Model of morphogenesis

Yue Wang[1#], Andrey Minarsky[2#], Robert Penner[1], Christophe Soulé[1], and Nadya Morozova[1,3,4*]

[1]IHÉS, France; [2]St-Petersburg Academic University, Russia; [3]UMR9198, I2BC, CNRS, France; [4]BIN RAS, St-Petersburg, Russia

***** corresponding author

# equal contribution

Yue Wang, Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440, Bures-sur-Yvette, France, tel. +33 01 60 92 66 59, yuewang@ihes.fr

Andrey Minarsky, National Research Academic University, Russian Academy of Science, Khlopina 8-3,194021, St.-Petersburg, Russia, tel. +7 812 232 0377, minarsky@school.ioffe.ru

Robert Penner, Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440, Bures-sur-Yvette, France, tel. +33 01 60 92 66 44, rpenner@ihes.fr

Christophe Soulé, Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440, Bures-sur-Yvette, France and Centre National de la Recherche Scientifique, Paris, France, tel. +33 01 60 92 66 21, soule@ihes.fr

Nadya Morozova, Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440, Bures-sur-Yvette, France and UMR9198, I2BC, Centre National de la Recherche Scientifique, Paris, France, tel. +33 01 60 92 01 89, morozova@ihes.fr

Abstract

We build a theoretical model of morphogenesis. This model describes cell fate in the developing organism using the notion of epigenetic code of each cell. Namely, given the epigenetic spectra of a cell and its neighboring cells, we can determine the corresponding cell event which it will perform. This means that the properties of a group of cells (comprising an embryo or its part) at any time point are also known, and thus, the evolution of an embryo can be described. By this strategy it is possible to establish the tissue, organ or embryo shapes at any time, starting from a zygote. As an essential part of the model, the formalization of the notion of cell potency is introduced, and the related properties are discussed.

Keywords: morphogenesis; modeling; signaling

Introduction

Morphogenesis of biological organisms is determined by the behavior of it cells, such as growth, division, differentiation, movement and death. These behaviors are affected by cell-cell interactions and cell-environment interactions (Hayashi et al. (2015), Sagner and Briscoe (2017)). The important components of morphogenesis machinery, such as morphogen gradients, electrical and mechanical signaling and differential gene expression are well described for many concrete developmental cases, being reflected in numerous mathematical models (Gilmour et al. (2017), Meinhardt (2013), Menshykau et al. (2014), Pauzi et al. (2018), McLaughlin and Levin (2018), Adjei and Heffernan (2015), Gordon and Gordon (2016)).

However, the conceptual gap between a set of particular mechanisms important for the process of morphogenesis and the creation of a concrete morphology of organisms is still not filled, thus representing an intriguing theoretical field (Levin (2011), Lobo et al. (2012), Wolpert and Flanagan (2016), Delile et al. (2017)).

The prevailing concept that the formation of a species-specific life form of organisms is a result of the implementation of the genetic program via differential gene expression does not explain the fine regulation of the shape of an organism and its parts, and, moreover, the coordinated development of the whole body. For the deterministic shaping all set of various intracellular processes important for morphogenesis should be coordinated and controlled by certain general law(s) (Wolpert and Flanagan (2015); Gilmour et al. (2017)).A fundamental theoretical question is: what are the “instructions” for building an organism or rebuilding its parts (during regeneration), where are they located, and what is the mechanism for their implementation?

Here we aim to uncover the conceptual laws underlying the creation of determined morphology (geometry) of organisms, applicable for both embryonic development and regeneration processes. The present work is a sequel to Morozova and Shubin (2013), Morozova and Penner (2015) and Minarsky et al. (2018) and is devoted to elaborating appropriate mathematical formalizations of the pattern formation process in morphogenesis.

In our previous works we conjecture the existence of an additional biological code (epigenetic code) which bears information about geometrical pattern of an organism and thus coordinates the cascades of molecular events implementing a pattern formation (e.g., differential gene expression, directed protein traffic, growth of cytoskeleton). We understand the term epigenetic in a broad sense, as any information in a cell additional to the genetic one, which can be inherited by cells and be involved in regulation of their cell fates in a tight interplay with the genetic code. By that, we can consider a wide spectrum of possible levels of epigenetic information. We assume that a vast set of all intracellular processes, important for morphogenesis, is governed and controlled by such an epigenetic code. For the time being, we do not aim to point out the exact molecular mechanisms, implementing this control, and a fortiori do not aim to model all set of different mechanisms (such as morphogen gradients or differential gene expression) and their possible interconnection with an epigenetic code signaling. Thus, we omit the details of the mechanisms, presenting a particular parts of implementation process in a concrete particular cases, and discuss a conceptual framework based on an epigenetic control theory.

The proposed theory of morphogenesis comprises the following fundamental hypotheses, which are revised from Minarsky et al. (2018):

– For each cell in an organism, the cell-surface distribution of chemical substances, which is called its (epigenetic) spectrum, governs morphogenesis.

– One cell affects its neighboring cells through signaling, which is determined by its epigenetic spectrum.

– Based on its own epigenetic spectrum and received signaling, each cell in an organism performs one of several possible cell events, such as spectrum change, movement, shape change including growth, mitotic division, and apoptosis.

– The rules that determine cell events are universal throughout nature.

It is important to emphasize that generally we consider a wide spectrum of possible epigenetic information, while the cell surface location of it is suggested only as a particular possibility, which seems to be the most attractive one, as discussed in Bessonov et al. (2019). The same applies for the suggestion of the encoding by special type of chemical substances, which we have conjectured in Minarsky et al. (2018) and Bessonov et al. (2019) just for the illustration of corresponding universal rules, leading to interconnection between such a code and spatial structure of an organism.

