A STOCHASTIC COMPARTMENTAL APPROACH TO MODELING AND SIMULATION OF CANCER SPHEROID FORMATION AND EVOLUTION

Mónica F. Bugallo*, Shishir Dash*, Galina Botchkina*, Marco Lops†, Petar M. Djurič*

* Stony Brook University, Stony Brook, NY 11794-2350 USA
† INPT/IRIT/ENSEEIHT - Toulouse, France & University of Cassino, Italy
monica@ece.sunysb.edu

ABSTRACT

In this paper we model and simulate a biological system describing the evolution of cancer stem cells into tumors. Starting from some basic hypotheses about the behavior of these cells, we develop a model that mimics the evolution of a system of cancer stem cells and show how random-set-theory naturally leads to a generation algorithm. Computer simulations demonstrate the potential of our approach by using simple random sampling rules and a lattice environment.

Index Terms— Random-set-theory, cancer stem cells, high dimensional systems.

1. INTRODUCTION

Human epithelial cancers remain incurable. The ineffectiveness of standard anti-cancer drugs has recently been attributed to the existence of rare, highly drug resistant tumor-driving cells, so-called cancer stem cells (CSCs) [1, 2].

Models for stem cell and CSC evolution have already been proposed [1, 3]. However, most of the existing models are based on deterministic approaches, which are known to be inaccurate in practical scenarios [4]. It is known that tumor cells acquire multidrug resistance as they assemble into multicellular spheroids [5]. Monte Carlo simulations have been used to investigate tumor growth [6], tumor response to radiotherapy [7], and ligand-receptor interactions [8]. Unfortunately, all these works consider a very simple model that does not accurately account for all the factors that involve cellular decisions.

In this paper we improve on existing models, which do not account for the stochastic dynamics of the cellular system [3] and do not accurately reflect the space-time-varying, nonlinear and three-dimensional nature of the system [9]. We model and simulate the CSC system using principles of random-set-theory (RST). The components of the system include CSCs and daughter cells (DCs), the latter being descendants of CSCs. In addition, the mathematical model and algorithm mirror and include the cellular communication that leads to different responses of the system [9, 10]. The objective is to employ the developed model and simulation method as a "in-silico" laboratory. The improved understanding will be used for design of better experiments, which in turn will provide more valuable information for furthering the advancement of our knowledge about the studied biological phenomena.

2. THE BIOLOGICAL SYSTEM

The processes that lead to formation and growth of tumors are mainly based on two types of cells, i.e. CSCs and DCs (other cells that are not CSCs). The former have two important properties: they can divide symmetrically or asymmetrically and they are long living. The latter can only divide asymmetrically and will eventually die. In particular, we assume that a CSC, denoted by $C_1$, can participate in one of the following events:

\[ C_1 \xrightarrow{p_{01,t}} \emptyset \]

\[ C_1 \xrightarrow{p_{02,t}} 2C_1 \]

\[ C_1 \xrightarrow{p_{03,t}} C_1 + C_{2,1} \]

where $\emptyset$ symbolizes death, $C_{2,1}$ is a DC of first generation, and $p_{01,t}$, $p_{02,t}$, and $p_{03,t}$ are probabilities at a time instant $t$ that the respective events will occur between $t$ and $t + 1$. Also, with $p_{04,t}$ representing the probability that $C_1$ is in a quiescent state.

\[ p_{01,t} + p_{02,t} + p_{03,t} + p_{04,t} = 1 \]

A DC of the $j$-th generation $C_{2,j}$, where $1 \leq j < l$, takes participation in one of the following events:

\[ C_{2,j} \xrightarrow{p_{j1,t}} \emptyset \]

\[ C_{2,j} \xrightarrow{p_{j2,t}} 2C_{2,j+1} \]

where $C_{2,j+1}$ represents a cell from the $j + 1$ generation, and the probabilities satisfy

\[ p_{j1,t} + p_{j2,t} = 1 \]

where $p_{j2,t}$ is the probability that $C_{2,j}$ is in a quiescent state.

The cells of the last generation $C_{2,l}$ can either die within the interval from $t$ to $t + 1$, i.e.,

\[ C_{2,l} \xrightarrow{p_{l1,t}} \emptyset \]
or stay alive with probability \( p_{1 \cdot t} \), where
\[
p_{1 \cdot t} + p_{2 \cdot t} = 1. \tag{9}
\]
It is evident that the whole system is per se open, in that no upper bound can be determined for the overall number of cells. It is thus evident that the structure of such system requires resorting to simulation tools where random collections of objects with random states may be easily handled, a scenario quite reminiscent of the concept of random finite sets, which has recently received much attention in information fusion for distributed tracking [11] and communications [12].

