

# Mathematical Identification of Critical Reactions in the Interlocked Feedback Model

Hiroyuki Kurata\*, Takayuki Tanaka, Fumitaka Ohnishi

Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Fukuoka, Japan

Dynamic simulations are necessary for understanding the mechanism of how biochemical networks generate robust properties to environmental stresses or genetic changes. Sensitivity analysis allows the linking of robustness to network structure. However, it yields only local properties regarding a particular choice of plausible parameter values, because it is hard to know the exact parameter values *in vivo*. Global and firm results are needed that do not depend on particular parameter values. We propose mathematical analysis for robustness (MAR) that consists of the novel evolutionary search that explores all possible solution vectors of kinetic parameters satisfying the target dynamics and robustness analysis. New criteria, parameter spectrum width and the variability of solution vectors for parameters, are introduced to determine whether the search is exhaustive. In robustness analysis, in addition to single parameter sensitivity analysis, robustness to multiple parameter perturbation is defined. Combining the sensitivity analysis and the robustness analysis to multiple parameter perturbation enables identifying critical reactions. Use of MAR clearly identified the critical reactions responsible for determining the circadian cycle in the *Drosophila* interlocked circadian clock model. In highly robust models, while the parameter vectors are greatly varied, the critical reactions with a high sensitivity are uniquely determined. Interestingly, not only the *per-tim* loop but also the *dclk-cyc* loop strongly affect the period of PER, although the *dclk-cyc* loop hardly changes its amplitude and it is not potentially influential. In conclusion, MAR is a powerful method to explore wide parameter space without human-biases and to link a robust property to network architectures without knowing the exact parameter values. MAR identifies the reactions critically responsible for determining the period and amplitude in the interlocked feedback model and suggests that the circadian clock intensively evolves or designs the kinetic parameters so that it creates a highly robust cycle.

Citation: Kurata H, Tanaka T, Ohnishi F (2007) Mathematical Identification of Critical Reactions in the Interlocked Feedback Model. PLoS ONE 2(10): e1103. doi:10.1371/journal.pone.0001103

## INTRODUCTION

A goal of systems biology is the generation of a coherent understanding of the mechanisms of how the integration of various components generates functional and robust systems in living organisms [1–3]. The importance for robustness is a functional criterion for performance of biochemical networks. Robustness is the ability to resume successful operation in the presence of signal and system uncertainties. Descriptions of the mechanisms by which biochemical networks are robust in the face of parameter uncertainty, environmental changes, and stochastic fluctuations have been published [3–6]. Different types of feedback loops and pathway redundancies are involved in enhanced robustness. Sensitivity analysis of mathematical models describing complex networks could allow for linking robustness properties to network structure by measuring the degree to which parametric perturbations change various target dynamics [7,8].

In principle, both molecular architecture and the values of kinetic parameters determine the robustness of dynamic systems. In biological systems, molecular structures are being built, but it is still hard to know the exact values of kinetic parameters *in vivo*. The values of kinetic parameters vary with time and environment and the measured values *in vitro* are often different from those *in vivo*. In most studies, a particular set of local kinetic parameters has been determined for convenience so that dynamic models reproduce target data. It is necessary to exclude the possibility that the calculated results are dependent on variations in the values assigned to kinetic parameters. The dependency of robustness on parameter values has been investigated by only a few reports. To compare some performances of alternative mathematical models, Alves *et al* [9] statistically searched their parameter values so that they make the other dynamic properties the same. Stelling *et al* [10] studied dynamic properties linked to network structure in the *per-tim* feedback loop model by systematically investigating the 2D-

parameter space and suggested some influential process determining the oscillator features. These previously presented random or systematic searches are a great step for approaching to global analysis, but they restricted the search space of parameters or the size of models due to calculation complexity and few criteria have been presented to determine whether the search is exhaustive.

The molecular mechanisms of how circadian clocks generate robust cycles have extensively been studied and negative feedback loops are found common structures for producing robust cycles [11–13]. An interlocked feedback model is typical architecture to provide robustness of the cycle and synchronization of two clocks to environmental perturbation or parameter uncertainty [14], while critical reactions responsible for determining cycle features such as amplitude and period remain to be seen due to experimental complexity. Generally both network architecture and its associated parameter values determine the system's function. For example,

.....  
**Academic Editor:** Nick Monk, University of Nottingham, United Kingdom

**Received** July 30, 2007; **Accepted** October 12, 2007; **Published** October 31, 2007

**Copyright:** © 2007 Kurata et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 2006, 18300098

**Competing Interests:** The authors have declared that no competing interests exist.

\* **To whom correspondence should be addressed.** E-mail: kurata@bio.kyutech.ac.jp

engineering design requires the tuning process of many parameters for producing an optimal and highly robust system. However, neither the optimal values of parameters nor the reactions critically responsible for a robust cycle are known in circadian clock systems.

Here, we propose a new strategy, named mathematical analysis for robustness (MAR), that combines robustness analysis with an evolutionary search that explores many plausible solutions for kinetic parameters producing the same dynamic properties. To search large parameter spaces without any human bias, the two-step evolutionary method is proposed and novel and rigorous criteria are introduced that characterize whether the search approaches an exhaustive level. MAR leads to an understanding of the global mechanism by which the *Drosophila* interlocked feedback system generates the robustness of the circadian cycle without insisting on the exact values of kinetic parameters.

## RESULTS

### Mathematical Analysis for Robustness algorithm

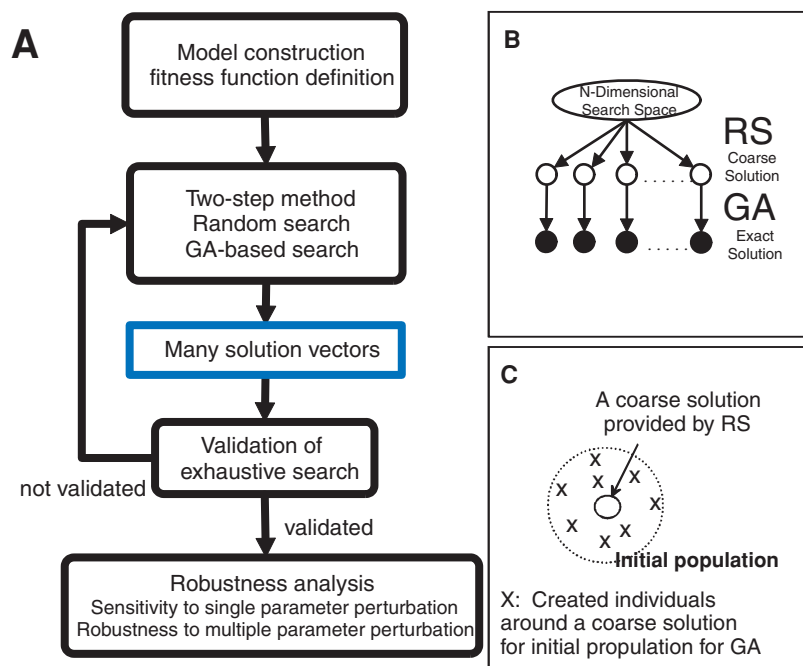
To link molecular architecture to robustness we propose MAR as shown in **Figure 1**. It consists of two stages: exhaustive search and robustness analysis. In the former stage, (1) we determine the target dynamic features that a mathematical model should simulate, (2) explore all possible solution vectors of kinetic parameters that reproduce the target features using the two-step evolutionary search, and (3) test the validity of exhaustive search (**Equations 3, 7**). In the latter stage, (4) we simulate the sensitivity of target features to perturbation of a single parameter (**Equation 8**), (5) calculate the robustness to perturbation of multiple parameters (**Equation 13**), and (6) identify the critical reactions responsible for determining target features by combining both the sensitivity and robustness analyses. Details of the algorithms are described in Materials and Methods.

### Validation of an exhaustive search method by using a theoretical model

We propose the two-step evolutionary method that consists of the random search and the subsequent evolutionary search as a fast and non-biased method (**Figure 1BC**). To demonstrate the advantage of the two-step method, we compared it with various evolutionary searches, such as Unimodal Normal Distribution Crossover (UNDX), UNDXm, Blend Crossover (BLX), and Simplex Crossover (SPX) [15–17], by using a theoretical model (**Text S1** and **Equations S1-S5** in **Text S1**). The two-step method was a best choice for a fast and non-biased search among available search method (**Table S1**). To test whether the search by the two-step method becomes saturated or covers the entire solution space, we investigated the convergence of the parameters spectrum width and the variability of solutions (*VarS*) (**Equations 3, 7**, **Figure S1**). Finally the sensitivity distributions were numerically calculated. The range of the simulated sensitivity was consistent with the theoretical range (**Equation S3** in **Text S1**) and the two-step method identified the potentially influential parameter (**Equation 10**, **Figure S2**). By using a theoretical model, the two-step method is demonstrated to explore the entire solution space at a fast speed. The parameter spectrum width and the variability of solutions are useful measures for the search process of MAR.

