Supporting Online Material

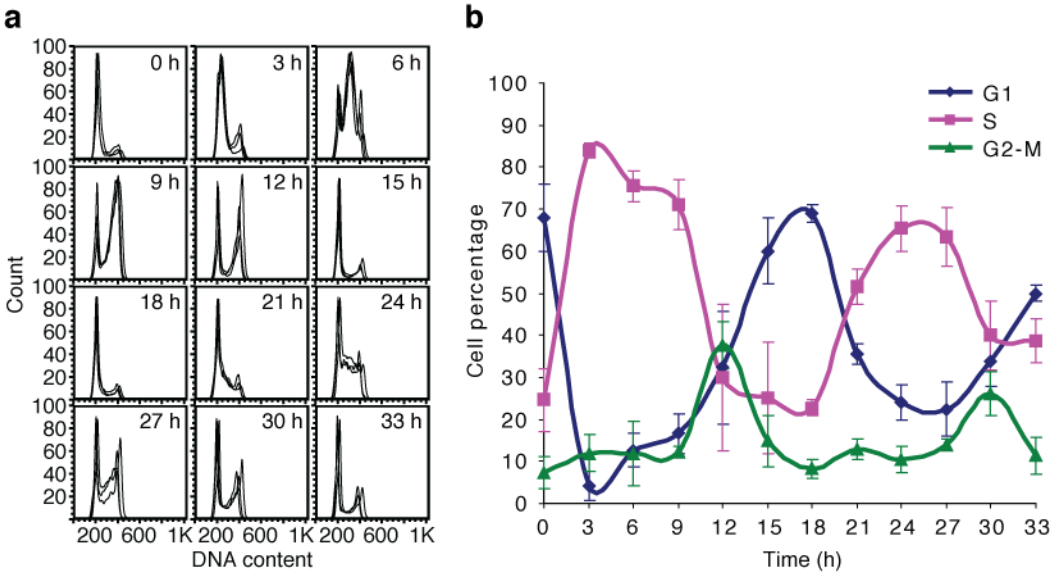
Supplementary Methods 1 – ChIP-seq

Plates (75cm^2) with ~6 million HaCaT cells were washed twice with PBS, then cross-linked with 0.5% formaldehyde at room temperature for 10 mins, following which cross-linking was quenched by the addition of glycine to 125 mM. Cells were again washed with PBS, then scraped from plates in PBS supplemented with 1mM EDTA and protease inhibitors. Cross-linked cells were collected by centrifugation at 1500 G, and resuspended in cold RIPA buffer (10 mM Tris pH 8.0, 1 mM EDTA, 140 mM NaCl, 1 % Triton X-100, 0.1% SDS, 0.1 % Na-Deoxycholate) supplemented with protease inhibitors at a density of 1 x 10^7 cells/ml. Lysed cells were then sonicated in TPX tubes (Diagenode) using a Bioruptor (Diagenode) for 12 mins on full power, then debris was removed by centrifugation in a chilled mircofuge at 16000 G for 15 mins. Soluble chromatin supernatant equivalent to 3.8 x 10^6 cells / IP was pre-cleared using 25 µl of a 1:1 mix of protein A and protein G Dynabeads (Invitrogen), then used for overnight immunoprecipitation using either 4 µg anti-H3K4me3 antisera (Diagenode s.a., Cat. # pAb-003-050; Lot # A2-002P) or 2 µg anti H3K27me3 (Upstate Biotechnology, Cat. # 07-449; Lot # DAM1387952). Antibody-protein complexes were immunoprecipitated with protein A/G Dynabeads, then washed 5 times with RIPA buffer supplemented with protease inhibitors, once with LiCl wash buffer (250 mM LiCl, 10 mM Tris pH 8.0, 1 mM EDTA, 0.5 % Igepal CA-630, 0.5 % Na-deoxycholate) supplemented with protease inhibitors, and once in TE. Immunoprecipitated DNA was isolated by successive RNaseA and proteinase K treatment, followed by purification over Genomic DNA cleanup & concentrator columns (Zymo Research Corp.). Two ng of immunoprecipitated material, or input chromatin isolated as above, was used for library preparation using TruSeq™ DNA Sample Preparation reagents (Illumina), and submitted to the Norwegian Sequencing Centre for sequencing on a HiSeq 2000 (Illumina). IP and control input libraries were clustered on a single lane of an Illumina flowcell using TruSeq™ SR Cluster Kit v3 reagents and subjected to 50 bases of sequencing using TruSeq™ SBS v3 reagents (Illumina). Image analysis and base calling was performed using Illumina's RTA software version 1.12.4.2. Demultiplexing and conversion from bcl file to fastq file was performed using CASAVA v1.8.2. Reads were filtered to remove those with low base call quality using Illumina's default chastity criteria.

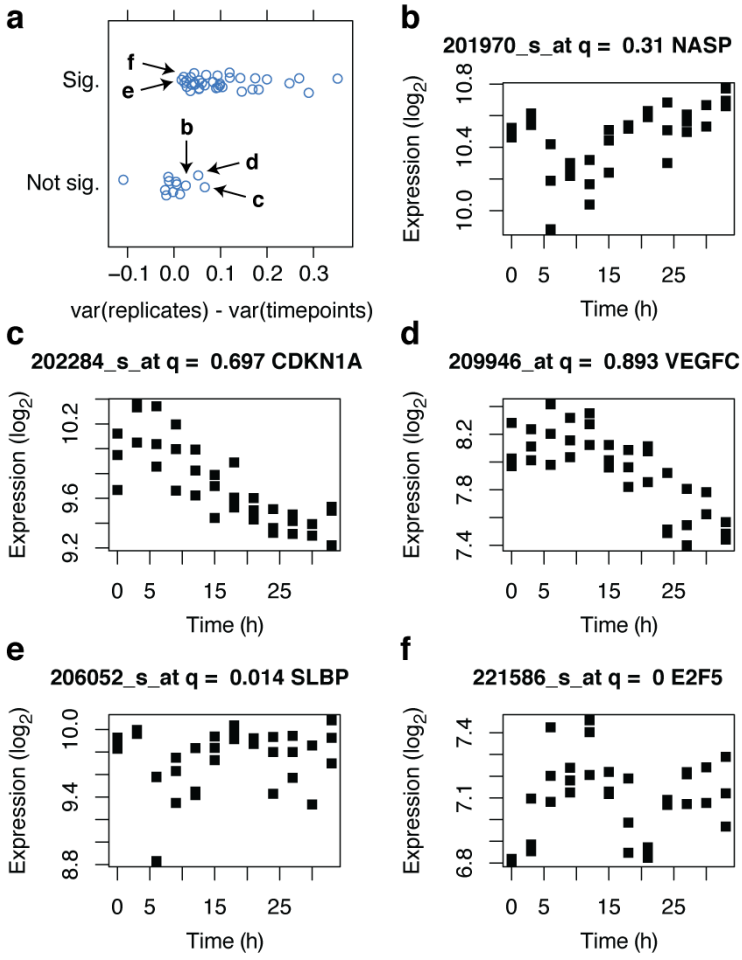
Supplementary Discussion 1 – A PLS model based only on the first 24 hours after release identifies additional cell cycle regulated genes at the expense of additional false positives

