# SYSTEM IDENTIFICATION AND ROBUSTNESS ANALYSIS OF THE CIRCADIAN REGULATORY NETWORK VIA MICROARRAY DATA IN ARABIDOPSIS THALIANA [*]

C. W. LI

*Systems Biology Group, Automatic Control and signal Laboratory, Hsinchu, 300, Taiwan*
*Email:d925908@oz.nthu.edu.tw*


W. C. CHANG

*Systems Biology Group, Automatic Control and signal Laboratory, Hsinchu, 300, Taiwan*
*Email:cwli@moti.ee.nthu.edu.tw*


B. S. CHEN [†]

*Department of Electrical Engineering, National Tsing Hua University, Hsinchu, 300, Taiwan*
*Email:bschen@ee.nthu.edu.tw*

The circadian regulatory network is one of the main topics of plant investigations. The intracellular interactions among genes in response to the environmental stimuli of light are related to the foundation of functional genomics in plant. However, the sensitivity analysis of the circadian system has not analyzed by perturbed stochastic dynamic model via microarray data in plant. In this study, the circadian network is constructed for *Arabidopsis thaliana* using a stochastic dynamic model with sigmoid interaction, activation delay, and regulation of input light taken into consideration. The describing function method in nonlinear control theory about nonlinear limit cycle (oscillation) is employed to interpret the oscillations of the circadian regulatory networks from the viewpoint that nonlinear network will continue to oscillate if its feedback loop gain is equal to 1 to support the oscillation of circadian network. Based on the dynamic model via microarray data, the system sensitivity analysis is performed to assess the robustness of circadian regulatory network via biological perturbations. We found that the circadian network is more sensitive to the perturbation of the *trans*-expression threshold, is more sensitive to the activation level of steady state, rather than the *trans*-sensitivity rate.

## 1 Introduction

Biological phenomena at different organismic levels have implicitly revealed some sophisticated systematic architectures of cellular and physiological activities. These architectures were built upon the biochemical processes before the emergence of proteome and transcriptome [1,2]; and most biological phenomena such as metabolism, stress response [3], and cell cycle are directly or indirectly influenced by genes and have been well studied on the molecular basis. Thus, the identification of a signal transduction pathway could be traced back to the genetic regulatory level. The rapid advances of

genome sequencing and DNA microarray technology make possible the quantitative analysis of signaling regulatory network besides the qualitative analysis [4].

In this study, The ARX dynamic system approach is applied to the circadian regulatory pathway of *Arabidopsis thaliana* with microarray data sets publicly available on the net [5]. According to the synchronously dynamic evolution of microarray data, we have successively identified the core signaling transduction from light receptors of phytochromes [6] and crytochromes [7] to the endogenous biological clock [8], which is coupled to control the correlatively physiological activity with paces on a daily basis. With the dynamic system approach, not only the regulatory abilities, but also the oscillatory frequency and the delays of regulatory activity were specified. Moreover, we design several simulation assays with the biological senses to mimic the biological experiments.

## 2 Dynamic System Description of Circadian Regulatory Model

We can consider any gene expression profile as a system response or output stimulated by some inputs from other gene expressions and environmental stimuli. According to this description, let $X_i(k)$ denote the expression profile of the i-th gene at time point k. Then the following general form of ARX difference equation is proposed to model the expression level of the $i$ -th gene as the synthesis of $n$ upstream genetic signals $\widetilde{X_i}$, $i = 1, 2, \cdots, n$ and an external input signal $u$ under their $\tau$ delays, (see figure 1)

$$
\begin{aligned}
X_i(k) = & d_{1,i1}\widetilde{X_1}(k-\tau_i) + d_{1,i2}\widetilde{X_2}(k-\tau_i) + \cdots + d_{1,ii}X_i(k-\tau_i) + \cdots + d_{1,in}\widetilde{X_n}(k-\tau_i) + \\
& d_{2,i1}\widetilde{X_1}(k-2\tau_i) + d_{2,i2}\widetilde{X_2}(k-2\tau_i) + \cdots + d_{2,ii}X_i(k-2\tau_i) + \cdots + d_{2,in}\widetilde{X_n}(k-2\tau_i) + \\
& \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
& d_{q,i1}\widetilde{X_1}(k-Q\tau_i) + d_{q,i2}\widetilde{X_2}(k-Q\tau_i) + \cdots + d_{q,ii}X_i(k-Q\tau_i) + \cdots + d_{q,in}\widetilde{X_n}(k-Q\tau_i) + \\
& b_i \cdot u(k-\tau_{iu}) + \varepsilon_i(k) \quad, i = 1, 2, \cdots, n
\end{aligned}
\tag{1}
$$