In this work we suggest that the mathematical formalization of cell events is in a form of linear operators, acting on epigenetic spectrum, and then search for possible principles by which the epigenetic spectrum can determine cell events. We propose an individual-collective model of cell behavior, meaning that if the epigenetic spectra and spatial information of all cells are known, then we can calculate the behavior of each cell at the next time point. Collectively the spatial (geometrical) and phenotypic property of each tissue at any time point can be determined

likewise. One of the main tasks of the future work is to determine the concrete forms of related operators through biological and computational experiments (Bessonov et al. (2019)).

Results

§1. Definitions and formalizations

Basic notions.

We choose a fixed coordinate system in R^3, then the spatial status of a cell is denote as *Se*=(*L*,*Or*,*Sh*), where *L* is a point in R^3, the location of cell; *Or* is a Cartesian coordinate structure based at *L*, defining the orientation of the cell; *Sh* is the shape function with respect to *Or*, thus if a cell moves or rotates, *Sh* does not change. *L*, *Or* and *Sh* together determine a subset of R^3, the space occupied by the cell.

We assume that in the simplified case an *epigenetic spectrum* of a cell *c* can be represented by a matrix showing the amount of each coding molecule in each spatial sector of the cell surface, namely a matrix $M(c)=(m_{ij})$, where m_{ij} is the number of molecule *i* located in the sector *j* of the surface of *c*.

The *cell state* of a cell is defined as (*Se*,*M*), describing both spatial status and epigenetic spectrum.

Each cell *c* has a *status of differentiation*, describing a cell type; we assume that this status can be determined from *M(c)*.

The *potency* of a cell is its ability to produce different cell types (Samsonraj et al. (2015)). It can be determined from the epigenetic spectrum of a cell. Its interpretation in our model is discussed in Section 3.

Cell events.

We formalize development using the notion of *cell events*, which describes the evolution of cell states. We define the following set of cell events which can occur to a cell *c*:

Cell division: one cell produces two cells.

Internal cell event: a cell changes its spectrum.

Growth: a cell changes its shape, but retains its spectrum.

Movement: a cell changes its location or orientation, but keeps shape and spectrum.

We will call the last three cell events as “one-cell event” type, since they do not change the cell number.

Stagnation: no change of a cell state (also can be considered as “one-cell event” with zero changes.

Apoptosis: programmed cell death.

There are two particular cases of internal cell event, which are of high importance for formalizing the development:

Differentiation: a cell adopts a new differential status with lower potency; occurs in a course of normal development.

Dedifferentiation: a cell adopts a new differential status with higher potency; occurs in a course of abnormal development, which should be considered in the model as a part of re-construction of proper determined morphology in abnormal situations, such as regeneration, transplantation, growth of an isolated cell in a cell culture, and other processes of changing cell fate.

Generally, it implies that the process of dedifferentiation represents the conversion of a partially or terminally differentiated cell to an earlier developmental stage (Bryant et al. (2002), McCusker et al. (2015), Kragl et al. (2009) and Tanaka et al. (2016)).

For example, during the process of regeneration after amputation, the new cells can be produced by division and differentiation of resident stem cells, and/or dedifferentiation of differentiated cells. In the case of limb regeneration, this leads to the formation of blastema (a mass of undifferentiated cells), which next regenerates the limb by cell migration, proliferation and differentiation. In other words, the process of dedifferentiation is that cells are reprogrammed to more embryonic-like states so as to acquire higher developmental potency.

However, the dedifferentiation can be a part of a normal developmental program, for example, in the cases of gamete formation, or in the process of metamorphosis.

Another mechanism, which has been reported for the regeneration of lost or damaged tissue is a process of transdifferentiation, in which a differentiated cell is converted into another cell type (Jopling et al. (2011), Kikuchi (2015), Frasch (2016), Cieslar-Pobuda et al. (2017)). Initially the term transdifferentiation (also referred as lineage reprogramming, or conversion, or metaplasia) was introduced to determine a process in which one specialized cell directly changes into another cell type without entering a pluripotent state. But further it was argued that in reality transdifferentiation occurs in two steps: first, dedifferentiation of cells, regressing to a point where they can switch lineages; and second, the activation of natural developmental program with differentiation into the new cell lineage. The question if direct (i.e. without reverting to pluripotency state) conversion of specialized cells to another cell type can occur or not still seems to be under a hot discussion.

Taking into consideration all these biological details, we will make a definition of cell event Dedifferentiation more precise and consider two distinct cases of cell events corresponding to it, namely: *Dedifferentiation* as an internal cell event, mapping a matrix of a cell to the matrix of its ancestor, thus assigning to it a differential status with higher potency; and

Transdifferentiation as any other type of internal cell event, which a cell can undergo under abnormal signaling, hence, it can lead to a cell with increased cell potency, the equal one or even with a decreased potency.

Figure 1 illustrates a set of possible cell events for a cell in a cell state (*Se*,*M*).

