Spatial Temporal Patterning and Dynamics of *E. Coli* Growth with Nutrient Variation

by

Changhan He

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2021 by the
Graduate Supervisory Committee:

Yang Kuang, Chair
Xiao Wang
Eric Kostelich
Xiaojun Tian
Abba Gumel

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

Synthetic biology (SB) has become an important field of science focusing on designing and engineering new biological parts and systems, or re-designing existing biological systems for useful purposes. The dramatic growth of SB throughout the past two decades has not only provided us numerous achievements, but also brought us more timely and underexplored problems. In SB's entire history, mathematical modeling has always been an indispensable approach to predict the experimental outcomes, improve experimental design and obtain mechanism-understanding of the biological systems.

Escherichia coli (*E. coli*) is one of the most important experimental platforms, its growth dynamics is the major research objective in this dissertation. Chapter 2 employs a reaction-diffusion model to predict the *E. coli* colony growth on a semi-solid agar plate under multiple controls. In that chapter, a density-dependent diffusion model with non-monotonic growth to capture the colony's non-linear growth profile is introduced. Findings of the new model to experimental data are compared and contrasted with those from other proposed models. In addition, the cross-sectional profile of the colony are computed and compared with experimental data.

E. coli colony is also used to perform spatial patterns driven by designed gene circuits. In Chapter 3, a gene circuit (MINPAC) and its corresponding pattern formation results are presented. Specifically, a series of partial differential equation (PDE) models are developed to describe the pattern formation driven by the MINPAC circuit. Model simulations of the patterns based on different experimental conditions and numerical analysis of the models to obtain a deeper understanding of the mechanisms are performed and discussed. Mathematical analysis of the simplified models, including traveling wave analysis and local stability analysis, is also presented and used to explore the control strategies of the pattern formation.

The interaction between the gene circuit and the host *E. coli* may be crucial and even greatly affect the experimental outcomes. Chapter 4 focuses on the growth feedback between the circuit and the host cell under different nutrient conditions. Two ordinary differential equation (ODE) models are developed to describe such feedback with nutrient variation. Preliminary results on data fitting using both two models and the model dynamical analysis are included.

To my Mom and Dad.

ACKNOWLEDGMENTS

I begin my thank you's, by acknowledging my advisor, Dr. Yang Kuang. It is my great pleasure to have such an incredible mentor, who has always provided guidance not only in research but also in principles and character. Dr. Kuang is very knowledgeable and patient, and always supportive of my new ideas. Working with him has greatly impacted many aspects of my life. I truly appreciate everything he has done for me.

I would also like to thank my committee members, Dr. Abba Gumel, Dr. Eric Kostelich, Dr. Xiao Wang, and Dr. Xiaojun Tian, and external reviewer, Dr. Kevin Flores, for their kind feedback and suggestions. Their involvement not only greatly improved my research and writing, but also brought me great suggestions on future directions. And a special thank you to Dr. John Nagy for his detailed comments during my thesis writing.

I am grateful to Dr. Xiao Wang and Dr. Xiaojun Tian, for their expertise and insight. I thoroughly enjoyed working with and learning from them. They are excellent collaborators and group leaders and every conversation with them was very inspiring.

I would like to thank the faculty members who have taught me here at ASU, I have learned a lot from you all. And I am especially thankful to the wonderful Graduate Program Coordinators here at the SoMMS department at ASU, Joelle Park and Jennifer May, for all that they do for the graduate students. I am also grateful for being part of the big ASU SoMSS community.

It is also my pleasure to meet and work with so many talented friends (Tin, Lifeng, Duane, Iboi, Elpi, Fan, Dylan, Penny, Fuqing, Juane, Samat, etc), who make my life in ASU so enjoyable.

I am also grateful as well to Dr. Xing Liang and Dr. Jianhong Wu, for guiding me towards the area of Math Biology during my undergraduate period.

I would like to express enormous appreciation and gratitude to my parents for their endless love and support, and for encouraging me to pursue my own interests, and for instilling an interest in science and problem-solving from a young age.

Lastly, but certainly not least, I would wholeheartedly like to thank my girlfriend, Fei, for her love and support.

I gratefully acknowledge the NIH grant 5R01GM131405 and Arizona State University for their financial support.

TABLE OF CONTENTS

Page

CHAPTER

APPENDIX

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Brief Introduction of Synthetic Biology

What is life? How does life evolve? These fundamental philosophical questions are also scientifically important. The curiosity about life inspired people to understand, transform and even redesign biological systems, which can be traced to the study of domestication and selective breeding by Charles Darwin in the 19th century. Meanwhile, biologists were recognizing the significance of cell theory due to the advances in microscopy. And later in the 20th century, some remarkable work has been made on cellular and molecular networks. The study of Francois Jacob and Jacques Monod on the *lactose* operon in *E. coli* led to a landmark publication in synthetic biology in 1961, which posited the existence of regulatory circuits that underpin the response of a cell to its environment (Monod and Jacob, 1961; Cameron *et al.*, 2014). In the early 2000s, more powerful genetic engineering capabilities and decreased experimental costs enabled scientists to study, redesign and build complex biological entities and systems, which forms a more clear area of synthetic biology. And during the past two decades, synthetic biology has undergone considerable growth in many directions, including synthetic gene circuit (Hasty *et al.*, 2002; Potvin-Trottier *et al.*, 2016), metabolic engineering (Martin *et al.*, 2003), multicellular pattern formation (Basu *et al.*, 2005; Asllani *et al.*, 2020) and synthesized genome (Gibson *et al.*, 2010). Such growth also greatly filled the gap between synthetic biology and biotechnology applications.

With the advances in technologies, scientists can now ask, study and examine questions that directly come from real life. For example, immunotherapies, as one of the most important cell-based therapies for fighting cancers, are mostly rely on engineering T cells. Recent studies showed that gene switches could enhance the safety of the treatment by terminating the therapy when severe side effects were detected (Jones *et al.*, 2014), or reactivating the system when needed (Wei *et al.*, 2012). Meanwhile, more and more companies were putting efforts into turning the experimental achievements into civil technologies. Seven years after the success of synthesizing 9.6 kbp (kilobase pairs) Hepatitis C virus genome in 2000 (Blight *et al.*, 2000), several companies developed the ability to synthesize up to 2000 bp genetic sequences in two weeks with affordable prices (Pollack, 2007).

As a highly multidisciplinary area, synthetic biology is also dependent on approaches that come from various subjects such as chemistry, mathematics, and computer science. And mathematical modeling has always been an indispensable way to study the mechanisms behind phenomena (Menn and Wang, 2019; Tian *et al.*, 2019; Zheng and Sriram, 2010). Models that are able to correctly predict the behavior of a system allow engineers to program new cellular behavior without having to perform large numbers of trial-and-error experiments (Chandran *et al.*, 2008). Many mathematical models have been published that are capable of reasonably describing biological systems and providing comparable simulation results to experimental observations, especially in the study of pattern formation. Basu *et al.* (2005) introduced an ordinary differential equation-based mathematical model to describe a synthetic multicellular system that could form ring-like patterns through fluorescent proteins. The model presented the oscillated expression of specific genes and ring formation dynamics with different sets of parameters. Liu *et al.* (2011) developed a reaction-diffusion model with chemotactic movement to present sequential strip patterns generated by *E. coli* colony. For more complicated patterns, reaction-diffusion models are always considered a powerful way to simulate and analyze the mechanisms, such as modeling the generation of flower-like bacteria colonies (Kawasaki *et al.*, 1997).

Although the modeling approach has become a significant way to improve our understanding of biological phenomena and reveal the mechanisms, there are still numerous underexplored problems. Our work is focused on the areas of bacteria pattern formation and circuit-host interactions. We continue to pursue the mechanisms of pattern formation by modeling and comparing engineered gene circuits. We also aim for providing a modeling framework for the circuit-host interactions with the understanding that the future results will help the experimental design and control. We believe that the data-verified mathematical model with a reasonable description of the biological system can provide a significant contribution to synthetic biology while targeting both the experimental design and the mechanisms.

1.2 Mathematics in Synthetic Biology

1.2.1 Cell Dynamics in Escherichia coli

Escherichia coli (*E. coli*) is a common bacteria abundant in human intestine. It is also one of the most studied bacteria and has been an excellent platform for biological research due to its simplicity and relatively low-cost in experiment. *E. coli* can move itself by rotating its flagella counter-clockwise or clockwise (Block *et al.*, 1983). More particularly, the bacterium "runs" in a straight line at constant velocity when all the flagellar motors rotate counterclockwise, and the bacterium "tumbles" without advancing when one or several or these motors change direction (Saragosti *et al.*, 2012). Since this process has high randomness, the movement of a large population of bacteria can be modeled by heat equation:

$$\frac{\partial N}{\partial t} = D\nabla^2 N, \tag{1.1}$$

where N represents the bacterial density and D is the diffusion coefficient.

In a nutrient gradient, movement of a bacterium can be more complicated since it will move preferentially towards higher concentration of the nutrient (Sourjik and Wingreen, 2012). For bacteria growing on an agar plate, this can be presented mathematically by introducing a chemotactic flux term (Leyva *et al.*, 2013):

$$J_c = \sigma n b^2 \chi(n) \nabla n, \tag{1.2}$$

where $n = n(x, y, t)$ and $b = b(x, y, t)$ represent the nutrient concentration and bacterial density, respectively. Parameter σ can be regarded as a constant measuring the hardness of the agar medium and $\chi(n)$ is the chemotactic sensitivity function with the form introduced by Lapidus and Schiller (1976):

$$\chi(n) = \frac{\chi_0 K_d}{(K_d + n)^2}, \tag{1.3}$$

where $\chi_0 > 0$ is a constant measuring the strength of the chemotaxis, and $K_d > 0$ is the receptor-ligand binding dissociation constant.

The combination of random walk-based diffusion and chemotaxis provides a comphrehensive discription of bacterial movement, and has already been widely used in the research of bacterial pattern formation (Brenner *et al.*, 1998; Leyva *et al.*, 2013; Li *et al.*, 2020). However, bacterial chemotaxis is more common in liquid or other soft media, where both the outermost and inner bacteria can move actively. When growth is occuring on a harder agar surface, only the outermost part of the colony grows by cell division and population mass increases, the inner bacteria cells are inactive (Ohgiwari *et al.*, 1992). Therefore, when considering bacterial growth on semi-solid agar , chemotaxis becomes negligible while the colony growth can still be appropriately presented (Kawasaki *et al.*, 1997; Satnoianu *et al.*, 2001).

As an important research objective and experiment platform, the dynamics of *E. coli*, including cell growth and colony morphogenesis, have always been a fundamental but pivotal topics. Mathematical models with solid theoretical foundation and testable hypotheses are useful tools to further our understanding on this area.

1.2.2 Modeling Synthetic Gene Circuit

Inspired from computer science and electronics, synthetic gene circuits have been designed to exhibit control over the flow of information in biological systems (Balagaddé *et al.*, 2008; Xiang *et al.*, 2018). This application uses specific genes with designed topology to perform logical functions, and has become an important method to modify cellular functions (Ausländer and Fussenegger, 2016), create cellular responses to environmental conditions (Antunes *et al.*, 2006), or influence cellular development (Prochazka *et al.*, 2017). Synthetic gene networks have evolved from simple proof-of-concept circuits to complex therapy-oriented networks over the past 20 years, and will accelerate the process towards broad practical applications (Wu and Wang, 2015).

Mathematical modeling complements the toolbox of synthetic biology by predicting and exploring the behavior of unbuilt gene circuits and multicellular systems (Klumpp *et al.*, 2009; Wang *et al.*, 2016). A representative modeling work is introduced by Gierer and Meinhardt (1972), which contains a short-range activator and a long-range inhibitor (Figure 1.1). This topology appears to be crucial for many important types of pattern formation and morphogenesis:

$$\frac{\partial a}{\partial t} = \rho \frac{a^2}{h} - \mu_a a + D_a \frac{\partial^2 a}{\partial x^2} + \rho_a, \tag{1.4a}$$

$$\frac{\partial h}{\partial t} = \rho a^2 - \mu_h h + D_h \frac{\partial^2 h}{\partial x^2} + \rho_h. \tag{1.4b}$$

Here $a(x,t)$ and $h(x,t)$ represent the concentration of activator and inhibitor at position x and time t, respectively, where $\{(x,t) : x \in \mathbb{R},\ t \in \mathbb{R}^+\}$. Parameters μ_a and μ_h are the decay rates, while ρ_a and ρ_h are the self-activation rates. The first term on the right of (1.4a) describes the poduction of the activator which is also slowed down

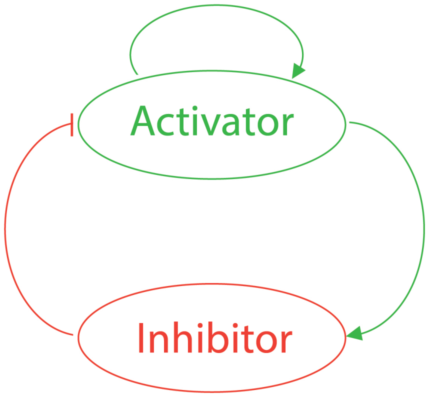

Figure 1.1: Topology of the gene circuit introduced by Gierer and Meinhardt (1972), which contains a short-range activator and a long-range inhibitor.

by the inhibitor, and the first term on the right of (1.4b) describes the production of inhibitor as a function of the activator concentration. This model also assumes that both the activator and inhibitor have constant diffusion rates D_a and D_h.

The Gierer-Meinhardt model is a fundamental work in modeling biological pattern formation. It provides us a bridge between the experimental observations and the molecular-genetic mechanisms and its rich dynamic properties have been well studied (Ruan, 1998; Song *et al.*, 2017). However, it is certainly a minimal model which described a very simple topology; either the gene circuit or the interactions between circuit components could be a lot more complicated.

The high complexity and sensitivity of the gene circuit make remaining experimental research challenging; thus developing suitable mathematical models can allow systematic description and analysis of such biological systems and greatly deepen our understanding on circuit functions and control strategies.

1.3 Two Modeling Approaches

A mathematical model in synthetic biology reseach which can be regarded as "a good model" always comes with the following two characteristics:

- The model can systematically describe the biological system and provide comparable simulation results.
- Dynamical properities can be explored via mathematical analysis of the model, which can help us understand the mechanisms.

However, accomplishing the above two requirements is challenging because numerical simulation and mathematical analysis are opposite in some sense. For example, a simple model with fewer equations and parameters will certainly be easier to analyze its dynamical properties, but may not be able to simulate experimental observations. On the other hand, a model with high complexity will be more powerful on numerical simulations while its mathematical analysis becomes rather difficult (Figure 1.2).

Admittedly, there is no universal way to solve such a dilemma and we need to trade-off based on specific questions we are facing. When modeling the expression of a synthetic gene circuit, we should systematically present the whole circuit which could lead to a high-complexity model. Then the mathematical analysis can proceed with reasonable simplifications. Otherwise, when we are not clear about the biological principles, such as the interactions between gene circuit and the host cell, we may start from a preliminary model which can capture experimental observations before taking other factors into account.

Even so, there are still many questions that need to be answered when understanding pattern formation. What are the necessary conditions for the gene circuit to generate stipe patterns? Will the bacteria colony expanding rate affect pattern formation? Does the interaction between the gene circuit and host cell play an important role during the generation of patterns? If appropriate models can be built to address these questions, they can also be used to improve our experimental design and predict results.

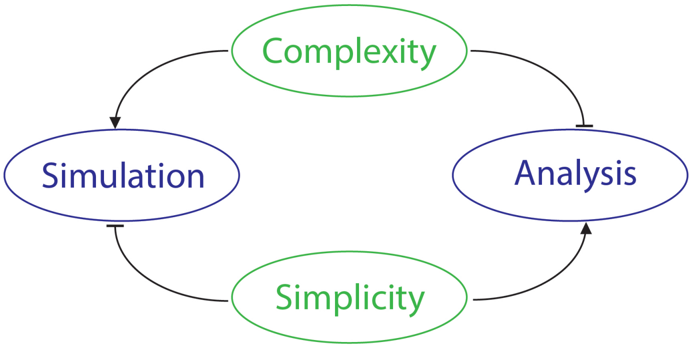

Figure 1.2: Schematics of the challenge of balancing the model complexity and simplicity.

Since some parts of my thesis have already been published, or are in the process of being published, and in accordance with Arizona State University policy, I will provide a detailed introduction about the contributions I made in each of these papers and projects. In addition, I will discuss the contributions made by each co-author and the unique attributes I made as well as my interpretation of the results.

1.4 A modeling framework on predicting *E. coli* colony growth

The article published in this section (He *et al.*, 2020), is an article in which I am the first author. This project concerns how bacterial colony growth is affected by multiple control factors such as temperature and amino acid concentration. The initial questions about why bacterial colonies will always reach a stable size and how they can be controlled were raised by Dr. Xiao Wang based on previous experimental observations. The corresponding experiments were designed and performed by Samat Bayakhmetov and Dr. Xiao Wang, which also verified that bacterial colony does not have a linear growth profile under different temperature or amino acid concentration. Previous mathematical models either only consider one control factor or not be able to capture the nonlinear growth profile. Our paper introduced a modeling framework for describing bacterial colony growth with multiple control factors. The one-dimensional reaction-diffusion-based model, which we introduced in this paper, is modified from Fisher's equation. Dr. Yang Kuang suggested that the density-dependent diffusion rate could be plausible to model the colony expansion. We considered a nonmonotonic density-dependent diffusion rate based on experimental observations, which showed that cells have low mobility when cell density is either too low or too high. We also introduced a density-dependent inhibition to the reaction term which mainly affects the low-density region, which was also inspired from the experimental observations that cell may not reach its maximum reproduction rate right after seeding on the agar plate. We then numerically analyze the model by fitting the experimental data on both colony radius and cross-sectional profiles. Our model provided much better fitting results for most of the cases compared to Fisher's equation and captured the general shape of the colony edge at different time points.

My part in this project was the model establishment, image processing, and part of numerical analysis. I developed a model with density-dependent diffusion rate and nutrient-dependent reaction rate according to Dr. Kuang's suggestions. And then Dr. Xiao Wang pointed out that there was sufficient nutrient throughout the experiment, and we were not clear about how the nutrient distribution would change, so I revised the model by replacing the nutrient limitation with a density-dependent inhibition, and only the revised model was included in the paper. I also used *MATLAB* to process the experimental images such as measuring the radius of the colony and plotting its cross-sectional profile based on the greyscale data. Duane Harris and I also estimated the parameters of our model in order to generate the best fit for the experimental data.

I wrote the majority of the paper, excepting the sections on experimental setup, imaging and microscopy (both written by Samat Bayakhmetov), and parameter estimations (written by Duane Harris). All authors assisted in proofreading and improving the paper.

1.5 Reaction-diffusion-based Mathematical Modeling on Circuit-driven Bacterial Patterning

The article in this section is an article in which I will be the second author, currently in preparation for submission. This multidisciplinary project included designing a synthetic gene circuit (MINPAC) and performing corresponding circuit-driven pattern formation experiments, processing the experimental microscopy images and developing mathematical models to describe the system, simulating the pattern formation process using our model, and quantitatively analyzing the system behaviors.

Because this work includes multidiscipline, there are many co-authors involved in this paper. I will specifically highlight the aspects of this project that I performed. For the experimental side, Dr. Fuqing Wu and Dr. Xiao Wang designed the MINPAC gene circuit and Dr. Fuqing Wu performed the experiments and collected microscopy images. Thai Ohnmacht, Qi Zhang, Xingwen Chen, performed the early testing of pattern formation using different agar concentrations; Dr. Kyle Allison, supervised single-cell imaging work for pattern formation. Dr. Fuqing Wu also performed control experiments for the modified MINPAC circuit in order to compare the function of the circuits. Samat Bayakhmetov repeated the experiments and Xingwen Chen performed the experiments estimating the diffusion rates of autoinducers. Dr. Xiao Wang oversaw the experimental portion of this project. I was involved in the processing of microscopy images, developing mathematical models, performing the simulations, and analyzing the model both numerically and mathematically.

In the image processing part, I measured the colony size via *MATLAB* and estimated the concentrations of fluorescent proteins within the colony. With the help of Duane Harris, we also measured the ring width via *MATLAB* to determine the number of rings.

Dr. Yang Kuang, Dr. Javier Baez, and Dr. Fuqing Wu initially proposed a reaction-diffusion-based model for the MINPAC system. I modified the solver which was initially written by Dr. Javier Baes in *MATLAB* and performed simulations for all sets of experiments, which agreed well with observations. I then numerically solve the system and compared the wave speed with the experimental colony expansion speed. Since the number of rings varied even with the same experimental conditions, Dr. Yang Kuang and Dr. Xiao Wang suggested that the randomness of the initial conditions could be the key factor. I then numerically analyzed the model dynamics and showed that the variation of a number of rings could be explained by the difference of the Poincare return time which depends on the initial points. Furthermore, Dr. Xiao Wang pointed out that the MINPAC circuit is composed of two topologically equivalent motifs. According to Dr. Xiao Wang's suggestion, I performed model-comparison simulation results showing that the MINPAC circuit harbors a greater robustness and amplitude against parameter perturbations to generate temporal oscillation. We then modified the model according to the control experiments. And similarly, I provided corresponding simulations which show that our models are plausible for all the control systems with reasonable modifications.

Since the model contains six equations, reasonable simplification is necessary before analyzing the model. Based on the experimental data, we assumed that the diffusion process of autoinducers through the membrane is much faster than the circuit expression and pattern formation. Thus we obtained a reduced model with four equations, and I then analyzed the equilibrium points and local stability of the system. With further simplification via quasi-steady-state assumption, I was able to provide traveling wave analysis for the reaction-diffusion system. Dr. Yang Kuang assisted me with the mathematical analysis part and noting that it was not included in the paper.

Overall, I initially wrote the majority of the mathematical modeling section in the supplementary material of the paper. All authors proofread and edited the paper as necessary.

1.6 Nutrient-modulating *E. Coli* Growth Feedback in Multi-nutrient Culture

The article published in this section (Melendez-Alvarez *et al.*, 2021), is an article in which I am the second author. This project studied the nutrient-modulating growth feedback from the gene circuit to the host cell. We first reported an unexpected damped oscillatory behavior of a self-activation gene circuit induced by nutrient-modulating growth feedback, then we developed a molecular mathematical model with nonmonotonic growth-rate regulation on gene production rate to describe such dynamics. The corresponding experiments were designed and performed by Juan Melendez-Alvarez, Dr. Rong Zhang, and Dr. Xiaojun Tian. The mathematical model was modified based on a previous paper (Zhang *et al.*, 2020). We used our model to fit experimental data and mathematically analyzed the rationality of the model. In addition, a model that includes the ribosome allocation toward gene production, cell growth, and cell maintenance was introduced and analyzed by Juan Melendez-Alvarez. Since I was not involved in this part, I did not include this portion of the work in my thesis.

My part in this project was the numerical and mathematical analysis of the model. I modified the program which was originally written by Juan Melendez-Alvarez. Inspired by the experimental data, Dr. Yang Kuang suggested that the nutrient-modulating growth feedback can be described by a growth-dependent nonlinear function. I then numerically tested and tuned the model, finding that the model could provide comparable fitting results of the experimental data by considering a non-monotonic growth-rate regulation. I then mathematically analyzed the model and showed that there were certain requirements for modeling the growth feedback.

Overall, I initially wrote the mathematical analysis part in supporting the information of the paper. All authors proofread and edited the paper as necessary.

1.7 Overview of the Thesis

In the following chapters, I will explore the spatiotemporal patterning and dynamics of *E. coli* growth through mathematical modeling and relevant analysis. By comparing our modeling results with experimental data, we will once more see the significance of a modeling approach. Chapter 2 introduces a modeling framework for describing *E. coli* colony growth under multiple control factors, which provides us a potential path to further investigate which are the most important control factors and how to mathematically describe the impact. Chapter 3 contains a detailed study on circuit-driven bacterial pattern formation. Our modeling approach shows that a plausible mathematical model can not only systematically describe the biological system and predict the experimental results, but also be a great platform to explore and explain the mechanisms. Meanwhile, the interaction between the synthetic gene circuit and the host cell can also be crucial and have direct impact on cell dynamics as well as circuit function. The corresponding work is included in Chapter 4. And the overall conclusions and discussion are reserved for Chapter 5.

Chapter 2

A MODELING FRAMEWORK ON PREDICTING *E. COLI* COLONY GROWTH

2.1 Abstract

Bacterial colony formations exhibit diverse morphologies and dynamics. A mechanistic understanding of this process has broad implications to ecology and medicine. However, many control factors and their impacts on colony formation remain underexplored. Here we propose a reaction-diffusion-based dynamic model to quantitatively describe cell division and colony expansion, where control factors of colony spreading take the form of nonlinear density-dependent function and the intercellular impacts take the form of density-dependent hill function. We validate the model using experimental E.coli colony growth data and our results show that the model is capable of predicting the whole colony expansion process in both time and space under different conditions. Furthermore, the nonlinear control factors can predict colony morphology at both center and edge of the colony.

2.2 Introduction

The process of bacterial colony formation provides an excellent experimental platform to study a broad range of biological phenomena such as biofilm (Hoffman *et al.*, 2005), multicellular interactions (Shapiro, 1995), and pattern formation (Basu *et al.*, 2005; Liu *et al.*, 2019). All of these processes are intimately intertwined with bacterial growth, so its control is an essential tool in regulating other processes within the colony. The formation of a bacterial colony cannot be simply regarded as the outcome of cell reproduction. The mechanisms behind it can be rather complicated and several recent works have provided us with valuable insights. For instance, Warren *et al.* (2019) thoroughly explained all of the forces and provided quantitative data on key elements that affect the spatiotemporal establishment of bacterial colonies on hard agar. They also were the first to emphasize the role of surface tension in radial cell expansion. In another study, Su *et al.* (2012) points out that the balance in interplays of the forces has a profound impact on the overall morphology of the colony.

All these control factors will influence the colony formation process and it can be difficult to determine which is the dominated one, since the dominated control factor may vary under different experiment conditions, during different formation stages, or even at different locations within the colony. Although under ideal growing conditions, some control factors may not play a drastic impact in a short period of time (such as nutrient limitation), they usually become unneglectable in long-term experiments. Therefore, focusing on one single control factor may not provide us a comprehensive analysis of the whole process. We shall be aware of the overall impact on the bacterial colony formation and start from a plausible, basic mathematical model before modeling multiple factors at the same time.

In this paper, I proposed a novel mathematical model with nonlinear density-dependent diffusion rate and growth rate to describe the colony growth under different control strategies. Corresponding experiments of *E. coli* colony growth on semi-solid agar under different conditions were designed and operated by Samat Bayakhmetov (Xiao Lab). The model is capable of describing the non-constant diffusion of the cells and the non-monotonic relationship between cell reproduction rate and cell density. Using this dynamic model, I not only compared the quantitative results with experimental data under different control strategies but also examined how the model performs on capturing the colony side spatial profile. Some of the work contained in this chapter has been previously published.

2.2.1 Previous Models

Although the morphology of bacterial colonies can be clearly observed in the experiments and certain control factors, including temperature and nutrient concentrations have been proposed (Shapiro, 1995; Basu *et al.*, 2005; Warren *et al.*, 2019; Su *et al.*, 2012), determining the key factors and study the mechanisms remain challenging. During the past two decades, many mathematical models were developed to describe bacterial colony growth and formation, and most of them were established based on Fishers equation (Mimura *et al.*, 2000; Kawasaki *et al.*, 1997; Leyva *et al.*, 2013).

Wakita *et al.* (1994) first used Fisher's equation to describe the colony formation:

$$\frac{\partial b}{\partial t} = D_b \nabla^2 b + (\varepsilon_0 - \mu b) b \tag{2.1}$$

here $b = b(r, t)$ represent the bacterial population density at time t at the spatial position r, ε_0 is the maximum growth rate, D_b is the diffusion coefficient and μ presents the competition among bacterial cells. The authors concluded that Fisher's equation is suitable for describing simple and homogeneously spreading colony patterns, which usually refers to soft agar medium with rich nutrients (Mimura and Tsujikawa, 1996).

Later in 2000, Mimura *et al.* (2000) introduced another reaction-diffusion model with the assumption that bacteria consist of active cells and inactive cells when modeling bacterial colony patterns on a semi-solid agar medium. Here the active cells will move, grow and perform cell division while inactive cells do nothing at all. Meanwhile, the authors also considered nutrient concentration as another variable. If we let $u(x, t)$ and $v(x, t)$ represent the density of active cells and nutrient concentration at time t and position x, respectively, then the model is:

$$\frac{\partial u}{\partial t} = d\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) + uv - a(u,v)u \tag{2.2a}$$

$$\frac{\partial v}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\partial v}{\partial x}\right) - uv \tag{2.2b}$$

where d is the diffusion coefficient of bacteria and $a(u,v)$ can be regarded as the conversion rate from active cells to inactive cells, which is depend on both the active bacterial density and nutrient concentration.

Furthermore, a chemotaxis model to describe the two-dimensional colony pattern formation was developed by Leyva *et al.* (2013).

$$n_t = D_n \bigtriangleup n - \kappa nb \tag{2.3a}$$

$$b_t = \nabla(\sigma nb\nabla b) + \theta\kappa nb - \nabla\left(\theta nb^2 \frac{\chi_0 K_d}{(K_d+n)^2}\nabla n\right) \tag{2.3b}$$

where $n = n(x,y,t)$ and $b = b(x,y,t)$ represent the concentration of the nutrient and the density of the bacterial cells, respectively. D_n is the diffusion coefficient of nutrient, κ is the nutrient consumption rate and θ is the nutrient intake rate. $\sigma = \sigma_0(1+\nabla)$ represents the combine effect of agar hardness (σ_0) and random diffusion of the cell (∇). Moreover, the chemotactic sensitivity function $\chi(n)$ takes the form of "receptor law" proposed by Lapidus and Schiller (1976):

$$\chi(n) = \frac{\chi_0 K_d}{(K_d+n)^2} \tag{2.4}$$

where $\chi_0 > 0$ is a constant measuring the strength of the chemotaxis and $K_d > 0$ is the receptor-ligand binding dissociation constant and represents the nutrient level needed for half of the receptor to be occupied. The authors analyzed the impact from both the nutrient concentration and the chemotactic sensitivity and presented model prediction of different colony patterns growing under different chemotactic sensitivity.

It is worthy to point out that although the above models are well-known and widely used, they do have their own limitations. Firstly, these models are mainly used to predict the colony patterns, which focus on predicting the final shape of the colony rather than the colony growth process. Secondly, both Wakita *et al.* (1994) and Mimura *et al.* (2000) assumed constant bacterial diffusion rate when modeling colony formation on semi-solid agar medium, which is not consistent with our experimental data showing that the colony grows at a faster speed in the first 48 hours and slows down afterward (Figure 2.1). Thirdly, neither the Logistic-form growth term introduced by Wakita *et al.* (1994) nor the linear growth term presented by Mimura *et al.* (2000) and Leyva *et al.* (2013) may not be the best way to describe cell reproduction. Based on experimental observations, the cell barely grows in the first couple hours after seeding on the agar plate, which indicates that the cell reproduction rate may not be a monotonic function of cell density. Detailed discussion will be provided in section 2.4. In addition, the model developed by Leyva *et al.* (2013) is a good example to predict the colony shape but may overestimate the impact of chemotactic and did not consider other possible factors that can affect cell reproduction rate, such as temperature.

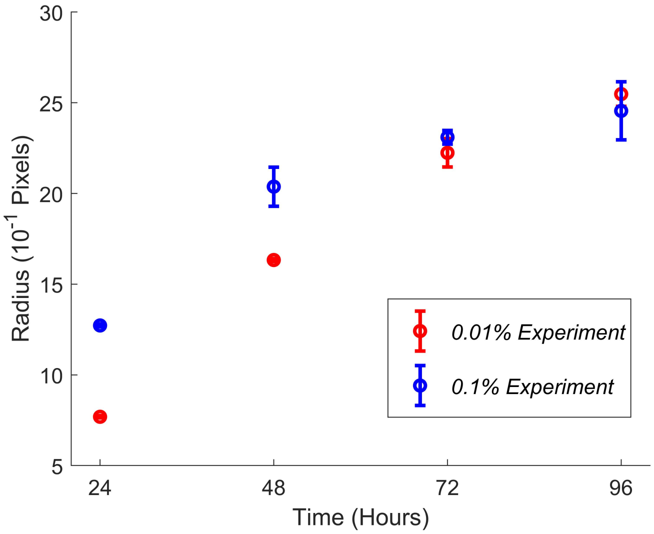

Figure 2.1: Experimental data of colony radius at different time with different AA. Red and blue colors represent normal (0.01%), and high (0.1%) AA, respectively.

