BIOLOGICAL NETWORKS UNDER OSCILLATING CONDITIONS AND IN EQUILIBRIUM

By

ZHEMING AN

A thesis submitted to the

Graduate School-Camden

Rutgers, The State University of New Jersey

In fulfillment of the requirements

For the degree of Master of Science

Graduate Program in Computational and Integrative Biology

Written under the direction of

Dr. Benedetto Piccoli

And approved by

_____

Benedetto Piccoli

_____

Kwangwon Lee

_____

Stephen Alessandrini

Camden, New Jersey

May 2018

THESIS ABSTRACT

Biological Networks Under Oscillating Conditions and in Equilibrium

by ZHEMING AN

Thesis Director:

Dr. Benedetto Piccoli

Two examples each representing a biological network under oscillating conditions and in equilibrium are studied in the thesis. In the first problem, we recall a new methodology for the simulation and control of large metabolic systems called LIFE (Linear in Flux Expressions) introduced in a recent paper [46].

The second problem focuses on the rhythmic behavior of an organism's biological system. Circadian rhythms are observed in most organisms on earth and are known to play a major role in successful adaptation to the 24-h cycling environment. Circadian phenotypes are characterized by a free-running period that is observed in constant conditions and an entrained phase that is observed in light-dark conditions. The relationship between the free-running period and phase of entrainment is of interest. Our analyses support the view that the circadian period of an organism is not the only predictor of the entrained phase.

Keywords: metabolic systems, Linear in Flux Expressions, quantitative systems pharmacology, Stoichiometric Matrix, kernel, period, phase of entrainment, mathematical model

Declaration of Authorship

I, ZHEMING AN, declare that this thesis titled, "BIOLOGICAL NETWORKS UNDER OSCILLATING CONDITIONS AND IN EQUILIBRIUM" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Dr. Benedetto Piccoli, Distinguished Professor, Joseph and Loretta Lopez Chair Professor of Mathematics, for the patient guidance and advice he has provided. I am very lucky to have him as my advisor who cares so much about my study. His guidance helped me throughout my research and composing of this thesis.

Also, I am highly thankful to Dr. Kwangwon Lee, Associate Professor in Biology Department, and Dr. Stephen Alessandrini, Adjunct Member of the Graduate Faculty, Department of Mathematical Sciences, for their valuable suggestions and indispensable help during my study.

I would like to thank my lab mates Sean McQuade and Nathaniel Merrill for their sympathetic attitude and unfailing help.

I also would like to thank Dr. Karim Azer and Ruth E. Abrams from Sanofi for their wonderful collaboration.

I feel a deep sense of gratitude to my parents who strongly support me through my life.

Contents

Abstract ii

Declaration of Authorship iii

Acknowledgements iv

Contents v

1 Introduction 1

2 Equilibria for Large Metabolic Systems and the LIFE Approach 7

2.1 Introduction 7

2.2 The LIFE Approach 9

2.2.1 LIFE approach for general metabolic systems 9

2.2.2 LIFE approach for Virtual Patients 13

2.3 Key Example 14

2.4 Results 16

2.5 Example 22

2.6 Conclusions 25

3 Experimental and Mathematical Analyses Relating Circadian Period and Phase of Entrainment in *Neurospora crassa* 28

3.1 Introduction 28

3.2 Materials And Methods 30

3.2.1 Strains and Growth conditions 30

3.2.2 Assessing Phenotype 30
3.2.3 τ Analysis 31
3.2.4 Ψ Analysis 32
3.2.5 Mathematical Model 32
3.2.6 Simulations 34
3.3 Results 35
3.4 Discussion 37

1 Introduction

This study concerns two main problems related to biological networks. In the first problem, firstly, we recall a new methodology for the simulation and control of large metabolic systems called LIFE (Linear in Flux Expressions) introduced by a recent paper [46]. In the last decades, great progress of drug development has been made by implementing the Quantitative Systems Pharmacology (QSP) approach which involves systems biology and pharmacology. In preclinical study, animal model test and in patients, QSP explores the mechanisms of pharmacodynamics of new and existing drugs by creating the mathematical and computational models at a systems level. It is an ideal tool for designing drugs and dosing regimens with a consideration of a patient's biology [55, 20]. We gain knowledge by testing a drug on a Virtual Patient (VP), which is a parametrizations of the QSP models reproducing the response to the drug corresponding to a specific patient [69, 22].

QSP models have been implemented in studying disease mechanisms and therapeutic effects for cholesterol, HIV [29, 45] and the others. It is important to note that the initial levels of metabolites (or more general, chemical compounds) of patients receiving treatment can vary greatly, and even patients with similar baseline levels may respond differently. These variations show that, for QSP models to effectively predict the clinical results of a drug, there must be a wide range of VPs representing the variety of patient responses [20, 69, 3]. QSP models generate VPs from typical parameters that have the greatest impact on the model and that vary across the patient population.

Metabolic networks comprise metabolites and fluxes among them.

These interactions could be modeled by graphs in which the nodes represent metabolites and the edges represent fluxes. The dynamics is modeled by a systems of Ordinary Differential Equations (ODEs). Traditional methods focuses on linearity of the systems of ODEs with respect to metabolites [24]. LIFE methodology relies on linearity with respect to the fluxes, thus allowing for non-linearity in metabolites. For given metabolite levels x one can define the kernel $K(x)$ as the null space of the stoichiometric matrix of metabolite reactions, seen as a linear map from the space of fluxes to the space of metabolites. This thesis focuses on understanding the map $x \to K(x)$ assigning to each metabolite level a subspace of the fluxes space, the latter is called Grassmannian. More precisely, we address two main problems: the first is understanding the intersection of the kernels corresponding to different levels of metabolites, while the second is finding all metabolite levels such that a given flux belongs to the corresponding kernels. A number of results for these two problems are presented in terms of properties of the graph representing the metabolic network. Such results are illustrated with an example from the human cholesterol metabolism and a simple toy network (Figure 1.1). We also present a practical application for the LIFE approach; it can be used for QSP simulations, particularly how VPs can be fit to clinical data using an optimization process.

The second problem focuses on the rhythmic behavior of biological system. Organisms, from bacteria to humans, experience predictable and unpredictable environmental changes. The ability of an organism to predict incoming stresses and resources could serve a selective advantage, and biological clocks underlie such an ability; the clock with an approximately 24-h period is called the circadian clock [6, 51, 63]. Among the features of a typical circadian clock [56, 17] are its free-running period

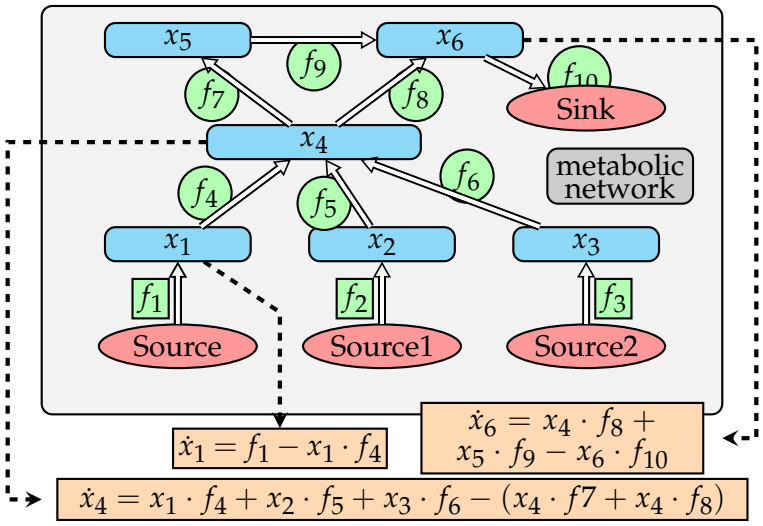

FIGURE 1.1: An example of a human cholesterol metabolic network with corresponding LIFE equations.

(τ), the time that the oscillator takes to complete 1 full cycle in constant conditions, and phase of entrainment (Ψ), the stable time difference between a defined phase of a behavioral or physiological rhythm and the cycling environmental signal (zeitgeber) that synchronizes it to 24 h (Figure 1.2).

Natural selection is predicted to favor a circadian period close to 24 h because this eases the process of entrainment to the daily solar cycle [56, 7, 63]. However, since Ψ is a feature expressed in natural conditions, it is the expected target of natural selection. Variation in Ψ among natural populations is likely to reflect differences in selection on circadian parameters in their progenitors [12, 7]. Since Ψ is critical to the evolution of the circadian clock, it is important to learn how Ψ is determined [62].

There have been significant efforts in theoretical and empirical studies to understand the mechanisms determining Ψ, while recent work has focused mostly on elucidating molecular mechanisms that determine τ

[62, 60]. There are several factors that influence Ψ when τ and the period of the entraining cycle do not match: zeitgeber strength or amplitude [4, 1], zeitgeber period [5, 47, 60], amplitude of the oscillator [38, 57], amplitude relaxation rate and coupling among oscillators [1], and photoperiod [60, 67].

A simple "rule" on the relationship between τ and Ψ is that a short τ leads to an earlier (advanced) Ψ and a long τ leads to a later (delayed) Ψ-has been supported by some dramatic examples. The short period mutant tau in the Syrian hamster showed an advanced Ψ [59, 42]; a short period allele of period in Drosophila melanogaster showed an advanced Ψ and a long period allele a delayed Ψ [26, 66]; a short period allele of frequency, $frq1$, in Neurospora crassa showed an advanced Ψ and a long period allele $frq7$ a delayed Ψ [60]; and even in human sleep disorders, with mutations in $hPer2$ in familial advanced sleep phase syndrome (FASPS) and in CRY in familial delayed sleep phase syndrome [71, 53].

However, this simple view does not explain all available data. The seminal work in explaining FASPS as the result of a short τ was based on the data from one patient [32]. In elderly individuals, advanced Ψ cannot be explained by a shortened τ with age [16, 15]. There are also examples in clock mutants that showed an altered Ψ without a change in τ [33, 27]. Although there are many genetic resources available to examine the relationship between Ψ and τ, most of the studies have focused on τ [41, 44, 74, 75].

We wished to systematically test the relationship between Ψ and τ using the model system Neurospora crassa and considered 3 different options: (1) association study among natural populations, if there is a

causative relationship between Ψ and τ; (2) biological aging or pharmaceutical alteration of τ; and (3) genetic alteration of τ.

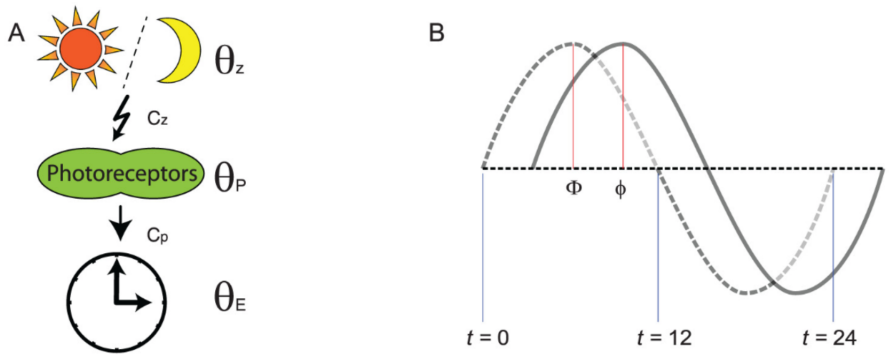

FIGURE 1.2: Mathematical model. (A) A cartoon describing the mathematical parameters in a biological system. The zeitgeber (θ_Z) entrains photosensitive protein (θ_P). Entrainment is active during a specific time of day, hence the photoreaction arrow. The endogenous clock (θ_E) is influenced by the state of the photosensitive protein (θ_P). The clock constantly aligns with the photosensitive protein, hence the straight arrow. (B) A plot showing the zeitgeber signal and the endogenous clock signal. The dotted wave line represents zeitgeber signal (θ_Z). The solid wave line represents endogenous signal (θ_E). The phase of entrainment (Ψ) between the 2 signals is calculated as $y = \Phi - \phi$. The blue lines indicate reference times of day, that is, $\theta_Z = 0 \Leftrightarrow t \equiv 0(mod24)$,$\theta_Z = \pi \Leftrightarrow t \equiv 12(mod24)$.

