# Discrete stochastic model for self-renewal and differentiation of progenitor cells

Muhammad Sahimi,[1] Ali Reza Mehrabi,[1] and Faramarz Naeim[2]
[1]*Department of Chemical Engineering, University of Southern California, Los Angeles, California 90089-1211*
[2]*Department of Pathology and Laboratory Medicine, School of Medicine, University of California, Los Angeles, California 90024*
(Received 4 November 1996)

We propose a discrete model for self-renewal and differentiation of hematopoietic stem cells based on the notion that, in a favorable environment, the commitment of the cells to a particular pair of progeny is a stochastic event with the possibility of either self-renewal or differentiation. Regulatory mechanisms are incorporated into the model, as is diffusion of the cytokines that carry the signals for such mechanisms. The model can produce chaotic states, and is shown to be capable of predicting some key features of the experimental data. [S1063-651X(97)50703-3]

PACS number(s): 87.10.+e, 87.22.−q

Dynamical processes in biological systems are some of the most complex phenomena in nature [1]. Traditionally, differential equations of population dynamics have been used [2] to model them. Such models provide information about the average properties, but cannot provide any insight into the effect of fluctuations and the spatial structure of the systems on the properties of biological phenomena, whereas such factors play a key role in the dynamics of the phenomena. Most appropriate for taking into account the effect of the fluctuations and the spatial structure of the environment are discrete or cellular automata models, in which a lattice site can take on a small number of states, and its evolution at the next time step depends on its present state and the environment around it. A few biological phenomena, such as the immune response, have already been modeled [3–7] using such models. An important and unsolved task is the development of a discrete model for predicting self-renewal and differentiation of progenitor cells, such as the bone marrow stem cells. Although several models have already been developed [8–12], they are too simple and usually take into account only one specific aspect of the problem, and thus they are not general enough, nor do they contain all the essential ingredients of the phenomenon, to be useful. In this paper we develop a discrete model which takes into account the effect of the most important factors that influence self-renewal and differentiation of hemapoietic stem cells. Our paper's goal is to study the *dynamics* of self-renewal or differentiation of the stem cells; their *spatial* distribution is the subject of a separate paper [13], and is not considered here.

The bone marrow is divided into irregular and interconnected regions by bone trabeculae, and consists of a complex hematopoietic cellular component that continuously undergoes self-replication and differentiation processes. The hematopoietic component is supported by a microenvironment composed of vascular structures, stromal cells, and a complex extracellular matrix (ECM). The stromal cells are the main generators of the ECM, and along with the accessory cells (such as *T*-lymphocytes and monocytes) are involved [14] in the production of the cytokines. The cytokines are soluble substrates that play a key role in the regulation of hematopoiesis by demonstrating stimulatory or inhibitory effects on them [15,16]. The bone marrow provides a proper spatial organization for cell-cell, cell-matrix, and cell-cytokine interactions [17]. This spatial organization and its average productivity are governed by the interactions between various components, which also determine the fate of a stem cell.

We consider a lattice in which a randomly selected fraction $f_1$ of its sites are occupied by the cytokines, a fraction $f_2$ represents the stem cells, another fraction $f_3$ contains the stromal cells, and the remaining fraction $f_4=1-(f_1+f_2+f_3)$ is the ECM. Each fraction $f_i$ is divided into four different subfractions $g_{ij}$. For example, the cytokines represent biological materials that favor the production and/or proliferation of various types of cells. We assume that a randomly selected fraction $g_{11}$ of the cytokine sites favor proliferation of the erythroid (E) cells, a fraction $g_{12}$ favors the megakaryocytes (Megs), another fraction $g_{13}$ favors the granulocyte-macrophage (GM) lineage, and the remaining fraction $g_{14}$ favors the self-renewal of the stem cells. Thus, due to the complexity of bone marrow, the parameter space of the model is quite large. But the values that we use for $f_i$'s and $g_{ij}$'s are biologically reasonable and believed to represent a typical bone marrow. For example, the stem cells have highly specific homing properties, self-renewal potential and multilineage differential capability. However, under normal conditions, only a small fraction of such cells enter the cell cycle that lead to the daily production of billions of end-stage mature hematopoietic cells.

