



HAL
open science

A circadian pattern of expression for one percent of all rat transcripts.

C. Grunchober, F. Delaunay, A. Puehlohofer, Gérard Triqueneaux, Vincent Laudet, T. Bartfai, P. Nef

► To cite this version:

C. Grunchober, F. Delaunay, A. Puehlohofer, Gérard Triqueneaux, Vincent Laudet, et al.. A circadian pattern of expression for one percent of all rat transcripts.. *Journal of Biological Chemistry*, 2001, 276 (50), pp.46751-46758. 10.1074/jbc.M107499200 . hal-00068385

HAL Id: hal-00068385

<https://hal.science/hal-00068385>

Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Circadian Regulation of Diverse Gene Products Revealed by mRNA Expression Profiling of Synchronized Fibroblasts*[§]

Received for publication, August 6, 2001, and in revised form, October 1, 2001
Published, JBC Papers in Press, October 11, 2001, DOI 10.1074/jbc.M107499200

Christophe Grundschober[‡], Franck Delaunay[§], Anja Pühlhofer[‡], Gerard Triqueneaux[§],
Vincent Laudet[§], Tamas Bartfai[‡], and Patrick Nef^{‡¶}

From the [‡]Central Nervous System Department, Hoffmann-La Roche, 4070 Basel, Switzerland and [§]CNRS UMR 5665,
Ecole Normale Supérieure, 69364 Lyon Cedex 07, France

Genes under a 24-h regulation period may represent drug targets relevant to diseases involving circadian dysfunctions. As a testing model of the circadian clock system, we have used synchronized rat fibroblasts that are known to express at least six genes in a circadian fashion. We have determined the expression patterns of 9957 transcripts every 4 h over a total period of 76 h using high density oligonucleotide microarrays. The spectral analysis of our mRNA profiling data indicated that ~2% (85 genes) of all expressed genes followed a robust circadian pattern. We have confirmed the circadian expression of previously known clock or clock-driven genes, and we identified 81 novel circadian genes. The majority of the circadian-regulated gene products are known and are involved in diverse cellular functions. We have classified these circadian genes in seven clusters according to their phase of cycling. Our pathway analysis of the mRNA profiling data strongly suggests a direct link between circadian rhythm and cell cycle.

The circadian rhythm allows organisms to anticipate daily environmental changes, and evolutionary conserved molecular clock and oscillator mechanisms have been observed in plants, fungi, and mammals (1). Circadian oscillators are generated by a small number of clock proteins that are regulated by self-sustained transcriptional-translational feedback loops, which drive the rhythmic expression of “clock-controlled” output genes. In mammals, the central pacemaker is located in the suprachiasmatic nucleus (SCN)¹ of the hypothalamus and is adjusted and entrained to the 24-h environmental light/dark cycle by photic stimuli via the retina. Surprisingly, confluent cultured fibroblast cells show an identical (2) circadian clock mechanism following synchronization with a 50% v/v serum shock treatment (3) and can be used as a simple model to study circadian gene expression. The discovery of a circadian clock in peripheral cells strongly suggests that this rhythmic mechanism controls basic cellular functions. So far the extent of gene expression following a circadian rhythm in cultured fibroblasts

is still unknown, as well as its overall nature. We report on fibroblast experiments using high density oligonucleotide microarray expression profiling (GeneChip®, Affymetrix) to determine the expression levels of 9957 rat genes over a 76-h period after synchronization.

EXPERIMENTAL PROCEDURES

Sample Preparation and Chip Hybridization—Rat 3Y1 embryonic fibroblasts (passage 30) (4) were grown in Dulbecco's modified Eagle's medium with 5% v/v fetal calf serum (Roche Molecular Biochemicals), 2 mM L-glutamine, 100 IU/ml penicillin, and 100 mg/ml streptomycin (Life Technologies, Inc.) at 37 °C in 7.5% CO₂ and were plated (in duplicate) on 15-cm dishes at 3 × 10⁶ cells/plate and grown in the same medium for 7 days. The cells were confluent for 4 days before serum synchronization. At this time (*t* = 0 h), cells were treated with 50% v/v horse serum (Roche Molecular Biochemicals) for 2 h and then kept in serum-free Dulbecco's modified Eagle's medium (3). Every 4 h, the content of two plates was lysed in 10 ml of RNazolB™ (AMS Biotechnology) and total RNA extracted following the manufacturer's protocol. Target preparation for chip hybridization was done as described (5); briefly, 20 μg of total RNA was used as starting material, and 30 μg of *in vitro* transcribed biotinylated target was fragmented, applied to the chip, and hybridized overnight at 45 °C. Four spiked controls were included in the hybridization mix (BioB, BioC, BioD, and Cre at 1.5, 5, 25, and 100 pM, respectively). We used Affymetrix high density arrays, custom made for Hoffmann-La Roche, testing a total of 12,204 rat sequences on 2 arrays: 2840 public sequences present in GenBank™ as of June, 1998, and 9364 proprietary ESTs (Incyte Pharmaceuticals). After washing at 45 °C and antibody amplification, the arrays were stained and scanned on a Hewlett-Packard scanner. The raw data were exported from the “GeneChip” software program (Affymetrix) and subsequently analyzed with ad hoc programs.

Expression Data Analysis—To allow for comparisons between arrays, the total signal intensities (sum of signal intensities of all individual genes on each chip) were normalized to the mean signal intensity of all chips.

To detect periodic variations in gene expression, we modeled the time series according to a second order autoregressive process (2, 6), and because outliers may be present, a robust estimation procedure for the coefficients was used (7, 8). Moreover, before applying the spectral analysis method, a robust linear regression (“least trimmed squares robust regression”) (9) was employed to remove any linear trend from the data corresponding to a drift of the base line. The spectral analysis was then applied to the residuals. For each time series, we calculated the frequency V_{\max} where the spectrum *f* reaches its maximum; with this method non-cycling transcripts would have a frequency of 0.

To select genes with a strong circadian expression pattern, the curves from the 4006 cycling time series identified by spectral analysis were plotted in Excel (Microsoft) (data not shown). The analysis was eased by the fact that it was a time course experiment. Changes in expression levels were regarded as significant, if they were present at successive time points in the same direction, followed by changes in the other direction over several subsequent time points. Genes were scored from 1 to 4 by four independent investigators. The highest mark (=4) went to genes with at least two peaks of maximal expression, a cycle period of 20–24 h, and high reproducibility between replicates. All the genes that got a mark between 2 and 4 were selected for further analysis

RNase Protection—The *rev-erba* probe corresponds to a 230-bp frag-

* This work was supported by ARC, LNCC, FRM, CNRS, MENRT, and ENS Lyon. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[§] The on-line version of this article (available at <http://www.jbc.org>) contains median hybridization values, first peak, period, quality, and Table II.

