Table II). Table II also shows the enrichment factor observed at that stringency as well as the number of genes that have been identified as containing that motif.

Next, we needed to devise a scoring scheme that would combine the individual predictions for each motif in a way that is most predictive for known E2F target genes. If the motifs were truly independent, just multiplying the enrichment factors for each motif found would represent such a scheme. However, some of the motifs are by definition not independent. For example, the motifs M00024, M00050, M00180, Kel et al., and Farnham are all E2F motifs. If a promoter contained an E2F binding site that is recognized by all these position weight matrices, the resulting score would be largely overestimating the significance of that finding. This situation can be corrected by grouping these five matrices together. Then, if any one of the matrices of that group identifies a promoter as containing an E2F binding site, we use only one enrichment factor (the biggest one) in the scoring scheme.

The situation with the remaining motifs is far less obvious, because we do not even know whether the motifs are similar or not. However, if two matrices are similar to one another, they share a large subset of sequences that are recognized by both matrices at the stringency fixed in Table II. This subset was explicitly identified for 50,000,000 random 25-mer sequences. Then, we asked whether the number of sequences that are recognized by both matrices is significantly larger than the random expectation. This approach identified four PWM families: E2F site (M00024, M00050, M00180, motif 5, motif Kel et al., motif Farnham), GC-rich (motif 7, motif 20, motif 33, motif 63, and motif 71), CCAAT box (motif 4), and CCAAT-like (motif 45) (see “Experimental Procedures” for details of the calculation). These families can be visualized in a cluster tree (Fig 5A).

Every motif shown in Table II is enriched in E2F target genes and it is reasonable to assume that the presence of more than one of these motifs makes it more likely for a gene to be regulated by E2F. We devised a scoring scheme for the quality of the in silico prediction based on the product of enrichment factors present in the promoter (see “Experimental Procedures” for details). By applying the in silico filter to the data base of 9652 promoters we assigned a score to every promoter in the data base. Most promoters have a score of one (i.e. no motif enriched in E2F target genes is present) but some promoters have significantly higher scores (data not shown). To estimate the error rate of the filter, once again we used known E2F target genes as defined in Fig. 2A as positive controls. By increasing the cutoff on the score (i.e. increasing the stringency of the filter) we obtained lists of genes that contain ever less false positives and ever more false negatives. Fig. 5B shows the prediction error of the in silico filter as a function of stringency. When the stringency is set between 1 and 7, the composition of the total error changes from being mainly false-positive errors to being mainly false-negative errors. From stringency 7 to 14, there is a plateau with few changes in the error rate. From stringency 14 onward the false negative rate slowly grows toward 100% with a concomitant increase in the fraction of known E2F target genes among the hits. This fraction reaches 18% at a stringency of 20.

We tested different stringencies of the in silico filter by combining the in silico derived gene lists with the microarray-derived gene lists and asked whether the fraction of known E2F target genes in the combined gene lists would grow. The results of these combinations using stringency 10 for the in silico filter are shown in Fig. 5, C and D. At this stringency, the presence of at least two different cis-acting sequence elements is required in the promoter because no single motif has an enrichment factor that is as big as 10. As can be seen from Fig. 5, C and D, in all cases, in silico filtering of the microarray-derived gene lists leads to a significant enrichment of known E2F target genes in the resulting combined gene lists. The enrichment is much more pronounced for E2F down-regulated genes and for p16/pRB up-regulated genes. The reason for this observation is unclear at the moment, but it may reflect a bias in the set of known E2F target genes toward cluster 1 genes (up-regulated by E2F, down-regulated by pRB and p16). It is worth noting that the genes shown in Fig. 5E have been chosen because they have not yet been reported as E2F target genes. Many (but not all) of the known target genes have been predicted by the in silico filter as well (e.g. MCM5, MCM6, MCM7, DHFR, RRM1, RRM2, RBL1, RFC3, and others). These genes are implicitly shown in Fig. 5, C and D. Some of the novel E2F target genes (e.g. BLM and BTG3) reported here have been predicted by the in silico filter and have also been verified by quantitative PCR (Fig. 3) without being shown in Fig. 5E.

