

Binghamton University

The Open Repository @ Binghamton (The ORB)

Graduate Dissertations and Theses

Dissertations, Theses and Capstones

5-10-2018

The role of criticality of gene regulatory networks on emergent properties of biological systems

Hyobin Kim

Binghamton University, hkim240@binghamton.edu

Follow this and additional works at: https://orb.binghamton.edu/dissertation_and_theses



Part of the [Biophysics Commons](#), [Morphology Commons](#), and the [Systems Engineering Commons](#)

Recommended Citation

Kim, Hyobin, "The role of criticality of gene regulatory networks on emergent properties of biological systems" (2018). *Graduate Dissertations and Theses*. 93.
https://orb.binghamton.edu/dissertation_and_theses/93

This Dissertation is brought to you for free and open access by the Dissertations, Theses and Capstones at The Open Repository @ Binghamton (The ORB). It has been accepted for inclusion in Graduate Dissertations and Theses by an authorized administrator of The Open Repository @ Binghamton (The ORB). For more information, please contact ORB@binghamton.edu.

THE ROLE OF CRITICALITY OF GENE REGULATORY NETWORKS ON
EMERGENT PROPERTIES OF BIOLOGICAL SYSTEMS

BY

HYOBIN KIM

BS, Pusan National University, 2009

MS, Korea Advanced Institute of Science and Technology, 2011

DISSERTATION

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Systems Science
in the Graduate School of
Binghamton University
State University of New York
2018

© Copyright by Hyobin Kim 2018

All Rights Reserved

Accepted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Systems Science
in the Graduate School of
Binghamton University
State University of New York
2018

May 10, 2018

Professor Hiroki Sayama, Chair and Faculty Adviser
Department of Systems Science and Industrial Engineering
Binghamton University

Professor Harold W. Lewis, Committee Member
Department of Systems Science and Industrial Engineering
Binghamton University

Professor Prahalad Rao, Committee Member
Department of Mechanical and Materials Engineering
University of Nebraska, Lincoln

Professor Kenneth Chiu, Outside Examiner
Department of Computer Science
Binghamton University

Abstract

The relationship between criticality of gene regulatory networks (GRNs) and dynamics of GRNs at a single cell level has been vigorously studied. However, the relationship between the criticality of GRNs and properties of multicellular organisms at a higher level has not been fully explored. Here we aim at revealing potential roles of the criticality of GRNs at a multicellular and hierarchical level, using a random Boolean network as a GRN. We perform three studies. Firstly, we propose a GRN-based morphogenetic model, and delve into the role of the criticality of GRNs in morphogenesis at a multicellular level. Secondly, we include an evolutionary context in our morphogenetic model by introducing genetic perturbations (e.g., mutations) to GRNs, and examine whether the role of the criticality of GRNs can be maintained even in the presence of the evolutionary perturbations. Also, we look into what the resulting morphologies are like and what kind of biological implications they have from the epigenetic viewpoint in morphology. Lastly, we present multilayer GRNs consisting of an intercellular layer and an intracellular layer. A network in an intercellular layer represents interactions between cells, and a network in an intracellular layer means interactions between genes. All the nodes of an intercellular network have identical intracellular GRNs. We investigate how the criticality of GRNs affects the robustness and evolvability of the multilayer GRNs at a hierarchical level, depending on cellular topologies and the number of links of an intercellular network. From the three studies, we found that the criticality of GRNs facilitated the formation of nontrivial morphologies at a multicellular

level, and generated robust and evolvable multilayer GRNs most frequently at a hierarchical level. Our findings indicate that the roles of the criticality of GRNs are hard to be discovered through the single-cell-level studies. It justifies the value of our research on the relationship between criticality of GRNs and properties of organisms in the context of multicellular settings.

Acknowledgements

First of all, I would like to express my deepest gratitude and respect to my adviser, Professor Hiroki Sayama for his continuous support for my doctoral research. His comments and tips during the period of my doctoral studies stimulated me to study the subjects in more depth, and helped me to grow as an independent researcher. I could not imagine having a greater adviser than him.

Also, I would like to thank my committee members, Professor Harold W. Lewis and Professor Prahalad Rao for their valuable comments and enlightening questions. In addition, I am grateful to my outside examiner, Professor Kenneth Chiu for his insightful feedback.

I would like to thank my fellow labmates and my friends sharing joys and sorrows together as well. Their kind words were great encouragement to me.

Last but not the least, I would like to express my gratitude to my family: my grandmother, my parents and my sister for their support and encouragement. I dedicate my doctoral dissertation to my family.

This material is based upon work supported by the National Science Foundation under Grant No. 1319152.

Table of Contents

List of Tables	ix
List of Figures	x
Chapter 1 Introduction	1
Chapter 2 Related Work	4
2.1 The Relationship Between Criticality of GRNs and Dynamics of GRNs at a Single Cell Level	4
2.2 The Relationship Between Properties of Intracellular GRNs and Properties of Multicellular Organisms	6
Chapter 3 Objective	9
Chapter 4 The Role of Criticality of Gene Regulatory Networks in Morphogenesis.....	10
4.1 Model: GRN-Based Morphogenetic Systems	10
4.1.1 Gene Regulatory Network (GRN)	11
4.1.2 Cell-Cell Interactions	14
4.1.3 Cellular Movements	18
4.2 Experiments	20
4.2.1 Measures for Morphogenetic Pattern Analysis	20
4.3 Results	27
Chapter 5 How the Criticality of Gene Regulatory Networks Affects the Resulting Morphogenesis under Genetic Perturbations.....	33
5.1 Experiments	33
5.1.1 Perturbations to GRNs and Robust & Evolvable GRNs	34
5.1.2 Measures to Investigate the Relationship Between GRNs and Expressed Cell Fates	35
5.2 Results	36
Chapter 6 Robustness and Evolvability of Multilayer Gene Regulatory Networks	44
6.1 Model: Multilayer GRNs	44

6.2 Experiments	49
6.2.1 Robustness and Evolvability Against Genetic Perturbations.....	50
6.3 Results	53
Chapter 7 Conclusions	57
Appendix A.....	61
Notes	63
References.....	64

List of Tables

Table 1. Parameters of the morphogenetic model and their values	21
Table 2. Parameters of the multilayer GRNs and their values	50
Table 3. Correlation coefficients between the properties of multilayer GRNs and intracellular GRNs	55

List of Figures

Figure 1. Simulation algorithm for our GRN-based morphogenetic model	11
Figure 2. Schematic diagrams for an example GRN and its state space	12
Figure 3. An example state space of a GRN where four cell fates are randomly assigned.....	14
Figure 4. Cell signaling for cell-cell interactions	15
Figure 5. An example showing how to calculate the average concentration of the signaling molecules neighboring cells produce and determine the state of receptor <i>gene2</i> of cell <i>i</i>	17
Figure 6. Six possible types of cell fate combinations between two cells	19
Figure 7. Network construction for the analysis of morphologies	22
Figure 8. Nontrivial morphology detection using KL divergence	23
Figure 9. An example showing how to calculate mutual information between cell fates of cells and their neighboring cells	25
Figure 10. An example showing how to calculate homogeneity of sizes of connected components (<i>homoSizeConnComp</i>)	27
Figure 11. Distributions of morphogenetic patterns according to the number of cells for $K = 1, 2, 3, 4$	28
Figure 12. Different morphogenetic patterns represented with networks for $K = 1, 2, 3, 4$	29
Figure 13. Comparison of means between groups ($K = 1, 2, 3, 4$) for 12 morphological measures	30
Figure 14. Colored correlation matrix for 12 morphological measures	31
Figure 15. Schematic diagrams illustrating the concept of a robust and evolvable GRN	34
Figure 16. Frequencies of robust & evolvable GRNs per group ($K = 1, 2, 3, 4$) and visualized spatial patterns	37

Figure 17. Comparison of means between groups ($K = 1, 2, 3, 4$) for 12 morphological measures	38
Figure 18. Correlation matrix for 12 morphological measures	39
Figure 19. Comparison of means between groups for basin and cell fate entropy computed from three cell fates	40
Figure 20. Topological properties of morphogenetic patterns for $K = 1, 2, 3, 4$	42
Figure 21. Feedback relationship in our morphogenetic process	42
Figure 22. A schematic diagram for example multilayer GRNs	45
Figure 23. Biological systems and intercellular networks	46
Figure 24. Schematic diagrams for explaining the dynamics of an intracellular GRN	47
Figure 25. Schematic diagrams for explaining the dynamics of multilayer GRNs	48
Figure 26. Four categories of the robustness and evolvability, and the relationship between the properties of multilayer GRNs and intracellular GRNs	52
Figure 27. Probabilities of generating robust & evolvable GRNs, robust GRNs, and evolvable GRNs depending on cellular topologies	54
Figure 28. Probabilities of generating robust & evolvable multilayer GRNs depending on the number of an intercellular network	56
Figure 29. Derrida plot representing dynamics of GRNs	61

Chapter 1 Introduction

Many complex biological structures including tissues and organs are formed through a developmental process from a single fertilized cell to a multicellular embryo [5]. The morphogenesis of those complex patterns are driven by a gene regulatory network (GRN) that exists within each cell and responds to cell-cell interactions [40, 41]. As a theoretical model of such GRNs, random Boolean networks (RBNs) were proposed by Kauffman [31]. In RBNs, genes (nodes) have binary states (either ON or OFF), whose dynamics are determined by a set of Boolean functions over the states of other genes. Although RBNs are a highly simplified model, they have been extensively utilized in artificial life and complex systems research [1, 3, 23, 24, 43, 50, 60, 66].

In the context of GRNs, the concept of *criticality* of RBNs has been discussed as a phase transition point between ordered and chaotic regimes for the dynamics of those networks [32, 33]. The criticality of GRNs has been recognized as a property which makes robustness and adaptability coexist in living organisms [2]. When perturbations are added to GRNs, ordered GRNs are so robust that they just sustain existing cellular functions. On the contrary, chaotic GRNs are so adaptable that they vigorously create new functions rather than conserving existing ones. Meanwhile, critical GRNs stably sustain their functions against the perturbations, and at the same time flexibly generate new phenotypes, which may help organisms to adapt to new environments because they have an optimal balance between robustness and adaptability.

Whereas many studies have been performed to elucidate the relationship between

criticality of GRNs and dynamics of GRNs at a single cell level based on RBNs [6, 48, 52, 57, 58, 61], the relationship between their criticality and properties of multicellular organisms at a higher level has not fully explored. Only a few studies have determined how the properties of intracellular GRNs influence the properties of organisms at a multicellular level, using RBNs as GRNs [13, 20, 45, 64]. Moreover, research on the relationship between the criticality of GRNs and properties of multicellular systems under genetic perturbations (e.g., mutations) has not been conducted yet, even though mutations do occur in cells of living organisms by stochasticity or environmental factors [4, 7, 8, 9, 11, 54, 59].

To investigate the potential roles of criticality of GRNs at a multicellular and hierarchical level, here we conduct three studies. Firstly, we propose a GRN-based morphogenetic model and reveal the role of criticality of GRNs in morphogenesis. Secondly, we include an evolutionary process in our morphogenetic model by introducing genetic perturbations to GRNs, and examine whether the role of the criticality of GRNs can be maintained even in the presence of the evolutionary perturbations. Also, we look into what the resulting morphologies are like and what kind of biological implications they have from the epigenetic viewpoint in morphology. Lastly, we present multilayer GRNs consisting of an intercellular layer and an intracellular layer, and delve into how the criticality of GRNs influences on the robustness and evolvability of the multilayer GRNs depending on cellular topologies and the number of links of an intercellular network.

The rest of the dissertation is structured as follows. In Chapter 2, we provide a brief literature review concerning the relationship between criticality of GRNs and

system properties. In Chapter 3, we mention the objective of our research. In Chapter 4, 5, and 6, we design the GRN-based morphogenetic model, the GRN-based morphogenetic model with genetic perturbations, and the multilayer GRNs, respectively. We, in each chapter, show experiments and results corresponding to the respective models mentioned above. In Chapter 7, we summarize and conclude our studies.

Chapter 2 Related Work

In this chapter, we provide two brief literature reviews. One is about the relationship between criticality of GRNs and dynamics of GRNs at a single cell level, and the other is about the relationship between properties of intracellular GRNs and properties of organisms at a multicellular level.

2.1 The Relationship Between Criticality of GRNs and Dynamics of GRNs at a Single Cell Level

Since the notion of criticality of GRNs based on RBNs was established by Kauffman [32, 33], many studies have been conducted on whether GRNs of living organisms are dynamically critical or not. The studies introduced in this section compare dynamic behaviors of RBNs in the critical regime with gene expression dynamics or dynamics of Boolean models of genetic networks, both of which are based on gene expression data of real living organisms [6, 48, 52, 57, 58, 61]. They demonstrate that the dynamics of the living organisms are consistent with those of critical RBNs, and thus conclude that the dynamics of living organisms are critical. Details are reviewed below.

Serra et al. [57] showed that the gene expression dynamics of *Saccharomyces cerevisiae* (commonly known as baker's yeast) were critical through the comparison of their gene expression data and critical RBNs in the perturbation avalanche analysis which measured the size of an avalanche, i.e., the number of genes that are affected by the knockout of a single gene. They found that the distribution of avalanche sizes of critical

RBNs approximated the distribution obtained experimentally on *Saccharomyces cerevisiae*. Also, Serra et al. [58] consolidated the result by deriving analytical approximations for the distribution of avalanches in RBNs. Similarly, Rämö et al. [52] showed that the gene expression dynamics of *Saccharomyces cerevisiae* were critical, using approximate formulas for the distributions of avalanche sizes. They demonstrated that the distributions of avalanche sizes of both critical RBNs and *Saccharomyces cerevisiae* followed power-law distributions with the same exponent value.

Shmulevich et al. [61] and Nykter et al. [48] showed that the gene expression dynamics of biological systems exhibited criticality, by applying quantitative measures used in data compression to the gene expression data of living organisms. Specifically, Shmulevich et al. [61] measured Lempel-Ziv (LZ) complexity of both the binarized gene expression data during HeLa cell cycle progression and RBNs in ordered, critical, and chaotic regimes. They found that the LZ value obtained from the gene expression dynamics of HeLa Cells was consistent with that of either ordered or critical dynamic behavior. Nykter et al. [48] calculated normalized compression distance (NCD) from the gene expression data of macrophage. They compared the measured values with the NCD values of ordered, critical, and chaotic RBNs. They found that the trajectory of NCD of macrophage corresponded with that of critical RBNs.

Balleza et al. [6] indicated that the dynamics of living organisms in four kingdoms operated close to criticality by examining the dynamics of Boolean models of examples belonging to four kingdoms in biology. Inferring interactions among genes from the gene expression data of *Saccharomyces cerevisiae* (yeast in kingdom fungi), *Escherichia coli* (bacteria in kingdom protista), *Bacillus subtilis* (bacteria in kingdom

protista), *Drosophila melanogaster* (insect in kingdom animalia), and *Arabidopsis thaliana* (plant in kingdom plantae), they implemented five Boolean networks. They displayed that the slopes of Derrida curves (see Appendix A) of the five Boolean networks were similar to that of a critical RBN, where a Derrida plot visualizes the dynamic behaviors of Boolean networks.

2.2 The Relationship Between Properties of Intracellular GRNs and Properties of Multicellular Organisms

In this section, we introduce studies which have explored the relationship between properties of intracellular GRNs and properties of organisms at a multicellular level [13, 20, 45, 64]. They all use theoretical multicellular models where all the cells have the same RBNs as GRNs in a discrete space like cellular automata. Details are reviewed below.

Flann et al. [20] and Mohamdlou et al. [45] studied the relationship between properties of GRNs and multicellular pattern complexity. Specifically, Flann et al. [20] examined how dynamics of GRNs influenced multicellular pattern complexity. They assumed epithelial cells as 20×20 square cell arrangement. They showed that the epithelium models with ordered and critical GRNs tended to generate the most information-rich patterns. Mohamdlou et al. [45] investigated how modularity of critical and chaotic GRNs affected multicellular pattern complexity. Using a lattice of 20×20 as an epithelial model, they found that modular connectivity of GRNs, especially feedback loops, increased the complexity of multicellular patterns.

Villani et al. [64] examined how a coupling strength (fraction of genes that are affected by neighboring cells) between cells influenced dynamics of GRNs. They used 20×20 square cell arrangement as an artificial tissue. Increasing the coupling strength, they measured the following three outcome variables in 1,000 simulation runs: the fraction of simulation runs where all the cells of the systems converge to the same attractor (α), the fraction of simulation runs where all the cells in the systems converge to some attractor (β), and the fraction of simulation runs where none of the cells converge to any attractor (γ). The higher the coupling strength was, the larger α and γ were, while the lower β was. They also found that increasing the coupling strength amplified the properties of GRNs, i.e., ordered GRNs became more ordered and disordered GRNs became more disordered.

Similarly, Damiani et al. [13] also studied how the strength of interaction affected attractors of multiple RBNs. They proposed multiple RBNs in 5×5 cellular automata, using RBNs in a critical regime. Based on the frequencies of different attractors of multiple RBNs, they calculated entropy and considered it as a measure to quantify diversity of cell behaviors. They showed that the diversity of cell behaviors was varied by the strength of interaction. Moreover, they found the value of interaction strength to maximize the cell behavior diversity, which corresponded to the percentage of genes related to cell signaling in an actual human cellular signaling network.

Because the existing models reviewed above all used a fixed set of neighbors in a discrete space like cellular automata, they are not realistic as a morphogenetic model to represent the developmental process from a single cell to a multicellular embryo. Moreover, none of them considered genetic perturbations which do occur in cells of

living organisms. In our studies, we reveal the potential roles of criticality of GRNs in the context of multicellular settings by developing a morphogenetic model which grows from a single cell in a continuous space and a multilayer GRN model with dynamic cellular topologies, and adding genetic perturbations to GRNs.