[¶] To whom correspondence should be addressed: Bldg. 69/410, F. Hoffmann-La Roche Ltd., CH-4070 Basel, Switzerland. Tel.: 41 61 688 08 28; Fax: 41 61 688 17 20; E-mail: Patrick.Nef@roche.com.

¹ The abbreviations used are: SCN, suprachiasmatic nucleus; Q-PCR, quantitative reverse transcriptase-PCR; CDK, cyclin-dependent kinase.

TABLE I
Clustered list of the 67 circadian gene transcripts with their known accession numbers

Genechip ID	Accession number	Description
Cluster 1 <i>n</i> = 6		
02685	Y07744	<i>Rattus norvegicus</i> UDP- <i>n</i> -acetyl-D-glucosamine-2-epimerase
06657	L11007	Rat cell division protein kinase 4 (cyclin-dependent kinase 4) 4 ESTs: 03258, 04719, 04930, 09766
Cluster 2 <i>n</i> = 16		
01429	M98327	Rat transfer RNA-valine synthetase
02682	Y00979	<i>R. norvegicus</i> β,β -enolase
03327	AB014722	rsalt-1
04760	U01024	Mouse phosphoribosylamine—glycine ligase
04871	AB012276	Mouse ATFx
05644	AF233449	Tumor necrosis factor receptor-associated factor 4
05727	Y17393	Prefoldin subunit 2
06391	X76040	Human mitochondrial ion protease homolog precursor
08649	CAB37992	dj465n24.2.1 (putative novel protein)
08919	AB016930	<i>Cricetulus griscus</i> phosphatidylglycerophosphate synthase
09468	AB012580	Mouse translation initiation factor eIF3 p66
09966	D16309	Rat cyclin D3
10509	X78683	<i>Mus musculus</i> B cell receptor-associated protein (BAP) 37
11561	U89431	<i>M. musculus</i> strain BALB/c unknown mRNA 2 ESTs: 09542, 10887
Cluster 3 <i>n</i> = 6		
00604	J03179	Rat D-binding protein
01084	M25804	Rat rev-crba- α protein
01682	U18729	<i>R. norvegicus</i> cytochrome <i>b</i> ₅₅₈ α -subunit
06965	AF039085	Cellugyrin
08023	AF123653	<i>Homo sapiens</i> FEZ1
09606	J03179	Rat D-binding protein
Cluster 4 <i>n</i> = 14		
00009	AB000507	<i>R. norvegicus</i> aquaporin 7
01171	M36421	Glutamate receptor 4 precursor (AMPA 4)
01978	U60063	<i>R. norvegicus</i> aldehyde dehydrogenase
05219	AF036249	Polymerase 1 and transcript release factor
08025	M37736	Mouse replication-dependent histone H2A.1
10275	AF181561	<i>R. norvegicus</i> pro-SAAS
11435	AF224721	Mouse secretory carrier membrane protein 4 7 ESTs: 03382, 04673, 05638, 06106, 07526, 10949, 12202
Cluster 5 <i>n</i> = 19		
00555	D88461	Rat n-wasp
00757	L11587	Rat leukocyte common antigen-related phosphatase
00885	L28801	Rat transcription factor iic α -subunit
02155	U89282	<i>R. norvegicus</i> telomerase protein component 1 (tip1)
02787	Z68145	<i>R. norvegicus</i> immunoglobulin λ -5 chain
03085	U35022	Rat <i>cis</i> -Golgi matrix protein
03114	AF227684	Mouse intracellular hyaluronan-binding protein p57
04528	D89801	Exon 12 of RING3, a nuclear serine-threonine kinase
06495	AF045458	<i>H. sapiens</i> serine/threonine kinase ULK1
06672	AF032872	Potassium channel regulatory protein
08011	AAC47068	Similar to K02E10.2 gene product (<i>Caenorhabditis elegans</i>)
08897	U08110	Mouse ran GTPase-activating protein 1
09837	AB028858	<i>M. musculus</i> mRNA for mDj8
09954	AB016532	Rat PER2
11173	AB024427	<i>M. musculus</i> Sid1669p
11290	X52583	Rat nuclear pore glycoprotein P 3 ESTs: 05409, 06971, 09001
Cluster 6 <i>n</i> = 8		
00426	D38101	Rat l-type voltage-dependent calcium channel α 1 subunit
00746	L10073	<i>R. norvegicus</i> 5-hydroxytryptamine receptor (5ht5b)
02692	Y08616	<i>R. norvegicus</i> fertilin- α
05028	D29683	Rat endothelin-converting enzyme
05235	AF116523	<i>M. musculus</i> periplakin (ppl)
06834	AB000777	CRY1 mouse photolyase/blue-light receptor
07847	U09399	Mouse glycine receptor β -subunit precursor (glrb) 1 EST: 08863
Cluster 7 <i>n</i> = 13		
01672	U17607	<i>R. norvegicus</i> ccaat binding transcription factor cbf subunit c
01996	U62821	<i>R. norvegicus</i> outer dense fiber protein (cdf2)
03108	AB030644	Tudor repeat associator with ptaire 2
04804	L06433	<i>R. norvegicus</i> c-HA-ras
05914	X67241	Rat guanine nucleotide-releasing protein (p140 ras-grf)
06960	AJ006340	26 S proteasome subunit p112
07510	AF051561	Rat Na-K-Cl cotransporter (Nkcc1)

TABLE I—continued

GeneChip® ID	Accession number	Description
08066	M19967	Rat annexin 1 (phospholipase a2 inhibitory protein)
08242	AF006086	Human arp2/3 complex 21-kDa subunit (p21-arc)
09818	AB017026	Mouse oxysterol-binding protein
09935	X79233	Ews (fragment)
10277	U40188	Human transcriptional activation factor TAF1132 1 EST: 03869
Not clustered $n = 4$		
00259	AJ222691	<i>R. norvegicus</i> dna polymerase δ (catalytic subunit)
06312	D63885	Lysophospholipase
07289	U91538	Similar to mouse vesicle trafficking protein sc22b 1 EST: 08264