3. MODELING THE SYSTEM USING RST

The system evolves from single cells to spheroids and aggregates (i.e., formations without CSCs). We consider that the formation of a spheroid is initiated by a CSC, i.e., one spheroid must contain at least one CSC. We also assume the hypothesis from the biologists that from one CSC we may potentially get one spheroid despite the number of symmetrical divisions that the original CSC will undergo. We use a compartmental approach, that is, we model the system as composed of two different compartments: the set of spheroids containing “useful” components (i.e., CSCs) and the set of junk aggregates with “useless” components (i.e., no CSCs),

\[
Z_t = X_t \cup J_t \tag{10}
\]

where \( Z_t \) is the total number of sets in the system at time \( t \), \( X_t \) is the set with the “useful” spheroids, i.e., those cell formations comprising a non-zero number of CSCs, while \( J_t \) is a set containing “useless” conglomerates, that is, cell aggregates that do not contain any CSC. The \( X_t \) compartment can be viewed as a disjoint union of singletons (spheroids)

\[
X_t = \bigcup_{k = 0}^{k_t} X_{k,t} \tag{11}
\]

where \( k_t = |X_t| \) denotes the cardinality of the set \( X_t \) and it varies with time (i.e., new spheroids can appear, or existing spheroids can disappear). The expression \( k_t = |X_t| = 0 \) denotes the absence of stem cells (i.e., the set of useful spheroids is empty). The singleton \( X_{k,t} \) is a descriptor of the \( k \)-th useful spheroid, and it can be expressed as:

\[
X_{k,t} = \begin{cases} 
  x_{k,1,t} \\
  x_{k,2,t} \\
  \vdots \\
  x_{k,t,t} \\
  \lambda_{k,t}
\end{cases} \tag{12}
\]

where \( x_{k,1,t} \) is the random variable modeling the number of CSCs in the \( k \)-th spheroid at time instant \( t \), \( x_{k,j,t}, j = 2, \ldots, n_{k_t} \) is the random variable modeling the number of DCs from the \( j \)-th generation in the \( k \)-th spheroid present at time instant \( t \), and \( \lambda_{k,t} \) represents the position of the group in the experiment domain. Note that \( X_{k,t} = \emptyset \) if \( x_{k,1,t} = 0 \), and that means that the \( k \)-th spheroid at that point disappears since no new stem cells can be generated. Therefore, \( X_{k,t} \) is a singleton-or-empty random set. At time \( t + 1 \), the singleton representing the \( k \)-th group has the following form

\[
X_{k,t+1} = \begin{cases} 
  x_{k,1,t+1} \\
  x_{k,2,t+1} \\
  \vdots \\
  x_{k,t,t+1} \\
  \lambda_{k,t+1}
\end{cases} \tag{13}
\]

We note that a spheroid can migrate to the junk set. Theoretically, a spheroid could lose all the CSCs which implies a switching to the set of junk aggregates, \( J_t \).

In order to derive the expressions that connect (12) and (13), we first note that we have two classes of cells with differentiated behaviors: the CSCs and the DCs. If at time instant \( t \) we have \( x_{k,1,t} \) CSCs, those cells can participate in one of the following actions: die (with probability \( p_{01,t} \)), divide symmetrically (with probability \( p_{02,t} \)), divide asymmetrically (with probability \( p_{03,t} \)), or remain idle (with probability \( p_{04,t} = 1 - p_{01,t} - p_{02,t} - p_{03,t} \)). Denoting by \( d_{k,1,t+1} \leq x_{k,1,t} \) the number of CSCs that will die in the interval \((t, t + 1)\) from spheroid \( X_{k,t} \), the number of possibly active CSCs (they may reproduce or do nothing) during \((t, t + 1)\) is \( x_{k,1,t} - d_{k,1,t+1} \). We represent as \( v_{k,1,t+1} \leq x_{k,1,t} - d_{k,1,t+1} \) the number of CSCs that will reproduce symmetrically in the interval \((t, t + 1)\) from spheroid \( X_{k,t} \), i.e., those CSCs that will produce new CSCs and as \( k_{1,t+1} \leq x_{k,1,t} - d_{k,1,t+1} - v_{k,1,t+1} \) the number of CSCs that will be in a quiescent state during the next period. We can then write that the number of CSCs in the system evolves with time according to the expression

\[
x_{k,1,t+1} = x_{k,1,t} - d_{k,1,t+1} + v_{k,1,t+1}. \tag{14}
\]

The previous equation means that the new number of CSCs is the amount of CSCs in the previous time instant minus the ones that died plus the new ones that were generated through symmetrical division.