where $\tilde{X}_j(k-q\tau_i)$, j=1,2,$\cdots$,n; $q = 1, 2, \cdots, Q$ is the upstream genetic signal transformed by $X_j(k)$ with the q-th order of $\tau_i$ delay and through a sigmoid activation function to denote the binding of transcription factor $X_j(k)$ on gene i, and the genetic kinetic parameter $d_{q,ij}$ denotes the regulation abilities of transcription factor $\tilde{x}_j(k)$ on gene i. Meanwhile, $u(k-\tau_{iu})$, which denotes the external input light with a delay $\tau_{iu}$ affecting $X_i(k)$, correlates with the output genetic expression $X_i(k)$ with the input kinetic parameters $b_i$. $\varepsilon_i(k)$ is the stochastic noise of current microarray data or the residue of the model. Here $\tau_i$ and $\tau_{iu}$, which are essential to the activation-time estimation, should be determined previously and will be discussed later. The ARX model (AutoRegressive with eXternal input), which admit a reformation to the linear regression model, is the special case of the ARMAX model (autoregressive moving average with exogenous input). Moreover, an oscillation will exist in circadian regulatory network by the feedbacks through other genes if these feedbacks are limited by sigmoid functions to avoid their unstable propagations, which will be discussed by describing function method [9] in the sequel.
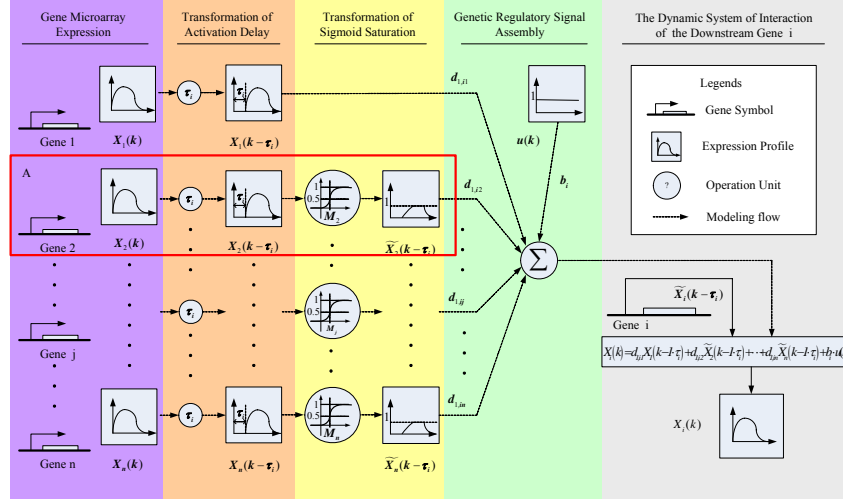
Figure 1.   Illustration of the dynamic system scheme using the ARX(1) model.   Block A represents the transformation of the genetic regulatory signal, $\tilde{X}_j(k - q\tau_i)$, for j=2 and q=1.

For the limited influence expression of $\tilde{X}_j(k - q\tau_i)$ (see Block A in Figure 1), the sigmoid function is chosen to express the nonlinear 'on' and 'off' activities of physically genetic interactions with parameters $\theta_j = \{\gamma, M_j, \tau_i\}$ as follows,

$$\widetilde{X_j}\left(k - q \cdot \tau_i\right) = \frac{1}{1 + e^{-\gamma\left(X_j\left(k - q \cdot \tau_i\right) - M_j\right)}}$$

(2)

where $\gamma$ is the trans-sensitivity rate, and $M_j$ is the trans-expression threshold derived from the mean of the j-tth gene's profile. $\gamma$ could determine the transition time of activation between the states of 'off' and 'on' from $X_j$ to $X_i$, for which a larger $\gamma$ is with a less transition time, to mimic the transient state of the genetic interaction on the *trans* level. $M_j$ can determine the threshold of the half activation level of $X_j$ to $X_i$, for which a larger $M_j$ is with a less activating ability, to mimic the steady state of the genetic interaction on the trans level. For the biological reason of small activation delay on mRNA level and less modeling complexity, we can reduce the order of the ARX model to no more than 2, Q=1 (i.e. ARX(1)) or Q=2 (i.e. ARX(2)) in Eq. (1).  We will determine an adequate order for our interesting system later.   And now we take the second order ARX model for illustration as follows,

$$\begin{aligned}
X_i\left(k\right) = &d_{1,i1}\widetilde{X_1}\left(k - \tau_i\right) + d_{1,i2}\widetilde{X_2}\left(k - \tau_i\right) + \cdots + d_{1,ii}X_i\left(k - \tau_i\right) + \cdots + d_{1,in}\widetilde{X_n}\left(k - \tau_i\right) + \\
&d_{2,i1}\widetilde{X_1}\left(k - 2\tau_i\right) + d_{2,i2}\widetilde{X_2}\left(k - 2\tau_i\right) + \cdots + d_{2,ii}X_i\left(k - 2\tau_i\right) + \cdots + d_{2,in}\widetilde{X_n}\left(k - 2\tau_i\right) + \\
&b_i \cdot u(k) + \varepsilon_i\left(k\right) \quad, i = 1, 2, \cdots, n
\end{aligned}$$