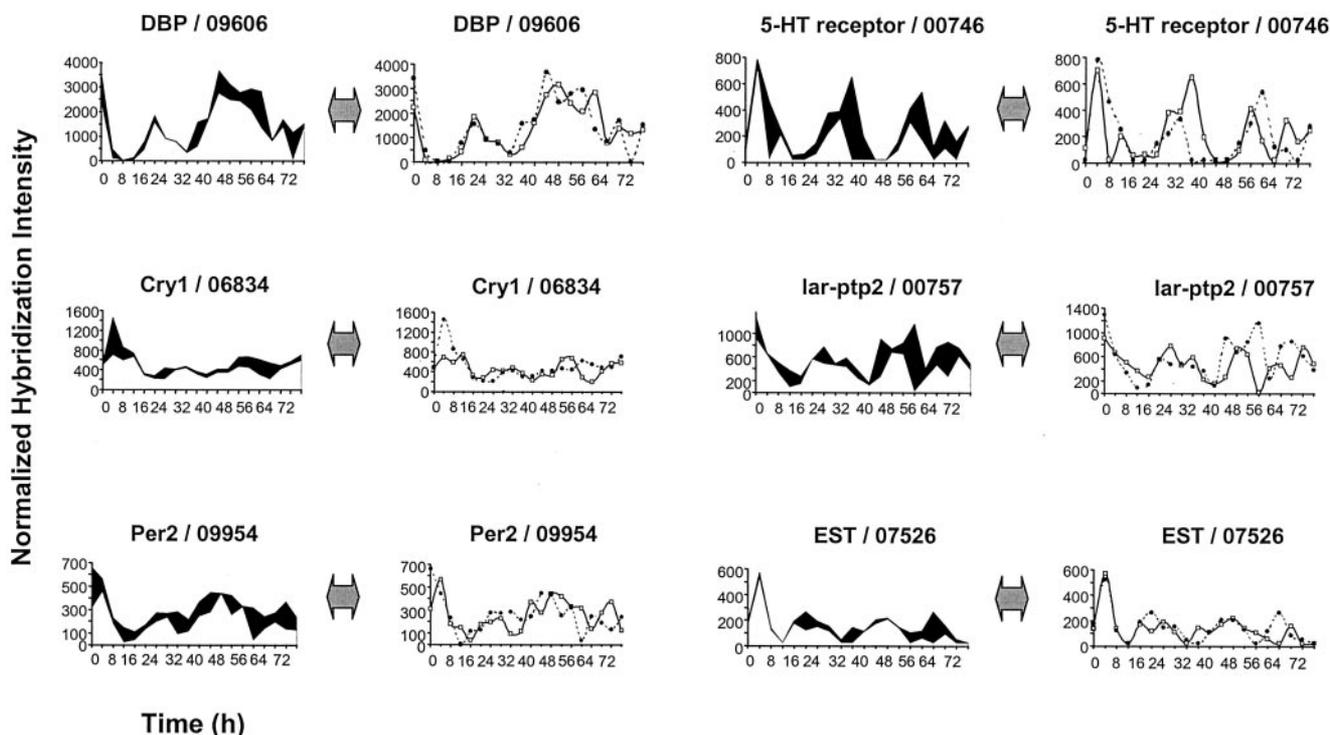


FIG. 1. Circadian expression profiles determined with synchronized rat fibroblasts using the GeneChip® microarray technology. Pairs of expression profiles (normalized hybridization intensity as a function of time (hour)) of selected transcripts are shown. For each pair, the left-hand panel shows the surface delimited by the maximal and minimal values obtained for each time point (duplicated measurements). The right-hand panel displays the same data, but all successive points belonging to one replicate are linked by a solid or dotted line. The 50% v/v serum shock synchronization was done between $t = 0$ and 2 h.

ment (residues 553–783) of the rat cDNA (GenBank accession number M25804) (10). The 36B4 control probe corresponds to a 119-bp fragment (residues 24–143) of the rat acidic ribosomal phosphoprotein P0 cDNA (GenBank accession number X15096). RNase protection assay was performed as described (11). For each time point, 10 μ g of total RNA (also used for chip target preparation) was hybridized overnight at 58 °C with 150 and 300 pg of the rev- α - and 36B4-labeled probes, respectively. Signals were quantified with a STORM 860 bio-analyzer (Molecular Dynamics).

Quantitative Reverse Transcriptase-PCR—For each sample, two cDNAs were synthesized from 2 μ g of DNase treated total RNA previously used for microarray target preparation. Quantitative PCR was performed with the ABI Prism 7700 and TaqMan probe or a Lightcycler using SYBR green detection. Primer and TaqMan probe sequences were the following: A, for rev- α sense, CTCCACAGTGCTGGGAGAGATC, and reverse, CTGGCGTAGACCATTTCAGTGC, probe ATGCCCGTCA-CCTT, and reverse, GGGTTCCTTCCATTTTGTCA, probe TGGGCAAC-TCTGTGGCGTGAAT; C, for dbp sense, CCGTGGAGGTGCTAATGAC-CCT, and reverse, CCTCTGAGAAGCGGTGCCT, probe TGAACCTGA-TCCGGCTGATCTTGCC; D, for glyceraldehyde-3-phosphate dehydrogenase sense, TGCCAAGTATGATGACATCAAGAAG, and reverse, TGCTGTTGAAGTCACAGGAGACA, probe TGGTGAAGCAGGCGGC-

CGAG. Primers used in the light cycler were the following: for ID08649 sense TGAAATCCATTACAGAAAGAGGCTT, and reverse CACCTGT-ATGCAATTCTCTGGGT; for ID6960 sense, TGGCACACACTGAAAT-GTCGT, and reverse TCAAAGACTCAGGACACGGGT; for RING3 sense, GTACCACCCTTCCCCTTT, and reverse, TCAAACAGGGC-CCTTAGAA; for ribosomal phosphoprotein sense CTCAGTGCTCAC-TCCATCA, and reverse, GGGGCTTAGTCAAGAGACC.

Samples contained 1 \times TaqMan Buffer A, 4% glycerol, 5 mM MgCl₂, 200 μ M each dATP, dCTP, and dGTP, 400 μ M dUTP, 0.05 units of AmpliTaq Gold polymerase, 0.01 units of AmpErase UNG, 200 nM primers, and 100 nM TaqMan probe in a 25- μ l volume or Faststart DNA master SYBR green I mix (Roche Molecular Biochemicals), with 3 mM MgCl₂, 500 nM primer each in a 20- μ l volume. The PCR conditions were as follow: with the ABI Prism 7700, 2 min at 50 °C, 10 min at 95 °C, then 40 cycles of 15 s at 95 °C, and 1 min at 60 °C. With the light cycler 10 min at 95 °C, then 10 s 95 °C, 5 s 55 °C, 15 s 72 °C for 40 cycles, followed by a melting curve analysis. The expression levels were normalized either to the levels of glyceraldehyde-3-phosphate dehydrogenase or to the rat acidic ribosomal phosphoprotein P0 (GenBank™ accession number X15096). Relative mRNA abundance was calculated using the comparative Ct method or a standard curve method as described (12).

