Molecular-timetable methods for detection of body time and rhythm disorders from single-time-point genome-wide expression profiles

Hiroki R. Ueda*†‡§, Wenbin Chen*, Yoichi Minami*, Sato Honma¶, Kenichi Honma¶, Masamitsu Iino†, and Seiichi Hashimoto*

*Molecular Medicine Laboratories, Institute for Drug Discovery Research, Yamanouchi Pharmaceutical Company, Limited, 21 Miyukigaoka, Tsukuba, Ibaraki 305-8585, Japan; †Department of Pharmacology, Graduate School of Medicine, University of Tokyo, 7-3-1 Bunkyo-ku, Tokyo 113-0033, Japan; ‡Laboratory for Systems Biology, Center for Developmental Biology, RIKEN, 2-2-3 Minatojiima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan; and §Department of Physiology, Hokkaido University Graduate School of Medicine, Sapporo 060-8638, Japan

Edited by Michael S. Waterman, University of Southern California, Los Angeles, CA, and approved May 27, 2004 (received for review March 17, 2004)

Detection of individual body time (BT) via a single-time-point assay has been a longstanding unfulfilled dream in medicine, because BT information can be exploited to maximize potency and minimize toxicity during drug administration and thus will enable highly optimized medication. To achieve this dream, we created a "molecular timetable" composed of >100 "time-indicating genes," whose gene expression levels can represent internal BT. Here we describe a robust method called the "molecular-timetable method" for BT detection from a single-time-point expression profile. The power of this method is demonstrated by the sensitive and accurate detection of BT and the sensitive diagnosis of rhythm disorders. These results demonstrate the feasibility of BT detection based on single-time-point sampling, suggest the potential for expression-based diagnosis of rhythm disorders, and may translate functional genomics into chronotherapy and personalized medicine.

DNA microarray | gene expression | chronotherapy

Diverse physiological and metabolic processes exhibit circadian rhythms, which are endogenous self-sustained oscillations with a period of ≈24 hours. In mammals, several clock genes, including *Clock*, *Bmal1/Mop3*, *Per1*, *Per2*, *Cry1*, *Cry2*, *Ckiε*, and *RevErbAα*, and clock-controlled transcription factors, including *Dbp, E4bp4, Dec1/Stra13, Dec2, Per3, Npas2, RevErbAβ, Rorα, Rorβ*, and *Rorγ*, regulate, at least in part, gene expression in central and/or peripheral clocks (1). Reflecting the temporal changes in gene expression in central and peripheral clocks (2–5), the potency and/or toxicity in drug administration depend on an individual's body time (BT) (6–10). It has been suggested that drug administration at the appropriate BT can improve the outcome of pharmacotherapy by maximizing potency and minimizing the toxicity of the drug (11), whereas drug administration at an inappropriate BT can induce severe side effects (12). Despite the effectiveness and importance of such BT-dependent therapy, termed "chronotherapy" (6–10), its clinical use has been obstructed by the lack of clinically applicable methods for BT detection. To address these difficulties, we attempted to create standard expression profiles, termed a "molecular timetable," composed of >100 "time-indicating genes" and their expression levels during the course of a day, and then to apply this timetable to BT detection.

Materials and Methods

Animals. To select time-indicating genes and construct their standard expression profiles, we analyzed the previously obtained genome-wide expression profiles (5) from pooled livers of four Balb/c mice (male) every 4 h over 2 d under 12-h light/12-h dark (LD) or constant-dark (DD) conditions. We independently sampled livers from eight individual Balb/c mice (male) at Zeitgeber time (ZT)12 ($n = 4$), ZT6 ($n = 1$), ZT18 ($n = 1$), Circadian time (CT)6 ($n = 1$), and CT18 ($n = 1$) to verify the capability of the molecular-timetable method. ZT is used for a timescale under LD conditions, whereas CT is used for a timescale under DD conditions. ZT0 represents lights on, and ZT12 represents lights off, whereas CT0 represents subjective dawn, and CT12 represents subjective dusk. We use the term "subjective," because there are no external time cues in the DD condition. We sampled livers from seven individual *Clock/Clock* homozygous mutant mice (male) at ZT12 ($n = 4$) or ZT8 ($n = 3$) and three individual Balb/c mice (male) at ZT8 ($n = 3$) to verify the feasibility of expression-based diagnosis of circadian rhythm disorders. We further sampled livers from three individual C3H mice (male) at ZT12 ($n = 3$) to verify the feasibility of the molecular-timetable method for individuals with heterogeneous genetic background. All these Balb/c, C3H, and *Clock/Clock* homozygous mutant mice were adapted under LD conditions for 2 weeks from 5 weeks postpartum and then sampled.

Microarray Experiments. Total RNA was prepared by using Trizol reagent (GIBCO/BRL). cDNA synthesis and cRNA labeling reactions were performed as described (5). Affymetrix high-density oligonucleotide arrays (Murine Genome Array U74A, Version 1.0, measuring 9,977 independent transcripts) were hybridized, stained, and washed according to the Technical Manual (Affymetrix). Affymetrix software was used to determine the average difference (AD) between perfectly matched probes and single-base-pair-mismatched probes. The AD of each probe was then scaled globally so that the total AD of each microarray was equal. The resulting AD values reflect the abundance of a given mRNA relative to the total RNA population and were used in all subsequent analyses.