### Mathematical model of a biochemical oscillator and exhaustive search

MAR is applied to the *Drosophila* circadian clock model as shown in **Figure 2**. We have derived mathematical equations from the interlocked-feedback model developed by Ueda et al. [14], as shown in **Tables 1–3**. The target feature of the circadian oscillator was set as the PER oscillation curves characterized by



**Figure 1. Schematic diagrams of exhaustive search and robustness analysis in MAR.** A: A flowchart for MAR. B: The proposed two-step search method that consists of a random search (RS) and a search by genetic algorithms (GAs). C: How to make the initial populations for a search by GAs. doi:10.1371/journal.pone.0001103.g001

the period and the top and bottom of the curves (amplitude). There are few reliable experimental data that measure the absolute concentrations for the circadian proteins in *Drosophila*, while they have been provided by most literatures as the relative value to a constant level molecule. Thus, we produce a mathematical but qualitative model that reproduces the intrinsic features of the circadian cycle and adjust it so that the simulated time courses of the proteins are consistent with those presented by commonly employed simulation models, e.g., the PER concentration varies from a few nM to a few dozen nM [18]. To explore the target PER oscillator, a fitness function was designed and twelve kinetic parameters were selected as the search parameters ( $S[1]$ ,  $A[1]=A[2]$ ,  $R[1]=R[2]$ ,  $V[1]$ ,  $V[2]$ ,  $D[1]$ ,  $P[1]$ ,  $T[1]$ ,  $K[1]$ , and  $K[2]$ ). The two-step method was employed. Details of the fitness function and searches are described in Method (Equations 14, 15) and the validity of these search parameters will be verified in the later section (Numerical sensitivity to single parameter perturbation). One thousand optimized PER oscillators were obtained by performing  $2 \times 10^5$  random searches and the subsequent GAs. We investigated the dynamic features of PER, dCLK and TIM, where the circadian model was optimized with regard to the PER cycle (Figures S3, S4). The period for dCLK or TIM was consistent with that for PER. The amplitude and mean concentration of PER oscillators varied over a narrow range of values and the ratio of the amplitude to the mean concentration was approximately 1.8. The amplitude and mean concentration of dCLK and TIM varied over a wider range of values than those for PER, and dCLK and TIM oscillated with ratios of 1.8 and 1.4,

respectively. These results indicate that they are good oscillators and reproduce the target features. According to experimental data regarding the circadian oscillator, the ratio of the amplitude to the mean concentration varies from 1.2 to 2, depending on experimental conditions. These values of the ration are reasonable within the range of experimental data.

### Characterization of exhaustive search

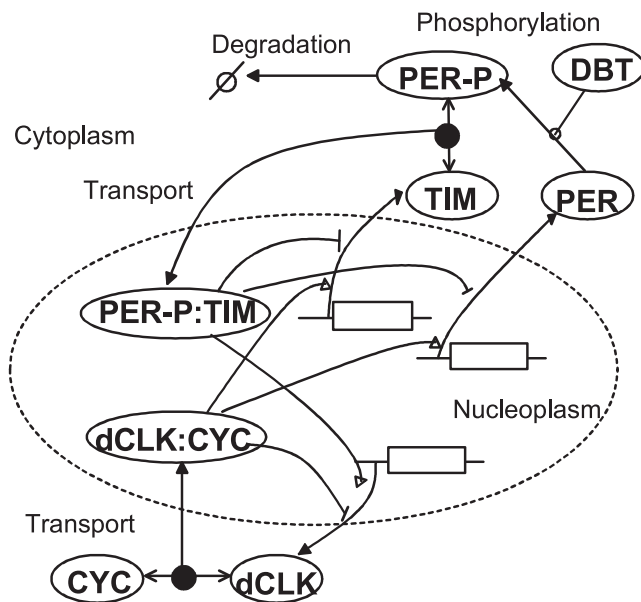
To characterize whether the search is sufficient or exhaustive we investigated the convergence for two measures: the parameter-spectrum width (Equation 3) and the variability in the feasible solution space of the parameter vectors (Equation 7). Figure 3 shows the changes in the parameter-spectrum width with respect to the number of solutions. The solutions were added by the progression of evolutionary searches. For all the search parameters, the parameter spectrum width increased with the number of solutions and then converged above a solution number of 400.

To reveal the variability of the 12-dimensional parameter vectors optimized to provide the similar time course for PER, we classified those solutions by using a hierarchical clustering with the average linkage method, where the parameter distances  $D(\mathbf{P}(i), \mathbf{P}(j))$  were calculated among the solutions (Equation 4). When the parameter spectra converged at a solution number of 400, we clustered the early 400 solutions into six super-balls, where the centroid for each cluster is employed as a representative vector.

Here we demonstrate whether the super-balls for those clusters are sufficient for containing the subsequently optimized solutions (from 401 to 1000). As shown in Figure 4, we plot  $VarS$  (Equation 7) with respect to the solution number, while the number of the employed clusters is varied. This figure indicates the convergence of the variability in solution vectors with respect to the solution number. By using  $VarS$ , we investigate if all the solutions belong to one of the six clusters.  $VarS$  is less than one for the early 400 solutions when the six clusters are employed. When solutions were freshly added by the progression of evolutionary searches, the value of  $VarS$  increased, showing that some new solutions are distant from the centroid of the employed clusters. An increase in the number of employed clusters decreased the  $VarS$  value, which confirms that the increase in the employed clusters enlarges the solution space. When all the six clusters were employed,  $VarS$  was saturated close to one. This shows that all the solutions can be assigned to one of the employed clusters, suggesting that the search process is saturated at a solution number of 400.

### Numerical sensitivity to single parameter perturbation

The period and amplitude sensitivities for the PER oscillator were simulated for 400 optimized solutions. The numerical sensitivity of period ( $S_T$ ) or amplitude ( $S_A$ ) to variations in a single kinetic parameter was calculated (Equation 16). Broad distributions of the absolute sensitivities occur as shown in Figure 5, where the minimum, maximum and mean values are plotted with respect to each kinetic parameter. The mean or minimum values indicate that this model provides a robust property of period and amplitude to parameter uncertainty. The period sensitivity was less than the amplitude sensitivity. Applying a threshold value of  $10^{-8}$  to the minimum sensitivities allowed the division of kinetic parameters into potentially influential and non-influential categories (Equation 10). Seventeen potentially influential parameters for period and amplitude were identified respectively, which were consistent between them. PER synthesis and degradation ( $S[1]$ ,  $S[2]$ ,  $D[1]$ ,  $D[2]$ ,  $D[3]$ ,  $D[6]$ ,  $D[7]$ ), phosphorylation and transport



Binding ( $\leftrightarrow$ ), protein synthesis or conversion ( $\rightarrow$ ), catalyzing the reaction ( $\rightarrow$ ), suppressing the protein synthesis ( $\dashv$ ), and activating the protein synthesis ( $\rightarrow$ ).

**Figure 2. A schematic diagram of the interlocked feedback system in the *Drosophila* circadian clock.** In the real biochemical reactions, the binding of PER-P:TIM to dCLK:CYC suppresses the transcription for PER and TIM and activates the transcription for dCLK. In this map, PER-P:TIM directly suppresses the transcription of PER and TIM and activates the transcription for dCLK. ":" and "-" indicate binding complexes and modification, respectively.

doi:10.1371/journal.pone.0001103.g002

**Table 1.** Mathematical equations for the interlocked-feedback system in *Drosophila*

Mathematical Equations	No
$\frac{d[mRNA(PER)]}{dt} = S[1] \cdot \frac{(\frac{dCLK:CYC(n)}{A[1]})^a + B[1]}{1 + (\frac{PER-TIM(n)}{R[1]})^r + (\frac{dCLK:CYC(n)}{A[1]})^a + B[1]} - D[1] \cdot [mRNA(PER)]$	M1
$\frac{d[PER]}{dt} = S[2] \cdot [mRNA(PER)] - P[1] \cdot \frac{[PER]}{K[1] + [PER]} - D[2] \cdot [PER]$	M2
$\frac{d[PER - P]}{dt} = P[1] \cdot \frac{[PER]}{K[1] + [PER]} - D[3] \cdot [PER - P]$	M3
$\frac{d[mRNA(TIM)]}{dt} = S[3] \cdot \frac{(\frac{dCLK:CYC(n)}{A[2]})^a + B[2]}{1 + (\frac{PER-P:TIM(n)}{R[2]})^r + (\frac{dCLK:CYC(n)}{A[2]})^a + B[2]} - D[4] \cdot [mRNA(TIM)]$	M4
$\frac{d[TIM]}{dt} = S[4] \cdot [mRNA(TIM)] - V[1] \cdot [PER - P] \cdot [TIM] + V[2] \cdot [PER - P : TIM] - D[5] \cdot TIM$	M5
$\frac{d[PER - P : TIM]}{dt} = V[1] \cdot [PER - P] \cdot [TIM] - V[2] \cdot [PER - P : TIM] - T[1] \cdot \frac{[PER - P : TIM]}{K[2] + [PER - P : TIM]} + T[2] \cdot \frac{[PER - P : TIM(n)]}{K[3] + [PER - P : TIM(n)]} - D[6] \cdot [PER - P : TIM]$	M6
$\frac{d[PER - P : TIM(n)]}{dt} = T[1] \cdot \frac{[PER - P : TIM]}{K[2] + [PER - P : TIM]} - T[2] \cdot \frac{[PER - P : TIM(n)]}{K[3] + [PER - P : TIM(n)]} - D[7] \cdot [PER - P : TIM(n)]$	M7
$\frac{d[mRNA(dCLK)]}{dt} = S[5] \cdot \frac{(\frac{PER-P:TIM(n)}{A[3]})^a + B[3]}{1 + (\frac{dCLK:CYC(n)}{R[3]})^r + (\frac{PER-P:TIM(n)}{A[3]})^a + B[3]} - D[8] \cdot mRNA(CLK)$	M8
$\frac{d[dCLK]}{dt} = S[6] \cdot [mRNA(dCLK)] - V[3] \cdot [dCLK] \cdot [CYC] + V[4] \cdot [dCLK : CYC] - D[9] \cdot [dCLK]$	M9
$\frac{d[dCLK : CYC]}{dt} = V[3] \cdot [dCLK] \cdot [CYC] - V[4] \cdot [dCLK : CYC] - T[3] \cdot \frac{[dCLK : CYC]}{K[4] + [dCLK : CYC]} + T[4] \cdot \frac{[dCLK : CYC(n)]}{K[5] + [dCLK : CYC(n)]} - D[10] \cdot [dCLK : CYC]$	M10
$\frac{d[dCLK : CYC(n)]}{dt} = T[3] \cdot \frac{[dCLK : CYC]}{K[4] + [dCLK : CYC]} - T[4] \cdot \frac{[dCLK : CYC(n)]}{K[5] + [dCLK : CYC(n)]} - D[11] \cdot [dCLK : CYC(n)]$	M11