We focus next on the first generation of DCs. These cells can only die, reproduce symmetrically towards the next generation (i.e., \( C_{2,1} p_{22,t} \)) or remain in a quiescent mode. The number of cells in this group at time instant \( t + 1 \) will be calculated as the number of cells of this same generation from the previous time instant that stayed idle (i.e., the cells that neither die nor divide towards the next generation) plus the cells that came from asymmetrical division of the CSCs, i.e.,

\[
x_{k,2,t+1} = (x_{k,1,t} - d_{k,1,t+1} - v_{k,1,t+1} - k_{1,t+1}) + k_{2,t+1}. \tag{15}
\]

For the next generations we can write a common update equation since after the first generation of DCs only death, idleness or symmetrical division towards future generation is possible. This can be formulated as follows:

\[
x_{k,j,t+1} = 2(x_{k,j-1,t} - d_{k,j-1,t+1} - v_{k,j-1,t+1} + k_{j-1,t+1}) + k_{j,t+1} \tag{16}
\]

The overall evolution of the system will be given by

\[
X_{k,t+1} = \begin{pmatrix}
  x_{k,1,t+1} \\
  x_{k,2,t+1} \\
  \vdots \\
  x_{k,t,t+1} \\
  \lambda_{k,t+1}
\end{pmatrix}
\]
where \( f(\cdot) \) is a function that determines the evolution of the location of the compound depending on the previous location and the amount of cells in the compound \( x_{k,t+1} = [x_{k,1,t+1}, \ldots, x_{k,d,t+1}] \).

The previous equations show that \( X_t \) is actually a Markov finite random set, and lends itself to being computer-generated. As an example, the model (which, however, neglects for simplicity reason a number of phenomena that can be observed) leads to an algorithm wherein \( X_{k,t+1} \) is generated from \( X_k \), as follows:

**CSC generation**

1. Generate \( d_{k,1,t+1} \sim Bi(p_{01,t}, x_{k,1,t})^2 \)
   (number of CSCs that die);  
2. Generate \( v_{k,1,t+1} \sim Bi\left(\frac{p_{02,t}}{p_{01,t} + p_{02,t} + p_{04,t}}, x_{k,1,t} - d_{k,1,t+1}\right) \)
   (number of CSCs that divide symmetrically);  
3. Update \( x_{k,1,t+1} = x_{k,1,t} - d_{k,1,t+1} + v_{k,1,t+1} \)
   (number of CSCs at \( t + 1 \));  
4. Generate \( i_{k,1,t+1} \sim Bi\left(\frac{p_{04,t}}{p_{03,t} + p_{04,t}}, x_{k,1,t} - v_{k,1,t+1}\right) \)
   (number of 1st generation that die – needed to obtain next generation).

**First generation of DCs**

1. Generate \( i_{k,2,t+1} \sim Bi(p_{33,t}, x_{k,2,t}) \)
   (number of 1st generation that are idle);  
2. \( x_{k,2,t+1} = (x_{k,1,t} - d_{k,1,t+1} - v_{k,1,t+1} - i_{k,1,t+1}) + i_{k,2,t+1} \)
   (number of 1st generation at \( t + 1 \));  
3. Generate \( d_{k,2,t+1} \sim Bi\left(\frac{p_{32,t}}{p_{31,t} + p_{32,t} + p_{34,t}}, x_{k,2,t} - i_{k,2,t+1}\right) \)
   (number of 1st generation that die – needed to obtain next generation).

\( j \)-th generation of DCs, \( j = 3, 4, \ldots, n_{k_i} \)

1. Generate \( i_{k,j,t+1} \sim Bi(p_{33,t}, x_{k,j,t}) \)
   (number of cells that are idle);  
2. \( x_{k,j,t+1} = 2(x_{k,j-1,t} - d_{k,j-1,t+1} - i_{k,j-1,t+1}) + i_{k,j,t+1} \)
   (number of cells of \( j \)-th generation at \( t + 1 \));  
3. Generate \( d_{k,j,t+1} \sim Bi\left(\frac{p_{32,t}}{p_{31,t} + p_{32,t} + p_{34,t}}, x_{k,j,t} - i_{k,j,t+1}\right) \)
   (number of 1st generation that die – needed to obtain generation \( j + 1 \)).

### 4. SIMULATION RESULTS

Section 3 described the post-migration cell-fates, but in order to accurately mimic in vitro and in vivo conditions, we included communication and migration capabilities [9, 13, 14]. The main goal was to track every evolution of a cancer spheroid starting from “seeds” of CSCs. Our simulation had two levels of organization: (a) individual cells, and (b) groups of cells (spheroids). Modeling was done in a 3D lattice of side \( n_{grid} \) points. Time units were days. The cellular automata rules are described as follows.

\[\text{Bi}(p, n)\] represents a binomial distribution with \( n \) trials and success probability \( p \).

