Published in final edited form as:
J Biol Dyn. 2011 January ; 5(1): 84–101. doi:10.1080/17513758.2010.491558.

A Dynamic Model for Functional Mapping of Biological Rhythms

Guifang Fu[*], **Jiangtao Luo**[†], **Arthur Berg**[†], **Zhong Wang**[†], **Jiahan Li**[*], **Kiranmoy Das**[*], **Runze Li**[*,†], and **Rongling Wu**[*,†,‡]

[*] Department of Statistics, Pennsylvania State University, University Park, PA 16802 USA

[†] Department of Public Health Sciences, Pennsylvania State College of Medicine, Hershey, PA 17033 USA

[‡] Center for Computational Biology, Beijing Forestry University, Beijing 100083, China

Abstract

Functional mapping is a statistical method for mapping quantitative trait loci (QTLs) that regulate the dynamic pattern of a biological trait. This method integrates mathematical aspects of biological complexity into a mixture model for genetic mapping and tests the genetic effects of QTLs by comparing genotype-specific curve parameters. As a way of quantitatively specifying the dynamic behavior of a system, differential equations have proven to be powerful for modeling and unraveling the biochemical, molecular, and cellular mechanisms of a biological process, such as biological rhythms. The equipment of functional mapping with biologically meaningful differential equations provides new insights into the genetic control of any dynamic processes. We formulate a new functional mapping framework for a dynamic biological rhythm by incorporating a group of ordinary differential equations (ODE). The Runge-Kutta fourth order algorithm was implemented to estimate the parameters that define the system of ODE. The new model will find its implications for understanding the interplay between gene interactions and developmental pathways in complex biological rhythms.

Keywords

Differential equation; Functional mapping; Quantitative trait loci; Biological rhythm

INTRODUCTION

Rhythmic phenomena that result from the planet's rotation around the sun represent one of the most striking manifestations of dynamic behavior in biological systems. Operating at all levels of organization, from molecular to ecosystem, rhythms equip organisms with a regulatory mechanism that allows them to adapt to environmental factors and evolve into a higher level of fitness (Webb 2003; Tauber et al. 2004). Also, physiological rhythms and oscillations in living cells have been thought to play a central role in regulating sleep-wake cycles, glucose, lipid, and drug metabolism, heart rate, stress and growth hormones, and immunity (Lowrey and Takahashi 2004; Turek et al. 2005; Gallego et al. 2006). Therefore, understanding the molecular and cellular mechanisms responsible for biological rhythms is crucial not only for unraveling the origin and dynamics of life, but also for designing efficient drug doses and schedules to prevent and control human diseases (Wijnen and Young 2006; Levi and Schibler 2007).

To whom all correspondence should be addressed: Rongling Wu, Department of Public Health Sciences, Pennsylvania State College of Medicine, Hershey, PA 17033 USA.

The past decades have witnessed a growing interest in developing quantitative models to estimate and test the pattern and inherent periodicity of rhythmic processes from a mechanistic perspective (Goldbeter 1996, 2002; Leloup et al. 2006). For example, differential equations of increasing complexity for the regulatory network producing circadian rhythms in the fly *Drosophila* predict the occurrence of sustained circadian oscillations of the limit cycle type (Leloup et al. 1999). When incorporating the effect of light, the models account for phase shifting of the rhythm by light pulses and for entrainment by light-dark cycles. The models also provide an explanation for the long-term suppression of circadian rhythms by a single pulse of light. Stochastic simulations were developed to test the robustness of circadian oscillations with respect to molecular noise (Gonze et al. 2002, 2003; Stelling et al. 2004a,b). The merits of mathematical and computational models are derived from their capacity to capture the most important information from chaotic systems and make a quantitative prediction and test of the behavior of the dynamic systems. By estimating key mathematical parameters that define biological rhythms, the origin, properties, and functions of the system can be well addressed.

It has been widely recognized that the normal function of biological rhythms is strongly correlated with the coordinated interactions between genes and their products at the molecular level. Several studies have identified several so-called clock genes and clock-controlled transcription factors through gene mutants in animal models (Takahashi 1993; Reppert and Weaver 2002). The expression of these clock genes and their interactive behavior in a rhythmic system may be different between individuals, although it is still not clear how much genetics contributes to these differences (Levi and Schibler 2007). With the advent of new technologies for genotyping single nucleotide polymorphisms (SNPs) that permit the rapid establishment of high-resolution haplotype maps (The International HapMap Consortium 2003, 2005, 2007), genome-wide quantitative trait loci (QTLs) responsible for individual differences in circadian biology can be identified and the detailed genetic mechanisms by which circadian behavior is originated and mediated can be well understood.

The dynamic characteristic of circadian rhythm will make it difficult to map its underlying QTLs. This difficulty arises from the fact that continuous time-dependent changes of the rhythm can only be observed at discrete time points. Part of this difficulty has now been solved by a dynamic mapping model, called *functional mapping*. This model was recently developed, aimed to map dynamic QTLs that control the pattern and rate of biological trajectories measured at a number of time points (Ma et al. 2002; Wu and Lin 2006). Functional mapping has emerged as a powerful statistical tool for detecting and mapping dynamic QTLs for stem growth and root growth in forest trees (Wu et al. 2003; Zhang et al. 2009), plant height in rice (Zhao et al. 2004a), tiller number increase in rice (Cui et al. 2006), biomass growth in soybeans (Li et al. 07), body mass growth in mice (Zhao et al. 2004b; Wu et al. 2005), and drug response (Lin et al. 2005). In this article, we will incorporate the principle of functional mapping to map genetic regulatory networks of biological rhythms with SNP markers. Specific purposes of this article are to integrate ordinary differential equations (ODE) into a mixture-based likelihood framework for functional mapping to study the genetic architecture and network of biological rhythms, explore the statistical properties of genotype-specific parameter estimates from observational dynamic and longitudinal data, and construct a quantitative platform for testing biologically meaningful hypotheses at the interplay among genetics, development, and oscillation. All these will make this research qualitatively beyond existing functional mapping in its mathematical and computational complexity as well as its potential application.

ODE FOR CIRCADIAN OSCILLATIONS

The molecular bases of circadian rhythms have been clarified during the past decade by experimental advances, first in *Drosophila* and *Neurospora*, and more recently in cyanobacteria, plants and mammals (Reppert and Weaver 2002; Dunlap 1999; Young and Kay 2001). All these studies conducted so far suggest that circadian rhythms originate from the negative feedback exerted by a protein on the expression of its gene (Hardin et al. 1990; Goldbeter 2002). In view of the large number of variables involved and of the complexity of feedback processes that generate oscillations, mathematical models are necessary to comprehend the transition from simple to complex oscillatory behavior and to delineate the conditions under which they arise (Goldbeter 2002). The first theoretical models for investigating the dynamic properties of circadian rhythms were borrowed from physics, such as the Van der Pol equations, derived for an electrical oscillator (Kronauer et al. 1982). Molecular models that incorporate the oscillatory mechanisms of circadian rhythms were initially proposed for circadian oscillations of the period (PER) protein and its mRNA in *Drosophila*. The PER protein behaves as a transcriptional regulator capable of influencing the expression of a variety of genes besides its own gene, *per*.