It has been proposed [17], based on *in vivo* and *in vitro* studies, that the commitment of a stem cell to a particular pair of progeny of given potentials is a random event, and thus, at the moment of commitment, there is the possibility of either self-renewal or differentiation into new types of cells. However, the fate of a stem cell is influenced by the environment around it, with the cytokine, the ECM, and the stromal cells all playing a role. Thus we assume that a stem cell commits itself only if the environment around it is favorable, which is an environment that contains at least three sites one each with the cytokine, the ECM, and a stromal cell. We restrict the environment around a stem cell to be the set of its nearest-neighbor sites. It is not difficult to expand this set and include the next-nearest neighbors, the third-nearest neighbors, and so on. Thus, at each time step, we decide how a stem cell evolves by checking the environment around it. If it is unfavorable, nothing will happen to the

 

stem cell; otherwise the stem cell evolves stochastically: it differentiates into an E cell, a GM cell, or the Megs, or it self-renews, with an evolution probability $p_e$ proportional to the total number of the nearest-neighbor sites that contain the cytokine, the ECM, and stromal cells that favor the production of the four types of cell. If a stem cell does evolve, its site is filled with another stem cell, so that the fraction of the stem-cell sites is constant.

Once all the stem-cell sites (SCS) are checked, the spatial organization of the system evolves. Normally, the spatial distributions of the ECM, the stromal and the stem cells vary with time very slowly, whereas, comparatively speaking, the cytokines diffuse relatively fast in the bone marrow, and redistribute themselves in the system. We have recently shown [13] that, diffusion of the cytokines is a crucial factor in the development of the spatial structure of the bone marrow. Thus, at each time step, after checking the SCS, we allow the cytokines to diffuse in the lattice, but do not change the spatial distributions of the other constituents of the system. The diffusion is simulated by a random walk: We consider a cytokine site $i$, pick one of its nearest neighbors $j$ at random, and move the cytokine to $j$, unless $j$ is occupied by a stem cell. Diffusion and redistribution of the cytokines change the environment around a stem cell, and thus an unfavorable environment can become favorable as the system evolves. Once diffusion of all the cytokines is complete, the process time is increased by one unit. The environment around the SCS is examined again for differentiation or self-renewal, the cytokines are allowed to diffuse again, the process time is increased by one unit, and so on. Each time a stem cell differentiates into another cell, $m$ units of that cell are produced, and thus we measure the concentrations of the cells in units of $m$.

Two other important biological facts must be incorporated in the model. One is that, after some time the produced cells either leave the bone marrow, or die. Thus, we assume that the average residence time of any cell in the system is $\tau_r$, i.e., the total concentration of any cell at any time $t$ is reduced by the amount of that cell that was produced at time $t-\tau_r$. In general, the residence time of a cell in the bone marrow is not necessarily the same as its average lifetime $\tau_\ell$, but for the sake of simplicity we take $\tau_\ell = \tau_r = \tau_1$ for all the cells. The second biological fact is that bone marrows respond to negative or positive regulatory feedbacks. That is, if at any time the concentration of a particular cell is too large, a negative regulatory mechanism is triggered that suppresses temporarily the production of that cell, while a positive regulatory feedback triggers normal production of the cell again, when its concentration becomes too small. We incorporate this fact into our model by calculating, at any time $t$, the *average* concentration of each type of cell, where the averaging is taken over the *last* $\tau_2$ time steps. Then, if the total concentration of a given type of cell at any time $t$ is *larger* than its average over the last $\tau_2$ time steps, we suppress its production temporarily. This is done by deactivating all the cytokine, ECM, and stromal cells that favor the production of that cell. In the subsequent time steps, such sites are not counted for calculating the evolution probability $p_e$ of the stem cells. Since the concentration of a cell at any time $t$ is reduced by the amount that was produced at time $t-\tau_1$, suppression of its production causes a reduction in its total concentration over the subsequent time steps. This concentration is monitored and compared with its average over the last $\tau_2$ time steps. Once the concentration of the cell becomes smaller than its average, the positive regulatory feedback triggers the activation of the deactivated sites. Since the presence of the newly produced cells does not affect self-renewal or differentiation of the stem cells, we only keep track of the time dependence of their concentrations, and ignore their spatial distribution.

