# 博士論文

# Evolution-Development Congruence in Pattern Formation Dynamics

## (パターン形成ダイナミクスにおける進化-発生対応)

香曽我部隆裕

# Contents

**1 Introduction** **1**
1.1 Preface: brief history on recapitulation . . . . . 1
1.2 Gene Regulation Network . . . . . 4
1.3 Pattern formation in reaction-diffusion systems . . . . . 5
1.4 Evolutionary simulation . . . . . 6
1.5 Temporal control of developmental processes . . . . . 7
1.6 Overview . . . . . 8

**2 Evolution-Development Congruence in Pattern Formation Dynamics** **11**
2.1 Introduction . . . . . 12
2.2 Methods . . . . . 17
2.3 Results . . . . . 23
2.4 Discussion . . . . . 45

**3 Boundary-Induced Pattern Formation from Temporal Oscillation: Spatial Map Analysis** **59**
3.1 Introduction . . . . . 60
3.2 Results . . . . . 61
3.3 Discussion . . . . . 70

**4 Time-scale separation of gene expression and evolvability** **79**
4.1 Introduction . . . . . 80
4.2 Method . . . . . 81
4.3 Result . . . . . 83
4.4 Discussion . . . . . 93

**5 Summary and Future issues** **95**

# Chapter 1

# Introduction

## 1.1 Preface: brief history on recapitulation

Let us consider development of animals. During development, embryo is transformed to a variety of shapes; from a fertilized egg to brastula, gasutula and further to some morphogenetic stages. As these transformation of embryo during development are commonly observed, we might take it for granted that development goes in that way. However, considering the goal of development is to form a viable adult form, development does not take straightforward path to shape an adult, but rather, involve roundabout, redundant process. For example, some organs that are once formed even vanish; Limb bud which is not necessary for snakes emerges in the development(Cohn and Tickle, 1999). To be extreme, organs do not necessarily have to emerge from anlagen. Development of animals could have been evolved in the way that start from making a miniature adult, then cell differentiations are induced to make organs and the size of the miniature would increase gradually. Teleologically speaking, that could be possible and embryo would not have to transform from one form to another during development. For example, jaws of land animals do not have to emerge from pharyngeal arches that is expected to be gill arch in fish development(Gilbert, 2014).

For some reason, development have not been evolved in such "anything-going" way. At least, developmental process that shapes basic body structure seems to be conserved during evolution. Those in the field of anatomy and taxonomy had noticed that embryonic changes in development is sometimes parallel to nesting structure of taxonomy, i.e., body structure of the embryo in the earlier stage often takes anatomical features that are shared by species in a particular phylum, and then as development goes, the shape of embryo diversifies to represent features that are shared in lower taxonomic domain; class, order, family, genus and species(Kuratani, 2017). Such developmental process that parallels to taxonomic order is called recapitulation (Shown in Fig. 1.1). Since the taxonomic order is generally believed to represent the evolutionary order, recapitulation then can be rephrased that developmental process parallels to evolutionary order.

**Figure 1.1. Schematic representation of recapitulation**
Illustration of taxonomy, shown together with ontogenic time course. Taxonomy has nesting structure such that a higher domain includes lower specific domains. When recapitulation occurs, embryonic change parallels to the descent in such taxonomic structure. The embryo of earlier stage shows anatomical features that defines higher domain. As ontogeny undergoes, embryo changes to show anatomical features of a lower domain. Correspondence between domains and ontogenic stages are represented by lines in the figure.

Basis of recapitulation was proposed almost 200 years ago, by Karl Ernst von Baer who proposed there exists a typical stage that every species in the same animal group has to pass through during ontogenesis(von Baer, 1828). That stage exists for each animal group, which is now known as phylum. Based on recent findings, it would be reasonable to think that common structure of body for each phylum is established at that stage. Nowadays the stage is known as the phylotypic stage. von Baer claimed that specie-specific traits do not emerge in advance of the phylotypic stage; development goes in the way that change from common to specific(von Baer's laws of development). [1] Remarkably, this developmental change from common to specie-specific shape parallels to the hierarchical taxonomic order, as mentioned just before. The similarity in developmental process of the animals in the same phylum is the basis of the recapitulation[2].

30 years later, Charles Darwin published "the origin of species" where he reveals that evolution occurs by means of natural selection on population with slightly different individuals(Darwin, 1859). Evolution was proposed to progress among as descents with modification. There he strongly suggested that conservation of body plan is the proof of the common ancestor. As von Baer suggested, although the appearance of each adult animals looks different, basic body plan is preserved, as long as those animals are in

[1] Embryonic stage earlier than phylotypic stage is not considered to be a proper subject of embryogenesis at his era(Sander and Schmidt-Ott, 2004), so that such earlier stages might not be regard as valid target of his laws.

[2] von Baer did not believe in evolution so that his claim was only on embryogenesis, at that time(Gould, 1977)

the same phylum. Current animals' body structures can be derived from basic body plan that emerges at the phylotypic stage with modifications. Conserved body plan and developmental change from common-to-specific has come to issues on evolutionary significance(Gould, 1977, Hall, 1999).

In response to such trend, Ernst Haeckel tried to integrate comparative embryology and evolution theory(Haeckel, 1866). He claimed that the conservation of body plan and developmental change from common to specific is the consequence of evolution and proposed the hypothesis why recapitulation occurs. The hypothesis is called as "biogenic law" which claims what happen in evolution is the addition of novel developmental process at the terminate of ontogenesis. Thus during the development of advanced animal, adult organisms of primitive species appear according to the evolutionary order. This idea is summarized as famous statement "ontogeny recapitulates phylogeny". He thought embryogenesis as recapitulated evolution from which von Baer's law of development is naturally derived. Even though, Haeckel represented recapitulation theory as a general law in nature, he did not show developmental mechanism that generate recapitulation. Thus recapitulation theory was criticized by experimental biologists who study embryology with empiric, physiology-oriented methodology(Gould, 1977). To these experimental biologists, recapitulation theory seems too dogmatic[3][4]. Recapitulation theory was never proven and then gradually declined. Unfortunately, recapitulation as phenomena was also dismissed together with the subject of study, even though the recapitulation theory and recapitulation as phenomena are different things.

Years later, with quantitative comparative analysis, existence of phylotypic stage has been clearer than ever, as well as its significance on the independence of each phylum and conservation of body plans(Hazkani-Covo et al., 2005, Irie and Sehara-Fujisawa, 2007, Domazet-Lošo and Tautz, 2010, Kalinka et al., 2010, Irie and Kuratani, 2011, Quint et al., 2012, Levin et al., 2012, Wang et al., 2013, Irie and Kuratani, 2014). It is observed that phylotypic stage is in the middle of embryogenesis and similarity among species of the same phylum converges at the stage. After the phylotypic stage, the similarity decreases as each embryo transforms to specie-specific state. Thus the phylotypic stage is the bottleneck of developmental similarity as is called as "hour-glass model" or "hour-glass hypothesis"(Sander et al., 1983, Duboule, 1994, Raff and Wolpert, 1996).

Hour-glass model refutes recapitulation theory: If Haeckel's claim were true as it is, the earlier the embryonic stage is, the more species in the same phylum would resemble each other [5], which is so far observed as not true. However, hour-glass hypothesis does not renounce recapitulation as phenomena since it can be said that diverging similarity after phylotypic stage is recapitulation. In this sense, hour-glass model does not necesarily deny recapitulation, but rather, can be regarded as the modern extension of it; now embryogenesis is not just a diverging process from common state to specific state in the

[3]Of course, it also seems too dogmatic to biologists nowadays

[4]However, I cannot help myself admiring Haeckel's passionate imagination on evolution.

[5]Haeckel dreamed the possibility of similarity in species from different phylum at the very early stages of ontogenesis(Sato, 2016)

order of taxonomy, but also includes a converging process to the common state in the very beggining. The question why recapitulation takes place still remains. [6]

There is a hypothesis to address the question of why recapitulation occurs, which is called developmental burden hypothesis (Riedl, 1978). In developmental burden hypothesis, organs that formed through each developmental process are regarded to give positional or temporal information that induces ensuing processes. In other words, preceding developmental processes lay the foundation of the following processes, and mutation on these preceding processes may bring about collapse of the following processes. Thus such preceding processes and their order in development are supposed to be conserved both in development and evolution. Then, recapitulation will take place.

The developmental burden may be a reasonable hypothesis. However, that is only a conceptual explanation, as practical questions to test the hypothesis remain to be unanswered. Is that really possible to implement such burdens in real biological system? Is there any general condition that recapitulation takes place? In what case recapitulation is broken? To answer these questions, the mechanism of recapitulation that is in the framework of modern developmental biology has to be investigated.

## 1.2 Gene Regulation Network

In search of mechanism of recapitulation, where can we start? Considering that what happens in evolution is modification on developmental process, it is good to start first how development is controlled by genes. Gene regulation network, then, provides a basic general framework for the study of controlling cellular functionality.

In order for cells to function a particular role, expressions of genes in the cells have to be coordinated. It does not only mean that all necessary genes have to be expressed, but also all unnecessary genes must not be expressed. This regulated expression of genes are controlled by transcription factors that have sequence-specific affinity(Jacob and Monod, 1961). Ordinarily, one transcription factor controls multiple genes as it has multiple binding sites in the genome. More importantly, such transcription factors are also regulated by other transcription factors(Spitz and Furlong, 2012). This cross-regulation of multiple genes in genome shapes a network as a whole that is called as gene regulation network(GRN, for short). Although GRN basically considers coordination of gene expressions in a single cell, some transcription factors are diffusible or encode proteins that constitute membrane, or trigger cell-cell signaling (like proteins involved in Delta-Notch signaling pathway). Hence GRN in a cell affects the gene regulation of neighboring cells which is compatible with the formulation by reaction-diffusion equations to be described

[6] Needless to say, question of why embryos of animals in the same phylum once converges to the phylotypic stage was posed at the time.

in the next section.

In late 20th century, genes that are responsible to make body structure were investigated, providing remarkable knowledge. First, only a small subset of the entire genes in the genome affects development. Second, these genes mainly encode transcription factors. Third, mutation on these small subset of genes bring unusual phenotypes that is out of one's body plan(Carrol, 1995). Most interestingly, these genes are widely conserved among different animal phyla(Gerhart and Kirschner, 1997). These genes are called toolkit genes or developmental toolkit. Knockout experiments of those toolkits on *Drosophila* and according bizarre phenotypes are well known(Carroll et al., 2009). For development work properly in harmony, coordination between toolkit genes must be important. Thus the GRN of toolkit genes are expected genomic basis of developmental process.

The mechanism of GRN function against various inputs has been investigated in systems biology(Alon, 2007). The dynamics of gene expression is basically based on GRN's structure. For example, it is well understood how GRN without backward loop can transform maternal morphogene gradient into different expression patterns depending on distinct spatial regions(Alon, 2007), which is GRN interpretation of Wolpert's "French flag model"(Wolpert, 1969). GRN with backward loop on the other hand can generate oscillatory gene expression as is also demonstrated experimentally(Elowitz and Leibler, 2000). Now it is widely accepted that GRN with maternal locational information can control embryonic patterning.

From evolutionary point of view, GRN is thought to have hierarchical structure. "Kernel" network structure that determines body plans for each phylum is considered to exist(Davidson and Erwin, 2006). The kernel, then is assumed to have been conserved in each phylum from its beginning and minor modifications/additions of other interactions at the downstream of the GRN give speciation within the phylum. In other words, the hierarchy in GRN structure and that of taxonomic categories are supposed to representat recapitulation. GRN structures of various animals are analyzed to test the hypothesis(Davidson et al., 2002, Stathopoulos and Levine, 2005, Imai et al., 2006). So far, however, evolution of such GRN of toolkit genes are not much understood, while GRN is still the format that is most fundamental and relevant tool to study mechanism of evolution of body plan.

## 1.3 Pattern formation in reaction-diffusion systems

There is another approach to the mechanism of development. This approach was first introduced by British mathematician Alan Turing. Turing studied a chemical system which has mainly two chemical components that interact each other. He assumed these chemicals undergo reactions and assumed to have diffusive ability. This setup is now widely known as "reaction-diffusion system"(Turing, 1952). He demonstrated that even such sim-

ple system can generate a variety of temporal and/or spatial patterns of chemicals that are organized from a uniform state. Especially, the system with an activator(the chemical activates the reaction of itself and the other) and an inhibitor(the chemical inhibits the reaction of itself and the other) is of interest. When the ratio off diffusion constants of activator and inhibitor satisfies a certain condition, a uniform state in space is unstable, and spatially periodic and temporally fixed pattern emerges as a stable state(Turing pattern).

Since then, pattern formation in reaction-diffusion system has been extensively studied theoretically on more details and with realistic applications(Gierer and Meinhardt, 1972, Murray, 1981), although its relevance to real development was still controversial. In 1996, Turing pattern was shown to be relevant to skin patterning of marine angelfish(Kondo and Asai, 1996). After that, pattern formation of reaction-diffusion system has been much popular among biologists. Study of the reaction-diffusion system assuming specific proteins is carried out with molecular-biologiy analysis(Kondo and Miura, 2010).

Reaction-diffusion system is also studied in evolutionary context. These studies basically assume GRNs that determines reaction, in which parameters or structures of GRN change with evolution. Solutions of a reaction-diffusion system sometimes change qualitatively with continuous change of parameters that is described by bifurcation theory. Such qualitative change via continuous change of parameters is supposed to be relevant to drastic change of pattern during evolution(Francois and Siggia, 2012). In fact, it is shown that parametric change in the Turing system induces fin-to-limb transition where conserved Turing mechanism underlie the development of both fin and limb(Onimaru et al., 2015). Reaction-diffusion system is not only the format of studying development, but also can provide a basic format to study the evolution of pattern formation.

## 1.4 Evolutionary simulation

Even though reaction-diffusion system with GRN framework provides a strong basis on the study of developmental mechanisms, it has not been much applied to study of evolution of developmental mechanisms. It is not just because applying these frameworks to experiment is difficult. Rather, experimental study of evolution itself is difficult, because evolutionary data is not complete. We have to rely on fossil records for evolution of developmental process. However they are very sparse, especially at the embryonic stages. Further, fossils usually do not stores genome data so that evolutionary change of GRN cannot be directly at hand. One-time-only nature of evolution makes things much worse. Evolutionary hypothesis that is derived from these fossil data cannot be tested by another evolution. We cannot experimentally setup the condition of passed time, and even if it were possible, it takes enormous time to study the evolution there. Then, we cannot test whether current developmental mechanism was accidentally selected in evolution or was selected for plausible reasons.

Computational evolutionary simulation can be a strong aid for this situation. Such computational simulation starts with population of individuals that carry randomly generated genotype that encodes spatial patterning to shape its phenotype. Fitness for each individual is determined by a particular function of phenotypes. The genotypes that encode phenotypes of higher fitness are selected and subject to mutation so that population can maintain variation of individuals. With these setup, evolution of pattern formation can be mimicked and one can study what type of mechanism for pattern formation will emerge under a particular fitness constraint on the phenotype. These computational simulations are advantageous as complete data through evolution are available. These data are *literally* complete as they keep the track of the pedigree as a chain of mother-to doughtersp. Analysis on such single-chain-phylogeny is the foundation of further species-wide comparison.As another advantage, these computational simulations are repeatable so that with multiple simulation runs, we can assess the plausibility of evolution of a particular mechanism that is found in a single-run simulation. Additionally, we can find possible mechanism that may not be selected in reality. Computational simulation can capture such mechanisms that are relevant to real biological system and/or plausible evolutionary trails.

Computational simulation was used to study evolution of GRN that governs segmentation. It was first exhibited in 2001 investigating mechanism of segmentation and its robustness(Salazar-Ciudad et al., 2001b,a). After that, plausible evolutional pathway of mechanism of segmentation was studied further. They suggest that simultaneous segmentation that observed in development of *Drosophila* and sequential segmentation of vertebrates may have the same origin of GRN(François et al., 2007), evolution of simultaneous segmentation may be induced by decreasing developmental time(Fujimoto et al., 2008), or segmentation precedes speciation of body parts(ten Tusscher and Hogeweg, 2011). The mechanism of segmentation found in these models can be relevant to understand that in real biological system, and with plausible hypothesis on evolution of segmentation(ten Tusscher, 2013). Computational evolutionary simulation is a hopeful tool for investigating mechanism of development and generating plausible hypothesis on evolution, that compensates weak point of experimental study of evolution.

## 1.5 Temporal control of developmental processes

So far, I summarized evolutionary/developmentally importance of recapitulation in section 1.1. To study the mechanism of recapitulation, frameworks of GRN and reaction-diffusion system would be relevant as described in section 1.2 and 1.3. Computational evolutionary simulation that involves GRN and reaction-diffusion system will be an appropriate direction to take. Still, we are missing one more thing. That is the control in developmental time course.

Recapitulation is the developmental phenomena that anatomical features of embryo appear in the order of evolutionary acquisition. There, it is assumed that different de-

velopmental mechanisms function to form different anatomical features. In development, each proper process works in the proper period of time, at a proper location, one and another. Each independent developmental mechanism could be captured by the frameworks of GRN or reaction-diffusion system. However, that is not enough to capture recapitulation as a whole procedure. To explain the mechanism of recapitulation, the mechanism that makes the order of developmental processes *in time* also has to be revealed. Recall again that, developmental processes do not evolve in anything-going way as I discussed in section 1.1, so that it is natural to assume that such mechanism of time control to make processes in order will lie at the as the center of origin of recapitulation.

Interestingly, there are not much studies on the mechanism of such temporal coordination of developmental processes. As an example of temporal control, sequential segmentation are thought to be controlled in time course by oscillation of gene expression coupled with morphogene gradient that shifts along with the growth of embryo(clock and wavefront model, (Cooke and Zeeman, 1976)). Evolutional simulation of segmentation suggested genes which indicate growth of embryo are in the upstream of GRN and those genes control temporal gene expression(François and Siggia, 2010). Still, relevance of such growth indicator to recapitulation has not been investigated at all. Moreover, why such time control will emerge during evolution remains as a deeper question to be solved.

## 1.6 Overview

In this thesis I will investigate the mechanism of recapitulation on the single chain phylogeny from ancestor to offsprings, that is obtained by evolutionary simulation. I evolved GRNs that govern pattern formation dynamics with diffusive interaction, to find that under the strong purifying selection, recapitulation in gene expression patterning between evolution and development on each chain. The term "evolution-development congruence" will be coined to address this of recapitulation in gene expression patterning.

This thesis consists of 5 chapters, including this introduction. Each chapter has different aims. **In chapter 2**, I seek for possible conditions and mechanisms for evolution-development congruence by use of computational simulation. With extensive simulations, I demonstrate that evo-devo congruence is not accidental, but rather naturally emerges under the model setup with strong purifying selection. The mechanism of pattern control in time and space are uncovered by network analysis and dynamical systems analysis. I will show that pattern control in space is encoded in the GRN structure with static morphogene gradient, while pattern formation dynamics are contorolled in time by slowly changing variable which spontaneously emerges during evolution. Exceptional cases where evolution-development congruence is broken are also found and explained with dynamical systems and network analysis.

**Chapter 3** is devoted to the detailed analysis on a pattern formation dynamics that I found in the evolutionary simulation in chapter 2. The appearance of the dynamics

looks similar to sequential segmentation of clock and wavefront model(Cooke and Zeeman, 1976). However, this pattern formation dynamics is different from the clock and wavefront model in that the dynamics do not need moving gradient. Instead, the dynamics are induced by boundary condition and diffusion at the boundary. Since the dynamics emerge from oscillating state, standard linear stability analysis for a fixed state does not work for analyzing the dynamics. I introduce nonlinear spatial map that is derived from the equation of stationary state, and demonstrate that the spatial map reproduces well the pattern , and is powerful to predict the stationary pattern.

The relevance of slow gene control of development to the evolution is investigated **in chapter 4**. With another model of evolutionary simulation, that are simpler versions of the model in chapter 2, I study whether the evolvability of slow gene corresponds to that of pattern formation itself. The result suggests that slow gene is relevant to evolvability of pattern formation. Moreover, it is observed that deletion of slow gene from evolved GRN leads to a larger damage to the pattern than other genes. Slow genes are shown to be emergent control parameter(s) in GRN.

**In chapter 5**, I will summarize the results, and discuss their relevance to biological systems, while providing promising directions for future studies.

# Chapter 2

# Evolution-Development Congruence in Pattern Formation Dynamics


### Summary

Here we uncover recapitulation, one of the oldest phenomena that suggests relationships between evolution and development by numerical evolution and unveil their origin in terms of dynamical systems theory. By representing developmental dynamics of spatially located cells with gene expression dynamics with cell-to-cell interaction under external morphogen gradient, gene regulation networks are evolved under mutation and selection with the fitness to approach a prescribed spatial pattern of expressed genes. For most numerical evolution experiments, evolution of pattern over generations and development of pattern by an evolved network exhibit remarkable congruence. Both in the evolution and development pattern changes consist of several epochs where stripes are formed in a short time, while for other temporal regimes, pattern hardly changes. In evolution, these quasi-stationary regimes are generations needed to hit relevant mutations, while in development, they are due to some gene expression that varies slowly and controls the pattern change. The morphogenesis is regulated by combinations of feedback or feedforward regulations, where the upstream feedforward network reads the external morphogen gradient, and generates a pattern used as a boundary condition for the later patterns. The ordering from up to downstream is common in evolution and development, while the successive epochal changes in development and evolution are represented as common bifurcations in dynamical-systems theory, which lead to the evolution-development congruence. Mechanism of exceptional violation of the congruence is also unveiled. Our results provide a new look on developmental stages, punctuated equilibrium, developmental bottlenecks, and evolutionary acquisition of novelty in morphogenesis.



*The content of the present chapter is published in Journal of Experimental Zoology Part B:Molecular Development Evolution. 2016 Jan;326(1):61-84..*

## 2.1 Introduction

The possible relationships between the development of multicellular organisms and their evolution have been the subject of intense research over a century. Recapitulation is one of the oldest phenomena that suggests such relationships. As I introduced in chapter 1.1, the basis of recapitulation was proposed by von Baer as laws of development, based on observations of development across species, which mainly claimed that the early embryo is mostly conserved across species, while embryonic changes through ontogeny move from a general form common to many species, to species-specific forms (von Baer, 1828). Then, Charles Darwin and other biologists of his time interpreted these laws as proof of evolution from a common ancestor (Darwin, 1859, Müller, 1869, Hall, 2000). Changes in embryos from a common to a specialized form are regarded as a reflection of evolutional history. Development has been studied in an evolutionary context, and so many biologists have searched for possible relationships between evolution and development. The most popular concept might be Ernst Haeckel's recapitulation theory, or simply the statement "Ontogeny recapitulates phylogeny" which assumes that development is literally the accelerated evolution, i.e., embryo in development is the same as one of an adult of its ancestors and during development, embryo changes from one ancestor to another according to the evolutional order (Haeckel, 1866).