doi:10.1371/journal.pone.0001103.t001

**Table 2.** Components used in the mathematical model

Component	Definition	Concentration
<i>mRNA(PER)</i>	mRNA for PER	1.0 [nM]
<i>PER</i>	PER	1.0 [nM]
<i>mRNA(TIM)</i>	mRNA for TIM	1.0 [nM]
<i>TIM</i>	TIM	1.0 [nM]
<i>mRNA(dCLK)</i>	mRNA(dCLK)	1.0 [nM]
<i>dCLK</i>	dCLK	1.0 [nM]
<i>PER-P</i>	Phosphorylated PER	1.0 [nM]
<i>PER-P:TIM</i>	Binding complex of PER-P and TIM	1.0 [nM]
<i>PER-P:TIM(n)</i>	Binding complex of PER-P and TIM in nucleus	1.0 [nM]
<i>dCLK:CYC</i>	Binding complex of dCLK and CYC	1.0 [nM]
<i>dCLK:CYC(n)</i>	Binding complex of dCLK and CYC in nucleus	1.0 [nM]
<i>CYC (constant)</i>	CYC	1.0 [nM]

The concentration shows the initial value and n indicates that molecules are located in nucleus. “:” represents a binding complex of more than one protein; “.” represents covalent modification of a protein. Since the concentration of CYC is fixed at a constant value, there is no differential equation for CYC in

**Table 1.**

doi:10.1371/journal.pone.0001103.t002

for PER (*P[1]*, *K[1]*, *T[1]*, *K[2]*), TIM synthesis and degradation (*S[3]*, *S[4]*, *D[4]*, *D[5]*), and binding of PER-P and TIM (*V[1]*, *V[2]*) are potentially influential. It indicates that the effect of every reaction involving the *per-time* loop cannot be cancelled in any possible models. It is reasonable because the *per-tim* loop is not redundant. By contrast, the kinetic parameters involving the *dclk-cyc* loop are not potentially influential. On the other hand, a maximum value of period sensitivity was more than 0.3 (7.2 h) with respect to some kinetic parameters, indicating that not only network structures but also parameter values are critical for generating a robust cycle. If an inappropriate set of kinetic parameters is provided, an oscillation can be damaged by even a small change in a single parameter. The interlocked feedback is able to provide robustness to the circadian rhythm when the kinetic parameters are appropriately tuned.

Numerical sensitivity analysis selected 17 potentially influential parameters within the provided 12-dimensional space. A question raises whether these parameters can be potentially influential in the entire 36 dimensional parameter space. Of course, such a complete search is impossible due to calculation explosion inherent of numerical analysis. Instead, an extended search is performed to investigate how an expansion of a search parameter space affects the choice of potentially influential parameters, where the number of search parameters is varied as 3, 6, 9, 12, 16, 25, and 36. (**Table S2** and **Figure S5**). The exploration with a search parameter number of  $\leq 9$ , a low-dimensional subspace, could not

**Table 3.** Kinetic parameters used in the mathematical model

Para-meter	Definition	Value	Class
S[1]	Transcription rate constant of PER	0.98 [nM h <sup>-1</sup> ]	S
S[2]	Translation rate constant of mRNA(PER)	0.98 [h <sup>-1</sup> ]	A
S[3]	Transcription rate constant of TIM	45 [nM h <sup>-1</sup> ]	A
S[4]	Translation rate constant of mRNA(TIM)	0.98 [h <sup>-1</sup> ]	A
S[5]	Transcription rate constant of dCLK	6.0 [nM h <sup>-1</sup> ]	A
S[6]	Translation rate constant of mRNA(dCLK)	23 [M <sup>-1</sup> ]	A
A[1]	DNA dissociation constant for <i>dCLK:CYC(n)</i>	0.14 [M]	S
A[2]	DNA dissociation constant for <i>dCLK:CYC(n)</i>	0.14 [M]	S
A[3]	DNA dissociation constant for <i>PER-P:TIM(n)</i>	0.21 [M]	A
R[1]	DNA dissociation constant for <i>PER-P:TIM(n)</i>	0.10 [M]	S
R[2]	DNA dissociation constant for <i>PER-P:TIM(n)</i>	0.10 [M]	S
R[3]	DNA dissociation constant for <i>dCLK:CYC(n)</i>	0.18 [M]	A
B[1]	Constant	0.71	A
B[2]	Constant	0.71	A
B[3]	Constant	0.71	A
V[1]	Association rate constant between PER-P and TIM	8.6 [nM <sup>-1</sup> h <sup>-1</sup> ]	S
V[2]	Dissociation rate constant between PER-P:TIM	0.28 [h <sup>-1</sup> ]	S
V[3]	Association rate constant between dCLK and CYC	0.095 [nM <sup>-1</sup> h <sup>-1</sup> ]	A
V[4]	Dissociation rate constant between dCLK:CYC	0.0038 [h <sup>-1</sup> ]	A
T[1]	Maximum rate constant for transportation of PER-P:TIM(cyt->nuc)	0.53 [nM h <sup>-1</sup> ]	S
T[2]	Maximum rate for transportation of PER-P:TIM(nuc->cyt)	0.0 [nM h <sup>-1</sup> ]	A
T[3]	Maximum rate for transportation of dCLK:CYC(cyt->nuc)	0.42 [nM h <sup>-1</sup> ]	A
T[4]	Maximum rate for transportation of dCLK:CYC(nuc->cyt)	0.0 [nM h <sup>-1</sup> ]	A
K[1]	Michaelis constant for phosphorylation of PER	42 [nM]	S
K[2]	Michaelis constant for transportation of PER-P:TIM	6.8 [nM]	S
K[3]	Michaelis constant for transportation of PER-P:TIM(nuc)	0.0 [nM]	A
K[4]	Michaelis constant for transportation of dCLK:CYC	2.3 [nM]	A
K[5]	Michaelis constant for transportation of dCLK:CYC(nuc)	0.0 [nM]	A
P[1]	Maximum rate constant for phosphorylation of PER	15 [nM h <sup>-1</sup> ]	S
D[1]	Degradation rate constant of mRNA(PER)	0.56 [h <sup>-1</sup> ]	S
D[2]	Degradation rate constant of PER	0.12 [h <sup>-1</sup> ]	A
D[3]	Degradation rate constant of PER-P	1.2 [h <sup>-1</sup> ]	A
D[4]	Degradation rate constant of mRNA(TIM)	0.60 [h <sup>-1</sup> ]	A
D[5]	Degradation rate constant of TIM	0.12 [h <sup>-1</sup> ]	A
D[6]	Degradation rate constant of PER-P:TIM	0.12 [h <sup>-1</sup> ]	A
D[7]	Degradation rate constant of PER-P:TIM(n)	0.12 [h <sup>-1</sup> ]	A
D[8]	Degradation rate constant of mRNA(dCLK)	2.3 [h <sup>-1</sup> ]	A
D[9]	Degradation rate constant of dCLK	0.12 [h <sup>-1</sup> ]	A
D[10]	Degradation rate constant of dCLK:CYC	0.12 [h <sup>-1</sup> ]	A
D[11]	Degradation rate constant of dCLK:CYC(n)	0.12 [h <sup>-1</sup> ]	A
a	Hill coefficient for activation	4	A
r	Hill coefficient for repressor	8	A

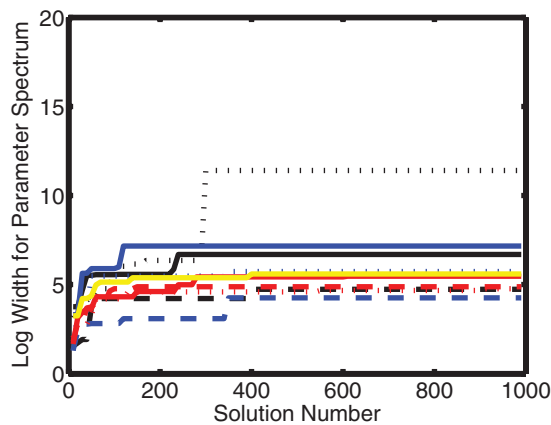
S: the kinetic parameters estimated by evolutionary searches, A: the values estimated/provided in the model.  
doi:10.1371/journal.pone.0001103.t003

obtain the consistent sets of potentially influential parameters. On the other hand, the analyses with a search parameter number of  $\geq 16$  selected the same parameter sets as the 12-dimensional search. Although the variability in the solutions does not necessarily converge less than one above a search parameter number of  $\geq 16$  due to an extremely large search space, this extended search suggests that the potentially influential parameters

are consistent. In conclusion, twelve search parameters are necessary and sufficient for finding them.