Chapter 3 Objective

We aim at revealing the potential roles of criticality of GRNs at a multicellular and hierarchical level. Specifically, using a GRN-based morphogenetic model, we elucidate the role of criticality of GRNs in morphogenesis at a multicellular level. Furthermore, adding genetic perturbations (e.g., mutations) to GRNs, we examine whether the role of the criticality of GRNs can be maintained even in the presence of the evolutionary perturbations. Also, we look into what the resulting morphologies are like and what kind of biological implications they have from the epigenetic viewpoint in morphology. Lastly, we delve into how the criticality of GRNs affects the robustness and evolvability of multilayer GRNs at a hierarchical level.

Chapter 4 The Role of Criticality of Gene Regulatory Networks in Morphogenesis

In this chapter, we propose a GRN-based morphogenetic model using a RBN as a GRN and Spring-Mass-Damper kinetics for cellular movements, and reveal the role of criticality of GRNs in morphogenesis.

4.1 Model: GRN-Based Morphogenetic Systems

We developed a computational model of morphogenetic processes of cell aggregation, in which all the cells have an identical intracellular GRN. Figure 1 shows the simulation algorithm for our model. The simulation starts with one seed cell. It imitates the process in which a single zygote divides and grows into multicellular form during embryonic development. Cells are equipped with a RBN as an intracellular GRN. Neighboring cells are detected within a fixed neighborhood radius. Through the interaction with neighbors, cells' fates are determined by the GRN. We assume that there are four fundamental cell fates in our model: proliferation, differentiation, apoptosis, and quiescence. Cells expressing proliferation, differentiation, or quiescence can switch their fates through cell-cell interactions. The cells are positioned in a two-dimensional continuous space by spring-mass-damper (SMD) kinetics. Until the termination condition of the simulation is satisfied, the initial seed cell grows into an aggregation, iterating the processes of finding neighboring cells and re-positioning cells in the space in each time step. The simulator of our model was implemented in Java.

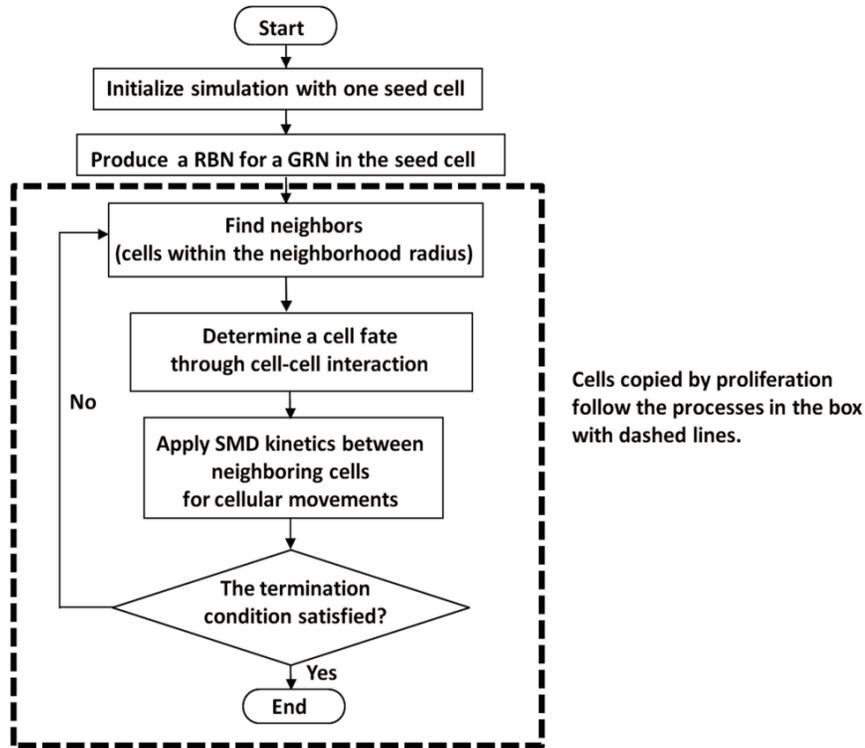
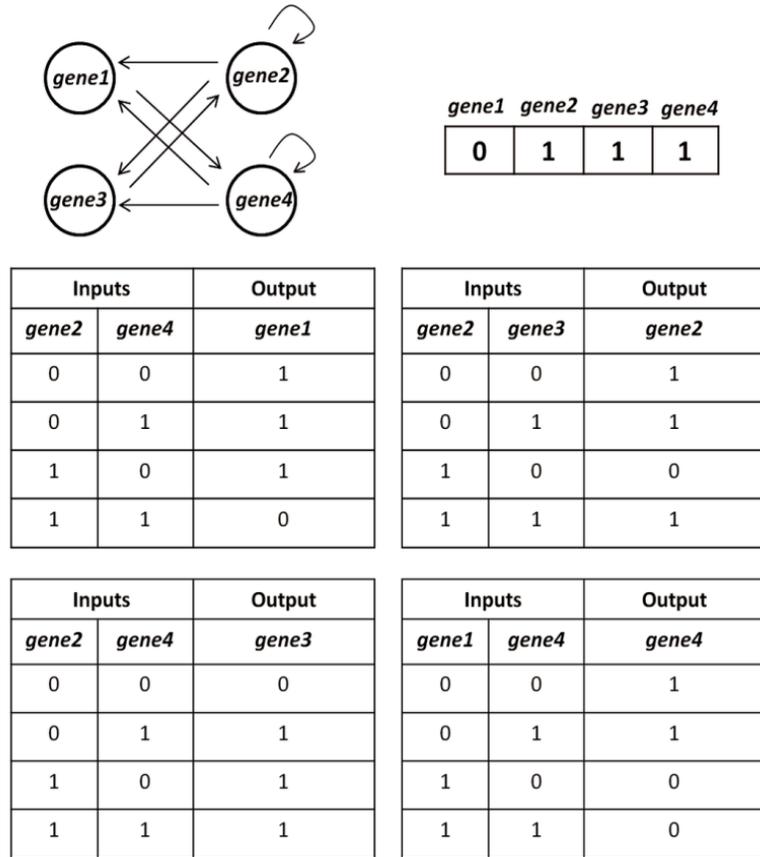


Figure 1. Simulation algorithm for our GRN-based morphogenetic model

4.1.1 Gene Regulatory Network (GRN)

A RBN (a.k.a., *NK* Boolean network) was suggested as a GRN model by Kauffman [31, 32, 33]. Here N is the number of nodes and K is the number of input links per node. A node represents a gene. The state of a node can be either ON (1, activated) or OFF (0, inhibited). The node state is determined by the states of input nodes and a Boolean function assigned to each node. A state space which is constructed from the topology of a RBN and assigned Boolean functions refers to the set of all the possible configurations and all the transitions among them. Figure 2 shows schematic diagrams for an example RBN and its state space. In the state space, stationary or cyclical

(a)



(b)

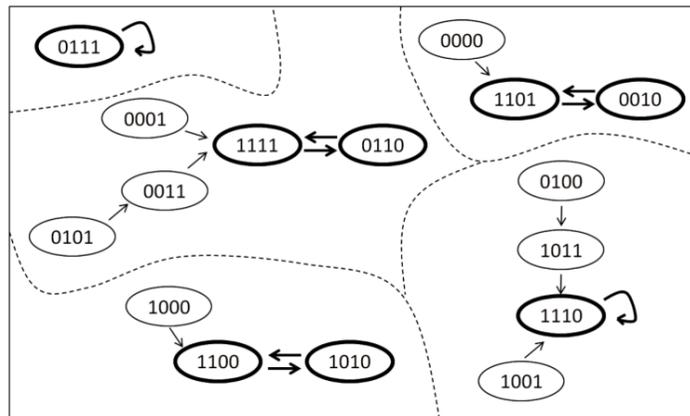


Figure 2. Schematic diagrams for an example GRN and its state space. (a) A RBN with $N = 4$, $K = 2$, and Boolean functions randomly assigned to each node. (b) State space of the RBN. The state space consists of $2^4 = 16$ configurations and transitions among them. The configurations with bold lines are attractors. Dashed lines draw boundaries for each basin of attraction.

configurations are defined as attractors, and the others are called basins of attraction of the attractors. The dynamics of RBNs are divided into three regimes depending on the structure of their state space: ordered, critical, and chaotic. Using node in-degree (K), internal homogeneity (p), or canalizing functions, the dynamics of RBNs can be systematically varied. The dynamics of a RBN are known to be determined by node in-degree (K), i.e., $K = 1$ is ordered, $K = 2$ is critical, and $K > 2$ is chaotic, on average [32, 33]. For our morphogenetic model, we use a RBN that consists of 16 nodes ($N = 16$) as an intracellular GRN. Adjusting node in-degree (K) from 1 to 4, we vary the dynamics of RBNs: ordered ($K = 1$), critical ($K = 2$), and chaotic ($K = 3, 4$).

In view of in vitro experimental data showing that attractors of GRNs represent cell types or cell fates, Huang et al. suggested a conceptual framework to explain stochastic and reversible transitions between cell fates using NK Boolean networks [10, 27, 28, 29]. Our morphogenetic systems are based on their framework. We randomly assign four cell fates to attractors of GRNs. Specifically, if there is only one attractor, proliferation is assigned to the attractor. If there are two attractors, proliferation and differentiation are randomly assigned to the two attractors. Likewise, if there are three attractors, proliferation, differentiation, and apoptosis are randomly assigned to those attractors. If there are four or more attractors, proliferation, differentiation, and apoptosis are randomly assigned to three attractors and quiescence is assigned to the rest of the attractors (Figure 3).

With regard to cellular behaviors, cells in proliferation are divided into two, and the daughter cells are placed within a fixed neighborhood radius (R) centering on the mother cells. Cells in differentiation are labeled as differentiated. Cells in apoptosis die

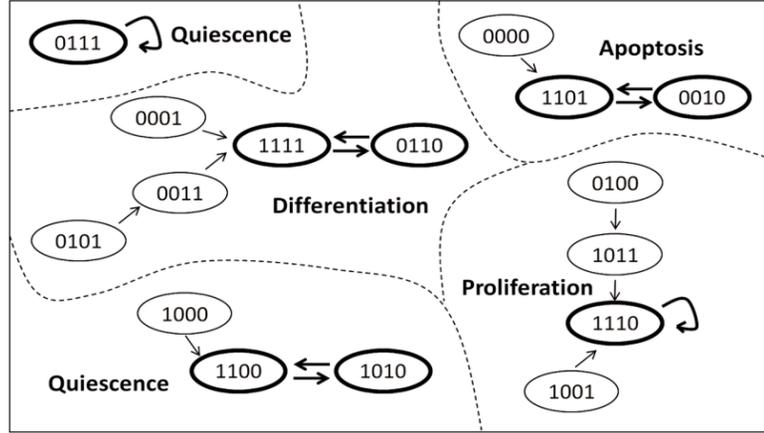


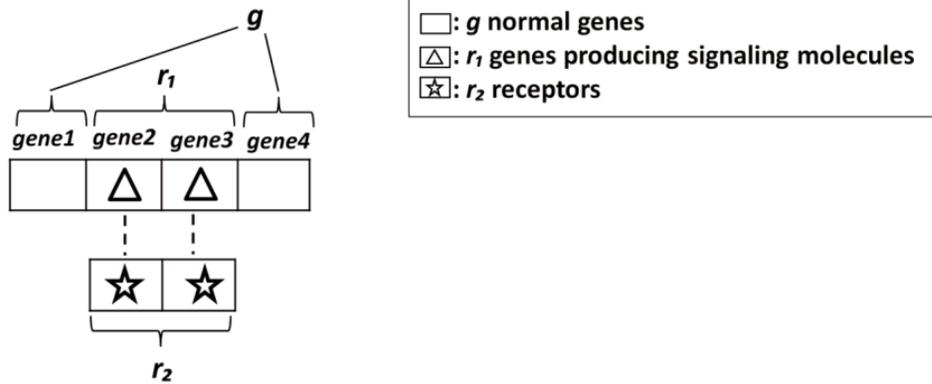
Figure 3. An example state space of a GRN where four cell fates are randomly assigned. (In actual simulations, 16 nodes were used. Thus, $2^{16} = 65,536$ configurations exist in the state space.)

and disappear from the space. Cell in quiescence do not show any behaviors.

4.1.2 Cell-Cell Interactions

Switching between cell fates occurs by perturbations of internal gene expression values of an intracellular GRN through cell-cell interactions. Our mechanism for cell-cell interactions is based on cell signaling of Damiani et al's multiple random Boolean networks on 2D cellular automata [12, 13]. In our model, an intracellular GRN has n genes, which consist of normal genes (g) and special genes (r) as shown in Figure 4 (a). The special genes (r) exist in pairs where genes producing signaling molecules (r_1) and receptors (r_2) are matched one to one. This one-to-one correspondence indicates signal transduction specificity by which certain signaling molecules respond to particular receptors. The special genes r_1 synthesize signaling molecules and release them. Then, those molecules bind to the corresponding receptors r_2 on cells within the neighborhood radius (R).

(a)



(b)

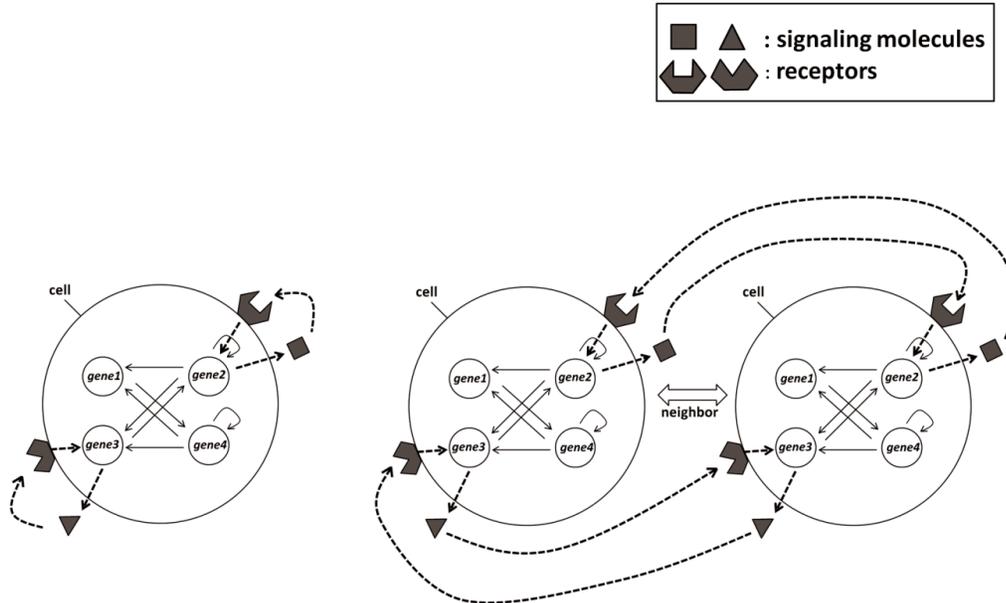


Figure 4. Cell signaling for cell-cell interactions. Schematic diagrams are examples illustrating the concept of cell signaling. (a) Assignment of genes in a GRN for cell signaling. (b) Two signal transduction mechanisms: autocrine (left) and paracrine (right).

The signal transduction has two mechanisms: autocrine and paracrine. Autocrine means that a cell produces signaling molecules that bind to receptors on the same cell. Paracrine means a cell produces signaling molecules that bind to receptors on its

neighboring cells. Figure 4 (b) illustrates the two mechanisms. In our model, when there are no neighboring cells, autocrine is used. When there are neighbors, paracrine is used.

The gene expression values of an intracellular GRN are updated as follows:

- Normal genes g : the states of the normal genes g are updated by the states of input nodes and randomly assigned Boolean functions.
- Genes producing signaling molecules r_1 : like the normal genes g , the states of the genes producing signaling molecules r_1 are updated by the states of input nodes and randomly assigned Boolean functions. If the states of r_1 are 1, the genes produce signaling molecules. If the states are 0, the genes do not synthesize signaling molecules.
- Receptors r_2 : the states of the receptors r_2 are determined by the average concentration of the signaling molecules within the neighborhood radius R . Figure 5 shows an example of calculating the average concentration of the signaling molecules from the neighboring cells and determining the state of receptor *gene 2* of cell i . Based on a certain threshold (τ_{th}), the state of receptor *gene 2* is updated. If the average concentration value is bigger than τ_{th} , the state becomes 1. Otherwise, it becomes 0.

The following steps are taken to update the gene expression values and to determine a cell fate from the change of the gene values:

- (1) Look into whether there are neighboring cells within the neighborhood radius R . If there are neighbors, paracrine signaling is used. Otherwise, autocrine signaling is used.

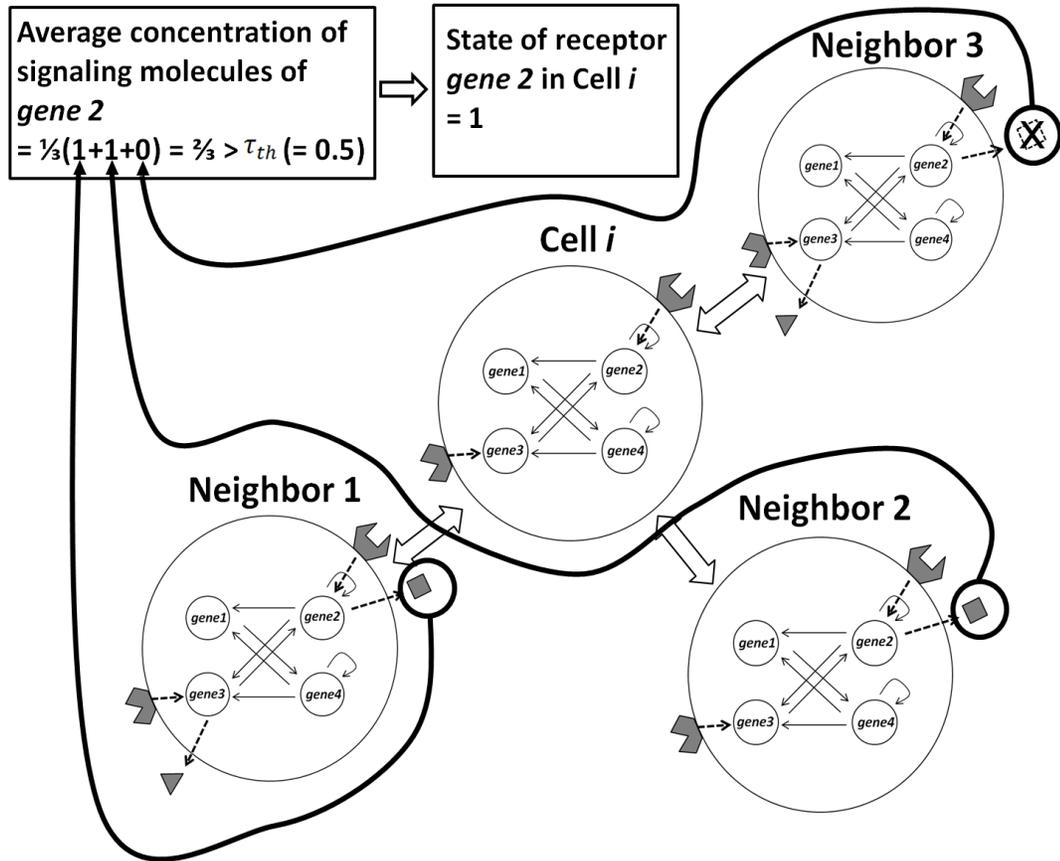


Figure 5. An example showing how to calculate the average concentration of the signaling molecules neighboring cells produce and determine the state of receptor *gene2* of cell *i*.

