GENE NETWORKS BEHAVIOR
IN A SERIES OF SUCCESSIVE CELL DIVISIONS

Galimzyanov A.V.
Department of Physicochemical Biology and Epigenetics, Ufa Research Center, RAS, Ufa, Russia,
e-mail: galim@anrb.ru

Key words: gene networks, cell division, modeling

SUMMARY

Motivation: When modeling the dynamics of the gene networks in developing cell ensembles, one should account for the processes of intercellular distribution of different molecular substances participating in the formation and regulation of functional regimes in cellular gene networks, i.e., determining the levels of gene activity in the cells.

Results: A computer model is constructed for the process of gene product distribution among daughter cells during the maternal cell division. An algorithm is proposed for traversing the tree of cell divisions with unique gene networks that makes it possible to obtain gene expression profiles in the cells of intermediate and final generations. Peculiar features are analyzed in the functioning of a gene network, which includes three cyclic digene systems with negative feedback, in successive cell generations.

Availability: The MAGeneNT software is available on request from the author.

INTRODUCTION

In successive cell generations, the transmission of extra-genome regulatory molecules, e.g., proteins, proceeds from ancestor to descendant. The gene-expression profile in the maternal cell before division determinates the initial data set of gene networks (GN) in the daughter cells. In this case the gene subnetwork dynamics can be either changed or preserved. In this connection in order to better appreciate the mechanisms controlling the development of living systems one should be able to trace the GN dynamics within cell lines.

METHOD AND ALGORITHM

Distribution model of substance’s molecules among daughter cells

The division of cell $c_{i,d}$ (i is the cell number in the d-th generation) gives rise to two daughter cells $c_{i_1,d+1}$ and $c_{i_2,d+1}$. Molecules of the products of the j-th gene enter cell $c_{i_1,d+1}$ with probabilities $p_{mj}$ (for each RNA molecule of the j-th gene) and $p_{pj}$ (for each protein molecule of the j-th gene) or they enter cell $c_{i_2,d+1}$ with probabilities $q_{mj}$ and $q_{pj}$. Each molecule has two possible outcomes (entry into one of the two daughter cells), the probabilities of outcomes therewith remain constant ($p = q = \frac{1}{2}$). Thus, for each cell $c_{i,d}$ ($i = 1, 2^d$, $d = 1, D$, where D is the number of generations under observation) and some substance $\mathcal{R}$ we have the Bernoulli distribution of $n$ trials, where $n$ is the amount of the molecules of substance $\mathcal{R}$ in a cell.
The number of successes \( S_n \) in \( n \) trials is a random value with the binomial distribution function. In our case it is interesting to compute not the probability of getting exactly \( k \) successes, but to find, with high probability (0.99), the range \((\alpha, \beta)\), in which value \( S_n \) lies. According to the de Moivre-Laplace limit theorem, for fixed \( x_1 \) and \( x_2 \) at \( n \to \infty \) we have:

\[
P\{np + x_1 \sqrt{npq} \leq S_n \leq np + x_2 \sqrt{npq} \} \to \Phi(x_2) - \Phi(x_1),
\]

where \( x_1 = (\alpha - np) h \), \( x_2 = (\beta - np) h \), \( h = 1 / \sqrt{npq} \), \( \Phi(x) = (1 / \sqrt{\pi}) \int_{-\infty}^{x} e^{-y^2 / 2} \, dy \) (the function is tabulated).

In the problem under consideration \( x_1 < 0, x_2 > 0, x_2 = |x_1| \), therefore \( \Phi(x_2) - \Phi(x_1) = 2\Phi(x_2) - 1 \); at \( P = 0.99 \): \( \Phi(x_2) = 0.995 \), \( x_2 = 2.61 \). Hence, the range \((\alpha, \beta)\) is defined, in which we randomly select the value of variation \( \Delta \), from the half-value concentration of substance \( R \) in a cell before division (at the end of life cycle).

**Algorithm “Greence” for walking through a tree of cell divisions**

The algorithm is based upon a ripple-through tracing cell lines from the original parent cell to all cells of the final generation (Fig. 1). A V-element is introduced into consideration to walk through a tree of cell divisions, with encountered nodes being discovered one by one. To discover node \( v_{i,d} \) (\( i \) is the number of a cell in the \( d \)-th generation) means to calculate the GN dynamics in two daughter cells \( c_{i,d+1} \) and \( c_{i,d+1} \) originating from parent cell \( c_{i,d} \) that corresponds to a given node. According to concentration levels of regulatory substances at the end of life cycles of cells \( c_{i,d+1} \) and \( c_{i,d+1} \), GN initial data is produced in their derivative cells, i.e., \( c_{i,2d+1} \), \( c_{i,2d+1} \) and \( c_{i,2d+2} \), \( c_{i,2d+2} \), respectively. The nodes \( (v_{i,d+1} \) and \( v_{i,d+1} \) \) that represent such daughter cells \( (c_{i,d+1} \) and \( c_{i,d+1} \) \) are said to be actualized. Next only one of the daughter cells (either \( c_{i,d+1} \) or \( c_{i,d+1} \)) is taken as a parent, and its representing node is discovered. The other node gains the status of a delayed alternative and simply waits for its turn. It will be discovered at a later time, and then the walk will begin through the branches running from this already discovered node to the nodes of the upper level. On reaching the final layer (cell generation), the V-element comes back along the same fragment of the ripple-through path already finished to the nearest delayed node (it is actualized) and begins to complete the delayed alternative ripple-through path.

