ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING

An Unbiased Sensitivity Analysis Reveals Important Parameters Controlling Periodicity of Circadian Clock

Zhi Xie,[1,2] Don Kulasiri,[1] Sandhya Samarasinghe,[1] Jiang Qian,[2]

[1]Centre for Advanced Computational Solutions (C-fACS), Lincoln University, Canterbury, New Zealand

[2]Wilmer Institute, Johns Hopkins School of Medicine, 600 North Wolfe Street, 844 Maumenee Building, Baltimore, Maryland 21287; telephone: +1-410-502-2955; fax: 1-410-502-5382; e-mail: zhi.xie@jhmi.edu

Received 8 May 2009; revision received 27 July 2009; accepted 8 September 2009

Published online 23 September 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22540

ABSTRACT: To assess the importance of model parameters in kinetic models, sensitivity analysis is generally employed to provide key measures. However, it is quite often that no information is available for a significant number of parameters in biochemical models. Therefore, the results of sensitivity analysis that heavily rely on the accuracy of parameters are largely ambiguous. In this study, we propose a computational approach to determine the relative importance of parameters controlling the performance of the circadian clock in *Drosophila*. While previous attempts to sensitivity analysis largely depend on the knowledge of model parameters which are generally unknown, our study depicts a consistent picture of sensitivity assessment for a large number of parameters, even when the values of these parameters are not available in vivo. The resulting parametric sensitivity analysis suggests that PER/TIM negative loop is critical to maintain the stable periodicity of the circadian clock, which is consistent to the previously experimental and computational findings. Furthermore, our analysis generates a rich hypothesis of important parameters in the circadian clock that can be further tested experimentally. This approach can also be extended to assess the sensitivity of parameters in any biochemical system where a large number of parameters have unknown values.

Biotechnol. Bioeng. 2010;105: 250–259.

KEYWORDS: sensitivity analysis; mathematical model; genetic regulatory networks; parameter estimation; circadian clock; systems biology

Introduction

Living organisms are exposed to multitude of environmental influences and many of them follow a daily periodic change. Consequently, many physiological processes in living organisms follow a daily periodicity, known as circadian rhythms, controlled by the mechanisms broadly encapsulated within the terms "circadian clocks" (Nitabach, 2005; van Gelder et al., 2003). In recent decades, many components and molecular mechanisms comprising circadian clocks have been discovered by studying the model organisms such as unicellular eukaryotes, fungi, plants, invertebrates, and mammals (Young and Kay, 2001). Along with the discovery of the components related to the circadian clock in *Drosophila*, a number of mathematical models have been proposed to describe interactions and controlling mechanisms among the molecular components (Goldbeter, 2002).

To assess relative importance of biochemical parameters with respect to the output of a system, two types of parametric sensitivity analysis are generally employed (Rand, 2008; Varma et al., 1999): local sensitivity analysis (LSA), and a more recent approach, global sensitivity analysis (GSA). In LSA, we allow one parameter to be varied one at a time and compute derivative vector ($S_i = \partial Y/\partial P_i$, where Y is the output of interest and P_i is the ith parameter) to obtain the sets of values for the parameters which would indicate the sensitive regions of each parameter. In contrast to LSA, GSA explores effects of simultaneous parameter variations of arbitrary magnitudes for a system (Rand, 2008). Since rate constants of diverse biochemical processes in vivo are more likely to vary simultaneously under varying environments, GSA is believed to be more appropriate for sensitivity analysis of biochemical systems (Rand, 2008; Stelling et al., 2004).

Because both LSA and GSA rely on the knowledge of biochemical parameters, a robust and reliable analysis of a biochemical system requires reliable quantitative information on these parameters. However, the lack of in vivo or in vitro measurements of the kinetic parameters makes parameter estimation one of major obstacles in studying

Correspondence to: Z. Xie

Additional Supporting Information may be found in the online version of this article.

biological networks and therefore weakens effectiveness of sensitive analyses (Mendes and Kell, 1998). In most of previous attempts to mathematically describe the circadian clocks, parameters were manually tuned to fit experimental data quantitatively aided by biological knowledge of system behaviors and phase plane analyses (Kulasiri and Xie, 2008; Kulasiri et al., 2008; Leloup and Goldbeter, 2003, 2000; Smolen et al., 2004; Tyson et al., 1999; Ueda et al., 2002; Vilar et al., 2002; Xie and Kulasiri, 2007). However, manual tuning is not only laborious but also gives no guarantee to "optimal" parameter values (Mendes and Kell, 1998), which makes any subsequent sensitivity analysis largely ambiguous. To surmount the limitation of this traditional approach, several global parameter estimation (GPE) approaches have been proposed to study biochemical systems in recent years. A recent comparative study of several GPE approaches suggested that although GPE approaches also cannot guarantee global optimality with certainty, their robustness makes them the best available candidate for estimating the parameters in a biochemical system (Moles et al., 2003).

In this research, we assess the relative importance of the parameters controlling the periodicities of the circadian clock naturally inherent in *Drosophila*. To reveal the important parameters in the system without a prior knowledge of parameter values, we propose a hybrid approach, named GPEGSA, which can be regarded as an "unbiased" approach that is independent of any particular choice of parameters. Our results show a number of parameters which are important to control periodicity of the circadian clock and suggested PER/TIM negative feedback loop is critical for maintaining stable circadian rhythms in *Drosophila*.

Table I. ODEs for the model.

$$\frac{dx_1}{dt} = a_2(1-x_1)x_{18} - a_8x_1 \quad (1)$$

$$\frac{dx_2}{dt} = a_3(1-x_2)x_{18} - a_9x_2 \quad (2)$$

$$\frac{dx_3}{dt} = a_1(1-x_3)x_{18} - a_7x_3 \quad (3)$$

$$\frac{dx_4}{dt} = a_4(1-x_4)x_{18} - a_{10}x_4 \quad (4)$$

$$\frac{dx_5}{dt} = a_5(1-x_6-x_5)x_{16} - a_{11}x_5 \quad (5)$$

$$\frac{dx_6}{dt} = a_6(1-x_6-x_5)x_{15} - a_{12}x_6 \quad (6)$$

$$\frac{dx_7}{dt} = a_{48}(a_{23}x_5 + a_{24}x_6 + a_{25}(1-x_6-x_5)) - a_{32}x_7 \quad (7)$$

$$\frac{dx_8}{dt} = a_{50}(a_{20}(1-(1-x_1)^{a45}) + a_{26}(1-x_1)^{a45}) - a_{34}x_8 \quad (8)$$