When studying the list of genes with previously published phases and the genes on that list not detected by our PLS model to be cell cycle expressed in HaCaT, we noticed that 7 of the 12 missing genes were previously published to be expressed in S phase (Supplementary Table 1). Closer inspection of the expression profiles indicated that some

of these genes appeared to have a cyclic expression pattern during the first cell division. We therefore created a new PLS model based on only the first 24 hours after release (Supplementary Fig. 3). This model identified several additional cell cycle genes, including 4 of the 12 genes missed by the PLS model based on all the 33 hours after release (Supplementary Fig. 4, Supplementary Table 1). Although the 24 hours PLS model gave an increased coverage of known cell cycle genes of about 8%, this was at the expense of about 9% additional significant cell cycle genes than the 33 hours model. Using all the available probes on the chip instead of only the probe that showed the highest variation for each Entrez gene further increased the coverage of known cell cycle genes to 42 of 48 genes, but also increased the total number of significant cell cycle genes to 2081 (data not shown). To limit the influence of potential false positives and focus our analyses on genes that showed consistent cyclic expression across multiple cell divisions, we therefore decided to use the 33 hours PLS model as basis for further analyses.

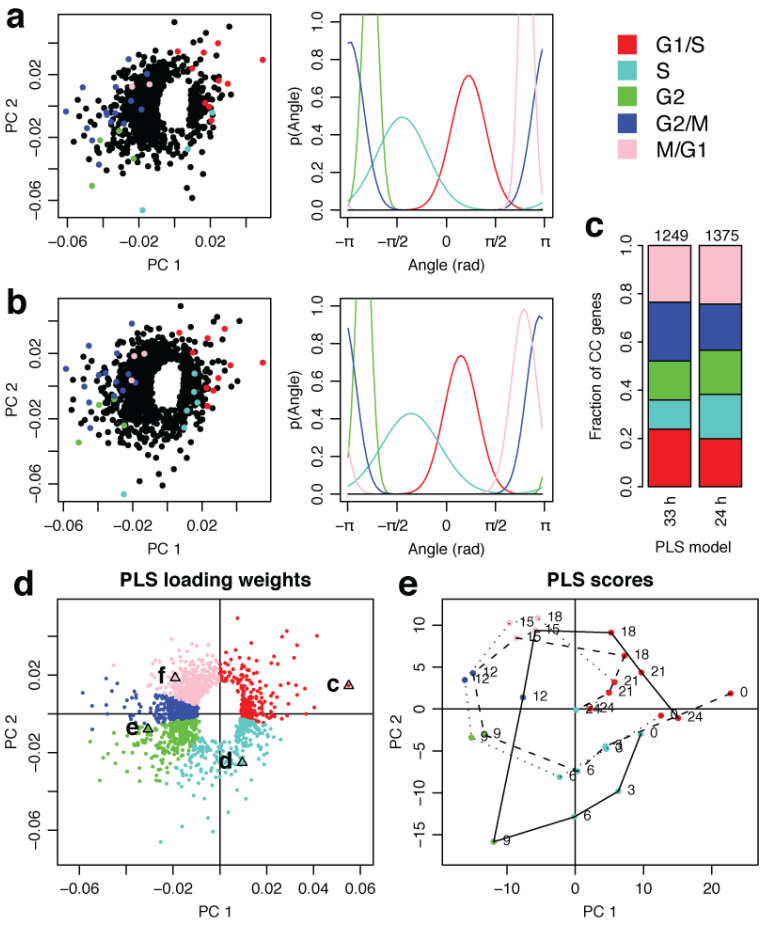


Supplementary Figure 1 - Synchrony of double thymidine blocked HaCaT cells. (**a**) Cell synchrony was monitored by flow cytometry of propidium iodide-stained cells. The figures show for each time point, superimposed profiles of three replicate synchronization experiments. Horizontal axes show DNA content (arbitrary units) and vertical axes show the number of events (cells) with the corresponding DNA content. (**b**) Percentage of cells assigned to G1, S, and G2/M phases for each of the time points analyzed. Values and error bars are averages and standard deviations (n = 3).



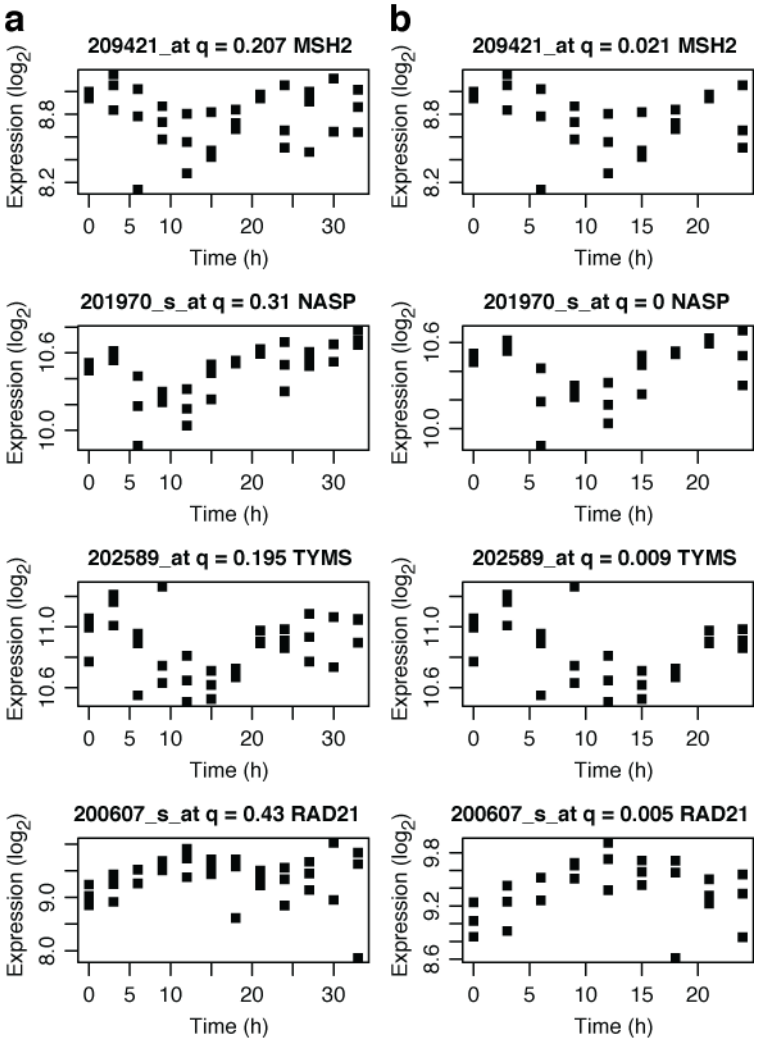
Supplementary Figure 2 - "Missing" cell cycle genes either show no cyclic expression or have more variance between replicates than "detected" cell cycle genes. (**a**) The figure shows the average expression variance across the 12 time points (var(replicates); variance averaged over the 3 replicates) minus the average expression variance across the 3 replicates (var(timepoints); variance averaged over the 12 time points) for the previously known cell cycle regulated genes (table S2) detected ("Sig.") and not detected ("Not sig.") to be significantly cell cycle regulated in HaCaT by this study. All the significant genes have higher average variance within a replicate than within time points. In contrast, 6 of the 12 "missing" genes (*BRCA1, BRCA2, DHFR, MSH2, RRM1*, and *RAD21*) have lower average variance within a replicate than within time points; that is, their expression was inconsistent and varied as much between the replicates as within the same replicate. Only three of the "missing" genes (*CDKN1A*, *NASP*, and *VEGFC*) had a difference in average variance that was greater than any of the significant genes. Arrows and labels show these three "missing" genes with the largest difference in average variance (**b-d**) and the two significant genes with the smallest difference in average variance (**e-f**). (b-f) Expression patterns for the five genes annotated in (a). The "missing" genes (b-d) do not show cyclic expression patterns

