

# Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors<sup>□</sup>

Michael L. Whitfield,\* Gavin Sherlock,\* Alok J. Saldanha,\* John I. Murray,\* Catherine A. Ball,\* Karen E. Alexander,<sup>†</sup> John C. Matese,\* Charles M. Perou,<sup>‡</sup> Myra M. Hurt,<sup>†</sup> Patrick O. Brown,<sup>§||¶</sup> and David Botstein\*<sup>¶</sup>

\*Departments of Genetics and <sup>§</sup>Biochemistry, <sup>||</sup>Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305; <sup>†</sup>Department of Biomedical Sciences, College of Medicine, Florida State University, Tallahassee, Florida 32306; and <sup>‡</sup>Department of Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Submitted January 8, 2002; Revised February 26, 2002; Accepted March 7, 2002  
Monitoring Editor: Mark J. Solomon

The genome-wide program of gene expression during the cell division cycle in a human cancer cell line (HeLa) was characterized using cDNA microarrays. Transcripts of >850 genes showed periodic variation during the cell cycle. Hierarchical clustering of the expression patterns revealed coexpressed groups of previously well-characterized genes involved in essential cell cycle processes such as DNA replication, chromosome segregation, and cell adhesion along with genes of uncharacterized function. Most of the genes whose expression had previously been reported to correlate with the proliferative state of tumors were found herein also to be periodically expressed during the HeLa cell cycle. However, some of the genes periodically expressed in the HeLa cell cycle do not have a consistent correlation with tumor proliferation. Cell cycle-regulated transcripts of genes involved in fundamental processes such as DNA replication and chromosome segregation seem to be more highly expressed in proliferative tumors simply because they contain more cycling cells. The data in this report provide a comprehensive catalog of cell cycle regulated genes that can serve as a starting point for functional discovery. The full dataset is available at <http://genome-www.stanford.edu/Human-CellCycle/HeLa/>.

## INTRODUCTION

Proper regulation of the cell division cycle is crucial to the growth and development of all organisms; understanding this regulation is central to the study of many diseases, most notably cancer. Many of the diverse molecular processes needed to duplicate a cell, such as DNA replication, monitoring of replication fidelity, and accurate segregation of chromosomes to daughter cells, are characteristically aberrant in cancer cells. These processes and their regulation have been extensively investigated at the molecular level (reviewed in Stillman, 1996; Nurse, 2000; Shah and Cleveland, 2000; Hinchcliffe and Sluder, 2001; Wittmann *et al.*,

2001). Characterization of the genome-wide transcriptional program of the cell division cycle in mammalian cells is a critical step toward understanding the basic cell cycle processes and their role in cancer.

The genome-wide transcriptional program during the cell cycle has been investigated in a wide range of organisms, including budding yeast (Cho *et al.*, 1998; Spellman *et al.*, 1998), bacteria (Laub *et al.*, 2000), primary human fibroblasts (Iyer *et al.*, 1999; Cho *et al.*, 2001), mouse fibroblasts (Ishida *et al.*, 2001), and human HeLa cells (Crawford and Piwnicka-Worms, 2001). In synchronized cultures of the budding yeast *Saccharomyces cerevisiae*, 800 genes were found to be periodically expressed (Spellman *et al.*, 1998) and in synchronized cultures of the bacterium *Caulobacter crescentus*, 553 periodically expressed genes were identified (Laub *et al.*, 2000). The human cell cycle was previously studied first by observing the genes whose expression changes when primary fibroblast cultures were stimulated with serum to reenter the cell cycle from a serum-deprived, resting state. Interpretation of these studies is complicated by the subsequent realization

<sup>□</sup> A complete data set for this article is available at [www.molbiol-cell.org](http://www.molbiol-cell.org). Article published online ahead of print. Mol. Biol. Cell 10.1091/mbc.02-02-0030. Article and publication date are at [www.molbiolcell.org/cgi/doi/10.1091/mbc.02-02-0030](http://www.molbiolcell.org/cgi/doi/10.1091/mbc.02-02-0030).

<sup>¶</sup> Corresponding authors. E-mail addresses: [botstein@genome.stanford.edu](mailto:botstein@genome.stanford.edu) or [pbrown@cmgm.stanford.edu](mailto:pbrown@cmgm.stanford.edu).

that the major gene expression program observed is more than a simple cell cycle response, but rather is combined with a prominent wound-healing response (Iyer *et al.*, 1999). By modifying this approach, Ishida *et al.* (2001) identified 578 genes that were induced upon serum stimulation and provided evidence of cell cycle regulation based on the pattern of expression after release from a cell cycle arrest with hydroxyurea, which blocks DNA replication. A study of the G2 DNA damage checkpoint in HeLa cells synchronized by a double thymidine block identified transcripts that showed differential expression from S to G2/M phases but failed to identify a specific transcriptional response associated with the G2 DNA damage checkpoint (Crawford and Piwnicka-Worms, 2001). Finally, ~700 genes were identified as cell cycle regulated in primary human fibroblasts synchronized by a double thymidine block, which attempts to avoid the confounding serum stimulation response altogether (Cho *et al.*, 2001).

Herein, we report the extension of these results by synchronizing HeLa cells not only by double thymidine block but also by two additional methods, a thymidine-nocodazole block and a physical method, mitotic shake-off. By combining data from experiments using these three different synchronization methods in five independent experiments, we identified >850 genes that are periodically expressed during the cell cycle. Using this list of genes we were able to show that most of the genes previously associated with the proliferative state of tumors are among the genes we identified as periodically synthesized during the human cell division cycle.

## MATERIALS AND METHODS

### Cell Culture, Synchronization, and RNA Preparation

HeLa S3 cells were plated ( $2 \times 10^6$  cells) in 150-mm tissue culture dishes in DMEM with 10% fetal bovine serum and 100 U of penicillin-streptomycin (Invitrogen, Carlsbad, CA). Cells were arrested in S phase by using a double thymidine block or in mitosis with a thymidine-nocodazole block essentially as described previously (Whitfield *et al.*, 2000) and in the supplemental material. Poly(A) RNA was prepared from cells collected at intervals (typically 1–2 h) by lysing cells directly on the plate with the FastTrack 2.0 mRNA isolation kit (Invitrogen). Synchrony was monitored by flow cytometry analysis of propidium iodide-stained cells (Stanford Shared FACS Facility, Stanford University School of Medicine, Stanford, CA).

Mitotic cells were collected every 10 min with an automated cell shaker (Eliassen *et al.*, 1998), stored on ice, and plated at 2-h intervals in fresh prewarmed media (at least  $10^6$  cells for each time point). Because only a small number of cells can be obtained by mitotic shake-off, total RNA was prepared using ULTRASPEC RNA isolation system (Biotecx, Houston, TX). The number of cells in S phase at each point was determined using a 5-Bromo-2'-deoxyuridine (BrdU) Labeling and Detection Kit I (Roche Applied Science, Indianapolis, IN).

Reference RNA was prepared from asynchronously growing HeLa cells using TRIzol (Invitrogen) and poly(A) RNA isolated by affinity chromatography on oligo-dT cellulose (Amersham Biosciences, Piscataway, NJ). Poly(A) RNA was used as a reference in all experiments except the mitotic shake-off, where total RNA was labeled.

### cDNA Synthesis and Microarray Hybridization

RNA from synchronous cells was reverse transcribed into Cy5-dUTP (Amersham Biosciences)-labeled cDNA, and reference RNA

reverse transcribed into Cy3-dUTP (Amersham Biosciences)-labeled cDNA by standard methods (Eisen and Brown, 1999) (Details are available at <http://genome-www.stanford.edu/Human-CellCycle/HeLa/>.) Total RNA samples from cells collected in the mitotic shake-off experiment, and total reference RNA were first amplified using a modified Eberwine protocol before cDNA synthesis (Wang *et al.*, 2000) then labeled cDNA was prepared from the amplified RNA.

Spotted cDNA microarrays, containing 22,692 elements representing ~16,332 different human genes or containing 43,198 elements representing ~29,621 genes (estimated by UNIGENE clusters), were manufactured in the Stanford Microarray Facility (<http://www.microarray.org>). Equal amounts of Cy5- and Cy3-labeled cDNA were hybridized to spotted cDNA microarrays and scanned using a GenePix 4000A Scanner (Axon Instruments, Union City, CA). Detailed protocols are available at <http://brownlab.stanford.edu/protocols.html/>.

### Data Processing

Data were extracted by superimposing a grid over each array using GenePix 3.0 software (Axon Instruments). Spots of poor quality, determined by visual inspection, were removed from further analysis. Data collected for each array were stored in the Stanford Microarray Database (SMD) and are available from SMD at <http://genome-www.stanford.edu/microarray/> (Sherlock *et al.*, 2001).

Only features with signal intensity at least 20% above background in both Cy5 and Cy3 channels and for which adequate quality data were obtained for at least 80% of the samples in a given time course were analyzed further. Data points that did not meet these criteria are blank in the primary data tables.  $\log_2(\text{Cy5/Cy3})$  was retrieved for each data point and used for all analysis, where (Cy5/Cy3) is the normalized ratio of the background-corrected intensities, as defined in SMD (Sherlock *et al.*, 2001).

Because of systematic differences between experiments (e.g., array batch, labeling methods, and synchronization methods) each time course was centered independently by filtering out the first, most significant eigengene (Alter *et al.*, 2000), which was a dominant, constant vector. Because singular value decomposition (SVD) requires a full data matrix, missing data points were estimated using a k-nearest neighbors algorithm (Troyanskaya *et al.*, 2001) with  $k = 12$ . These imputed values were used throughout the analysis and but were restored to “unknown” status in the figures and left blank in the primary data tables (<http://genome-www.stanford.edu/Human-CellCycle/HeLa/>).

### Identification of Periodically Expressed Transcripts

A Fourier Transform (eqs. 1–3) was applied to the data for each clone in an experiment (Spellman *et al.*, 1998), and the resulting vector ( $C$ , eq. 3) of the sine ( $A$ ) and cosine ( $B$ ) coefficients was stored, where  $T$  is the cell cycle period,  $t$  is the time after release,  $\phi$  is the phase offset, and  $\text{ratio}(t)$  is the normalized Cy5/Cy3 expression ratio at time  $t$ . The value of  $\phi$  was initially set to 0. The values obtained for  $C$  were determined over a range of 40 values of  $T$  equally spaced 1 h above and below the estimated cell cycle period and the resulting values averaged and stored.

$$A = \sum \sin(2\pi(t/T) + \phi) \log_2(\text{ratio}(t)) \quad (1)$$

$$B = \sum \cos(2\pi(t/T) + \phi) \log_2(\text{ratio}(t)) \quad (2)$$

$$C = \langle A, B \rangle \quad (3)$$

$$D = \sqrt{A^2 + B^2} \quad (4)$$

The optimal cell cycle period was determined by finding the value of  $T$  where the largest numbers of genes pass an arbitrary magnitude cutoff ( $D$ , eq. 4). Fourier transforms were applied to the data series for each gene (eqs. 1–3), with equally spaced values of  $T$ , from

**Table 1.** Periodicity scores of selected genes

Sample periodicity scores for representative cell cycle regulated genes. The cutoff for a cell cycle regulated gene was established by balancing the false positive and false negative rates. We used a cutoff that gave an estimated 0.15–0.75% false positive rate and 6.5% false negative rate, resulting in 1333 clones. The maximum fold change is the maximum amplitude observed in any of the five synchronizations. The average fold change is the average amplitude across all five synchronizations.

Rank	Gene	Periodicity score	Max fold change	Avg. fold change
1	STK15	58.8	16	10
2	PLK	56.0	11	8.6
7	CENPE	42.6	11	6.8
17	CKS2	35.1	11	5.5
25	CCNA2	31.2	5.9	4.5
43	CDC6	25.3	3.8	3.0
50	CCNB1	24.3	7.3	4.1
54	E2F1	23.0	4.9	3.9
95	PCNA	16.9	3.4	2.7
105	CIT	16.2	7.0	4.0
135	GOLGIN-67	13.6	4.8	3.1
182	ORC 1	11.2	3.7	2.2
355	PRIM1	7.26	2.6	2.1
499	VCL	5.80	3.4	2.7
607	BRCA1	5.17	2.5	2.1
618	CCNE2	7.18	3.1	2.5
1127	PSEN1	3.65	2.1	1.8
1333	CDC42	3.29	2.3	2.0
<b>1334</b>	<b>cutoff</b>	<b>3.286</b>	—	—
1358	ZNF162	3.25	1.9	1.8
1780	YY1	2.75	3	2.3

0 to 40 h in 15-min increments. The number of genes whose magnitude ( $D$ , eq. 4) exceeded the arbitrary cutoff of 3, 5, or 7 was plotted and a period ( $T$ ) was chosen that maximized the numbers of genes exceeding our arbitrary thresholds. In most cases, the determined value of  $T$  was consistent with the data obtained by flow cytometry.

Because each experiment does not start at exactly the same point in the cell cycle, an offset ( $\phi$ , eqs. 1 and 2) was calculated for each dataset relative to the first double thymidine arrest. The magnitudes of the Fourier transform ( $D$ , eq. 4) for the 1000 highest scoring clones were summed using different values of  $\phi$ , equally spaced between 0 and  $2\pi$ . The offset that gave the highest average combined magnitude ( $D$ , eq. 4) between the two datasets for these 1000 genes was then used. The Fourier transform was then repeated on the remaining datasets with the following values of  $T$  and  $\phi$ : Thy-Thy 2 ( $T = 15.5$  h,  $\phi = 0.5$  rad), Thy-Thy 3 ( $T = 15.4$ ,  $\phi = 0.0$  rad), Thy-Noc ( $T = 18.5$ ,  $\phi = 3.2$  rad), and mitotic selection ( $T = 24.5$ ,  $\phi = 3.5$  rad). The vectors  $C$  (eq. 3) for all five datasets were then summed and the genes ranked according to the magnitude ( $D$ , eq. 4) of their combined vectors. Note, the Thy-Noc and mitotic shake-off experiments, which arrest cells in mitosis, have offsets of approximately half a cycle ( $\pi$  radians) from Thy-Thy 1, which arrest cells at G1/S.

Because the gene expression profiles of many cell cycle genes do not precisely match sine and cosine curves, the expression profile of each gene was correlated to an idealized vector obtained from known genes expressed in each cell cycle phase (G1/S, S, G2, and G2/M) as defined in Figure 2A. Using a standard Pearson correlation, each gene received a peak correlation score defined as the highest absolute value correlation between one of the four idealized vectors and its expression profile (Spellman *et al.*, 1998). The absolute value of the peak correlation was used to scale the magnitude of the vector ( $C$ , eq. 2) generating a “periodicity score” for each gene (Table 1).

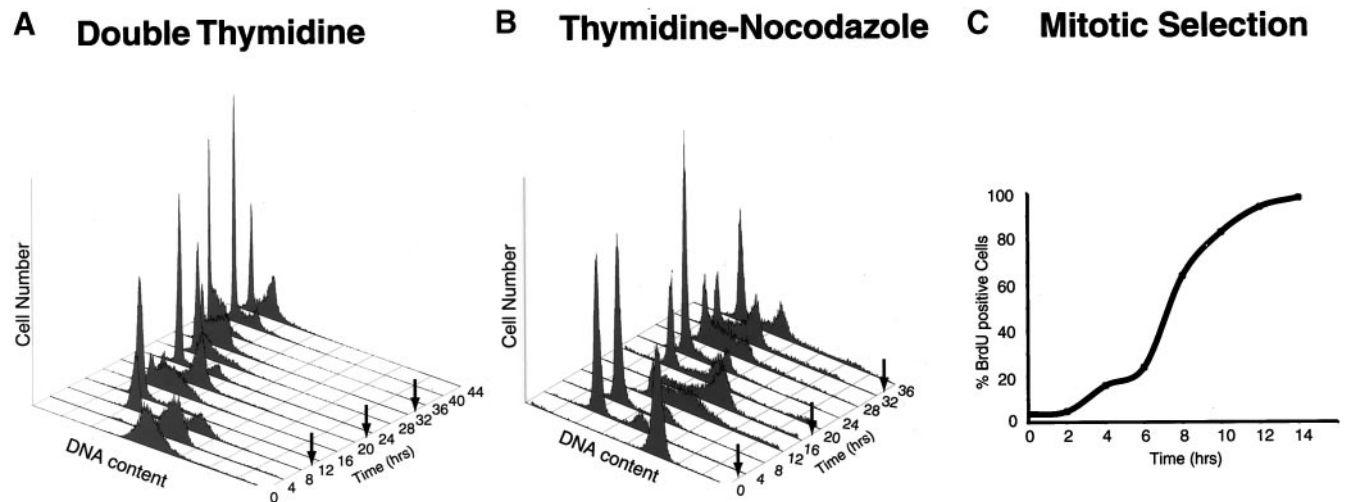
To estimate the minimum periodicity score for a cell cycle regulated gene, the analysis described above was repeated on randomized data. The data were randomized either within rows only, or within both rows and columns, for each of the five datasets starting

with the imputed, SVD-centered data. The Fourier transform and correlations were applied using the previously calculated values of  $T$  and  $\phi$ ; the resulting vectors ( $C$ , eq. 3) were combined for each dataset. The magnitude ( $D$ , eq. 4) of the Fourier transform was scaled by each “gene’s” peak correlation to one of the four ideal expression profiles. This analysis, including the data randomization, was repeated 10 times and the scores combined by averaging the score for each of the highest scoring “genes” from each randomization, followed by the second highest, third highest, etc. The estimated false positive rate at a given periodicity score is the number of genes that obtain at least that score in the randomized data. We chose a minimum periodicity score of 3.29, which gave us 1333 clones at an initial false positive estimate of 1% when the data were randomized in rows. Repeating this analysis 10 times gave an estimated 10 false positives (0.75%; periodicity scores of 5.18–3.33) when the data were randomized only within rows, and two false positives (0.15%; periodicity scores of 3.72 and 3.30) when the data were randomized in both rows and columns.