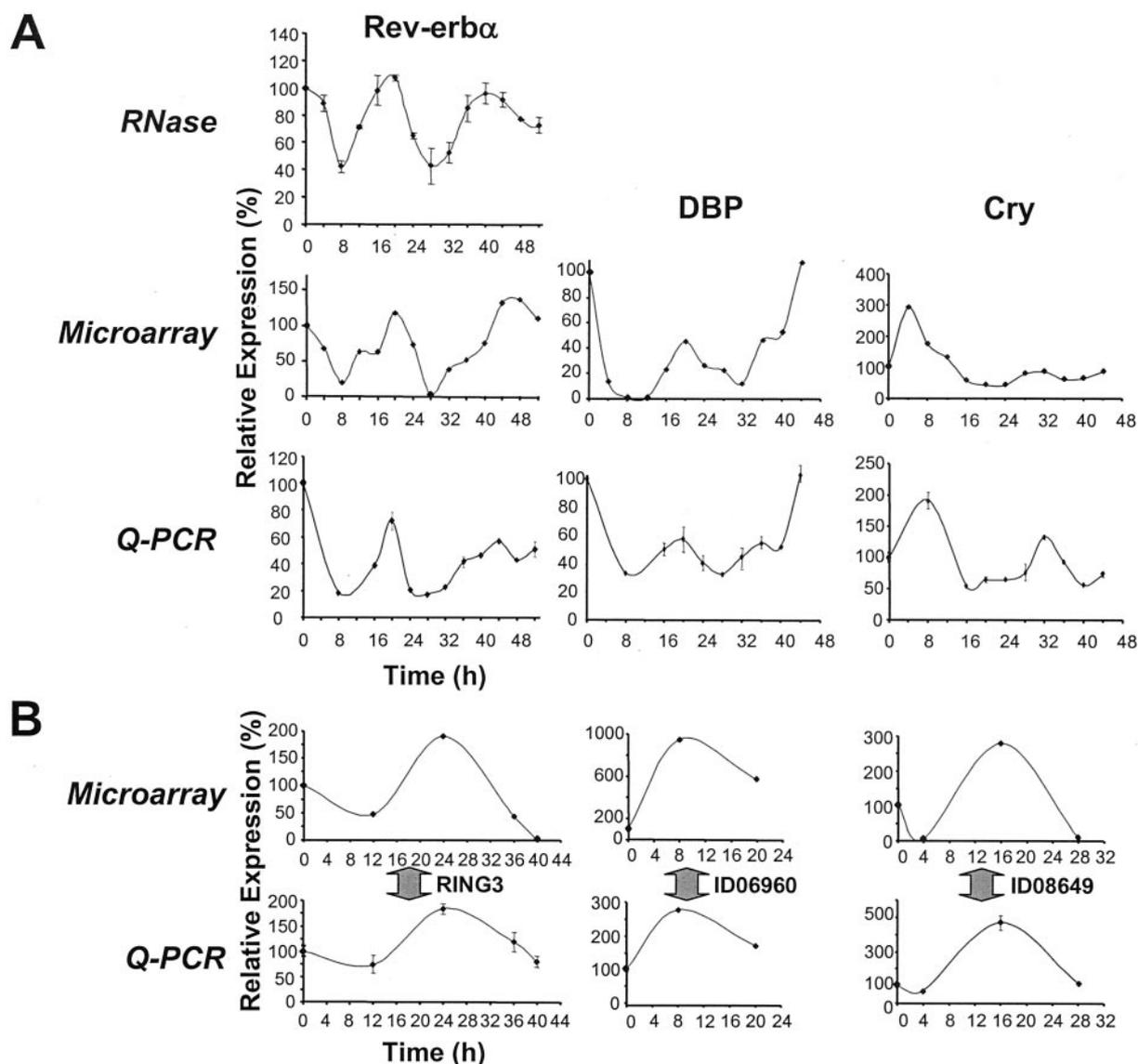


FIG. 2. Qualitative assessment of the GeneChip® microarray technology by comparisons with well established approaches such as RNase protection and Q-PCR. *A*, good reproducibility between the different technologies is shown by the normalized expression profiles of *rev-erba* expression as measured by RNase protection, microarrays, and Q-PCR. *dbp* and *cry1* were also confirmed by Q-PCR. Note that the Q-PCR time points 4 and 12 h were not determined. *B*, comparison of the expression profile obtained by microarray and Q-PCR for RING3 (ID04528), the 26 S proteasome subunit (ID06960), and a putative protein (ID08649). The Q-PCR time points correspond to the first minimum and maximum expression level for a given gene determined by the microarray technology.

RESULTS

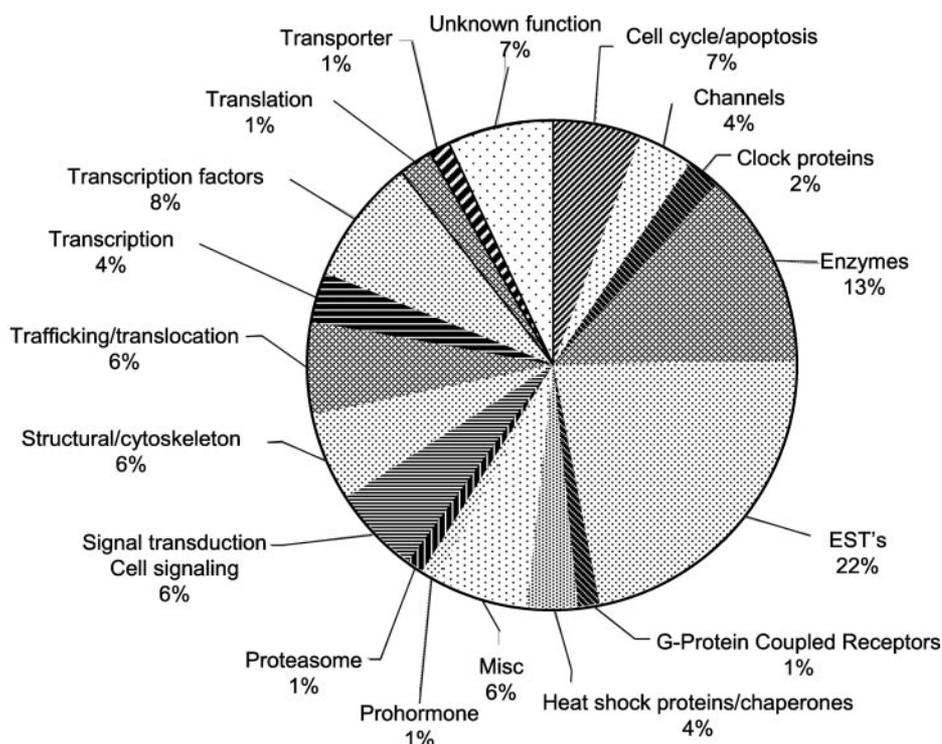
Confluent, non-dividing 3Y1 fibroblasts, were synchronized for 2 h with 50% v/v horse serum. After removal of the medium, fibroblasts were harvested every 4 h (independent duplicates), and total RNA was extracted and cRNA target prepared for hybridization on the microarrays. The 12,204 probe sets present on the microarrays do not represent an equivalent number of genes, because some probe sets possibly test the same gene. Between June 1998 and April 2001 3418 ESTs could be matched to newly sequenced genes, and from these ESTs 24% were matching the same gene. If this percentage is extrapolated to all ESTs, our array included an estimated number of 9957 independent genes.