### Robustness to multiple parameter perturbation

There are many possible models that generate circadian oscillators. It raises a question of which models are close to a real

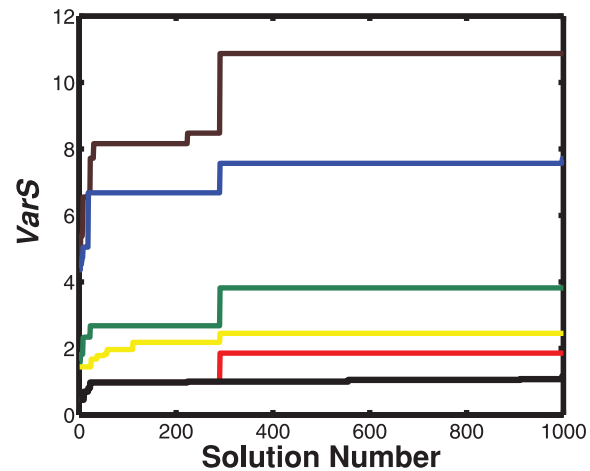


**Figure 3. Convergence of the logarithmic width of the parameter spectrum for 12 search parameters.** One thousand solutions for optimized kinetic parameters were obtained by using the repetition of the two-step search. The black solid line is  $S[1]$ , the black dotted line  $A[1]=A[2]$ , the black chain  $R[1]=R[2]$ , the blue solid line  $V[1]$ , the blue dotted line  $V[2]$ , the blue chain  $D[1]$ , the red solid line  $P[1]$ , the red dotted line  $T[1]$ , the yellow solid line  $K[1]$ , and the red chain  $K[2]$ . doi:10.1371/journal.pone.0001103.g003

biochemical oscillator or a question of which reactions feature robust oscillators. Robustness to multiple parameter perturbation is a key criterion for model selection. Here we define the CVs of the period and amplitude distributions generated by randomly-perturbed parameters  $CV\_Period(i)$  and  $CV\_Amplitude(i)$  (Equation 13), which characterize the robustness of the period and amplitude to simultaneous changes in all kinetic parameters, assuming that the oscillators are always exposed to uncertainty or changes in parameters under strict environmental stresses. An oscillator with a small value of CV is a highly robust model, because the change in period or amplitude is suppressed to be small against random changes in all kinetic parameters.

For all 400 solution vectors (oscillator models), the period and amplitude were simulated 10,000 times by randomly varying all kinetic parameters within a range from 1/1.5-fold to 1.5-fold, resulting in 400 distributions of the simulated period and amplitude. The means and CVs almost converged at 10,000 time simulations. The frequency distributions of the period and amplitude for a certain parameter vector are exemplified as shown in Figure 6. We plotted the frequency distribution of the CVs of period and amplitude for 400 solution vectors as shown in Figure 7. The CVs of period changed from 0.08 to 0.44. The mean periods were around 24 h (data not shown). A CV of 0.08 indicates 1.9 h, which significantly changes the circadian cycle. On the other hand, the CVs of amplitude changed above 0.4, which was larger than the CVs of period. The mean amplitudes were around 10 nM, which is reasonable because it is within the range of commonly simulated values [18]. This figure indicates that an appropriate selection of kinetic parameters decreases the CVs, i.e., a tuning of kinetic parameters enhances robustness to random changes in all kinetic parameters. It is interesting if a biological system adjusts kinetic parameter values so that robustness is most enhanced.

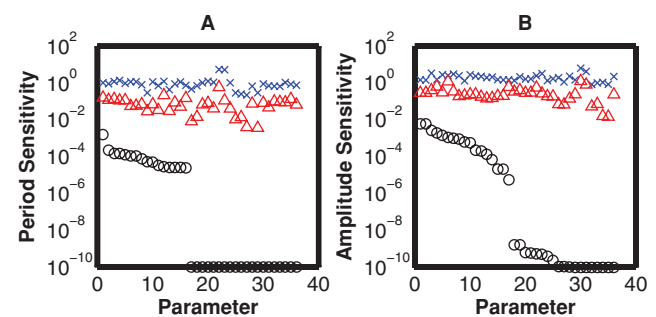
To investigate the mechanism of how the circadian clock models provide a robust property to the period and amplitude, the relationship between the CVs of period and amplitude was plotted (Figure S6). Significant linear correlation was not observed between the CVs of period and amplitude. The mechanism that provides a robust property to the period is suggested to be different from that to the amplitude.



**Figure 4. Convergence for the variability in the solutions of kinetic parameters.** Six clusters were generated from using the early 400 solutions. The number of the employed clusters is changed from 1 to 6.  $VarS$  (Equation 7) is plotted with respect to the solution number.  $VarS$  of less than one shows the parameter solutions are included in the employed clusters. The brown line is one cluster employed, the blue line two clusters, the green line three clusters, the yellow four clusters, the red line five clusters, and the black line six clusters. doi:10.1371/journal.pone.0001103.g004

### Critical parameters in highly robust oscillators

MAR combines the numerical sensitivity analysis with the CV analysis to identify the critical reactions in the interlocked feedback model. To explore the parameters that show a high sensitivity of period or amplitude, we investigated how the robustness of the oscillator models,  $CV\_Period(i)$  or  $CV\_Amplitude(i)$  (Equation 13), is correlated to the sensitivity distributions or to the kinetic parameter vectors. The period and amplitude sensitivity distributions were classified into 6 clusters by using hierarchical clustering with the average linkage (Matlab), respectively. They were named sensitivity clusters of 1 to 6. In the same way the parameter solution vectors were classified into 6 clusters, named parameter clusters of 1 to 6. As shown in Figure 8, the cluster index that



**Figure 5. Distributions of period and amplitude sensitivities.** For 400 solutions, we simulated the period and amplitude sensitivity by changing the value of each kinetic parameter by 1.5-fold. A change ratio of 1.5 was validated (data not shown). The absolute sensitivities are employed. The resulting distributions exclude the models whose oscillations are abolished by parameter perturbations ( $|sensitivity| > 10$ ). The sensitivities with a value of less than  $10^{-10}$  are set to  $10^{-10}$ . The black circles indicate the minimum, the blue crosses the maximum, and the red triangles the mean, respectively. The kinetic parameters are sorted according to the minimum value of the absolute sensitivities in the descending order. Use of a threshold value of  $10^{-8}$  separates the critically influential parameters. (A) Period sensitivity, (B) Amplitude sensitivity. doi:10.1371/journal.pone.0001103.g005

each model belongs to is plotted with respect to the model sorted in the ascending order of the CVs of period or amplitude.

For a high CV period (CV period order  $>300$ ) the period sensitivity distributions were varied because they belonged to all the clusters. By contrast, in highly robust models with a small CV value (CV period order  $<50$ ) their sensitivity distributions were consistent because most of the models belonged to the one cluster, while the parameter vectors were greatly varied (they belonged to six clusters). For a high CV amplitude (CV amplitude order  $>300$ ) the amplitude sensitivity distributions were varied. In highly robust models with a small CV value (CV amplitude order  $<30$ ), their sensitivity distributions were consistent because most of the models belonged to the one cluster, while the parameter vectors were different. The period and amplitude sensitivity distributions and their associated parameter vectors are illustrated (**Figures S7 and S8**). These demonstrate that the period or amplitude sensitivity distributions are very consistent in highly robust models, while the parameter vectors are greatly varied. In other words the critical reactions are uniquely determined in highly robust models despite the variability of the kinetic parameters.

Next, the critical parameters were selected from the sensitivity distributions of both period and amplitude. For ten highly robust models, we calculated the means of period and amplitude absolute sensitivities and normalized them by their maximum values, as shown in **Figure 9**. Twelve critical parameters showing a high sensitivity were selected that determine period and amplitude, respectively. The critical parameters responsible for period are  $D[7]$ ,  $V[2]$ ,  $T[3]$ ,  $D[9]$ ,  $D[6]$ ,  $P[1]$ ,  $D[1]$ ,  $D[10]$ ,  $V[3]$ ,  $D[11]$ ,  $K[1]$ ,  $D[4]$ ; those responsible for amplitude are  $R[1]$ ,  $D[7]$ ,  $T[1]$ ,  $K[2]$ ,  $D[1]$ ,  $D[4]$ ,  $R[2]$ ,  $P[1]$ ,  $V[2]$ ,  $S[1]$ ,  $S[2]$ ,  $D[2]$ . (The parameters are sorted in the rank order of the absolute sensitivity value.) They feature highly robust oscillators (**Figures S9**).

Degradation ( $D[1], D[2], D[6], D[7], D[9], D[10], D[11]$ ) is the key reaction that determines both the period and amplitude of the PER oscillator. Since the CVs of period were hardly correlated with those of amplitude (**Figure S6**) and the period sensitivity distributions were different from the amplitude ones, the mechanisms providing robust properties to period and amplitude are different. This is supported elsewhere [10]. Here, the kinetic parameters are classified according to their biochemical processes. The amplitude of PER can be determined mainly by reactions in the *per-tim* loop ( $S[1]$ ,  $S[2]$ ,  $R[1]$ ,  $R[2]$ ,  $T[1]$ ,  $K[2]$ ,  $D[1]$ ,  $K[1]$ ,  $D[4]$ ), especially the synthesis and transport of PER and TIM ( $S[1]$ ,  $S[2]$ ,  $R[1]$ ,  $R[2]$ ,  $T[1]$ ,  $K[2]$ ). Both  $T[1]$  and  $K[2]$  are critical, indicating transport cannot be represented as a single linear reaction. It suggests that a nonlinear reaction plays a significant role in determining the amplitude. On the other hand, the period of PER is strongly affected not only by the *per-tim* loop but also by the *dclk-cyc* loop ( $D[9]$ ,  $D[10]$ ,  $D[11]$ ,  $V[3]$ ,  $T[3]$ ), although the *dclk-cyc* loop is neither potentially influential nor does affect the amplitude of PER. It shows that the *dclk-cyc* loop plays a critical role in determining the PER period. The phosphorylation reaction ( $P[1]$ ,  $K[1]$ ) would play a major role in determining the period rather than the amplitude. Both  $P[1]$  and  $K[1]$  are critical, which suggests that a nonlinear reaction is necessary for determining the period.