$$\frac{dx_9}{dt} = a_{51}(a_{21}(1-(1-x_2)^{a45}) + a_{26}(1-x_2)^{a45}) - a_{35}x_9 \quad (9)$$

$$\frac{dx_{10}}{dt} = a_{52}(a_{22}(1-(1-x_4)^{a46}) + a_{26}(1-x_4)^{a46}) - a_{36}x_{10} \quad (10)$$

$$\frac{dx_{11}}{dt} = a_{49}(a_{19}(1-(1-x_5)^{a47}) + a_{26}(1-x_5)^{a47}) - a_{33}x_{11} \quad (11)$$

$$\frac{dx_{12}}{dt} = a_{29}x_8 - a_{14}x_{12}x_{13} + a_{17}x_{17} - a_{39}x_{12} \quad (12)$$

$$\frac{dx_{13}}{dt} = a_{30}x_9 - a_{14}x_{12}x_{13} + a_{17}x_{17} - a_{40}x_{13} \quad (13)$$

$$\frac{dx_{14}}{dt} = a_{27}x_7 - a_{13}x_{14}x_{20} + a_{16}x_{18} - a_{37}x_{14} \quad (14)$$

$$\frac{dx_{15}}{dt} = a_{31}x_{10} - a_{41}x_{15} \quad (15)$$

$$\frac{dx_{16}}{dt} = a_{28}x_{11} - a_{38}x_{16} \quad (16)$$

$$\frac{dx_{17}}{dt} = a_{14}x_{12}x_{13} - a_{17}x_{17} - a_{15}x_{17}x_{18} + a_{18}x_{19} - a_{42}x_{17} \quad (17)$$

$$\frac{dx_{18}}{dt} = a_{13}x_{14}x_{20} - a_{16}x_{18} - a_{15}x_{17}x_{18} + a_{18}x_{19} - a_{43}x_{18} \quad (18)$$

$$\frac{dx_{19}}{dt} = a_{15}x_{17}x_{18} - a_{18}x_{19} - a_{44}x_{19} \quad (19)$$

Materials and Methods

Construction of Circadian Clock Model

The structure of the genetic regulatory network of the circadian clock considered in the present study has been reported previously (Xie and Kulasiri, 2007). The ordinary differential equations describing the model are listed in Table I. The rate constant parameters and components of the model are given in Figure 1 and Table II. The model is briefly described in the Supplementary Material.

GPEGSA Approach

To analyze the parametric sensitivity of the systems without knowing the parameter values, the proposed hybrid approach (GPEGSA) is outlined in Table III. This approach is a sequential combination of GPE, GSA and relevant statistical analyses. First, GPE was used to search possible parameter solutions of the circadian clock model in the parameter space confined by biological and biochemical knowledge. Secondly, GSA was used to analyze the sensitivity of parameters for all the parameter sets resulted from GPE. Finally, statistical analyses were performed for the results from GSA to find the most important parameters in the circadian clock.

Global Parameter Estimation

Parameter Constraints

We constrained some of the parameters to the appropriate values and the ranges based on biological knowledge to-date and insights to simplify the estimation. First, we assumed the following to reduce the number of parameters to be estimated from 44 to 38: (1) The binding and unbinding rates of CLK/CYC dimer to *per* gene were the same as those of *tim* gene. The reasoning behind this assumption is that the genome analysis revealed that the upstream of both *per* and *tim* transcription initiation sites contains a consensus E-box DNA motif (CACGTG), which is a known target for CLK and CYC, a family of basic helix-loop-helix domain (bHLH) transcription factors (TFs) (Hao et al., 1997). As the peaks and troughs of *per* and *tim* mRNA concentrations are similar according to the experiments, we hypothesized that the TF dimers, CLK/CYC, had similar kinetic affinities on the *per* and *tim* promoters producing similar binding and

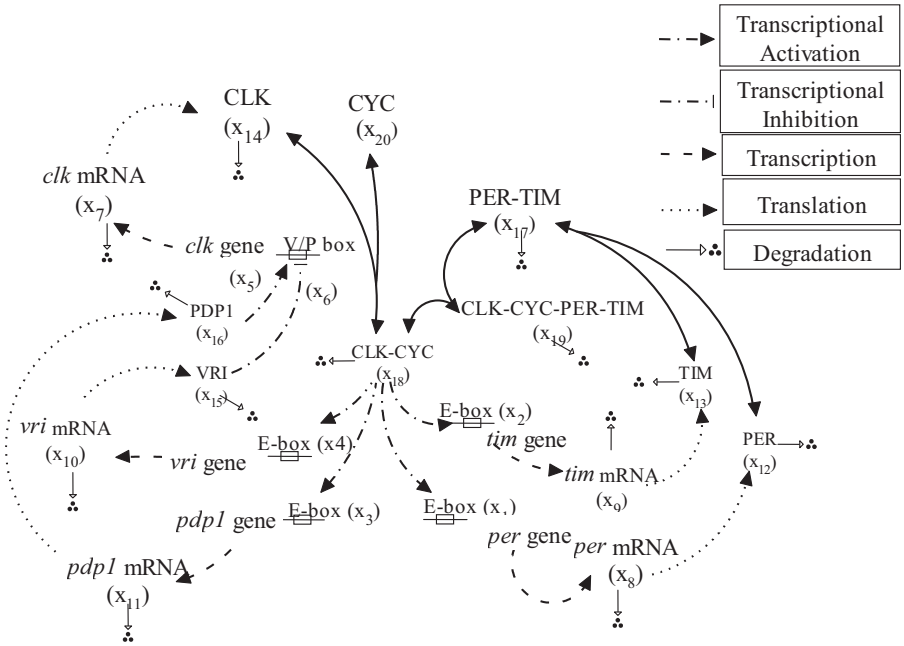

Figure 1. Schematic diagram of the structure of the model. The molecular components in the system and their corresponding mathematical symbols used for the equations in Table I are denoted.

unbinding rates to the respective promoters, thus making transcription rates similar as well. (2) A number of experiments showed that the concentrations of PER and TIM had highly similar profiles and these results suggested that the oscillating PER and TIM levels resulted from the oscillating levels of *per* and *tim* mRNAs (Sehgal et al., 1995). We therefore assumed that the translation rates of *per* and *tim* mRNAs were the same. (3) The kinetics of degradation of PER and TIM were assumed to be the same since experimental evidence showed that the degradation of both PER and TIM are all regulated by the product of the *slimb* gene in the condition of constant darkness (Grima et al., 2002). In addition, several other kinases are involved in the regulatory processes of PER and/or TIM such as DBT, GSK-3, or CKII.