In this thesis, Lee lab first examined the correlation between τ and Ψ in 3 different populations of *N. crassa*, with different genetic complexities-natural accessions, near-isogenic strains produced by back-crosses, and classical single gene mutants-and found no simple linear relationship between τ and Ψ in all populations studied. Then, Lee lab generated 14 isogenic F1 populations by crossing classical period mutants to a common female strain and analyzed 2 populations that violated the rule. Finally, to characterize the parameter(s) that determine the relationship between Ψ and τ, we developed a mathematical model. There are previous deterministic models characterizing Ψ based on multiple factors in the literature [1, 67]. Our mathematical model can describe all possible relationships between τ and Ψ; furthermore, the model is dynamic,

showing unstable Ψ's gradually becoming stable over a few days.

2 Equilibria for Large Metabolic Systems and the LIFE Approach

2.1 Introduction

The rate of drug development has increased in recent years. With the improved understanding of the clinically relevant differences among patient biology, there is a growing need to develop treatments in the context of a specific patient. Quantitative Systems Pharmacology (QSP) is an ideal tool for designing drugs and dosing regimens with a consideration of a patient's biology [55, 20]. In QSP, mathematical models of biological systems are implemented in-silico. The effects of a treatment can vary between individual patients. The reasons for this variability are not yet well understood, however with QSP models we may gain understanding by testing a drug on a Virtual Patient (briefly VP), an in-silico representation of a person's response to a drug.

QSP models can be tested on several VPs that sample the space of patient biological networks, and the result of a simulation using a VP better predicts the response to drug for a patient with similar biology to the VP. QSP models have been used for applications in modeling cholesterol, HIV, and arthritis among others [68, 29, 45]. These predictions compare the expected effectiveness of the drug with the current established methods as well as predict the safety of new drug candidates. It is important to note that the initial levels of metabolites (or more general, chemical compounds) of patients receiving treatment can vary greatly, and even

patients with similar baseline levels may respond differently. These variations show that, for QSP models to effectively predict patient response to a drug, there must be a wide range of VPs that represent the variety of patient responses [20, 69, 3].

QSP models generate VPs from several different parameterizations. The purpose of multiple parameterizations is to produce a wide range of responses that closely matches the range of clinical data. The parameters selected are typically parameters that have the greatest impact on the model and that vary across the patient population. These parameters will be chosen to fit the desired range, and then qualified by using a clinical dataset to test the model [68, 29, 3, 21]. The parameters used in QSP models are often assumed to have little to no correlation, or correlation is ignored completely [68, 29, 3]. Some patients may initially have similar baseline chemical compound levels yet respond very differently to treatment, and understanding how a perturbation of parameters effects the system will help predict patient response.

Traditional methods focuses on linearity of the systems of Ordinary Differential Equations (ODEs) [24], representing the metabolic network, with respect to the state of metabolites. LIFE methodology relies on linearity with respect to the fluxes, thus allowing for non-linearity in metabolites. Now, for given metabolite levels x one can define the kernel $K(x)$ as the null space of the stoichiometric matrix of metabolite reactions, seen as a linear map from the space of fluxes to the space of metabolites. This paper focuses on understanding the map $x \to K(x)$ from the space of metabolites to the space of subspaces of fluxes, also called Grassmannian. More precisely, we address two main problems: the first is understanding the intersection of the kernels corresponding

to different levels of metabolites, while the second is finding all metabolite levels such that a given flux belongs to the corresponding kernels. A number of results for these two problems are presented in terms of properties of the graph representing the metabolic network. Such results are illustrated with an example from the human cholesterol metabolism and a simple toy network. We also present a practical application for the LIFE approach; it can be used for QSP simulations, particularly how VPs can be fit to clinical data using an optimization process.

2.2 The LIFE Approach

2.2.1 LIFE approach for general metabolic systems

A biochemical system consists of compounds and chemical reactions. In modeling biochemical systems, the stoichiometric matrix can be formed from the reaction network of the system. This matrix is constructed such that every column represents a chemical reaction and every row represents a chemical compound [50]. The fluxes are the chemical reaction rates. The stoichiometric matrix transforms the flux vector into a vector containing the derivatives of compounds and is dependent on the state of the system.

The LIFE method begins by precisely defining a network of biochemical reactions as a graph, where the nodes are reactants/products of metabolic reactions; the edge labels represent reaction rates. A system of ODEs which governs the quantities of biochemical compounds (x) and

fluxes (f) is written as

$$\frac{dx}{dt} = F(x, f),$$
$$\frac{df}{dt} = G(x, f).$$

The dynamics of the state of the compounds is a relatively fast process compared to the dynamics of fluxes in a biochemical system [35, 24]. We can assume $G(x, f) \approx 0$. In the LIFE method, we consider

$$\frac{dx}{dt} = S(x)f \quad x \in \mathbb{R}^n, f \in \mathbb{R}^m. \tag{2.1}$$

$S(x)$ is called the *Stoichiometric Matrix*. Note that the system is linear in fluxes, but $S(x)$ is not necessarily linear in x. The kernel of $S(x)$ is denoted by $K(S(x))$ (or $K(x)$ for brevity), it is a subspace of the space of fluxes, and a mapping of the state to the corresponding kernel can be written $x \in \mathbb{R}^n \to K(x) \in \mathbb{R}^m$.

$$K(x) = \{f : S(x)f = 0\}.$$

With analysis of the kernel of $S(x)$ in (2.1), we investigate fluxes that keep the system in equilibrium. Generally, biological systems are composed of subsystems which have a faster dynamics than the larger system, and it is reasonable to assume that the state x is not far from equilibrium [24].

We use generalized idea of a *directed graph*, where we allow inflows to a graph from a general source, and outflows from the graph to a general sink. We say *graph* for brevity in this paper.

Definition 2.2.1 *The* indegree *of a node is the number of directed edges for which the node is the terminal vertex. The* outdegree *of a node is the number*

of directed edges for which the node is the initial vertex.

Definition 2.2.2 *A* source *of a graph is a directed edge with a node representing a compound only at the terminal end; the initial vertex has outdegree 1, indegree 0 and is not represented in our system. This is equivalent to an exchange reaction entering the system [50]. A* sink *of a graph is a directed edge with a node representing a compound only at the initial end. The terminal vertex has indegree 1 and outdegree 0, and is not represented in the system. This is equivalent to an exchange reaction leaving the system.*

Definition 2.2.3 *The stoichiometric matrix depends on a state variable x, and is denoted $S(x)$, (or S for brevity). $(S)_{ij} = s_{ij}$ can be defined from a graph. If the edge f_j has initial vertex x_i and terminal vertex x_k, then*

$$\begin{cases} s_{ij} = -x_i \\ s_{kj} = x_i \, . \end{cases}$$

If the edge f_j is a source with terminal vertex x_k,

$$s_{kj} = 1 \, .$$

If the edge f_j is a sink with initial vertex x_i,

$$s_{ij} = -x_i \, .$$

Definition 2.2.4 *A* Weakly Connected Component *of a graph is a maximum subgraph such that an undirected path exists between every pair of nodes. A graph is weakly connected if there exists such a path between every pair of nodes.*

Definition 2.2.5 *The* Grassmannian $G(k,V)$ *is the k-dimensional linear subspace of a space of dimension* V.

The kernel of dimension d of a system is a subset of the $Grassmannian(d,m)$. We study the map $x \to K(x)$ as it relates to perturbations of stable systems.

Lemma 2.2.1 *Let* x *be the initial state for system* (2.1), $f \in K(x)$. *Assume that all eigenvalues of the Jacobian matrix of the system at* x *have negative real part. Then there exists* $\epsilon > 0$ *such that if* $y = x + \delta$, $|\delta| < \epsilon, y(\cdot)$ *is the solution starting at* y

$$\lim_{t \to +\infty} y(t) \in K^{-1}(f).$$

Proof 2.2.1 *The assumption on the eigenvalues of the Jacobian matrix imply the system is Lyapunov stable at* x, *see theorem 4.1.2 of [8], which implies* $\lim_{t\to+\infty} S(y(t))f = 0$ *; we conclude* $\lim_{t\to 0} y(t) \in K^{-1}(f)$.

Lemma 2.2.1 motivates our investigation of $K^{-1}(f)$ and will determine candidate states to which a stable system will return after a perturbation.

Two main problems are investigated in this work.

- Problem 1: Characterize the intersections of the kernel for different states. For $x \neq y$ determine the intersection of the kernels $K(x) \cap K(y)$.
- Problem 2: Given $x, f \in K(x)$ compute $K^{-1}(f)$.

By exploring the map $x \to K(x)$ we will characterize $K(x) \cap K(\tilde{x})$ for some perturbation of the state, $\tilde{x} \neq x$. We show that for a fixed state x, $K(x) \cap K(\tilde{x})$ can have any dimension depending on $\tilde{x}$. That is, for some $\tilde{x}$ there is a $dim(K(x) \cap K(\tilde{x})) = 1$, and for some other $\tilde{x}$, $dim(K(x) \cap K(\tilde{x})) = 2$, etc.

2.2.2 LIFE approach for Virtual Patients

Traditional QSP approaches perturb the fluxes of a system and analyze the response. The LIFE method also perturbs the system, but assumes a steady state prior to the perturbation. We can utilize $K(x)$ to simulate VPs with the LIFE method, and optimize the fluxes to simulate metabolite trajectories that approximate clinical data. We find parameterizations of the system which minimize the distance between compound trajectories and measurements (Fig. 2.1). The first step of the procedure to fit our LIFE model is to generate a flux from $K(x)$. We sample coefficients to use for a linear combination of kernel basis vectors. The coefficients of the basis for the kernel are hereafter called *parameters* and are denoted $a_i, i \in \{1,\dots,k\}$ for $k = dim(K(x))$. Different parameters give a different sample from the kernel, and different samples produce different trajectories over time. We calculate the trajectory of metabolites $x(t)$ according to system (2.1), as well as solutions to a variational system: $v_i \in \mathbb{R}^n$ for $i \in \{1,\dots,m\}$. More precisely, for $\hat{f}$, a flux sampled from $K(x)$, and $\frac{dx}{dt} = S(x) \cdot \hat{f}$

$$\frac{dv_i}{dt} = F_i(x, v_i, \hat{f})$$
$$F_i(y, z, \hat{f}) = (D(S(x) \cdot f) \cdot z)\Big|_{x=y, f=\hat{f}} + \frac{\partial (S(x) \cdot f)_{|x=y}}{\partial f_i}$$

We calculate trajectories of x and v by using a fourth order Runge-Kutta scheme. We use v to calculate the gradient of the cost function and use steepest decent method for minimizing this cost. For measurement

of metabolites at time $t_j, (\bar{x}_j)$, we have

$$J = \sum_j \|x(t_j, a, x_0) - \bar{x}_j\|^2 \tag{2.2}$$

$$\frac{\partial J}{\partial f_i} = \sum_j \langle 2(\hat{x} - \bar{x}(t_j)), v_i(t_j) \rangle \tag{2.3}$$

where $v_i(t_j)$ is a value of v_i corresponding to time t_j. The derivative of the cost (2.3) (see proposition 1 and 2 in [10]) is used to selected new parameters a_i.

The steepest decent method can be implemented to update the coefficients.

$$a_i^{new} = a_i^{old} - gJ_{f_j} \qquad \text{for } i \in \{1, \ldots, k\} \tag{2.4}$$

where $k = dim(K(x))$, with $g \in \mathbb{R}^+$ an optimization parameter.

The LIFE method requires that we sample fluxes from the kernel $K(x)$ of our stoichiometric matrix S, and that all metabolites are positive values. Therefore we consider the intersection of the positive orthant with the $K(x)$. A convex combination of kernel basis vectors with positive entries will achieve this goal, however, it will only describe a subset of the kernel, in general. This problem was recognized by Palsson, [50]. In future work we will investigate this further.

2.3 Key Example

A graph of a simple metabolic network is shown (Fig. 2.2). In this network are six metabolites $x = x_i, i \in \{1, ..., 6\}$. $\dot{x}_i$ indicates the derivative of metabolite, x_i. The fluxes are fixed and our system models the dynamics of the metabolites with a given flux.

In Fig. 2.2, the fluxes inside rectangles $\{f_1, f_2, f_3\}$ represent constant rates, whereas those in circles are first order rates. Specifically, the amount of x_1 molecules increases at a rate of f_1 per hour. Linearity in the flux space of the LIFE method facilitates the description for the kernel. We utilize fluxes from the kernel to analyze the system at steady state.

