A cellular automata model for cell differentiation

H.S. Silva, M.L. Martins

Departamento de Física, Universidade Federal de Viçosa, Viçosa, Minas Gerais 36571-000, Brazil

Received 26 July 2002; received in revised form 29 October 2002

Abstract

Developmental processes generating spontaneously coordinated and inhomogeneous spatiotemporal patterns with differentiated cell types are one of the main problems in modern biology. In this paper a cellular automata model for cell differentiation is proposed. It takes into account the time evolution of the gene networks representing each cell, cell–cell interactions through gene couplings and cell division. Our computer simulations show that a society with differentiated cell types exhibiting the main features observed in biological morphogenesis emerges from a marginally stable regime at the edge of chaos.

© 2002 Elsevier Science B.V. All rights reserved.

PACS: 87.10.+e; 87.18.La; 87.18.Hf

Keywords: Cell differentiation; Morphogenesis; Cellular automata

1. Introduction

All multicellular organisms consist of many different cell types generated from single fertilized eggs through the course of a complex functional assembly of cells into tissues and organs. Morphogenesis involves several processes such as cell migration, signalling and cell–cell recognition, and gene regulation occurring under precise spatiotemporal coordination. At the end of morphogenesis, the cells, which contain the same genome and complex metabolic networks, exhibit differentiated patterns of gene expression [1]. Since cell differentiation in animals and plants is a consequence of the ability of their genes to influence each other, there is a neat connection between the mechanisms of gene control and differentiation processes [2]. The rapidly expanding knowledge about ontogeny suggests that morphogenesis proceeds via universal or conserved...
signalling pathways, regulatory mechanisms, and effector genes in both vertebrates and invertebrates.

In morphogenesis, cells choose between proliferation and differentiation at every cell division. Cell proliferation involves the exact duplication and segregation of genetic information in the form of nucleic acid sequences as well as the equal division of cytoplasmic contents. Consequently, the generated progeny cells are essentially identical to their parents. In turn, cell differentiation uses regulatory signals and asymmetric segregation of cell-fate determinants into descendent cells to produce progeny cells distinct from cells of the previous generation. Such differentiated cells will comprise tissues and organs having specific functions. The interplay between “reproductive invariance” and “structural teleonomy” is the essence of morphogenesis. These two properties separate living organisms from the inorganic world [3].

Since the seminal work of Turing [4] in which a reaction–diffusion mechanism was proposed, morphogenesis has continuously attracted the attention of scientists interested in pattern formation and self-organization. In the last few years, several mathematical models based on coupled maps [5], ordinary differential equations [6–9], and cellular automata (CA) [10] have been investigated. In particular, the “isologous diversification theory”, proposed by Kaneko and Yomo [11] as a general mechanism of spontaneous cell differentiation, includes cell metabolic networks, preserved by mitotic division, and cell interactions mediated by the medium contacting with these cells. According to this theory, cell differentiation emerges from the amplification of tiny differences among cells through the chaotic dynamics of the entire system (intracellular chemical networks and cell–cell interactions).

In this paper, we present a simple CA model for cell differentiation at the early stages of embryonic development. The paper is organized as follows. In Section 2 the CA rules implemented in our model are presented. In Section 3 the simulational results are discussed. Finally, a summary is given in the last section. Our main result is that only in the marginal regime of the intracellular dynamics, where the repertoire of distinct attractors (diversity) is maximal, cell differentiation has a moderate rate. Moreover, only in this regime a small fraction of undifferentiated (stem cells) and primitive spatially ordered tissues are generated at the intercellular level.

2. The cellular automata model

In a previous paper, a CA for the dynamics of single gene networks [12] was introduced. This model exhibited a phase diagram in which a marginal region is localized between order and chaos. This phase has stable attractors (cell types) endowed with the necessary flexibility to allow mutations and therefore natural evolution, a central result shown in Fig. 1. Furthermore, the observed power laws in the marginal regime are in the range suggested by the biological data concerning the cell cycle length (average periods) and number of differentiated cells (distinct attractors) in an organism. Hence, an extension of this CA including cell–cell interactions seems to be appropriated as a simple model to study the emergence of differentiation in a cell society.
Fig. 1. Phase diagram of the CA sensitivity to initial conditions as obtained in Ref. [12] without cell–cell interactions. The system presents three phases: frozen, marginal and chaotic, depending on whether the $t \to \infty$ Hamming distance $H$ vanishes, remains of the same size, or approaches an independent finite value for almost all $H(0) \to 0$, respectively. The normalized Hamming distance $H$ (or damage) is defined as $H(t) = (1/N) \sum_{i=1}^{N} |\sigma_i'(t) - \sigma_i(t)|$, i.e., the fraction of gene activities ($\sigma_i'$) in the replica system that differ from their counterparts ($\sigma_i$) in the original system. The data correspond to 1000 different nets or single cells with $N = 400$ genes regulated by $K = 8$ input genes. Each net was tested with one random initial state.