The false positive estimate, calculated above, is likely to underestimate the true false positive rate because it does not take into account genes that received a high Fourier score because they exhibited a sinusoidal pattern in only part of a time course. To filter out genes that did not show periodic expression, autocorrelations for each 1333 clones were calculated (eq. 5). The autocorrelation  $A$  is equal to the summation over all times  $t$  of the product of the ratio at  $t$  multiplied by the ratio at a time  $t + T$ , where  $T$  is the cell cycle period determined by Fourier analysis. If the data for a gene repeats with a period  $T$ , the autocorrelation will be high.

$$A = \sum \text{ratio}(t) * \text{ratio}((t + T) \% n) \quad (5)$$

Autocorrelation scores were calculated for Thy-Thy 3 and the Thy-Noc experiments because they represent multiple cell cycles and because points were taken at equally spaced intervals throughout each time course. In experiment 3 autocorrelations were calculated for periods ( $T$ , eq. 5) of 15, 16, and 17 h. In experiment 4 autocor-



**Figure 1.** Synchrony of HeLa cells. (A) Cells were arrested at the beginning of S phase by using a double thymidine block, and cell synchrony was monitored by flow cytometry of propidium iodide-stained cells. Flow cytometry data were collected for each of the three independent double thymidine blocks performed in this study; data are shown only for the second double thymidine arrest (Thy-Thy2), although equivalent synchrony was obtained in each of the three experiments. The number of cells (arbitrary units) is plotted against DNA content for time points at 4-h intervals for 44 h; an arrow indicates the time of mitosis, as estimated from the flow cytometry data. Upon release from the thymidine block, >95% of the cells progressed into S phase (0–4 h), entered G2 phase (5–6 h), underwent a synchronous mitosis at 7–8 h, and reentered S phase after completing one full cell cycle at 14–16 h. Typically two to three additional synchronous cell cycles were obtained. (B) Cells were arrested in mitosis by blocking first in thymidine followed by release and then blocking in nocodazole. After release from the nocodazole block, most of the cells (>75%) divided synchronously within 2 h of release from the arrest, entered S phase by 10–12 h after release, and completed the next synchronous mitosis by 18–20 h, ultimately completing two full cell cycles. (C) Cells were selected by mitotic shake-off, and the percentage of cells in S phase was measured by BrdU incorporation. The percentage of cells in S phase is plotted against the time after plating; >95% of the cells entered S phase by 12 h.

relations were calculated for periods of 16, 18, and 20 h. The score for each gene in a given time course was taken to be the maximum of the three autocorrelations. The final autocorrelation score assigned to each gene was the sum of the scores calculated for each of the two time courses.

Autocorrelations were used as a filter to remove genes that showed transient expression despite receiving a high periodicity score. 199 Genes with a negative autocorrelation (a negative autocorrelation indicates the measured ratios do not repeat every cell cycle) were eliminated from the initial set of putative cell cycle-regulated genes. Autocorrelations were also calculated for data randomized in rows, whereupon few genes received negative autocorrelations in the randomized data, indicating that the negative scores are unlikely to occur by chance. The distribution of autocorrelation scores is shown in Supplemental Figure 16.

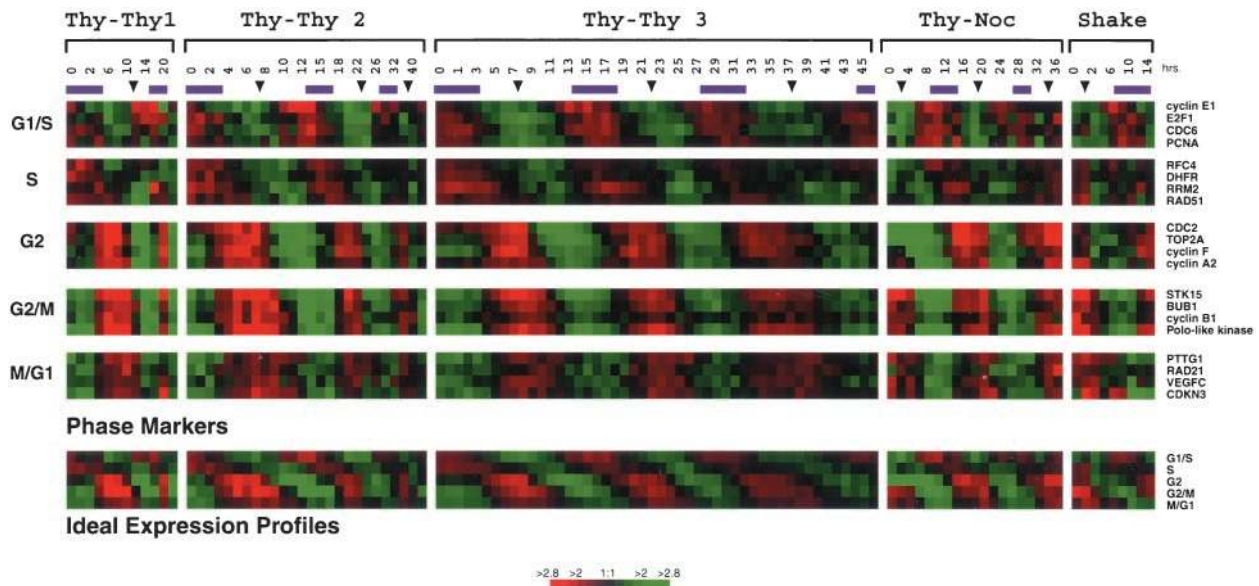
Our final list contains 1134 clones that correspond to 874 UNIGENE clusters (UNIGENE build 143, released November 9, 2001, 21 clones not found in UNIGENE, 66 map to more than one UNIGENE cluster). The data for all 1134 clones as well as the primary data are available at <http://genome-www.stanford.edu/Human-CellCycle/HeLa/>.

## RESULTS

The sensitivity with which one can detect periodic activities in synchronized cell cultures depends almost entirely on the degree to which cells can be synchronized. In particular, cells that fail to begin cycling promptly upon release will contribute noise in the analysis. We chose the well-studied epithelial cell line, HeLa S3, derived from a cervical carcinoma (Puck *et al.*, 1956), specifically because a high degree of

synchrony can be achieved with diverse methods (Knehr *et al.*, 1995; Whitfield *et al.*, 2000). HeLa cells can be synchronized so that 95% of the cells reenter the cell cycle, whereas the comparable figure for primary cells (such as foreskin fibroblasts or mammary epithelial cells) is usually reported to be between 60 and 80% (Tobey *et al.*, 1988; Fonagy *et al.*, 1993; Stampfer *et al.*, 1993). Three different methods were used to obtain synchronous populations; in each case synchrony was monitored either by flow cytometry or BrdU incorporation (Figure 1). A double thymidine block was first used to arrest cells at the G1/S boundary (Adolph and Phelps, 1982), providing the best synchrony in G2 and M phases of the cell cycle (Figure 1A). This protocol provided a somewhat less robust synchrony of the G1-to-S transition, because cells tend to differ in the rate at which they pass through G1. To improve our resolution in the G1-to-S period, we used another method, in which HeLa cells were first arrested by thymidine, released, and subsequently arrested in mitosis with the antimicrotubule drug nocodazole (Figure 1B) (Zieve *et al.*, 1980). In this case, better synchrony was obtained in G1 and S phases, because the synchronous release was at M. The final synchronization method involved no drugs; instead, HeLa cells at metaphase were collected by a physical method, mitotic shake-off (Schneiderman *et al.*, 1972), by using an automated cell shaker (Eliassen *et al.*, 1998). More than 95% of the cells collected by mitotic shake-off reentered the cell cycle and progressed into S phase as monitored by incorporation of BrdU (Figure 1C). In





**Figure 2.** Gene expression of well-characterized cell cycle genes. The expression profiles are shown for 20 well-characterized genes whose expression peaks in each phase of the cell cycle. Data are shown for each of the five time courses (114 total arrays). The purple bars above each column represent S phase and the arrows indicate mitoses as estimated by flow cytometry or BrdU labeling. In each time course, RNA was collected for points (typically every 1–2 h) for 30 h (Thy-Thy1), 44 h (Thy-Thy2), 46 h (Thy-Thy3), 36 h (Thy-Noc), or 14 h (Shake) after the synchronous arrest, at the intervals indicated. The four genes representing each cell cycle phase were averaged to create five different “ideal expression profiles” (G1/S, S, G2, G2/M, and M/G1).

all cases, with the exception of the mitotic shake-off, two to three synchronous cell cycles were obtained.

### Identification of Cell Cycle-regulated Transcripts

To identify cell cycle-regulated transcripts, RNA was isolated from HeLa cells at intervals (typically 1–2 h) after release from a synchronous arrest (Figure 2A). Cy5- or Cy3-labeled cDNA was synthesized using standard protocols and hybridized to cDNA microarrays containing either 22,692 features representing ~16,332 different human genes (Thy-Thy 1 and Thy-Thy 2) or 43,198 features representing ~29,621 different genes (Thy-Thy 3, Thy-Noc, and Mitotic Shake-off) as estimated by UNIGENE clusters.

The data from each of the synchronization experiments required separate analysis before combination into a single dataset as described in MATERIALS AND METHODS. To identify those transcripts that were cell cycle regulated, a periodicity score was obtained for each clone using a Fourier transform and correlation to known cell cycle genes as described for yeast (Spellman *et al.*, 1998). Representative genes and their periodicity scores are shown in Table 1; the range was 0–58.8.

The minimum score for a gene we designated as “cell cycle-regulated” was determined by estimating the false positive rate by using randomized data as described in MATERIALS AND METHODS. We chose a threshold score of 3.29 that gave an estimated 10 false positives (0.75%; periodicity scores of 5.18–3.33) when the data were randomized only within rows (i.e., within each gene), and two false positives (0.15%; periodicity scores of 3.72–3.30) when rows and columns were randomized. Using this relatively conser-

vative threshold, we identified 1333 clones as cell cycle regulated.

A fraction of the genes with relatively high Fourier scores varied in a quasi-sinusoidal manner during part of the time series but clearly were not periodically expressed. This quasi-sinusoidal variation resulted from variation in gene expression at the beginning of each time course, probably as a result of the synchronization procedure. Possible sources of this variation are the serum response, resulting from the addition of fresh growth media upon release of the cells from the arrest, or simply a stress response resulting from the cell cycle block. To filter out genes that did not oscillate across multiple cell cycles, each clone was assigned an autocorrelation score that describes whether the expression ratio at a given time is a good predictor for the value of the expression ratio one cell cycle later. Thus, genes with a consistent pattern of periodic expression every cell cycle have positive autocorrelation scores, whereas genes with low or negative scores do not have consistent periodic expression. Using this method, 199 of the initial 1333 clones received negative autocorrelation scores and were removed from the list, leaving 1134 clones representing 874 different genes.

To assess the quality of the criteria for a periodically expressed gene, the false negative rate was estimated by using a list of known cell cycle-regulated genes compiled from the literature. The known genes were limited to those regulated at the mRNA level during a continuous human cell cycle, as determined by traditional methods (e.g., Northern blot, S1 nuclease assay, or RNase protection; Table 2). Many “known” cell cycle genes are not included in this list

because their regulation has been demonstrated only during the resting-to-growing transition, not in cells synchronized independent of serum, or because data for the regulation of the mRNA could not be found, despite a sometimes extensive literature on protein levels, localization, and function during the cell cycle (e.g., CENPE; Wood *et al.*, 1997; Abrieu *et al.*, 2000). Of the 49 known cell cycle-regulated genes in our list, three were not measured in our analysis. Of the remaining 46, 52.2% (24/46) are found in the top 120 scoring clones, and 93.5% (43/46) are included in the top 850 clones, yielding a false negative rate of 6.5%. Although most of the known genes are represented in the top 850 clones, we believe there is still significant information to be obtained between our 850 and 1134 clones, which include duplicate clones of many of the known genes (e.g., the histones and RAD51) and many genes expressed during mitosis.

### Classification of Transcripts by Cell Cycle Phase

Figure 2 presents data for well-studied genes that showed peak expression at specific points in the cell cycle. The times of S phase and mitosis, as estimated from the flow cytometry and BrdU data (Figure 1), are indicated in Figure 2. Maximum expression of the mitotic cyclins (CCNE1, CCNA2, and CCNB1) was observed in the known temporal order and at the expected times (Pines and Hunter, 1989; Pines and Hunter, 1990; Lew *et al.*, 1991). Cyclin E1 expression at the G1/S boundary was accompanied, as expected, by E2F1 and the DNA replication factors CDC6 and PCNA (Morris and Mathews, 1989; Ohtani *et al.*, 1995; Yan *et al.*, 1998). Many DNA metabolism genes were expressed at the beginning of S phase; these are represented here by RFC4, DHFR, RRM2, and RAD51. In our data, two distinct groups of mitotic genes could be distinguished. Transcripts of some genes, such as CCNA2, TOP2A, Cyclin F, and CDC2, peaked in G2, whereas others, such as CCNB1, BUB1, STK15, and PLK1, peaked in mitosis. The differences in expression of the G2 and M phase genes are most evident in the third double thymidine arrest experiment in which samples were taken every hour for 46 h and at the earliest times after the thymidine-nocodazole arrest (Figure 2). We have followed our data, rather than precedent, in recognizing that many, if not all, the genes expressed during the physical act of mitosis continue to be expressed into the G1 period and have thus labeled these genes M/G1. These genes are represented in Figure 2 by RAD21, PTTG1, VEGFC, and CDKN3.

All 1134 clones identified as periodically expressed were sorted by the point in the cell cycle when they showed peak expression as calculated from the sine and cosine components for the Fourier transform (Figure 3A). The periodic nature of the expression patterns is evident from the alternating red (strong expression relative to the asynchronous reference) and green (weak expression relative to the asynchronous reference). It is notable that these patterns persist across multiple cell cycles in both Figures 2 and 3.

To minimize unnecessary reassignment of genes relative to the published literature, each gene was assigned a cell cycle phase (G1/S, S, G2, G2/M, or M/G1) by their peak correlation to an idealized expression profile generated from the well-studied genes selected for Figure 2. These assignments are necessarily somewhat arbitrary because a gene assigned to G1/S can be immediately adjacent to a gene assigned to S phase. By assigning phase in this way, 211

(18.6%) of the clones were maximally expressed during the G1/S transition, and 221 (19.5%) were maximally expressed in S phase. Genes in each of these phases have a known role in replication initiation and DNA metabolism. More than 50% of the clones showed peak expression in G2 and M phases of the cell cycle; 239 clones (21.1%) peaked in G2, 273 clones (24.1%) at G2/M, and 190 clones (16.8%) at M/G1. Genes with maximal expression levels in G2 and M phases have roles in chromosome metabolism, surveillance of mitotic processes, and cell adhesion. Interestingly, relatively few genes show peak expression between M/G1 and G1/S, as judged from the distribution of arctangent values calculated from the Fourier transform (Supplemental Figure 15).

In most cases the phase assignments we made for the known genes are in agreement with the published literature (Table 2). Several genes that were previously characterized to peak in a specific cell cycle phase were assigned to an adjacent phase. Examples are CDC45L and CDKN3, both reported to peak at G1/S, whereas in our data they fell into the S phase and M/G1 groups, respectively (Table 2). The difference in assignments may represent differences in experimental systems, methods of measurement, or resolution in the different experiments.

The replication-dependent human histone mRNAs are not polyadenylated (Marzluff, 1992) and thus behave sporadically in all experiments except the mitotic shake-off experiment where total RNA was labeled. Because of this inconsistent behavior, they were incorrectly assigned a phase of G1/S by our method. Because the data obtained from mitotic selection and extensive literature on the cell cycle regulation of the histone genes indicated they are expressed in S phase, the phase has been assigned manually, rather than by correlation to the ideal vector (indicated by an asterisk [\*] in Table 2 and the supplemental data).