The idea of recapitulation theory, of course, is not acceptable today, as the idea is too dogmatic without mechanistic explanation(See section 1.1). There are also many counterexamples against it(Gould, 1977). However, as modern thesis, possible recapitulation at the gene expression pattern needs to be tested. It is frustrating that although existence of recapitulation-like phenomena has been found, contemporary evolutional-developmental biology still cannot provide general concept for it which replaces old recapitulation theory(Gould, 1977, Hall, 1999, Richardson and Keuck, 2002).

Previously, this type of research was hindered by a lack of quantitative arguments. To transcend the century-long controversies associated with this research, efforts are being made to quantitatively analyze the evolution-development ('evo-devo') relationship by using gene expression and genome sequence data. In particular, it is becoming clear that there exists a developmental bottleneck among species in some phylum, where differences in gene expressions among several species from the same progenitor decreases at the same developmental stage (Hazkani-Covo et al., 2005, Irie and Sehara-Fujisawa, 2007, Domazet-Lošo and Tautz, 2010, Kalinka et al., 2010, Irie and Kuratani, 2011, Quint et al., 2012, Levin et al., 2012, Wang et al., 2013). These studies are consistent with the developmental hourgrass hypothesis that refutes recapitulation theory(Sander et al., 1983, Duboule, 1994, Raff and Wolpert, 1996).

In spite of these advances, however, recapitulation as phenomena, as well as its origin, if it exists, remains to be unveiled owing to the limitations in available data on developmental processes along the evolutionary course. Species-wide comparison is made using phylogenic trees with branchings to different species, as schematically shown in Figure 2.1. In contrast, one can make a comparison over species along a single phylogenetic chain from ancestor to offspring. Ontogenies are compared across ancestors along a single evolutionary chain, as shown in Figure 2.1. This comparison is hardly possible in practice, as fossil data usually do not include information on developmental processes. However, such comparison, if available, gives more straightforward information on relationship between development and evolution, and will provide a basis for species-wide comparison. With such single-chain comparisons, one can gain insight of possible mechanisms that may give evo-devo relationship.

There has arisen alternative approach in evo-devo, which uses in-silico evolution. It adopts dynamical-systems theory for development, as a process to generate pattern dynamics for protein concentrations (expressions). In spite of simplification needed to model the developmental and evolutionary processes, it has the following advantages One can trace the phylogenic tree of every single individuals through the course of evolution. Dynamical process for development (embryogenesis) of ancestors are available, as well as their genomic information. By taking advantage of these features, and applying heuristic approaches based on extensive numerical simulations, it will be possible to uncover evolution-development relationship, and examine the validity of the hypotheses proposed so far. Indeed, several numerical evolution of developmental process has been recently carried out. The models in these studies consider the spatial arrangements and behaviors of cells that are subject to morphogenic input, while protein expression levels within the cells change over time by intra-cellular gene expression dynamics and also by cell-to-cell interaction. The developmental processes of cellular states are represented by these gene expression dynamics to form a spatial pattern of expressions, while the gene regulation networks associated with these dynamics evolve through modifications by genome changes. By establishing a fitness condition of making pattern with particular features, the evolution of the network to generate such a pattern can be studied quantitatively. Indeed, with the fitness condition of rewording segment number, recent studies of in silico evolution have suggested basic mechanisms for stripe formation through development and accessed plausible evolutionary scenarios(Salazar-Ciudad et al., 2001b,a, François et al., 2007, Fujimoto et al., 2008, ten Tusscher and Hogeweg, 2011).

**Figure 2.1. Schematic representation of single-chain-phylogeny and comparison across species in phylogenic tree**
Schematic representation of the comparisons along single-chain phylogeny and across species. In the phylogenic tree shown here, the currently existing species represented by the right-end circles, are diverted from a common ancestor. Branching from a common ancestor leads to establishment of some new species, while some are terminated by extinction. The comparison of developmental processes across species is made over the existing species. On the other hand, a single phylogenic chain, which we focused in this study, is given as the line from the common ancestor to the offspring in concern. From the species in concern, ancestors are uniquely traced back. The comparison of developmental processes along this chain is possible at least in theory or simulations, which provides fundamental information on possible relationship between development and evolution, including recapitulation.

### Our strategy

Following the earlier theoretical studies for the development to form stripes, we focus here on uncovering recapitulation at gene expression level and understand them in dynamical-systems and network analysis. Cells are allied in 1-dimensional space, each of which consists of a set of proteins. The cellular state is thus represented by a set of protein concentrations whose temporal changes are governed by intra-cellular gene regulation network and cell-to-cell interaction via diffusion of some proteins. (Note that the diffusive interaction is rather general in developmental process. For example, pattern formation by mechanical stress can also be represented by equations of the reaction-diffusion type (Murray, 2002). We introduced the fitness so that the gene expression of a given output gene in space approaches a prescribed spatial pattern (which is not necessarily periodic in space as in the earlier studies). Accordingly, our evolution is under strongly purifying selection. With this fitness condition, we can trace how a certain pattern is evolved on single-chain phylogeny of ancestor to descendant, which is different than species-wide comparison in experiments (See Fig. 2.1). Comparing the developmental processes to shape a given gene expression pattern through an evolutionary course under mutation, we examine the developmental dynamics of offspring achieving high fitness and evolutionary change in the patterns. We found parallelism between evolution and development along the single-chain phylogeny, as is named as "evo-devo congruence", which is observed for the majority of simulation instances. We will show that both development and evolution consist of a few epochs that rapidly change to form new stripes, and the quasi-stationary regimes between rapid changes. In this thesis, other than this chapter, we analyze evo-devo relationship in terms of dynamical-systems theory and network analysis(Strogatz and Balachandran, 1994, Hirsch et al., 2012, Alon, 2007, Kaneko, 2006). For people who may not be familiar with the theory, here we provide brief description of key terms as glossary table so that one can come back to it later. In addition, we also prepare a column for the relationship between mechanisms of pattern formation and structure of gene regulation network that govern them.

## Glossary

*Dynamical systems*:
Mathematical formulation of temporal change of state variables. Dynamical system formulates the temporal change of state by how a point in the state space moves according to the fixed rule for time development. The rule gives a flow on the state space. Application to Biology: Consider a cell whose state is represented by a set of (protein) concentrations of k species, $x_1(t), ..., x_k(t)$, i.e, expression level of each gene. Then the cellular state is represented by a point in a k-dimensoinal state space, each of which axis gives expression level of each gene (protein). According to the reaction process involving gene regulation, the cellular state changes in time giving a trajectory in this state space, as formulated by a dynamical system.

*Attractor*:
After some time, the state in dynamical systems is attracted into, and remains at, a certain portion of state space. This region is called attractor. Sometimes, temporal change in the state ceases after sometime, i.e., a fixed state is reached. This is called fixed point attractor. For some attractor, the state oscillates periodically in time known as limit-cycle attractor.

*Nullcline*:
A line (or plane) in the state space in which one variable does not change in time, i.e., the flow in the direction of one axis is null. Nullcline is useful in depicting the flows in the state space, while fixed-points are obtained as the intersection of all nullclines.

*Bifurcation*:
Dynamical systems include control parameter(s) that deterimine the rule of temporal change. Often, the attractor changes continuously against the change in parameter, without qualitative/topological change. However, when a parameter reaches some value, qualitative change in the attractor occurs; for example change from one fixed-point attractor to another, or change from fixed-point to limit-cycle attractor. Such change is termed as bifurcation.

*Pattern*:
Besides the intra-cellular dynamics of a cell, cells that are aligned in space interact with each other via diffusion. Thus the protein concentration forms a pattern in space, which changes in time. This pattern dynamics is represented by coupled dynamical systems in which cells with intracellular dynamical systems are spatially allied.

*Feedforward (network)*:
With the gene regulation network, each protein expression mutually influences, with activation or inhibition. When the expression of gene A influences B but there is no direct or indirect influence from B to A, (i.e., there is no loop between A and B), the structure from A to B is called as feedforward.

*Feedback (network)*:
In contrast, if the two genes influence on each other' s expression, i.e., if there is a loop between two genes, it is termed as feedback network. When A activates B and B inhibits A (or vice versa), the feedback structure can generate oscillatory gene expression, as adopted in segmentation clock.

*Epoch (the term introduced in the present thesis)*:
The short period of time in which pattern changes drastically, where the expression of output gene (state) exhibits a distitct change. in development: In the present system, some variables change slowly in time, during a quasi-stationary regime. Its time span is much longer than the epoch with the drastic change. If this slow change influences the output pattern, it can be regarded to work as a control parameter. Then the drastic change of output occurs as a result of bifucation against this control parameter. in evolution: Mutation causes the change in gene regulation network, leading to the change in the reaction path and accordingly the rule of dynamical systems. With this change, the pattern shows a drastic change, which, as will be shown, is represented also as bifurcation against a change in the reaction path as a control parameter.

The organization of the present chapter is as follows. In the present chapter, after presenting the numerical evidences for evolution-development congruence, while we explain its origin in terms of dynamical systems. We will be show that both development and evolution adopt the same type of bifurcation, while the structure in gene regulation networks to support the correspondence is understood as a combination of upstream feedforward-network with downstream feedforward- or feedback-networks. In rare examples, however, evo-devo correspondence is found to be violated, whose mechanisms are also elucidated. After first examining extensive numerical results to support the above conclusion, its relevance to biological development and evolution will be discussed.

## 2.2 Methods

**Table 2.1.** Parameters and initial settings for the simulation

| | |
|---|---|
| Numbers of network nodes | 16 (fixed) |
| Total cell number (space size) | 96 (fixed) |
| Numbers of average edges in the initial network | 16 |
| Range of $\gamma$ (randomly selected) | [0:1] |
| Range of $D$ (randomly selected) | [0:1] |
| Range of $\theta$ (randomly selected) | [0:1] |
| Degree of nonlinearity of the reaction term $\beta$ | 40 |
| Population size | 100 (fixed) |
| Mutation rate $\mu$ | 1/256 |
| Degree of evolutionary fluctuation of parameter $\sigma$ | 0.01 |

### Gene Regulation Network(GRN) model for pattern formation

A cell's state is represented by the expression levels of k genes/proteins, $x_i(l,t)$, involving the protein expression levels of the $i$-th gene in the $l$-th cell at time $t$, representing $N$ genes ($i = 1, ., N$) and $M$ cells, aligned in a one-dimensional space. A protein expressed from each gene either activates, inhibits, or does not influence, the expression of other genes, in addition to itself. For simplicity, we assumed that the change in the $i$-th protein expression level is given by the equation:

$$\frac{\partial x_i(l,t)}{\partial t} = \gamma_i(F(i,l,t)) - x_i(l,t)) + D_i \frac{\partial^2 x_i(l,t)}{\partial l^2} \tag{2.1}$$

with

$$F(i,l,t) = f(\sum_j J_{i,j} x_j(l,t) - \theta_i) \tag{2.2}$$

where the term $-x_i(l,t)$ in (1) provides a measure of the degradation of the $i$-th protein with $\gamma_i$ as its rate (Glass and Kauffman, 1973, Mjolsness et al., 1991, Salazar-Ciudad et al., 2001a,b). The expression level is scaled so that the maximum level is unity. The function $f(x)$ is similar to a step function, where the function approaches 1 as $x$ is increased to a positive side, and approaches 0 as $x_i(l,t)$ is decreased to a negative side: In other words, if the term $\sum_j J_{i,j} x_j(l,t)$ is sufficiently larger than the threshold $\theta_i$, then $F(i,l,t) \sim 1$, which indicates that the gene is fully expressed, and if it is smaller than the $\theta_i$, then $F(i,l,t) \sim 0$, which indicates that the gene expression is suppressed. Here, we chose, $f(x) = 1/(1+e^{-\beta x})$, where $\beta$, which was set to 40, denoting the sensitivity of the expression at the threshold. Roughly speaking, it is proportional to the Hill coefficient.

The gene regulation network was introduced to our model based on work reported in earlier studies. In Figure 2.2, each node of the network represents a gene, and the edge of the network represents the interaction between genes, given by $N \times N$ matrix $\boldsymbol{J} = \{J_{i,j}\}$: where $J_{i,j}$ is 1, if gene $j$ activates the expression of the gene $i$, $-1$ if it suppresses the expression, and 0 if there is no connection. All cells have an identical regulatory network, with the same parameter values, which are determined by genetic sequence in the genome.

Finally, the last term in Eq. (1) shows the diffusion of a protein, between neighboring cells, with $D_i$ as the diffusion constant. For the majority of the simulations described here, we set $M = 96$, and $N = 16$, while preliminary simulations adopting larger values for these did not alter the conclusion in the present chapter.

**Initial/Boundary Condition**

As an initial condition, the expression levels of all genes were set to 0. Furthermore, external morphogens, which are denoted as the proteins 0 and 1, are supplied externally. Fixed linear morphogens are induced from both sides for cellular use, so that $x_0(l,t) = x_0(l) = (M-l)/M$ and $x_1(l,t) = x_1(l) = l/M$. We also evaluated a case involving a gradient with an exponential dependence in space, as $C\exp(-l/\xi)$, but this condition did not alter the conclusions presented in this study. Discrete Neumann boundary conditions were adopted at both ends for this study, i.e., $x(1) = x(2)$ and $x(M) = x(M-1)$.

### Initial condition for evolutionary dynamics

Gene interaction matrix $J$ is generated with the ratio that $J_{i,j} = 1 : J_{i,j} = 0 : J_{i,j} = -1$ $=1/N : 8/N : 1/N$.

The parameters, $\gamma_i$, $\theta_i$ and $D_i$ are randomly generated from [0:1] (See also table 2.1).

### Algorithm to Define an Epoch

Both in evolution and development, if the change in expression level of the output gene at a given site exceeds 0.9 within 50 time steps, it is regarded as an epoch.

### Definition of Fitness

To study the evolution of morphogenesis, we imposed a fitness condition to generate a given specific target pattern, for the expression of a given output gene. By setting a target pattern as $T(l)$, the fitness $f_i$ was defined as the sum of the distance between this target pattern and output gene expression at each cell, as:

$$f_i = \sum_l 1 - |T(l) - x_{output}(l)| \tag{2.3}$$

where $l$ is a cellular index. From the equation, the smaller the distance, the higher the fitness. Here the output gene pattern was defined after a given transient time, which was chosen to be large enough to reach a stationary pattern. For each genotype (i.e., GRN and a set of parameter values), the fitness was thus computed, after simulating each set of pattern dynamics.

For this analysis, we chose 100 individuals with different genotypes. Among these individuals, those who had higher fitness values were preferred to be selected for the next generation. Selection of the individual $i$ with a fitness value of $f_i$ was defined by

$$p_i = \frac{e^{f_i}}{\sum_k e^{f_k}} \tag{2.4}$$

The denominator summation of index k is aimed for all individuals in the population.

To generate the offspring, each genotype was slightly modified. A path in GRN was added, eliminated, or its sign was flipped with the probability $1/N^2$. Also, the parameter values $\gamma_i$, $D_i$, $\theta_i$ were modified by adding a random number from a Gaussian distribution $\eta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{x^2}{2\sigma^2})$, while restricting these values to the set [0,1]. We set $\sigma = 0.01$.

**extracting working network**

Although the GRN is uniform for each cell in an individual, working gene expression dynamics differ from cell to cell. Additionally, the expression level of most genes changes discretely between epochs, indicating that the regulation of the output gene is different between epochs. Thus, it is difficult to describe gene expression regulation in one cell for a given epoch, because the whole GRN has too many edges, which may work for other cells during other epochs. To identify which regulations are essential at a certain location and time, we extracted part of the GRN that works at a certain location during a certain period by adopting the following systematic method.

1. We fixed a cell to analyze and trace the dynamics of the input to the output gene in that cell. The developmental time was divided into epochs by checking the changes in the input to the target, which are responsible for the epoch. In practice, we defined the time span during which the input goes out of the range $[-2/\beta : 2/\beta]$, which is the dynamic range for the reaction term in the reaction diffusion equation (see equation (1) and (2)). Epochs are defined as the time span.

2. Fix the epoch to analyze and trace the dynamics of other genes during the epoch. If a gene is not expressed at all throughout the period, the regulation from and to the gene does not work. Hence, both input and output edges to such genes are eliminated.

3. Even if a gene is expressed for a certain time during the period, it does not necessarily mean that paths connecting to the gene are essential for the dynamics of the output gene' s expression. To check this point, we examine gene $j$'s expression as the input of another gene $i$ with $J_{ij} \neq 0$. We then, checked whether changes in gene $j$'s expression contributed to the expression of the gene $i$ as an input, while the latter's change stayed within the dynamic range $[-2/\beta : 2/\beta]$. If this contribution, defined below, was larger than a certain threshold value, then the edge $j \to i$ is assumed to work. Otherwise the edge was eliminated.

The specific procedure is as follows: Consider an edge $j \to i$. For a given cell, let $\Delta x_k(t)$ as the expression change of gene $k$.

$$if \ \ |\sum_k J_{ik}\Delta x_k(t)| \ > 0 \ \ and \ -2/\beta < \sum_k J_{ik}\Delta x_k(t) < 2/\beta,$$

$$check\ whether \ \ \frac{J_{ij}\Delta x_j(t)}{\sum_k J_{ik}\Delta x_k(t)} > threshold$$

where $\sum_k J_{ik}\Delta x_k(t)$ is the net input change of gene $i$. The threshold was 0.01. Delete all the edges $j \to i$ that do not satisfy the above condition.

4. Last, delete all the edges on genes that did not have a route from morphogens or to the output gene.

**overlap ratio**

The overlap between two networks $\boldsymbol{A} = \{a_{i,j}\}$ and $\boldsymbol{B} = \{b_{i,j}\}$ was computed as follows.

$$overlap = \frac{\sum \theta(\boldsymbol{A}, \boldsymbol{B}, i, j)}{min(\sum \phi(\boldsymbol{A}, i, j), \sum \phi(\boldsymbol{B}, i, j))} \tag{2.5}$$

where

$$\theta(\boldsymbol{A}, \boldsymbol{B}, i, j) = \begin{cases} 1 & \text{if } a_{i,j} = b_{i,j} \\ 0 & \text{else} \end{cases}, \quad \phi(\boldsymbol{A}, i, j) = \begin{cases} 0 & \text{if } a_{i,j} = 0 \\ 1 & \text{else} \end{cases}$$

The numerator is the number of common edges between the two networks, while the denominator is the size of the smaller network. The overlap is 1 when a smaller network is completely included in the larger network and is 0 when the two networks do not have common edges.

## Column: mechanism for regulation of morphogenesis by gene regulation network

Let us briefly review the relationship between the basic mechanisms of pattern formation and the structure of the network that controls them. As we defined reaction-diffusion dynamics in the model bt eq.(2.1) with (2.2), reaction term of the dynamics has a certain threshold value $\theta_i$ for each gene. The $\theta_i$, which roughly corresponds to dissociation constant in Hill function, determines the threshold for step-function-like reaction term, i.e, gene $i$ will be expressed when the input for the gene is larger than $\theta_i$, and otherwise it is not expressed.

**Feedforward network structure and regulated gene expression:**
Since the reaction term works in the way described above, static gradient of morphogene expressions can be interpreted to spatial stripe gene expression by GRN of genes with different $\theta_i$. Fig. 1C shows a schematic representation of regulation of spatial stripe gene expression by GRN of genes with different $\theta_i$. There is a gene network in the left end of the figure, which is composed of 5 genes and gene expression dynamics of each gene is depicted as 5 subfigures in the right side. Gene expression level is indicated by color bar in the right end.

**Figure 1C**

Since the gene network has no backward loop, we can easily understand the action of each gene. First, there is activation from morphogene to gene A. Then the expression of gene A will arise the region where morphogene gene expression is larger than $\theta_A$, as shown in the diagram A. Similarly, gene B also activated by morphogene but gene B is also inhibited by gene A. Thus the expressed region of gene B is subtraction from the region where morphogene gene expression is larger than $\theta_B$ by the region where gene A is expressed. If $\theta_A > \theta_B$ is satisfied, expression of gene B gives a stripe as in the diagram B. Stripe gene expression in the time-space is regulated in this way. In the same way, expression of gene C and gene D have two stripes.

This gene regulation mechanism with morphogene gradient and difference in $\theta_i$ is called *Feedforward*. The network structure has hierarchy from top(morphogene) to bottom and there is no way backward as we see in Fig.1C.

**Feedback network structure and regulated gene expression**

In contrast, if directed edges of gene interactions in gene network compose a loop, such gene network is termed as gene network with feedback. Gene network with feedback can generate oscillatory gene expression. Fig. 2C shows the typical network structure and oscillatory gene expression that is generated by the network. In the network structure, gene A activates B and itself, while B inhibits A. Thus at first gene A activated by itself. After gene A is sufficiently expressed, gene B comes to be expressed and the input to gene A gradually decreases because of the inhibition from gene B. Then gene A starts to decrease and following it, gene B also starts to decrease. Then it turns back to the first stage. The feedback structure can generate oscillatory gene expression in this way. In the following manuscript, we will see how the temporal oscillation generated by this network structure turns into spatial stripes.

**Figure 2C**

## 2.3 Results

We numerically evolved gene regulatory networks governing development in order to study evo-devo relationship. Here each organism consisted of M (= 96) cells aligned in one-dimensional space, where maternal factors were supplied from each end of the space. Each cell had N (= 16) genes (proteins) whose expression dynamics were governed by expression levels of other genes through a given gene regulatory network (GRN), while interaction between neighboring cells was mediated via diffusion of expressed proteins. These conditions defined the developmental dynamics of the study. We prepared 100 individuals with slightly modified GRNs (See Table 2.1 in the Method section for further information for parameters and initial settings for the simulation). After each gene expression level reached a stationary value through development, we computed fitness from the expression of a prescribed output gene.

Fitness was defined as the difference between this output expression pattern in space and a prescribed target pattern, with the highest fitness values defined by the best match. We used a genetic algorithm to select the individuals with higher fitness by introducing mutations in the GRN (See Figure 2.2 for schematic representation and Methods for details).

**Figure 2.2. Schematic representation of simulation procedure.**
(**A**): There are 100 individuals in a pool for each generation. Each individual consists of 96 uniform cells, which share a common GRN, while the GRN differs slightly between individuals. (**B**): Each individual develops from the same initial state in which genes are not expressed(i.e., with $x \sim 0$) except for genes receiving the maternal gradient. Over time, individuals develop into stable states. Colors of cells indicate the expression level of the output genes; yellow is high, gray is low. After reaching a stable state, the expression pattern of the output gene was compared with the predefined target pattern. The fitness level was then elevated as the stable expression of the output gene approached the target pattern (see Methods for detail). (**C**): After the fitness of every individual was calculated, the population for the next generation was created. Each individual was selected as a mother with a probability proportional to its fitness. In the figure, the degree of red color indicates the fitness. (**D**): The GRN of a daughter is slightly different from the mother's, with a given mutation rate. The mutation involves deletion or addition of paths in the mother's GRN, and a change in characteristic parameters in expression dynamics and the diffusion constant (see Methods for details).