(3)

Consequently, the vector difference form underlined in this equation is applied to $m$ time points in order.

$$\underline{X_i} = d_{1,i1}\widetilde{X_{1,\tau_i}} + \cdots + d_{1,ii}\underline{X_{i,\tau_i}} + \cdots + d_{1,in}\widetilde{X_{n,\tau_i}} +$$
$$d_{2,i1}\widetilde{X_{1,2\tau_i}} + \cdots + d_{2,ii}\underline{X_{i,2\tau_i}} + \cdots + d_{2,in}\widetilde{X_{n,2\tau_i}} + b_i \cdot \underline{u} + \underline{\varepsilon_i}$$

(4)

where $\underline{x_i} = \begin{bmatrix} X_i(k_1) \\ X_i(k_2) \\ \vdots \\ X_i(k_m) \end{bmatrix}, \underline{X}_{n,q} = \begin{bmatrix} X_n(k_1 - q) \\ X_n(k_2 - q) \\ \vdots \\ X_n(k_m - q) \end{bmatrix}, \tilde{\underline{X}}_{n,q} = \begin{bmatrix} \tilde{X}_n(k_1 - q) \\ \tilde{X}_n(k_2 - q) \\ \vdots \\ \tilde{X}_n(k_m - q) \end{bmatrix}, \underline{u}_{\tau_{iu}} = \begin{bmatrix} u(k_1) \\ u(k_2) \\ \vdots \\ u(k_m) \end{bmatrix}, \underline{\varepsilon_i} = \begin{bmatrix} \varepsilon_i(k_1) \\ \varepsilon_i(k_2) \\ \vdots \\ \varepsilon_i(k_m) \end{bmatrix}$, and m denotes the number of time

points. $\tau_i$ is the specific activation delay.

In the next step, to estimate the kinetic parameters $d_{q,in}$, $q = 1,2$ ; $n = 1,2,\cdots$ , and $b_i$ , the formula Eq. (4) should be translated into the difference matrix equation as follows,

$$Y_i = A_i\Omega_i + E_i$$

(5)

where $Y_i = \underline{X_i}$ , $\Omega_i = \begin{bmatrix} d_{1,i1} & \cdots & d_{1,in} & d_{2,i1} & \cdots & d_{2,in} & b_i \end{bmatrix}^T$ , and $E_i = \underline{\varepsilon_i}$ are in vector forms, while $A = \begin{bmatrix} \tilde{\underline{X}}_{1,\tau_i} & \cdots & \tilde{\underline{X}}_{n,\tau_i} & \tilde{\underline{X}}_{1,2\tau_i} & \cdots & \tilde{\underline{X}}_{n,2\tau_i} & \underline{u}_{\tau_{iu}} \end{bmatrix}$ is a matrix.

We assume that each element in the stochastic noise vector, $\varepsilon_i(k_l)$ , $i = \{1,\cdots,m\}$ ,is an independent random variable with a normal distribution with zero mean and variance $\sigma^2$ , which is unknown and needs to be estimated. Thus, we will estimate the parameter $\hat{\Omega}_i$ using the maximum likelihood method.

The maximum likelihood estimate of $\sigma^2$ is the estimate of noise covariance. Substituting Eq. (8) into Eq. (7) yields,

$$L(\Omega_i, \sigma^2) = -\frac{m}{2}\ln\left[2\pi\sigma^2\right] - \frac{m}{2}$$

(6)

where $\sigma^2 = \frac{1}{m}\sum_{l=1}^{m}\left[Y_i - A_i\Omega_i\right]^T\left[Y_i - A_i\Omega_i\right]$

Therefore, we can find the maximum likelihood estimation of $\Omega_i$ by minimizing the value of $\sigma^2$ . From Eq. (6) best choice of parameter vector $\Omega_i$ to minimize $\sigma^2$ using the least-squares method is obtained as follows [10],

$$\widehat{\Omega_i} = \left(A_i^T A_i\right)^{-1} A_i^T Y_i$$

(7)

After the parameter estimation in Eq. (7), substituting $\Omega_i$ in Eq. (7) into stochastic model in Eq. (3) lead to the estimated circadian regulatory network equations.