DISCUSSION

We report here the identification of several novel E2F target genes by combining the results of five independent microarray screens (induction of transcriptional activity of E2F1, E2F2, E2F3, induction of expression of p16, and pRB/CDK) and show that the E2F binding site along with other motifs is enriched in the upstream regulatory regions of these genes. A computational approach to the identification of E2F target genes has been described previously (21). This work employed the E2F site to search for novel targets. We show that this E2F motif is enriched 4.4-fold in our set of E2F target genes (Table II). A similar enrichment factor was observed for sets of E2F target...
genes that were identified by chromatin immunoprecipitation or in *bona fide* E2F targets. Therefore, a search based on the E2F motif only is likely to be associated with a high rate of error. In this report, we show that other motifs are co-enriched in the promoters of E2F target genes. It is, therefore, possible to build a more selective *in silico* filter. Genes identified by this *in silico* filter were indeed regulated in our microarray screens (Fig. 5).

To our knowledge, this is the first report that identifies a biologically meaningful motif (the E2F site) in the promoters of mammalian genes that are co-regulated in microarray screens. The methodology works well in yeast (31, 32). In the mammalian system, several difficulties need to be overcome that are mainly related to the quality of the promoter database and the level of noise in microarray screens. The promoter database used in this report is based on sequence annotations performed at the University of California Santa Cruz (Golden Path) and an additional cleaning step that ensures that the 5′-end of the reference sequence cDNA coincides with the exon 1 sequence reported in the genome browser. The noise level of the microarray data was lowered by cross-comparing five independent microarray screens. With these data on hand, we were able to distinguish general motifs found in most promoters from the motifs that are specifically enriched in E2F target genes using the E2F binding site as a positive control. The enriched motifs were grouped into four families, GC-box, CCAAT-box, CCAAT-like box, and E2F binding sites. As opposed to the motif specificity search performed by the AlignACE-ScanACE package (31), our approach takes advantage of biological data rather than statistical arguments to distinguish general from specific motifs. Thus, the utility of our approach is 2-fold. Motifs that are specifically enriched in groups of co-regulated genes may reveal the involvement of a transcription factor binding to that motif in the regulation of these genes. On the other hand, enrichment of a set of motifs in the promoters of co-regulated genes can be used as an *in silico* filter to identify novel members of the co-regulated cluster. This is useful, for instance, when the tissue of interest is of low abundance or hard to work with.

Previously, we reported that E2F target genes cover a wide range of functions, namely that they are involved in the regulation of apoptosis, cell cycle, differentiation, and development. Independent studies have confirmed the relevance of E2F in the regulation of apoptosis via transcription of APAF1 and several caspases (33, 34). However, several published reports hinted toward underrepresentation of cell cycle and checkpoint regulators in our gene lists (8–11, 35). The reason for these discrepancies is unclear but may be related to the fact that we

**Fig. 5.** Identification of E2F target genes using *in silico* filtering of microarray data. A, similarity of PWMs visualized by cluster analysis (see “Experimental Procedures” for details). B, error profile of the *in silico* filter. C, plot of the number of genes versus the fraction of known E2F target genes in the indicated gene lists. Every gene list is shown before and after application of the *in silico* filter. D, fraction of known E2F target genes in each gene list before (blue bar) and after (brown bar) application of the *in silico* filter. E, microarray data of some novel E2F-target genes that have been predicted *in silico* and have been found regulated in at least one microarray experiment.
studied proliferating U-2 OS as opposed to quiescent cells used by other groups. Furthermore, U-2 OS cells express low levels of BRG1 whose activity contributes to pRB-mediated repression of transcription (30, 36). In this system, many E2F target genes may be expressed at levels that cannot be increased significantly by induction of E2F activity. Moreover, the use of asynchronous cells may render signals from genes that are sensitive to E2F-mediated regulation only at specific points in the cell cycle less robust.

We observed that induction of pRBΔCDK was capable of reverting the changes induced by activation of E2F (5). In particular, induction of pRBΔCDK expression led to decreased levels of CCNE1 and EGR1 mRNAs, two genes that are induced by activation of E2F, while the levels PAI1 and CTGF mRNAs (whose expression levels are down-regulated by E2F activity) were induced. These results suggest that the pRB-E2F pathway-specific genes can be identified by cross-comparison of gene lists that are specific for activators of the pathway (E2F1–3) with gene lists derived from suppressors of the pathway (pRB, p16). Importantly, cross-comparison of gene lists reduces the level of noise but enhances the significance of weak signals. Consequently, regulators of the cell cycle, DNA replication, DNA repair, and checkpoints constitute a major fraction of pRB-E2F pathway-specific genes defined here, in agreement with other reports (8–11, 35).