through both cell cycles, whereas the significant genes do (e-f). (b) *NASP* showed an expression pattern consistent with being up-regulated during S phase and down-regulated during G2/M phase in the first cell cycle but not in the second; that is, *NASP* was down-regulated in the first G2/M phase (9 and 12 hours after release), but not in the second G2/M phase (27 and 30 hours after release). (c-d) *CDKN1A* and *VEGFC* were highly expressed in the first G1/S transition and S phase, but were not up-regulated in the second cell cycle. Figure titles show the probe id, the false discovery rate (q), and the gene name.

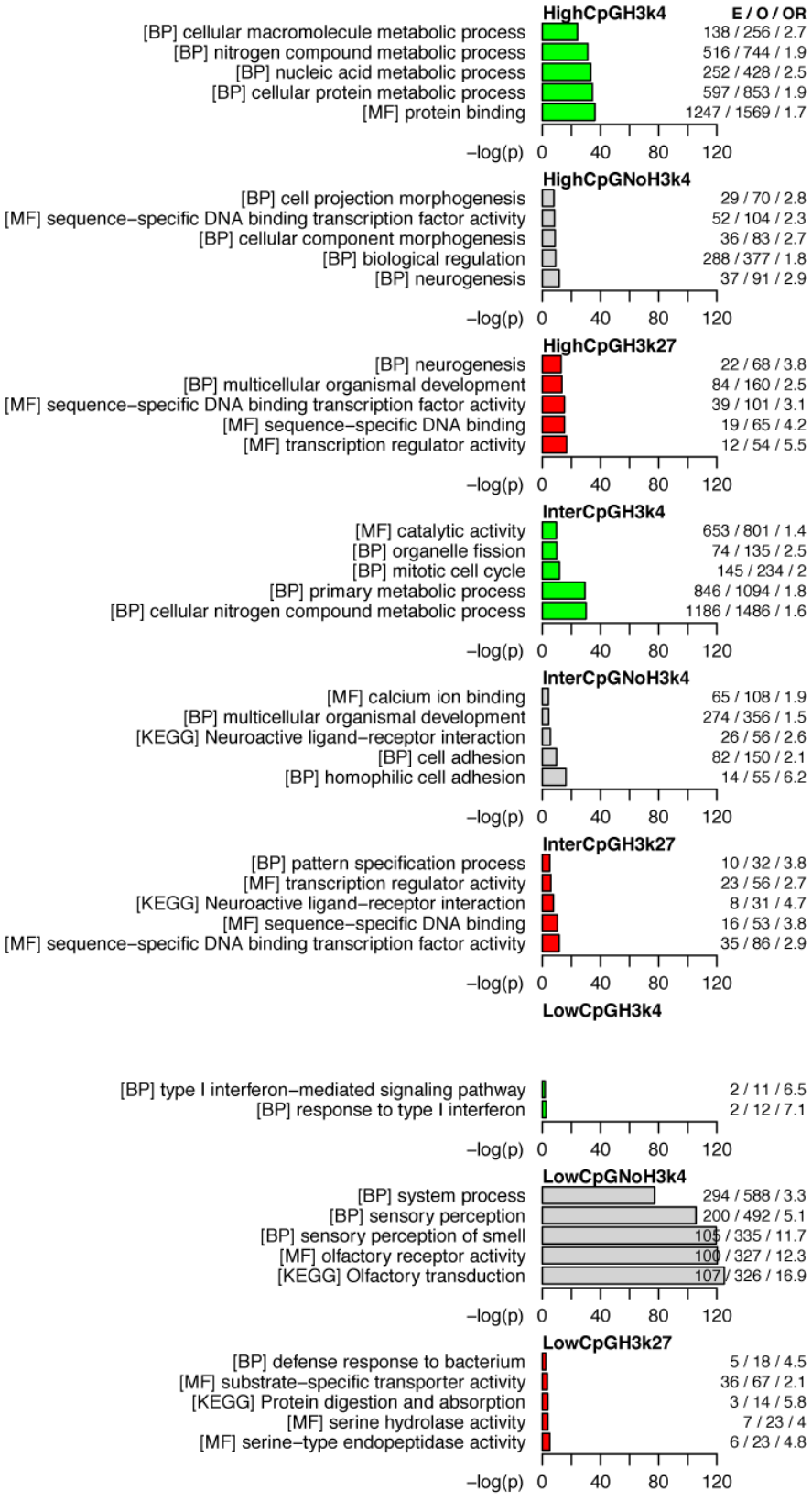


Supplementary Figure 3 – A cyclic PLS model based on the first 24 hours after release identifies additional genes expressed in S phase in HaCaT. (**a-b**) Known cell cycle genes cluster according to their published cell cycle phase. (Left) Loadings plot showing significant genes ($q < 0.05$; permutation test) and the genes with known cell cycle expression (Supplementary Table 2) in the cyclic PLS models based on all the 33 hours (a) and only the first 24 hours (b) after release from double thymidine block. The genes with known cell cycle expression are color-coded according to their known cell cycle phase (red, G1/S; turquoise, S; green, G2; blue, G2/M; and pink, M/G1). (Right) Phase angle probability distributions for the five cell cycle phases estimated based on the phase angles of the known cell cycle genes in the loadings plot (see Methods). Horizontal axis shows the phase angle (radians); vertical axis shows the probability. Note that because the horizontal axis represents a linearization of a circle, the distributions "wrap around" from π to -π. (**c**) The 24 hours PLS model identifies more genes as significant cell cycle genes than the 33 hours PLS model and the largest increase is for genes expressed in S phase. The bar graphs show the relative distribution of significant genes ($q < 0.05$) in the five cell cycle phases for the 33 hours and 24 hours PLS models. The numbers above the bars show the total number of significant genes in each model. Note that the 24 hours model gave poorer separation of the G2/M and M/G1 phases than the 33

