

University of Massachusetts Medical School

eScholarship@UMMS

Open Access Articles

Open Access Publications by UMMS Authors

2002-05-31

Gene profiling of cell cycle progression through S-phase reveals sequential expression of genes required for DNA replication and nucleosome assembly

Caroline M. J. van der Meijden
University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.

Follow this and additional works at: <https://escholarship.umassmed.edu/oapubs>



Part of the [Cancer Biology Commons](#), [Genetics and Genomics Commons](#), [Medical Genetics Commons](#), and the [Oncology Commons](#)

Repository Citation

van der Meijden CM, Lapointe DS, Luong MX, Peric-Hupkes D, Cho BC, Stein JL, Van Wijnen AJ, Stein GS. (2002). Gene profiling of cell cycle progression through S-phase reveals sequential expression of genes required for DNA replication and nucleosome assembly. Open Access Articles. Retrieved from <https://escholarship.umassmed.edu/oapubs/360>

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Open Access Articles by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

Gene Profiling of Cell Cycle Progression through S-Phase Reveals Sequential Expression of Genes Required for DNA Replication and Nucleosome Assembly¹

Caroline M. J. van der Meijden, David S. Lapointe, Mai X. Luong, Daniel Peric-Hupkes, Brian Cho, Janet L. Stein, Andre J. van Wijnen, and Gary S. Stein²

Department of Cell Biology and Cancer Center [C. M. J. v. d. M., M. X. L., D. P.-H., B. C., J. L. S., A. J. v. W., G. S. S.] and Information Services [D. S. L.], University of Massachusetts Medical School, Worcester, Massachusetts 01655

ABSTRACT

The ordered expression of genes after growth factor stimulation in G₁ supports the onset of DNA replication. To characterize regulatory events during S-phase when cell cycle progression has become growth factor independent, we have profiled the expression of over 7,000 human genes using GeneChip DNA microarray analysis. HeLa cells were synchronized at the beginning of S-phase by thymidine/aphidicolin block, and RNA populations were analyzed throughout the S and G₂ phases. Expression of genes involved in DNA replication is maximal during early S-phase, whereas histone mRNAs peak at mid S-phase. Genes related to cell proliferation, including those encoding cyclins, oncoproteins, growth factors, proteins involved in signal transduction, and DNA repair proteins, follow distinct temporal patterns of expression that are functionally linked to initiation of DNA replication and progression through S-phase. The timing of expression for many genes in tumor-derived HeLa cells is highly conserved when compared with normal cells. In contrast, a number of genes show growth phenotype-related expression patterns that may directly reflect loss of stringent growth control in tumor cells. Our data reveal there is a core subset of cell growth-related genes that is fundamental to cycling cells irrespective of cell growth phenotype.

INTRODUCTION

Cell proliferation is regulated by the intricate interplay of cell growth regulators, signaling cascades, and mediators of cell cycle progression including cyclins, cdk³, tumor suppressors, and oncoproteins. As cells progress through the cell cycle, functionally distinct classes of proteins are synthesized and degraded at specific transitions. This cell cycle-dependent turnover and the stage-specific requirements for distinct biochemical activities necessitate a continuous regulation of RNA and protein levels. Tumor cells exhibit defects in cell growth control, such as anchorage-independent growth, reduced growth factor requirements, and increased proliferative potential. These changes are at least in part mediated by altered activities of oncoproteins and tumor suppressors. However, mitotic division demands that cells maintain their ordered progression through the cell cycle. Therefore, we anticipate that there is a fundamental set of genes with an expression pattern that is conserved between normal and tumor cells.

Cell cycle control during G₁ phase is mediated in part by interactions between retinoblastoma-related tumor suppressor proteins (p105, p130, and p107; Ref. 1) and the E2 promoter binding factor (E2F) class of transcription factors (E2F1 to 6; Ref. 2), which are

regulated by the actions of distinct cyclins (cyclins D, E, and A) and cdk³ (cdk4/6 and cdk2; Ref. 3–5). Consistent with the important role of E2F in late G₁, many cell growth-regulated genes have E2F binding sites. However, genes for other cell cycle-regulated proteins whose expression is up-regulated immediately after G₁ (e.g., histones H1, H2A, H2B, H3, and H4) do not have E2F sites (6). Thus, there may be multiple gene-regulatory programs involved in the control of E2F-dependent and -independent genes during the G₁ and S phases of the cell cycle.

During S-phase when the genome is replicated, newly synthesized DNA is organized into nucleosomes through interactions with histone proteins. This functional coupling is reflected by the tight linkage between the timing of DNA synthesis and histone gene expression (6, 7). The elevation of nucleotide pools, necessary to support DNA synthesis, and histone-dependent packaging of nascent DNA into nucleosomes are interdependent processes. These processes may very well be sequentially executed, because accumulation of histone mRNA and synthesis of histone proteins are only necessary after the appearance of nascent DNA.

To obtain a genome-wide understanding of the temporal pattern of gene expression in relation to DNA synthesis, we assessed RNA populations in cells synchronized at the G₁-S-phase transition. Two principal findings of our study are that: (a) subsets of genes are successively expressed during the DNA synthesis phase of the cell cycle; (b) there is a core group of cell growth-related genes that is required for cell cycle progression in both normal and tumor cells.

MATERIALS AND METHODS

Cell Culture and Cell Cycle Synchronization. HeLa S3 cells were maintained at a density of 3×10^5 cells/ml in minimal essential medium (Life Technologies, Inc., Rockville, MD), supplemented with 5% calf serum (Atlanta Biologicals, Norcross, GA), 2% horse serum (Nova-Tech, Inc., Grand Island, NE), L-glutamine (Sigma Chemical Co., St. Louis, MO), and penicillin/streptomycin (Sigma). Rapidly dividing HeLa cells were blocked in early S-phase by a 16-h thymidine block (Sigma), followed by a 10-h release, and a subsequent 16-h block with aphidicolin (kindly provided by the Drug Synthesis & Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute). FACS analyses were performed to determine cell cycle phase distribution. The rate of DNA synthesis was determined by [³H]thymidine incorporation as described previously (8).

Isolation of mRNA and Hybridization Analyses. Total cellular RNA was isolated with Trizol (Life Technologies, Inc.) and with RNeasy kits (Qiagen, Valencia, CA) according to the manufacturers' instructions. RNA was quantitated by spectrophotometry (DU640B; Beckman Coulter, Fullerton, CA), and the quality of the RNA was assessed by electrophoresis in a 1% agarose gel. Size-fractionated RNA was transferred to a Hybond-N⁺ nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ) by capillary transfer overnight. Using [³²P]dCTP (NEN) and Rediprime II (Amersham Pharmacia Biotech), DNA fragments encoding histone H4 (H4,3–2 SK+) and GAPDH (phGAPDH) were labeled, purified, and hybridized to the RNA blots at 42°C. For microarray analyses, total RNA (5 μg) was reverse transcribed into double-stranded cDNA (SuperScript Choice system; Life Technologies, Inc.)

Received 1/2/02; accepted 4/2/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ This work was supported by NIH Grant GM32010.

² To whom requests for reprints should be addressed, at Department of Cell Biology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655.

³ The abbreviations used are: cdk, cyclin-dependent kinase; FACS, fluorescence-activated cell sorter; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; DHFR, dihydrofolate reductase.

with the T7-dT24 primer (Genset Corp., La Jolla, CA). After phenol-chloroform extraction, the cDNA was transcribed into biotinylated cRNA (Enzo, Farmingdale, NY), which was purified (RNeasy; Qiagen) and fragmented according to the manufacturer's instructions (Affymetrix Expression Analysis Technical manual rev. 4, Santa Clara, CA). We used test chips and the HU6800 GeneChip probe arrays and followed the standard hybridization procedures with the control oligonucleotide B2 (Genset). The arrays were washed in the Affymetrix GeneChip Fluidics Station 400 and scanned using the Affymetrix GeneChip array scanner.

Bioinformatic Analyses of Gene Expression. The signals from each microarray were analyzed, normalized, and converted to a numerical output using the Affymetrix GeneChip software. The data generated by the different arrays were globally scaled to the target average intensity of 1500. The average expression value for each gene across the arrays was used to normalize the mRNA intensities. Adjusted data were subjected to further analysis using Xcluster.⁴ Expression data from two HeLa cell synchronies were analyzed with Xcluster using log scaling and Pearson's correlation coefficients to assess the similarity of expression profiles during the time course, and data were grouped using the *k*-means (*k* = 10) and self-organizing map options of Xcluster. Data were further analyzed by functional groupings, for which we used both Microsoft Excel and Microsoft Access. The HeLa cell microarray data were compared with data generated from normal human fibroblasts.⁵ Several criteria were used in this comparison: (a) we only included genes that exhibited a relative level of expression above a threshold (average difference value, >500); (b) expression patterns with values that fluctuated <1.5-fold during the time course were considered constitutive; and (c) to be considered similar between the two cell types, the peaks or nadirs of the gene expression pattern should occur within a timeframe of 2 h. For additional information, see our website.

RESULTS AND DISCUSSION

Genome-wide Expression Profiling Reveals Distinct Gene Regulatory Programs during the Cell Cycle. We postulate that global gene-regulatory programs are sequentially expressed during S-phase to support initiation and progression of chromosome replication. In addition, temporally and functionally distinct programs may be operative in the subsequent cell cycle stages to mediate chromosomal segregation and formation of daughter cells. We analyzed genome-wide expression patterns using GeneChip microarrays to assess RNA expression levels during a HeLa cell cycle. We evaluated the microarray data with Affymetrix expression analysis algorithms that monitor the dataset for the presence or absence of transcripts and determine their relative expression levels. To validate the accuracy and biological relevance of our assays by an independent approach, we performed Northern blot analyses with the same RNA samples used for the microarray analyses (Fig. 1). We found that the RNA expression profiles for histone H4, which is a well-established and highly specific S-phase-related parameter, give quantitatively comparable results with both methods. The distribution of cells in the cell cycle was monitored by FACS analyses (Fig. 2A) and reveals that the fraction of cells in S-phase is proportional to histone gene expression. The FACS analyses also demonstrate that timepoints chosen reflect enrichment for cells in early S-phase (1.5 h), mid S-phase (3 h), late S-phase (4.5–6 h), G₂-M phases (6–9 h), and early G₁ phase (9 h).