## 3. Assay of the Model

### 3.1. Assay of ARX System Model

The assays of the ARX system model are divided into four categories. The first is the confirmation of the oscillation frequency of circadian regulatory network by the oscillatory characteristics of the dynamic circadian regulatory model; the second is the input stimulus changes; the third is the trans disturbance; and the last is about the cis perturbation. For each pair of gene expressions from both the biological assay and the simulation, we calculated the Pearson correlation coefficient between the genes' mRNA

expression profiles of $X_i(k)$ *in vivo* and $\widehat{X}_i(k)$ *in silico* at all time points $k = k_1, k_2, \cdots, k_m$ as follows.

$$r\left(x_i, \widehat{x}_i\right) = \frac{\sum_{l=1}^{m} X_i(k_l)\widehat{X}_i(k_l) - \left(\frac{\sum_{l=1}^{m} X_i(k_l)\sum_{l=1}^{m}\widehat{X}_i(k_l)}{m}\right)}{\sqrt{\left[\sum_{l=1}^{m} X_i^2(k_l) - \frac{\left[\sum_{l=1}^{m} X_i(k_l)\right]^2}{m}\right]\left[\sum_{l=1}^{m}\widehat{X}_i^2(k_l) - \frac{\left[\sum_{l=1}^{m}\widehat{X}_i(k_l)\right]^2}{m}\right]}}$$

(8)

To measure the period of the time-course expression profile, the power spectrum, which has different magnitudes in different frequencies (the reciprocal of periods), is employed to detect which frequency has the largest magnitude. First, we should take the Discrete Fourier Transform of $X_i(k)$ for $k = k_1, k_2, \cdots, k_m$ as follows,

$$X_i(w) = \sum_{l=1}^{m} X_i(k_l)e^{-jwk}$$

(9)

where w is the radian frequency.
Then we can detect the frequency with the maximum magnitude,

$$\omega_i = \arg\max_\omega |X_i(\omega)| = \frac{2\pi}{T_i}$$

(10)

where $T_i$ is the period of $X_i(k)$ and can be determined from the reciprocal of the detected frequency $\omega_i$. Furthermore, the measure of mean expression of $X_i(k)$ is important for distinguishing the deviation of expression profile under different assays as follows,

$$M_i = \frac{1}{m}\sum_{l=1}^{m} X_i(k_l)$$

(11)

### 3.1.1. Determination of system order

In this study, the formulated ARX model should be first assigned with a proper modeling order and an activation delay to analyze the experimental expression data of microarray. According to Eq. (1), we compared the first-order (Q=1) ARX model (i.e. ARX(1)) and the second-order (Q=2) ARX model (i.e. ARX(2)) with different activation delays $\tau$ as shown in Fig. 2a. We exploited the mean similarity between the raw expression and the simulation of all 16 system genes, which is measured by Pearson correlations, to evaluate the performance of the network model. Owing to the least difference at 0.5-hr delay between ARX(1) and ARX(2), we would prefer the more flexible ARX(1) model with a 0.5-hr activation delay as the system model for the circadian regulatory network. Consequently, the simulation expressions of the derived circadian network model are shown in Fig. 2b, $\tilde{x}_j(k - q\tau_i)$, for j=2 and q=1. The detection of the static structural characteristics will help reconstruct their hidden significance of cis connectivity as in the signaling transduction network of Fig. 3.
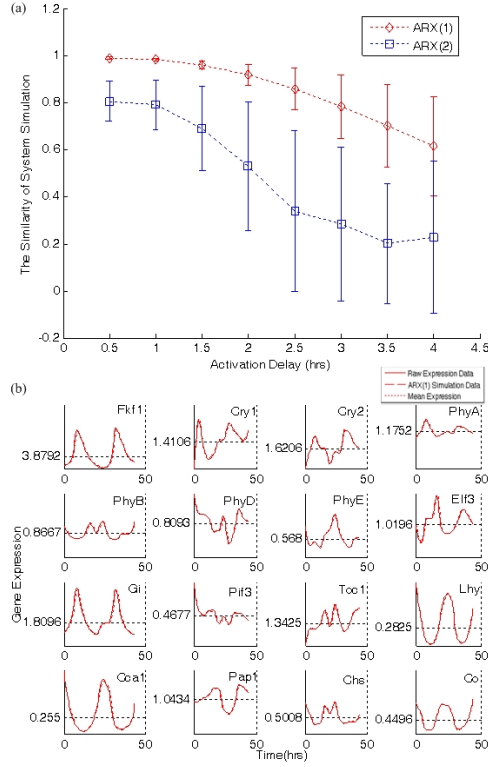
Figure 2. ARX system modeling with determination of system modeling order and activation delay. (a) The average similarity (measured by Pearson correlation) of all system genes under different activation delays. (b) The dynamic data fitting of 16 genes in the circadian network with ARX(1) model and 0.5-hr activation delay.
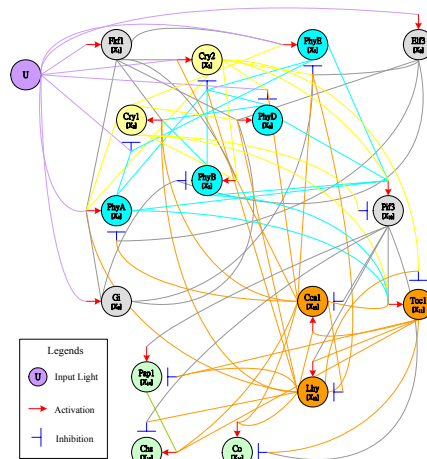


Figure 3. Signaling transduction network of system genes and input light in the circadian network of *Arabidopsis*. The colored circles indicate the system genes with their names and notations of $X_1 \sim X_{16}$.