For this example, $S(x)$ is a 6×10 matrix, f is a vector composed of ten rate constants, from Fig. 2.2. A similar method for modeling biochemical networks is explained in [50], however S is not dependent on x in [50].

$$S(x) = \begin{pmatrix} 1 & 0 & 0 & -x_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -x_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -x_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & x_3 & -x_4 & -x_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_4 & 0 & -x_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_4 & x_5 & -x_6 \end{pmatrix}. \tag{2.5}$$

Stochiometric matrix (2.5) governs the metabolites shown in Fig. 2.2. We may write this system of six ODEs from our sample in matrix form (2.5). One advantage of writing our system this way is we can calculate the kernel of the flux space for large systems. The kernel of $S(x)$ is a set

of flux vectors. We call an element of this set $\hat{f}$

$$\hat{f} = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$$

$$= a_1 \begin{bmatrix} x_6 \\ 0 \\ 0 \\ \frac{x_6}{x_1} \\ 0 \\ 0 \\ 0 \\ \frac{x_6}{x_4} \\ 0 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{x_5}{x_4} \\ \frac{-x_5}{x_4} \\ 1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} -x_3 \\ 0 \\ x_3 \\ \frac{-x_3}{x_1} \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + a_4 \begin{bmatrix} -x_2 \\ x_2 \\ 0 \\ \frac{-x_2}{x_1} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2.6}$$

The kernel for equation (2.5) is given in equation (2.6). Note that there are four free variables, a_1, a_2, a_3, a_4, in this kernel for any fixed set of metabolite levels, x.

2.4 Results

Lemma 2.4.1 *Let $S \in M_{n\times m}$, $n < m$, be a stoichiometric matrix and G the associated directed graph. Assume G to be weakly connected with no sources or sinks. Denote by s_i the ith row of S. Then we have,*

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n \iff \sum_{i=1}^{n} \alpha_i s_i = \vec{0}.$$

Proof 2.4.1 *Because G has no sources and sinks S will have exactly two nonzero elements in each column. This is because each column represents a flow from*

one node to another.

$\Longleftarrow$) *Fix a column j and let a,b be the rows with nonzero entries. Consider a linear combination of the rows of S such that*

$$\sum_{i=1}^{n} \alpha_i s_i = \vec{0}. \tag{2.7}$$

Recall from definition (2.2.3) *we have $s_{a,j} = -s_{b,j}$. Because a,b are the only nonzero entries in column j, the jth entry of $\sum_{i=1}^{n} \alpha_i s_i$ satisfies*

$$\alpha_a s_{a,j} + \alpha_b s_{b,j} = 0 \implies \alpha_a = \alpha_b.$$

Because G is weakly connected there exists a path between any pair of nodes. Select two arbitrary nodes in the graph G and label them v,v'. Let W be the path between v and v' and label the nodes on the path W as $v = v_1, v_2, \ldots, v_{p-1}, v_p = v'$. Let j_i be the edge connecting v_i and v_{i+1}. Then for any i, the j_ith column satisfies

$$\begin{cases} s_{i,j_i} = -s_{i+1,j_i} \\ s_{k,j_i} = 0, \qquad \text{for } k \neq i, i+1. \end{cases}$$

Assume (2.7) *then,*

$$\alpha_i s_{i,j_i} + \alpha_{i+1} s_{i+1,j_i} = 0 \implies \alpha_i = \alpha_{i+1}.$$

Because j_i can represent any edge on path W, we have

$$\alpha_1 = \alpha_2, \alpha_2 = \alpha_3, \ldots, \alpha_{p-1} = \alpha_p \implies \alpha_1 = \alpha_2 = \cdots = \alpha_p.$$

Because v,v' were arbitrary nodes,

$$\sum_{i=1}^{n} \alpha_i s_i = \vec{0} \implies \alpha_1 = \alpha_2 = \cdots = \alpha_n. \tag{2.8}$$

$\Longrightarrow$) *We assume that* $\alpha_1 = \alpha_2 = \cdots = \alpha_n$. *As before, fix a column* j *and let* a, b *be the rows with nonzero entries. From definition* (2.2.3) *we have* $s_{a,j} = -s_{b,j}$. *Now consider the* j*th column of* $\sum_{i=1}^{n} \alpha_i s_i$,

$$\sum_{i=1}^{n} \alpha_i s_{i,j} \tag{2.9}$$

which simplifies to

$$\alpha_a s_{a,j} + \alpha_b s_{b,j} = \alpha_a s_{a,j} - \alpha_b s_{a,j} = (\alpha_a - \alpha_b) s_{a,j} = 0$$

This is true for each column, which gives us $\sum_{i=1}^{n} \alpha_i s_i = \vec{0}$.

Proposition 2.4.1 *Let* $S \in M_{n\times m}$, $n < m$, *be a stoichiometric matrix and* G *the associated directed graph. Assume* G *to be weakly connected with no sources or sinks. Then we have,*

$$Rank(S) = n - 1.$$

Proof 2.4.2 *The* ($\Longrightarrow$) *of lemma 2.4.1 implies* $Rank(S) < n$. *Now we show that* $Rank(S) \geq n - 1$. *Consider the submatrix* S^* *constructed by removing the nth row from* S. *Then for* s_i^* *the* i*th row of* S^*,

$$\sum_{i=1}^{n-1} \alpha_i s_i^* = \left(\sum_{i=1}^{n} \alpha_i s_i \right)_{\big|\alpha_n = 0}. \tag{2.10}$$

by (2.10) *and lemma 2.4.1, it follows that*

$$\sum_{i=1}^{n} \alpha_i s_i^* = \vec{0} \implies \alpha_i = 0 \text{ for all } i \in \{1, \ldots, n-1\}.$$

Therefore, $Rank(S^*) = n - 1 \implies n - 1 \leq Rank(S) < n$ *and so* $Rank(S) = n - 1$.

A system with the properties of proposition 2.4.1 also satisfies the zero deficiency theorem of [19], which implies it has one equilibrium solution.

Proposition 2.4.2 *Let $S \in M_{n\times m}$, $n < m$, be a stoichiometric matrix and G the associated directed graph. Assume G to be weakly connected with at least one source and no sinks. Then we have,*

$$Rank(S) = n.$$

Proof 2.4.3 *First we show that for a graph G with a single source and no sinks, that for S, the stoichiometric matrix for G, $Rank(S) = n$. Let the source be called f_{m+1} and the terminal vertex of f_{m+1} be called x_1. Let G^* be the subgraph of G without the source, and S^* be the matrix for G^*. S^* is a submatrix of S excluding the column containing the source. We have $rank(S^*) = n - 1$. We can use elementary row operations to change a row in S^* without changing the rank of S^*. We replace the first row of S^* with $\sum_{i=1}^{n} s_i^* = \vec{0}$ and call this new matrix S^1, $rank(S^1) = rank(S^*) = n - 1$. Similarly, if we append a column of zeros to the right side of S_1^* the rank will not change. We call the matrix with the added column S^2, $rank(S^2) = rank(S^1) = rank(S^*) = n - 1$. S^2 is S with the first row of S set to $\vec{0}$. Now we replace the first row of S^2 with $(s_{1,1}, s_{1,2}, \ldots, s_{1,m-1}, s_{1,m} = 1)$ which gives us S. Because the first row is independent to all others:*

$$rank(S) = rank(S^2) + 1 = n.$$

For ease of proof the graph contained no sinks. However, adding sinks to the graph will not change the rank of the S. This is because S is already full rank and adding a sink will append a new column to S. A graph with

sources and no sinks is not realistic as it will have continuous accumulation of metabolites.

Proposition 2.4.3 *Let S be the stoichiometric matrix and G the associated directed graph. Assume G to be weakly connected with no sources or sinks. Consider the kernel of S, $K(x)$, and assume that $\tilde{x} = cx$ for some $c \in \mathbb{R}$. Then we have, $K(x) = K(\tilde{x})$.*

Proof 2.4.4 *Let $\hat{f} \in K(x)$, and $\check{f} \in K(\tilde{x})$ then*

$$\begin{aligned} S(x)\hat{f} = 0 &\implies cS(x)\hat{f} = S(\tilde{x})\hat{f} = 0 \\ S(\tilde{x})\check{f} = 0 &\implies cS(x)\check{f} = 0. \end{aligned} \tag{2.11}$$

Proposition 2.4.4 *Let S be a stoichiometric matrix for a graph containing a directed path along three nodes, and the middle node has only one incoming and one outgoing edge. For a state x and different state $\tilde{x}$, if $\tilde{x} \neq cx$ for $c \in \mathbb{R}$ then*

$$K(x) \cap K(\tilde{x}) = \{\vec{0}\}. \tag{2.12}$$

Proof 2.4.5 *G has a directed path along three nodes, initial node x_1, middle node x_2, terminal node x_3; call edge connecting x_1 to x_2 as f_1 the other edge is f_2. Then the second row of S is $s_2 = (-x_1, x_2, 0, \ldots, 0)$ and*

$$\begin{aligned} &S(x)f = 0 \implies f_1 x_1 = f_2 x_2 \implies f_1 = f_2 \frac{x_2}{x_1} \\ &S(\tilde{x})f = 0 \implies f_1 \tilde{x}_1 = f_2 \tilde{x}_2 \implies f_1 = f_2 \frac{\tilde{x}_2}{\tilde{x}_1} \\ &f \in K(x) \cap K(\tilde{x}), f \neq \{\vec{0}\} \implies f_2 \frac{x_2}{x_1} = f_2 \frac{\tilde{x}_2}{\tilde{x}_1} \implies \tilde{x} = cx. \end{aligned}$$

Proposition 2.4.5 *Let $S \in M_{n\times m}$, $n < m$, be a stoichiometric matrix and G the associated directed graph. Assume G to be weakly connected with one source and no sinks. Let $S^* \in M_{n\times m-1}$, be a submatrix of S where the source is*

removed. (WLOG let the source in G be represented by the last column of S). Consider the kernels of S and S^*, $K(x)$ and $K^*(x)$ respectively and let B^* be a basis of $K^*(x)$. Let B be the collection of vectors such that each $b \in B$ is equal to a $b^* \in B^*$ with a 0 appended as the last entry for each vector. Then B is a basis for $K(x)$.

Proof 2.4.6 *We prove:*

1. *for $b \in B, Sb = \vec{0}$ and so $b \in K(x)$.*
2. *B is an independent set with number of elements equal to dimension of $K(x)$.*

Let e be an $n \times 1$ column vector containing a single 1 and the other entries 0.

$$S_{n\times(m+1)}b = \left(S^*|e\right)\left(\begin{array}{c} b^* \\ \hdashline 0 \end{array}\right).$$

Let b_j be the jth entry of vector b. For A_i, the ith entry of the solution to Sb.

$$A_i = \sum_{j=1}^{m} S_{ij}b_j = \sum_{j=1}^{m-1} S^*_{ij}b^*_j + S_{i\,m}\cdot 0 = 0.$$

Appending a 0 to each vector of a linearly independent set gives an linearly independent set. From propositions 2.4.1 and 2.4.2 we know that

$$Rank(S^*(x)) = n-1, Rank(S(x)) = n.$$

The dimension of each kernel is the same, i.e. $dim(K^*(x)) = (m-1) - (n-1)$ and $dim(K(x)) = m - n$. The cardinality of B^* = cardinality of $B = dim(K(x)) = dim(K^*(x))$ because B is a basis and both kernels have the same dimension (though the dimension of their ambient space differs), we conclude that B is a basis for $K(x)$.

2.5 Example

Here we show a complete solution to problem 1. In this section we explore the kernel of an example network. The initial state of the kernel will be characterized, and the intersection of this kernel with the kernels of perturbed metabolic states will be analyzed.