Differentiation and development are complex processes that are governed by a multitude of pathways. Mutual cross-talk between a number of these pathways has been well documented (20, 37–48). The genes reported here provide starting points to investigate mechanistic links between some of these pathways and the pRB-E2F pathway. Most signal transduction pathways are composed of a ligand, a receptor, one or more signal transducers, a transcription factor, and target genes. Some of the target genes have feedback functions to monitor the activity of the pathway, whereas others perform some biological function, often by influencing the activity of other pathways at one or more levels mentioned above (49). The nature of the pRB-E2F target genes identified here illustrates this notion.

For example, TGFBR2 and FST are genes involved in TGF-β signaling. The TGF-β superfamily of ligands comprises TGF-β, bone morphogenetic proteins, and activins, which are involved in signaling. The TGF-

REFERENCES

1. Sherr, C. J. (1996) Science 274, 1672–1677.
2. Hanahan, D. and Weinberg, R. A. (2000) Cell 100, 57–70.
3. Morris, E. J., and Ewen, M. J. (1999) Nat. Rev. Cancer 5, 2022.
4. Dyson, N. (1998) Genes Dev. 12, 2345–2362.
5. Muller, M., Bracken, A. P., Vergnaud, G., Morona, M. C., Christians, F., Grasselli, E., Prosperini, E., Vigo, E., Oliner, J. D., and Helin, K. (2001) Genes Dev. 15, 267–285.
6. Helin, K. (1998) Curr. Opin. Genet. Dev. 8, 28–35.
7. Trimparchi, J. M., and Lees, J. A. (2002) Nat. Rev. Mol. Cell Biol. 3, 11–20.
8. Ishida, S., Huang, E., Zuzan, H., Spang, R., Leone, G., West, M., and Nevin, J. S. (2001) Mol. Cell. Biol. 21, 4684–4699.
9. Ben-Efraim, H., Takashashi, Y., Vollet, T., Terragni, J., Young, R. A., and Dyllacht, B. (2002) Genes Dev. 16, 245–256.
10. Polager, S., Kalma, Y., Berkovich, E., and Ginsberg, D. (2002) Oncogene 21, 437–446.
11. Weintmann, A. S., Yan, S. P., Osseroer, M. J., Huang, T. H., and Farnham, P. J. (2002) Genes Dev. 16, 235–244.
12. Marky, M. P., Angus, S. P., Strohecker, M. W., Williams, S. L., Gunawardena, R. W., Arnow, B. J., and Knudsen, E. S. (2002) Cancer Res. 62, 6578–6597.
13. Wells, J., Yan, S. P., Chelly, M., Huang, T., and Farnham, P. J. (2003) Oncogene 22, 1445–1456.
14. Zhu, L., van den Heuvel, S., Helin, K., Fattey, A., Ewen, M., Livingston, D., Dynson, N., and Harlow, E. (1993) Genes Dev. 7, 1111–1125.
15. Lukas, J., Parry, D., Aangaard, L., Mann, D. J., Bartkova, J., Strauss, M., Peters, G., and Bartek, J. (1995) Nature 375, 503–506.
16. Lukas, J., Peteresen, B. O., Holm, K., Bartek, J., and Helin, K. (1996) Mol. Cell. Biol. 16, 1047–1057.
17. Jiang, H., Chou, H. S., and Zhu, L. (1998) Mol. Cell. Biol. 18, 5284–5290.
18. Lukas, J., Sorensen, C. S., Lukas, C., Santoniu-Rugiu, E., and Bartek, J. (1999) Oncogene 18, 3930–3935.
19. Liebhart, D. J., Dong, H., Young, M. C., Follette, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E. L. (1996) Nat. Biotechnol. 14, 1675–1680.
20. Fambrough, D., McClure, K., Kazlauskas, A., and Lander, E. S. (1999) Cell 97, 727–741.
21. Kael, A. E., Kel-Margulis, O. V., Farnham, P. J., Bartley, S. M., Wingender, E., and Zhang, M. Q. (2001) J. Mol. Biol. 309, 99–120.
22. Zheng, Y., Eaton, E. N., Pincet, M., Roberts, J. M., Lundberg, A. S., Gifford, A., Sarstedt, C., and Weinberg, R. A. (1996) Oncogene 12, 1173–1180.
23. Schulze, A., Zerfass, K., Spitzkovsky, D., Middendorf, S., Berges, J., Helin, K., Janssen-Durr, P., and Henglein, B. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 11264–11268.
24. Lam, E. W., and Watson, R. J. (1993) EMBO J. 12, 2705–2713.
25. Schlabach, S., Halperin, T., Vidal, M., and Nevin, J. S. (2002) EMBO J. 21, 5775–5786.
26. Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D. M., and Nakata, Y. (2002) Science 296, 1127–1130.
27. Chen, C. R., Kang, Y., Siegel, P. M., and Massague, J. (2002) Cell 109, 19–32.
28. Kozmiczak, M., Krek, W., and Nagamine, Y. (2000) Mol. Cell. Biol. 20, 2014–2022.