Based on the negative control exerted by the PER protein on the expression of *per*, a molecular model governed by a set of five ordinary differential equations was derived (Goldbeter 1995). In the model (Fig. 1), the per gene is first expressed in the nucleus and transcribed into per mRNA. The latter is transported into the cytosol, where it is translated into the PER protein, P_0. The PER protein undergoes multiple phosphorylation, from P_0 into P_1 and from P_1 into P_2. These modifications, catalyzed by a protein kinase, are reverted by a phosphatase. The fully phosphorylated form of the protein is marked up for degradation and transported into the nucleus in a reversible manner. The nuclear form of the protein (P_N) represses the transcription of the gene. In the model, the temporal variation of the concentrations of mRNA (R) and of the various forms of the regulatory protein – cytosolic (P_0, P_1, P_2) or nuclear (P_N) – is governed by the following system of kinetic equations

$$\begin{aligned}
\frac{dR}{dt} &= v_s \frac{K_I^n}{K_I^n + P_N^n} - v_m \frac{R}{K_m + R} \\
\frac{dP_0}{dt} &= k_s M - V_1 \frac{P_0}{K_1 + P_0} + V_2 \frac{P_1}{K_2 + P_1} \\
\frac{dP_1}{dt} &= V_1 \frac{P_0}{K_1 + P_0} - V_2 \frac{P_1}{K_2 + P_1} - V_3 \frac{P_1}{K_3 + P_1} + V_4 \frac{P_2}{K_4 + P_2} \\
\frac{dP_2}{dt} &= V_3 \frac{P_1}{K_3 + P_1} - V_4 \frac{P_2}{K_4 + P_2} - v_d \frac{P_2}{K_d + P_2} - k_1 P_2 + k_2 P_N \\
\frac{dP_N}{dt} &= k_1 P_2 - k_2 P_N .
\end{aligned} \tag{1}$$

The model assumes that per mRNA is synthesized in the nucleus and transfers to the cytosol, where it accumulates at a maximum rate v_s and is degraded by an enzyme at maximum rate v_m and Michael constant K_m. The rate of synthesis of the PER protein, proportional to R, is characterized by an apparent first-order rate constant k_s. Parameters V_1, …, V_4 and K_1, …, K_4 are the maximum rate and Michaelis constant of the kinase(s) and phosphatase(s) involved in the reversible phosphorylation of P_0 into P_1 and P_1 into P_2, respectively. The fully phosphorylated form (P_2) is degraded by an enzyme at maximum rate v_d and Michaelis constant K_d, and transported into the nucleus at a rate characterized by the apparent first-order rate constant k_1. Transport of the nuclear, bisphosphorylated form of PER (P_N) into the cytosol is characterized by the apparent first-order rate constant k_2. The negative feedback exerted by nuclear PER on per transcription is described by an equation of the Hill type, in which n denotes the degree of cooperativity, and K_I denotes the threshold constant for repression.

STRUCTURE OF FUNCTIONAL MAPPING

Genetic Design

Suppose there is a sample of N subjects drawn randomly from a natural population in Hardy-Weinberg equilibrium (HWE). In this sample, genome-wide SNP markers are genotyped, aimed to identify QTLs affecting biological rhythms. Consider a particular SNP marker, with two alleles M (with a probability of p) and m (with a probability of $1 - p$). Let M_1, M_2, and M_3 denote the observations of three marker genotypes MM, Mm, and mm, respectively. We assume that this marker is associated with a QTL with two alleles A (with a probability of q) and a (with a probability of $1 - q$) through the linkage disequilibrium of D. The marker and QTL form four haplotypes MA, Ma, mA, and ma, with the frequencies denoted as p_{11}, p_{10}, p_{01}, and p_{00}, respectively. It is shown that $p_{11} = pq + D$, $p_{10} = p(1 - q) - D$, $p_{01} = (1 - p)q - D$, $p_{00} = (1 - p)(1 - q) + D$ where $\max(-pq, -(1 - p)(1 - q)) \le D \le \min(p(1 - q), (1 - p)q)$. By testing the significance of D and estimating its value, we can determine the probabilities of each of the QTL genotypes, AA (denoted as 1), Aa (denoted as 2), and aa (denoted 3), given a particular marker genotype.

For each subject, the relative concentrations of mRNA (R) and four different forms of the regulatory protein, cytosolic (P_0, P_1, P_2) or nuclear (P_N), are measured at a series of time points during a daily light-dark cycle. Assuming that different subjects use different measurement schedules, we have $\mathbf{t}_i = (t_{i1}, \ldots, t_{iT_i})$ for subject i. Thus, five sets of serial measurements for mRNA and protein variables are expressed as

$$\mathbf{y}_{ki}=[y_{ki}(t_{i1}),\ldots,y_{ki}(t_{iT_i})],$$

where

$$k=\begin{cases}1 & \text{for the concentration of mRNA } R\\ 2 & \text{for the concentration of protein } P_0\\ 3 & \text{for the concentration of protein } P_1\\ 4 & \text{for the concentration of protein } P_2\\ 5 & \text{for the concentration of protein } P_N\end{cases}$$

We use $\mathbf{z}_i=(\{\mathbf{y}_{ki}\}_{k=1}^{5})$ to denote all the five variables.

Likelihood

For simplicity, we neglect covariate effects, although they can be incorporated into the model. The trait phenotypes of subject i measured at time $t_{i\tau}$ can be expressed by a five-variate linear model, i.e.,

$$y_{ki}(t_{i\tau})=\sum_{j=1}^{3}\xi_{ij}u_{kj}(t_{i\tau})+e_{ki}(t_{i\tau})+\varepsilon_{ki}(t_{i\tau}), \quad (2)$$

where ξ_{ij} is an indicator variable describing a possible QTL genotype j for subject i ($j = 1$ for AA, 2 for Aa, and 3 for aa) which is defined as 1 if a particular genotype is observed and 0 otherwise; $u_{kj}(t_{i\tau})$ is the genotypic mean of variable k for QTL genotype j at time $t_{i\tau}$; and

$e_{ki}(t_{i\tau})$ and $\varepsilon_{ki}(t_{i\tau})$ are the permanent and random error for variable k, respectively, which together follow a multivariate normal distribution with mean 0 and covariance matrix $\mathbf{\Sigma}_{ki}$.

The likelihood of N subjects can be formulated as

$$L(\mathbf{\Delta})=\prod_{i=1}^{N}\left[\sum_{j=0}^{2}\pi_{j|i}f_j(\mathbf{z}_i)\right], \tag{3}$$

where $\mathbf{\Delta}$ is the unknown vector specifying the likelihood, which will be defined below, $\pi_{j|i}$ is the conditional probability of QTL genotype j given the marker genotype of subject i, and $f_j(\mathbf{z}_i)$ is a multivariate normal distribution with mean vector

$$\mathbf{u}_{j|i}=\{\mathbf{u}_{kj|i}\}_{k=1}^{5}=\{u_{kj|i}(t_{i1}),\ldots,u_{kj|i}(t_{iT_i})\}_{k=1}^{5}, \tag{4}$$

and covariance matrix

$$\sum_i=\begin{pmatrix}\sum_{1i} & \sum_{12i} & \sum_{13i} & \sum_{14i} & \sum_{15i}\\ \sum_{21i} & \sum_{2i} & \sum_{23i} & \sum_{24i} & \sum_{25i}\\ \sum_{31i} & \sum_{3i} & \sum_{33i} & \sum_{34i} & \sum_{35i}\\ \sum_{41i} & \sum_{4i} & \sum_{43i} & \sum_{44i} & \sum_{45i}\\ \sum_{51i} & \sum_{5i} & \sum_{53i} & \sum_{54i} & \sum_{5i}\end{pmatrix} \text{ with } \sum_{kk'i}=\sum_{k'ki}. \tag{5}$$

The unknown $\mathbf{\Delta}$ contains three components: the mixture proportion ($\pi_{j|i}$) which is subject-dependent; the parameters ($\mathbf{\Psi}_u$) that model the mean vector (4); and the parameters ($\mathbf{\Psi}_v$) that model the covariance (5).

Modeling Mixture Proportions

In a natural population at HWE, the frequency of a joint marker and QTL diplotype can be expressed as the product of the frequencies of the two haplotypes derived from different parents that constitute the diplotype. For the genotypes which are homozygous at one or two loci, the diplotype frequency is the same as the genotype frequency (Table 1). The double heterozygote *AaBb* contains two possible diplotypes *AB*|*ab* and *Ab*|*aB*, where the haplotypes derived from maternal and paternal parents are separated by the vertical lines. Thus, the total frequency of genotype *AaBb* is the sum of the frequencies of these two diplotypes. Since unobservable QTL genotypes can be inferred from observed marker genotypes due to the marker-QTL association, we will derive the the conditional probability of a QTL genotype given a marker genotype using the joint genotype frequencies from Table 1. For subject i carrying a particular marker genotype, its conditional probability is denoted by $\pi_{j|i}$, which is the mixture proportion of the likelihood (3).