$S(x)$ is the stoichiometric matrix associated to the graph in Fig. 2.3

$$S(x) = \begin{pmatrix} -x_1 & 0 & 0 & x_4 & -x_1 & 0 \\ x_1 & -x_2 & 0 & 0 & 0 & -x_2 \\ 0 & x_2 & -x_3 & 0 & x_1 & 0 \\ 0 & 0 & x_3 & -x_4 & 0 & x_2 \end{pmatrix}.$$

From proposition 2.4.1 we have $rank(S) = 3$, which implies the dimension of the kernel is 3. The basis for the kernel is

$$(v_1|v_2|v_3) = \begin{pmatrix} 0 & -1 & \frac{x_4}{x_1} \\ -1 & -\frac{x_1}{x_2} & \frac{x_4}{x_2} \\ -\frac{x_2}{x_3} & 0 & \frac{x_4}{x_3} \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

A perturbation of $x \to \tilde{x}$ will alter the basis vectors and thus change the kernel $K(x) \to K(\tilde{x})$. $K(\tilde{x})$ may have some non trivial intersection with $K(x)$. Any flux in the perturbed kernel can be represented by the perturbed basis vectors. For all $f \in K(\tilde{x}) : f = \tilde{\lambda}_1\tilde{v_1} + \tilde{\lambda}_2\tilde{v_2} + \tilde{\lambda}_3\tilde{v_3}$ where each $\tilde{v_i}$ represents a perturbed basis vector and each $\tilde{\lambda_i} \in \mathbb{R}$. A flux

$f \in K(x) \cap K(\tilde{x})$ can be found as a solution to the following equation:

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \tilde{\lambda}_1 \tilde{v_1} + \tilde{\lambda}_2 \tilde{v_2} + \tilde{\lambda}_3 \tilde{v_3}.$$

Comparing the equation by components, we have the conditions that must be satisfied for any flux in the intersection.

$$\lambda_1 = \tilde{\lambda}_1, \quad \lambda_2 = \tilde{\lambda}_2, \quad \lambda_3 = \tilde{\lambda}_3 \tag{2.13}$$

$$\begin{cases} \lambda_1 \frac{x_4}{x_1} - \lambda_2 & = \tilde{\lambda}_1 \frac{\tilde{x}_4}{\tilde{x}_1} - \tilde{\lambda_2} \\ \lambda_1 \frac{x_4}{x_2} - \lambda_2 \frac{x_1}{x_2} - \lambda_3 & = \tilde{\lambda}_1 \frac{\tilde{x_4}}{\tilde{x}_2} - \tilde{\lambda}_2 \frac{\tilde{x_1}}{\tilde{x}_2} - \tilde{\lambda_3} \\ \lambda_1 \frac{x_4}{x_3} - \lambda_3 \frac{x_2}{x_3} & = \tilde{\lambda}_1 \frac{\tilde{x}_4}{\tilde{x}_3} - \tilde{\lambda}_3 \frac{\tilde{x}_2}{\tilde{x}_3}. \end{cases} \tag{2.14}$$

With (2.13), we simplify system (2.14) to

$$\frac{x_4}{x_1} = \frac{\tilde{x}_4}{\tilde{x}_1} \tag{2.15}$$

$$\lambda_1 \left(\frac{x_4}{x_2} - \frac{\tilde{x_4}}{\tilde{x_2}} \right) = \lambda_2 \left(\frac{x_1}{x_2} - \frac{\tilde{x_1}}{\tilde{x_2}} \right) \tag{2.16}$$

$$\lambda_1 \left(\frac{x_4}{x_3} - \frac{\tilde{x_4}}{\tilde{x_3}} \right) = \lambda_3 \left(\frac{x_2}{x_3} - \frac{\tilde{x}_2}{\tilde{x}_3} \right). \tag{2.17}$$

Depending on which of the conditions are met the dimension of the intersection $dim(K(x) \cap K(\tilde{x}))$ can be determined. Different perturbations of x will be considered that satisfy only some of these conditions. The following cases ((I) through (V)) show results specific to which condition are satisfied.

(I) Let $\tilde{x}$ be a perturbation such that (2.15) is not satisfied, the intersection will be trivial and $dim(K(x) \cap K(\tilde{x})) = 0$.

(II) Let $\tilde{x}$ be a perturbation which satisfies (2.15) and

$$\left(\frac{x_4}{x_2} - \frac{\tilde{x}_4}{\tilde{x}_2} \right) \left(\frac{x_1}{x_2} - \frac{\tilde{x}_1}{\tilde{x}_2} \right) \neq 0, \left(\frac{x_4}{x_3} - \frac{\tilde{x}_4}{\tilde{x}_3} \right) \left(\frac{x_2}{x_3} - \frac{\tilde{x}_2}{\tilde{x}_3} \right) \neq 0.$$

This allows equations (2.16) and (2.17) to be arranged in the following manner.

$$\lambda_2 = \lambda_1 \frac{\frac{x_4}{x_2} - \frac{\tilde{x}_4}{\tilde{x}_2}}{\frac{x_1}{x_2} - \frac{\tilde{x}_1}{\tilde{x}_2}}, \qquad \lambda_3 = \lambda_1 \frac{\frac{x_4}{x_3} - \frac{\tilde{x}_4}{\tilde{x}_3}}{\frac{x_2}{x_3} - \frac{\tilde{x}_2}{\tilde{x}_3}}.$$

This shows a relationship where both λ_2 and λ_3 depend on λ_1 and the metabolites x_i. λ_1 is the only free variable and so $dim(K(x) \cap K(\tilde{x})) = 1$.

(III) Let $\tilde{x}$ satisfy (2.15). And also let $\left(\frac{x_4}{x_3} - \frac{\tilde{x}_4}{\tilde{x}_3}\right)\left(\frac{x_2}{x_3} - \frac{\tilde{x}_2}{\tilde{x}_3}\right) \neq 0$, $\frac{x_4}{x_2} - \frac{\tilde{x}_4}{\tilde{x}_2} = 0$, $\frac{x_1}{x_2} - \frac{\tilde{x}_1}{\tilde{x}_2} = 0$. Then (2.16) is satisfied regardless of the value of λ_2. λ_2 is a free variable in addition to λ_1 while λ_3 is still dependent on the state $\tilde{x}$ and λ_1. With two free variables $dim(K(x) \cap K(\tilde{x})) = 2$.

(IV) Let $\tilde{x}$ satisfy (2.15) and also let

$$\frac{x_4}{x_3} - \frac{\tilde{x}_4}{\tilde{x}_3} = 0, \frac{x_2}{x_3} - \frac{\tilde{x}_2}{\tilde{x}_3} = 0 \tag{2.18}$$

$$\left(\frac{x_4}{x_2} - \frac{\tilde{x}_4}{\tilde{x}_2}\right)\left(\frac{x_1}{x_2} - \frac{\tilde{x}_1}{\tilde{x}_2}\right) \neq 0. \tag{2.19}$$

Upon further inspection, however, we find that (2.15) and (2.18) implies $\left(\frac{x_4}{x_2} - \frac{\tilde{x}_4}{\tilde{x}_2}\right) = \left(\frac{x_1}{x_2} - \frac{\tilde{x}_1}{\tilde{x}_2}\right) = 0$ which contradicts (2.19). Thus the perturbation given by case (IV) doesn't exist.

(V) Let $\tilde{x}$ satisfy (2.15). And let

$$\frac{x_4}{x_2} - \frac{\tilde{x}_4}{\tilde{x}_2} = \frac{x_1}{x_2} - \frac{\tilde{x}_1}{\tilde{x}_2} = \frac{x_4}{x_3} - \frac{\tilde{x}_4}{\tilde{x}_3} = \frac{x_2}{x_3} - \frac{\tilde{x}_2}{\tilde{x}_3} = 0.$$

Equations (2.15), (2.16) and (2.17) are satisfied for any value of λ_1, λ_2 and λ_3. With three free variables $dim(K(x) \cap K(\tilde{x})) = 3$. Fig. 2.4 shows states $\tilde{x}$ for which $K(x) \cap K(\tilde{x})$ is non trivial. The reference state x is shown, and the entire space represents other states $\tilde{x}$ such that equation (2.15) is satisfied.

2.6 Conclusions

Our method exploits the linearity of some metabolic systems with respect to fluxes. Namely, we have demonstrated that there are systematic relationships among fluxes at equilibrium. In our method, model parameters correspond to kernel basis vectors, as opposed to individual fluxes. We've presented results concerning the rank of our stoichiometric matrix, from which the dimension of the kernel is deduced. Then, we demonstrated that the kernels of $S(x)$ and $S(\tilde{x})$ may have non-trivial intersection for perturbed $\tilde{x}$, and thus perturbations do not necessarily exit the kernel.

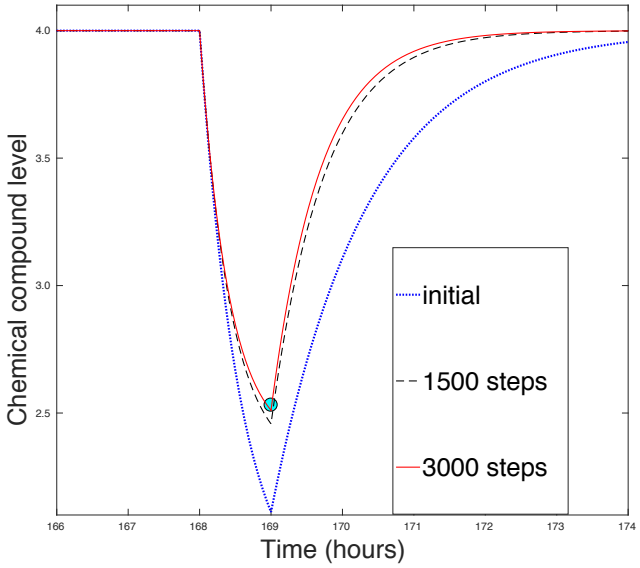

FIGURE 2.1: An example of optimization algorithm performed with the LIFE method. Each step of the procedure minimizes cost in (2.2) according to measurements (cyan dot).

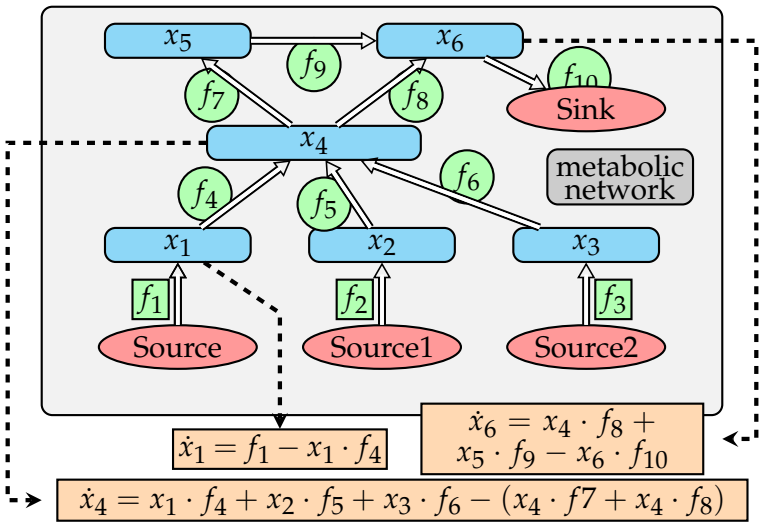

FIGURE 2.2: An example of a human cholesterol metabolic network with corresponding LIFE equations.

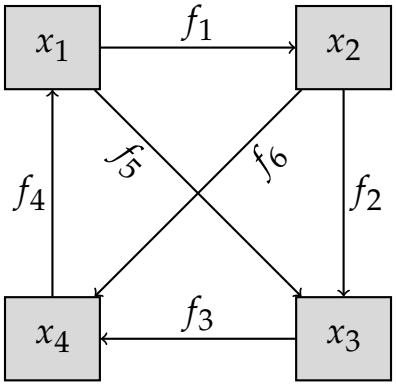

FIGURE 2.3: A directed graph representing a biochemical system.

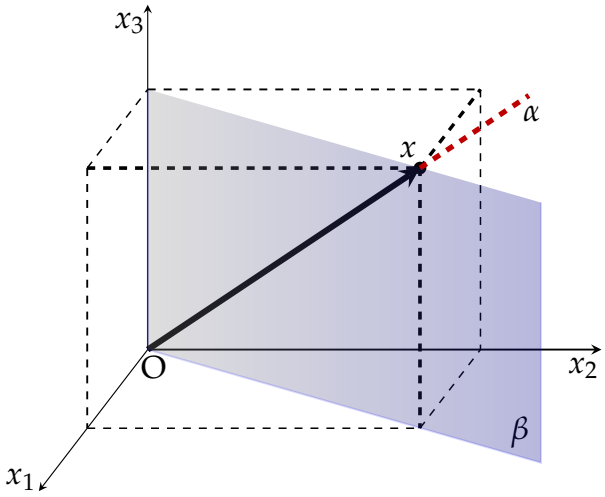

FIGURE 2.4: A 3-D representation of the metabolic state space which highlights states $\tilde{x}$ with nontrivial intersections with the kernel of the initial state (represented by x). A three dimensional representation is appropriate because (2.15) implies that x_1 and x_4 are proportional. The line α represents case (IV) where $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ are proportional to x_1, x_2, x_3 respectively. States $\tilde{x}$ on this line will have kernels such that $dimension(K(x) \cap K(\tilde{x})) = 3$. The plane β represents case (III) where only $\tilde{x}_1, \tilde{x}_2$ are proportional to x_1, x_2.

