Two Discrete Stochastic Cellular Automata Models of Cancer Stem Cell Proliferation

Andrew Q. Ninh

Abstract—**Cancer stem cells (CSCs) are cancer cells that exhibit stem cell-like properties. They are immune to standard chemotherapy and are often implicated for relapse and metastasis. Modeling of CSC-caused relapse is difficult as organisms tend to die before the relapse can be studied, and thus in silico models are ideal but are in development. Two kinds of CSC-induced tumor growth were modeled mathematically and visually using the mass-action and spatial models. Mathematical models of population growth and a better understanding of cancer stem cell population dynamics and neural networks can be achieved by applying discrete stochastic models, automata theory, and cellular automaton. Due to its wide range of possibilities, cellular automata theory opens up new field of mathematical applications in cancer modeling and providing a bridge between bioinformatics and individualized cancer modeling.**

Index Terms—**Cancer stem cell, cellular automata, tumorigenesis, stochastic, gompertzian growth.**

I. INTRODUCTION

Cancer Stem cells (abbreviated as CSCs) are cancerous cells that exhibit properties similar to normal stem cells. This means that CSCs are multipotent and are able to differentiate into cancer cells and can undergo self-renewal. CSCs essentially are tumorigenic, meaning they are capable of creating tumors, a quality other cancerous cells do not possess. Another quality of CSCs is immortality; whereas other cells have a limited number of times they can divide (Hayflick limit) CSCs are able to divide indefinitely [1].

One of the main problems of CSCs in cancer treatment is that they are generally unaffected by chemotherapy used to kill most differentiated cancer cells (which make up most of the tumor). CSCs generally make up about 1-3% of a tumor [2]. Thus, following chemotherapy, CSCs left behind would be able to replenish a tumor and cause a relapse of the cancer [3]. In addition, tumor modeling and understanding relapse due to CSCs are currently ill understood because most organisms with relapse cancers *in vitro* die before they can be further studied.

If unheeded, CSCs can, in theory, cause continual relapses of a tumor, and are capable of metastasis – the migration cancerous cells including CSCs to other organs or tissues in the body to create new tumors (carcinogenesis). By applying discrete stochastic models, automata theory, and cellular automaton programming to create more accurate models of population growth and a better understanding of population dynamics and neural networks.

Due to discrepancies in the presence of CSCs, differences in various types of cancers, and variations in individuals, there are few general mathematical models that describe CSC-induced tumor proliferation. However, using compartmental methods and predictive mathematical models as well as cellular automaton, CSC-induced tumorigenesis is possible [4]-[6].

II. MATERIALS AND METHODS

The Java programs were written on the BlueJ integrated development system (IDE) and resultant data points were graphed on Mathematica. The cellular automaton models were plotted using Mathematica's and ArrayPlot functions for analysis.

Two programs were written on BlueJ to simulate the mass action and spatial cell growth models. Because the Random package provided by Oracle is not as cryptographically secure, the SecureRandom package was used instead when determining stochastic processes.

Because the programs need to account for both stem and differentiated stem cell, each automaton has three states: empty (value of 0), stem (value of 1), progenitor (value of 1.5), and differentiated (value of 2) (see Fig. 1). Cancer stem cells primarily undergo self-renewal (mitosis) or divide into a progenitor cell which still has CSC qualities but is slightly more differentiated. This adds a value of 0.5 to the CSC value. These progenitor cells can divide or differentiate further by adding 0.5 to create a differentiated cell. Differentiated cells can continue to proliferate or undergo apoptosis as a method of cellular control and balance. However, due to the cancerous nature of cancer cells, these cells divide or differentiate indefinitely and have their proliferation controls (such as the Hayflick limit) and apoptosis processes inhibited.
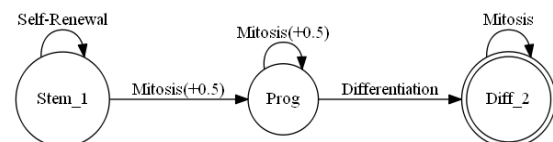

Fig. 1. Schematic diagram of the cancer stem cell induced tumor growth process. The circle (stem_1) represents a stem cell, the smaller circle (prog) represents a progenitor cell, and the double circle (diff_2) represents a differentiated cell.