Modeling the Covariance Structure

To well focus on the integration of ODE into functional mapping, we will use a simple approach to model the covariance matrix. As shown by equation (1), the interrelationships among five different variables have been specified by ODE. For this reason, we can reasonably assume that the residual errors in equation (2) are independent among five variables, in which case a subject-specific covariance matrix is expressed as

$$\sum_i = \begin{pmatrix} \sum_{1i} & 0 & 0 & 0 & 0 \\ 0 & \sum_{2i} & 0 & 0 & 0 \\ 0 & 0 & \sum_{3i} & 0 & 0 \\ 0 & 0 & 0 & \sum_{4i} & 0 \\ 0 & 0 & 0 & 0 & \sum_{5i} \end{pmatrix},$$

where $\mathbf{\Sigma}_{ki}$ ($k = 1, 2, 3, 4, 5$) is a ($T_i \times T_i$) matrix.

Equation (2) contain two types of errors, permanent and random. It is possible that these two types of errors make a different contribution to the overall variances and covariances. Let η_k and $1 - \eta_k$ denote the relative contributions of the permanent and random errors for variance k, respectively. Thus, a subject-specific residual covariance matrix for variable k can be partitioned into the two parts as follows:

$$\sum_{ki} = \eta_k \sum_{ki}^{e} + (1 - \eta_k) \sum_{ki}^{\varepsilon}. \quad (6)$$

Since random errors arise stochastically at individual time points, they can be assumed to be independent among different times. Further, the random variances are assumed to be time-invariant, denoted by σ_k^2 for variable k. The covariance matrix due to the permanent errors may show an autocorrelation (Zimmerman and Nunez-Anton 2001). The permanent covariance ($\sum_{ki}^{e}$) structure can be modeled by many structural approaches including parametric stationary autoregressive (AR) model (Diggle et al. 2002), parametric non-stationary structured antedependence (SAD) model (Gabriel 1962; Zimmerman and Nunez-Anton 2001), parametric Ornstein-Uhlenbeck (OU) process (Gillespie 1996), nonparametric model (Yap et al. 2008), and semiparametric model (Fan et al. 2007).

For each of these approaches, covariance matrix $\sum_{ki}^{e}$ is structured by a limited number of parameters. The closed forms for the determinant and inverse of $\sum_{ki}^{e}$ if it is modeled by a parametric approach, such as AR or SAD, can be derived (Ma et al. 2002; Zhao et al. 2005), facilitating the computation of the model. In this study, we will use a simple AR approach to model the covariance matrix because our focus is on the implementation of ODE into functional mapping. For each variable, the AR model uses two parameters, variance σ_k^2 and correlation ρ_k. Thus, all the parameters that model the residual covariance are arrayed in $\mathbf{\Psi}_v = (\{\sigma_k^2, \rho_k, \eta_k\}_{k=1}^{5})$. Note that η_k describes the relative size of the variances due to e_{ki} and ε_{ki}. Note that the AR model assumes the stationarity of variances and covariances. If these two assumptions are violated, other approaches should be used (see above for a discussion of this issue).

Modeling the Mean Vector

Modeling the mean vector (4) for the dynamic changes of the concentrations of mRNA and various forms of protein for different QTL genotypes by differential equations (1) is the most innovative and difficult part of this study. A clear understanding of the mathematical properties of these differential equations will provide a prerequisite for estimating genotype-specific parameters that define biological rhythms from observational data.

For a rhythmic system governed by ODE (1), we use $\Psi_u=(\{v_{s_j}, v_{m_j}, k_{s_j}, v_{d_j}, k_{1j}, k_{2j}, K_{I_j}, K_{d_j}, K_{1j}, K_{2j}, K_{3j}, K_{4j}, n_j, V_{1j}, V_{2j}, V_{3j}, V_{4j}, K_{m_j}\}_{j=1}^{3})$ to define QTL genotype-specific differences in the dynamics of the system. Some of these parameters describe the oscillation behavior of the system, whereas the others determine the phase response curves under outside stimuli. These parameters have no analytical solution. We implemented a numerical approach for estimating parameters that define QTL genotype-specific differential equations based on the Runge-Kutta fourth algorithm (Press et al. 1992).

Let $\mathbf{z}_{j|i}=(\{\mathbf{y}_{k(j|i)}\}_{k=1}^{5})$ denote the vector of five time-dependent variables (R, P_0, P_1, P_2, P_N) for subject i who carries QTL genotype j. Rewrite the differential equation for variable k in the ODE system (1) as

$$g_{k(j|i)}(t, \mathbf{z}_{j|i})=\frac{dy_{k(j|i)}}{dt}, \quad \text{for } k=1, \ldots, 5.$$

We define the $\mathbf{z}_{k(j|i)}$ value in iterative step l by $\mathbf{z}_{k(j|i)}^{\{l\}}$. Based on the Runge-Kutta scheme, we have

$$\mathbf{z}_{k(j|i)}^{\{l+1\}}=\mathbf{z}_{k(j|i)}^{\{l\}}+\frac{1}{6}(\mathbf{RK}_{1(j|i)}+2\mathbf{RK}_{2(j|i)}+2\mathbf{RK}_{3(j|i)}+\mathbf{RK}_{4(j|i)}), \tag{7}$$

where $RK_{1(j|i)}=(\{RK_{1k(j|i)}\}_{k=1}^{5})$, $RK_{2(j|i)}=(\{RK_{2k(j|i)}\}_{k=1}^{5})$, $RK_{3(j|i)}=(\{RK_{3k(j|i)}\}_{k=1}^{5})$, and $RK_{4(j|i)}=(\{RK_{4k(j|i)}\}_{k=1}^{5})$ and

$$\begin{aligned}
\mathbf{RK}_{1k(j|i)}&=hg_{k(j|i)}(t^{\{l\}}, \mathbf{z}_{j|i}^{\{l\}}),\\
\mathbf{RK}_{2k(j|i)}&=hg_{k(j|i)}(t^{\{l\}}+\tfrac{1}{2}h, \mathbf{z}_{j|i}^{\{l\}}+\tfrac{1}{2}\mathbf{RK}_{1(j|i)}),\\
\mathbf{RK}_{3k(j|i)}&=hg_{k(j|i)}(t^{\{l\}}+\tfrac{1}{2}h, \mathbf{z}_{j|i}^{\{l\}}+\tfrac{1}{2}\mathbf{RK}_{2(j|i)}),\\
\mathbf{RK}_{4k(j|i)}&=hg_{k(j|i)}(t^{\{l\}}+\tfrac{1}{2}h, \mathbf{z}_{j|i}^{\{l\}}+\tfrac{1}{2}\mathbf{RK}_{3(j|i)}).
\end{aligned}$$

The Runge-Kutta fourth order algorithm with step size $h = 0.1$ is used to approximate the solution in high accuracy given a trial set of parameter values and initial conditions.

Estimation

The parameters that define the likelihood (3) are obtained by differentiating the likelihood with respect to each parameter, letting the derivative equal to zero, and then solving the log-likelihood equations. We implemented the EM algorithm to estimate the parameters. The E step is designed to calculate the posterior probability with which subject i has QTL genotype j given its marker and phenotypic information, expressed as

$$\Pi_{ij}=\frac{\pi_{j|i}f_j(z_i)}{\sum_{j'=1}^{3}\pi_{j'|i}f_{j'}(z_i)}. \tag{8}$$

Using the calculated posterior probabilities, the M step is derived to solve the haplotype frequencies expressed as

$$\begin{aligned}
\widehat{p}_{11} &= \frac{1}{2N}\left[\sum_{i=1}^{N_1}(2\Pi_{i1}+\Pi_{i2})+\sum_{i=1}^{N_2}(\Pi_{i1}+\theta\Pi_{i2})\right],\\
\widehat{p}_{10} &= \frac{1}{2N}\left[\sum_{i=1}^{N_1}(\Pi_{i2}+2\Pi_{i3})+\sum_{i=1}^{N_2}(\Pi_{i3}+(1-\theta)\Pi_{i2})\right],\\
\widehat{p}_{01} &= \frac{1}{2N}\left[\sum_{i=1}^{N_3}(2\Pi_{i1}+\Pi_{i2})+\sum_{i=1}^{N_2}(\Pi_{i1}+(1-\theta)\Pi_{i2})\right],\\
\widehat{p}_{00} &= \frac{1}{2N}\left[\sum_{i=1}^{N_3}(\Pi_{i2}+2\Pi_{i1})+\sum_{i=1}^{N_2}(\Pi_{i3}+\theta\Pi_{i2})\right],
\end{aligned} \tag{9}$$

where $\theta = p_{11}p_{00}/(p_{11}p_{00} + p_{10}p_{01})$. The iteration including the E (8) and M steps (9) are repeated until the estimates converge to stable values. These stables values are the maximum likelihood estimates (MLEs) of parameters.