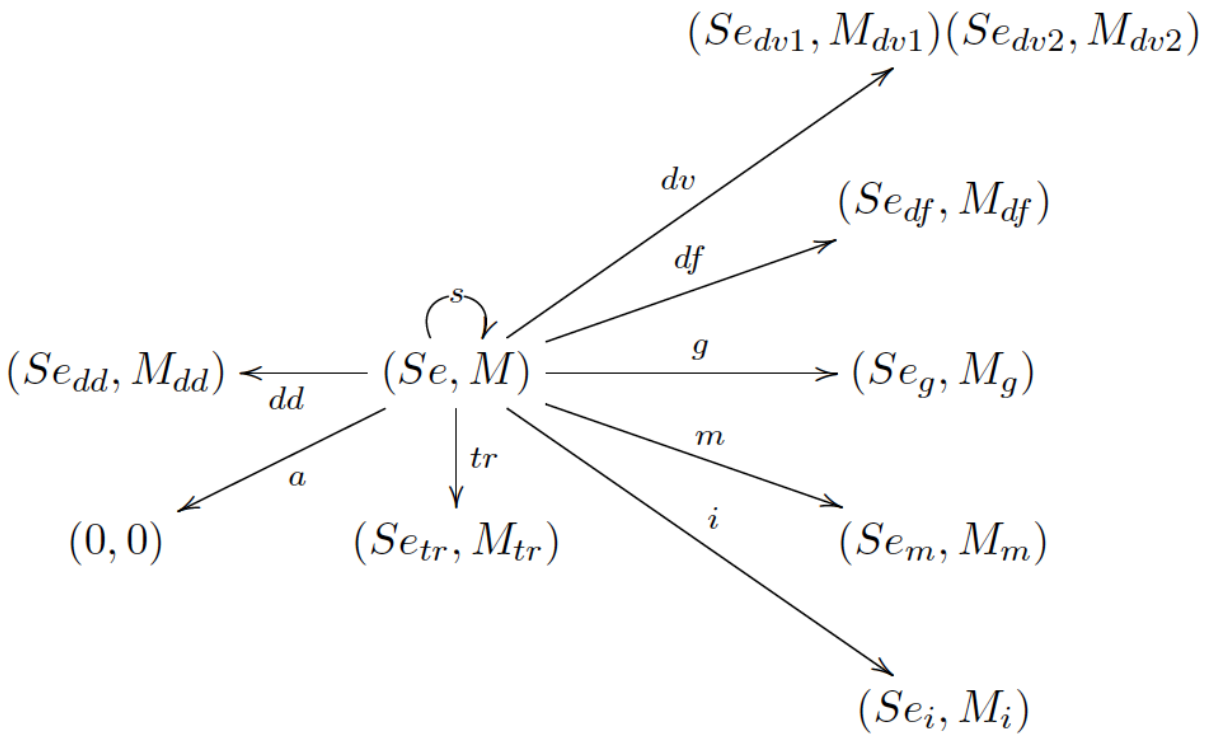

Figure 1. A set of possible cell events for a cell in cell state (*Se*,*M*). Cell event division (dv) results in two cells; cell events differentiation (df), growth (g), movement (m), internal cell event (i) produce one resulting cell; cell events dedifferentiation (dd), transdifferentiation (tr), stagnation (s) and apoptosis (a) are the particular cases, discussed in the text.

Operators of cell events:

To describe the evolution of epigenetic spectrum, we introduce four linear operators A_1, A_2, A_3, A_4. Each operator maps a spectrum matrix into another spectrum matrix. Linear means that for a matrix M, each entry of AM is a linear combination of entries of M. Thus if the size of M is m*n, then A has a representation as an (m*n)*(m*n) matrix. We do not assume that the operators A_1, A_2, A_3, A_4 are invertible.

Operators A_1 and A_2 correspond to cell event division, A_3 is for one-cell events in normal situations, and A_4 is for one-cell events in abnormal situations. In normal situations, depending on the particular spectrum M of a cell c, a one-cell event can be differentiation, growth, movement, internal cell event or stagnation. In abnormal situations, the internal cell event can be dedifferentiation or transdifferentiation.

We assume that once a spectrum of a cell c meets the conditions, corresponding to programmed cell death (apoptosis), this cell immediately starts the apoptosis procedure. These conditions can be represented by a subset $\boldsymbol{M_{ap}}$ of a set of all matrices $\boldsymbol{M}$, and the resulting AM matching $\boldsymbol{M_{ap}}$ can occur under action of each of the operators A_1, A_2, A_3, A_4, depending on the matrix M.

To describe the evolution of spatial status $Se=(L,Or,Sh)$, we introduce operators B_1, B_2, B_3, B_4, corresponding to operators A_1, A_2, A_3, A_4. Each spatial operator B_i maps a spatial status $Se=(L,Or,Sh)$ into another spatial status $Se=(L',Or',Sh')$.

Operators of cell division A_1 and A_2, producing two matrices M_1 and M_2 corresponding to two daughter cells, are related due to the fact that according to our conjecture (Minarsky et al. (2018), Bessonov et al. (2019)) the matrices M_1 and M_2 should have certain interrelation, for example, by the rule of complementarity (Bessonov et al. (2019)). Here we do not discuss the concrete form of such interrelation, but state a general description of it.

Formalization of development as a graph

We formalize the development process as a graph (rooted tree) starting from a zygote, where each vertex corresponds to a cell state (*Se*,*M*) at a time point, and each edge corresponds to a cell event (Figure 2). The x-axis is time, and the y-axis corresponds to cell states, including spectrum, location, orientation and shape, therefore the dimension of the y-axis is very high. We postulate that for the normal development of a particular organism there exists a graph (tree) of optimal cell states and cell events (TOE), which can be reconstructed from the matrix M_z of the zygote by application of operators optimal for each matrix at each moment of time; the slice of this tree at any moment *t* will then make up the shape of an organism at this time.

The tree illustrates the relationship between any two cell states at different time points. Also, at each time point, we can take a slice, and obtain the states of all cells at this time point, from which we can reconstruct a *graph of adjacency*, where vertices of cell states corresponding to the directly neighboring cells, are connected by the edges (Figure 1, in the box).