Because the finite state machines studied are CSCs, these cell population control methods are ignored and cells are left to proliferate or differentiate at constant rates defined by stochastic processes.

A. Mass-Action Model

The mass-action model is a cancer growth model proposed

by Dr. Ting-Chao Chou in 2011 [7]. The mass-action model is based off of the chemical and epidemiological laws of mass-action in which any individual in a homogenous mixture has an equal probability of interacting with any other individual in the grid (see Fig. 2).

Essentially, the mass-action model allows for a randomly chosen cell to place its resultant offspring from self-renewal or differentiation on another randomly chosen location on the grid. These processes are determined stochastically.
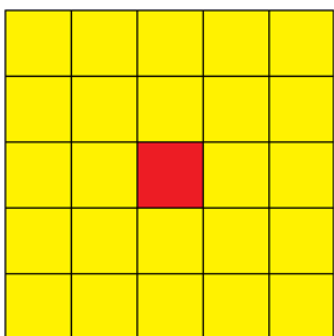

Fig. 2. The red cell is a randomly chosen cell and the yellow cells represent all locations the cell can interact with.

B. Spatial Model

The spatial model is another commonly used model in cancer modeling [8], [9]. The spatial model differs from the mass-action model in that a randomly chosen cell can only place its offspring on a random location in its neighborhood. The selected neighborhood for a cell in the spatial model is a Moore neighborhood with range $r=1$; this is also known as a Chebyshev distance of 1. The spatial model functions similarly to the mass-action model but a randomly chosen cell is only able to interact with randomly chosen cells in its neighborhood.
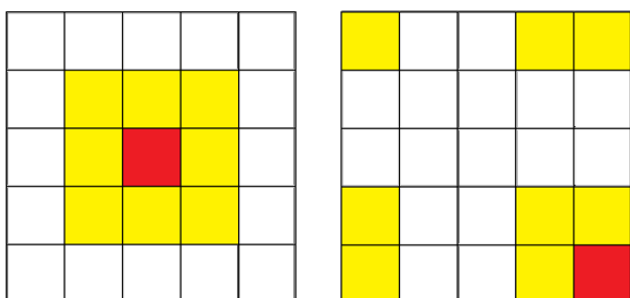

Fig. 3. The red cell is a randomly chosen cell and the yellow cells represent the red cell's Moore neighborhood with range r=1. The diagram to the left demonstrates the torus-nature of the spatial model.

The Moore neighborhood contains the cell and the surrounding cells with Chebyshev distance 1 (see Fig. 3). The area of the Moore neighborhood can be described by the area equation $(2r+1)^2$. These numbers, such as 9, 25, 49, etc., are odd squares. In the spatial model, with a neighborhood with range $r=1$, the total area is 9. In addition, the spatial model is able to wrap over and around itself to create a torus. With no axis of rotation, the torus mathematically degenerates to represent a sphere (Fig. 3).

The Moore neighborhood is defined by the following set:

$$N^{M}_{(x_o,y_o)} = \left\{(x,\,y) : |x - x_0| \le r, |y - y_0| \le r\right\} \quad (1)$$

III. RESULTS

After hundreds of plots describing 10,000 generations of cell processes were averaged using the law of large numbers and graphed on Mathematica, it was noted that the differentiated cancer cells (which naturally make up the majority of the tumor) had the highest population and exhibited a Gompertzian growth model. The progenitor cell population made up about 5% of the total population and exhibited a far more gradual Gompertz function (varying r parameter). The CSC population, as expected, remained at about 3% of the total population and followed the Gompertz function as well.

The Gompertz function is defined by the equation

$$N(t) \;=\; K\,\exp\left(\log\!\left(\frac{N_0}{K}\right)\exp\!\left(-\alpha t\right)\right) \quad (2)$$

in which $N(t)$ is the function describing the size of the tumor at time t, $N(0)$ would thus be the initial size of the tumor, K is the carrying capacity, and α is the proliferation constant of the cancerous cells.