Secondly, the initial ranges of parameters are sought to confine within physiologically meaningful ranges. In a previous study, a similar random sampling search algorithm was used where all the parameters were arbitrarily set to be bounded between 0 and 10, with units in nM and hours (Locke et al., 2005). In this study, instead of giving a constant initial range for all the parameters, we searched more meaningful parameter ranges in the literature based on functional similarities of genes, mRNAs and proteins. Any values of parameters found in literature are called reference values. There are some previous studies suggesting that transcription rates, degradation rates of mRNA and protein in gene expression vary several orders in magnitude (Hargrove et al., 1991). Therefore, in the absence of any other specific information, we set the lower and higher boundaries of our target parameters to one-tenth of the lowest reference value and 10 times the highest reference value, respectively. The ranges of parameters used for the parameter estimation is summarized in Table IV and the reasoning behind the parameter ranges is provided in the Supplementary Material.

Random Number Generator

Even though we reduced 44 parameters to 38 parameters, testing all possible combinations of 38 unknown parameters for potential parameter values within their ranges was still computationally infeasible. Assuming that only 10 candidate values for each parameter, the total combinations would be as high as 10^{38}. To overcome this difficulty, we generated a number of parameter sets, where one parameter set consisted of 38 parameters, using random sampling from the parameter ranges in Table IV. To minimize the gaps between the preceding random numbers, one million parameter sets were produced. Random numbers were generated using the Sobol algorithm, which is known to be useful in computational problems where numbers are computed on a grid without a prior knowledge of how the grid should be (Press, 1997).

Selection of Acceptable Parameter Sets

We first defined 23–25 h as the period range of rhythms for wild-type (WT) *Drosophila* (Hardin, 2005). We also defined 19–29 h as the period range of rhythms for mutant *Drosophila* (Dunlap, 1999), although in a few cases ultra-long (up to 33 h) or ultra-short (down to 16 h) mutants have also been observed (Konopka et al., 1994; Rothenfluh et al., 2000). The goal of parameter estimation is to find out parameter sets which can produce WT circadian rhythms.

Table II. Biochemical meaning of parameters.

Parameters	Biochemical meaning
a_1	Binding rate of CLK/CYC to an E-box in *pdp1* promoter
a_2	Binding rate of CLK/CYC to an E-box in *per* promoter
a_3	Binding rate of CLK/CYC to an E-box in *tim* promoter
a_4	Binding rate of CLK/CYC to an E-box in *vri* promoter
a_5	Binding rate of PDP1 to the V/P box in *clk* promoter
a_6	Binding rate of VRI to the V/P box in *clk* promoter
a_7	Unbinding rate of CLK/CYC to an E-box in *pdp1* promoter
a_8	Unbinding rate of CLK/CYC to an E-box in *per* promoter
a_9	Unbinding rate of CLK/CYC to an E-box in *tim* promoter
a_{10}	Unbinding rate of CLK/CYC to an E-box in *vri* promoter
a_{11}	Unbinding rate of PDP1 to the V/P box in *clk* promoter
a_{12}	Unbinding rate of VRI to the V/P box in *clk* promoter
a_{13}	Association rate of CLK/CYC dimer
a_{14}	Association rate of PER/TIM dimer
a_{15}	Association rate of CLK/CYC/PER/TIM complex
a_{16}	Dissociation rate of CLK/CYC dimer
a_{17}	Dissociation rate of PER/TIM dimer
a_{18}	Dissociation rate of CLK/CYC/PER/TIM complex
a_{19}	Transcription rate of activated *pdp1* gene
a_{20}	Transcription rate of activated *per* gene
a_{21}	Transcription rate of activated *tim* gene
a_{22}	Transcription rate of activated *vri* gene
a_{23}	Transcription rate of activated *clk* gene
a_{24}	Transcription rate of repressed *clk* gene
a_{25}	Transcription rate of *clk* gene (binding neither PDP1 nor VRI)
a_{26}	Transcription rate of deactivated *per*, *tim*, *vri* or *pdp1* gene
a_{27}	Translation rate of *clk* mRNA
a_{28}	Translation rate of *pdp1* mRNA
a_{29}	Translation rate of *per* mRNA
a_{30}	Translation rate of *tim* mRNA
a_{31}	Translation rate of *vri* mRNA
a_{32}	Degradation rate of *clk* mRNA
a_{33}	Degradation rate of *pdp1* mRNA
a_{34}	Degradation rate of *per* mRNA
a_{35}	Degradation rate of *tim* mRNA
a_{36}	Degradation rate of *vri* mRNA
a_{37}	Degradation rate of CLK protein
a_{38}	Degradation rate of PDP1 protein
a_{39}	Degradation rate of PER protein
a_{40}	Degradation rate of TIM protein
a_{41}	Degradation rate of VRI protein
a_{42}	Degradation rate of PER/TIM dimer
a_{43}	Degradation rate of CLK/CYC dimer
a_{44}	Degradation rate of CLK/CYC/PER/TIM complex
a_{45^*}	Number of E-boxes in *per* or *tim* promoter
a_{46^*}	Number of E-boxes in *vri* promoter
a_{47^*}	Number of E-boxes in *pdp1* promoter
a_{48^*}	Concentration of *clk* promoter
a_{49^*}	Concentration of *pdp1* promoter
a_{50^*}	Concentration of *per* promoter
a_{51^*}	Concentration of *tim* promoter
a_{52^*}	Concentration of *vri* promoter

Constants are marked with "*." The values of the constants: a_{45}: 5; a_{46}: 4; a_{47}: 5; a_{48}–a_{53}: 0.003185 nM.

To obtain more possible acceptable parameter sets, we set the initial parameter estimation criterion to mutant rhythm instead. For those parameter sets producing mutant rhythms but out of the range of WT, we could rescale the oscillations to WT period by multiplying each differential equation with the same appropriate scaling factor. The period was measured after 500-h simulation time to eliminate potential transition states. Amplitudes were not taken into account for selection criteria because the concentrations of the circadian clock related mRNAs and proteins are not known and only relative concentration abundance was measured in vivo.

Table III. Procedure for the GPEGSA approach.