Such extended model is built upon the dynamics of the internal gene networks of the cells coupled by local cell–cell interactions. Through these couplings the activities of certain genes of a cell regulate the expression of another gene subset in their neighboring cells. Since all intercellular interactions are short-ranged, the model naturally incorporates the spatial or positional information considered as a key ingredient in morphogenesis. Any cell has both the gene expression pattern and the structure of the gene network preserved under mitotic division unless of rare mutations. The model appears appropriated to describe the regulative development, found in vertebrates for example. In such developmental process early embryonic cells are multipotent and differentiation is mainly determined by positional information and other extracellular signals independent on lineage. The main advantage of the model is the use of binary variables and simple discrete evolution rules, facilitating the analysis of the structure and stability of the cell attractors. However, the model, in its present form, does not include cell migrations caused by differential cell adhesion, a feature which plays an important role in morphogenesis [10].
The CA rules implemented in our model are the following.

2.1. Intracellular dynamics

As in Ref. [12], the genome of each single cell \( q (q=1,2,\ldots) \) is composed by a set of \( N \) genes characterized by binary states \( \sigma_i^{(q)} \), \( i=1,2,\ldots,N \). When \( \sigma_i^{(q)} = 1 \) the gene \( i \) is active for transcription and the proteins it codifies are produced by the cell \( q \). On the other hand, when \( \sigma_i^{(q)} = 0 \) the gene \( i \) is inactive and its products are not synthesized by the cell \( q \). Each gene \( i \) in a cell \( q \) has \( K \) intracellular regulatory input genes. These inputs include the gene itself and \( K-1 \) other input genes chosen at random among either its nearest and next-nearest neighbors, with probability \( 1-p_2 \), or all the other remaining genes in the same cell, with probability \( p_2 \). Thus, as biologically observed, a given gene can be regulated by either its neighbors or distant DNA sequences, whose proteins, produced in the cytoplasm, diffuse towards the cell nucleus.

Concerning only the intracellular interactions, the activity of each gene \( i \) in a cell \( q \) is updated through the function

\[
\sigma_i^{(q)}(t+1) = \text{sgn} \left[ J_{ii} \sigma_i^{(q)}(t) + \sum_{l=1}^{K-1} J_{ij_l(i)} \sigma_{j_l(i)}^{(q)}(t) \right]
\]

which depends only of the previous states of its regulatory elements. Here \( J_{ij(i)} \) is the coupling constant representing the regulatory action of the \( j_l(i) \) \( (l=1,2,\ldots,K-1) \) input on gene \( i \) and \( J_{ii} \) is the self-regulation. \( \text{sgn}(x) = 0 \) if \( x \leq 0 \) and \( \text{sgn}(x) = 1 \) if \( x > 0 \). All the gene states are simultaneously updated.

The intracellular couplings \( J_{ij} \) (and the self-interactions \( J_{ii} \)) are chosen at random according to the distribution

\[
P(J_{ij}) = \frac{1-p_1}{2} \left[ \delta(J_{ij}-J) + \delta(J_{ij}+J) \right] + p_1 \delta(J_{ij}),
\]

where \( \delta(x) \) is Dirac’s delta function and \( J=1 \). Therefore, each bond \( J_{ij} \) is activatory (+1) or inhibitory (−1), with probability \( (1-p_1)/2 \), or diluted \( (J_{ij}=0) \) with probability \( p_1 \).

Once the \( K \) intracellular inputs of each gene and the corresponding interactions \( J_{ij} \) are chosen at the beginning (the fertilized egg), the cell genome (intracellular input genes and couplings) is fixed forever.