- (2) Determine the states of the receptors r_2 through comparisons of the average concentrations of signaling molecules within the neighborhood radius R and the threshold value τ_{th} .
- (3) Activate or inhibit genes that have the receptors as input nodes in an intracellular GRN based on the states of the receptors r_2 . If the states of the receptors are 1, the states of genes become ON (1, activated). Otherwise, the states become OFF (0, inhibited).
- (4) Check the attractor that the updated gene states finally converge to.
- (5) Express the cell fate that is assigned to the attractor for cellular behaviors.

(6) Assign the states of the attractor as gene expression values for the next time step.

In the case of cyclical attractors, the states of the attractor that the updated gene states in (4) first reach become gene expression values for the next time step.

4.1.3 Cellular Movements

Our mechanism for cellular movements is based on Doursat's approach [18]. We determine cells' positions in each time step through Spring-Mass-Damper (SMD) kinetics. Specifically, we assume that cells within the neighborhood radius R are connected by a spring with spring constant k and equilibrium length l , and a damper with damping coefficient c between each other. When cell A's position is $P_A = (x_A, y_A)$ and cell B's position is $P_B = (x_B, y_B)$, the equation for cellular movements is as follows:

$$m\ddot{P}_{AB} = -k \left(1 - \frac{l}{\|P_{AB}\|} \right) P_{AB} - c\dot{P}_{AB}$$

where

$$P_{AB} = P_B - P_A = (\delta \cos \theta, \delta \sin \theta),$$

$$\delta = \|P_{AB}\|, \quad \theta = \arctan \left(\frac{y_B - y_A}{x_B - x_A} \right),$$

$$\|P_{AB}\| = \|P_B - P_A\| = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}$$

Because we neglect the effect of inertia, we replace $m\ddot{P}_{AB}$ with zero. Then, we finally obtain the following position update equation at each time step $\Delta t = 1$:

$$\Delta P_B = -\Delta P_A = \frac{\Delta P_{AB}}{2} = \frac{-k}{2c} \left(1 - \frac{l}{\|P_{AB}\|} \right) P_{AB}$$

We can obtain different shapes of spatial patterns by the above position updating rule, allowing physical interactions such as pushing or adhesion.

To acquire much more diverse spatial patterns, we introduce the dependence of parameters k , l , and c on cell fates and perturbations to the cell position (x, y) . In the case of k , l , and c , we determine the values depending on six possible types of cell fate combinations between two cells: $[proli-proli]$, $[proli-diff]$, $[proli-qui]$, $[diff-qui]$, $[diff-diff]$, and $[qui-qui]$, where *proli* is proliferation, *diff* is differentiation, and *qui* is quiescence (Figure 6). Here apoptosis is excluded because cells due to apoptosis disappear from the space. In each simulation run, the parameter values of k , l , and c are randomly chosen in certain ranges given in Table1. For the perturbations, small perturbation values are added to the updated cell positions.

When the dependence of k , l , and c on cell fates and perturbations to the cell positions are introduced, the final position of cell A whose neighbor is cell B is as follows:

$$P_A(t + 1) = P_A(t) + (\Delta P_A)_{[\alpha-\beta]} + \omega_{[\alpha-\beta]}$$

where α is cell A's cell fate, β is cell B's cell fate, and ω is the perturbation to the updated coordinate of cell A. This positional updating is performed for all the neighboring cells of cell A.

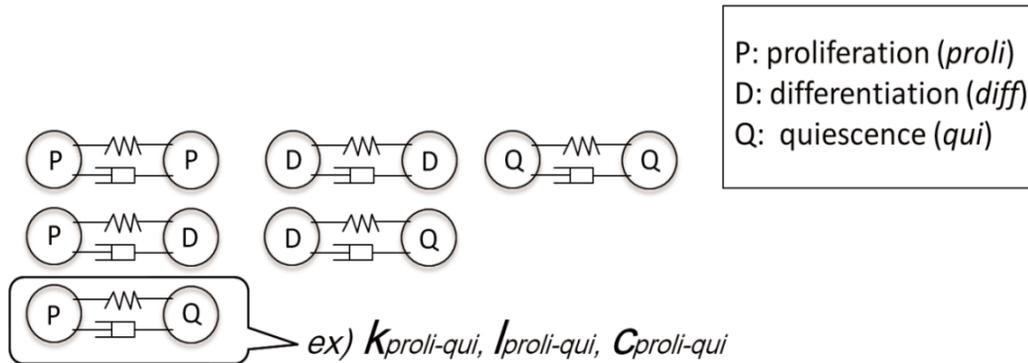


Figure 6. Six possible types of cell fate combinations between two cells.

4.2 Experiments

We performed 10,000 independent simulation runs for each value of K , i.e., $K=1, 2, 3$, and 4. Specifications of parameters for the simulations were the following:

- Space: the cells were positioned in a 2D continuous 700×700 (in arbitrary unit) square area.
- Limitation of cell population: the population growth was limited up to 200 cells to keep computational loads reasonable in each run.
- Simulation termination condition: the simulations were terminated when the time step t reached 1,000 or there was no cell remaining in the space due to apoptosis.
- Parameter values: the values of parameters concerning GRNs, cell-cell interactions, and cellular motions are given in Table 1. The number of special genes (r) was determined according to Damiani et al.'s model [13] and biological evidence [49]. Damiani et al. showed that the diversity of cellular behaviors was maximized when the coupling strength (fraction of genes that are affected by neighboring cells) was around 0.1 [13]. This value is also similar to the ratio of the number of genes related to cell signaling to the number of human genes [49]. In our model, the coupling strength is set to 0.125, because the number of special genes (r) is 2 and the number of nodes of a GRN (N) is 16.

4.2.1 Measures for Morphogenetic Pattern Analysis

To compare how morphogenetic patterns are different between groups $K=1, 2, 3$, and 4, we used the 12 measures (i - xii) described below [55, 56]. Among those measures,

Table 1. Parameters of the morphogenetic model and their values

Model	Parameter	Value
GRN	Number of nodes (N)	16
	Number of in-degree per node (K)	1, 2, 3, 4
Cell-Cell Interactions	Neighborhood radius (R)	30
	Number of special genes (r)	2
	Threshold of signaling molecules (τ_{th})	0.5
Cellular Movements	Spring constant (k)	$k \in \text{unif}(0, 1) \subset \mathbb{R}$
	Spring equilibrium length (l)	$l \in \text{unif}(0, 100) \subset \mathbb{R}$
	Damper coefficient (c)	$c \in \text{unif}(0, 200) \subset \mathbb{R}$

vi - xii are measures regarding network topology. To apply them to our morphologies, we constructed a network from each morphogenetic pattern by connecting each cell to other cells within the neighborhood radius R . Figure 7 shows an example morphology and a network constructed from it using our network construction method. Such network construction allowed for detection of topological differences more effectively. All the morphogenetic measures were obtained from the final configuration of each simulation.

i. Number of cells (*numOfCells*)

This is the total number of cells in a morphogenetic pattern.

ii. Average distance of cells from center of mass (*massDistance*)

This is the mean of Euclidean distances between each cell position and the center

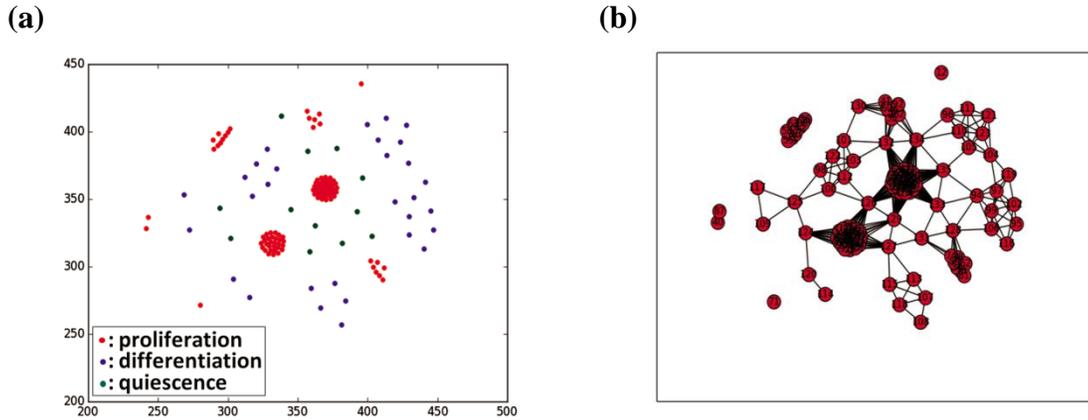


Figure 7. Network construction for the analysis of morphologies. (a) Snapshot of a morphogenetic pattern. (b) Network constructed using our network construction method from (a).

of mass (\bar{x}, \bar{y}) , that is, the point with the average coordinates of all the cells.

iii. Average pairwise distance (*pairDistance*)

This is the mean of Euclidean distances between two randomly sampled cells' positions. The mean was calculated based on 10,000 pairs, which were sampled with replacement.

iv. Kullback-Leibler divergence between pairwise particle distance distributions of a morphogenetic pattern and a random pattern (*kld*)

This quantifies how nontrivial morphogenetic patterns are, compared to randomly distributed patterns. It was calculated as the Kullback-Leibler (KL) divergence between pairwise particle distance distributions of a morphogenetic pattern (Figure 8 (a)) and a randomly distributed pattern (Figure 8 (b)) made of the same number of cells within the same spatial dimensions. Each pairwise

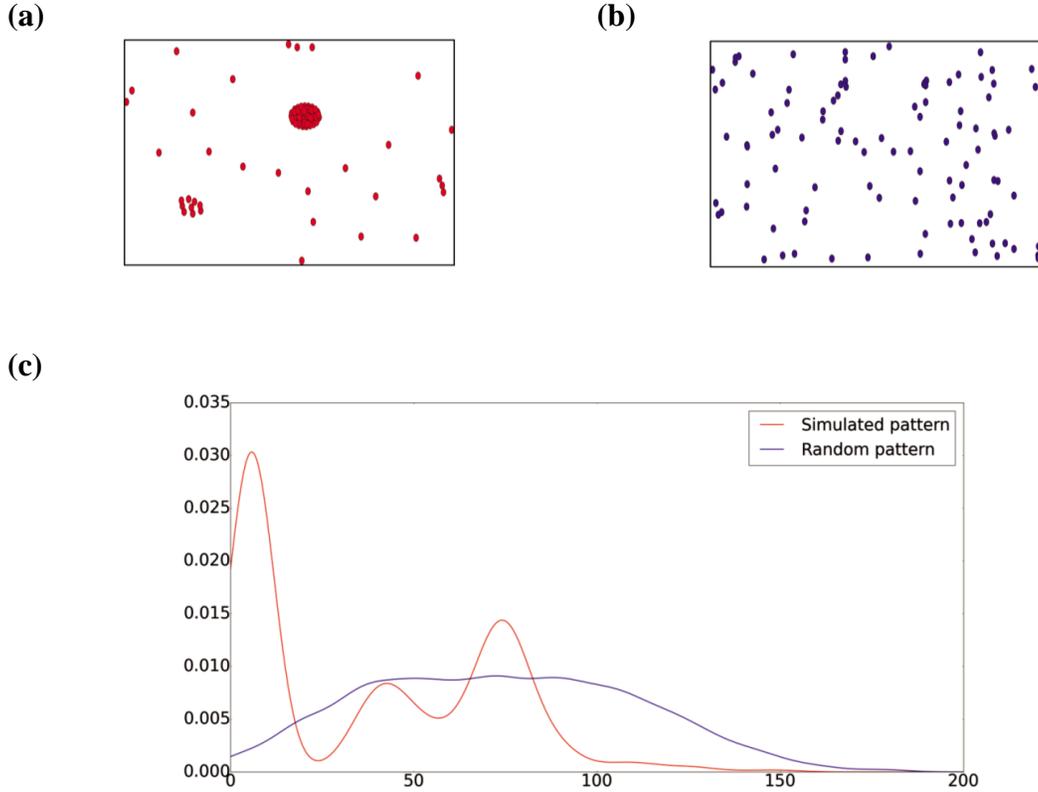


Figure 8. Nontrivial morphology detection using KL divergence. (a) A morphogenetic pattern acquired from a simulation. (b) A random pattern obtained from a uniform distribution. (c) Pairwise particle distance distributions of a simulated pattern and a random pattern. The curves are estimated by Gaussian kernel density estimation.

particle distance distribution was obtained through 10,000 random sampling with replacement of a pair of coordinates of cells (Figure 8 (c)). Thus, the larger kld is, the more structured (nonrandom) the morphogenetic pattern is.

Both $pairDistance$ and kld used pairwise particle distances. $pairDistance$ measures a rough size of a morphogenetic pattern, while kld quantifies nontriviality of its morphology. Two morphogenetic patterns may have similar $pairDistance$ values but very different kld values at the same time.

v. Mutual information of cell fates between neighboring cells (*MI*)

This examines nonlinear correlation of cell fates between neighboring cells in a morphogenetic pattern. It was calculated using the frequencies of three neighboring cell fates (except for apoptosis, because cells expressing apoptosis die and disappear from the space). Figure 9 shows an example calculating *MI* in a morphogenetic pattern. Counting the frequencies of combinations of fates of neighboring cells, *MI* captures how much informational correlation would exist between the fate of a cell and that of its neighbors. If there was only one cell remaining, the value of *MI* was set to 0.

vi. Average clustering coefficient (*avgCluster*)

This explains how densely connected the nodes (cells) are to each other in a network. The clustering coefficient $C(i)$ of node v_i in a network is defined as follows:

$$C(i) = \frac{|\{e_{jk}: v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)/2}$$

Here, e_{jk} is a link that connects node v_j and node v_k within the set of neighboring cells N_i around a node v_i , E is a set of links in the network, and k_i is the degree of v_i (i.e., the size of N_i). The denominator is the total number of possible node pairs within node v_i 's neighborhood. The numerator is the number of actually connected node pairs among them. The average clustering coefficient is given by

$$\bar{C} = \frac{\sum_i C(i)}{n},$$

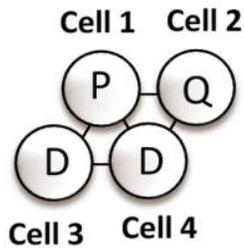
where n is the total number of nodes.

vii. Link density (*linkDensity*)

This describes the density of connections in a network. For a network G composed of nodes n and m links, the link density $\rho(G)$ is given by

$$\rho(G) = \frac{m}{\frac{n(n-1)}{2}} = \frac{2m}{n(n-1)}$$

for an undirected network, where m is the number of links.



P: proliferation
 D: differentiation
 Q: quiescence

$$\begin{array}{cccc}
 \text{Cell 1} & \text{Cell 2} & \text{Cell 3} & \text{Cell 4} \\
 \text{X} = [& \text{P, P, P, Q, Q, D, D, D, D, D, D} &] \\
 \text{Y} = [& \text{Q, D, D, P, D, P, D, P, Q, D} &]
 \end{array}$$

$L = 10$ (length of information)

$$MI = (H(X) + H(Y) - H(XY)) / (\log L)$$

Figure 9. An example showing how to calculate mutual information between cell fates of cells and their neighboring cells. The value of computed mutual information was divided by $\log L$ for the purpose of normalization

viii. Number of connected components (*numConnComp*)

This is the number of connected components in a network. A connected component is a subgraph in which there is a path between every pair of nodes. A single isolated cell was also considered one connected component by itself.

ix. Average size of connected components (*meanSizeConnComp*)

This is the mean of the numbers of nodes in each connected component in a network. When there was no connected component, the value was set to 0.

x. Homogeneity of sizes of connected components (*homoSizeConnComp*)

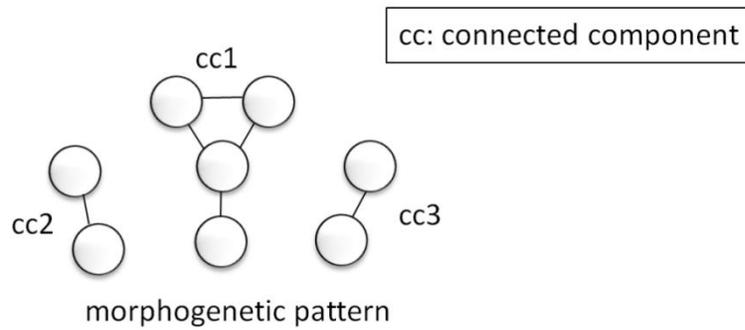
This quantifies how similar the sizes of connected components are in a network. This measure was calculated as one minus the normalized entropy in the distribution of sizes of connected components. Figure 10 shows an example calculating *homoSizeConnComp* in a morphogenetic pattern. When there was only one connected component, the value was set to 1.

xi. Size of the largest connected components (*sizeLarConnComp*)

This is the maximum size of the connected components in a network.

xii. Average size of connected components smaller than the largest one (*meanSizeSmaller*)

This is the mean of the sizes of connected components except for the largest connected component in a network. If there was only one connected component



list of size of connected components: $X = [2, 4, 2]$

length of information: $L = 3$

normalized entropy: $H(X)/\log L = \frac{-(1/3)*\log(1/3)-(2/3)*\log(2/3)}{\log(3)} = 0.58$

$homoSizeConnComp = 1 - (H(X)/\log L) = 1 - 0.58 = 0.42$

Figure 10. An example showing how to calculate homogeneity of sizes of connected components (*homoSizeConnComp*). The value of computed entropy $H(X)$ was divided by $\log L$ for the purpose of normalization.

in the network, the value was set to 0.