A general overview of microarray data that reveals sequential expression of groups of genes at different substages of the cell cycle is provided in Fig. 2B. Analysis of the microarray data indicates that ~40% of the genes (2,846) have signal intensities that are significantly above background. On the basis of the temporal pattern of expression during this time course, each gene was assigned to a cluster by using the *k*-means clustering algorithm (*k* = 10; Fig. 2B). The clusters were organized according to their temporal expression pat-

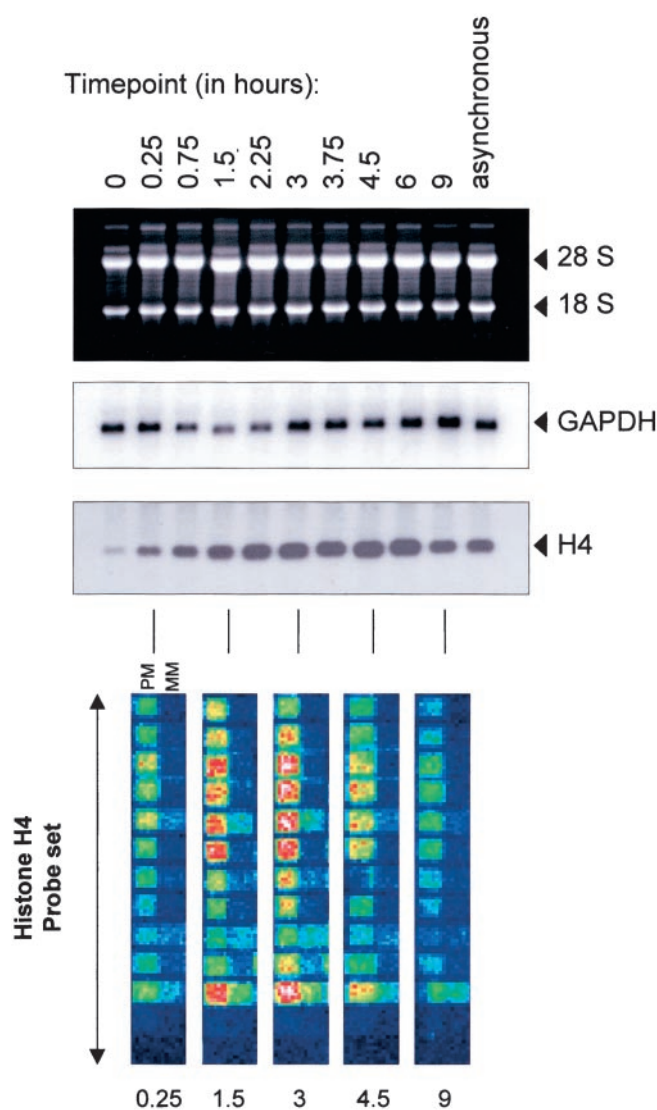


Fig. 1. Northern analyses confirm GeneChip probe array results. A Northern blot of cellular RNAs from synchronized HeLa cells shows *GAPDH* and *histone H4* mRNA levels (top panel). The probe sets representing *histone H4* for the different experiments are depicted (lower panel). Each square represents a cell containing millions of oligonucleotides either a perfect match (PM) to the gene or containing a mismatch (MM) in the sequence. The average of the differences between PM and MM is the value used for further analyses and corresponds to the mRNA level for a particular gene in a given sample. The color scheme used (lower panel) ranges from white, red, yellow, green, blue, and black to indicate the transition from high signal intensities to the absence of signal.

tern, with lower cluster numbers assigned to groups of genes with the highest expression levels in early S-phase and higher cluster numbers assigned to groups of genes that peak in G₂-M to G₁ phase. We found that many clusters are enriched in genes that correspond with specific, cell cycle stage-related functions (representative genes are indicated on the right of Fig. 2B).

The correlation between temporal clusters and functional groupings was further assessed by scoring the frequency of occurrence of distinct groups of genes in each cluster (Table 1). We found several groups of genes involved in, for example, signal transduction, protein synthesis, and protein degradation, that are more prevalent in certain clusters. Inspection of the temporal expression pattern for each cluster suggests that there are substages during S-phase. For example, cluster 2 represents genes that are predominantly expressed in early S-phase, whereas genes in cluster 6 are expressed at highest levels in mid S-phase (Fig. 3). About half of the

⁴ Internet address: <http://genome-www.stanford.edu/~sherlock/cluster.html>.

⁵ Internet address: <http://www.salk.edu/docs/labs/chipdata/>.

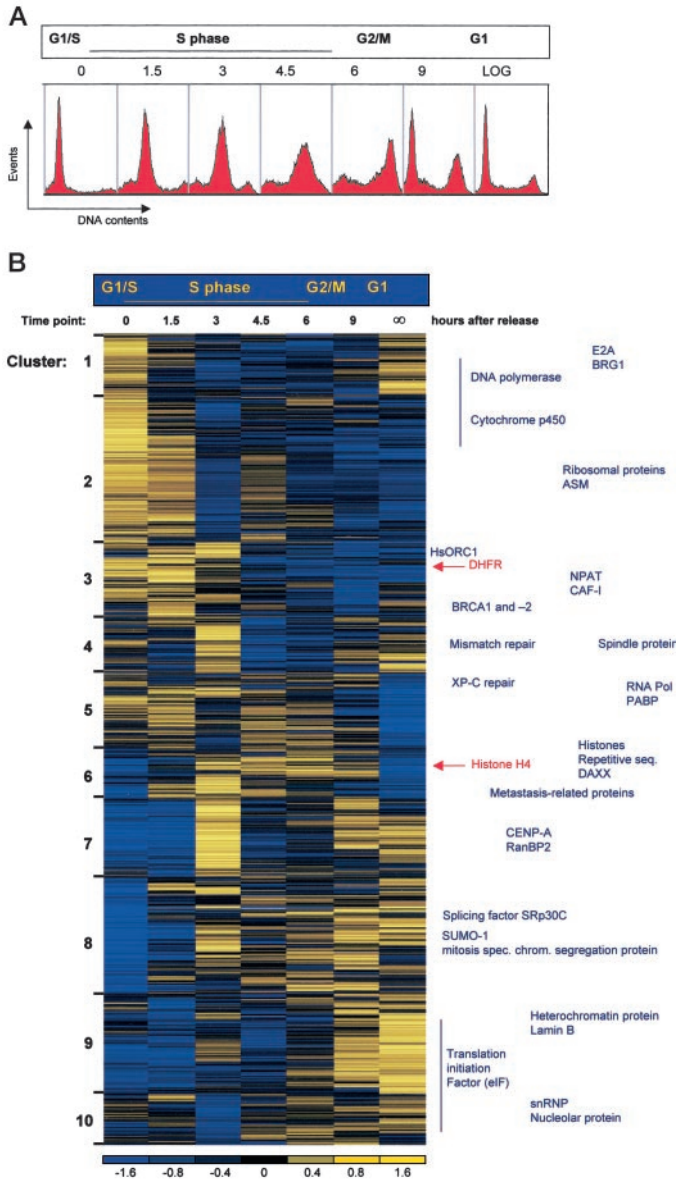


Fig. 2. Temporal pattern of expression reveals differences in timing of peak expression for particular groups of genes during S-phase. *A*, FACS analyses confirm the synchronization of the cells. *B*, shown is a summary bin generated by combining 10 bins (*i.e.*, cluster) obtained through *k*-means clustering. Every horizontal line represents a gene, and columns are representatives of different timepoints taken, with [∞] representing samples generated from log-phase cells. Numbers on the left indicate different clusters. Shown on the right are particular genes found in different clusters. The yellow color indicates a high level of expression, black represents no change, and blue indicates the lowest expression as shown on the bottom. The transition of yellow from the left top to the right bottom of the bin reflects succession in maximal expression levels.

expressed genes exhibit temporally regulated expression patterns, in contrast with genes encoding proteins such as cytoplasmic β -actin and most proteasomal proteins, which show barely any fluctuation. Overall, a succession of gene expression occurs over time, with particular genes or groups of genes appearing at particular stages of the cell cycle phases.

If the groups with particular peaks in expression represent distinct gene-regulatory programs required for either initiation or progression of chromosome replication, we would expect to find a difference in the timing of expression for the genes *DHFR* and *histone H4*. Indeed, we found that *DHFR* and *histone H4* are present in different clusters (Fig. 2*B*), with *DHFR* expression increasing before the up-regulation of *histone H4*. This finding suggests that there are distinct substages

in S-phase to accommodate biological requirements for the duplication of the genome.

Induction of Genes Involved in DNA Replication and Early Response Genes in Early S-Phase. To further elucidate temporal transitions during S-phase, we examined the genes that are expressed at the G₁-S boundary and the subsequent entry into S-phase. We addressed whether genes whose expression is up-regulated at early S-phase are involved in DNA replication and/or in reinitiating proliferation. We found an increase in transcripts for the proteins MCM6, ORC1, ORC2, and the kinases *cdc6* and *cdc7* (Fig. 2*B*, clusters 1 and 3; Fig. 3). These proteins all participate in the initial stages of DNA replication, which encompass recognition of the origin of replication and assembly of an initiation complex (10). The subsequent initiation of DNA synthesis is mediated by DNA primase and DNA polymerase α , followed by DNA polymerase δ (11). The corresponding gene expression patterns for these DNA replication enzymes (shown in Fig. 4*A*) are comparable with those for genes involved in origin recognition. Enzymes involved in generating the cellular supply of nucleotides are also expressed in early S-phase (*e.g.*, *DHFR* in Fig. 2*B*). Taken together, many enzymes and proteins involved in DNA synthesis are present in early S-phase.

We found a dramatic increase in the presence of transcripts involved in cell proliferation (see proliferation category in Table 1) in the group of genes that are maximally expressed at 0–1.5 h after release (*e.g.*, cluster 2) as well as in gene groups that are expressed at high levels in subsequent stages. We assessed which specific genes exhibit high expression in early S-phase. Genes present at release or up-regulated immediately thereafter included a number of early response genes (*ATF-3*, *SRF*, *ERF-2*, *IEX-1*, and *HSP-70*), transcriptional regulatory factors (SWI/SNF, ATRX, and others; Ref. 12), cytokines (MDNCF and BSF-2), nuclear orphan receptors (LXR- α , TR3, Ear-3, and MINOR), and kinases (*e.g.*, *c-fes* and *cdk10*; Fig. 3). Other genes that are induced in early S-phase include those encoding cell surface proteins, such as cytokine receptors (*e.g.*, *frizzled*) and transport-related proteins (*e.g.*, *rab13* and *TWIK*; Table 1). We also note that there is a moderate increase in metabolic proteins [*e.g.*, cyclooxygenase (*Cox-2*), and acid sphingomyelinase] when the cell cycle proceeds (Fig. 2*B*). Therefore, in this early stage of S-phase, expression of members from several signal transduction pathways as well as expression of metabolic enzymes is prominent.

