On relationship between proliferation and transition rates of multicells

Hiroshi Yoshida
Human Genome Center, Institute of Medical Science, University of Tokyo

Hirokazu Arai
(F) Fujitsu Laboratories Ltd./CREST JST

Shigeki Oki
(株) 富士通科学ソリューション

Katsuhisa Horimoto
産業技術総合研究所・生命情報科学研究センター
Computational Biology Research Center, AIST

Abstract
The development of a multicellular organism is a dynamic process. Starting with one or a few cells, the organism develops into different types of cells with distinct functions to survive. We have constructed a simple model with a cell number increase and a cell-type order conservation to assess conditions for cell-type diversity. This model is based on probabilistic Lindenmayer system for three types of cells. In the present model, we have successfully derived rigorous relations between the proliferation and transition rates for cell-type diversity by using an algebraic operation, quantifier elimination (QE). Surprisingly, three modes for the proliferation and transition rates emerge against various ratios of the initial cells to the developed cells. Furthermore, it has been revealed that the high cell-type diversity originates from the order conservation. During the developing process of multicellular organisms, the complex but explicit relations exist between the cell-type diversity and the development rates.

1 Introduction
In a multicellular organism, a single cell—an egg—or a group of cells develops into a certain pattern with a variety of cell types (Gilbert, 2003). These different cell types are created through cell differentiation, which starts with an initial type, and then the cell changes into several intermediate types before differentiating into the final type. The process of cell differentiation can be shown as a cell lineage. One representative of a real cell lineage is the development of blood cells, wherein a stem cell is capable of extensive proliferation, creating more stem cells as well as more differentiated cellular progeny.

The theoretical study of cell differentiation and morphogenesis was pioneered by Alan Turing, who showed that a reaction–diffusion system can produce an inhomogeneous, stable pattern (Turing, 1952). Independent of initial conditions, concentrations of chemicals form a stripe or wave pattern, and this pattern formation process is robust against perturbations. Turing’s theory provides the basis for a dynamical system for morphogenesis and potentiality of cell differentiation. Embryogenesis with an increase of cell numbers was, however, not studied, and the intracellular dynamics were not sufficiently complex. In fact, resource chemicals are transported into the cell, and a complex catalytic reaction network within the cell changes the cell’s state over time. Genes are expressed and repressed in response to these intracellular dynamics. Kauffman proposed that each cell type should be regarded as an attractor of such intracellular dynamics (Kauffman, 1993), where each cell type is represented as an attracting state of a genetic network. Again, morphogenetic processes with cell differentiation were not studied. By considering Turing’s study and intracellular dynamics, together with the cell division process to increase the cell numbers, Kaneko and Yomo proposed isologous diversification (Kaneko and Yomo, 1997, 1999). This allows spontaneous cell differentiation through cell division processes and cell–cell interactions. These studies have provided a basis for the cell-type diversity of a multicellular organism. However, the

*phiroshi@ims.u-tokyo.ac.jp

*anai@jp.fujitsu.com

*torii@strand.sug.fujitsu.com

*horimoto@cbrc.jp
explicit relevance of the proliferation and transition rates between cell types to cell—type diversity has not been studied.

Apart from the approach above, Lindenmayer system (abbreviated as L—system) is a parallel rewriting system that was originally introduced to model the development of multicellular organisms (Lindenmayer, 1968a,b). Indeed, L—system is used for modelling the development process of various organisms (Yoshida et al., 2005c). Furthermore, probabilistic aspects are introduced into L—system, termed probabilistic L—system (Eichhorst and Ruskey, 1981; Eichhorst and Savich, 1980). The probabilistic L—system can take account of the influences of proliferation and transition rates, depending on the cell types.

The aim of this work is the derivation of rigorous relations between proliferation and transition rates for high cell—type diversity with conservation rule. For this purpose, we have constructed a model based on probabilistic L—system with interactions and have analyzed it by using quantifier elimination (abbreviated as QE). The derivation allows us to understand the explicit relations between the cell—type order conservation rule and high cell—type diversity over multicellular organisms.