Most evolved networks, after few thousand generations, were capable of generating predefined target patterns. An example of the developmental time course to shape such a pattern is given in Figure 2.3A, where the space-time diagram of the expression level of the output gene is displayed with the horizontal axis as the developmental time and the vertical axis as the cellular index (i.e., spatial axis). As shown, the target pattern (Figure 2.3C) is shaped after several developmental stages for stripe formations.(Unless otherwise mentioned, development after evolution is plotted for the fittest individual at 2000th generation.) Next, we examined how the output gene pattern had evolved, by tracing the final output pattern of the ancestors successively.

The output pattern after development of the ancestor at each generation is plotted

in Figure 2.3B, where the color code and spatial axis are identical to those in Figure 2.3A, while the horizontal axis represents the generation (evolution time) in Figure 2.3B. The similarity between the developmental (Figure 2.3A) and evolutionary (Figure 2.3B) spacetime diagrams are clearly discernible.

For reference, we have also plotted the developmental course at intermediate (1, 300, 750, 2000) generations in Figure 2.3D. With successive generations, novel stripes are acquired, moving the system towards the target pattern.

**Figure 2.3. An example of space-time diagram of evolution and development.** (**A**): The expression level of the output gene is shown with developmental time as the horizontal axis and cell index (spatial position) as the vertical axis. The expression level of the output gene of the corresponding cell at a given time is color coded, (side bar) with black indicating the lowest and yellow indicating the highest expression levels. Development consists of a few epochs with rapid changes in the pattern, separated by quasi-stationary regimes with little change in the pattern, until the target pattern is shaped by development.
(**B**): The spacetime diagram of the evolutionary course, corresponding to (**A**). The expression level of the final output gene (at time=2000) is shown with evolutionary generation as the horizontal axis and cell index (spatial position) as the vertical axis. This figure shows how the pattern is acquired through evolution. At each generation, the final pattern of the direct ancestor of the next generation is shown. The evolution of the developed output pattern consists of quasi-static regimes sandwiched by epochs with rapid change resulting from mutation, until the target pattern is evolved.
(**C**): The predefined target pattern adopted in the present simulation.
(**D**): Space-time diagram of the developmental process for several generations in (**B**). Each figure shows the development of the ancestral expression pattern at each generation, 0(D1), 300(D2), 750(D3), and 2000(D4). For reference, these generations are each marked by a red triangle at the top or bottom in (**B**).

## Development with epochs that correspond to those derived through evolution

Correspondence between developmental and evolutionary space-time diagrams was commonly observed in our simulations (Figure 2.3). Additional examples are provided in

**Figure 2.4. Four additional examples of evo-devo congruence.**
(**A**)-(**D**): Each row shows space-time diagrams of evolution and development, in the same way as Figure 2.3, although the target patterns are different. See supplemental information for additional examples.

Figure 2.4 and in Supplemental Figure S2.1.

It is remarkable that the pattern formation progressed in a stepwise manner, with respect to both evolution and development. Each stripe emerges not gradually, but discretely at some step in development and in evolution. More interestingly, the correspondence between evolutionary and developmental diagrams is supported by the correspondence of epochs in the two diagrams. This correspondence is valid for a large portion of our simulations. Furthermore, we generally observed good agreement between development and evolution modes, based on the topology of stripe formation: i.e. how later stripes branched from earlier stripes (see Figs.2.3, 2.4 and Supplemental Figure S2.1).

To quantitatively evaluate the correspondence between evolutionary and developmental space-time diagrams, we measured the overlap between the diagrams of the output expression levels. The procedure to compute the overlap is shown in Figure 2.5A. For

**Figure 2.5. Quantitative analysis of the degree of evo-devo congruence.**
(**A**): Schematic diagram illustrating quantitative analysis of the similarity between evolution and development. The differences between development and evolution were computed by subtracting expression levels at each pixel. By taking the absolute value of difference, and averaging the space-time pixels, the average difference was computed . To avoid over-estimating similarity, the region before the emergence of the first stripe and after the final pattern was ignored for both development and evolution. For example, the gray-masked region of the development and evolution figures does not include data for the calculation. If one stripe is completely shifted in time, is approximately 8%.
(**B**): Histogram of the distribution of the $\Delta$ values. The abscissa is the $\Delta$ value computed via the procedure described in (**A**). The ordinate is the frequency of touch $\Delta$ values determined by bin size 2 . Distribution was obtained from 500 runs with different target patterns.

both diagrams, we took only the temporal regime in which the pattern formation progressed, i.e., we discarded both the early stages where the output gene was not expressed in all cells (i.e. $x_i(t,l) \sim 0$ for all $i$), and the final stage after emergence of the stationary pattern, when no additional changes were observed. The distances between the output expression levels for both the diagrams were then averaged over all space-time pixels, $\Delta$, thus allowing us to compute the differences between the two diagrams. The distribution of $\Delta$ from approximately 500 evolution trials for different target patterns is shown in Figure 2.5B, with a peak distribution located at approximately 8%. Note that if the difference between the two diagrams is one stripe over all of the space-time pixels, $\Delta$ here is evaluated to be 8.2% (Fig. 2.5B). Hence, the peak value in the distribution is mostly just one stripe difference over all space-time. Thus for most examples, the space-time diagrams between developmental and evolutionary processes show remarkable similarity. These results suggest that the correspondence between evolution and development is not an accident, but is a general outcome for most evolution samples.

Furthermore, the distribution of $\Delta$ over time, as shown in Figure 2.6, shows that deviation is more frequent as earlier time scale. At later stages, the congruence is higher, which suggests that degree of changes decreases at later stage as dynamics depend on network structures pre-existing from earlier stage.

**Figure 2.6. Time distribution of $\Delta$, the difference between evolution and development.**
We computed the difference $\Delta$ in space-time pixel. Then we obtained the distribution of $\Delta$ over all simulation results, by using the histogram with the time bin of 1/10 of rescaled time between 0 (initial) and 1 (final). Difference in evolution and development is more frequent at earlier stages.

Note that if the difference between the two diagrams is one stripe over all of the spacetime pixels, $\Delta$ here is evaluated to be 8.2% (Figure 2.5B). Hence, the peak value in

the distribution is mostly just one stripe difference over all spacetime. Thus for most examples, the spacetime diagrams between developmental and evolutionary processes show remarkable similarity. These results suggest that the correspondence between evolution and development is not an accident, but is a general outcome for most evolution samples. We therefore began an investigation to determine why this evo-devo congruence holds so frequently.

### Mechanisms of evo-devo congruence

### i) emergence of slowly changing gene expressions

During each epoch, patterns change within a short span for both evolution and development, whereas the pattern remains quasi-stationary between epochs (Figs.3 and 4). Evo-devo congruence is a direct consequence of the congruence of these epochs between evolution and development. Thus, in order to clarify the mechanisms of evo-devo congruence, firstly we must study how such epochal pattern formation is possible for both evolution and development.

The epoch formation in evolutionary courses is trivial. Since the evolution undergo with strongly purifying selection, only neutral mutations or beneficial mutations remain during evolution. Generally, appearance of the relevant mutations to beneficial to pattern change require some generations. Furthermore, accumulation of neutral mutations is often needed for the beneficial mutation to work (indeed, for 80% of epochs, they are needed). Thus the evolutionary course of the developed pattern consists of a quasi-stationary regime and requires several epochs to change the stripe pattern. This epochal pattern change in evolution has previously been coined as "punctuated equilibrium" (Eldredge and Gould, 1972).

Correspondingly, we have observed a long quasi-stationary period in development between epochs of rapidly changing stripe patterns; however, this is not self-evident. In this section, we focus on how epochal pattern formation is possible in development. Indeed, we will show that slow gene expression of certain genes is essential to generate a developmental epoch and unveil the origin and function of such slow changes in gene expression.

In Figure 2.7, we show the change in gene expression at the site marked in Figure 2.7A, which presents the space-time diagram of the output gene expression. Figure 2.7B shows the expression dynamics in the marked cells in Figure 2.7A. Each colored line indicates the expression dynamics of a given gene. As shown, the expression of the output gene (red line) switched between on and off at the time denoted by the crossed red lines in Figure 2.7A. This switch corresponds to an epoch. As seen in Figure 2.7B, expression of a subset of genes changes over time, before and after an epoch. Even though the expression

**Figure 2.7. An example of developmental dynamics with several epochs for a single cell.**
(**A**): Developmental space-time diagram of the output expression level. The developmental time up to 1200 in Figure 2.3 is zoomed, in order to clearly distinguish epochs. There are four epochs associated with one cell, marked by crossing of the red lines. (**B**): Gene expression dynamics at the cell highlighted by red horizontal line in (**A**). The time series of the expression level is plotted where the line color corresponds to that of network node in (**C**), which representing each gene. Most genes change their expression level within each epoch, except for the gene S represented by the blue line. (**C**): The working networks that function to switch the output expression at each epoch. These networks are derived from analyses on gene expression dynamics(see method). Arrow edges indicate positive regulation while the headed edges indicate negative regulation.

change in most genes, including the output gene, was rather fast to support the epoch, there exists a subset of genes whose expression level changes slowly over time (blue line).

To understand the role of genes with slowly varying expression, a core part of the GRN, which is responsible for stripe formation at each epoch, was extracted (see methods). The core network at each epoch is termed the "working network", as shown in Figure 2.7C.

From the working network, it is clear that slowly changing gene expression serves as a control variable for the switch in output gene expression dynamics. The input to the output gene was beyond or below its threshold level at the times marked by crossed red lines in Figure 2.7A. This change was driven by the on/off switch of gene A, seen in Figure 2.7C, while the switching in gene A was mainly driven by the slow change in input, denoted by gene S in Figure 2.7C. The slow change in the expression of gene S served as a parameter cue, providing the timing of the epoch.

The morphogen concentration acting on cells was fixed, as indicated by the black horizontal line in Figure 2.7B. Morphogens initially activated the output gene that provides the first epoch. However, they simultaneously activated gene S, whose expression level increased at a slower rate than that of the output gene. When it exceeded the threshold for the expression level of gene A, the suppression of the output gene was dominant, leading to decreased output gene expression and thereby generating the second epoch. Indeed, the third and fourth epochs were controlled by a slowly varying input in the same manner (data not shown).

We examined several other examples, and found that the working network after evolution always includes a gene with slowly varying expression at the corresponding epoch. The slow gene expression did not give a direct input for the output gene, but gave an input to a gene that gives an input to the target (i.e. gene A and gene B, Figure 2.7). We then carried out statistical analyses to confirm that control via slow expression changes is a general outcome of evolution.

In our model, the input term for each expression has a dynamic range given by the threshold ($\theta_i$)and the slope ($\beta$) of the input-output relationship(see method). If the input for a gene is out of the dynamic range (i.e., $-2/\beta < Input - \theta < 2/\beta$), the expression of the gene is either 0 or 1. Thus, the time span required for the input to pass through the dynamic range provides an indicator of the time-scale for the control of the input. We computed the timescale of the input to both the output gene and the other genes.

In Figure 2.8, the evolutional change of these two time scales are plotted (the latter time scale is the slowest change among all genes that have a path to the output gene). As shown, the time scales are nearly equal at the first generation, indicating the absence

**Figure 2.8. Evolution of the time scale for the input for the output and other genes**

By taking genes whose expression level change between on and off for each developmental epoch, the time scales are computed as the time span that the input for the gene passes through the dynamic range during each development. The red square gives the time span for the input of the output gene, and the blue triangle(green circle) denotes the average (the largest) of the time span among the genes that have a path to the output gene, respectively. The time spans are computed from the average of 500 samples of evolution simulations.

of slow expression control. Throughout evolution, the timescale of the output gene was not altered, which also supported the notion of epochs with a short time span of change. In contrast, the time scales of other genes increased, such that the ratio of the target timescale to that of others decreased, reaching 1/5 of the average. Thus the relative timescale of input gene expression to the change in target expression was slower during evolution. Hence the results of Figure 2.8 support the emergence of epochs by slow genetic control.

**The origin of slowness in expression**

Questions remained regarding the origins of such slow expression dynamics. Following analysis of all examples, we concluded that the origin of slowness could be attributed to the following two mechanisms:

(i) The existence of genes with small rate constants $\gamma_i$ associated with expression change. The expression dynamics in our model includes a parameter $1/\gamma_i$, representing the time constant for change. Hence if some gene $i$ has a small $\gamma_i$ value, expression changes slowly. While this itself may appear rather trivial, it should be noted that the rate parameters $\gamma_i$'s after evolution are distributed by gene $i$, and some genes have smaller $\gamma_i$ values. Therefore, the expression levels of genes with small $\gamma_i$ values function as a slow variable. Indeed, in the example presented in Figure 2.7, $\gamma_i$ for the gene $S$ is 0.063; a full order of magnitude smaller than the others. Through evolution, genes with distinctively small $\gamma_i$ values appear, even though we initially established nearly uniform $\gamma_i$ values for all genes.

(ii) Expression levels near the threshold. The expression dynamics here have a threshold $\theta_i$. If the input to the gene is larger (or smaller) than $\theta_i$, it is expressed (or suppressed), respectively. When the input term from other genes to the gene $i$ is close to $\theta_i$, then, the expression level can be balanced at an intermediate value between 0 and 1. Indeed, if the deviation of input from $\theta_i$ is smaller than $1/\beta$, the inverse of sensitivity, then the expression level of $x_i(l,t)$ is no longer attracted to 0 or 1. In this case, this stationary state is less stable than those closer to 0 or 1 (see Supporting text S1.1 for the mathematical explanation using the Jacobian matrix). Hence, the time-scale around this fixed point is longer in duration.

This slow relaxation to the stationary state as a single-cell dynamics is extended through the entire space, mediated by the diffusion interactions with other cells. With diffusion, the slow expression change of a certain cell can propagate spatially to other sites, to change their expression levels slowly.

### ii) mechanisms for pattern formation and their dependency

Now, we show how stripes(valleys) are formed in developmental process here, based on gene regulation dynamics, cell-to-cell diffusion, and morphogen gradient. Through extensive analysis of 500 samples of the evolved pattern-formation, we confirmed that the stripe formation process is reduced to only two basic mechanisms in gene expression dynamics with corresponding GRN structures. In fact, these two mechanisms have previously been identified and studied extensively, which are known as feedforward and feedback regulations(Salazar-Ciudad et al., 2001a,b, Alon, 2007, François et al., 2007, Fujimoto et al., 2008, Cotterell and Sharpe, 2010, ten Tusscher and Hogeweg, 2011).

#### Feedforward

The classic mechanism for stripe formation, which was analyzed in the segmentation process in Drosophila, is feedforward regulation. This mechanism has been analyzed both theoretically and experimentally(von Dassow et al., 2000, Jaeger et al., 2004, Ishihara et al., 2005). In this case, a gene 'reads' the morphogen gradient for spatial information, to establish an 'on/off' response under a given threshold level, so that the gene is expressed on the one side of space, and non-expressed on the other side. Another 'downstream' gene receives positive (or negative) input from this gene, and negative (or positive) input from the morphogen, then responds to create another segmentation in space, if the threshold parameters satisfy a suitable condition. By combining this feedforward regulation, more stripes are formed for the downstream gene. The corresponding GRN does not require feedback regulation, or cell-to-cell interaction by diffusion; only unidirectional, feedforward regulation from morphogen input to downstream genes is required. This feedforward regulation frequently exists in our evolved GRN, and is used to generate at least some stripes.

#### Feedback oscillation within a boundary

The other mechanism for stripe formation, commonly observed in the present simulations, takes advantage of feedback regulation to produce a temporal oscillation in the expression level. This temporal oscillation at a single-cell level is then fixed into a spatial periodic pattern by the diffusion among cells. A typical core network structure and expression dynamics are shown in Figure 2.9A. Here, gene A activates the expression of both itself and gene B, while gene B suppresses the expression of gene A. Since this network is just a typical negative feedback loop, it produces a temporally oscillating expression when the parameter values for the expression dynamics are appropriate. Now, with the diffusion of B under an appropriate boundary condition, this temporal oscillation is fixed into a spatially periodic pattern (see Figure 2.9B).

**Figure 2.9. Feedback oscillation and its fixation.**
(**A**)**Without boundary:** Minimal network for oscillatory expression with the time series of the expression for a specific cell. Gene A activates the expression of gene B and itself, and gene B suppresses A. In the plotted time series, developmental time is plotted as the abscissa, and the expression levels of A (red) and B (blue) are plotted as the ordinate. (**B**)**With a boundary:** The input from gene M, which was influenced by the maternal factor, was included in the oscillatory network. The space-time diagram of genes M, A, and B illustrate how oscillatory expressions of gene A and gene B were fixed to form stripes. Gene M was expressed near the boundary.

Consider a case where the input from the morphogen M suppresses the activation of A at the boundary. Again, without input from M at the boundary, there appears temporal oscillation in the expression of A and B. At this boundary, the expression of gene B is also suppressed. Then the protein B at the adjacent upper site diffuses to this boundary. Subsequently, the expression of B at the site is decreased, so that the suppression of A is relaxed(Figure 2.9B, bottom). Then the protein of gene A is fixed to a higher level, instead of oscillating. This also leads to an increase in the expression of B. At the next upper sites, oscillation still remains.

When the expression of A is low, the diffusion of protein B from the lower site suppresses the activation of A, so that the increase of the expression of A no longer occurs. Thus, the expression level of A is fixed at a lower level. With this process, temporal

oscillation of one period is mapped into one spatial stripe. The same fixation process is repeated with the subsequent oscillation at further upper sites, since at the nearest lower site, the expression of gene A is fixed to a lower level. Thus, the temporal oscillation is recursively fixed to a spatially periodic pattern. With this mechanism, the stripe pattern in space is formed and fixed. This mechanism will be discussed in **chapter 3**, but we give a brief remark below.

This mechanism is analogous to the classic Turing pattern in which case the suppression of B to itself is necessary to exclude a spatially homogeneous, temporally oscillating state. Here, the diffusion of the inhibitor gene B works in the same way as the Turing pattern(Turing, 1952), but the mechanism here adopts temporally oscillatory dynamics, and is understood as Turing-Hopf bifurcation(De Wit et al., 1996). Meinhard' s model to transform oscillation into segmentation, in which the fixation is driven by a morphognen gradient (Meinhardt, 1982), in contrast to our mechanism driven by the fixed boundary.In an interacting 2-cell system, this differentiation from the oscillatory state is understood as, a saddle node bifurcation on invariant cycle (SNIC). SNIC in a globally coupled dynamical system has been studied as a mechanism for cell differentiation from a stem cell(Goto and Kaneko, 2013).

We note two points. The mechanism here uses the suppression at a boundary to fix a pattern, instead of the inhibition of B to itself, and thus the boundary condition is important. Next, the mechanism also resembles a classic wavefront mechanism, but is different. In the wavefront mechanism, the temporal oscillation is fixed into a spatial periodic pattern through input from the morphogen gradient and the growth of the system size. In the mechanism, however, diffusion (or cell-to-cell interaction) is not essential, and external manipulation by the morphogen gradient for all cells leads to fixation forming a stripe from the oscillation. In our case, such external manipulation exists only at the boundary, and further stripe formation progresses spontaneously by the diffusive cell-to-cell interaction. Indeed, this is a new class of pattern formation that induced by boundary condition.

#### Stripe formation order by the combination of the two mechanisms

All of the potential evolved stripe formation processes in our model could be generated by a combination out of four possible ways of combining these two developmental mechanisms, sequentially. However, for the feedback mechanism to work, the boundary depending on the morphogen has to be established in advance, to fix the temporal oscillations to spatial stripes. Thus, the feedforward mechanism to read the external morphogen is needed to produce the boundary. Otherwise, no stripe will be formed, so that such networks will not remain in the evolutionary simulation. Hence, we consider only two combinations: feedforward-feedforward and feedforward-feedback. Indeed, these two cases are the bases for all of the possible developmental processes evolved in our simulation.

**Sequential Feedforward mechanisms**

Stripe formations involving the combination of feedforward mechanisms have been extensively studied. In some examples, the developmental processes evolved here are achieved by sequentially combining feedforward processes, where cell-to cell interaction is not needed. Consider a new feedforward mechanism, added at some point downstream from an upstream feedforward circuit. As long as the upstream mechanism is not affected by the downstream mechanism (which is true if there is no feedback from the latter to the former), the stripe formation progresses first by the upstream mechanism, and then, at a later epoch, the stripe is generated by the downstream mechanism.

This ordering in the developmental process agrees with the order of evolution, since the downstream mechanism is acquired later in the evolutionary course. Hence, in this simple sequential feedforward mechanism, the evo-devo correspondence is a natural outcome. Since the evolved GRN typically has slower gene expression dynamics that control the downstream expression as already shown, the stripe formation will occur sequentially in developmental time, with some delay, in agreement with evolutionary time course.

The evolved network illustrated in Figure 2.10A consists of a combination of sequential feedforward networks and a downstream feedback network. Through evolution, first the feedforward network via gene 10 and gene 3 (see Figure 2.10D) is acquired at around the 10th generation. Then, a domain in the middle space is shaped in development as shown in Figs. 9B and 9E. Later, at the 88th generation, another feedforward network via gene 12 is attached downstream through evolution (Figure 2.10F). With this attached network, a domain is shaped in the interior of the earlier domain as seen in Figs.9F and 9G, right after the earlier domain formation is shaped. The shaping of domains is successfully completed at an early stage of development. This leads to the evo-devo congruence. Later, this modified domain in the middle works as a boundary condition for the subsequent feedback network to be discussed.

**Figure 2.10. Example of evo-devo congruence with network structures**
(**A**): An example of a core part of the GRN at the 2000th generation, evolved to achieve the target pattern. From the maternal factors, the feedforward networks is surrounded by magenta, while the network module for feedback oscillation, consisting of genes 12 and 15, is surrounded by blue. Here, genes and paths that are not essential to the output pattern formation were eliminated. (**B,C**):Space-time diagrams of the output gene expression for development (**B**) and evolution (**C**) or the GRN are displayed together to show the degree of similarity between them. The vertical axis denotes the space (cell index), and the horizontal axis denotes either evolutionary generation (evolution) or developmental time (development). For this example, the $\Delta$ value is 8.0%. (**D**): An example of the core part of the GRN at the 10th generation (i.e., very early generation) and the corresponding space-time expression diagrams of the output. (**E**). Feedforward structure of the GRN is evolved at this early stage of evolution. The vertical axis of the phase diagram denotes the space (cell index), and the horizontal axis denotes developmental time. This expression is observed at a very early stage of development in (**B**), at approximately the 10th generation. (**F**): The network structure at the 88th generation. Through evolution, feedforward structures are sequentially acquired in the downstream region of the core part of the GRN. (**G**): Developmental space-time diagram of the expression of the output gene for the network **F**. This expression profile provides the top and bottom stripes in (**B**). (**H**): Developmental space-time diagram of the output expression of the 2000th generation where the feedback oscillation module is eliminated. Without feedback, only part of (**B**) is generated.(caption continues to the next page)

(**I**):Developmental space-time diagram of the expression of gene 12, one of the feedback modules in (**A**), which produces a spatially homogeneous and temporally periodic oscillation if constant activation is applied by gene 3. The combination of this feedback oscillation and the boundary condition provided by gene 12 shown here produces the three internal stripes in (**B**).