### Overview of Periodically Expressed Gene Clusters

Hierarchical clustering of transcripts with similar expression patterns often groups genes according to the processes in which they participate; these processes can often be inferred from the annotations of the known genes in each group (Eisen *et al.*, 1998). To begin to identify the processes that are represented in our clusters, we annotated the known genes in our clusters with gene ontology (GO) terms (Ashburner *et al.*, 2000) for biological process from LocusLink (Pruitt and Maglott, 2001). GO, initially developed using the model organisms *S. cerevisiae*, *Drosophila melanogaster*, and the mouse *Mus musculus*, provides a controlled vocabulary to describe the functions of known genes. GO terms are divided into three different categories: 1) molecular function, 2) biological process, and 3) cellular component. Examination of the lists of clustered genes and their annotations shows that within our set of 1134 clones there are clusters of genes that have similar or identical annotations (Figure 3B). In what follows, GO terms are italicized and terms we have applied provisionally herein because of the incomplete application of GO to human genes are marked with an asterisk (\*). The DNA replication clusters include genes associated with the following process annotations: *DNA replication*, expressed in G1/S (early DNA replication; e.g., the components of the prereplication complex and ORC1); *DNA metabolism* (including its daughter processes *DNA repair* and *DNA recombination*) and *nucleotide metabolism*, expressed in S (late

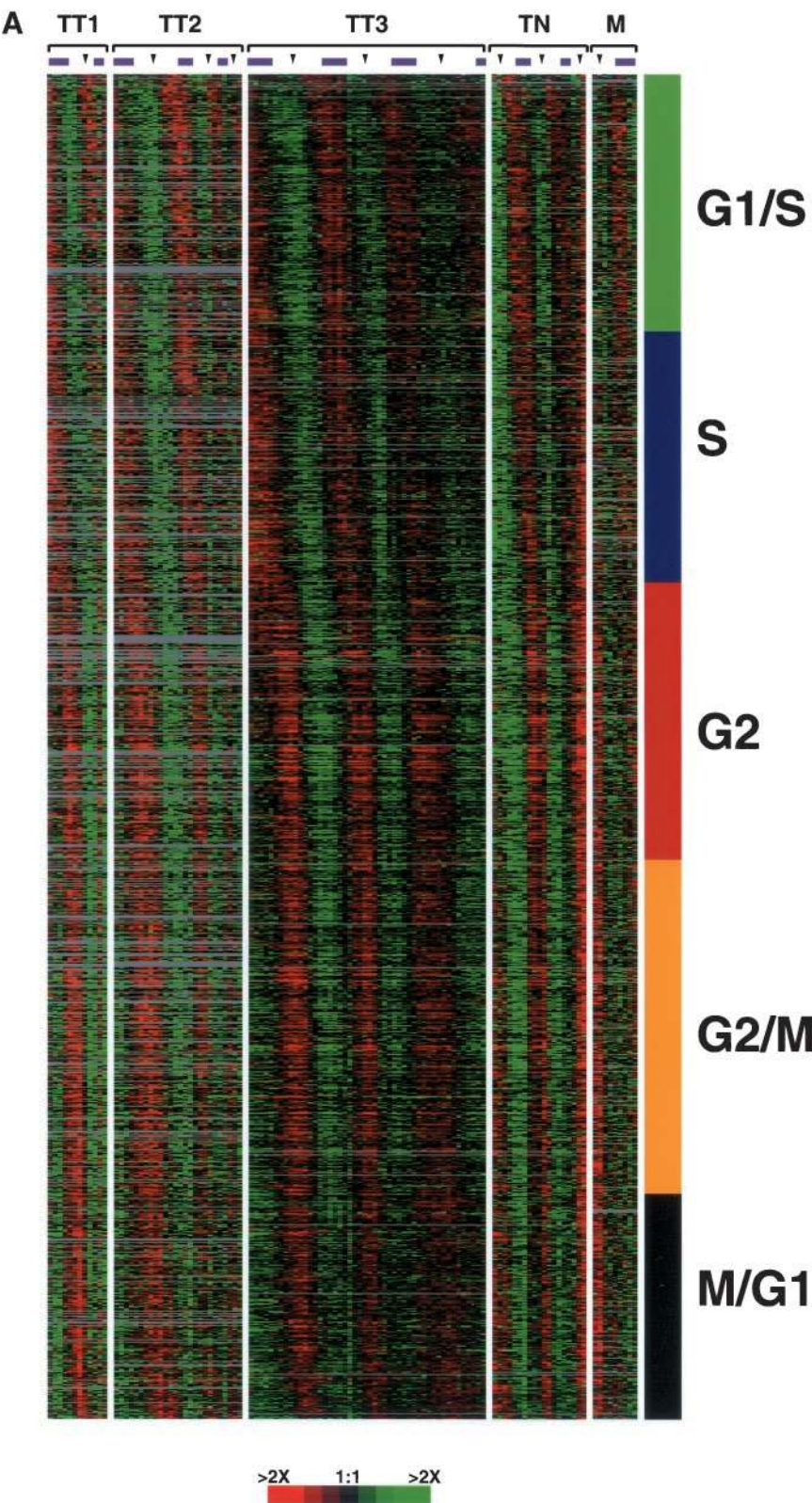
**Table 2.** Known cell cycle-regulated genes

A list of genes was compiled from the literature that had been shown to be cell cycle regulated by traditional methods in cells synchronized independent of serum stimulation. In each case we have listed the primary reference describing the cell cycle regulation of the gene and the predicted cell cycle phase. The cell cycle phase assigned in our study is shown on the right. \*, histone genes were inappropriately assigned to G1/S because they are not polyadenylated and hence behaved sporadically in all experiments except the mitotic selection experiments where total RNA was labeled. Due to extensive literature on the cell cycle regulation of the histone genes and their S phase expression in the mitotic shake-off experiment, the phase was reassigned manually to S phase.

Gene symbol	Accession	Published phase	Reference	Assigned phase
E2F5	NM_001951	G1	(Sardet <i>et al.</i> , 1995; Grosheva <i>et al.</i> , 2001)	G2/M
CCNE1	M74093	G1/S	(Lew <i>et al.</i> , 1991)	G1/S
CCNE2	NM_004702	G1/S	(Zariwala <i>et al.</i> , 1998)	G1/S
CDC25A	NM_001789	G1/S	(Jinno <i>et al.</i> , 1994)	G1/S
CDC45L	NM_003504	G1/S	(Saha <i>et al.</i> , 1998)	S phase
CDC6	NM_001254	G1/S	(Leone <i>et al.</i> , 1998; Yan <i>et al.</i> , 1998)	G1/S
<b>CDKN1A, p21</b>	<b>NM_000389</b>	<b>G1/S</b>	<b>(Hiyama <i>et al.</i>, 1998)</b>	ND
CDKN3	NM_005192	G1/S	(Gyuris <i>et al.</i> , 1993)	M/G1
E2F1	NM_005225	G1/S	(Leone <i>et al.</i> , 1998)	G1/S
MCM2	NM_004526	G1/S	(Leone <i>et al.</i> , 1998)	G1/S
MCM6	NM_005915	G1/S	(Leone <i>et al.</i> , 1998)	G1/S
NPAT	NM_002519	G1/S	(Ma <i>et al.</i> , 2000; Zhao <i>et al.</i> , 2000)	G1/S
PCNA	NM_002592	G1/S	(Morris and Mathews, 1989)	G1/S
SLBP	NM_006527	G1/S	(Whitfield <i>et al.</i> , 2000)	G1/S
BRCA1	NM_007295	S phase	(Rajan <i>et al.</i> , 1996)	S phase
<b>BRCA2</b>	<b>NM_000059</b>	<b>S phase</b>	<b>(Rajan <i>et al.</i>, 1996; Vaughn <i>et al.</i>, 1996)</b>	ND
<b>CCNG2</b>	<b>NM_004354</b>	<b>S phase</b>	<b>(Horne <i>et al.</i>, 1996)</b>	ND
CDKN2C	NM_001262	S phase	(Okuda <i>et al.</i> , 1995)	G2
DHFR	NM_000791	S phase	(Slansky and Farnham, 1996)	S phase
Histone H1	multiple genes	S phase	(Heintz <i>et al.</i> , 1983; Plumb <i>et al.</i> , 1983; Harris <i>et al.</i> , 1991)	S phase*
Histone H2A	multiple genes	S phase	(Heintz <i>et al.</i> , 1983; Plumb <i>et al.</i> , 1983; Harris <i>et al.</i> , 1991)	S phase*
Histone H2B	multiple genes	S phase	(Heintz <i>et al.</i> , 1983; Plumb <i>et al.</i> , 1983; Harris <i>et al.</i> , 1991)	S phase*
<b>Histone H3</b>	<b>multiple genes</b>	<b>S phase</b>	<b>(Heintz <i>et al.</i>, 1983; Plumb <i>et al.</i>, 1983; Harris <i>et al.</i>, 1991)</b>	ND
Histone H4	multiple genes	S phase	(Heintz <i>et al.</i> , 1983; Plumb <i>et al.</i> , 1983; Harris <i>et al.</i> , 1991)	S phase*
MSH2	NM_000251	S phase	(Marra <i>et al.</i> , 1996)	G1/S
NASP	NM_002482	S phase	(Richardson <i>et al.</i> , 2000)	G1/S
RRM1	NM_001033	S phase	(Bjorklund <i>et al.</i> , 1990)	S phase
RRM2	NM_001034	S phase	(Bjorklund <i>et al.</i> , 1990)	S phase
TYMS	NM_001071	S phase	(Kaczmarek, 1986)	S phase
CCNA2	NM_001237	G2	(Pines and Hunter, 1990)	G2
CCNF	NM_001761	G2	(Bai <i>et al.</i> , 1994)	G2
CENPF	NM_016343	G2	(Liao <i>et al.</i> , 1995)	G2/M
TOP2A	NM_001067	G2	(Goswami <i>et al.</i> , 1996)	G2
BIRC5	NM_001168	G2/M	(Li and Altieri, 1999)	G2/M
BUB1	NM_004336	G2/M	(Ouyang <i>et al.</i> , 1998)	G2/M
BUB1B	NM_001211	G2/M	(Davenport <i>et al.</i> , 1999)	G2/M
CCNB1	NM_031966	G2/M	(Pines and Hunter, 1989)	G2/M
CCNB2	NM_004701	G2/M	(Jackman <i>et al.</i> , 1995)	G2/M
CDC2	NM_001786	G2/M	(Dalton, 1992; Zwicker <i>et al.</i> , 1995)	G2
CDC20	NM_001255	G2/M	(Weinstein <i>et al.</i> , 1994; Weinstein, 1997)	G2/M
CDC25B	NM_021874	G2/M	(Jinno <i>et al.</i> , 1994)	G2/M
CDC25C	NM_001790	G2/M	(Lucibello <i>et al.</i> , 1995; Zwicker <i>et al.</i> , 1995)	G2
CDKN2D, p19	NM_001800	G2/M	(Hirai <i>et al.</i> , 1995)	G2/M
CENPA	NM_001809	G2/M	(Shelby <i>et al.</i> , 1997)	G2/M
CKS1	BC001425	G2/M	(Richardson <i>et al.</i> , 1990)	G2
CKS2	NM_001827	G2/M	(Richardson <i>et al.</i> , 1990)	G2/M
PLK	NM_005030	G2/M	(Holtrich <i>et al.</i> , 1994; Lee <i>et al.</i> , 1995)	G2/M
STK15	NM_003600	G2/M	(Shindo <i>et al.</i> , 1998)	G2/M
<b>STK6</b>	<b>NM_003158</b>	<b>G2/M</b>	<b>(Kimura <i>et al.</i>, 1997)</b>	NM
<b>RACGAP1</b>	<b>NM_013277</b>	<b>G2/M</b>	<b>(Hirose <i>et al.</i>, 2001)</b>	NM
<b>RAB6KIFL</b>	<b>NM_005733</b>	<b>G2/M</b>	<b>(Fontijn <i>et al.</i>, 2001)</b>	NM

ND, not detected as cell cycle regulated; NM, not measured in our microassays.





**Figure 3.** Gene expression of the human cell cycle. (A) The 1134 periodically expressed clones are ordered by the point in the cell cycle where they show peak expression, as calculated from the sine and cosine coefficients of the Fourier transform. Each gene was assigned a cell cycle phase based on its correlation to an idealized expression profile defined in Figure 2. Color bars on the right represent the phase in which the majority of the genes are expressed (G1/S, green; S phase, blue; G2, red; G2/M, orange; and M/G1, black); this color scheme is used throughout the remainder of the figures. (B) The 1134 clones that have similar expression profiles are grouped together by hierarchical clustering. The biological theme of each cluster is shown on the right, and the color of the bar indicates the cell cycle phase represented by the genes. The cell adhesion and chromosome metabolism clusters have been labeled a, b, and c as each cluster contains similar GO annotations. Both figures, with all gene names and accession numbers, are available as Supplemental Figure 9 at the Mol. Biol. Cell Web site and at a Web site maintained by the authors (<http://genome-www.stanford.edu/Human-CellCycle/HeLa/>).



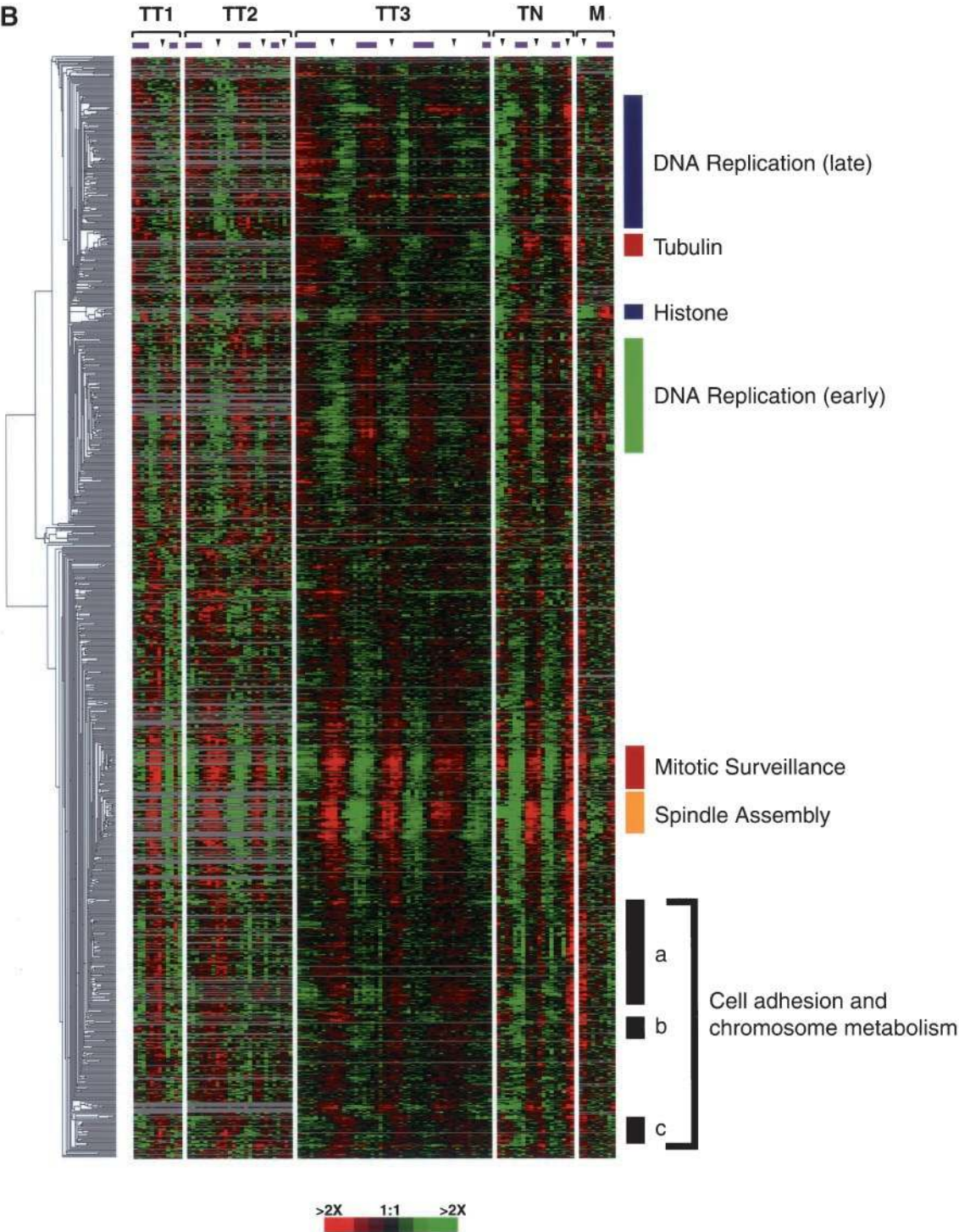


Figure 3 (legend on facing page).