Several mRNAs were present at high levels at the G₁-S boundary but exhibited a transient decrease in expression during S-phase. Among these were envoplakin (a cytoskeletal linker protein), the growth arrest-specific gene (*gas1*), the transcription coregulators *Smad1* and *Smad6*, and *BARD1* (Fig. 3). The temporal expression of these genes may be G₁ or G₁-S transition specific or correlated with the synchronization procedure. Thus, there are multiple gene-regulatory programs that are initiated in early S-phase, and only a subset directly supports the onset and progression of DNA replication.

Maximal Expression of Genes Involved in DNA Repair and Histone Genes Is Linked to Maximal DNA Synthesis. At mid S-phase, we found a second category of genes that are maximally expressed when the DNA synthesis rate is highest, which defines this substage of S-phase. DNA replication may be coupled to mechanisms that monitor the integrity of nascent DNA. Consistent with this concept, we found the highest levels of transcripts for proteins involved in DNA repair (*e.g.*, mismatch repair gene) when cells progress toward mid S-phase (Fig. 2*B*, clusters 3, 4, and 5). The coordinate up-regulation of histone RNA levels at 4.5–6 h into S-phase (Fig. 2*B*, cluster 6) accommodates the high demand for core histones (*i.e.*, H2A, H2B, H3, and H4; Refs. 7 and 9), which are involved in the formation of chromatin as DNA is synthesized. Preceding the up-regulation of these two classes of genes, we find high expression for

Table 1 Correlation between temporal classes and functional groupings

Upon obtaining the different clusters, the data are organized by function. Occurrence of genes from the different functional groups is counted for each cluster. When comparing these values with the total number of genes in each cluster, several functional groups of genes appear to be more prevalent in certain clusters (e.g., signal transduction and nuclear components).

Group	Cluster									
	1	2	3	4	5	6	7	8	9	10
Genes in the cluster	210	489	224	195	225	202	239	361	295	170
Cell cycle	4	6	4	1	3	1	3	4	4	2
Chromatin	1	2	2		3	8	4	3		
Mitotic		11	1	6	2		1	2	1	1
Structural	1	18	3	2	8	1	10	6	2	5
Nuclear	26	42	36	23	28	21	28	64	58	25
DNA replication	2	2	1	2	1		1	2		
Transport	1	7	5	5	3	2	1	2	2	1
DNA repair	1		2	3	2			1	1	1
RNA/protein synthesis	2	49	2	0	8	0	17	9	4	3
Signal transduction	32	59	31	19	21	16	37	27	17	16
Kinases	13	36	14	6	11	6	21	13	10	10
Phosphatases	4	9	4	5	3	4	10	0	1	2
Receptors	10	18	11	8	6	5	5	12	2	2
Proliferation	1	18	14	10	15	12	14	15	13	11
Degradation	4	11	2	1	3	5	10	8	5	3

several gene regulators involved in chromatin assembly [such as nuclear protein, ataxia-telangiectasia locus (*NPAT*) and the chromatin assembly factor p150 (*CAF*); Fig. 2B, cluster 3]. Thus, when cells exhibit peak rates of DNA synthesis, we found maximal expression of

transcripts for proteins that monitor and restore newly replicated DNA and chromatin.

E2F-dependent versus E2F-independent Gene Regulation. Our studies indicate that there are temporal differences in RNA expression

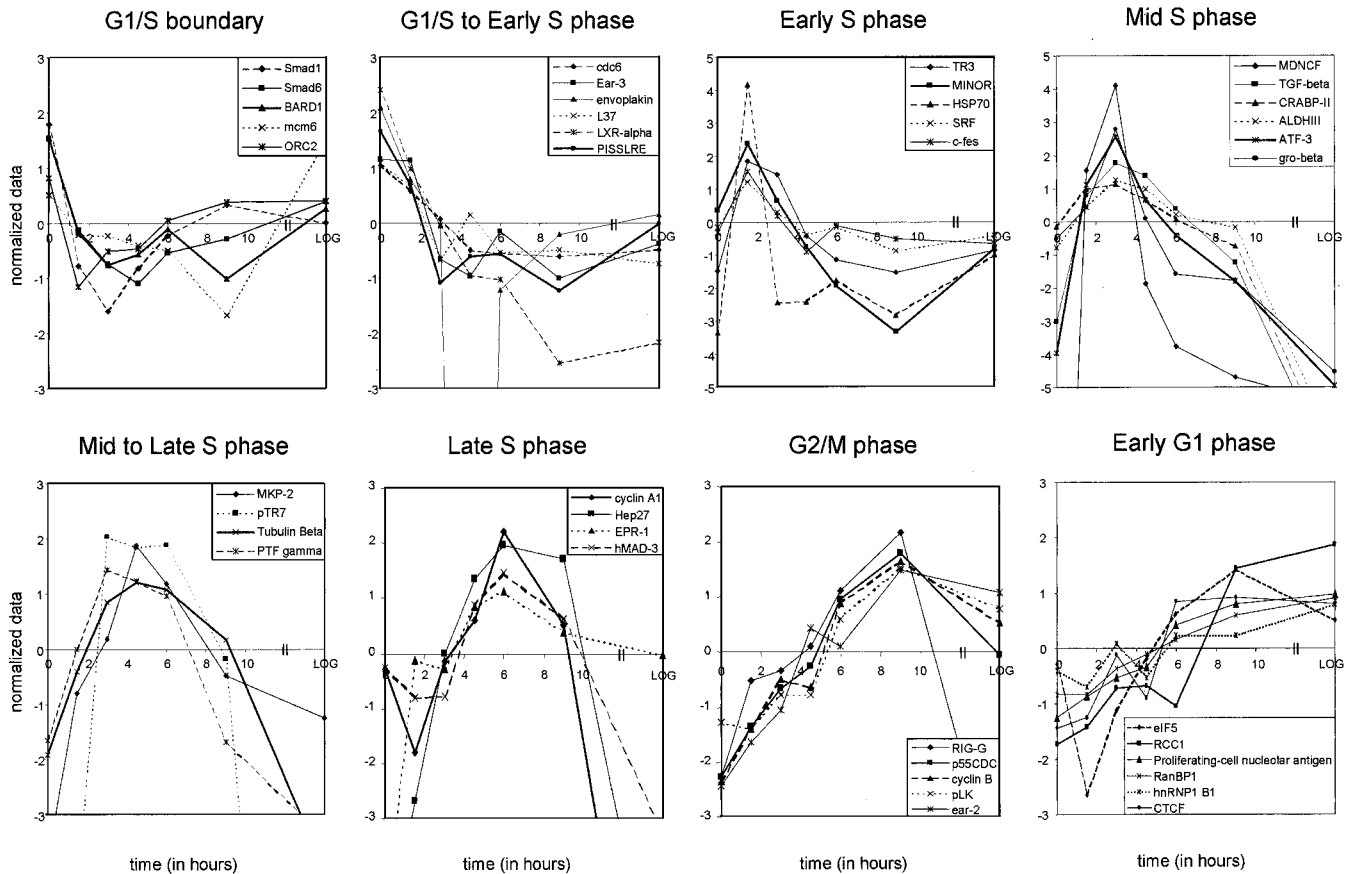


Fig. 3. Temporal expression patterns reveal substages in S-phase. Particular groups of genes appear to be highly expressed at particular times throughout the synchrony. Examples of such genes are given in these plots. *G1/S boundary*, Smad 1 (Sma and Mad), Smad 6, BARD1 (BRCA1-associated RING domain protein), mcm6 (mini chromosome maintenance protein), and ORC2 (origin recognition complex). *G1/S to Early S phase*, cdc6 (cell division cycle protein), Ear-3 (v-erbA-related protein), envoplakin, ribosomal protein L37, LXR- α , and PISSLRE [cyclin-dependent kinase (CDC2-like) 10 (CDK10)]. *Early S phase*, TR3, MINOR (mitogen-inducible nuclear orphan receptor), HSP70 (heat shock protein 70), SRF (serum response factor), and c-fes. *Mid S phase*, MDNCF (monocyte-derived neutrophil chemotactic factor), TGF- β (transforming growth factor- β), CRABP-II (cellular retinoic acid-binding protein type II), ALDHIII (aldehyde dehydrogenase), ATF-3 (activating transcription factor 3), and gro- β (growth regulated protein β). *Mid to Late S phase*, MKP-2 (mitogen-activated protein kinase phosphatase 2), pTR7 (repetitive sequence), tubulin- β , and PTF γ [(PSE)-binding transcription factor]. *Late S phase*, cyclin A1, Hep27, EPR-1 (effector cell protease receptor-1), and Smad3. *G2/M phase*, RIG-G (retinoic acid-induced gene G), p55CDC (cdc20), cyclin B, pLK (polo-like kinase), and Ear-2. *Early G1 phase*, eIF5 (eukaryotic translation initiation factor 5), RCC1 (regulator of chromosome condensation 1), proliferating cell nuclear antigen, RanBP1 (Ran-binding protein 1), hnRNP1 B1 (heterogeneous nuclear ribonucleoprotein), and CTCF (CCTC-binding factor).