### Feedback-oscillation mechanism attached downstream of the feedforward network

The Upstream feedforward network is indeed necessary for the feedback mechanism to work as already explained. In development, the stripe formation by the feedback mechanism cannot work without a boundary, and only after the appropriate boundary condition is generated by the feedforward mechanism. On the other hand, the feedforward circuit is first acquired in the earlier stage of evolution to increase fitness, and later the feedback-oscillation is obtained to create further stripes using the former feedforward stripe as a boundary. Hence, we again observed good agreement in the time courses of stripe formation development and evolution, as long as the latter feedback mechanism does not influence the former feedforward mechanism. Slower expression change of controlling genes as already discussed works for the separation of two epochs.

An example of evo-devo congruence caused by feedback-oscillation downstream of the feedforward mechanism is displayed in Figure 2.10A. Corresponding space-time diagrams of evolution and development are presented in Figs.9B and 9C. Evo-devo congruence is detected, in particular between the third and fourth upper stripes. These two stripes are generated by the oscillation-fixation mechanism generated by the feedback loop (Figure 2.10A), attached downstream of gene 3, which is a component of the feedforward network from a maternal morphogen. This feedback module is inhibited by two morphogens and gene 5, so that this oscillation does not start without an input for activation. The only activation input for this feedback module is gene 3, which is expressed only in a domain restricted by the upstream feedforward network. Thus, the oscillation starts after the expression level of the gene 3 is sufficiently high (Figure 2.10I), and thus is bounded within the domain, maintaining the expression of gene 3 (Figure 2.10H). Following the mechanism discussed in the next section, a stripe is generated in this domain. This feedback oscillation is regulated by the upstream feedforward network but does not disturb upstream feedforward expression.

## iii) Parallelism between the working GRNs of evolution and development

The results in the last section suggest that the ordering of working networks over epochs is in agreement with development and evolution, and that both progress from feedforward-based networks to networks including feedback loops in addition to feedforward networks.

We examined the validity of this ordering statistically by using all the data in our simulated evolutions.

We first examined whether a working network includes a feedback loop, and computed the fraction of purely feedforward networks that do not include a feedback loop at each epoch. As shown in Figure 2.11A, this feedforward ratio is close to 1 at the first epoch ($\sim$ 0.85), and it decreases in later evolutionary epochs. Nevertheless, as the working network size for the $i$-th epoch $k_i$ increases, the probability of a network without feedback loops is expected to decrease. This probabilistic decrease is estimated as $f^{k_i/k_1}$, which is also shown in Figure 2.11A. The observed rate decrease without feedback loops is much higher than this estimate. Hence, the fraction of feedforward networks is significantly higher during the first epoch, while the fraction with feedback loops increases faster than was estimated by using the increase in network size. In comparison, we also plotted the fraction of pure feed-forward networks from random networks of corresponding size (Network size were computed after randomly generated networks were processed such that genes that was not included in a path from the morphogen to the target gene were removed). The fraction in the evolved network was much larger. Thus, the feed-forward network was preferentially selected.

**Figure 2.11. Feedforward ratio**
(**a**):The red circle denotes the fraction of working networks that do not include a feedback loop, plotted as a function of evolutional epoch. The fraction is computed from 500 samples of evolution simulations. For reference, the probability estimated from the value at the first epoch only as a result of the increase in the network size is plotted as the blue square, while the triangle denoted such probability computed from $2 \times 10^6$ random networks of the corresponding size, generated only under the constraint that there are paths from the morphogens, and to the target.The average size of the network computed from the simulation is also plotted as cross with the second vertical axis. (**b**):Ratio of the feedforward network as a function of each developmental epoch, computed in the same manner as (**a**) and the same use of symbols.

We also checked whether the ancestral network is conserved by computing the fraction of networks preserved in later epochs (see methods). This overlap ratio is shown in Figure

2.12A as a function of evolutionary epochs. Values were higher than 0.75 over the epochs, such that over 75% of the ancestral working networks were preserved during evolution. In summary, the results shown in Figure 2.11A and 2.12A indicate that working networks mostly consisted of feedforward networks, and were well conserved throughout evolution. Additionally, later in evolution, feedback loops were added.

The fraction of pure feed-forward networks in the working network is plotted against the developmental epochs in Figure 2.11B. The network size was large at the first epoch, such that the ratio at the first epoch was small. Still, the decay of the fraction is much larger than expected by the probability calculations due to the increase in network size.

Finally, we checked the overlap ratio of the working networks between evolution and development. As shown in Figure 2.12B, the overlap remains high throughout the epochs, indicating that the working networks in evolution and development correspond with each other. Thus, the same pattern formation dynamics are adopted in the same order between evolution and development.

**Figure 2.12. Network overlap**
(**A**)The overlap of the working network at each epoch with its ancestral network. The overlap of the 1st(○), the 2nd(△) and the 3rd(□) epochs with the offspring network at later epochs are plotted. See method for the definition of the overlap, which is computed as a statistical average from 500 samples of evolution. Dashed line is the average overlap between randomly generated network and the same network that underwent 2000 sttif of random mutations. (**B**)The overlap of the working networks at the 1st(○), the 2nd(△), the 3rd(□) and the 4th(×) *developmental* epoch with the ancestral networks at each *evolutional* epoch represented by the abscissa. Dashed line is the average overlap between randomly generated network and the same network that unnderwent 2000 sttif of random mutations.

### iv) A slowly varying expression level works as a bifurcation parameter to produce a developmental epoch

So far we have uncovered the existence of slow expression change working as a control for the output gene expression and the combination of feedforward and feedback networks. These are important for evo-devo correspondence in pattern formation and gene-expression ordering, but we need to understand how these two features lead to fast switch-like change in target patterns at epochs, and how these are correlated in development and evolution. Here we describe this congruence of fast switching dynamics in evolution and development, in terms of bifurcation in dynamical-systems theory.

Consider the example of the network presented in Figure 2.7, with a slowly changing expression of Gene S. When the expression level of gene S (slow variable) increases slowly and reaches a certain level, the expression level of gene A increases from $\sim 0$ to $\sim 1$. Input changes to gene B may then lead to bifurcation. Here the morphogen (gene M) activates gene S and gene B, while gene S activates gene A, and gene A subsequently inhibits gene B. If the expression level of gene S is smaller than the total activation input to gene B, the dynamics of expressions of gene A and B are given by the flow presented in Figure 2.13 (upper left). The nullcline of gene B forms z-like structure in this phase space, which crosses the perpendicular nullcline of gene A, at coordinates near (0,1). As the expression level of gene S increases, the nullcline for the expression of gene A moves horizontally, so that the fixed point at $(x_A, x_B) \sim (1, 0)$ disappears and is replaced by the fixed point at $(x_A, x_B) \sim (0, 1)$, as seen in Figure 2.13 (bottom left). Thus, the bifurcation between fixed point attractors occurs with a change in the expression level of gene S as the bifurcation parameter.

#### Bifurcation behind Evo-Devo correspondence

During this study, we observed that the developmental process consisted of a quasi-stationary regime and epochs to form new stripes, due to bifurcations resulting from slowly changing expressions as parameters. This is relevant to achieving evo-devo congruence, since the evolutionary process also consists of a quasi-stationary regime prior to the emergence of a relevant mutation capable of increasing fitness within a relatively short time span. Indeed, such mutations change the gene expression dynamics drastically to form a new stripe, which again is regarded as a bifurcation. At a certain generation in the evolution, a mutation occurs to add an inhibition path from gene S to gene A (Figure 2.13). This mutation occurs in a discrete manner: Whether a path exists or not, it is not represented as a continuous change in a parameter value. However, we can introduce a continuous strength parameter that changes from 0 to $\pm 1$, which can be regarded as a bifurcation parameter. Then with this continuous change, an on/off discrete change appears at a certain value of path strength that depends on the threshold of gene A. Dynamics of the expression of gene A and B are represented in the two-dimensional state space in Fig-

**Figure 2.13. Evolution and development as bifurcation.**
The network structure where the expression level of gene S changes slowly (left). Phase space diagrams plotting the expression levels of gene A (horizontal access) and gene B (vertical access). The blue line represents the nullcline of gene A, and the red line represents the nullcline of gene B (right). The green circle denotes the final stable cell state from the initial conditions in each of the diagrams.
**Development(left column):** Expression level of the slow variable works as a bifurcation parameter. While the expression of gene S is lower than the threshold of gene A, the stable fixed point can be found at approximately ($x_A \sim 0$,$x_B \sim 1$)(upper left). As development progresses, the expression level of gene S increases, and after the expression of gene 1 exceeds the threshold of gene A, the nullcline of gene A shifts slightly to the right, indicating a higher value (lower left). Gene A inhibits the expression of gene B, so that the fixed point is changed to ($x_A \sim$1,$x_B \sim 0$).
**Evolution(right column):** Phase diagram representing the expression levels of gene A (horizontal axis) and gene B (vertical axis). The blue line represents the nullcline of gene A and the red line represents the nullcline of gene B. The activation strength from gene A to gene B is regarded as a continuous value here. If the activation strength is low, the expression level of gene 1 is low, so that the stable fixed point is observed at approximately ($x_A \sim 0, x_B \sim 1$)(upper right). When the strength is sufficiently large, the expression level of gene 1 assumes a higher value so that the fixed point is observed at approximately ($x_A \sim 1, x_B \sim 1$)(lower right). Note that by comparing these two columns, a strong correspondence is observed in bifurcation between evolution and development.

ure 2.13(right column). At a lower strength in the path, the nullcline of gene B expression changes so that the former stable fixed$(1, 0)$ point exhibits a saddle-node bifurcation, to move to another fixed point (0,1). Hence, the mutational change in the network leads to a bifurcation. As seen in Figure 2.13, this bifurcation through the evolutionary process agrees with that observed during development.

After examining hundreds of numerical evolution simulations, the results were summarized as follows:
*Development*:
slow change during expression works as a bifurcation parameter, and bifurcation in the expression dynamics generates a novel state, which gives rise to an epoch.
*Evolution*:
search for mutation resulting in relevant change to a new state. Epoch in evolution is also generated by the same bifurcation. In this way, evo-devo correspondence is achieved through bifurcation.

### Violation of Evo-Devo correspondence

Although evo-devo correspondence was frequently observed and was discussed as a natural outcome of the combination of network motifs for development, small, but non-negligible portions of the simulation runs exhibited deviation from this evo-devo correspondence. Examples of such exceptional cases are presented in Figure 2.14A, 2.15A, 2.16A.

In the example of 2.14A, the developmental and evolutionary diagrams differ distinctly, not only in the timing of the formation of the second and third upper stripes, but also in the topology in their branching. During the course of evolution, there is a drastic change in the final pattern, at approximately 1272-1273 generations. Here, only a single mutation occurred in a GRN (addition of a single path). In this example, the feedback oscillation of gene 5 was responsible for the output gene expression, in particular for the second and third stripes, while the expression of gene 6, which lies upstream of gene 5, acted as a boundary for the feedback oscillation, which also contributed to the expression of the output gene.

In Figure 2.14B, the gene expression dynamics of the selected genes 5 and 6, as well as the output gene, are displayed for generations before and after this mutation, in the left and right rows, respectively. Here, the mutation occurred upstream of gene 6, and reduced the range in which the gene was expressed, accordingly. The expression around sites 60-70 was subsequently suppressed, allowing for the formation of an additional stripe near site 70, while at lower sites (around site 60) the expression level continued to oscillate, forming a stripe much later. Hence the temporal ordering of the formation of the second (near site 80) stripe, and that of the third (near site 70) stripe was reversed by this mutation. Indeed, before the mutation, the third and fourth stripes were generated together (while the second stripe did not exist), and after this mutation, the second and fourth stripes

were generated together, and the third stripe was shaped later. Thus, the ordering and topological branchings of stripes were altered by the mutation, which led to a violation of the evo-devo correspondence.

To summarize, the violation of the correspondence was due to an upstream expression change resulting from mutation, which caused a change in the boundary condition for the feedback oscillation of the downstream expression gene. We have studied several other examples that showed a violation of evo-devo correspondence, and confirmed that differences in the topology in stripe branchings between development and evolution is caused by mutation upstream of the feedforward mechanism acting as a boundary of the feedback mechanism as shown in Figure 2.15 and 2.16.

## 2.4 Discussion

### Summary:

The potential relationship between phenotypic dynamics to shape phenotypes and evolution in genotypes has been the focus of the evo-devo field, since the time when genetic assimilation was first proposed by Waddington (Waddington, 1957). The relationship has been investigated in RNA evolution (Ancel et al., 2000) and gene expression dynamics (Ciliberti et al., 2007, Kaneko, 2007a), as also summarized in recent reviews (Wagner, 2005, Kaneko, 2006, Soyer, 2012). Still, studies to establish relationships between multicellular pattern formation dynamics and evolutionary processes that shape the pattern remain premature both in theory and experimentally.

Here, we carried out extensive simulations to evolve gene regulatory networks subject to strongly purifying selection, in order to generate a predefined target pattern for the expression of a given output gene. The main results of the present chapter are summarized as follows:

#### 1. Epochs of development as bifurcation by slow expression dynamics:

The developmental course of the expressions of the output gene, after evolution, consisted of a few epochs characterized by rapid temporal change in gene expression and a quasi-stationary regime with slow temporal change between the epochs. The slow quasi-stationary regime is due to expression levels of some genes that vary slowly over time, while the rapid drastic change is due to a bifurcation in the expression dynamics. The slowly varying expressions always emerge as a result of evolution, and they work as *bifurcation parameters* to control the fast change in the expression of the output gene.

**Figure 2.14. Violation of evo-devo congruence 1.**
(**A**): **Evolution:** The expression level of the final output gene (at time=2000) is shown with the generation (horizontal axis) and cell index (vertical axis). The color scale is presented as a side bar, as in Figure 3B. According to the figure, the second upper stripe is acquired at the most recent stage of evolution, and the first, third and fourth upper stripes branch from the same root, so that the second stripe emerges from the first upper valley. **Development:** Space-time diagram of the expression with developmental time (horizontal axis) and cell index (vertical axis). The third upper stripe emerges at the most recent stage of development. Unlike evolution, the first, second and fourth upper stripes branch from the same root, and the third stripe emerges from the second upper valley. Here, evo-devo congruence is topologically violated. (**B**): Developmental diagrams plotted for generations 1272 and 1273. These genes show drastic change in their expression between the two generations. Gene 6 provides a feedforward regulation to the output, and inhibits the expression of gene 5, which is a component of the feedback loop to generate oscillatory expression. A mutation, which adds a path to gene 6, occurs between the two generations, which inhibits the expression of gene 6. Through this mutation, the expression of gene 6 is suppressed, thus shrinking the resulting stripe, and producing an additional stripe for gene 5. With this change, the ordering of the expression of the output gene is altered.

**Figure 2.15. Network analysis of the extra example of the violation of evo-devo congruence 2.**
In development, the third and fourth upper stripes stem from the some root, while in evolution the top three stems originate from the same root. This branching change occurs during generation 1549 where the second and third branches are clearly stabilized. In this case, unlike the former examples, topological changes in branching occur sequentially during three generations. A time-space diagram of the output and gene 7 are presented in Figure B. Gene 7 exhibits two stripes from generation 1549, which are driven by a feedforward mechanism. Due to the boundary effect of gene 7, the upper three stripes are generated in the output gene. Then, during generation 1550, part of the feedforward mechanism upstream of gene 7 is deactivated, which enhances the region expressed by the feedback oscillation mechanism. As a result, the upper four stripes that emerge share the same oscillation mechanism. At generation 1551, mutation occurs upstream of gene 7, so that the morphogen comes to inhibit the remaining feedforward mechanism. Before the mutation, gene 7 exhibits weak temporal expression in cell sites 30-85. After the mutation, this temporal expression is inhibited so that the expression region is restricted to cell sites 60-85. Due to this change, the third and the fourth upper stripes emerge faster than the first and second stripes, while the third stripe, generated in advance of the first two, provided a boundary to generate the second stripe. To summarize, the upper 4 stripes were generated by the feedback oscillation mechanism, but the change in the boundary condition due to mutation in the upstream feedforward mechanism introduced the branching combination.

**Figure 2.16. Network analysis of an additional example of the violation of evo-devo congruence 3.**
In evolution, the three central stripes are acquired nearly simultaneously, and two additional stripes are subsequently acquired independently. However, in development, at the final evolved generation, the 1st and 4th stripes were generated from the same root at the same time, and subsequently the 2nd and 3rd stripes were generated from a common root. These two branchings follow the oscillation-fixation mechanism. Only the bottom stripe is generated independently. The developmental order of stripe formation was acquired, between generations 1104 and 1105. Genes that exhibited relevant change are displayed in Figure B. The expression of an upstream gene (green in the GRN figure below) and the downstream gene (red in the GRN figure) are plotted at the upper and lower columns, respectively. In this example, the feedforward mechanism worked only temporally, as shown in the transient expressed before time step = 100 (Figure C). This temporal expression region also corresponded to the region of feedback oscillation. Due to the mutation, the spatial domain of the transient expression was extended upward. Violation of evo-devo congruence was therefore induced by this expansion of the transient expression region.

If we mutate the timescale of the slow gene $\gamma_i$ to a smaller value, development can be accelerated, but with a further change, the original pattern is collapsed by bifurcation in the dynamics leading to destruction or branching of stripes. In fact, such bifurcation could also occur by changing the time scale of a gene with faster time scale, but it is less frequent and needs much larger change in the time scale. Thus, the pattern is more sensitive to change in slower genes. Hence, acquiring slower genes accelerates the evolution to novel pattern formation. Indeed, it is reported that the existence of a gene with slower expression, termed as "timer gene", controls the evolution of somitogenesis, and discussed its relevance to short-germ/long-germ transition(François and Siggia, 2010).

**2. Punctuated equilibrium in the evolution of morphology as bifurcation:**

Likewise, the evolutionary course of expression dynamics consists of a few epochs with a drastic change, and a quasi-stationary regime between the epochs. The drastic change is again represented by a bifurcation, which is caused by mutations in the gene regulation network.

**3. Evo-devo congruence through common bifurcations:**

In most cases, we observed good correspondence between development and evolution, with regard to spatiotemporal dynamics from a uniform state to a target pattern. We observed good agreement between development and evolution when evaluating epoch changes from one pattern to another, as well as the ordering of epochs. Indeed, the same bifurcations occurred for both, and thus the evo-devo congruence was due to the common bifurcation at each epoch.

**4. The combination of feedforward and feedback gene regulation networks to support developmental epochs:**

The combination of feedforward and feedback modules in gene regulation networks provides successive bifurcations at epochs. The upstream feedforward network converted the external gradient of the maternal factor into an output pattern, while the feedback loop converted the temporal oscillation of gene expression into a spatial stripe, under a given boundary condition provided by the feedforward expression dynamics. The evo-devo correspondence was preserved as long as the upstream feedforward network was maintained.

**5. Violation of evo-devo correspondence through modification of upstream feedforward regulation under downstream feedback mechanism:**

In rare examples of our simulated runs, we observed violations of evo-devo congruence. These violations were always associated with a structure of the upstream feedforward network and a downstream feedback loop, in which modification of the upstream feedforward network changed the boundary condition of the downstream feedback.

This then raised questions as to why the sequential feedforward network was excluded therein, and whether the violation always involved the feedforward-feedback combination. The feedforward mechanism reads the morphogen gradient of a maternal factor, so that the feedforward-feedforward process transfers spatial information of the maternal gradient sequentially, from upstream to downstream. The flow of spatial information was unidirectional, so that the downstream genes could not generate new stripes on their own. Since each stripe location was defined by the expression of the upstream genes, the downstream genes could not translate their stripe location in parallel. For the violation of evo-devo congruence to occur without the loss of fitness, two mutations, one to delete a stripe and one to add a stripe, had to occur at the same time, otherwise, downstream stripe formation would be damaged, and fitness would decrease. As two such simultaneous mutations are less probable, the violation of evo-devo congruence under feedforward-feedforward network rarely occurred. Conversely, in the case of the feedforward-feedback network, the downstream feedback loop maintained the stripe formation mechanism, and the upstream changes affected only the boundary condition. Hence, as a result of a single mutation, the stripe position could be shifted without destroying it. In this instance, only a single mutation was needed, which is why the violation of evo-devo congruence we observed was always in association with the feedforward-feedback network rather than through a sequential feedforward network.

**Relevance of our results to developmental and evolutionary biology**

Now we discuss the relevance of our results to evolution and development of biological patterns, corresponding to the points noted above.

(**1**) Note that the developmental process evolved in our simulation involved slow change in concentrations of some chemical controlling the dynamics. Slow gradual changes in the concentrations of several chemicals are known to play an important role in the developmental process, which may involve some signal molecules, hormones, and morphogens (Carroll et al., 2009). The developmental process is generally believed to consist of successive stages, each of which involves time spans with slow gradual change, and epochs involving drastic change leading to the next stage. Novel dynamical processes are necessary for such epochal changes(Carroll et al., 2009). This empirical facts in development are consistent with our observations, while our bifurcation scheme provided an interpreta-

tion for commonly observed developmental stages. Because processes that generate such drastic changes are not fully understood so far in developmental biology, it will be relevant to analyze such changes in terms of dynamical systems, in particular, by bifurcation against slow change in some concentration of chemicals.

(**2**) The existence of a quasi-stationary regime and rapid change are often discussed in evolution in terms of punctuated equilibrium (Eldredge and Gould, 1972). Indeed morphological changes observed through fossil data have suggested these temporal modes throughout the course of evolution. Our results suggest that such temporal modes can be explained as bifurcation. Indeed, research has suggested that novel developmental events are acquired in evolution as a result of bifurcations (i.e., *evolution as bifurcation*)(François and Siggia, 2012, Jaeger et al., 2012).

From fossil data it is difficult to confirm our theoretical consequence that the acquisition of morphological novelty in evolution is achieved by bifurcation in developmental dynamics. Alternatively, we expect that a novel morphological pattern may be achieved by imposing suitable changes in gene expression dynamics that might correspond to evolutionary change. For example, by introducing a hormone and over-expressing a single gene, Freitas et al succeeded in inducing fin distal expansion and fin fold reduction in zebrafish, which conceivably represented a prototype of vertebrate appendages(Freitas et al., 2012). The induced epigenetic change leads to a novel gene expression pattern, thereby generating a stripe pattern. This may correspond to bifurcation of a spatial pattern due to genetic change of expression dynamics in our study.

(**3**) Correspondence between evolution and development through common bifurcation is difficult to check directly from experimental data, since the morphology is not easily traced through an evolutionary course, while the comparison of phylogeny and ontogeny usually involves of the morphology only of present organisms diverged from common ancestral species (Figure 2.1). However, if the morphological novelty is a result of bifurcation, different novel morphologies are expected to be diverged from a common ancestral pattern, through different bifurcations. This viewpoint is consistent with von Baer's third law of embryology, which claims that a common basic morphological feature of the group emerges in advance of special features for each species. If we assume that ancestral features are basic for the group, our result suggests that von Baer's third law is valid for strongly purifying selection during evolution.