For all of the above measures, if there were no cells remaining in the space, their values were set to 0.

4.3 Results

Figure 11 shows distributions of the morphogenetic patterns based on the number of cells at the end of each simulation: *larger than one cell*, *single cell*, and *no cell*. We found that the larger K is, the more frequent the cases of *no cell* and *single cell* are. That is, the number of morphogenetic patterns which consist of more than one cell decreases

as K increases. These distributions of morphogenetic patterns are due to the fact that greater values of K make it more likely for GRNs to have more than two attractors so apoptosis can occur more frequently. Figure 12 shows different spatial patterns of each group acquired from randomly sampled 20 simulations. The trend of the distributions in Figure 11 is visually confirmed in Figure 12. Figure 13 summarizes the 12 measures of spatial pattern characteristics, where Kruskal-Wallis and Nemenyi (as post-hoc analysis) tests were conducted to detect statistically significant differences among the four groups ($K = 1, 2, 3, 4$). For the measures except for MI and kld , the average values decreased as K increased. Based on the statistical tests, we found that the values of kld and MI were highest at $K = 2$ (Figure 13 (c) and (l)). To investigate correlations between the 12 measures, we obtained a correlation matrix (Figure 14). Seeing the row of $numOfCells$, we found that most of the measures were highly correlated to $numOfCells$.

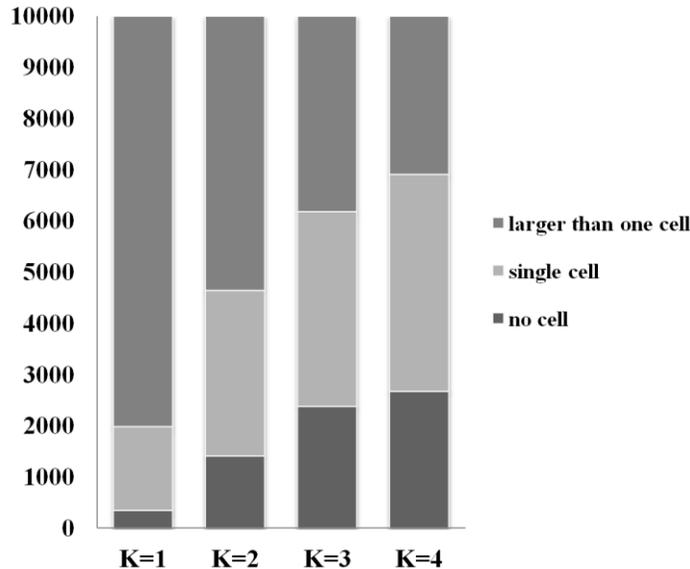
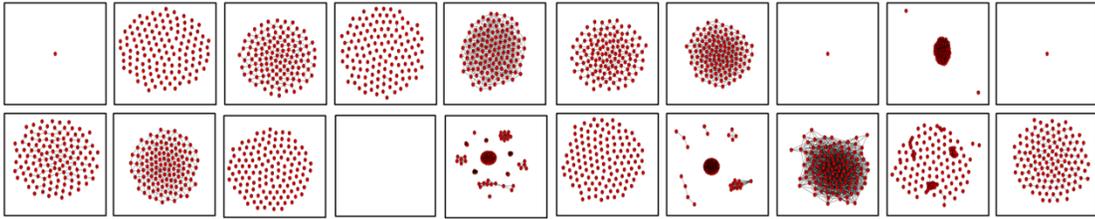
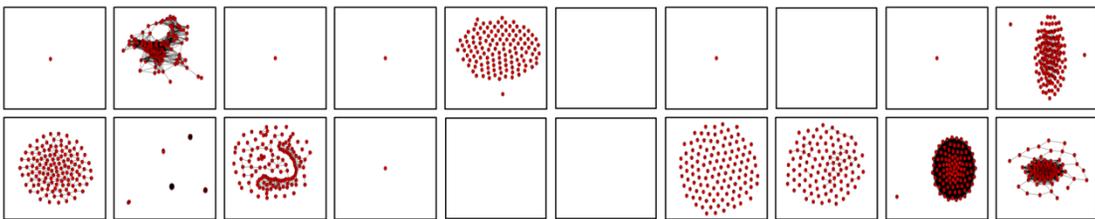


Figure 11. Distributions of morphogenetic patterns according to the number of cells for $K = 1, 2, 3, 4$.

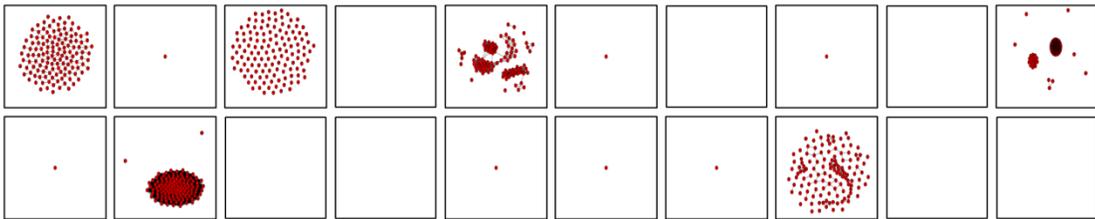
(a)



(b)



(c)



(d)

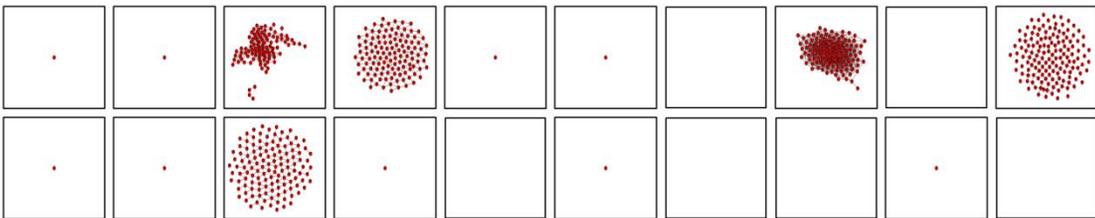


Figure 12. Different morphogenetic patterns represented with networks for $K = 1, 2, 3, 4$. The patterns are acquired from randomly sampled 20 simulations. (a) $K = 1$. (b) $K = 2$. (c) $K = 3$. (d) $K = 4$.

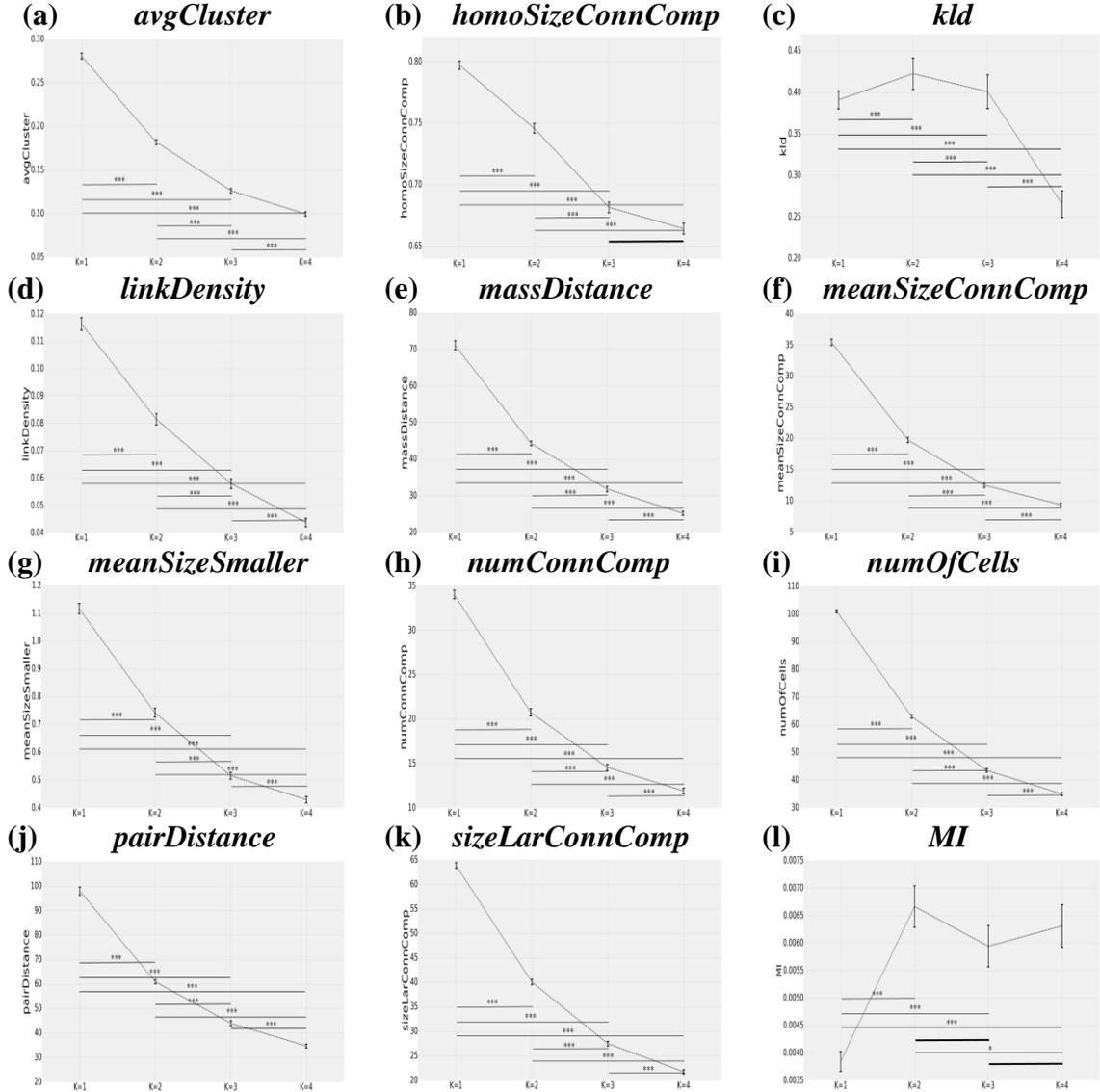


Figure 13. Comparison of means between groups ($K = 1, 2, 3, 4$) for 12 morphological measures (Kruskal-Wallis test: $p < 2.2e-16$, Nemenyi test (post-hoc): ‘.’: $p < 1.0$, ‘.’: $p < 0.1$, ‘*’: $p < 0.05$, ‘**’: $p < 0.01$, ‘***’: $p < 0.001$). In the case that there is no difference between two groups, a bold line without an asterisk is presented in the plot. (a) Average clustering coefficient (*avgCluster*). (b) Homogeneity of sizes of connected components (*homoSizeConnComp*). (c) KL divergence between pairwise particle distance distributions of morphogenetic pattern and a random pattern (*kld*). (d) Link density (*linkDensity*). (e) Average distance of cells from center of mass (*massDistance*). (f) Average size of connected components (*meanSizeConnComp*). (g) Average size of connected components smaller than the largest one (*meanSizeSmaller*). (h) Number of connected components (*numConnComp*). (i) Number of cells (*numOfCells*). (j) Average pairwise distance (*pairDistance*). (k) Size of the largest connected component (*sizeLarConnComp*). (l) Mutual information between cell fates of cells and their neighboring cells (*MI*).

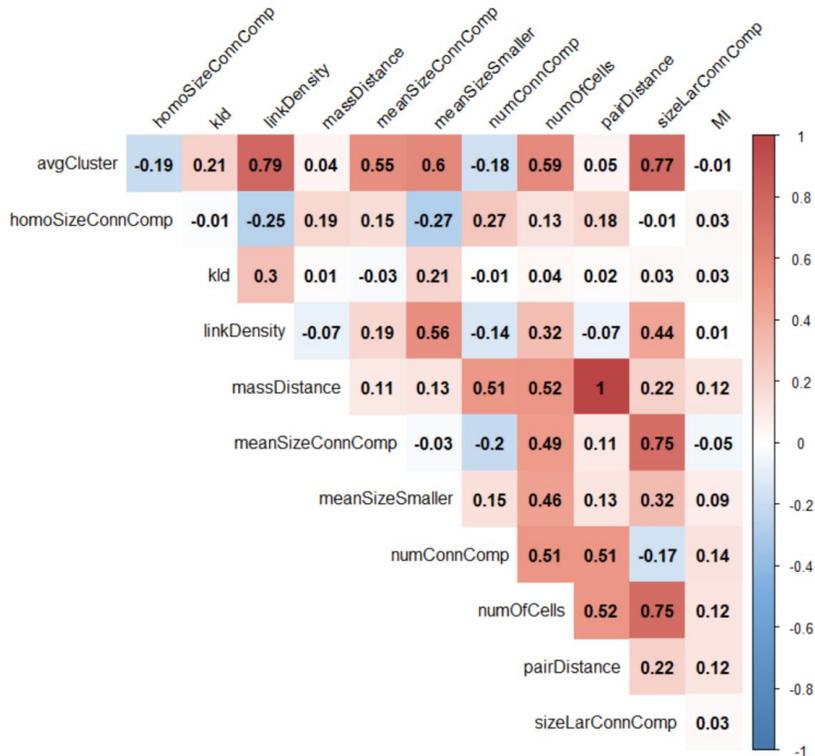


Figure 14. Colored correlation matrix for 12 morphological measures.

Here, the most notable measure is *kld*. We used KL divergence as a measure for detecting nontrivial spatial patterns. In Figure 13 (c), *kld* was largest at $K = 2$ unlike the intuition that the more patterns of larger than one are, the more nontrivial patterns are produced, which means that nontrivial morphogenetic patterns can be generated most frequently when the properties of GRNs are critical. It can arise from that the group of $K = 1$ gets to have many homogeneous and circular patterns by the influence of one cell fate. In *MI* of Figure 13 (l), the value was lowest at $K = 1$ despite the most number of cells, which implies there were many patterns where cell states had one cell fate, especially proliferation. In this case, because one kind of parameters of SMD kinetics (k , l , c of

[*proli - proli*]) between cells were applied, homogeneous and circular patterns were generated.

We will explain the creation of the nontrivial morphogenetic patterns at the criticality in more detail in the next chapter, using morphogenetic systems where evolutionary perturbations are added.

Chapter 5 How the Criticality of Gene Regulatory Networks Affects the Resulting Morphogenesis under Genetic Perturbations

In Chapter 4, we presented a GRN-based morphogenetic model and revealed the role of the criticality of GRNs in morphogenesis [34, 35]. The results in Chapter 4 include all kinds of randomly generated systems. However, real biological systems are products of evolution and therefore the results in Chapter 4 may have been affected by the inclusion of lots of biologically irrelevant data. Thus, in this chapter, we assume that biologically relevant GRNs are robust and evolvable [14, 38, 47, 51, 62, 65], and filter biologically irrelevant GRNs based on the criterion. We continue to use the same model as the one used in Chapter 4. Using the morphogenetic systems with robust and evolvable GRNs against genetic perturbations (e.g., mutations), we examine whether the role of the criticality of GRNs can be maintained even in the presence of the evolutionary perturbations. In addition, we investigate what the resulting morphologies are like and what kind of biological implications they have from the epigenetic viewpoint in morphology.

5.1 Experiments

We performed 10,000 independent simulation runs for $K = 1, 2, 3, 4$. Specifications of parameters for the simulations were the same as those in Chapter 4. In this section, we just describe updated parts (perturbations to GRNs and basin & cell fate entropy). The 12 measures in 4.2.1. of Chapter 4 are identically applied here.

5.1.1 Perturbations to GRNs and Robust & Evolvable GRNs

In this chapter, we introduced genetic perturbations changing the topology of GRNs in our morphogenetic model. Specifically, we assumed that the genetic perturbation was due to a germinal mutation occurring in a pre-zygotic cell, which is a small scale mutation at a genetic level. The germinal mutation is passed on to offspring, and it is present in all resulting cells during embryo development [21, 25]. We perturbed the intracellular GRN of a seed cell at the initial time step by adding, deleting, or switching one regulatory link between a pair of genes [17, 26, 42]. Because cells were duplicated through the process of cell division from the perturbed GRN in the seed cell, all the cells composing a morphogenetic pattern had the same perturbed GRNs.

Such a small regulatory link perturbation did not significantly change the average number of input links per node (K). For example, if $K = 2$, the total number of links of a GRN is 32, because the node size is 16. If one regulatory link is deleted as a genetic perturbation, the GRN consists of 31 links, making the value of $K = 1.94$.

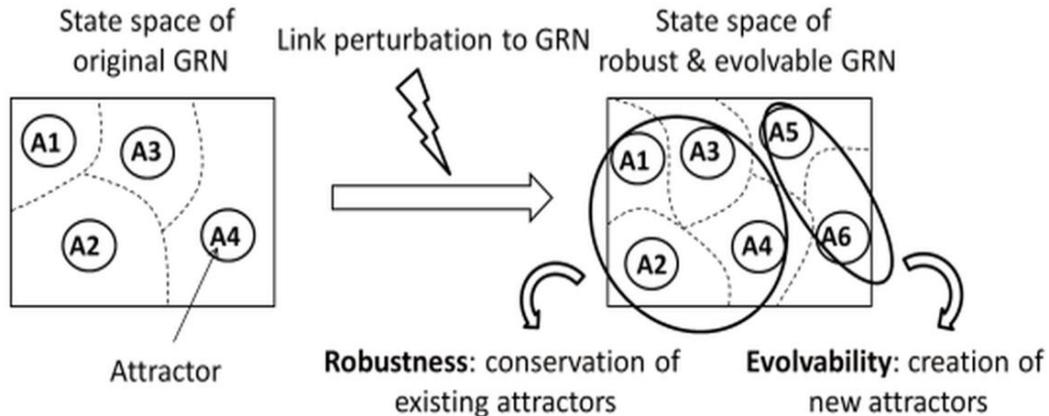


Figure 15. Schematic diagrams illustrating the concept of a robust and evolvable GRN.