DNA replication; e.g., RAD51, DNA polymerases, and nucleotide metabolism enzymes); and *chromatin assembly/disassembly* (e.g., the histones), also expressed in S. The mitotic clusters include *mitotic chromosome segregation* (the  $\alpha$ - and

$\beta$ -tubulins) expressed in G2; *spindle assembly*, expressed during G2 (e.g., kinesins and centrosome duplications genes); *mitotic checkpoint* and *nuclear division*, expressed at G2/M (e.g., the checkpoint genes BUB1 and CENPE); *sister chroma-*



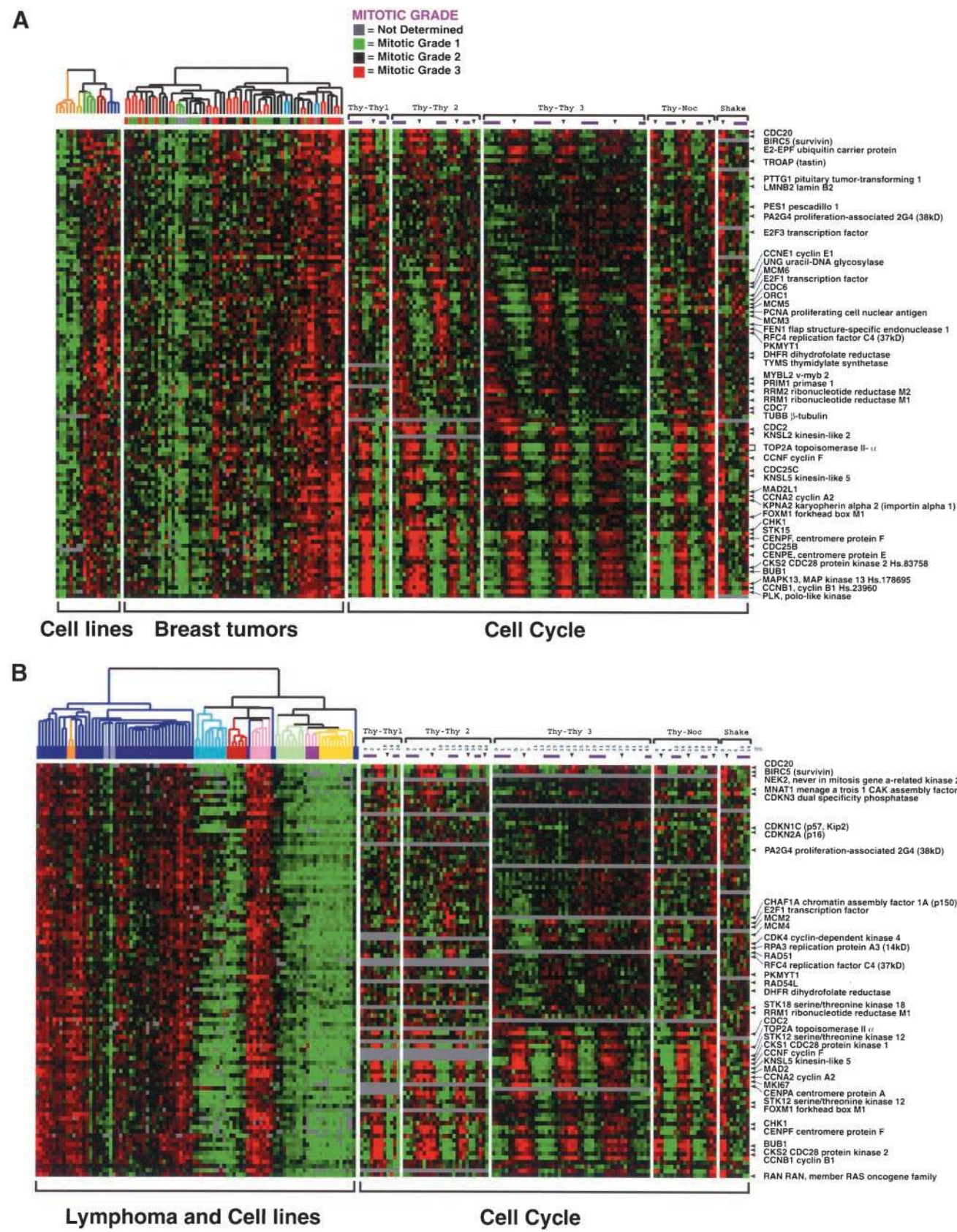


Figure 4.

*tid cohesion* (e.g., PTTG1 and genes of the cohesin complex STAG1, RAD21); and *cell adhesion* (e.g., CTNND1 and vinculin) and components of the RAS signal transduction pathway expressed during the physical act of mitosis, M/G1. The first point to be made about these clusters (shown in more detail in Figures 6–8 and the supplemental data, which contain the complete cluster diagram with gene names) is that each represents a reasonably complete and specific list of the genes necessary for the major, essential processes that must occur every cell division cycle. The second point to be made about these is that, in addition to genes of known function, there are numerous genes of heretofore uncharacterized function that are now implicated in these processes.

### Tumor Proliferation Cluster

Recently, several studies of global gene expression patterns in human tumors and tumor cell lines have been published (Perou *et al.*, 1999, 2000; Alizadeh *et al.*, 2000; Ross *et al.*, 2000). Each of these studies identified a prominent cluster of genes whose expression was correlated with the rate of proliferation of the tumors or cell lines under study. As we show now, half or more of the genes that comprise this “proliferation cluster” are the same genes that we have found to be periodically expressed in HeLa cells.

To compare explicitly the genes of the proliferation cluster with the 874 periodically expressed genes described here, we took the list of genes in each of the proliferation clusters identified in the studies of breast tumors (Perou *et al.*, 2000) and lymphoma (Alizadeh *et al.*, 2000), and extracted the patterns of gene expression for each from the dataset described here. These patterns, lined up by peak expression during the cell cycle as in Figure 3, are presented for 112 genes measured here and in breast cancer (Figure 4A) and

for 96 genes measured here and in lymphoma (Figure 4B). As the figures show, 62% (69/112) of the genes in the breast tumor proliferation cluster and 45% (43/96) of the lymphoma proliferation cluster and are among the 874 periodically expressed genes we detected in synchronized HeLa cell cultures.

The simplest interpretation of this result is that the genes whose transcript levels correlate with the proliferation rate of the tumors are expressed only in cycling cells, presumably as a consequence of regulation that allows transcription only during the appropriate stage of the cell cycle and not at any other time, including quiescence. Even if this idea is correct, it need not apply to all genes that are periodically expressed during the cell cycle. It seems quite possible that some genes that are periodically expressed in cycling cells might not appear in the proliferation clusters of tumors because they might be expressed under some circumstances in noncycling cells or strongly regulated not only by the cell cycle but also by related factors, and thus show little association with tumor proliferation. One might even anticipate that some periodically expressed genes might be regulated in ways (e.g., strong expression in differentiated nonreplicating cells) that result in reduced transcription during active cell division cycles.

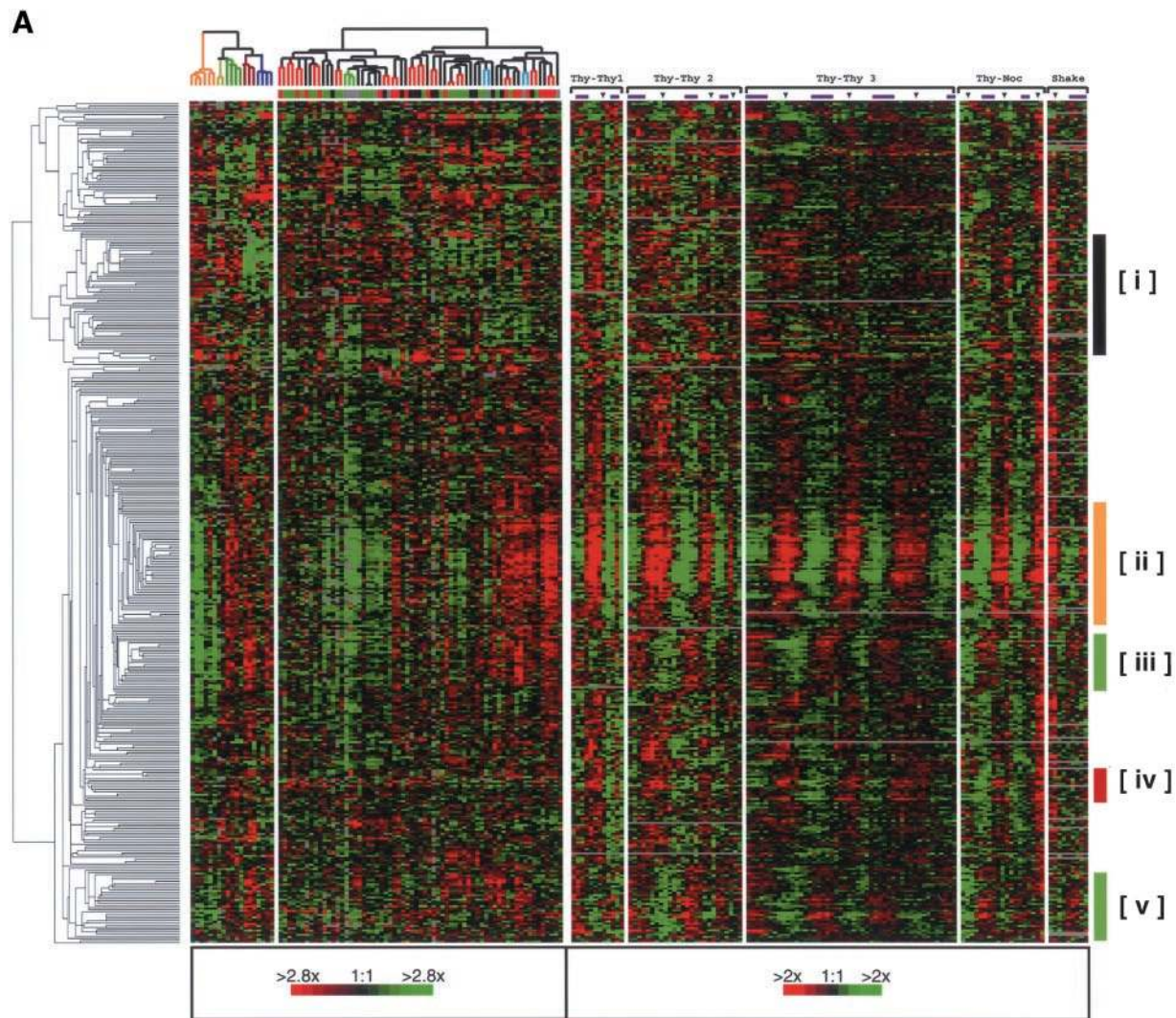
To explore further the relationship between periodic expression during the cell cycle and expression in proliferative tumors, we have extracted the data for every clone that was both identified as periodically expressed in this study and also measured in the breast tumor study, regardless of its pattern there (Perou *et al.*, 2000). This yielded 386 clones, of which <25% appeared in the proliferation cluster. Hierarchical clustering of this set of 386 using the combined data from both studies revealed, as anticipated, not only periodically expressed genes that are highly correlated with tumor proliferation (a more comprehensive proliferation cluster, as described above; Figure 5B, ii and iii) but also periodically expressed genes that show heterogeneous expression in tumors (i.e., no obvious association with tumor proliferation; Figure 5B, v), and periodically expressed genes whose expression is apparently uncorrelated with the proliferative state of the tumors (Figure 5B, i and iv).

Correlation between periodic expression in the cell cycle and strong expression in highly proliferative tumors can be accounted for simply because there is an increase in the number of cycling cells (as described above), but other more specific explanations can also be proposed. Many genes that we have found to be periodically expressed have already been implicated directly in tumorigenesis and their expression has been reported to coincide with cellular transformation and oncogenic potential (reviewed in Bishop, 1991; Hunter, 1997). Genes of this kind are common in clusters whose expression is highly correlated with tumor proliferation in our study, including some expressed primarily in the G2 and M phases (Figure 5B, ii) and others whose expression peaks in G1 and S phases (Figure 5B, iii).

Among the periodically expressed genes whose expression is heterogeneous or uncorrelated with tumor proliferation are genes for a variety of cell cycle processes, including genes for cell-cell adhesion (calponin-2, smoothelin, and vinculin) (Figure 5B, i). Expression of some of these genes, such as vinculin, in transformed cells has been reported to decrease tumorigenicity (Rodriguez Fernandez *et al.*, 1992).

**Figure 4 (facing page).** Genes associated with the proliferative state of tumors are cell cycle regulated. (A) Cell cycle regulation of the breast tumor proliferation cluster. The expression patterns of the 112 clones in the breast tumor proliferation cluster (Perou *et al.*, 2000) were examined in the cell cycle dataset, along with their expression in breast tumors and cell lines. The clones are ordered by their peak expression during the cell cycle (as in Figure 3A) such that each row corresponds to a single clone. The order of the breast tumor samples based on hierarchical clustering by using 1753 genes is as described previously (Perou *et al.*, 2000). The mitotic grade of the tumors is indicated (a higher mitotic grade is indicative of a more proliferative tumor). Of the 112 genes in the proliferation cluster, 69 were identified as cell cycle regulated in this study. (B) Cell cycle regulation of lymphoma proliferation cluster. Because the lymphoma study was performed on the Lymphochip Microarray, which is comprised of a different clone set than the Stanford Human Microarray used for this study, the genes on the two arrays were linked via their LocusLink ID (Pruitt and Maglott, 2001). We selected 154 clones from the previously published lymphoma proliferation cluster (Alizadeh *et al.*, 2000), and 19 clones without LocusLink IDs were discarded. Of the remaining 135, 128 clones with matching LocusLink IDs were measured in our cell cycle study. Duplicate clones were removed leaving 96 genes. Of these, 43 were identified as cell cycle regulated in this study. As above, the genes (i.e., rows) are ordered by peak expression during the cell cycle. The order of the tumors, based on hierarchical clustering by using a larger data set, is as published previously (Alizadeh *et al.*, 2000). This figure with all gene names is available as Supplemental Figure 10.





**Figure 5.** Expression of cell cycle-regulated genes in breast tumors. The tumor samples have been ordered by hierarchical clustering using 1753 genes as described previously (Perou *et al.*, 2000). (A) Expression profiles are shown of 386 of the 1134 cell cycle-regulated clones that were also present in the breast cancer study (Perou *et al.*, 2000). The clones are ordered by their similarity of gene expression profiles across both data sets by hierarchical clustering. (B) Detailed views of the clusters identified in Figure 5A. A subset of genes is expressed at lower levels in the more proliferative tumors (NODE379X, cluster i; NODE303X, cluster iv) and includes genes that peak at M/G1, S, and G2 phases. Genes for core cell cycle processes that peak at G2 and M phases (NODE142X, ii) or at G1 and S phases (NODE133X, iii) are highly expressed in the more proliferative tumors; many of these genes are the same genes found in the breast tumor proliferation cluster. A subset of cell cycle regulated genes is expressed in a heterogeneous manner (NODE345X, v). An enhanced figure with all gene names is available as Supplemental Figure 11.

Another group of cell cycle-regulated genes that are not correlated with tumor proliferation includes genes necessary for withdrawal from the cell cycle and apoptosis (the cyclin-dependent kinase [cdk] inhibitor CDKN2C (p18) and Caspase 3; Figure 5B, iv). Finally, a group of G1 and S phase genes shows heterogeneous expression in tumors (Figure 5B, v), including DHFR, E2F1, CDC6, CCNE1, CHAF1A, TOP3A, ORC1, and BRCA1; these data suggest that the regulation of each of these genes is more complex than simple restriction of transcription to a particular phase of the cell division cycle.

### Biological Roles of Periodically Expressed Gene Clusters

The close similarity between the set of genes whose expression is associated with tumor proliferation and those that are periodically expressed invites a closer examination of the biological roles of the gene clusters detected in Figure 3B, in light of the specific features of both the cell cycle and tumor biology. In what follows, we survey roles of the genes of known function in each of the prominent clusters. It should be recalled, however, that each of these clusters contains, in



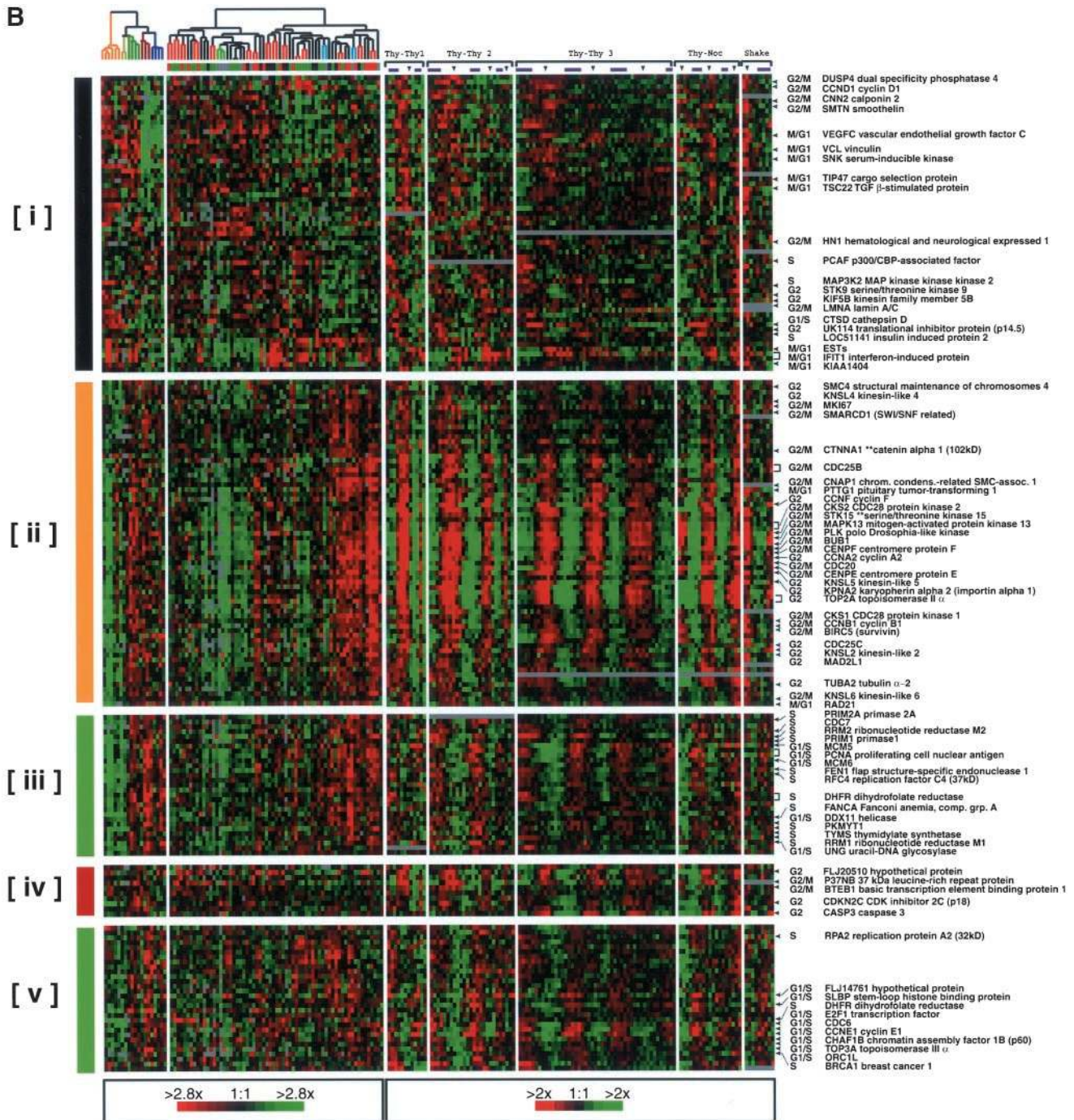


Figure 5 (legend on facing page).

addition to the genes of known function, genes whose biological role remains to be determined (expressed sequence tags; ESTs). The numbers of such uncharacterized genes are noted in the legends to Figures 6–8 and, of course, the detailed position of each of them can be found in the supplementary information and on the Web site.