## 3.2. Sensitivity Analysis of Circadian System

The sensitivity measure of the circadian system for the analysis of robustness can also be derived from the system model. For illustration, we would rearrange Eq. (3) into the following difference matrix equation,

$$\underline{Y}(k) = D_1 \underline{\tilde{Y}}(k-\tau) + D_2 \underline{Y}(k-\tau) + Bu(k) \tag{12}$$

where $D_1 = \begin{bmatrix} 0 & d_{1,12} & \cdots & \cdots & d_{1,1n} \\ d_{1,21} & 0 & d_{1,23} & \cdots & \vdots \\ \vdots & d_{1,32} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & d_{1,(n-1)n} \\ d_{1,n1} & \cdots & \cdots & d_{1,n(n-1)} & 0 \end{bmatrix}, D_2 = \begin{bmatrix} d_{2,11} & 0 & \cdots & \cdots & 0 \\ 0 & d_{2,22} & 0 & \cdots & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & d_{2,nn} \end{bmatrix}$ $\underline{Y}(k) = \begin{bmatrix} X_1(k) \\ X_2(k) \\ \vdots \\ X_n(k) \end{bmatrix}, \underline{Y}(k-\tau) = \begin{bmatrix} X_1(k-\tau_1) \\ X_2(k-\tau_2) \\ \vdots \\ X_n(k-\tau_n) \end{bmatrix}$ $\underline{\tilde{Y}}(k-\tau) = \begin{bmatrix} \tilde{X}_1(k-\tau_1) \\ \tilde{X}_2(k-\tau_2) \\ \vdots \\ \tilde{X}_n(k-\tau_n) \end{bmatrix}$

$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \underline{U} = \begin{bmatrix} u(k) \\ u(k) \\ \vdots \\ u(k) \end{bmatrix}$ and n is the number of genes.

### 3.2.1. Circadian clock frequency assay

While we obtain the oscillation frequencies $w_i$ of circadian network by the intersection in Eq. (10), we will compare with the oscillation frequencies calculated by Eq. (9) and (10) to validate the accuracy of the proposed dynamic model in the sequel.

A dynamic system with saturation (or sigmoid function) nonlinear feedback will lead to oscillation (limit cycle) [9]. This oscillation phenomenon can be interpreted by the theory of the describing function, which will be used to describe the circadian regulatory network of *Arabidopsis thaliana*. According to Eq. (12), we get

$$\underline{Y}(k) = (I - z^{-\tau}D_2)^{-1}D_1\underline{\tilde{Y}}(k-\tau) + (I - z^{-\tau}D_2)^{-1}Bu(k) \tag{13}$$

where $z^{-\tau} = \begin{bmatrix} z^{-\tau_1} & 0 & \cdots & 0 \\ 0 & z^{-\tau_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & z^{-\tau_n} \end{bmatrix}$, and $z^{-\tau_i}$ denotes delay operator of $\tau_i$.

If the oscillation (limit cycle) occurs in circadian network, then the sigmoid function $\tilde{Y}(k)$ in Eq. (2) can be approximated by the describing function N(A) as [9]

$$\underline{\tilde{Y}}(k) = \begin{bmatrix} \tilde{Y}_1(k) \\ \tilde{Y}_2(k) \\ \vdots \\ \tilde{Y}_n(k) \end{bmatrix} \approx N(A)\underline{Y}(k) \tag{14}$$

where the describing function matrix $N(A) = \begin{bmatrix} z^{-\tau_1} & 0 & \cdots & 0 \\ 0 & z^{-\tau_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & z^{-\tau_n} \end{bmatrix}$, and $N_i(A_i)$ denotes the describing

function of the i-th gene of oscillation and $A_i$ denotes the amplitude of oscillation of the i-th gene of gene j is free of oscillation, then the corresponding $N_j(A_j) = 0$. From Eq. (13) and (14), we can approximate the circadian network as

$$\underline{Y}(k) = (I - z^{-\tau}D_2)^{-1}D_1N(A)z^{-\tau}\underline{Y}(k) + (I - z^{-\tau}D_2)^{-1}Bu(k) \tag{15}.$$