Time-Indicating Genes. To select time-indicating genes whose expression exhibits circadian rhythmicity with high amplitude, the expression profile of each gene was analyzed through two filters, one for circadian rhythmicity and the other for high amplitude. To extract genes with circadian rhythmicity under both LD and DD conditions, we first calculated the correlation over time between 12-point time courses under LD (or DD) conditions and cosine curves of defined periods and phases. We prepared cosine curves of 24-h periodicity with peaks from 0 to 24 h in increments of 10 min, yielding a total of 144 test cosine curves, and calculated the correlation value of the best-fitted cosine curve for each probe set. We selected the probe sets

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: LD, light/dark; DD, constant dark; ZT, Zeitgeber time; CT, Circadian time; BT, body time(s).

§To whom correspondence should be addressed. E-mail: uedah-tky@umin.ac.jp.

whose best correlation values in LD and DD conditions were both above the cutoff correlation value of 0.8. To further extract genes with high amplitude under both LD and DD conditions, we calculated the coefficient of variation of the expression levels under LD (or DD) conditions, which is defined as its SD divided by the average of expression levels. We selected the probe sets whose coefficients of variation in LD and DD conditions were both above the cutoff value of 0.15. Based on two successive filtrations, 168 genes (182 probes) in the liver were identified.

Estimation of Molecular Peak Time. To estimate the peak time of cycling genes, we tested for correlation over time between the 12-point time course of each gene and the 24-h period cosine curves with different peak times at 10-min intervals. We estimated the peak time of each cycling gene from the peak time of the best-fitted cosine curve and defined it as the molecular peak time.

Molecular-Timetable Method. We first normalized an expression level, X_i, of the time-indicating gene i ($i = 1 \ldots N$) using its average μ_i and SD σ_i in the molecular timetable. The normalized expression level Y_i is described as follows: $Y_i = (X_i - \mu_i)/\sigma_i$. We created an expression profile $\{t_i, Y_i\}$ ($i = 1 \ldots N$) composed of the molecular peak time t_i of gene i and its normalized expression level Y_i. To estimate the BT of an expression profile, we calculated the correlation over genes ($i = 1 \ldots N$) between an expression profile $\{t_i, Y_i\}$ and a 24-h cosine curve $\{t_i, \sqrt{2}\ Cos(2\pi(t_i - b)/24)\}$ with a certain phase b ($0 \leq b < 24$). The amplitude of the cosine curve is set to $\sqrt{2}$, so that the SD of the normalized expression level Y_i matches the SD of a continuous cosine waveform ($\sqrt{1/24 \int_0^{24} Cos^2(2\pi(t - b)/24)\, dt} = 1/\sqrt{2}$). We prepared 24-h cosine curves with a phase b from 0 to 24 h in increments of 10 min. We then selected the best-fitted cosine curve that gave the best correlation value c. We noted that the best correlation value c is always positive, because we calculated the maximum value of correlation between an expression profile and 144 test cosine curves. We also noted that the phase of the best-fitted cosine curve (b_c) indicates an estimated BT.

Statistical Significance in the Detection of BT. To evaluate the statistical significance of BT estimation, we generated a random expression profile $\{t_i, Y_r\}$ ($i = 1 \ldots N$), where Y_r represents a random variable following the distribution of Y_i and then calculated the correlation value c^r and the phase b_c^r of the best-fitted cosine curve as described above. We repeated this procedure 10,000 times to create the distribution of correlation value c^r and the phase b_c^r (e.g., Fig. 4 *A* and *B*, which is published as supporting information on the PNAS web site). The probability (termed P_r value) that a random expression profile has a best-fitted cosine curve giving correlations equal to or greater than those of the real expression profiles was determined from the distribution of correlation value (c^r) from 10,000 random expression profiles.

Estimation of Measurement Noise. To estimate the measurement noise of eight Bablc/C expression profiles (see *Animals* above) at ZT12 ($n = 4$), ZT6 ($n = 1$), ZT18 ($n = 1$), CT6 ($n = 1$), and CT18 ($n = 1$), we first calculated the difference d_i between a real and an estimated expression level of gene i ($i = 1 \ldots N$). d_i is defined as $d_i \equiv Y_i - \sqrt{2}\ Cos(2\pi(t_i - b_c)/24)$, where b_c represents estimated BT. We then calculated the SD of d_i over all time-indicating genes and defined it as the measurement noise. Measurement noises of wild-type expression profiles range from 86% to 108% (95 ± 8% for mean ± SD).

Statistical Significance in the Detection of Circadian Rhythm Disorders. To evaluate the statistical significance of circadian rhythm disorders, we generated a control expression profile with 100% measurement noise $\{t_i, \sqrt{2}\ Cos(2\pi(t_i/24)) + d_r)\}$ ($i = 1 \ldots N$), where d_r represents a random variable following the distribution of d_i, and then calculated the correlation value c^c and the phase b_c^c of the best-fitted cosine curve, described above. We repeated this procedure 10,000 times to create the distribution of correlation value c^c and the phase b_c^c (e.g., Fig. 4 *A* and *B*). The probability (termed P_c value) that a control expression profile has a best-fitted cosine curve giving a correlation equal to or lower than those of the observed expression profiles was determined from the distribution of the correlation value c^c from 10,000 control expression profiles.