2.3 Experimental Design

2.3.1 Experiment Introduction

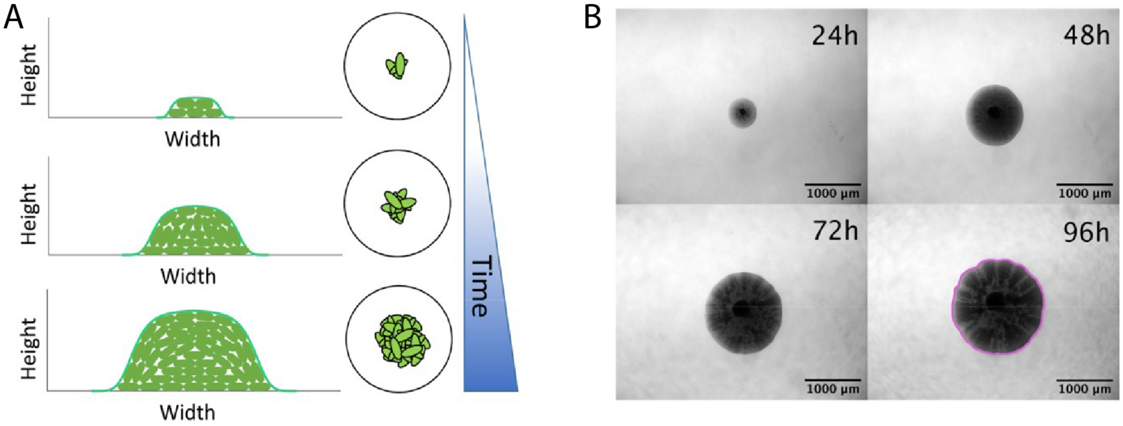

Figure 2.2: Schematics and experimental observations of the growth and morphology of a bacterial colony - Figure taken from He *et al.* (2020). **(A)** Schematics of a bacterial colony growth and its cross-sectional profile. Cells are indicated by small green bars and black circles indicate the edges of an agar plate. Left column: cross-sectional profiles of a bacterial colony at various time. **(B)** Phase contrast images of an *E. coli* colony growing on a semi-solid agar plate taken every 24 hours after seeding ($t = 0$). Darker color implies higher cell density and the pink trace at 96h colony indicates the edge detected by computer program.

To describe the morphology and dynamics of a growing bacterial colony, we focus on its radial and vertical growth as illustrated by schematics in Figure 2.2A. The colony begins to grow after seeding, eventually taking on a circular shape when viewed from above and resembling a semielliptical arch when viewed from the side. Microscopy images (Figure 2.2B) show that a clear circular pattern can usually be observed 24 hours after seeding. From here, a colony continues to grow in size while maintaining its circular shape, which is consistent with previous findings (Warren *et al.*, 2019). We observe that a colony usually has a grey edge and a dark core in the earliest images, which are taken at the 24th hour. As the colony grows, the grey edge becomes barely visible as the size of the dark core increases. The strong contrast between the colonies and the agar surfaces indicates that the colonies have sharp edges.

There are two sets of experiments performed by Samat Bayakhmetov (Xiao Lab). The first set of experiments were performed at two different T, 30°C and 37°C. Each colony was inoculated as a single cell growing on a semi-solid agar plate for up to 4 days. The second set of experiments was performed with two different AA, 0.01% and 0.1% with similar procedures. The clear circular patterns could be observed after 24 hours, so the first data point in each experiment was collected at the 24th hour and then every 24 hours. There were four colonies collected in the first set experiment (with different T) and six colonies collected in the second set experiment (with different AA concentrations).

2.3.2 Image Processing

A *MATLAB* program was created to detect the colony edge (Figure 2.3 Step 1) using these phase contrast images. Cropping and adjusting the contrast of the images may need to be denoise. Then the colony center and the radius were found by fitting the detected colony edge with a circle function (Figure 2.3 Step 2). This provides an efficient and accurate method to measure its size. In each set of experiments, the average radii of all colonies were considered as the final data point. In addition, since the greyscale data can be used as a reporter of the cell density, the cross-sectional profile of the colony can be presented by plotting the greyscale data from the colony center 2.3 Step 3 and 4).

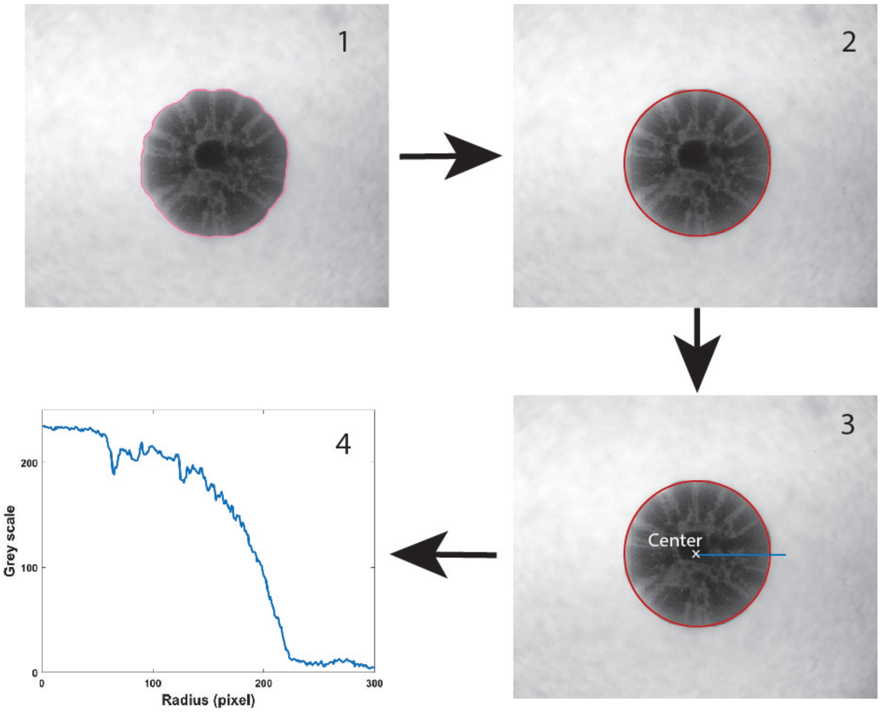

Figure 2.3: Schematics of the image processing. Step 1: Detecting the colony edge by using the *MATLAB Image Processing Toolbox*. The pink trace indicates the detected colony edge. Step 2: Fitting the pink trace by a circle function, which can determine the radius and center of the colony. Step 3 and 4: Plotting the grayscale data of the phase contrast image starting from colony center to the edge, which provides the cross-sectional profile of the colony.

2.4 Model Introduction

2.4.1 Fisher's Equation

The growth of *E. coli* colonies on homogenous semi-solid agar consist of reproduction and expansion, which always presents circular patterns and can be systematically modeled by one-dimensional reaction-diffusion equations. Many mathematical models were developed upon Fishers equation during the past two decades, which takes the following form:

$$\frac{\partial N}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial N}{\partial x}\right) + \mu\left(1 - \frac{N}{K}\right)N \tag{2.5}$$

where $N = N(x,t)$ is the cell density at time t and spatial position x (distance to colony center), which can also be regarded as the colony height. D is diffusion coefficient which describe the speed of colony expansion, μ is the maximum reproduction rate and K is the maximum cell density.

2.4.2 Model with Nutrient

The bacteria colony growth is controlled by heterogeneous factors coming from a cells metabolism and external environment. Recent research shows that cell reproduction can be affected by nutrient concentration (Matz and Jürgens, 2003) and acetate accumulation (Shiloach and Fass, 2005), while colony expansion on the agar surface can be slowed down by physical friction (Warren *et al.*, 2019), especially when cell density is high. Based on experimental observations, an *E. coli* colony grows at a non-constant speed, which is consistent with previous findings that colony expansion of mobile cells is density-dependent (Liu *et al.*, 2011; Fu *et al.*, 2012). Particularly, the diffusion process of *E. coli* is very slow at the beginning since they have plenty of space and nutrients, and a colony always has a faster expansion rate during the 24th to 72nd hour (Figure 2.1), which implies that treating cell diffusion rates as constant may be an oversimplification and misleading since constant diffusion rate will lead to a linear colony growth profile. Therefore, presenting the diffusion rate as a non-monotonic function of cell density N becomes reasonable, noticing that density-dependent diffusion has been introduced and validated in relevant fields such as modeling Glioblastoma growth (Stepien *et al.*, 2015) and modeling *E. coli* growth (Liu *et al.*, 2011). Here we describe the diffusion rate using a nonlinear density-dependent function:

$$D(N) = \frac{a_1 N^{m_1}}{N^{m_2} + 1} \tag{2.6}$$

with the assumption $0 < m_1 < m_2$ which ensures the diffusion is slow when the cell density is either too low or too high.

We assume phosphorous (denoted as P) to be the most limiting nutrient with an initial concentration $P(0) = P_0$, and g represents the cell quota. The concentration of phosphorous will constantly decrease and evert adverse influence on a cell's reproduction, which is modeled by $P^{n_1}/(P^{n_2}+h)$ and h is the Michaelis constant. Meanwhile, since the larger colony and higher cell density will result in stronger physical friction and lower nutrient uptake rate, we propose a function

$$f(x) = \frac{1}{1+c_1 x} \tag{2.7}$$

to describe the inhibition of cell growth caused by this interaction where x is the distance to the colony center. Based on the Logistic growth model with maximum cell density K, we can derive the following reaction-diffusion equation:

$$\frac{\partial N}{\partial t} = \underbrace{\frac{\partial}{\partial x}\left(\frac{a_1 N^{m_1}}{N^{m_2}+1}\frac{\partial N}{\partial x}\right)}_{\text{cell diffusion}} + \underbrace{\frac{P^{n_1}}{P^{n_2}+h}}_{\substack{\text{nutrient}\\ \text{limitation}}} \underbrace{\frac{1}{1+c_1 x}}_{\substack{\text{physical}\\ \text{friction}}} \underbrace{\mu\left(1-\frac{N}{K}\right)N}_{\text{Logistic growth}} \tag{2.8a}$$

$$P = P_0 - \underbrace{gN}_{\substack{\text{nutrient}\\ \text{uptake}}} \tag{2.8b}$$

This model provides a reasonable approximation of the non-constant spreading of an *E. coli* colony incorporating several control factors such as cell density and nutrient limitation.

2.4.3 Revised Model

Firstly, without losing the non-monotonicity, we simplify the diffusion function $D(N)$ by assuming $m_1 = 1$ and $m > 1$ as follow:

$$D(N) = \frac{a_1 N}{N^m + 1} \tag{2.9}$$

the profile of (2.9) is presented in (Figure 2.4), which shows that the diffusion rate is low when cell density is either too low or too high.

Meanwhile, the greyscale data of a phase contrast image can be used to estimate the cell density. With assumptions that the whole colony surface is a calotte and the distribution of the cell inside the colony is uniform (Warren *et al.*, 2019), we can compare the total population of the cell by calculating the area of the colonys half-side profile. Furthermore, we can estimate the relative reproduction rate of a bacteria colony at different time points. Based on the experimental data, we found out that the total population of the cell shared a similar profile of the colony size, as it grew slowly during the first 24 hours after seeding, then reached a much higher growth rate in the next 24 hours (Figure 2.5B). Such profile indicates that the reproduction rate may not be a monotone function of cell density N.

Based on the above analysis, we modified the Logistic growth by considering a density-dependent inhibition described by Michaelis-Menten kinetics, which takes the following form:

$$F(N) = \frac{N}{K_1 + N} \tag{2.10}$$

where K_1 is the Michaelis constant. Note that this inhibition effect is strong when cell density is low and weak when cell density is high, this profile shows that such inhibition mainly affects at the early stage of bacterial colony growth (Figure 2.4).

Combining the two regulations, our revised model takes the following form:

$$\frac{\partial N}{\partial t} = \frac{\partial}{\partial x}\left(D(N)\frac{\partial N}{\partial x}\right) + \mu F(N)\left(1 - \frac{N}{K_2}\right)N \tag{2.11}$$

where the first term on the right side of (2.11) describes cell diffusion as well as colony expansion, with a density-dependent diffusion rate; the second term modeled cell growth with a non-monotonic growth rate about the density N.

By comparing the simulation results of Fisher's equation and our revised model, we can conclude that our model can capture the bacterial colony's non-linear growth profile, while Fisher's equation can only present a linear growth profile which is mainly due to the constant diffusion rate.

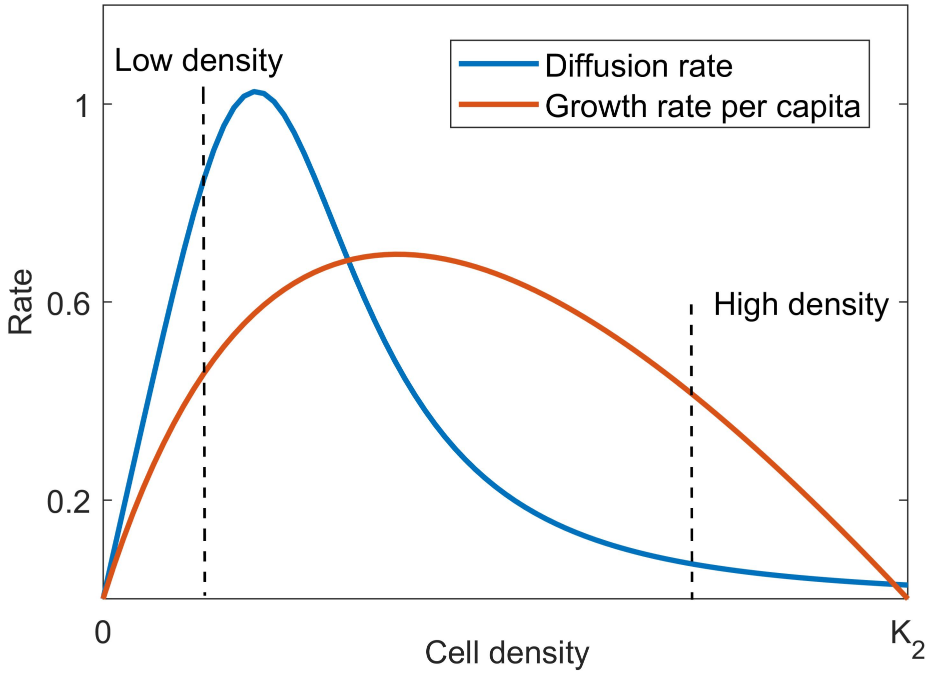

Figure 2.4: An example of density-dependent diffusion rate ($D(N)$) and growth rate per capita ($\mu F(N)(1 - N/K_2)$), showing the nonlinearity of the rates. The ranges of low density and high density are randomly picked. X-axis represents the cell density and Y-axis refers to the rates. Parameters values are: $a_1 = 1.8$, $m = 4$, $K_1 = 2$, $K_2 = 4$, $\mu = 2.6$.

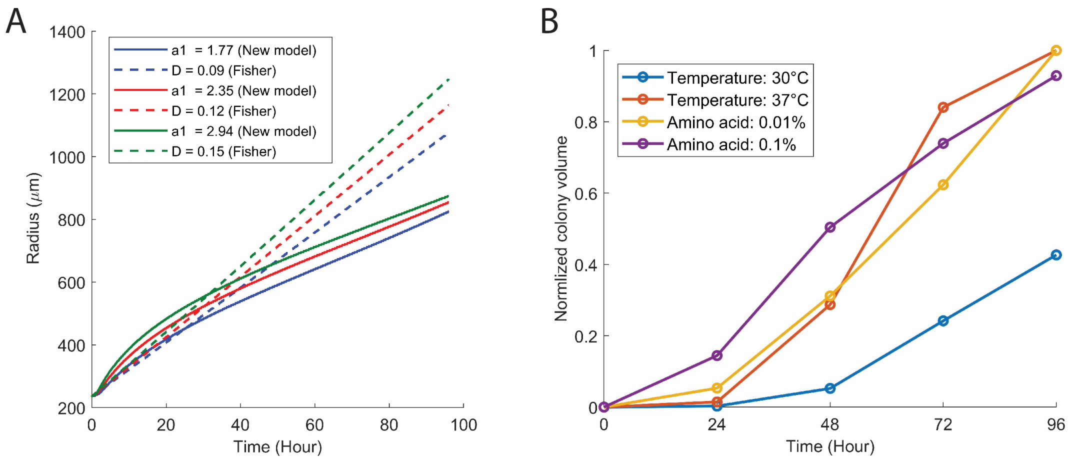

Figure 2.5: Comparison of Fishers equation and our new model via simulations from He *et al.* (2020). **(A)** Comparison of Fishers equation and our new model via simulations. Y-axis represents colony radius in micrometer, while X-axis is time points starting from 0h to 96h. The dashed curves show the change of a colony radius at various times with different diffusion rates D. The solid curves represent the change of a colony radius at various times with different values of a_1. **(B)** Normalized colony volumes under different experimental conditions at various times. Y-axis represents normalized colony volumes computed from experimental data, while X-axis is time points at which colony was imaged, starting from 0h to 96h. One colony copy was randomly picked in each set of experiments, and the normalized colony volume was computed based on the greyscale.

2.5 Simulation

With the comprehensive discussion of the model derivation, the numerical analysis of the model becomes indispensable. Fitting the experimental data can help us to know if the model is capable of predicting colony growth under different conditions. Noting that our control strategies are tuning of culture temperature (T) and amino acid concentration (AA) in the growth media, respectively. Although we did not use any parameters or functions to represent temperature nor amino acid concentration, we still know that the change of these conditions will affect both cell diffusion and reproduction, and based on earlier findings (Toennies, 1964; Naganuma *et al.*, 2018; Ratkowsky *et al.*, 1982), both factors are modeled to be directly proportional to the growth rate and thus can be represented by a_1 and μ in (2.11) in our model. Thus we are fitting a_1 and μ in our simulations. Based on the experimental observations, the cell density at the colony center is stable at the end of the experiment, even under different conditions. This implies that the maximum cell density K_2 can be regarded as a constant. Moreover, since K_1 is the Michaelis constant and m determines the profile of function $D(N)$, they are all fixed in the simulations. Parameters $m = 2.0$, $K_1 = 63.0$, and $K_2 = 189.0$ are fixed.

For comparison, we also present experimental data fitting of Fishers equation by varying the parameters D and μ, in (2.5) via *MATLAB*. Despite the simplicity of Fishers equation, it is still an appropriate example for comparison because of several reasons: (1). Fisher's equation maintains the linear growth profile of bacterial colony and Logistic growth term with are two representative characteristics of many previous models. (2). The complexity of Fisher's equation and the new model is close compared to other models with nutrient concentration or chemotactic. Although the new model has two more parameters than Fisher's equation, we only fit two parameters for both of them. (3). Since we are trying to describe the colony growth under multiple control factors by using a preliminary model, it is better to choose a basic model for comparison rather than a complicated mode.

Here we first presented the fitting results of the colony radii data, which shows that our model can produce the simulation results that agree well with the experimental data and are much better than the results of Fisher's equation. We then provided the simulation results of the colony's cross-sectional profile and compared them with the experimental data, which also shows that our model can capture the general shape of the colony edge at different time points.

2.5.1 Radii Date Fitting

The first set of experiments were performed at two different T, 30°C and 37°C. The clear circular patterns could be observed after 24 hours, so the first data point of this experiment is $t = 24h$ and experimental data was collected every 24 hours. Four colonies were recorded at each T, and the average radii of all colonies were computed at every time point for the purposes of data fitting. We set the initial radius for each colony smaller than its first data point and fit the experimental data with the new model by varying the parameters a_1 and μ, in (2.11) via *MATLAB*. These results are shown in Figure 2.6 Left. At 37°C, the colony radius shares a non-constant increasing speed while the colony radius increases at an approximately constant speed at 30°C. At each phase, the colonies at 37°C always have higher colony expansion rates than colonies at 30°C. Here, our model provides comparable fitting results and captures both the non-constant increase of the colony radius at 37°C and constant increase at 30°C. This is because under lower T, the bacterial colony will grow slower and have a more linear growth profile Pipe and Grimson (2008). Note that Fishers equation can only capture the constant increase profile in low T cases, but has much higher fitting errors in high T cases. The insert box of Figure 2.6A presents a comparison of the simulations and experimental data for the 37°C case at time $t = 48h$, showing that the experimental results are stable and the simulation results are reasonably accurate.

The second set of experiments was performed with two different AA, 0.01% and 0.1%. In both cases, the first data point was at $t = 24h$ and the experimental data was collected every 24 hours. There were six colonies for each case and the average radius of all colonies for each case was computed at each time point for the purposes of data fitting. Similar simulations were performed via *MATLAB* and are shown in Figure 2.6 Right.

Surprisingly, colonies with higher AA always have a lower colony expansion rate when compared at the same time period to colonies with lower AA. The key reason for this could be that colonies with higher AA grow much faster in the first 24 hours, which leads to larger colony size and a higher cell density, which subsequently slows down the growth. Based on the experimental data, AA may only change the speed at which an *E. coli* colony grows. Thus, both cases show a non-constant increasing speed of the colony radius. Our model captures this profile and presents comparable fitting results with a much smaller fitting error than Fisher's equation (Table 2.1), which only presents a linear increase of the colony radius. Noting that the fitting results of the two models are comparable for the colony growing under 30°C, which has the lowest expanding rate and linear-like growth profile. This implies that Fisher's equation can only capture the linear growth profile which may only occur in some specific cases.

Fitting error	T:30°C	T:37°C	AA:0.01%	AA:0.1%
Fisher	53.42	288.76	223.92	301.83
New model (revised)	86.28	105.68	24.23	88.22

Table 2.1: Fitting errors of two models under different conditions.

To compare and contrast the data fitting performance of the above two models, we also included two standard model selection criteria: the Akaike information criterion (AIC) (Akaike, 1998) and the Bayesian information criterion (BIC) (Schwarz *et al.*, 1978). Both AIC and BIC are very important estimators of the goodness of fit (Miao *et al.*, 2009, 2012), and deal with the trade-off between the goodness of fit and the model simplicity, while the model with the lowest AIC or BIC is preferred. Under the assumption that the model error is independent and normally distributed, the above criteria can be written as follow:

$$AIC = n\ln(\frac{RSS}{n}) + 2K \tag{2.12a}$$

$$BIC = n\ln(\frac{RSS}{n}) + K\ln(n) \tag{2.12b}$$

where RSS refers to the residual sum of squares, K is the number of parameters, and n is the sample size. More importantly, for small sample size, a variation of AIC which contains a bias correction term may be more appropriate:

$$AICc = n\ln(\frac{RSS}{n}) + \frac{2nK}{n-K-1} \tag{2.13}$$

where AICc represents "corrected Akaike information criterion".

Here we have two parameters for both Fisher's equation and our revised model, for each set of experiments we have four data points, which means $n = 2$ and $K = 4$. Since $K > n/40$ in our case, the AICc instead of AIC should be used (Miao *et al.*, 2009). And recall the formula of RSS:

$$RSS = \sum_{i=1}^{K}(r_i - \hat{r}_i)^2 \tag{2.14}$$

where K is the total number of data points, t_i is the time of the i^{th} data point. Here r_i represent the radii values at time t_i produced from the model, and $\hat{r}_i$ correspond to the experimental observations at time t_i. By comparing the AICc of both two models, which are shown in Table 2.2, we can conclude that our new model has better fitting quality in most cases with lower AICc scores.

2.5.2 Cross-sectional Profile

In addition to radial colony growth, we also quantitatively analyzed the spatial profile of the colonies. We estimated the cell density by using the greyscale data

AICc	T:30°C	T:37°C	AA:0.01%	AA:0.1%
Fisher	29.14	35.89	34.87	36.07
New model (revised)	31.06	31.87	25.98	31.15

Table 2.2: AICc scores of two models under different conditions.

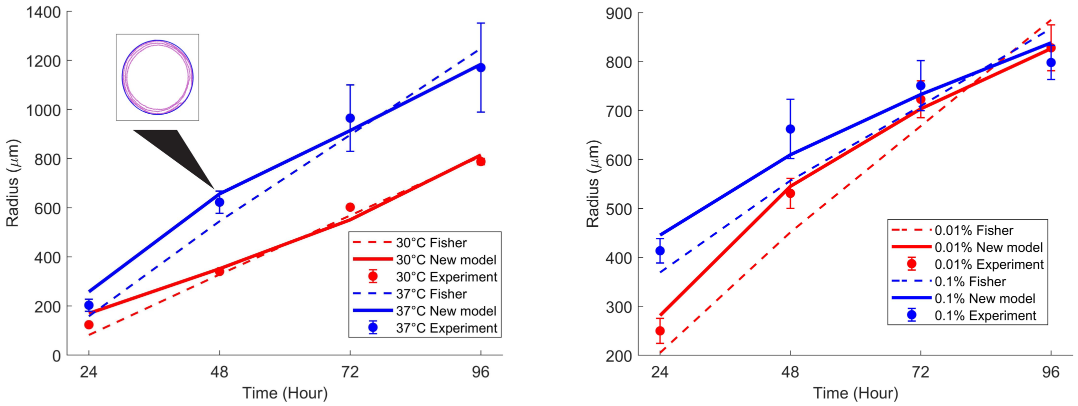

Figure 2.6: Fitting results of two models under different conditions from He *et al.* (2020). Y-axis represents colony radius in micrometer, while X-axis is time points at which colony was imaged, starting from 24h to 96h. Blue and red curves represent simulated results of the new model, dashed lines show the results of Fishers equation, the dots are averaged experimental data while error bars indicate the standard error between exact data and their averages. **Left:** Fitting results of two models under different T. Red and blue colors represent low (30°C), and normal (37°C) T, respectively. In the box, blue circle is the fitted simulation, while purple circles are the circumference of the colonies taken from the raw images. **Right:** Fitting results of two models with different AA. Red and blue colors represent normal (0.01%), and high (0.1%) AA, respectively.

of a phase contrast image. Since cell densities can be viewed as the colony height (from colony surface to agar plate), the greyscale data can be used as a reporter of experimental density, so plotting the greyscale data ranging from the colony center to the edge can be regarded as an approximation of the spatial profile. To minimize the background noise, we exported the greyscale data ranging from the center to the leftward edge and flip the image along the vertical axis. In this paper, we have chosen one of the colonies growing at 30°C and have plotted its profiles at $t = 72h$ and $t = 96h$. Note that due to different noise levels, the threshold for the colony edge may vary. The corresponding simulation results of Figure 2.7 are also presented. Rather than aiming at predicting the exact shape of the colony edge, we are focusing on developing a framework to capture the general shape of the colony boundary. Since the exact profile of a bacterial colony may vary in different experiments, the exact fitting would likely require much more detailed modeling of the internal mechanics of colony formation and is left for future explorations.

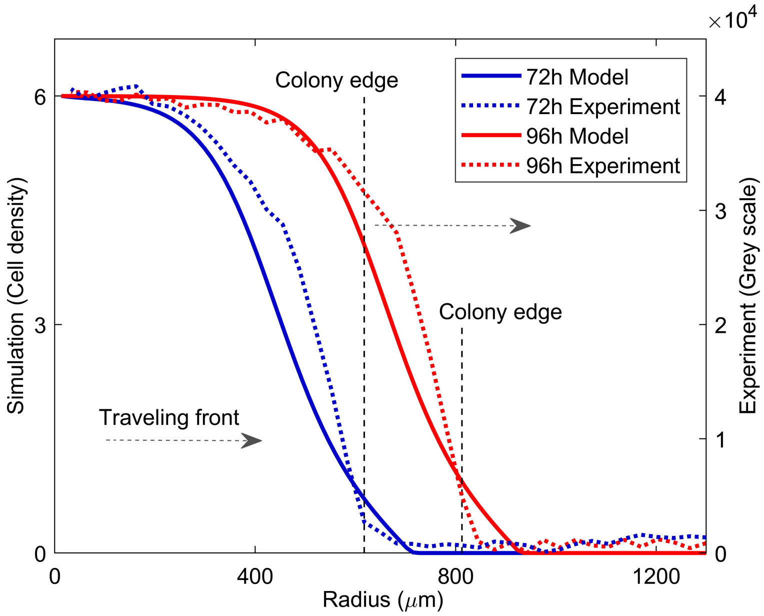

Figure 2.7: Comparison of quantified temporal and spatial greyscale of experimental observations and simulated cross-sectional colony profiles - Figure taken from He *et al.* (2020). Dashed curves represent the greyscale of a bacterial colony (T: 30°C) starting from the colony center (detected by computer program) at different times. The simulated colony profiles starting from colony center of the corresponding time are shown in solid curves. Black dashed lines indicate the location of colony edges and grey arrows indicated the colony spreading direction.

2.5.3 Parameter Estimation Methods

The two control factors we focus on are T and AA. It is clear that the *E. coli* cell will obtain both a higher diffusion rate and higher reproduction rate if growing under higher T or AA, which will result in higher values of a_1 and μ in our model. Therefore a_1 and μ will be the fitting parameters. Noting that although the new model has two more parameters, we are only fitting two of them which is the same with Fisher's equation.

We first solve the model by using the *pdepe* package, a visible threshold is a setup based on the threshold used in processing the phase contrast image to determine the radius of the simulation result.

Then to find the values of a_1 and μ that best fit the observed radii data, we use the simplex method from Lagarias *et al.* (1998), which is implemented in *MATLAB* (version 2018b) function *fminsearch*. The objective function begins by performing a simulation assuming radial symmetry on an infinite domain. The initial condition is created by looking at the initially observed colony radius and we define the no-flex boundary condition in our simulations. Any spatial nodes that are closer to the colony center than this initial radius receive a cell density value equal to the carrying capacity, while all other spatial nodes are set to zero. Next, the colony radius generated by the model is found by taking the radius of the farthest spatial node from the center that has a cell density greater than the visibility threshold, equal to 0.03 in this case. Finally, the error is calculated using the least-square formula given by:

$$Error = \frac{1}{2}\sum_{i=1}^{K}(r_i - \hat{r}_i)^2 \tag{2.15}$$

where K is the total number of data points, t_i is the time of the i^{th} data point. Here r_i represent the radii values at time t_i produced from the model, and $\hat{r}_i$ correspond to the experimental observations at time t_i.

2.6 Dicsuccion and Conclusion

In this study, we proposed a partial differential equation model describing the growth of *E. coli* colonies on semi-solid agar starting from a single cell with multiple control factors. For each set of conditions, we compared the model simulations with corresponding experimental results. *E. coli* are common bacteria that can reproduce on semi-solid agar surfaces and after a certain period of time, form a circular colony. They can also propel themselves by means of long hair-like appendages known as flagella. Our experiments show that as *E. coli* colonies grow, their profiles evolve in a qualitatively similar way despite being subjected to differing T or AA. In the first 24 hours, the radial growth is slow in comparison to the vertical growth. However, over the course of the next 48 to 72 hours, the radial growth speeds up before it slows down. The key factors that control colony growth include physical friction among cells and between the cells and the agar surface and inhibition of cell reproduction potentially caused by agar dehydration. Also, cell colonies present similar growth profiles throughout the experiments, but there are notable differences in colony radius. However, when changing the AA, the final sizes of all colonies were very close. This phenomenon might imply that T affects both the growth speed and final size of a colony, while nutrient concentration mainly affects the growth speed. Meanwhile, it is necessary to point out that although T and AA are independent variables, the mechanisms of how they will affect cells metabolism and movement are still unclear. Fortunately, the outcomes of tuning these variables are predictable and can be described via cell density. The density-dependent functions presented in our reaction-diffusion-based model can be a possible approach. Furthermore, our model can not only capture these characteristics but also produce simulation results comparable to the experimental data, which shows that the interactions of these control factors can be well-described by nonlinear density-dependent functions. Moreover, the model can provide comparable simulation results on a colonys spatial (cross-sectional) profile at various times.