In the M step, the mean vector in $f_j(z_i)$ is modeled by the ODE (1) with parameters $\mathbf{\Psi}_u$ solved with the Runge-Kutta fourth order algorithm. The AR(1)-structured covariance in $f_j(z_i)$ is specified by parameters $\mathbf{\Psi}_v$ estimated by the simplex algorithm. Both Runge-Kutta fourth order and simplex algorithms are embedded within the EM algorithm described above.

Hypothesis Tests

Whether there is a significant QTL that controls circadian rhythms can be tested. The hypotheses for this test are formulated as

$$\begin{aligned}
H_0:&\quad \mathbf{\Psi}_{u_j} \equiv \mathbf{\Psi}_u,\ (j=1,2,3)\\
H_1:&\quad \text{At least one of the equalities above does not hold,}
\end{aligned}$$

where the H_0 corresponds to the reduced model, in which the data can be fit by a single oscillation, and the H_1 corresponds to the full model, in which there exist three QTL genotype-specific oscillations to fit these data. The test statistics for testing the hypothesis above is calculated as the log-likelihood ratio (LR) of the reduced to the full model.

$$LR = -2[\log L(\tilde{\mathbf{\Psi}}_u, \tilde{\mathbf{\Psi}}_v) - \log L(\widehat{\mathbf{\Delta}})],$$

where the tildes and hats denote the MLEs of the unknown parameters under the H_0 and H_1, respectively. An empirical approach for determining the critical threshold is based on permutation test, as advocated by Churchill and Doerge (1994).

We can also test the significance of the genetic effect of a QTL on individual parameters defining the ODE (1) individually or jointly, depending on the purpose of the study. For example, if we are interested in testing how the QTL affects the capacity of a rhythmic system to yield the limit cycle, we can formulate the following hypotheses about the Hill coefficient n:

$$H_0: n_j \equiv n\,(j=1,2,3)\ \text{vs.}\ H_1: n_j \neq n.$$

This test can determine how a gene controls the oscillation behavior of the system. Furthermore, the period, amplitude, and phase of the oscillation can all be tested accordingly.

NUMERICAL RESULTS FROM COMPUTER SIMULATION

The aim of our simulation experiment is to estimate all the unknown parameters using the simulated data and then compare the estimated parameters with the true ones that were used to generate the data. Thus, computer simulation will permit us to verify the proposed model. The study design for computer simulation includes a sample size of $N = 400$ subjects randomly derived from an HWE human population. A segregating QTL for circadian rhythms has three genotypes, *AA* (1), *Aa* (2), and *aa* (3), each with a different set of rhythmic curves derived from the biochemical mechanism shown in Fig. 1. A few empirical studies with Drosophila (Goldbeter 1995) reported estimated values for a total of 18 parameters that define a system of ODE (1). We choose the ODE-parameter values within the ranges of these estimates. For three different QTL genotypes, the following sets of parameter values are used:

$$\begin{aligned}\boldsymbol{\Psi}_{u_j} &= (\{v_{s_j}, v_{m_j}, k_{s_j}, v_{d_j}, k_{1j}, k_{2j}, K_{I_j}, K_{d_j}, K_{1j}, K_{2j}, K_{3j}, K_{4j}, n_j, V_{1j}, V_{2j}, V_{3j}, V_{4j}, K_{m_j}\}) \\ &= \begin{cases} (0.76, 0.65, 0.38, 0.95, 1.9, 1.3, 1, 0.2, 2, 2, 2, 2, 4, 3.2, 1.58, 5, 2.5, 0.5), & \text{for } j=1 \\ (0.82, 0.5, 0.28, 0.6, 2.3, 1, 1.6, 0.5, 2.2, 2.2, 2.2, 2.2, 6, 2.2, 1.38, 8, 1.5, 0.7), & \text{for } j=2 \\ (1, 0.3, 0.48, 0.75, 1.5, 0.8, 0.6, 0.7, 2.5, 2.5, 2.5, 2.5, 10, 4.2, 1.78, 12, 0.5, 1.5). & \text{for } j=3 \end{cases}\end{aligned}$$

The additional rules for choosing these parameters include: first, they should assure the three QTL genotypes different from each other; second, for each set of parameters, the system generates the stable oscillation solutions; and third, the period of each oscillation (i.e. the limit cycle) is close to 24h. Figure 2 shows the appropriateness of the set of parameter values for three different QTL genotypes chosen above, with which a system constructed by a five-variable ODE (1) has sustained oscillations. When plotting the time evolution of per mRNA as a function of different types of PER protein, these oscillations evolve toward a limit cycle (i.e., a closed curve) (Fig. 2).

The above genotype-specific ODE parameter values were used to simulate time-specific phenotypic values for each variable by assuming that five variables follow a multivariate normal distribution with mean vector specified by the ODE (1) and covariance specified by an AR(1) model. Assume that residual errors are purely due to permanent errors. Four different schemes are used with different heritabilities (0.1 and 0.4) and different numbers of time points (5 and 25). We assume that all subjects have the same measure schedules in terms of the length of time intervals and the number of time points although the model allows subject-specific measure schedules. In each simulation scheme, different variables have the same heritability defined at a middle time point.

We simulated a marker associated with the QTL through a gametic linkage disequilibrium. The allele frequencies at the marker were used to simulate the distribution of three marker genotypes in the sampled population. Within each marker genotype, QTL genotypes are distributed following the conditional probabilities calculated from Table 1 which are expressed in terms of haplotype frequencies.

The simulated data were analyzed by the model proposed, with results shown in Tables 2 to 5. For each scheme, the same simulation was repeated 200 times, allowing the calculation of the means and standard errors of parameter estimates. In general, all the 18 parameters that define a 5-dimensional system of ODE (1) can well be estimated with reasonable accuracy and precision even when the rhythm has a modest heritability (0.1) and is measured at a modest number of time points (5) (Table 2). Figure 3 compares the ODE curves estimated from the model with the true curves, suggesting that they are broadly consistent even when the number of time points and heritability are modest. The estimation precision of all these parameters can be improved quickly with increasing heritability and increasing time points (Tables 3–5; Fig. 4). The general guidance obtained from simulation results is that, if the rhythm has a small heritability, more time points should be used. On the other hand, for a highly heritable rhythm, a small number of time points could be sufficient. The model also shows reasonable power to detect a significant rhythm QTL when the heritability and the number of time points are both modest. As expected, the power increases exponentially with increasing heritability and time points. In practice, adequate power can be obtained by either increasing the heritability to 0.4 for the study in which only fewer time points (say 5) can be available, or increasing the number of time points to 25 for a less heritable rhythm trait (say 0.1).

The estimates of covariance-structuring parameters will affect the inference of ODE parameters. It is shown that the AR(1) parameters can be precisely estimated (Tables 2–5). Increasing heritability and time points can improve the estimates of these parameters. The estimates of the allele frequencies at the putative QTL and its gametic linkage disequilibrium can well be estimated in all the simulation schemes (Tables 2–5). This is partly due to the use of a closed form derived to estimate haplotype frequencies (equation 9).