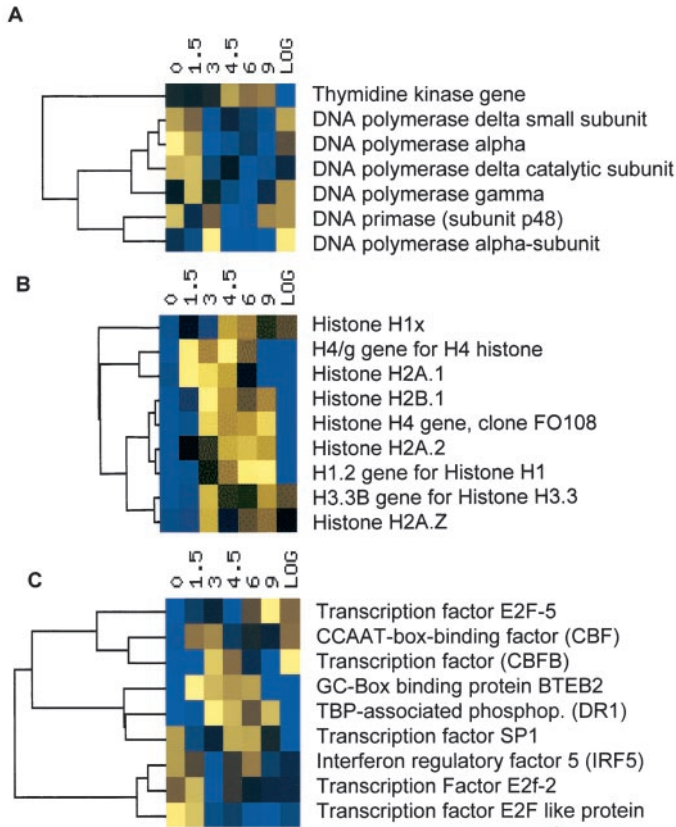


Fig. 4. Differential expression of E2F-dependent and -independent genes at the G₁-S transition. To further distinguish between expression patterns for genes involved in DNA replication and chromatin assembly, genes of these categories were grouped together, with one representative bin for genes encoding proteins involved in DNA synthesis (E2F dependent; A) and one bin for histones (E2F independent; B). Dendrograms indicate relatedness in expression patterns. C, fluctuations of transcripts for several transcription factors, which may be involved in the E2F-dependent or -independent expression control, based on the characteristic promoter elements. The color scheme used is shown in Fig. 2B.

of proteins involved in DNA synthesis *versus* transcripts encoding histone proteins. The first group of transcripts is present at relatively high levels at the G₁-S boundary, whereas the second group does not peak until at least 1.5 hours into S-phase (Figs. 4, A and B). This difference in timing is at least in part regulated through transcriptional mechanisms, as reflected by particular promoter characteristics for the different groups. The predominant regulatory elements that govern G₁-S-phase-specific expression and corresponding transcription factors have been identified previously. Of these recognition sequences, the E2F site is one element critical for the passage through the R point and G₁-S transition, and many genes involved in DNA replication are E2F responsive (2, 13, 14). Promoters of these genes have at least one E2F site as well as one or more auxiliary elements (*e.g.*, Sp1 sites) but in general lack TATA and CCAAT boxes. It is remarkable that histone gene promoters almost without exception lack E2F sites. Instead, promoters of the histone gene family are controlled by TATA and CCAAT boxes, Sp1 sites, and histone subtype-specific elements (*e.g.*, binding sites for IRF, Oct1, CDP, and HiNF P; Refs. 6 and 9).

Distinctions in promoter organization and gene expression profiles of DNA replicative genes predict that the cognate transcription factors may also have differences in expression. The gene expression patterns of transcription factors that are either known or presumed to interact with G₁-S-phase-related promoter elements (E2F and Sp1 sites, CCAAT and TATA boxes, and others) were clustered by Pearson's correlation and analyzed with *k*-means (Fig. 4C). The factors we analyzed appear to separate into three temporally distinct groups.

Three of these genes (*i.e.*, E2F-2, an E2F-like protein, and IRF-5) have expression patterns resembling those of several DNA polymerases (Fig. 4, A and C). A second group of factors [*i.e.*, Sp1, basic transcription element binding protein (BTEB2), and down-regulator of transcription 1 (DR1); Refs. 15 and 16] is regulated in parallel with the histone gene subtypes. Hence, our data suggest that the expression of at least a subset of transcription factors present in S-phase correlates with the expression of potential downstream target genes.

Transcripts Encoding Protein-modifying Factors and Nuclear Components Are Temporally Linked with Mitosis. After completion of S-phase, cells prepare for mitotic division, which requires condensation, separation, and redistribution of the chromosomes, as well as major turnover and/or posttranslational modification of proteins. The mitotic phase is preceded by the G₂ checkpoint, which is essential to monitor correct duplication of genetic information, accurate structural formation, and competency for division and completion of the cell cycle (17). We found maximal expression for several transcriptional repressors as cells progressed from mid S-phase to early G₁ (*e.g.*, DR1, C/EBP γ , TGIF, and CTCF; Figs. 3 and 4C), which may be related to the dramatic decrease in RNA synthesis that occurs during mitosis (18). We found that transcripts encoding proteins involved in ubiquitination and degradation are present at later stages of the cell cycle (Table 1, cluster 7). For example, maximal transcript levels for several genes encoding proteins involved in protein degradation, such as small SUMO and the cullin proteins Hs-cul-1 and Hs-cul-4A, were found at 4.5–6 h after release (Fig. 2B). The increased expression of these protein-inactivating factors indicates that cells have completed S-phase and are preparing for cell division. Similarly, other enzymes mediating protein modification are predominantly present in the same temporal cluster (Table 1, cluster 7; *e.g.*, phosphatases and kinases). RNA encoding proteins known to be involved in mitotic regulation or mitosis are successively expressed with maximal expression at mid S-phase [*e.g.*, kinesin-like spindle protein (HKSP) and β -tubulin], late S-phase (*e.g.*, cyclin F and PIN1), or the G₂-M phases (*e.g.*, p55cdc and cyclin B; Figs. 2B and 3). We conclude that several division-specific proteins show an increase in RNA expression during the mitotic period.

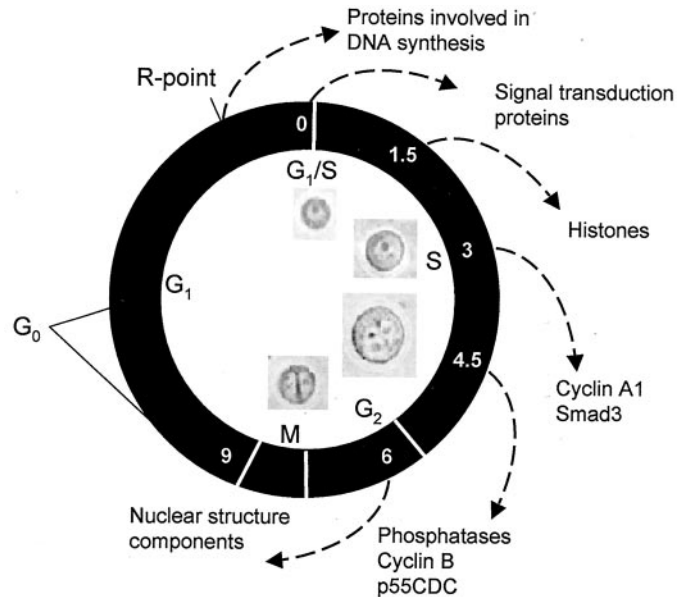


Fig. 5. Summary of gene expression profiles during the cell cycle. Shown is a representation of the cell cycle, with the light micrographs indicating different cell shapes and sizes that are prominent at different stages of the cell cycle. The appearance of RNAs for particular groups of proteins is indicated; maximal expression generally occurs at the adjacent timepoints.