To focus only on radial expansion, it is worth noting that here we have omitted vertical growth because recent studies have shown that these two processes are independent (Farrell *et al.*, 2013). Moreover, the forces that govern each propagation factor are distinct and can be incorporated separately. Key parameters that were tested are AA and T (Toennies, 1964; Naganuma *et al.*, 2018). These factors are tunable and have dramatic effects on colony expansion (Ratkowsky *et al.*, 1982). The default settings for T and AA were 37°C and $0.01\% g/mL$ respectively. By varying these key parameters, we were able to either accelerate or decelerate the growth rate. Specifically, we found that decreasing the T to 30°C doubles the time it takes for the colony to proliferate. On the other hand, increasing AA by a factor of 10 leads to an escalated initial growth rate (Figure 2.6 Right). After 24 hours, we see that the colony with a high AA has a diameter that is twice as large as the control. However, this difference diminishes as both colonies look similar in size by the end of 4th day. There are many mechanisms contributing to these visible changes, including nutrient availability, cell-agar friction, surface tension, and adhesion, as well as dehydration (Ratkowsky *et al.*, 1982; Jayathilake *et al.*, 2017; Grant *et al.*, 2014; Cole *et al.*, 2015). Most of these points were taken into consideration while designing our experiments. For instance, to decrease mechanical forces between agar and cells, we prepared a semi-solid agar with 0.3% agarose by weight, so that the colony could grow with less friction and adhesion, but still maintains a proper shape. Also, water tanks were placed inside the incubators to minimize evaporation from the surfaces of the plates.

The model we propose here, which produces quantitative results that are comparable to the experimental data, is not only a reasonable example of integrating *E. coli* colony expansion with multiple control factors, but also a potentially powerful tool that can be used to help develop strategies for controlling colony growth under different conditions and for interpreting the mechanisms of other biological phenomena such as spatial pattern formation. In future studies, parameters for the experimental variables (T and AA) in the model could be included explicitly, and impacts to other density-dependent factors besides the maximum growth rate and diffusion rate could be considered, both would require additional experimental validation. With consideration of more complex effects of cell growth and metabolism as well as environmental factors, the model can be used to depict both the radial and vertical growth of other microbe colonies. This will provide valuable mathematical insight and help improve experimental designs. Furthermore, this model can be a framework to systematically present spatial pattern formation. The growth of a bacterial colony will not only interact with circuit expression (Wu *et al.*, 2018, 2017) and pattern formation (Basu *et al.*, 2005; Cao *et al.*, 2016), but will also control the pattern boundary. Developing a mechanistic model that builds on this colony growth model can be an interesting way to study the variable boundary problem seen in pattern formation.

Chapter 3

REACTION-DIFFUSION-BASED MATHEMATICAL MODELING ON CIRCUIT-DRIVEN BACTERIAL PATTERNING

3.1 Abstract

The reaction-diffusion-based (RD-based) clock and wavefront model has long been proposed as the mechanism underlying biological pattern formation of repeated and segmented structures including somitogenesis. However, systematic molecular-level understanding of the mechanism remains elusive, largely due to the lack of suitable experimental systems to probe RD quantitatively *in vivo*. Here we design a synthetic gene circuit that couples gene expression regulation (reaction) with quorum sensing (diffusion) to guide bacterial cells self-organizing into stripe patterns at both microscopic and colony scales. An experimentally verified mathematical model confirms that these periodic spatial structures are emerged from the integration of oscillatory gene expression as the molecular clock and the outward expanding diffusions as the propagating wavefront. Furthermore, our paired model-experiment data illustrate that the RD-based patterning is sensitive to initial conditions and can be modulated by external inducers to generate diverse patterns, including multiple-stripe patterns, target-like patterns, and ring patterns with reversed fluorescence. Powered by our synthetic biology setup, we also test different topologies of gene networks and show that network motifs enabling robust oscillations are foundations of sequential stripe pattern formation. These results verified close connections between gene network topology and resulting RD-driven pattern formation, offering an engineering approach to help understand biological development.

3.2 Introduction

3.2.1 Somitogenesis in Vertebrates

Somitogenesis is one of the significant developmental features during embryonic development, which forms a series of repetitive segments (or somites) along the anterior-posterior axis. Somites are bilateral pairs of mesoderm, which refers to the middle layer of an early-development embryo. Somites of vertebrates will subdivide into sclerotomes, myotomes, and dermatomes which contribute to the formation of vertebrae, rib cage, skeletal muscle, dermis, and skin. More interestingly, the somitogenesis occurs with a certain periodicity, for example, the average somite formation rate of a corn snake is one pair every 100 min, compared to rates of one pair every 30, 90, and 120 min in zebrafish, chicken and mouse (Gomez *et al.*, 2008). This periodicity comes from the oscillatory expression of particular genes which has been known as the "segmentation clock", while the wavefront progresses slowly in an anterior-to-posterior direction. Thus, the somitogenesis process is coordinated by a "clock and wavefront" mechanism.

The introduction of the clock and wavefront mechanism can be traced back to Cooke and Zeeman (1976), and great progress has been made during the past 40 years on understanding somitogenesis from different aspects. Haddon *et al.* (1998) examined the functions of multiple *delta* genes in somite development by gene expression experiments, Vermot and Pourquié (2005) reported the effect of retinoic acid on somitogenesis in chick embryos, and Naiche *et al.* (2011) concluded that specific fibroblast growth factors (*Fgf4* and *Fgf8*) comprise the wavefront activity based on experimental observations. And during the past several years, lots of relevant work has brought us valuable insights into this mechanism (Yabe and Takada, 2016; Diaz-Cuadros *et al.*, 2020) as well as modeling approaches (Lewis, 2003; Cotterell *et al.*, 2015). Meanwhile, having a systematic molecular-level understanding of the mechanism can provide us valuable insights into the key factors of gene oscillatory expression and the control of wavefront. However, due to the lack of suitable experimental systems and mathematical models to probe the clock and wavefront mechanism quantitatively *in vivo*, this area remains elusive.

The study of somitogenesis on a systematic molecular level requires appropriate biological systems which are capable of generating repeated patterns. The bacterial stripe patterning which is guided by synthetic gene circuits has become a suitable platform to study the clock and wavefront mechanism. In the experiments, designed gene circuits with fluorescent proteins were transformed into bacteria cells before seeding on the agar plates, and due to the difference of fluorescent protein concentrations, the stripe patterns can be observed via live single-cell time-lapse fluorescence microscopy. Although we have already known that the cause of sequential stripe patterns is the oscillatory expression of certain genes, we are still not clear what is the key factor for the synthetic circuit to have this oscillatory expression and how to control it. Therefore, developing an appropriate mathematical model is pivotal in studying the mechanism of sequential stripe patterning, which can also be a valuable platform to study pattern formation in a more complicated environment.

3.2.2 Reaction-diffusion-based Pattern Formation

Turing's seminal work first proposed reaction-diffusion (RD) as the "chemical basis of morphogenesis" over six decades ago (Turing, 1990). It provides a general theoretical foundation of pattern formation via RD mechanisms. Two decades later, RD-driven clock and wavefront (CW) mechanism was hypothesized as the mechanism underlying the formation of repeated and segmented structures such as somites in development (Cooke and Zeeman, 1976). Since then, although RD-driven pattern formation has been demonstrated or identified in chemical, physical, and ecological systems (Lengyel and Epstein, 1992; Kondo and Miura, 2010), its much-hypothesized role in multicellular pattern formation hasn't been fully studied biologically nor mathematically. This is largely due to the lack of suitable model systems to test such hypotheses. For example, somite development requires precise temporal and spatial coordination between a heterogeneous web of intracellular responses and intercellular communications, both under the control of complex gene regulation networks and influences of universal gene expression stochasticity. Such complexity poses a great challenge to fully understand the mechanistic basis of somite formation *in vivo*. Engineered microbes carrying rationally designed gene circuits provide an effective venue to study this problem from the bottom up. Previous studies using synthetic circuits have demonstrated the formation of the predefined patterns, cell motility-based stripe formation, and scale-invariant ring pattern formation (Basu *et al.*, 2005; Liu *et al.*, 2011). However, gene network directed RD based clock and wavefront pattern formation, despite its importance in developmental biology and extensive theoretical studies (Gomez and Arcak, 2017; Chen *et al.*, 2014; Baker *et al.*, 2006), has not been experimentally realized.

Meanwhile, the reaction-diffusion model is widely used in modeling colony formation and multicellular pattern formation. However, mathematical modeling which is based on the gene circuit topology to demonstrate the circuit-driven pattern formation is rarely discussed. Having a dynamical model developed from the circuit topology to study the pattern formation will bring us huge advantages in understanding the mechanisms of the oscillatory gene expression.

3.3 Experimental Design

3.3.1 Circuit Introduction

Past studies have suggested that nonlinear multistable systems could also direct spatiotemporal pattern formation when coupled with external diffusion process (Capasso *et al.*, 2012; Leppänen *et al.*, 2004; Tuszyński *et al.*, 1991). Following this strategy to achieve a multicellular pattern formation, Wang et al. designed and constructed a mutually inhibitory network with positive autoregulation and communications (MINPAC) (Wu *et al.*, 2019) by expanding their previously demonstrated quadra-stable gene circuit Wu *et al.* (2017) with added quorum-sensing modules to enable intercellular communications.

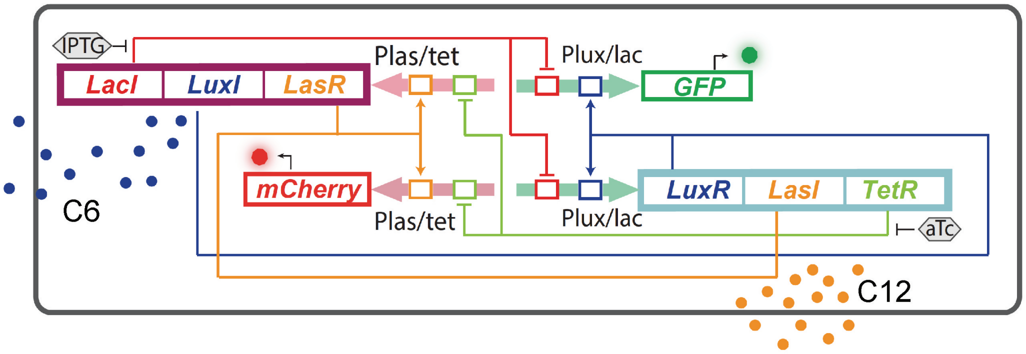

Figure 3.1: Experimental design of the MINPAC network - Figure taken from Wu *et al.* (2019).

Specifically, the MINPAC topology (Figure 3.1) is built upon two hybrid promoters *Plas/tet* and *Plux/lac*, which harbor high nonlinearity and inducibility. *Protein group 1* contains three different proteins *LasR, LacI, LuxI* and *Protein group 2* contains *LasI, LuxR, TetR*. *Plas/tet* drives *LasR, LacI* and *LuxI* expression whereas *Plux/lac* regulates transcription of *LuxR, LasI* and *TetR*. *LasI* and *LuxI* are synthases that catalyze the synthesis of autoinducer 3-oxo-C12-HSL (C12) and 3-oxo-C6-HSL (C6), respectively. The two small autoinducers can diffuse out of and into cells to mediate cell-cell communication and coordinate population behaviors on a spatial domain. *LasR* and *LuxR* activate *Plas/tet* and *Plux/lac* in the presence of C12 and C6, respectively, forming positive autoregulations. *IPTG* inhibits the repressive effect of *LacI* on *Plux/lac*, and *aTc* counteracts *TetR* inhibition on *Plas/tet*, forming the mutual inhibitions. Green fluorescent protein (GFP) and mCherry protein serve as the corresponding reporters of *Plux/lac* and *Plas/tet* activities in living cells.

To help us understand the circuit topology, we can simplify the MINPAC network by denoting the two hybrid promoters as *Promoter 1* and *Promoter 2*, respectively. Meanwhile, *LasR, LacI, LuxI* are produced from the same promoter *Plas/tet* and have similar production terms. We assume the three proteins have similar degradation rates and similar dynamics and use *LuxI* to represent the rest two and denoted by U. Similarly, we use *LasI* to represent *LuxR* and *TetR* expression dynamics from *Plux/lac* and denoted by A. In addition, since the two small autoinducers can diffuse out of and into cells, we use C_i and C_e to represent internal C6 and external C6, respectively. Similarly, H_i and H_e refer to internal C12 and external C12, respectively. The simplified MINPAC network is shown in Figure 3.2.

3.3.2 Experiment Introduction

To investigate whether MINPAC could direct single cells to self-organize into spatial patterns, Wang et al. transformed the circuit into *E. coli* cells and serially diluted cell cultures into single cells before seeding on a semi-solid minimal M9 medium. The first microscopy image was taken at the 15th hour, when a light, a small colony can be observed, the second picture was taken at the 24th hour and then every 24 hours. The phase contrast images (Figure 3.3 Last row) were taken without any filters while GFP and mCherry images (Figure 3.3 First and second row) were taken with corre-

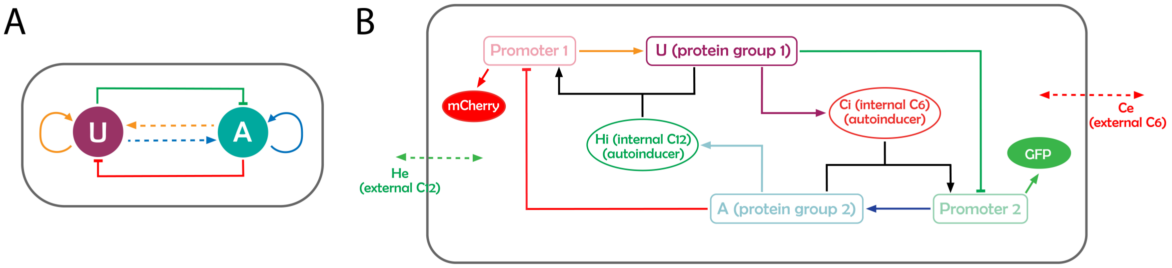

Figure 3.2: A simplified MINPAC network - Figure adapted from Wu *et al.* (2019). **(A)**. Abstract diagram of MINPAC topology where U and A mutually inhibit each other (T-bars) and auto-activate (arrowheads) itself, meanwhile U and A can mutually activate through small autoinducer mediated intercellular communication (dashed arrowheads). **(B)**. *Promoter 1* can be activated by U and repressed by A. U synthesizes C_i (internal C6) binding with A to activate *Promoter 2*, while A synthesizes H_i (internal C12) binding with U to activate *Promoter 1*. The C6 and C12 which diffuse out of the cell are denoted C_e (external C6) and H_e (external C12), respectively. GFP and mCherry serve as reporters for *Promoter 1* and *Promoter 1*, respectively.

sponding filters. After merging the GFP and mCherry images together we obtained the colored images with the ring patterns observed (Figure 3.3 Third row). Time-lapse colony imaging results show that the single colony has no obvious pattern at 15 hr and exhibits a weak yellow flat disk, suggesting cells express either GFP or mCherry are distributed without order (Figure 3.3). After 24 hr, cells in the colony started to differentially and orderly express GFP and mCherry and self-organize into a stable double-ring pattern of an outer GFP-ring and inner mCherry disk at 48 hr (Figure 3.3), with a small temporary yellow ring between these two rings. The double-ring pattern is stable with time. Fluorescence quantification also confirms higher GFP expression for cells on the edge of the colony and higher mCherry expression for cells in the center.

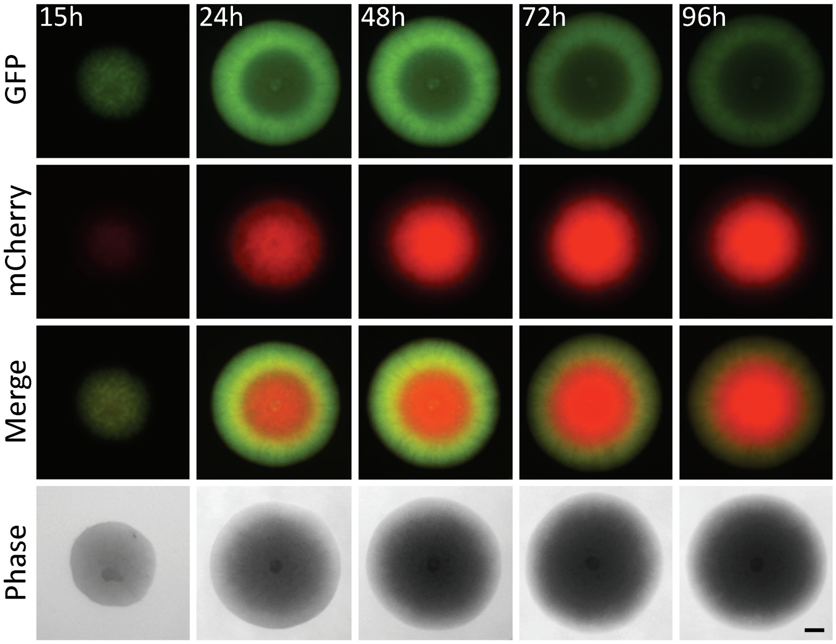

Figure 3.3: MINPAC cells self-organized double-ring pattern at colony scale. The first microscopy image was taken at the 15th hour, while a weak circular colony can be observed. The second microscopy image was taken at the 24th hour and then every 24 hours. The fourth row are phase contrast images without any filters. The first two rows are images with corresponding filters to indicate the intensity of GFP and mCherry. The third row refers to the images merged by the first two rows. Scale bar represents $100\mu m$. Magnification: 2x. - Figure taken from Wu *et al.* (2019).

3.4 Full System

3.4.1 Model Introduction

To develop a quantitative and mechanistic understanding of the MINPAC-directed ring patterning process, we developed a partial differential equation (PDE) model based on the reaction-diffusion process involving the regulation, production, and diffusion of morphogens C6 and C12.

Firstly, we notice that *Plux/lac* activity is determined by the relative concentrations of *LacI* and *LuxR*-C6, which is a complex of *LuxR* protein and intracellular C6 (C_i). Similarly, *Plas/tet* dynamics is determined by the relative concentrations of *TetR* and *LasR*-C12, which is a complex of *LasR* protein and intracellular C12 (H_i). Based on the simplified topology, the *Plux/lac* activity can be presented as a hill function of A and C_i, while *Plas/tet* dynamics can be described as a hill function of U and H_i. Meanwhile, we use hill functions to describe the repressions. Including the self-activation terms and degradation terms, the differential equations modeling the dynamics of U and A are (3.1a) and (3.1b), respectively.

Secondly, the biosynthesis of C_i and H_i primarily depends on U and A, respectively. And both C_i and H_i can diffuse out of and into cells, and bacterial cells further respond to the autoinducers when their concentrations exceed a certain threshold. Based on this, we described the C_i and H_i dynamics by (3.1c) and (3.1d), respectively.

Finally, since the external C6 (C_e) and C12 (H_e) can diffuse across the cell colony on the M9 agarose medium, we use (3.1e) and (3.1f) describe C_e and H_e dynamics.

The full system can be written as following and the biological meaning of the parameters are included in Table 3.1:

$$\frac{\partial U}{\partial t} = \underbrace{\beta_1}_{\text{self-activation}} + \underbrace{\frac{k_1(UH_i)^{n_1}}{1+(UH_i)^{n_1}}}_{\text{expression in the presence of } H_i} \underbrace{\frac{1}{1+A^{m_1}}}_{\text{inhibition}} - \underbrace{d_1U}_{\text{degradation}} \tag{3.1a}$$

$$\frac{\partial A}{\partial t} = \underbrace{\beta_2}_{\text{self-activation}} + \underbrace{\frac{k_2(AC_i)^{n_2}}{1+(AC_i)^{n_2}}}_{\text{expression in the presence of } C_i} \underbrace{\frac{1}{1+U^{m_2}}}_{\text{inhibition}} - \underbrace{d_2A}_{\text{degradation}} \tag{3.1b}$$

$$\frac{\partial C_i}{\partial t} = \underbrace{\frac{k_3U^{n_3}}{K_c^{n_3}+U^{n_3}}}_{\text{expression governed by } U} - \underbrace{d_3C_i}_{\text{degradation}} + \underbrace{D_c(C_e-C_i)}_{\text{intercellular diffusion}} \tag{3.1c}$$

$$\frac{\partial H_i}{\partial t} = \underbrace{\frac{k_4A^{n_4}}{K_h^{n_4}+A^{n_4}}}_{\text{expression governed by } U} - \underbrace{d_4H_i}_{\text{degradation}} + \underbrace{D_h(H_e-H_i)}_{\text{intercellular diffusion}} \tag{3.1d}$$

$$\frac{\partial C_e}{\partial t} = \underbrace{-D_c(C_e-C_i)}_{\text{intercellular diffusion}} - \underbrace{d_5C_e}_{\text{degradation}} + \underbrace{D_1\frac{\partial^2 C_e}{\partial x^2}}_{\text{outer diffusion}} \tag{3.1e}$$

$$\frac{\partial H_e}{\partial t} = \underbrace{-D_h(H_e-H_i)}_{\text{intercellular diffusion}} - \underbrace{d_6H_e}_{\text{degradation}} + \underbrace{D_2\frac{\partial^2 H_e}{\partial x^2}}_{\text{outer diffusion}} . \tag{3.1f}$$

It is noteworthy that although *Plux/lac* activity is activated by the complex of *LuxR* and internal C6, the quorum-sensing mechanism is cell population density-dependent. In other words, *Plux/lac* can be activated only when the local environmental C6 reaches a certain threshold. Thus, it is the external C6 (C_e) and C12 (H_e) that determine the dynamics of MINPAC as well as the patterning process. Furthermore, since mCherry and GFP are two reporters of the hybrid promoters *Plas/tet* and *Plux/lac*, so we can use external C6 (C_e) and C12 (H_e) to equivalently simulate mCherry and GFP dynamics.

Symbol	Definition
β_1, β_2	Basal expression rates from *Plas/tet* and *Plux/lac*, respectively
k_1, k_2, k_3, k_4	Maximum production rates
K_c, K_h	Half maximal effective concentrations for C6 and C12 productions, respectively
D_c, D_h	Diffusion rates of the diffusion through the cell membrane
D_1, D_2	Diffusion rates of the diffusion across the colony on the medium
n_1, n_2, n_3, n_4, m_1, m_2	Hill coefficients
d_1, d_2, d_3, d_4, d_5, d_6	Degradation rates

Table 3.1: Parameter Explanations of the Model (3.1).

3.4.2 Simulation Results

Since multiple programming platforms have been used to simulate pattern formation under different conditions, a detailed introduction of the numerical methods will be provided in Section 3.11. Noting that there are many parameters, some are fixed throughout all simulations included in this chapter. Here we fixed all the Hill coefficients to $m_1 = 4$, $m_2 = 4$, $n_1 = 2$, $n_2 = 4$, $n_3 = 3$ and $n_4 = 2$, which model the nonlinearity of the network interactions. We also fixed all the degradation rates to $d_1 = 1.19$, $d_2 = 1.19$, $d_3 = 0.56$, $d_4 = 0.8$, $d_5 = 20$, $d_6 = 20$. The higher degradation rates of external C6 and C12 are due to the possibility that autoinducers can diffuse out of the colony which will not be considered anymore. Furthermore, it is well known that the two diffusible signaling molecules with different diffusion constants are one of the fundamental requirements to generate Turing patterns. In our system, we use C6 and C12 to mediate the intercellular communication by diffusing in the colony and semi-solid medium. Previous studies indicate that the diffusion rates for C6 and C12 are very similar (less than 1.5 fold difference) (Basu *et al.*, 2005; Liu *et al.*, 2011) and given that their similar chemical structures (C12 only has six more carbons than C6), we here assume they have the same diffusion coefficient $D_1 = D_2$.

The first set of simulations is generated from both *NetLogo* and *MATLAB* (Figure 3.4). We used *NetLogo* to simulate *E. coli* cell's reproduction and random movement, coupling with the reaction part of the model (3.1) to describe the circuit expression and *MATLAB* to plot the average data of three rounds of *NetLogo* simulations. The color of a cell is determined by the concentration of C6 and C12 and the whole process starts from one single cell with a random color. Our simulation shows that there is no obvious ring patterns in the early-stage colony while the ring patterns will finally be presented (Figure 3.4 Top and bottom left). And the average result of three rounds of simulations (Figure 3.4 Bottom right) shows that the colony will always have a mCherry core and GFP outer-ring. These results illustrated that *E. coli* cell transformed with MINPAC circuit can be a robust platform generating ting patterns. In the *NetLogo* simulations, we assume the cell moves every 10 units of time and divide every 20 units of time, while the rest parameters are $b_1 = 0.8$, $b_2 = 0.5$, $k_1 = 640$, $k_2 = 700$, $k_3 = 80$, $k_4 = 105$, $K_c = 70$ and $K_h = 82$. We also regarded this set of parameters as the "standard set" for comparing with parameters in other simulations.

The experimental observations show that there were different ring patterns under

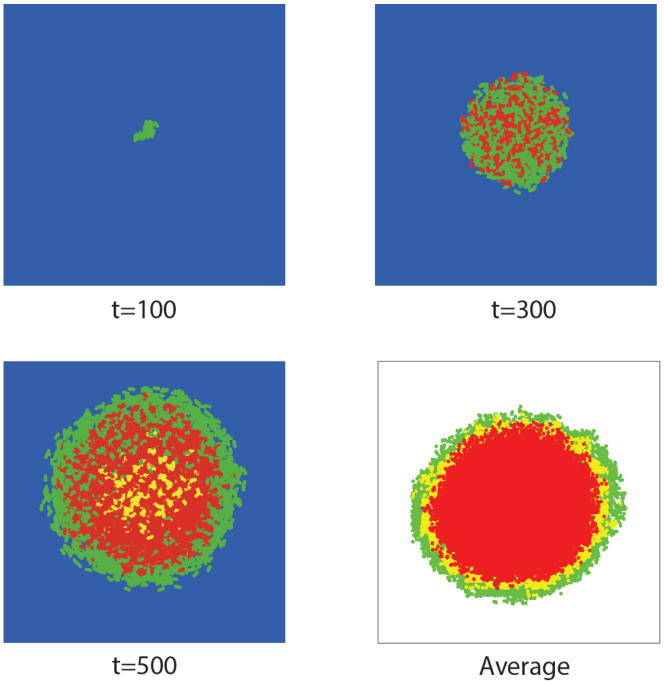

Figure 3.4: Simulation results at different time points by *NetLogo* and the average results. Red cell indicates higher mCherry concentration and green cell indicates higher GFP concentration, while yellow cell indicates close concentrations of both two fluorescent proteins. Top and bottom right: Simulation of colony growth and circuit expression starting from a single cell with cell's random walk at different time points. Bottom right: Average results of three rounds of simulations via *NetLogo*.

the same conditions without any inductions (Figure 3.5 First column). We hypothesize that it is due to random variations of the initial concentrations of intracellular proteins and autoinducers (detailed discussion is provided in Section 3.6), which is corresponding to the initial condition of the model. Taking the parameters used in (Figure 3.4) with different initial conditions and set the diffusion rates $D_c = D_h = 4$ and $D_1 = D_2 = 800$, our simulations presented comparable results for both two cases, showing that our model can capture such variation by only changing the initial conditions.

To further examine the patterns controllability, we applied external inducers to perturb the regulations of MINPAC and hence pattern formation. C6, when applied

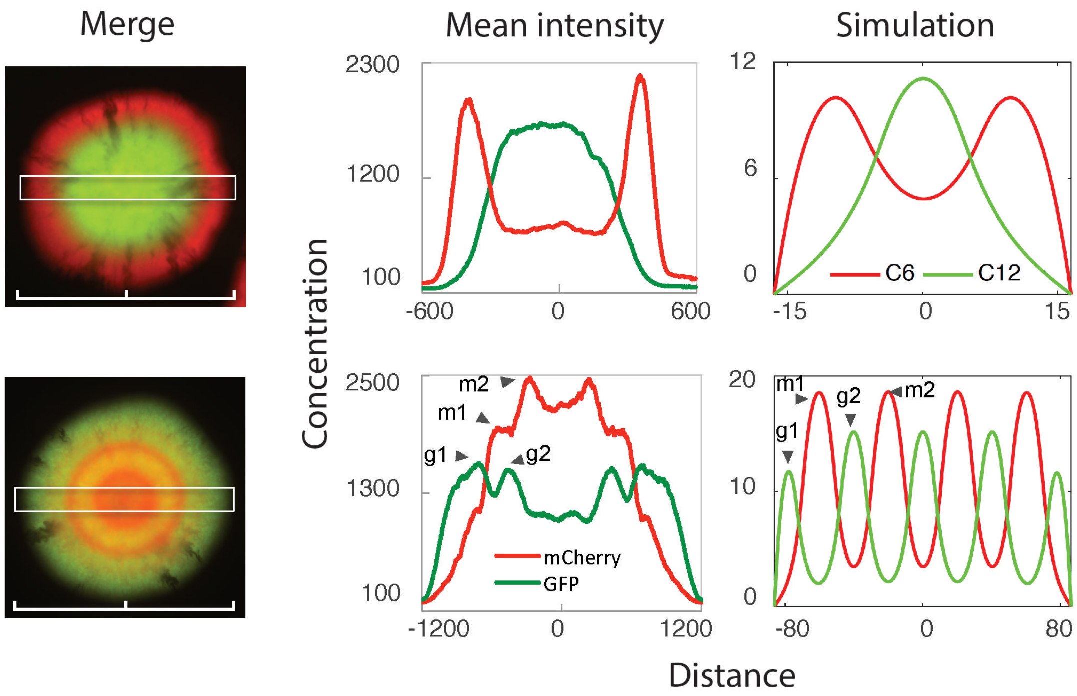

Figure 3.5: Two observations generated by MINPAC circuit and the corresponding simulation results - Figure taken from Wu *et al.* (2019). Top: a ring pattern with a GFP core and a mCherry outer ring; bottom: a multi-ring pattern with mCherry core. Middle: mean fluorescence intensities across the center of the corresponding colonies. Right: Model simulations recapitulate experimental patterns.

externally (at the beginning of the experiment), would promote GFP expression and also *LasI* and *TetR* production, which could both activate and inhibit mCherry expression. So the net impact of C6 induction is nonlinear and nontrivial. We also assumed that the application of external inducers would influence the original diffusion rates of C6 and C12 going through the cell membrane to some extent (owning to the limitation of the cell membranes molecule transport). For example, C6 addition on the medium leads to a slight decrease in the diffusion rate of C12 through the cell membrane (D_h). Using the PDE model to simulate C6 application, it is predicted that we can expect a multiple GFP-mCherry ring pattern when MINPAC is induced with external C6 (Figure 3.6 Top). Comparing with the "standard set", the only changed parameter is $D_h = 3.5$.

On the other hand, the IPTG (which inhibits the repressive effect of *LacI* on *Plux/lac*) induction can modulate the strength of mutual inhibition in the circuit. IPTG counteracts *LacI*s inhibition on *Plux/lac*, leading to more *LasI* expression and intracellular C12 production. Simulating these changes by perturbing corresponding parameters, the model predicts a target-like mCherry ring with an outer GFP ring pattern, which is further verified by the experimental data (Figure 3.6 Bottom). Comparing with the “standard set”, we set $D_c = D_h = 3.5$, and $k_4 = 128$ which represented the higher C12 production rate, while $b_1 = 0.4$ and $b_2 = 0.45$ indicated slightly lower self-activation rates due to the induction.