## DISCUSSION

The objective of MAR is to determine critical reactions without insisting on the exact values of kinetic parameters, thereby linking robust properties to specific molecular architectures. MAR seeks global and firm results that do not depend on particular parameter values. It is a powerful tool because it is still very hard to measure full kinetics *in vivo*. The key technique is to combine robustness analysis

with a two-step evolutionary search that explores many solutions for kinetic parameters showing the similar dynamic features.

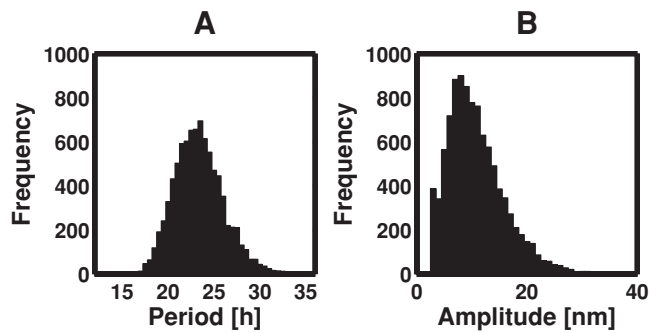
Compared with the previous method that systematically analyzed two-dimensional parameter space in the *per-time* feedback model [10], MAR enables more precise analysis of more complex models, the interlocked feedback model that contains the *per-tim* and *dclk-cyc* loops. MAR is advantageous because it enables one to explore a higher dimensional and much larger space of the kinetic parameters of a complex model. Searching a larger parameter space is demonstrated to make clearer separation of potentially influential parameters (**Table S2** and **Figure S5**). Notice that a low dimensional search may find a false set of potentially influential parameters. In the circadian model, a search parameter number of  $\geq 12$  should be used to obtain the true set of potentially influential reactions in a large parameter space. This shows the extensive search is really required to obtain global and firm results from mathematical analysis.

The potentially influential parameters, which become consistent with progression of search, means that the effect of them cannot be reduced very small or less than a certain threshold value in any possible models, but they do not necessarily indicate critical parameters responsible for target instances. The potentially influential parameters would be effective in testing whether the search is exhaustive rather than in exploring critical reactions.

The central contribution here is that MAR provides the results compatible with biological knowledge or lead to new hypotheses on the function of cellular networks with realistic complexity. By searching wide parameter space, MAR gives global results that a local simulation cannot obtain. The identification of the relationship among the robustness to multiple parameter perturbation (**Equation 13**), the sensitivity distribution, and the parameter solution vectors could provide a basis to assume a potential mechanisms generating robustness. The robustness to multiple parameter perturbation greatly depends on the values of kinetic parameters. Actually, appropriate sets of kinetic parameters greatly enhance the robustness of period or amplitude, while their sensitivity values converge to a unique distribution. In highly robust models the critical reactions are uniquely determined despite the variability of the parameter values.

While the phosphorylation of PER and the transport and synthesis of PER and TIM are critical for period and for amplitude, respectively, degradations are key reactions for both period and amplitude. This outcome agrees with previous simulation studies suggesting that circadian performance is greatly affected by changes in degradation [8,10]. Interestingly, the *dclk-cyc* loop plays a major role in determining the PER period, although the *dclk-cyc* loop is neither potentially influential nor does affect the amplitude of PER. The mechanism generating robustness to period is found different from that to amplitude. These are supported by a general theory that the robustness of one function does not imply that of other functions due to robustness tradeoff[5,6]. Biochemical oscillators should be evolved to provide a robust property to different dynamic functions such as period and amplitude.

Engineering design conventionally consists of the investigation of functional specification and the subsequent basic and detail design. Specification is the act of understanding of a product to be modelled or of defining what a function has to achieve rather than how it has to do it. Basic design carries out the identification, classification and selection of constraints and then determines the design architecture of how the system obtains its specification. The process of detail design manipulates the system's parameters so as to satisfy the specification. In analogy to engineering design, evolution of a circadian clock is understandable in *Drosophila*. The



**Figure 6. Frequency distributions of the period and amplitude in a randomly perturbed oscillator model.** The period (A) and amplitude (B) were simulated 10,000 times while all kinetic parameters were randomly varied within a range from 1/1.5-fold to 1.5-fold. The oscillators with a long period (>36 h) are excluded, resulting in the remove of the oscillators with small amplitude. doi:10.1371/journal.pone.0001103.g006

functional specification can be to provide robustness of the circadian cycle to various perturbations. In basic design under the constraint that the system must use genetic circuits within a cell, an interlocked feedback system is created as network architecture so that two clocks are consistently harmonized with environmental changes such as light pulse and they show robustness of the circadian period against various perturbations[14]. Generally such robustness is provided by molecular architecture such as feedback [3,5,6]. Then the process of detail design further evolves kinetic parameters so as to optimize the specification in the given network structure. In terms of engineering design concept, the circadian clock is suggested to not only evolve an interlocked feedback system but also intensively design kinetic parameters for further enhanced robustness.

## MATERIALS AND METHODS

### General dynamic model

Dynamic models for biochemical networks are formulated as ordinary differential equations:

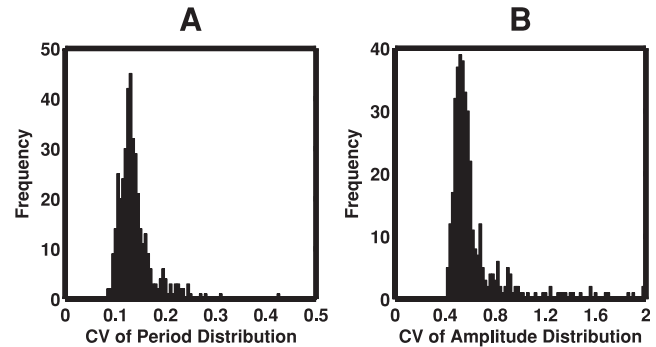
$$\frac{dy}{dt} = \mathbf{F}(t; \mathbf{y}; \mathbf{p}), \quad (1)$$

where  $t$  is the time,  $\mathbf{y}$  is the vector whose elements are the differential variables for molecular concentrations, and  $\mathbf{p}$  is the vector of kinetic parameters.

### Evolutionary search for all possible solutions

The target features are determined that the dynamic model should reproduce. A fitness function must be designed so that the mathematical model reproduces the target dynamics and then all possible solutions are explored to satisfy the fitness function.

Excessive computational requirements make it impossible to explore the entire space of kinetic parameters in large-scale dynamic models. Techniques that address this problem should provide ways to reduce the number of search parameters combined with the method that allows searches within a large parameter space to proceed efficiently. In general, an evolutionary search seeks out a global minimum of certain fitness functions based on the heuristic assumption that best solutions will be found in the regions of the parameter space by using the genetic operations of selection, crossover, and mutation. However, the objective of MAR is not to find a global minimum, but to explore all possible solutions of kinetic parameter vectors that produce the target dynamics.



**Figure 7. Frequency of the CVs of period and amplitude distributions in randomly perturbed oscillator models.** For 400 solution vectors (oscillator models), the period and amplitude were simulated 10,000 times by randomly varying all kinetic parameters within a range from 1/1.5-fold to 1.5-fold, resulting in 400 distributions of the simulated period and amplitude. The CVs were calculated for every period and amplitude distribution and the frequency distributions of the CVs for period (A) and amplitude (B) are plotted for 400 oscillator models. doi:10.1371/journal.pone.0001103.g007

MAR presents the two-step evolutionary search that combines a random search with a search by GAs [17], as shown in **Figure 1**. First, the random search explores a large parameter space without any human biases and finds multiple coarse solutions showing a good fitness value. It is not necessary to find any solutions providing highest fitness values. The coarse solutions are employed to make initial populations for the subsequent GAs. Second, each initial population for GAs is created around one of the coarse solution vectors. The search by GAs focuses on the space surrounding the coarse solution and is applied to each initial population independently to find plausible solution vectors that show a high fitness value, i.e., provide the target features. The search is performed by various crossover methods, resulting in obtaining local solutions around the coarse solutions. The  $i$ -th resultant solution vector of kinetic parameters  $\mathbf{P}(i)$  is given by:

$$\mathbf{P}(i) = (p(i,1), p(i,2), p(i,3), \dots, p(i,N)), \quad (2)$$

where  $p(i, k)$  is the value of the  $k$ -th parameter of the  $i$ -th solution vector and  $N$  is the number of search parameters.