The present paper is organized as follows. Firstly, in Section 2, we have a brief view of our previous model and results (Yoshida et al., 2005b), wherein the cell—type order conservation rule has appeared spontaneously. In Section 3, we introduce a model of a multicellular organism consisting of one-dimensional cells. This model postulates the cell—type order conservation rule as one of the rewriting rules. We make a brief explanation on QE method in Section 4. Results of algebraic computation by using QE are given in Section 5, which describes rigorous relations between the proliferation and transition rates. In Sections 5.2 and 5.3, these relations have revealed that the cell—type order conservation rule plays a key role in high cell—type diversity. We will also have a brief discussion on relevance of our results to the specific relation between proliferation and transition rates which has been observed in our previous work (Yoshida et al., 2005b).

2 Previous study

In this section, we have a brief view of our previous work (Yoshida et al., 2005b), which is the basis of the construction and analysis of the model in this work.

In a multicellular organism, a single cell—an egg—correctly develops into a prospectively determined pattern. This morphogenesis is robust against environmental perturbations, and the same pattern is always generated from an egg. In other words, recursive reproduction is repeated. At the same time, the developmental process in a multicellular organism produces a variety of cell types. The compatibility of these two points is surprising, because 'recursive production' is the reproduction of the same pattern of an individual cell, while 'cell—type diversity' is the existence of various patterns, namely, various cell types, within an individual. The question we addressed in our previous work was the selection of initial cell(s), to allow for compatibility between recursive production and cell—type diversity.

We present our previously developed model of a multicellular organism in Fig. 1. Within each cell, catalytic and autocatalytic chemical reactions maintain the cell itself and synthesize some chemicals for the cell membrane. Our numerical results have indicated that by starting with an initial object, consisting of both the chaotic cell type with diverse chemicals and the regular dynamics cell type with less chemical diversity, the recursive production of a multicellular organism with cell—type diversity has been realized. In addition, the recursive production, a remarkable regeneration pattern, which is analogous to the intercalary regeneration in cockroach legs (See Fig. 2), and planarian and salamander limb blastema (Gilbert, 2003), has been observed in our previous work (Yoshida et al., 2005b). Starting with the two cells corresponding to $I_1$ and $I_n$, the regeneration pattern corresponding to $I_1I_2\ldots I_n$ has been eventually produced as illustrated in Fig. 3. Here, such regeneration phenomena can be described as the following rewriting rule, named a cell—type order conservation rule:

$$I_iI_j \rightarrow I_iI_{i+1}\cdots I_{j-1}I_j \quad (j > i + 1).$$

This rewriting rule appears as an interaction term in the L—system which will be introduced in the next section.
3 Model

In this section, we present a simple model of a multicellular organism in which the cell lineage can be represented as a line, that is, only sequential differentiation occurs. Our model is schematically illustrated in Fig. 4. We assume that cell differentiation starts with an initial type, $I_1$, and then the cell differentiates into several intermediate types $I_2 \rightarrow I_3 \rightarrow \ldots \rightarrow I_{n-1}$ before differentiating into the final type, $I_n$. The proliferation and transition rates of cell type $i$ ($1 \leq i \leq n$) are defined as follows:

\[
I_i \rightarrow \begin{cases} 
I_i & p_{ii} \\
I_{i+1} & p_{i,i+1} \\
I_{i-1} & 1 - p_{ii} - p_{i,i+1}
\end{cases} \quad (1 \leq i < n), \\
I_n \rightarrow \begin{cases} 
I_n' & p_{nn} \\
I_{n} & 1 - p_{nn}
\end{cases}
\]

(2)

with $0 \leq p_{ii} \leq 1$ ($1 \leq i \leq n$), $0 \leq p_{i,i+1} \leq 1$ ($1 \leq i < n$), $p_{ii} + p_{i,i+1} \leq 1$ ($1 \leq i \leq n$). In addition to the rewriting rules above, we further adopt a rewriting rule, a cell-type order conservation rule: $I_jI_j \rightarrow I_jI_{j+1}\ldots I_{j-1}I_j$ ($j > i + 1$), which guarantees the contiguity of cell types, shown in Section 2.