(1) Global parameter estimation
  (1a) Set parameter constraints to reduce parameter space if possible
  (1b) Define parameter ranges according to biological knowledge
  (1c) Generate a large number of parameter sets within the ranges using Sobol algorithm
  (1d) Run the model using the parameter sets, if the parameter set produces a desired output, save the parameter set for the Step 2
(2) Global sensitivity analysis
  (2a) For a parameter set from Step 1d, select the parameter to be examined and define the parameter variation ranges
  (2b) Generate a number of parameter sets within the range using Latin Hypercube Sampling
  (2c) Fit each parameter set into the model, test if the output produced is "behavior" given criteria
  (2d) Compare distribution of "behavior" and "non-behavior" parameter sets; evaluate sensitivity of parameters
  (2e) Go to Step 2a and repeat the Step 2 for all the parameter sets in Step 1d
(3) Statistical analysis of sensitivities for all the parameter sets

Global Sensitivity Analysis

We used a well-established approach, regionalized sensitivity analysis (RSA), for GSA. In RSA all the parameters vary simultaneously within defined ranges. Then the parameter sets were split into two groups according to the selected objective function. The statistical difference of the two groups was calculated to indicate the sensitivity of the model performance to the parameter variation. We refer to the original paper (Hornberger and Pear, 1981) for the technical details of RSA. A summary of RSA procedure employed in this study is given below.

Table IV. Initial parameter ranges used for parameter estimation.

Parameters	Running range	Note
a_1–a_6	0.0072–7.2	Protein–DNA binding
a_7–a_{12}	0.144–72	Protein–DNA unbinding
a_{13}–a_{18}	3.6–108	Protein–protein binding and unbinding
a_{19}–a_{23}	0.2–200	Transcription (activated)
a_{24}, a_{26}	0.0001–0.01	Transcription (repressed and deactivated)
a_{25}	0.0001–1	Transcription (*clk*)
a_{27}–a_{31}	0.2–200	Translation
a_{32}–a_{36}	0.007–0.3	Degradation of mRNA
a_{37}–a_{41}	0.011–2.3	Degradation of protein
a_{42}–a_{43}	0.011–2.3	Degradation of dimmer
a_{44}	0.055–11.5	Degradation of tetramer

Unit: nM and hour.

(1) We selected the parameters to be tested for sensitivity analysis and defined the ranges of the selected parameters to be varied. Here we chose all the parameters related to the rate constants of biochemical processes (Table II). The parameters a_{45} (the number of E-boxes in *per* or *tim* promoter), a_{46} (E-boxes in *vri* promoter), and a_{47} (E-boxes in *pdp1* promoter) are related to the structure of the model and were not considered to be subjected to variation. The parameters were examined within a range of ±30% of their reference values. The variation of 30% was taken since the later results indicated that 30% variations gave us a reasonable number of parameter sets which can produce sustained oscillations.
(2) We generated vectors of parameters from random distributions within defined ranges. For each acceptable parameter set, we generated 5,000 sets of parameters. We used Latin Hypercube Sampling (LHS) to generate sample random parameter vectors since LHS guarantees that individual parameter ranges are evenly covered (McKay et al., 1979).
(3) We ran the model using 5,000 sets of parameters from LHS and determined whether a parameter vector was "behavior" by examining whether the set produced sustained oscillations for all the components in the system with period of WT rhythm ±5 h. The circadian rhythms beyond this range were defined as "non-behavior."
(4) For each parameter, we compared the cumulative distributions of the parameter values associated with the "behavior" and "non-behavior." If the two distributions were not statistically different, the parameter was classified as insensitive; otherwise, the parameter was classified as sensitive. Kolmogorov–Smirnov (KS) test was used to evaluate the statistical difference between the two distributions and P value less than 0.05 was defined as "significantly different" (Stephens, 1970).

Local Sensitivity Analysis

We also performed LSA in this study for comparison with GSA. Similar to GSA, only period sensitivity was considered for LSA, which was defined by $S(\tau; p_j) = \mathrm{d}\tau(t, p_j)/\partial p_j$, where $\tau(p)$ defines the period of the system for a given parameter p. Furthermore, normalized period sensitivity (Sn) is defined by

$$Sn(\tau; p_j) = \frac{p_j}{\tau(p_j)} S(\tau; p_j) \tag{20}$$

Both period sensitivity and normalized period sensitivity have been used in previous research for analyzing period sensitivity of oscillations in chemical and biochemical systems (Stelling et al., 2004; Varma et al., 1999), and we used normalized period sensitivity in this study.

Computational Implementation

All the simulations were performed with Matlab. ODEs were solved using "ode23s" solver. Sobol algorithm was implemented using the program written by John Burkardt (http://people.scs.fsu.edu/~burkardt/m_src/sobol/sobol.html). LHS random sampling was implemented using "LHS" function in the statistics toolbox in Matlab. We used "kstest2" function in the statistics toolbox in Matlab to implement KS test.

Results

Determination and Characterization of the Parameter Sets

To search the possible parameter sets for the model, we first generated one million parameter sets using quasi-random algorithm, each set represented a vector of random parameter values. By setting the initial acceptable periods of 19–29 h, we were able to obtain 47 parameter sets, where 24 sets produced damped oscillations and 23 sets produced sustained oscillations. Among the 23 sets, 5 are with WT rhythms (23–25 h period) and the periods of the rest 18 sets could be rescaled to WT. Finally, these 23 parameter sets were used for further analysis (values listed in Supplementary Table I). Principal component analysis showed the distribution of all the parameter sets in the first three principle axes (Fig. 2), where these three principal axes define 85.55% of the entire parameter space. As shown in Figure 2, 23 parameter sets do not cluster in particular positions in the space defined by the three principal axes, instead they widely disperse throughout the parameter space, illustrating that a large parameter space confined by the parameter constrains has been searched.

Sensitive Parameters Revealed by GSA

Having known the possible parameter sets for the model, we performed GSA for all the 23 sets to determine the effects of changes in model parameters on the period performance of the circadian system. The resulting KS tests comprised of a 23×44 matrix in which P values from 44 parameters in each parameter set for all the sets were recorded (Supplementary Table II). To illustrate how the P value reflects the sensitivity of parameters, we arbitrarily chose the parameter set 4 and plotted the cumulative distributions for its most sensitive parameter, a_{36} (degradation rate of *vri* mRNA and $P = 1.34\mathrm{E}-7$), and its most insensitive parameter, a_{25} (basal transcription rate of *clk* gene and $P = 0.96$) (Fig. 3). The cumulative distribution of a_{36} showed a distinct difference for the acceptable and