### G1 and S Clusters

The “early DNA replication” cluster contains genes expressed in late G1 phase with their expression continuing into S phase. The GO process annotations that apply to most of these genes are *DNA replication* (including well-character-



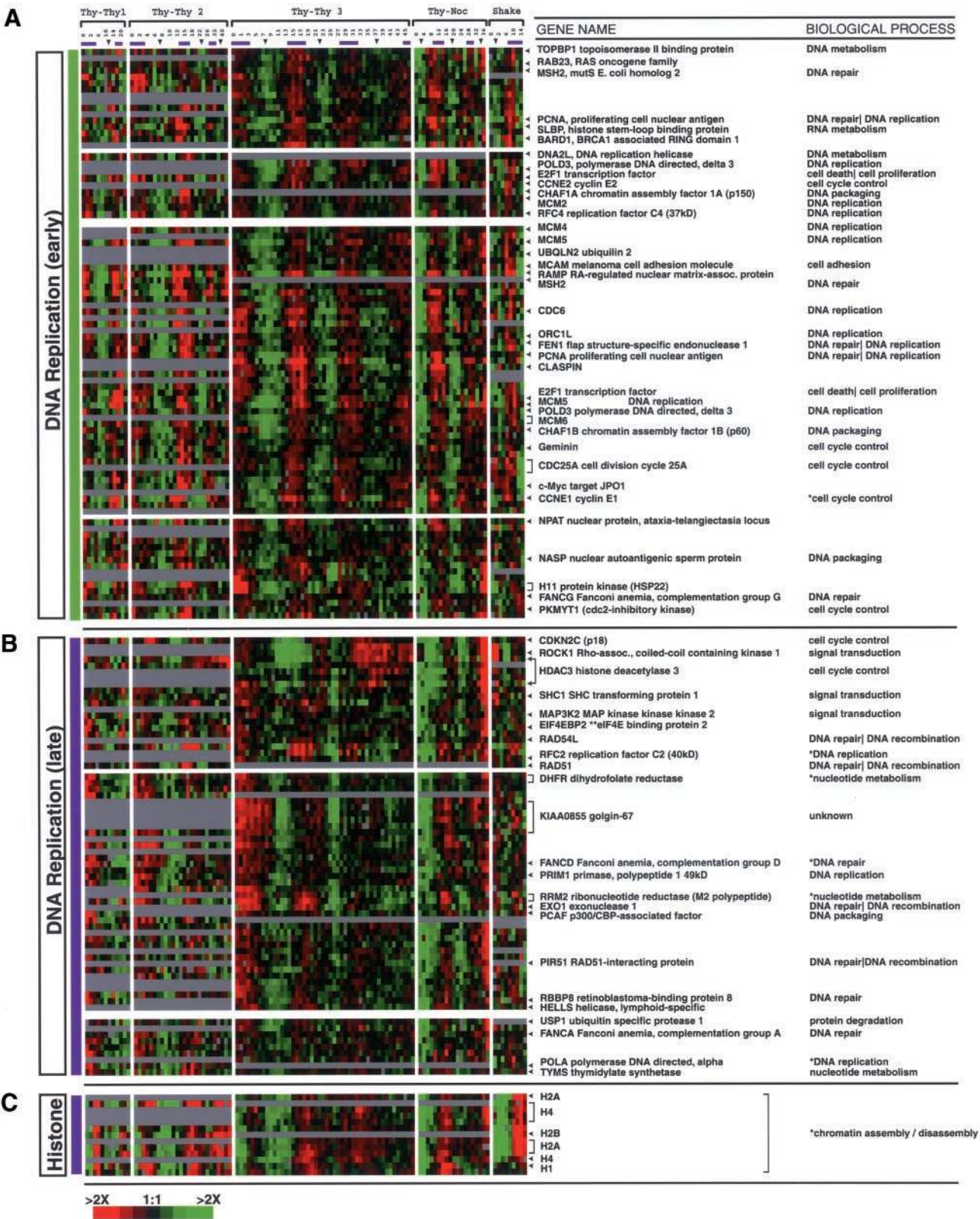


Figure 6.



ized genes encoding RFC4, DNA polymerase delta 3, DDX11, and geminin), *DNA packaging* (including CHAF1A, CHAF1B, and PCAF), and *DNA repair* (including MSH2, FEN1, and PCNA). Many of the genes in the DNA replication category are components of prereplicative complex and include CDC6, ORC1L, MCM2, MCM4, MCM5, and MCM6 (Lei and Tye, 2001); this group of genes is consistently expressed at relatively high levels in diverse types of tumors and has been proposed as diagnostic markers by Laskey and coworkers (Williams *et al.*, 1998). The cluster also includes genes that fall into the GO category *cell cycle control* and are necessary for entry into S phase (in this group are Cyclin E1, E2F1, CDC25A, and Cyclin E2).

The "late DNA replication" cluster (Figure 6B) contains genes necessary for the continued synthesis of DNA, including *nucleotide metabolism* (including TYMS, DHFR, RRM1, and RRM2) and additional *DNA replication* genes (DNA polymerase alpha and theta; Primase 1 and 2A). Also present are genes for *DNA repair* and *DNA recombination* (RAD54 and RAD51). Genes involved in DNA repair were also expressed in S phase in a study of the cell cycle in primary human fibroblasts, and it was reported that these genes were also induced when cells were treated with UV radiation (Cho *et al.*, 2001).

Histone synthesis is required during DNA replication to package newly synthesized chromatin and is tightly regulated during the cell cycle (Marzluff and Pandey, 1988; Schumperli, 1988). The replication-dependent histone genes are coordinately regulated and the expression of their mRNAs is tightly restricted to S phase of the cell cycle by both transcriptional and posttranscriptional mechanisms (Harris *et al.*, 1991; Eliassen *et al.*, 1998). The "histone" cluster (Figure 6C), expressed during S phase, contains the core histones H2A, H2B, and H4 and the linker histone H1, which we have assigned to the GO category of *chromatin assembly/disassembly*. These mRNAs are not polyadenylated and hence are measured poorly and inconsistently in all experiments except mitotic selection where total RNA was labeled. It should be noted that the several H2A genes will cross-hybridize, as will the H2B, H4, and H1 genes, but significant cross-hybridization between the different genes (e.g., H2A and H2B) is unlikely. Histone H3 mRNA, although cell cycle-regulated, was not found in our study, possibly be-

cause of preferential labeling of the polyadenylated, non-cell cycle-regulated variant H3.3 (Wells and Kedes, 1985).

Two regulators of mammalian histone mRNA synthesis are present in the G1/S cluster (Figure 6A), the stem-loop binding protein (SLBP) necessary for histone pre-mRNA processing (Wang *et al.*, 1996a; Dominski *et al.*, 1999), and nuclear protein mapped to the AT locus (NPAT), which was recently identified as an activator of histone gene transcription and is hypothesized to be a chromatin remodeling factor (Ma *et al.*, 2000; Zhao *et al.*, 2000). The SLBP was previously shown to be cell cycle regulated by both translational and posttranslational mechanisms; the magnitude of the changes observed herein in the mRNA level (approximately twofold) is consistent with Whitfield *et al.* (2000).

Many of the genes expressed at G1/S and S phase are known E2F targets. A study of the cell cycle and the E2F transcription factors in mouse embryo fibroblast, using microarrays, identified both G1/S and S phase genes, as well as genes expressed at G2 and M phases, as targets of the E2F transcription factor (Ishida *et al.*, 2001). Ishida *et al.* (2001) identified the histone SLBP as induced by E2F, suggesting that histone synthesis may be linked to the Rb/E2F pathway. Furthermore, NPAT is a cyclin E/cdk2 substrate and its overexpression promotes S phase entry (Zhao *et al.*, 1998). Although NPAT itself has not been implicated in tumorigenesis, the role of chromatin remodeling in cancer is becoming increasingly clear (Archer and Hodin, 1999).

## G2 and M Clusters

The mitotic genes form three distinct clusters, one of which includes genes that encode  $\alpha$ - and  $\beta$ -tubulins; the other two contain well-characterized genes that encode proteins involved in assembly and surveillance of the mitotic spindle (Figure 3B). The remaining G2 and M phase genes that fall outside of these core clusters also have functions concerned with the GO processes *mitosis*, and its daughter processes (e.g., *mitotic chromosome segregation*, *spindle assembly*, *mitotic checkpoint*, etc).

The tubulin cluster includes  $\alpha$ - and  $\beta$ -tubulins (TUBA1, TUBA2, TUBA3, TUBB, and TUBB2) as well as BUB3, which is a mitotic checkpoint regulator (Figure 7A). (Note that there may be cross-hybridization among the several  $\alpha$ -tubulin genes or between the several  $\beta$ -tubulin genes, but cross-hybridization between  $\alpha$ - and  $\beta$ -tubulin genes is unlikely.) Coordinate regulation of tubulin synthesis at the level of mRNA abundance is evident herein, even though a large component of the regulation of these genes is posttranscriptional (Cleveland, 1989). Yeast BUB3 was isolated as a multicopy suppressor of a conditional mutation in  $\alpha$ -tubulin (Guenette *et al.*, 1995), the coexpression of hBUB3 with the coordinately regulated tubulins is consistent with its role of monitoring microtubule assembly as cells enter mitosis.

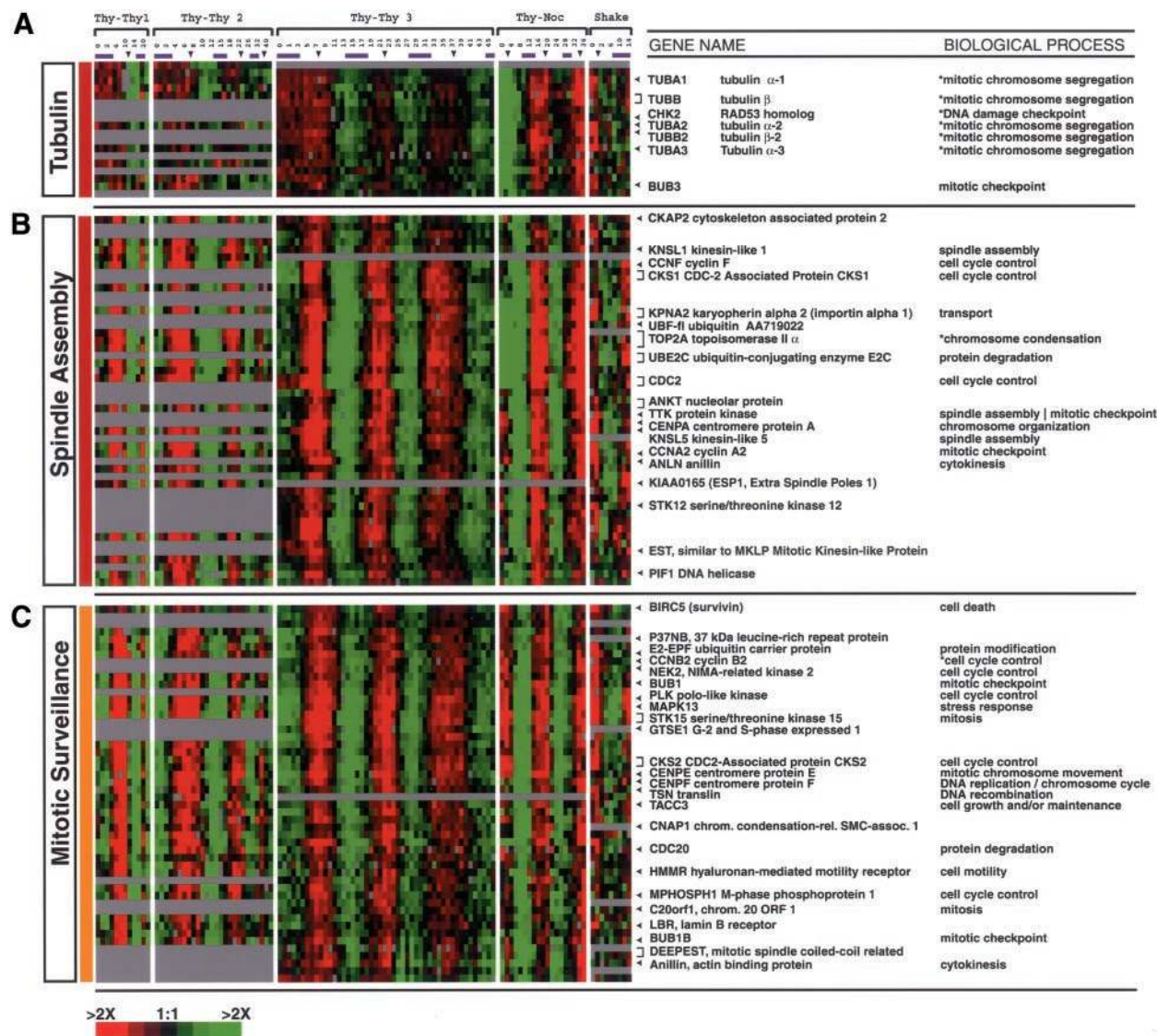
The second G2 phase cluster contains genes predominantly for organization of the mitotic spindle, including kinesins and the TTK kinase (human homolog of the *S. cerevisiae* MPS1 spindle checkpoint gene; Figure 7B; Abrieu *et al.*, 2001). Five different kinesins are present: KNSL1, KNSL2, KNSL4, KNSL5, and KNSL6. KNSL6 is expressed at the boundary between G2 and M phase genes and in our study has been assigned a phase of G2/M. In addition to Cyclin A2, CDC2, CKS1, Cyclin F, and numerous other regulatory factors are found in this cluster. Also specifically

**Figure 6 (facing page).** G1 and S phase clusters. Shown are detailed views of the cell cycle clusters indicated in Figure 3B. Only a portion of each cluster is shown; the full cluster with all gene names, phase assignments, and annotations can be viewed in Supplemental Figure 12. The names of selected genes and their corresponding rows (indicated by an arrow) are shown to the right. Some genes had multiple clones on the arrays that clustered together (e.g., CDC25A and Golgin-67) and are indicated by a bracket. The GO biological process annotation is shown for each gene. Due to the incomplete application of GO to the human genome, we have assigned the biological process to some genes provisionally; an asterisk (\*) precedes these terms. (A) Early DNA replication cluster (NODE1077X) contains 202 clones (94 ESTs and hypothetical proteins); 88% have been assigned a phase of G1/S. (B) Late DNA replication cluster (NODE793X) contains 139 clones (66 ESTs and hypothetical proteins); 80% show peak expression in S phase. (C) Histone gene cluster (NODE 516X) contains 13 clones expressed during S phase; 12 of the clones are histone genes.

expressed in G2 is ESP1 or human Separin, which is released from PTTG1 (human homolog of the yeast securin PDS1) during anaphase and promotes chromatid separation (Zou *et al.*, 1999). Finally, it may be worth noting that Importin alpha, also in this cluster, has recently been shown to inhibit spindle assembly by binding a protein that promotes spindle formation, Repp86 (*Xenopus* TPX2) (Gruss *et al.*, 2001; Nachury *et al.*, 2001; Wiese *et al.*, 2001). This inhibition is relieved by the action of the small GTPase RAN when bound to GTP (present in the M/G1 cluster; Figure 8), which is proposed to surround M phase chromosomes (Carazo-Salas *et al.*, 1999; Hetzer *et al.*, 2000), thus allowing spindle forma-

tion only in the vicinity of chromosomes. The temporal succession of the transcripts for these two proteins suggests a possible role in regulating the mitotic spindle.