The normalized mean intensity and standard deviation of hybridization signals for each gene was calculated for each time point. To perform expression-profiling analysis of the experimental data, we used a methodology based on spectral analysis of time series as follows: 4006 probe sets gave a periodic hybridization signal as analyzed by the autoregressive

process; 5936 probe sets gave a constant hybridization signal over all time points; and 2263 had hybridization signals below the detection threshold indicative of non-expressed genes or undetectable transcripts. Because the above numerical analysis was based on the mean hybridization intensities (without considering standard deviations), a subsequent visual inspection of the data (see "Experimental Procedures") was indispensable to further identify transcripts with robust circadian expression patterns. By using these combined data analysis approaches, we have identified 86 transcripts with clear rhythmic and circadian expression profiles as shown in Table I (see also the Supplemental Material containing median hybridization values, first peak, period, and quality and Table II or the raw data excel file *supp_cycling_raw_data.xls*).

Among the 86 transcripts, 20 were known published sequences and 66 were ESTs from a private data base (Incyte). The EST sequences were used to search the public data base GenBank™ for further identification and updating; 22 ESTs were identical to sequences previously published in the data

FIG. 3. Functional classes within the 85 cycling transcripts. The gene products belong to 16 known functional classes, two additional groups represented ESTs or proteins with unknown functions. Percentage (%) represents the fraction of the 85 cycling genes that belong to a given class.



bases (BLAST E value = 0) and 25 were similar to sequences in the data bases (BLAST E value ranging from $1E-178$ to $3E-37$, median = $5E-142$). Therefore, we could identify 67 cycling transcripts present in the public data base as follows: 61 coding for proteins with known function and 6 with unknown function. The 19 remaining sequences were ESTs without known sequence homology (Table I).

Four genes known to have an ~24-h cycle of expression both *in vivo* and in synchronized cell lines (3) are present on the microarray used here. In addition to *rev-erba* and *dbp*, we identified *cry1* and *per2* transcripts after EST identification against known sequences. These two genes were recently shown to be essential protein components of the circadian clock (13, 14). As shown in Fig. 1 and 2A, the four transcripts were cycling in our experiment. Among other cycling transcripts, we found two independent ESTs encoding the same product, DBP. The hybridization signals for both probe sets were almost identical, showing a common cycling expression pattern (data not shown). *dbp* was therefore tested twice with two separate probe sets, hence our identification of 86 sequences with circadian expression corresponds to 85 independent genes. Our data on five known cycling transcripts demonstrated the validity of our experimental set-up and our expression data analysis and indicated a good reproducibility of the microarray methodology used here.

Are the microarray data comparable with those obtained using other techniques such as RNase protection assay or real time quantitative reverse transcriptase-PCR (Q-PCR)? By RNase protection assays and Q-PCR, the *rev-erba* transcript was shown to cycle with a pattern very similar to the one obtained from the microarray analysis (Fig. 2A), and the same was true for *dbp* and *cry1* analyzed by Q-PCR (Fig. 2A). By using Q-PCR, we further confirmed the cycling pattern of eight independent transcripts as shown in Fig. 2B for RING3, ID06960, ID08649 and similar data not shown for *per2*, cyclin D3, ID01672, ID04871, and ID08242. This comparison clearly demonstrated that microarray data were comparable with those obtained with well established methodologies.

The most striking observation from our experimental data

analysis was that circadian patterns of expression were observed for numerous transcripts that encode proteins belonging to several distinct functional classes. We found proteins known to be involved in housekeeping functions, transcription, translation, protein degradation, signaling, cell cycle, transport, and trafficking. These cycling transcripts encoded proteins such as enzymes, structural proteins, transcription factors, receptors, or ion channels (Fig. 3 and Supplemental Material). Are there common features among these cycling gene transcripts such as their period or amplitude levels that could reflect direct interactions, common physiological responses, or specific pathway activations? Hierarchical clustering is widely used to extract patterns from gene expression profiles and to group genes accordingly to reduce the complexity of the data. We clustered the 85 cycling transcripts based on their expression profiles between $t = 0$ and 48 h by a hierarchical correlation method (15). In this experiment, genes displayed the most robust cyclic expression pattern during this time window, and we hypothesized that this would lead to the creation of the most biologically relevant clusters. As shown in Fig. 4, A and B, seven main clusters accounting for 82 genes were identified. Four transcripts had expression patterns too different to belong to any cluster (Fig. 4A and Table I). The individual expression profiles representative of each cluster are shown in Fig. 4B, whereas the median values of the relative hybridization intensities of all transcripts in each cluster are shown in Fig. 4C. This allows the visualization and comparison of the seven cycling patterns. Because clusters 4 and 5 are very similar except for a sharp peak at $t = 4$ h in cluster 4, probably due to the serum addition, they could be considered as a single profile. In this case the first peak appearing after $t = 8$ h in each of the six different clusters is shifted from one to the next by a time lag of 4 h, representing all possible phases of a 24-h rhythm. As expected, gene transcripts known to cycle in phase like *rev-erba* and *dbp* were clustered together in cluster 3.