DISCUSSION

Although there is a widely accepted view that the normal function of biological rhythms is strongly correlated with the genes that control them, the detailed genetic mechanisms by which circadian behavior is originated and mediated are poorly understood (Takahashi 2004). Several studies have identified so-called clock genes and clock-controlled transcription factors through gene mutants in animal models (Reppert and Weaver 2002; Takahashi 2004). The implications of these detected genes in clinical trials will hold a great promise for the determination of an individualized optimal body time for drug administration based on an patient's genetic makeup. It has been suggested that drug administration at the appropriate body time can improve the outcome of pharmacotherapy by maximizing potency and minimizing the toxicity of the drug (Levi et al. 1997), whereas drug administration at an inappropriate body time can induce severe side effects (Ohdo et al. 2001). In practice, body time-dependent therapy, termed "chronotherapy" (Labrecque and Belanger 1991), can be optimized by implementing the patient's genes that control expression levels of his/her physiological variables during the course of a day.

With the completion of the Human Genome Project, it has been possible to draw a comprehensive picture of genetic control for the functions of the biological clock and, in the ultimate, integrate genetic information into routine clinical therapies for disease treatment and prevention. To achieve this goal, there is a pressing need to develop powerful statistical and computational algorithms for detecting genes or quantitative trait loci (QTLs) that determine circadian rhythms. In this article, we present such a statistical model for mapping specific rhythm QTLs. This model integrates two recent developments from different fields. One is the mathematical formulation of circadian rhythms by a system of differential equations which allow the quantitative investigation and prediction of the behavior of rhythmic oscillations (Goldbeter 1996, 2002; Leloup et al. 2006). Because they are derived

directly from molecular mechanisms for biological clocks, these equations have proven to show robust capacity to explore the origin and cause of an oscillatory system. The second is the statistical derivation of functional mapping aimed to detect the dynamic expression of a QTL in developmental courses (Ma et al. 2002; Wu and Lin 2006). Although the role of functional mapping in detecting rhythm QTLs has been recognized in our previous publication (Liu et al. 2007), we only used a moment approach to solve differential equations, thus affecting many mathematical inferences from differential equations. The new model proposed in this article has fully capitalized on the mathematical properties of differential equations and can be used to map the QTLs responsible for rhythmic oscillations that are generated by complex cellular feedback processes composed of a large number of variables.

The new model integrates biological mechanisms into a statistical framework for genetic mapping, allowing a number of hypothesis tests to be made at the interplay between genetics and developmental disorders. Through simulation studies under different schemes, the model was found to perform reasonably well even when the level of heritability for biological rhythms is modest. Numerical simulations also provide a practical guidance about the design of genetic studies used to test molecular control of rhythmic oscillations. When the rhythm has a moderately large heritability, the number of time points can be as less as five, to characterize specific QTLs for the trait. Although we have no real data available, it is anticipated that the model will find an immediate application to solving practical problems. In deed, given the growing availability of gene expression and proteomic profile data, the model will help to stimulate an insightful understanding of the genetic cause and origin of rhythmic oscillations.

The new model can further be integrated with two areas to attract more complicated and more realistic genetic problems. First, the model can be extended to consider a higher dimension of biological systems. By accounting for the effect of light on the circadian system induced by the degradation of the timeless (or TIM) protein, an extended, ten-variable model was proposed in which the negative regulation is exerted by a complex of proteins formed by PER and TIM (Leloup et al. 1999). In mammals, the role of TIM as a partner for PER is played by the cryptochrorne (CRY) protein, and light operates by inducing gene expression rather than protein degradation as in *Drosophila* (Reppert and Weaver 2002). An extended model governed by a system of 16 kinetic equations was constructed by Leloup and Goldbeter (2003). The integration of these high-dimensional problems with our model will make it possible to elucidate a detailed picture of genetic regulatory network for biological rhythms.

Second, to clearly explain the idea, the current model was derived for a single marker analysis. However, complex biological traits may be controlled by a web of genes each interacting with others and environmental factors in a coordinated manner. Genome-wide association studies (GWAS) provide a powerful tool for unraveling the genetic network of trait control. Also, genes may affect a trait through the combination of their alleles on the same chromosome (i.e., the haplotype) (Judson et al. 2000; Bader 2001; Rha et al. 2007; Wu and Lin 2008). The incorporation of our model into GWAS and haplotype analysis will provide an unprecedented opportunity to study the genetic atlas of biological rhythms.

Acknowledgments

This work is partially supported by a Joint grant DMS/NIGMS-0540745 to RW. RL's research is supported by NIDA, NIH grants R21 DA024260 and R21 DA024266. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDA or the NIH.

References

Bader JS. The relative power of snps and haplotype as genetic markers for association tests. Pharmacogenomics 2001;2:11–24. [PubMed: 11258193]

Diggle, PJ.; Heagerty, P.; Liang, KY.; Zeger, SL. Analysis of Longitudinal Data. Oxford University Press; UK: 2002.

Dunlap JC. Molecular bases for circadian clocks. Cell 1999;96:271–290. [PubMed: 9988221]

Fan J, Huang T, Li RZ. Analysis of longitudinal data with semiparametric estimation of covariance function. J Am Stat Assoc 2007;35:632–641. [PubMed: 19707537]

Gabriel KR. Ante-dependence analysis of an ordered set of variables. Trans Roy Soc Edinb 1962;33:201–212.

Gallego M, Eide EJ, Woolf MF, Virshup DM, Forger DB. An opposite role for tau in circadian rhythms revealed by mathematical modeling. Proc Natl Acad Sci USA 2006;103:10618–10623. [PubMed: 16818876]

Gillespie DT. Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. Physic Rev E 1996;54:2084–2091.

Goldbeter A. A model for circadian oscillations in the *Drosophila* period protein (PER). Proc Roy Soc Lond B 1995;261:319–324.

Goldbeter, A. Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behavior. Cambridge University Press; UK: 1996.

Goldbeter A. Computational approaches to cellular rhythms. Nature 2002;420:238–245. [PubMed: 12432409]

Gonze D, Halloy J, Goldbeter A. Stochastic versus deterministic models for circadian rhythms. J Biol Physics 2002;28:637–653.

Gonze D, Halloy J, Leloup JC, Goldbeter A. Stochastic models for circadian rhythms: influence of molecular noise on periodic and chaotic behavior. C R Biologies 2003;326:189–203. [PubMed: 12754937]

Hardin PE, Hall JC, Rosbash M. Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature 1990;343:536–540. [PubMed: 2105471]

Judson R, Stephens JC, Windemuth A. The predictive power of haplotypes in clinical response. Pharmacogenomics 2000;1:15–26. [PubMed: 11258593]

Kronauer RE, Czeisler CA, Pilato SF, Moore-Ede MC, Weitzman ED. Mathematical model of the human circadian system with two interacting oscillators. Am J Physiol 1982;242:R3–R17. [PubMed: 7058927]

Labrecque G, Belanger PM. Biological rhythms in the absorption, distribution, metabolism and excretion of drugs. Pharmacol Therapy 1991;52:95–107.

Leloup JC, Gonze D, Goldbeter A. Limit cycle models for circadian rhythms based on transcriptional regulation in *Drosophila* and *Neurospora*. J Biol Rhythms 1999;14:433–448. [PubMed: 10643740]

Leloup, JC.; Gonze, D.; Goldbeter, A. Computational models for circadian rhythms: Deterministic versus stochastic approaches. In: Kriete, A.; Eils, R., editors. Computational Systems Biology. Elsevier Academic Press; Burlington, San Diego, CA: 2006. p. 249-291.