To obtain morphogenetic systems that have only biologically relevant GRNs, we focused on robust and evolvable GRNs among the perturbed GRNs. In our model, if the GRN conserved its existing attractors and created new attractors simultaneously after the perturbation, we considered the GRN as a robust and evolvable GRN (Figure 15) [2]. This is because it means that existing cellular functions such as proliferation and differentiation were maintained and at the same time new cellular functions emerged.

5.1.2 Measures to Investigate the Relationship Between GRNs and Expressed Cell Fates

To investigate the relationship between intracellular GRNs and expressed cell fates, we calculated the basin entropy and cell fate entropy from the sizes of basins of attractions and cell fates distributed in a morphogenetic pattern. We thought that the numbers of actually expressed cell fates in a morphology might be proportional to the basin sizes of attractors where each cell fate was assigned. We calculated basin and cell fate entropies to look into whether or not our expectation would be correct. As in the computation of *MI*, only three cell fates (proliferation, differentiation, and quiescence) were considered (i.e., apoptosis was ignored).

i. Basin entropy

$$H_{basin} = - \sum_{\rho} P_{\rho} \cdot \log_2 P_{\rho}$$

where P_{ρ} is the size of the basin of the attractor ρ (to which proliferation, differentiation, or quiescence was assigned) divided by the sum of sizes of all the basins except for the basin size of the attractor for apoptosis. Thus,

$$\sum_{\rho} P_{\rho} = 1.$$

Originally, basin entropy was suggested by Krawitz as a measure of the complexity of information that a system can store in NK Boolean networks [39]. We used it as a measure to examine the versatility of the three cell fates (proliferation, differentiation, quiescence).

ii. Cell fate entropy

$$H_{cell\ fate} = - \sum_f P_f \cdot \log_2 P_f$$

where P_f is the number of cells expressing a cell fate f (proliferation, differentiation, quiescence), divided by the numbers of cells (except for those in apoptosis) in a morphogenetic pattern at the final time step. Hence,

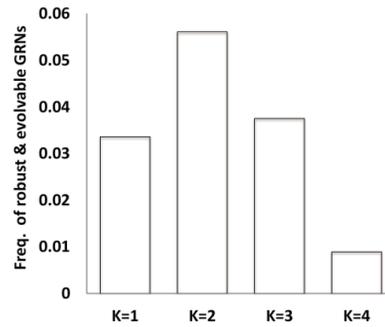
$$\sum_f P_f = 1.$$

In the case that there were no cells expressing a fate (proliferation, differentiation, quiescence), its log value was set to 0. Also, when there was no cell in the space, the value was set to 0.

5.2 Results

Figure 16 (a) shows probabilities of producing robust and evolvable GRNs against perturbations for $K = 1, 2, 3,$ and 4 in 10,000 simulation runs. We found that robust and evolvable GRNs were generated with the highest probability at $K = 2$. Figure 16 (b) shows samples of visualized morphogenetic patterns produced by robust and

(a)



(b)

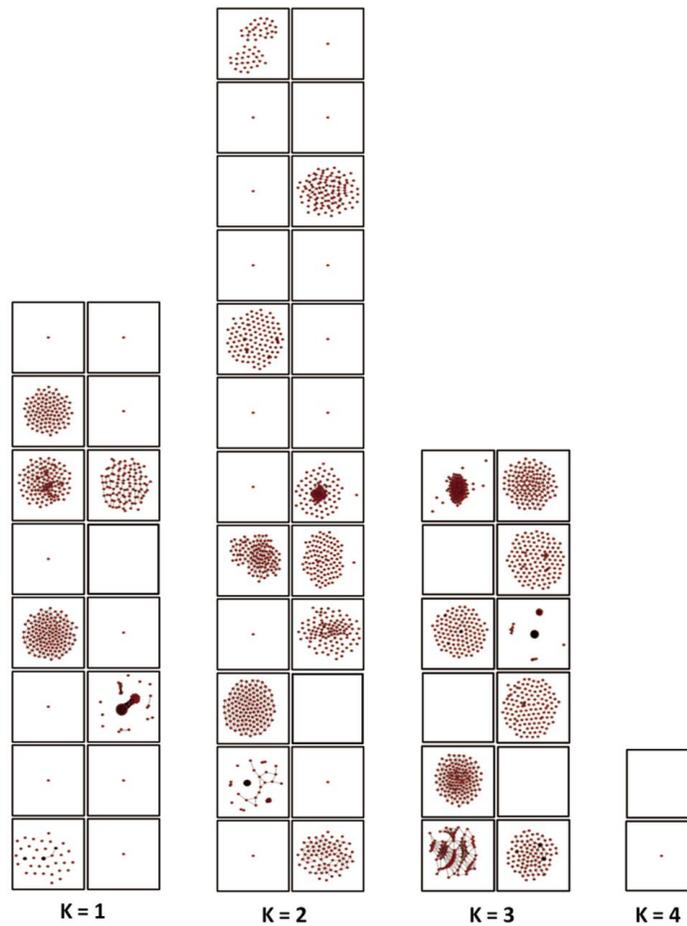


Figure 16. Frequencies of robust & evolvable GRNs per group ($K = 1, 2, 3, 4$) and visualized spatial patterns. (a) Probabilities of generating robust and evolvable GRNs for $K = 1, 2, 3, 4$ in 10,000 simulation runs. (b) Different morphogenetic patterns obtained from robust and evolvable GRNs for $K = 1, 2, 3, 4$. The numbers of the patterns were counted from robust and evolvable GRNs produced in 500 simulation runs.

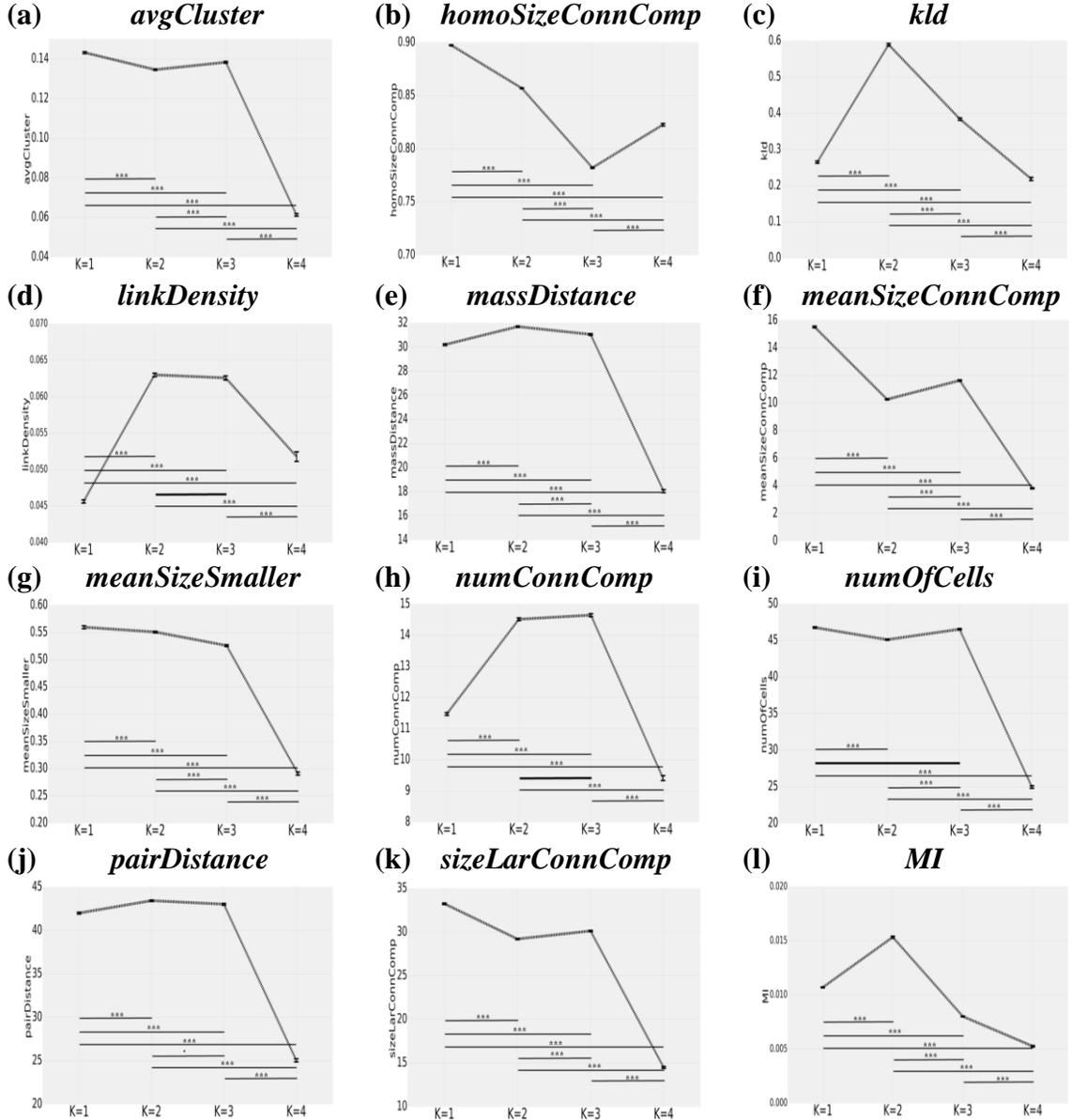


Figure 17. Comparison of means between groups ($K = 1, 2, 3, 4$) for 12 morphological measures (Kruskal-Wallis test: $p < 2.2e-16$, Nemenyi test (post-hoc): ' ': $p < 1.0$, '.': $p < 0.1$, '*': $p < 0.001$). In the case that there is no difference between two groups, a bold line without an asterisk is presented in the plot. (a) Average clustering coefficient (*avgCluster*). (b) Homogeneity of sizes of connected components (*homoSizeConnComp*). (c) KL divergence between pairwise particle distance distributions of morphogenetic pattern and a random pattern (*kld*). (d) Link density (*linkDensity*). (e) Average distance of cells from center of mass (*massDistance*). (f) Average size of connected components (*meanSizeConnComp*). (g) Average size of connected components smaller than the largest one (*meanSizeSmaller*). (h) Number of connected components (*numConnComp*). (i) Number of cells (*numOfCells*). (j) Average pairwise distance (*pairDistance*). (k) Size of the largest connected component (*sizeLarConnComp*). (l) Mutual information between cell fates of cells and their neighboring cells (*MI*).**

evolvable GRNs for each K .

We calculated the 12 morphological measures of morphogenetic patterns generated by robust and evolvable GRNs. Because the robust and evolvable GRNs were generated with different probabilities for different values of K , we applied bootstrap sampling 1,000 times to the values of the 12 measures for comparison between groups ($K = 1 - 4$) with unequal sample sizes. Figure 17 indicates the comparison of means between groups for the measures, where Kruskal-Wallis and Nemenyi (as post-hoc analysis) tests were performed to show statistically significant differences among the groups.

Furthermore, we produced a correlation matrix to investigate correlations between the 12 measures (Figure 18). We found that the following six measures were highly correlated with *numOfCells*: *avgCluster*, *massDistance*, *meanSizeConnComp*, *meanSizeSmaller*, *pairDistance*, and *sizeLarConnComp*. These correlations were found in

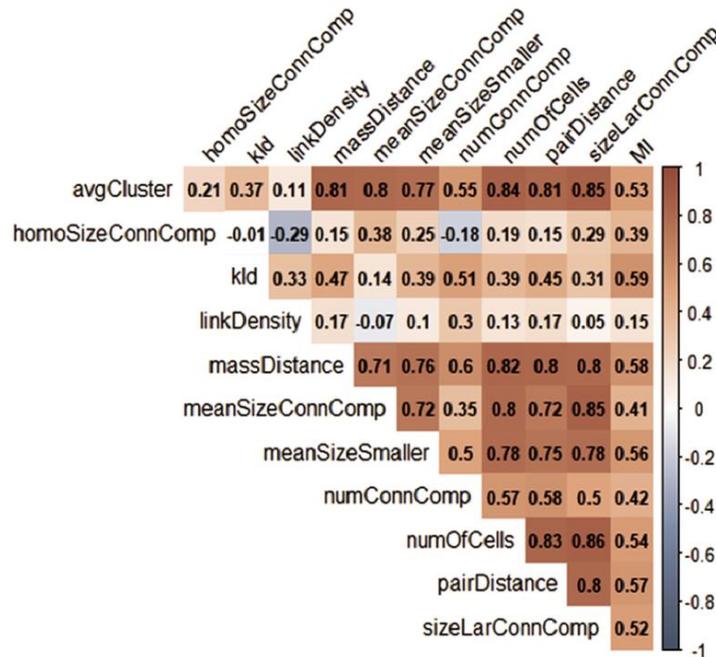


Figure 18. Correlation matrix for 12 morphological measures.

Figure 17 as well. The values of *numOfCells* at $K = 1, 2, 3$ were similar but fell sharply at $K = 4$. This trend was also shown in the six measures highly correlated with *numOfCells*. Meanwhile, *kld*, *MI*, *homoSizeConnComp*, *linkDensity*, and *numConnComp* showed different trends. In Figure 17 (c), *kld* was highest at $K = 2$, which means that nontrivial morphogenetic patterns were generated most frequently when the GRNs were critical under the genetic perturbations. This result demonstrates that the role of criticality of GRNs is maintained even in the presence of evolutionary perturbations.

In addition, from *MI*, *homoSizeConnComp*, *linkDensity*, and *numConnComp*, we found two interesting properties of the nontrivial morphologies at the criticality. Firstly, certain combinations of cell fates between neighboring cells occurred most frequently. In Figure 17 (l), *MI* was highest at $K = 2$. It indicates that the fate of a cell is strongly correlated with the fate of its neighboring cells in a morphogenetic pattern generated at the criticality. To examine the relationship between intracellular GRNs and expressed cell fates, we measured basin entropy and cell fate entropy (Figure 19). Our original expectation was that if the basins of attraction for the three cell fates were most evenly

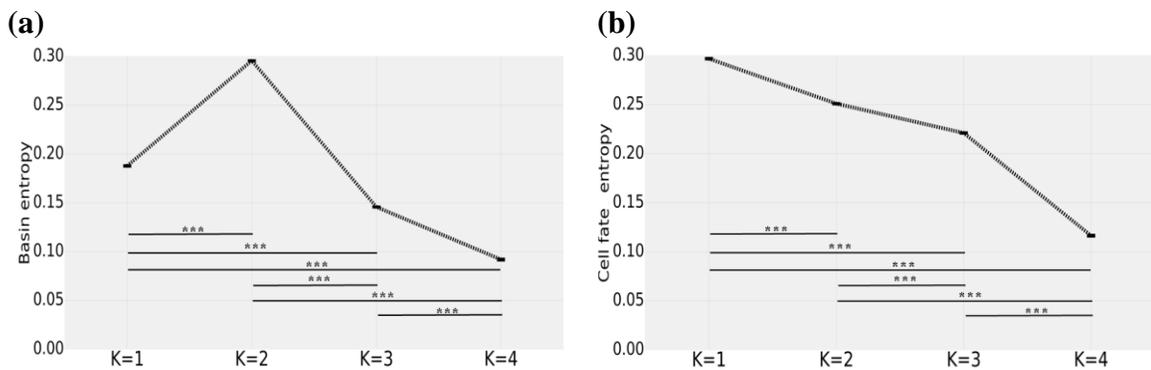


Figure 19. Comparison of means between groups for basin and cell fate entropy computed from three cell fates (proliferation, differentiation, quiescence). (a) Average basin entropy for $K = 1, 2, 3, 4$. (b) Average state entropy of cell fates performed in a simulation at the final time step for $K = 1, 2, 3, 4$.

distributed at $K = 2$, the expressions of different cell fates would be maximally balanced in a morphogenetic pattern. However, the cell fate entropy was highest at $K = 1$ although the basin entropy was highest at $K = 2$. This means that the distribution of cell fates in a morphology was not a simple reflection of the basin sizes of a GRN at a single cell level, but more like an emergent property at a multicellular level obtained through the developmental process involving cell-cell interactions.

Secondly, the nontrivial morphologies emerged typically in topologically homogeneous cell clusters. In Figure 17 (b), (d), (h), $K = 1$ showed relatively high *homoSizeConnComp*, low *linkDensity*, and low *numConnComp* values, on average, compared to the corresponding measures at $K = 2, 3$, and 4. Here we simply express the observations qualitatively as (high, low, low). Similarly, in the same order, $K = 2$ showed (high, high, high), $K = 3$ showed (low, high, high), and $K = 4$ showed (low, low, low). These can be interpreted as follows: The morphologies at $K = 1$ consisted of homogeneous large-size connected components which had a shape like a long chain. The morphologies at $K = 2$ were composed of homogeneous-size connected components where cells were interconnected. The morphologies at $K = 3$ were composed of heterogeneous-size connected components where cells were interconnected. The morphologies at $K = 4$ were composed of heterogeneous-size small connected components that had a shape like a short chain. Figure 20 summarizes the typical topological properties for $K = 1 - 4$ schematically.

In our morphogenetic model, there is a feedback relationship (Figure 21). Interactions with neighboring cells determine a cell fate. Depending on the cell fates, the parameters of SMD kinetics are applied. The cells are positioned by SMD kinetics. The

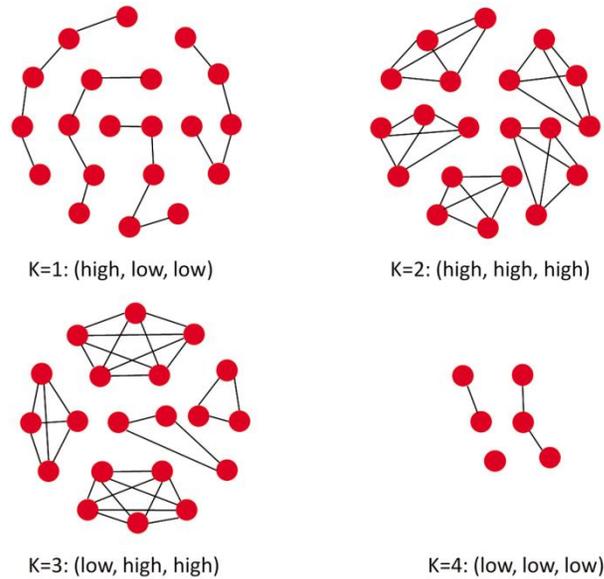


Figure 20. Topological properties of morphogenetic patterns for $K = 1, 2, 3, 4$. “low” and “high” mean the relative values against K in the order of (b) *homoSizeConnComp*, (d) *linkDensity*, (h) *numConnComp* in Figure 17.