There are two rhythms, one is circadian rhythm and another is diurnal rhythm. The first term with gain equal to 1 on the right hand side of Eq. (16) is the response for circadian rhythm; and the second term for diurnal rhythm, which is controlled by diurnal cycling of light and dark u(k) and some photoreceptor genes are of this case. Since the oscillation exists in the circadian network, by control theory, the closed loop gain should be lossless in order to support the oscillation, i.e.

$$D_1 N(A) = z^\tau - D_2 \tag{16}$$

At frequency domain, we can get

$$e^{jw_i \tau_i} - d_{2,ii} = \sum_{j \neq i}^{n} d_{1,ij} N_j(A_j) \tag{17}$$

where $e^{jw\tau} = \begin{bmatrix} e^{-jw_1\tau_1} & 0 & \cdots & 0 \\ 0 & e^{-jw_2\tau_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{-jw_N\tau_N} \end{bmatrix}$

For example, for gene PhyE, $e^{jw_7\tau_7} - d_{2,77} = -0.1579 - 0.1226i$ and $\sum_{j \neq 7}^{N} d_{1,7j} N_j(A_j) = -0.1339 - 0.1253i$ ,which matches the oscillation condition in Eq. (17). For gene Lhy, $e^{jw_{12}\tau_{12}} - d_{2,1212} = -0.2181 - 0.1045i$ and, $\sum_{j \neq 12}^{N} d_{1,12j} N_j(A_j) = -0.2144 - 0.0589i$ which roughly matches the oscillation condition of describing function in the nonlinear circadian system. By describing analysis of nonlinear oscillation [9], the intersection of $e^{jw_i\tau_i} - d_{2,ii}$ and $\sum_{i \neq j}^{N} d_{ij} N(A_i)$ in Eq. (17) implies the occurrence of oscillation and the $A_i$ and $w_i$ at the interaction point are the oscillation amplitude and oscillation frequency.

### 3.2.2. Trans-perturbation assay

As in the description of Eq. (2), $\gamma$ is the trans-sensitivity rate which is related to the transition time of trans-activaton and $M_j$ is the trans-expression threshold that determines the saturating transformation level of expression. We also induce the corresponding sensitivity in the following,

$$\frac{\Delta Y}{\Delta F} \simeq D_l \frac{\Delta \tilde{Y}}{\Delta F} \tag{18}$$

where $F = \{\gamma, M\}$ and $\frac{\Delta \tilde{Y}}{\Delta \gamma} = \frac{-Ye^{-\gamma(Y-M)}}{(1 + e^{-\gamma(Y-M)})^2}$ . Hence we could discuss the sensitivity on the trans level like the input sensitivity.

### 3.2.2.1. Trans-sensitivity rate $\gamma$ simulation of gene

In a similar way as in input perturbation, we changed $\gamma$ from 100% to 0% (-100%) and 200% (+100%) of system genes in pathway to compare with their sensitivities to $\gamma$, as shown in Table 2A. We also average the three measure indexes of each gene, which are shown in Fig. 4.1
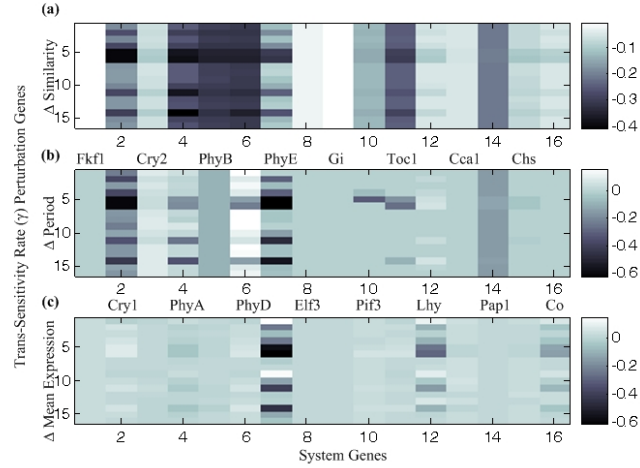
Figure 4.1. Deviation representations of the system genes under the perturbation of trans-sensitivity rate $\gamma$. The perturbation is performed in the vertical axis and the responses of 16 system genes are shown in the horizontal axis, and the colored bars of degree are on the right-hand side of each inset. (a) $\Delta$ Similarity (measured by the Pearson correlation), (b) $\Delta$ Period, and (c) $\Delta$ Mean expression.