Levi F, Schibler U. Circadian rhythms: Mechanisms and therapeutic implications. Ann Rev Pharmacol Toxicol 2007;47:593–628. [PubMed: 17209800]

Levi F, Zidani R, Misset JL. Randomised multicentre trial of chronotherapy with oxaliplatin, fluorouracil, and folinic acid in metastatic colorectal cancer. Lancet 1997;350:681–686. [PubMed: 9291901]

Li HY, Huang ZW, Wu S, Gai JY, Zeng YR, Wu RL. A conceptual framework to mapping quantitative trait loci governing ontogenetic allometry. PLoS ONE 2007;2(8):e1245. [PubMed: 18043752]

Lin M, Hou W, Li HY, Johnson JA, Wu RL. Modeling interactive quantitative trait nucleotides for drug response. Bioinformatics 2007;23:1251–1257. [PubMed: 17392331]

Liu T, Liu XL, Chen YM, Wu RL. A computational model for functional mapping of genes that regulate intra-cellular circadian rhythms. Theor Biol Med Model 2007;4:5. [PubMed: 17261199]

Leloup JC, Goldbeter A. Toward a detailed computational model for the mammalian circadian clock. Proc Natl Acad Sci USA 2003;100:7051–7056. [PubMed: 12775757]

Lowrey PL, Takahashi JS. Mammalian circadian biology: Elucidating genome-wide levels of temporal organization. Ann Rev Genom Hum Genet 2004;5:407–441.

Ma CX, Casella G, Wu RL. Functional mapping of quantitative trait loci underlying the character process: A theoretical framework. Genetics 2002;161:1751–1762. [PubMed: 12196415]

Ohdo S, Koyanagi S, Suyama H, Higuchi S, Aramaki H. Changing the dosing schedule minimizes the disruptive effects of interferon on clock function. Nat Med 2001;7:356–360. [PubMed: 11231636]

Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. Numerical Recipies. Cambridge University Press; UK: 1992.

Reppert SM, Weaver DR. Coordination of circadian timing in mammals. Nature 2002;418:935–941. [PubMed: 12198538]

Rha SY, Jeung HC, Choi YH, Yang WI, Yoo JH, Kim BS, Roh JK, Chung HC. An association between RRM1 haplotype and gemcitabineinduced neutropenia in breast cancer patients. Oncologist 2007;12:622–630. [PubMed: 17602053]

Scheper TO, Klinkenberg D, Pennartz C, Pelt JV. A mathematical model for the intracellular circadian rhythm generator. J Neuroscience 1999;19:40–47.

Stelling J, Gilles ED, Doyle FJ III. Robustness properties of circadian clock architectures. Proc Natl Acad Sci USA 2004;101:13210–13215. [PubMed: 15340155]

Stelling J, Sauer U, Szallasi Z, Doyle FJ III, Doyle J. Robustness of cellular functions. Cell 2004;118:675–685. [PubMed: 15369668]

Takahashi JS. Gene regulation. Circadian clocks *à* la CREM. Nature 1993;365:299–300. [PubMed: 8397337]

Takahashi JS. Finding new clock components: Past and future. J Biol Rhy 2004;19:339–347.

Tauber E, Last KS, Olive PJ, Kyriacou CP. Clock gene evolution and functional divergence. J Biol Rhy 2004;19:445–458.

The International HapMap Consortium. The International HapMap Project. Nature 2003;426:789–796. [PubMed: 14685227]

The International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437:1229–1320.

The International HapMap Constortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–862. [PubMed: 17943122]

Turek FW, Joshu C, Kohsaka A, Lin E, Ivanova G, McDearmon E, Laposky A, Losee-Olson S, Easton A, Jensen DR, Eckel RH, Takahashi JS, Bass J. Obesity and metabolic syndrome in circadian clock mutant mice. Science 2005;308:1043–1045. [PubMed: 15845877]

Wang ZH, Wu RL. A statistical model for high-resolution mapping of quantitative trait loci determining human HIV-1 dynamics. Stat Med 2004;23:3033–3051. [PubMed: 15351959]

Webb AAR. The physiology of circadian rhythms in plants. New Phytologist 2003;160:281–303.

Wijnen H, Young MW. Interplay of circadian clocks and metabolic rhythms. Ann Rev Genet 2006;40:409–448. [PubMed: 17094740]

Wu RL, Ma CX, Hou W, Corva P, Medrano JF. Functional mapping of quantitative trait loci that interact with the *hg* gene to regulate growth trajectories in mice. Genetics 2005;171:239–249. [PubMed: 15965258]

Wu RL, Lin M. Functional mapping–A new tool to study the genetic architecture of dynamic complex traits. Nat Rev Genet 2006;7:229–237. [PubMed: 16485021]

Wu, RL.; Lin, M. Statistical and Computational Pharmacogenomics. Chapman & Hall/CRC; London: 2008.

Wu RL, Ma CX, Yang MCK, Chang M, Santra U, Wu S, Huang M, Wang M, Casella G. Quantitative trait loci for growth in *Populus*. Genet Res 2003;81:51–64. [PubMed: 12693683]

Yap J, Fan J, Wu RL. Nonparametric modeling of covariance structure in functional mapping of quantitative trait loci. Biometrics 2009;65:00–00.

Young MW, Kay SA. Time zones: A comparative genetics of circadian clocks. Nat Rev Genet 2001;2:702–715. [PubMed: 11533719]

Zhang B, Tong CF, Yin TM, Zhang XY, Zhuge Q, Huang MR, Wang MX, Wu RL. Detection of quantitative trait loci influencing growth trajectories of adventitious roots in Populus using functional mapping. Tree Genet Genom 2009;5:539–552.

Zhao W, Chen YQ, Casella G, Cheverud JM, Wu RL. A nonstationary model for functional mapping of complex traits. Bioinformatics 2005;21:2469–2477. [PubMed: 15769837]

Zhao W, Ma CX, Cheverud JM, Wu RL. A unifying statistical model for QTL mapping of genotype-sex interaction for developmental trajectories. Physiol Genom 2004;19:218–227.

Zhao W, Zhu J, Gallo-Meagher M, Wu RL. A unified statistical model for functional mapping of genotype × environment interactions for ontogenetic development. Genetics 2004;168:1751–1762. [PubMed: 15579721]

Zimmerman DL, Nunez-Anton V. Parametric modeling of growth curve data: An overview (with discussion). Test 2001;10:1–73.

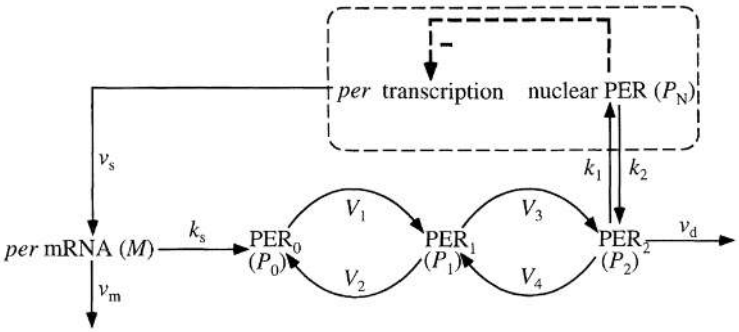

Figure 1.
A molecular model for explaining circadian oscillations in *Drosophila* based on negative autoregulation of the per gene by its protein product PER. Adapted from Goldbeter (1995).

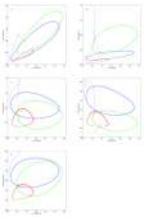

Figure 2.
Limit cycles generated by sustained oscillations of per mRNA R and PER protein P_0 (**A**), P_1 (**B**), P_2 (**C**), P_N (**D**), and the total amount of PER protein (**E**), respectively, for different QTL genotypes, *AA* in green, *Aa* in blue, and *aa* in red, according to ODE (1).

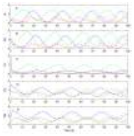

Figure 3.
True curves (solid) of different QTL genotypes, *AA* in green, *Aa* in blue, and *aa* in red, for the time evolution of per mRNA R (**A**), PER protein P_0 (**B**), P_1 (**C**), P_2 (**D**), and P_N (**E**), respectively, in a comparison with the corresponding curves (broken) estimated from 5 time points and 0.1 heritability.