**Figure 1.** Walking through the graph of cell divisions with the use of recursion. Designations: circle with dark inset refers to the \( V \)-element, white circle to undiscovered node, dark circle to discovered node; shaded circle to delayed node; marks (*) near the circles show a memorized route to delayed nodes; \( d = 1, D \) is the number of generation.
The $V$-element is an argument and output of the recursive function that determines the direction of its next step and the type of operations to be performed on the encountered node (whether it should be actualized, discovered, memorized as a point of the route or cut off). The use of recursion makes it unnecessary to store the information on each cell of the generation preceding a new one. The $V$-element “brings” initial data for a new step and also leaves them in the delayed nodes of an alternative route (not more than the number of generations); subtrees with all nodes discovered and without delayed nodes are cut off. The approach described above enables one to calculate the dynamics of intracellular GN in successive cell generations up to 30 in number comparable with Hayflick’s limit (~50) for a number of eukaryotic cell divisions (Hayflick, Moorhead, 1961).

IMPLEMENTATION AND RESULTS

The model and algorithm are software realized in the concept of object-oriented programming and, as computing modulus, included in the software package MAGeneNT (Galimzyanov, 2005). The MAGeneNT system is intended for analyzing the dynamics of control gene networks based on the method of generalized threshold models (GTM) (Tchuraev, 1991) and makes it possible to describe the processes of positional information distribution, functioning of intracellular GN with the account for the values of regulatory substance concentrations, transmittance in a series of successive generations of the cells of extra-genomic regulatory molecules as well as intercellular interactions.

The approach developed by this study is applied to gene network TCDS$^{-3}$, which includes three cyclic digene systems with negative feedback, CDS$^{-3}$, being a genetic toggle switch (Fig. 2). The TCDS$^{-3}$ model is constructed in terms of GTM formalism (Tchuraev, 1991). The model takes into account the following parameters: $m_i(t)$ and $r_i(t)$ are the concentrations of mRNA and proteins, respectively, expressed as the numbers of molecules per cell ($j = 1, 7$); $a_{ij}$ is the unit intensities of transcription (the promoter force) (molec./cell × min$^{-1}$); $a_{2j}$ is the unit intensities of translation (molec./cell × min$^{-1}$); $b_{1j}$ is the rates of mRNA degradation (molec./cell × min$^{-1}$); $b_{2j}$ is the rate of protein degradation (molec./cell × min$^{-1}$); $P_{ij}$ constants describing threshold concentrations of protein product of $i$-th gene, that are necessary to inhibit the synthesis of the transcripts of $j$-th genes (these constants are determined by the affinity of regulatory substances to the protein-binding sites and by the degree of multimer); $T$ is the cell cyclic duration (min); $K = 10$ is the number of cell generations. The parameter values for the typical CDS$^{-3}$: $a_{11} = 2.3$, $a_{12} = 1.3$; $a_{21} = 1.5$, $a_{22} = 1.0$; $b_{11} = 0.3$, $b_{12} = 0.22$; $b_{21} = 0.013$, $b_{22} = 0.0115$, $P_{12} = 10$, $P_{21} = 9$, $T = 36$. Initial conditions: $m_1(t_0) = 7$, $m_2(t_0) = 6$, $r_1(t_0) = 300$, $r_2(t_0) = 250$. The kinetic curves for the molecular components (mRNA and proteins) of the GN were calculated individually for every cell of each generation. The initial protein and mRNA concentrations for each gene in two daughter cells were taken to be half the concentrations of these substances in the parent cell at the end of the life cycle (before division), with the account for random deviation according to the binomial law.

Figure 2. Scheme for the gene network TCDS$^{-3}$. The negative interaction among genes (transcription repression) is denoted with a dotted line, and the positive interaction (transcription activation) with a solid line.
In the experiments in silico we obtained quantitative relations among the cells being heterogeneous by the TCDS\(^{−1}\) functional states, in a series of generations. As shown in Table 1, after a series of cell divisions cell subpopulations with four different epigenotypes occur and persist in succeeding generations.

<table>
<thead>
<tr>
<th>d</th>
<th>Functional states</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S₁ = (0; 0; 1; 0; 0; 0; 0)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>S₂ = (0; 1; 1; 0; 0; 1; 0)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>S₃ = (0; 1; 0; 1; 0; 0; 1)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>S₄ = (0; 1; 0; 0; 1; 1; 0)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>S₅ = (0; 1; 0; 1; 1; 0; 0)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>S₆ = (0; 1; 0; 0; 0; 1; 0)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>S₇ = (0; 1; 0; 0; 0; 0; 1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>S₈ = (0; 1; 0; 0; 0; 0; 0)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>S₉ = (0; 1; 0; 0; 0; 0; 0)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>S₁₀ = (0; 1; 0; 0; 0; 0; 0)</td>
<td></td>
</tr>
</tbody>
</table>

The functional states: \(Sₖ\) \((k = 1,2,3,4,5,6,7,8,9,10,11,12)\) is functional state of j-th gene \((1 – \text{active}, 0 – \text{non-active})\) \((j = 1,7); d – \text{is the number of generation.}\)

**DISCUSSION**

Earlier experiments in silico were carried out on mathematical model of CDS\(^{−1}\), being a dynamic epigene, that is a cyclic system of genes with more than one inherited functional state, or epigenotype (Tchuraev et al., 2006). A new property of dynamic epigenes was confirmed, i.e., metastability of some epigenotypes, which was predicted theoretically (Tchuraev, 2006) and found in experiments in vivo (Stupak E., Stupak I., 2006). These metastable states realize one of possible mechanisms of primary differentiation of the gene activity patterns that occurs during ontogenesis of multicellular organisms. It is showed on TCDS\(^{−1}\) model, that even systems in which CDS\(^{−1}\) are typical elements, can possess this property.

**REFERENCES**