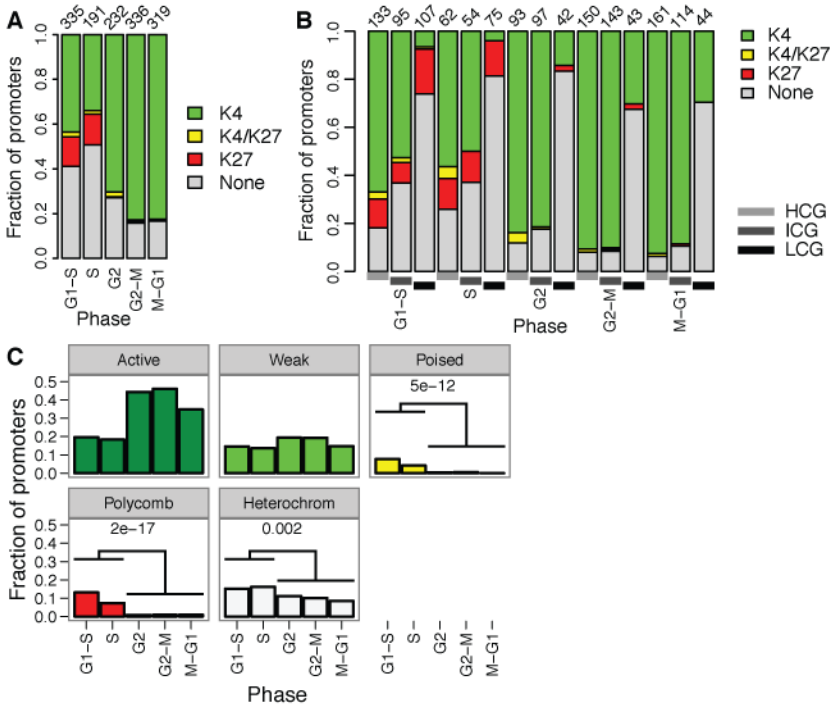
hours model (compare phase angle probability distributions in (a) and (b)), which resulted in the 24 hours model having fewer genes in G2/M phase than the 33 hours model had. (**d-e**) Loadings (d) and scores (e) plots for the cyclic PLS model based on only the first 24 hours after release from double thymidine block. The solid, dashed, and dotted lines in (e) show the trajectories for the three biological replicates. Labeled triangles in (d) show the location of the four selected genes in Fig. 1: (c) *CCNE2*; (d) *CDKN2C*; (e) *CCNA2*; and (f) *CDKN3*. See Figure 1 for the four genes' expression profiles and additional details regarding loadings and scores plots.



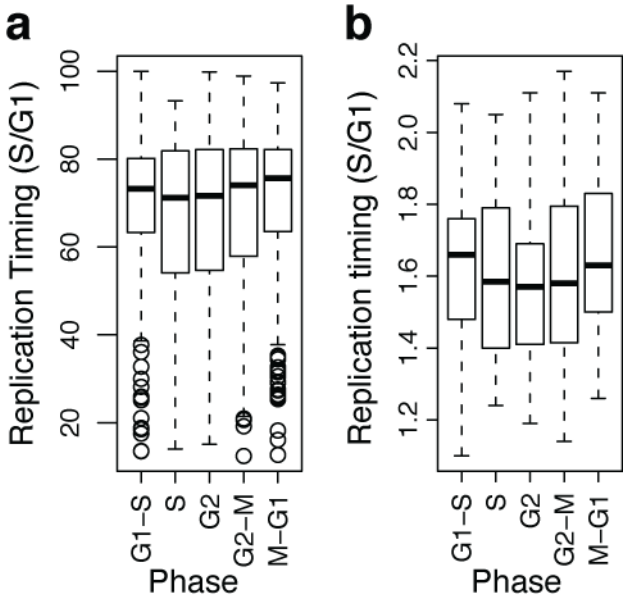
Supplementary Figure 4 – Expression patterns of the four known cell cycle regulated genes identified by the 24 hours PLS model (b) but missed by the 33 hours PLS model (a). Figure titles show the probe id, the false discovery rate (q), and the gene name. Except for *RAD21*, the additional genes detected by the 24 hours model were predominantly expressed in G1/S or S phase.


HighCpGH3k4
E / O / OR
[BP] cellular macromolecule metabolic process
138 / 256 / 2.7
[BP] nitrogen compound metabolic process
516 / 744 / 1.9
[BP] nucleic acid metabolic process
252 / 428 / 2.5
[BP] cellular protein metabolic process
597 / 853 / 1.9
[MF] protein binding
1247 / 1569 / 1.7
−log(p) 0 40 80 120
HighCpGNoH3k4
[BP] cell projection morphogenesis
29 / 70 / 2.8
[MF] sequence−specific DNA binding transcription factor activity
52 / 104 / 2.3
[BP] cellular component morphogenesis
36 / 83 / 2.7
[BP] biological regulation
288 / 377 / 1.8
[BP] neurogenesis
37 / 91 / 2.9
−log(p) 0 40 80 120
HighCpGH3k27
[BP] neurogenesis
22 / 68 / 3.8
[BP] multicellular organismal development
84 / 160 / 2.5
[MF] sequence−specific DNA binding transcription factor activity
39 / 101 / 3.1
[MF] sequence−specific DNA binding
19 / 65 / 4.2
[MF] transcription regulator activity
12 / 54 / 5.5
−log(p) 0 40 80 120
InterCpGH3k4
[MF] catalytic activity
653 / 801 / 1.4
[BP] organelle fission
74 / 135 / 2.5
[BP] mitotic cell cycle
145 / 234 / 2
[BP] primary metabolic process
846 / 1094 / 1.8
[BP] cellular nitrogen compound metabolic process
1186 / 1486 / 1.6
−log(p) 0 40 80 120
InterCpGNoH3k4
[MF] calcium ion binding
65 / 108 / 1.9
[BP] multicellular organismal development
274 / 356 / 1.5
[KEGG] Neuroactive ligand−receptor interaction
26 / 56 / 2.6
[BP] cell adhesion
82 / 150 / 2.1
[BP] homophilic cell adhesion
14 / 55 / 6.2
−log(p) 0 40 80 120
InterCpGH3k27
[BP] pattern specification process
10 / 32 / 3.8
[MF] transcription regulator activity
23 / 56 / 2.7
[KEGG] Neuroactive ligand−receptor interaction
8 / 31 / 4.7
[MF] sequence−specific DNA binding
16 / 53 / 3.8
[MF] sequence−specific DNA binding transcription factor activity
35 / 86 / 2.9
−log(p) 0 40 80 120
LowCpGH3k4
[BP] type I interferon−mediated signaling pathway
2 / 11 / 6.5
[BP] response to type I interferon
2 / 12 / 7.1
−log(p) 0 40 80 120
LowCpGNoH3k4
[BP] system process
294 / 588 / 3.3
[BP] sensory perception
200 / 492 / 5.1
[BP] sensory perception of smell
105 / 335 / 11.7
[MF] olfactory receptor activity
100 / 327 / 12.3
[KEGG] Olfactory transduction
107 / 326 / 16.9
−log(p) 0 40 80 120
LowCpGH3k27
[BP] defense response to bacterium
5 / 18 / 4.5
[MF] substrate−specific transporter activity
36 / 67 / 2.1
[KEGG] Protein digestion and absorption
3 / 14 / 5.8
[MF] serine hydrolase activity
7 / 23 / 4
[MF] serine−type endopeptidase activity
6 / 23 / 4.8
−log(p) 0 40 80 120