The third mitotic cluster contains genes that peak at G2/M and centers on Cyclin B2 (Figure 7C); Cyclin B1 falls just outside this cluster. Genes that have known functions in the mitotic spindle checkpoint show peak expression at G2/M, including BUB1, BUB1B, CDC20, and CENPE. Three genes in this cluster (STK15, PLK1, and NEK2) have been shown to have roles in centrosome duplication, whose improper expression is a potential cause of tumor aneuploidy (Lengauer *et al.*, 1998). It is worth noting that PLK1 and STK15 had the



**Figure 7.** G2 and M phase clusters. Clusters of genes from Figure 3B involved in mitotic processes are shown in more detail. The annotation is as described in Figure 6. (A) Tubulin gene cluster (NODE481X) contains 18 clones (six ESTs) that peak at G2 of the cell cycle. The coordinate expression of the  $\alpha$ - and  $\beta$ -tubulins is notable. (B) Spindle assembly cluster (NODE180X) contains 49 clones (22 ESTs); 88% have been assigned a phase of G2, and many of the known genes are involved in establishment and maintenance of the mitotic spindle. (C) Mitotic surveillance cluster (NODE191X) contains 50 clones (13 ESTs) that primarily peak in G2/M. The full clusters with all gene names, phase assignments, and annotations can be viewed in Supplemental Figure 13.



very highest periodicity scores (Table 1), and thus are the most periodically synthesized mRNAs among the cell cycle-regulated transcripts.

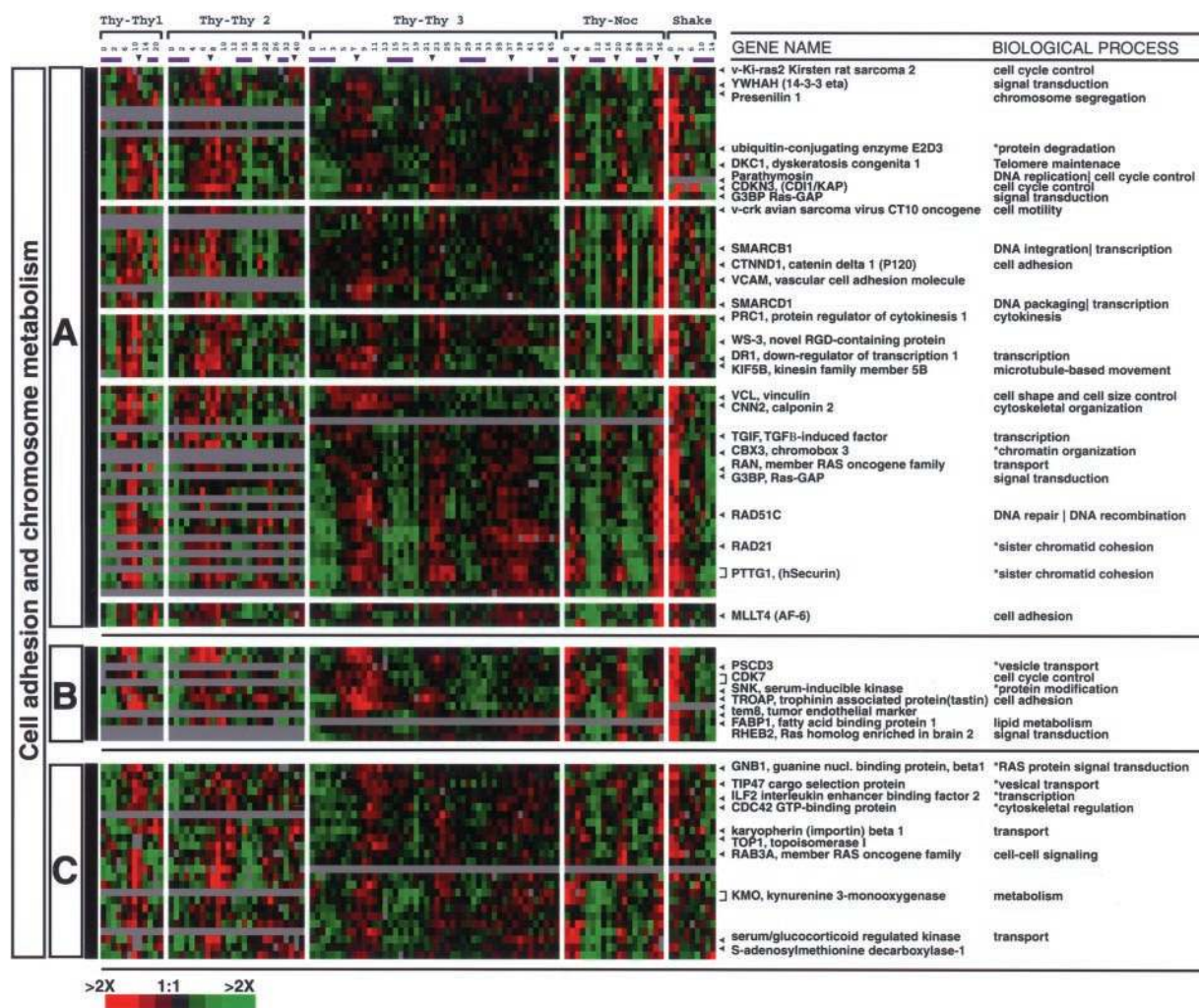
### M/G1 Clusters

Three groups of genes whose expression peaks in mitosis merit a more detailed examination. We classified these genes as M/G1 because they are expressed during the physical act of mitosis and their transcripts persist into G1 phase. Some of these genes are known to have roles in cell adhesion, chromosome remodeling, and membrane trafficking as opposed to roles specifically associated with the progress of the cell division cycle.

The M/G1 cluster contains genes known to function in actin cytoskeleton remodeling and *cell adhesion* (Figure 8). As

a synchronous population of cells proceeds through the cell cycle, visible changes occur to the cellular morphology. The most notable of these changes is the dramatic remodeling of cell shape that can be observed in cell culture during mitosis, when cells detach from the plate and surrounding cells, undergo cytokinesis, and reattach to the plate as they enter the next cell cycle (Schneiderman *et al.*, 1972). A cluster of M/G1 genes is expressed during this transition and contains genes whose products have been localized to the structures that connect the actin cytoskeleton to the membrane or have been implicated in the control of cell-cell contacts.

The Ras GTPase can induce perturbations in cell-cell contacts when overexpressed in its constitutively activated form (Kinch *et al.*, 1995; Zhong *et al.*, 1997; Yamamoto *et al.*, 1999). Included in the M/G1 cluster is Kirsten Ras-2 (KRAS2),



**Figure 8.** M/G1 clusters. Three clusters of M/G1 genes from Figure 3B are shown in more detail. Each contains genes involved in cell adhesion, chromosome metabolism and RAS signal transduction and is annotated as described in Figure 6. Cluster A, (NODE568X) contains 104 clones (33 ESTs and hypothetical proteins); 37% show peak expression in G2/M and 63% show peak expression in M/G1. Cluster B, (NODE688X) is comprised of 12 clones (three ESTs), including CDK7, the cdk of the cdk-activating complex (CAK); 92% show peak expression in M/G1. Cluster C, (NODE988X) which peaks in late M/G1, contains 25 clones (eight ESTs and hypothetical proteins); 96% peak in M/G1. The full clusters with all gene names, phase assignments and annotations can be viewed in Supplemental Figure 14.



regularly mutated in many malignancies, including lung adenocarcinoma, mucinous adenoma, ductal carcinoma of the pancreas, and colon tumors (Barbacid, 1990) (Figure 8A). Ras is believed to act through downstream targets, including CDC42 (which appears in this cluster; Figure 8C) as well as the Rho GTPases (which we do not detect as cell cycle-regulated). These GTPases regulate both the assembly of actin filaments and the assembly of the adhesion complexes. Among the M/G1 genes that have been localized to cell-cell adherens junctions are catenin delta 1 (CTTND1, also known as p120), which regulates cadherin clustering at adherens junctions through Rho GTPases (Anastasiadis and Reynolds, 2001); presenilin 1, which binds E-cadherin at cell-cell junctions (Georgakopoulos *et al.*, 1999) and has also been shown to localize to the nuclear membrane, interphase kinetochores, and centrosomes, suggesting a function in chromosome segregation (Li *et al.*, 1997); vinculin, involved in attaching actin to the membrane (Jockusch and Isenberg, 1981; Wilkins and Lin, 1982); calponin 2 (CNN2), which contains an actin binding site and is localized to adherens junctions with vinculin (Masuda *et al.*, 1996); and MLLT4 (known also as AF-6 or afadin), which is a component of tight junctions (Yamamoto *et al.*, 1997). Taken together, these results suggest that the regulation of transcripts during the cell cycle plays a role in orchestrating the morphological changes that occur during mitosis. Expression of these genes during the physical act of mitosis may be necessary to prepare the cell to reestablish contact with the surrounding milieu.

Genes involved in chromosome architecture and remodeling are also expressed at M/G1 (Figure 8A). These include presenilin 1, discussed above, and parathymosin, which is localized to sites of early DNA replication (Vareli *et al.*, 2000) and binds to the linker histone H1 (Kondili *et al.*, 1996). The dyskeratosis gene (DKC1) is implicated in telomere maintenance, binds to H/ACA class of snoRNPs, and has been shown to associate with the telomerase RNA that acts as a template to add sequence repeats to the ends of chromosomes (Mitchell *et al.*, 1999). Two members of the SWI/SNF complex (SMARCB1 and SMARCD1), present in the M/G1 cluster, function in chromatin remodeling and transcriptional regulation (Wang *et al.*, 1996b; Muchardt and Yaniv, 2001). Other M/G1 genes with roles in chromosome architecture and remodeling include human securin (PTTG1), and RAD21, both genes involved in sister chromatid cohesion (Zou *et al.*, 1999; Hoque and Ishikawa, 2001).

Finally, expression of CDK7 (MO15) the catalytic subunit of the CDK-activating kinase (CAK) also peaks at M/G1 (Figure 8B). CDK7 is a catalytic subunit of CAK, necessary for the phosphorylation and full activity of CDK2 and CDC2 (Fisher and Morgan, 1994). CAK activity and CDK7 protein have been reported to be constant during the cell cycle (Matsuoka *et al.*, 1994; Tassan *et al.*, 1994) although data on the mRNA regulation have not previously been reported. The CAK cyclin subunit Cyclin H (Fisher and Morgan, 1994) does not oscillate in our dataset even though we do see oscillation of the mRNA for the mitotic cyclins.

In summary, we identified 874 genes that show periodic expression in the cell division cycle of a human cell line. We found a relationship between genes associated with proliferation of tumors and those that are periodically expressed during the cell division cycle. Many of the genes that are periodically expressed have well-characterized functions as-

sociated with the cell division cycle, and a large number, >450 clones, are either uncharacterized ESTs or hypothetical proteins whose pattern of regulation during the cell division cycle now points to specific directions for further investigation.

## DISCUSSION

We identified 874 genes that show periodic expression across the human cell cycle in a well-studied cancer cell line (HeLa). This system was chosen because a high degree of synchrony and low background of noncycling cells can be obtained by multiple methods. Herein, 1134 clones (representing 874 UNIGENE clusters) passed a minimum set of objective quantitative criteria for a periodically expressed gene. Although both the Rb and P53 tumor suppressors are inactivated in HeLa cells as a result of binding of the E6/E7 proteins of the human papillomavirus (Scheffner *et al.*, 1991), genes involved in basic cell cycle processes such as DNA replication, chromosome segregation, and cell adhesion still showed periodic expression.

Each gene was assigned a cell cycle phase as described above (Figures 2 and 3). Approximately 38% of the genes identified as cell cycle regulated were assigned a phase of G1/S or S phase, whereas 45% were assigned a phase of G2 or G2/M. In mammalian cells, G2 and M phases are short (4 h of each 14–16-h cell cycle) relative to G1 and S phase, yet almost half of the periodically expressed genes peak during this interval. One potential explanation is that we were more able to detect genes expressed at G2 and M phases because three of our five synchronization methods arrest cells in S phase, which provides more robust synchrony in G2 and M phases. This possibility now seems particularly unlikely because a similar distribution is observed in the proliferation clusters, which have now been observed in a variety of tumors independently.

We recognize that our microarrays do not completely represent all expressed genes in the human genome, limiting the completeness of our survey. For example, three of the 49 genes known from the literature to be cell cycle-regulated were not represented on our arrays (see above and Table 2). Subject only to these omissions, we conclude, based on the similarities found between cells synchronized in S phase, cells synchronized in M phase, cells selected by mitotic shake-off, and the tumor proliferation clusters, that we have identified a comprehensive list of cell cycle-regulated human genes.

### Comparison to Previous Surveys of Periodically Expressed Genes in Animal Cells

Genome-wide transcriptional profiles of the cell cycle in animal cells have been reported previously from experiments carried out on fibroblasts and epithelial cells. A study of the response of primary human fibroblast to serum revealed not only a cell cycle response but also an equally prominent wound-healing response (Iyer *et al.*, 1999). This made clear the shortcomings of serum as a synchronizing agent if the goal is to find periodically expressed genes. Herein, we have avoided the use of serum "starvation" entirely.

A study of the cell cycle and E2F transcription factors in mouse embryo fibroblasts identified 578 cell cycle-regulated genes and E2F targets (Ishida *et al.*, 2001). Many of the genes identified in the mouse study were also identified as cell cycle-regulated in HeLa cells herein. Among these are TYMS, RRM1, RRM2, Cyclin E1, MCM3, MCM7, PCNA, TOP2A, FEN1, RAD51, SLBP, Cyclin A2, Cyclin B1, Cyclin B2, KI-67, and importin alpha 2. Those identified as E2F targets but not identified as cell cycle regulated in our study include thymidine kinase 1 (TK), CDK2, DNA ligase I, and RB. It should be noted that not all E2F targets are expected to be cell cycle regulated in our study of HeLa cells because some genes show regulation primarily during the resting to growing transition rather than in a continuous cell cycle.

A study of the G2 DNA damage checkpoint in HeLa cells revealed a delay in the expression of mitotic genes when the cells are exposed to ionizing radiation (e.g., Cyclin B1, CKS2, TTK, STK15, CDC20, and CENPA) (Crawford and Piwnicka-Worms, 2001). Many of these genes were classified as G2/M in our study. It is worth pointing out that similar results were obtained herein for G2 genes when cells are arrested in mitosis by nocodazole; G1/S, S, and G2 genes show relatively low expression, whereas G2/M and M/G1 genes show relatively high expression (e.g., 0 h Thy-Noc samples in Figures 2 and 6–8). The opposite is observed in cells arrested in S phase with thymidine; relatively high expression of G1/S and S phase genes, and relatively lower expression of G2, G2/M, and M/G1 genes (Figures 2 and 6–8). The exception is the replication-dependent histone mRNAs, which, although expressed during S phase, are rapidly and coordinately degraded when DNA synthesis is inhibited (Harris *et al.*, 1991). Together, these observations suggest that the gene expression program observed during a cell cycle arrest is strongly influenced by the point in the cell cycle where the arrest occurs; the genes normally expressed in that phase are usually but not always, relatively overexpressed.

Another study of the cell cycle in primary human fibroblasts synchronized by a double thymidine block, identified ~700 cell cycle-regulated genes (Cho *et al.*, 2001) that we have mapped to 595 unique UNIGENE clusters. Surprisingly, only 96 genes were identified as cell cycle-regulated in both studies (these comparisons are available in the supplemental data). Of the remaining 499 genes that were identified as cell cycle-regulated by Cho *et al.* (2001), but not in this study, only 50 were not measured on our microarrays. Of the 778 genes identified as cell cycle regulated in this study, but not by Cho *et al.* (2001), only 109 were not measured on their microarrays. We have no ready explanation of this difference in results, except to note that there are differences in the cell lineage, the microarray technology, and in the analysis methods. As we have shown, 69 of the genes that were cell cycle-regulated in our study are nevertheless associated with the proliferative state of tumors and cell lines, whereas only 29 of the genes reported to be cell cycle-regulated by Cho *et al.* (2001) appear in the tumor proliferation clusters. We suspect that the most significant differences may well be in the degree of cell synchrony achieved and the percentage of cells that reentered the cell cycle after removal of the synchronizing block. Alternatively, HeLa cells, derived from a cervical carcinoma, may more closely resemble cancers in their cell cycle program of gene expres-

sion and differ significantly from the corresponding program in a normal fibroblast.

Summing up the comparisons with previous work, we believe that we have detected most of the genes expressed periodically through the HeLa cell division cycle. It will be important to compare this list of 874 genes to similarly comprehensive studies of more normal human cell types in the future.