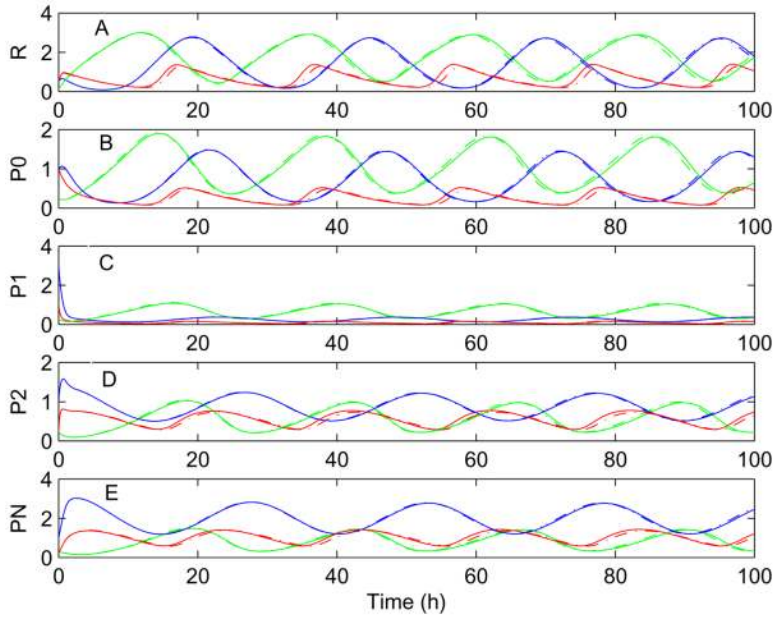

Figure 4.

True curves (solid) of different QTL genotypes, *AA* in green, *Aa* in blue, and *aa* in red, for the time evolution of per mRNA R (**A**), PER protein P_0 (**B**), P_1 (**C**), P_2 (**D**), and P_N (**E**), respectively, in a comparison with the corresponding curves (broken) estimated from 25 time points and 0.4 heritability.

Table 1

Joint genotype frequencies at the marker and QTL in terms of gametic haplotype frequencies.

		AA	*Aa*	*aa*	
Genotype	**Diplotype**	***A*|*A***	***A*|*a* + *a*|*A***	***a*|*a***	**Observations**
MM	*M*|*M*	p_{11}^2	$2p_{11}p_{10}$	p_{10}^2	N_1
Mm	*M*|*m*	$2p_{11}p_{01}$	$2p_{11}p_{00} + 2p_{10}p_{01}$	$2p_{10}p_{00}$	N_2
mm	*m*|*m*	p_{01}^2	$2p_{01}p_{00}$	p_{00}^2	N_3

Table 2

The means and standard errors of the MLEs of three types of parameters ($\mathbf{\Psi}_u$, $\mathbf{\Psi}_v$, $\pi_{j|i}$) defining the likelihood obtained from 200 simulation replicates for QTL mapping of biological rhythms defined by a system of ODE (1). The rhythmic phenotypic traits were simulated at 5 time points for a heritability of 0.1.

	AA		*Aa*		*aa*	
$\mathbf{\Psi}_u$	**True**	**MLE(SE)**	**True**	**MLE(SE)**	**True**	**MLE(SE)**
v_s	0.76	0.7840(0.0041)	0.82	0.8544(0.0068)	1	1.0331(0.0107)
v_m	0.65	0.7182(0.0145)	0.5	0.5238(0.0061)	0.3	0.3001(0.0060)
k_s	0.38	0.3842(0.0061)	0.28	0.2849(0.0043)	0.48	0.5372(0.0125)
v_d	0.95	0.9699(0.0048)	0.6	0.6268(0.0195)	0.75	0.7980(0.0195)
k_1	1.9	1.9834(0.0250)	2.3	2.3297(0.0141)	1.5	1.4866(0.0086)
k_2	1.3	1.3203(0.0043)	1	1.0647(0.0183)	0.8	0.9210(0.0221)
K_I	1	1.0196(0.0041)	1.6	1.6560(0.0057)	0.6	0.6600(0.0233)
K_d	0.2	0.1986(0.0066)	0.5	0.5872(0.0171)	0.7	0.7164(0.0185)
K_1	2	2.0565(0.0244)	2.2	2.1909(0.0110)	2.5	2.4761(0.0089)
K_2	2	2.0122(0.0049)	2.2	2.2219(0.0068)	2.5	2.5354(0.0080)
K_3	2	2.0046(0.0304)	2.2	2.2536(0.0325)	2.5	2.5944(0.0241)
K_4	2	2.0607(0.0114)	2.2	2.1864(0.0124)	2.5	2.5024(0.0111)
n	4	4.0259(0.0117)	6	6.009(0.008)	10	9.9999(0.0124)
V_1	3.2	3.2070(0.0082)	2.2	2.2935(0.0174)	4.2	4.2400(0.0103)
V_2	1.58	1.6049(0.0070)	1.38	1.3657(0.0090)	1.78	1.7846(0.0092)
V_3	5	5.1046(0.0279)	8	8.0127(0.0114)	12	11.9973(0.0110)
V_4	2.5	2.5153(0.0101)	1.5	1.5182(0.0158)	0.5	0.6120(0.0330))
K_m	0.5	0.5193(0.0075)	0.7	0.7180(0.0049)	1.5	1.5323(0.0087)

$\mathbf{\Psi}_v$	**True**	**MLE(SE)**
σ_1	1.62	1.6599(0.0050)
σ_2	0.11	0.1168(0.0041)
σ_3	0.60	0.5100(0.0208)
σ_4	0.29	0.1633(0.0100)

Ψ_v	True	MLE(SE)
σ_5	0.13	0.1427(0.0051)
ρ_1	0.8	0.7546(0.0213)
ρ_2	0.7	0.7289(0.0096)
ρ_3	0.6	0.8057(0.0222)
ρ_4	0.6	0.9142(0.0233)
ρ_5	0.7	0.7347(0.0123)

π_{ji}	True	MLE(SE)
p	0.6	0.6005(0.0016)
q	0.7	0.6674(0.0064)
D	0.05	0.0451(0.0026)

Table 3

The means and standard errors of the MLEs of three types of parameters ($\mathbf{\Psi}_u$, $\mathbf{\Psi}_v$, $\pi_{j|i}$) defining the likelihood obtained from 200 simulation replicates for QTL mapping of biological rhythms defined by a system of ODE (1). The rhythmic phenotypic traits were simulated at 5 time points for a heritability of 0.4.

	AA		*Aa*		*aa*	
$\mathbf{\Psi}_u$	**True**	**MLE(SE)**	**True**	**MLE(SE)**	**True**	**MLE(SE)**
v_s	0.76	0.7661(0.0027)	0.82	0.8249(0.0017)	1	1.0162(0.0043)
v_m	0.65	0.6698(0.0027)	0.5	0.5108(0.0028)	0.3	0.2989(0.0073)
k_s	0.38	0.3775(0.0024)	0.28	0.2836(0.0010)	0.48	0.4796(0.0009)
v_d	0.95	0.9513(0.0005)	0.6	0.6069(0.0109)	0.75	0.7523(0.0036)
k_1	1.9	1.9105(0.0036)	2.3	2.3330(0.0195)	1.5	1.4989(0.0064)
k_2	1.3	1.2995(0.0003)	1	1.0031(0.0012)	0.8	0.8239(0.0028)
K_I	1	1.0016(0.0037)	1.6	1.6239(0.0033)	0.6	0.5956(0.0008)
K_d	0.2	0.1996(0.0023)	0.5	0.5229(0.0068)	0.7	0.7218(0.0181)
K_1	2	2.0782(0.0231)	2.2	2.1849(0.0099)	2.5	2.4951(0.0071)
K_2	2	1.9959(0.0052)	2.2	2.2464(0.0144)	2.5	2.5277(0.0073)
K_3	2	2.0595(0.0176)	2.2	2.2039(0.0047)	2.5	2.5047(0.0074)
K_4	2	2.0201(0.0087)	2.2	2.2215(0.0052)	2.5	2.5381(0.0058)
n	4	4.0206(0.0033)	6	6.0116(0.0036)	10	10.022(0.0043)
V_1	3.2	3.2133(0.0051)	2.2	2.2653(0.0219)	4.2	4.2478(0.0197)
V_2	1.58	1.6257(0.0134)	1.38	1.3668(0.0063)	1.78	1.7923(0.0065)
V_3	5	5.0864(0.0274)	8	8.0220(0.0069)	12	12.025(0.0042)
V_4	2.5	2.5502(0.0171)	1.5	1.5145(0.0056)	0.5	0.5196(0.0057)
K_m	0.5	0.5271(0.0078)	0.7	0.7016(0.0036)	1.5	1.5307(0.0097)