Currently, a popular topic in the evo-devo field is to examine the existence of phylotypic stages and the validity of the developmental hourglass (Hazkani-Covo et al., 2005). Several recent studies have investigated these topics for different species including Drosophila (Kalinka et al., 2010), vertebrates (Irie and Kuratani, 2011), plants(Quint et al., 2012), Caenorhabditis (Quint et al., 2012) and the soft-shell turtles (Wang et al., 2013). As previously mentioned, our study was unable to provide direct evidence to support the developmental hourglass, since it was not intended for species-wide comparison

but for a single chain in the phylogeny to unveil the basic mechanism for the evo-devo congruence. Also, all individuals in our model were subject to the same initial conditions, with suppressed expression of genes under the same external morphogen gradient, which could not adequately mimic the conditions that would be observed for real multicellular organisms (Sander and Schmidt-Ott, 2004). Despite these deficiencies, the results of this study may have relevance to species-wide comparison also. Our results suggest that evolutionary branching from common ancestral pattern to generate diverse morphological patterns occurs through bifurcation in dynamical systems from common ancestral pattern. Diversification from the bottleneck in evolution and development can be understood accordingly, which may give a basis for the hourglass model.

(**4**) The importance of feedforward and feedback regulations in development has now gained more extensive recognition. The relevance of the successive combination of feedforward networks has been recognized in long-germ segmentation processes in Drosophila, together with theoretical analyses (von Dassow et al., 2000, Jaeger et al., 2004, Ishihara et al., 2005). On the other hand, the relevance of a feedback loop to form temporal oscillations has been revealed for several decades (Goodwin et al., 1963, Chance et al., 1967, Cooke and Zeeman, 1976, Horikawa et al., 2006, Masamizu et al., 2006). In vertebrates, Pourquié discovered that somite genesis is achieved by mapping this temporal oscillation to a spatial stripe formation, where a wavefront model is applicable (Pourquié, 2003). Our mechanism to fix the temporal oscillation to the spatial pattern is similar to the wavefront model, but has a critical difference. In the wavefront model, a combination of the morphogen gradient, size growth, and oscillatory dynamics forms the stripe pattern, while, in our case, a combination of cell-to-cell interaction with diffusion and oscillation leads to stripe formation under the boundary conditions provided by upstream feedforward gene regulation. This distinction will be experimentally verifiable by determining whether the cell-to-cell interaction is essential to stripe formation. Detailed analysis of this boundary induced pattern formation is in chapter 3.

Here we also uncovered the importance of the feedforward-feedback combination, to read external morphogen information leading to robust stripe formation in a bounded domain, where the boundary condition to determine the domain is supported by the upstream feedforward network. Complex gene regulatory networks in the present organisms often include a combination of feedforward-feedback networks (Carroll et al., 2009), although their functional roles have not been fully uncovered. It will be important to elucidate the role of the feedforward network as a boundary-maker and the role of the feedback loop in robust patterning, as suggested here.

(**5**) Experimental confirmation of the violation of evo-devo congruence through modification of the upstream feedforward network, with a conserved downstream feedback loop, is expected to be difficult, considering limitations in available evolutionary data. Therefore, we propose to examine whether morphological novelty arises as a result of modification of upstream feedforward regulation under feedback regulation. While direct examination

of our feedforward-feedback hypothesis from the data is difficult, it is possible to evaluate the hypothesis by externally destroying the upstream feedforward network while retaining the feedback loop.

# Supplementary Materials

**Figure S2.1 Space-time diagrams of evolution, development and differences between the two.** (**A**)-(**H**) Eight additional space-time comparisons between evolution and development. Each consists of a space-time diagram of

development, evolution, preset target pattern and difference between evolution and development. Space-time diagrams of evolution and development and their target pattern are plotted in the same way as in Figs.3 and 4. Calculated values of the difference $\Delta$ are shown below the diagram. For A-D, the evolution and development corresponded well, while a clear violation was observed for H.

**Text S1.1 Mathematical analysis of the slow variable**

Here we demonstrated that slow relaxation emerges even without small the differences in $\gamma_i$. We considered the case that $\gamma_i$ is uniform for all $i$. For simplicity we set here $\gamma_i = 1$. The stability of the fixed point was given by eigenvalues of the Jacobian matrix $W_{ij}$ where the diagonal component $W_{ii}$ is given by $-1$ and the off-diagonal component $W_{ij}$ is given by $J_{ij}\beta exp(-\beta X_i)/(1+exp(-\beta X_i))^2$ where $X_i = \sum_j J_{ij}(x_j - \theta_j)$. If $x_j$'s are close to 0 or 1, their deviation from $\theta_j$ is sufficiently larger than the detection threshold $1/\beta$, the off-diagonal elements are close to zero, and the eigenvalues are given by -1 (or -$\gamma_i$ if it is not 1). When $x_j$'s takes on intermediate values closer to $\theta_i$, the off-diagonal elements assume larger values, and the eigenvalues are shifted from -1, either upwards or downwards. Hence, some exponents approach zero. As long as the real components of the eigenvalues are negative, the fixed point remains stable, but the stability is weaker, with the exponent closer to zero. This results in an increase in the timescale of the relaxation, given by the inverse of the real component of the eigenvalue. With this mechanism, the slowly changing variable is generated even without small $\gamma_i$.

# Chapter 3

# Boundary-Induced Pattern Formation from Temporal Oscillation: Spatial Map Analysis


## Abstract

Boundary-induced pattern formation from a spatially uniform state is investigated using one-dimensional reaction-diffusion equations. The temporal oscillation is successively transformed into a spatially periodic pattern, triggered by diffusion from the fixed boundary. We introduced a spatial map, whose temporal sequence, under selection criteria from multiple stationary solutions, can completely reproduce the emergent pattern, by replacing the time with space. The relationship of the pattern wavelength with the period of oscillation is also obtained. The generality of the pattern selection process and algorithm is discussed with possible relevance to biological morphogenesis.



*The content of the present chapter was accepted by europhysics letters*

## 3.1 Introduction

Spatial patterns are ubiquitous in non-equilibrium systems. The process of spontaneous pattern formation has been extensively studied as one of the main issues in non-linear dynamics. Topic areas in pattern formation include fluid, solid-state, optical, geophysical, chemical, and biological systems(Nicolis and Prigogine, 1977, Cross and Hohenberg, 1993, Walgraef, 1997, Winfree, 2001, Kapral and Showalter, 2012). In particular, the reaction-diffusion system, which was pioneered by Turing(Turing, 1952), has been a focus of non-equilibrium studies. In his celebrated study, Turing showed how a spatial pattern or temporal rhythm is spontaneously formed from a spatially homogeneous and temporally stationary state, which is unstable against perturbations. Turing classified such pattern formation processes into six cases, one of which is known as the *Turing pattern*; stationary periodic pattern with the finite wavelength given by linear stability analysis(See also chapter 1.3). While Turing originally proposed his theory as a model of morphogenesis, it has been applied to general pattern formation dynamics beyond developmental biology(Segel and Jackson, 1972, Lenfyel and Epstein, 1991).

Turing also noted that, in a reaction-diffusion system, spontaneous pattern formation from temporally dynamic states is also possible. (Turing, 1952). Destabilization of a spatially homogeneous and temporally oscillatory state by diffusive interaction often results in spatiotemporal dynamics such as waves, spirals, and turbulence(Kuramoto, 2012). Spatiotemporal dynamics by combination of Turing instability for pattern formation and Hopf bifurcation for temporal oscillation have been studied as Turing–Hopf bifurcation(De Wit et al., 1996, Meixner et al., 1997, Baurmann et al., 2007).

However, an important factor that has not been fully explored is the influence of boundary conditions on spatial patterns. In particular, pattern formation from temporally dynamic states may be crucially influenced by the introduction of a fixed boundary. The influence may be globally propagated to alter the pattern dynamics. (See also (Pomeau and Zaleski, 1981, Fujimoto and Kaneko, 2001) for the relevance of boundary conditions in pattern dynamics.) In the present letter, we are concerned with such boundary-induced pattern formation, namely, how a fixed boundary condition destabilizes the stable, temporally periodic, spatially uniform state and transforms it into a stationary spatial pattern. Indeed, the pattern dynamics of feedback oscillaiton in cooperation with feedforward mechanism in chapter 2 is an example of such spatial pattern formation.

The question of boundary-induced pattern formation is not only of theoretical interest but also of experimental interest in developmental biology. In the somitogenesis of vertebrate development, temporal oscillation in protein expression is fixed into a spatial pattern(Palmeirim et al., 1998, Pourquie, 2001). In chick development, the gene expression of c-hairy1, which initially oscillates in the presomitic mesoderm, is fixed into a striped pattern along aligned cells. So far, such spatial pattern formation has been studied by introducing external inputs that move in space with time development (*clock*

*and wavefront*)(Cooke and Zeeman, 1976). Diffusive interaction has been introduced by Meinhardt in addition to external input using a spatial gradient(Meinhardt, 1982). In his theoretical study, the use of an external input is essential to destabilize the homogeneous state for pattern formation. Some recent experiments, however, suggested that an external input may not be essential, and intrinsic instability due to diffusion might result in somitogenesis(Dias et al., 2014, Cotterell et al., 2015). If boundary-induced pattern formation is possible without imposing an external input throughout the space, it will provide a plausible mechanism for vertebrate somitogenesis.

Here, we demonstrate that temporal oscillation in a spatially one-dimensional reaction-diffusion system is fixed into a stationary periodic spatial pattern by introducing fixed boundary condition. In this case, in contrast to the celebrated Turing pattern, the wavelength of the generated pattern cannot be obtained using linear stability analysis around the fixed point. Instead, to predict the selected pattern, we introduced a one-dimensional spatial map, whose attractor gives the one-dimensional pattern by replacing time with space. The generality of this oscillation fixation and the pattern selection mechanism will be discussed.

## 3.2 Results

Now, consider a spatially one-dimensional reaction-diffusion system of two components $X$ and $Y$. We assume that the diffusion of $X$ is much faster than that of $Y$, and the diffusion of the latter is neglected for simplicity (the formalism to be discussed is valid even without this approximation). Then, the equation is written as

$$\begin{aligned} \frac{\partial X}{\partial t} &= f(X,Y) + D\frac{\partial^2 X}{\partial x^2} \\ \frac{\partial Y}{\partial t} &= g(X,Y) \end{aligned} \tag{3.1}$$

where $f(X,Y)$ and $g(X,Y)$ are the reaction functions for $X$ and $Y$, $D$ is the diffusion constant of $X$, and the attractor of the dynamical system without the diffusion term is a limit cycle. [1].

[1]When this limit cycle is generated by a Hopf bifurcation from the fixed point $(X^*, Y^*)$, the eigenvalues of the Jacobi matrix around the fixed point are complex with a positive real part and a non-zero imaginary part such that $a + d > 0$ and $-2 < \frac{d-a}{\sqrt{-bc}} < 2$(Turing, 1952), where

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial X} & \frac{\partial f}{\partial Y} \\ \frac{\partial g}{\partial X} & \frac{\partial g}{\partial Y} \end{pmatrix}_{X=X^*, Y=Y^*} \tag{3.2}$$

.

We consider the case in which the spatially uniform oscillatory state is stable against perturbations such that under Neumann or periodic boundary conditions, the spatially uniform oscillation is the attractor of the system. On the other hand, under a fixed boundary condition, the variable $X$ close to the boundary cannot oscillate, which may destabilize the oscillatory attractor. We have found that a fixed periodic spatial pattern is often generated in this case.

As a specific example, consider the following system (see also 2.2 and 2.3, "mechanisms of patten formation and their dependency),

$$\begin{aligned} f(X,Y) &= \frac{1}{1+e^{-\beta(Y-1/2)}} - X \\ g(X,Y) &= \frac{1}{1+e^{-\beta(Y-X)}} - Y, \end{aligned} \tag{3.3}$$

which describes the protein expression dynamics with two genes, where $X$ inhibits the expression of $Y$ and $Y$ activates the expressions of both $X$ and $Y$(Mjolsness et al., 1991, Goto and Kaneko, 2013). If $\beta > 8$ and $D = 0$, this system has one unstable fixed point $(X^*, Y^*) = (1/2, 1/2)$, and the limit cycle is an attractor (see Fig.3.1A). Indeed, from initial conditions close to a spatially homogenous state, a uniform, limit-cycle state is reached if the boundary condition is Neumann or periodic. In contrast, if a fixed boundary condition is adopted for at least one end, i.e. $X(0,t) = X_0$, temporal oscillation is replaced by a fixed pattern [2] (see Fig. 3.2A and 3.2B where $\frac{\partial X}{\partial x}|_{x=L} = 0$ is adopted at the other end). Note that a pattern of the same wavelength is organized independently of $X_0$ as long as $X_0$ is between [0,1](see Fig. 3.3). Here, oscillation ceases in the vicinity of $x = 0$, which works as a boundary for $(X(x), Y(x))$ for larger $x$. Accordingly, the oscillation successively ceases for larger $x$ and is replaced by a fixed, periodic pattern.

Note that this pattern formation dynamics takes place far from the uniform fixed-point solution. Hence, the standard linear stability analysis for Turing pattern concerning around the uniform fixed-point solution cannot explain the wavelength observed here. (see Supplemental Figure 3.1).

Meanwhile, the emergent stationary pattern has to satisfy

$$\begin{aligned} f(X,Y) + D\frac{\partial^2 X}{\partial x^2} &= 0 \\ g(X,Y) &= 0 \end{aligned} \tag{3.4}$$

Eq.(3.4) describes nullclines of $X$ and $Y$ of the emergent stationary pattern at each location. As the nullcline of $X$ is changed by the diffusion term from the original ordinary

[2] A boundary condition for $Y$ is not required as diffusion in $Y$ is neglected.

**Figure 3.1.** (**A**) The nullclines $f(X,Y)=0$ (red) and $g(X,Y)=0$ (black) as well as the limit cycle orbit (green) are plotted in the state space for the model (3.3), for the case $D=0$. Flow in the state space is also illustrated by arrows. (**B**) State space plot of nullclines of $X$ ($f(X,Y)+D\frac{\partial^2 X}{\partial x^2}=0$, black) and $Y$ ($g(X,Y)=0$, red) are plotted at some location $x$. Flow in the state space is also illustrated as arrows. The nullcline of $X$ is shifted horizontally from the $D=0$ case with the spatial gradient at each location (See also Fig.3.3). The series of cross-points of the nullclines at $Y\sim 1$ or $Y\sim 0$ at each location give the emergent stationary pattern orbit (green).

**Figure 3.2.** Time development of the model (3.3) under a fixed boundary condition, $X(0)=0$. (**A**) $X(x,t)$ is displayed with a color scale given by the side bar. The abscissa is time $t$, and the ordinate is space $x$. (**B**) The generated stationary pattern $X(x)$ and $Y(x)$. The abscissa is the concentration of $X(x)$ or $Y(x)$, and the ordinate is space $x$. $\beta=40$ and $D=1.6\times 10^{-3}$.

**Figure 3.3.** Stationary pattern $(X_{st}(l), Y_{st}(l))$ of the reaction diffusion dynamics (3.3) and the predicted pattern of $(X_l, Y_l)$ obtained from the spatial map (3.6). The predicted pattern (given by $\circ$) agrees well with the stationary pattern (lines). Three patterns with boundary values of $X_0 = 0.0$ (purple), $X_0 = 0.5$ (green), and $X_0 = 1.0$ (yellow) are considered. $\beta = 40$ and $D = 4.0 \times 10^{-4}$.

**Figure 3.4.** (**A**) The nullcline of $Y$, given by $g(X,Y) = 0$ (red solid). Crossing-points of $g(X,Y)$ and $X = X_l$(black dashed) show candidate solutions of spatial map at $l$(green circles). (**B**) State space plot of the nullcline of $Y$(red solid) and the nullcline of $X$ when $D = 0$(black solid). The two dashed lines are nullclines of $X$ assuming $Y_l = Y_l^+$ (the upper one with the crossing-point $(X_l, Y_l^+)$) and that assuming $Y_l = Y_l^-$(the lower one with the crossing-points $(X_l, Y_l^-)$). Distance between each dashed lines and the original nullcline $F(X,Y) = 0$ corresponds with the diffusion terms, i.e., $D^+ = D^{'}|X_{l+1}^+ - 2X_l + X_{l-1}|$ or $D^- = D^{'}|X_{l+1}^- - 2X_l + X_{l-1}|$.

differential equations, under the diffusion term $D\frac{\partial^2 X}{\partial x^2}$, novel fixed points can appear.

At each location, crossing-points of two nullclines indicate such novel fixed points. Stationary pattern is described as the chain of such novel fixed points along space. Actually, for a given stationary pattern, we confirmed that the stationary state $(X, Y)$ is same as a novel stable fixed point given by calculated diffusion term at each location(See Fig.3.4B). Thus, under the diffusion term at each location, the stationary pattern is computed.

By adopting spatial discretization, this condition is given by

$$\begin{aligned} &f(X_l, Y_l) + D^{'}(X_{l+1} - 2X_l + X_{l-1}) = 0 \\ &g(X_l, Y_l) = 0, \end{aligned} \tag{3.5}$$

where $l$ is discretized space index with $a$ as the discretized unit length, i.e. $x = al$ and $X_l = X(x/a)$, $Y_l = Y(x/a)$. Here, $D^{'}$ is the rescaled diffusion constant, which satisfies $D' = D/a^2$. From these equations, we can derive the spatial map as follows:

$$\begin{aligned} &X_{l+1} = 2X_l - X_{l-1} - \frac{f(X_l, Y_l)}{D'} \\ &Y_l = g_{X_l}^{-1}(0), \end{aligned} \tag{3.6}$$

where $g_{X_l}^{-1}$ is the inverse function of $g(X, Y)$, given that $X$ is fixed to $X_l$, which corresponds to the cross-points of the two nullclines. With these equations, $X_{l+1}$ can be determined from $X_l$, $Y_l$ and $X_{l-1}$, while $Y_l$ is determined as $g_{X_l}^{-1}(0)$ (for the use of a "spatial map" in the analysis of a spatial pattern, see also (Aubry, 1978, Willeboordse and Kaneko, 1995)). Hence, the spatial pattern is determined as a time-series of the mapping (3.6) by replacing $l$ to spatial sequence.

Here, however, multiple branches in the solution can exist for $Y_l = g_{X_l}^{-1}(0)$, as shown in Fig.3.4A. All of these give potential, stationary solutions, but not all of the candidate solutions $(X_l, Y_l)$ are stable in terms of dynamical systems (3.1). Furthermore, we need to choose such solution that is attracted from the uniform oscillatory state, according to the original reaction-diffusion equation (3.1). These properties give constraint on the selection of the branch $Y_l = g_{X_l}^{-1}(0)$.

Consider the case of eq.(3.3). Fig.3.4A shows the candidate solutions at given $l$. There are three candidate solutions of $Y_l$, namely $Y_l^+, Y_l^{mid}$, and $Y_l^-$, given by different branches of $g_{X_l}^{-1}(0)$. Here, $(X_l, Y_l^{mid})$ is an unstable fixed point, as can be directly seen in the flow in Fig.3.1B. Thus, there remain two candidates for the emergent stationary pattern, i.e. $(X_l, Y_l^+)$ or $(X_l, Y_l^-)$. Both these candidate solutions are stable fixed points, as seen in

Fig.3.1B and analyzed in supplementary text ST3.1.

Then, the constraint that the fixed-point at $(X_l, Y_l)$ is attracted from a uniform oscillatory state under the boundary condition determines the solution to be selected: As the pattern emerges from a uniform state and $X$ is under diffusion to homogenize the solution, such a solution with a smaller diffusion term $|X_{l+1} - 2X_l + X_{l-1}|$ should be selected. From eq.(3.5), this diffusion term is given by $|f(X_l, Y_l)|$, which is identical with the distance from the original nullcline of $X$, i.e. $f(X, Y) = 0$. Then, as shown in Fig.3.4B, once a given branch (say $Y^+$) is selected, the diffusion term $|X^+_{l+1} - 2X_l + X_{l-1}|$ is smaller than that of the other branch (i.e. $|X^-_{l+1} - 2X_l + X_{l-1}|$) up to a certain lattice point $l^*$ (to be shown below). Owing to the continuity of $\frac{\partial^2 X}{\partial x^2}$, the solution at the same branch (e.g. $Y^+_{l+1}$) will be selected at the next location(see Fig.3.4B). Now, by taking the same branch (say + in Fig.3.4B), $|f(X_l, Y_l)|$, and accordingly, the distance between $(X_l, Y_l)$ and the original $X$-nullcline increases. Then, at a certain lattice point $l = l^*$, $|f(X_l, Y^+_l)|$ turns to be equal to $|f(X_l, Y^-_l)|$, beyond which the latter is smaller. Hence, beyond $l = l^*$, $|X^+_{l+1} - 2X_l + X_{l-1}|$ exceeds $|X^-_{l+1} - 2X_l + X_{l-1}|$. Then, $Y^+_l$ is no longer selected but $Y^-_l$ is selected, due to the proximity to uniformity.

Here, it should be noted that by the original dynamics given by eq. (3.3), $X, Y$ moves along the limit-cycle attractor. Then, starting from $(X^+_l, Y^+_l)$, $(X, Y)$ is directed toward $(X^-_l, Y^-_l)$. Hence, the $-$ branch of solution that gives rise to a smaller diffusion term is accessible by the original dynamics such that the selected branch switches from + to -. Then, the minus branch remains to be selected due to the continuity, until $|X^-_{l+1} - 2X_l + X_{l-1}|$ exceeds $|X^+_{l+1} - 2X_l + X_{l-1}|$. Repeating the above procedure, the emergent periodic pattern is predicted.

As the present model has point symmetry with respect to $(X, Y) = (0.5, 0.5)$, the condition for selection of the solution branch is simply written as

$$Y_l = \begin{cases} Y^+_l & (|D'(X^+_{l+1} - 2X_l + X_{l-1}| < 0.5) \\ Y^-_l & (otherwise) \end{cases} \tag{3.7}$$

An example of the pattern obtained from the spatial map with the above selection criterion is shown in Fig. 3.3, which completely agrees with the numerically obtained pattern. The above procedure works independently of the parameter values, demonstrating its validity.

Note that the spatial periodic pattern is obtained as an attractor of the spatial map (3.6). Hence, independently of the initial condition in the map, the same periodic pattern is selected as long as the initial condition in the spatial map belongs to the basin of the above attractor. The initial condition in the spatial map corresponds to the value of fixed boundary. Hence, a pattern of the same wavelength is reached independently of the boundary value, while the phase of wave pattern is different, which is consistent with the

numerical result (see Fig. 3.3).

This procedure also suggests that there can be a relationship between the period of oscillation $T$ and the wavelength $\lambda$. First, by rescaling the spatial scale, the wavelength $\lambda$ is proportional to $\sqrt{D}$. (The period is scaled by $\tau$ if $f$ and $g$ are similarly scaled by $1/\tau$, but this is already set to unity in (3.1).) Then, we have found $\lambda/\sqrt{D} \propto T$. In fact, we numerically confirmed that one stripe is generated in correspondence with one period of oscillation from a uniform limit-cycle state, except for the vicinity of the Hopf bifurcation (See Fig.3.2A).