2.2. Cell division

Any cell divides by mitosis into two almost identical cells each time it traverses its attractor cycle. After mitosis, the daughter cell inherits the same intracellular gene couplings and regulatory inputs of its mother cell. In turn, the gene activity pattern of the daughter cell may differ from those of its mother by a small number of random mutations, which occur with a probability \( \sim 10^{-4} \). Since the model assumes that cells grow in a two-dimensional square lattice, every new daughter cell occupies, at random, a site of its mother neighborhood eventually displacing other cells along this direction by one lattice constant.
Fig. 2. Scheme of cell–cell interaction coupling a target gene $i$ in a cell $q$ to one of the genes in a neighbor cell $p$ which correspond to the intracellular regulatory inputs of $i$. Only one intercellular coupling can be established by each target gene in a cell after each mitotic division. Also, at each mitotic round the set of target genes is the same in both mother and daughter cells. All of them are randomly chosen with a probability $p_3 = 1/N$. In contrast, for a given target gene the neighbor cell and the connected input gene are randomly and independently selected by the mother and daughter cells.

2.3. Cell–cell interactions

Following each mitotic division, both mother and daughter cells establish intercellular gene couplings with cells in their Moore neighborhoods. The same target genes, selected with a probability $p_3 = 1/N$ among the genome, are involved in these couplings in both cells. Any target gene builds up a single intercellular coupling according to the following procedure, depicted in Fig. 2. Firstly, a neighbor cell is drawn at random with equal chance. Secondly, in this selected neighbor cell one among the $K - 1$ genes correspondent to the intracellular inputs of the target gene is randomly chosen and coupled to the target. Thirdly, the strength of this intercellular interaction is equal to the correspondent one in the intracellular network. Thus, every new intercellular coupling introduces an additional term $J_{ij}(i)\sigma_{jl}(p)$ in Eq. (1), linking a target gene $i$ in cell $q$ to an inducer gene $j_l(i)$ with activity $\sigma_{jl}(p)$ in a neighbor cell $p$. The inducer genes can be thought of as positional regulatory network, since their products mediate local intercellular interactions. This entire procedure is repeated after every mitotic division and the establishment of a duplicated intercellular coupling is forbidden.
In order to take into account the intercellular couplings, the activity of a gene $i$ in a cell $q$ is updated using the function

$$\sigma_{i}^{(q)}(t + 1) = \text{sgn} \left[ J_{ii}^{(q)} \sigma_{i}^{(q)}(t) + \sum_{l=1}^{K-1} J_{ij_{l}(i)}^{(q)} \sigma_{j_{l}(i)}^{(q)}(t) \right. $$

$$\left. + \sum_{\text{intercellular couplings} \ l} J_{ij_{l}(i)}^{(p)} \sigma_{j_{l}(i)}^{(p)}(t) \right]$$

in which the second sum involves all the eventual intercellular couplings to gene $i$ established along the successive mitotic divisions of cell $q$.

Finally, it is important to notice that each intercellular coupling preserves the regulatory network present in the genome of the egg cell. Also, since each intercellular coupling is chosen randomly and independently, the mother and the daughter cells have, in general, distinct local nets of interacting cells.

3. Results

In all the simulations, any initial states of the genes were equally probable and each gene was regulated by $K = 8$ intracellular input genes.

The intracellular dynamics exhibits, as demonstrated in Ref. [12], attractors (limit cycles) which are or strongly sensitive, or sensitive, or insensitive to the initial conditions. In consequence, the CA parameter space is partitioned into three phases: chaotic, marginal and frozen, respectively, as seen in Fig. 1. In the chaotic regime the average period of the attractors increases exponentially with the genome size, the wide majority of the genes oscillates between the inactive and active states, and even initial minimal perturbations on the gene expression patterns trigger large cascade of mutations involving 20% or more of the genome. In contrast, in the frozen phase we observed a power law increase of the average period of the attractors, the overwhelming majority of the genes is fixed in inactive or active states, and the final damage generated by minimal perturbations is very small (%6 of the genome size). Finally, in the marginal regime, localized between order and chaos, the average period of the attractors increases as a power law, a frozen core of genes begins to percolate and the subset of oscillating genes is just splitting in separated islands. Although minimal perturbations on the activity state of some key genes can trigger moderate cascade of mutations involving more than 10% of the genome, the rule is small changes (<10%) with low frequencies (\~{}1%), as shown in Fig. 3. In order to estimate the fraction of these key genes, each one of the genes in fixed regulatory networks was submitted to a minimal structural perturbation by changing a single of its regulatory couplings. Our results reveal that the average number of these key genes vanishes in the frozen phase, constitute a large fraction of the genome (\geq{} 40%) in the chaotic phase, but is small (<20%) in the marginal regime. These genes represent major regulatory sequences, such as the segmentation and homeotic genes, and play a critical role in cell differentiation and morphogenesis.
Fig. 3. Frequency distribution of damages $H(T)$ generated by minimal perturbations at the chaotic, marginal and frozen phases of the CA. The damage is measured as the long-time Hamming distance between two gene expression patterns which differ initially by changing the activity of a single gene. The data correspond to $T = 800$ and 100 different nets or single cells with $K = 8$ input genes, $p_2 = 0.50$ and $N = 2500$ genes. Each net was tested with 100 random initial states.