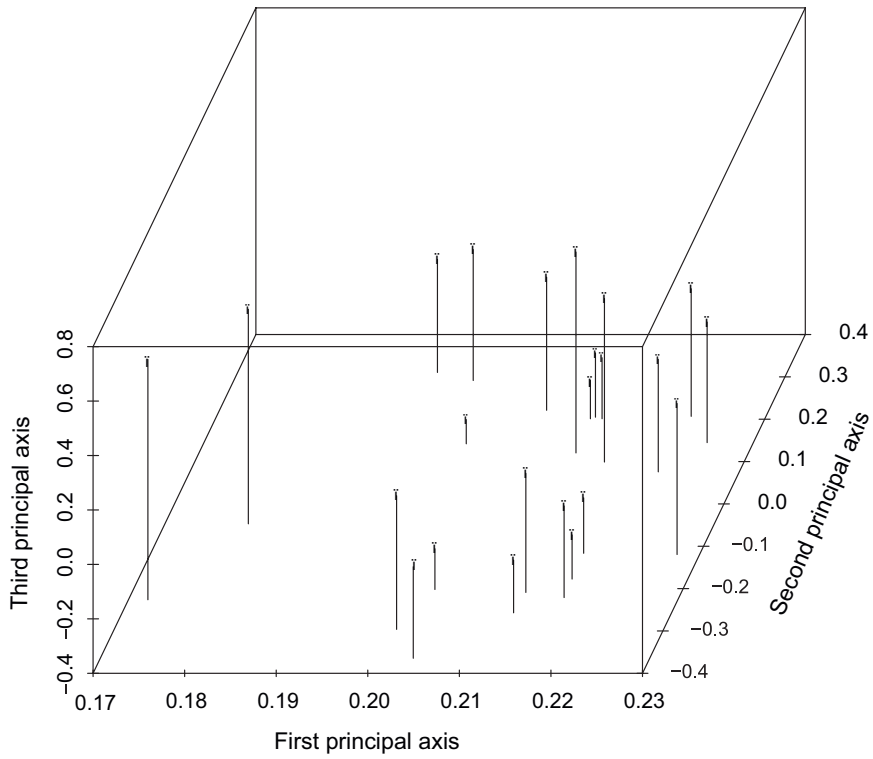

Figure 2. Principal component analysis for 23 parameter sets in the parameter space. The first three principal axes are displayed.

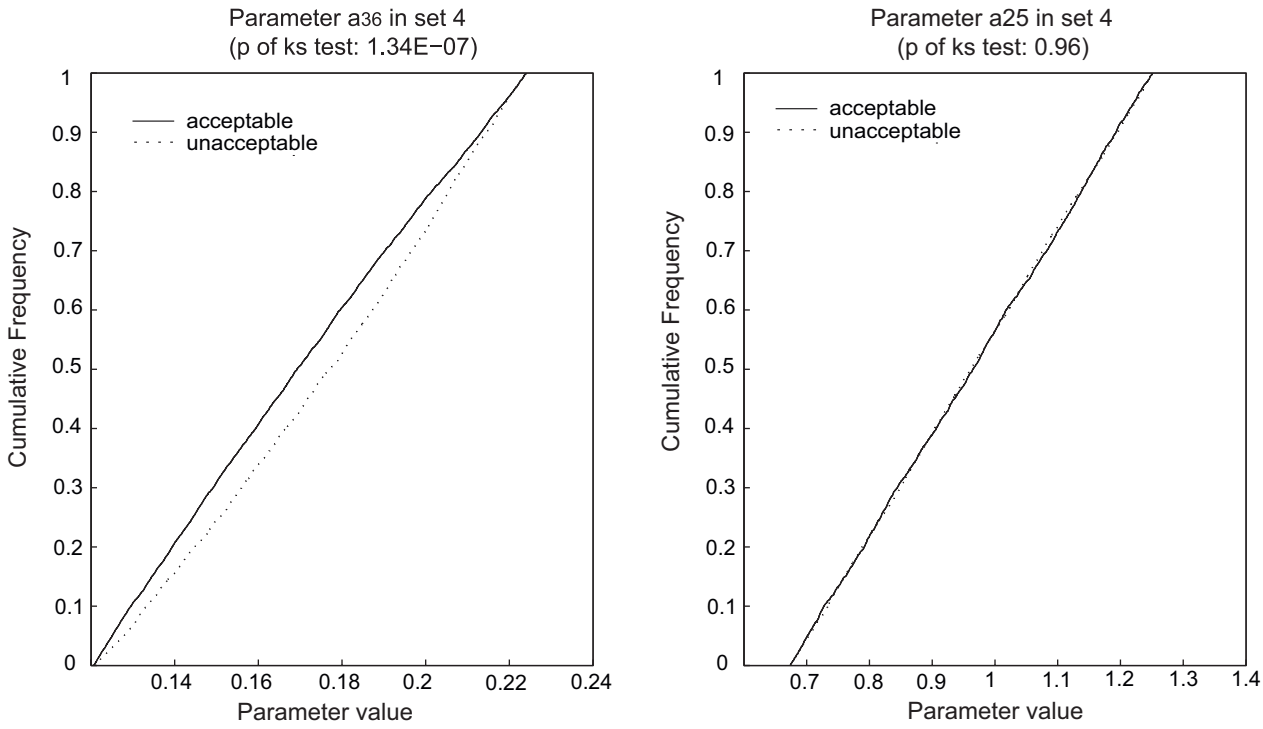

Figure 3. Cumulative frequency distribution of two representative parameters. The most sensitive parameter (a_{36}) and the most insensitive parameter (a_{25}) in the parameter set 4 were plotted. *P* values from the KS test are denoted in the figure title.

unacceptable outputs, indicating that the variations of this parameter resulted distinguishing behaviors of the system, whereas a_{25} had nearly similar cumulative distribution curves for the acceptable and unacceptable outputs, indicating that changes in a_{25} had little influence to the system output.