Sensitivity and Specificity of Molecular-Timetable Methods. We can calculate the sensitivity and specificity of the molecular timetable method in the presence of a certain level of measurement noise (100% or 200% measurement noise). First, we generated 10,000 control or random expression profiles in the presence of a certain measurement noise level and then calculated the distribution of c^c or c^r from 10,000 control or random expression profiles as described above. We determined the P_c or P_r value at the certain threshold value of correlation (c) from the distribution of c^c or c^r, respectively. Sensitivity (S_c) is defined as $S_c(c) \equiv 1 - P_c(c)$, indicating the probability of true positives or the probability of circadian rhythmicity (control expression profile), correctly identified by the test as meeting a certain threshold value of correlation. Specificity (S_r) is defined as $S_r(c) \equiv 1 - P_r(c)$, indicating the probability of true negatives or of circadian rhythm disorder (random expression profile), correctly identified by the test as not meeting a certain threshold value of correlation. There are tradeoffs between sensitivity and specificity, because sensitivity monotonically decreases, whereas specificity monotonically increases, with the threshold correlation value. This tradeoff relationship is plotted as the receiver operating characteristic (ROC) curve (e.g., Fig. 4*C*). The ROC curve is defined as $\{S_r(c), S_c(c)\}$ ($0 \leq c \leq 1$).

Performance of Molecular-Timetable Methods with Different Numbers of Time-Indicating Genes with Simulated Expression Profiles. Performance of the molecular-timetable method with N time-indicating genes ($N \in \{3, 5, 10, 20, 30, 50, 100, 182\}$) was calculated with simulated expression profiles in the presence of 100% or 200% measurement noise. First, we randomly selected N from 168 time-indicating genes. Then, we generated control $\{t_i, \sqrt{2}\ Cos(2\pi(t_i/24)) + d_r\}$ and random expression profiles $\{t_i, Y_r\}$ in the presence of 100% measurement noise, where d_r and Y_r represent random variables following the distributions of d_i and Y_i, respectively. In the presence of 200% noise, we generated control $\{t_i, \sqrt{2}\ Cos(2\pi(t_i/24)) + 2d_r\}$ and random expression profiles $\{t_i, 2Y_r\}$. We repeated these procedures 10,000 times and then calculated the distribution of the correlation value (c^c and c^r) and the phase (b_c^c and b_c^r) from 10,000 control and random expression profiles and determined the P_c and P_r values, sensitivity S_c, and specificity (S_r) at a certain threshold value of correlation, as described above (Fig. 5, which is published as supporting information on the PNAS web site).

Performance of Molecular-Timetable Methods with Different Numbers of Time-Indicating Genes with Real Expression Data. Performance of the molecular-timetable method with N time-indicating genes ($N \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150\}$) was calculated with real expression profiles. To select N time-indicating genes from 168 time-indicating genes in the order of high-amplitude circadian rhythmicity, we first calculated the sum of correlation in LD, correlation in DD, coefficient of variation in LD, and coefficient of variation in DD for each gene. We then selected the top N time-indicating genes in the order of this sum.