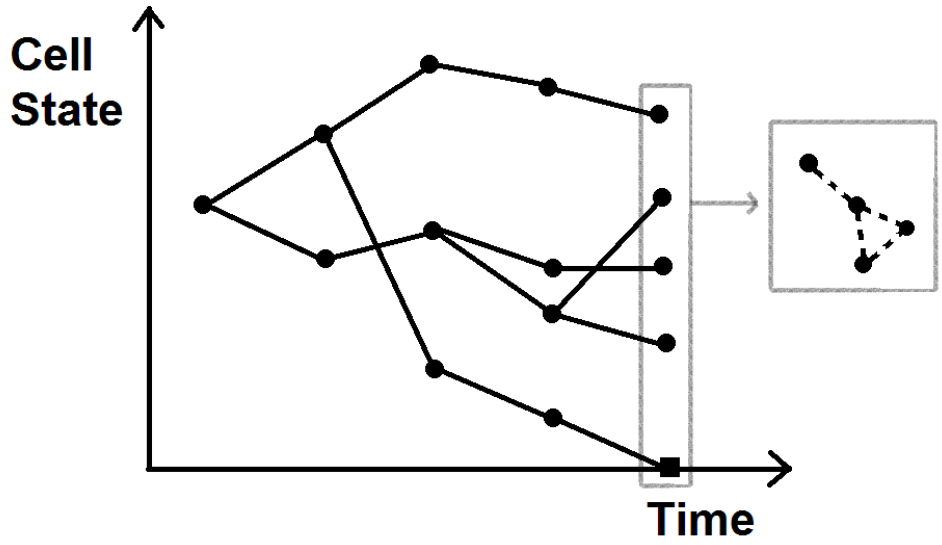

Figure 2. Developmental graph (rooted tree) starting from a zygote. Each vertex corresponds to a cell state at a time point, where the square vertex represents a dead cell. Each edge is a cell event, possibly division, differentiation, dedifferentiation, growth, movement, internal cell event, stagnation, or apoptosis. In the box, the *graph of adjacency* reconstructed for the last time slice is shown.

Signaling.

Cells exchange information between each other and with the environment, and such information influence the cell behaviors. This phenomenon is called *signaling* and has been widely discussed in the literature (Bando et al. (2018), Byambaragchaa et al. (2018), Pusapati et al. (2018)).

Among many different types of considered signals the most frequently mentioned one is the signaling by biochemical substances, such as hormones or morphogens, which can provide distant regulation of cell behavior, but there are also reports about electrical and mechanical signaling, and many types of cell-to-cell communication via extracellular proteins.
The regulation by morphogens, the molecules forming the gradients of their concentration in a local part of a body and thus inducing various specific cellular responses (e.g. extracellular proteins, such as FGF, BMP, TGF-β, Wnt, Hh; metabolites such as retinoic acid in animals or auxin in plants (Sagner and Briscoe (2017), Lander (2013), Cunningham and Duester (2015), McCusker et al. (2015), Tuazon and Mullins (2015), Paque and Weijers (2016)) has been simulated in numerous mathematical models (Meinhardt (2013), Menshykau et al. (2014), Pauzi et al. (2018)).

The role of the electrical (Law and Levin (2015), Funk (2015), McLaughlin and Levin (2018)) and the mechanical (Adjei and Heffernan (2015), Gordon and Gordon (2016)) signals in the process of morphogenesis were also reflected in many mathematical models.

Though these important components of signaling are well described for many concrete cases, no one of them can be considered as the main instructive signaling, governing the formation of the determined morphology of an organism.

In the framework of our model we define *signaling* as the transmission of epigenetic spectrum of a cell to a collection of its neighbors (Minarsky et al. (2018)).

We assume that this signaling can account for changes in cell fates as within a normal course of development, so in the crises like regeneration, transplantation, cell isolation, etc. For the time being, we do not aim to point out the exact molecular mechanisms involved in implementation of this signaling, and a fortiori do not aim to model all set of different types of signals and signaling mechanisms in living organisms and their interconnection with this "epigenetic signaling". We assume that all other types of signaling mechanisms mentioned above can be governed by signaling cascades, started from the epigenetic code varying, and represent in different local cases local implementation of the main instructions, governed by this code.

Thus, we assert that the *signal* received by a cell c is a matrix $S(c)$ which can be written as:

$$S(c)=\Sigma_k G(c,k)M(k),$$

where the summation is taken over all neighbors k of the cell c. Here $G(c,k)$ is a non-negative number, from the *graph of influence* of a cell c, which corresponds to the *graph of adjacency of cell c* supplemented with a non-negative real number for each edge, determined by the properties of cell c and cell k. We do not require that $G(c,k)$ equals $G(k,c)$.

After signaling, we define a new matrix

$$M(c)^S=M(c)+h(c)S(c),$$

where the positive real number $h(c)$ is called the *sensitivity* of a cell c to signals.

The format of $S(c)$ can be more complicated. For example, we can calculate $S(c)$ sector by sector.

§2. Model of Development

A cell c with epigenetic spectrum M can divide to produce two new cells with epigenetic spectra T_1 and T_2, or perform a one-cell event in normal situations to produce one new cell with epigenetic spectrum T_3, or perform a one-cell event in abnormal situations to produce one new cell with epigenetic spectrum T_4, or die. These four new epigenetic spectra form a set $\{T_1,T_2,T_3,T_4\}$, which is called the *target image* of c. This set consists of all four possible epigenetic spectra that might appear after one cell event. As mentioned, the matrix for apoptosis can be a particular case of any spectrum from T_1,T_2,T_3,T_4.

We conjecture the existence of a group of rules, obeyed by nature, due to which (1) the spectrum of a cell c and (2) a signaling from its neighboring cells determine which spectrum or spectra in target image will be chosen.

We introduce three models, where the difference lies in the construction of target image.

1. Minimal model.

This is a model with a fully defined set of possible cell states. For this version of the theory, we assume that the set of possible spectra of a cell which occurs after one cell event is determined only by the spectrum of a cell M and by the operator A_i. Thus, the target image is $\{T_1,T_2,T_3,T_4\}=\{A_1M,A_2M,A_3M,A_4M\}$.

Thus, the resulting matrix $M'(c)$ is the one which is pre-determined for the normal development of a particular organism for the proper shape formation (belonging to a graph (tree) of optimal cell events (TOE)).