4 Method

The key point in this work is the usage of QE, which is one of the subjects in computer algebra (Caviness and Johnson, 1998). QE deals with first-order formulae, which consist of polynomial equations, inequalities, quantifiers ($\exists, \forall$) and Boolean operators. QE computes an equivalent quantifier-free formula for a given first-order formula over the real closed field. For example, for the input $\forall x(x^2 + bx + c > 0)$, QE outputs the equivalent quantifier-free formula $b^2 - 4c < 0$. It follows from this that we can obtain a condition for unquantified variables that makes the input formula true by QE. We can also obtain the maximum value by adding one extra value, $e$, as follows:

$\exists x\exists y (x^2 + y^2 \leq 1, y \geq x^2, y \geq e)$.

4: Schematic representation of our model. Cell differentiation proceeds as follows: $I_1 \rightarrow I_2 \rightarrow \ldots \rightarrow I_n$. 
For this formula, QE outputs \( e \leq (\sqrt{5} - 1)/2 \), which shows that the maximum value of \( y \) is \((\sqrt{5} - 1)/2 \).

Recently, by using this ability, we can perform symbolic–numeric optimization for the biochemical kinetic model (Orii et al., 2005a,b) and algebraic computation for the multicell development model (Yoshida et al., 2005a).

5 Results and discussion

5.1 Analysis for growth matrix in probabilistic L-system

Now, we calculate the growth matrix \( M \) of the two contiguous cell types \( I_lI_i, I_lI_{i+1}, I_{i+1}I_i \) \( (1 \leq i < n - 1) \), which enables us to estimate the composition of \( I_lI_k \) \( (k = \ell - 1, \ell, \ell + 1) \) at step \( m \). It should be noted that other two contiguous cell types (e.g., \( I_lI_{i+2} \)) never appear at any steps in virtue of the cell-type order conservation rule. We could use the growth matrix of a longer sequence. For the present study, however, this simple matrix with the two contiguous cell types is sufficient to characterize the diversity of cell-type composition.

If one starts with \( I_1I_1 \), then the composition at step \( m \) can be calculated by the following formula:

\[
(1, 0, 0, \ldots)M^m. 
\] (3)

Here, we have studied the case of \( n = 3 \), showing the existence of three cell types. For the sake of simplicity, let \( A, B \) and \( C \) denote \( I_1, I_2 \) and \( I_3 \), respectively, in what follows. In this case \((n = 3)\), the growth matrix \( M \) is:

\[
\begin{pmatrix}
2p_{1,1} + (1 - p_{1,2})^2 & (1 - p_{1,2})p_{1,2} & (1 - p_{1,2})p_{1,2} & p_{1,2} & 0 & 0 & 0 \\
p_{1,1} & 1 - p_{1,2} & 0 & p_{1,2} + p_{2,2} - p_{1,2}p_{2,3} & p_{2,3} & 0 & 0 \\
0 & 0 & 1 - p_{1,2} & p_{1,2} + p_{2,2} - p_{1,2}p_{2,3} & 0 & p_{2,3} & 0 \\
0 & 0 & 0 & 2p_{2,2} + (1 - p_{2,3})^2 & (1 - p_{2,3})p_{2,3} & (1 - p_{2,3})p_{2,3} & p_{2,3} + p_{2,3} \\
0 & 0 & 0 & 0 & p_{2,2} & 1 - p_{2,3} & 0 \\
0 & 0 & 0 & 0 & 0 & p_{2,2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & p_{2,3} + p_{2,3}
\end{pmatrix}
\]

with its eigenvalues:

\[
1 - p_{1,2}, 1 + 2p_{1,1} - p_{1,2}, (1 - p_{1,2})^2, 1 - p_{2,3}, 1 + 2p_{2,2} - p_{2,3}, (1 - p_{2,3})^2 \text{ and } 1 + 2p_{2,3}. 
\] (4)

Let \( S \) denote a diagonal matrix with the eigenvalues above in the order of (4). The features of the growth matrix \( M \) can be summarized as follows:

- Only \( 1 + 2p_{1,1} - p_{1,2}, 1 + 2p_{2,2} - p_{2,3} \) and \( 1 + 2p_{2,3} \) can possibly have values of greater than 1. If the eigenvalues differ from one another, then \( M \) can be divided into \( P S P^{-1} \), where \( P \) is a regular matrix.