To determine a cutoff correlation value in the detection of

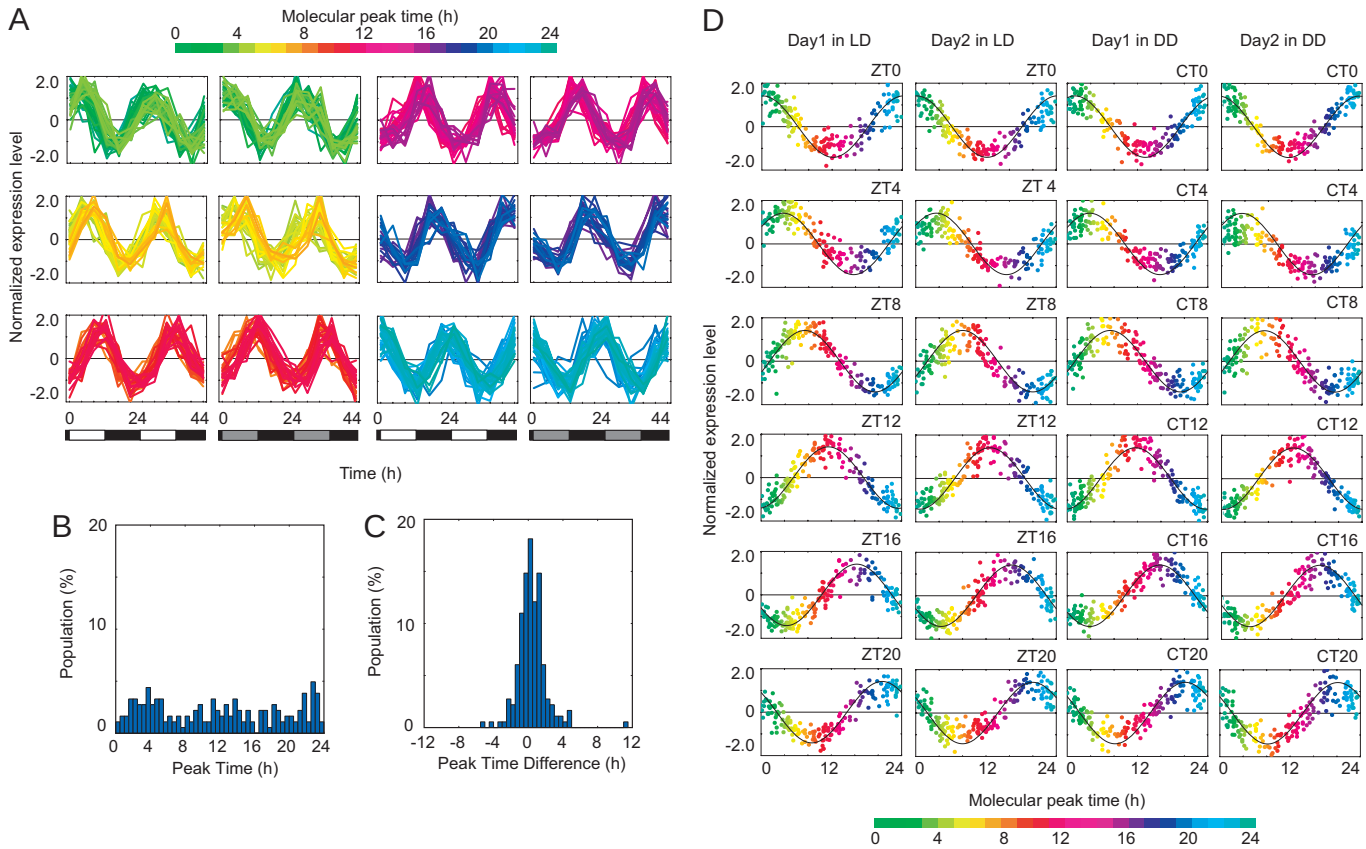

Fig. 1. Time-indicating genes in the mouse liver. (*A*) Temporal expression data of 168 time-indicating genes (182 probes) in the mouse liver. The colors, in ascending order from green to red to blue, represent the molecular peak time of time-indicating genes (the color code is represented above the diagrams). An expression level of each time-indicating gene is subtracted by its average and divided by its SD over 12-point time courses. The blocked horizontal bars below the diagrams correspond to the LD schedule. White blocks indicate periods of exposure to light, whereas black blocks correspond to periods of darkness, and gray blocks indicate the period of subjective light under constant darkness. (*B*) Distribution of the average of molecular peak time between LD and DD conditions. Average molecular peak times are distributed from 0 to 24 h. (*C*) Distribution of differences of molecular peak times between LD and DD conditions. Molecular peak times in LD and DD conditions are similar to each other. (*D*) Expression profiles of 168 time-indicating genes in the mouse liver at different ZT under LD conditions or CT under DD conditions. The best-fitted cosine curve is represented, and its peak indicates the estimated BT.

circadian rhythmicity and circadian rhythm disorders, we first calculated the receiver operating characteristic (ROC) curve of the molecular-timetable method with N time-indicating genes in the presence of 100% measurement noise, as described above. We then determined a cutoff correlation value c ($0 \leq c \leq 1$) as the value yielding equal sensitivity and specificity in the ROC curve [$S_r(c) = S_c(c)$]. We noted that c is a monotonically decreasing function of N.

To calculate sensitivity, specificity, and accuracy with real expression data, we examined 11 wild-type Balb/c expression profiles at ZT12 ($n = 4$), ZT6 ($n = 1$), ZT18 ($n = 1$), CT6 ($n = 1$), CT18 ($n = 1$), and ZT8 ($n = 3$), and 7 *Clock*/*Clock* mutant expression profiles at ZT12 ($n = 4$) and ZT8 ($n = 3$). We calculated sensitivity, which is defined as the percent of wild-type Balb/c expression profiles ($n = 11$), with a correlation value higher than a cutoff. We also calculated specificity, which is defined as the percent of *Clock*/*Clock* mutant expression profiles ($n = 7$), with a correlation value lower than a cutoff. To evaluate the accuracy of molecular-timetable methods, we calculated mean estimation error in BT detection, which is defined as the average of absolute differences between sampling times and estimated BT of wild-type Balb/c expression profiles ($n = 11$). Calculated sensitivity, specificity, and mean errors are plotted along with N time-indicating genes (Fig. 6, which is published as supporting information on the PNAS web site).

Results and Discussion

Recently, we, along with others, performed genome-wide gene expression analyses using high-density DNA microarrays to identify clock-controlled genes in the mouse central [suprachiasmatic nucleus (SCN)] and peripheral (liver) clocks (2–5). We analyzed the previously obtained genome-wide expression profiles from four pooled livers every 4 h over 2 d under LD or DD conditions and found 168 time-indicating genes (Fig. 1 and Table 1, which is published as supporting information on the PNAS web site) whose expression exhibits high circadian rhythmicity and whose peak time, termed molecular peak time, can therefore indicate the time of day (see *Materials and Methods*). Importantly, the molecular peak times of time-indicating genes were distributed over 24 h (Fig. 1*B*). Moreover, each time-indicating gene exhibited similar expression patterns under LD and DD conditions (Fig. 1*A*), and the molecular peak times of each gene under LD and DD conditions were similar to each other (Fig. 1*C*), suggesting that expression profiles of time-indicating genes can indicate BT, the endogenous state of the circadian clock.

To verify this possibility, we attempted to extract the BT information from the expression profiles of 168 time-indicating genes (see *Materials and Methods*). As expected, BT information can be extracted from the expression profiles of 168 time-indicating genes with high accuracy and estimation errors from 0.0 to 1.3 h (0.4 ± 0.3 h for mean ± SD, Fig. 1*D*). For instance, at ZT0 (the beginning of day) or CT0 (the beginning of a

subjective day), dawn-indicating genes (green), whose molecular peak times are approximately ZT0 or CT0, are highly expressed, whereas dusk-indicating genes (red), whose molecular peak times are approximately ZT12 or CT12, are expressed at a low level (Fig. 1*D*). At ZT12 or CT12, on the other hand, dawn-indicating genes are expressed at a low level, whereas dusk-indicating genes are highly expressed (Fig. 1*D*). In both cases, the normalized expression levels of day-indicating genes (yellow-green to orange) or night-indicating genes (purple to blue) gradually change from higher to lower expression levels or from lower to higher levels along their molecular peak times. Such gradual and cyclic changes in expression along the molecular peak time line allow us to extract the BT from the expression profiles via cosine curve fitting (Fig. 1*D*). The peak of the best-fitted cosine curve to an expression profile of 168 time-indicating genes indicates the estimated BT. For example, the BT extracted from expression profiles at ZT0 on the first day under LD or at CT0 on the first day under DD are BT0.2 and BT1.3, respectively. We termed this the molecular-timetable method (see *Materials and Methods*).