The heatmap image in Figure 4 shows that in each parameter set, the system was very sensitive to the variations of some parameters whereas some parameters had little or no influence to the system's output. We further defined parameters as "sensitive" if its P values of the KS tests being lower than 0.05, or "insensitive" otherwise. As shown in the bar plot on the upper axis of the image, each parameter set contains different number of sensitive parameters, ranging from 12 to 27. However, if we examine the sensitivity of each parameter in different parameter sets, we can see that some parameters show a surprisingly consistent picture as shown in the bar plot on the right axis of the image. In particularly, the parameter a_{36} (degradation rate of *vri* mRNA) were classified as "sensitive" 21 times out of 23 tests. Notably, a_{24}, a_{25} (repressed and basal transcription rate of *clk* gene) and a_{26} (transcription

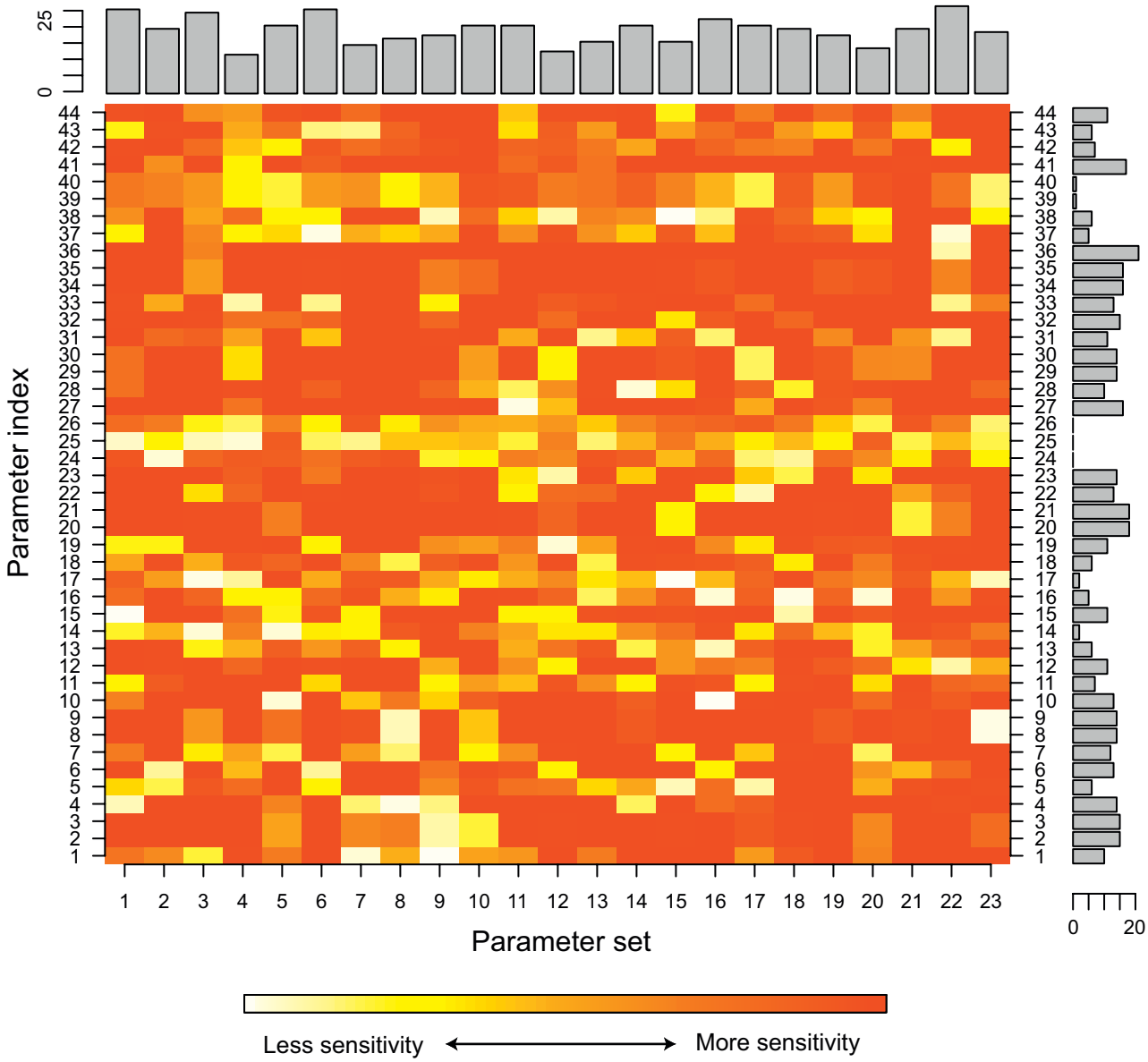

Figure 4. Global parameter sensitivity for the circadian clock system. The heatmap image contains KS test values for all the 44 parameters in all the 23 parameter sets. The parameter is classified as "sensitive" if its P value from KS test is less than 0.05. The bar-plot on the right y-axis shows the counts of each parameter classified as "sensitive" over 23 parameter sets. The bar-plot on the upper x-axis shows the number of sensitive parameters in each parameter set. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

rate of deactivated *per*, *tim*, *vri*, or *pdp1* gene) were consistently classified as "insensitive" although their parameter values considerably changed within their defined ranges.

Functional Analysis of Sensitive Parameters

We next evaluated the biological context of the identified sensitive parameters. A parameter is defined as "sensitive" if its median P value of 23 tests is significant (<0.05). Seventeen identified sensitive parameters and their corresponding biochemical meanings are listed in Table V. Surprisingly, all the sensitive parameters are associated with *vri*, *clk*, *per*, and *tim* molecules, including their genes, mRNAs, and proteins whereas none of the parameters regulating *pdp1* molecules are classified as "sensitive." Given the fact that the current circadian clock contains *per*/*tim* negative feedback loops, *vri* negative feedback loop, and *pdp1* positive feedback, our sensitivity analysis suggested that the negative feedback loops play central roles in controlling the period of the circadian rhythms. It is also notable that all the controlling parameters in the protein DNA interactions in the *per*/*tim* feedback loop were all identified as sensitive. We also classified the parameter sensitivity according to different biochemical functions in Table V, including protein–DNA interactions, transcription, translation, degradation of mRNAs, degradation of proteins and protein–protein interactions. The results show that periodicity was consistently more sensitive to mRNAs degradations than protein degradations where the degradation rates of four mRNAs out of five were identified as "sensitive" whereas the degradation rate of only one protein (VRI) appeared in the sensitive list. Furthermore, the circadian system was sensitive to the perturbations of the degradation rates of both *vri* mRNA and protein, and the stability of *vri* mRNA seemed to be important than that of VRI protein.

Comparison of LSA and GSA

Because the sensitivity analyses for circadian clock systems in the previous investigation were carried out by LSA (Leloup and Goldbeter, 2000; Smolen et al., 2001, 2004), we also examined whether LSA gave similar indication as GSA did (Supplementary Table III). The ranks of parameter sensitivity given by LSA against those of GSA were shown in Figure 5, where the ranks were calculated according to median sensitivity values over the results from 23 parameter sets. If the ranks of a parameter given by two investigated approaches are similar, they should be aligned along the 45° (diagonal) line. Figure 5 illustrates two trends of the results from both approaches. On the one hand, some parametric sensitivity given by the two approaches is very similar; in particular, both approaches gave the same result for parameters a_7, a_{10}, and a_{42} (see Table II for their biochemical meaning). On the other hand, the resulting rankings of some parameters are significantly different. For example, parameter a_{36} was ranked to be the most sensitive by GSA whereas it was ranked 16th by LSA.