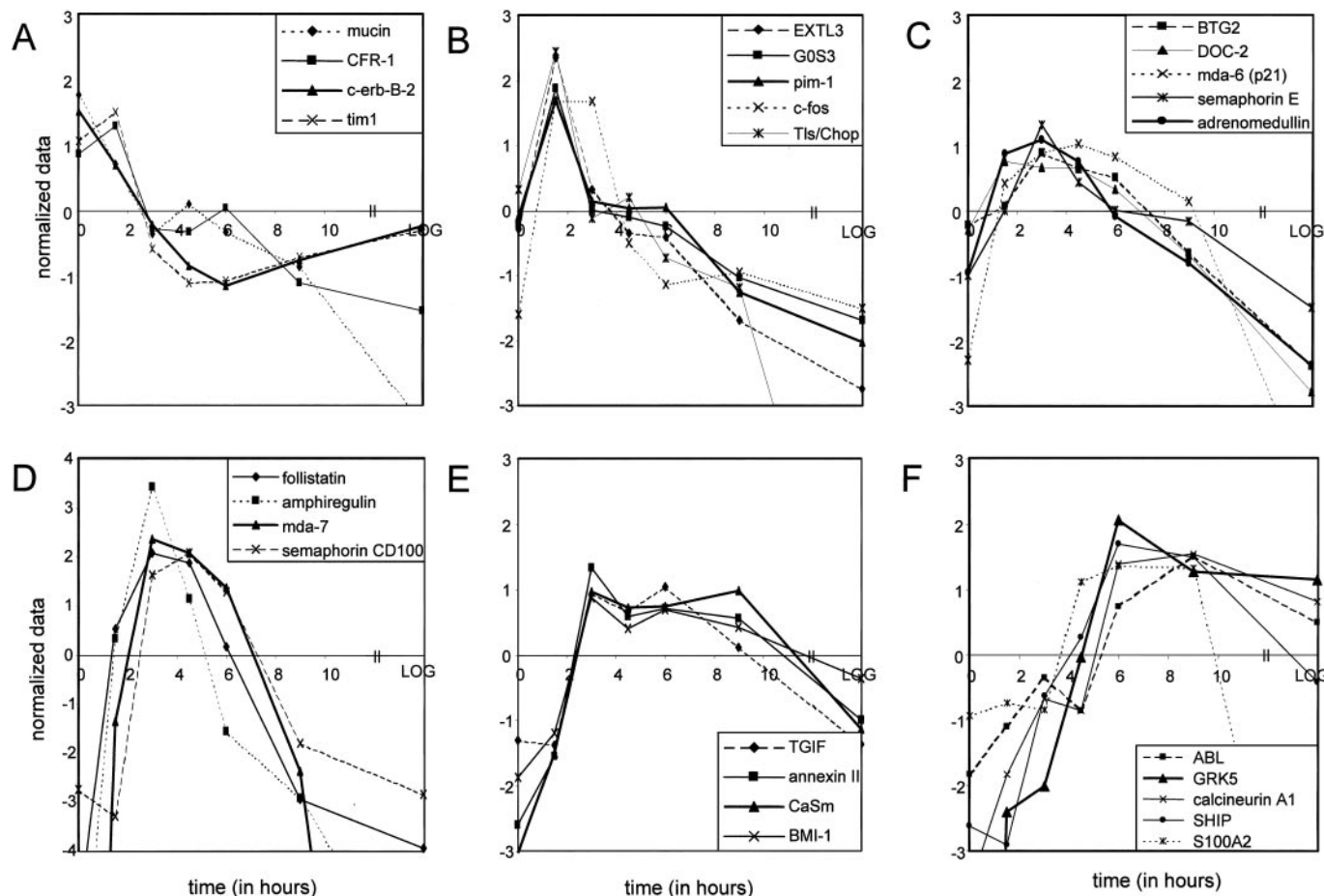


Fig. 6. Expression levels for mRNAs of cancer-related proteins show distinct functional and temporal patterns. The proteins shown are selected by their prominent peaks in expression at particular times during the cell cycle as well as their reported correlation to tumor growth. A, pancreatic mucin, CFR-1 (cystein-rich fibroblast growth factor receptor), c-erb-B-2, tim1 (21, 22, 24). B, EXTL3 (EXT like protein 3), G0S3 (G_0 - G_1 switch gene 3), pim-1, c-fos, and Tls/Chop (translocation liposarcoma/C/EBP homologous protein; Refs. 23, 25, and 32). C, BTG2 (B-cell translocation gene proteins), DOC-2 (differentially expressed in ovarian carcinoma-2), mda-6 (melanoma differentiation associated protein 6 = p21), semaphorin E, and adrenomedullin (29, 31, 33, 38). D, follistatin, amphiregulin, mda-7 (melanoma differentiation associated protein 7), and semaphorin CD100 (35–37). E, TGIF (5[′]TG3[′] interacting factor), annexin II, CaSm (cancer-associated Sm-like oncoprotein), and BMI-1 (27, 28, 30). F, Abl, GRK5 (G protein-coupled receptor kinase-5), calcineurin A1, SHIP (SH2-containing inositol phosphatase), and S100A2 (39–41). The graphs show maximal RNA expression of signal transduction molecules in early S-phase, proteins involved in cell growth and tumor malignancy from mid to late S-phase, and potential signal transduction attenuators during G_2 -M.

Table 2 Comparison of gene expression in two cell types by functional groupings.

Group	Expressed in		
	Both cell types	Primary fibroblast	HeLa tumor cell
Number of genes	1837	1091	987
Cell cycle	19	2	6
Chromatin	12	0	13
Mitotic	12	0	7
Structural	123	43	31
Nuclear	246	49	137
DNA replication	10	1	2
Transport	29	20	18
DNA repair	12	2	1
RNA/protein synthesis	68	1	7
Signal transduction	180	96	116
Kinases	87	24	51
Phosphatases	26	8	14
Receptors	67	64	51
Proliferation	94	28	40
Degradation	22	1	5

The data obtained from the two microarray assays are organized by function for cell type-specific expression as well as conserved expression. Occurrence of genes from the different functional groups is scored. Genes from groups like Receptors and Transport show little similarities between the two cell types. In contrast, the groups Cell cycle, DNA replication, DNA repair, RNA/protein synthesis and Degradation contain mostly genes that are expressed in both the tumor-related HeLa cells and the normal fibroblasts.

The redistribution of chromatin and synthesis of the nuclear envelope in the daughter cells mark the final stage of mitosis and beginning of G_1 (20). During these phases, we observed a striking rise in nuclear proteins (Table 1, clusters 8 and 9). Representative genes expressed during this transition encoded proteins involved in nuclear trafficking (*e.g.*, RCC1, RanBP1 and 2, importin β , and nucleopore protein), nuclear structure (*e.g.*, the nuclear matrix protein matrin 3 and lamins A and B), and nucleolar organization (*e.g.*, the proliferating cell nucleolar antigen; Fig. 3). We found a significant increase in the majority of RNA-processing factors and translation initiation factors in early G_1 . For example, we found that most heterogeneous nuclear ribonucleoproteins, small nuclear ribonucleoprotein particles, the nuclear cap binding protein, and translation initiation factors show peak expression at 9 h (Fig. 2B). In summary, our gene profiling studies have revealed the sequential expression during S-phase of RNA encoding proteins such as growth signal-responsive proteins, proteins involved in DNA synthesis, and proteins involved in chromatin assembly (*e.g.*, histones), followed by maximal expression in the later cell cycle phases of proteins involved in cell division and nuclear structure (Fig. 5).

Table 3 Comparison of microarray data from HeLa tumor-related cells with cell cycle data from normal fibroblasts reveals a highly conserved subset of genes

ID	Name	Description
A. Genes with similar cell cycle-related expression patterns in normal fibroblasts and tumor-related HeLa cells		
Cell cycle		
X54942	<i>CKS1</i>	CDC28 protein kinase 2
X05360	<i>CDC2</i>	Cell division cycle 2
U05340	<i>CDC20</i>	Cell division cycle 20, <i>Saccharomyces cerevisiae</i> homologue
L22005	<i>CDC34</i>	Cell division cycle 34
U14518	<i>CENPA</i>	Centromere protein A (17kD)
X51688	<i>CCNA2</i>	Cyclin A
M25753	<i>CCNB1</i>	Cyclin B1
Z36714	<i>CCNF</i>	Cyclin F
X80230	<i>CDK9</i>	Cyclin-dependent kinase 9 (CDC2-related kinase)
U40343	<i>CDKN2D</i>	Cyclin-dependent kinase inhibitor 2D (p19, inhibits CDK4)
M13194	<i>ERCC1</i>	Excision-repair cross-complementing rodent repair deficiency
D64142	<i>H1FX</i>	H1 histone family, member X
X14850	<i>H2AFX</i>	H2A histone family, member X
X60486	<i>H4FG</i>	H4 histone family, member G
J05614	<i>PCNA</i>	Human proliferating cell nuclear antigen (<i>PCNA</i>) gene
D26361	<i>KIAA0042</i>	Kinesin superfamily member DUnc 104
U37426	<i>KNSL1</i>	Kinesin-like 1
X67155	<i>KNSL5</i>	Kinesin-like 5 (mitotic kinesin-like protein 1)
D84557	<i>MCM6</i>	Minichromosome maintenance deficient (mis 5, <i>Schizosaccharomyces pombe</i>) 6
D21063	<i>MCM2</i>	Minichromosome maintenance deficient (<i>Saccharomyces cerevisiae</i>) 2 (mitotin)
D38073	<i>MCM3</i>	Minichromosome maintenance deficient (<i>Saccharomyces cerevisiae</i>) 3
U01038	<i>PLK</i>	Polo (<i>Drosophila</i>)-like kinase
D38551	<i>RAD21</i>	RAD21 (<i>S. pombe</i>) homologue
M87339	<i>RFC4</i>	Replication factor C (activator 1) 4 (37kD)
M91670	<i>E2-EPF</i>	Ubiquitin carrier protein
U73379	<i>UBCH10</i>	Ubiquitin carrier protein E2-C
U58090	<i>Hs-cul-4A</i>	
Signaling		
X89985	<i>BCL7B</i>	B-cell CLL/lymphoma 7B
U72066	<i>CHIP</i>	CtBP interacting protein CHIP
M60278	<i>DTR</i>	Diphtheria toxin receptor
U15932	<i>DUSP5</i>	Dual specificity phosphatase 5
L38487	<i>ESRRA</i>	Estrogen-related receptor α
U18018	<i>ETV4</i>	ets variant gene 4 (E1A enhancer-binding protein, E1AF)
U80947		Glucocorticoid receptor
U75679	<i>HBP</i>	Hairpin binding protein, histone
S81914	<i>IER3</i>	Immediate-early response 3
X69111	<i>ID3</i>	Inhibitor of DNA binding 3, dominant-negative helix-loop-helix protein
X58377	<i>IL11</i>	Interleukin 11
U56816	<i>PKMYT1</i>	Tyrosine- and threonine-specific cdc2-inhibitory kinase
X60188	<i>MAPK3</i>	Mitogen-activated protein kinase 3
X80692	<i>MAPK6</i>	Mitogen-activated protein kinase 6
L08246	<i>MCL1</i>	Myeloid cell leukemia sequence 1 (BCL2-related)
D87953	<i>NDRG1</i>	<i>N-myc</i> downstream regulated
U90426	<i>DDXL</i>	Nuclear RNA helicase, DECD variant of DEAD box family
U33053	<i>PRKCL1</i>	Protein kinase C-like 1
U47077	<i>PRKDC</i>	Protein kinase, DNA-activated, catalytic polypeptide
X89750	<i>TGIF</i>	TG-interacting factor (TALE family homeobox)
M86699	<i>TTK</i>	TTK protein kinase
U20647	<i>ZNF151</i>	Zinc finger protein 151 (pHZ-67)
Miscellaneous		
M91029	<i>AMPD2</i>	Adenosine monophosphate deaminase 2 (isoform L)
X76228	<i>ATP6E</i>	ATPase, H ⁺ transporting, lysosomal (vacuolar proton pump)
M94345	<i>CAPG</i>	Capping protein (actin filament), gelsolin-like
U81556	<i>OS4</i>	Conserved gene amplified in osteosarcoma
U55206	<i>GGH</i>	γ -Glutamyl hydrolase
X62691		Glycine cleavage system T-protein
L37368		<i>Homo sapiens</i> mRNA for ribosomal protein
X12458		Human (clone E5.1) RNA-binding protein mRNA
X76771	<i>PMP22</i>	Human peripheral myelin protein-22 (<i>PMP22</i>) gene
L42572	<i>IMMT</i>	Inner membrane protein, mitochondrial (mitofilin)
U28386	<i>KPNA2</i>	Karyopherin α 2 (RAG cohort 1, importin α 1)
M61916	<i>LAMB1</i>	Laminin, β 1
X82895	<i>MPP2</i>	Membrane protein, palmitoylated 2
J04031	<i>MTHFD1</i>	Methylenetetrahydrofolate dehydrogenase
X07315	<i>PP15</i>	Nuclear transport factor 2 (placental protein 15)
U59752	<i>PSCD2</i>	Pleckstrin homology, Sec 7 and coiled/coil domains 2
U53204	<i>PLEC1</i>	Plectin 1, intermediate filament binding protein, 500kD
X78136	<i>PCBP2</i>	Poly(rC)-binding protein 2
U04636	<i>PTGS2</i>	Prostaglandin-endoperoxide synthase 2
X55954	<i>RPL23</i>	Ribosomal protein L23
L19527	<i>RPL27</i>	Ribosomal protein L27
U46751	<i>SOSTM1</i>	Sequestosome 1
X99585	<i>SMT3H2</i>	SMT3 (suppressor of mif two 3, yeast) homologue 2
U81375	<i>SLC29A1</i>	Solute carrier family 29 (nucleoside transporters)
J04611	<i>G22P1</i>	Thyroid autoantigen 70kD (Ku antigen)
M90657	<i>TMF4SF1</i>	Transmembrane 4 superfamily member 1
U04810	<i>TROAP</i>	Trophinin-associated protein (tastin)
M19714		Tropomyosin, α
X89398	<i>UNG</i>	Uracil-DNA glycosylase
D13633	<i>KIAA0008</i>	
D31885	<i>KIAA0069</i>	
D43948	<i>KIAA0097</i>	

^a The results presented in this study with tumor cells are compared with those generated by a synchrony of normal diploid cells (19).⁵ Genes that are expressed in both studies are listed. Grouped together are: (A) genes with similar cell cycle-related expression patterns (conserved), (B) genes that are expressed in both cell lines but with altered expression patterns (different, constitutive, or not determined), or (C) genes that are expressed in the normal fibroblast yet are not detected the HeLa cancer-related cell line (absent).