$\mathbf{\Psi}_v$	**True**	**MLE(SE)**
σ_1	0.66	0.6884(0.0032)
σ_2	0.04	0.0535(0.0009)
σ_3	0.26	0.2541(0.0025)
σ_4	0.12	0.0384(0.0036)

Ψ_v	True	MLE(SE)
σ_5	0.05	0.0711(0.0014)
ρ_1	0.8	0.8381(0.0066)
ρ_2	0.7	0.6842(0.0021)
ρ_3	0.6	0.6325(0.0156)
ρ_4	0.6	1.2782(0.0220)
ρ_5	0.7	0.6881(0.0055)

π_{ji}	True	MLE(SE)
p	0.6	0.6007(0.0016)
q	0.7	0.6675(0.0053)
D	0.05	0.0449(0.0028)

Table 4

The means and standard errors of the MLEs of three types of parameters ($\mathbf{\Psi}_u$, $\mathbf{\Psi}_v$, $\pi_{j|i}$) defining the likelihood obtained from 200 simulation replicates for QTL mapping of biological rhythms defined by a system of ODE (1). The rhythmic phenotypic traits were simulated at 25 time points for a heritability of 0.1.

	AA		*Aa*		*aa*	
$\mathbf{\Psi}_u$	**True**	**MLE(SE)**	**True**	**MLE(SE)**	**True**	**MLE(SE)**
v_s	0.76	0.8381(0.0197)	0.82	0.8144(0.0063)	1	1.0219(0.0118)
v_m	0.65	0.6175(0.0064)	0.5	0.5416(0.0045)	0.3	0.3399(0.0103)
k_s	0.38	0.4788(0.0154)	0.28	0.4003(0.0145)	0.48	0.6048(0.0213)
v_d	0.95	0.9770(0.0114)	0.6	0.6116(0.0118)	0.75	0.7373(0.0112)
k_1	1.9	1.8099(0.0145)	2.3	2.3421(0.0057)	1.5	1.5106(0.0089)
k_2	1.3	1.4338(0.0241)	1	0.9724(0.0087)	0.8	0.8045(0.0076)
K_I	1	1.1447(0.0253)	1.6	1.5791(0.0086)	0.6	0.6049(0.0077)
K_d	0.2	0.2350(0.0033)	0.5	0.5382(0.0086)	0.7	0.7181(0.0149)
K_1	2	1.9889(0.0069)	2.2	2.1682(0.0141)	2.5	2.5024(0.0201)
K_2	2	2.0762(0.0180)	2.2	2.2345(0.0061)	2.5	2.5094(0.0108)
K_3	2	2.2132(0.0297)	2.2	2.2819(0.0176)	2.5	2.5359(0.0125)
K_4	2	1.9631(0.0117)	2.2	2.2302(0.0114)	2.5	2.5335(0.0086)
n	4	4.2178(0.0331)	6	6.0304(0.0099)	10	10.027(0.0073)
V_1	3.2	3.3059(0.0121)	2.2	2.3107(0.0215)	4.2	4.2026(0.0086)
V_2	1.58	1.5224(0.0146)	1.38	1.3658(0.0113)	1.78	1.7979(0.0112)
V_3	5	4.9326(0.0200)	8	7.9923(0.0129)	12	12.032(0.0162)
V_4	2.5	2.6169(0.0250)	1.5	1.5364(0.0159)	0.5	0.5405(0.0126)
K_m	0.5	0.6791(0.0291)	0.7	0.6618(0.0091)	1.5	1.5079(0.0083)

$\mathbf{\Psi}_v$	**True**	**MLE(SE)**
σ_1	2.09	2.0976(0.0041)
σ_2	0.94	0.9523(0.0007)
σ_3	0.49	0.1734(0.0079)
σ_4	1.17	1.1747(0.0017)

Ψ_v	True	MLE(SE)
σ_5	2.96	2.9837(0.0023)
ρ_1	0.8	0.8363(0.0048)
ρ_2	0.7	0.6927(0.0036)
ρ_3	0.6	1.1266(0.0135)
ρ_4	0.6	0.6319(0.0016)
ρ_5	0.7	0.6924(0.0037)

π_{ji}	True	MLE(SE)
p	0.6	0.6017(0.0015)
q	0.7	0.6876(0.0034)
D	0.05	0.0322(0.0015)

Table 5

The means and standard errors of the MLEs of three types of parameters ($\mathbf{\Psi}_u$, $\mathbf{\Psi}_v$, $\pi_{j|i}$) defining the likelihood obtained from 200 simulation replicates for QTL mapping of biological rhythms defined by a system of ODE (1). The rhythmic phenotypic traits were simulated at 25 time points for a heritability of 0.4.

	AA		*Aa*		*aa*	
$\mathbf{\Psi}_u$	**True**	**MLE(SE)**	**True**	**MLE(SE)**	**True**	**MLE(SE)**
v_s	0.76	0.7636(0.0017)	0.82	0.8218(0.0007)	1	0.9996(0.0009)
v_m	0.65	0.6501(0.0006)	0.5	0.4991(0.0016)	0.3	0.2893(0.0078)
k_s	0.38	0.3853(0.0017)	0.28	0.2856(0.0030)	0.48	0.4875(0.0068)
v_d	0.95	0.9546(0.0008)	0.6	0.6028(0.0009)	0.75	0.7340(0.0069)
k_1	1.9	1.8951(0.0019)	2.3	2.3021(0.0006)	1.5	1.5008(0.0022)
k_2	1.3	1.3003(0.0008)	1	1.0023(0.0013)	0.8	0.7919(0.0062)
K_I	1	0.9977(0.0015)	1.6	1.6022(0.0007)	0.6	0.5802(0.0069)
K_d	0.2	0.2067(0.0013)	0.5	0.5021(0.0011)	0.7	0.7018(0.0024)
K_1	2	1.9959(0.0009)	2.2	2.1979(0.0009)	2.5	2.4990(0.0009)
K_2	2	2.0002(0.0047)	2.2	2.1978(0.0016)	2.5	2.4989(0.0009)
K_3	2	2.0402(0.0204)	2.2	2.2069(0.0064)	2.5	2.4989(0.0009)
K_4	2	2.0432(0.0226)	2.2	2.1978(0.0011)	2.5	2.4993(0.0007)
n	4	3.9996(0.0013)	6	5.9981(0.0013)	10	9.9985(0.0008)
V_1	3.2	3.2054(0.0066)	2.2	2.2031(0.0006)	4.2	4.1985(0.0012)
V_2	1.58	1.5740(0.0013)	1.38	1.3787(0.0037)	1.78	1.7797(0.0008)
V_3	5	4.9924(0.0025)	8	7.9986(0.0011)	12	11.999(0.0009)
V_4	2.5	2.5285(0.0170)	1.5	1.4980(0.0016)	0.5	0.4990(0.0008)
K_m	0.5	0.5104(0.0068)	0.7	0.6946(0.0031)	1.5	1.4993(0.0019)

$\mathbf{\Psi}_v$	**True**	**MLE(SE)**
σ_1	0.85	0.8607(0.0009)
σ_2	0.38	0.3884(0.0014)
σ_3	0.19	0.1576(0.0059)
σ_4	0.48	0.4801(0.0008)

Ψ_v	True	MLE(SE)
σ_5	1.21	1.2125(0.0011)
ρ_1	0.8	0.8226(0.0048)
ρ_2	0.7	0.7787(0.0162)
ρ_3	0.6	0.5770(0.0306)
ρ_4	0.6	0.7083(0.0155)
ρ_5	0.7	0.7572(0.0087)

π_{ji}	True	MLE(SE)
p	0.6	0.6004(0.0014)
q	0.7	0.6969(0.0022)
D	0.05	0.0618(0.0016)