Although all genes found in a given cluster were roughly cycling according to a circadian rhythm, they did not belong to one particular protein functional class (Table I). On the other hand, a given cluster may contain functionally related genes or

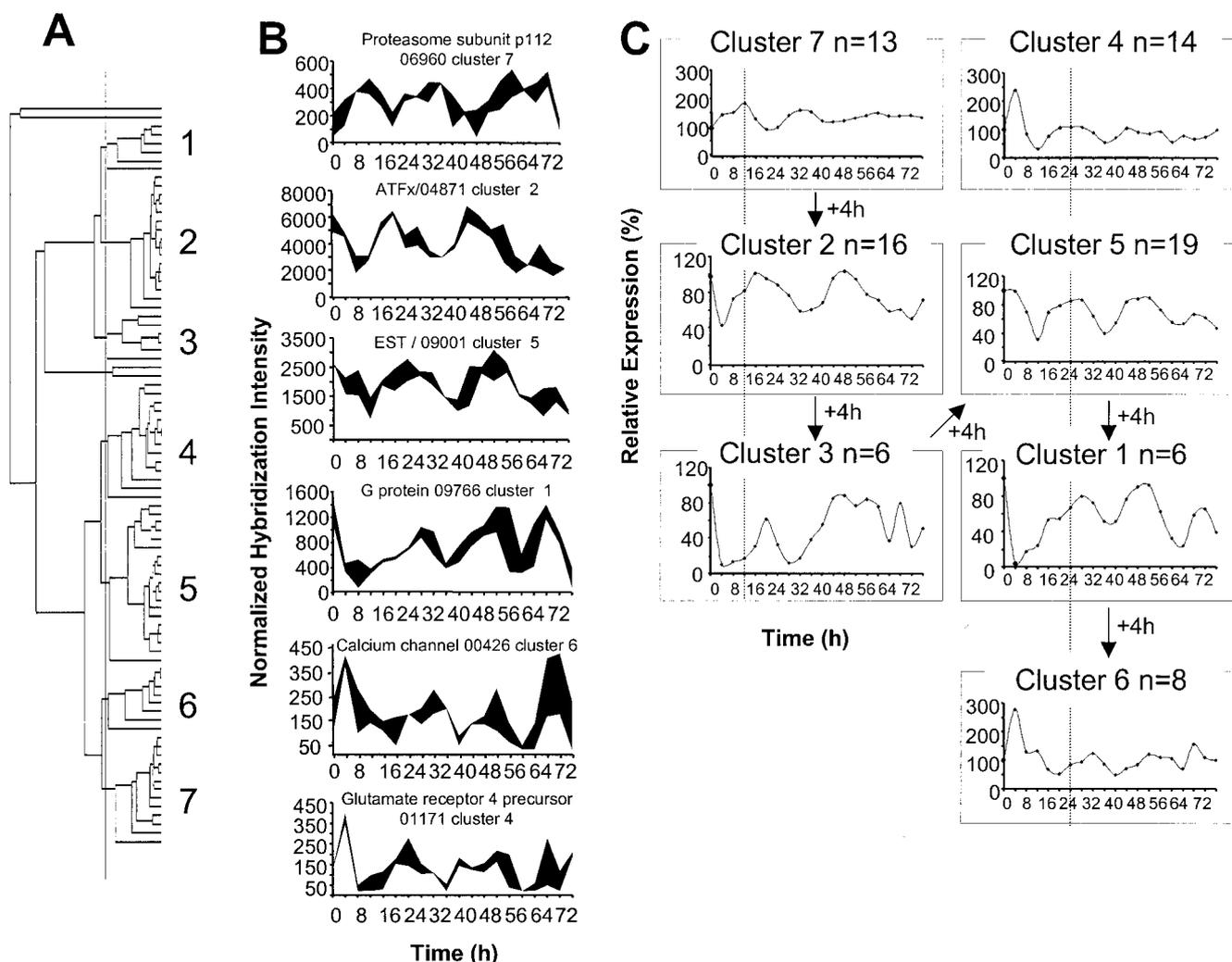


FIG. 4. Clustering of cycling transcripts using hierarchical correlation. *A*, the hierarchical tree shows the clustering of the 85 transcripts based on their circadian expression profiles between $t = 0$ and 48 h (the complete time series are displayed in *B* and *C*). The vertical line in *A* indicates the cut-off correlation value (0.63) used to reveal seven clusters. *B*, expression profiles of representative members for 6 clusters. Normalized hybridization intensity in function of time (hour) is shown. Samples were measured every 4 h for a period of 76 h. *C*, for each time point, the median relative (100% set at $t = 0$ h) expression level of all transcripts in a given cluster is shown, with n equal to the number of circadian transcripts in a given cluster. The 50% v/v serum shock was performed between $t = 0$ and 2 h.

genes whose products may have a higher likelihood of interaction. Here are some examples of such cycling transcripts. Cluster 7 contained both the transcription factor CBF-C (ID01672) that binds to the common CCAAT motif present in several mammalian promoters (16) and the transcription activator TAFII32 (ID10277), part of TFIID, a transcription initiation factor that binds TATA regulatory sequences (17). Cluster 3 contains two previously known clock or clock-driven genes (*rev-erba*, *dbp*), as well as the transcription factor FEZ1 (ID08023), a leucine-zipper protein.

Other genes were also expressed in a circadian fashion. Cluster 4 includes the glutamate receptor 4 (ID01171) transcript, and cluster 6 contains the serotonin 5-HT_{5B} receptor transcript (ID00746). The leukocyte common antigen-related phosphatase (ID00757 in cluster 5) is known to regulate the action of insulin by dephosphorylation of the insulin receptor (18). The endothelin-converting enzyme (ID05028 in cluster 6) is known to produce active endothelin that serves as a potent vasoconstrictor and possesses mitogenic properties, causing proliferation and hypertrophy of vascular smooth muscle, cardiac myocytes, bronchial smooth muscle, and fibroblasts (19). Our data suggest that these functions might be circadian. Similarly, the recently identified granin-like neuroendocrine peptide precursor

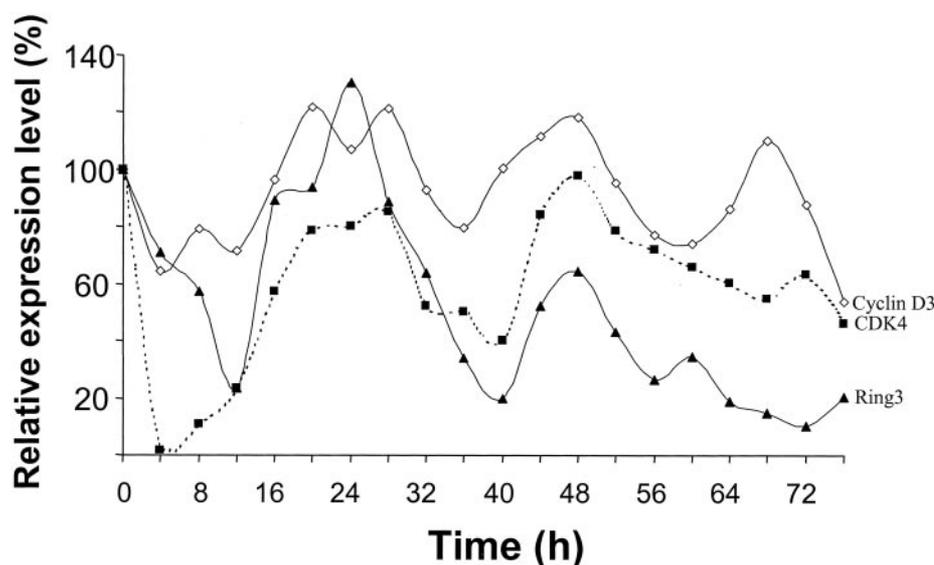
(proSAAS) (20), present as ID10275 in cluster 4, could be a circadian-regulated endogenous inhibitor of prohormone convertase 1, the protease responsible for the processing of many neuroendocrine peptides such as proopiomelanocortin, the precursor of several signaling substances.