3 Experimental and Mathematical Analyses Relating Circadian Period and Phase of Entrainment in *Neurospora crassa*

3.1 Introduction

Organisms, from bacteria to humans, experience predictable and unpredictable environmental changes. The ability of an organism to predict incoming stresses and resources could serve a selective advantage, and biological clocks underlie such an ability; the clock with an approximately 24-h period is called the circadian clock [6, 51, 63]. Among the features of a typical circadian clock [56, 17] are its free-running period (τ), the time that the oscillator takes to complete 1 full cycle in constant conditions, and phase of entrainment (Ψ), the stable time difference between a defined phase of a behavioral or physiological rhythm and the cycling environmental signal (zeitgeber) that synchronizes it to 24 h.

There have been significant efforts in theoretical and empirical studies to understand the mechanisms determining Ψ, while recent work has focused mostly on elucidating molecular mechanisms that determine τ [62, 60]. A simple "rule" on the relationship between τ and Ψ-that a short τ leads to an earlier (advanced) Ψ and a long τ leads to a later (delayed)

Ψ-has been supported by some dramatic examples.

However, this simple view does not explain all available data. The seminal work in explaining FASPS as the result of a short τ was based on the data from 1 patient [32]. In elderly individuals, advanced Ψ cannot be explained by a shortened τ with age [16, 15]. There are also examples in clock mutants that showed an altered Ψ without a change in τ [33, 27].

We wished to systematically test the relationship between Ψ and τ using the model system Neurospora crassa and considered 3 different options: (1) association study among natural populations, if there is a causative relationship between Ψ and τ; (2) biological aging or pharmaceutical alteration of τ; and (3) genetic alteration of τ.

In previous study using the first approach, Lee lab could not observe meaningful correlations between Ψ and τ among 564 F1 strains in 3 mapping populations (supplementary figure 1 in [34]). However, Lee lab could not make a definitive conclusion since they analyzed only 3 mapping populations generated from 6 ecotype parents. The second approach could be interesting, and there are reports showing that τ is altered based on an organism's age, nutritional condition, and pharmaceutical chemical treatment [13, 36, 54]. However, the value of Ψ in individuals with such altered τ might reflect confounding effects because of other age-related or chemical-caused secondary effects and may not lead to a definitive conclusion. Thus, Lee lab chose to use the genetic approach to test the universality of the simple rule relating Ψ and τ (i.e., shorter τ, advanced Ψ; longer τ, delayed Ψ).

In the current study, Lee lab first examined the correlation between τ and Ψ in 3 different populations of N. crassa, with different genetic complexities-natural accessions, near-isogenic strains produced by backcrosses, and classical single gene mutants-and found no simple linear

relationship between τ and Ψ in all populations studied. Then, Lee lab generated 14 isogenic F1 populations by crossing classical period mutants to a common female strain and analyzed 2 populations that violated the rule. Finally, to characterize the parameter(s) that determine the relationship between Ψ and τ, we developed a mathematical model. There are previous deterministic models characterizing Ψ based on multiple factors in the literature [1, 67], but we sought a model that could account for our new experimental results.

3.2 Materials And Methods

3.2.1 Strains and Growth conditions

Lee lab obtained 75 natural ecotypes and 14 other classical mutants from the Fungal Genetics Stock Center (*www.fgsc.net*). *N*453 is the *BC*3 (backcross third generation) between *FGSC* 4715 and *FGSC* 4720. *FGSC* 4720 has been a recurring female strain for back-crosses. Strains used in this study were cultured as previously described [40]. The list of strains used in this study and the analyzed data set to make figures are available in the supplementary data.

3.2.2 Assessing Phenotype

N. crassa, a model organism for the study of circadian rhythm, expresses its rhythm by alternating areas of asexual spore formation (conidiation) with areas of thinner hyphal growth without spore formation when growing on solid agar medium. Areas of spore formation are known as "bands", and thinner areas without spore formation are known as "inter-bands". The observed overt rhythm has a periodicity of about

24 h in constant environmental conditions (free-running period). The rhythm is also reset (entrained) by environmental cues in a cycling environment, which allows the internal clock to align its internal time to the local time. The overt clock phenotypes including τ and Ψ were measured using the inverted race tube assay [52]. Race tubes were incubated in constant light (LL) for 12 h at room temperature. After confirming normal mycelial growth in the race tube, tubes were transferred to an $I-36L$ Percival Scientific (Perry, IA) growth chamber and incubated an additional 12 h under LL. For all experiments, temperature was set at 25 °C. After the 24-h LL treatment, the light was off for the rest of the experiment for the τ measurement. The growing front was marked at the light-to-dark transition and on the last day of the experiment. In the race tube experiment for the Ψ phenotype, the light condition was a light 12 h:dark 12 h (LD) cycle. The growing front of the culture in the race tube was marked every 24 h at the time when the light-to-dark transition occurred. The fluence rate was $250\mu E/m^2/sec$ in LL. Light sources were white fluorescent bulbs and incandescent bulbs (Osram Sylvania, Danvers, MA). In both τ and Ψ experiments, tubes were randomly positioned within the chamber to reduce the possibility of positional effects. In each experiment, 3 replicates of each progeny were assayed. Lee lab repeated the experiment to generate data from at least 3 biological replicates for each strain.

3.2.3 τ Analysis

For the analysis of τ phenotype, individual τ estimates of F1 progenies of each produced after 4 to 5 days of consecutive conidial banding, using the fast Fourier transform nonlinear least-squares program [58, 77] and the program Chrono [64].

3.2.4 Ψ Analysis

The reference phase of individual genotype/progeny was the conidial band center. Thus, the Ψ of individual progeny was determined based on the time elapsed to reach the band center within a day. The time when cultures were transferred to the dark cycle is, by definition, CT12 (dusk). Thus, in these experiments, the time in the band center of an individual was calculated by the following formula: zeitgeber (ZT) Ψ = (growth to band center/overall growth) $\times 24 + 12$. For example, if a conidial band occurs at 180 mm and the total growth after light-dark transition is 280 mm, ZT $\Psi = 24 \times (180/280) + 12 = 27.43$. By convention, ZT is always expressed as 24 ZT h. For example, ZT Ψ27.43 is expressed as ZT 3.43(27.43˜24) instead of ZT 27.43. Lee lab used the program Chrono to calculate Ψ [64].

3.2.5 Mathematical Model

We model the relationship between Ψ and τ with a set of ODEs. Because our focus is on this relationship, we do not model the range of entrainment (i.e., the range of zeitgeber periods to which an oscillator is able to entrain), which depends also on oscillator amplitude and zeitgeber strength [1, 23]. Our model consists of 3 oscillators: the periodic signal representing the environmental rhythm (zeitgeber, θ_Z), the periodic expression of a light-sensitive protein (θ_P), and the periodic expression of an endogenous clock (θ_E). The zeitgeber entrains the light-sensitive protein at a specific time of day, for instance, during the time when light-to-dark transition occurs (dusk). The model identifies the start of an oscillator's period with 0 and the completion of a full cycle with 2π.

The zeitgeber oscillator represents the diurnal cycle, and dawn is represented as $\theta_Z = 0$, while dusk is $\theta_Z = \pi$ ($\pi = 12$ h, for the zeitgeber cycle). This is halfway through the zeitgeber cycle beginning at dawn. In our model, we specify how close the zeitgeber must be to dusk with a parameter denoted ϵ, which is the radius of the time window that defines when the zeitgeber is "close enough" to dusk for entrainment to occur. For instance, with parameter $\epsilon = \frac{\pi}{24}$, our model activates entrainment 30 min before and after dusk, that is, when the zeitgeber angle satisfies $|\theta_Z - \pi| < \epsilon = \frac{\pi}{24}$. The model is given by,

$$\begin{cases} \frac{d\theta_Z}{dt} &= \frac{2\pi}{24} \\ \\ \frac{d\theta_P}{dt} &= \frac{2\pi}{\tau} + \Psi(\theta_Z)(\theta_Z - \theta_P) \\ \\ \frac{d\theta_E}{dt} &= \frac{2\pi}{\tau} + C_P(\theta_P(t - t_0) - \theta_E(t)) \end{cases} \tag{3.1}$$

$$\Psi(\theta_Z) = \begin{cases} C_Z & if |\theta_Z - \pi| < \epsilon \\ 0 & otherwise. \end{cases} \tag{3.2}$$

In the system above, we show the 3 oscillators and the differential equations that govern them; τ is the endogenous period of the biological clock. Entrainment is modeled in 2 stages; each stage synchronizes 2 clocks together by a term in the style of the Kuramoto model [37]. The first stage of entrainment is described by the term $\Psi(\theta_Z)(\theta_Z - \theta_P)$, which synchronizes θ_P with θ_Z. The second stage is described by the term $C_P(\theta_P(t - t_0) - \theta_E(t))$, which synchronizes θ_E with θ_P. We call C_Z "the entrainment strength" and C_P "the alignment strength." The zeitgeber (θ_Z) entrains the light-sensitive protein (θ_P) at dusk. The endogenous

clock (θ_E) constantly aligns to this light-sensitive protein. The entrainment of the endogenous clock to the zeitgeber is a $2-$step process, and in our model, the endogenous clock does not directly sense the zeitgeber but instead senses an intermediate rhythmic molecule. The t_0 term in the equation for $\frac{d\theta}{dt}$ represents an time offset in the alignment of the endogenous clock (θ_E) to the light-sensitive protein (θ_P). Supplementary Document 1 provides a more detailed description of the model.

3.2.6 Simulations

In our simulations, we visualize the periodic signals over time of the endogenous clock and the zeitgeber by plotting $sin(\theta_E)$ and $sin(\theta_Z)$, respectively. We measure the phase of entrainment (Ψ) between the endogenous signal and the zeitgeber by comparing the time corresponding to the peak signal (shown in Fig. 3.1B). The phase of entrainment is defined as,

$$\Psi = \Phi - \phi \tag{3.3}$$

where Φ is the phase of the zeitgeber and ϕ is the phase of the entrained biological oscillation. Note that $\Psi < 0$ implies a delayed phase, and $\Psi > 0$ implies an advanced phase. For every day (peak of the zeitgeber), we measure Ψ (Fig. 3.1B) by observing the time coordinate of the closest peak of the endogenous clock to the zeitgeber peak. When we run our simulations, we observe how Ψ changes from day to day. Plots of time versus Ψ are shown in Fig. 3.1C. The Ψ between the zeitgeber and the endogenous clock changes over consecutive days and then finally stabilizes to a stable Ψ. The code to generate Fig. 3.1C is available in the supplementary data.

3.3 Results

Lee lab measured τ and Ψ among 75 natural N. crassa ecotypes and found a weak correlation between them, Pearson correlation $r = 0.3, p < 0.01$ (Fig. 3.2 A).

Since these natural ecotypes are adapted to different habitats and have different genetic backgrounds, Lee lab measured circadian traits among 9 near-isogenic lines (NIL) *N*453 (see the Materials and Methods section). In theory, these NIL should share 93.75% of the recurrent female's genome. Lee lab still found no apparent correlation between τ and Ψ (Pearson correlation $r = -0.18, p = 0.59$, Fig. 3.2B). Next, Lee lab measured τ and Ψ of the 14 classical mutants of τ. There was an apparent negative correlation between τ and Ψ among mutants (Pearson correlation $r = -0.64, p = 0.01$; Fig. 3.2 C). However, this apparent negative correlation is due to the convention of ZT hours expressed as discontinuous numbers (e.g., ZT 1 = ZT 25). When the Ψ data are expressed as continuous numbers, the correlation between τ and Ψ among all the tested mutants is reduced (Pearson correlation $r = 0.46, p = 0.09$, Fig. 3.2 C).

To further test the relationship between τ and Ψ using the *F*1 populations, Lee lab crossed the classical mutants to the common strain *FGSC* 4720. Lee lab chose 2 classical mutant alleles, a short period mutant *cys*-9 (*FGSC* 2160) and a long period mutant *prd*-1 (*FGSC* 4902), for further analysis.