### Comparison to Yeast Cell Cycle

Comparison of the results of this study with our previous study of the yeast cell division cycle (Spellman *et al.*, 1998) produced interesting results. When the “well-characterized” (i.e., those that encode curated protein sequences in RefSeq) and periodically expressed human gene sequences were compared with the yeast genome sequence, ~18% (155 genes) had putative orthologs (Ball, Whitfield, and Botstein, unpublished data). Of these, only ~26% (41) were clearly cell cycle regulated in yeast (Spellman *et al.*, 1998), suggesting regulation at the level of transcription is not necessarily a conserved property. An example of a gene periodically expressed in human cells but not in yeast is the human CDC2 kinase (homolog of *S. cerevisiae* CDC28 kinase). Yeast CDC28 is not regulated at the transcriptional level but rather the kinase activity of CDC28 is regulated (Lee and Nurse, 1987). It is likely that many of the periodically expressed genes in human cells that do not have periodically expressed counterparts in yeast are subject to multiple layers of regulation (e.g., phosphorylation and proteolysis) as is already known for many well-studied cell cycle genes. Among the genes periodically expressed in both species are many involved in basic processes such as DNA replication, repair, and metabolism, and mitosis. However, as noted by Spellman *et al.* (1998), many of the genes that were periodically synthesized in yeast are not obviously involved in these basic processes but instead are genes involved in bud emergence and bud growth, which take place at particular points in the cell cycle because of the particular biology of budding yeasts. Likewise, it may be that many genes are periodically expressed in an animal cell for reasons that do not apply to yeast. Spellman *et al.* (1998) reasoned, partly from this observation, that parsimony is likely to explain much of the observed cell cycle regulation; genes are periodically expressed primarily because there is special need for the gene products at particular points in the cell cycle. For human cells, the periodic synthesis of the cytoskeleton and cell adhesion-associated genes is a particularly interesting case in point. The overall shape of an animal cell generally changes, and tends to lose contact with its substrate, when it forms the mitotic spindle and undergoes cell division. Reasoning in parallel to the yeast example, we suppose that animal cells might periodically synthesize cytoskeletal genes and adhesion factors because of the particular need to change shape in mitosis, and subsequently reattach to substrate, a need not shared by the yeasts.

### Cell Cycle and Cancer

Comparison of the 874 genes that we identify as cell cycle regulated in HeLa cells to studies of human breast tumors shows that almost 70% of the transcripts in the proliferation cluster are cell cycle regulated. We showed above that many

genes involved in basic cell cycle processes are also more highly expressed in more proliferative tumors. However, the correlation we found may reflect only the reality that proliferative tumors contain a large proportion of actively cycling cells.

There are nevertheless periodically expressed genes (e.g., KRAS2, PTTG1, STK15, and PLK) (Der *et al.*, 1982; Pei and Melmed, 1997; Smith *et al.*, 1997; Zhou *et al.*, 1998) that are likely to contribute more directly to tumor phenotypes. For example, the two highest scoring genes in our study, STK15 and Polo-like kinase (PLK), show peak expression at G2/M, are highly expressed in proliferative tumors, and have transforming activity in NIH3T3 cells (Holtrich *et al.*, 1994; Smith *et al.*, 1997; Bischoff *et al.*, 1998; Zhou *et al.*, 1998; Takai *et al.*, 2001). STK15 has been shown to be amplified in human colon cancers and in cell lines derived from many other kinds of human tumors (Bischoff *et al.*, 1998; Zhou *et al.*, 1998). High expression has also been observed in the absence of amplification and contributes to abnormal centrosome numbers in cell lines (Zhou *et al.*, 1998). PLK has also been implicated in centrosome maturation; injection of PLK into either HeLa cells or human foreskin fibroblasts resulted in reduced centrosome size and abnormal chromatin distribution (Lane and Nigg, 1996). Amplification or overexpression of STK15 and/or PLK has been postulated as a potential cause of aneuploidy in human tumors (Lengauer *et al.*, 1998). The high expression of genes involved in centrosome duplication may contribute to the chromosomal translocations and aneuploidy found in HeLa cells.

Although there is significant overlap between the genes found in the proliferation cluster and those that we have identified as cell cycle-regulated, not all cell cycle-regulated genes are highly expressed in the more proliferative tumors. Some cell cycle-regulated genes are expressed in a heterogeneous pattern with no clear relationship to proliferation in the breast tumors studied (e.g., CDC6, CCNE1, ORC1, and BRCA1). Some cell cycle-regulated genes are expressed at lower levels in the more proliferative tumors (calponin 2, smoothelin, vinculin, and actin filament-associated protein). Many of the genes expressed at lower levels in tumors are genes with peak expression at M/G1 and have roles in cell-cell adhesion and regulation of the actin cytoskeleton. The strongly decreased expression of these genes in tumors is often associated with increased invasion, metastasis, and ultimately, poor prognosis (Rodriguez Fernandez *et al.*, 1992; Rudiger, 1998; Engers and Gabbert, 2000). It is highly likely that the regulation of these genes is responsive to many other factors in addition to the cell cycle. Alternatively, the relative low level of expression of some of these genes may actually be advantageous in the proliferative tumors.

The genes identified in this study include the molecular targets of several different classes of chemotherapeutic agents (antimetabolite targets TYMS, RRM1 and 2, and DHFR; tubulin targeted by antimicrotubule drugs, TOP1 and TOP2 inhibitors, etc.) (Ratain, 1997). The list of genes that are strongly expressed in proliferative tumors, especially those previously uncharacterized, may prove to be a useful source of additional drug targets of this kind.

In conclusion, it is important to note that we have extracted only a fraction of the information inherent in this dataset. The entire dataset is now freely available for any purpose ([http://genome-www.stanford.edu/Human-](http://genome-www.stanford.edu/Human-CellCycle/HeLa/)

[CellCycle/HeLa/](http://genome-www.stanford.edu/Human-CellCycle/HeLa/)). We are confident that others will find it a useful source for further discoveries and insights into the cell cycle and cancer.

## ACKNOWLEDGMENTS

We acknowledge members of the Botstein and Brown laboratories for helpful discussions; Max Diehn, Ash Alizadeh, and Jennifer Boldrick for assistance in the collection of time points; Orly Alter and Olga Troyanskaya for discussions of analysis methods; the Stanford Functional Genomics Core Facility for the production of microarrays; and Tim Stearns and May C. Chen for critical reading of the manuscript. M.L.W. is supported by a National Research Service Award Postdoctoral Fellowship from the National Human Genome Research Institute (HG00220-02) and by funds from the Scleroderma Research Foundation. J.I.M. is a Howard Hughes Medical Institute Predoctoral Fellow. M.M.H. is supported by a grant from the American Cancer Society, Florida Division (F99FSU-3) with prior support from the National Institutes of Health (GM-46768). This work was supported by grants from the National Cancer Institute to D.B. and P.O.B. (CA-77097 and CA-85129). P.O.B. is an Associate Investigator of the Howard Hughes Medical Institute.

## REFERENCES

- Abrieu, A., Kahana, J.A., Wood, K.W., and Cleveland, D.W. (2000). CENP-E as an essential component of the mitotic checkpoint in vitro. *Cell* 102, 817–826.
- Abrieu, A., Magnaghi-Jaulin, L., Kahana, J.A., Peter, M., Castro, A., Vigneron, S., Lorca, T., Cleveland, D.W., and Labbe, J. (2001). Mps1 is a kinetochore-associated kinase essential for the vertebrate mitotic checkpoint. *Cell* 106, 83–93.
- Adolph, K.W., and Phelps, J.P. (1982). Role of non-histones in chromosome structure. Cell cycle variations in protein synthesis. *J. Biol. Chem.* 257, 9086–9092.
- Alizadeh, A.A., *et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403, 503–511.
- Alter, O., Brown, P.O., and Botstein, D. (2000). Singular value decomposition for genome-wide expression data processing and modeling. *Proc. Natl. Acad. Sci. USA* 97, 10101–10106.
- Anastasiadis, P.Z., and Reynolds, A.B. (2001). Regulation of Rho GTPases by p120-catenin. *Curr. Opin. Cell Biol.* 13, 604–610.
- Archer, S.Y., and Hodin, R.A. (1999). Histone acetylation and cancer. *Curr. Opin. Genet. Dev.* 9, 171–174.
- Ashburner, M., *et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29.
- Bai, C., Richman, R., and Elledge, S.J. (1994). Human cyclin F. *EMBO J.* 13, 6087–6098.
- Barbacid, M. (1990). ras oncogenes: their role in neoplasia. *Eur. J. Clin. Invest.* 20, 225–235.
- Bischoff, J.R., *et al.* (1998). A homologue of *Drosophila* aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J.* 17, 3052–3065.
- Bishop, J.M. (1991). Molecular themes in oncogenesis. *Cell* 64, 235–248.
- Bjorklund, S., Skog, S., Tribukait, B., and Thelander, L. (1990). S-phase-specific expression of mammalian ribonucleotide reductase R1 and R2 subunit mRNAs. *Biochemistry* 29, 5452–5458.
- Carazo-Salas, R.E., Guarguaglini, G., Gruss, O.J., Segref, A., Karsenti, E., and Mattaj, I.W. (1999). Generation of GTP-bound Ran



- by RCC1 is required for chromatin-induced mitotic spindle formation. *Nature* 400, 178–181.
- Cho, R.J., *et al.* (1998). A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* 2, 65–73.
- 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. (2001). Transcriptional regulation and function during the human cell cycle. *Nat. Genet.* 27, 48–54.
- Cleveland, D.W. (1989). Gene regulation through messenger RNA stability. *Curr. Opin. Cell Biol.* 1, 1148–1153.
- Crawford, D.F., and Piwnica-Worms, H. (2001). The G2 dna damage checkpoint delays expression of genes encoding mitotic regulators. *J. Biol. Chem.* 276, 37166–37177.
- Dalton, S. (1992). Cell cycle regulation of the human cdc2 gene. *EMBO J.* 11, 1797–1804.
- Davenport, J.W., Fernandes, E.R., Harris, L.D., Neale, G.A., and Goorha, R. (1999). The mouse mitotic checkpoint gene Bub1b, a novel Bub1 family member, is expressed in a cell cycle-dependent manner. *Genomics* 55, 113–117.
- Der, C.J., Krontiris, T.G., and Cooper, G.M. (1982). Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. *Proc. Natl. Acad. Sci. USA* 79, 3637–3640.
- Dominski, Z., Zheng, L.X., Sanchez, R., and Marzluff, W.F. (1999). Stem-loop binding protein facilitates 3'-end formation by stabilizing U7 snRNP binding to histone pre-mRNA. *Mol. Cell Biol.* 19, 3561–3570.
- Eisen, M.B., and Brown, P.O. (1999). DNA arrays for analysis of gene expression. *Methods Enzymol.* 303, 179–205.
- Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95, 14863–14868.
- Eliassen, K.A., Baldwin, A., Sikorski, E.M., and Hurt, M.M. (1998). Role for a YY1-binding element in replication-dependent mouse histone gene expression. *Mol. Cell Biol.* 18, 7106–7118.
- Engers, R., and Gabbert, H.E. (2000). Mechanisms of tumor metastasis: cell biological aspects and clinical implications. *J. Cancer Res. Clin. Oncol.* 126, 682–692.
- Fisher, R.P., and Morgan, D.O. (1994). A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. *Cell* 78, 713–724.
- Fonagy, A., Swiderski, C., Wilson, A., Bolton, W., Kenyon, N., and Freeman, J.W. (1993). Cell cycle regulated expression of nucleolar antigen P120 in normal and transformed human fibroblasts. *J. Cell Physiol.* 154, 16–27.
- Fontijn, R.D., Goud, B., Echard, A., Jollivet, F., van Marle, J., Pannekoek, H., and Horrevoets, A.J. (2001). The human kinesin-like protein RB6K is under tight cell cycle control and is essential for cytokinesis. *Mol. Cell Biol.* 21, 2944–2955.
- Georgakopoulos, A., *et al.* (1999). Presenilin-1 forms complexes with the cadherin/catenin cell-cell adhesion system and is recruited to intercellular and synaptic contacts. *Mol. Cell* 4, 893–902.
- Goswami, P.C., Roti Roti, J.L., and Hunt, C.R. (1996). The cell cycle-coupled expression of topoisomerase II $\alpha$  during S phase is regulated by mRNA stability and is disrupted by heat shock or ionizing radiation. *Mol. Cell Biol.* 16, 1500–1508.
- Grosheva, I., Shtutman, M., Elbaum, M., and Bershadsky, A.D. (2001). p120 catenin affects cell motility via modulation of activity of Rho-family GTPases. A link between cell-cell contact formation and regulation of cell locomotion. *J. Cell Sci.* 114, 695–707.
- Gruss, O.J., Carazo-Salas, R.E., Schatz, C.A., Guarguaglini, G., Kast, J., Wilm, M., Le Bot, N., Vernos, I., Karsenti, E., and Mattaj, I.W. (2001). Ran induces spindle assembly by reversing the inhibitory effect of importin  $\alpha$  on TPX2 activity. *Cell* 104, 83–93.
- Guenette, S., Magendantz, M., and Solomon, F. (1995). Suppression of a conditional mutation in  $\alpha$ -tubulin by overexpression of two checkpoint genes. *J. Cell Sci.* 108, 1195–1204.
- Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75, 791–803.
- Harris, M.E., Böhni, R., Schneiderman, M.H., Ramamurthy, L., Schümperli, D., Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. *Mol. Cell Biol.* 11, 2416–2424.
- Heintz, N., Sive, H.L., and Roeder, R.G. (1983). Regulation of human histone gene expression: kinetics of accumulation and changes in the rate of synthesis and in the half-lives of individual histone mRNAs during the HeLa cell cycle. *Mol. Cell Biol.* 3, 539–550.
- Hetzer, M., Bilbao-Cortes, D., Walther, T.C., Gruss, O.J., and Mattaj, I.W. (2000). GTP hydrolysis by Ran is required for nuclear envelope assembly. *Mol. Cell* 5, 1013–1024.
- Hinchcliffe, E.H., and Sluder, G. (2001). "It takes two to tango": understanding how centrosome duplication is regulated throughout the cell cycle. *Genes Dev.* 15, 1167–1181.
- Hirai, H., Roussel, M.F., Kato, J.Y., Ashmun, R.A., and Sherr, C.J. (1995). Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Mol. Cell Biol.* 15, 2672–2681.
- Hirose, K., Kawashima, T., Iwamoto, I., Nosaka, T., and Kitamura, T. (2001). MgcRacGAP is involved in cytokinesis through associating with mitotic spindle and midbody. *J. Biol. Chem.* 276, 5821–5828.
- Hiyama, H., Iavarone, A., and Reeves, S.A. (1998). Regulation of the cdk inhibitor p21 gene during cell cycle progression is under the control of the transcription factor E2F. *Oncogene* 16, 1513–1523.
- Holtrich, U., Wolf, G., Brauninger, A., Karn, T., Bohme, B., Rubsamen-Waigmann, H., and Strebhardt, K. (1994). Induction and down-regulation of PLK, a human serine/threonine kinase expressed in proliferating cells and tumors. *Proc. Natl. Acad. Sci. USA* 91, 1736–1740.
- Hoque, M.T., and Ishikawa, F. (2001). Human chromatid cohesin component hRad21 is phosphorylated in M phase and associated with metaphase centromeres. *J. Biol. Chem.* 276, 5059–5067.
- Horne, M.C., Goolsby, G.L., Donaldson, K.L., Tran, D., Neubauer, M., and Wahl, A.F. (1996). Cyclin G1 and cyclin G2 comprise a new family of cyclins with contrasting tissue-specific and cell cycle-regulated expression. *J. Biol. Chem.* 271, 6050–6061.
- Hunter, T. (1997). Oncoprotein networks. *Cell* 88, 333–346.
- Ishida, S., Huang, E., Zuzan, H., Spang, R., Leone, G., West, M., and Nevins, J.R. (2001). Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. *Mol. Cell Biol.* 21, 4684–4699.
- Iyer, V.R., *et al.* (1999). The transcriptional program in the response of human fibroblasts to serum. *Science* 283, 83–87.
- Jackman, M., Firth, M., and Pines, J. (1995). Human cyclins B1 and B2 are localized to strikingly different structures: B1 to microtubules, B2 primarily to the Golgi apparatus. *EMBO J.* 14, 1646–1654.
- Jinno, S., Suto, K., Nagata, A., Igarashi, M., Kanaoka, Y., Nojima, H., and Okayama, H. (1994). Cdc25A is a novel phosphatase functioning early in the cell cycle. *EMBO J.* 13, 1549–1556.
- Jockusch, B.M., and Isenberg, G. (1981). Interaction of  $\alpha$ -actinin and vinculin with actin: opposite effects on filament network formation. *Proc. Natl. Acad. Sci. USA* 78, 3005–3009.