Fig. 4 shows that, considering the ensemble of all genes networks, the number of different attractors or cell types increases as a power law of the number of genes in all the three CA phases. Moreover, as shown in Fig. 5, the marginal regime exhibits the maximum number of differentiated cell types for a fixed genome size, one of the main results obtained by our simulations. Therefore, the marginal phase is characterized by stable attractors endowed with the necessary flexibility to allow mutations and by the greater diversity of cell types upon which natural evolution operates.

Henceforth, our results for early morphogenesis will be focused. Starting from a fertilized egg, successive cell divisions occur, and give rise to local nets of cell–cell interactions. Tiny differences among gene expression patterns generated either by rare mutations in the daughter cells or in response to intercellular signalling can be dynamically ampliﬁed, leading cells to differentiate from their parents. Cell differentiation corresponds to a transition from a given attractor to another accessible neighboring cell type. Such transition is preceded by a progressive synchrony loss of the cells along their ancestor limit cycle. Clearly, the dynamical transition associated to cell differentiation, and consequently morphogenesis, depends on the stability of the underlying intracellular dynamics of the fertilized egg. In Fig. 6 are shown typical differentiation processes occurring in each one of the three phases present in Fig. 1. In the frozen phase, cell differentiation is highly constrained and almost all cells traverse in phase a common attractor. In contrast, the instability of the intracellular dynamics in the chaotic regime leads to a high rate of cell differentiation. This is due to the fast dephasing between cells of the same type and the large number of neighboring cell types.
Fig. 4. Typical log–log plot of the average number of different cell types as a function of the genome size $N$ for the three distinct CA phases. The data correspond to 100 different nets or single cells with $K = 8$ input genes and $p_2 = 0.50$, each one tested with 5000 random initial states. Each cell has been evolved until attains its limit cycle.

Fig. 5. Average number of distinct dynamical attractors, interpreted as differentiated cell types, along a fixed $p_2$ cut traversing the three phases of the intracellular dynamics. As the genome size increases, a neat maximum becomes gradually evident for $p_1$ values inside the marginal phase. The data correspond to 100 different genomes or single cells with $K = 8$ input genes and $p_2 = 0.50$, each one tested with 5000 random initial states of gene expression. Again, each cell has been evolved until attains its limit cycle.

accessible to any attractor. Since in the marginal phase each cell type has only a few accessible neighboring cell types, differentiation progress at moderate rates as the number of cells in the embryo increases. At this point it is worthwhile to mention that, in contrast to the reaction networks considered in Refs. [6–8] for which less than 10%
Fig. 6. Differentiated cell types as a function of the total number of cells during the embryonic development in the three phases of the intracellular dynamics. The data correspond to a zygote with $N = 25$ genome size and $p_2 = 0.50$. The cell differentiation sequences for *T. torematicus* and *C. elegans* are shown in the inset.

show cell differentiation, virtually all of the random gene networks operating at the chaotic and marginal phases of our model exhibits differentiation. The reason is that, in addition to a large number of positive and negative feedback reactions, these gene networks are replete of autocatalytic interactions.