Moreover, we identified four cycling transcripts coding for proteins known to play a role in the regulation of the cell cycle. We found, in different clusters, the replication-dependent histone H2A1 (ID08025 in cluster 4), cyclin D3 (ID09966 in cluster 2), cell division protein kinase 4 CDK4 (ID06657 in cluster 1), and RING3 (ID04528 linked to cluster 5). Although the last three transcripts were not clustered together because of different profiles between $t = 0$ and 12 h, they clearly showed the same phase of expression for the first two peaks at 24 and 48 h (Fig. 5).

DISCUSSION

What is the overall percentage of all expressed genes showing a circadian pattern of expression? Accurate gene expression studies rely on the maximal sensitivity and reproducibility of the methodology used to detect mRNA levels. The microarray protocol used here allowed us, in principle, to detect mRNA transcripts present at a frequency of 1/150,000, or equivalent to

FIG. 5. **Circadian expression of 3 cell cycle genes.** Normalized (100% set at $t = 0$ h) mean hybridization intensity of cyclin D3 (open diamond), CDK4 (black square), and RING3 (black triangle) gene expressions are shown. All follow a clear circadian pattern of expression.



2–6 mRNA molecules/cell, as estimated by the hybridization intensity of spiked transcripts (data not shown). Therefore, it should be able to detect the vast majority of transcripts that are typically found at a concentration of 1–10 copies/cell. The presence or absence of a given transcript in a sample was calculated using the present call algorithm of the Affymetrix software. 5,815 sequences, representing 4,361 independent genes, were called present in at least half of the time points. Because 85 independent transcripts are found cycling, it represents a frequency of $\sim 2\%$ of all present genes. By knowing how many genes are expressed in a single cell, one can extrapolate the total number of cycling transcripts. By using the SAGE technology, it has been previously estimated that a single cell expresses about 13,000 genes (21). Therefore, our data indicate that one should expect a total of about 260 genes to follow a circadian rhythm of expression after synchronization.

A recent paper (22) reported the identification of 51 circadian transcripts in liver cells *in vivo* by a differential display technique. 14 of these transcripts were present on our microarrays: four transcripts had hybridization signals below the detection threshold, and nine showed a transient expression peak between time 0 and 24 h as described by Kornmann *et al.* (22), but in our experiment they remained stable for the remaining observation period. Only *dbp* displayed a cyclic expression pattern for more than one cycle. This may indicate that these genes have a circadian expression profile *in vivo* that is not maintained in cultured cells.

The main difficulty we encountered in this study was to define statistically the cycling transcripts. We tested two methods to detect periodic signals, spectral analysis and autocorrelation. These methods were developed and are adequate to analyze electronic signals where at least 1000 time points can easily be sampled, but they show shortcomings in biological systems characterized by a higher variability, varying amplitude of the signal, and a small number of time points available for analysis. While detecting our positive controls, *rev-erba*, *dbp*, *cry*, and *per* as periodically oscillating, these methods were unable to score them as statistically significant. Therefore, we used these mathematical methods as periodicity detector that had to be followed by a qualitative analysis to take into account the standard deviation of the measurement. This may have resulted in a conservative estimation of the number of cycling transcripts.

As the clustering was obtained from a phenomenological data analysis, further molecular studies such as comparisons of

gene promoter regulations will be necessary to reveal why those transcripts are co-regulated in a circadian manner. Although different levels of mRNA transcript do not always translate into equivalent changes in protein levels, our clustering data suggest that gene products with similar profiles might have a greater likelihood of interactions with each other than with those whose transcripts are cycling with a different phase. Also of interest is the circadian expression of the glutamate receptor 4 transcript in cluster 4, which is known to be expressed in the SCN (23) and of the serotonin 5-HT_{5B} receptor transcript in cluster 6. A circadian pattern of expression for the 5HT_{2C} receptor could be shown in the hippocampus (24) as well as for the GluR2/3 subunits of the α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptor in the SCN at the protein level (25). Glutamate and serotonin are thought to mediate and modulate light resetting of the mammalian clock, respectively (26, 27). Circadian expression of these receptors may indicate that the input component of the clock is a target of the circadian oscillator. The most remarkable common physiological feature was observed with four genes implicated in the regulation of the cell cycle. Cyclin D3, CDK4 and RING3 showed the same pattern of expression between 12 and 56 h. Cyclin D3 and CDK4 are known to interact directly (cyclin-D3 activating CDK4) and to play an important role in the G₁/S transition phase in the cell cycle (28). The cyclin-D3 protein amount is regulated at the transcriptional level, and its half-life is 25 min (29). Circadian rhythm and the cell cycle have been linked *in vivo*. In humans, circadian variation are observed in the number of progenitor cells detected in bone marrow (30), as well as in the level of expression of cell cycle-associated proteins (cyclins) in biopsy samples of the oral mucosa (31). A role for CDK4 on circadian rhythms was suggested previously (32), but it could not be determined if the effect on phase shifts of the circadian rhythm by CDK inhibitors were due to the direct inhibition of CDK or due to a general inhibition of transcription. RING3 is a novel serine-threonine kinase located in the nucleus, which can transactivate cyclin (D11, A, and E) and dihydrofolate reductase genes (33). RING3 is implicated in cell cycle, and its expression pattern is circadian and synchronous with those of cyclin D3 and CDK4 in our experiment. Together, these results indicate that transcripts of several proteins involved in the control of the cell cycle are regulated, in non-dividing confluent cells, according to a circadian rhythm.