In the minimal model, neighboring cells can affect the choice in the target image through signaling, but the target image itself is independent from signaling.

2. Normal model.

In this model, the deterministic dependence of the descendent (advanced) cell state on the state of the ancestor is also preserved. However, here we assume that for a given cell c with a cell state $(Se(c),M(c))$, among all possible matrices M^S(c) after signaling there exists a determined *normal signal* $N(c)$ and the matrix $M(c)^N$ corresponding to it:

$$M(c)^N=M(c)+h(c)N(c).$$

For example, we can assume that in the case of normal development the spatial arrangement of surrounding cells should not change. This means that each cell should have special types of interrelations with its neighbor cells (e.g. by the already mentioned “rule of complementarity”) because each pair of cells are produced by one mother cell. We can state that if a pair of cells has such interrelations, then the number $G(c,k)$ on the corresponding edge of the *graph of adjacency* should be equal to zero. In this case for the normal development the *normal signal* $N(c)$ will be equal to zero, which means that the resulting matrix $M(c)^N$ will be the one which is pre-determined for the normal development of a particular organism (belonging to TOE).

Thus, in this model the target image relies on the normal signaling:

$\{T_1,T_2,T_3,T_4\}=\{A_1M^N,A_2M^N,A_3M^N,A_4M^N\}$.

The “normal” model needs the signaling-dependent rules of choice of the cell event similar to those in the “minimal” model.

3. Direct model.

In this model, signaling (neighboring cells) directly influences the resulting spectrum. The target image is $\{T_1,T_2,T_3,T_4\}=\{A_1M^S,A_2M^S,A_3M^S,A_4M^S\}$.

Rules of choice

The *rules of choice* are applied to each of the suggested models for selecting the evolution result in target image set $\{T_1,T_2,T_3,T_4\}$ with taking into account the signaling received by a cell.

We define $M_0=(A_1M^S+A_2M^S+A_3M^S)/3$ as the *auxiliary image* of *c*. The auxiliary image contains the information of signaling. To make a choice in the target image, we compare the auxiliary image with the target image, and choose the best match. This is how the signaling from neighboring cells influence the result of cell event.

The match is described by the distance between spectra. For two matrices *M* and *M'*, we define the distance between them by $d(M,M')=||M-M'||_F$, where $||*||_F$ is the Frobenius norm. To calculate Frobenius norm, sum up the square of all entries, then take its square root.

We suggest the following *rules of choice:*

First, we calculate the distances $d_1=d(M_0,T_1)$, $d_2=d(M_0,T_2)$, $d_3=d(M_0,T_3)$. If $(d_1+d_2)/2<d_3$, then M_0 matches more with T_1 and T_2, and the cell divides into two cells with spectra T_1 and T_2. If $(d_1+d_2)/2>d_3$, then M_0 matches more with T_3, and the cell becomes another cell with spectrum T_3.

If both $(d_1+d_2)/2$ and d_3 are larger than a chosen threshold (which can depend on the cell sensitivity), then the results of division (T_1,T_2) and one-cell event in normal situations (T_3) do not match the auxiliary image, then this indicates an abnormal situation, manifesting in abnormal signaling (Figure 3).

As a reply for abnormal signaling a cell can undergo internal cell event (which includes dedifferentiation or transdifferentiation), “de-division”, apoptosis or stagnation, where

- transdifferentiation is the internal cell event resulting in the matrix *M* other than pre-determined for normal development (belonging to a graph (tree) of optimal cell events (TOE)); a cell acquires another cell state with spectrum T_4.

-dedifferentiation is the case of internal cell event resulting in the matrix *M* matching with one of the preceding states on the TOE; a cell acquires another cell state with spectrum T_4.

- de-division is the division of a cell with at least one daughter cell with matrix *M* fitting with a transdifferentiated or dedifferentiated matrix of a cell, corresponding to a particular case of a spectrum T_2. This is possible in some special cases in the direct model.

Also, in the direct model, since the signaling directly affects the target image, there exists a possibility that the result of A_3 corresponds to dedifferentiation or transdifferentiation.

When $((d_1+d_2)/2,d_3)$ is very close to the boundaries of these three cases, namely, when $(d_1+d_2)/2 \approx d_3$, $(d_1+d_2)/2$ or d_3 almost meets the threshold, the result is randomly chosen from either side. For example, if $(d_1+d_2)/2>d_3$, both are much smaller than the threshold, but $(d_1+d_2)/2-d_3$ is very small, the result is more likely to be T_3, and less likely to be T_1 and T_2. If it is exactly on the boundary, the chance is half-half for each side.

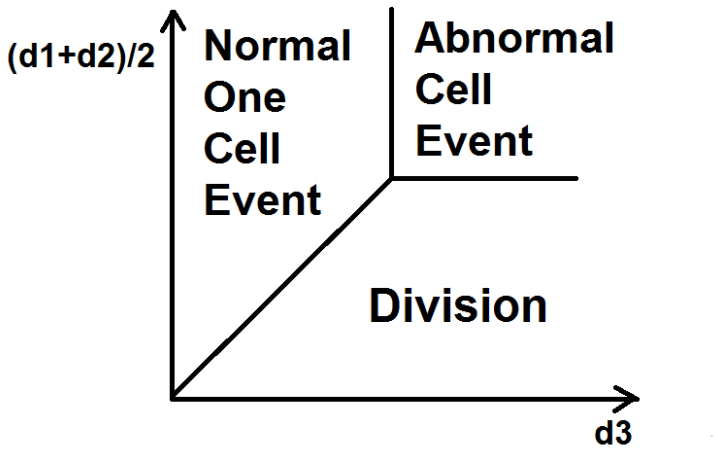

Figure 3: Rules of choice for a cell event. If $(d_1+d_2)/2<d_3$, then the cell divides into two cells under operators A_1 and A_2. If $(d_1+d_2)/2>d_3$, then the cell acquires another cell state under operator A_3. If both $(d_1+d_2)/2$ and d_3 are larger than a threshold, then the cell acquires an abnormal cell state under operator A_4.