Table 3 *Continued*

ID	Name	Description
B. Genes expressed in normal fibroblasts and HeLa cells but with different cell cycle-related expression patterns ^a		
Different		
D16562	<i>CAST</i>	ATP synthase
X54941	<i>CKS1</i>	CDC28 protein kinase 1
X55740		cDNA coding for 5 nucleotidase (EC 3.1.3.5)
X96753	<i>CSPG4</i>	Chondroitin sulfate proteoglycan 4
U09579	<i>p21/Cip1</i>	Cyclin-dependent kinase inhibitor 1A (p21)
X06256		Fibronectin receptor α subunit
M23379	<i>RASA1</i>	GTPase-activating protein ras p21 (RASA)
L28010	<i>HnRNP F</i>	Heterogeneous nuclear ribonucleoprotein H
U47105		Human H105e3 mRNA
X81788	<i>ICT1</i>	Immature colon carcinoma transcript 1
M22919		<i>MLC</i> gene (non-muscle myosin light chain)
X92720	<i>PCK2</i>	Phosphoenolpyruvate carboxykinase
Y00971	<i>PRPS2</i>	Phosphoribosyl pyrophosphate synthetase 2
D84110	<i>RBP-MS/ty</i>	RNA-binding protein gene with multiple splicing
J03161	<i>SRF</i>	Serum response factor
J05243	<i>SPTAN1</i>	Spectrin, α
Constitutive		
D90209	<i>ATF4</i>	Activating transcription factor 4
AF006084	<i>ARPC1B</i>	Arp2/3 protein complex subunit p41-Arc (ARC41)
D16217	<i>CAST</i>	Calpastatin
M55268	<i>CSNK2A2</i>	Casein kinase 2
X52142	<i>CTP A</i>	CTP synthase
X66364	<i>CDK5</i>	Cyclin-dependent kinase 5
U53446	<i>DAB2</i>	Disabled homologue 2 (mitogen responsive)
U10362	<i>GP36B</i>	Endoplasmic reticulum glycoprotein
M62994	<i>FLNB</i>	Filamin B
U74612	<i>FOXM1</i>	Forkhead boxM1
L37033	<i>FKHR</i>	Forkhead domain protein
U46023	<i>Xq28</i>	Gene mutated in X-linked myotubular myopathy
U10860		Guanosine 5'-monophosphate synthase
M94630	<i>HNRPD</i>	Heterogeneous nuclear ribonucleoprotein D
X53586		Integrin α 6
X51345	<i>JUNB</i>	<i>junB</i> proto-oncogene
D23673	<i>KIAA0864</i>	KIAA0864
U63743	<i>KNSL6</i>	Kinesin-like 6
Y00764		Mitochondrial hinge protein
Z24725	<i>MIG2</i>	Mitogen inducible 2
Z97054	<i>SMC1</i>	Mitosis-specific chromosome segregation protein
M69066	<i>MSN</i>	Moesin
U02493	<i>NONO</i>	Non-Pou domain-containing octamer binding protein
X12492	<i>NFIC</i>	Nuclear factor I/C (CCAAT-binding TF)
HG3415-HT3598 at		Poliovirus receptor
L76517	<i>PSEN1</i>	Presenilin 1
M77836	<i>PYCR1</i>	Pyroline-5-carboxylase, β polypeptide
L19527	<i>RPL27</i>	Ribosomal protein L27
D38555	<i>SEC24C</i>	SEC24C
U18543		Zinc-finger protein mRNA
Not determined		
M19645		78-kDa glucose-regulated protein (GRP78)
X53416	<i>filamin</i>	Actin-binding protein
U33286	<i>CAS</i>	Chromosome segregation gene homologue
L35240		Enigma gene
K03195	<i>HepG2</i>	Glucose transporter gene
L12723	<i>hsp70</i>	Heat shock protein 70
D50405	<i>RPD3L1</i>	Homologous to RPD3, a transcription factor in <i>Saccharomyces cerevisiae</i>
D79997	<i>KIAA0175</i>	KIAA0175
U14391		Myosin-1C
M95178		Non-muscle α -actinin
X79439		Notch 3
U24152	<i>PAK1</i>	p21/Cdc42/Rac1-activated kinase 1
X70476		Subunit of coatomer complex

Expression Patterns for Genes with Cancer-related Properties.

Our microarray analyses have established that there are substages in S-phase with distinct gene expression programs. Because HeLa cells are derived from a cervical carcinoma, we evaluated the presence of cancer-related and cell growth-regulatory genes in our dataset and analyzed their expression patterns in relation to S-phase substages (Fig. 6). In early S-phase, we found high expression of cytokine receptors involved in tumor proliferation (21, 22), as well as a number of oncoproteins (23–26). These expression patterns parallel the increased expression of transcripts encoding other cell signaling-related proteins that are maximally expressed in early S-phase (see Fig. 3). At mid to late S-phase, we observed peak expression for several genes encoding proteins with mitogenic properties that are involved in metastasis (27), maintenance of the transformed phenotype (28), or

induction of proliferation (29, 30). We also found increased expression of genes encoding several tumor suppressor proteins or proteins with antiproliferative properties. Thus, although many proteins with oncogenic properties are expressed from mid to late S-phase, a similar peak of expression occurs for numerous growth inhibitors and tumor suppressors (31–38). From late S-phase to G₂-M, we found enhanced expression of several genes encoding proteins involved in calcium and inositol signaling. Moreover, some of these proteins have also been associated with negative effects on cell growth and/or involvement in apoptosis (39–41). Hence, we have characterized expression of signal transduction molecules in early S-phase, proteins involved in cell growth and tumor-related proteins from mid to late S-phase, and potential signal transduction attenuators during G₂-M.

To assess whether the cancer-related genes described above play a

Table 3 *Continued*

ID	Name	Description
C. Genes that are expressed in normal fibroblasts but not detected in HeLa tumor-related cells		
Cell cycle		
L47276		α -Topoisomerase truncated form
X59798	<i>CCND1</i>	Cyclin D1 (PRAD1; parathyroid adenomatosis 1)
M60974	<i>GADD45A</i>	Growth arrest and DNA-damage-inducible, α
U28749	<i>HMGIC</i>	High-mobility group protein isoform I-C
X85137	<i>KNSL1</i>	Kinesin-like 1
X62153	<i>MCM3</i>	Minichromosome maintenance deficient (<i>Saccharomyces cerevisiae</i>) 3
X74331	<i>PRIM2A</i>	Primase, polypeptide 2A (58kD)
J04088	<i>TOP2A</i>	Topoisomerase (DNA) II α (170kD)
Signaling		
X84213	<i>BAK1</i>	BCL2-antagonist/killer 1
X52541	<i>EGR1</i>	Early growth response 1
L26336	<i>HSPA2</i>	Heat shock 70kD protein 2
M63838	<i>IF116</i>	Interferon, γ -inducible protein 16
X04602	<i>IL6</i>	Interleukin 6 (interferon, beta 2)
U20734	<i>JUNB</i>	<i>jumB</i> proto-oncogene
S57212	<i>MEF2C</i>	MADS box transcription enhancer factor 2, polypeptide C
M12174	<i>ARHB</i>	<i>ras</i> homologue gene family, member B
S82240	<i>ARHE</i>	<i>ras</i> homologue gene family, member E
U27655	<i>RGS3</i>	Regulator of G-protein signalling 3
U27768	<i>RGS4</i>	Regulator of G-protein signalling 4
X82877	<i>RSC1A1</i>	Regulatory solute carrier protein, family 1, member 1
U72206	<i>ARHGEF2</i>	rho/rac guanine nucleotide exchange factor (GEF) 2
D43968	<i>RUNX1</i>	runx-related transcription factor 1 (oncogene)
X64652	<i>Mssp-1</i>	Single-stranded DNA-binding protein
U14134		Transcription factor IIIa
X66087	<i>MYBL1</i>	v- <i>myb</i> avian myeloblastosis viral oncogene homologue-like 1
M92843	<i>ZFP36</i>	Zinc finger protein homologous to Zfp-36 in mouse
Miscellaneous		
M19650	<i>CNP</i>	2',3'-Cyclic nucleotide 3' phosphodiesterase
M36430	<i>ARF1</i>	ADP-ribosylation factor 1
M57763	<i>ARF6</i>	ADP-ribosylation factor 6
M80899	<i>AHNAK</i>	AHNAK nucleoprotein (desmoyokin)
L42563	<i>ATPALI</i>	ATPase, Na ⁺ /K ⁺ transporting, α polypeptide-like 1
HG2290-HT2386_at		Calcitonin
M19311	<i>CALM2</i>	Calmodulin 2 (phosphorylase kinase, δ)
M30448	<i>CSNK2B</i>	Casein kinase 2, β polypeptide
U13737	<i>CASP3</i>	Caspase 3, apoptosis-related cysteine protease
U32114		Caveolin-2 mRNA
X03934	<i>CD3D</i>	CD3D antigen, δ polypeptide (TiT3 complex)
M23323	<i>CD3E</i>	CD3E antigen, ϵ polypeptide (TiT3 complex)
U76764	<i>CD97</i>	CD97 antigen
M27436	<i>F3</i>	Coagulation factor III (thromboplastin, tissue factor)
X15880	<i>COL6A1</i>	Collagen, type VI, $\alpha 1$
X52022	<i>COL6A3</i>	Collagen, type VI, $\alpha 3$
D26535	<i>DLST</i>	Dihydrolipoamide S-succinyltransferase
U53445	<i>DOC1</i>	Down-regulated in ovarian cancer 1
U00802	<i>DBN1</i>	Drebin 1
U14187	<i>EFNA3</i>	Ephrin-A3
U43916	<i>EMP1</i>	Epithelial membrane protein 1
M24485	<i>GSTP1</i>	Glutathione S-transferase π
L38503	<i>GSTT2</i>	Glutathione S-transferase $\theta 2$
X79537		Glycogenin
X76942	<i>GOLGA4</i>	Golgi autoantigen, golgin subfamily a, 4
M65028	<i>hnRNP A/E</i>	Heterogeneous nuclear ribonucleoprotein A/B
D16593	<i>HPCA1</i>	Hippocalcin
U54804	<i>HAS2</i>	Hyaluron synthase
U29343	<i>HMMR</i>	Hyaluronan-mediated motility receptor (RHAMM)
L41919	<i>HIC1</i>	Hypermethylated in cancer 1
X57809	<i>IGL</i>	Immunoglobulin λ locus
X90763	<i>KRTHA5</i>	Keratin, hair, acidic, 5
D31764	<i>KIAA0064</i>	KIAA0064 gene product
D50925	<i>KIAA0135</i>	KIAA0135 protein
X79683	<i>LAMB2</i>	Laminin, $\beta 2$ (laminin S)
U65416	<i>MICB</i>	MHC class I polypeptide-related sequence B
J02854	<i>MYRL2</i>	Myosin regulatory light chain 2, smooth muscle isoform
M94167	<i>NRG1</i>	Neuregulin 1
D86425	<i>NID2</i>	Nidogen 2
U24183	<i>PFKM</i>	Phosphofructokinase, muscle
K03021	<i>PLAT</i>	Plasminogen activator, tissue
U09937	<i>PLAUR</i>	Plasminogen activator, urokinase receptor
M21574	<i>PDGFRA</i>	Platelet-derived growth factor receptor, α polypeptide
Z11585		Potassium channel protein (Gb:Z11585)
U25988	<i>PSG11</i>	Pregnancy-specific β -1-glycoprotein 11
D28235	<i>PTGS2</i>	Prostaglandin-endoperoxide synthase 2
U19948	<i>PDI</i>	Protein disulfide isomerase
X17042	<i>PRG1</i>	Proteoglycan core protein
X14684	<i>SSB</i>	Sjögren syndrome antigen B (autoantigen La)
U19495	<i>SDF1</i>	Stromal cell-derived factor 1
D79984	<i>SUPT6H</i>	Suppressor of Ty (<i>Saccharomyces cerevisiae</i>) 6 homologue
M32304	<i>TIMP2</i>	Tissue inhibitor of metalloproteinase 2
M99439	<i>TLF4</i>	Transducin-like enhancer of split 4, homologue of Drosophila E(sp1)
M55153	<i>TGM2</i>	Transglutaminase 2
D78577	<i>YWHAH</i>	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein