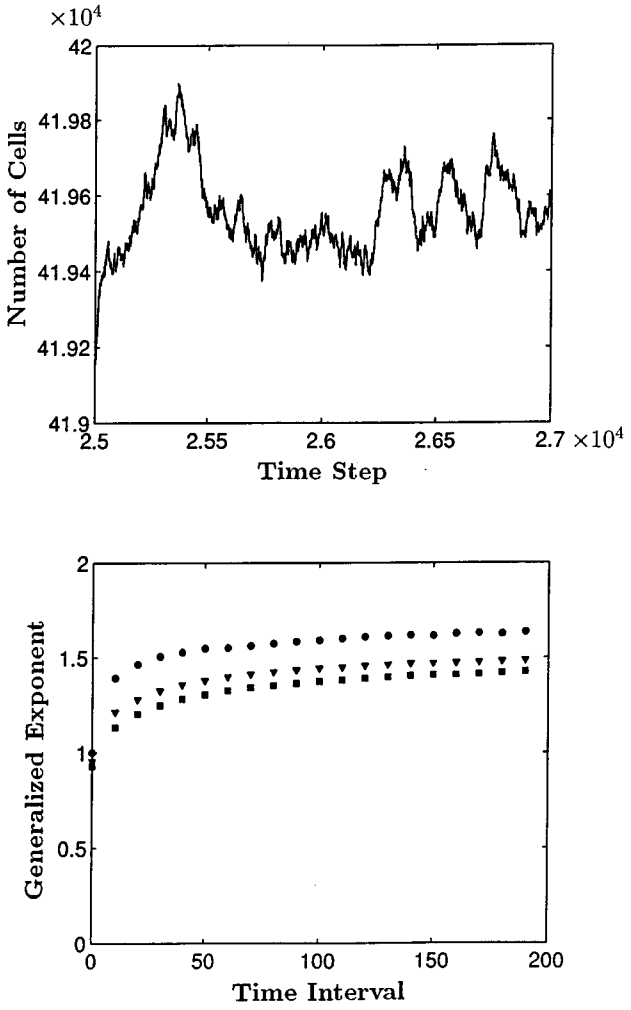


FIG. 1. Time variations (in Monte Carlo steps) of the predicted number of the GM cells (top) and its fractal analysis (bottom) which shows, from top, $D_0$, $D_1$, and $D_2$.

Most of our simulations were carried out with $50\times50\times50$ cubic lattices, which proved to provide statistically reliable results. The simulations parameters were $f_1=0.45$ (the cytokines), $f_2=0.30$ (the ECM), $f_3=0.15$ (the stromal cells), and $f_4=0.10$ (the stem cells). We also used $g_{i1}=0.28$ for the production of the E cells, $g_{i2}\simeq0.7$ for the GM cells, $g_{i3}\simeq0.01$ for the Megs, and $g_{i4}\simeq0.01$ for self-renewal of the stem cells, where $i=1-3$ (see above). These numbers are believed to be a reasonable representation of an actual bone marrow [17].

Figure 1 shows results for the time dependence of the concentration $C(t)$ of the GM cells for $\tau_1=5000$ and $\tau_2=500$. If we start the simulation with no GMs, then initially their concentration increases and attains a large maximum