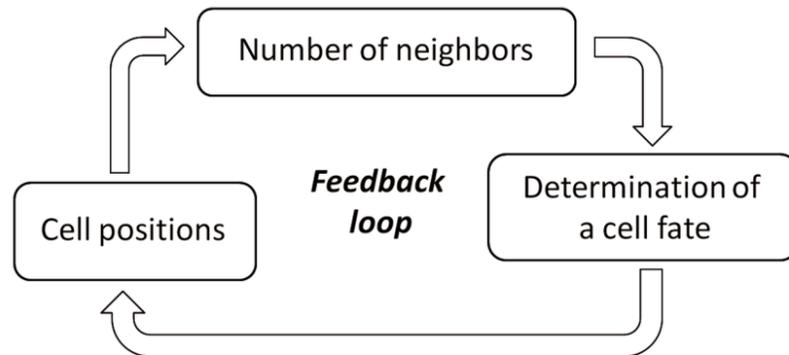


Figure 21. Feedback relationship in our morphogenetic process.

positions of cells influence the number of neighboring cells. In the feedback relationship, we found that the nontrivial morphologies were produced most frequently when the GRNs were critical under the genetic perturbations. Besides, the nontrivial morphologies

at criticality had the most frequent occurrence of certain combinations of cell fates between neighbors, and were composed of topologically homogeneous cell clusters. Because the parameter values of SMD kinetics determining cells' positions depended on the cell fates, the more frequent those combinations of cell fates between neighbors were, the more likely to be applied the same SMD parameter values were among the cells. Thus, the most frequent combinations of cell fates between neighbors would naturally produce more homogeneous-size connected components where cells were interconnected. Such spatial arrangements of multicellular patterns due to the criticality of GRNs are not easily predictable from a single cell level.

Chapter 6 Robustness and Evolvability of Multilayer Gene Regulatory Networks

In Chapter 4 and 5, because we used particle-based morphogenetic models with SMD kinetics, it was difficult to assess the robustness and evolvability of the whole system. Thus, we present a more formal hierarchical network model and investigate how criticality of GRNs affects the robustness and evolvability of the whole system, the hierarchical network depending on cellular topologies and the number of links of an intercellular network in this chapter.

6.1 Model: Multilayer GRNs

We present multilayer GRNs consisting of an intercellular layer and an intracellular layer. A network in an intercellular layer represents interactions between cells, and a network in an intracellular layer indicates interactions between genes (Figure 22). All the nodes of an intercellular network have identical RBNs as intracellular networks. Our multilayer GRNs are divided into two types depending on cellular topologies. One is the multilayer GRNs having fixed cellular topologies, and the other is the multilayer GRNs having cellular topologies that are randomly changed in each simulation run. The multilayer GRNs with static cellular topologies are assumed as epithelial cells based on the existing models representing epithelium as the square arrangement of cells having a fixed set of neighbors (the adjacent neighboring cells: north, south, east, west) [20, 45] (Figure 23. (a)). The multilayer GRNs with dynamic cellular

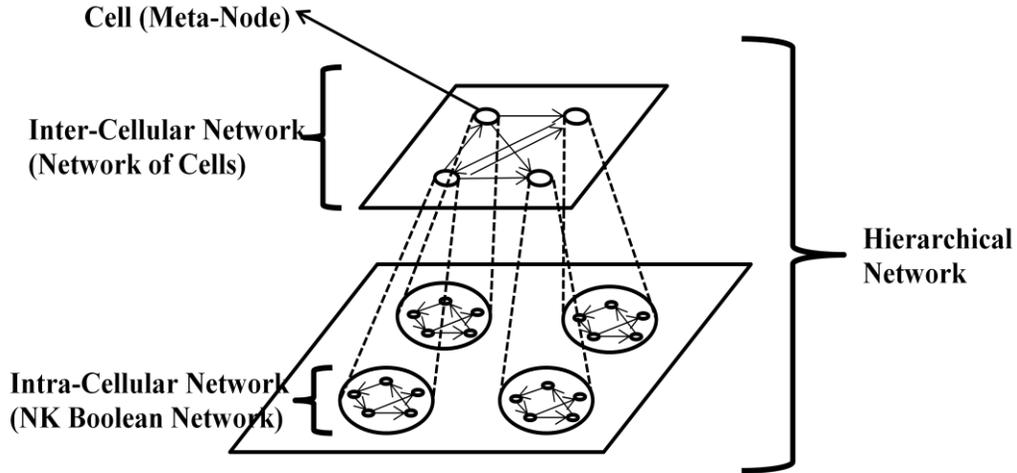
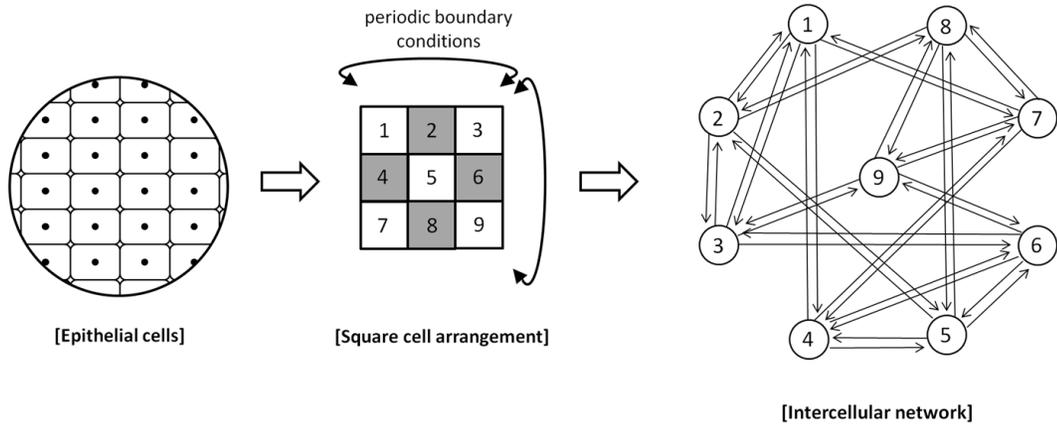


Figure 22. A schematic diagram for example multilayer GRNs with $N_{inter} = 4$, $N_{intra} = 5$, $K_{intra} = 1$. (In actual simulations, $N_{inter} = 9$, $N_{intra} = 6$ nodes were used.)

topologies are assumed as a developing embryo (Figure 23. (b)). This assumption is based on biological evidence showing that the topology of the intercellular network keep changing because of cellular movements and cell growth during embryonic development [30]. For the two types of multilayer GRNs, we generate multilayer GRNs taking ordered, critical, and chaotic intracellular GRNs by adjusting node in-degree (K).

The dynamics of multilayer GRNs as the whole system at a hierarchical level are determined by the dynamics of intracellular GRNs (the input nodes of each gene and the assigned Boolean functions to the genes) and the topology of the intercellular network (the neighboring cells for the interactions between cells). In our multilayer GRNs, we implement cell signaling for the interactions between cells, following Villani et al.'s coupled RBN model [64]. In an intracellular GRN, a certain gene is assigned to communicate with neighboring cells. This gene is called communicating gene. The communicating gene is activated if any of the communicating genes of neighboring cells

(a)



(b)

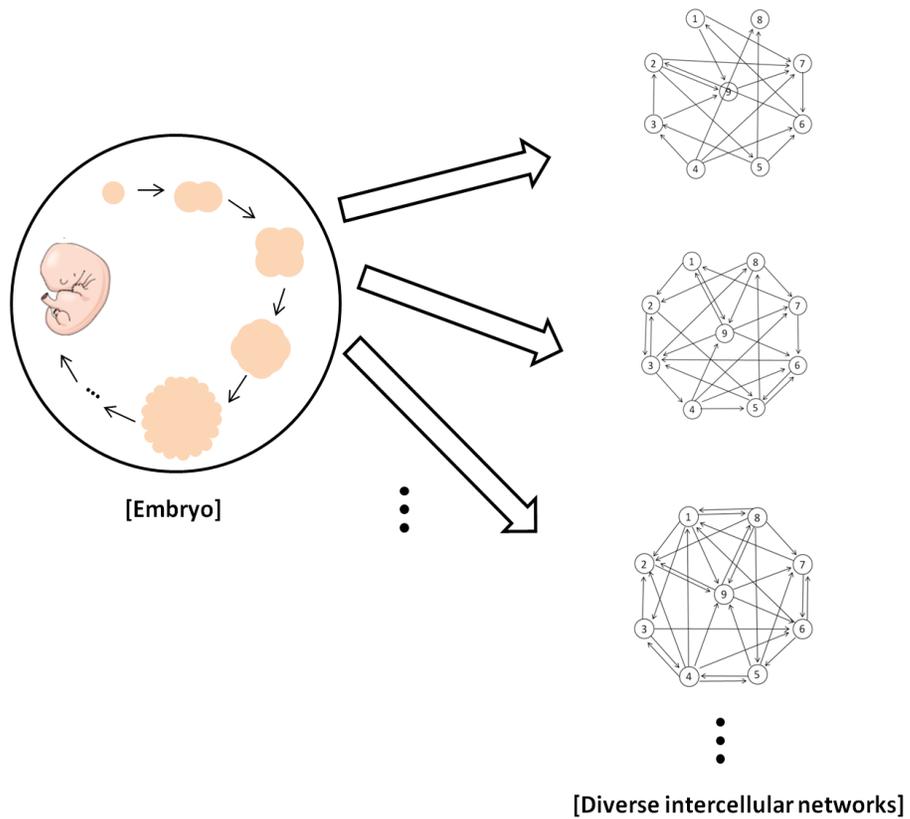


Figure 23. Biological systems and intercellular networks. (a) Epithelial cells having static cellular topologies. (b) Embryo having dynamic cellular topologies.

are activated. The states of the other genes except for the communicating gene are updated by the input nodes of each gene and randomly assigned Boolean functions to the genes in the intracellular GRN.

Figure 24. (a) shows an example GRN. In the GRN, *gene2* is a communicating gene. The assigned Boolean functions to *gene1* and *gene2* are shown in Figure 24. (b).

Figure 25. (a) illustrates example multilayer GRNs where each cell has the network of Figure 24. (a) as its intracellular GRN. *cell A* and *cell B* are a neighbor to each other.

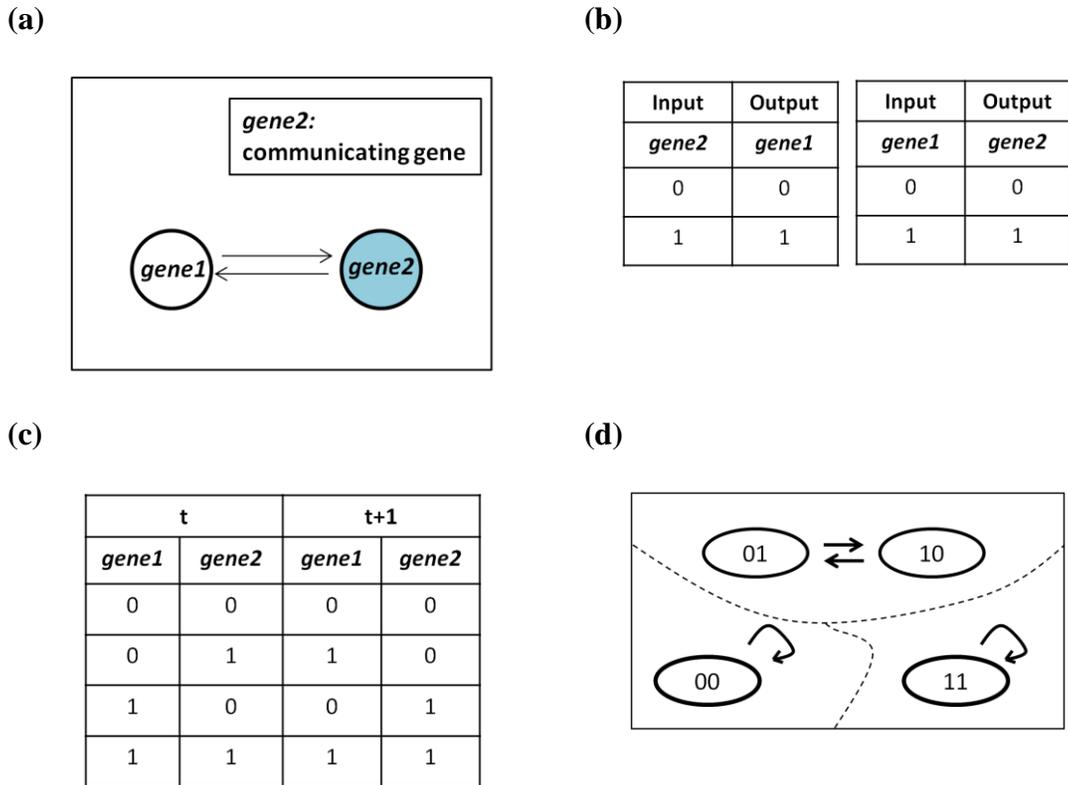


Figure 24. Schematic diagrams for explaining the dynamics of an intracellular GRN. (a) An intracellular GRN with $N_{intra} = 2$, $K_{intra} = 1$. (b) Boolean functions randomly assigned to each node. (c) State transition table of the intracellular GRN. (d) State space of the intracellular GRN. The state space consists of $2^2 = 4$ configurations and transitions among them. The configurations with bold lines are attractors. Dashed lines draw boundaries for each basin of attraction.

Figure 25. (b) shows the state transition table representing the dynamics of the multilayer GRNs. For example, we assume that initial states of genes are 0110 at time t in the state transition table. For the communicating node, *gene 2* of *cell A* becomes 0 at the next time step because *gene 2* of *cell B* at t is not activated. On the contrary, *gene 2* of *cell B* becomes 1 because *gene 2* of *cell A* is activated. For the non-communicating node, *gene 1* of *cell A* becomes 1 at $t + 1$ by the state of *gene2* in *cell A* and the assigned Boolean

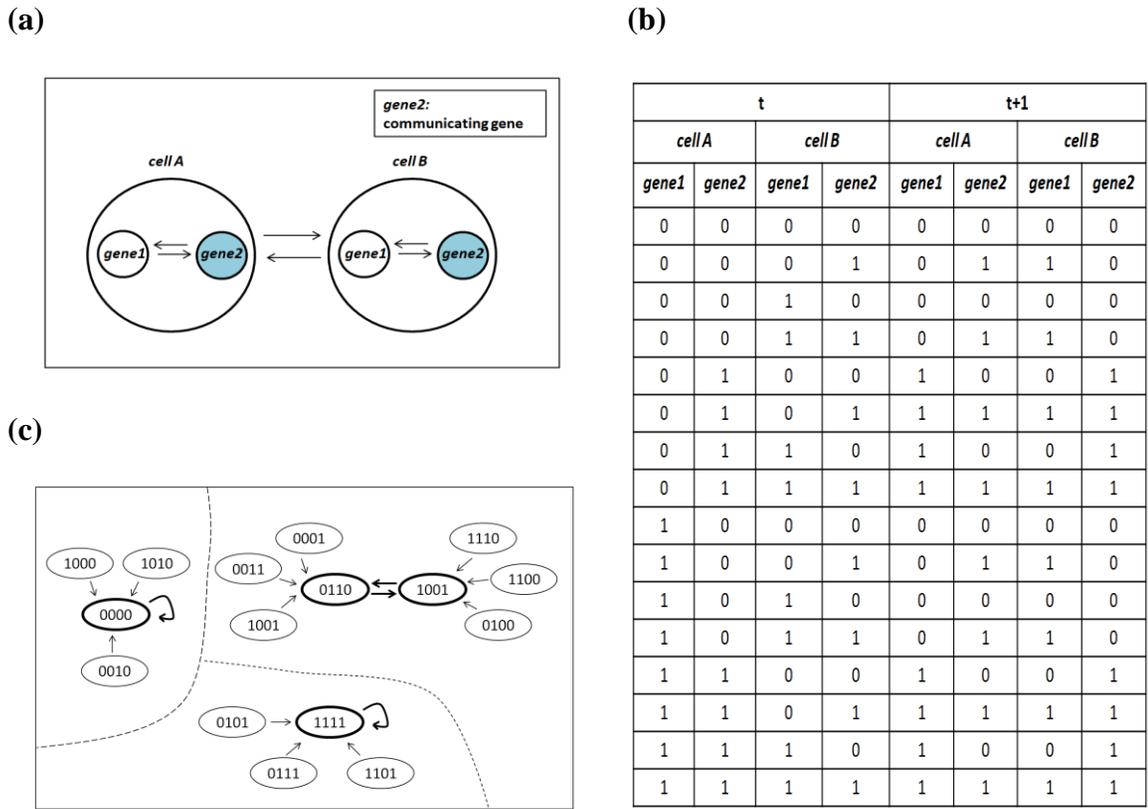


Figure 25. Schematic diagrams for explaining the dynamics of multilayer GRNs. (a) An Intercellular GRN with $N_{inter} = 2$ containing the intracellular GRNs of Figure 24. (b) State transition table of the multilayer GRNs. (c) State space of the multilayer GRNs. The state space consists of $2^{N_{intra} \times N_{inter}} = 2^{2 \times 2} = 16$ configurations and transitions among them. The configurations with bold lines are attractors. Dashed lines draw boundaries for each basin of attraction. (Because $N_{inter} = 9$, $N_{intra} = 6$ nodes were used in actual simulations, the state space size of the multilayer GRNs is $2^{6 \times 9} = 2^{54}$.)

function to *gene1*. Similarly, *gene 1* of *cell B* becomes 0 at $t + 1$ by the state of *gene 2* in *cell B* and the assigned Boolean function. As the result, the initial states 0110 finally become 1001 at $t + 1$. In this way, all the values of the state transition table of the multilayer GRNs can be filled in.