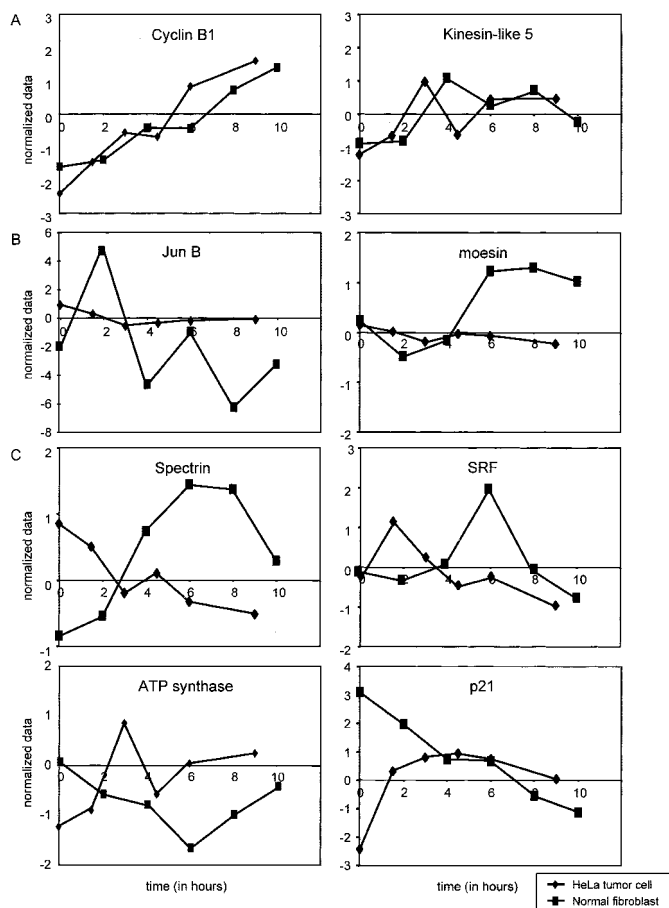


Fig. 7. Examples of differences in expression patterns between normal and tumor cells. A, *cyclin B1* and *kinesin-like 5* genes are examples of two genes with similar cell cycle related expression patterns in both cell types. B, *jun B* and *moesin* are two genes that exhibit cell cycle-related fluctuations in expression pattern, yet are constitutively expressed in the HeLa cell cycle. C, Spectrin, SRF, ATP synthase, and p21^{CIP1} are all expressed with cell cycle-related fluctuations in both cell types, which are not similar.

role in aberrant cell growth, we compared our results with gene expression data obtained from diploid fibroblasts synchronized in S-phase (19).⁶ The majority of the transcripts, encoding oncoproteins, proliferation-inducing proteins, and potential signal transduction attenuators, were not expressed in normal cells, suggesting a role in tumor growth. Most of those genes that were expressed in normal cells exhibited differences in temporal expression patterns between normal and tumor cells, such as elevated expression levels or loss of fluctuation.⁷ These data reveal that the timing of the maximal expression of cancer-related genes is specific for particular cell cycle stages. Our findings are consistent with potential roles for these proteins in the regulation of cell cycle progression or in the timing of certain tumor-related processes during the cell cycle.

Comparison of Normal and Tumor Cells Revealed a Core Group of Cell Growth-related Genes. Gene expression analysis for synchronized normal fibroblasts has been presented (19) using the same type of microarrays. To evaluate the similarities and differences in gene expression in normal and tumor cells, we analyzed the global expression data from the two studies. The results of this comparison reveal that two-thirds (*i.e.*, 1,837) of the total number of genes expressed in HeLa cells are also expressed in normal fibroblasts and

vice versa. A large number of genes exhibit higher expression levels in the cancer cell and often lose their cell cycle-related pattern. This group mostly included transcription factors, structural proteins, and several proteins involved in the process of chromosomal segregation. We scored the occurrence of genes in several functional groups and compiled the results (Table 2). Two groups (Receptors and Transport) include many genes that exhibit cell type-specific expression and may represent phenotypic markers. In contrast, genes in five groups (*i.e.*, Cell cycle, DNA replication, DNA repair, RNA/protein synthesis, and Degradation) are expressed in both normal and tumor cells. Thus, despite significant phenotypic differences in the cell types, there is a striking conservation in the expression of multiple functional classes of genes. For additional information, see our website.

We further compared the two datasets to identify genes involved in the cell cycle for which expression patterns are conserved, genes for which patterns are altered, and genes that are not expressed in the cancer cell. Of the 222 cell cycle genes identified in the human diploid fibroblast study, 81 genes (36%) are expressed below the level of detection in the tumor cell cycle. The most pronounced difference is that several RNAs for proteins involved in cell adhesion and cell structure are not expressed in the HeLa tumor cell, as well as many genes encoding signaling molecules. Therefore, the differences between normal and tumor cells with respect to cell structure and growth characteristics are reflected by gene expression profiles.

In contrast to these genes that are not expressed, the group of cell cycle-regulated genes identified in the diploid fibroblast cells contains 141 (64%) genes that were also expressed in the synchronized HeLa cells. Table 3 provides an overview of the similarities and differences in the expression patterns of cell cycle-specific genes identified in the two studies. The data reveal that the expression patterns for many genes (45 of 141) are different, and most of these genes are constitutively expressed in the tumor cell line (for example, transcription factors and genes involved in mitosis; see Fig. 7). Only a limited number of genes exhibited temporal expression patterns in the cancer cells that were fluctuating yet different from those in the normal cell type. Among these genes were the tumor repressor p21, the transcription factor serum response factor, the protein spectrin (involved in cell shape), and metabolic proteins such as ATP synthase (Table 3 and Fig. 7). Thus, although a number of genes show alterations in the temporal regulation of RNA levels, the expression of the majority of this subgroup of genes is no longer cell cycle regulated.

The expression patterns of the remaining genes (82 of 141) exhibit similar fluctuations in expression during the cell cycle. Many genes that display cell cycle-controlled expression are key cell cycle regulatory factors, including kinesins, MCMs, histones, replication factor C, and several cyclins and *cdcs* (Table 3 and Fig. 7). Hence, a relatively small fraction (37%) of the genes that are cell cycle regulated in the diploid fibroblast cell type are also expressed and similarly modulated during the tumor cell cycle. These genes may constitute a core set of regulatory genes fundamental to cell cycle progression.

In conclusion, in this study we have characterized the sequential expression of genes that regulate or mediate genome replication. We identified groups of genes that are maximally expressed at specific substages during S-phase and G₂. These genes encode cell growth and/or cancer-related proteins that may play specific roles at critical cell cycle transitions. We propose that the classically defined S and G₂ phases of the cell cycle can be subdivided into distinct substages based on diagnostic gene expression profiles. The second principal finding that emerged from our studies is the identification of a core set of genes that is similarly regulated in both normal and tumor cells and

⁶ Internet address: <http://www.salk.edu/docs/labs/chipdata/primary.html>.

⁷ For additional information, see our web site: <http://www.umassmed.edu/cellbio/labs/steinlab>.

may support the fundamental mechanisms required for cell cycle progression.

ACKNOWLEDGMENTS

We thank Dr. Feng He for experimental advice, the members of the laboratory for stimulating discussions, and Rosa Mastrototaro for technical assistance.