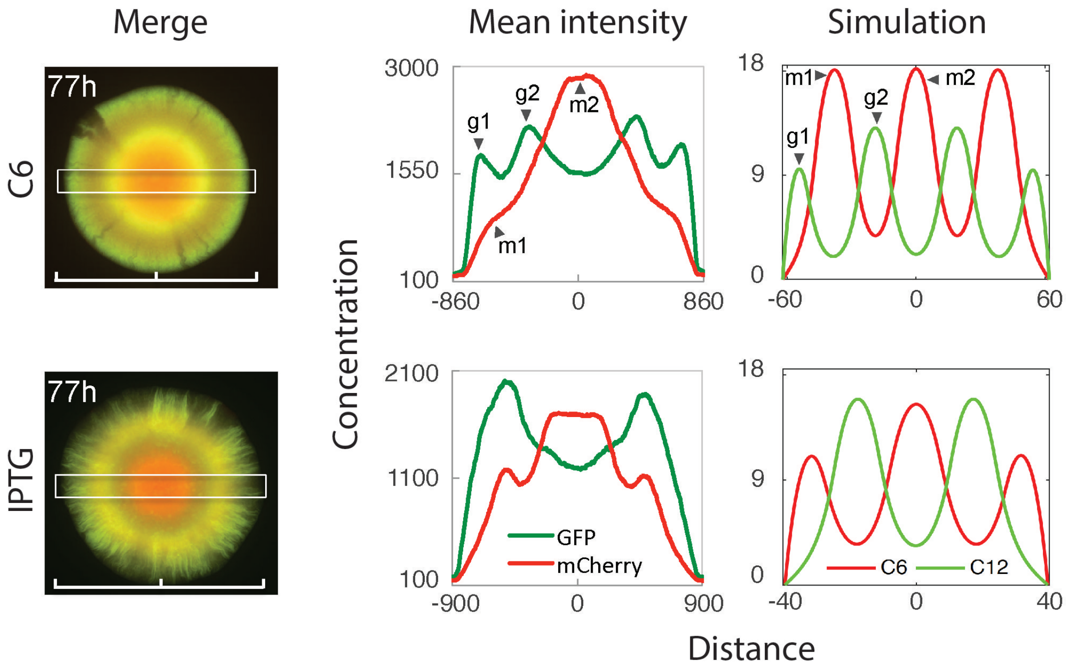

Figure 3.6: Two observations generated by MINPAC circuit and the corresponding simulation results - Figure taken from Wu *et al.* (2019). Top: a multi-ring pattern with mCherry core under external inducers C6; bottom: a multi-ring pattern with mCherry core under external IPTG. Middle: mean fluorescence intensities across the center of the corresponding colonies. Right: Model simulations recapitulate experimental patterns.

The above simulations and experiment results suggest that the ring patterns we observed are the outcomes of the spatiotemporal interaction of oscillatory dynamics owing to the network topology and the movement stemming from the diffusion process. After a short time, the solution of the reaction-diffusion system approaches the form of a traveling wave. The traveling-wave-like solution will move forward at a speed asymptotically constant. If the speed is about one unit of length per unit of time, then the wavefront resembles the mirror image of an oscillatory trajectory of the reaction system with a small initial value. A faster wave speed will stretch such oscillatory trajectory while a slower wave speed will compact it. The multiple peaks of such oscillatory trajectories give rise to the observed ring patterns. With our PDE model, we simulated the temporal dynamics of C12 on the spatial scale and new peaks emerged periodically at the wavefront (Figure 3.7 B) Experimentally, ring patterns with multiple stripes were also observed sequentially by time-lapse imaging of large colonies (Figure 3.7 A), as the model predicted. Collectively, these results suggest that the ring patterns we observed are the outcomes of the spatiotemporal interaction of oscillatory dynamics owing to the network topology and the movement stemming from the diffusion process.

3.5 Control Systems

The above study gives the following naive question: what is the necessary condition for circuit topology to generate the sequential ring patterns?

Although it is still challenging for us to find and verify the necessary condition experimentally, we can still argue whether MINPAC has a relatively concise topology that can drive the sequential ring patterns by comparing it with other modified circuits, which are called “control systems”.

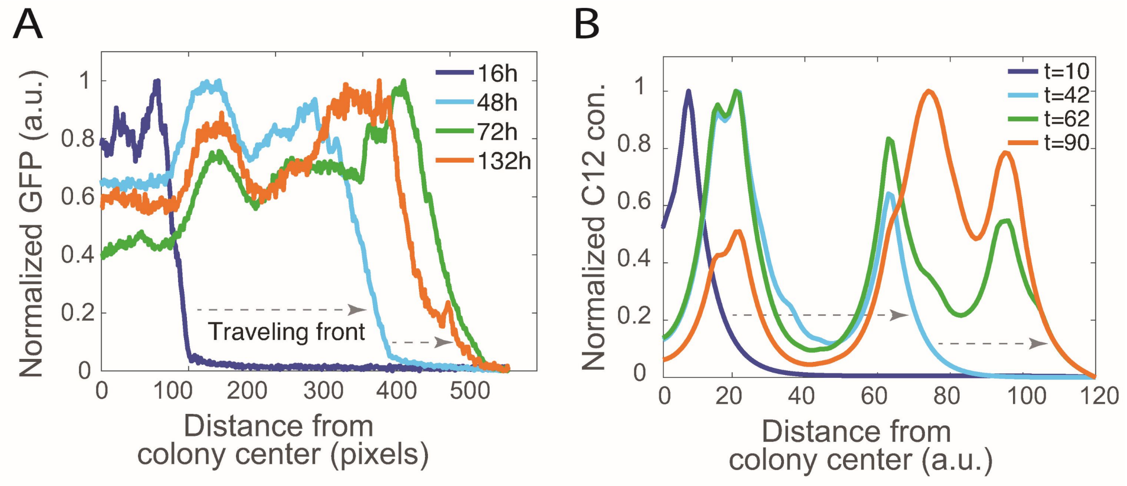

Figure 3.7: **A**. Quantified temporal and spatial fluorescence intensities of the multiple GFP ring-forming colony. **B**. Normalized external C12 concentration, directly correlated with experimental GFP intensities, of a pattern-growing colony with time and space from the PDE model simulation. Starting from the center of a colony, colored curves represent C12 concentrations along the colony radius at different time points. Grey arrows indicate the traveling direction of the wavefront - Figure taken from Wu *et al.* (2019).

3.5.1 Model Introduction

Since the circuit directed cell-cell communication is established as a viable strategy to generate RD-based and tunable patterns, we employ this method to study the fundamental question of relationship between gene network topology and resulting multicellular pattern. Wu *et al.* (2019) first designed a perturbed MINPAC topology, where the intercellular $U - A$ communication modules are replaced by intercellular auto-activations of U and A (Figure 3.8 A). Thus we only need to change the reaction

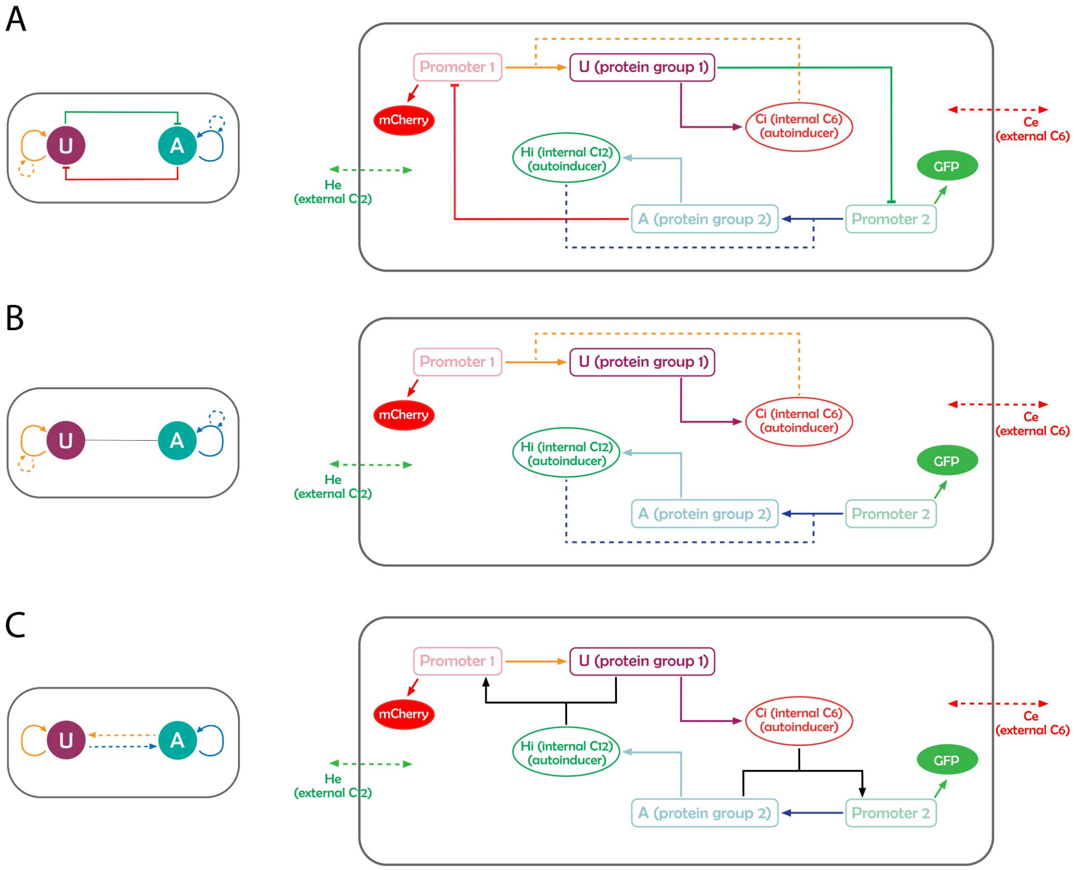

Figure 3.8: Three control circuits abstract diagrams (left column) and topology (right column) - Figure adapted from Wu *et al.* (2019). **A**. The intercellular $U - A$ communications are replaced by intercellular auto-activation of U and A. No specific pattern is observed experimentally. **B**. The mutual inhibition is further removed. **C**. The mutual inhibition is removed but keeping the other regulatory edges.

parts for U and A based on specific circuit topology which takes the following form:

$$\frac{\partial U}{\partial t} = \beta_1 + \underbrace{\frac{k_1(UC_i)^{n_1}}{1+(UC_i)^{n_1}}}_{\text{intercellular auto-activations}} \frac{1}{1+A^{m_1}} - d_1 U \tag{3.2a}$$

$$\frac{\partial A}{\partial t} = \beta_2 + \underbrace{\frac{k_2(AH_i)^{n_2}}{1+(AH_i)^{n_2}}}_{\text{intercellular auto-activations}} \frac{1}{1+U^{m_2}} - d_2 A \tag{3.2b}$$

In the second control circuit, the mutual inhibition module is removed compared to the first control circuit, namely a circuit with two positive feedback motifs (Figure 3.8 B). So the model can be written as:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1(UC_i)^{n_1}}{1+(UC_i)^{n_1}} - d_1 U \tag{3.3a}$$

$$\frac{\partial A}{\partial t} = \beta_2 + \frac{k_2(AH_i)^{n_2}}{1+(AH_i)^{n_2}} - d_2 A \tag{3.3b}$$

The third control is a sub-network of MINPAC, where the mutual inhibition is removed but having all the other regulatory edges (Figure 3.8 C). So the model can be described as:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1(UH_i)^{n_1}}{1+(UH_i)^{n_1}} - d_1 U \tag{3.4a}$$

$$\frac{\partial A}{\partial t} = \beta_2 + \frac{k_2(AC_i)^{n_2}}{1+(AC_i)^{n_2}} - d_2 A \tag{3.4b}$$

It is worth noting that although the parameter symbols in the three control circuits are the same to MINPAC model, their values may be different, especially for the production rates of U (k_1) and A (k_2) and promoter leakages (β_1 and β_2) because of the distinct architectures and molecular regulations on the promoters. For example, the basal expression in the second circuit (Figure 3.8 B) should be larger than in the first circuit (Figure 3.8 A) and MINPAC circuit (Figure 3.2) owning to a lack of repressors and direct positive autoregulation from U and A.

3.5.2 Simulation Results

The experimental result of the first control circuit (Figure 3.9 A) shows that although there is still autoinducer diffusion, this circuit mitigates the interactions and dependency between U and A and would remarkably change the intrinsic dynamics. Our model simulation shows that there is no specific pattern but a reddish colony, which is consistent with the microscopy image (Figure 3.9 A, third column). In the second control circuit (Figure 3.9 B), since the mutual inhibition module is removed,

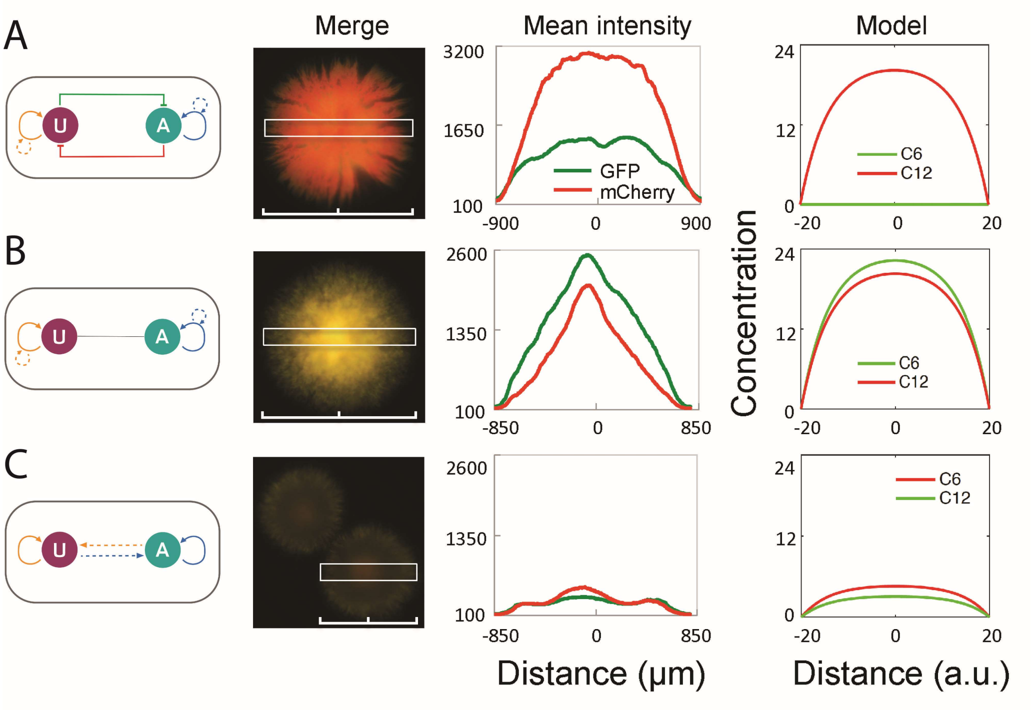

Figure 3.9: Three control circuits topology and directed patterns - Figure adapted from Wu *et al.* (2019). All the circuits are constructed with the same molecular components in MINAPC. First column: abstract diagrams of three control circuits. Second column: microscopy images taken at 72 hr. Third column: Mean fluorescence intensities across the center of the ring patterns. Fourth column: Model simulations of the three control circuits show consistency to experimental results. **A**. No specific pattern is observed experimentally. **B**. Strong GFP and mCherry are simultaneously expressed and merged fluorescence is yellow. **C**. A weak yellow core and the outer ring is observed.

the circuit consists of two positive feedback motifs. Our model shows that both C6 and C12 expressions are very strong and close (Figure 3.9 A, fourth column), which is verified by the experimental data showing strong GFP and mCherry expression, and a yellow fluorescent colony was observed. Lastly, in the third control circuit (Figure 3.9 C), which has a mutual-activation topology, provided a weak yellow target-like ring pattern with low GFP and mCherry expression. The corresponding model confirms the observation by showing that the system has low-C6 and low-C12 expression. This result is consistent with previous theoretical studies that demonstrated that a mutual-activation circuit with autoregulation is multistable, and harbors a big parameter space for low-low state (Guantes and Poyatos, 2008; Del Vecchio *et al.*, 2017). Taken together, our model can provide a systematic description of these three control circuits under corresponding modifications. It is also necessary to point out that we only changed the production rates of U (k_1) and A (k_2) and promoter leakages (β_1 and β_2) in simulations due to the distinct architectures and molecular regulations, all other parameters are consistent with previous simulations.

3.6 Quantitative Analysis

The comparison with different control models implies that MINPAC topology contains the necessary structure of generating ring patterns, which has also been numerically verified and mathematically modeled. To further investigate how MINPAC directs the generation of ring patterns, we carried out the following study about two naive questions based on our observations: 1. Is MINPAC a robust topology that can generate ring patterns? 2. With same experimental conditions, why does the number of ring varies? Corresponding analysis based on system dynamics will deepen our understanding of controlling the circuit function.

3.6.1 Balance Test

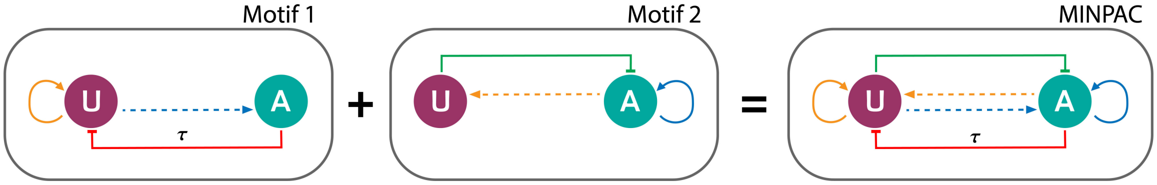

Figure 3.10: Illustration of the MINPAC composition of two symmetric positive-plus-negative oscillator motifs. Parameter τ is used to describe the strength of one negative feedback (node U inhibits node A). - Figure adapted from Wu *et al.* (2019)

From a network topology point of view, MINPAC is composed of two topologically equivalent motifs where a self-activating node activates the other node and it, in turn, inhibits the self-activating node (Figure 3.10), each forming a robust positive-plus-negative oscillator topology (Tsai *et al.*, 2008; Danino *et al.*, 2010; Chen *et al.*, 2015). Here, we compared their capability in generating ring patterns by examining the robustness of their oscillating reaction part (i.e. the ODE part), which could drive an organized pattern formation across the expanding colony. It is easy to know that a fully symmetric MINPAC topology would rapidly go to stable steady states without oscillation and to numerically describe an asymmetric MINPAC topology, we introduce a new parameter τ to control the strength of the negative feedback from A to U (Figure 3.10), which also models the imbalance between the two motifs:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1 (UH_i)^{n_1}}{1 + (UH_i)^{n_1}} \frac{1}{1 + \tau A^{m_1}} - d_1 U \tag{3.5}$$

Since C6 and C12 are small molecules that diffuse fast comparing to the colony expansion, and the volume of the autoinducer is very small comparing to the *E. coli* cell. Therefore we can regard the intercellular diffusion (diffusion through the cell membrane) as a transient process comparing to pattern formation. We apply this approximation by assuming $C := C_i = C_e, \quad H := H_i = H_e$. Thus, the reaction part of the simplified full system with parameter τ takes the following form:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1 (UH)^{n_1}}{1 + (UH)^{n_1}} \frac{1}{1 + \tau A^{m_1}} - d_1 U \tag{3.6a}$$

$$\frac{\partial A}{\partial t} = \beta_2 + \frac{k_2 (AC)^{n_2}}{1 + (AC)^{n_2}} \frac{1}{1 + U^{m_2}} - d_2 A \tag{3.6b}$$

$$\frac{\partial C}{\partial t} = \frac{k_3 U^{n_3}}{K_c^{n_3} + U^{n_3}} - d_3 C \tag{3.6c}$$

$$\frac{\partial H}{\partial t} = \frac{k_4 A^{n_4}}{K_h^{n_4} + A^{n_4}} - d_4 H \tag{3.6d}$$

while the simplified model of motif 1 (Figure 3.10) is:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1(UH)^{n_1}}{1+(UH)^{n_1}}\frac{1}{1+\tau A^{m_1}} - d_1U \tag{3.7a}$$

$$\frac{\partial A}{\partial t} = \beta_2 + \frac{k_2C^{n_2}}{1+C^{n_2}} - d_2A \tag{3.7b}$$

$$\frac{\partial C}{\partial t} = \frac{k_3U^{n_3}}{K_c^{n_3}+U^{n_3}} - d_3C \tag{3.7c}$$

$$\frac{\partial H}{\partial t} = \frac{k_4A^{n_4}}{K_h^{n_4}+A^{n_4}} - d_4H \tag{3.7d}$$

we then comparing their capability of generating ring patterns by simulating the oscillatory expression of C6 and C12 with different values of τ. To maintain the parameter consistency, we used the parameters in Figure 3.5, and kept both initial boundary conditions the same for (3.6) and (3.7). Our model-comparison results show that oscillation from one-motif topology is highly dependent on the strength of its negative feedback (τ) and more importantly, the two-motif MINPAC harbors greater robustness and amplitude against parameter perturbations to generate temporal oscillation (Figure 3.11). Such robustness enhances the likelihood of observing our desired phenotypic outputs from the synthetic gene circuit.

3.6.2 Limit Cycle

Meanwhile, even a macroscopic reaction-diffusion system could still be highly sensitive to initial conditions due to the nonlinearity of the network interactions, evidenced by diverse patterns shown in Figure 3.5, some colonies self-organize into a reversed double-ring pattern with GFP accumulating in the inner ring and mCherry on the outer ring (top). A more complicated pattern is also observed, in which two GFP rings alternating with two mCherry rings, forming a multiple GFP-mCherry ring pattern (bottom). Given that these different patterns emerge from the same MINPAC circuit operating in the same cells and under the same conditions, we hypothesize that it is due to random variations of the initial concentrations of intracellular proteins

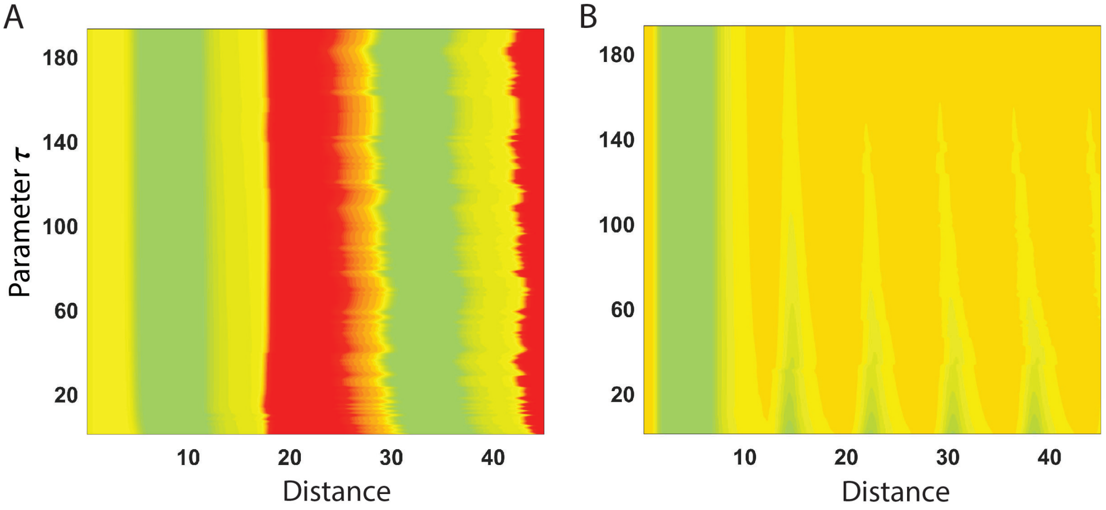

Figure 3.11: Simulation results of the full system (**A**) and motif 1 (**B**) with different values of parameter τ. The X-axis represents the distance to colony center, while Y-axis refers to parameter τ values. The color is corresponding to the concentration of C6 and C12.

and autoinducers. To computationally test this hypothesis, we tested various initial conditions of the PDE but kept all the parameters the same. The model indeed reproduces the experimental patterns (Figure 3.5). Furthermore, these differences in the patterns suggest the system is not at a steady state and, instead, is evolving towards a steady state. The initial condition determines the starting point of the MINPAC system, which will go through a temporal non-oscillating spiral (blue line in Figure 3.12) and finally approach oscillation periods (starting from the red curve in Figure 3.12). The choice of the initial point (i.e. initial conditions) can significantly affect the length of the temporal non-oscillating spiral, and quantitative simulations show that the oscillatory system, with different initial points, could require significantly different times, so-called Poincaré return time, to approach the first stable limit cycle (Figure 3.12). Thus, the initial condition and resulting approach-time variances lead to diverse patterns with different stripes (besides colony size). These results illustrate that initial conditions play an important role in shaping the formation of biological patterns, which is consistent with recent theoretical analysis (Gomez and Arcak, 2017; Salazar-Ciudad *et al.*, 2000).

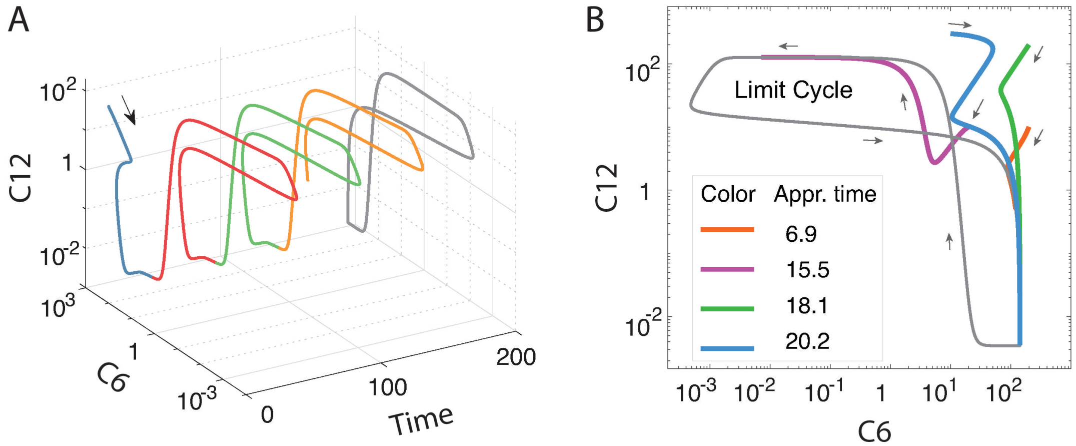

Figure 3.12: Limit cycle of the system. - Figure taken from Wu *et al.* (2019). **A**. A trajectory of a random initial point (black arrow) going to oscillation periods (red, green and yellow curves) simulated from MINPAC reaction term. The grey "butterfly" curve illustrates the limit cycle. **B**. Approaching time for different initial conditions. Colored curve shows the trajectory before stable oscillations and the approaching time is calculated for the solution going from its starting point to the stable limit cycle (grey curve).

3.7 Traveling Wave Analysis of a Simplified Two-dimensional Model

3.7.1 Simplification

Since C6 and C12 are small molecules that diffuse fast comparing to the colony expansion, and the volume of the autoinducer is very small comparing to the *E. coli* cell. Therefore we can regard the intercellular diffusion (diffusion through the cell membrane) as a transient process comparing to pattern formation. We reapply the approximation that $C := C_i = C_e, \quad H := H_i = H_e$, which leads to the following simplified model:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1(UH)^{n_1}}{1+(UH)^{n_1}}\frac{1}{1+A^{m_1}} - d_1U \tag{3.8a}$$

$$\frac{\partial A}{\partial t} = \beta_2 + \frac{k_2(AC)^{n_2}}{1+(AC)^{n_2}}\frac{1}{1+U^{m_2}} - d_2A \tag{3.8b}$$

$$\frac{\partial C}{\partial t} = \frac{k_3U^{n_3}}{K_c^{n_3}+U^{n_3}} - d_3C + D_1\frac{\partial^2 C}{\partial x^2} \tag{3.8c}$$

$$\frac{\partial H}{\partial t} = \frac{k_4A^{n_4}}{K_h^{n_4}+A^{n_4}} - d_4H + D_2\frac{\partial^2 H}{\partial x^2}. \tag{3.8d}$$

Furthermore, based on the experimental results (Figure 3.3), there was no obvious ring pattern observed until 24 hours, while a clear circular bacterial colony was observed since 15 hours. The time gap between these two observations suggests that the pattern formation process will take longer time than the cell reproduction and colony formation, which makes the quasi-steady state assumption plausible. Here we assume:

$$0 = \beta_1 + \frac{k_1(UH)^{n_1}}{1+(UH)^{n_1}}\frac{1}{1+A^{m_1}} - d_1U \tag{3.9a}$$

$$0 = \beta_2 + \frac{k_2(AC)^{n_2}}{1+(AC)^{n_2}}\frac{1}{1+U^{m_2}} - d_2A \tag{3.9b}$$

Therefore, we can solve U and A as functions of C and H: denote $U = g_1(C, H)$ and $A = g_2(C, H)$. And (3.8c)-(3.8d) can be rewrite as follow:

$$\frac{\partial C}{\partial t} = \frac{k_3g_1^{n_3}}{K_c^{n_3}+g_1^{n_3}} - d_3C + D_1\frac{\partial^2 C}{\partial x^2} \tag{3.10a}$$

$$\frac{\partial H}{\partial t} = \frac{k_4g_2^{n_4}}{K_h^{n_4}+g_2^{n_4}} - d_4H + D_2\frac{\partial^2 H}{\partial x^2} \tag{3.10b}$$

noticing that n_3 and n_4 are hill coefficients and without lossing the nonlinearity, we take $n_3 = n_4 = 1$. And for simplicity, we denote $C := u$, $H := v$, then (3.10) will take the following form:

$$\frac{\partial u}{\partial t} = \frac{k_3g_1(u,v)}{K_c+g_1(u,v)} - d_3u + D_1\frac{\partial^2 u}{\partial x^2} \tag{3.11a}$$

$$\frac{\partial v}{\partial t} = \frac{k_4g_2(u,v)}{K_h+g_2(u,v)} - d_4v + D_2\frac{\partial^2 v}{\partial x^2} \tag{3.11b}$$

With the given assumptions, we can now discuss the above reduced model that focuses on the spatial dynamics of two autoinducers, which are the key factors of pattern formation. In addition, all parameters are positive to keep the model biologically realistic.

We then nondimensionlize the system using the characteristic length $\sqrt{D_1/d_3}$ and the characteristic time $1/d_1$ so that $x = \sqrt{D_1/d_3}\hat{x}$ and $t = \hat{t}/d_3$, which leads to:

$$\frac{\partial u}{\partial \hat{t}} = \frac{K_3 \cdot g_1(u,v)}{K_c + g_1(u,v)} - u + \frac{\partial^2 u}{\partial \hat{x}^2} \tag{3.12a}$$

$$\frac{\partial v}{\partial \hat{t}} = \frac{K_4 \cdot g_2(u,v)}{K_h + g_2(u,v)} - dv + D\frac{\partial^2 v}{\partial \hat{x}^2} \tag{3.12b}$$

where $K_3 = k_3/d_3$, $K_4 = k_4/d_3$, $d = d_4/d_3$ and $D = d_3 D_2/D_1$.

3.7.2 A Special Case

Since the expression of U governs the production of C_i, higher C_i concentration will corresponding to higher U concentration. Meanwhile, higher U concentration will result in stronger inhibition to A. Therefore we can assume that $g_1(u,.)$ is an increasing function and $g_2(u,.)$ is a decreasing function. Similarly, we assume that $g_1(.,v)$ is a decreasing function and $g_2(.,v)$ is an increasing function. Without losing the biologically realistic meaning of the model, $g_1(u,v)$ and $g_2(u,v)$ are all positive and bounded for $u, v \in (0, +\infty)$ and $g_1(0,v) = g_2(u,0) = 0$.

Here we consider the following g_1 and g_2 functions which satisfiy the above conditions and also maintain the nonlinearity:

$$g_1(u,v) = \frac{u}{r_1 u + v + r_3} \tag{3.13a}$$

$$g_2(u,v) = \frac{v}{r_2 v + u + r_4} \tag{3.13b}$$

substituting (3.13) into (3.12) and drop the hat gives the following system:

$$\frac{\partial u}{\partial t} = \frac{k_5 u}{k_7 u + v + r_3} - u + \frac{\partial^2 u}{\partial x^2} \tag{3.14a}$$

$$\frac{\partial v}{\partial t} = \frac{k_6 v}{k_8 v + u + r_4} - dv + D\frac{\partial^2 v}{\partial x^2} \tag{3.14b}$$

noticing that for convenience, we denote $K_3/K_c = k_5$, $K_4/K_h = k_6$, $(K_c r_1 + 1)/K_c = k_7$ and $(K_h r_2 + 1)/K_h = k_8$. Since k_5/r_3 and k_6/r_4 refer to the maximum expression rates of C6 and C12, the positivity of $(k_5 - r_3)$ and $(k_6 - dr_4)$ avoids C6 and C12 to have an exponential-decay profile, which is not consistent with the experimental results, it is necessary to assume that $(k_5 > r_3)$ and $(k_6 > dr_4)$.