### 4.1. Basics

**Cell growth:** Each cell occupied one grid point at most. Proliferation capacity (maximum number of generations \( l \)) was finite for DCs and infinite for CSCs. Also, CSCs were assumed to be immortal \((p_{01,t} = 0)\).

**Cell-to-cell signaling:** Based on the inter-cellular signaling method implemented in [9], we assumed that at each time step each cell could emit a chemoattractant, whose concentration at some grid point changed inversely with the distance from that point from the cell’s center. The surface chemoattractant concentration for each cell was fixed at \( \alpha \). Each cell could detect (via surface sensors) chemical gradients resulting from the total chemoattractant emitted by all other cells.

**Cell migration:** At each time step cells could migrate either in the direction of a chemoattractant gradient, with the number of steps depending on the chemo-gradient force, or randomly, i.e., independently from the chemical-gradient. For each cell, the ability to migrate was limited by the number of neighboring cells \( n_a \). If a cell was part of a large aggregate (“spheroid”) then the entire spheroid would move in the average chemoattractant direction. However, newly generated cells had a small probability \((p_{nd,t})\) of separating from the aggregate altogether instead of moving with it \((p_{01,t})\). For each cell, the ratio \( p_{nd,t}/p_{01,t} \) was directly proportional to the number of attachments \( n_a \) of that cell, i.e., cells with higher number of attachments were less likely to separate from the spheroid. Thus \( p_{nd,t}/p_{01,t} \) varied from cell to cell and at each time step. If such cells chose to detach, then they would randomly migrate for some steps in a manner similar to that in [13].

**Spheroid formation:** At each time step, each cell checked if there were active cells in its immediate neighborhood (colliding cells). In case of existence of colliding cells, the current cell and all colliding cells were updated as belonging to this spheroid. On the other hand, if all the colliding cells happened to be originally unattached, then a new spheroid was created. The current cell and all colliding cells were updated as belonging to the new spheroid.

### 4.2. Results

We considered an initial scenario with two CSCs at the positions \((25, 25, 25)\) and \((60, 60, 60)\) in a 3D lattice with width \( n_{grid} = 100 \) and \( l = 6 \). Post-migration probabilities were
\{p_{01,t} = 0.0, p_{02,t} = 0.05, p_{03,t} = 0.65, p_{04,t} = 0.3\} \text{ and } \{p_{11,t} = 0.1, p_{12,t} = 0.6, p_{13,t} = 0.3\}.

In order to show how spheroid evolution varied with changes in parameters, we modified the model parameters related to cell-signaling (\(\alpha\)) and detachment (the ratio \(p_{ni,t}/p_{nd,t}\)). We defined three scenarios: a control case where we kept \(\alpha = 9\) and \(p_{ni,t}/p_{nd,t} = n_\alpha/0.3\); a low detachment setup, with \(p_{ni,t}/p_{nd,t} = n_\alpha/0.03\); and a higher cell-to-cell talk setting, where \(\alpha = 18\).

Figure 1 shows a comparison of the total cell populations for these three scenarios. Figure 2 shows snapshots of the 3D lattice for the low-detachment simulation at different times.

As expected, the low-detachment scenario leads to lower cell populations than in the control case, since more cells “decide” to crowd around proliferating cells and prevent cell division (due to lack of space). This leads to more cells staying quiescent or dying instead of proliferating and the result is a lower cell count.

The high-cell-talk scenario results in cell populations midway between the control and the low-detachment cases. Cells can detect higher levels of chemoattractant, which means that they tend to congregate at higher rates, leading to more limitations on proliferation space. However, because of the higher migration rates, many cells also tend to find more free space around them, thus leading to higher cell-division rates. Therefore, the two effects balance out to some extent and cell populations are not unduly affected by this parameter change.

5. CONCLUSIONS

The present contribution is aimed at shedding some light on the possible applications on random-set-theory for analysis and simulation of complex systems, epitomized here by the evolution process of cancer cells starting from cancer stem cells. Our final goal is to exploit the properties of disjoint unions of random sets and the Bayes recursions on random sets to obtain the posterior distribution of \(X_t\).

It is important to remark that the flexibility of random-set-theory allows the inclusion of different effects (such as, e.g., migration effects from different spheroids or aggregates, un-determinacy of the next spheroid or aggregate location from the previous one and so on) in a very easy and user-friendly way.

6. REFERENCES