The Gompertz function is characterized by a more gradual approach from the inflection point to carrying capacity (future asymptote) than from initial growth to the inflection point.
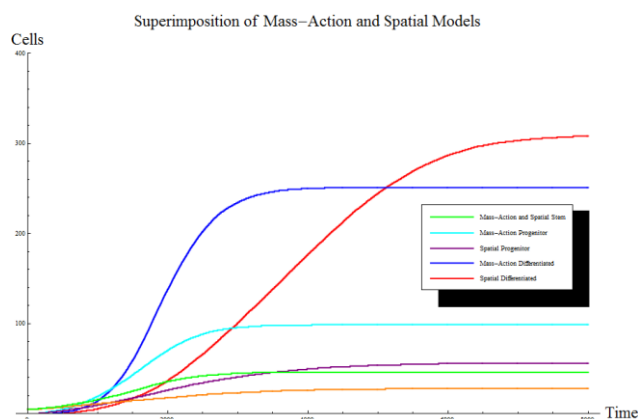

Fig. 4. A superimposition of the mass-action and spatial CSC, progenitor, and differentiated cancer cell plots.
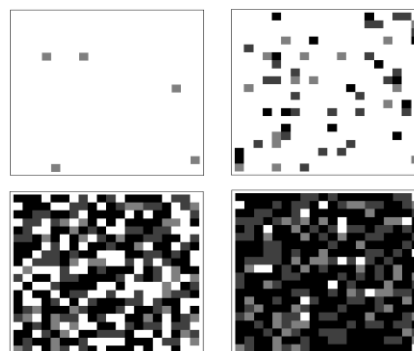

Fig. 5. CA model describing the progression of a tumor starting with 5 CSCs in the mass-action model through stochastic processes.

Fig. 4 represents the graphical plots of the mass-action and spatial CSC, progenitor, and differentiated cancer cell populations.

The cellular automaton arrays were also printed out over various intervals of generations (after every 500 generations)

and plotted on Mathematica to visualize and compare the growth of CSC-induced tumors in the mass-action (Fig. 5) and spatial (Fig. 6) models.
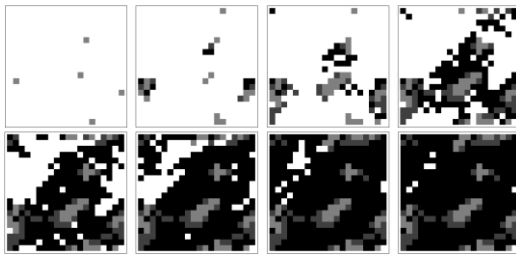

Fig. 6. CA model describing the progression of a tumor starting with 5 CSCs in the spatial model through stochastic processes.

IV. Discussion

Ultimately, the mass-action and spatial models provide relatively similar results in that differentiated cells (similar to clinical scenarios) exhibited the largest population and comprised of the majority of the tumor.

It was noticed that although the spatial model reached equilibrium at a slower rate, it reached a higher carrying capacity. This is because the Moore neighborhood confines the CSC and progenitor growth by forcing these cancer stem cells to create a cell niche and thus any cell from these cell niches would have to be differentiated cancer cells. Thus, the majority of these cells exist in cell niches of CSCs surrounded by progenitor cells.

Though the spatial model provides a closer Gompertzian growth model and cellular automaton imaging to lifelike scenarios, this does not completely discredit the mass-action model as it may still be used as a base comparison for variations on the spatial model (such as applying a von Neumann neighborhood or using an extended Moore neighborhood). In respect, these variations on the spatial model may better fit CSC-induced tumor growths of various cancers. However, further clinical trials may have to be done.

CSCs exhibit a classic Gompertz tumor growth model with a much more gradual slope to carrying-capacity than the Gompertz growth models of differentiated and progenitor populations. Though this does not make much difference in the outcome, this provides interesting future research in CSC size relative to tumor size or cancer type. In addition, other growth models such as the Janoschek, Weibull, and Richards curves may be fitted against these stochastic plots for comparison.