It is worthy to point out that the above reaction-diffusion system can be regarded as a plausible simplification of the original model, and its dynamical behavior can still demonstrate the biological mechanism.

3.7.3 Minimal Wave Speed

We seek a traveling wave solution of the form:

$$u(\xi) = u(x - ct), \quad v(\xi) = v(x - ct), \quad \xi = x - ct$$

where $c \geq 0$ is the wave speed of the traveling wave, functions u and v are defined on the interval $(-\infty, +\infty)$. Substitution of above forms into (3.14) gives the following system:

$$0 = \frac{k_5 u}{k_7 u + v + r_3} - u + \frac{d^2 u}{d\xi^2} + c\frac{du}{d\xi} \tag{3.15a}$$

$$0 = \frac{k_6 v}{k_8 v + u + r_4} - dv + D\frac{d^2 v}{d\xi^2} + c\frac{dv}{d\xi} \tag{3.15b}$$

Then linearising the system ahead of the wave (about $u = v = 0$) and assuming $u(\xi) = Ae^{-r\xi}$, $v(\xi) = Be^{-r\xi}$ to leading order, leads to:

$$\left(r^2 - cr + \frac{k_5}{r_3} - 1\right) A = 0 \tag{3.16a}$$

$$\left(Dr^2 - cr + \frac{k_6}{r_4} - d\right) B = 0 \tag{3.16b}$$

since u, v must be positive, we require A, B to be positive and r to be real, so that $c \geq 2\sqrt{k_5/r_3 - 1}$ and $c \geq 2\sqrt{D(k_6/r_4 - d)}$. By giving the condition that $(k_5 - r_3)$ and $(k_6 - dr_4)$ are all positive, the minimum speed of the wave is:

$$c_{min} = \min\left(2\sqrt{\frac{k_5}{r_3} - 1}, 2\sqrt{D(\frac{k_6}{r_4} - d)}\right) \tag{3.17}$$

The comparison of the theoretical minimal wave speed and simulated wave speed is shown in Figure 3.13, which illustrated the theoretical minimal wave speed as a reasonable lower bound.

To obtain an approximate wave profile, we use the method introduced by Canosa (1973). Rescaling the wave coordinate as $z = -\xi/c$, which leads to:

$$\frac{k_5 u}{k_7 u + v + r_3} - u + \frac{1}{c^2}\frac{d^2u}{dz^2} - \frac{du}{dz} = 0 \tag{3.18a}$$

$$\frac{k_6 v}{k_8 v + u + r_4} - dv + \frac{D}{c^2}\frac{d^2v}{dz^2} - \frac{dv}{dz} = 0 \tag{3.18b}$$

we can further simplify the above system by assuming that $1/c^2$ is very small and obtain the reduced system:

$$\frac{du}{dz} = \frac{k_5 u}{k_7 u + v + r_3} - u \tag{3.19a}$$

$$\frac{dv}{dz} = \frac{k_6 v}{k_8 v + u + r_4} - dv \tag{3.19b}$$

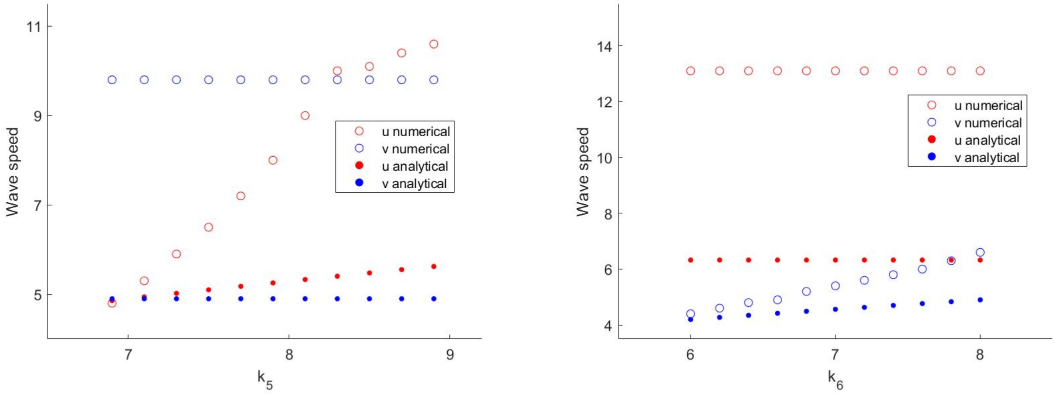

Figure 3.13: Comparison of simulated wave speed and theoretical minimum wave speed when varying parameter k_5 and k_6 with all other parameters values fixed. The simulated wave speed is shown in the circles while the theoretical minimum wave speed is shown in the solid dots. The fixed parameters are: $r_3 = 1$, $r_4 = 1$, $d = 0.5$, $D_1 = 1$, and $D_2 = 0.8$. Left case: $k_6 = 8$, $k_7 = 6$, $k_8 = 10$. Right case: $k_5 = 11$, $k_7 = 12$, $k_8 = 5$.

3.7.4 Equilibrium Points and Nullclines of the Reduced System

Similar to Stepien *et al.* (2018), we discuss the equilibrium points and nullclines of the reduced system (3.19). The components of its Jacobian are:

$$J_{11}(u,v) = \frac{k_5(v+r_3)}{(k_7u+v+r_3)^2} - 1 \tag{3.20a}$$

$$J_{12}(u,v) = -\frac{k_5u}{(k_7u+v+r_3)^2} \tag{3.20b}$$

$$J_{21}(u,v) = -\frac{k_6v}{(k_8v+u+r_4)^2} \tag{3.20c}$$

$$J_{22}(u,v) = \frac{k_6(u+r_4)}{(k_8v+u+r_4)^2} - d \tag{3.20d}$$

and thus at the equilibrium point $(u_1^*, v_1^*) = (0,0)$, the Jacobian is

$$J(u_1^*, v_1^*) = J(0,0) = \begin{pmatrix} \frac{k_5-r_3}{r_3} & 0 \\ 0 & \frac{k_6-dr_4}{r_4} \end{pmatrix} \tag{3.21a}$$

the eigenvalues of $J(0,0)$ are positive,

$$\lambda_1^1 = \frac{k_5 - r_3}{r_3} > 0, \quad \lambda_1^2 = \frac{k_6 - dr_4}{r_4} > 0 \tag{3.22}$$

implying that the equilibrium point $(0,0)$ is a source.

With the following condition

$$k_5 - r_3 > \frac{k_6 - dr_4}{k_8 d} \tag{3.23}$$

$$\frac{k_6 - dr_4}{d} > \frac{k_5 - r_3}{k_7} \tag{3.24}$$

we can determine the existence and local stability of the other three equilibrium points by computing their Jacobian:

$$J(u_2^*, v_2^*) = J\left(0, \frac{k_6 - dr_4}{k_8 d}\right) = \begin{pmatrix} \frac{k_5}{v_2^* + r_3} - 1 & 0 \\ -d\frac{v_2^*}{k_8 v_2^* + r_4} & \frac{dr_4}{k_8 v_2^* + r_4} - d \end{pmatrix} \tag{3.25}$$

$$J(u_3^*, v_3^*) = J\left(\frac{k_5 - r_3}{k_7}, 0\right) = \begin{pmatrix} \frac{r_3}{k_7 u_3^* + r_3} - 1 & -\frac{u_3^*}{k_7 u_3^* + r_3} \\ 0 & \frac{k_6}{u_3^* + r_4} - d \end{pmatrix} \tag{3.26}$$

$$\begin{aligned} J(u_4^*, v_4^*) &= J\left(\frac{k_8(k_5 - r_3) - \frac{1}{d}(k_6 - dr_4)}{k_7 k_8 - 1}, \frac{\frac{k_7}{d}(k_6 - dr_4) - (k_5 - r_3)}{k_7 k_8 - 1}\right) \\ &= \begin{pmatrix} -\frac{k_7 u_4^*}{k_7 u_4^* + v_4^* + r_3} & -\frac{u_4^*}{k_7 u_4^* + v_4^* + r_3} \\ -\frac{d v_4^*}{k_8 v_4^* + u_4^* + r_4} & -\frac{d k_8 v_4^*}{k_8 v_4^* + u_4^* + r_4} \end{pmatrix} \end{aligned} \tag{3.27}$$

from (3.23)-(3.24) we know that the eigenvalues of $J(u_2^*, v_2^*)$ are:

$$\begin{aligned} \lambda_2^1 &= \frac{k_5}{v_2^* + r_3} - 1 = \frac{1}{v_2^* + r_3}\left(k_5 - r_3 - \frac{k_6 - dr_4}{k_8 d}\right) > 0, \\ \lambda_2^2 &= \frac{dr_4}{k_8 v_2^* + r_4} - d = -\frac{k_8 v_2^*}{k_8 v_2^* + r_4} < 0 \end{aligned} \tag{3.28}$$

and similarly, the eigenvalues of $J(u_3^*, v_3^*)$ are:

$$\begin{aligned} \lambda_3^1 &= \frac{r_3}{k_7 u_3^* + r_3} - 1 = -\frac{k_7 u_3^*}{k_7 u_3^* + r_3} < 0, \\ \lambda_3^2 &= \frac{k_6}{u_3^* + r_4} - d = \frac{1}{u_3^* + r_4}\left(k_6 - dr_4 - \frac{d(k_5 - r_3)}{k_7}\right) > 0 \end{aligned} \tag{3.29}$$

meanwhile, since $k_7 k_8 > 1$ we have $tr(J(u_4^*, v_4^*)) < 0, \quad det(J(u_4^*, v_4^*)) < 0$, which implies that $J(u_4^*, v_4^*)$ has two negative eigenvalues $\lambda_4^{1,2} < 0$.

Based on above discussion, we can conclude that the equilibrium point $(0, 0)$ is a source, while two equilibrium points (u_2^*, v_2^*) and (u_3^*, v_3^*) are saddle, and the interior equilibrium point (u_4^*, v_4^*) is a sink.

Next, we derive the nullclines of the system. The vertical motion nullclines are given by the condition

$$\frac{k_5 u}{k_7 u + v + r_3} - u = 0 \tag{3.30}$$

thus we have two vertical motion nullclines:

$$l_1 : u = 0 \quad and \quad l_2 : v = -k_7 u + (k_5 - r_3) \tag{3.31}$$

similarly we can derive the horizontal motion nullclines:

$$l_3 : v = 0 \quad and \quad l_4 : v = -\frac{1}{k_8}u + \frac{k_6 - dr_4}{dk_8} \tag{3.32}$$

and it is easy to show that the intersection of l_2 and l_4 is (u_4^*, v_4^*).

3.7.5 Positively Invariant Region and Phase Portrait

Lemma 3.7.1. *If $k_5 > r_3$ and $k_6 > dr_4$, let Ω be the open region bounded by the lines $\{(u, v) : u = 0\}$, $\{(u, v) : v = 0\}$ and $\{(u, v) : v = -\frac{d(k_5 - r_3)}{k_6 - dr_4}u + (k_5 - r_3)\}$ (Figure 3.14. Dashed line). Then Ω is positively invariant.*

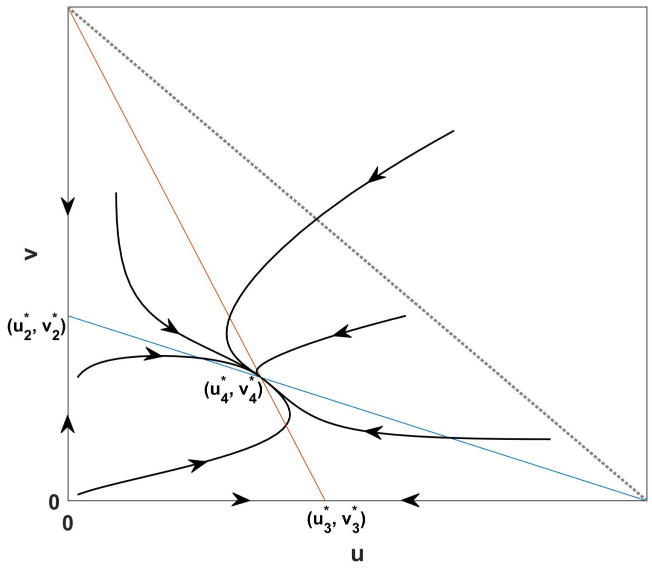

Figure 3.14: Phase portrait of the system. The red line refers to l_2 and blue line refers to l_4.

Proof. Along the line $\{(u,v) : u = 0\}$ we have $du/dz = 0$ and

$$\frac{dv}{dz} > 0 \ \ when \ \ v \in (0, \frac{k_6 - dr_4}{dk_8}), \quad \frac{dv}{dz} < 0 \ \ when \ \ v \in (\frac{k_6 - dr_4}{dk_8}, \infty) \tag{3.33}$$

thus the vertical motion along this axis will move towards the equilibrium point (u_2^*, v_2^*) and no flow will across the line.

Along the line $\{(u,v) : v = 0\}$ we have $dv/dz = 0$ and

$$\frac{du}{dz} > 0 \ \ when \ \ u \in (0, \frac{k_5 - r_3}{k_7}), \quad \frac{du}{dz} < 0 \ \ when \ \ u \in (\frac{k_5 - r_3}{k_7}, \infty) \tag{3.34}$$

thus the vertical motion along this axis will move towards the equilibrium point (u_3^*, v_3^*), and no flow will across the line.

Along the line $\{(u,v) : v = -\frac{d(k_5-r_3)}{k_6-dr_4}u + (k_5 - r_3)\}$ we have:

$$\begin{aligned}
\frac{du}{dz} &= \frac{u}{k_7u+v+r_3}\left((k_5-r_3) - k_7u + \frac{d(k_5-r_3)}{k_6-dr_4}u - (k_5-r_3)\right) \\
&= \frac{u}{k_7u+v+r_3}\frac{dk_7u}{k_6-dr_4}\left(-\frac{k_6-dr_4}{d} + \frac{k_5-r_3}{k_7}\right) < 0
\end{aligned} \tag{3.35a}$$

$$\begin{aligned}
\frac{dv}{dz} &= \frac{v}{k_8v+u+r_4}\left((k_6-dr_4) - dk_8v + \frac{k_6-dr_4}{k_5-r_3}v - (k_6-dr_4)\right) \\
&= \frac{u}{k_7u+v+r_3}\frac{dk_8v}{k_5-r_3}\left(\frac{k_6-dr_4}{dk_8} - (k_5-r_3)\right) < 0
\end{aligned} \tag{3.35b}$$

so the flow is down and to the left across the line. Also notice that here is no motion at (u_2^*, v_2^*) , (u_3^*, v_3^*) nor $(0,0)$. At the corner $(0, k_5-r_3)$, the flow is directly downward and at the corner $((k_6-dr_4)/d, 0)$ the flow is directly leftward. Thus Ω is positively invariant. □

In order to determine the phase portrait of the system, we discuss the motion along the line l_2 (Figure 3.14 Red line) and l_4 (Figure 3.14 Blue line). Along the line l_4, we have $dv/dz = 0$ and:

$$\begin{aligned}
\frac{du}{dz} &= \frac{u}{k_7u+v+r_3}\left((k_5-r_3) - k_7u - v\right) \\
&= \frac{u}{k_7u+v+r_3}\left((k_5-r_3) - \frac{k_6-dr_4}{k_8d} - \frac{k_7k_8-1}{k_8}u\right) \\
&= \frac{u}{k_7u+v+r_3}\frac{k_7k_8-1}{k_8}\left(u_4^* - u\right)
\end{aligned} \tag{3.36a}$$

so on the top left part ($u \in (0, u_4^*)$), the flow across l_4 from the left, and on the bottom right part ($u \in (u_4^*, (k_6-dr_4)/d)$), the flow across l_4 from the right. Similarly we know that the flow across l_2 from the left when ($u \in (0, u_4^*)$), and across l_2 from the right when ($u \in (u_4^*, u_3^*)$)

3.7.6 No Periodic Orbits

Let $f_1(u,v)$ be the right-hand side of (3.19a) and $f_2(u,v)$ be the right-hand side of (3.19b). If we define

$$B(u,v) = \frac{1}{uv}, \tag{3.37}$$

where $u, v \in \Omega \setminus (\{u : u = 0\} \cup \{v : v = 0\})$. Then

$$\frac{\partial(Bf_1)}{\partial u} + \frac{\partial(Bf_2)}{\partial v} = -\frac{k_5 k_7}{v(k_7 u + v + r_3)^2} - \frac{k_6 k_8}{u(k_8 v + u + r_4)^2} < 0, \tag{3.38}$$

so the expression does not change sign in almost everywhere of the invariant region Ω. Based on the Dulac's criterion, we can conclude that there are no periodic orbits within the closed positively invariant region Ω.

Recall that there is an interior sink point (u_4^*, v_4^*) within the positively invariant region Ω, there is a traveling wave solution connecting the $(0,0)$ and (u_4^*, v_4^*).

3.8 Traveling Wave Analysis of a Simplified Four-dimensional Model

Model (3.8) is a plausible description of the MINPAC circuit, which may provide us reasonable mathematical explanations of some experimental observations. In this section, we first simplified (3.8), which maintains the nonlinearity and the systematical description of MINPAC, and then we present some preliminary results of its traveling wave analysis.

3.8.1 Simplification

Since n_i $(i = 1, 2, 3, 4)$ and m_j $(j = 1, 2)$ are all hill coefficients and without lossing the nonlinearity, we take $n_i = 1$ $(i = 1, 2, 3, 4)$ and $m_j = 1$ $(j = 1, 2)$, which lead to:

$$\frac{\partial U}{\partial t} = \beta_1 + \frac{k_1 UH}{1+UH}\frac{1}{1+A} - d_1 U \tag{3.39a}$$
$$\frac{\partial A}{\partial t} = \beta_2 + \frac{k_2 AC}{1+AC}\frac{1}{1+U} - d_2 A \tag{3.39b}$$
$$\frac{\partial C}{\partial t} = \frac{k_3 U}{K_c + U} - d_3 C + D_1 \frac{\partial^2 C}{\partial x^2} \tag{3.39c}$$
$$\frac{\partial H}{\partial t} = \frac{k_4 A}{K_h + A} - d_4 H + D_2 \frac{\partial^2 H}{\partial x^2}. \tag{3.39d}$$

Next, we nondimensionlize the system by denoting $x = \sqrt{D_1/d_3}\hat{x}$ and $t = \hat{t}/d_3$, and the above system can be rewritten as follow:

$$\frac{\partial U}{\partial \hat{t}} = b_1 + \frac{K_1 UH}{1+UH}\frac{1}{1+A} - d_5 U \tag{3.40a}$$
$$\frac{\partial A}{\partial \hat{t}} = b_2 + \frac{K_2 AC}{1+AC}\frac{1}{1+U} - d_6 A \tag{3.40b}$$
$$\frac{\partial C}{\partial \hat{t}} = \frac{K_3 U}{K_c + U} - C + \frac{\partial^2 C}{\partial \hat{x}^2} \tag{3.40c}$$
$$\frac{\partial H}{\partial t} = \frac{K_4 A}{K_h + A} - d_7 H + D \frac{\partial^2 H}{\partial \hat{x}^2}, \tag{3.40d}$$

where $b_i = \beta_i/d_3,\ K_i = k_i/d_3\ i = (1, 2),\ d_5 = d_1/d_3,\ d_6 = d_2/d_3,\ d_7 = d_4/d_3$ and $D = d_3 D_2/D_1$ are all positive.

3.8.2 Minimal Wave Speed

Assuming the traveling wave solution has variable $\xi = \hat{x} - c\hat{t}$, where $c \geq 0$ is the wave speed of the traveling wave, and solutions are defined on the interval $(-\infty, +\infty)$.

Substituting $\xi = \hat{x} - c\hat{t}$ into (3.40) gives the following system:

$$
\begin{aligned}
0 &= b_1 + \frac{K_1 UH}{1+UH}\frac{1}{1+A} - d_5 U + c\frac{\partial U}{\partial \xi} && \text{(3.41a)}\\
0 &= b_2 + \frac{K_2 AC}{1+AC}\frac{1}{1+U} - d_6 A + c\frac{\partial A}{\partial \xi} && \text{(3.41b)}\\
0 &= \frac{K_3 U}{K_c + U} - C + \frac{\partial^2 C}{\partial \xi^2} + c\frac{\partial C}{\partial \xi} && \text{(3.41c)}\\
0 &= \frac{K_4 A}{K_h + A} - d_7 H + D\frac{\partial^2 H}{\partial \xi^2} + c\frac{\partial H}{\partial \xi}, && \text{(3.41d)}
\end{aligned}
$$

We linearising the system ahead of the wave, which refers to the steady state $(U^*, A^*, C^*, H^*) = (b_1/d_5, b_2/d_6, 0, 0)$, and the following linearized system can be derived:

$$
\begin{aligned}
c\frac{\partial U}{\partial \xi} &= d_5(U - \frac{b_1}{d_5}) - B_1 H && \text{(3.42a)}\\
c\frac{\partial A}{\partial \xi} &= d_6(A - \frac{b_2}{d_6}) - B_2 C && \text{(3.42b)}\\
\frac{\partial^2 C}{\partial \xi^2} + c\frac{\partial C}{\partial \xi} &= -B_3(U - \frac{b_1}{d_5}) + C && \text{(3.42c)}\\
D\frac{\partial^2 H}{\partial \xi^2} + c\frac{\partial H}{\partial \xi} &= -B_4(A - \frac{b_2}{d_6}) + d_7 H, && \text{(3.42d)}
\end{aligned}
$$

where

$$
\begin{aligned}
B_1 &= \frac{K_1 U^*}{1+A^*} = \frac{d_6}{d_5}\frac{K_1 b_1}{b_2 + d_6} && \text{(3.43a)}\\
B_2 &= \frac{K_2 A^*}{1+U^*} = \frac{d_5}{d_6}\frac{K_2 b_2}{b_1 + d_5} && \text{(3.43b)}\\
B_3 &= \frac{K_3 K_c}{(K_c + U^*)^2} = \frac{K_3 K_c}{(K_c + \frac{b_1}{d_5})^2} && \text{(3.43c)}\\
B_4 &= \frac{K_4 K_h}{(K_h + A^*)^2} = \frac{K_4 K_h}{(K_h + \frac{b_2}{d_6})^2}, && \text{(3.43d)}
\end{aligned}
$$

and we assume the solutions take the form:

$$U(\xi) = x_1 e^{-\gamma\xi} + \frac{b_1}{d_5} \tag{3.44a}$$

$$A(\xi) = x_2 e^{-\gamma\xi} + \frac{b_2}{d_6} \tag{3.44b}$$

$$C(\xi) = x_3 e^{-\gamma\xi} \tag{3.44c}$$

$$H(\xi) = x_4 e^{-\gamma\xi}, \tag{3.44d}$$

substiting (3.44) into (3.42) will derive the following system:

$$0 = (d_5 + c\gamma)x_1 - B_1 x_4 \tag{3.45a}$$

$$0 = (d_6 + c\gamma)x_2 - B_2 x_3 \tag{3.45b}$$

$$0 = -B_3 x_1 + (-\gamma^2 + c\gamma + 1)x_3 \tag{3.45c}$$

$$0 = -B_4 x_2 + (-D\gamma^2 + c\gamma + d_7)x_4. \tag{3.45d}$$

(3.45) is a linear equations system which can be written as $M\vec{x} = \vec{0}$, where $\vec{x} = (x_1, x_2, x_3, x_4)^T$ and

$$M = \begin{pmatrix} c\gamma + d_5 & 0 & 0 & -B_1 \\ 0 & c\gamma + d_6 & -B_2 & 0 \\ -B_3 & 0 & -c\gamma^2 + c\gamma + 1 & 0 \\ 0 & -B_4 & 0 & -Dc\gamma^2 + c\gamma + d_7 \end{pmatrix} \tag{3.46}$$

knowing that for any $c \geq 0$, $rank(M) = 4$, so system (3.45) always has real solutions. Since x_i $(i = 1, 2, 3, 4)$ are all positive, the following two conditions mush hold:

$$-c\gamma^2 + c\gamma + 1 > 0 \tag{3.47a}$$

$$-Dc\gamma^2 + c\gamma + d_7 > 0. \tag{3.47b}$$

Meanwhile, in order to have a traveling wave solution, we shall find conditions to ensure (U^*, A^*, C^*, H^*) is a stable steady state. For simplification we first define $\hat{U} = U - \frac{b_1}{d_5}$ and $\hat{A} = A - \frac{b_2}{d_6}$, and denote $U := x_1$, $A := x_2$, $C := x_3$, $H := x_4$, we then introduce $\frac{\partial C}{\partial \xi} := x_3$ and $\frac{\partial H}{\partial \xi} := x_6$, then (3.42) can be rewrite as following:

$$\frac{\partial x_1}{\partial \xi} = \frac{d_5}{c}x_1 - \frac{B_1}{c}x_4 \tag{3.48a}$$

$$\frac{\partial x_2}{\partial \xi} = \frac{d_6}{c}x_2 - \frac{B_2}{c}x_3 \tag{3.48b}$$

$$\frac{\partial x_3}{\partial \xi} = x_5 \tag{3.48c}$$

$$\frac{\partial x_4}{\partial \xi} = x_6 \tag{3.48d}$$

$$\frac{\partial x_5}{\partial \xi} = -B_3x_1 + x_3 - cx_5 \tag{3.48e}$$

$$\frac{\partial x_6}{\partial \xi} = -\frac{B_4}{D}x_2 + \frac{d_7}{D}x_4 - \frac{c}{D}x_6. \tag{3.48f}$$

The Jacobian matrix of (3.48) is:

$$J = \begin{pmatrix} \frac{d_5}{c} & 0 & 0 & -\frac{B_1}{c} & 0 & 0 \\ 0 & \frac{d_6}{c} & -\frac{B_2}{c} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -B_3 & 0 & 1 & 0 & -c & 0 \\ 0 & -\frac{B_4}{D} & 0 & \frac{d_7}{D} & 0 & -\frac{c}{D} \end{pmatrix} \tag{3.49}$$

by computing the eigenvalue of J we can derive the conditions for stability and the upper bound of γ, since it is bounded by the spectral radius of J:

$$\gamma \leq max\{\mid \lambda_i \mid, i = (1,2,3,4,5,6)\}. \tag{3.50}$$

3.9 Local Stability Analysis of the Linearized Systems

Based on system (3.8), we assume $n_1 = n_2 = n_3 = n_4 = m_1 = m_2 = 1$, which will not lose the nonlinearity. In addition, we assume that the self-activate rates are very small, which means $\beta_1 = \beta_2 = 0$. Thus we can derive following linearization:

$$\frac{k_1(UH)^{n_1}}{1+(UH)^{n_1}} = \frac{k_1(UH)}{1+(UH)} \doteq \gamma(UH). \tag{3.51a}$$

$$\frac{k_2(AC)^{n_2}}{1+(AC)^{n_2}} = \frac{k_2(AC)}{1+(AC)} \doteq \mu(AC). \tag{3.51b}$$

$$\frac{k_3U^{n_3}}{K_c^{n_3}+U^{n_3}} = \frac{k_3U}{K_c+U} \doteq \alpha U. \tag{3.51c}$$

$$\frac{k_4A^{n_4}}{K_h^{n_4}+A^{n_4}} = \frac{k_4A}{K_h+A} \doteq \beta A. \tag{3.51d}$$

where α, β, γ and μ are all positive constants.

3.9.1 Full System

The linearized full system model takes the following form:

$$\frac{\partial U}{\partial t} = \frac{\gamma(UH)}{1+A} - d_1U \tag{3.52a}$$

$$\frac{\partial A}{\partial t} = \frac{\mu(AC)}{1+U} - d_2A \tag{3.52b}$$

$$\frac{\partial C}{\partial t} = \alpha U - d_3C + D_1\frac{\partial^2 C}{\partial x^2} \tag{3.52c}$$

$$\frac{\partial H}{\partial t} = \beta A - d_4H + D_2\frac{\partial^2 H}{\partial x^2}. \tag{3.52d}$$

then we can study the equilibrium points of system (3.52) by considering the reduced ODEs system:

$$\frac{\partial U}{\partial t} = \frac{\gamma(UH)}{1+A} - d_1U \tag{3.53a}$$
$$\frac{\partial A}{\partial t} = \frac{\mu(AC)}{1+U} - d_2A \tag{3.53b}$$
$$\frac{\partial C}{\partial t} = \alpha U - d_3C \tag{3.53c}$$
$$\frac{\partial H}{\partial t} = \beta A - d_4H. \tag{3.53d}$$

To make the system biological realistic, we have the condition that $\min\{\gamma, \mu, \alpha, \beta\} \leq \max\{d_1, d_2, d_3, d_4\}$ which ensures the positive equilibrium point $E^* = (U^*,\ A^*,\ C^*,\ H^*)$ where

$$U^* = \frac{d_2d_3}{\alpha\mu - d_2d_3} \tag{3.54a}$$
$$A^* = \frac{d_1d_4}{\gamma\beta - d_1d_4} \tag{3.54b}$$
$$C^* = \frac{\alpha d_2}{\alpha\mu - d_2d_3} \tag{3.54c}$$
$$H^* = \frac{\beta d_1}{\beta\gamma - d_1d_4}. \tag{3.54d}$$

the Jacobian matrix at E^* is

$$J(E^*) = \begin{pmatrix} 0 & -\frac{d_1U^*}{1+A^*} & 0 & \frac{d_1d_4}{\beta}\frac{U^*}{A^*} \\ -\frac{d_2A^*}{1+U^*} & 0 & \frac{d_2d_3}{\alpha}\frac{A^*}{U^*} & 0 \\ \alpha & 0 & -d_3 & 0 \\ 0 & \beta & 0 & -d_4 \end{pmatrix} \tag{3.55}$$

And the *characteristic polynomial* $F(\lambda)$ is:

$$\begin{aligned} F(\lambda) = \lambda^4 + (d_3 + d_4)\lambda^3 + \left(d_3 d_4 - \frac{d_1 d_2 U^* A^*}{(1 + A^*)(1 + U^*)}\right)\lambda^2 \\ + \frac{D\left(\frac{A^*}{d_4} + \frac{U^*}{d_3}\right)}{(1 + A^*)(1 + U^*)}\lambda - \frac{D}{(1 + A^*)(1 + U^*)} \end{aligned} \tag{3.56}$$

where $D = d_1 d_2 d_3 d_4$.

We can show that the Jacobian matrix of this system at steady state has a unique positive (real) eigenvalue and at least one negative (real) eigenvalue. And the system will cause transitional oscillation if there exists complex eigenvalues.

3.9.2 Control System 1

Similarly, we can derive the reduced ODEs system:

$$\frac{\partial U}{\partial t} = \gamma(UH) - d_1 U \tag{3.57a}$$

$$\frac{\partial A}{\partial t} = \mu(AC) - d_2 A \tag{3.57b}$$

$$\frac{\partial C}{\partial t} = \alpha U - d_3 C \tag{3.57c}$$

$$\frac{\partial H}{\partial t} = \beta A - d_4 H. \tag{3.57d}$$

The positive equilibrium point E_1^* can be easily calculated:

$$E_1^* = (U^*,\, A^*,\, C^*,\, H^*) = (\frac{d_2 d_3}{\alpha\mu},\, \frac{d_1 d_4}{\gamma\beta},\, \frac{d_2}{\mu},\, \frac{d_1}{\gamma}) \tag{3.58}$$

the Jacobian matrix at E_1^* is

$$J(E_1^*) = \begin{pmatrix} 0 & 0 & 0 & \frac{\gamma d_2 d_3}{\alpha\mu} \\ 0 & 0 & \frac{\mu d_1 d_4}{\gamma\beta} & 0 \\ \alpha & 0 & -d_3 & 0 \\ 0 & \beta & 0 & -d_4 \end{pmatrix} \tag{3.59}$$

And the *characteristic polynomial* $F_1(\lambda)$ is:

$$F_1(\lambda) = \lambda^4(\lambda + d_3)(\lambda + d_4) - D \tag{3.60}$$

where $D = d_1 d_2 d_3 d_4$.