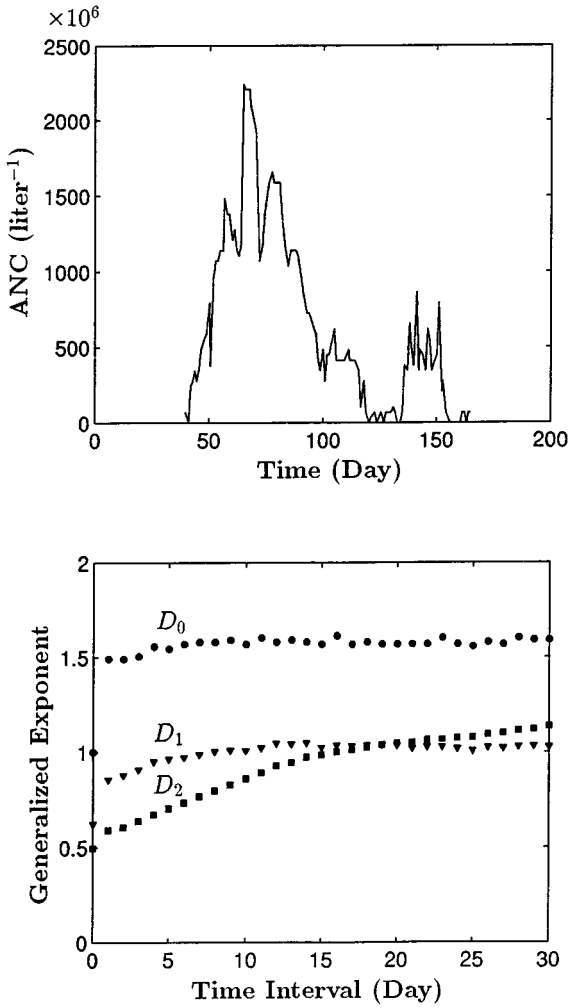


FIG. 2. Time variations (in Monte Carlo steps) of the absolute neutrophil count [19] (ANC) (top) and its fractal analysis (bottom).

relatively quickly, beyond which it reaches a quasi-steady state and varies with time around its steady-state value. Similar results are obtained for the other types of cells. If we use other values of the fractions $f_i$ and $g_{ij}$ in the simulations, we obtain qualitatively similar results, except that, e.g., a larger $g_{ij}$ implies larger concentrations of the cell type $j$, and thus larger fluctuations in its concentration.

The concentration fluctuations are reminiscent of a chaotic state. To characterize the chaotic state, consider the discrete time series $\{C(i)\}$, $i=1,2,\ldots,$ such that $C(i)$ and $C(i+1)$ are related by the functional equation, $C(i+1)=h[C(i)]$. We assume that the series has reached a quasi-steady state. An important feature of chaotic systems is their memory loss, i.e., their behavior at any time $t$ is independent of that at much earlier times [18]. To characterize the memory loss, we first consider the values of $C(t)$ in the domain $1,2,\ldots,i$, and divide the range of $C$ into $N$ equal intervals with a length $\epsilon=[\max(C)-\min(C)]/N$. Suppose that $p_i$ is the probability that a value of $C(t)$ is in the $i$th interval of the range of $C(t)$, and that $p_{ij}(t)$ is the joint probability of finding $C(i)$ in the $i$th interval and $C(i+t)$ in the $j$th interval. For each time interval $t$, we calculate a set of joint probabilities $p_{ij}(t)$ and compute time-dependent generalized exponents $D_q(\epsilon,t)=\ln[\Sigma_i\Sigma_j p_{ij}^q(t)]/[(q-1)\ln\epsilon]$ for