To demonstrate the capability of this molecular-timetable method, we attempted to infer BT from independent samples. We obtained fresh liver samples from four individual mice at ZT12, which was thought to be one of the noisiest time points, because lights out at that time resets the phase of circadian clocks in species, such as mice, with a free running period shorter than 24 h. We then measured the expression profiles of 168 time-indicating genes (Fig. 2*A*). Using the molecular-timetable method, we significantly detected the circadian rhythmicity in all expression profiles of these samples ($P < 0.0001$, Fig. 2*A*), and the estimated BT indicated BT12.8, BT10.2, BT11.2, and BT11.7, respectively (Fig. 2*A*). These results suggest that BT can be accurately inferred from individual expression profiles with estimation errors from 0.3 to 1.8 h (1.0 ± 0.6 h for mean ± SD, Fig. 2*A*). To further demonstrate the capability of the molecular-timetable method, we obtained new liver samples from individual mice at ZT6, ZT18, CT6, and CT18, all times that were different from the time points used to construct the molecular timetable, and measured the expression profiles of 168 time-indicating genes (Fig. 2*B*). We significantly detected the circadian rhythmicity in all expression profiles of ZT6, ZT18, CT6, and CT18 samples ($P < 0.0001$, Fig. 2*B*) and found that the estimated BT indicated BT6.7, BT19.3, BT5.7, and BT20.0, respectively (Fig. 2*B*). These results suggest that BT can be accurately inferred from expression profiles at time points that were not used to construct the molecular timetable, with estimation errors from 0.3 to 2.0 h (1.1 ± 0.7 h for mean ± SD, Fig. 2*B*).

The molecular-timetable method can be applied for the diagnosis of circadian rhythm disorders (see *Materials and Methods*). To demonstrate the feasibility of expression-based diagnosis of circadian rhythm disorders, we sampled livers at ZT12 from four individual *Clock*/*Clock* homozygous mutant mice, which were known to have altered circadian behavioral rhythms (13, 14), and then measured the expression profiles of 168 time-indicating genes (Fig. 3*A*). We significantly detected the rhythm disorders in all expression profiles of *Clock*/*Clock* homozygous mutant samples at ZT12 ($P < 0.0001$, Fig. 3*A*). To further demonstrate the feasibility of the expression-based diagnosis of circadian rhythm disorders, we simultaneously sampled livers from three individuals of wild-type (+/+) mice and *Clock*/*Clock* mutant mice at ZT8 and measured the expression profiles for 168 time-indicating genes (Fig. 7, which is published as supporting information on the PNAS web site). We significantly detected the circadian rhythmicity in all expression profiles of wild-type (+/+) samples ($P < 0.0001$, Fig. 7) and found that the estimated BT indicated BT7.3, BT6.7, and BT9.5 (Fig. 7), which were around sampling time ZT8. On the other hand,