- Kaczmarek, L. (1986). Protooncogene expression during the cell cycle. *Lab. Invest.* 54, 365–376.
- Kimura, M., Kotani, S., Hattori, T., Sumi, N., Yoshioka, T., Todokoro, K., and Okano, Y. (1997). Cell cycle-dependent expression and spindle pole localization of a novel human protein kinase, Aik, related to Aurora of *Drosophila* and yeast Ipl1. *J. Biol. Chem.* 272, 13766–13771.
- Kinch, M.S., Clark, G.J., Der, C.J., and Burridge, K. (1995). Tyrosine phosphorylation regulates the adhesions of ras-transformed breast epithelia. *J. Cell Biol.* 130, 461–471.
- Knehr, M., Poppe, M., Enulescu, M., Eickelbaum, W., Stoehr, M., Schroeter, D., and Paweletz, N. (1995). A critical appraisal of synchronization methods applied to achieve maximal enrichment of HeLa cells in specific cell cycle phases. *Exp. Cell Res.* 217, 546–553.
- Kondili, K., Tsolas, O., and Papamarcaki, T. (1996). Selective interaction between parathymosin and histone H1. *Eur. J. Biochem.* 242, 67–74.
- Lane, H.A., and Nigg, E.A. (1996). Antibody microinjection reveals an essential role for human polo-like kinase 1 (Plk1) in the functional maturation of mitotic centrosomes. *J. Cell Biol.* 135, 1701–1713.
- Laub, M.T., McAdams, H.H., Feldblyum, T., Fraser, C.M., and Shapiro, L. (2000). Global analysis of the genetic network controlling a bacterial cell cycle. *Science* 290, 2144–2148.
- Lee, M.G., and Nurse, P. (1987). Complementation used to clone a human homologue of the fission yeast cell cycle control gene *cdc2*. *Nature* 327, 31–35.
- Lee, K.S., Yuan, Y.L., Kuriyama, R., and Erikson, R.L. (1995). Plk is an M-phase-specific protein kinase and interacts with a kinesin-like protein, CHO1/MKLP-1. *Mol. Cell. Biol.* 15, 7143–7151.
- Lei, M., and Tye, B.K. (2001). Initiating DNA synthesis. from recruiting to activating the MCM complex. *J. Cell Sci.* 114, 1447–1454.
- Lengauer, C., Kinzler, K.W., and Vogelstein, B. (1998). Genetic instabilities in human cancers. *Nature* 396, 643–649.
- Leone, G., DeGregori, J., Yan, Z., Jakoi, L., Ishida, S., Williams, R.S., and Nevins, J.R. (1998). E2F3 activity is regulated during the cell cycle and is required for the induction of S phase. *Genes Dev.* 12, 2120–2130.
- Lew, D.J., Dulic, V., and Reed, S.I. (1991). Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. *Cell* 66, 1197–1206.
- Li, F., and Altieri, D.C. (1999). The cancer antiapoptosis mouse survivin gene: characterization of locus and transcriptional requirements of basal and cell cycle-dependent expression. *Cancer Res.* 59, 3143–3151.
- Li, J., Xu, M., Zhou, H., Ma, J., and Potter, H. (1997). Alzheimer presenilins in the nuclear membrane, interphase kinetochores, and centrosomes suggest a role in chromosome segregation. *Cell* 90, 917–927.
- Liao, H., Winkfein, R.J., Mack, G., Rattner, J.B., and Yen, T.J. (1995). CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis. *J. Cell Biol.* 130, 507–518.
- Lucibello, F.C., Truss, M., Zwicker, J., Ehlert, F., Beato, M., and Muller, R. (1995). Periodic Cdc25C transcription is mediated by a novel cell cycle-regulated repressor element (CDE). *EMBO J.* 14, 132–142.
- Ma, T., *et al.* (2000). Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev.* 14, 2298–2313.
- Marra, G., Chang, C.L., Laghi, L.A., Chauhan, D.P., Young, D., and Bolland, C.R. (1996). Expression of human MutS homolog 2 (hMSH2) protein in resting and proliferating cells. *Oncogene* 13, 2189–2196.
- Marzluff, W.F. (1992). Histone 3' ends: essential and regulatory functions. *Gene Expression* 2, 93–97.
- Marzluff, W.F., and Pandey, N.B. (1988). Multiple levels of regulation of histone mRNA concentrations. *Trends Biochem. Sci.* 13, 49–52.
- Masuda, H., Tanaka, K., Takagi, M., Ohgami, K., Sakamaki, T., Shibata, N., and Takahashi, K. (1996). Molecular cloning and characterization of human non-smooth muscle calponin. *J. Biochem.* 120, 415–424.
- Matsuoka, M., Kato, J.Y., Fisher, R.P., Morgan, D.O., and Sherr, C.J. (1994). Activation of cyclin-dependent kinase 4 (cdk4) by mouse MO15-associated kinase. *Mol. Cell. Biol.* 14, 7265–7275.
- Mitchell, J.R., Wood, E., and Collins, K. (1999). A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* 402, 551–555.
- Morris, G.F., and Mathews, M.B. (1989). Regulation of proliferating cell nuclear antigen during the cell cycle. *J. Biol. Chem.* 264, 13856–13864.
- Muchardt, C., and Yaniv, M. (2001). When the SWI/SNF complex remodels . . . the cell cycle. *Oncogene* 20, 3067–3075.
- Nachury, M.V., Maresca, T.J., Salmon, W.C., Waterman-Storer, C.M., Heald, R., and Weis, K. (2001). Importin beta is a mitotic target of the small GTPase Ran in spindle assembly. *Cell* 104, 95–106.
- Nurse, P. (2000). A long twentieth century of the cell cycle and beyond. *Cell* 100, 71–78.
- Ohtani, K., DeGregori, J., and Nevins, J.R. (1995). Regulation of the cyclin E gene by transcription factor E2F1. *Proc. Natl. Acad. Sci. USA* 92, 12146–12150.
- Okuda, T., Hirai, H., Valentine, V.A., Shurtleff, S.A., Kidd, V.J., Lahti, J.M., Sherr, C.J., and Downing, J.R. (1995). Molecular cloning, expression pattern, and chromosomal localization of human CDKN2D/INK4d, an inhibitor of cyclin D-dependent kinases. *Genomics* 29, 623–630.
- Ouyang, B., Lan, Z., Meadows, J., Pan, H., Fukasawa, K., Li, W., and Dai, W. (1998). Human Bub1: a putative spindle checkpoint kinase closely linked to cell proliferation. *Cell Growth Differ.* 9, 877–885.
- Pei, L., and Melmed, S. (1997). Isolation and characterization of a pituitary tumor-transforming gene (PTTG). *Mol. Endocrinol.* 11, 433–441.
- Perou, C.M., *et al.* (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. USA* 96, 9212–9217.
- Perou, C.M., *et al.* (2000). Molecular portraits of human breast tumors. *Nature* 406, 747–752.
- Pines, J., and Hunter, T. (1989). Isolation of a human cyclin cDNA: evidence for cyclin mRNA and protein regulation in the cell cycle and for interaction with p34cdc2. *Cell* 58, 833–846.
- Pines, J., and Hunter, T. (1990). Human cyclin A is adenovirus E1A-associated protein p60 and behaves differently from cyclin B. *Nature* 346, 760–763.
- Plumb, M., Stein, J., and Stein, G. (1983). Coordinate regulation of multiple histone mRNAs during the cell cycle in HeLa cells. *Nucleic Acids Res.* 11, 2391–2410.
- Pruitt, K.D., and Maglott, D.R. (2001). RefSeq and LocusLink. NCBI gene-centered resources. *Nucleic Acids Res.* 29, 137–140.
- Puck, T.T., Marcus, P.I., and Cieciura, S.J. (1956). Clonal growth of mammalian cells in vitro. *J. Exp. Med.* 103, 273–284.



- Rajan, J.V., Wang, M., Marquis, S.T., and Chodosh, L.A. (1996). Bca2 is coordinately regulated with Bca1 during proliferation and differentiation in mammary epithelial cells. *Proc. Natl. Acad. Sci. USA* 93, 13078–13083.
- Ratain, M.J. (1997). Pharmacology of cancer chemotherapy. In: *Cancer: Principles and Practice of Oncology*, ed. V.T. DeVita, S. Hellman, S.A. Rosenberg, Philadelphia: Lippincott-Raven Publishers, 375–512.
- Richardson, R.T., Batova, I.N., Widgren, E.E., Zheng, L.X., Whitfield, M., Marzluff, W.F., and O'Rand, M.G. (2000). Characterization of the histone H1-binding protein, NASP, as a cell cycle-regulated somatic protein. *J. Biol. Chem.* 275, 30378–30386.
- Richardson, H.E., Stueland, C.S., Thomas, J., Russell, P., and Reed, S.I. (1990). Human cDNAs encoding homologs of the small p34Cdc28/Cdc2-associated protein of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Genes Dev.* 4, 1332–1344.
- Rodriguez Fernandez, J.L., Geiger, B., Salomon, D., Sabanay, I., Zoller, M., and Ben Ze'ev, A. (1992). Suppression of tumorigenicity in transformed cells after transfection with vinculin cDNA. *J. Cell Biol.* 119, 427–438.
- Ross, D.T., et al. (2000). Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* 24, 227–235.
- Rudiger, M. (1998). Vinculin and  $\alpha$ -catenin: shared and unique functions in adherens junctions. *Bioessays* 20, 733–740.
- Saha, P., Thome, K.C., Yamaguchi, R., Hou, Z., Weremowicz, S., and Dutta, A. (1998). The human homolog of *Saccharomyces cerevisiae* CDC45. *J. Biol. Chem.* 273, 18205–18209.
- Sardet, C., Vidal, M., Cobrinik, D., Geng, Y., Onufryk, C., Chen, A., and Weinberg, R.A. (1995). E2F-4 and E2F-5, two members of the E2F family, are expressed in the early phases of the cell cycle. *Proc. Natl. Acad. Sci. USA* 92, 2403–2407.
- Scheffner, M., Munger, K., Byrne, J.C., and Howley, P.M. (1991). The state of the p53 and retinoblastoma genes in human cervical carcinoma cell lines. *Proc. Natl. Acad. Sci. USA* 88, 5523–5527.
- Schneiderman, M.H., Dewey, W.C., Leeper, D.B., and Nagasawa, H. (1972). Use of the mitotic selection procedure for cell cycle analysis. Comparison between the X-ray and cycloheximide G2 markers. *Exp. Cell Res.* 74, 430–438.
- Schumperli, D. (1988). Multilevel regulation of replication-dependent histone genes. *Trends Genet.* 4, 187–191.
- Shah, J.V., and Cleveland, D.W. (2000). Waiting for anaphase: Mad2 and the spindle assembly checkpoint. *Cell* 103, 997–1000.
- Shelby, R.D., Vafa, O., and Sullivan, K.F. (1997). Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell Biol.* 136, 501–513.
- Sherlock, G., et al. (2001). The Stanford Microarray Database. *Nucleic Acids Res.* 29, 152–155.
- Shindo, M., Nakano, H., Kuroyanagi, H., Shirasawa, T., Mihara, M., Gilbert, D.J., Jenkins, N.A., Copeland, N.G., Yagita, H., Okumura, K. (1998). cDNA cloning, expression, subcellular localization, and chromosomal assignment of mammalian aurora homologues, aurora-related kinase (ARK) 1 and 2. *Biochem. Biophys. Res. Commun.* 244, 285–292.
- Slansky, J.E., and Farnham, P.J. (1996). Transcriptional regulation of the dihydrofolate reductase gene. *Bioessays* 18, 55–62.
- Smith, M.R., Wilson, M.L., Hamanaka, R., Chase, D., Kung, H., Longo, D.L., and Ferris, D.K. (1997). Malignant transformation of mammalian cells initiated by constitutive expression of the polo-like kinase. *Biochem. Biophys. Res. Commun.* 234, 397–405.
- Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* 9, 3273–3297.
- Stampfer, M.R., Pan, C.H., Hosoda, J., Bartholomew, J., Mendelsohn, J., Yaswen, P. (1993). Blockage of EGF receptor signal transduction causes reversible arrest of normal and immortal human mammary epithelial cells with synchronous reentry into the cell cycle. *Exp. Cell Res.* 208, 175–188.
- Stillman, B. (1996). Cell cycle control of DNA replication. *Science* 274, 1659–1664.
- Takai, N., Miyazaki, T., Fujisawa, K., Nasu, K., Hamanaka, R., and Miyakawa, I. (2001). Expression of polo-like kinase in ovarian cancer is associated with histological grade and clinical stage. *Cancer Lett.* 164, 41–49.
- Tassan, J.P., Schultz, S.J., Bartek, J., and Nigg, E.A. (1994). Cell cycle analysis of the activity, subcellular localization, and subunit composition of human CAK (CDK-activating kinase). *J. Cell Biol.* 127, 467–478.
- Tobey, R.A., Valdez, J.G., and Crissman, H.A. (1988). Synchronization of human diploid fibroblasts at multiple stages of the cell cycle. *Exp. Cell Res.* 179, 400–416.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R.B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics* 17, 520–525.
- Vareli, K., Frangou-Lazaridis, M., van, d.K., I, Tsolas, O., and van Driel, R. (2000). Nuclear distribution of prothymosin alpha and parathymosin: evidence that prothymosin alpha is associated with RNA synthesis processing and parathymosin with early DNA replication. *Exp. Cell Res.* 257, 152–161.
- Vaughn, J.P., Cirisano, F.D., Huper, G., Berchuck, A., Futreal, P.A., Marks, J.R., and Iglehart, J.D. (1996). Cell cycle control of BRCA2. *Cancer Res.* 56, 4590–4594.
- Wang, E., Miller, L.D., Ohnmacht, G.A., Liu, E.T., and Marincola, F.M. (2000). High-fidelity mRNA amplification for gene profiling. *Nat. Biotechnol.* 18, 457–459.
- Wang, Z.-F., Whitfield, M.L., Ingledue, T.I., Dominski, Z., and Marzluff, W.F. (1996a). The protein which binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev.* 10, 3028–3040.
- Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B.R., and Crabtree, G.R. (1996b). Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev.* 10, 2117–2130.
- Weinstein, J. (1997). Cell cycle-regulated expression, phosphorylation, and degradation of p53Cdc. A mammalian homolog of CDC20/Fizzy/slp1. *J. Biol. Chem.* 272, 28501–28511.
- Weinstein, J., Jacobsen, F.W., Hsu-Chen, J., Wu, T., and Baum, L.G. (1994). A novel mammalian protein, p55CDC, present in dividing cells is associated with protein kinase activity and has homology to the *Saccharomyces cerevisiae* cell division cycle proteins Cdc20 and Cdc4. *Mol. Cell. Biol.* 14, 3350–3363.
- Wells, D., and Kedes, L. (1985). Structure of a human histone cDNA: evidence that basally expressed histone genes have intervening sequences and encode polyadenylated mRNAs. *Proc. Natl. Acad. Sci. USA* 82, 2834–2838.
- Whitfield, M.L., Zheng, L.X., Baldwin, A., Ohta, T., Hurt, M.M., and Marzluff, W.F. (2000). Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. *Mol. Cell. Biol.* 20, 4188–4198.

- Wiese, C., Wilde, A., Moore, M.S., Adam, S.A., Merdes, A., and Zheng, Y. (2001). Role of importin-beta in coupling Ran to downstream targets in microtubule assembly. *Science* 291, 653–656.
- Wilkins, J.A., and Lin, S. (1982). High-affinity interaction of vinculin with actin filaments in vitro. *Cell* 28, 83–90.
- Williams, G.H., Romanowski, P., Morris, L., Madine, M., Mills, A.D., Stoeber, K., Marr, J., Laskey, R.A., and Coleman, N. (1998). Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc. Natl. Acad. Sci. USA* 95, 14932–14937.
- Wittmann, T., Hyman, A., and Desai, A. (2001). The spindle: a dynamic assembly of microtubules and motors. *Nat. Cell Biol.* 3, E28–E34.
- Wood, K.W., Sakowicz, R., Goldstein, L.S., and Cleveland, D.W. (1997). CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome alignment. *Cell* 91, 357–366.
- Yamamoto, T., Harada, N., Kano, K., Taya, S., Canaani, E., Matsuura, Y., Mizoguchi, A., Ide, C., and Kaibuchi, K. (1997). The Ras target AF-6 interacts with ZO-1 and serves as a peripheral component of tight junctions in epithelial cells. *J. Cell Biol.* 139, 785–795.
- Yamamoto, T., Taya, S., and Kaibuchi, K. (1999). Ras-induced transformation and signaling pathway. *J. Biochem.* 126, 799–803.
- Yan, Z., DeGregori, J., Shohet, R., Leone, G., Stillman, B., Nevins, J.R., and Williams, R.S. (1998). Cdc6 is regulated by E2F and is essential for DNA replication in mammalian cells. *Proc. Natl. Acad. Sci. USA* 95, 3603–3608.
- Zariwala, M., Liu, J., and Xiong, Y. (1998). Cyclin E2, a novel human G1 cyclin and activating partner of CDK2 and CDK3, is induced by viral oncoproteins. *Oncogene* 17, 2787–2798.
- Zhao, J., Dynlacht, B., Imai, T., Hori, T., and Harlow, E. (1998). Expression of NPAT, a novel substrate of cyclin E-CDK2, promotes S-phase entry. *Genes Dev.* 12, 456–461.
- Zhao, J., Kennedy, B.K., Lawrence, B.D., Barbie, D.A., Matera, A.G., Fletcher, J.A., and Harlow, E. (2000). NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes Dev.* 14, 2283–2297.
- Zhong, C., Kinch, M.S., and Burridge, K. (1997). Rho-stimulated contractility contributes to the fibroblastic phenotype of Ras-transformed epithelial cells. *Mol. Biol. Cell* 8, 2329–2344.
- Zhou, H., Kuang, J., Zhong, L., Kuo, W.L., Gray, J.W., Sahin, A., Brinkley, B.R., and Sen, S. (1998). Tumor amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. *Nat. Genet.* 20, 189–193.
- Zieve, G.W., Turnbull, D., Mullins, J.M., and McIntosh, J.R. (1980). Production of large numbers of mitotic mammalian cells by use of the reversible microtubule inhibitor nocodazole. Nocodazole accumulated mitotic cells. *Exp. Cell Res.* 126, 397–405.
- Zou, H., McGarry, T.J., Bernal, T., and Kirschner, M.W. (1999). Identification of a vertebrate sister-chromatid separation inhibitor involved in transformation and tumorigenesis. *Science* 285, 418–422.
- Zwicker, J., Lucibello, F.C., Wolfrum, L.A., Gross, C., Truss, M., Engeland, K., and Muller, R. (1995). Cell cycle regulation of the cyclin A, cdc25C and cdc2 genes is based on a common mechanism of transcriptional repression. *EMBO J.* 14, 4514–4522.