### 3.2.2. Trans-expression threshold $M_j$ simulation of gene

We varied $M_j$ to 100% lower (-100%) and higher (+100%) than the original mean expression of the j-th gene respectively and compared with their sensitivities of $M_j$, which are shown in Table 2B; and their average measure indexes are shown in Fig. 4.2.
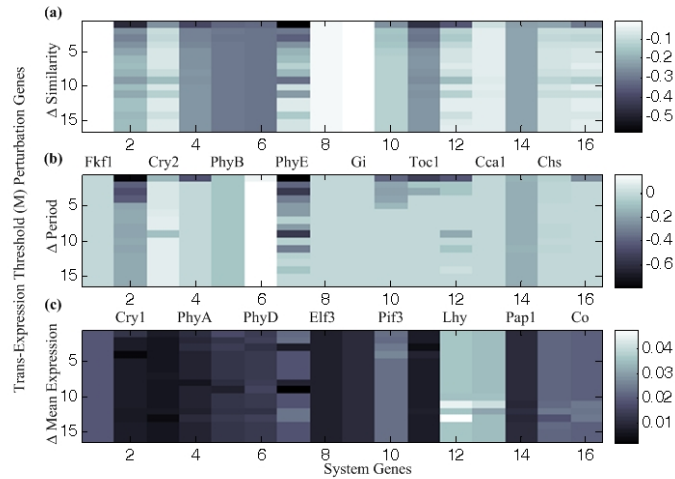


Figure 4.2. Deviation representations of the system genes under the perturbation of trans-expression threshold $M$. The perturbation is performed in the vertical axis and the responses of 16 system genes are shown in the horizontal axis, and the colored bars of degree are on the right-hand side of each inset. (a) $\Delta$ Similarity (measured by the Pearson correlation), (b) $\Delta$ Period, and (c) $\Delta$ Mean expression.

Table 2. The sensitivities of the system genes in the circadian regulatory network under different perturbations due to the *input* light, *trans* level, and *cis* level. The bold values represent significant sensitivities (less robustness). In general, the fact that sensitivities are not large implies that the circadian regulatory network is robust enough.

| Perturbation | Type of Sensitivity | Specification | The Sensitivities of the system genes in the circadian regulatory network | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Flfl | Cry1 | Cry2 | PhyA | PhyB | PhyD | PhyE | EfD | Gi | PrD | Toc1 | Lhy | Ccsl | Pap1 | Cha | Co |
| Trans level | A. Trans-sensitivity rate (γ) | +0 | 0.0112 | **0.017** | 0.0119 | 0.008 | 0.0104 | **0.0129** | -0.008 | 0.0112 | 0.0112 | 0.0113 | **0.013** | 0.0059 | 0.0106 | 0.0116 | 0.01 | 0.0069 |
| | | -100% | 0.0112 | **0.017** | 0.011 | 0.0096 | 0.0112 | **0.0128** | -0.008 | 0.0112 | 0.0112 | 0.0115 | **0.013** | 0.0055 | 0.011 | 0.0115 | 0.0096 | 0.0062 |
| | | +100% | 0.0112 | **0.017** | 0.0121 | 0.0073 | 0.0102 | **0.013** | -0.008 | 0.0112 | 0.0112 | 0.0114 | **0.013** | 0.0066 | 0.0106 | 0.0117 | 0.0105 | 0.0076 |
| | | Mean | 0.0112 | **0.017** | 0.0117 | 0.0083 | 0.0106 | **0.0129** | -0.008 | 0.0112 | 0.0112 | 0.0114 | **0.013** | 0.006 | 0.0107 | 0.0116 | 0.01 | 0.0069 |
| | B. Trans-expression level (M) | +0 | 0.0112 | -0.068 | 0.0088 | 0.0474 | 0.026 | 0.00403 | **0.267** | 0.0112 | 0.0112 | 0.0236 | -0.031 | **0.122** | 0.0116 | 0.0074 | 0.0228 | **0.105** |
| | | -100% | 0.0112 | 0.01 | -0.007 | 0.06 | 0.0301 | -0.0003 | **0.272** | 0.0112 | 0.0112 | -0.008 | -0.031 | **0.11** | 0.0011 | 0.0104 | -0.002 | **0.095** |
| | | +100% | 0.0112 | -0.047 | -0.043 | 0.033 | 0.0211 | 0.00812 | **0.233** | 0.0112 | 0.0112 | 0.0156 | 0.0002 | **0.124** | -0.002 | 0.0106 | 0.0139 | **0.107** |
| | | Mean | 0.0112 | -0.035 | -0.014 | 0.0468 | 0.0257 | 0.00394 | **0.257** | 0.0112 | 0.0112 | 0.0103 | -0.021 | **0.119** | 0.0037 | 0.0095 | 0.0116 | **0.102** |

## 4. Results

In the perturbation of *trans*-sensitivity rate ($\gamma$), we will discuss whether the transition rate, which determines the transition time of one gene binding to or interacting with another one, affects the system gene's expression in this model system. It seems that the similarity (Fig. 4.1a) remains unchanged for most system genes except Cry1 [$X_2$], PhyA [$X_4$], PhyD [$X_6$], and PhyE [$X_7$] [11]. If we consider the periodic variation in Fig. 4.1b, Cry2's [$X_3$] period is lengthened about 10%, whilst that of Cry1 [$X_2$] and Pap1 [$X_{14}$] are shortened about 20%, respectively. The diversity and sensitivity of period due to perturbation of the transition time are evident as in Fig. 4.1c. The mean expressions of system genes are almost unaltered but PhyE [$X_7$] [10] is reduced. Because the largest difference in the mean sensitivity of each gene of Table 2A is about 0.025, we would conclude that the *trans*-sensitivity rate, which determines the transition time indicating the transient state of *trans* activation, has less influence on the circadian system.