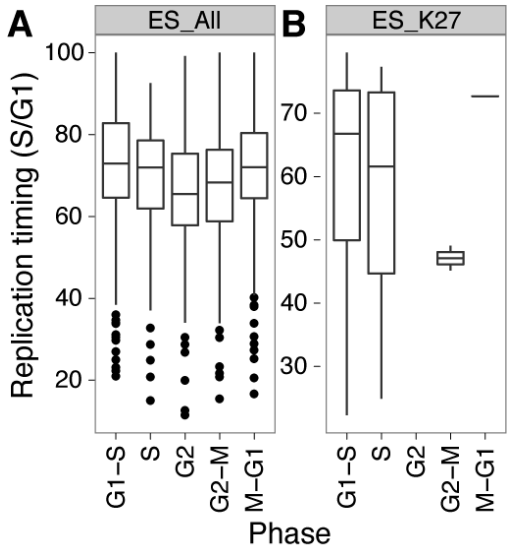
Supplementary Figure 5 – High-CpG and consistent H3K4me3 are associated with housekeeping functions whereas low-CpG and no or inconsistent H3K4me3 are associated with cell type-specific functions. The barplots show the five most significantly KEGG pathways (KEGG), or GO molecular function (MF) or biological process (BP) terms overrepresented among all genes with high, intermediate, or low CpG promoters and consistent H3K4me3, H3K27me3, or lack of consistent H3K4me3 modifications. Horizontal axes show Benjamini-Hochberg corrected p-values (–log10); numbers on the right show the expected (E) and observed (O) number of genes and the corresponding odds ratios (OR). Cell type-specific functions such as "Olfactory transduction" are strongly overrepresented among LCG/NoH3K4me3 genes whereas housekeeping functions such as "cellular metabolic process" are overrepresented among HCG/H3K4me3 genes.



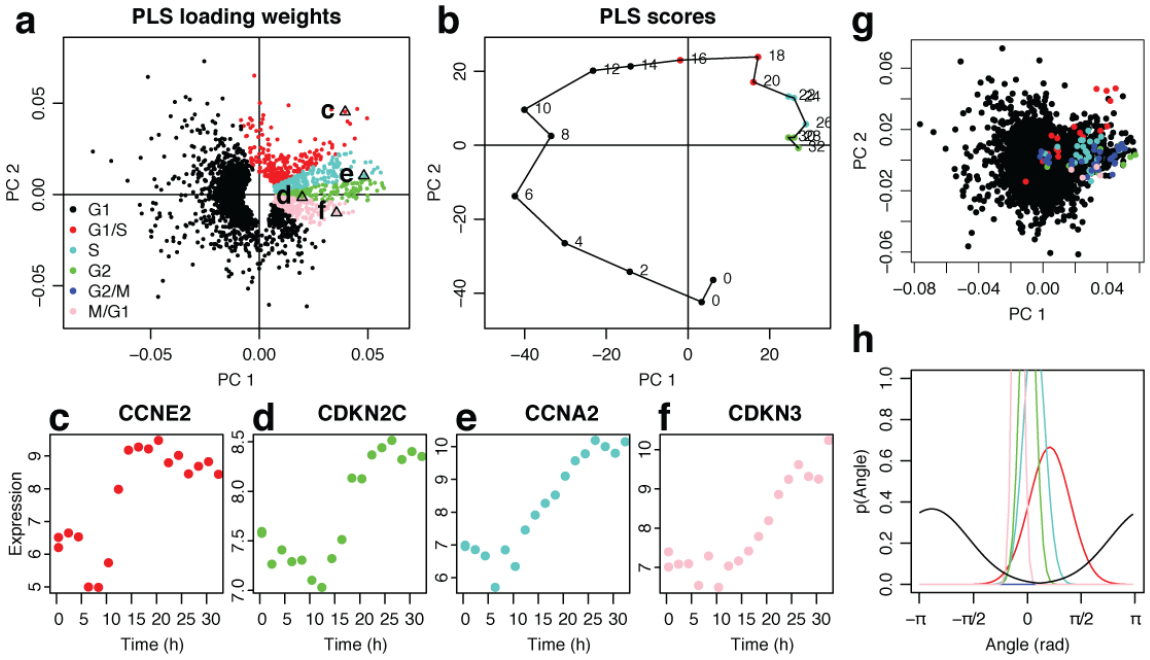
Supplementary Figure 6 - G1/S and S phase genes are consistently enriched for H3K27me3 marks within multiple cell lines. (**A**) Distribution of consistent H3K4me3 and H3K27me3 marks in promoters of genes upregulated in G1/S, S, G2, G2/M, and M/G1 phases; see Fig. 4. (**B**) Fractions of low-CpG (LCG), intermediate-CpG (ICG), and high-CpG (HCG) promoters for genes upregulated in the five cell cycle phases. (**C**) Promoter chromatin states for cell cycle genes within the NHEK keratinocyte cell line show that poised and Polycomb-regulated promoters are strongly enriched whereas promoters with heterochromatin regions are weakly enriched among genes expressed during DNA replication. The bar graphs show the fraction of promoters that contain regions classified by a published "Chromatin-state" Hidden Markov Model (1) to be Active, Weak, and Poised promoters and Polycomb-repressed and Heterochromatic regions. Lines and numbers within the graphs indicate chromatin states (Poised, Polycomb, and Heterochrom) that are enriched among genes expressed during DNA replication (G1/S and S phases; numbers are p-values from one-tailed binomial tests).



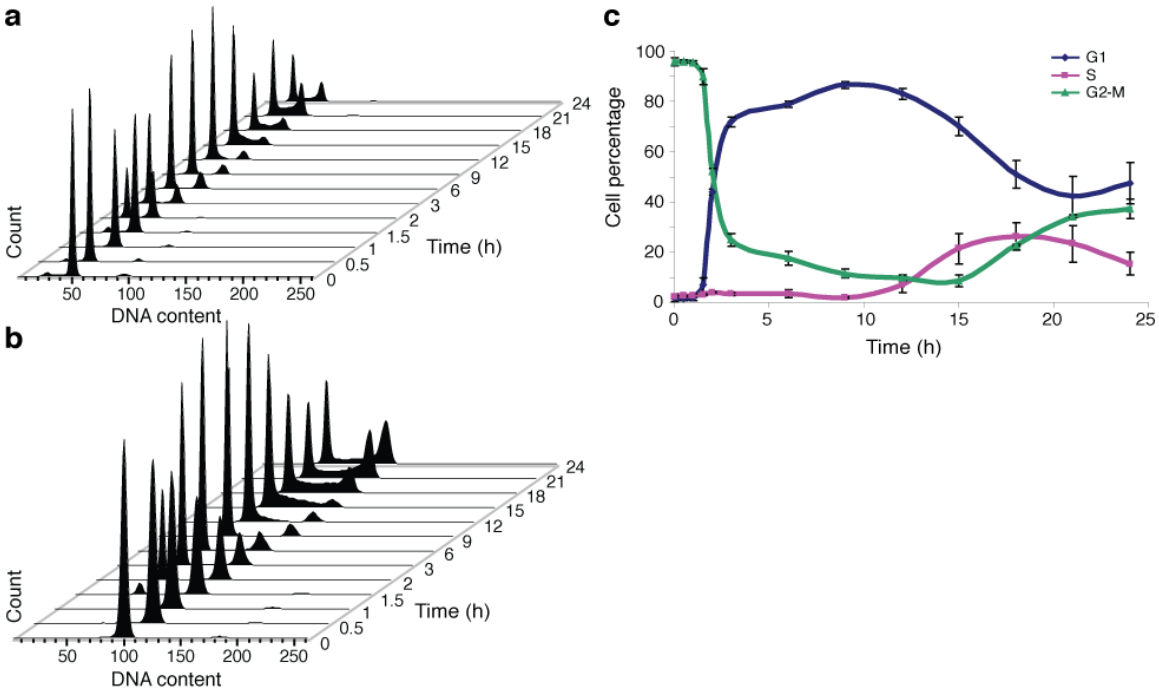
Supplementary Figure 7 – Data from different cell lines show that HaCaT G1/S genes tend to be replicated earlier than S genes. The box plots show the replication timing of cell cycle genes from HaCaT in (**a**) basophilic erythroblasts (2) and (**b**) lymphoblastoid cells (3). The replication timing is the median S/G1 ratio of regions overlapping each gene; data from sequencing-based TimEX (**a**) are normalized S/G1 ratios (see (2)); data from microarray-based hybridization experiments (**b**) are probe S/G1 ratios (see (3)). For (**a**) early and late replicating regions have values close 100 and 0, respectively, whereas for (**b**) early and late replicating regions have values close to 2 and 1, respectively. Boxes, horizontal black line, and circles show the first and third quartiles, the median, and outliers; whiskers show the most extreme datapoints up to 1.5 times the interquartile range from the box.