The set of possible resulting spectra $\{M_{dv1}, M_{dv2}, M_{df}, M_{dd}, M_g, M_m, M_i, M, 0\}$ (as presented in Figure 1) is much larger than the target image set $\{T_1,T_2,T_3,T_4\}$, because, as it was already discussed, T_3 from A_3 can match M_{df}, M_g, M_m, M_i and M, and T_4 from A_4 can match M_{dd}, $M_{i.}$ and M.

When the change of epigenetic spectrum is determined, the change of spatial status is correspondingly determined. If M becomes A_1M and A_2M, then Se=(L,Or,Sh) becomes $Se_1=B_1(L,Or,Sh)$ and $Se_2=B_2(L,Or,Sh)$, if M becomes A_3M *or* A_4M, then Se becomes correspondingly $Se_3=B_3(L,Or,Sh)$ or $Se_4=B_4(L,Or,Sh)$.

If a cell acquires another cell state through operators (A_3,B_3) or (A_4,B_4), then one can use the change of cell state (Se,M) to determine the type of cell event which happens: differentiation, dedifferentiation, growth, movement, internal cell event, stagnation or apoptosis.

By this model, from a zygote we can calculate the spectrum matrix, location, shape, orientation of all cells at all time points. Then using this information, we can reconstruct the developmental graph and corresponding tissue, organ and organism shape at any time.

§3. Formalization of biological concepts.

Based on the proposed model, we suggest the formalization of several important biological concepts.

3.1. Cell potency, stem cells and differentiated cells

The cells in organism differ in their *cell potency* which is the ability to produce different cell types. The cells which have a high potency and can develop in many ways are called *stem cells*. It is important to distinguish two different types of stem cells: embryonic stem cells, making up an early embryo starting from the zygote, and stem cells of adult organisms which are located in the specialized tissues of adult organisms. The potency of embryonic stem cells is very high, which manifests itself in the ability to produce many types of differentiated tissues under different signaling. This phenomenon is called totipotency or pluripotency of embryonic stem cells. In a course of development the potency of the most cells in an organism decreases, and the terminally differentiated cells in adult organism normally have zero potency (can not develop or divide), thus they do not change with time. According to the proposed model, this implies that for each step, A_3 is the chosen operator, and the spectrum matrix does not change much. In the language of dynamical systems, this cell type is an attractor under the action

of operator A_3. Rigorously speaking, there exists one spectrum matrix M^* in this cell type, which is a fixed point of A_3, namely $A_3M^*=M^*$. Further assume that the norm of A_3 induced by the Frobenius norm is not greater than 1. Then for any spectrum matrix M and signaling S, we have

$$||A_3M^S-M^*||_F=||A_3(M^S-M^*)||_F\leq||M^S-M^*||_F,$$

which means that the distance between M^S and M^* does not increase. M^* has a neighborhood, in which any M satisfies

$$||A_3M-M^*||_F<||M-M^*||_F,$$

which means that without signaling, M is attracted to M^* after iteration of A_3. This attracting neighborhood, as a subset of all possible spectrum matrices, corresponds to a fully differentiated cell type. For each cell in this type, if the chosen operator is A_3, and the received signaling is not too strong such that M^S is still in this subset, then after the action of A_3, the result is still of this type. In this type, the rank of any spectrum matrix corresponds to zero potency.

For a stem cell of adult organisms, the general behavior is to divide asymmetrically, producing a stem cell of the same type, and a more differentiated cell. We can describe such stem cell similarly in our model. Assume there exists a spectrum matrix $M^\#$ of this stem cell type, which is fixed under operator A_1, namely $A_1M^\#=M^\#$. Similarly assume that the norm of A_1 induced by the Frobenius norm is not greater than 1. Then $||A_1M^S-M^\#||_F\leq||M^S-M^\#||_F$, and $M^\#$ has a neighborhood, in which any M satisfies $||A_1M-M^\#||_F<||M-M^\#||_F$. This attracting neighborhood corresponds to a stem cell type in adult organisms. For each cell in this type, if the chosen operators are A_1 and A_2, and the received signaling is not too strong such that M^S is still in this subset, then the result of A_1 is still of this type, while the result of A_2 might be of a more differentiated type.

3.2 Formalization of cell potency

In our model we propose that the potency *P(c)* of a cell *c* is positively related to the *rank* of its matrix (epigenetic spectrum). It can be expressed as:

$$P(c)=f\{rk[M(c)]\},$$

where *f* is an increasing function.

This allows to obtain an important quantification of the phenomenon of the cell potency and to impose additional conditions (restrictions) on the operators of cell events.

According to this assumption, with each cell event occurring in normal development the rank of resulting new matrix(es) should decrease or at least not increase. For the "fully differentiated" cell, which has potency equal to zero, the rank of its matrix should be the minimal possible one. We conjecture the existence of such a minimal rank, and assume that all types of fully differentiated cells in an organism should have matrices of this rank. Moreover, we conjecture that all matrices with this minimal rank can be divided into characteristic classes, each of which corresponds to one type of specialized cells.

We will take into account the asymmetry between operators of divisions by the request that operator A_1 should not increase the rank of matrix *M*, while operator A_2 can decrease, increase the rank of matrix *M* or leave it unchanged. Thus, with each division, at least one of the descendants will be with the decreased potency. For the "minimal" model it will always be true, for the "normal" and "direct" model it will be true for normal and weak (close to normal) signals. The situation, when as the result of A_2 acting on *M* the rank of the resulting matrix is increased corresponds to “de-division” (production of a dedifferentiated cell in a course of division) in the case of abnormal signaling.