As shown in the Mathematica cellular automaton model of the spatial cancer cell populations, the CSCs generated progenitor cells surrounding them which then divided into more progenitor cells or eventually differentiated into differentiated cancer cells. This models the lifelike case of cancer stem cell niches which remain after chemotherapy and differentiated cancer cells become necrotic [10].

The Mathematica cellular automaton arrays of the mass-action cancer cell populations show randomness. However, this randomness might disappear as certain conserved structures may be found or lists of Turing machines at different generations may be compared to see whether these models follow any elementary cellular automaton "rules" proposed by Stephen Wolfram.

V. Conclusion

The mass-action and spatial models of microbiological population dynamics were applied in cancer stem cell modeling. The two cell compartmental models of cell population dynamics demonstrate that cancer progenitor cell and differentiated cancer populations exhibit a Gompertzian growth model which is typical in tumor modeling. Interestingly, cancer stem cell populations remained at life-like percentages and also resembled the Gompertz growth model. It was also observed that the total progenitor and cancer stem cells reached a dynamic equilibrium at which cells could no longer proliferate but progenitor cells continued to differentiate. Thus, the differentiated cell populations exhibited a double Gompertz function whereas the stem-progenitor populations returned to the 1-3% CSC population in a tumor. It is important to note that the current model includes all progenitor cells as cancer stem cells rather than as their individual graphs.

The mass action and spatial models can be expanded to become three dimensional grids and thus neighborhoods would become three dimensional as well. This can be accomplished as the current Java code is written as a Turing machine, meaning information is stored on a one dimensional array and one cell on the array is able to interact with and change other cells. Thus in languages with 2D arrays such as Java and Python, a 3D model is possible with arrays of Turing machines. However, when coding in the C++ language, it is possible to use 3D arrays, so a Turing machine in this case would not be necessary.

The 3D Moore neighborhood takes the shape of a cube with 27 locations including the chosen cell and can be defined by the set

$$N^{M}_{(x_o, y_o, z_o)} = \left\{(x, y) : |x - x_0| \leq r, |y - y_0| \leq r, |z - z_0| \leq r\right\} \tag{3}$$

The model can be transposed into a von Neumann neighborhood with range $r=1$ and, more interestingly, range $r=2$ (Fig. 7). Because the range of a von Neumann neighborhood is diagonal, the distance in which a cell can interact is known as the Manhattan distance. The von Neumann neighborhood is defined by the equation

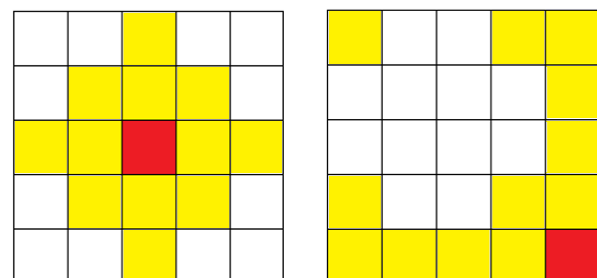$$N^{V}_{(x_o, y_o)} = \left\{(x, y) : |x - x_0| + |y - y_0| \leq r\right\} \tag{4}$$

Fig. 7. The red cell is a randomly chosen cell and the yellow cells represent the red cell's von Neumann neighborhood with range $r=2$. The diagram to the left demonstrates the torus-nature of the spatial model.

Cell motility may be applied to these models as the process is very similar to the current cellular processes except a cell

removes itself from its current location and places itself on another location in a defined neighborhood. Study in cell motility would help in the moving of cancer cells that is implicated in causing metastasis.

Using cellular automaton arrays, various stochastic trials may be run and compared with one another to see if there are any conserved structures (similar to those in Conway's Game of Life) may be found in the models.

As aforementioned, each individual Turing machine can be printed and graphed by generations using Mathematica's ArrayPlot function and compared with Stephen Wolfram's proposed rules and codes of elementary cellular automaton. Currently, it is expected that due to the stochastic nature of the current models, these cellular automaton models may be classified as being Class 3 automata in that they will appear to be random states.