The simulations reveal that, during the developmental process, a fraction of the embryo cells does not reach a definite limit cycle. In the marginal regime, this fraction is very small ($\sim 1\%$ of the cells) and increases slowly. However, this fraction has a fast and irregular increase in the chaotic phase up to a significant number ($\sim 20\%$) of the cells. These undifferentiated cells can be interpreted as stem cells because some of them are eventually attracted by one of several distinct cell types through cell–cell interactions. In addition, at both marginal and chaotic phases new attractors unobserved in the dynamics of isolated cells emerge during the developmental process, as already observed by Jackson et al. [13]. Consequently, instead of being fully determined by its internal dynamics, cell fate depends on the other cells, providing additional support to the concept of “partial attractor” stabilized only through cell–cell interactions [6]. Also, both the cell cycle lengths and the fractions of oscillating genes decrease as the number of cells in the embryo increases, as seen in Fig. 7. Therefore, during morphogenesis the intracellular signalling dynamically enhances the global stability of the embryo by reducing the complexity of the generated cell types. The decreasing number of oscillating genes translates in more fixed and cell-specific patterns of gene expression as, indeed, observed in real systems [14]. Hence, at the level of individual cells, intercellular interactions may continuously enlarge the ordered region (the marginal and frozen phases) as the number of embryo cells increase. Indeed, the fraction of frozen genes
Fig. 7. Fraction of oscillating genes as a function of the total number of cells during the embryonal development in the three phases of the intracellular dynamics. The data correspond to an egg cell with $N = 25$ genes and $p_2 = 0.50$.

Fig. 8. Typical spatial patterns of cells in embryos developing in the (a) frozen, (b) marginal, and (c) chaotic phases of the intracellular dynamics. These patterns contain 500 cells and the parameters used were $K = 8$ and $p_2 = 0.50$.

is the essential feature conferring stability to the intracellular dynamics, as claimed in Refs. [2,12]. From Fig. 7 it can be inferred that the changes in the phase diagram of Fig. 1 for an isolated cell are mainly due to the slow shrinking of the chaotic region. But these changes are expected to become tiny for increasing genome sizes, since the number of intercellular couplings per cell remains small as compared to the fraction of oscillating genes in the chaotic phase. In turn, cell–cell interactions provide, as shown in Fig. 8, the positional information needed to the self-organization of the tissues in
the embryo. In the frozen regime cell differentiation does not occur; in contrast, in the chaotic phase there is an excessive and spatially random differentiation. However, in the marginal regime, primitive spatially ordered patterns of distinct cell types emerge even in the absence of differential cell adhesion.

At last, we situate our central result (Fig. 6) within the biological context. The early embryonic development is characterized by nearly synchronous cell divisions in which most cells take part, as well as the absence of cell motility, therefore, gastrulation movements. This embryonic stage lasts for 12 divisions in *Xenopus laevis*, nine in *Drosophila*, five to nine in sea urchins [9], and five divisions in the worm *C. elegans* [1]. The cell differentiation maps of the sea urchin *T. toreumaticus* and *C. elegans*, shown in the inset of Fig. 6, are surprisingly similar to the results for the marginal regime. Indeed, after the third division, the embryo of *T. toreumaticus* contains eight identical blastomeres, which differentiate into three cell types (eight mesomeres, four small micromeres and four large macromeres) at the next (fourth) division. Then, after the fifth division, at a 32-cell stage, a new cell type (large micromere) is generated, and these four cell types proliferate until the ninth division (∼300 cells). In turn, the first five divisions of *C. elegans* embryogenesis generate six founder blastomeres (AB, MS, E, C, D and P4) at 30-cell stage. Thus, in the marginal regime at the edge of chaos, even disordered internal gene networks coupled by varying intercellular interactions can exhibit the main features observed in biological morphogenesis.

4. Conclusions

Our CA model demonstrates that, in the chaotic and marginal phases, cells differentiate through the interplay between intracellular dynamics and cell–cell interactions. As the cell number increases, new local intercellular gene couplings are established, dephasing the internal dynamics in each cell, and settling on a transition to a new attractor. However, only in the marginal regime, where the repertoire of distinct attractors is maximal, cell differentiation has a moderate rate, generates primitive spatially ordered tissues, and a small fraction of undifferentiated or stem cells. This central result contrasts with the global chaotic dynamics (or “open chaos”) scenario for cell differentiation proposed by Kaneko and Yomo [7,8,11], which requires a chaotic underlying intracellular dynamics.

Acknowledgements

This work was partially supported by the Brazilian Agencies CAPES and CNPq.

References