Discussion

The dynamic behavior of the circadian clock is affected by topology of the molecular components and their

Table V. Summary table of the sensitive parameters.

Parameters	Median P of KS test	Parameter description	Molecular component	Biochemical process
a_{36}	9.77E − 05	Degradation rate of *vri* mRNA	*vri*	dmRNA
a_{41}	2.14E − 04	Degradation rate of VRI protein	*vri*	dprotein
a_{23}	4.44E − 04	Transcription rate of activated *clk* gene	*clk*	Tsc
a_{27}	2.64E − 03	Translation rate of *clk* mRNA	*clk*	Tl
a_{20}	6.81E − 03	Transcription rate of activated *per* gene	*per*	Tsc
a_{21}	6.81E − 03	Transcription rate of activated *tim* gene	*tim*	Tsc
a_{29}	7.32E − 03	Translation rate of *per* mRNA	*per*	Tl
a_{30}	7.32E − 03	Translation rate of *tim* mRNA	*tim*	Tl
a_{32}	8.04E − 03	Degradation rate of *clk* mRNA	*clk*	dmRNA
a_6	1.23E − 02	Binding rate of VRI to the V/P box in *clk* promoter	*vri*	PDI
a_{22}	1.38E − 02	Transcription rate of activated *vri* gene	*clk*	Tsc
a_{34}	1.59E − 02	Degradation rate of *per* mRNA	*per*	dmRNA
a_{35}	1.59E − 02	Degradation rate of *tim* mRNA	*tim*	dmRNA
a_2	1.60E − 02	Binding rate of CLK/CYC to an E-box in *per* promoter	*per*	PDI
a_3	1.60E − 02	Binding rate of CLK/CYC to an E-box in *tim* promoter	*tim*	PDI
a_8	1.77E − 02	Unbinding rate of CLK/CYC to an E-box in *per* promoter	*per*	PDI
a_9	1.77E − 02	Unbinding rate of CLK/CYC to an E-box in *tim* promoter	*tim*	PDI
a_4	1.92E − 02	Binding rate of CLK/CYC to an E-box in *vri* promoter	*vri*	PDI
a_{10}	3.59E − 02	Unbinding rate of CLK/CYC to an E-box in *vri* promoter	*vri*	PDI

dmRNA, degradation of mRNA; dprotein, degradation of proteins; Tsc, transcription; Tl, translation; PDI, protein–DNA interaction.
Molecular component: the upstream molecule of the parameter. Biochemical process: the functional category of the biochemical process where the parameter controls.

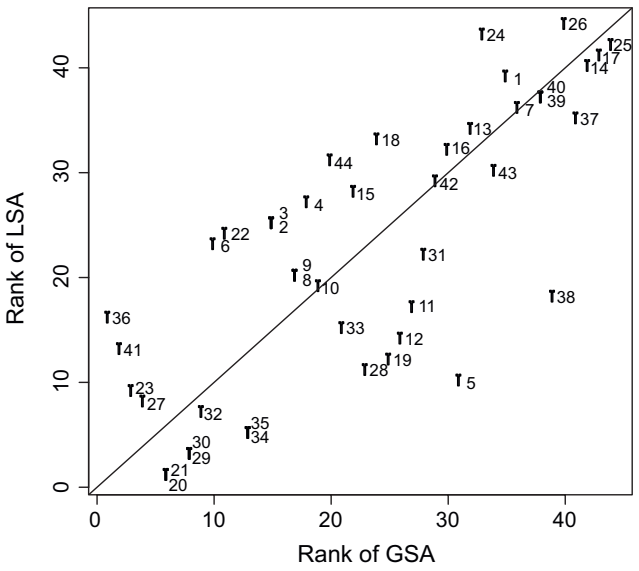

Figure 5. Ranks for LSA and GSA. The *x*- and *y*-axes denote the ranks of sensitivity calculated by GSA and LSA, respectively, ranging from 1 to 44. The lower value from GSA and the higher value from LSA indicate more sensitive parameters (Supplementary Material).

biochemical parameters. Sensitivity analysis can determine the relationships between system behavior and parameter variations. However, the knowledge about the strength of parameters in the circadian clock is generally absent. Even some measurements of biochemical parameters are available from literature, they are often estimated within an order of magnitude under different experimental conditions. In this study, we have systematically investigated the influence of biochemical parameters on the period behavior of clock system based on our proposed computational approach, without a prior knowledge of parameter values.

We have tested one million possible parameter sets from the parametric ranges derived from the biochemical and biological literature. Although intensive searching the parameter space costs extensive computational time, around 1 week in a personal computer, it ensured that we obtained to receive much more reliable parameter sets than the previous models where manual tuning were usually used. Out of one million initial parameter sets, 47 produced oscillations with periods of 19–29 h. Among them, 24 oscillations were damped and 23 oscillations were sustained which were subsequently used as inputs for GSA. Principal components analysis revealed that 23 parameter sets are widely dispersed throughout the parameter space, illustrating that the interconnectivity of parameters, instead of particular values of individual parameters, specify model outputs.

The results from the sensitivity analyses for all the 23 parameter sets identified 19 important parameters controlling the system's periodic outputs. Functional analysis of these parameters revealed that the negative feedback loops in the circadian clock system played a more critical role than positive feedback loop in terms of period. This is consistent with the finding from a previous theoretical model that a single negative feedback alone can maintain the periodicity of a clock system (Smolen et al., 2001). In particular, all the parameters in the protein DNA interactions involved in the *per*/*tim* feedback loop are identified to be important. This computationally proved that the transcriptional regulation of the *per*/*tim* feedback loop is the main force to alter the period of the system, as a similar conclusion drawn from a previous mathematical model (Smolen et al., 2004). It is also notable that the functional analysis for the sensitive parameters suggested that the periodicity of the system was more sensitive to mRNA degradation variations than that of protein. Given the fact that the mRNA degradation rates are on average lower than the protein degradation rates in the model system, one may ask whether the higher sensitivity of the system to mRNA variations is caused by the lower value of mRNA degradation rates. Comparing mRNA degradation rate to its protein degradation rate for each molecular component in the 23 parameter sets, we observed 17 cases where the degradation rates of a mRNA was higher than its protein. Among 14 of the 17 cases, the system was found to be more sensitive to mRNA variation than that of the protein. Therefore, it is likely that the structure of the model, instead of the lower mRNA degradations rates, decides that the mRNA degradation is more important than protein degradation for controlling the periodicity of this system.