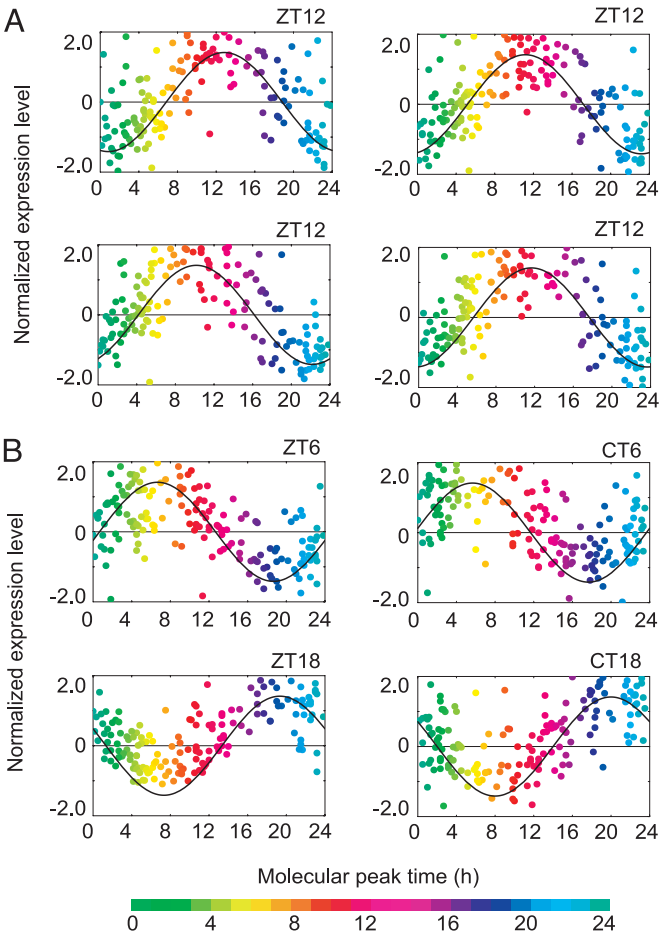

Fig. 2. Significant and quantitative detection of BT from individual expression profiles at ZT12 (*A*) or ZT6, ZT18, CT6, and CT18 (*B*). The colors, in ascending order from green to red to blue, represent the molecular peak time of time-indicating genes (the color code is represented below the diagrams). An expression level of each time-indicating gene is subtracted by its average and divided by its SD in the molecular timetable. The best-fitted cosine curve is represented, and its peak indicates the estimated BT, BT12.8 (*A Upper Left*), BT10.2 (*A Lower Left*), BT11.2 (*A Upper Right*), and BT11.7 (*A Lower Right*), and BT6.7 (*B Upper Left*), BT19.3 (*B Lower Left*), BT5.7 (*B Upper Right*), and BT20.0 (*B Lower Right*).

we significantly detected rhythm disorders in all expression profiles of *Clock*/*Clock* homozygous mutant samples at ZT8 ($P < 0.0001$, Fig. 7). These results suggest that the molecular-timetable method can be applied not only for BT detection but also for the detection of circadian rhythm disorders.

In clinical situations, methods for BT detection should be applicable for populations with heterogeneous genetic backgrounds. To demonstrate the capability of the molecular-timetable method for individuals with heterogeneous genetic backgrounds, we attempted to apply the molecular-timetable method for other inbred strains with different genetic backgrounds from the original strain used for the construction of the molecular timetable. We sampled livers at ZT12 from three individual C3H mice, which were different from the original Balb/c strains used to construct the molecular timetable, and measured expression profiles for 168 time-indicating genes (Fig. 3*B*). We significantly detected the circadian rhythmicity in all expression profiles of C3H samples ($P < 0.0001$, Fig. 3*B*) and found that the estimated BT indicated BT12.3, BT11.3, and BT11.3, respectively (Fig. 3*B*). These results suggest that BT can be accurately inferred from the expression profiles of individuals

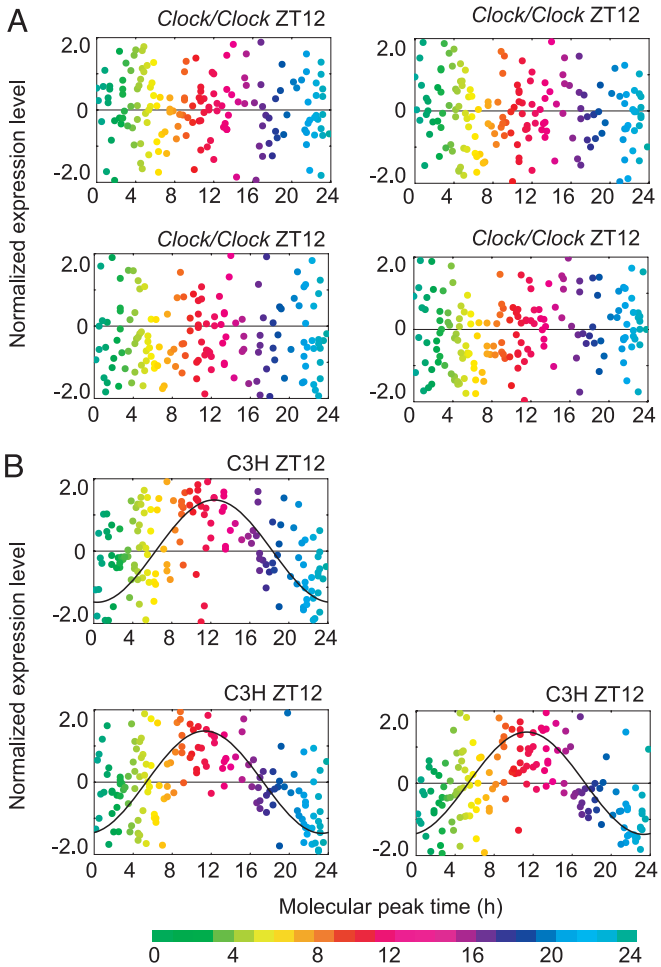

Fig. 3. Significant detection of BT and rhythm disorders from expression profiles of *Clock*/*Clock* mutant mice at ZT12 (*A*) and C3H mouse at ZT12 (*B*). The colors, in ascending order from green to red to blue, represent the molecular peak time of time-indicating genes (the color code is represented below the diagrams). An expression level of each time-indicating gene is subtracted by its average and divided by its SD in the molecular timetable. The best-fitted cosine curve is represented, and its peak indicates the estimated BT, BT12.3 (*B Upper Left*), BT11.3 (*B Lower Left*), and BT11.3 (*B Lower Right*).

with heterogeneous genetic backgrounds, with estimation errors from 0.3 to 0.7 h (0.6 ± 0.2 h for mean ± SD, Fig. 3*B*).

In this study, we used mouse livers to construct the molecular timetable and demonstrated the feasibility of the molecular-timetable method based on single-time-point sampling for detection of BT and diagnosis of circadian rhythm disorders. Other tissues besides the liver can be used to construct molecular timetables, because there are circadian oscillators scattered throughout the body in various tissues and organs, including the mouse suprachiasmatic nucleus (SCN), liver, skeletal muscle, lung, cornea, kidney, heart, aorta, skin, oral mucosa, and white blood cells (2–5, 15–19). Genome-wide expression analyses in some of these tissues have been performed and have demonstrated that between 2% and 10% of the analyzed genes exhibit circadian oscillations in RNA expression levels (2–5). In fact, we analyzed the previous genome-wide expression data for the mouse SCN (5) and found that BT can be represented in the expression profiles of SCN time-indicating genes (Table 2, which is published as supporting information on the PNAS web site), with low estimation errors from 0.3 to 1.8 h (1.0 ± 0.4 h for mean ± SD, Fig. 8, which is published as supporting information on the PNAS web site).