One-cell events in normal situations (A_3) do not increase the rank of matrix M (the potency of a cell). Thus

$$f\{rk[A_i(M)]\} \leq f\{rk(M)\},\ i=1,3.$$

Dedifferentiation, which increases the cell potency, is a possibility of one-cell events in abnormal situations (A_4). Therefore operator A_4 can decrease, increase cell potency or leave it unchanged.

It can happen that

$$rk(A_3M^N) > rk(M),$$

which means that the target image contains a matrix with higher potency than the matrix of a given cell. This corresponds to the cell event dedifferentiation in normal situations, which is necessary, for example, for the formation of gametes, especially an ovule.

3.3. Inverse potency law and cell sensitivity

In Minarsky et al. (2018) we conjectured that the strength of the signal emitted by a cell *c* is inversely related to its potency *P(c)*. We can formalize this conjecture in a following way .

Let us consider a cell *c* with spectrum *M(c)*. Its corresponding matrix with signaling is

$$M^S(c)=M(c)+h(c)\Sigma_k G(c,k)M(k),$$

where *h(c)* is its sensitivity to incoming signals, and the summation takes over all neighbors of *c*. The strength of signal emitted by *c* is

$$t_0=||M(c)||_F.$$

The strength of received signal is

$$t=h(c)||\Sigma_k G(c,k)M(k)||_F.$$

We will assume that the cell potency is positively related to its sensitivity. We consider two cases: stem cells and fully differentiated cell.

For a stem cell, its potency (thus sensitivity) is high, therefore the strength of emitted signal t_0 should be relatively small compared with the strength of received signal t.

For a fully differentiated cell, its potency (thus sensitivity) is low, therefore the strength of emitted signal t_0 should be relatively large compared with the strength of received signal t.

3.4. Cellular tissue

Minimal and normal models make possible to define *cellular tissue*. Let us define F as a sequence of operators A_i, i=1,2,3,4. If we denote the spectrum of a zygote by M_z, then FM_z will be the spectrum of the cell produced by operator (event) sequence F from the zygote. Then we can define the cellular tissue of FM_z by $S_F = \{M: M=KFM_z, \forall K\}$, where S_F consists of all possible cell spectra produced from FM_z, K is a sequence of operators A_i, i=1,2,3,4. If the spectrum of a cell is in S_F, then this cell belongs to the biological tissue produced by the cell with spectrum FM_z and its descendants.

Cellular tissue can be defined also in the direct model, though the signaling makes the expression more complicated. Since signaling directly influences target image in the direct model, we cannot define cellular tissue likewise.

Discussion and Conclusion

In this paper, we propose a model of determining cell properties, which can be used to describe the developmental and regeneration processes.

One important comment which should be discussed is the modeling of a cell response to environmental changes (e.g. growth factors, oxygen, nutrients, etc.). We assume that this can be taken into account by proposition of “internal cell event” during which a code (matrix) of a cell is changed without any external cell event for a cell. Internal cell events are regarded as the necessary steps in realization of the determined developmental program, and they are included in the model exactly to be able to reflect a response of a cell to a set of biochemical factors, which should come to a cell at this step in a case of normal development (and as a result, the matrix will be changed). But it is assumed that internal cell events can also occur as a response to the environmental changes. One example of it is presented in Bessonov et al. (2019), when it is proposed as a rule that a cell changes its matrix as a response to the changed information from its neighbor cells. For the time being the details of other types of abnormal environmental changes are not taken into account.

Another important comment is that in the framework of our model, the process of carcinogenesis can be understood as a loss by cancer cells the ability to maintain a proper content of coding matrices, which in turn causes the disturbance of corresponding signaling cascades, governed by the code. A possible explanation is that some mutations that lead to cancer progression may influence the pathways responsible for the correct turnover of coding molecules, thus breaking from normal cell behavior. Another possibility may be a loss by cancer cells of the ability to accept the correct signals from the environment due to the pathologically lost sensitivity.

Our current model focuses on the dynamics of single cells. It is possible to develop mathematical results under this framework to better describe developmental processes on more macroscopic levels, such as gastrulation, axis formation and regeneration.

Another goal is to supplement the model with the experimental data. First, we can measure the quantities of different membrane molecules and cell behaviors, and use causal inference methods to determine the members of epigenetic spectrum, since epigenetic spectrum should be the distribution of membrane molecules that best describe the cell behavior. Next step is to determine the concrete forms of operators A_i, B_i.

Acknowledgments

The work of Yue Wang and Andrey Minarsky are supported by Simons Foundation (IHÉS program of mathematical biology). The work of N. Morozova was carried out within the framework of the state assignment to Komarov Botanical Institute RASN°AAAA-A18-118051590112-8. The research of A. Minarsky was supported by Ministry of Science and Higher Education of Russian Federation (assignment 1.9788.2017/BCh).

Author Disclosure Statement

No competing financial interests exist.

References

Adjei, S.A. and Heffernan, N.T. 2015. Improving learning maps using an adaptive testing system: PLACEments. In International Conference on Artificial Intelligence in Education (517-520). Springer, Cham.

Bando, T., Mito, T., Hamada, Y., et al. 2018. Molecular mechanisms of limb regeneration: insights from regenerating legs of the cricket Gryllus bimaculatus. Int. J. Dev. Biol. 62(6-7-8), 559-569.

Bessonov, N., Butuzova, O., Minarsky, A., et al. 2019. Morphogenesis Software based on Epigenetic Code Concept. Comput. Struct. Biotechnol. J. 17, 1203-1216.

Bryant, S.V., Endo, T., Gardiner, D.M. (2002). Vertebrate limb regeneration and the origin of limb stem cells. Int. J. Dev. Biol. 46, 887–896.