It may be worth noting that special conditions in which the eigenvalues are exactly the same have no particular physical or biological importance.

- \((1, 0, 0, 0, 0, 0, 0)P \) is

\[
\begin{pmatrix}
0, 2, -\frac{(1 - p_{1,2})p_{1,2}}{p_{1,1}}, 0, e_5, e_6, e_7
\end{pmatrix}
\]

where \( e_5, e_6 \) and \( e_7 \) are non-zero values.

- The 5th and 7th rows of \( P^{-1} \), corresponding to the eigenvalues \( 1 + 2p_{2,2} - p_{2,3} \) and \( 1 + 2p_{2,3} \) have zero elements at the \( AA, AB, BA \) and \( AA, AB, BA, BB, BC, CB \) columns, respectively.

Taking these results into account, we can obtain one of the necessary conditions that \( AA, AB, BA, BB, BC, CB \) and \( CC \) are well mingled as \( m \) approaches infinity, in other words, as the chain of cells becomes sufficiently long:

\[
1 + 2p_{1,1} - p_{1,2} > 1 \text{ and } 1 + 2p_{1,1} - p_{1,2} > 1 + 2p_{2,2} - p_{2,3} > 1 + 2p_{2,3}. 
\] (5)

In addition, we assume the following constraints:

\[
\gamma N(AA) = N(BB) = N(CC) \quad \text{and} \quad \gamma N(AB) = N(BC), 
\]

where \( N(XY) \) denotes the number of sequence \( XY \) as \( m \) approaches infinity and \( \gamma \) denotes that the ratio of the initial cells to the developed cells is \( 1/\gamma \). Notice that \( N(AB) = N(BA) \) and \( N(BC) = N(CB) \) always hold true because of constructions of the rewriting rules (2).
Under the condition (5) and the assumption of constraint (6), let \( m \) approach infinity, and the following equations are derived:

\[
N(AB) = \frac{\gamma(p_{1,2} - p_{2,3})(1 - p_{1,2} - p_{2,3})}{\gamma(p_{1,2} - p_{2,3}) + p_{2,3}}.
\]

\[
N(BC) = N(CB) = \gamma N(AB),
\]

\[
N(BB) = \gamma,
\]

\[
p_{1,1} = p_{1,2}(1 - p_{1,2})(p_{3,2} + \gamma(p_{1,2} - p_{2,3}))/[2\gamma(p_{1,2} - p_{2,3})(1 - p_{1,2} - p_{2,3})],
\]

\[
p_{2,2} = \{p_{1,2}p_{2,3}(-1 - p_{1,2})(p_{3,2} + p_{2,3} - p_{1,2}p_{2,3}^2) + (p_{1,2}^3 - 5p_{2,3})(2 - p_{1,2})(1 - p_{2,3})p_{2,3}^3 - p_{1,2}^5 (1 - 2p_{2,3}) + p_{1,2}p_{2,3}^2(-1 + 2(1 - p_{2,3})(p_{3,2} + p_{2,3}) + p_{1,2}p_{2,3}(5 - 9p_{2,3} + 6p_{2,3}^2)
\]

\[-p_{1,2}^2(2 + p_{2,3} - 7p_{2,3}^2 + 4p_{2,3}^3)\gamma + (p_{1,2} - p_{2,3})^2 (1 - p_{2,3})(2 - p_{1,2} - p_{2,3})\gamma^2) ]

\[
+(2(p_{1,2} - p_{2,3})(-1 + p_{1,2} + p_{2,3})\gamma((-2 + p_{1,2})p_{1,2}\gamma - p_{2,3}(1 - (2 - p_{2,3})\gamma)))
\]