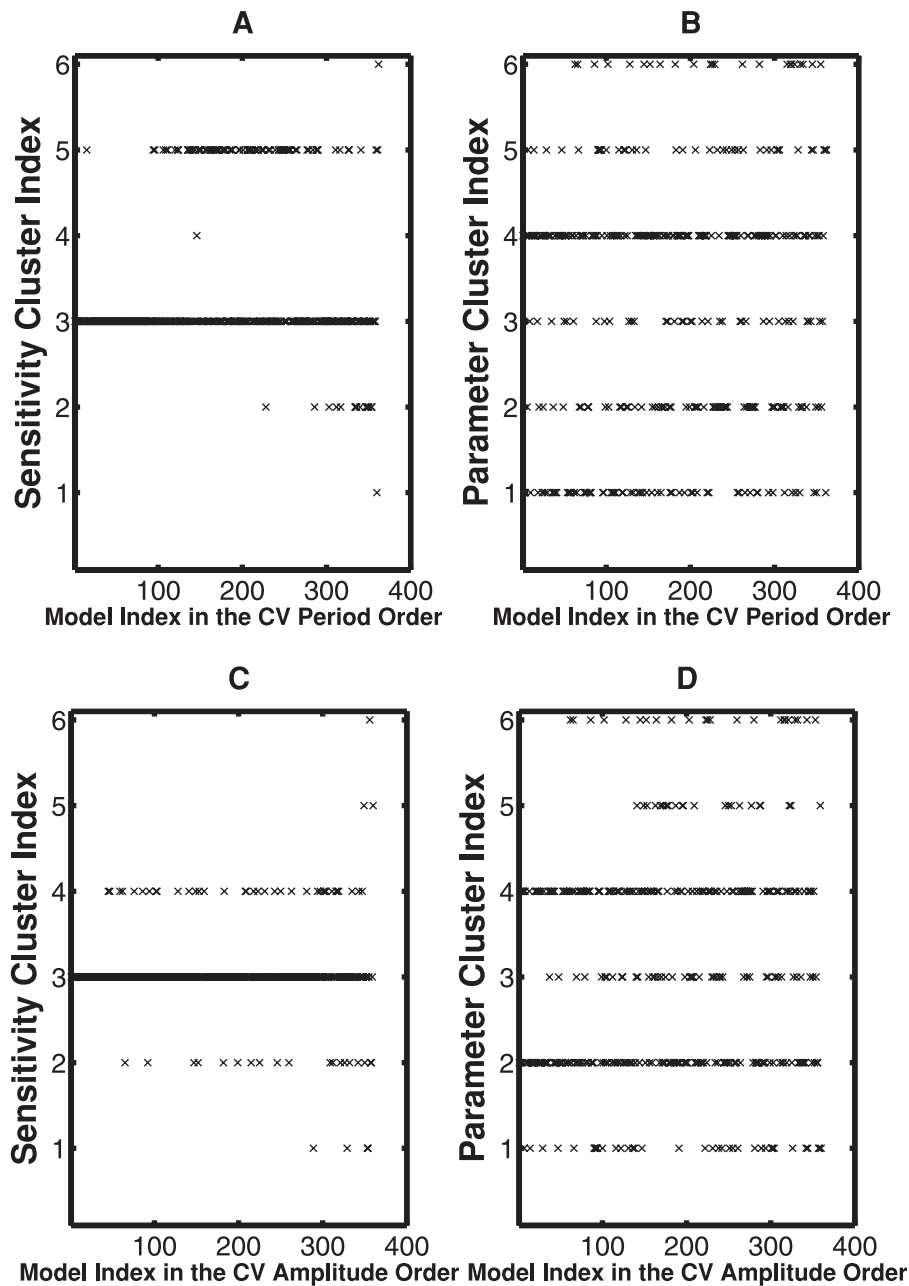
### Characterization of exhaustive search

The proposed two-step method explores multiple plausible solution vectors for kinetic parameters that generate target features or satisfy a fitness function. It is necessary to characterize whether the search is saturated or exhaustive. Here saturation means that the variability in all possible solutions is large enough to cover almost the entire solution space. Since numerical methods are not theoretically able to guarantee the exhaustive search, we present practically useful criteria: a parameter spectrum width and the variability in the solution vectors for parameters.

First, a parameter spectrum is defined as the allowable range of each parameter. The range is from the minimum value to the maximum for the resultant parameter solutions. The logarithmic width of the parameter spectrum for each kinetic parameter is defined by:

$$\begin{aligned} & \text{Log Width for Parameter Spectrum } (k) \\ & = \max_i \{\log_{10} p(i,k)\} - \min_i \{\log_{10} p(i,k)\} \quad (i = 1, 2, \dots, La) \quad (3) \end{aligned}$$





**Figure 8. Classification of the sensitivity distribution and the parameter solution vectors.** The sensitivity distributions and their associated parameter solution vectors for all the oscillator models that reproduce the circadian cycle are classified into six clusters. The six clusters for the sensitivity distributions and their associated parameter vectors are named sensitivity cluster from 1 to 6 and parameter cluster from 1 to 6, respectively. The models are sorted in the ascending order of the CV value, where the left is highly robust models because their CVs are small. The sensitivity and parameter cluster indexes that each model belongs to are plotted with respect to the model index. The sensitivity cluster (A) and parameter cluster (B) are for the models sorted according to the CV period value. The sensitivity cluster (C) and parameter cluster (D) are for the models sorted according to the CV amplitude value. Note that the number of models is less than 400 because some sensitivity distributions with a very large value ( $|\text{sensitivity}| > 10$ ) are excluded.  
doi:10.1371/journal.pone.0001103.g008

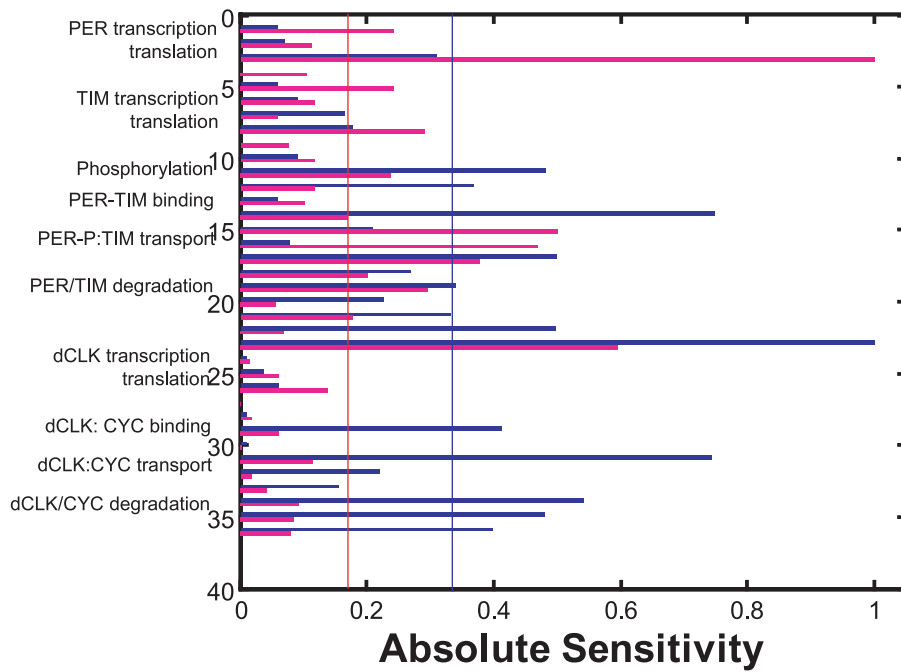
where  $L_s$  is the number of all the solutions. The parameter spectrum width can be regarded as the indicator for a large-space search. Convergence of the width with respect to the number of generated solution vectors suggests that the search is approaching saturation. The number of the solutions at which the spectrum width begins to saturate is defined as  $L_s$ .

Second, to reveal the variability in the solution vectors that show the target dynamic behaviours, the distance between two  $N$ -dimensional solution vectors ( $\mathbf{P}(i), \mathbf{P}(j)$ ) is defined by:

$$(D(\mathbf{P}(i), \mathbf{P}(j))) = \sqrt{\sum_{k=1}^N \{\log_{10} p(i,k) - \log_{10} p(j,k)\}^2} \quad (i \neq j), \quad (4)$$

$(i, j = 1, 2, \dots, L_s).$

We classify the solution vectors from 1 to  $L_s$  by using hierarchical clustering with the average linkage. The number of clusters



**Figure 9. A distribution of the mean absolute sensitivity for period and amplitude in highly robust models.** The blue and red bars are the period and amplitude sensitivities, respectively. The blue line is the threshold for selecting the critical parameters responsible for period and the red for amplitude. The absolute sensitivity values for ten highly robust models that show the smaller CV values are averaged. The resulting mean sensitivity is normalized by the maximum value. 1:*S*[1], 2:*A*[1], 3:*R*[1], 4:*B*[1], 5:*S*[2], 6:*S*[3], 7:*A*[2], 8:*R*[2], 9:*B*[2], 10:*S*[4], 11:*P*[1], 12:*K*[1], 13:*V*[1], 14:*V*[2], 15:*T*[1], 16:*K*[2], 17:*D*[1], 18:*D*[2], 19:*D*[4], 20:*D*[5], 21:*D*[3], 22:*D*[6], 23:*D*[73], 24:*S*[5], 25:*A*[3], 26:*R*[3], 27:*B*[3], 28:*S*[6], 29:*V*[3], 30:*V*[4], 31:*T*[3], 32:*K*[4], 33:*D*[8], 34:*D*[9], 35:*D*[10], 36:*D*[11]. doi:10.1371/journal.pone.0001103.g009

(*M*) is determined by inspecting the dendrogram. Each cluster is named *C*(*j*), (*j* = 1, 2, ..., *M*). Assuming that the clusters form *N*-dimensional super-balls, we define the radius of each cluster as follows:

$$Rcluster(j) = \max\{D(\mathbf{P}(i), \mathbf{G}(j)) | \mathbf{P}(i) \in C(j) \wedge i = 1, 2, \dots, L_s\}, \quad (5)$$

where  $\mathbf{G}(j) = (g(j,1), g(j,2), g(j,3), \dots, g(j,N))$  is the centroid for *C*(*j*) and *g*(*j*, *k*) is the *k*-th parameter of the centroid vector. The distances between the centroid and intra-cluster parameter vectors are calculated and the maximum distance is set to the radius of the cluster ball.