By assuming that $d_1 = d_2 = s > 0$, $d_3 = d_4 = t > 0$, the above equation has the following solutions:

$$\lambda_{1,2} = \frac{1}{2}(-t \pm \sqrt{t(4s + t)}) \tag{3.61a}$$

$$\lambda_{3,4} = \frac{1}{2}(-t \pm \sqrt{t^2 - 4st)}) \tag{3.61b}$$

which means the Jacobian matrix of this system at steady state will have four real eigenvalues (one positive and three negative) when $t > 4s$, otherwise it will have two real eigenvalues (one positive and one negative) and two complex eigenvalues with negative real part when $t < 4s$.

3.9.3 Control System 2

The reduced ODEs model for the second control system is:

$$\frac{\partial U}{\partial t} = \gamma(UC) - d_1 U \tag{3.62a}$$

$$\frac{\partial A}{\partial t} = \mu(AH) - d_2 A \tag{3.62b}$$

$$\frac{\partial C}{\partial t} = \alpha U - d_3 C \tag{3.62c}$$

$$\frac{\partial H}{\partial t} = \beta A - d_4 H. \tag{3.62d}$$

The positive equilibrium point E_2^* can be easily calculated:

$$E_2^* = (U^*,\, A^*,\, C^*,\, H^*) = (\frac{d_1 d_3}{\alpha\mu},\, \frac{d_2 d_4}{\gamma\beta},\, \frac{d_1}{\mu},\, \frac{d_2}{\gamma}) \tag{3.63}$$

the Jacobian matrix at E_2^* is

$$J(E_2^*) = \begin{pmatrix} 0 & 0 & \frac{\gamma d_1 d_3}{\alpha} & 0 \\ 0 & 0 & 0 & \frac{\mu d_2 d_4}{\beta} \\ \alpha & 0 & -d_3 & 0 \\ 0 & \beta & 0 & -d_4 \end{pmatrix} \tag{3.64}$$

And the *characteristic polynomial* $F_2(\lambda)$ is:

$$F_2(\lambda) = \lambda^4 + (d_3 + d_4)\lambda^3 + (d_3 d_4 - d_2 d_4 - d_1 d_3)\lambda^2 - (\frac{D}{d_1} + \frac{D}{d_2})\lambda + D \tag{3.65}$$

where $D = d_1 d_2 d_3 d_4$. And the above equation has the following solutions:

$$\lambda_{1,2} = \frac{1}{2}(-d_3 \pm \sqrt{d_3(4d_1 + d_3)}) \tag{3.66a}$$

$$\lambda_{3,4} = \frac{1}{2}(-d_4 \pm \sqrt{d_4(4d_2 + d_4)}) \tag{3.66b}$$

which means the jaccobian matrix of this system at steady state will always have four real eigenvalues. And the steady state will be either node or saddle, which will not cause transitional oscillation.

3.9.4 Control System 3

The reduced ODEs model for the third control system is:

$$\frac{\partial U}{\partial t} = \frac{\gamma(UC)}{1 + A} - d_1 U \tag{3.67a}$$

$$\frac{\partial A}{\partial t} = \frac{\mu(AH)}{1 + U} - d_2 A \tag{3.67b}$$

$$\frac{\partial C}{\partial t} = \alpha U - d_3 C \tag{3.67c}$$

$$\frac{\partial H}{\partial t} = \beta A - d_4 H \tag{3.67d}$$

To make the system biological realistic, we have the condition that $E = \gamma\mu\alpha\beta > D = d_1 d_2 d_3 d_4$ which ensures the positive equilibrium point $E_3^* = (U^*,\ A^*,\ C^*,\ H^*)$ where

$$U^* = \frac{\mu\beta d_1 d_3 + D}{E - D} \tag{3.68a}$$

$$A^* = \frac{\alpha\gamma d_2 d_4 + D}{E - D} \tag{3.68b}$$

$$C^* = \frac{\alpha}{d_3}\frac{\mu\beta d_1 d_3 + D}{E - D} \tag{3.68c}$$

$$H^* = \frac{\beta}{d_4}\frac{\alpha\gamma d_2 d_4 + D}{E - D} \tag{3.68d}$$

and the Jacobian matrix at E_3^* is

$$J(E_3^*) = \begin{pmatrix} 0 & -\frac{d_1^2 d_3}{\alpha\gamma} & \frac{d_1 d_3}{\alpha} & 0 \\ -\frac{d_2^2 d_4}{\beta\mu} & 0 & 0 & \frac{d_2 d_4}{\beta} \\ \alpha & 0 & -d_3 & 0 \\ 0 & \beta & 0 & -d_4 \end{pmatrix} \tag{3.69}$$

And the *characteristic polynomial* $F_3(\lambda)$ is:

$$\begin{aligned} F_3(\lambda) = \lambda^4 + (d_3 + d_4)\lambda^3 + (d_3 d_4 - d_2 d_4 - d_1 d_3 - \frac{D}{E} d_1 d_2)\lambda^2 \\ - (d_2 d_3 d_4 + d_1 d_3 d_4 + \frac{D}{E}(d_1 d_2 d_4 + d_1 d_2 d_3))\lambda + D - \frac{D^2}{E}, \end{aligned} \tag{3.70}$$

and the eigenvalues are the solutions of the following equation:

$$\lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4 = 0, \tag{3.71}$$

where

$$a_1 = d_3 + d_4 > 0 \tag{3.72a}$$

$$a_2 = d_3d_4 - d_2d_4 - d_1d_3 - \frac{D}{E}d_1d_2 \tag{3.72b}$$

$$a_3 = -(d_2d_3d_4 + d_1d_3d_4 + \frac{D}{E}(d_1d_2d_4 + d_1d_2d_3)) < 0 \tag{3.72c}$$

$$a_4 = D - \frac{D^2}{E} > 0. \tag{3.72d}$$

Noting that the values of d_i $(i = 1, 2, 3, 4)$ were all fixed in our simulations:

$$(d_1, d_2, d_3, d_4) = (1.19, 1.19, 0.56, 0.8), \tag{3.73}$$

we take this parameter set as an example and set $D = 3$ and $E = 6$, we can compute the eigenvalues: $\lambda_{1,2} = -1.4975 \pm 0.3849i$, $\lambda_3 = 1.0194$, and $\lambda_4 = 0.6155$.

This result also shows that the system with above given parameter values will cause transitional oscillation.

3.10 Numerical Methods

According to the previous discussion, the simulations were generated by different programs or even different platforms. In this section, we will introduce the numerical methods based on each simulation plot.

Part of Figure 3.4 (top and the bottom left) was generated by *NetLogo* (a programming language and integrated development environment for agent-based modeling). The simulation starts with one single cell at the center of the domain while cells were set to move every 10 units of time and divide every 20 units of time. The "movement" of the cell, which also represented the diffusion of the C6 and C12, was defined by the following two steps: "turning head at a random degree" and "moving forward one step". The reaction part of the model (3.1) was used to compute the concentration of C6 and C12 inside the cell and determine the cell's color. Other parameters were fixed beforehand. The whole process included circuit expression (which determined the cell's color) and cell movement (diffusion). Multiple rounds of simulations were performed and the average results were provided by *MATLAB* (Figure 3.4 Bottom right). In this part, the computational domain was the uniform grid, the location of each cell was determined by its center. The color of each grid block was determined by the most color cell.

Figure 3.5, Figure 3.6 and Figure 3.9 were generated from similar programs via *MATLAB*, thus we only discuss Figure 3.5 as an example. We first numerically solved the PDE model (3.1) with *Pdepe* package in *MATLAB*, which has been used to solve initial-boundary value problems for systems of parabolic and elliptic PDEs in the one spatial variable x and time t. The model contains two parts, ODE part (3.1a)-(3.1d) and PDE part (3.1c)-(3.1f). We first used *ode45* package to solve the ODE term ($U,\ A,\ C_i,\ H_i$) with defined initial conditions and given parameter sets. Then we use the solutions of U and A to solve the PDE part ($C_i,\ H_i,\ C_e,\ H_e$) by using the *bvp5c* package. Here we assume that the range of X-axis is large enough so that all variables are 0 at the boundary, thus the initial condition for the PDE part is $[1, 1, 0, 0]$.

Figure 3.7 B was generated by *MATLAB*, we solved model (3.1) by using the *Pdepe* package with zero boundary conditions, given initial conditions and parameter sets. Then we selected multiple time points and ploted the normalized C12 concentration.

Figure 3.11 was generated by *MATLAB*. We solved the reaction part of the model by using *ode45* package with given initial conditions and parameter values. Let the X-axis represents the distance to the colony center, we determined the color at each point by comparing the concentrations of C_i and H_i. Then we plotted the results under different values of parameter τ. Therefore we could compare the robustness of the oscillation presented by each model.

Figure 3.12 was generated by *MATLAB*. We solved the reaction part of the model by using *ode45* package with given initial conditions and parameter values. By running the simulation long enough, we obtained the limit cycle of the system. For a random initial point, we first found the time point that the solution reaches (i.e. close enough to) the limit cycle and then the Poincaré return time was also recorded. For different initial points, we compared their corresponding Poincare return time and illustrated that the longer Poincare return time will lead to later oscillation which can result in fewer rings.

3.11 Dicsuccion and Conclusion

Biological pattern formation requires complex gene regulation networks and accurate cell-cell coordination. Indeed, coordinated cell population behavior in response to self-regulated morphogen kinetics is a common phenomenon in development (Kondo and Miura, 2010; Kicheva *et al.*, 2012; Scholes and Isalan, 2017). Here, we introduced a novel synthetic gene network MINPAC which is capable of directing engineered single cells to form self-organized tunable patterns with multiple rings. A reaction-diffusion-based mathematical model (the full-system model) was developed to describe the gene expression and autoinducers' diffusion, in order to simulate the pattern formation process. The model simulations showed that we were able to predict the ring patterns for bacterial colonies growing under different experimental conditions by tuning the corresponding parameters or initial conditions. By comparing the traveling wave profile of the solution with the experimental data, we confirmed that our model is able to simulate the wave profile of the autoinducers' concentrations at different time points, which directly causes the ring patterns.

To further investigate the mechanisms of how MINPAC drives the pattern formation, three control circuits were designed and their mathematical models were also established by modifying the full-system model. All of the control models provided simulations that agreed well with our experimental observations. These results show that our model can not only systematically describe the whole gene network topology, but is also plausible to be applied to similar circuits with reasonable modifications, which implies that this modeling work can potentially be a significant tool for studying the mechanisms and control of synthetic gene circuits.

Besides simulating the pattern formation process, the model also provided valuable insights into the system dynamics. Since MINPAC is composed of two topologically equivalent motifs, and each forming a robust positive-plus-negative oscillator topology. By comparing the model simulations of the full-system with one sub-motif with parameter perturbations, it is clear that MINPAC is a much more robust topology to generate temporal oscillation than the sub-motif. Furthermore, we noticed that the number of rings varied in different sets of experiments with the same growing conditions, which could also be biologically interesting. By examining the system's phase portrait, we found that this variation can be explained by the change of Poincare return time, which mainly depends on the location of the initial point.

In addition, we also provided mathematical analysis of the model with reasonable simplifications. Although we haven't done a full-system analysis yet, we were still able to perform the traveling wave analysis of the reduced reaction-diffusion model, and local stability analysis of all models' reaction parts. This analysis work showed the connections between pattern formation and circuit topology and could shed light on the molecular mechanisms of somitogenesis and biological pattern formation.

Since the MINPAC circuit is novel and the mechanism of circuit-driven pattern formation is still underexplored, our work in this project is unique and can be regarded as a modeling framework for studying spatial pattern formation. Our PDE model simulations and experimental measurements strongly support that the observed ring patterns are driven by a reaction-diffusion-based oscillatory gene network with propagating wavefront, the so-called clock and wavefront mechanism. It is noteworthy to point out that we used one single PDE model to recapitulate and predict all the MINPAC-directed biological patterns. Furthermore, we maintained the parameter consistency in all the simulations, including different sets of experiments and different circuit topologies, as well as for the quantitative analysis. This could be extremely challenging due to the complexity of the model and the randomness of the experimental results. The success of doing so also demonstrated the rationality of our modeling approach, which could be applied to broad topics and contribute to a better understanding of the natural developmental processes, and facilitate the engineering of synthetic tissues in the future.

Synthetic gene circuit has become one of the most important applications in synthetic biology research, and due to its complexity, the progress in modeling the gene network and predicting the circuit function is still not enough for us to obtain full control of the system. There are lots of underexplored topics such as the interactions between the gene circuit and the host cell, which could exert a non-neglectable impact on the system. In the next chapter, I will introduce my work on nutrient-modulating growth feedback from the host cell to circuit expression.

Chapter 4

NUTRIENT-MODULATING *E. COLI* GROWTH FEEDBACK IN MULTI-NUTRIENT CULTURE

4.1 Abstract

The study of circuit-host interaction has always been fundamental and crucial throughout the development of synthetic biology. The unwanted interactions can exert an adverse effect on gene circuits and sometimes may cause the circuit failure to function as expected. Meanwhile, how to systematically present the interaction by the mathematical model is still underexplored. In this paper, we focus on the growth-mediated feedback between *E. coli* and synthetic gene circuit in different growth environments. We first introduced our previous dynamical model with a non-monotonous function presenting the interaction. Our results showed that the model can predict the oscillatory expression of the circuit. Furthermore, we modify the model with a multi-nutrient Droop equation to describe the cell's response to different nutrient components. And the model is capable of capturing both the behaviors of the circuit and host. In addition, some dynamical analysis of the model is also provided. Our study suggests that the growth-mediated feedback can be represented by non-monotonous function and the host organism's behavior in multi-nutrient environments can be described by the extended Droop model.

4.2 Introduction

4.2.1 Circuit-cell Interaction

During the study of engineered biology systems, understanding and predicting their behaviors become critical but essential (Blanchard *et al.*, 2018). One of the major challenges is the underexplored circuit-host interactions, which may occur in different ways including cell growth, nutrient competition, metabolic burden, and can potentially impact the circuit behaviors. Understanding how the interactions are established and affect the gene circuit will largely help us to design and control robust gene circuits. Recent studies have provided many valuable insights into the circuit-host interactions for both controlling them in the experiments and predicting the impact quantitatively. For example, available tools and principles for avoiding unwanted interactions when engineer microbial gene circuits have been reviewed by Bradley *et al.* (2016). And a combination of deterministic modeling, stochastic simulation, and Fokker-Planck equation formalism has been applied in studying the roles of circuit-host coupling in shaping circuit dynamics by Blanchard et al. Blanchard *et al.* (2018). Nevertheless, in practice, studying the quantitative behaviors and mechanisms of the interactions usually have higher operability than eliminating them when engineering the circuits. Thus, mathematical modeling has become an important methodology, due to its power of systematic representing and quantitative predicting of biological systems (Lee *et al.*, 2007). Recent studies of Zhang *et al.* (2020) and Melendez-Alvarez *et al.* (2021) on how growth-mediated feedback affects the behavior of gene circuits are good examples.

However, the experimental data (which provided by Tian's lab) also shows that if the liquid culture is the mixture of two different nutrient mediums (which refer to LB and M9 in this paper), the two-stage growth profile of the cell is observed (Figure 4.4

blue dots). Differs to the cell growth in multiple limiting nutrients, this two-stage growth profile may cause by the cell's nutrient preference since the cells have a higher growth rate during the first growth stage. Meanwhile, the mixture of the two different cultures makes it almost impossible to model the consumption of these two cultures separately. In all, this is a rarely considered problem that can be clearly observed from experiments and could lead to very interesting mathematical problems. Also, the study of the nutrient preference can help us understand the cell's dynamics in the multi-nutrient environment and hence improve the experimental design on controlling the cell's growth rate.

In this study, we first consider nonlinear growth feedback and propose a modified model based on Zhang *et al.* (2020). This modified model is capable of capturing both the circuit and host behaviors under different nutrient conditions, and the following analysis provides fair explanations of the mechanisms and long-term behavior of the system. Thus we conclude that the growth feedback between circuit and host is more likely to be a nonlinear function. Furthermore, we proposed a modified model based on the model introduced by Melendez-Alvarez *et al.* (2021). In the new model, we considered an instantaneous trade-off of *E. coli* cells in the multi-nutrient environment, which consists of two different media LB and M9. We assumed that the *E. coli* cells would consume LB first before the instantaneous trade-off occur. Since LB is a higher quality media comparing to M9, but they are not completely different, we used a single nutrient equation with two different cell quotas representing the cell growth with LB and M9 respectively. Using this dynamical model, we look at how different nutrient conditions affect this biological system, and how the growth feedback performs when the LB to M9 proportion varies.

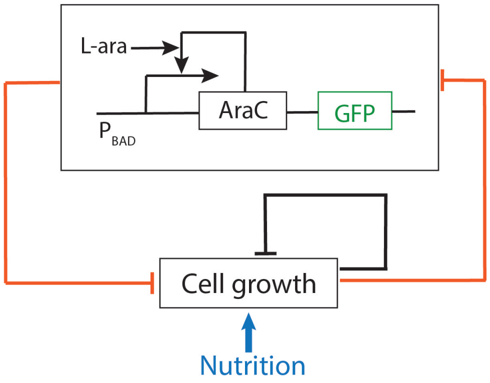

Figure 4.1: Diagram of interactions between the self-activation (SA) gene circuit and the host cell growth, which is modulated by the nutrient. In the SA circuit, transcriptional factor AraC forms a dimer and binds to promoter P_{BAD} in the presence of stimulus L-ara, and thus activates the expression of itself. Here GFP is used as the reporter. - Figure taken from Melendez-Alvarez *et al.* (2021).

4.2.2 Introducrion of the Self-activated Circuit (SA circuit) and Experiment

A simple synthetic self-activation (SA) gene circuit was built by Zhang *et al.* (2020) and also the research objective in Melendez-Alvarez *et al.* (2021), in which the transcription factor AraC activates the expression of itself by binding to its promoter P_{BAD}. Reporter gene green fluorescent protein (GFP) was used to visualize the dynamics of AraC (Figure 4.1). The design is similar to the previously reported positive feedback synthetic gene circuit (Becskei *et al.*, 2001).

Before the cells were transferred into varied nutrient culture media they were cultured with LB broth (LuriaBertani broth, Sigma-Aldrich). The experimental procedure for each biological replicate of the self-activated circuit (SA circuit) induction was carried out in the following way: On day one, SA circuit plasmid was transformed into *E. coli* cells which were grown on LB plate e with 50 μg/mL kanamycin overnight at 37°C. On day two in the morning, one colony was picked and inoculated into 400 μL LB medium with 25 g/mL chloramphenicol and was grown to OD 1.0 (measured in 200 μL volume in 96-well plate by plate reader for absorbance at 600 nm) in a 5 mL culture tube in the shaker. The cells were then diluted 1000 folds into 2 mL fresh LB medium supplemented with $1.25 \times 10^{-3}\%$ of L-ara, and grew in a 15 mL culture tube with 250rpm at 37°C for 16h. On day three, cells inducted in the last step were 100-fold diluted into each culture medium of the varied nutrient levels supplemented with the desired concentration of L-ara and antibiotics. Then, three technical replicates of 200 μL culture mix for each culture medium of the varied nutrient levels were load onto a 96-well plate, which was immediately placed onto the plate-reader to start the measurement.

4.2.3 Previous Model

A dynamic model of self-activation switch circuit with linear growth feedback is recently introduced by Zhang *et al.* (2020). This model successfully incorporates the growth feedback into the system and takes the experimental data into account, which is an important achievement in establishing a powerful method for studying circuit-host interactions. In this model, AraC is the transcriptional factor that activates the expression of itself and the reporter green fluorescent protein (GFP) in the presence of stimulus L-ara. A represents the concentration of AraC, which is co-expressed with GFP. Without the growth feedback, this self-activation switch can be modeled as

$$\frac{dA}{dt} = \underbrace{f(A)}_{\text{expression}} - \underbrace{dA}_{\text{dilution}} \tag{4.1}$$

Symbol	Definition
k_0	baisc production rate of the gene circuit
k_1	L-ara induced production rates
d	dilution rate of AraC
C_{max}, C_{min}	the maximum and minimum affinities of AraC dimers binding on the promoter, respectively
$Lara$	concentration of L-ara
n	Hill coefficient
K	Michaelis constant

Table 4.1: Parameter explanations of the model (4.1).

here the expression function $f(A)$ takes the following form:

$$f(A) = \underbrace{k_0}_{\text{basic production}} + \underbrace{k_1\left(\frac{S_aA^2}{S_aA^2+1}\right)}_{\text{L-ara induced production}} \tag{4.2a}$$

$$S_a = C_{min} + \underbrace{(C_{max} - C_{min})\left(\frac{Lara^n}{Lara^n + K^n}\right)}_{\text{activation rate of promoter}} \tag{4.2b}$$

The definition of parameters are provided in 4.1. Here $\left(\frac{Lara^n}{Lara^n+K^n}\right)$ describes the activation rate of promoter by Lara. dA measures the dilution process of AraC. Furthermore, since AraC and *E. coli* have a potential competitive relationship, a higher growth rate of *E. coli* will result in less available nutrients for AraC. So the authors consider the growth of *E. coli* will directly cause dilution of AraC, and the mathematical model is revised by coupling cell growth.

$$\frac{dA}{dt} = f(A) - dA - \underbrace{\frac{dN}{dt}\frac{1}{N}A}_{\text{dilution caused by cell growth}} \tag{4.3a}$$

$$\frac{dN}{dt} = \underbrace{k_g\left(\frac{1}{A/J+1}\right)\left(1-\frac{N}{N_{max}}\right)}_{\text{cell growth rate}} N \tag{4.3b}$$

where $N(t)$ is the *E. coli* cell density at time t and k_g, N_{max} represents the maximum growth rate and maximum cell density, respectively. The growth rate of *E. coli* is dependent on the circuit expression and J is defined as the overload parameter of the gene circuit to the growth rate. Notice that the important assumption of this model is the growth of *E. coli* will directly cause dilution of AraC, which will take the form of $\left(\frac{dN}{dt}\frac{1}{N}A\right)$. This may be true in single-nutrient culture such as M9, where either the cell or the circuit can reach a high growth/expression rate, and the dynamics of the system is relatively simple. In a more complicated environment, with well-mixed M9 and LB culture solutions, the interplay between *E. coli* growth and gene circuit may act nonlinearly, and this will aid directly in the development of our model.

4.3 Model Formulations

4.3.1 Growth-mediated Feedback Model (Model I)

Based on the observation during the experiments, we find the growth of the host does not always exert an adverse influence on the gene circuit expression. When the *E. coli* grow fast, they will consume most of the available nutrients inside the cell and may leave insufficient nutrients for the gene circuit. On the other hand, once the cell density reaches the maximum, the growth rate will maintain close to zero and cells are quiescent which implies that not enough promoters are produced by

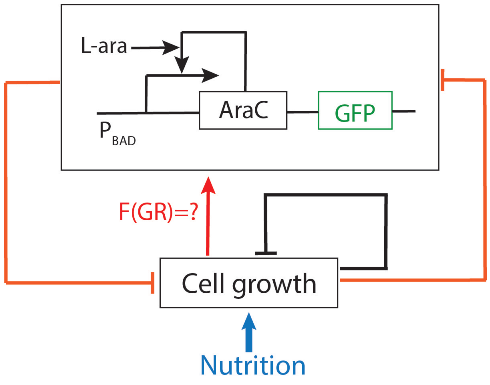

Figure 4.2: The diagram of the model by considering the regulation of the production rate of the circuit by growth rate. - Figure adapted from Melendez-Alvarez *et al.* (2021).

the cell. Otherwise the growth of *E. coli* will promote the expression of gene circuit via high available nutrient concentration and promoter production. However, the model introduced by Zhang *et al.* (2020) may not be flexible enough to allow for such dynamics, and a system with non-monotonous growth feedback is perhaps optimal because such a system can reproduce the complexity of circuit dynamics responding to the cell growth. Upon this consideration, we propose nonlinear function $F(GR)$ about cell growth rate per capita $GR = \frac{dN}{dt}\frac{1}{N}$ to describe the interference. Thus a mathematical model was developed by Melendez-Alvarez *et al.* (2021) to describe the circuit expression with the feedback which depends on the host cell's growth rate per capita, which takes the following form:

$$\frac{dA}{dt} = \underbrace{\left(k_0 + k_1\frac{S_aA^2}{S_aA^2+1}\right)}_{\text{circuit expression}} \quad \underbrace{F(GR)}_{\text{growth feedback function}} \quad -(d+GR)A \tag{4.4a}$$

$$\frac{dN}{dt} = k_g\left(\frac{1}{A/J+1}\right)\left(1-\frac{N}{N_{max}}\right)N. \tag{4.4b}$$

where the parameter definitions are the same as (4.3).

This mechanistic model provides us another tool to study the circuit-host interaction in a more precise way, especially the circuit dynamics in different phases.

However, when the M9-based culture solution is mixed with a small proportion of LB solution, *E. coli* will appear two-stage growth profile, which may not be correlated to our current model. The two-stage growth profile implies that *E. coli* cell will respond to the different nutrient environments when growing, and the quasi-steady-state argument could lead to the negligence of this change. As a result, taking the nutrient concentration as a variable and presenting the cell's response to different cultures become necessary for establishing a more comprehensive model.

4.3.2 Growth-mediated Feedback Model in Multinutrient Environment (Model II)

In a variable environment, cells will respond to the changes and can present complicated dynamics. Recent studies show that both microbes and bacteria will regulate their nutrient uptake mechanisms in order to adapt to the change of available resources, which can be presented by the extended Droop growth model (Klausmeier *et al.*, 2007; Grover, 2003). Hwa's recent work shows that the use of regulatory functions can describe the status shift (Erickson *et al.*, 2017). And Klausmeier points out that very fast or very slow acclimation of the cell may be better than intermediate speed acclimation (Klausmeier *et al.*, 2007). Since each one of our experiments is done within 24 hours and cell growth is relatively fast, we assume the responsibility to be instantaneous.

Although the main ingredients of M9 and LB cultures are fixed, at what speeds they are consumed during the experiments, or are there any specific ingredient's concentration that can reflect the response of the cells still remain unclear. Here we use variable P representing the nutrient concentration related to LB culture solutions and a higher initial value $P(0)$ means a higher proportion of LB solution. About the cell's response to the change of nutrient concentration, we assume that the adaptation to a lower P level can be described as changing of minimal cell quota. Thus a double-

Droop growth model with different cell quotas and different maximum growth rates is considered in our model. Moreover, the regulatory functions $w_1(P)$ and $w_2(P)$, which depend on the nutrient concentration, are considered to be instantaneous. Based upon these assumptions, a system of three ODEs is proposed as follow:

$$\frac{dA}{dt} = \left(k_0 + k_1 \frac{S_a A^2}{S_a A^2 + 1}\right) F(GR) - (d + GR)A \tag{4.5a}$$

$$\frac{dN}{dt} = k_g \left(\frac{1}{A/J + 1}\right) \left(w_1(P)g_1(P) + w_2(P)g_2(P)\right) N \tag{4.5b}$$

$$\frac{dP}{dt} = -\gamma(P - q_2)N. \tag{4.5c}$$

where $-\gamma(P-q_2)$ describes the nutrient uptake by *E. coli* with maximum uptake rate γ. Here $P(0) = P_0 > q_1 > q_2 > 0$ and $r_1 > r_2$, and the growth terms are functions of nutrient level:

$$g_1(P) = r_1 \left(1 - \frac{q_1}{P}\right), \tag{4.6}$$

$$g_2(P) = r_2 \left(1 - \frac{q_2}{P}\right). \tag{4.7}$$

Meanwhile, the regulatory functions should be continuous and differentiable with following conditions: $w_1'(P) \geq 0, \quad w_1(P_0) = 1$, and

$$w_1(P) = 0, \quad P \in [q_2, q_1], \tag{4.8a}$$

$$w_2(P) = 1 - w_1(P). \tag{4.8b}$$

and here we consider the feedback function as the following form:

$$F(GR) = w_1(P)F_1(GR) + w_2(P)F_2(GR), \tag{4.9}$$

where $F_1(GR)$ and $F_2(GR)$ are positive and bounded for all $GR \geq 0$, and $F_1(0) = F_2(0) = 1$ ensures that there is no growth feedback if the host cells are not growing.

4.4 Simulation

4.4.1 Growth-mediated Feedback Model (Model I)

Several functions of $F(GR)$ were tested to fit the experimental data and to find the correct phenomenological dependence of the synthetic gene production rate on growth rate. The function $F(GR)$ represents the regulation of synthesis rate of AraC by the growth rate. We concluded that the nonmonotonic function $F(GR) = (aGR^{n_1} + b)/(cGR^{n_2} + 1)$ where $n_1 < n_2$ provided the best fitting results compared to linear function and monotonic function (Figure 4.1). The best fitted parameters with this model, unless otherwise mentioned, are: $a = 0.0079$, $b = 1$, $k_0 = 0.4488$, $k_1 = 8.9770$, $S_a = 1$, $d = 4.4885$, $k_g = 0.9634$, $J = 2.8066$ for the linear case. While for the monotonic case, the best fitted parameters are $a = 727.7$, $b = 1$, $c = 277.5$, $k_0 = 0.0289$, $k_1 = 0.5779$, $S_a = 1$, $d = 0.2889$, $k_g = 0.9065$, $J = 8.1291$. And for the nonmonotonic case are: $a = 7.549$, $b = 1$, $c = 6.7002$, $n_1 = 0.5$, $n_2 = 2$, $k_0 = 0.0514$, $k_1 = 1.0288$, $S_a = 1$, $d = 0.5144$, $k_g = 1.2540$, $J = 2.1230$.

4.4.2 Growth-mediated Feedback Model in Multinutrient Environment (Model II)

To describe the instantaneous nutrient switch, here we define the regulatory functions w_1 and w_2 as following:

$$w_1(P) = 0.5(sign(P - p_1) + 1) \tag{4.10}$$

$$w_2(P) = 0.5(sign(p_1 - P) + 1), \tag{4.11}$$

which explain that in high-quiality nutrient environment ($q_1 < P < P_0$), $w_1(P) = 1$ and $w_2(P) = 0$, otherwise when $q_2 < P < q_1$, $w_1(P) = 0$ and $w_2(P) = 1$.

Also, in high-quality nutrient environment, the host cells are growing very fast which may cause negative feedback to circuit expression. Thus we introduce the

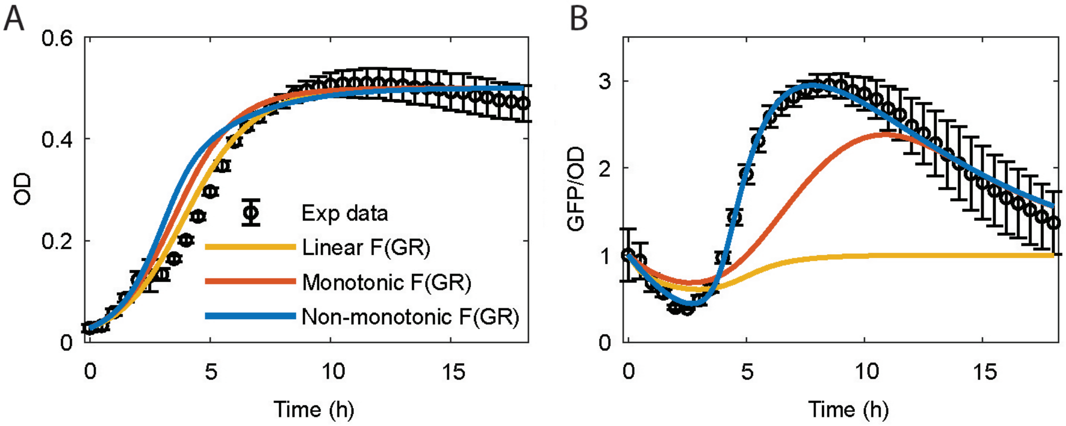

Figure 4.3: Fitting of the model to the dynamics of the host cell growth **(A)** and the circuit gene expression **(B)**. Linear (yellow lines), monotonic (orange lines), and nonmonotonic (blue lines) functions were used to test the regulation of the production rate of the gene circuit by growth rate ($F(GR)$) and only the nonmonotonic function enable the models to fit the experimental data perfectly - Figure taken from Melendez-Alvarez *et al.* (2021).

following feedback functions F_1 and F_2:

$$F_1(GR) = \frac{1}{K \cdot GR + 1} \tag{4.12}$$

$$F_2(GR) = \frac{aGR^{n_1} + b}{cGR^{n_2} + 1}. \tag{4.13}$$

The simulation results are shown in Figure 4.4, while two different data sets are used to represent different nutrient conditions. The model is capable of providing comparable fitting results with the following parameter values: $J = 2.123$, $k_0 = 0.0514$, $k_g = 1.2540$, $S_a = 1$, $d = 0.4813$, $c = 6.7$, $n_1 = 0.5$, $n_2 = 2$, $q_2 = 0.25$, which were fixed in both two simulations. For case 1 (LB proportion is 20%), $k_1 = 1.7678$, $a = 10.6323$, $\gamma = 2.2911$, $q_1 = 0.7234$, $K = 0.0862$, $r_1 = 4.6700$, $r_2 = 1.6330$, and for case 2 (LB proportion is 40%), $k_1 = 1.5842$, $a = 5.3552$, $\gamma = 1.3008$, $q_1 = 0.7247$, $K = 0.0918$, $r_1 = 4.6606$, $r_2 = 1.0499$.