\[
+(2(p_{1,2} - p_{2,3})(1 - p_{1,2} - p_{2,3})(-1 + 2 + p_{1,2})\gamma)
\]

\[
= p_{1,2}^2p_{2,3}(-1 + 2(1 - p_{2,3})(p_{3,2} + p_{2,3}) + (1 - p_{2,3})p_{2,3}(1 - 2p_{1,2} + p_{2,3})\gamma^2)
\]

\[
+(2(p_{1,2} - p_{2,3})(-1 + p_{1,2} + p_{2,3})\gamma((-2 + p_{1,2})p_{1,2}\gamma - p_{2,3}(1 - (2 - p_{2,3})\gamma))).
\]

(7)

In the equations above, \( N(AA) \) is normalized, i.e., \( N(AA) = 1 \). Thus, \( N(XY), (X, Y \in \{A, B, C\}), p_{1,1}, \ p_{2,2} \) and \( p_{3,3} \) can explicitly be represented as functions of \( p_{1,2} \) and \( p_{2,3} \).

Notice that as all of \( N(AB) = N(BA), N(BC) = N(CB), N(CC) \) approach 1, the diversity of the composition approaches the highest.

### 5.2 QE analysis for relations of the proliferation and transition rates

Now, let us calculate some relations between the proliferation and transition rates. Firstly, we have determined the maximum values of \( N(AB) \) by the following QE procedure:

\[
3p_{1,2} - 3p_{2,3}, \text{subject to the constraints: (5), (6) and (7), } N(AB) \geq \epsilon > 0
\]

(8)

The QE procedure (8) outputs the following inequalities: \( 0 < \epsilon < (\sqrt{17} + 1)/8 \sim 0.64039, (\sqrt{881} - 9)/40 \sim 0.517041 \) and \( (\sqrt{89801} - 99)/400 \sim 0.50167 \) when \( \gamma = 1, 10 \) and 100, respectively. Thus, we have determined the maximum values as seen in Section 4. To sum up, we have obtained the following most diverse composition:

\[
(AA, AB, BA, BB, BC, CB, CC) = (1, f(\gamma), f(\gamma), f(\gamma), f(\gamma), f(\gamma), \gamma)
\]

(9)

with \( f(1) = (\sqrt{17} + 1)/8 \), \( f(10) = (\sqrt{881} - 9)/40 \) and \( f(100) = (\sqrt{89801} - 99)/400 \). Thus, by QE method, we have obtained the exact maximum value, pruning huge numbers of candidates for the maximum effectively.

It may be worth noting that it is difficult to calculate the interval of \( \epsilon \) under the complicated constraints (5), (6) and (7) having many equations and inequalities, but QE method can actually calculate such interval.

Furthermore, QE method has let us know the rigorous relations between the proliferation and transition rates when the maximum values above are satisfied. The relations for \( \gamma = 1, 10 \) and 100 are shown in Fig. 5. For instance, the relation when \( \gamma = 10 \) is the following:

- **Mode I:**
  \[
  p_{3,3} = \text{the first root of the equation for } x,
  190p_{1,2}^3 - 490p_{1,2}^2 + 200p_{1,2} + (-391p_{1,2} + 681p_{1,2}^2 - 100p_{1,2}x = (200 + 120p_{1,2} - 310p_{1,2}^2)x^2 + (-310 + 100p_{1,2})x^3 + 110x^4 = 0 \quad (0 < p_{1,2} < p_0),
  \]
  where \( p_0 \) is approximately 0.293122 and exactly the first root of the equation for \( x, \)

  \[
  399 - 3274x + 9188x^2 - 10232x^3 + 3920x^4 = 0.
  \]

- **Mode II:**
  \[
  p_{2,3} = \left( 1 + 78p_{1,2} - \sqrt{1 + 36p_{1,2} - 76p_{1,2}^2} \right)/20 \quad (p_0 \leq p_{1,2} < 2/5).
  \]