Bioinformatics can potentially be bridged with CA modeling as each gene can be considered an individual rule for every cell on the grid. Genes may be turned into boolean structures which would then be applied as a rule for each individual cell [11], [12]. This should create more individualized cell models and be directly applicable in clinical situations by creating individualized tumor models. This could also potentially help with tumor prediction in early stages of detection as well as predicting metastasis.

REFERENCES

[1] H. Sugihara, "Cell kinetic and genetic lineage analyses of cancer development," Thesis paper. Dept. of Pathology (Division of Molecular and Diagnostic Pathology), Shiga University of Medical Science, Shiga Prefecture, Japan, n.d.

[2] University of Michigan Stem Cell Research. (2011). Cancer Stem Cell Research Introduction Defining Stem Cells. [Online]. Available: http://www.cancer.med.umich.edu/research/stemcells_two.shtml.

[3] G. Vogel. (2012). Cancer Stem Cells can Fuel Tumor Growth. *ScienceNOW*. [Online]. Available: http://news.sciencemag.org/sciencenow/2012/08/cancer-stem-cells-can-fuel-tumor.html.

[4] R. Ganguly and I. K. Puri, "Mathematical model for the cancer stem cell hypothesis," *Cell Proliferation*, vol. 39, no. 1, pp. 3-14, 2006.

[5] R. Ganguly and I. K. Puri, "Mathematical model for chemotherapeutic drug efficacy in arresting tumour growth based on the cancer stem cell hypothesis," *Cell Proliferation*, vol. 40, no. 3, pp. 338-354, 2007.

[6] D. D. Phan and J. S. Lowengrub, "A discrete cellular automaton model demonstrates cell motility increases fitness in solid tumors," *The UCI Undergraduate Research Journal*, pp. 55-66, 2010.

[7] T. C. Chou, "The mass-action law based algorithm for cost-effective approach for cancer drug discovery and development," *American Journal of Cancer Research*, vol. 1, no. 7, pp. 925-954, 2011.

[8] E. A. Reis, L. B. L. Santos, and S. T. R. Pinho. (2008). A cellular automata model for avascular solid tumor growth under the effect of therapy. [Online]. Available: http://arxiv.org/pdf/0806. 1063v1.pdf

[9] K. A. Rejniak and A. R. A. Anderson, "Hybrid Models of Tumor Growth," *NIH Public Access*, vol. 3, no. 1, pp. 115-125, 2012.

[10] T. Borovski, F. de S. e Mello, L. Vermeulen, and J. P. Medema, "Cancer stem cell niche: the place to be," *Cancer Research*, vol. 71, pp. 634, 2011.

[11] B. V. Vasic, "An information theoretic approach to constructing robust boolean gene regulatory networks," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 9, no. 1, pp. 52-65, 2012.

[12] R. Sanchez, E. Morgado, and R. Grau, "A genetic code Boolean structure. I. The meaning of Boolean deductions," *Bulletin of Mathematical Biology*, vol. 67, no. 1, pp. 1-14, 2005.

Andrew Q. Ninh was born in Newport Beach, California on September 9, 1995. Andrew Ninh is currently enrolled as an undergraduate student in California State University, Fullerton located in Fullerton, California.

He has done research at the University of California, Irvine located in Irvine, California on cellular automaton modeling and population dynamics of embryonic stem cells. He has presented his research at the 14th and 15th annual Biomedical Engineering Bio-Tech competitions sponsored by the IEEE Los Angeles BME society and currently has a few publications including a data mining paper on the distribution and density of palindromic sequences in the SMAD4 gene and a published abstract comparing two cellular automata neighborhoods in modeling embryonic stem cell populations. His research interests include bioinformatics, comparative genomics, mathematical biology, and data mining.

Mr. Ninh is a member of the Institute of Electrical Engineers' Engineering in Medicine and Biology society (IEEE EMBS) and the National Eagle Scout Association.