Third, to investigate whether the solution vectors,  $\{\mathbf{P}(i) | i = 1, 2, \dots, L_s\}$ , can be assigned to one of the existing cluster balls, we define the relative minimum distance between  $\mathbf{P}(i)$  and the centroid of each cluster as follows:

$$MinD(\mathbf{P}(i)) = \min \left\{ \frac{\sqrt{\sum_{k=1}^N (\log_{10} p(i,k) - \log_{10} g(j,k))^2}}{Rcluster(j)} \mid j = 1, 2, \dots, M \right\}, \quad (6)$$

(*i* = 1, 2, ..., *L<sub>s</sub>*)

where the relative distance between  $\mathbf{P}(i)$  and each  $\mathbf{G}(j)$  is calculated and the relative minimum distance is selected as *MinD*( $\mathbf{P}(i)$ ). *MinD*( $\mathbf{P}(i)$ ) determines whether a solution vector  $\mathbf{P}(i)$  belongs to one of the existing cluster balls or not. When *MinD*( $\mathbf{P}(i)$ )

is less than one,  $\mathbf{P}(i)$  belongs to one of the existing clusters. When *MinD*( $\mathbf{P}(i)$ ) is greater than one, it does not any cluster.

Finally, to display the convergence of the variability in the solution vectors, we define *VarS*(*L*) as the maximum distance of  $\{MinD(\mathbf{P}(i)) | i = 1, 2, \dots, L_s\}$ :

$$VarS(L) = \max\{MinD(\mathbf{P}(i)) | i = 1, 2, \dots, L_s\}. \quad (7)$$

*VarS*(*L*) is defined as an indicator of whether the *L* parameter solutions fall into one of the cluster-balls generated by the *L<sub>s</sub>* solutions. When *VarS*(*L*) is less than one, the *L* solution vectors are assigned to one of the existing clusters, indicating that the cluster-balls contain all *L* solutions. In contrast, when *VarS*(*L*) is greater than one, these clusters are not able to contain all *L* solutions. Convergence of *VarS* to less than one indicates that the cluster-balls are able to cover the parameter space spanned by the *L* solution vectors. In other words, *L<sub>s</sub>* is large enough to provide the variability in the solutions and the evolutionary search is suggested to approach saturation at a solution number of *L<sub>s</sub>*.

### Numerical sensitivity to single parameter perturbation

It is important to note which structural characteristics of specific molecular networks are responsible for specific instances of robustness. The numerical sensitivity of a target instance to variations in a single kinetic parameter *p*(*i*, *k*) is defined as follows:

$$S(i,k) = \frac{\Delta \ln Target(\mathbf{P}(i))}{\Delta \ln p(i,k)}, \quad (i = 1, 2, \dots, L_s) (k = 1, 2, \dots, N). \quad (8)$$

Sensitivity analysis contributes to finding critical reactions for

determining the target feature, in which kinetic parameters showing high sensitivities are critical.

In MAR the potentially influential parameters for determining a specific target feature are defined. The minimum sensitivity values provided by:

$$\min S(k) = \min\{S(i,k) | i=1,2,\dots,Ls\}, \quad (9)$$

are compared for all kinetic parameters and the potentially influential parameters ( $C$ ) are selected by:

$$C = \{k | \min S(k) > \sigma\}, \quad (10)$$

where  $k$  is the index of the selected parameters and  $\sigma$  is the threshold value that is determined from the distribution of  $\min S(k)$ . The threshold value depends on mathematical models or target instances. The numerical sensitivity for the potentially influential parameters cannot be reduced less than a threshold value of  $\sigma$ , while that for the other parameters can be reduced.

### Robustness to multiple parameter perturbation

There are many possible models that produce the target instance. It raises a question of which models are close to a real biochemical model. Robustness to multiple parameter perturbation is a key criterion for model selection. The intracellular environment consistently varies with time and with changes in external stresses, which would provide uncertainty to kinetics for all intracellular molecules. To characterize the changes in a target instance with respect to simultaneous changes in all kinetic parameters,  $Target(\mathbf{p}_r(i))$  are simulated by randomly varying all kinetic parameters of  $\mathbf{p}_r(i) (i=1,2,\dots,Ls)$  within a specific range given by  $\rho$ , where  $\mathbf{p}_r(i) = (p_r(i,1), p_r(i,2), \dots, p_r(i,N)) (r=1,2,\dots,Nr)$  is the randomly-perturbed parameter vector and  $Nr$  indicates the repetition number of the random simulations. The component of  $\mathbf{p}_r(i)$  is given by:

$$p_r(i,k) = \gamma(r,i,k) \cdot p(i,k), \quad (k=1,2,\dots,N)$$

where  $\gamma(r,i,k)$  is the uniform random value within a range of:

$$\frac{1}{1+\rho} < \gamma(r,i,k) < 1 + \rho. \quad (11)$$

By simulating  $Target(\mathbf{p}_r(i))$ , the frequency distribution of  $Target(\mathbf{p}_r(i))$  occurs. To characterize its distribution, we define the mean and coefficient of variance (CV) by:

$$\text{mean\_Target}(i) = \frac{1}{Nr} \sum_{r=1}^{Nr} Target(\mathbf{p}_r(i)) \quad (12)$$

$$CV\_Target(i) = \frac{\sqrt{\frac{\sum_{r=1}^{Nr} (Target(\mathbf{p}_r(i)) - \text{mean\_Target}(i))^2}{Nr-1}}}{\text{mean\_Target}(i)}, \quad (13)$$

respectively. A small value of CV indicates that a distribution is narrow and the model provides a robust property to simultaneous changes in all kinetic parameters.

### Link of reactions to robustness

Finally MAR combines the sensitivity distribution  $S(i,k)$  ( $k=1,2,\dots,N$ ) and  $CV\_Target(i)$  with respect to  $\mathbf{p}(i)$  to explore the

critical reactions responsible for determining the target instance. For highly robust models with a small value of  $CV\_Target(i)$ , the parameters showing a high sensitivity are selected as critical parameters. Use of statistical methods investigates how the selected critical parameters feature highly robust models, thereby linking reactions to robustness.

### Biochemical model

Circadian rhythms are defined as endogenously generated 24-h variations in behaviour and physiology that allow organisms to adapt to the varying environmental demands of the solar cycle. In *Drosophila*, oscillating levels of period (*per*) and timeless (*tim*) constitute the fly's molecular clock. The circadian feedbacks are composed of two interlocked negative feedback loops: a *per-tim* loop, which is activated by *Drosophila* clock (*dclk*) and repressed by PER:TIM, and a *dclk-cyc* loop, which is repressed by dCLK:CYC and derepressed by PER:TIM [12]. A loop that connects *per-tim* to *dclk-cyc* represents a positive feedback loop. **Figure 2** shows a schematic diagram of the circadian clock, with the notation of the reactions described elsewhere [19]. PER and TIM form dimers and then inhibit transcription of their own genes. Expression of the *per* and *tim* genes is regulated by a pair of transcription factors, dCLK and CYC, whose activity is decreased by forming a complex with PER:TIM [20]. Phosphorylation and proteolysis of PER make a time-delayed negative feedback, generating oscillations. PER is phosphorylated by DBT (encoded by the double-time (*dbt*) gene), which is present at roughly constant levels during the rhythm [21]. Phosphorylated PER (PER-P) is readily susceptible to degradation, while PER-P is stabilized by dimerization with TIM. The dimerized PER-P:TIM is transported to the nucleus. Transcription of *dclk* and *cyc* is regulated by PER-P:TIM-mediated release of dCLK:CYC-dependent repression.

### Mathematical model

We have derived mathematical equations from the interlocked-feedback model developed by Ueda et al. [14]. In Ueda's model, while PER directly binds to TIM, it is not considered that phosphorylated PER binds to TIM. Thus, we present the modified mathematical model that includes the mechanism for DBT-phosphorylated PER binding to TIM (**Tables 1–3**). DBT does not appear in the model as it is assumed that the DBT concentration is constant. The modified model is described by 11 ordinary differential equations. The CADLIVE system is employed for numerical simulation (<http://www.cadlive.jp>) [17].

### Optimization for the mathematical model

The dynamic model of the circadian system has 36 kinetic parameters to search as shown in **Table 3**, where  $T[2]$ ,  $T[4]$ ,  $K[3]$ , and  $K[5]$  are set to zero to reduce the number of search parameters. Twelve parameters ( $S[1]$ ,  $A[1]$ ,  $A[2]$ ,  $R[1]$ ,  $R[2]$ ,  $V[1]$ ,  $V[2]$ ,  $D[1]$ ,  $P[1]$ ,  $T[1]$ ,  $K[1]$ ,  $K[2]$ ) that are related to feedback architectures were explored using the evolutionary search in order to generate stable oscillators for PER. These parameters involve the transcription, phosphorylation, transport, and degradation of PER and the binding between PER-P and TIM. The values of the remaining parameters are estimated or provided from the previous model [14].