Now, following the vector field $(f(X,Y), g(X,Y))$ in the state space (See Fig. 3.1A), $(X(t), Y(t))$ oscillates with the period $T$. Here, within one period of oscillation, $f(X,Y)$ increases and takes local maximum, and decreases to change the sign, and then takes local minimum, and increases to change the sign, as in the state space in Fig.3.1B.

On the other hand, when a stripe is formed, the "spatial orbit" $(X(x), Y(x))$ against $x$ is shaped within the state space. Recalling that the switch to the other branch occurs when $|f(X,Y)|$ takes a local maximum, $f(X,Y)$ changes along with the spatial sequence as follows: By starting from the + branch, $f(X,Y) > 0$ decreases to take the local minimum, and then increases to take a local maximum, and then $f(X,Y)$ switches its sign. Entering the - branch, $|f(X,Y)|$ first decreases to local minimum, and then increases to take the local maximum to switch to the + branch. Hence, although the timing of sign change is different, $f(X,Y)$ and its derivative experience one cycle of sign change both through the limit cycle and spatial orbit.

With the parameter change, the (magnitude of) vector $(f, g)$ in the state space changes, which changes the angular velocity along the limit cycle as well as that of the "spatial orbit" in the same way. Hence, the period $T$ and wavelength $\lambda$ (or $\lambda/\sqrt{D}$) change in proportion. Of course, this is a rough estimate, but this proportionality approximately holds against changes in the parameter values $\beta$ (and $D$), as shown in Fig. 3.5.

For smaller values of $\beta$, deviation from the proportionality is discernible, where spatial pattern exhibits period-doubling (see Supplementary Figure SF3.2). For a much smaller value of $\beta$, the pattern predicted by the spatial map can be quasiperiodic or chaotic, where, however, such spatial pattern that is fixed in time turns to be unstable in the time evolution of the original reaction-diffusion dynamics. Then, the attractor is no longer a fixed-point in time but is replaced by spatiotemporal dynamics.

The fixation of the temporal to the spatially periodic pattern as well as the analysis using the spatial map can be generally applied to a spatially one-dimensional reaction-diffusion equation that has a uniform limit-cycle attractor. As another illustration, we have also examined the so-called Brusselator(Nicolis and Prigogine, 1977):

**Figure 3.5.** Relationship between the scaled wavelength $\lambda/\sqrt{D}$ of the emergent pattern and the period of the limit cycle orbit $T$ is shown by changing parameter $\beta$. Results from different $D$ values are also plotted with different symbols. The relationship from the Brusselator eq. (3.8) is also plotted from the obtained data by varying $B$. $A$ is fixed to 1.

$$\begin{aligned} f(X,Y) &= BY - XY^2 \\ g(X,Y) &= A - (B+1)X + XY^2 \end{aligned} \tag{3.8}$$

In this case, a uniform periodic oscillation is replaced by a spatial pattern by applying the fixed boundary condition $X(0) = X_0$, for a certain range of parameters. Selection of the branch for $Y_l$ in the spatial map is much easier for this case than eq.(3.3). One can just choose the same branch as long as the branch exists and it is in the basin of the limit cycle, i.e. the solution satisfies $X_l > 0$ and $Y_l > 0$. In fact, the switch of the branch due to the selection of smaller $|X_{l+1} - 2X_l + X_{l-1}|$ need not to be taken into account in this case because the flow in the state space goes toward the fixed point of the original branch, and the switch cannot occur starting from the limit-cycle attractor (This is because the rotation along the limit cycle of the Brusselator model is counterclockwise).

Again, the pattern predicted through spatial map and above selection criteria of branches agrees well with the emergent pattern (see Supplemental Figure 3.3). The approximate proportionality between the period and wavelength is worse than the model (2) (see Fig. 3.5), possibly because the correspondence between the flows in the spatial map and original limit cycle is not so good in the state space.

## 3.3 Discussion

In the present chapter, we have studied the formation of a periodic pattern from a uniform oscillatory state induced by the boundary conditions. The emergent pattern is predicted as an attractor of the spatial map under the selection principle. We have confirmed the generality of this pattern formation in reaction-diffusion systems. The periodic oscillation exists in a two-component reaction system with negative feedback. That is, in a system with an activator and an inhibitor, for a certain range of parameters, the spatially periodic pattern that exists as an attractor replaces the uniform oscillatory state, triggered by the fixed boundary condition, if the diffusion of the inhibitor is sufficiently large.

We also note that the pattern formation process as well as the nullcline analysis is still valid even for $D_Y \neq 0$. As long as $D_Y$ is small, the present stripe formation will be preserved under perturbative analysis. The analysis of spatial map can be applicable generally, although the inverse function that includes both $X$ and $Y$ would make the analysis more difficult.

Pattern formation generally depends on the boundary and initial conditions. In the spatial map, the former is represented by the initial condition in the map, while the latter is considered by the proximity to uniformity, which indicates that the initial condition is not far from the uniform oscillatory state. If the initial condition is far from uniformity, local inhomogeneity can grow. In the state-space representation, whether such growth occurs is determined according to the size of the spatial diffusion term, specifically whether the spatial diffusion term is large enough to go across the nullclines and to induce a switch to the other branch. In other words, the condition close to uniformity represents the absence of such inhomogeneity.

On the other hand, for such initial conditions that allow the local growth in inhomogeneity, the selection criterion of the solution branch is the proximity to a given initial condition, which leads to a requested switch to a different branch of fixed-points. In other words, the choice of the fixed-point solution corresponding to the initial condition can predict the emergent pattern. Conversely, control of an emergent pattern is thus possible by manipulating the initial condition.

Experimental confirmation of the present pattern formation will be possible in reaction-diffusion systems. In particular, relevance of boundary condition to pattern selection should be of importance because the real experimental system is finite and often under a fixed boundary condition. By carefully examining the boundary effects, we can confirm the present pattern formation mechanism, where the period-wavelength relationship (Fig.3.5) will be confirmed.

As mentioned in section 3.1, spatial pattern formation based on temporal oscillation is often observed in somitoogenesis(Palmeirim et al., 1998, Pourquie, 2001) as well as in the numerical evolution of that models somitogenesis(Fujimoto et al., 2008, Kohsokabe and Kaneko, 2016). Historically, reaction-diffusion system, in particular cell-cell interaction was not thought to be relevant to somitogenesis. This was due to the experimental observation that during somitogenesis, somite can scale its size in proportion to the size of embryo so that total number of somites are conserved(Cooke, 1975), whereas Turing pattern with proportional conservation was unknown at that time(Ishihara and Kaneko, 2006). Somitogenesis was thought to be controlled by interplay of oscillation of gene expressions(clock) and long tail gradients from both anterior and posterior end of embryo which determine the location where the oscillation ceases(wavefront).

However, recent studies suggest that cell-cell interaction could be important to somitogenesis rather than chemical gradient. Accordingly relevance of reaction-diffusion pattern formation dynamics to somitogenesis has been recognized(Cotterell et al., 2015, Kohsokabe and Kaneko, 2016). In the numerical simulation in chapter 2, the present interaction-induced fixation of oscillation is often observed (see also Supplementary figure SF3.4, SF3.5), where the boundary condition is provided by upper-flow gene regulation network. Considering the simplicity in our mechanism, which only requires diffusion of the inhibitor and a fixed boundary, we expect that it could be adopted in biological development, which will be confirmed by examining cell-cell interaction and boundary effects.

## Supplemental Materials

**SF3.1** (**A**) The real part of the Jacobi matrix eigenvalues around the uniform solution for each wavenumber $k$ for model (2), when $\beta > 8$. The eigenvalues are positive for $0 < k < k_1$ (where the imaginary part is non-zero) and for $k > k_2$ (where the imaginary part is zero). The imaginary part of the eigenvalue vanishes at $k = \frac{1}{4\pi}\sqrt{\frac{\beta}{D}} < k_2$. In the blue region, both Turing and Hopf instability exist, and only Turing instability exists in the red region. (**B**) The emergent wavenumber $k_s$ (green, +), marginal wavenumber for the blue region in (A) $k_1$(blue), and marginal wavenumber for the red region in (A) $k_2$ (red) are plotted as a function of $\beta$. $D = 1.0 \times 10^{-4}$.

**SF3.2** An example of the emergent spatial pattern which exhibits period-doubling. The pattern $X(x)$ is plotted as a function of $x$. $\beta = 44.583$ and $D = 1 \times 10^{-3}$.

**SF3.3** Stationary pattern $(X_{st}(l), Y_{st}(l))$ of the reaction-diffusion equation (3.1) with reaction term (3.8) and the predicted pattern of $(X_l, Y_l)$ obtained from spatial map (3.6). The predicted pattern (given by $\circ$) agrees well with the stationary pattern (lines). Three patterns with boundary values of $X_0 = 0.0$ (purple), $X_0 = 1.5$ (green), and $X_0 = 3.0$ (yellow) are considered. $A = 1$, $B = 3$ and $D = 4.0 \times 10^{-4}$.

**SF3.4 Detailed analysis on feedback oscillation.**
(**A**): network structure that represents reaction diffusion system shown in eq.(3.3). This network structure is as same as that is presented in Figure 2.9. (**B**): Phase diagram of the expression dynamics. Two nullclines of gene expression cross at a single, unstable fixed point, and the cell state will oscillate on a limited cycle. Green circles represent the cell state at time step intervals of 5, within a single cycle. The distance between two nullclines is shortest at the upper right and lower left corners so that cell state changes are slower at these corners. (**C**): Time profile of the feedback oscillation for a specific cell. The abscissa represents developmental time and the ordinate is expression level. Gene A is plotted as a red line, while gene B plotted as a Blue line.

**SF3.5 Temporal change of flow in the phase space over cells.**

The oscillation fixation mechanism is revealed through comparison of flow temporal changes in the phase space over cells. The central cells of the first two stripes (cell indices 12 and 28) and valleys (cell indices 4 and 20).

(1)**t=0**:
Both gene A and gene B assume a null value, and gene B is inhibited by gene M within the first 8 cells. With these two initial conditions, flow in the phase space where cell index = 4 is different from other cells. At a stable fixed point therein, both the expressions of gene A and gene B are low (i.e., a low-low state ). This fixed point is the root of the first valley.
(2)**t=80**:
As development begins, expression of non-inhibited cells begin to oscillate and move towards a state where both expressions of gene A and gene B are high (i.e., a high-high state), while cells of the first valley maintain the slow-low state. Thus, protein A diffuses from the first stripe to the first valley. Due to the incoming diffusion of protein expression of gene A, at cell index 4, the nullcline of gene A slides to the right, so that the expression of gene A assumes a higher value. Correspondingly, at cell index 8, the nullcline of gene A slides to the left, and crosses the nullcline of gene B to create a fixed point.
(3)**t=100**:
The expression of cell index 4 is constrained to the newly formed high-high fixed point. However, the expressions of cells at index 20 and 28 continued to oscillate.
(4)**t=180**:
Through the oscillation, cell index 20 approaches a low-low state for the second time, and at this time, protein B at the first stripe diffuses to the second valley. Thus, nullclines slide in the cells at indices 12 and 20 to the left and right, respectively. The nullcline of gene B then crosses, at a low-low state, in the cell at index 20.
(5)**t=200**:
The expression level at cell index 28 continued to oscillate while that at cell index 20 was constrained at the newly formed low-low state fixed point, similar to cell index 12 at t=80. Protein B subsequently diffused from the cell index 28 to the second valley, which resulted in the emergence of the second stripe. In this way stripes were shaped from the oscillation.

**ST3.1 Stability of stationary patterns obtained from spatial map with eq.(3.3)**

As discussed in the main text, the stationary pattern satisfies eq.(3.4), which is obtained as the stationary condition of eq.(3.1). Now we study the linear stability of such fixed-point solution eq.(3.4) at each location, with reaction terms are given by eq.(3.3). The Jacobian matrix for each fixed point is given by

$$J = \begin{pmatrix} \frac{\partial f}{\partial X} & \frac{\partial f}{\partial Y} \\ \frac{\partial g}{\partial X} & \frac{\partial g}{\partial Y} \end{pmatrix} = \begin{pmatrix} -1 & \frac{\beta e^{-\beta(Y-1/2)}}{\{1+e^{-\beta(Y-1/2)}\}^2} \\ \frac{-\beta e^{-\beta(Y-X)}}{\{1+e^{-\beta(Y-X)}\}^2} & -1+\frac{\beta e^{-\beta(Y-X)}}{\{1+e^{-\beta(Y-X)}\}^2} \end{pmatrix} \tag{3.9}$$

As for the +branch or −branch for the spatial map, $Y$ is approximately 0 or 1, respectively, as $f(X,Y)$ and $g(X,Y)$ is close to the step function.

Thus, for both + and − branches, the Jacobian matrix is approximately given by

$$J \sim \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.10}$$

Now, the evolution of small perturbation $(\Delta X(x), \Delta Y(x))$ is written as:

$$\begin{aligned} \frac{\partial \Delta X}{\partial t} &= J_{XX}\Delta X + J_{XY}\Delta Y + D\frac{\partial^2 \Delta X}{\partial x^2} \\ \frac{\partial \Delta Y}{\partial t} &= J_{YX}\Delta X + J_{YY}\Delta Y \end{aligned} \tag{3.11}$$

$$\begin{aligned} \frac{\partial \Delta X}{\partial t} &= -\Delta X + D\frac{\partial^2 \Delta X}{\partial x^2} \\ \frac{\partial \Delta Y}{\partial t} &= -\Delta Y \end{aligned} \tag{3.12}$$

Applying Fourier's transformation that $\Delta X(x,t) = \int dk \xi(k,t) e^{-ikx}$ and $\Delta Y(x,t) = \int dk \eta(k,t) e^{-ikx}$,

$$\begin{aligned} \frac{\partial \xi(\omega,t)}{dt} &= (-1-k^2 D)\xi(\omega,t) \\ \frac{\partial \eta(\omega,t)}{dt} &= -\eta(\omega,t) \end{aligned} \tag{3.13}$$

Thus, the stationary pattern with $+,-$ branches derived by the spatial map is stable as long as $\beta$ is sufficiently large.

On the other hand, for the case $Y_l = Y_l^{mid}$, unlike the solution on +branch and −branch, the real part of the eigenvalues of the Jacobi matrix is positive, as $X_l \sim Y_l^{mid}$.

# Chapter 4

# Time-scale separation of gene expression and evolvability


## Abstract

Evolutionary importance of slow gene expression in development is investigated. By numerical evolution to make different cell types under given input gradient, it is shown that gene expression with much slower time scale than others evolve as also observed in the simulations of chapter 2. The time-scale of slow genes increases together with the fitness increase. Control of development by such slow gene expression is enhanced. Even though such slow changes in genetic expression is infrequent among the total number of interactions between genes, few genes with such slow control gene show a strong influence on the fitness as demonstrated by numerical knockout experiment. Mechanism for such slow gene control is analyzed, and its relevance to biology is discussed.


*The content of the present chapter is in prep*

## 4.1 Introduction

For developmental processes to work in proper order, developmental control in time is important. For example, consider the boundary induced pattern formation dynamics analyzed in chapter 3. This dynamics can be implemented in the model of chapter 2, in a way that oscillation of gene expression as a feedback loop in GRN, while gene expression pattern generated by the feedforward loop in the upstream of the feedback loop sets the boundary condition for the downflow expression dynamics. In order for such dynamics to function properly, the gene expression for establishing the boundary condition should be expressed earlier than the oscillatory expression, i.e. the expression dynamics at the feedforward network has to work much earlier than that by the feedback loop. Otherwise, oscillation would propagate to the region beyond the expected domain. This inappropriate order of gene expressions would result in pattern alteration.

In chapter 2, we showed that slow gene expressions in development work as control parameters. With the gradual change, these slow gene expressions can encode the order of gene expression in the form of bifurcation. Such order determines subset of genes that are expressed during a certain epoch, that constitutes working GRN, which is the key concept for the analysis of recapitulation at the network level as discussed in chapter 2. Such, slow genes do not directly act on the target gene that is a marker gene for a phenotype, but rather act on the other genes that give input to the target gene. Timescale of expression change is separated between the target gene and the slow genes, so that the pattern change can be regarded as bifurcation between attractors with slow gene expression as emergent bifurcation parameters.

In chapter 2 we analyzed how slow input genes control their temporal development. Now, although the emergence of slow control in evolution is a general trend of the model therein, plausibility of the emergence of the slow control gene is not fully understood as yet. As discussed in chapter 2, modifying coupling with such slow gene expressions will easily alter the output behavior that they control compared by modification of other genes. Thus slow gene is expected to facilitate evolution. In this chapter, we will investigate the generality of the emergence of slow gene expression and its relationship to evolvability. We adopt a model which is a simpler version of the model in chapter 2. In the model, GRN also evolves to satisfy a particular fitness condition relevant to spatial gene expression patterning. First, we study whether slow gene expression would evolve, and then analyze how it is relevant to increase the fitness. Knock-out of such slow genes has a stronger influence to the fitness, suggesting the control of gene-expression pattern by few slow genes.

## 4.2 Method

### gene expression dynamics

A cell' s state is represented by the expression levels of genes/proteins, $x_i(l,t)$, involving the protein expression levels of the $i$-th gene in the $l$-th cell at time $t$, representing $N$ genes ($i = 1, ., N$) and $M$ cells, aligned in a one-dimensional space. A protein expressed from each gene either activates, inhibits, or does not influence the expression of other genes or itself.

Equation for gene expression dynamics is as same as that of the chapter 2 except for this model has no diffusion term.

$$\frac{dx_i(l,t)}{dt} = \gamma_i(f(\sum J_{i,j}x_j(l,t) - \theta_i) - x_i(l,t)) \tag{4.1}$$

with

$$f(x) = \frac{1}{1+e^{-\beta x}} \tag{4.2}$$

where the term $-x_i(l,t)$ in (4.1) gives the degradation of $i$-th protein with $\gamma_i$ as its rate (Glass and Kauffman, 1973, Mjolsness et al., 1991, Salazar-Ciudad et al., 2001a,b). The expression level is scaled so that the maximum level is unity. The function $f(x)$ is similar to a step function, which approaches 1 as $x$ is increased to a positive side, and approaches 0 as $x$ is decreased to a negative side: In other words, if the term $\sum_j J_{i,j}x_j(l,t)$ is sufficiently larger than the threshold $\theta_i$, then $f(\sum J_{i,j}x_j(l,t) - \theta_j) \sim 1$, which indicates that the gene is fully expressed, and if it is smaller than $\theta_i$, then $f(\sum J_{i,j}x_j(l,t) - \theta_j \sim 0$, which indicates that the gene expression is suppressed. Here, we chose, $f(x) = 1/(1+e^{-\beta x})$, where $\beta$, which was set to 40, denoting the sensitivity of the expression at the threshold. Roughly speaking, it is proportional to the Hill coefficient.

As in chapter 2 .each node of the network represents a gene, and the edge of the network represents the interaction between genes, given by $N \times N$ matrix $\boldsymbol{J} = \{J_{i,j}\}$: where $J_{i,j}$ is 1, if gene $j$ activates the expression of the gene $i$, $-1$ if it suppresses the expression, and 0 if there is no connection.

For the simulations described here, we set $M = 8$, and $N = 10$, while preliminary simulations adopting larger values show that the conclusion in the present chapter is not altered.

### Initial condition for pattern formation dynamics

As an initial condition, the expression levels of all genes were set to 0. Furthermore, external morphogens, which are denoted as the proteins 0 and 1, are supplied externally. Fixed linear morphogens are induced from both sides, so that $x_0(l,t) = x_0(l) = (M-l)/M$ and $x_1(l,t) = x_1(l) = l/M$. Although this model has a spatial index, there is no diffusive interactions so that the boundary conditions are not necessary to be imposed. The spatial index just denotes the level of inuts $x_0$ and $x_1$.

### Initial condition for evolutionary dynamics

Gene interaction matrix $J$ is randomly generated with the ratio that $J_{i,j} = 1 : J_{i,j} = 0 : J_{i,j} = -1 = 1 : 8 : 1$.

The parameter values for the initial generation for each individual are given as follows. All $\gamma_i$ and $\theta_i$ are set to 0.5. Except for difference in GRN structure, there is no difference between individuals in the first place. Unless otherwise mentioned, presented examples are the result of evolutionary simulation under the mutational change imposed only on the timescale of gene $i$ $\gamma_i$, besides the changes in the path of GRN.

### Definition of Fitness

Fitness is defined as the number of cell types that one GRN can generate from eight cells aligned spatially. There are 3 target genes (out of 10 genes) namely $x_2$, $x_3$, and $x_4$. ON/OFF expression of these 3 genes gives $2^3$ cell types. Thus the lowest fitness is 1, i.e., only a single gene expression pattern exists among the eight cells, while the highest fitness is 8, in which case all cell types are different to take all $2^3$ possibilities. To eliminate the possibility of oscillatory expression dynamics, the fitness for a cellular state that is no attracted to a certain fixed state after given (sufficient) developmental time development is set to zero.

For this analysis, we chose 100 individuals with different GRNs. Among these individuals, those which have higher fitness values are selected for the next generation. Selection of the individual $i$ with a fitness value of $f_i$ is defined by

$$p_i = \frac{10^{f_i}}{\sum_k 10^{f_k}} \tag{4.3}$$

The denominator summation of index k is aimed for all individuals in the population.

To generate the offspring, each genotype is slightly modified. A path in GRN is added, eliminated, or its sign is flipped with the probability $1/N^2$. Also, according to

each simulations setup, the parameter values $\gamma_i$, $\theta_i$ were modified by adding a random number from a Gaussian distribution $\eta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{x^2}{2\sigma^2})$, while restricting these values within [0,1]. We set $\sigma = 0.01$.

### ON/OFF determination

If a gene's expression level is higher than 0.99, the gene is regarded to be at the ON state, while if it is lower than 0.01, the gene is regarded to be in the OFF state.

### Computation of the time-scale

In the model, the input term for each expression has a dynamic range given by the slope of the input-output relationship $\beta$, as $[-2/\beta : 2/\beta]$(See also the subsection "gene expression dynamics"). If the input for a gene is out of this range, the expression of the gene is basically fixed at either 0 or 1. The time span for the input to pass through this dynamic range is an indicator for the slow control of the input.

### Timescale ratio

In chapter 2, timescale separation was shown in Fig. 2.8 where time span of the input change of both the target gene and other genes that gives input to the target gene are shown together. The target gene is controlled by faster genes while other input genes are controlled by slow genes. The indicator of timescale separation between genes is obtained by dividing one timescale by another. We define the timescale ratio as the ratio of the timescale of the input changes in non-target genes to that of the target genes.

## 4.3 Result

We prepare 100 individuals with slightly different GRNs (See Method for further information of parameters and initial settings for the simulation) while the fitness is given by the stationary expression pattern of target genes. As mentioned above, three genes in GRN are assumed as these target genes that determine the cellular state, where the combination of on/off states of 3 genes represent $2^3$ cell types. The fitness is defined as the number of cell types that the GRN provides under the input, with 8 as the maximal value.