The molecular-timetable method can be also applied to organisms other than the mouse because a wide range of organisms, including human, rat, *Drosophila*, *Arabidopsis*, *Neurospora*, and cyanobacteria, are known to have circadian clocks (20), and genome-wide expression analyses in some of these organisms have shown that a substantial portion of the analyzed genes exhibit circadian oscillations in RNA expression levels (21–27). Actually, we analyzed the previously obtained genome-wide expression data of *Drosophila* heads (24) and found that BT can be represented in the expression profiles of 113 time-indicating genes of *Drosophila* (Table 3, which is published as supporting information on the PNAS web site) with low estimation errors from 0 to 1.5 h (0.4 ± 0.4 h for mean ± SD, Fig. 9, which is published as supporting information on the PNAS web site). We applied this molecular timetable for the detection of BT and the diagnosis of circadian rhythm disorders in *Drosophila*. We simultaneously sampled heads from wild-type (+/+) and *Drosophila Clock* homozygous mutant (28) (*dClock*/*dClock*) flies at ZT13 and obtained expression profiles for 113 time-indicating genes (Fig. 10, which is published as supporting information on the PNAS web site). We significantly detected the circadian rhythmicity in all expression profiles of wild-type samples (+/+) at ZT13 ($P < 0.0001$, Fig. 10) and found that the estimated BT indicated BT13.7 and BT14.5, respectively (Fig. 10). On the other hand, rhythm disorders were significantly detected from *dClock*/*dClock* expression profiles ($P < 0.0001$, Fig. 10). These results suggest that the molecular-timetable method can be applied for the detection of BT and circadian rhythm disorders in organisms other than mouse.

Performance of the molecular-timetable method is characterized by three measures specificity, sensitivity, and estimation error (Fig. 4 and see *Materials and Methods*). The first measure, sensitivity, is defined as the percent of control expression profiles (blue) higher than threshold of correlation in cosine curve fitting (Fig. 4*A*). The second measure, specificity, is defined as the percent of random expression profiles (red) lower than threshold of correlation (Fig. 4*A*). There is a tradeoff between sensitivity and specificity, because changing the threshold of correlation influences these measures in opposite directions, and thus performance of molecular-timetable method depends on segregation between control and random expression profiles (Fig. 4*C*). The third measure, estimation error, is defined as the difference between estimated BT and true BT (blue, Fig. 4*B*). These results show that the molecular-timetable method with 168 time-indicating genes in mouse liver is a remarkably specific, sensitive, and accurate method.

In this study, we devised a method for BT detection utilizing single-time-point data for multiple molecules. To elucidate the importance of the use of many time-indicating genes, we assessed the specificity, sensitivity, and estimation error of the molecular-timetable method with different numbers of time-indicating genes in the presence of measurement noise, which is inevitable in clinical situations (see *Materials and Methods*). In the presence of 100% measurement noise, the application of the molecular-timetable method with 100 (or 150) time-indicating genes had the capability of detecting the BT with >99.99% (99.99%) sensitivity and 99.99% (99.99%) specificity and with low estimation errors 0.30 ± 0.24 h for mean ± SD (0.25 ± 0.21 h for mean ± SD, Fig. 5). Even in the presence of 200% measurement noise, use of the molecular-timetable methods with 100 (or 150) time-indicating genes still had the capability of detecting BT with 98.58% (99.72%) sensitivity and 98.58% (99.72%) specificity and with low estimation errors 0.62 ± 0.48 h for mean ± SD (0.50 ± 0.39 h for mean ± SD, Fig. 5). On the other hand, the molecular-timetable method with three (or five) time-indicating genes failed to specifically, sensitively, or accurately detect BT or rhythm disorders, suggesting the importance of the use of many time-indicating genes (Fig. 5). The use of >100 time-indicating

genes can confer with high sensitivity, specificity, and accuracy, to the molecular timetable, even in the presence of high measurement noise. Interestingly, we noted that the molecular-timetable method, with as few as 60 time-indicating genes, if selected appropriately can sensitively and accurately detect BT and sensitively diagnose rhythm disorders in real expression profiles (Fig. 6).

Collectively, we constructed a molecular timetable and devised a specific, sensitive and accurate method for detection of BT and rhythm disorders from a single-time-point expression profile using this molecular timetable. The power of the molecular-timetable method can be demonstrated by the quantitative and accurate detection of BT from individual expression profiles and the accurate detection of circadian rhythm disorders in *Clock*/*Clock* homozygous mutant mice. These results demonstrate the feasibility of BT detection based on single-time-point sampling, suggest the capacity for the expression-based diagnosis of circadian rhythm disorders, and may lead to the development of chronotherapy and personalized medicine.