6.2 Experiments

We conducted the following two computational experiments:

- (1) Robustness and evolvability of multilayer GRNs and intracellular GRNs depending on cellular topologies:

In multilayer GRNs with static cellular topologies, cells always had the fixed four neighboring cells in each simulation. Meanwhile, in multilayer GRNs with dynamic cellular topologies, cells had different neighboring cells in each simulation because the topology of an intercellular network was randomly determined based on the number of links randomly chosen between 1 and 81.

When intracellular GRNs were ordered ($K = 1$), critical ($K = 2$) and chaotic ($K = 3$), we assessed the robustness and evolvability of multilayer GRNs and intracellular GRNs for static and dynamic cellular topologies. Here we omitted $K = 4$ to simplify the experimental process.

- (2) Robustness and evolvability of multilayer GRNs depending on the number of links of an intercellular network:

For multilayer GRNs taking ordered ($K = 1$), critical ($K = 2$) and chaotic ($K = 3$) intracellular GRNs, increasing the number of links of an intercellular network

from 10 to 80 by 10, we assessed the robustness and evolvability of multilayer GRNs with dynamic cellular topologies.

We performed 1,000 independent simulation runs for each group ($K = 1, 2, 3$) of multilayer GRNs. The parameter values for the simulations are given in Table 2.

Table 2. Parameters of the multilayer GRNs and their values

Model	Parameter	Value
Intracellular Layer	Number of nodes (N_{intra})	6
	Number of in-degree per node (K_{intra})	1, 2, 3
	Internal homogeneity (p)	0.5
	Number of communicating nodes (N_c)	1
Intercellular Layer	Number of nodes (N_{inter})	9
	Number of links (L_{inter})	(1) static: $L_{inter} = 36$ dynamic: $L_{inter} \in \text{unif}(1, 81) \subset \mathbb{Z}$ (2) $L_{inter} \in \{10, 20, 30, 40, 50, 60, 70, 80\} \subset \mathbb{Z}$

6.2.1. Robustness and Evolvability Against Genetic Perturbations

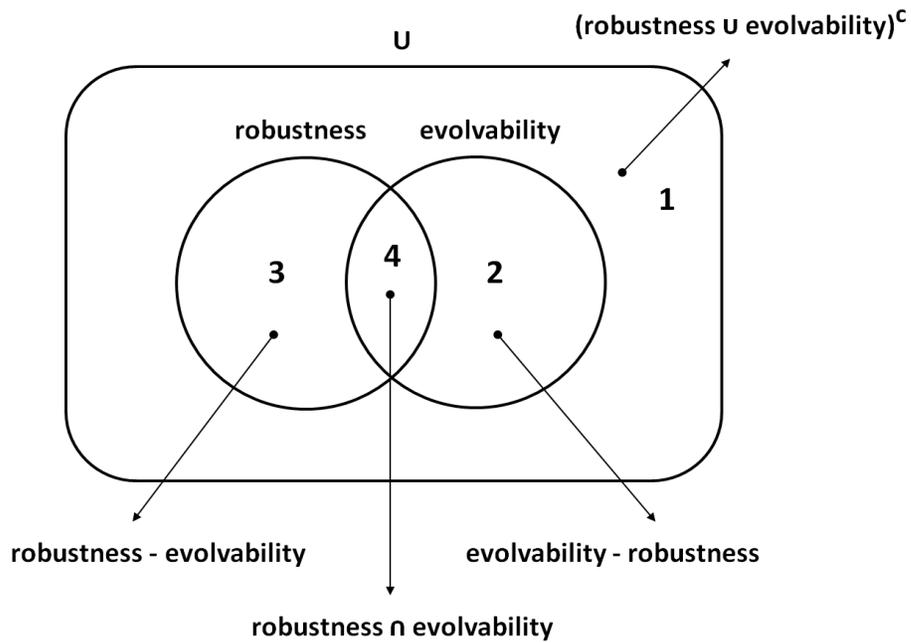
In Chapter 5, we added a germinal mutation occurring in a pre-zygotic cell to our GRN-based morphogenetic model. In this chapter, adding another kind of mutation, we

investigated the role of the criticality of GRNs under the different form of mutation. Here we introduced somatic mosaic mutations occurring in post-zygotic cells to our multilayer GRNs. Somatic mutations, which are known to be present in actual embryo development, are not inherited genetic alterations in the course of cell division [19, 22, 44]. A mosaic means that two or more populations of cells with different genotypes exist in one individual developed from a single zygotic cell [63]. In our multilayer GRNs consisting of nine cells, we perturbed an intracellular GRN in one cell. We added, deleted, or switched one regulatory link between a pair of genes [17, 26, 42]. As a result, different intracellular GRNs came to exist within one system consisting of nine cells.

We measured the robustness and evolvability of multilayer GRNs at a hierarchical level and intracellular GRNs at a single cell level against the genetic perturbations. Figure 26. (a) shows four categories depending on the properties. The categories are as follows:

- $robustness \cap evolvability$ (4): If existing attractors were conserved and new attractors were created simultaneously against the genetic perturbation, the GRN was considered as a robust & evolvable GRN [2].
- $robustness - evolvability$ (3): If only existing attractors were conserved without new attractors being created, the GRN was regarded as a robust GRN.
- $evolvability - robustness$ (2): If new attractors were created without existing attractors being conserved, the GRN was regarded as an evolvable GRN.
- $(robustness \cup evolvability)^c$ (1): The GRN which did not belong to any categories above was included into this category.

(a)



(b)

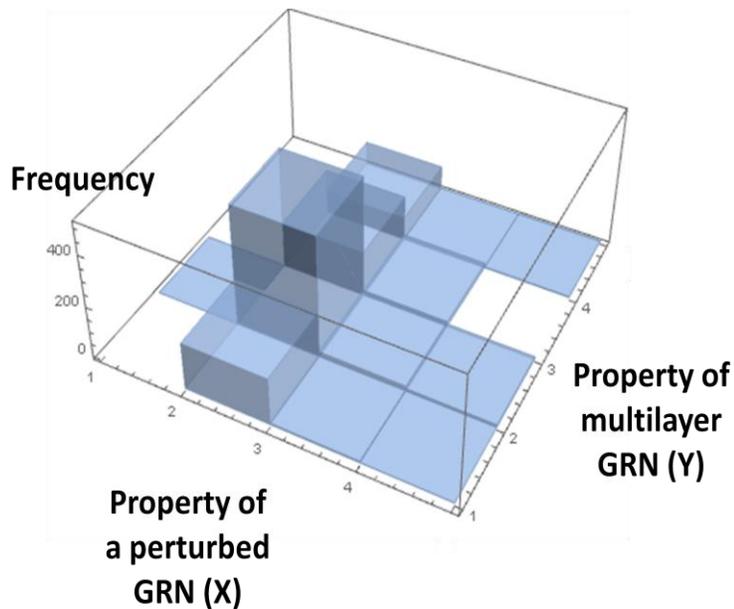


Figure 26. Four categories of the robustness and evolvability, and the relationship between the properties of multilayer GRNs and intracellular GRNs. (a) A Venn diagram representing different sets depending on the changes of attractors. (b) An example 3D histogram illustrating the degree of correlation between properties of multilayer GRNs and intracellular GRNs.

To investigate linear dependence between the properties of multilayer GRNs and intracellular GRNs, we assigned integer values to the four categories in order of 4, 3, 2, and 1. Using the integers representing the properties of intracellular GRNs and multilayer GRNs as coordinates (X, Y), we made a 3D histogram (Figure 26. (b)). We thought that the properties of multilayer GRNs at a hierarchical level might be the same as the properties of intracellular GRNs at a single cell level ($Y = X$). For example, if a perturbed intracellular GRN is robust against the genetic perturbation, the multilayer GRNs containing the perturbed GRN would be also robust. We calculated correlation coefficients between the properties of multilayer GRNs and intracellular GRNs to look into whether or not our expectation would be correct.

When finding the attractors of multilayer GRNs at a hierarchical level, we focused on the attractors with the largest basins of attraction, which is for keeping computational loads reasonable. Because the state space size of the multilayer GRNs is 2^{54} , it is not feasible to explore all the state space. Thus, we used 10,000 randomly chosen initial states to find the attractors with the largest basins of attraction. The number of the initial states was determined based on studies identifying the attractors of large-scale Boolean networks [2, 37].

6.3 Results

Figure 27 shows probabilities of generating robust & evolvable GRNs, robust GRNs, and evolvable GRNs depending on cellular topologies. We focused on multilayer GRNs and intracellular GRNs which had robustness and evolvability simultaneously

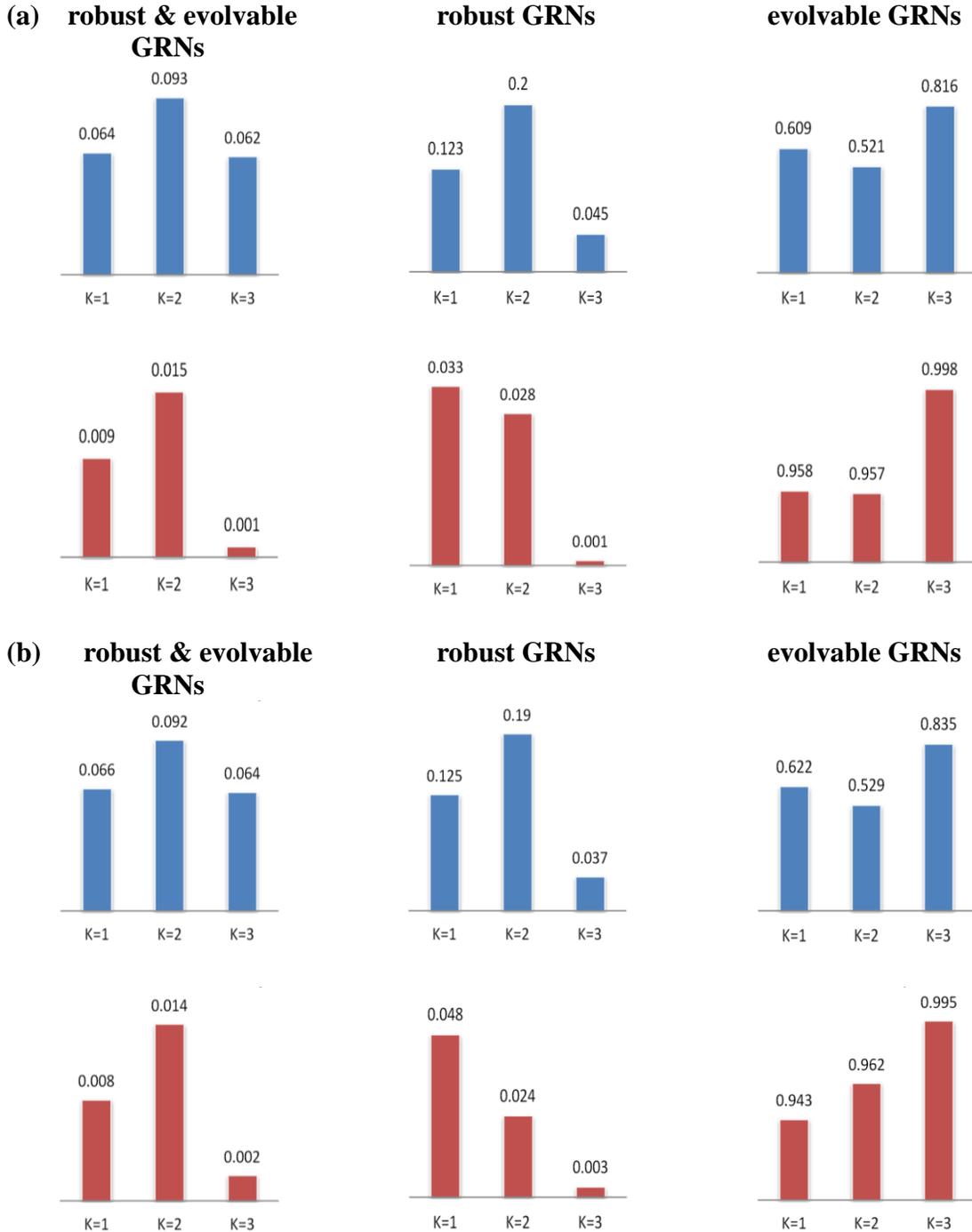


Figure 27. Probabilities of generating robust & evolvable GRNs, robust GRNs, and evolvable GRNs depending on cellular topologies. The blue graphs represent the robustness and evolvability of perturbed intracellular GRNs at a single cell level, and the red ones represent the robustness and evolvability of multilayer GRNs containing the perturbed intracellular GRNs at a hierarchical level. (a) Robustness and evolvability of multilayer GRNs with static cellular topologies. (b) Robustness and evolvability of multilayer GRNs with dynamic cellular topologies.

against the genetic perturbations based on the studies showing that living organisms exhibit robustness and evolvability [14, 38, 47, 51, 62, 65]. We found that the robust & evolvable multilayer GRNs and intracellular GRNs were generated with the highest probabilities at $K = 2$. In addition, the multilayer GRNs with both static and dynamic cellular topologies had this trend in common. It means that the criticality of GRNs promotes not only the generation of robust & evolvable intracellular GRNs at a single cell level but also the production of robust & evolvable multilayer GRNs at a hierarchical level. Furthermore, the role of the criticality of GRNs maintains in both epithelial cells with static cellular topologies and a developing embryo with dynamic cellular topologies.

To investigate correlation between the robustness and evolvability of multilayer GRNs and intracellular GRNs, we computed correlation coefficients in Table 3. All the values were smaller than 0.17, which indicates that there are almost no correlation or a very weak correlation between the properties of multilayer GRNs and intracellular GRNs. That is, although an intracellular GRN is robust against the genetic perturbation, multilayer GRNs can be not robust. The properties of multilayer GRNs are not simply determined by the properties of intracellular GRNs at a single cell level. Thus, the properties of multilayer GRNs obtained at a hierarchical level must be understood as emergent properties.

Table 3. Correlation coefficients between the properties of multilayer GRNs and intracellular GRNs

	$K = 1$	$K = 2$	$K = 3$
Static cellular topologies	0.059	0.168	0.041
Dynamic cellular topologies	0.141	0.134	0.102

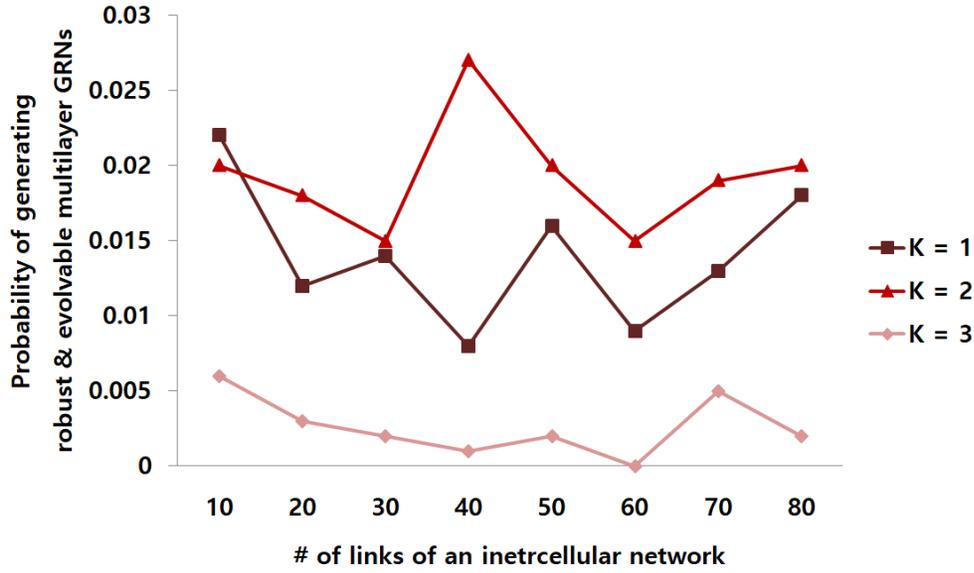


Figure 28. Probabilities of generating robust & evolvable multilayer GRNs depending on the number of links of an intercellular network.

Figure 28 shows probabilities of generating robust & evolvable multilayer GRNs depending on the number of links of an intercellular network. Increasing the number of links of an intercellular network from 10 to 80 by 10, we observed how the probabilities were varied. As the number of links of an intercellular network grew, the probabilities did not monotonically increase or decrease but fluctuated. In addition, when compared to the probabilities at $K = 1$ and $K = 3$, robust & evolvable multilayer GRNs were produced with the higher probability at $K = 2$. Especially, the probability of generating robust & evolvable multilayer GRNs reached the maximum when the number of links of an intercellular network was around 40 (link density ≈ 0.5). It means that the degree of interactions between cells can maximize the generation of robust & evolvable multilayer GRNs by amplifying the effect of the criticality of GRNs.

Chapter 7 Conclusions

In this dissertation, we explored the roles of the criticality of GRNs at a multicellular and hierarchical level, using our GRN-based morphogenetic model and multilayer GRNs.

In Chapter 4, we proposed GRN-based morphogenetic systems using Kauffman's RBNs as intracellular GRNs and SMD kinetics for cellular movements to show self-organized spatial patterns during the developmental process. Varying the properties of GRNs from ordered ($K = 1$), through critical ($K = 2$), to chaotic ($K = 3, 4$) regimes, we simulated our morphogenetic model. As a result, the simulations demonstrated that nontrivial morphogenetic patterns were produced most frequently in the morphogenetic systems with critical GRNs. Our finding indicates that the criticality of GRNs plays an important role in facilitating the formation of nontrivial morphogenetic patterns in the GRN-based morphogenetic systems.