There were 2 sub-populations with different periods as expected (Fig. 3.3 A,C). More interestingly, there were statistically significant differences in the Ψ's of the sub-populations (Fig. 3.3 B,D; Table 3.1). The short τ progeny in *N*264 had a delayed Ψ (Fig. 3.3 B), and the long τ

progeny in *N*272 had an advanced Ψ (Fig. 3.3 D). These populations represent a stable τ and Ψ relationship that does not obey the simple rule that is thought to relate τ and Ψ (i.e., shorter τ, advanced Ψ; longer τ, delayed Ψ). To understand the parameter space that would produce the experimental relationship, we turned to mathematical modeling.

Our mathematical model (see the Materials and Methods section) can describe all possible relationships between τ and Ψ; furthermore, the model is dynamic, showing unstable Ψ's gradually becoming stable over a few days (Fig. 3.1 C). These 2 aspects of the model suggest that we have a simplified version of mechanisms that govern the periodic behavior of endogenous clocks upon entrainment. The endogenous clock is sensitive to the choice of "the entrainment strength" (C_Z) and "the alignment strength" (C_P; Fig. 3.1 D and Materials and Methods section). We concluded that combinations of the entrainment strength and the alignment strength are key parameters that can lead to a variety of stable τ and Ψ relationships.

TABLE 3.1: **Statistical Analysis**

	Source	SS	df	MS	F	Prob>F
*N*264, τ	Groups	14.641	1	14.641	17.288	0.00317
	Error	6.775	8	0.864		
	Total	21.41604	9			
*N*264, Ψ	Groups	17.360	1	17.360	8.366	0.02012
	Error	16.600	8	2.075		
	Total	33.961	9			
*N*272, τ	Groups	17.080	1	17.080	92.940	4.85E-06
	Error	1.653	9	0.183		
	Total	18.733	10			
*N*272, τ	Groups	13.647	1	13.647	1.906	0.20062
	Error	64.410	9	7.156		
	Total	78.057	10			

3.4 Discussion

Although the relationship between τ and Ψ is of fundamental interest to chronobiologists, there has been no systematic empirical or mathematical study done to address the possibility, as shown here, of a short τ/delayed Ψ and a long τ/advanced Ψ.

One possible explanation for such a relationship between τ and Ψ is that it reflects a different directionality and/or sensitivity of different stages of the circadian oscillator to a zeitgeber. The phase response curve (PRC) has been an important tool to account for Ψ. The PRC represents differential responses of a circadian oscillator to an outside stimulus applied at different phases of its circadian cycle [31]. At the cellular level, τ is quite elastic depending on nutritional and environmental conditions [28, 73, 54]. And there are different stages of a circadian oscillation that

respond differentially to genetic and pharmacological perturbations [54]. A mutation in one clock component may yield a particular τ and Ψ relationship in one context but another relationship in a different one. As an example, a mutation of $CKI\sigma$ that caused $FASPS$ in humans caused a shorter τ in mice and a longer τ in Drosophila [74]. In a typical limit cycle model, the zeitgeber is explained as a unidirectional force moving a state variable to a predictable isochron. Neurospora has been a good tool explaining entrainment mechanisms by light and temperature [17]. However, it is also possible that the same zeitgeber signal may cause different directional changes at the molecular level. For example, light as a zeitgeber entrains fungal and mammalian clocks by inducing the negative regulators [43, 17], whereas the same zeitgeber entrains the fly clock by degrading the negative regulators [14]. It is possible that the different sensitivity of circadian oscillators at different stages to a particular zeitgeber may create a diverse set of τ and Ψ relationships.

Although the PRC is a powerful tool, it is not a sufficient one to account for all τ and Ψ relationships [62, 60]. Ψ is influenced by many different factors [4, 5, 38, 57, 62, 1, 60, 67]. Quantitative trait locus (QTL) analysis is one way of estimating the number of genetic factors underlying a complex trait. In one study with 3 mapping populations of Neurospora, the authors found 16 QTL for τ and 27 QTL for Ψ, in which 7 loci were co-localized for both τ and Ψ [34]. It is not surprising that there are more genetic factors responsible for the variation of Ψ than that of τ. Thus, it may be challenging to derive a unified and simple model that explains all entrainment data; however, such a model could provide new insights on fundamental questions such as the evolution of clocks and on applied aspects of clock studies such as diagnosing or intervening in sleep disorders.

Another process that might account for variation in the relationship between τ and Ψ is the "rate of entrainment", the kinetics of signal transduction, by which a zeitgeber signal reaches the oscillator. In our mathematical model, this could be an example of alignment strength, C_P (discussed more later). There are 2 known examples of Ψ-influencing clock mutants in the Neurospora clock; both are blue-light receptors, $wc-1$ and *vvd* [33, 27]. In these mutants, Ψ is altered while τ is not, demonstrating the feasibility of different Ψ's when the input pathway of a circadian clock is altered.

Lee lab's investigation of a short period mutant $cys-9$ ($FGSC$ 2160) and a long period mutant $prd-1$ ($FGSC$ 4902) revealed associated Ψ's that were counterintuitive. The short period gene $cys-9$ is an $NADPH$-dependent thioredoxin reductase [48]. The authors in their initial report suggested that the clock phenotype could be caused by altered protein modification of potential clock transcription factors or kinases or by altered house-keeping metabolism [48].

Although its short τ clock phenotype was reported 20 years ago, there has been no follow-up study on the gene since the initial report. There are now ample data supporting the existence of circadian oscillations in cellular redox conditions both in prokaryote and eukaryote organisms [70] In the prokaryote clock, the redox change in the form of the adenosine triphosphate (ATP)/adenosine diphosphate ratio can replace the phase shift of the clock by light [65]. In a similar manner, the redox condition plays an important role in circadian clock and light entrainment in *Neurospora* [76, 25]. So one possible explanation for the $cys-9$ phenotype could be that an altered cellular redox condition by the $cys-9$ mutation results in a delayed Ψ. The long period clock gene $prd-1$ is an ATP-dependent DEAD-box RNA helicase [18,

2]. DEAD-box helicases are involved in many cellular processes involving RNA metabolism [61]. In addition to a function unwinding RNA [9], DEAD-box helicases are involved in different functions at different stages of circadian regulation. In the Neurospora clock, FRQ-interacting RNA helicase's primary role is to mediate a proper contact among FRQ, $CK1a$, and WCC [30, 39, 11]. TOGR1, a DEAD-box helicase in the rice plant, is involved in temperature-dependent and clock-controlled plant growth through stabilizing rRNA homeostasis [72]. $DDX5$, an ortholog of $PRD-1$, is associated in the PER complex and regulates per RNA metabolism and its clock function [49]. As in the case of $CYS-9$, $PRD-1$ appears to be involved in housekeeping metabolic processes and not a part of the transcription/translation feedback loop oscillator (TTFL); $prd-1$ is not regulated by light or the TTFL [18]. Although $PRD-1$ is clearly a part of the metabolic oscillator, the biochemical function of $PRD-1$ is not clear yet [18, 2], and it is too early to speculate on how $PRD-1$ contributes to shape the τ/Ψ relationship.

We have constructed a mathematical model to describe the complex relationships between τ and Ψ that we observe in the current study. This model is general in the sense that it does not describe particular molecules that regulate rhythmic expression in the endogenous clock or participate in entrainment. Rather, it describes biological rhythms that synchronize with a zeitgeber and how such signals affect the endogenous clock. In our model, the endogenous clock does not directly sense the zeitgeber. Instead, there is an intermediate quantity that responds to the light-to-dark transition by synchronizing to the zeitgeber as the Kuramoto model describes. We considered parameters in our mathematical model, entrainment strength (C_Z) and the alignment strength of the oscillator to the light receptor (C_P). Zeitgeber strength has been included

in other models [1, 67]. The alignment strength (C_P) could reflect several other factors, including amplitude of the oscillator and coupling of the oscillators.

Some parameterizations of this model show long τ with a delayed Ψ and short τ with an advanced Ψ, as traditional models do. Our proposed model in this work is the result of starting with a simpler model. We began our investigation by producing a very simple set of ODEs with 2 oscillators. In this first model, the zeitgeber entrained the endogenous clock the same way that the zeitgeber entrains the light-sensitive protein in our tri-oscillator model. This direct synchronization of the 2 oscillators showed stable entrainment only with short τ/advanced Ψ and long τ/delayed Ψ relationships (this first model is shown in the supplementary material). To build a model that could explain our experimental observations, we added a third oscillator in our proposed model, which is the light-sensitive protein. The addition of an intermediate protein that would be entrained by the zeitgeber, which in turn would serve as a regulator of the endogenous clock, was able to describe the wider range of behavior. Our mathematical model requires more optimization to include other known factors to explain the full spectrum of τ/Ψ relationships.

In summary, genetic data generated in Lee lab demonstrate variation in the relationship between τ and Ψ, and the mathematical model provides working hypotheses for further characterization of molecular mechanisms of entrainment. Our data also imply that misaligned Ψ's (e.g., in human sleep disorders) may have etiologies distinct from an altered τ.

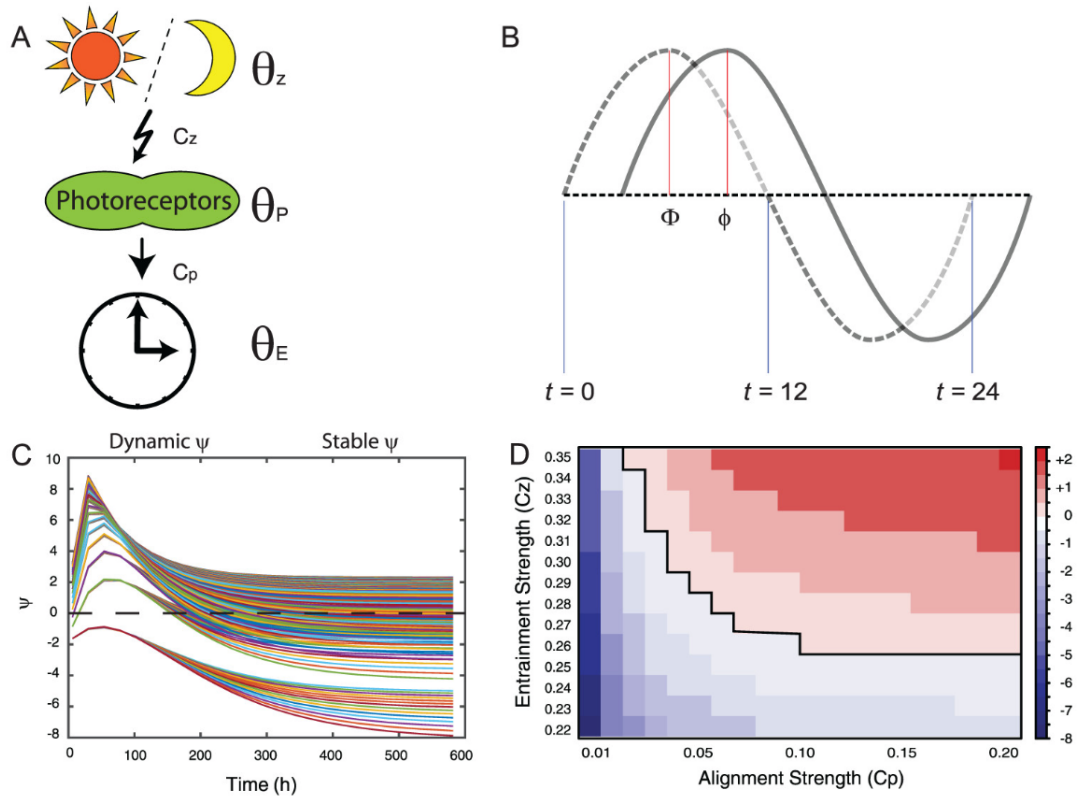

FIGURE 3.1: Mathematical model. (A) A cartoon describing the mathematical parameters in a biological system. The zeitgeber (θ_Z) entrains photosensitive protein (θ_P), and the time derivative for the state of this protein ($\frac{d\theta_P}{dt}$) is a function of the entrainment strength (C_Z). Entrainment is active during a specific time of day, hence the photoreaction arrow. The endogenous clock (θ_E) is influenced by the state of the photosensitive protein (θ_P), and the time derivative of the clock ($\frac{d\theta_E}{dt}$) is a function of the alignment strength (C_P). The clock constantly aligns with the photosensitive protein, hence the straight arrow. (B) A plot showing the zeitgeber signal and the endogenous clock signal. The dotted wave line represents zeitgeber signal (θ_Z). The solid wave line represents endogenous signal (θ_E). The phase of entrainment (Ψ) between the 2 signals is calculated as $y = \Phi - \phi$. The blue lines indicate reference times of day, that is, $\theta_Z = 0 \Leftrightarrow t \equiv 0(mod24), \theta_Z = \pi \Leftrightarrow t \equiv 12(mod24)$. (c) A plot showing how Ψ gradually changes over time (dynamic Ψ) and eventually stabilizes after several days (stable Ψ). (d) A heat map of the simulations in panel *c* showing the values of the entrainment strength (C_Z) and alignment strength (C_P) that produce various stable Ψ's over time.