Finally, the comparison study between GSA and the conventional LSA suggested that for some parameters, these two different approaches drew significantly different conclusions. As GSA renders more realistic perturbations of biological systems, GSA may give more reliable information regarding the sensitivity of these parameters.

Overall, we have demonstrated the feasibility of GPEGSA approach to measure the functional influence of biochemical parameters on the periodic outputs of a circadian clock system in this study. Hence, we are able to identify critical parameters and to better understand the mechanisms of the complex circadian clock system. The results verified the consistency of the analysis with the current knowledge of biology and provided indications for the most informative parts in the circadian clock, which can be used as initial guesses in experiment design for accurate parameter estimation. As our current knowledge of components in the circadian clock expands, the model structure used in this study is likely to be modified in the near future. Our proposed sensitivity analysis framework provides a high degree of flexibility by additional components into the model without knowing their quantitative relationship to the other components.

We are grateful to anonymous reviewers for helpful comments. Z. Xie also thanks the Agriculture and Life Science Division, Lincoln University, for providing writing scholarship.

References

Dunlap JC. 1999. Molecular bases for circadian clocks. Cell 96(2):271–290.

Goldbeter A. 2002. Computational approaches to cellular rhythms. Nature 420(6912):238–245.

Grima B, Lamouroux A, Chelot E, Papin C, Limbourg-Bouchon B, Rouyer F. 2002. The F-box protein Slimb controls the levels of clock proteins Period and Timeless. Nature 420(6912):178–182.

Hao H, Allen DL, Hardin PE. 1997. A circadian enhancer mediates PER-dependent mRNA cycling in Drosophila melanogaster. Mol Cell Biol 17(7):3687–3693.

Hardin PE. 2005. The circadian timekeeping system of Drosophila. Curr Biol 15(17):714–722.

Hargrove JL, Hulsey MG, Beale EG. 1991. The kinetics of mammalian gene expression. BioEssays 13(12):667–674.

Hornberger GM, Pear RC. 1981. An approach to the preliminary analysis of environmental system. J Environ Manag 12:7–18.

Konopka RJ, Hamblen-Coyle MJ, Jamison CF, Hall JC. 1994. An ultrashort clock mutation at the period locus of Drosophila melanogaster that reveals some new features of the fly's circadian system. J Biol Rhythms 9(3–4):189–216.

Kulasiri D, Xie Z. 2008. Modelling circadian rhythms in Drosophila and investigation of VRI and PDP1 feedback loops using a new mathematical model. Math Model Nat Phenom 3(2):1–26.

Kulasiri D, Nguyen L, Samarasinghe S, Xie Z. 2008. A review of systems biology perspective on genetic regulatory networks with examples. Curr Bioinformatics 3:197–225.

Leloup JC, Goldbeter A. 2000. Modeling the molecular regulatory mechanism of circadian rhythms in Drosophila. Bioessays 22(1):84–93.

Leloup J-C, Goldbeter A. 2003. Toward a detailed computational model for the mammalian circadian clock. Proc Natl Acad Sci USA 100(12):7051–7056.

Locke JCW, Southern MM, Kozma-Bognar L, Hibberd V, Brown PE, Turner MS, Millar AJ. 2005. Extension of a genetic network model by iterative experimentation and mathematical analysis. Mol Syst Biol 1(1):13.

McKay MD, Beckman RJ, Conover WJ. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2):239–245.

Mendes P, Kell D. 1998. Non-linear optimization of biochemical pathways: Applications to metabolic engineering and parameter estimation. Bioinformatics (Oxford, England) 14(10):869–883.

Moles CG, Mendes P, Banga JR. 2003. Parameter estimation in biochemical pathways: A comparison of global optimization methods. Genome Res 13(11):2467–2474.

Nitabach MN. 2005. Circadian rhythms: Clock coordination. Nature 438(7065):173–175.

Press WH. 1997. Numerical recipes in C: The art of scientific computing. Cambridge; New York: Cambridge University Press. xxvi, 994 p.

Rand DA. Mapping global sensitivity of cellular network dynamics: Sensitivity heat maps and a global summation law. J R Soc Interface 5(Suppl 1):S59–S69.

Rothenfluh A, Abodeely M, Price JL, Young MW. 2000. Isolation and analysis of six timeless alleles that cause short- or long-period circadian rhythms in Drosophila. Genetics 156(2):665–675.

Sehgal A, Rothenfluh-Hilfiker A, Hunter-Ensor M, Chen Y, Myers MP, Young MW. 1995. Rhythmic expression of timeless: A basis for promoting circadian cycles in period gene autoregulation. Science 270(5237):808–810.

Smolen P, Baxter DA, Byrne JH. 2001. Modeling circadian oscillations with interlocking positive and negative feedback loops. J Neurosci 21(17):6644–6656.

Smolen P, Hardin PE, Lo BS, Baxter DA, Byrne JH. 2004. Simulation of Drosophila circadian oscillations, mutations, and light responses by a model with VRI, PDP-1, and CLK. Biophys J 86(5):2786–2802.

Stelling J, Gilles ED, Doyle FJ III. 2004. Robustness properties of circadian clock architectures. Proc Natl Acad Sci USA 101(36):13210–13215.

Stephens MA. 1970. Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables. J R Stat Soc [Ser B] 32(1):115–122.

Tyson JJ, Hong CI, Thron CD, Novak B. 1999. A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM. Biophys J 77(5):2411–2417.

Ueda H, Hirose K, Iino M. 2002. Intercellular coupling mechanism for synchronized and noise-resistant circadian oscillators. J Theor Biol 216(4):501–512.

van Gelder RN, Herzog ED, Schwartz WJ, Taghert PH. 2003. Circadian rhythms: In the loop at last. Science 300(5625):1534–1535.

Varma A, Morbidelli M, Wu H. 1999. Parametric sensitivity in chemical systems. Cambridge: Cambridge University Press. xvi, 342 p.

Vilar JM, Kueh HY, Barkai N, Leibler S. 2002. Mechanisms of noise-resistance in genetic oscillators. Proc Natl Acad Sci USA 99(9):5988–5992.

Xie Z, Kulasiri D. 2007. Modelling of circadian rhythms in Drosophila incorporating the interlocked PER/TIM and VRI/PDP1 feedback loops. J Theor Biol 245(2):290–304.

Young MW, Kay SA. 2001. Time zones: A comparative genetics of circadian clocks. Nat Rev Genet 2(9):702–715.