Based on the two-step method we randomly vary the values of 12 parameters in logarithmic space to find parameter solutions that indicate some oscillations, where a basis parameter vector is determined that produces a circadian oscillator and the value of each basis parameter is  $10^{-2}$  to  $10^2$ -fold varied. Such solutions will

be employed as the initial population for the subsequent search by GAs. In the random search, an oscillation curve of PER is sampled after a simulation time of 300 h and the fitness function is provided by:

$$F_{random}(X_{min}, X_{max}) = \frac{X_{max} - X_{min}}{X_{max} + X_{min}}, \quad (14)$$

where  $X_{min}$  is the minimum value of an oscillation curve of PER and  $X_{max}$  is its maximum. At this stage, it is important to find the coarse solutions showing some oscillations because GAs cannot evolve non-oscillatory solutions into stable oscillators. The initial population with 100 individuals is created around the coarse solution that shows some oscillatory behaviors.

Next, to search for the solutions that show stable oscillations with a 24-h period and large amplitude, GAs are carried out in logarithmic space, where each coarse solution obtained by the random search is set to the basis parameters and the value of each basis parameter is  $10^{-1}$  to 10-fold varied. Twelve kinetic parameters are optimized where the generations and population are set to 100 and 100, respectively. The Unimodal Normal Distribution Crossover (UNDX) is employed as crossover [17], while mutation is not used. Twenty elites are selected for each generation. The cycle features provide the fitness function:

$$F_{GA}(T, X_{min}, X_{max}) = - \left| \frac{X_{max} - X_{min}}{X_{max} + X_{min}} - 0.9 \right| - 0.05 |X_{max} - X_{min} - 10|, \quad (15)$$

under  $|T - 24| < 1$ ,

where  $T$  is the period. **Equation 15** is an experience formula that generates the 24-h cycles whose amplitude is large enough for the average of the oscillatory curves. A target PER oscillator is that the ratio of the amplitude to the mean concentration of PER (the first term in **Equation 15** is 1.8 ( $0.9 \times 2$ ), the amplitude (the second term) is 10 nM, and the period is between 23 and 25 h.

### Robustness analysis for the biochemical model

The numerical sensitivity of period ( $S_T$ ) or amplitude ( $S_A$ ) to variations in a single kinetic parameter is calculated by:

$$S_T(i, k) = \frac{\Delta \ln Period}{\Delta \ln p(i, k)}, \quad S_A(i, k) = \frac{\Delta \ln Amplitude}{\Delta \ln p(i, k)}, \quad (16)$$

$(i = 1, 2, \dots, L_s) \quad (k = 1, 2, \dots, N)$

where  $L_s$  and  $N$  are set to 400 and 36, respectively. Each kinetic parameter was changed by 1.5-fold. In robustness analysis to multiple parameter perturbation, characterized by the CVs of period and amplitude, the period and amplitude are simulated 10,000 times by randomly varying all kinetic parameters within a range from 1/1.5-fold to 1.5-fold ( $\rho = 0.5$ ) for  $L_s$  models. The means and CVs:  $mean\_Period(i)$ ,  $CV\_Period(i)$ ,  $mean\_Amplitude(i)$ , and  $CV\_Amplitude(i)$ , are calculated for all the distributions ( $i = 1, 2, \dots, L_s$ ).

### Computation

Since a search parameter space is very large, the obgrid system (RIKEN, Yokohama) is used to greatly increase a calculation speed, where dozens of computers are available simultaneous-

ly [22]. Dynamic simulation programs are written in C. The Runge-Kutta method is employed.

### Statistical analysis

A hierarchical clustering method with the average distance is used to classify parameter solution vectors. The Matlab functions, linkage and dendrogram, are employed to determine the number of clusters. Other statistical analyses are also carried out by Matlab.

## SUPPORTING INFORMATION

### TextS1

Found at: doi:10.1371/journal.pone.0001103.s001 (0.08 MB PDF)

### TableS1

Found at: doi:10.1371/journal.pone.0001103.s002 (0.13 MB PDF)

### TableS2

Found at: doi:10.1371/journal.pone.0001103.s003 (0.11 MB PDF)

### FigureS1

Found at: doi:10.1371/journal.pone.0001103.s004 (0.11 MB PDF)

### FigureS2

Found at: doi:10.1371/journal.pone.0001103.s005 (0.09 MB PDF)

### FigureS3

Found at: doi:10.1371/journal.pone.0001103.s006 (0.10 MB PDF)

### FigureS4

Found at: doi:10.1371/journal.pone.0001103.s007 (0.12 MB PDF)

### FigureS5

Found at: doi:10.1371/journal.pone.0001103.s008 (0.14 MB PDF)

### FigureS6

Found at: doi:10.1371/journal.pone.0001103.s009 (0.11 MB PDF)

### FigureS7

Found at: doi:10.1371/journal.pone.0001103.s010 (0.10 MB PDF)

### FigureS8

Found at: doi:10.1371/journal.pone.0001103.s011 (0.10 MB PDF)

### FigureS9

Found at: doi:10.1371/journal.pone.0001103.s012 (0.12 MB PDF)

## ACKNOWLEDGMENTS

### Author Contributions

Conceived and designed the experiments: HK. Performed the experiments: HK TT FO. Analyzed the data: HK TT. Contributed reagents/materials/analysis tools: HK TT FO. Wrote the paper: HK.

## REFERENCES

1. Savageau MA (1971) Concepts relating the behavior of biochemical systems to their underlying molecular properties. *Arch Biochem Biophys* 145: 612–621.
2. Savageau MA (1971) Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* 229: 542–544.
3. Csete ME, Doyle JC (2002) Reverse engineering of biological complexity. *Science* 295: 1664–1669.
4. El-Samad H, Kurata H, Doyle JC, Gross CA, Khammash M (2005) Surviving heat shock: control strategies for robustness and performance. *Proc Natl Acad Sci U S A* 102: 2736–2741.
5. Kurata H, El-Samad H, Iwasaki R, Ohtake H, Doyle JC, et al. (2006) Module-based analysis of robustness tradeoffs in the heat shock response system. *PLoS Comput Biol* 2: e59.
6. Stelling J, Sauer U, Szallasi Z, Doyle FJ 3rd, Doyle J (2004) Robustness of cellular functions. *Cell* 118: 675–685.
7. Ma L, Iglesias PA (2002) Quantifying robustness of biochemical network models. *BMC Bioinformatics* 3: 38.
8. Bagheri N, Stelling J, Doyle FJ 3rd (2007) Quantitative performance metrics for robustness in circadian rhythms. *Bioinformatics* 23: 358–364.
9. Alves R, Savageau MA (2000) Extending the method of mathematically controlled comparison to include numerical comparisons. *Bioinformatics* 16: 786–798.
10. Stelling J, Gilles ED, Doyle FJ 3rd (2004) Robustness properties of circadian clock architectures. *Proc Natl Acad Sci U S A* 101: 13210–13215.
11. Leloup JC, Goldbeter A (2000) Modeling the molecular regulatory mechanism of circadian rhythms in *Drosophila*. *Bioessays* 22: 84–93.
12. Glossop NR, Lyons LC, Hardin PE (1999) Interlocked feedback loops within the *Drosophila* circadian oscillator. *Science* 286: 766–768.
13. Cheng P, Yang Y, Liu Y (2001) Interlocked feedback loops contribute to the robustness of the *Neurospora* circadian clock. *Proc Natl Acad Sci U S A* 98: 7408–7413.
14. Ueda HR, Hagiwara M, Kitano H (2001) Robust oscillations within the interlocked feedback model of *Drosophila* circadian rhythm. *J Theor Biol* 210: 401–406.
15. Ono I, Kobayashi S (1997) A Real-coded Genetic Algorithm for Function Optimization Using Unimodal Normal Distribution Crossover. *Proc of 7th Int Conf on Genetic Algorithms*. pp 246–253.
16. Eshelman LJ, Schaffer JD (1993) Real-coded genetic algorithms and interval schemata. In: Whitley LD, ed. *Foundations of Genetic Algorithms*. San Francisco: Morgan Kaufmann. pp 187–202.
17. Kurata H, Masaki K, Sumida Y, Iwasaki R (2005) CADLIVE Dynamic Simulator: Direct Link of Biochemical Networks to Dynamic Models. *Genome Res* 15: 590–600.
18. Smolen P, Baxter DA, Byrne JH (2001) Modeling circadian oscillations with interlocking positive and negative feedback loops. *J Neurosci* 21: 6644–6656.
19. Kurata H, Matoba N, Shimizu N (2003) CADLIVE for constructing a large-scale biochemical network based on a simulation-directed notation and its application to yeast cell cycle. *Nucleic Acids Res* 31: 4071–4084.
20. Stanewsky R (2003) Genetic analysis of the circadian system in *Drosophila melanogaster* and mammals. *J Neurobiol* 54: 111–147.
21. Suri V, Hall JC, Rosbash M (2000) Two novel doubletime mutants alter circadian properties and eliminate the delay between RNA and protein in *Drosophila*. *J Neurosci* 20: 7547–7555.
22. Kimura S, Kawasaki T, Hatakeyama M, Naka T, Konishi F, et al. (2004) OBIYagns: a grid-based biochemical simulator with a parameter estimator. *Bioinformatics* 20: 1646–1648.

© 2007 Kurata et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (the “License”), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.