- **Mode III:**
  \[
  p_{2,3} = \left( 20 - 9p_{1,2} - \sqrt{400 - 1960p_{1,2} + 2480p_{1,2}^2} \right)/40 \quad (0 < p_{1,2} < 2/5).
  \]
5: Relations between the proliferation $p_{ij}$ and transition rates $p_{i,i+1}$ when the maximum values above are satisfied. The three lines: black line, dashed line, gray line denote the relations in the case that $\gamma$ is 1, 10 and 100, respectively with the cell-type order conservation rule; in contrast, the dot-dashed line denotes that in the case of $\gamma = 10$ without the cell-type order conservation rule. (b) Modes I, II and III correspond to the three curves (or lines) into which the points where the curve is not smooth separate the whole region. Mode I include the origin. (d) is the magnified graph of (c) around (1, 1). Note in (c) and (d), the line $1+2p_{1,1} - p_{1,2} = 1+2p_{2,2} - p_{2,3}$ is much the same as the gray curve.

6: Relations between the the points which deviates from the highest cell-diversity curve by 0.01 (a) and 0.05 (b).

Modes I, II and III show the existence of 3 stages, in which the cell-types are highly diverse. We have also observed the existence of 3 stages in the case that $\gamma$ is 1, 10 or 100.

We have focused on the case of $\gamma = 10$ because in our previous simulation (Yoshida et al., 2005b), the constraint (6) over $N(XY)$, $(X, Y \in \{A, B, C\})$ has been observed. We have also calculated the relation between the proliferation and transition rates when $N(AB)$ is the maximum without the cell-type order conservation rule (1) in order to evaluate the effect of the conservation rule. It is observed in Figs. 5 (c) and (d) that with the cell-type order conservation rule, the $(1+2p_{1,1} - p_{1,2}, 1+2p_{2,2} - p_{2,3})$ curve (dashed and gray) is close to the line $1+2p_{1,1} - p_{1,2} = 1+2p_{2,2} - p_{2,3}$; by contrast, the curve of $\gamma = 10$ without the conservation rule (the dot-dashed line) is separate from $1+2p_{1,1} - p_{1,2} = 1+2p_{2,2} - p_{2,3}$.

5.3 Numerical analysis for relation between cell-type diversity and order conservation rule

Furthermore, we have evaluated robustness of high cell-diversity when $\gamma$ is 10 with and without the cell-type order conservation rule. This evaluation has been performed by deriving the relation between the proliferation and transition rates which deviates by 0.01 and 0.05 from the highest cell-diversity curve. As illustrated in Figs. 6 (a) and (b), the set of points (gray) without the conservation rule is more separate from the original set than the set (black) with the rule.

This fact show that without the cell-type order conservation rule, the relation between the proliferation and transition rates wherein high cell-diversity is realized has less robustness against deviation. Taking the results
in Section 5.2 and the results above into account, one can safely state that the cell–type order conservation rule plays a key role in high cell–type diversity.

Lastly, it is possible to compare the results in this work with the specific relation between the proliferation and transition rates which has been observed (but not written) in our previous work (Yoshida et al., 2005b). There, the following relation has been observed:

\[
1 + 2p_{ij} - p_{ij+1} \sim 1 + 2p_{ij} - p_{ij+1}, \ (i \neq j)
\]

(10)

when the cell differentiation proceeds as \( I_1 \rightarrow I_2 \rightarrow I_3 \rightarrow \ldots \rightarrow I_n \). On the other hand, Fig. 5(d) shows that such a relation of \( i = 1, j = 2 \) appear and that this relation would disappear without the cell–type order conservation rule as mentioned in Section 5.2. We are now in a position to state that when high cell–diversity is assumed, the relation (10) is satisfied if and only if the cell–type order conservation rule appears.

6 Conclusion

One of the remarkable features in the present study is that the rigorous relations have been derived over the L–system with interactions with the aid of quantifier elimination. Indeed, the derived relations between the cell–type diversity and the cell–type order conservation have revealed that the cell–type diversity appears robustly if and only if the cell–type order conservation rule exists. Although the present model is assumed to be composed of three cell–types, the present approach, the combination of discrete model and algebraic computation, will shed some light on important role of cell–type order conservation rule over multicellular organisms.

参考文献