After sufficient generations, 8 cell types are generated under a given input according to the evolved GRN. Fig.4.1 is an example of the evolutionary simulation result.

**Figure 4.1. An example of evolutionary simulation result**
(A)Expression dynamics of the target genes are shown with developmental time as horizontal axis and cell index as the vertical axis. Each different color represents a different cell type. As in the examples in chapter 2, pattern formation consists of several epochs, i.e., rapid change of the expression level of target genes between quasi-stationary states. (B) Evolutional history of the individual represented in 4-1(A). The target expression level is shown with generations as horizontal axis and cell index as the vertical axis. At each generation, cell-type pattern of ancestors are plotted.

### Emergence of slow-time scale in gene expression as a result of evolution

As in examples in chapter 2, both developmental and evolutional courses consist of several epochs to form pattern with 8 cell types. In Fig. 4.1(A), epochs do not appear within a short span of time, but rather, some epochs appear only at a late stage of development, which suggests existence of slow time scale that controls the development. It is confirmed by checking the timescale in the change of input value to each gene, as is also adopted in the analysis in chapter 2.

Two examples of the evolution of the fitness and time-scale ratio are shown in Fig. 4.2A. We compute the time scales of expression-level change of each gene. Plotted are the average of such time scale for genes that give input to the target genes, to that for the target genes (see also Method). This time scale ratio is an indicator of the time-scale separation; higher the time scale ratio is, the larger the control by slow gene expressions is. Both timescale ratio and averaged fitness gradually increase during evolution. Interestingly, the timescale ratio increases in the same manner with the averaged fitness during the evolution.

To examine the generality of the result over different simulation runs, we computed the change in average of fitness and time-scale ratio over different runs. As the generation that reaches the maximal fitness depends on each run, we first rescale the evolutionary genera-

**Figure 4.2. Evolution of timescale separation and fitness**
(A) Ratio of timescale of input change of target genes to that of non-arget genes(blue) and fitness(green) are plotted according to the rescaled evolutionary time course. Two examples of evolution runs. Timescale ratio gradually increases with the evolutionary process, and increases in a step manner before the maximal fitness is achieved (B) The average of time-scale ratio and fitness over different simulation runs. The average is taken by rescaling the evolutionary time at 1 when the maximal fitness is reached. After the maximal fitness is achieved the timescale ratio does no more increase substantially. For comparison, evolutionary change of the timescale ratio(cyan) and averaged fitness(orange) are also plotted from the evolutionary simulation just by mutation but without fitness

.

tion, so that the time 0 represents the starting generation and 1 represents the generation at which the maximal fitness is reached .i.e., the value 8. Then, we plotted the average fitness and time-scale ratio over runs in Fig. 4.2B. Now the correspondence between the two increases is clear. In fact, this agreement is not a coincidence, but is a result of di-

rected evolution under given fitness condition, as data from evolutionary simulation just by mutation but without fitness does not show such trend of agreement(Fig.4.2B,below). This agreement suggests strong connection between fitness and timescale separation.

We then check the evolutionary change in the distribution of the time-scale in the input to each gene, over runs and over genes, as is shown in Fig.4.3. At the initial generation, both for target and non-target genes have a sharp distribution of the timescale. This distribution corresponds to the initial condition of $\gamma_i$ given by the uniform distribution (Fig. 4.3A and 4.3C). For both the target and non-target genes, distribution of the timescale shows several sharp peaks. These peaks are due to self-feedback regulation that is explained later in the mechanism section.

With the progress in evolution, the distributions are broader due to the divergence of $\gamma_i$ from the initial condition. This spread of distribution occurs even when there is no fitness is imposed so that this is not due to directed evolution.However, by comparing Fig.4.3B and 4.3D, after evolution, it is clear that the distribution of timescale of non-target genes is broader than that of target genes. In particular, the distribution is broader at the timescale of input change larger than 10. The timescale separation in genes are thus observed.

**Figure 4.3. Distribution of the timescale of input to each gene, over genes and simulation runs**
Distribution of timescale of input change for both target and non-target genes, at the initial and final generation (i.e., when the maximal fitness is reached). The horizontal axis is logarithmic scale of timescale of input to each gene and the vertical axis is its frequency. (A)(C)The distribution at the generation before evolution. (B)(D) The distribution at the generation when the fitness reaches the maximal value 8. The distribution of the target (A)(B) and non-target (C)(D) genes.

### Slow gene as a minority

So far, we confirmed that the separation of timescale of input is a general trend. Inputs to genes that are not targets change in a slower time scale. However, the peak positions of distributions between Fig.4.3 (C) and (D) do not show significant difference between target and non- target genes. However, the frequency distribution of inputs to non-target genes is extended to slower timescale. For example those with the values larger than 10 is about 20% for non-targets, while they are rare for target genes. Then, in Fig.4.4, plotted are the distribution of the largest timescale of input change for each simulation. This largest one is always supported by the non-target genes, As seen, the cases in which the largest timescale is less than 10 is very rare, (about less than 5%), while the time-scale for target genes are less than 10 for most cases (and typically one or two). Hence the time-scale separation is achieved with the evolution. Recall that in Fig. 4.2, the average time-scale ratio is larger than 2, when the maximal fitness is achieved. In Fig 4.4., the time scale of the slowest gene is typically 20 or more (ranging to hundred). Hence the slowest gene typically has a time-scale of 10 times longer than the averaged time scale. Fig.4.4 suggests that although they may be not so common among total gene interactions, in most cases, there exist a few genes that have a much longer timescale than that of majority.

**Figure 4.4. Distribution of largest timescale of input change**
Distribution of the largest timescale of input change for each simulation. The horizontal axis gives the timescale of input change in a logarithmic scale while the vertical axis gives the frequency.

### Developmental mechanism of slow genes

There are two mechanisms for making slow gene expression that are as same as those found in chapter 2. Slow gene emerges if $\gamma_i$, the timescale of gene $i$ is small, or the case when $\sum J_{i,j}X_j - \theta_i \sim 0$, the input to gene $i$ is nearly equal to the expression threshold of gene $i$. Slow genes that give input with longer timescale to other genes emerge due to random mutation on $\gamma_i$, or gene interaction $J_{i,j}$. [1] In the results in chapter 2, slow gene was important to make order in gene expression, for proper pattern formation. This is also the case here.

Similarly to the result of chapter 2, the timescale of input of target genes is shorter than that of non-target genes after the spontaneous timescale separation(Fig. 4.2C and 4.2D). We checked thousands of GRNs for locations of slow genes and encoded dynamics within them. Such slow gene expression controls order of expression as in the case in chapter 2. Fig.4.5 and Fig. 4.6 demonstrates a mechanism of cell type control by slow gene. In both figures, the adopted GRN is common. We assume that external morphogene gradients are induced from different directions(see Methods) . Input0 and input1 in GRN denotes such morphogene (with spatial gradient). We assume here that the effect of inhibition form input 0 in Fig 4.6 is stronger than that in Fig 4.5. Although network structure shown here is a bit complicated, the dynamics that encode is simple. In Fig.4.5, due to weaker inhibition from input0, nullcline of X shifts slightly below in the initial condition(Fig4.5A). Then, as the gene expression level of slow genes is increased, nullcline of X is shifted in the underneath direction further, which leads to saddle-node bifurcation in Fig. 4.5C. After saddle-node bifurcation, the cell state flows on the nullcline of Y so that terminal cell state is $(X, Y) = (1, 0)$ as is shown in Fig.4.5E. In Fig. 4.6, the inhibition of $X$ by input0 is stronger than that in Fig.4.5, so that nullcline of X is slightly up-shifted compared with Fig.4.5A. As development goes, slow gene expression increase, however, nullcline of $X$ is not shifted further to make saddle-node bifurcation. Then as gene expression of slow gene increases much more, gene A is activated(Fig.4.6E). Then activated by gene A, nullcline of $Y$ is shifted in lateral drastically, which brings a different saddle-node bifurcation. After that cell state flows in the phase space to settle in newly born attractor $(X, Y) = (1, 1)$.

Slow gene expression with external morphogene gradient can control different cell types at different location in the manner shown here. Once a network structure is achieved for this mechanism to work, further alteration of network do not needed. Just timescale separation of slow gene from gene A is needed. If slow gene and gene A has the same timescale, nullclines are shifted both for $X$ and $Y$, so that there would be only one attractor $(X, Y) = (1, 1)$. Recall, during evolutionary simulation, $\theta_i$, the threshold of

[1] There are also some minor changes of these mechanisms that involve self-feedback gene regulation. When there is a positive feedback loop on gene $i$ and $\sum J_{i,j}X_j - \theta_i \sim 0$, the input value for gene $i$ gradually increases. Timescale of this gradual increase is different than the ordinary case without self-feedback. Similarly, when there is a negative feedback and $\sum J_{i,j}X_j - \theta_i$ is not large enough, different time-scales emerge. Peaks in Fig.4.2A and 4.2C are due to such self-feedbacks.

**Figure 4.5. Mechanism of cell type control by a slow gene I**
Phase space analysis of the mechanism that is relevant to cell type control by slow genes for the GRN shown in the left column. (A)-(E) is the time laps of phase space with the horizontal axis is the expression of X and the vertical axis is that of Y. (A) represents the earliest phase space and (E) represents the terminate state. Blue line is the nullcline of X and orange line is the nullcline of Y. Cell state is shown by a green circle. It follows flow in the gene expression dynamics that is denoted by arrows in the phase space. Fitness of GRN after gene deletion is shown. The horizontal axis is the rescaled evolutionary time and the vertical axis is the ratio of fitness after deletion to the fitness before deletion. The result of deletion of genes with largest timescale for each simulation is plotted in blue square, while the deletion of genes with smallest timescales is plotted in green circle.

**Figure 4.6. Mechanism of cell type control by a slow gene II**
Phase space analysis of the same mechanism in Fig. 4.5, at different location. Time laps of phase spaces are represented in the same way as Fig. 4.5. The difference is on a initial condition that is represented in (A). Nullcline of X shift slightly avobe than that in Fig.4.5A due to input change(location change).

gene expression is not altered, so that now gene expression of feedforward loop do not much control cell types. It is natural for this gradient and slow-gene-expression type cell type control is achieved and thus slow gene expression naturally emerges.

In summary, with location information by morphogene gradient, slow gene can determine order of gene expression, i.e., which shift of nullclines in the phase space would bring bifurcation of an attractor. Once the order is determined, under self positive-feedback loop, the cell state is fixed. This mechanism does not work without timescale separation. We checked this by numerically changing each gene $\gamma_i$ for evolved population to initial uniform state. Some cellular state is then altered by the change of time-scale of slow gene, leading to loss of fitness.

### Numerical knockout experiment of genes

Then the question whether such slow genes are important for controlling gene expression or not is tested. We numerically knockout each gene in GRN. For each knockout experiment, all incoming/going-out edges to a given gene are deleted and then check the fitness change by such GRN without that gene.

Example of such deletion experiment is shown in Fig. 4.7. Since gene 0 and gene 1

**Figure 4.7. Example of gene deletion experiment**
Gene expression dynamics of the represented network is shown with developmental time as the horizontal axis and cellular index as the vertical axis. Pattern formation dynamics of GRN which loses gene 5(A), gene 7(B), gene 8(C) and gene 9(D) are shown together.

are maternal morphogene gradients and gene 2-4 is target genes, knockout of these genes

results in large damage in expression patterns for most cases. On the other hand, importance of genes without such role i.e., gene 5-9, can vary on each case. Some genes have less/more importance on gene expression control than others. For example, in Fig.4.7, deletion of gene 7 or gene 9 do not alter much output gene expression patterns, while the knockout of gene 8 brings a drastic change in gene expression pattern.

We checked such variation on the importance of genes for controlling gene expression dynamics and their relationship to the time-scale. At each generation in the evolution, we performed the knock-out for the slowest and fastest genes (that are neither target nor input). The ratio of the fitness after knockout to that before the knockout is plotted in Fig. 4.8. In the figure there it is clearly shown that at every time point in evolution, knockout mutants of slowly controlled gene have much smaller fitness than the knockout mutants of genes controlled by faster genes. Slow gene is more important for gene expression control than other genes. It should be also noted that even in the very beginning of evolution, knockout of slow genes decrease the fitness to a larger extent than knockout of faster genes. The ratio of the fitness reduction by slow genes to target genes remains roughly at the same level, which may suggest that responsible gene regulation pathway by slow genes to make certain gene expression patterning is acquired in the beginning of evolution and then preserved. Slow genes, in this sense, are drivers of later evolution.

**Figure 4.8. Fitness decrease with numerical knockout experiment**
Fitness of GRN after numerical knockout experiment is shown. The horizontal axis is the rescaled evolutionary time course and the vertical axis is the ratio of fitness after deletion to the fitness before deletion. The result of deletion of genes with largest input timescale for each simulation is plotted in blue square, while the deletion of genes with smallest input timescales is plotted in green circle.

## 4.4 Discussion

### Slow gene control as minority control

The emergence of slow control genes is a general trend observed in all evolutionary simulation results. Slow control gene expressions have strong positive correlation with the fitness, so that the increases in its time-scale and the fitness show good correspondence. Genes with slow expression changes play important roles for the cell-type differentiation. Numerical knockout of input from such slow gene shows a larger change in gene expression. Slow control genes work as a bifurcation parameter for the regulation in the evolved GRN, even though (or because) the number of such slow genes is few.

Control of a system by a minority component was proposed as minority control(Kaneko and Yomo, 2002, Kaneko, 2007b), in which origin of molecules that carry genetic information was discussed accordingly. This discussion can be extended so that minority control of slow components will be relevant in evolution. In the early stage of evolution, all genes have similar roles for the evolvability. Then as evolution goes, interaction between genes become more complex, then there will appear several cores in a reaction network that has more interactions between other genes. Mutation on that genes affects all other genes that have interaction between the core gene so that mutation at the core genes influence to all genes that have interaction with cores. The core genes take up high evolvability than other genes in that sense. So that evolution is accelerated by having such small number of core genes. ,

We checked that slow gene control is minority among total genetic interactions, and has an important role to the fitness. Mutation of gene interaction pathway from such slow gene brings a larger alteration of gene expression. Hence it is expected that such slow genes work as minority controllers for gene expression. Furthermore, as the time scale is larger and separated than others, its change can influence globally on the average. (If the time scale is similar, its influence could be easily averaged out). In fact the change to the gene leads to bifurcation in gene expression dynamics. Hence, slow minority gene is a consequence of evolution but also works as a driver for evolution. This may be the reason why such slow control is observed in the present chapter as well as in chapter 2.

### Evolution-Development congruence

In this simulation, we cannot find good evolution-development congruence as observed in chapter 2. Evolution-development congruence might be directed by the fitness condition for the target pattern depending on the absolute location. Fitness condition with the target pattern at absolute location induces constraints on the gene expression so that once a particular gene expression process is fixed to make a gene expression of a part of target pattern, then it is hard to alter it. Unlike such fitness condition in chapter

2, the fitness condition in this chapter has no such constraint so that genetic regulation once acquired could be changed and reused to form other part of gene expression, so that there are less purifying selection on gene regulation pathways compared with the fitness condition of chapter 2. The one may expect that with more complex fitness condition that brings hierarchical structure of gene interaction, evo-devo congruence may emerge.

At the moment we have not uncovered the mathematical framework to characterize such constraint condition. In contrast emergence of slow gene control could be much more general and valid among different fitness conditions. Phenomena in evolution of pattern formation can be seen as a process of forming a certain structure in GRN, which is determined by a certain fitness condition, where the process will be generally guided by slow gene control.

# Chapter 5

# Summary and Future issues

### Summary

The possible relationship between developmental dynamics to shape phenotypes and evolution in genotypes has been the focus of the evolutionary developmental biology. Recapitulation was one of the oldest such possible relationship that was never proven. We investigated the possibility for recapitulation to take place at gene expression level and analyzed the mechanism of recapitulation by dynamical systems and systems biological approach.

First, we checked whether recapitulation would take place at gene expression level in chapter 2. With extensive computational simulation on pattern formation, we found both pattern formation dynamics in development and evolution progress in succession with epochs. Such epochs often correspond between evolution and development which are the realizations of recapitulation at gene expression level we sought. We coined the phenomena as "evolution-development congruence".

Analyzing developmental control in time and space, we uncovered the mechanisms for evolution-development congruence. Developmental control in time is implemented as slow gene expression. In development, actually working network is expanded in succession at each epoch. Such expansion of network corresponds to what had happened during evolution from the ancestral network. Considering such slow control gene as external parameter, development can be regarded as a sequence of bifurcations according to the parameter change. The evolution-development congruence is then interpreted as correspondence of bifurcations between evolution and development. Here slow gene control spontaneously emerges during evolution, implying that slow gene expression is the result of adaptation.

Developmental control in space is implemented in GRN structures and the resultant dynamics. We found a novel class of pattern formation that is induced by diffusive inter-

action and fixed boundary condition. The acquisition of the dynamics in evolution can be divided into two parts: first, the acquisition of boundary-making pattern formation dynamics, then the acquisition of oscillatory dynamics that change into periodic stripes with the diffusive interaction. The acquisition order would not be reversed. Also, expression order is often as same as the acquisition order since the boundary condition has to be set earlier than the start of oscillation. Pattern formation dynamics thus has tendency to recapitulate that underlie in the evolution-development congruence we found here.

Few exceptional examples in the collapse of evo-devo congruence is also explained by this pattern formation dynamics. When boundary condition is altered by mutation on GRN, it can results in total reformation of pattern formation.

The mathematical analysis on the dynamics to transform oscillatiroy expression into spatial pattern is executed in chapter 3, where we can predict the emergent spatial pattern by spatial map derived from discretization of stationary condition.

Generality of the emergence of slow gene and their relationship to evolution are investigated in chapter 4, with a simpler model than that of chapter 2. There, the emergence of slow gene strongly correlates with the gain in fitness. Also, numerical knockout experiment of such slow gene results in higher damage to pattern formation dynamics than other faster genes.We also checked that slow genes are minority among total genes in GRN, and slow minority control as discussed in the study of the origin of genetic informationKaneko and Yomo (2002), Kaneko (2007b) where, those components that is minority in a system carry information of the system. These results suggest the separation of timescales in gene expressions is relevant to evolution.

The main results of this thesis are summarized as follows:

1. **Epochs of development as bifurcation by slow expression dynamics**

2. **Punctuated equilibrium in the evolution of morphology as bifurcation**

3. **Evo-devo congruence through common bifurcations**

4. **The combination of feedforward and feedback gene regulation networks to support developmental epochs**

5. **Violation of evo-devo correspondence through modification of upstream feedforward regulation under downstream feedback mechanism**

6. **Minority slow genes take control of pattern formation**

## Discussion

### Mechanism of developmental burden

In the chapter 1, we introduced hypothesis on recapitulation as "developmental burden"(Riedl, 1978). Developmental burden assumes that the organs that formed through each developmental process are regarded to give positional or temporal information that induces ensuing processes. The order of preceding developmental processes and following processes were never altered since the following processes cannot work alone; they need foundation laid by the preceding processes, and thus the recapitulation take place. There is no explicit way to implement the developmental burden in the structure of GRN itself. However, now we can see that the dynamics we found in chapter 2 and that are also analyzed in chapter 3 provides an implementation of developmental burden. There, the upstream feedforward mechanism gives boundary condition(locational information) for the following downstream oscillatory mechanism while this order of the mechanisms is determined by slow gene expression. This interplay of upstream feedforward mechanism for boundary condition and downstream feedback for oscillation with the aid of slow gene gives an implementation of developmental burden.

### Why recapitulation could be observed?

Throughout chapter 2, we looked for recapitulation in gene expression level and a possible mechanism that underlies it. In the model, evolution progress in a step-wise manner that corresponds to the view of "punctuated equilibrium". It is more remarkable that development also progresses in a step-wise manner, which is due to slow gene control as we have seen. Slow gene controls successive expansion of working GRN and the GRN expansion during development occurs in a similar manner with that through the evolution. This result means, once a way to gain a higher fitness is acquired, it is hardly lost during evolution. Evolution goes in the way that conserves a fitter ancestral GRN structure(with its regulating dynamics) and then adds beneficial mutations to it. Clearly, this type of evolution occurs under strong purifying selection.

Here, slow genes can work both as a stabilizer of pattern formation dynamics of ancestral origin and a primer for beneficial mutations. We showed that GRN that acquires slow gene control by evolution works for later epochs in development. That means, after ancestral GRN are fully expressed, then the subnetwork in GRN by newly added mutations works. In the term of dynamical systems, this corresponds to removal of unnecessary perturbations for the ancestral system to work. In other words, with slow gene control,

evolved network can easily reproduce in a transient at least some ancestral pattern. Accordingly, if newly added mutation is neutral, the fitness equivalent with the ancestor is guaranteed. After reproduction of ancestral pattern formation, then newly added mutations start to work. Again, without slow gene control, such newly added mutations can bring novel gene expression pattern at the earliest stage of development, which would be the perturbation to the ancestral system and result in a different sequence in pattern formation that does not guarantee fitness equivalent as the ancestor.

Our model in chapter 2 was good in both inducing purifying selection and emergence of slow gene control. In other words, in the fitness condition in which most of mutations are neutral, there would not be recapitulation since there are no reason to conserve the ancestral network structure. Previous studies set fitness condition that might induce weaker purifying selection such as the number of stripes(Salazar-Ciudad et al., 2001b,a, François et al., 2007, Fujimoto et al., 2008, ten Tusscher and Hogeweg, 2011). This might be the reason why the recapitulation was not observed in their study. Indeed, the model in chapter 4 did not show the recapitulation. Model setting, especially the fitness condition might be critical for recapitulation to take place. The key for recapitulation might be whether fitness condition can induce strong purifying selection or not. However there is no clear indicator of purifying selection for the fitness condition. Further studies must be needed.

### Future Issues

The present studies are initial steps towards resolving the larger issue of evo-devo correspondence. Even within the present model, a number of issues remain to be clarified, as follows:

(i) Even though we have confirmed that our result is independent of the details in the model, such as the cell number, gene number, model parameters, and the form of the external morphogen profile, further study is necessary to confirm the universal applicability of our results. In particular, dependence on the fitness condition is of interest. In the earlier theoretical studies on the evolution of segmentation, the evo-devo congruence was not noted explicitly (Salazar-Ciudad et al., 2001b,a, François et al., 2007, Fujimoto et al., 2008, ten Tusscher and Hogeweg, 2011). In these studies, fitness condition does not include the information on the absolute location of segments, as, for example, given by the number of segments or cell types. Indeed, we found that the congruence is much worse when just the number of stripes is adopted as fitness function. In our case, the fitness given as a function of fixed (non-periodic) spatial pattern. How much does the evo-devo congruence depend on the strength of purifying selection?