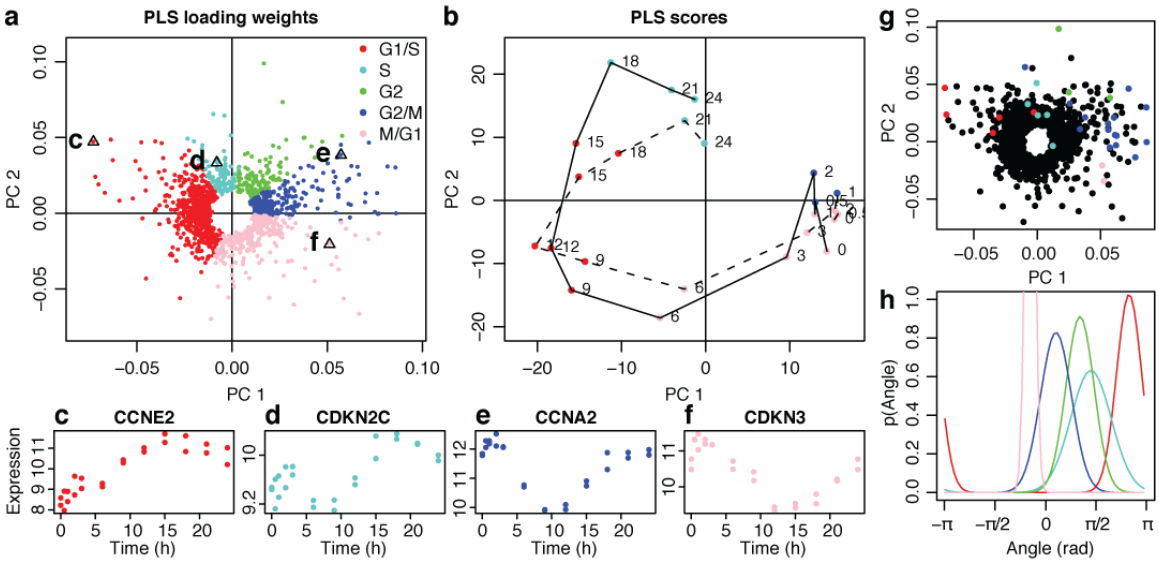
Supplementary Figure 8 – Replication timing for genes residing in regions with stable replication timing. (**A**) Replication timing within all cell cycle genes that do not overlap regions found by Hansen and colleagues to exhibit replication timing plasticity (4). The data is a subset of Figure 4E where all transcripts that overlap regions with replication timing plasticity have been removed (regions of replication timing plasticity were defined in Supplementary Table S2 in (4)) (**B**) Replication timing within the subset of genes from (**A**) that are marked with H3K27me3 within their promoters. (**A**-**B**) Replication timing is the median S/G1 ratio within embryonic stem cells (2) of regions overlapping each gene. See Supplementary Figure 7 for details on the box-and-whisker plots.



Supplementary Figure 9 – Partial least squares (PLS) regression reanalysis of serum starved primary foreskin fibroblasts identifies cell cycle genes. (**a-b**) Loadings (a) and scores (b) plots for a cyclic PLS model of the gene expression profiles from Bar-Joseph and colleagues' primary foreskin fibroblast experiment (5). The period used was 30 hours estimated based on the cells' DNA content profile (Fig. 1a in (5)). Labeled triangles in (a) show the location of the four selected genes from Figure 1 and panels (**c-f**) show their expression profiles. The genes are (c) *CCNE2*; (d) *CDKN2C*; (e) *CCNA2*; and (f) *CDKN3*. See Figure 1 for additional details regarding loadings and scores plots. (**g**) Loadings plot showing the genes with known cell cycle expression (Supplementary Table 2) in the PLS model (a). (**h**) Phase angle probability distributions for four of the five cell cycle phases. Specifically, we excluded the G2/M phase from the model, as the experiment, in contrast to the HaCaT double thymidine block experiment (Fig. 1; Supplementary Figure 3a), did not give sufficient temporal resolution to accurately separate the genes in the G2/M phase from the G2 and M/G1 phases; compare the patterns for the genes with known cell cycle expression in Supplementary Figure 3a with their pattern in (g). In Supplementary Figure 3a, these known genes are well separated, whereas in the primary foreskin fibroblast data, the genes are clustered at the end of the series and the G2/M phase genes completely overlap the G2 and M/G1 genes. Also shown in (h) is the phase angle probability distribution for the G1 phase, as estimated by the cells' DNA content profile. These phase angles had a marked lack of known cell cycle genes (see g), and were therefore excluded from the model. The final cell cycle model, which included the G1/S, S, G2, and M/G1 phases, consisted of 627 distinct genes and included 60% of the genes originally reported by Bar-Joseph and colleagues. See Methods and Supplementary Figure 3 for additional details.



Supplementary Figure 10 – Synchrony by nocodazole and mitotic shake-off blocked HeLa cells. (**a** and **b**) Cell synchrony was monitored by flow cytometry of propidium iodide-stained cells. The figures show for each of the two biological replicates, DNA profiles at each time point of the synchronization experiment. Horizontal axes show DNA content (arbitrary units) and vertical axes show the number of events (cells) with the corresponding DNA content. (**c**) Percentage of cells assigned to G1, S, and G2/M phases for each of the time points analyzed. Values and error bars are averages and standard deviations (n = 2).