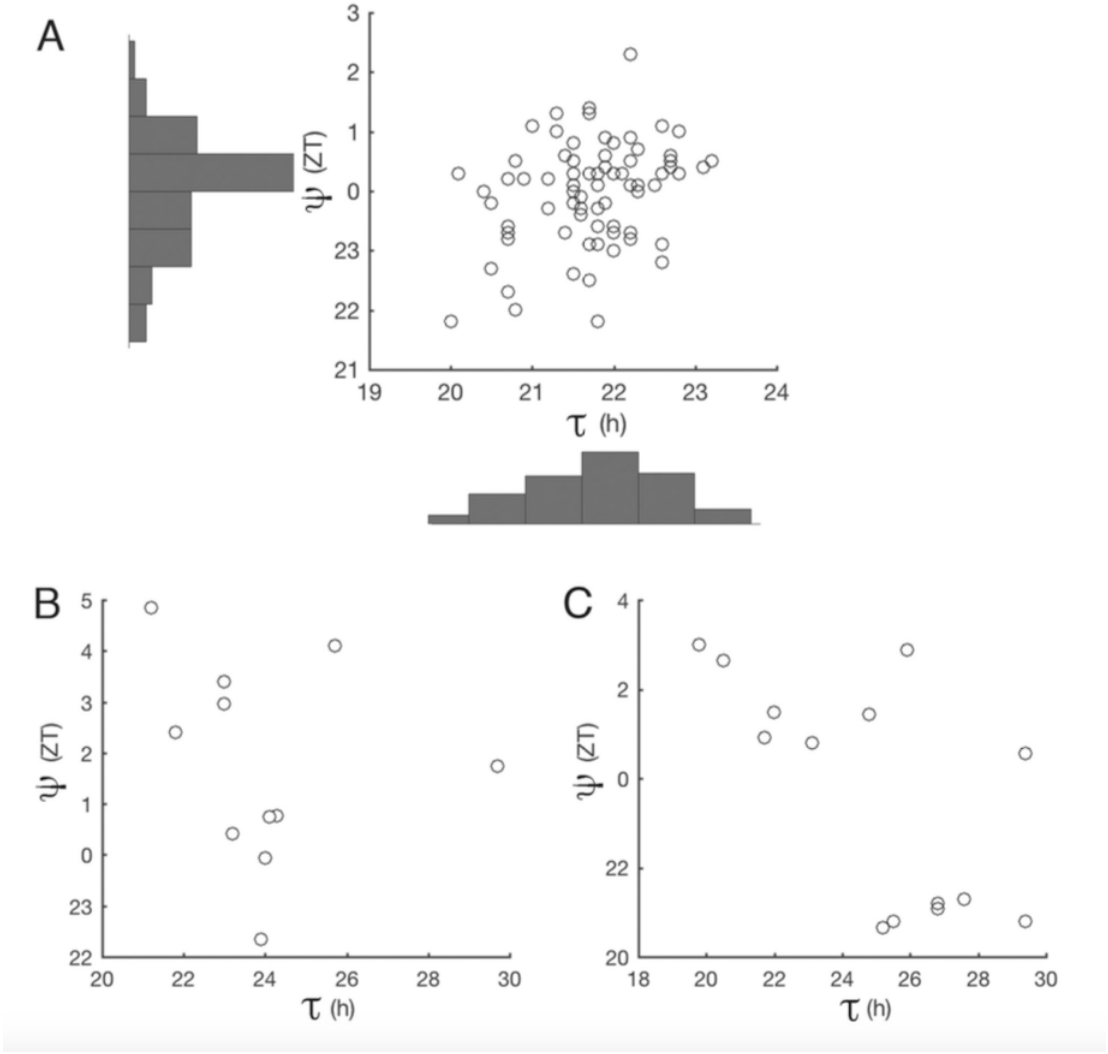

FIGURE 3.2: τ/Ψ correlation in 75 natural ecotypes (A), 9 *BC*3 near-isogenic lines (B), and 14 classical mutants (C).

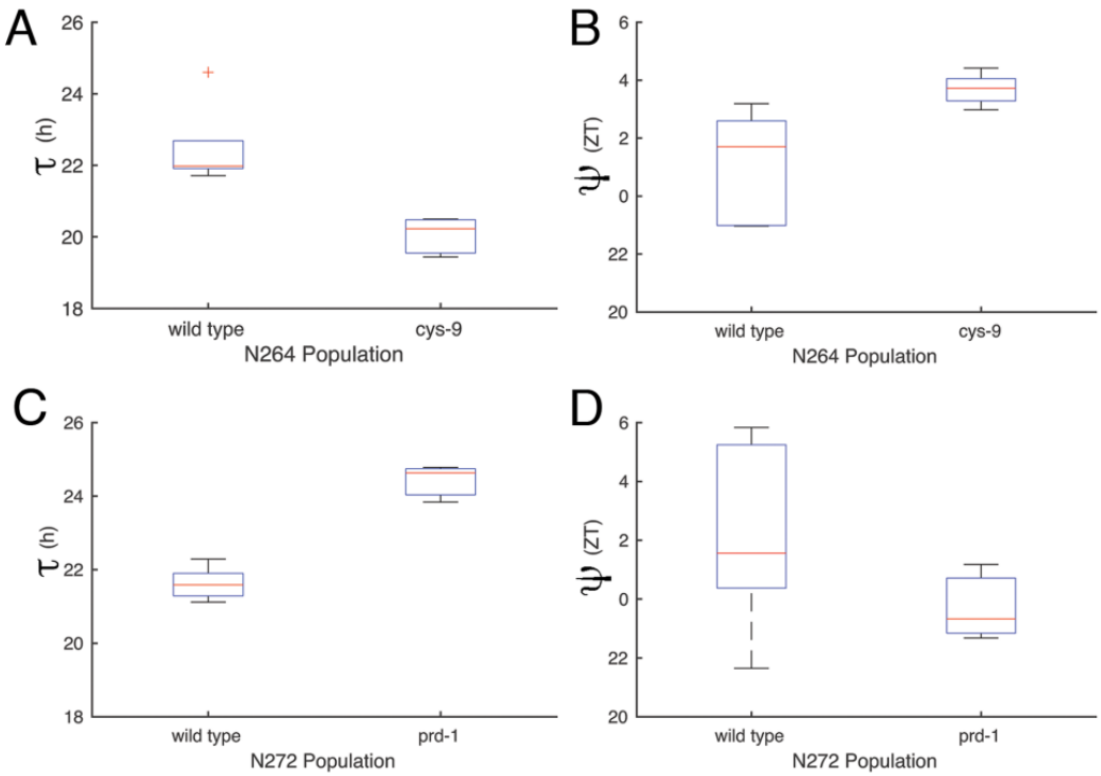

FIGURE 3.3: A significant correlation between τ and Ψ of $F1$ populations; a short period (*cys*-9, $N264$) and a long period (*prd*-1, $N272$) mutant. τ (A) and Ψ (B) of $N264$ progeny. τ (C) and Ψ (D) of $N272$ progeny.

Bibliography

[1] Ute Abraham, Adrián E Granada, Pål O Westermark, Markus Heine, Achim Kramer, and Hanspeter Herzel. Coupling governs entrainment range of circadian clocks. *Molecular systems biology*, 6(1):438, 2010.

[2] Keyur Adhvaryu, Ghazaleh Firoozi, Kamyar Motavaze, and Patricia Lakin-Thomas. Prd-1, a component of the circadian system of neurospora crassa, is a member of the dead-box rna helicase family. *Journal of biological rhythms*, 31(3):258–271, 2016.

[3] RJ Allen, Theodore R Rieger, and Cynthia J Musante. Efficient generation and selection of virtual populations in quantitative systems pharmacology models. *CPT: pharmacometrics & systems pharmacology*, 5(3):140–146, 2016.

[4] Jürgen Aschoff. Exogenous and endogenous components in circadian rhythms. In *Cold Spring Harbor symposia on quantitative biology*, volume 25, pages 11–28. Cold Spring Harbor Laboratory Press, 1960.

[5] Jürgen Aschoff and Hermann Pohl. Phase relations between a circadian rhythm and its zeitgeber within the range of entrainment. *Naturwissenschaften*, 65(2):80–84, 1978.

[6] Joseph Bass and Mitchell A Lazar. Circadian time signatures of fitness and disease. *Science*, 354(6315):994–999, 2016.

[7] Grigory Bordyugov, Ute Abraham, Adrian Granada, Pia Rose, Katharina Imkeller, Achim Kramer, and Hanspeter Herzel. Tuning the phase of circadian entrainment. *Journal of The Royal Society Interface*, 12(108):20150282, 2015.

[8] Alberto Bressan and 1968 Piccoli, Benedetto. *Introduction to the mathematical theory of control*. Springfield, MO : American Institute of Mathematical Sciences, 2007. Includes bibliographical references (p. [305]-309) and index.

[9] Veronica F Busa, Maxwell J Rector, and Rick Russell. The dead-box protein cyt-19 uses arginine residues in its c-tail to tether rna substrates. *Biochemistry*, 56(28):3571–3578, 2017.

[10] Antonio Cappuccio, Filippo Castiglione, Benedetto Piccoli, and Valerio Tozzi. Evaluation of hiv-1 and cd4+ t cell dynamic parameters in patients treated with genotypic resistance testing-guided haart. *Current HIV research*, 6(4):363–369, 2008.

[11] Karen S Conrad, Jennifer M Hurley, Joanne Widom, Carol S Ringelberg, Jennifer J Loros, Jay C Dunlap, and Brian R Crane. Structure of the frequency-interacting rna helicase: a protein interaction hub for the circadian clock. *The EMBO journal*, page e201694327, 2016.

[12] Chiarina Darrah, Bethan L Taylor, Kieron D Edwards, Paul E Brown, Anthony Hall, and Harriet G McWatters. Analysis of phase of luciferase expression reveals novel circadian quantitative trait loci in arabidopsis. *Plant physiology*, 140(4):1464–1474, 2006.

[13] Subhuti Dharmananda. Studies of the circadian clock of neurospora crassa: light-induced phase shifting. 1982.

[14] Christine Dubowy and Amita Sehgal. Circadian rhythms and sleep in drosophila melanogaster. *Genetics*, 205(4):1373–1397, 2017.

[15] Jeanne F Duffy and Charles A Czeisler. Age-related change in the relationship between circadian period, circadian phase, and diurnal preference in humans. *Neuroscience letters*, 318(3):117–120, 2002.

[16] Jeanne F Duffy, Derk-Jan Dijk, Elizabeth B Klerman, and Charles A Czeisler. Later endogenous circadian temperature nadir relative to an earlier wake time in older people. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 275(5):R1478–R1487, 1998.

[17] Jay C Dunlap and Jennifer J Loros. Making time: Conservation of biological clocks from fungi to animals. *Microbiology spectrum*, 5(3), 2017.

[18] Jillian M Emerson, Bradley M Bartholomai, Carol S Ringelberg, Scott E Baker, Jennifer J Loros, and Jay C Dunlap. period-1 encodes an atp-dependent rna helicase that influences nutritional compensation of the neurospora circadian clock. *Proceedings of the National Academy of Sciences*, 112(51):15707–15712, 2015.

[19] Martin Feinberg. The existence and uniqueness of steady states for a class of chemical reaction networks. *Archive for Rational Mechanics and Analysis*, 132(4):311–370, 1995.

[20] CM Friedrich. A model qualification method for mechanistic physiological qsp models to support model-informed drug development. *CPT: pharmacometrics & systems pharmacology*, 5(2):43–53, 2016.

[21] K Gadkar, N Budha, A Baruch, JD Davis, P Fielder, and S Ramanujan. A mechanistic systems pharmacology model for prediction of ldl cholesterol lowering by pcsk9 antagonism in human dyslipidemic populations. *CPT: pharmacometrics & systems pharmacology*, 3(11):1–9, 2014.

[22] Hugo Geerts, Athan Spiros, and Patrick Roberts. Assessing the synergy between cholinomimetics and memantine as augmentation therapy in cognitive impairment in schizophrenia. a virtual human patient trial using quantitative systems pharmacology. *Frontiers in pharmacology*, 6:198, 2015.