(ii) In the gene expression dynamics after the evolution, we found that there always exists a slowly varying gene expression level that works as a control parameter. Thus far, we have not uncovered the conditions responsible for the emergence of this slower mode, which controls other expressions that are relevant to fitness. Both in chapter 2 and 4, by modifying coupling with this slower mode, the output behavior that it controls is more readily altered, so that evolution can be facilitated. In chapter 4, degree of timescale separation between target and non-target genes increases with strong correlation to fitness. Moreover, numerical knockout of slow gene expression results stronger influence on pattern formation than other genes. Therefore, the emergence of slowly varying expression(s) may be evolutionary advantageous and developmentally important. Further investigation of the generality of the emergence of this slow variable and its function on other models will be needed.

(iii)We have not observed the classic Turing mechanism (Turing, 1952) in the developmental process by evolved networks in chapter 2. Under the influence of a morphogen gradient, it may be natural to use the maternal information effectively with respect to evolution. Instead of Turing mechanism, pattern formation from oscillatory gene expression under fixed boundary condition and diffusive interaction is commonly observed, as analyzed in chapter 3. However, in both condition, there assumed no noise in gene expressions. It is then still an open question whether without the external information (but by imposing the boundary condition instead), the Turing-pattern mechanism can evolve dominantly. under noise in gene expression.
(iv)How do evolutionary reachability of the target and complexity in the developmental process depend on the predefined target pattern? It may be expected that as the target

pattern is more complex, it takes more time to evolve GRN to produce such patterns, and development involves more epochs, but is there a way to quantitatively characterize such complexity?

(v) We have explained evolution-development congruence in terms of the correspondence of epochs, but it is not clear whether quantitative congruence exists beyond this level. For example, does the time span for the quasi-stationary regime between two epochs correlate between development and evolution? In other words, if the evolutionary search time to generate a relevant mutation for the next epoch is longer, then, is the quasi-stationary regime before the epoch also longer in development? Our preliminary results suggest that this correlation exists for cases where small $\Delta$ (in chapter 2) values are observed, while further analysis is required to clarify the conditions and mechanisms for such congruence.

(vi) Evolved pattern in chapter 2 was made by GRN with hierarchic structure, i.e, feedback structure in the downstream of feedforward structure. Feedback structure can generate gene expression pattern next to the gene expression pattern that generated by upstream feedforword structure. As is analyzed in chapter 3, gene expression generated by the upstream feedforward structure sets boundary condition for the downstream feedback structure. Thus, in the hierarchic structure, once upstream feedforward structure determines the location of gene expression, other parts of GRN are not necessary to determine absolute location of gene expression again, but are needed to determine relative location to the gene expression regulated by the upstream structure. In other words, gene expression first generated by the upstream feedforward structure acts as a basis for following pattern formation. Gene expression pattern governed by the downstream feedback structure could serve boundary condition of further downstream feedback structure. This further downstream feedback structures also consult for its expression location on upmost feedforward structure. Thus if upmost feedforward structure were altered, following gene expression regulated by downstream structure would also be altered, which might bring collapse of total patterning. This increasing duty of upmost feedforward structure and gene expression control resembles the concept of developmental burden as introduced in the end of chapter 1.1. This framework of feedforward to feedback structure and hierarchical gene expression can be tested a theoretical model for developmental burden. Emergence of developmental burden through evolution and its quantitative analysis will be an important issue for future studies

(vii)Extension of our model for higher spatial dimensions, introduction of size (cell number) growth through development, and inclusion of recombination in a genetic algorithm will be important in future studies. Furthermore, it is important to note that real morphogenesis in multi-cellular organisms is far more complex than these models, and is not necessarily governed only by the reaction-diffusion mechanism. Cell rearrangement under mechanical stress could also play an important role, and inclusion of development mechanisms will be important. Still, we also note that previous research has indicated that macroscopically represented, stress-induced pattern formation can also be represented by

equations of the reaction-diffusion type (Murray, 2002). Hence, the present conclusions on evo-devo congruence through bifurcations may be applicable beyond development based on a reaction-diffusion system.

(viii)Last but not least, the implications of our single-chain-phylogeny study on species-wide comparison have to be explored. For example, by adding population division and speciation process with imposing different target pattern to the model, species-wide extension will be available. That future study will be important not only for the validation of our results, but also for further understanding of evo-devo relationship in species-wide comparison.

## Concluding remark

In contrast to recent advances in experiments aimed towards analyzing the evo-devo relationship at a quantitative level, theoretical studies based on dynamical systems and statistical physics are still in their infancy. While we acknowledge that our current model may be oversimplified, we hope that the present work can act as a springboard to launch future cooperative efforts in the field of evo-devo between theories and experiments.

# Acknowledgement

This thesis is absolutely the result of warm supports and guidances from persons concerned inside and outside the laboratory.

First and foremost, I am deeply grateful to my supervisor professor Dr. Kunihiko Kaneko for the continuous support of my Ph.D research, for his patience, motivation, and immense knowledge. I appreciate all his contributions of time, ideas and funding to make my Ph.D. It is said that "the great teacher inspires". I do not know if he did consciously or unconsciously, every time I met and discussed with him, I did get inspired and got some confidence in my work that was my motivation. Being his student is my pride and days I spent studying with him are just priceless in my life.

Besides my supervisor, I would also like to thank assistant professors in the laboratory, Dr. Tetsuhiro Hatakeyama, Dr. Atsushi Kamimura, Dr. Nen Saito, Dr. Nobuto Takeuchi and Dr. Shuji Ishihara for their insightful comments, but also for critical questions which motivated me to widen my research from various perspectives. Off the laboratory, to me, they are caring lab seniors. It was really helpful that they showed me how good and hard being as an young researcher.

Associate professor Dr. Naoki Irie always giving me not only insightful but also encouraging comments to my work, that also enhanced my motivation. I would like to appreciate.

All my work was done by computation system of the laboratory. I would like to thank all the laboratory system administrators for every day maintenance of the system.

I thank my fellow labmates, Hironobu Noguchi, Yusuke Himeoka, Yoshiya Matsubara, Kenji Shinoda, Shunpei Yamamoto, Yuuma Fujimoto and Jonathan Young for the stimulating discussions and for all the fun I have had in the last three years. I enjoyed my graduate school life with good colleagues. I would also like to thank lab secretary Motoko Miura, for always caring me and did cumbersome administrative procedures for me, many times.

I would like to thank my best friends, Michiko Kurikawa, Takane Saito, Akihiro Yoshimura and Taro Yoshida for accepting the way I am. From my undergraduate school days, they relieved me a lot especially when I was nervous and not in a good shape. It is really heartbreaking those days never come back. R.I.P., Takane. You are my best friend.

Lastly, I would like to thank my family: my parents and to my sister and brother for supporting me spiritually throughout writing this thesis and my life in general.

# References


Alon, U. (2007). Introduction to Systems Biology: And the Design Principles of Biological Networks, volume 10. CRC press. pages 5, 15, 34

Ancel, L. W., Fontana, W., et al. (2000). Plasticity, evolvability, and modularity in rna. Journal of Experimental Zoology, 288(3):242–283. pages 45

Aubry, S. (1978). The new concept of transitions by breaking of analyticity in a crystallographic model. In Solitons and Condensed Matter Physics, pages 264–277. Springer. pages 66

Baurmann, M., Gross, T., and Feudel, U. (2007). Instabilities in spatially extended predator–prey systems: spatio-temporal patterns in the neighborhood of turing–hopf bifurcations. Journal of Theoretical Biology, 245(2):220–229. pages 60

Carrol, S. B. (1995). Homeotic genes and the evolution of arthropods and chordates. Nature, 376(6540):479–485. pages 5

Carroll, S. B., Grenier, J., and Weatherbee, S. (2009). From DNA to diversity: molecular genetics and the evolution of animal design. Wiley. com. pages 5, 50, 52

Chance, B., Pye, K., and Higgins, J. (1967). Waveform generation by enzymatic oscillators. Spectrum, IEEE, 4(8):79–86. pages 52

Ciliberti, S., Martin, O. C., and Wagner, A. (2007). Robustness can evolve gradually in complex regulatory gene networks with varying topology. PLoS Comput Biol, 3(2):e15. pages 45

Cohn, M. J. and Tickle, C. (1999). Developmental basis of limblessness and axial patterning in snakes. Nature, 399(6735):474–479. pages 1

Cooke, J. (1975). Control of somite number during morphogenesis of a vertebrate, xenopus laevis. pages 71

Cooke, J. and Zeeman, E. C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. Journal of theoretical biology, 58(2):455–476. pages 8, 9, 52, 61

Cotterell, J., Robert-Moreno, A., and Sharpe, J. (2015). A local, self-organizing reaction-diffusion model can explain somite patterning in embryos. Cell Systems, 1(4):257–269. pages 61, 71

Cotterell, J. and Sharpe, J. (2010). An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. Molecular systems biology, 6(1):425. pages 34

Cross, M. C. and Hohenberg, P. C. (1993). Pattern formation outside of equilibrium. Reviews of modern physics, 65(3):851. pages 60

Darwin, C. (1859). On the origin of the species by natural selection. pages 2, 12

Davidson, E. H. and Erwin, D. H. (2006). Gene regulatory networks and the evolution of animal body plans. Science, 311(5762):796–800. pages 5

Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.-H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., et al. (2002). A genomic regulatory network for development. science, 295(5560):1669–1678. pages 5

De Wit, A., Lima, D., Dewel, G., and Borckmans, P. (1996). Spatiotemporal dynamics near a codimension-two point. Phys. Rev. E, 54:261–271. pages 36, 60

Dias, A. S., de Almeida, I., Belmonte, J. M., Glazier, J. A., and Stern, C. D. (2014). Somites without a clock. Science, 343(6172):791–795. pages 61

Domazet-Lošo, T. and Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. Nature, 468(7325):815–8. pages 3, 12

Duboule, D. (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate bauplan and the evolution of morphologies through heterochrony. Development, 1994(Supplement):135–142. pages 3, 12

Eldredge, N. and Gould, S. J. (1972). Punctuated equilibria: an alternative to phyletic gradualism. Models in paleobiology, 82:115. pages 29, 51

Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature, 403(6767):335–338. pages 5

François, P., Hakim, V., and Siggia, E. D. (2007). Deriving structure from evolution: metazoan segmentation. Mol Syst Biol, 3:154. pages 7, 13, 34, 98, 99

François, P. and Siggia, E. D. (2010). Predicting embryonic patterning using mutual entropy fitness and in silico evolution. Development, 137(14):2385–2395. pages 8, 49

Francois, P. and Siggia, E. D. (2012). Phenotypic models of evolution and development: geometry as destiny. Current opinion in genetics & development, 22(6):627–633. pages 6

François, P. and Siggia, E. D. (2012). Phenotypic models of evolution and development: geometry as destiny. Curr Opin Genet Dev, 22(6):627–33. pages 51

Freitas, R., Gómez-Marín, C., Wilson, J. M., Casares, F., and Gómez-Skarmeta, J. L. (2012). Hoxd13 contribution to the evolution of vertebrate appendages. Developmental cell, 23(6):1219–1229. pages 51

Fujimoto, K., Ishihara, S., and Kaneko, K. (2008). Network evolution of body plans. PLoS ONE, 3(7):e2772. pages 7, 13, 34, 71, 98, 99

Fujimoto, K. and Kaneko, K. (2001). Sensitive boundary condition dependence of noise-sustained structure. Physical Review E, 63(3):036218. pages 60

Gerhart, J. and Kirschner, M. (1997). Cells, embryos, and evolution: Toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability. Number 575.21 GER. Blackwell Science Malden. pages 5

Gierer, A. and Meinhardt, H. (1972). A theory of biological pattern formation. Kybernetik, 12(1):30–39. pages 6

Gilbert, S. (2014). Developmental Biology. Sinauer. pages 1

Glass, L. and Kauffman, S. A. (1973). The logical analysis of continuous, non-linear biochemical control networks. Journal of Theoretical Biology, 39(1):103–129. pages 18, 81

Goodwin, B. C. et al. (1963). Temporal organization in cells. a dynamic theory of cellular control processes. Temporal organization in cells. A dynamic theory of cellular control processes. pages 52

Goto, Y. and Kaneko, K. (2013). Minimal model for stem-cell differentiation. Phys. Rev. E, 88:032718. pages 36, 62

Gould, S. J. (1977). Ontogeny and phylogeny. Harvard University Press. pages 2, 3, 12

Haeckel, E. (1866). Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie. Allgemeine Anatomie der Organismen. Number 1 in Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie. Allgemeine Anatomie der Organismen. Reimer. pages 3, 12

Hall, B. K. (1999). Evolutionary developmental biology. Springer. pages 3, 12

Hall, B. K. (2000). Balfour, garstang and de beer: the first century of evolutionary embryology. American Zoologist, 40(5):718–728. pages 12

Hazkani-Covo, E., Wool, D., and Graur, D. (2005). In search of the vertebrate phylotypic stage: a molecular examination of the developmental hourglass model and von baer's third law. J Exp Zool B Mol Dev Evol, 304(2):150–8. pages 3, 12, 51

Hirsch, M. W., Smale, S., and Devaney, R. L. (2012). Differential equations, dynamical systems, and an introduction to chaos. Academic press. pages 15

Horikawa, K., Ishimatsu, K., Yoshimoto, E., Kondo, S., and Takeda, H. (2006). Noise-resistant and synchronized oscillation of the segmentation clock. Nature, 441(7094):719–723. pages 52

Imai, K. S., Levine, M., Satoh, N., and Satou, Y. (2006). Regulatory blueprint for a chordate embryo. Science, 312(5777):1183–1187. pages 5

Irie, N. and Kuratani, S. (2011). Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. Nat Commun, 2:248. pages 3, 12, 51

Irie, N. and Kuratani, S. (2014). The developmental hourglass model: a predictor of the basic body plan? Development, 141(24):4649–4655. pages 3

Irie, N. and Sehara-Fujisawa, A. (2007). The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information. BMC Biology, 5(1):1. pages 3, 12

Ishihara, S., Fujimoto, K., and Shibata, T. (2005). Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. Genes to Cells, 10(11):1025–1038. pages 34, 52

Ishihara, S. and Kaneko, K. (2006). Turing pattern with proportion preservation. Journal of theoretical biology, 238(3):683–693. pages 71

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. Journal of molecular biology, 3(3):318–356. pages 4

Jaeger, J., Irons, D., and Monk, N. (2012). The inheritance of process: a dynamical systems approach. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 318(8):591–612. pages 51

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., et al. (2004). Dynamic control of positional information in the early drosophila embryo. Nature, 430(6997):368–371. pages 34, 52

Kalinka, A. T., Varga, K. M., Gerrard, D. T., Preibisch, S., Corcoran, D. L., Jarrells, J., Ohler, U., Bergman, C. M., and Tomancak, P. (2010). Gene expression divergence recapitulates the developmental hourglass model. Nature, 468(7325):811–4. pages 3, 12, 51

Kaneko, K. (2006). Life: an introduction to complex systems biology. Springer. pages 15, 45

Kaneko, K. (2007a). Evolution of robustness to noise and mutation in gene expression dynamics. PLoS One, 2(5):e434. pages 45

Kaneko, K. (2007b). Question 9: Minority control and genetic takeover. Origins of Life and Evolution of Biospheres, 37(4-5):465–468. pages 93, 96

Kaneko, K. and Yomo, T. (2002). On a kinetic origin of heredity: minority control in a replicating system with mutually catalytic molecules. Journal of theoretical biology, 214(4):563–576. pages 93, 96

Kapral, R. and Showalter, K. (2012). Chemical waves and patterns, volume 10. Springer Science & Business Media. pages 60

Kohsokabe, T. and Kaneko, K. (2016). Evolution-development congruence in pattern formation dynamics: bifurcations in gene expression and regulation of networks structures. Journal of Experimental Zoology Part B:Molecular and Developmental Evolution, 326B(1):61. pages 71

Kondo, S. and Miura, T. (2010). Reaction-diffusion model as a framework for understanding biological pattern formation. science, 329(5999):1616–1620. pages 6

Kuramoto, Y. (2012). Chemical oscillations, waves, and turbulence, volume 19. Springer Science & Business Media. pages 60

Kuratani, S. (2017). Evolutionary Morphology: Bauplan and Embryonic Development of Vertebrates(in Japanese). The University of Tokyo Press. pages 1

Lenfyel, I. and Epstein, I. R. (1991). Modeling of turing structures in the chlorite—iodide—malonic acid—starch reaction system. Science, 251(4994):650–652. pages 60

Levin, M., Hashimshony, T., Wagner, F., and Yanai, I. (2012). Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. Developmental cell, 22(5):1101–1108. pages 3, 12

Masamizu, Y., Ohtsuka, T., Takashima, Y., Nagahara, H., Takenaka, Y., Yoshikawa, K., Okamura, H., and Kageyama, R. (2006). Real-time imaging of the somite segmentation clock: revelation of unstable oscillators in the individual presomitic mesoderm cells. Proceedings of the National Academy of Sciences, 103(5):1313–1318. pages 52

Meinhardt, H. (1982). Models of biological pattern formation, volume 6. Academic Press, London. pages 61

Meixner, M., De Wit, A., Bose, S., and Schöll, E. (1997). Generic spatiotemporal dynamics near codimension-two turing-hopf bifurcations. Physical Review E, 55(6):6690. pages 60

Mjolsness, E., Sharp, D. H., and Reinitz, J. (1991). A connectionist model of development. Journal of theoretical Biology, 152(4):429–453. pages 18, 62, 81

Müller, F. (1869). Facts and arguments for Darwin. John Murray. pages 12

Murray, J. D. (1981). A pre-pattern formation mechanism for animal coat markings. Journal of Theoretical Biology, 88(1):161–199. pages 6

Murray, J. D. (2002). Mathematical biology, volume 2. springer. pages 15, 101

Nicolis, G. and Prigogine, I. (1977). Self-organization in nonequilibrium systems. John Wiley & Sons, New York. pages 60, 68

Onimaru, K., Kuraku, S., Takagi, W., Hyodo, S., Sharpe, J., and Tanaka, M. (2015). A shift in anterior–posterior positional information underlies the fin-to-limb evolution. eLife, 4:e07048. pages 6

Palmeirim, I., Dubrulle, J., Henrique, D., Ish-Horowicz, D., and Pourquié, O. (1998). Uncoupling segmentation and somitogenesis in the chick presomitic mesoderm. Developmental genetics, 23(1):77–85. pages 60, 71

Pomeau, Y. and Zaleski, S. (1981). Wavelength selection in one-dimensional cellular structures. Journal de Physique, 42(4):515–528. pages 60

Pourquie, O. (2001). The vertebrate segmentation clock. Journal of anatomy, 199(1-2):169–175. pages 60, 71

Pourquié, O. (2003). The segmentation clock: converting embryonic time into spatial pattern. Science, 301(5631):328–330. pages 52

Quint, M., Drost, H.-G., Gabel, A., Ullrich, K. K., Bönn, M., and Grosse, I. (2012). A transcriptomic hourglass in plant embryogenesis. Nature, 490(7418):98–101. pages 3, 12, 51

Raff, R. A. and Wolpert, L. (1996). The shape of life-genes, development and evolution of animal forms. Genetical Research, 68(3):261. pages 3, 12

Richardson, M. K. and Keuck, G. (2002). Haeckel's abc of evolution and development. Biological Reviews, 77(4):495–528. pages 12

Riedl, R. (1978). Order in living organisms: a systems analysis of evolution. John Wiley & Sons. pages 4, 97

Salazar-Ciudad, I., Newman, S. A., and Solé, R. V. (2001a). Phenotypic and dynamical transitions in model genetic networks. i. emergence of patterns and genotype-phenotype relationships. Evol Dev, 3(2):84–94. pages 7, 13, 18, 34, 81, 98, 99

Salazar-Ciudad, I., Solé, R. V., and Newman, S. A. (2001b). Phenotypic and dynamical transitions in model genetic networks. ii. application to the evolution of segmentation mechanisms. Evol Dev, 3(2):95–103. pages 7, 13, 18, 34, 81, 98, 99

Sander, K. et al. (1983). The evolution of patterning mechanisms: gleanings from insect embryogenesis and spermatogenesis. Development and evolution, pages 137–159. pages 3, 12

Sander, K. and Schmidt-Ott, U. (2004). Evo-devo aspects of classical and molecular data in a historical perspective. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 302(1):69–91. pages 2, 52

Sato, K. (2016). Haeckel and his dream of evolution(in Japanese). Kousakusha. pages 3

Segel, L. A. and Jackson, J. L. (1972). Dissipative structure: an explanation and an ecological example. Journal of Theoretical Biology, 37(3):545–559. pages 60

Soyer, O. S. (2012). Evolutionary systems biology, volume 751. Springer Science & Business Media. pages 45

Spitz, F. and Furlong, E. E. (2012). Transcription factors: from enhancer binding to developmental control. Nature Reviews Genetics, 13(9):613–626. pages 4

Stathopoulos, A. and Levine, M. (2005). Genomic regulatory networks and animal development. Developmental cell, 9(4):449–462. pages 5

Strogatz, S. H. and Balachandran, B. (1994). Nonlinear dynamics and chaos with applications to biology, chemistry and engineering. pages 15

ten Tusscher, K. (2013). Mechanisms and constraints shaping the evolution of body plan segmentation. The European Physical Journal E, 36(5):1–12. pages 7

ten Tusscher, K. H. and Hogeweg, P. (2011). Evolution of networks for body plan patterning; interplay of modularity, robustness and evolvability. PLoS Comput Biol, 7(10):e1002208. pages 7, 13, 34, 98, 99

Turing, A. M. (1952). The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 237(641):37–72. pages 5, 36, 60, 61, 99

von Baer, K. E. (1828). Ueber entwicklungsgeschichte der thiere. Beobachtung und Reflexion, 3. pages 2, 12

von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. (2000). The segment polarity network is a robust developmental module. Nature, 406(6792):188–192. pages 34, 52

Waddington, C. H. (1957). The strategy of the genes. A discussion of some aspects of theoretical biology. Allen & Unwin. pages 45

Wagner, A. (2005). Robustness, evolvability, and neutrality. FEBS letters, 579(8):1772–1778. pages 45

Walgraef, D. (1997). Spatio-temporal Pattern Formation with Examples from Physics, Chemistry and Materials Science. Springer. pages 60

Wang, Z., Pascual-Anaya, J., Zadissa, A., Li, W., Niimura, Y., Huang, Z., Li, C., White, S., Xiong, Z., Fang, D., Wang, B., Ming, Y., Chen, Y., Zheng, Y., Kuraku, S., Pignatelli, M., Herrero, J., Beal, K., Nozawa, M., Li, Q., Wang, J., Zhang, H., Yu, L., Shigenobu, S., Wang, J., Liu, J., Flicek, P., Searle, S., Wang, J., Kuratani, S., Yin, Y., Aken, B., Zhang, G., and Irie, N. (2013). The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. Nat Genet, 45(6):701–6. pages 3, 12, 51

Willeboordse, F. H. and Kaneko, K. (1995). Pattern dynamics of a coupled map lattice for open flow. Physica D: Nonlinear Phenomena, 86(3):428–455. pages 66

Winfree, A. T. (2001). The geometry of biological time, volume 12. Springer Science & Business Media. pages 60

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. Journal of theoretical biology, 25(1):1–47. pages 5

*Page number indicates the pages where the article are cited in this thesis