Acknowledgments—We thank Elena Colli and Elena Prosperini for excellent technical assistance. We thank Liang Zhu, Jiri Bartek, and Jiri Lukas for the generous gift of reagents.
Gene Expression Profiling of the pRB Pathway

46137

29. Cho, R. J., Huang, M., Campbell, M. J., Deng, H., Steinmetz, L., Sapinoso, L., Hampton, G., Elledge, S. J., Davis, R. W., and Lockhart, D. J. (2001) Nature Genet. 27, 48–54
30. Zhang, H. S., Postigo, A. A., and Dean, D. C. (1999) Cell 97, 53–61
31. Hughes, J. D., Estep, P. W., Tavazoie, S., and Church, G. M. (2000) Proc. Natl. Acad. Sci. USA 97, 15839–15843
32. Roth, F. P., Hughes, J. D., Estep, P. W., and Church, G. M. (1998) Nat. Biotechnol. 16, 133, 143–150
33. Stanelle, J., Stiewe, T., Theseling, C. C., Peter, M., and Putzer, B. M. (2002) Nucleic Acids Res. 30, 1859–1867
34. Wang, A. K., Shanahan, F., Chen, Y., Lian, L., Ha, P., Hendricks, K., Ghaffari, S., Iyer, D., Penn, B., Woodland, A. M., Smith, R., Salada, G., Carillo, A., Laity, K., Gupte, J., Swedlund, B., Tavtigian, S. V., Teng, D. H., and Lees, E. (2000) Cancer Res. 60, 6171–6177
35. Sherr, C. J., and McCormick, F. (2002) Cancer Cell 2, 103–112
36. Crabtree, G. R., and Olson, E. N. (2002) Mol. Cell. Biol. 22, 7775–7782
37. Bodner, A. D., and Grompe, M. (2003) J. Biol. Chem. 278, 109,499–109,509
38. Basch, R. S., and Goldstein, G. (1975) Cell. Immunol. 10, 231–240
39. Schlesinger, D. H., Goldstein, G., Scheid, M. P., and Boyse, E. A. (1975) Cell. Immunol. 22, 271–286
40. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A., and Thomas, G. (1992) Nature 359, 165–168
41. Copeland, N. G., Cohen, J. E., and Eisenman, R. N. (1995) EMBO J. 14, 5646–5659
42. Basch, R. S., and Goldstein, G. (1975) Cell. Immunol. 10, 218–228
43. Schlesinger, D. H., Goldstein, G., Scheid, M. P., and Boyse, E. A. (1975) Cell 5, 367–370
44. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A., and Thomas, G. (1992) Nature 359, 165–168
45. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
46. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
47. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
48. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
49. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
50. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
51. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
52. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
53. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
54. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
55. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
56. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
57. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
58. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
59. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
60. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
61. May, W. A., Lessnick, S. L., Braun, R. S., Klemasz, M., Lewis, B. C., Lumsford, L. B., Hromas, R., and Denny, C. T. (1993) Mol. Cell. Biol. 13, 7393–7398
Identification of Target Genes of the p16<sup>INK4A</sup>-pRB-E2F Pathway
Richard Vernell, Kristian Helin and Heiko Müller

J. Biol. Chem. 2003, 278:46124-46137.
doi: 10.1074/jbc.M304930200 originally published online August 15, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M304930200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2003/09/10/M304930200.DC1

This article cites 70 references, 21 of which can be accessed free at
http://www.jbc.org/content/278/46/46124.full.html#ref-list-1