Byambaragchaa, M., Kim, D.J., Kang, M.H. et al. 2018. Site specificity of eel luteinizing hormone N-linked oligosaccharides in signal transduction. Gen. Comp. Endocrinol. 268, 50-56.

Cieslar-Pobuda, A., Knoflach, V., Ringh, M.V, et al. (2017). Transdifferentiation and reprogramming: Overview of the processes, their similarities and differences. BBA - Mol. Cell Res. 1864, 1359–1369.

Cunningham, T.J. and Duester, G. 2015. Mechanisms of retinoic acid signalling and its roles in organ and limb development. Nat. Rev. Mol. Cell Biol. 16(2), 110-123.

Delile, J., Herrmann, M., Peyriéras, N., et al. 2017. A cell-based computational model of early embryogenesis coupling mechanical behaviour and gene regulation. Nat. Commun. 8, 13929.

Frasch, M. 2016. Dedifferentiation, Redifferentiation, and Transdifferentiation of Striated Muscles During Regeneration and Development. Curr. Top. Dev. Biol. 116, 331–355.

Funk, R.H.W. 2015. Endogenous electric fields as guiding cue for cell migration. Front. Physiol. 6, 143.

Gilmour, D., Rembold, M., Leptin, M. 2017. From morphogen to morphogenesis and back. Nature 541: 311-320.

Gordon, N.K., and Gordon, R. 2016. Embryogenesis Explained. World Scientific, Hackensack, NJ.

Hayashi, S., Yokoyama, H., and Tamura, K. 2015. Roles of Hippo signaling pathway in size control of organ regeneration. Dev. Growth Differ. 57, 341-351.

Jopling, C., Boue, S., and Izpisua Belmonte, J.C. 2011. Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. Nat. Rev. Mol. Cell Biol. 12, 79–89.

Kikuchi, K. 2015. Dedifferentiation, Transdifferentiation, and Proliferation: Mechanisms Underlying Cardiac Muscle Regeneration in Zebrafish. Curr. Pathobiol. Rep. 3, 81–88.

Kragl, M., Knapp, D., Nacu, E., et al. 2009. Cells keep a memory of their tissue origin during axolotl limb regeneration. Nature 460(7251), 60-67.

Lander, A.D. 2013. How cells know where they are. Science, 339(6122), 923-927.

Law, R. and Levin, M. 2015. Bioelectric memory: modeling resting potential bistability in amphibian embryos and mammalian cells. Theor. Biol. Med. Model. 12(1), 22.

Levin, M. 2011. The wisdom of the body: future techniques and approaches to morphogenetic fields in regenerative medicine, developmental biology and cancer. Regen. Med. 6(6), 667-673.

Lobo, D., Beane, W.S. and Levin, M. 2012. Modeling planarian regeneration: a primer for reverse-engineering the worm. PLoS Comput. Biol. 8(4), e1002481.

McCusker, C., Bryant, S.V., Gardiner, D.M. (2015). The axolotl limb blastema: cellular and molecular mechanisms driving blastema formation and limb regeneration in tetrapods. Regeneration 2, 54-71.

McLaughlin, K.A. and Levin, M. 2018. Bioelectric signaling in regeneration: mechanisms of ionic controls of growth and form. Dev. Biol. 433(2), 177-189.

Meinhardt, H. 2013. From hydra to vertebrates: Models for the transition from radial- to bilateral-symmetric body plans, 207-224. In Capasso V., Gromov M., Harel-Bellan A., et al. eds. Pattern Formation in Morphogenesis. Springer, Berlin, Heidelberg.

Menshykau, D., Blanc, P., Unal, E. et al. 2014. An interplay of geometry and signaling enables robust lung branching morphogenesis. Development, 141(23), 4526-4536.

Minarsky, A., Morozova, N., Penner, R. et al. 2018. Theory of morphogenesis. J. Comput. Biol. 25(4), 444-450.

Morozova, N. and Penner, R. 2015. Geometry of morphogenesis, 331-350. In Mondaini R.P. eds. BIOMAT 2014: International Symposium on Mathematical and Computational Biology.

Morozova, N. and Shubin, M. 2013. The geometry of morphogenesis and the morphogenetic field concept, 255-282. In Capasso V., Gromov M., Harel-Bellan A., et al. eds. Pattern Formation in Morphogenesis. Springer, Berlin, Heidelberg.

Paque, S. and Weijers, D. 2016. Q&A: Auxin: the plant molecule that influences almost anything. BMC Biol. 14(1), 67.

Pauzi, M.N.M., Darus, M. and Siregar, S. 2018. Second Hankel determinant for a class defined by modified Mittag-Leffler with generalized polylogarithm functions. J. Math. Comp. Sci. 18(4), 453-459.

Pusapati, G.V., Kong, J.H., Patel, B.B., et al. 2018. CRISPR screens uncover genes that regulate target cell sensitivity to the morphogen sonic hedgehog. Dev. Cell. 44(1), 113-129.

Sagner, A., and Briscoe, J. 2017. Morphogen interpretation: concentration, time, competence, and signaling dynamics. Wiley Interdiscip. Rev. Dev. Biol. 6(4), e271.

Samsonraj, R.M., Rai, B., Sathiyanathan, P., et al. 2015. Establishing criteria for human mesenchymal stem cell potency. Stem Cells. 33(6), 1878-1891.

Tanaka, E.M. 2016. The molecular and cellular choreography of appendage regeneration. Cell 165(7), 1598-1608.

Tuazon, F.B. and Mullins, M.C. 2015. Temporally coordinated signals progressively pattern the anteroposterior and dorsoventral body axes. Semin. Cell Dev. Biol. 42: 118–133.

Wolpert, D.M. and Flanagan, J.R. 2016. Computations underlying sensorimotor learning. Curr. Opin. Neurobiol. 37, 7-11.