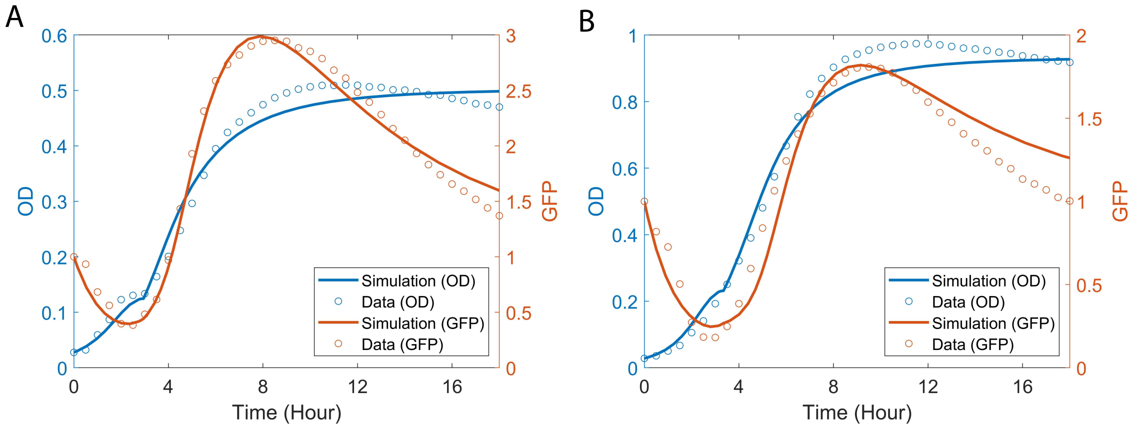

Figure 4.4: Fitting results of two different nutrient proportion. Right Y-axis represents optical density which can also be understood as cell density, while left Y-axis represents average GFP level, and X-axis is time points. Blue and red curves represent simulated results of the model and the dots are averaged experimental data. **(A)** LB proportion is 20%. **(B)** LB proportion is 40%.

4.5 Numerical Method

We used similar methods in performing data fitting for the above two models so here we take Figure 4.3 as an example. Since the self-activated circuit plasmid was transformed into *E. coli* which were grown on LB plate one day before being inoculated into the liquid culture, we deduced that the circuit expression had already reached the steady state. Thus we first used *ode23* (an ODE solver in *MATLAB*) to solve the model, the initial value of A was set to be a positive random value around the first data point of GFP/OD while the initial value of N was the first data point of OD (optical density). This step helped us finding the steady state of the circuit gene expression. Then we again solved the system by using *ode23*, this time the initial value of A was set to be its steady state while the initial value of N was still the first data point of OD. After solving the system we used the *fminsearch* search function to find parameter values that best fit the experimental data. Here the error is calculated using the least-square formula given by:

$$Error = \sum_{i=1}^{M} \left(\frac{(N_i - \hat{N}_i)}{\hat{N}_i} \right)^2 + \sum_{i=1}^{M} \left(\frac{(A_i - \hat{A}_i)}{\hat{A}_i} \right)^2 \tag{4.14}$$

where M is the total number of data points, $\hat{N}_i$ is the i^{th} data point of OD and $\hat{A}_i$ is the i^{th} data point of GFP/OD. Here N_i and A_i are generated by the model.

4.6 Analysis

4.6.1 Exploring the Growth-mediated Feedback Function (Model I)

For simplification, we introduce following notations: $A := u$, $N := v$ and $N_{max} := v_m$. Thus model (4.4) can be written as:

$$\frac{du}{dt} = \left(k_0 + k_1 \frac{S_a u^2}{S_a u^2 + 1} \right) F(GR) - (d + GR)u, \tag{4.15a}$$

$$\frac{dv}{dt} = k_g \left(\frac{1}{u/J + 1} \right) \left(1 - \frac{v}{v_m} \right) v. \tag{4.15b}$$

Proposition 1. *There is no peak of AraC (i.e. no local maximum point of u with $t \in (0, T)$, where T is the ending time) when $F(GR) = 1$.*

Proof. Prove by contradiction. Assuming there is at least one local maximum point $u_p = u(t_p)$ with $t_p \in (0, T)$, which leads to $\frac{du}{dt} \mid_{t=t_p} = 0$. Thus, there exist t_1, t_2 close to t_p such that $t_1 < t_p < t_2$ and $u_1 = u_2 < u_p$ (notice that we denote $u(t_i) := u_i$ and $v(t_i) := v_i$ for $i = 1, 2$) and

$$\frac{du}{dt} \mid_{t=t_1} > 0 > \frac{du}{dt} \mid_{t=t_2}, \tag{4.16}$$

which is equivalent to

$$f(u_1) - (d + GR \mid_{t=t_1}) u_1 > 0 > f(u_2) - (d + GR \mid_{t=t_2}) u_2. \tag{4.17}$$

Since $f(u)$ and du are monotonous about u, we have $f(u_1) = f(u_2)$ and $du_1 = du_2$, which means the following inequality should also hold:

$$GR \mid_{t=t_1} < 0 < GR \mid_{t=t_2}, \tag{4.18}$$

On the other hand, since GR is non-negative and v is non-decreasing over time, and $v_1 \leq v_2$ holds, which bring us

$$\begin{aligned} GR\mid_{t=t_1} &= k_g\left(\frac{1}{1+u_1/J}\right)\left(1-\frac{v_1}{v_m}\right) \\ &= k_g\left(\frac{1}{1+u_2/J}\right)\left(1-\frac{v_1}{v_m}\right) \\ &\geq k_g\left(\frac{1}{1+u_2/J}\right)\left(1-\frac{v_2}{v_m}\right) \\ &= GR\mid_{t=t_2} \end{aligned} \tag{4.19}$$

which conflicts with (4.18). □

This proposition is corresponding to the previous model (Zhang *et al.*, 2020), where the regulation of the production rate by the growth rate per capita is not considered.

Proposition 2. *There is no peak of AraC when $F(GR)$ is amonotonous decreasing function.*

Proof. Similarly, we assume that there is at least one local maximum point $u_p = u(t_p)$ with $t_p \in (0, T)$, $\frac{du}{dt}\mid_{t=t_p} = 0$ and we prove by contradiction. Thus, there exist t_1, t_2 close to t_p such that $t_1 < t_p < t_2$ and $u_1 = u_2 < u_p$. And also we have:

$$\frac{du}{dt}\mid_{t=t_1} > 0 > \frac{du}{dt}\mid_{t=t_2}. \tag{4.20}$$

However, since $v_1 \preceq v_2$ we have

$$GR\mid_{t=t_1} \geq GR\mid_{t=t_2}, \tag{4.21}$$

so

$$F(GR)\mid_{t=t_1} \geq F(GR)\mid_{t=t_2}, \tag{4.22}$$

which provided us

$$\frac{du}{dt}\mid_{t=t_1} \leq \frac{du}{dt}\mid_{t=t_2}, \tag{4.23}$$

and this is conflicting with (4.23). □

Proposition 3. *If $F(GR) = aGR + 1$ where a is a positive constant (i.e. the grouth feedback function is a linear increasing function of the cell growth rate per capita). Then the following two situations will not occur at the same time:*

- *There is a local minimum u_{min} with $t_{min} \in (0, T)$ and $u_{min} < u_0$.*
- *There is a local maximum u_{max} with $t_{max} \in (0, T)$ and $u_{max} > u_0$.*

here $u_0 = u(0)$ is the initial concentration.

Proof. We prove this proposition by contradiction. Assuming the above two situations both hold, which means there is at least one local minimum point and one local maximum point. We then assume that among all these extrema, the first one is a local minimum which is denoted as $u_v = u(t_v)$ with $t_v \in (0, T)$ and $u_v < u_0$ (Noting that the proof will be similar if the first one is a local maximum point). And we denote the local maximum (which following u_v) as u_p with $t_p \in (0, T)$ and $u_p > u_0$. Thus, we have $0 < t_v < t_p < T$ and $u_v < u_0 < u_p$.

So there exists a time t_* where $t_* \in (t_v, t_p)$ such that $u(t_*) := u_* = u_0$ and

$$\begin{aligned}\frac{du}{dt}\ |_{t=t_*} &= f(u_*)(a \cdot GR\ |_{t=t_*} + 1) - (d + GR\ |_{t=t_*})u_* \\ &= (af(u_*) - u_*) \cdot GR\ |_{t=t_*} + (f(u_*) - du_*) \\ &= (af(u_0) - u_0) \cdot GR\ |_{t=t_*} + (f(u_0) - du_0) \\ &> 0.\end{aligned} \tag{4.24}$$

Since u_0 is a steady state of the system without growth-mediated feedback (i.e. $GR = 0$), we have $f(u_0) - du_0 = 0$, which leads to:

$$(af(u_0) - u_0) \cdot GR\ |_{t=t_*} > 0. \tag{4.25}$$

Meanwhile, since the local minmum point u_v is also the first local extremum, we have

$$\begin{aligned} \frac{du}{dt} \mid_{t=0} &= f(u_0)(a \cdot GR \mid_{t=t_*} +1) - (d + GR \mid_{t=0})u_0 \\ &= (af(u_0) - u_0) \cdot GR \mid_{t=t_0} +(f(u_0) - du_0) \\ &< 0, \end{aligned} \tag{4.26}$$

and this is conflicting with (4.25). □

Based on the above discussion we know that, the following forms of growth feedback function $F(GR)$ can not capture the experimental data:

- $F(GR)$ is a constant function.
- $F(GR)$ is a linear function.
- $F(GR)$ is a monotonic decreasing function.

4.6.2 Equilibrium Points and Local Stability (Model II)

Before the analysis, we shall make the following assumptions:

- Since the base expression rate k_0 is very low, we assume that $k_0 = 0$.
- The nutrient level refers to normalized concentration and $P(0) = P_0 > q_2$.
- We set $S_a = 1$, since it was fixed in previous simulations.

And model (4.5) can be rewrite as following:

$$\frac{dA}{dt} = k_1 \frac{A^2}{A^2+1}(w_1(P)F_1(GR) + w_2(P)F_2(GR)) - (d + GR)A \tag{4.27a}$$

$$\frac{dN}{dt} = k_g \left(\frac{1}{A/J+1}\right)(w_1(P)g_1(P) + w_2(P)g_2(P))\, N \tag{4.27b}$$

$$\frac{dP}{dt} = -\gamma(P - q_2)N. \tag{4.27c}$$

Theorem 4.6.1. *If* $A(0) \geq 0$, $N(0) \geq 0$, *and* $P(0) > q_2$, *then* A, N, *and* P *will remain nonnegative for* $t > 0$.

Proof. Firstly, it is easy to know that P with $P(0) > q_2$ is bounded below by q_2 and bounded above by $P(0)$. Then since $g_i(P)$ and $w_i(P)$ $i = (1, 2)$ are all nonnegative with $q_2 < P < P(0)$, we have $dN/dt \geq 0$ and $N(t) \geq 0$ for all $t > 0$. Thirdly, since $A(0) \geq 0$ and $dA/dt \mid_{A=0} = 0$, we have $A(t) \geq 0$ for all $t > 0$. □

Theorem 4.6.2. *The system is bounded if* $A(0) \geq 0$, $N(0) \geq 0$, *and* $P(0) > q_2$.

Proof. From Theorem 4.6.2 we know that P is bounded. Because P is strictly decreasing and will finally reach and stay at the minimum q_2, so there exist time $t_q > 0$ such as $P(t) > q_2$ for $0 \leq t < t_q$ and $P(t) = q_2$ for all $t \geq t_q$. Also from (4.27b) we know that

$$w_1(q_2)g_1(q_2) + w_2(q_2)g_2(q_2) = 0, \tag{4.28}$$

which leads to $dN/dt = 0$ for all $t \geq t_q$. Since $N(t)$ is continous, $N(t)$ is bounded above by $N(t_q)$.

Meanwhile, $GR = 0$ for all $t \geq t_q$ will lead to

$$w_1(P)F_1(0) + w_2(P)F_2(0) = 1, \tag{4.29}$$

so for $t \geq t_q$, (4.27a) is equivalent to

$$\frac{dA}{dt} = k_1 \frac{A^2}{A^2 + 1} - dA, \tag{4.30}$$

so that $A(t)$ is bounded for $t \geq t_q$. Also since $A(t)$ is continous and $A(0) \geq 0$, it is bounded for $0 \leq t < t_q$, thus $A(t)$ is bounded for $t \geq 0$. □

Noting that $N(t)$ represents the host cell density at time t, and $N(0) = 0$ means there is no cell in the liquid culture, thus $N(t) = 0$ will hold for all $t > 0$. In this case there will be no circuit expression ($A(t) = 0$ for all $t \geq 0$) and $P(t) = P(0)$ for all $t \geq 0$. To keep the system biologically realistic, we only consider the steady state with $N^* > 0$ (N-positive equilibira).

Theorem 4.6.3. *In this theorem, we only discuss the equilibrium* (A^*, N^*, P^*) *with* $N^* > 0$*, while* $P^* = q_2$ *is the only equilibrium for* P*, and the value of* N^* *is also the maximum of* N *which means* $N^* = N(t)$ *for* $t \geq t_q$ *(*t_q *is defined in Theorem 4.6.2). And for* A^* *we have following conclusions:*

- *When* $d > k_1/2$*, there is a unique* N*-positive equilibrium* $A_0^* = 0$*.*
- *When* $d = k_1/2$*, there are two* N*-positive equilibria* $A_0^* = 0$*,* $A_1^* = 1$*.*
- *When* $d < k_1/2$*, there are three* N*-positive equilibria* $A_0^* = 0$*,* $A_{1,2}^* = \dfrac{1}{2d}(k_1 \pm \sqrt{k_1^2 - 4d^2})$*.*

Proof. The steady state (A^*, N^*, P^*) can be solved from the following system:

$$0 = k_1 \frac{A^2}{A^2+1}(w_1(P)F_1(GR) + w_2(P)F_2(GR)) - (d + GR)A \tag{4.31a}$$

$$0 = k_g \left(\frac{1}{A/J+1}\right)(w_1(P)g_1(P) + w_2(P)g_2(P))\, N \tag{4.31b}$$

$$0 = -\gamma(P - q_2)N. \tag{4.31c}$$

With the condition that $N^* > 0$, we have $P^* = q_2$, and $GR = 0$. So (4.31a) is equivalent to:

$$\begin{aligned} 0 &= k_1 \frac{A^2}{A^2+1} - dA \\ &= \frac{A}{A^2+1}(-dA^2 + k_1 A - d), \end{aligned} \tag{4.32}$$

noting that k_1 and d are positive, so if $d > k_1/2$, $-dA^2 + k_1 A - d < 0$ always holds and $A_0^* = 0$ will be the only solution of (4.32). And if $d = k_1/2$, $-dA^2 + k_1 A - d = 0$

will have a unique positive solution $A_1^* = 1$ and $A_0^* = 0$, $A_1^* = 1$ are the solutions of (4.32). If $d < k_1/2$, $-dA^2 + k_1 A - d = 0$ will have two positive solutions $A_{1,2}^* = \frac{1}{2d}(k_1 \pm \sqrt{k_1^2 - 4d^2})$ and there will be three positive equilibria for A: $A_0^* = 0$, and $A_{1,2}^* = \frac{1}{2d}(k_1 \pm \sqrt{k_1^2 - 4d^2})$. □

We then derive the Jacobian matrix of (4.27). Since $F(GR) = w_1(P)F_1(GR) + w_2(P)F_2(GR)$ and $GR = (1/N)(dN/dt)$, we have $GR = GR(A, P)$ and $F(GR) = F(A, P)$. Knowing that $P^* = q_2$ holds for all N-positive equilibria, so $GR(A, P^*) = 0$ and $F(A, P^*) = F(0) = 1$. Futhermore,

$$\begin{aligned} \left.\frac{dGR(A,P)}{dA}\right|_{P=q_2} &= k_g(w_1(P)g_1(P) + w_2(P)g_2(P))\frac{d}{dA}\left(\frac{1}{A/J+1}\right)\Big|_{P=q_2} \\ &= 0 \end{aligned} \tag{4.33}$$

$$\begin{aligned} \left.\frac{dF(A,P)}{dA}\right|_{P=q_2} &= w_1(P)\frac{dF_1(GR)}{dGR}\left.\frac{dGR}{dA}\right|_{P=q_2} + w_2(P)\frac{dF_2(GR)}{dGR}\left.\frac{dGR}{dA}\right|_{P=q_2} \\ &= 0, \end{aligned} \tag{4.34}$$

and the Jacobian matrix of N-positive equilibria can be written as following:

$$J(A^*, N^*, P^*) = \begin{pmatrix} k_1\frac{2A^*}{(A^{*2}+1)^2} - d & 0 & \frac{df_1}{dP}\,|_{(A^*,N^*,P^*)} \\ 0 & 0 & \frac{df_2}{dP}\,|_{(A^*,N^*,P^*)} \\ 0 & 0 & -\gamma N^* \end{pmatrix} \tag{4.35}$$

so the eigenvalues are: $\lambda_1 = k_1\frac{2A^*}{(A^{*2}+1)^2} - d$, $\lambda_2 = 0$, and $\lambda_3 = -\gamma N^*$.

Theorem 4.6.4. *With the initial conditions $A(0) \geq 0$, $N(0) \geq 0$, and $P(0) > q_2$, the stability of N-positive equilibria can be conluded as following:*

- *When $d > k_1/2$, $A_0^* = 0$ is a globally stable N-positive equilibrium.*

- *When $d = k_1/2$, there are two positive N-positive equilibria $A_0^* = 0$, $A_1^* = 1$ and A_0^* is locally asymptotically stable.*

- *When $d < k_1/2$, there are three positive N-positive equilibria $A_0^* = 0$, $A_{1,2}^* = \dfrac{1}{2d}(k_1 \pm \sqrt{k_1^2 - 4d^2})$ and $A_{0,2}^*$ are locally asymptotically stable.*

Proof. Recall t_q which was defined in Theorem 4.6.2, we know that $N(t) = N^*$ and $P(t) = P^* = q_2$ for all $t \geq t_q$, so the stability of the N-positive equilibria can be determined by the stability of A.

When $d > k_1/2$, $A_0^* = 0$ is the only N-positive equilibrium and the corresponding eigenvalue is $k_1 \frac{2A_0^*}{(A_0^{*2}+1)^2} - d = -d < 0$. So $(A_0^*, N^*, P^*) = (0, N^*, q_2)$ is a globally stable N-positive equilibrium.

When $d = k_1/2$, there are two positive N-positive equilibria $A_0^* = 0$, $A_1^* = 1$ and the corresponding eigenvalues are $-d$ and 0, respectively. So $(A_0^*, N^*, P^*) = (0, N^*, q_2)$ is locally asymptotically stable.

When $d < k_1/2$, there are three positive N-positive equilibria $A_0^* = 0$, $A_{1,2}^* = \dfrac{1}{2d}(k_1 \pm \sqrt{k_1^2 - 4d^2})$, and the signs of corresponding eigenvalues are negative, positive, and negative, respectively. So $(A_0^*, N^*, P^*) = (0, N^*, q_2)$ and $(A_2^*, N^*, P^*) = \left(\dfrac{1}{2d}(k_1 - \sqrt{k_1^2 - 4d^2}), N^*, q_2\right)$ are locally asymptotically stable. □

4.7 Dicsuccion and Conclusion

Growth-mediated feedback between the host cell and gene circuits is almost ineliminable during the experiments and will potentially affect the system's behaviors. Meanwhile, the multi-nutrient environment can also interfere the host cell growth. Although recent research works yielded insights into the circuit-host interactions and some mathematical models were developed, there is still a lack of a model that described the growth-mediated feedback in the multi-nutrient environment.

Here we introduced a modified model based on the work of Melendez et al. Melendez-Alvarez *et al.* (2021). The models we propose here, are not only examples of modeling growth-mediated feedback with nutrient factor, but also allow for a broader range of circuit-host interactions. This framework can potentially raise new questions that are rarely discussed via mathematical models.

Chapter 5

CONCLUSIONS AND FURTHER DIRECTIONS

This thesis attempts to deepen our understanding of circuit-driven bacterial pattern formation and the connection between *E. coli* growth and synthetic gene circuit under different conditions using a mathematical modeling approach. In this chapter, I will summarize the conclusions of each of the projects included in previous chapters. Additionally, I will discuss how this research work has shaped and will shape my future research, thoughts, and ideas.

In Chapter 2, we focused on *E. coli* colony growth and studied the colony growth process under different experimental conditions. As one of the most common biological phenomena, colony growth can be affected by multiple control factors and presents a nonlinear growth profile that has been rarely discussed. Since the mechanisms of these control factors are unclear and current models may not be able to capture the growth profile, we developed a modeling framework by introducing a one-dimensional reaction-diffusion model, to describe the colony growth under multiple control factors which will affect both cell movement and reproduction. Our simulations showed that the model is capable of capturing this nonlinear growth profile under different growing conditions with much lower fitting errors when compared with Fisher's equation. Furthermore, the model provided comparable results on simulating the colony's cross-sectional profile.

Despite the insights that this model has afforded us, there is still much to be done. Future work may include traveling wave analysis of the model. This will allow us to determine the minimal wave speed and how the wave speed will change if parameters vary. The density-dependent diffusion function may generate singularities which means that approximation of the diffusion function may be needed.

Although the model did not directly measure any specific control factors such as temperature or nutrient concentration, it provided an important framework to start with. Based on this model, we can introduce other variables such as nutrient level, or experimentally verified functions to measure any specific control factors. Such a modeling approach is a significant way to study underexplored topics with minimal hypotheses.

In Chapter 3, we performed a multi-perspective analysis of circuit-driven pattern formation. Bacterial stripe patterning is commonly used to study somitogenesis at a molecular level, and the latter can also be explained by the clock and wavefront mechanism. Here we first introduced a novel synthetic gene circuit (MINPAC), which can direct spatiotemporal pattern formation within the bacterial colony. A reaction-diffusion-based model was developed to describe the circuit topology and external diffusion process (full-system model). The model demonstrated oscillatory expression of the autoinducers that mimic pattern formation progression. With the change of corresponding parameters and initial conditions, the model was capable of simulating ring patterns under different conditions, which agreed well with the experimental data.

To investigate the relationship between gene network topology and the resulting multicellular pattern, several perturbed MINPAC topologies were introduced (control systems). And to match these control systems, the control models were modified from the full-system model. Without loss of parameter consistency, all the control models provided comparable simulation results with the experimental observations, which showed that none of the control systems could generate obvious ring patterns. These results indicated that the multiple-ring pattern is unique to MINPAC circuit.

We then presented a quantitative analysis to explain several important questions based on the experimental results. Firstly, the MINPAC topology can be regarded as a composition of two topologically equivalent sub-motifs, and each of them forms a robust positive-plus-negative oscillator topology. We wanted to determine how this coupling affects the robustness of oscillations. We then performed model-comparison simulations showing that the two-motif MINPAC harbors greater robustness and amplitude against parameter perturbations to generate temporal oscillation which enhances the likelihood of observing our desired phenotypic outputs from the synthetic gene circuit. Secondly, the cause of why does the total number of the ring varies even with the same experimental settings was also discussed. We hypothesized that this phenomenon is due to random variations of the initial concentrations of intracellular proteins and autoinducers. This hypothesis was supported by the phase portrait of the system which pointed out that such randomness could change the location of the initial point then change the Poincare return time.

Besides the above study, I also analyzed the reduced reaction-diffusion system and derived the theoretical minimal wave speed. The reaction part of the model was used to analyze the local stability at equilibrium points. The theoretical results also supported the experimental observations.

An ongoing aspect of this research of this project, which will improve the model, is combining our pattern formation model with the colony growth model we introduced in Chapter 2. This will allow us to simulate the pattern formation process with the moving boundary which mimics the colony growth. The new model will be able to capture the color differences between each ring rather than providing the oscillation with the same amplitude. Further investigation is required to analyze model behavior.

This project forced me to study a biological question from different aspects, which also shows the power of a modeling approach. Systematically describing the biological system and providing comparable simulation results can be regarded as the first step of such an approach. Using mathematical models to explain the phenomena observed during experiments and verify the hypotheses are also very important. Because of this project, I have now reached the point where I will raise questions from experimental observations and formulate corresponding analyses based on the model.

In Chapter 4, which differs from previous research work, we used mathematical models to describe the impact of host cell's growth on circuit expression, which is unneglectable and can significantly change circuit function. This work can be regarded as an extension of previous work (Zhang *et al.*, 2020), in which a simple self-activation circuit was transformed into *E. coli* cell and grew in liquid culture. An ODE model was first introduced by Zhang *et al.* (2020) with negative feedback which drives the protein degradation rate. However, more recent experiments showed an unexpected damped oscillatory behavior of a self-activation gene circuit induced by nutrient-modulating growth feedback in a multi-nutrient environment, which cannot be described using previous models. Here we modified a previous ODE model by introducing a growth feedback function that governs circuit expression. We computationally tested different types of functions and concluded that the non-monotonic function provided the best fits for the experimental data. I further analyzed the model, and the necessary conditions for the feedback function to fit the experimental data were proved.

Furthermore, changing the proportion of nutrient components will also affect a cell's growth profile, which can be observed as "two-stage" growth. Then a modified ODE model was also presented; here we used the Droop model with regulatory functions to describe the "trade-off" process of *E. coli.* The new model not only captured the oscillatory behavior of the gene circuit but also showed the "two-stage" growth profile of the host cell under different nutrient conditions. Further analysis of the model was also presented including local stability of the equilibrium points, which supported the existence of the oscillatory behavior.

Of course, the second model is still only an approximation of the “trade-off” process with several hypotheses that need to be verified experimentally. Unlike the cell growth under nutrient limitations, where a minimal function is usually used to model the nutrient-dependent growth rate, here we assume the existence of the nutrient preference. Although this is highly possible based on the experimental data, we are still unclear about what is the key nutrient component that causes such a phenomenon, or how this “trade-off” process will affect circuit expression. Thus, with further experiments, our model can be improved by specifying the regulatory functions and the consumption of two different nutrients.

In all, because of the advanced experimental design and engineering technique and the complexity of cell dynamics coupled with gene circuit expression, more and more questions shall be raised while mathematics will continue to play a large role in biological understanding and practice. Mathematical models can not only be used to describe and explain the biological mechanisms but more importantly, they can be used to predict experimental outcomes and help to improve the experimental design. Although there is still a gap between mathematical models to the actual experimental design, the work performed in this thesis aims to assist the future mathematicians who strive to achieve this goal.

REFERENCES

Akaike, H., “Information theory and an extension of the maximum likelihood principle”, in “Selected papers of hirotugu akaike”, pp. 199–213 (Springer, 1998).

Antunes, M. S., S.-B. Ha, N. Tewari-Singh, K. J. Morey, A. M. Trofka, P. Kugrens, M. Deyholos and J. I. Medford, “A synthetic de-greening gene circuit provides a reporting system that is remotely detectable and has a re-set capacity”, Plant biotechnology journal **4**, 6, 605–622 (2006).

Asllani, M., T. Carletti, D. Fanelli and P. K. Maini, “A universal route to pattern formation in multicellular systems”, The European Physical Journal B **93**, 7, 1–11 (2020).

Ausländer, S. and M. Fussenegger, “Engineering gene circuits for mammalian cell–based applications”, Cold Spring Harbor perspectives in biology **8**, 7, a023895 (2016).

Baker, R. E., S. Schnell and P. K. Maini, “A clock and wavefront mechanism for somite formation”, Developmental biology **293**, 1, 116–126 (2006).

Balagaddé, F. K., H. Song, J. Ozaki, C. H. Collins, M. Barnet, F. H. Arnold, S. R. Quake and L. You, “A synthetic escherichia coli predator–prey ecosystem”, Molecular systems biology **4**, 1, 187 (2008).

Basu, S., Y. Gerchman, C. H. Collins, F. H. Arnold and R. Weiss, “A synthetic multicellular system for programmed pattern formation”, Nature **434**, 7037, 1130–1134 (2005).

Becskei, A., B. Séraphin and L. Serrano, “Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion”, The EMBO journal **20**, 10, 2528–2535 (2001).

Blanchard, A. E., C. Liao and T. Lu, “Circuit-host coupling induces multifaceted behavioral modulations of a gene switch”, Biophysical journal **114**, 3, 737–746 (2018).

Blight, K. J., A. A. Kolykhalov and C. M. Rice, “Efficient initiation of hcv rna replication in cell culture”, Science **290**, 5498, 1972–1974 (2000).

Block, S. M., J. E. Segall and H. C. Berg, “Adaptation kinetics in bacterial chemotaxis.”, Journal of bacteriology **154**, 1, 312–323 (1983).

Bradley, R. W., M. Buck and B. Wang, “Tools and principles for microbial gene circuit engineering”, Journal of Molecular Biology **428**, 5, 862–888 (2016).

Brenner, M. P., L. S. Levitov and E. O. Budrene, “Physical mechanisms for chemotactic pattern formation by bacteria”, Biophysical journal **74**, 4, 1677–1693 (1998).

Cameron, D. E., C. J. Bashor and J. J. Collins, “A brief history of synthetic biology”, Nature Reviews Microbiology **12**, 5, 381–390 (2014).

Canosa, J., “On a nonlinear diffusion equation describing population growth”, IBM Journal of Research and Development **17**, 4, 307–313 (1973).

Cao, Y., M. D. Ryser, S. Payne, B. Li, C. V. Rao and L. You, “Collective space-sensing coordinates pattern scaling in engineered bacteria”, Cell **165**, 3, 620–630 (2016).

Capasso, V., M. Gromov, A. Harel-Bellan, N. Morozova and L. L. Pritchard, *Pattern Formation in Morphogenesis: problems and mathematical issues*, vol. 15 (Springer Science & Business Media, 2012).

Chandran, D., W. Copeland, S. Sleight and H. Sauro, “Mathematical modeling and synthetic biology”, Drug Discovery Today: Disease Models **5**, 4, 299–309 (2008).

Chen, L., J. Noorbakhsh, R. M. Adams, J. Samaniego-Evans, G. Agollah, D. Nevozhay, J. Kuzdzal-Fick, P. Mehta and G. Balázsi, “Two-dimensionality of yeast colony expansion accompanied by pattern formation”, PLoS Comput Biol **10**, 12, e1003979 (2014).

Chen, Y., J. K. Kim, A. J. Hirning, K. Josić and M. R. Bennett, “Emergent genetic oscillations in a synthetic microbial consortium”, Science **349**, 6251, 986–989 (2015).

Cole, J. A., L. Kohler, J. Hedhli and Z. Luthey-Schulten, “Spatially-resolved metabolic cooperativity within dense bacterial colonies”, BMC systems biology **9**, 1, 1–17 (2015).

Cooke, J. and E. C. Zeeman, “A clock and wavefront model for control of the number of repeated structures during animal morphogenesis”, Journal of theoretical biology **58**, 2, 455–476 (1976).

Cotterell, J., A. Robert-Moreno and J. Sharpe, “A local, self-organizing reaction-diffusion model can explain somite patterning in embryos”, Cell systems **1**, 4, 257–269 (2015).

Danino, T., O. Mondragón-Palomino, L. Tsimring and J. Hasty, “A synchronized quorum of genetic clocks”, Nature **463**, 7279, 326–330 (2010).

Del Vecchio, D., H. Abdallah, Y. Qian and J. J. Collins, “A blueprint for a synthetic genetic feedback controller to reprogram cell fate”, Cell systems **4**, 1, 109–120 (2017).