Supplementary Figure 11 – Nocodazole, mitotic shake-off, and partial least squares (PLS) regression identifies HeLa cell cycle genes. (**a-b**) Loadings (a) and scores (b) plots for a cyclic PLS model of the gene expression profiles from HeLa cells synchronized by nocodazole treatment and mitotic shake-off. The period used was 25 hours estimated based on the cells' DNA content profile (Supplementary Fig. 10). Labeled triangles in (a) show the location of the four selected genes from Figure 1 and panels (**c-f**) show their expression profiles. The genes are (c) *CCNE2*; (d) *CDKN2C*; (e) *CCNA2*; and (f) *CDKN3*. See Figure 1 for additional details regarding loadings and scores plots. (**g**) Loadings plot showing significant genes ($q < 0.05$; permutation test) and the genes with known cell cycle expression (Supplementary Table 2) in the PLS model (a). (**h**) Phase angle probability distributions for the five cell cycle phases. See Methods and Supplementary Figure 3 for additional details.

Supplementary Table 1 - List of genes with previously published phases. The genes are from Table 2 in (6). "Predicted (33 h)" are the cell cycle phases assigned by this study's main PLS model for the HaCaT cells, which is based on all the time points up to 33 h after release; "Predicted (24 h)" are the cell cycle phases assigned by a PLS model based on only the time points up to 24 h after release (in HaCaT); "Whitfield" are the cell cycle phases reported based on microarray analyses of synchronized HeLa cells (6); "Actual" are previously reported cell cycle phases as compiled by (6).

Gene	Predicted (33 h)	Predicted (24 h)	Whitfield	Actual
E2F5	G2	G2	G2/M	G1
CCNE1	G1/S	G1/S	G1/S	G1/S
CCNE2	G1/S	G1/S	G1/S	G1/S
CDC25A	G1/S	G1/S	G1/S	G1/S
CDC45L	S	S	S	G1/S
CDC6	G1/S	G1/S	G1/S	G1/S
CDKN1A	ND	ND	ND	G1/S
E2F1	G1/S	G1/S	G1/S	G1/S
MCM2	G1/S	G1/S	G1/S	G1/S
MCM6	G1/S	G1/S	G1/S	G1/S
NPAT	ND	ND	G1/S	G1/S
PCNA	G1/S	G1/S	G1/S	G1/S
SLBP	G1/S	G1/S	G1/S	G1/S
BRCA1	ND	ND	S	S
BRCA2	ND	ND	ND	S
CCNG2	S	S	ND	S
CDKN2C	S	S	G2	S
DHFR	ND	ND	S	S
MSH2	ND	G1/S	G1/S	S
NASP	ND	G1/S	G1/S	S
RRM1	ND	ND	S	S
RRM2	G1/S	S	S	S
TYMS	ND	S	S	S
CCNA2	G2	G2	G2	G2
CCNF	G2	G2	G2	G2
CENPF	G2	G2	G2/M	G2
TOP2A	G2	G2	G2	G2
BIRC5	M/G1	ND	G2/M	G2/M
BUB1	G2/M	G2/M	G2/M	G2/M
BUB1B	G2/M	G2/M	G2/M	G2/M
CCNB1	G2/M	G2/M	G2/M	G2/M
CCNB2	M/G1	M/G1	G2/M	G2/M
CDK1	G2	G2	G2	G2/M
CDC20	M/G1	M/G1	G2/M	G2/M
CDC25B	G2/M	G2/M	G2/M	G2/M

CDC25C	G2/M	G2	G2	G2/M
CDKN2D	ND	ND	G2/M	G2/M
CENPA	G2	G2	G2/M	G2/M
CKS1B	G2/M	G2/M	G2	G2/M
CKS2	G2/M	G2/M	G2/M	G2/M
PLK1	G2/M	G2/M	G2/M	G2/M
AURKA	G2	G2	G2/M	G2/M
RACGAP1	G2/M	G2/M	NM	G2/M
KIF20A	G2/M	M/G1	NM	G2/M
RAD21	ND	G2/M	M/G1	M/G1
PTTG1	M/G1	M/G1	M/G1	M/G1
VEGFC	ND	ND	ND	M/G1
CDKN3	M/G1	M/G1	M/G1	M/G1

ND: Not determined; the gene was not detected as having a significant cell cycle expression pattern. NM: Not measured; the gene was not present on the microarray used by (6).

Supplementary Table 2 - Cell cycle genes common for foreskin fibroblasts (FF) (5), HeLa (6), and HaCaT are strongly enriched for cell cycle-related terms. The table shows all KEGG pathways and GO molecular function (MF), biological process (BP), and cellular component (CC) terms that are overrepresented ($p < 0.05$) among the 125 cell cycle Entrez genes common for FF, HeLa, and HaCaT. The table is sorted based on the Benjamini-Hochberg-corrected p-value (p) of a hypergeometric test for overrepresentation, which for the GO terms was conditioned on the structure of the GO graph (7). Exp. and Count are the expected and observed number of genes; OR is the corresponding odds ratio; Size is the size of the background group used in the test. Rows shaded in gray indicate terms that were not significant ($p < 0.05$) among any of the other subsets of cell cycle genes common or exclusive for FF, HeLa, or HaCaT (see fig. S6).

Ontology	Term	ID	p	OR	Exp.	Count	Size
BP	organelle fission	GO:0048285	2e-45	43.1	2.0	44	250
BP	cell division	GO:0051301	1e-33	34.6	1.7	35	230
BP	M phase	GO:0000279	3e-33	30.0	2.2	37	309
BP	cell cycle process	GO:0022402	8e-31	27.3	2.4	36	359
KEGG	Cell cycle	04110	3e-30	72.9	1.1	26	129
BP	mitotic cell cycle	GO:0000278	1e-29	28.0	2.1	34	309
BP	mitosis	GO:0007067	1e-26	40.1	1.1	26	157
CC	non-membrane-bounded organelle	GO:0043228	2e-19	6.6	17.3	61	2351
CC	nucleoplasm	GO:0005654	1e-17	9.2	5.7	36	770
CC	organelle lumen	GO:0043233	2e-17	6.6	12.3	50	1670
BP	DNA replication	GO:0006260	3e-16	22.8	1.2	20	158
CC	intracellular organelle	GO:0043229	2e-15	8.7	62.3	104	8480
BP	cellular component organization	GO:0016043	2e-14	6.5	13.4	48	1960
CC	intracellular	GO:0005622	9e-14	18.2	76.7	111	10435
CC	membrane-bounded organelle	GO:0043227	2e-12	5.4	55.3	95	7530
CC	organelle part	GO:0044422	7e-12	7.2	9.8	36	2115
BP	chromosome organization	GO:0051276	1e-11	8.9	3.7	25	460
CC	condensed chromosome	GO:0000793	1e-11	25.5	0.7	13	91
KEGG	DNA replication	03030	3e-10	48.9	0.3	9	36
CC	spindle	GO:0005819	2e-9	29.6	0.4	10	65
CC	chromosome	GO:0005694	1e-8	12.4	1.4	14	221
BP	phosphoinositide-mediated signaling	GO:0048015	2e-8	22.4	0.6	11	78
CC	condensed chromosome kinetochore	GO:0000777	3e-8	39.8	0.3	8	38
BP	DNA packaging	GO:0006323	7e-8	16.5	0.9	12	112
BP	spindle organization	GO:0007051	1e-7	138.5	0.1	6	12
BP	regulation of cyclin-dependent protein kinase activity	GO:0000079	2e-7	28.1	0.4	9	52
BP	chromosome segregation	GO:0007059	5e-7	24.9	0.5	9	59
CC	nucleus	GO:0005634	9e-7	4.9	12.8	33	3157
CC	microtubule cytoskeleton	GO:0015630	1e-6	12.6	1.0	11	180
MF	protein binding	GO:0005515	3e-6	3.6	44.1	71	6555