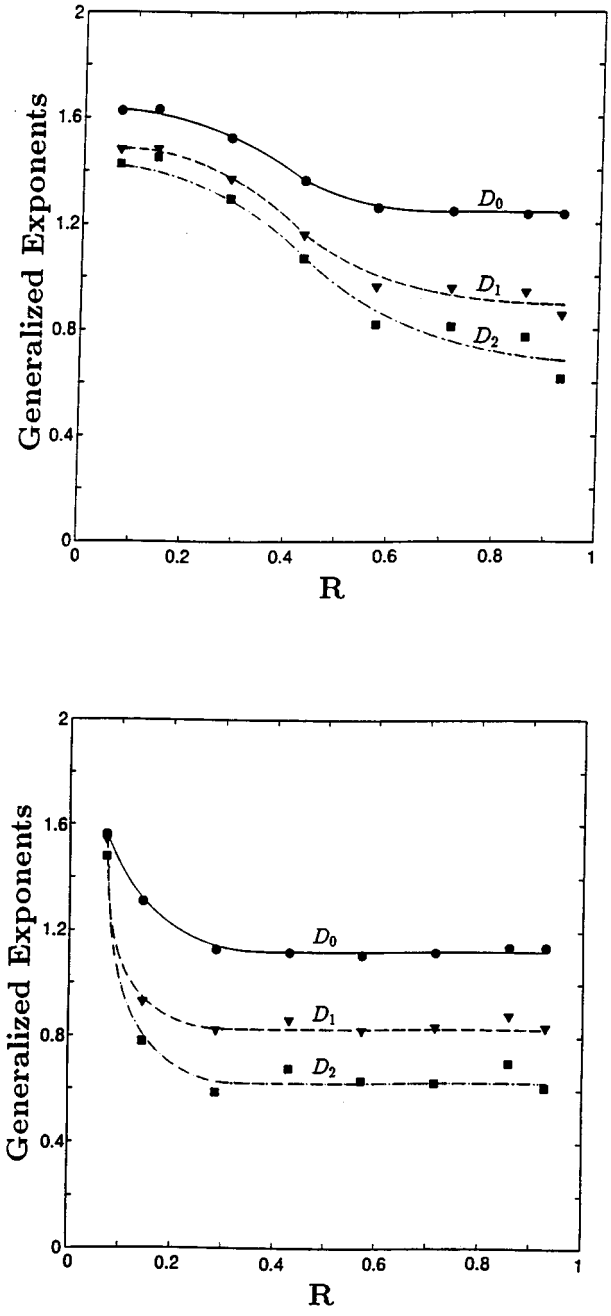


FIG. 3. The dependence of the generalized exponents $D_q(\epsilon,t)$ on $R=\tau_2/\tau_1$ for the Megs (top) and the GM cells (bottom).

$q\neq 1$, and $D_1=[\Sigma_i\Sigma_j p_{ij}\ln p_{ij}]/\ln\epsilon$ for $q=1$. $D_q$ characterizes a chaotic system [18] such that for $q=0$, 1, and 2 it represents, respectively, the fractal, information, and correlation dimensions of the data set. $D_2$ is related to the scaling of the two-point correlation function, and for $q>2$ $D_q$ is related to higher correlations in the data.

To calculate $D_q(\epsilon,t)$ we fix $t$ and construct a table with $N^2$ elements for $p_{ij}(t)$. We then find the intervals to which $C(i)$ and $C(i+t)$ belong. If, e.g., they belong to the $i$th and $j$th intervals, respectively, we add 1 to the corresponding $(i,j)$ cell in the table. Then $p_{ij}(t)$ is the value assigned to each cell normalized by the total number of pairs $[C(i),C(i+t)]$ that we find in the time series. By varying $t$ from 0 to a given value, we calculate $D_q(\epsilon,t)$. If $C(i)$ and

$C(i+t)$ are not related to each other, then $p_{ij}(t)=p_i p_j$, in which case [18] $D_q(\ell,t)=2D_q(\ell,0)=2\lim_{\ell\to 0}D_q(\ell,0)$. We use this principle to characterize the degree of chaos in the time variations of $C(t)$ and its memory loss, since [18] if $D_q(\ell,t)$ is doubled after a certain time, then the time series has lost its memory and behaves totally chaotically. However, if $D_q(\ell,t)$ is not doubled as $t$ becomes large, the time series is only partially chaotic.

Figure 1 shows the results for the GM cells (similar results are obtained for the Megs and the E cells). The fractal dimensions start with low values and, for short time intervals, increase. As the time intervals *increase*, $D_q$ attains its asymptotic values. In particular, $D_0$ starts out at a value of unity and for large time intervals reaches a value of about 1.6, close to completely chaotic state. However, $D_q$ does not attain values twice its initial values, characteristic of completely chaotic systems, since the system has memory and is influenced by regulatory mechanisms, as we always calculate the average cell concentrations over a finite time interval $\tau_2$ to trigger the negative or positive regulatory feedback mechanisms.

How can the model be tested against experimental data? Because of the wide physiological variability from experiment to experiment, absolute numbers of the produced cells are of limited quantitative value. A sensible test of the model is to see whether the data show trends toward a system with a degree of chaos as predicted by our model. We consider a typical case study reported by Boulad *et al.* [19], and comment on other possible scenarios. They reported on a patient who was treated for embryonal rhabdomyosarcoma with bone marrow transplantation followed by administration of recombinant human granulocyte-macrophage colony-stimulating factor (rHuGM-CSF) on day 25 posttransplant for a 28 day course of intravenous rHuGM-CSF at 250 $\mu$g/m$^2$ per day. Figure 2 shows his absolute neutrophil count (which are the end product of the GM cells) during the first 150 days of treatment, and its fractal analysis. These are similar to those shown in Fig. 1, and in particular, the generalized exponent $D_0\simeq 1.6$, similar to that of our model.