We thank Akira Matsumoto, Teiichi Tanimura, Miho Kawamura, Hisanori Wakamatsu, Yasufumi Shigeyoshi, Mamoru Nagano, Ken-ichi Nakahama, and Akihito Adachi for fly and mouse GENECHIP analysis. We also thank Tomoko Kojima, Toshie Katakura, and Hiromi Urata for technical assistance and John B. Hogenesch, Shigehiro Ohdo, Hideyuki Terazono, and Hideo Iwasaki for stimulating discussions. This work was performed as a part of a research and development project of the Industrial Science and Technology Program supported by the New Energy and Industrial Technology Development Organization (NEDO).

1. Reppert, S. M. & Weaver, D. R. (2002) *Nature* **418,** 935–941.
2. Akhtar, R. A., Reddy, A. B., Maywood, E. S., Clayton, J. D., King, V. M., Smith, A. G., Gant, T. W., Hastings, M. H. & Kyriacou, C. P. (2002) *Curr. Biol.* **12,** 540–550.
3. Panda, S., Antoch, M. P., Miller, B. H., Su, A. I., Schook, A. B., Straume, M., Schultz, P. G., Kay, S. A., Takahashi, J. S. & Hogenesch, J. B. (2002) *Cell* **109,** 307–320.
4. Storch, K. F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F. C., Wong, W. H. & Weitz, C. J. (2002) *Nature* **417,** 78–83.
5. Ueda, H. R., Chen, W., Adachi, A., Wakamatsu, H., Hayashi, S., Takasugi, T., Nagano, M., Nakahama, K., Suzuki, Y., Sugano, S., *et al*. (2002) *Nature* **418,** 534–539.
6. Halberg, F. (1969) *Annu. Rev. Physiol.* **31,** 675–725.
7. Reinberg, A. & Halberg, F. (1971) *Annu. Rev. Pharmacol.* **11,** 455–492.
8. Reinberg, A. & Smolensky, M. (1983) *Biological Rhythms and Medicine* (Springer, New York).
9. Lemmer, B., Scheidel, B. & Behne, S. (1991) *Ann. N. Y. Acad. Sci.* **618,** 166–181.
10. Labrecque, G. & Belanger, P. M. (1991) *Pharmacol. Ther.* **52,** 95–107.
11. Levi, F., Zidani, R. & Misset, J. L. (1997) *Lancet* **350,** 681–686.
12. Ohdo, S., Koyanagi, S., Suyama, H., Higuchi, S. & Aramaki, H. (2001) *Nat. Med.* **7,** 356–360.
13. Antoch, M. P., Song, E. J., Chang, A. M., Vitaterna, M. H., Zhao, Y., Wilsbacher, L. D., Sangoram, A. M., King, D. P., Pinto, L. H. & Takahashi, J. S. (1997) *Cell* **89,** 655–667.
14. King, D. P., Zhao, Y., Sangoram, A. M., Wilsbacher, L. D., Tanaka, M., Antoch, M. P., Steeves, T. D., Vitaterna, M. H., Kornhauser, J. M., Lowrey, P. L., *et al*. (1997) *Cell* **89,** 641–653.
15. Yamazaki, S., Numano, R., Abe, M., Hida, A., Takahashi, R., Ueda, M., Block, G. D., Sakaki, Y., Menaker, M. & Tei, H. (2000) *Science* **288,** 682–685.
16. Yamazaki, S., Straume, M., Tei, H., Sakaki, Y., Menaker, M. & Block, G. D. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10801–10806.
17. Nonaka, H., Emoto, N., Ikeda, K., Fukuya, H., Rohman, M. S., Raharjo, S. B., Yagita, K., Okamura, H. & Yokoyama, M. (2001) *Circulation* **104,** 1746–1748.
18. Bjarnason, G. A., Jordan, R. C., Wood, P. A., Li, Q., Lincoln, D. W., Sothern, R. B., Hrushesky, W. J. & Ben-David, Y. (2001) *Am. J. Pathol.* **158,** 1793–1801.
19. Tomoda, A., Joudoi, T., Kawatani, J., Ohmura, T., Hamada, A., Tonooka, S. & Miike, T. (2003) *Chronobiol. Int.* **20,** 893–900.
20. Dunlap, J. C. (1999) *Cell* **96,** 271–290.
21. Harmer, S. L., Hogenesch, J. B., Straume, M., Chang, H. S., Han, B., Zhu, T., Wang, X., Kreps, J. A. & Kay, S. A. (2000) *Science* **290,** 2110–2113.
22. Claridge-Chang, A., Wijnen, H., Naef, F., Boothroyd, C., Rajewsky, N. & Young, M. W. (2001) *Neuron* **32,** 657–671.
23. McDonald, M. J. & Rosbash, M. (2001) *Cell* **107,** 567–578.
24. Ueda, H. R., Matsumoto, A., Kawamura, M., Iino, M., Tanimura, T. & Hashimoto, S. (2002) *J. Biol. Chem.* **277,** 14048–14052.
25. Lin, Y., Han, M., Shimada, B., Wang, L., Gibler, T. M., Amarakone, A., Awad, T. A., Stormo, G. D., Van Gelder, R. N. & Taghert, P. H. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 9562–9567.
26. Ceriani, M. F., Hogenesch, J. B., Yanovsky, M., Panda, S., Straume, M. & Kay, S. A. (2002) *J. Neurosci.* **22,** 9305–9319.
27. Duffield, G. E., Best, J. D., Meurers, B. H., Bittner, A., Loros, J. J. & Dunlap, J. C. (2002) *Curr. Biol.* **12,** 551–557.
28. Allada, R., White, N. E., So, W. V., Hall, J. C. & Rosbash, M. (1998) *Cell* **93,** 791–804.