CC	centrosome	GO:0005813	6e-6	10.4	1.2	11	175
MF	ATP binding	GO:0005524	1e-5	3.7	10.8	31	1426
BP	G2 phase of mitotic cell cycle	GO:0000085	1e-5	513.4	0.0	4	5
BP	cell cycle	GO:0007049	2e-5	16.3	0.7	9	175
MF	microtubule motor activity	GO:0003777	2e-5	16.6	0.6	8	76
MF	adenyl nucleotide binding	GO:0030554	4e-5	3.4	11.5	31	1518
CC	chromosome passenger complex	GO:0032133	6e-5	Inf	0.0	3	3
MF	nucleoside binding	GO:0001882	6e-5	3.3	11.7	31	1544
BP	DNA repair	GO:0006281	1e-4	7.9	1.8	12	230
BP	cellular response to stimulus	GO:0051716	2e-4	4.5	5.3	20	656
CC	midbody	GO:0030496	4e-4	56.0	0.1	4	14
BP	regulation of mitosis	GO:0007088	4e-4	24.2	0.3	6	40
CC	microtubule	GO:0005874	4e-4	7.4	1.5	10	216
BP	cell proliferation	GO:0008283	6e-4	3.8	7.3	23	914
KEGG	p53 signaling pathway	04115	8e-4	12.7	0.6	6	69
CC	cytosol	GO:0005829	8e-4	3.4	7.6	22	1035
CC	chromosome, centromeric region	GO:0000775	1e-3	22.2	0.3	5	41
BP	positive regulation of nuclear division	GO:0051785	1e-3	32.3	0.2	5	25
CC	DNA replication factor C complex	GO:0005663	1e-3	138.7	0.0	3	6
MF	purine ribonucleotide binding	GO:0032555	1e-3	2.9	13.4	31	1772
BP	mitotic sister chromatid segregation	GO:0000070	1e-3	66.6	0.1	4	12
KEGG	Mismatch repair	03430	1e-3	27.0	0.2	4	23
BP	microtubule-based movement	GO:0007018	1e-3	11.0	0.8	8	104
BP	establishment of chromosome localization	GO:0051303	1e-3	388.6	0.0	3	4
BP	nucleosome assembly	GO:0006334	2e-3	13.4	0.6	7	77
CC	outer kinetochore of condensed chromosome	GO:0000940	2e-3	104.0	0.1	3	7
BP	DNA replication initiation	GO:0006270	3e-3	52.3	0.1	4	14
BP	regulation of organelle organization	GO:0033043	3e-3	10.0	0.9	8	117
CC	spindle microtubule	GO:0005876	3e-3	31.1	0.2	4	22
CC	nuclear chromosome	GO:0000228	4e-3	8.8	0.9	7	120
BP	chromatin assembly or disassembly	GO:0006333	5e-3	9.2	1.0	8	124
MF	DNA-dependent ATPase activity	GO:0008094	6e-3	16.3	0.4	5	47
BP	mitotic cell cycle spindle assembly checkpoint	GO:0007094	1e-2	95.3	0.1	3	7
MF	5'-flap endonuclease activity	GO:0017108	1e-2	Inf	0.0	2	2
BP	cytokinesis	GO:0000910	2e-2	17.5	0.3	5	42
CC	nucleolus	GO:0005730	2e-2	3.4	5.0	15	675
MF	hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	GO:0016818	2e-2	3.4	5.4	16	712
BP	negative regulation of mitosis	GO:0045839	2e-2	76.3	0.1	3	8
MF	nucleotide binding	GO:0000166	2e-2	2.4	16.3	32	2156
CC	cytoplasm	GO:0005737	2e-2	2.0	50.4	70	6857
CC	spindle pole centrosome	GO:0031616	2e-2	275.0	0.0	2	3

BP	regulation of ubiquitin-protein ligase activity	GO:0051438	2e-2	11.3	0.6	6	75
BP	positive regulation of cell cycle	GO:0045787	2e-2	15.8	0.4	5	46
KEGG	Small cell lung cancer	05222	2e-2	8.0	0.7	5	86
BP	mitotic cell cycle checkpoint	GO:0007093	3e-2	25.6	0.2	4	25
CC	kinesin complex	GO:0005871	3e-2	32.0	0.1	3	16
BP	cellular biopolymer metabolic process	GO:0034960	3e-2	2.2	43.9	65	5477
BP	regulation of catalytic activity	GO:0050790	4e-2	3.4	5.7	17	716
BP	regulation of phosphate metabolic process	GO:0019220	4e-2	4.4	3.1	12	383
BP	cellular macromolecular complex subunit organization	GO:0034621	4e-2	4.3	3.1	12	385
BP	establishment or maintenance of microtubule cytoskeleton polarity	GO:0030951	5e-2	Inf	0.0	2	2
BP	DNA replication-dependent nucleosome assembly	GO:0006335	5e-2	Inf	0.0	2	2
BP	positive regulation of mitotic metaphase/anaphase transition	GO:0045842	5e-2	Inf	0.0	2	2
BP	mitotic chromosome movement towards spindle pole	GO:0007079	5e-2	Inf	0.0	2	2
BP	mitotic spindle elongation	GO:0000022	5e-2	Inf	0.0	2	2
BP	DNA replication, removal of RNA primer	GO:0043137	5e-2	Inf	0.0	2	2
BP	regulation of kinase activity	GO:0043549	5e-2	5.1	2.2	10	272

References

1. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M. *et al.* (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, **473**, 43-49.
2. Desprat, R., Thierry-Mieg, D., Lailler, N., Lajugie, J., Schildkraut, C., Thierry-Mieg, J. and Bouhassira, E.E. (2009) Predictable dynamic program of timing of DNA replication in human cells. *Genome Res*, **19**, 2288-2299.
3. Woodfine, K., Fiegler, H., Beare, D.M., Collins, J.E., McCann, O.T., Young, B.D., Debernardi, S., Mott, R., Dunham, I. and Carter, N.P. (2004) Replication timing of the human genome. *Hum Mol Genet*, **13**, 191-202.
4. Hansen, R.S., Thomas, S., Sandstrom, R., Canfield, T.K., Thurman, R.E., Weaver, M., Dorschner, M.O., Gartler, S.M. and Stamatoyannopoulos, J.A. (2010) Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc Natl Acad Sci U S A*, **107**, 139-144.
5. Bar-Joseph, Z., Siegfried, Z., Brandeis, M., Brors, B., Lu, Y., Eils, R., Dynlacht, B.D. and Simon, I. (2008) Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc Natl Acad Sci U S A*, **105**, 955-960.
6. Whitfield, M.L., Sherlock, G., Saldanha, A.J., Murray, J.I., Ball, C.A., Alexander, K.E., Matese, J.C., Perou, C.M., Hurt, M.M., Brown, P.O. *et al.* (2002) Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell*, **13**, 1977-2000.
7. Falcon, S. and Gentleman, R. (2007) Using GOstats to test gene lists for GO term association. *Bioinformatics*, **23**, 257-258.