However, the analysis of other data [20] indicate that, the limiting values of $D_q$ can vary between 0.5 and 1.7. In general, the system's behavior is controlled by *both* $\tau_1$ and $\tau_2$. A large $\tau_2$ implies a more regulated system, since $\tau_2$ represents the memory of the system. $\tau_1$ also influences the behavior of the system, since if the cells leave the system too soon, the fluctuations in the concentrations increase. Thus, the limiting values of $D_q$ may be controlled by $R=\tau_2/\tau_1$. We show in Fig. 3 the dependence of $D_q$ on $R$ for the GM cells and the Megs. For low values of $R$, i.e., large $\tau_1$s and small $\tau_2$s, the system is almost completely chaotic, whereas for large values of $R$, i.e., large values of $\tau_2$, $D_q$ is small, in agreement with the data [20].

An important issue in most biological systems [1] is the role of the competition between the regulatory mechanisms and the stochasticity in the development of their temporal and spatial organizations. Our model predicts that this competition is in fact the main controlling factor in the time evolution of differentiation and proliferation of progenitor cells by forcing the system to hover between chaos $[D_q(\ell,t)\simeq 2\,D_q(\ell,0)]$ and order $[D_q(\ell,t)\simeq D_q(\ell,0)]$. Moreover, the intensity of chaos in biological systems can be characterized by the generalized exponents $D_q$ which, as we showed above, can be measured experimentally. Elsewhere [13] we have shown that this competition also organizes the spatial structure of bone marrow in a fractal manner with a well-defined fractal dimension, thus leading to unified description of the spatial and temporal distributions of the cells in bone marrow.


[1] S. A. Kauffman, *The Origins of Order* (Oxford University Press, New York, 1993).

[2] See, e.g., *Theoretical Immunology*, edited by A. S. Perelson (Addison-Wesley, New York, 1988).

[3] M. Kauffman, J. Urbain, and R. Thomas, J. Theor. Biol. **114**, 527 (1985).

[4] G. Weisbuch and H. Atlan, J. Phys. A **21**, L189 (1988).

[5] R. B. Pandey and D. Stauffer, J. Stat. Phys. **61**, 235 (1990); R. B. Pandey, J. Phys. A **23**, 4321 (1990).

[6] D. Chowdhury, M. Sahimi, and D. Stauffer, J. Theor. Biol. **152**, 263 (1991).

[7] M. Sahimi and D. Stauffer, Phys. Rev. Lett. **71**, 4271 (1993); D. Stauffer and M. Sahimi, J. Theor. Biol. **166**, 289 (1994); D. Stauffer, Int. J. Mod. Phys. C **5**, 513 (1994).

[8] G. Van Zant and E. Goldwasser, Blood **53**, 946 (1979).

[9] T. Nakahata, A. J. Gross, and M. Ogawa, J. Cell. Physiol. **113**, 455 (1982).

[10] R. Mehr and Z. Agur, BioSystems **26**, 231 (1992).

[11] M. Ogawa, Blood **81**, 2844 (1993).

[12] J. E. Till, E. A. McCulloch, and L. Siminovitch, Proc. Natl. Acad. Sci. USA **51**, 29 (1964).

[13] F. Naeim, F. Moatamed, and M. Sahimi, Blood **87**, 5027 (1996).

[14] A. Kimura, O. Katoh, H. Hyodo, and A. Kuramoto, Br. J. Haematol. **72**, 486 (1989).

[15] M.A.S. Moore, Blood **78**, 1 (1991).

[16] K. Kaushansky and P. A. Karplus, Blood **82**, 3229 (1993).

[17] For a review see, F. Naeim, *Hematol. Pathol.* **9**, 107 (1995).

[18] H.G.E. Hentschel and I. Procaccia, Physica D **8**, 835 (1983); P. M. Gade and R. E. Amritkar, Phys. Rev. Lett. **65**, 389 (1990).

[19] F. Boulad *et al.*, *Bone Marrow Transplant.* **13**, 661 (1994).

[20] M. Sahimi, A. R. Mehrabi, and F. Naeim (unpublished).