In another perturbation of *trans*-expression threshold $M$, there are five genes of Cry1 [$X_2$], Cry2 [$X_3$], PhyD [$X_6$], Pif3 [$X_{10}$] and Toc1 [$X_{11}$] with perceptible variations, which have the same behavior in the measures of similarity and period (see Figs. 4.2a and 4.2b). Owing the largest difference in the mean sensitivity of each gene of Table 2B being close to 0.29, the circadian network is more sensitive to the perturbation of the *trans*-expression threshold $M$, is more sensitive to the activation level of steady state, rather than the *trans*-sensitivity rate $\gamma$.

## 5. Discussion

In our dynamic system approach applied to the circadian network using ARX, we not only can identify the regulatory abilities via ARX(1) with activation delays, but also indicate the regulatory strength from the input-light signal. The greatest importance of the proposed dynamic model is the convenience of the consequent system analysis, for example, sensitivity analysis, to gain more insight about the circadian regulatory network. There are some shortcomings in our study. First, although the time-course microarray data are available, its lower samplings will distort the real changes of gene expressions, especially for fast dynamic evolution. A more sampling experiment with respect to the intrinsic turnover rate is expected for a more precise analysis. Second, we formulate our ARX circadian network model using the biological knowledge of the correlations

between the circadian genes. In the circadian regulatory network, it is enough to reconstruct the system because of its simulation similarity approaching 0.99 in Fig. 2a.

In the near future, as the system modeling algorithms are further developed, we expect this dynamic system approach to have immense impact in elucidating the underlying molecular mechanisms of network in a variety of organisms besides the circadian network in Arabidopsis thaliana, especially after the maturation of the protein chips.

## Acknowledgments

## References

1. J. R. Kettman, J. R. Frey and I. Lefkovits. Proteome, transcriptome and genome: top down or bottom up analysis. *Biomol. Eng.*, 18:207-212, 2001.
2. J. Scheel, M.C. Von Brevern, A. Horlein, A. Fischer, A. Schneider and A. Bach. Yellow pages to the transcriptome. *Pharmacogenomics*, 3:791-807, 2002.
3. S. Motaki, K. Ayako, Y. S. Kazuko and S. Kazuo. Molecular response to drought, salinity and frost: common and different paths for plant protection. *Cur. Opn. Biotechnol.,* 14:194-199, 2003.
4. T. S. Hughes, B. Abou-Khalil, P. J. Lavin, T. Fakhoury, B. Blumenkopf and S. P. Donahue. Visual field defects after temporal lobe resection: a prospective quantitative analysis. *Neurology.*, 53:167-72, 1999.
5. S. L. Harmer, J. B. Hogenesch, M. Straume, H. S. Chang, B. Han, T. Zhu, X. Wang, J. A. Kreps and S. A. Kay. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science*, 290:2110-2113, 2000.
6. J. J. Casal. Phytochromes, cryptochromes, phototropin: photoreceptor interactions in plants. *Photochem. Photobiol.*, 71:1-11, 2000.
7. C. Fankhauser and D. Staiger. Photoreceptors in Arabidopsis thaliana: light perception, signal transduction and entrainment of the endogenous clock. *Planta.*, 216:1-16, 2002.
8. D. Alabadi, T. Oyama, M. J. Yanovsky, F. G. Harmon, P. Mas and S. A. Kay. Reciprocal regulation between TOC1 and LHY/CCA1 within the *Arabidopsis* circadian clock. *Science*, 293:880-883, 2001.
9. J. J. E. Slotine and W Li. Applied Nonlinear Control, Prentice Hall, Englewood Cliffs, NJ, 1991.
10. R. Johansson. *System Modeling and Identification*, PrenticeHall, Englewood Cliffs, NJ, 1993.
11. M. J. Aukerman, M. Hirschfeld, L. Wester, M. Weaver, T. Clack. R. M. Amasino and R. A. Sharrock. A deletion in the *PHYD* gene of the Arabidopsis Wassilewskija ecotype defines a role for phytochrome D in red/far-red light sensing. *Plant Cell.*, 9**:**1317-1326, 1997.