In Chapter 5, we introduced genetic perturbations that change the interactions between genes (e.g., mutations) to our morphogenetic model. We looked into whether the role of the criticality of GRNs reported in Chapter 4 could be maintained even in the presence of evolutionary perturbations. Also, we investigated what the resulting morphologies were like and what kind of biological implications they had from the epigenetic viewpoint in morphology. We found that nontrivial morphologies were generated most frequently when the GRNs were critical under the genetic perturbations, which was consistent with the previous result obtained from morphogenetic systems

without evolutionary perturbations. Moreover, we found that the nontrivial morphologies at the criticality tended to be made of topologically homogeneous cell clusters due to the spatial arrangements in which certain combinations of cell fates between neighboring cells occurred most frequently. Based on these findings, we conclude that the criticality of GRNs facilitates the formation of nontrivial morphologies by adjusting the spatial arrangements of cells in GRN-based morphogenetic systems, even under the genetic perturbations.

Our findings in Chapter 4 and Chapter 5 have implications from an epigenetic viewpoint. Researchers in epigenesis have suggested that heterogeneous and complex features emerge from homogeneous and less complex components through the interactions among them [46, 53]. In our model, we showed that the nontrivial morphologies were produced most frequently at criticality, typically with topologically homogeneous cell clusters. Thus, the result not only supports the theory of epigenesis in developmental biology, but also implies that highly structured tissues or organs in morphogenesis of multicellular organisms might stem from cell aggregation with critical GRNs.

In Chapter 6, we presented multilayer GRNs consisting of an intercellular layer and an intracellular layer. We obtained probabilities of generating robust & evolvable multilayer GRNs and intracellular GRNs against genetic perturbations, varying the properties of intracellular GRNs with $p = 0.5$ from ordered ($K = 1$), through critical ($K = 2$), to chaotic ($K = 3$) regimes. We found that the robust & evolvable multilayer GRNs and intracellular GRNs were generated with the highest probabilities at $K = 2$ for both static and dynamic cellular topologies. Especially, the probability of generating robust &

evolvable multilayer GRNs reached the maximum when the link density of an intercellular network was around 0.5. Our finding means that the criticality of GRNs at a single cell level promotes the production of robust & evolvable multilayer GRNs at a hierarchical level. In addition, the effect of the criticality of GRNs can be amplified by the degree of interactions between cells.

Through the three studies in Chapter 4, 5, and 6, we found that the criticality of GRNs facilitated the formation of nontrivial morphologies at a multicellular level, and generated robust and evolvable multilayer GRNs with the highest probability at a hierarchical level. Our findings demonstrate that the roles of the criticality of GRNs are hard to be discovered through the single-cell-level studies by showing that the formation of nontrivial morphologies and the generation of robust & evolvable multilayer GRNs must be understood as not predictable properties at a single cell level but emergent properties at a higher system level. It justifies the value of our research on the relationship between criticality of GRNs and properties of organisms in the context of multicellular settings.

The present studies have limitations. The properties of morphogenetic patterns and the robustness & evolvability of multilayer GRNs were explored only using our artificial model based on RBNs as GRNs. To make our findings more relevant to real biological systems, we need to develop more biologically plausible models, using empirically obtained biological Boolean networks.

For future work, we plan to look into the spatial and temporal distribution of cells during the growing processes from the seed cell to the cell aggregation to fully account for why the nontrivial morphologies were produced most frequently in

morphogenetic systems with critical GRNs. Also, we will examine how not only the attractors of multilayer GRNs but also the basins of the attraction are changed by the genetic perturbations to thoroughly explain why the robust & evolvable multilayer GRNs were generated with the highest probability when intracellular GRNs were critical.

Appendix A

The dynamics of RBNs are divided into three regimes depending on the structure of state space: ordered, critical, and chaotic. One can know which dynamics RBNs have by drawing a Derrida plot [15, 16]. Steps to draw a Derrida plot are as follows:

- (1) Randomly choose two initial states $S_1(t)$ and $S_2(t)$ that are close to each other in the state space of a RBN.
- (2) Obtain $S_1(t + 1)$ and $S_2(t + 1)$.
- (3) Calculate Hamming distances $X = H(t) = |S_1(t) - S_2(t)|$ and $Y = H(t + 1) = |S_1(t + 1) - S_2(t + 1)|$.
- (4) Plot coordinate (X, Y) .
- (5) Iterate the above steps several times.

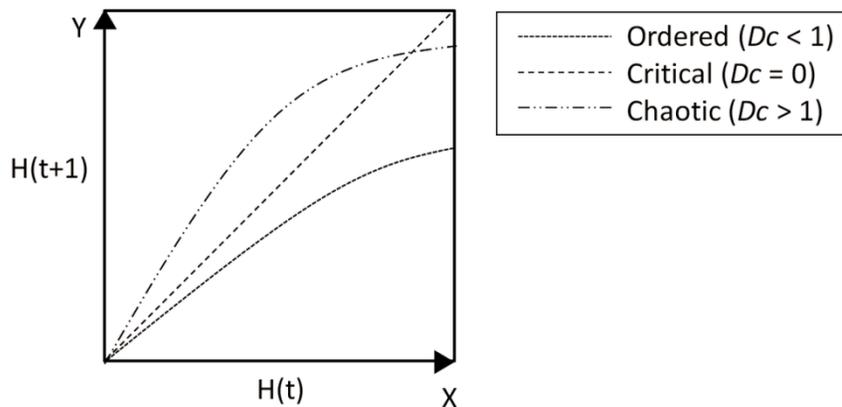


Figure 29. Derrida plot representing dynamics of GRNs.

In a Derrida plot, the Derrida coefficient (D_c) is defined as $\log s$, where s is the slope of the Derrida curve at the origin. $D_c < 0$ means that the dynamics of GRNs are ordered. $D_c = 0$ indicates critical dynamics, and $D_c > 0$ represents chaotic dynamics (Figure 29).

Notes

The materials presented in the dissertation are based on the published work [34, 35, 36].

References

1. Aldana, M. (2003). Boolean dynamics of networks with scale-free topology. *Physica D: Nonlinear Phenomena*, 185(1), 45-66.
2. Aldana, M., Balleza, E., Kauffman, S., & Resendiz, O. (2007). Robustness and evolvability in genetic regulatory networks. *Journal of Theoretical Biology*, 245(3), 433-448.
3. Aldana, M., Coppersmith, S., & Kadanoff, L. P. (2003). Boolean dynamics with random couplings. In E. Kaplan, J. E. Marsden, & K. R. Sreenivasan (Eds.), *Perspectives and Problems in Nonlinear Science* (pp. 23-89). New York: Springer.
4. Aminetzach, Y. T., Macpherson, J. M., & Petrov, D. A. (2005). Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. *Science*, 309(5735), 764-767.
5. Ball, P. (2009). *Shapes: Nature's patterns: A tapestry in three parts* (1st ed.). New York: Oxford University Press.
6. Balleza, E., Alvarez-Buylla, E. R., Chaos, A., Kauffman, S., Shmulevich, I., & Aldana, M. (2008). Critical dynamics in genetic regulatory networks: Examples from four kingdoms. *PLOS ONE*, 3(6), e2456.
7. Bertram, J. S. (2000). The molecular biology of cancer. *Molecular Aspects of Medicine*, 21(6), 167-223.
8. Burrus, V., & Waldor, M. K. (2004). Shaping bacterial genomes with integrative and conjugative elements. *Research in Microbiology*, 155(5), 376-386.
9. Carroll, S. B., Grenier, J. K., & Weatherbee, S. D. (2013). *From DNA to diversity: Molecular genetics and the evolution of animal design* (2nd ed.). Oxford: John Wiley & Sons.
10. Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E., & Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*, 453(7194), 544-547.
11. Chen, J., Miller, B. F., & Furano, A. V. (2014). Repair of naturally occurring mismatches can induce mutations in flanking DNA. *eLife*, 3, e02001.
12. Damiani, C. (2013). Modelling the influence of cell signaling on the dynamics of gene regulatory networks. In P. Lecca (Eds.), *Biomechanics of Cells and Tissues* (pp. 103-130). Dordrecht: Springer.

13. Damiani, C., Serra, R., Villani, M., Kauffman, S. A., & Colacci, A. (2011). Cell–cell interaction and diversity of emergent behaviours. *IET Systems Biology*, 5(2), 137-144.
14. de Visser, J. A. G. M., Hermisson, J., Wagner, G. P., Meyers, L. A., Bagheri-Chaichian, H., Blanchard, J. L., ... & Gibson, G. (2003). Perspective: Evolution and detection of genetic robustness. *Evolution*, 57(9), 1959-1972.
15. Derrida, B., & Pomeau, Y. (1986). Random networks of automata: A simple annealed approximation. *Europhysics Letters*, 1(2), 45.
16. Derrida, B., & Stauffer, D. (1986). Phase transitions in two-dimensional Kauffman cellular automata. *Europhysics Letters*, 2(10), 739.
17. Ding, S., & Wang, W. (2011). Recipes and mechanisms of cellular reprogramming: A case study on budding yeast *Saccharomyces cerevisiae*. *BMC Systems Biology*, 5(1), 50.
18. Doursat, R. (2008). Programmable architectures that are complex and self-organized: From morphogenesis to engineering. In S. Bullock, J. Noble, R. Watson, & M. A. Bedau (Eds.), *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems* (pp. 181-188). Winchester: MIT Press.
19. Fitzgerald, P. H., Donald, R. A., & Kirk, R. L. (1979). A true hermaphrodite dispermic chimera with 46, XX and 46, XY karyotypes. *Clinical Genetics*, 15(1), 89-96.
20. Flann, N. S., Mohamadlou, H., & Podgorski, G. J. (2013). Kolmogorov complexity of epithelial pattern formation: The role of regulatory network configuration. *BioSystems*, 112(2), 131-138.
21. Foulkes, W. D., & Real, F. X. (2013). Many mosaic mutations. *Current Oncology*, 20(2), 85.
22. Freed, D., Stevens, E. L., & Pevsner, J. (2014). Somatic mosaicism in the human genome. *Genes*, 5(4), 1064-1094.
23. Gershenson, C. (2012). Guiding the self-organization of random Boolean networks. *Theory in Biosciences*, 131(3), 181-191.
24. Gershenson, C., Kauffman, S. A., & Shmulevich, I. (2006). The role of redundancy in the robustness of random Boolean networks. In L. M. Rocha, L. S. Yaeger, M. A. Bedau, D. Floreano, R. L. Goldstone, & A. Vespignani (Eds.), *Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems* (pp. 35-42). Cambridge: MIT Press.
25. Griffiths, A. J., Wessler, S. R., Lewontin, R. C., Gelbart, W. M., Suzuki, D. T., & Miller, J. H. (2005). *An Introduction to Genetic Analysis* (8th ed.). New York: W. H. Freeman & Co.

26. Han, B., & Wang, J. (2007). Quantifying robustness and dissipation cost of yeast cell cycle network: The funneled energy landscape perspectives. *Biophysical Journal*, 92(11), 3755-3763.
27. Huang, S. (1999). Gene expression profiling, genetic networks, and cellular states: An integrating concept for tumorigenesis and drug discovery. *Journal of Molecular Medicine*, 77(6), 469-480.
28. Huang, S., & Ingber, D. E. (2000). Shape-dependent control of cell growth, differentiation, and apoptosis: Switching between attractors in cell regulatory networks. *Experimental Cell Research*, 261(1), 91-103.
29. Huang, S., Eichler, G., Bar-Yam, Y., & Ingber, D. E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical Review Letters*, 94(12), 128701.
30. Jackson, M. D., Duran-Nebreda, S., & Bassel, G. W. (2017). Network-based approaches to quantify multicellular development. *Journal of The Royal Society Interface*, 14(135), 20170484.
31. Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22(3), 437-467.
32. Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution* (1st ed.). New York: Oxford University Press.
33. Kauffman, S. A. (1996). *At home in the universe: The search for the laws of self-organization and complexity* (reprint ed.). New York: Oxford University Press.
34. Kim, H., & Sayama, H. (2017). Criticality of gene regulatory networks and the resulting morphogenesis. In C. Knibbe, G. Beslon, D. P. Parsons, D. Misevic, J. Rouzaud-Cornabas, N. Bredeche, S. Hassas, O. Simonin, & H. Soula (Eds.), *ECAL 2017: Proceedings of the 14th European Conference on Artificial Life* (pp. 245-246). Cambridge: MIT Press.
35. Kim, H., & Sayama, H. (2017). *The role of criticality of gene regulatory networks in morphogenesis*. Available at <https://arxiv.org/abs/1708.07882>.
36. Kim, H., & Sayama, H. (2018). How criticality of gene regulatory networks affects the resulting morphogenesis under genetic perturbations, *Artificial Life*, 24(2), pp. 85-105.
37. Kim, J., Vandamme, D., Kim, J. R., Munoz, A. G., Kolch, W., & Cho, K. H. (2014). Robustness and evolvability of the human signaling network. *PLOS Computational Biology*, 10(7), e1003763.
38. Kirschner, M., & Gerhart, J. (1998). Evolvability. *Proceedings of the National Academy of Sciences*, 95(15), 8420-8427.

39. Krawitz, P., & Shmulevich, I. (2007). Basin entropy in Boolean network ensembles. *Physical Review Letters*, 98(15), 158701.
40. Lander, A. D. (2007). Morpheus unbound: Reimagining the morphogen gradient. *Cell*, 128(2), 245-256.
41. Lander, A. D. (2011). Pattern, growth, and control. *Cell*, 144(6), 955-969.
42. Li, F., Long, T., Lu, Y., Ouyang, Q., & Tang, C. (2004). The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences*, 101(14), 4781-4786.
43. Lizier, J. T., Prokopenko, M., & Zomaya, A. Y. (2008). The information dynamics of phase transitions in random Boolean networks. In S. Bullock, J. Noble, R. Watson, & M. A. Bedau (Eds.), *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems* (pp. 374-381). Cambridge: MIT Press.
44. Marchi, M. D., Carbonara, A. O., Carozzi, F., Massara, F., Belforte, L., Moldsatti, G. M., ... & Palestro, G. (1976). True hermaphroditism with XX/XY sex chromosome mosaicism: Report of a case. *Clinical Genetics*, 10(5), 265-272.
45. Mohamadolu, H., Podgorski, G. J., & Flann, N. S. (2016). Modular genetic regulatory networks increase organization during pattern formation. *BioSystems*, 146, 77-84.
46. Moss, L. (2004). *What genes can't do* (1st ed.). Cambridge: MIT Press.
47. Nehaniv, C. (2003). Evolvability. *BioSystems*, 69, 77-81.
48. Nykter, M., Price, N. D., Aldana, M., Ramsey, S. A., Kauffman, S. A., Hood, L. E., Yli-Harja, O., & Shmulevich, I. (2008). Gene expression dynamics in the macrophage exhibit criticality. *Proceedings of the National Academy of Sciences*, 105(6), 1897-1900.
49. Papin, J. A., Hunter, T., Palsson, B. O., & Subramaniam, S. (2005). Reconstruction of cellular signalling networks and analysis of their properties. *Nature Reviews Molecular Cell Biology*, 6(2), 99.
50. Paul, U., Kaufman, V., & Drossel, B. (2006). Properties of attractors of canalizing random Boolean networks. *Physical Review E*, 73(2), 026118.
51. Poole, A. M., Phillips, M. J., & Penny, D. (2003). Prokaryote and eukaryote evolvability. *BioSystems*, 69(2), 163-185.
52. Rämö, P., Kesseli, J., & Yli-Harja, O. (2006). Perturbation avalanches and criticality in gene regulatory networks. *Journal of Theoretical Biology*, 242(1), 164-170.

53. Robert, J. S. (2004). *Embryology, epigenesis and evolution: Taking development seriously* (1st ed.). Cambridge and New York: Cambridge University Press.
54. Rodgers, K., & McVey, M. (2016). Error-prone repair of DNA double-strand breaks. *Journal of Cellular Physiology*, 231(1), 15-24.
55. Sayama, H. (2014). Four classes of morphogenetic collective systems. In H. Sayama, J. Rieffel, S. Risi, R. Doursat, & H. Lipson (Eds.), *Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems* (pp. 320-327). Cambridge: MIT Press.
56. Sayama, H., & Wong, C. (2011). Quantifying evolutionary dynamics of swarm chemistry. In T. Lenaerts, M. Giacobini, H. Bersini, P. Bourguine, M. Dorigo, & R. Doursat (Eds.), *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems* (pp. 729-730), Cambridge: MIT Press.
57. Serra, R., Villani, M., & Semeria, A. (2004). Genetic network models and statistical properties of gene expression data in knock-out experiments. *Journal of Theoretical Biology*, 227(1), 149-157.
58. Serra, R., Villani, M., Graudenzi, A., & Kauffman, S. A. (2007). Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data. *Journal of Theoretical Biology*, 246(3), 449-460.
59. Sharma, S., Javadekar, S. M., Pandey, M., Srivastava, M., Kumari, R., & Raghavan, S. C. (2016). Homology and enzymatic requirements of microhomology-dependent alternative end joining. *Cell Death & Disease*, 6(3), e1697.
60. Shmulevich, I., & Kauffman, S. A. (2004). Activities and sensitivities in Boolean network models. *Physical Review Letters*, 93(4), 048701.
61. Shmulevich, I., Kauffman, S. A., & Aldana, M. (2005). Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proceedings of the National Academy of Sciences*, 102(38), 13439-13444.
62. Stelling, J., Sauer, U., Szallasi, Z., Doyle, F. J., & Doyle, J. (2004). Robustness of cellular functions. *Cell*, 118(6), 675-685.
63. Strachan, T., Read, A. P. (1999). *Human Molecular Genetics* (2nd ed.). New York: Wiley-Liss.
64. Villani, M., Serra, R., Ingrami, P., & Kauffman, S. A. (2006). Coupled random Boolean network forming an artificial tissue. In S. El Yacoubi, B. Chopard, & S. Bandini (Eds.), *Proceedings of the International Conference on Cellular Automata (ACRI 2006)* (pp. 548-556). Berlin: Springer.
65. Wagner, A. (2005). *Robustness and evolvability in living systems* (1st ed.). Princeton: Princeton University Press.

66. Willadsen, K., Triesch, J., & Wiles, J. (2008). Understanding robustness in random Boolean networks. In S. Bullock, J. Noble, R. Watson, & M. A. Bedau (Eds.), *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems* (pp. 694-701). Cambridge: MIT Press.