We conclude that, in synchronized fibroblasts, 2% of all gene transcripts show a circadian pattern of expression. The 85

independent cycling transcripts described and clustered here encode a surprisingly large number of functional protein classes, some representing potentially interesting drug targets for diseases involving the circadian clock such as depression, jetlag, insomnia, and other sleep disorders. Our data analysis also suggests a molecular link or pathway between the cell cycle and endogenous circadian cellular clock functions.

Acknowledgment—We thank M. Hürzeler and C. Keller from AICOS Technologies AG (Basel) for the spectral analysis and C. Broger and M. Neeb (Hoffmann-La Roche) for excellent bioinformatic tools to perform data analysis. We thank J. R. Martin and M. Robinson-Rechard for the critical reading of the manuscript.

REFERENCES

- Dunlap, J. C. (1999) *Cell* **96**, 271–290
- Yagita, K., Tamanini, F., van Der Horst, G. T., and Okamura, H. (2001) *Science* **292**, 278–281
- Balsalobre, A., Damiola, F., and Schibler, U. (1998) *Cell* **93**, 929–937
- Kimura, G., Itagaki, A., and Summers, J. (1975) *Int. J. Cancer* **15**, 694–706
- Rogge, L., Bianchi, E., Biffi, M., Bono, E., Chang, S. Y., Alexander, H., Santini, C., Ferrari, G., Sinigaglia, L., Seiler, M., Neeb, M., Mous, J., Sinigaglia, F., and Certa, U. (2000) *Nat. Genet.* **25**, 96–101
- Box, G. E. P., Jenkins, G. M., and Reinsel, C. (1994) *Time Series Analysis: Forecasting and Control*, pp. 1–598, Prentice Hall, Inc., Englewood Cliffs, NJ
- Martin, R. D. (1980) *Directions in Time Series*, pp. 228–254, Institute of Mathematical Statistics, Hayward, CA
- Martin, R. D. (1981) *Applied Time Series Analysis II*, pp. 683–759, Academic Press, New York
- Rousseeuw, P. J., and Leroy, A. M. (1987) *Robust Regression and Outlier Detection*, pp. 1–329, Wiley Interscience, New York
- Lazar, M. A., Hodin, R. A., Darling, D. S., and Chin, W. W. (1989) *Mol. Cell. Biol.* **9**, 1128–1136
- Modjtahedi, N., Haddada, H., Lamonerie, T., Lazar, E., Lavialle, C., and Brison, O. (1992) *Int. J. Cancer* **52**, 483–490
- Anonymous (1997) *User Bulletin Number 2 ABI Prism 7700 Sequence Detection System*, PerkinElmer Life Sciences
- Miyamoto, Y., and Sancar, A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6097–6102
- Zheng, B., Larkin, D. W., Albrecht, U., Sun, Z. S., Sage, M., Eichele, G., Lee, C. C., and Bradley, A. (1999) *Nature* **400**, 169–173
- Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14863–14868
- Maitly, S. N., and de Crombrughe, B. (1998) *Trends Biochem. Sci.* **23**, 174–178
- Hisatake, K., Ohta, T., Takada, R., Guermah, M., Horikoshi, M., Nakatani, Y., and Roeder, R. G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 8195–8199
- Ren, J. M., Li, P. M., Zhang, W. R., Sweet, L. J., Cline, G., Shulman, G. I., Livingston, J. N., and Goldstein, B. J. (1998) *Diabetes* **47**, 493–497
- Miyauchi, T., and Masaki, T. (1999) *Annu. Rev. Physiol.* **61**, 391–415
- Fricke, L. D., McKinzie, A. A., Sun, J., Curran, E., Qian, Y., Yan, L., Patterson, S. D., Courchesne, P. L., Richards, B., Levin, N., Mzhavia, N., Devi, L. A., and Douglass, J. (2000) *J. Neurosci.* **20**, 639–648
- Velculescu, V. E., Madden, S. L., Zhang, L., Lash, A. E., Yu, J., Rago, C., Lal, A., Wang, C. J., Beaudry, G. A., Ciriello, K. M., Cook, B. P., Dufault, M. R., Ferguson, A. T., Gao, Y., He, T. C., Hermeking, H., Hiraldo, S. K., Hwang, P. M., Lopez, M. A., Luderer, H. F., Mathews, B., Petroziello, J. M., Polyak, K., Zawel, L., and Kinzler, K. W. (1999) *Nat. Genet.* **23**, 387–388
- Kornmann, B., Preitner, N., Rifat, D., Fleury-Olela, F., and Schibler, U. (2001) *Nucleic Acids Res.* **29**, e51 (NAR methods online)
- Stamp, J. A., Piggins, H. D., Rusak, B., and Semba, K. (1997) *Brain Res.* **756**, 215–224
- Holmes, M. C., French, K. L., and Seckl, J. R. (1995) *Brain Res. Mol. Brain Res.* **28**, 186–192
- Chambille, I. (1999) *Brain Res.* **833**, 27–38
- Ding, J. M., Chen, D., Weber, E. T., Faiman, L. E., Rea, M. A., and Gillette, M. U. (1994) *Science* **266**, 1713–1717
- Pickard, G. E., and Rea, M. A. (1997) *Biol. Cell* **89**, 513–523
- Sherr, C. J. (1995) *Trends Biochem. Sci.* **20**, 187–190
- Reisman, D., and Thompson, E. A. (1995) *Mol. Endocrinol.* **9**, 1500–1509
- Smaaland, R., Laerum, O. D., Sothorn, R. B., Sletvold, O., Bjerknes, R., and Lote, K. (1992) *Blood* **79**, 2281–2287
- Bjarnason, G. A., Jordan, R., and Sothorn, R. (1999) *Am. J. Pathol.* **154**, 613–622
- Sankrithi, N., and Eskin, A. (1999) *J. Neurochem.* **72**, 605–613
- Denis, G. V., Vaziri, C., Guo, N., and Faller, D. V. (2000) *Cell Growth Differ.* **11**, 417–424

Circadian Regulation of Diverse Gene Products Revealed by mRNA Expression Profiling of Synchronized Fibroblasts

Christophe Grundschober, Franck Delaunay, Anja Pühlhofer, Gerard Triqueneaux, Vincent Laudet, Tamas Bartfai and Patrick Nef

J. Biol. Chem. 2001, 276:46751-46758.

doi: 10.1074/jbc.M107499200 originally published online October 11, 2001

Access the most updated version of this article at doi: [10.1074/jbc.M107499200](https://doi.org/10.1074/jbc.M107499200)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Supplemental material:

<http://www.jbc.org/content/suppl/2001/12/07/276.50.46751.DC1>

This article cites 28 references, 10 of which can be accessed free at <http://www.jbc.org/content/276/50/46751.full.html#ref-list-1>