[23] Adrián E Granada, Grigory Bordyugov, Achim Kramer, and Hanspeter Herzel. Human chronotypes from a theoretical perspective. *PLoS One*, 8(3):e59464, 2013.

[24] Jeremy Gunawardena. A linear framework for time-scale separation in nonlinear biochemical systems. *PloS one*, 7(5):e36321, 2012.

[25] Norbert Gyöngyösi and Krisztina Káldi. Interconnections of reactive oxygen species homeostasis and circadian rhythm in neurospora crassa. *Antioxidants & redox signaling*, 20(18):3007–3023, 2014.

[26] Melanie J Hamblen-Coyle, David A Wheeler, Joan E Rutila, Michael Rosbash, and Jeffrey C Hall. Behavior of period-altered circadian rhythm mutants ofdrosophila in light: Dark cycles (diptera: Drosophilidae). *Journal of insect behavior*, 5(4):417–446, 1992.

[27] Christian Heintzen and Yi Liu. The neurospora crassa circadian clock. *Advances in genetics*, 58:25–66, 2007.

[28] Erik D Herzog, Sara J Aton, Rika Numano, Yoshiyuki Sakaki, and Hajime Tei. Temporal precision in the mammalian circadian system: a reliable clock from less reliable neurons. *Journal of biological rhythms*, 19(1):35–46, 2004.

[29] I Hosseini and F Mac Gabhann. Mechanistic models predict efficacy of ccr5-deficient stem cell transplants in hiv patient populations. *CPT: Pharmacometrics & Systems Pharmacology*, 5(2):82–90, 2016.

[30] Jennifer M Hurley, Luis F Larrondo, Jennifer J Loros, and Jay C Dunlap. Conserved rna helicase frh acts nonenzymatically to support the intrinsically disordered neurospora clock protein frq. *Molecular cell*, 52(6):832–843, 2013.

[31] Carl Hirschie Johnson. Phase response curves: what can they tell us about circadian clocks. *Circadian clocks from cell to human*, pages 209–249, 1992.

[32] Christopher R Jones, Scott S Campbell, Stephanie E Zone, Fred Cooper, Alison DeSano, Patricia J Murphy, Bryan Jones, Laura Czajkowski, and Louis J Ptček. Familial advanced sleep-phase syndrome: A short-period circadian rhythm variant in humans. *Nature medicine*, 5(9):1062, 1999.

[33] Krisztina Káldi, Beatriz Herreros González, and Michael Brunner. Transcriptional regulation of the neurospora circadian clock gene wc-1 affects the phase of circadian output. *EMBO reports*, 7(2):199–204, 2006.

[34] Tae-Sung Kim, Benjamin A Logsdon, Sohyun Park, Jason G Mezey, and Kwangwon Lee. Quantitative trait loci for the circadian clock in neurospora crassa. *Genetics*, 177(4):2335–2347, 2007.

[35] David J Klinke and Stacey D Finley. Timescale analysis of rule-based biochemical reaction networks. *Biotechnology progress*, 28(1):33–44, 2012.

[36] Akira Kohsaka, Aaron D Laposky, Kathryn Moynihan Ramsey, Carmela Estrada, Corinne Joshu, Yumiko Kobayashi, Fred W Turek, and Joseph Bass. High-fat diet disrupts behavioral and molecular circadian rhythms in mice. *Cell metabolism*, 6(5):414–421, 2007.

[37] Yoshiki Kuramoto. Cooperative dynamics of oscillator communitya study based on lattice of rings. *Progress of Theoretical Physics Supplement*, 79:223–240, 1984.

[38] Patricia L Lakin-Thomas, Stuart Brody, and Gary G Coté. Amplitude model for the effects of mutations and temperature on period and phase resetting of the neurospora circadian oscillator. *Journal of biological rhythms*, 6(4):281–297, 1991.

[39] Linda Lauinger, Axel Diernfellner, Sebastian Falk, and Michael Brunner. The rna helicase frh is an atp-dependent regulator of ck1a in the circadian clock of neurospora crassa. *Nature communications*, 5:3598, 2014.

[40] Kwangwon Lee, Jay C Dunlap, and Jennifer J Loros. Roles for white collar-1 in circadian and general photoperception in neurospora crassa. *Genetics*, 163(1):103–114, 2003.

[41] Jennifer J Loros and Jay C Dunlap. Genetic and molecular analysis of circadian rhythms in n eurospora. *Annual Review of Physiology*, 63(1):757–794, 2001.

[42] Phillip L Lowrey, Kazuhiro Shimomura, Marina P Antoch, Shin Yamazaki, Peter D Zemenides, Martin R Ralph, Michael Menaker, and Joseph S Takahashi. Positional syntenic cloning and functional

characterization of the mammalian circadian mutation tau. *Science*, 288(5465):483–491, 2000.

[43] Phillip L Lowrey and Joseph S Takahashi. Genetics of circadian rhythms in mammalian model organisms. In *Advances in genetics*, volume 74, pages 175–230. Elsevier, 2011.

[44] Sebastian Martinek, Susan Inonog, Armen S Manoukian, and Michael W Young. A role for the segment polarity gene shaggy/gsk-3 in the drosophila circadian clock. *Cell*, 105(6):769–779, 2001.

[45] Mark T Mc Auley, Darren J Wilkinson, Janette JL Jones, and Thomas BL Kirkwood. A whole-body mathematical model of cholesterol metabolism and its age-associated dysregulation. *BMC systems biology*, 6(1):130, 2012.

[46] Sean T McQuade, Ruth E Abrams, Jeffrey S Barrett, Benedetto Piccoli, and Karim Azer. Linear-in-flux-expressions methodology: Toward a robust mathematical framework for quantitative systems pharmacology simulators. *Gene Regulation and Systems Biology*, 11:1177625017711414, 2017.

[47] Martha Merrow, Michael Brunner, and Till Roenneberg. Assignment of circadian function for the neurospora clock gene frequency. *Nature*, 399(6736):584, 1999.

[48] Kiyoshi Onai and Hideaki Nakashima. Mutation of the cys-9 gene, which encodes thioredoxin reductase, affects the circadian conidiation rhythm in neurospora crassa. *Genetics*, 146(1):101–110, 1997.

[49] Kiran Padmanabhan, Maria S Robles, Thomas Westerling, and Charles J Weitz. Feedback regulation of transcriptional termination by the mammalian circadian clock period complex. *Science*, 337(6094):599–602, 2012.

[50] Bernhard Palsson. *Systems biology*. Cambridge university press, 2015.

[51] Satchidananda Panda. Circadian physiology of metabolism. *Science*, 354(6315):1008–1015, 2016.

[52] Sohyun Park and Kwangwon Lee. Inverted race tube assay for circadian clock studies of the neurospora accessions. *Fungal Genetics Reports*, 51(1):12–14, 2004.

[53] Alina Patke, Patricia J Murphy, Onur Emre Onat, Ana C Krieger, Tayfun Özçelik, Scott S Campbell, and Michael W Young. Mutation of the human circadian clock gene cry1 in familial delayed sleep phase disorder. *Cell*, 169(2):203–215, 2017.

[54] Andrew P Patton, Johanna E Chesham, and Michael H Hastings. Combined pharmacological and genetic manipulations unlock unprecedented temporal elasticity and reveal phase-specific modulation of the molecular circadian clock of the mouse suprachiasmatic nucleus. *Journal of Neuroscience*, 36(36):9326–9341, 2016.

[55] Violeta I Pérez-Nueno. Using quantitative systems pharmacology for novel drug discovery. *Expert opinion on drug discovery*, 10(12):1315–1331, 2015.

[56] Colin S Pittendrigh. Temporal organization: reflections of a darwinian clock-watcher. *Annual review of physiology*, 55(1):17–54, 1993.

[57] Colin S Pittendrigh, Walter T Kyner, and Tsuguhiko Takamura. The amplitude of circadian oscillations: temperature dependence, latitudinal clines, and the photoperiodic time measurement. *Journal of Biological Rhythms*, 6(4):299–313, 1991.

[58] Jeffrey D Plautz, Martin Straume, Ralf Stanewsky, Creston F Jamison, Christian Brandes, Harold B Dowse, Jeffrey C Hall, and Steve A Kay. Quantitative analysis of drosophila period gene transcription in living animals. *Journal of biological rhythms*, 12(3):204–217, 1997.

[59] Martin R Ralph and Michael Menaker. A mutation of the circadian system in golden hamsters. *Science*, 241(4870):1225–1227, 1988.

[60] Jan Rémi, Martha Merrow, and Till Roenneberg. A circadian surface of entrainment: varying t, τ, and photoperiod in neurospora crassa. *Journal of biological rhythms*, 25(5):318–328, 2010.

[61] Sanda Rocak and Patrick Linder. Dead-box proteins: the driving forces behind rna metabolism. *Nature reviews Molecular cell biology*, 5(3):232, 2004.

[62] Till Roenneberg, Serge Daan, and Martha Merrow. The art of entrainment. *Journal of Biological Rhythms*, 18(3):183–194, 2003.

[63] Till Roenneberg and Martha Merrow. The circadian clock and human health. *Current biology*, 26(10):R432–R443, 2016.

[64] Till Roenneberg and Walter Taylor. Automated recordings of bioluminescence with special reference to the analysis of circadian rhythms. 2000.

[65] Michael J Rust, Susan S Golden, and Erin K O'Shea. Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science*, 331(6014):220–223, 2011.

[66] DS Saunders, SW Gillanders, and RD Lewis. Light-pulse phase response curves for the locomotor activity rhythm in period mutants of drosophila melanogaster. *Journal of Insect Physiology*, 40(11):957–968, 1994.

[67] Christoph Schmal, Jihwan Myung, Hanspeter Herzel, and Grigory Bordyugov. A theoretical study on seasonality. *Frontiers in neurology*, 6:94, 2015.

[68] Brian J Schmidt, Fergal P Casey, Thomas Paterson, and Jason R Chan. Alternate virtual populations elucidate the type i interferon signature predictive of the response to rituximab in rheumatoid arthritis. *BMC bioinformatics*, 14(1):221, 2013.

[69] Peter K Sorger, Sandra RB Allerheiligen, Darrell R Abernethy, Russ B Altman, Kim LR Brouwer, Andrea Califano, David Z D'Argenio, Ravi Iyengar, William J Jusko, Richard Lalonde, et al. Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic mechanisms. In *An NIH white paper by the QSP workshop group*, volume 48. NIH Bethesda, MD, 2011.

[70] Alessandra Stangherlin and Akhilesh B Reddy. Regulation of circadian clocks by redox homeostasis. *Journal of Biological Chemistry*, 288(37):26505–26511, 2013.

[71] Kong L Toh, Christopher R Jones, Yan He, Erik J Eide, William A Hinz, David M Virshup, Louis J Ptáček, and Ying-Hui Fu. An hper2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science*, 291(5506):1040–1043, 2001.

[72] Dong Wang, Baoxiang Qin, Xiang Li, Ding Tang, Yu'e Zhang, Zhukuan Cheng, and Yongbiao Xue. Nucleolar dead-box rna helicase togr1 regulates thermotolerant growth as a pre-rrna chaperone in rice. *PLoS genetics*, 12(2):e1005844, 2016.

[73] David K Welsh, Seung-Hee Yoo, Andrew C Liu, Joseph S Takahashi, and Steve A Kay. Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Current Biology*, 14(24):2289–2295, 2004.

[74] Ying Xu, Quasar S Padiath, Robert E Shapiro, Christopher R Jones, Susan C Wu, Noriko Saigoh, Kazumasa Saigoh, Louis J Ptáček, and Ying-Hui Fu. Functional consequences of a ckiδ mutation causing familial advanced sleep phase syndrome. *Nature*, 434(7033):640, 2005.

[75] Shai Yerushalmi, Esther Yakir, and Rachel M Green. Circadian clocks and adaptation in arabidopsis. *Molecular ecology*, 20(6):1155–1165, 2011.

[76] Yusuke Yoshida, Hideo Iigusa, Niyan Wang, and Kohji Hasunuma. Cross-talk between the cellular redox state and the circadian system in neurospora. *PloS one*, 6(12):e28227, 2011.

[77] Tomasz Zielinski, Anne M Moore, Eilidh Troup, Karen J Halliday, and Andrew J Millar. Strengths and limitations of period estimation methods for circadian data. *PloS one*, 9(5):e96462, 2014.