REFERENCES

- Lipinski, M. M., and Jacks, T. The retinoblastoma gene family in differentiation and development. *Oncogene*, *18*: 7873–7882, 1999.
- Ishida, S., Huang, E., Zuzan, H., Spang, R., Leone, G., West, M., and Nevins, J. R. Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. *Mol. Cell. Biol.*, *21*: 4684–4699, 2001.
- Ford, H. L., and Pardee, A. B. The S phase: beginning, middle, and end: a perspective. *J. Cell Biochem. Suppl.*, *30–31*: 1–7, 1998.
- Pavletich, N. P. Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *J. Mol. Biol.*, *287*: 821–828, 1999.
- Wu, C. L., Zukerberg, L. R., Ngwu, C., Harlow, E., and Lees, J. A. *In vivo* association of E2F and DP family proteins. *Mol. Cell. Biol.*, *15*: 2536–2546, 1995.
- Stein, G. S., Stein, J. L., van Wijnen, A. J., and Lian, J. B. Regulation of histone gene expression. *Curr. Opin. Cell Biol.*, *4*: 166–173, 1992.
- Osley, M. A. The regulation of histone synthesis in the cell cycle. *Annu. Rev. Biochem.*, *60*: 827–861, 1991.
- Stein, G. S., Stein, J. L., Lian, J. B., Last, T. J., Owen, T., and McCabe, L. Synchronization of normal diploid and transformed mammalian cells. *In: J. E. Celis (ed.)*, *Cell Biology: A Laboratory Handbook*, Vol. 1, pp. 282–287. San Diego: Academic Press, Inc., 1994.
- van der Meijden, C. M. J., Vaughan, P. S., Staal, A., Albig, W., Doenecke, D., Stein, J. L., Stein, G. S., and van Wijnen, A. J. Selective expression of specific histone *H4* genes reflects distinctions in transcription factor interactions with divergent *H4* promoter elements. *Biochim. Biophys. Acta*, *1442*: 82–100, 1998.
- Fujita, M. Cell cycle regulation of DNA replication initiation proteins in mammalian cells. *Front. Biosci.*, *4*: D816–D823, 1999.
- Foiani, M., Lucchini, G., and Plevani, P. The DNA polymerase α -primase complex couples DNA replication, cell-cycle progression and DNA-damage response. *Trends Biochem. Sci.*, *22*: 424–427, 1997.
- Picketts, D. J., Higgs, D. R., Bachoo, S., Blake, D. J., Quarrell, O. W., and Gibbons, R. J. ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum. Mol. Genet.*, *5*: 1899–1907, 1996.
- Rossow, P. W., Riddle, V. G., and Pardee, A. B. Synthesis of labile, serum-dependent protein in early G1 controls animal cell growth. *Proc. Natl. Acad. Sci. USA*, *76*: 4446–4450, 1979.
- Yen, A., and Pardee, A. B. Arrested states produced by isoleucine deprivation and their relationship to the low serum produced arrested state in Swiss 3T3 cells. *Exp. Cell Res.*, *114*: 389–395, 1978.
- Sogawa, K., Imataka, H., Yamasaki, Y., Kusume, H., Abe, H., and Fujii-Kuriyama, Y. cDNA cloning and transcriptional properties of a novel GC box-binding protein, BTEB2. *Nucleic Acids Res.*, *21*: 1527–1532, 1993.
- Kim, S., Na, J. G., Hampsey, M., and Reinberg, D. The Dr1/DRAP1 heterodimer is a global repressor of transcription *in vivo*. *Proc. Natl. Acad. Sci. USA*, *94*: 820–825, 1997.
- Nigg, E. A. Mitotic kinases as regulators of cell division and its checkpoints. *Nat. Rev. Mol. Cell. Biol.*, *2*: 21–32, 2001.
- Salem, C., El Alfy, M., and Leblond, C. P. Changes in the rate of RNA synthesis during the cell cycle. *Anat. Rec.*, *250*: 6–12, 1998.
- Cho, R. J., Huang, M., Campbell, M. J., Dong, H., Steinmetz, L., Sapinoso, L., Hampton, G., Elledge, S. J., Davis, R. W., and Lockhart, D. J. Transcriptional regulation and function during the human cell cycle. *Nat. Genet.*, *27*: 48–54, 2001.
- Moir, R. D., Yoon, M., Khuon, S., and Goldman, R. D. Nuclear lamins A and B1. Different pathways of assembly during nuclear envelope formation in living cells. *J. Cell Biol.*, *151*: 1155–1168, 2000.
- Hensel, F., Brandlein, S., Eck, M., Schmidt, K., Krenn, V., Kloetzer, A., Bachi, A., Mann, M., Muller-Hermelink, H. K., and Vollmers, H. P. A novel proliferation-associated variant of CFR-1 defined by a human monoclonal antibody. *Lab. Investig.*, *81*: 1097–1108, 2001.
- Yamamoto, T., Ikawa, S., Akiyama, T., Semba, K., Nomura, N., Miyajima, N., Saito, T., and Toyoshima, K. Similarity of protein encoded by the human *c-erbB2* gene to epidermal growth factor receptor. *Nature (Lond.)*, *319*: 230–234, 1986.
- Tulchinsky, E. Fos family members: regulation, structure and role in oncogenic transformation. *Histol. Histopathol.*, *15*: 921–928, 2000.
- Chan, A. M., McGovern, E. S., Catalano, G., Fleming, T. P., and Miki, T. Expression cDNA cloning of a novel oncogene with sequence similarity to regulators of small GTP-binding proteins. *Oncogene*, *9*: 1057–1063, 1994.
- Adelmant, G., Gilbert, J. D., and Freytag, S. O. Human translocation liposarcoma-CCAAT/enhancer binding protein (C/EBP) homologous protein (TLS-CHOP) oncoprotein prevents adipocyte differentiation by directly interfering with C/EBP β function. *J. Biol. Chem.*, *273*: 15574–15581, 1998.
- Lilly, M., Sandholm, J., Cooper, J. J., Koskinen, P. J., and Kraft, A. The PIM-1 serine kinase prolongs survival and inhibits apoptosis-related mitochondrial dysfunction in part through a bcl-2-dependent pathway. *Oncogene*, *18*: 4022–4031, 1999.
- Emoto, K., Sawada, H., Yamada, Y., Fujimoto, H., Takahama, Y., Ueno, M., Takayama, T., Uchida, H., Kamada, K., Naito, A., Hirao, S., and Nakajima, Y. Annexin II overexpression is correlated with poor prognosis in human gastric carcinoma. *Anticancer Res.*, *21*: 1339–1345, 2001.
- Schweinfest, C. W., Graber, M. W., Chapman, J. M., Pappas, T. S., Baron, P. L., and Watson, D. K. CaSm: an Sm-like protein that contributes to the transformed state in cancer cells. *Cancer Res.*, *57*: 2961–2965, 1997.
- Miller, M. J., Martinez, A., Unsworth, E. J., Thiele, C. J., Moody, T. W., Elsasser, T., and Cuttitta, F. Adrenomedullin expression in human tumor cell lines. Its potential role as an autocrine growth factor. *J. Biol. Chem.*, *271*: 23345–23351, 1996.
- van der Lugt, N. M., Domen, J., Linders, K., van Roon, M., Robanus-Maandag, E., te Riele, H., van der Valk, M., Deschamps, J., Sofroniew, M., and van Lohuizen, M. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. *Genes Dev.*, *8*: 757–769, 1994.
- Zhou, J., and Hsieh, J. T. The inhibitory role of DOC-2/DAB2 in growth factor receptor-mediated signal cascade. DOC-2/DAB2-mediated inhibition of ERK phosphorylation via binding to Grb2. *J. Biol. Chem.*, *276*: 27793–27798, 2001.
- van Hul, W., Wuys, W., Hendrickx, J., Speleman, F., Wauters, J., De Boule, K., Van Roy, N., Bossuyt, P., and Willems, P. J. Identification of a third EXT-like gene (*EXTL3*) belonging to the *EXT* gene family. *Genomics*, *47*: 230–237, 1998.
- Mok, S. C., Chan, W. Y., Wong, K. K., Cheung, K. K., Lau, C. C., Ng, S. W., Baldini, A., Colitti, C. V., Rock, C. O., and Berkowitz, R. S. DOC-2, a candidate tumor suppressor gene in human epithelial ovarian cancer. *Oncogene*, *16*: 2381–2387, 1998.
- Sakamuro, D., Elliott, K. J., Wechsler-Reya, R., and Prendergast, G. C. BIN1 is a novel MYC-interacting protein with features of a tumour suppressor. *Nat. Genet.*, *14*: 69–77, 1996.
- Jiang, H., Lin, J. J., Su, Z. Z., Goldstein, N. I., and Fisher, P. B. Subtraction hybridization identifies a novel melanoma differentiation associated gene, *mda-7*, modulated during human melanoma differentiation, growth and progression. *Oncogene*, *11*: 2477–2486, 1995.
- Welt, C. K., Lambert-Messerlian, G., Zheng, W., Crowley, W. F., Jr., and Schneyer, A. L. Presence of activin, inhibin, and follistatin in epithelial ovarian carcinoma. *J. Clin. Endocrinol. Metab.*, *82*: 3720–3727, 1997.
- Plowman, G. D., Green, J. M., McDonald, V. L., Neubauer, M. G., Distchech, C. M., Todaro, G. J., and Shoyab, M. The amphiregulin gene encodes a novel epidermal growth factor-related protein with tumor-inhibitory activity. *Mol. Cell. Biol.*, *10*: 1969–1981, 1990.
- Prevot, D., Morel, A. P., Voeltzel, T., Rostan, M. C., Rimokh, R., Magaud, J. P., and Corbo, L. Relationships of the antiproliferative proteins BTG1 and BTG2 with CAF1, the human homolog of a component of the yeast CCR4 transcriptional complex: involvement in estrogen receptor α signaling pathway. *J. Biol. Chem.*, *276*: 9640–9648, 2001.
- Wicki, R., Franz, C., Scholl, F. A., Heizmann, C. W., and Schafer, B. W. Repression of the candidate tumor suppressor gene *S100A2* in breast cancer is mediated by site-specific hypermethylation. *Cell Calcium*, *22*: 243–254, 1997.
- Tiruppathi, C., Yan, W., Sandoval, R., Naqvi, T., Pronin, A. N., Benovic, J. L., and Malik, A. B. G protein-coupled receptor kinase-5 regulates thrombin-activated signaling in endothelial cells. *Proc. Natl. Acad. Sci. USA*, *97*: 7440–7445, 2000.
- Scharenberg, A. M., El Hillal, O., Fruman, D. A., Beitz, L. O., Li, Z., Lin, S., Gout, I., Cantley, L. C., Rawlings, D. J., and Kinet, J. P. Phosphatidylinositol-3,4,5-trisphosphate (PtdIns-3,4,5-P3)/Tec kinase-dependent calcium signaling pathway: a target for SHIP-mediated inhibitory signals. *EMBO J.*, *17*: 1961–1972, 1998.