Diaz-Cuadros, M., D. E. Wagner, C. Budjan, A. Hubaud, O. A. Tarazona, S. Donelly, A. Michaut, Z. Al Tanoury, K. Yoshioka-Kobayashi, Y. Niino *et al.*, “In vitro characterization of the human segmentation clock”, Nature **580**, 7801, 113–118 (2020).

Erickson, D. W., S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland and T. Hwa, "A global resource allocation strategy governs growth transition kinetics of escherichia coli", Nature **551**, 7678, 119–123 (2017).

Farrell, F., O. Hallatschek, D. Marenduzzo and B. Waclaw, "Mechanically driven growth of quasi-two-dimensional microbial colonies", Physical review letters **111**, 16, 168101 (2013).

Fu, X., L.-H. Tang, C. Liu, J.-D. Huang, T. Hwa and P. Lenz, "Stripe formation in bacterial systems with density-suppressed motility", Physical review letters **108**, 19, 198102 (2012).

Gibson, D. G., J. I. Glass, C. Lartigue, V. N. Noskov, R.-Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie *et al.*, "Creation of a bacterial cell controlled by a chemically synthesized genome", science **329**, 5987, 52–56 (2010).

Gierer, A. and H. Meinhardt, "A theory of biological pattern formation", Kybernetik **12**, 1, 30–39 (1972).

Gomez, C., E. M. Özbudak, J. Wunderlich, D. Baumann, J. Lewis and O. Pourquié, "Control of segment number in vertebrate embryos", Nature **454**, 7202, 335–339 (2008).

Gomez, M. M. and M. Arcak, "A tug-of-war mechanism for pattern formation in a genetic network", ACS synthetic biology **6**, 11, 2056–2066 (2017).

Grant, M. A., B. Wacław, R. J. Allen and P. Cicuta, "The role of mechanical forces in the planar-to-bulk transition in growing escherichia coli microcolonies", Journal of The Royal Society Interface **11**, 97, 20140400 (2014).

Grover, J. P., "The impact of variable stoichiometry on predator-prey interactions: a multinutrient approach", The American Naturalist **162**, 1, 29–43 (2003).

Guantes, R. and J. F. Poyatos, "Multistable decision switches for flexible control of epigenetic differentiation", PLoS Comput Biol **4**, 11, e1000235 (2008).

Haddon, C., L. Smithers, S. Schneider-Maunoury, T. Coche, D. Henrique and J. Lewis, "Multiple delta genes and lateral inhibition in zebrafish primary neurogenesis", Development **125**, 3, 359–370 (1998).

Hasty, J., D. McMillen and J. J. Collins, "Engineered gene circuits", Nature **420**, 6912, 224–230 (2002).

He, C., S. Bayakhmetov, D. Harris, Y. Kuang and X. Wang, "A predictive reaction-diffusion based model of e. coli colony growth control", IEEE Control Systems Letters (2020).

Hoffman, L. R., D. A. D'Argenio, M. J. MacCoss, Z. Zhang, R. A. Jones and S. I. Miller, "Aminoglycoside antibiotics induce bacterial biofilm formation", Nature **436**, 7054, 1171–1175 (2005).

Jayathilake, P. G., P. Gupta, B. Li, C. Madsen, O. Oyebamiji, R. González-Cabaleiro, S. Rushton, B. Bridgens, D. Swailes, B. Allen *et al.*, "A mechanistic individual-based model of microbial communities", PloS one **12**, 8, e0181965 (2017).

Jones, B. S., L. S. Lamb, F. Goldman and A. Di Stasi, "Improving the safety of cell therapy products by suicide gene transfer", Frontiers in pharmacology **5**, 254 (2014).

Kawasaki, K., A. Mochizuki, M. Matsushita, T. Umeda and N. Shigesada, "Modeling spatio-temporal patterns generated bybacillus subtilis", Journal of theoretical biology **188**, 2, 177–185 (1997).

Kicheva, A., M. Cohen and J. Briscoe, "Developmental pattern formation: insights from physics and biology", Science **338**, 6104, 210–212 (2012).

Klausmeier, C., E. Litchman and S. A. Levin, "A model of flexible uptake of two essential resources", Journal of theoretical biology **246**, 2, 278–289 (2007).

Klumpp, S., Z. Zhang and T. Hwa, "Growth rate-dependent global effects on gene expression in bacteria", Cell **139**, 7, 1366–1375 (2009).

Kondo, S. and T. Miura, "Reaction-diffusion model as a framework for understanding biological pattern formation", science **329**, 5999, 1616–1620 (2010).

Lagarias, J. C., J. A. Reeds, M. H. Wright and P. E. Wright, "Convergence properties of the nelder–mead simplex method in low dimensions", SIAM Journal on optimization **9**, 1, 112–147 (1998).

Lapidus, I. R. and R. Schiller, "Model for the chemotactic response of a bacterial population", Biophysical journal **16**, 7, 779–789 (1976).

Lee, T. J., C. Tan, D. Tu and L. You, "Modeling cellular networks", Bioinformatics: An Engineering Case-Based Approach pp. 151–172 (2007).

Lengyel, I. and I. R. Epstein, "A chemical approach to designing turing patterns in reaction-diffusion systems.", Proceedings of the National Academy of Sciences **89**, 9, 3977–3979 (1992).

Leppänen, T., M. Karttunen, R. Barrio and K. Kaski, "Morphological transitions and bistability in turing systems", Physical Review E **70**, 6, 066202 (2004).

Lewis, J., "Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator", Current Biology **13**, 16, 1398–1408 (2003).

Leyva, J. F., C. Málaga and R. G. Plaza, "The effects of nutrient chemotaxis on bacterial aggregation patterns with non-linear degenerate cross diffusion", Physica A: Statistical Mechanics and its Applications **392**, 22, 5644–5662 (2013).

Li, X., F. Gonzalez, N. Esteves, B. E. Scharf and J. Chen, "Formation of phage lysis patterns and implications on co-propagation of phages and motile host bacteria", PLoS computational biology **16**, 3, e1007236 (2020).

Liu, C., X. Fu, L. Liu, X. Ren, C. K. Chau, S. Li, L. Xiang, H. Zeng, G. Chen, L.-H. Tang *et al.*, "Sequential establishment of stripe patterns in an expanding cell population", Science **334**, 6053, 238–241 (2011).

Liu, W., J. Cremer, D. Li, T. Hwa and C. Liu, "An evolutionarily stable strategy to colonize spatially extended habitats", Nature **575**, 7784, 664–668 (2019).

Martin, V. J., D. J. Pitera, S. T. Withers, J. D. Newman and J. D. Keasling, "Engineering a mevalonate pathway in escherichia coli for production of terpenoids", Nature biotechnology **21**, 7, 796–802 (2003).

Matz, C. and K. Jürgens, "Interaction of nutrient limitation and protozoan grazing determines the phenotypic structure of a bacterial community", Microbial Ecology **45**, 4, 384–398 (2003).

Melendez-Alvarez, J., C. He, R. Zhang, Y. Kuang and X.-J. Tian, "Emergent oscillation induced by nutrient-modulating growth feedback", bioRxiv (2021).

Menn, D. and X. Wang, "Modeling gene networks to understand multistability in stem cells", in "Computational Stem Cell Biology", pp. 173–189 (Springer, 2019).

Miao, H., C. Dykes, L. M. Demeter and H. Wu, "Differential equation modeling of hiv viral fitness experiments: model identification, model selection, and multimodel inference", Biometrics **65**, 1, 292–300 (2009).

Miao, H., X. Jin, A. S. Perelson and H. Wu, "Evaluation of multitype mathematical models for cfse-labeling experiment data", Bulletin of mathematical biology **74**, 2, 300–326 (2012).

Mimura, M., H. Sakaguchi and M. Matsushita, "Reaction–diffusion modelling of bacterial colony patterns", Physica A: Statistical Mechanics and its Applications **282**, 1-2, 283–303 (2000).

Mimura, M. and T. Tsujikawa, "Aggregating pattern dynamics in a chemotaxis model including growth", Physica A: Statistical Mechanics and its Applications **230**, 3-4, 499–543 (1996).

Monod, J. and F. Jacob, "General conclusions: teleonomic mechanisms in cellular metabolism, growth, and differentiation", in "Cold Spring Harbor symposia on quantitative biology", vol. 26, pp. 389–401 (Cold Spring Harbor Laboratory Press, 1961).

Naganuma, T., Y. Iinuma, H. Nishiwaki, R. Murase, K. Masaki and R. Nakai, "Enhanced bacterial growth and gene expression of d-amino acid dehydrogenase with d-glutamate as the sole carbon source", Frontiers in microbiology **9**, 2097 (2018).

Naiche, L., N. Holder and M. Lewandoski, "Fgf4 and fgf8 comprise the wavefront activity that controls somitogenesis", Proceedings of the National Academy of Sciences **108**, 10, 4018–4023 (2011).

Ohgiwari, M., M. Matsushita and T. Matsuyama, “Morphological changes in growth phenomena of bacterial colony patterns”, Journal of the Physical Society of Japan **61**, 3, 816–822 (1992).

Pipe, L. Z. and M. J. Grimson, “Spatial-temporal modelling of bacterial colony growth on solid media”, Molecular BioSystems **4**, 3, 192–198 (2008).

Pollack, A., “How do you like your genes? biofabs take orders”, New York Times **12**, 2007 (2007).

Potvin-Trottier, L., N. D. Lord, G. Vinnicombe and J. Paulsson, “Synchronous long-term oscillations in a synthetic gene circuit”, Nature **538**, 7626, 514–517 (2016).

Prochazka, L., Y. Benenson and P. W. Zandstra, “Synthetic gene circuits and cellular decision-making in human pluripotent stem cells”, Current Opinion in Systems Biology **5**, 93–103 (2017).

Ratkowsky, D. A., J. Olley, T. McMeekin and A. Ball, “Relationship between temperature and growth rate of bacterial cultures.”, Journal of bacteriology **149**, 1, 1–5 (1982).

Ruan, S., “Diffusion-driven instability in the gierer-meinhardt model of morphogenesis”, Natural Resource Modeling **11**, 2, 131–141 (1998).

Salazar-Ciudad, I., J. Garcia-Fernández and R. V. SOLÉ, “Gene networks capable of pattern formation: from induction to reaction–diffusion”, Journal of theoretical biology **205**, 4, 587–603 (2000).

Saragosti, J., P. Silberzan and A. Buguin, “Modeling e. coli tumbles by rotational diffusion. implications for chemotaxis”, PloS one **7**, 4, e35412 (2012).

Satnoianu, R., P. K. Maini, F. Sánchez-Garduno and J. Armitage, “Travelling waves in a nonlinear degenerate diffusion model for bacterial pattern formation”, (2001).

Scholes, N. S. and M. Isalan, “A three-step framework for programming pattern formation”, Current opinion in chemical biology **40**, 1–7 (2017).

Schwarz, G. *et al.*, “Estimating the dimension of a model”, Annals of statistics **6**, 2, 461–464 (1978).

Shapiro, J. A., “The significances of bacterial colony patterns”, Bioessays **17**, 7, 597–607 (1995).

Shiloach, J. and R. Fass, “Growing e. coli to high cell densitya historical perspective on method development”, Biotechnology advances **23**, 5, 345–357 (2005).

Song, Y., R. Yang and G. Sun, “Pattern dynamics in a gierer–meinhardt model with a saturating term”, Applied Mathematical Modelling **46**, 476–491 (2017).

Sourjik, V. and N. S. Wingreen, “Responding to chemical gradients: bacterial chemotaxis”, Current opinion in cell biology **24**, 2, 262–268 (2012).

Stepien, T. L., E. M. Rutter and Y. Kuang, “A data-motivated density-dependent diffusion model of in vitro glioblastoma growth”, Mathematical Biosciences & Engineering **12**, 6, 1157 (2015).

Stepien, T. L., E. M. Rutter and Y. Kuang, “Traveling waves of a go-or-grow model of glioma growth”, SIAM Journal on Applied Mathematics **78**, 3, 1778–1801 (2018).

Su, P.-T., C.-T. Liao, J.-R. Roan, S.-H. Wang, A. Chiou and W.-J. Syu, “Bacterial colony from two-dimensional division to three-dimensional development”, PloS one **7**, 11, e48098 (2012).

Tian, X.-J., M. V. Ferro and H. Goetz, “Modeling ncrna-mediated circuits in cell fate decision”, Computational Biology of Non-Coding RNA pp. 411–426 (2019).

Toennies, G., “Differential responses to amino acids in bacterial growth”, (ACS Publications, 1964).

Tsai, T. Y.-C., Y. S. Choi, W. Ma, J. R. Pomerening, C. Tang and J. E. Ferrell, “Robust, tunable biological oscillations from interlinked positive and negative feedback loops”, Science **321**, 5885, 126–129 (2008).

Turing, A. M., “The chemical basis of morphogenesis”, Bulletin of mathematical biology **52**, 1, 153–197 (1990).

Tuszyński, J., M. Otwinowski and J. Dixon, “Spiral-pattern formation and multistability in landau-ginzburg systems”, Physical Review B **44**, 17, 9201 (1991).

Vermot, J. and O. Pourquié, “Retinoic acid coordinates somitogenesis and left–right patterning in vertebrate embryos”, Nature **435**, 7039, 215–220 (2005).

Wakita, J.-i., K. Komatsu, A. Nakahara, T. Matsuyama and M. Matsushita, “Experimental investigation on the validity of population dynamics approach to bacterial colony formation”, Journal of the Physical Society of Japan **63**, 3, 1205–1211 (1994).

Wang, L.-Z., F. Wu, K. Flores, Y.-C. Lai and X. Wang, “Build to understand: synthetic approaches to biology”, Integrative Biology **8**, 4, 394–408 (2016).

Warren, M. R., H. Sun, Y. Yan, J. Cremer, B. Li and T. Hwa, “Spatiotemporal establishment of dense bacterial colonies growing on hard agar”, ELife **8**, e41093 (2019).

Wei, P., W. W. Wong, J. S. Park, E. E. Corcoran, S. G. Peisajovich, J. J. Onuffer, A. Weiss and W. A. Lim, “Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells”, Nature **488**, 7411, 384–388 (2012).

Wu, F., C. He, X. Fang, J. Baez, T. Ohnmacht, Q. Zhang, X. Chen, K. R. Allison, Y. Kuang and X. Wang, “A synthetic biology approach to sequential stripe patterning and somitogenesis”, bioRxiv p. 825406 (2019).

Wu, F., R.-Q. Su, Y.-C. Lai and X. Wang, "Engineering of a synthetic quadrastable gene network to approach waddington landscape and cell fate determination", Elife **6**, e23702 (2017).

Wu, F. and X. Wang, "Applications of synthetic gene networks", Science progress **98**, 3, 244–252 (2015).

Wu, F., Q. Zhang and X. Wang, "Design of adjacent transcriptional regions to tune gene expression and facilitate circuit construction", Cell systems **6**, 2, 206–215 (2018).

Xiang, Y., N. Dalchau and B. Wang, "Scaling up genetic circuit design for cellular computing: advances and prospects", Natural computing **17**, 4, 833–853 (2018).

Yabe, T. and S. Takada, "Molecular mechanism for cyclic generation of somites: lessons from mice and zebrafish", Development, growth & differentiation **58**, 1, 31–42 (2016).

Zhang, R., J. Li, J. Melendez-Alvarez, X. Chen, P. Sochor, H. Goetz, Q. Zhang, T. Ding, X. Wang and X.-J. Tian, "Topology-dependent interference of synthetic gene circuit function by growth feedback", Nature Chemical Biology pp. 1–7 (2020).

Zheng, Y. and G. Sriram, "Mathematical modeling: bridging the gap between concept and realization in synthetic biology", Journal of Biomedicine and Biotechnology **2010** (2010).

APPENDIX A

CO-AUTHOR PERMISSIONS

I certify that my co-authors, Samat Bayakhmetov, Duane Harris, Dr. Yang Kuang and Dr. Xiao Wang have given me permission, in writing, to include all material in my PhD thesis for Chapter 1.

I certify that my co-authors, Dr. Fuqing Wu, Xin Fang, Dr. Javier Baez, Thai Ohnmacht, Qi Zhang, Xingwen Chen, Dr. Kyle R Allison, Dr. Yang Kuang and Dr. Xiao Wang have given me permission, in writing, to include all material in my PhD thesis for Chapter 2.

I certify that my co-authors, Juan Melendez-Alvarez, Dr. Rong Zhang, Dr. Yang Kuang and Dr. Xiaojun Tian have given me permission, in writing, to include all material in my PhD thesis for Chapter 3.

APPENDIX B

COMPUTER CODE FOR CHAPTER 2

Due to limited space, only the main code is listed in Appendix B. However, all code is available on the author's Github (https://github.com/Changhan93/Chapter2).

```
% This code is used to estimate parameter a1 and mu and generate Figure 2.6

% Written by Duane Harris and Changhan He in 2020.

clear;
close all;
clc;

% Importing experimental data data
D = importdata('Data0820.mat');


% CHOOSE RADIUS DATA AND SAVE FILE:

%   Colony 1_37:
% C = D(1:end-1,2)*0.05;
% S = 'Params_C1_37_Fit.mat';

%   37 Degrees:
% C = D(2:5,1:4);
% C = 0.1*mean(C)';
% S = 'Params_37_Degrees_Fit_2.mat';
S = 'Params_37_Degrees_Fit_2.mat';

% Choosing fitting model
f = @PDE_Function;

% Setting visibility threshold
v = 0.03;

% Loading parameters
load('Guess_37D_Params.mat');

% P(3) and P(24) are a1 and mu, respectively
```

```
I = [3 24];

% Setting optimization options
opt = optimset('Display','Iter',... % Display Info
             'TolFun',1e-6,...   % Error Tolerance for Function
             'TolX',1e-6);       % Error Tolerance for Variables

% Set Initial Parameter
P0 = P;
X0 = P(I);

% Find Local Minimum:
[X,F,E,O] = fminsearch(@err,X0,opt,x,t,f,P,I,K,C,v);
P(I) = X;
save(S,'x','t','P','y0','K');

%Visualizing results

close all;

% Solve PDE with Initial Guess X0
Y0 = f(t,x,y0,P0);
U0 = Y0(:,:,5);

% Solve PDE with Minimizer X
Y  = f(t,x,y0,P);
%Y  = PDE_Function(t,x,y0,P);
U  = Y(:,:,5);

% Calculate Radii
R0 = rad(U0,x,v); % Radii for Initial Guess
R  = rad(U,x,v); % Radii for Minimizer

% Plotting figures
figure;
plot(t,C,'bx-',t,R,'ro-',t,R0,'ro--','linewidth',2)
title('Parameter Estimation Results (C1L)');
```

```
legend('Exact','Approx','Initial Guess','location','best');
xlabel('Time'); ylabel('Radius');
xticks(t); grid on;

% Objective Function:
function E = err(X,x,t,f,P,I,K,C,v)

    % Solve PDE
    P(I) = X;                % Replace Variable Parameters
    [~,ind] = min(abs(x-C(1))); % Find Index of Initial Colony Size
    y0 = zeros(5,length(x)); % Set Initial Conditions to Zero
    y0(5,1:ind) = P(K);      % Set Initial Col. Density to K for x<=C(1)
    Y = f(t,x,y0,P); % Solve PDE
    %Y = PDE_Function(t,x,y0,P); % Solve PDE
    N = Y(:,:,5);            % Extract Colony Radii
    R = rad(N,x,v);          % Calculate Radii for Specified Times
    E = sum((C-R).^2)/2;     % Calculate Least Squares Error

end

function R = rad(u,x,v)
% Given a solution, this function calculates the visible radius of the
% colony at each time step
% INPUT:
%   u = Matrix = Solution of the PDE
%   x = Vector = Spatial Domain of the PDE
%   v = Scalar = Visibility Threshold of the Colony
% OUTPUT:
%   R = Vector = Colony Radius at Each Time Step
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    R = zeros(size(u,1),1);   % Allocate Memory for Radius Values
    for i=1:size(u,1)
        R(i) = max(x(u(i,:)>v)); % Calculate Radius at time i
    end
end
```

APPENDIX C

COMPUTER CODE FOR CHAPTER 3

Due to limited space, only the main code is listed in Appendix C. However, all code is available on the author's Github (https://github.com/Changhan93/Chapter3).

```
% This code is used to simulate the oscillatory expression of C6 and C12
% without add-on drugs. (Figure 3.5 (Top))

% To generate other simulation results (such as Figure 3.6 and Figure 3.9),
% the user may need to adjust corresponding parameters and system
% equations.

% Written by Changhan He and Javier Baez in 2018.

function [Ce, He] = simulation_pattern_Fig1_Case1(N)
close all;
clear all;

% Setting up the domain size
N=30;
M=1000;ex = 0;

x = linspace(-N,N,M); %
xm = linspace(-N-ex,N,length(x)+ex);
xm2 = linspace(-N,N,M);

% Solving the system with given initial conditions
% Different initial conditions may give different number of rings
[~,x1] = ode45(@odesystem,xm,[1 10 6 10]);

% Global Variables to pass to spatial steady state Patternode
global Ci
global Hi
Ci = @(x) interp1(xm2,x1(ex+1:end,3), x);
Hi = @(x) interp1(xm2,x1(ex+1:end,4), x);

solinit = bvpinit(linspace(-N,N,M),[6 10 0 0]); % [1 1 100 1]
% Now solve the problem with
sol = bvp5c(@Patternode, @Patternbc,solinit);
```

```
% Evaluate the numerical solution at 100 equally spaced points

% Plotting simulation results
y = deval(sol,x);
figure(1)
plot(x,y(1,:),'r','LineWidth',2)
hold on
plot(x,y(3,:),'g','LineWidth',2)
legend('Ce','He')
xlabel('x')

Ce = y(1,:);
He = y(3,:);

% -------------------------------------------------------------
function dydx = Patternode(x,u) % Spatial Steady State
Ce = u(1); He = u(3);
global Ci
global Hi

d4= 20;
d6= 20;
Dh = 4;
Dc = 4;

Dn = 800;
dydx = [u(2);
       -(-Dc*(Ce-Ci(x)) - d4*Ce)/(Dn);
       u(4);
       -(-Dh*(He-Hi(x)) - d6*He)/(Dn)];

% -------------------------------------------------------------
function res = Patternbc(ya,yb) % Boundary Conditions
  res = [ ya(1)
         yb(1)
         ya(3)
         yb(3) ];
```

```
% -------------------------------------------------------------
function [s] = odesystem(t,x) % Ode System
I = x(1); A=x(2); Ci = x(3); Hi = x(4);
% Parameters
k1= 640; % production rate of LuxI
k2= 700; % production rate of LasI
k3= 80; % production rate of Ci
k4= 105; % production rate of Hi
k5= 1;
k6= 1;
kc= 70; % inhition efficiency of ATC
kh= 82; % inhition efficiency of IPTG
m1= 4;
m2= 4;
b1= 0.8;
b2= 0.5;
d1= 1.19;
d2= 1.19;
d3= .56;
d5= .8;
n1= 2;
n2= 4;
n3= 3;
n4= 2;

% System of Equations
s = [b1 + k1*(I*Hi).^n1/(1+k5*(I*Hi).^n1)*(1/(1+A.^m1)) - d1*I;...
    b2 + k2*(A*Ci).^n2/(1+k6*(A*Ci).^n2)*(1/(1+I.^m2)) - d2*A;...
    k3*I.^n3/(kc.^n3 + I.^n3) - d3*Ci;...
    k4*A.^n4/(kh.^n4 + A.^n4) - d5*Hi];
```

```
% This code is used to generate traveling wave profiles
% The traveling wave profiles can be used to generate Figure 3.7 B by
% plotting the concentrations at selected time points

% Written by Changhan He in 2019.

function pdex1
m=0;
global k1 k2 k3 k4 kc kh m1 m2 b1 b2 d1 d2 d3 d4 d5 d6 n1 n2 n3 n4 Dh Dc Dn

% Parameters
k1= 580; % production rate of LuxI
k2=  700; % production rate of LasI
k3= 80;  % production rate of Ci
k4= 105; % production rate of Hi
kc= 70;  % inhition efficiency of ATC
kh= 82;  % inhition efficiency of IPTG
m1= 4;
m2= 4;
b1= 0.8;
b2= 0.5;
d1= 1.19;
d2= 1.19;
d3= .56;
d5= .8;
n1= 2;
n2= 4;
n3= 3;
n4= 2;
d4= 20;
d6= 20;
Dh = 4;
Dc = 4;
Dn = 800;

% Setting up the domain size
```

```
x=0:1:240;
t=0:1:100;

sol = pdepe(m,@pdefun,@pdex1ic,@pdex1bc,x,t);
u1=sol(:,:,1);
u2=sol(:,:,2);
u3=sol(:,:,3);
u4=sol(:,:,4);
u5=sol(:,:,5);
u6=sol(:,:,6);

% Plotting figures
figure;
surf(x,t,u1)
title('I/u1(x,t)')
xlabel('Distance x')
ylabel('Time t')

figure;
surf(x,t,u2)
title('A/u2(x,t)')
xlabel('Distance x')
ylabel('Time t')

figure;
surf(x,t,u3)
title('Ci/u3(x,t)')
xlabel('Distance x')
ylabel('Time t')

figure;
surf(x,t,u4)
title('Hi/u4(x,t)')
xlabel('Distance x')
ylabel('Time t')

figure;
```

```
surf(x,t,u5)
title(’Ce/u5(x,t)’)
xlabel(’Distance x’)
ylabel(’Time t’)

figure;
surf(x,t,u6)
title(’He/u6(x,t)’)
xlabel(’Distance x’)
ylabel(’Time t’)

%-------------PDE---------------------
% Systems of equations
function [c,f,s] = pdefun(x,t,u,du)
global k1 k2 k3 k4 kc kh m1 m2 b1 b2 d1 d2 d3 d4 d5 d6 n1 n2 n3 n4 Dh Dc Dn

s1 = b1 + k1*(u(1)*u(4)).^n1/(1+(u(1)*u(4)).^n1)*(1/(1+u(2).^m1)) -
   d1*u(1);
s2 = b2 + k2*(u(2)*u(3)).^n2/(1+(u(2)*u(3)).^n2)*(1/(1+u(1).^m2)) -
   d2*u(2);
s3 = k3*u(1).^n3/(kc.^n3 + u(1).^n3) - d3*u(3);
s4 = k4*u(2).^n4/(kh.^n4 + u(2).^n4) - d5*u(4);
s5 = -Dc*(u(5)-u(3)) - d4*u(5);
s6 = -Dh*(u(6)-u(4)) - d6*u(6);

c = [1;1;1;1;1;1]; %Ds
f = [0;0;0;0;Dn*du(5);Dn*du(6)];
s = [s1;s2;s3;s4;s5;s6];

%--------------IC-----------------
function u0 = pdex1ic(x)
% global k1 k2 k3 k4 kc kh m1 m2 b1 b2 d1 d2 d3 d4 d5 d6 n1 n2 n3 n4 Dh Dc
   Dn
k = IC(x);
u0 = [k;k;k;k;0;0];

%--------------BC--------------------
```

```
function [pl,ql,pr,qr] = pdex1bc(xl,ul,xr,ur,t)
% global k1 k2 k3 k4 kc kh m1 m2 b1 b2 d1 d2 d3 d4 d5 d6 n1 n2 n3 n4 Dh Dc Dn
pl = [ul(1);ul(2);ul(3);ul(4);ul(5);ul(6)];
ql = [0;0;0;0;0;0];
pr = [ul(1);ul(2);ul(3);ul(4);ul(5);ul(6)];
qr = [0;0;0;0;0;0];

%--------------FIC-------------------
% This initial condition describes the colony growing at the center of the
% domain with a small size
function fic = IC(x)
 fic=0.0001*(x>=0&x<59)+100*(x>=59&x<=60)+0.0001*(x>60&x<=120);
```

APPENDIX D

COMPUTER CODE FOR CHAPTER 4

Due to limited space, only the main code is listed in Appendix D. However, all code is available on the author's Github (https://github.com/Changhan93/Chapter4).

```
% This code is used to estimate parameter kr1, a, gama, q1, K, r1, r2.
    (Figure 4.4 (Left))

% Written by Changhan He and Juan Melendez-Alvarez in 2021.

clc
clear
global kr0 kr1 dr
global kg J
global c1 c2 c
global a
global gama q1 q2 r1 r2 K

% Fixed parameters
J=2.123;
kr0=0.0514;
kg=1.2540;
c1=1;
c2=0;
c=6.7;
q2=0.25;
dr=0.4813;

% Initial guess for the fiting parameters
% Here we select the best fitted parameters as the initial guess
Para0=[1.76784, 10.6323, 2.2911, 0.7234, 0.0862, 4.67, 1.633];

% Load and scale experimental data
rol=importdata('M9LB.mat');
Time_EXP=rol.time;
rol.GFP_Lara(:,6) = rol.GFP_Lara(:,6)*0.001;
rol.OD_Lara(:,6) = rol.OD_Lara(:,6)*1;
GFP_EXP20=(rol.GFP_Lara(:,6))/(rol.GFP_Lara(1,6));
OD_EXP20=rol.OD_Lara(:,6);
```

```
Score_Best=1000000;
Para_best=Para0;
Score0=Score_Best;

%-------------------------------------------------------------------------
% Iteration time = 5000
for step=1:5000

    step
    pause(.000000000000001)

    Para1=Para0.*(.95+.1*rand(size(Para0)));

    temp=num2cell(Para1);
    [kr1, a, gama, q1, K, r1, r2]=deal(temp{:});

    % Solve the system in two steps
    % Step 1: solving the system with random AraC initial concentration so
    % that the circuit expression will reach its steady state
    % Step 2: Setting the steady state of circuit expression as the initial
    % condition
    sol20=ode23s(@ODESystem_new,[0 20],[0.1 OD_EXP20(1) 1]);
    sol20=ode23s(@ODESystem_new,[0 20],[sol20.y(1,end) OD_EXP20(1) 1]);

    OD_Sim=deval(sol20,Time_EXP,2);
    GFP_Sim=deval(sol20,Time_EXP,1);
    GFP_Sim=GFP_Sim/GFP_Sim(1);

    % Compute fitting error
    Score1=5*sum(((OD_Sim'-OD_EXP20)./OD_EXP20).^2
    +1*((GFP_Sim'-GFP_EXP20)./GFP_EXP20).^2);

    if Score1<Score0 || rand < .01
        Para0=Para1;
        Score0=Score1;
    end
```

```
    if Score1<Score_Best

        Score_Best=Score1;
        Para_best=Para1;

        C20=sol20.y(1,:);
        N20=sol20.y(2,:);
        P20=sol20.y(3,:);

        % Plotting figures
        subplot(3,2,3)
        plot(sol20.x,N20,'linewidth',2)
        hold on
        plot(Time_EXP, OD_EXP20,'o')
        ylabel('N')
        xlabel('Time')
        hold off

        subplot(3,2,4)
        plot(sol20.x,C20/C20(1),'linewidth',2)
        hold on
        plot(Time_EXP, GFP_EXP20,'o')
        hold off
        ylabel('AraC')
        xlabel('Time')

        subplot(3,2,5)
        plot(sol20.x,P20)
        ylabel('Nutrient')
        xlabel('Time')

    end
end
```

```
% This code describes Model II with nutrient preference.

% Written by Changhan He in 2021.

function dy=ODESystem(t,y)

global kr0 kr1 dr
global kg J
global c1 c2 c
global a
global gama q1 q2 r1 r2 K

AarC=y(1);
N=y(2);
P=y(3);

gA = 1/(AarC/J+1)+c2 ;
fA = kr0+kr1*(AarC^2)/(AarC^2+c1);
W1 = 0.5*(sign(P-q1)+1);
W2 = 0.5*(sign(q1-P)+1);
G1 = W1*(r1*(1-q1/P));
G2 = W2*(r2*(1-q2/P));
GR = kg* gA * ( G1 + G2 );
F1 = 1/(K*GR+1);
F2 = (a*GR^0.5+1)/(c*GR^2+1);

dN=GR*N;
dAarC=fA*( W1*F1 + W2*F2 )-(dr+GR)*AarC; %AarC
dP=-gama*N*(P-q2)^1.2;

dy=[dAarC; dN; dP];

end
```