Modeling Multicellular and Tumorous Existence with Genetic Cellular Automata

Armand Bankhead III[1], Nancy Magnuson[2] and Robert B. Heckendorn[1]
[1]Bioinformatics and Computational Biology; University of Idaho, Moscow, ID 38843
[2]School of Molecular Biosciences; Washington State University, Pullman, Washington 99164
bank2192@uidaho.edu, heckendo@uidaho.edu

Abstract

We model a population of cells using cellular automata with genetically-based rules. As in actual multicellular systems, each cell's state is based on environment and genetics. With imperfect reproduction and the accumulation of random mutations, tumorous behavior naturally emerges in a stochastic manner. We validate our model by reproducing results used to confirm other published models. We also demonstrate that our model exhibits both homeostatic and tumorous behavior using metrics based on clinical biopsy diagnosic techniques.

Introduction

Multicellular organisms are composed of cells that co-exist despite resource limitations (ie. glucose, oxygen). Unlike unicellular organisms, such as bacteria or eubacteria, multicellular organisms possess proto-oncogenes that activate or inhibit reproduction to benefit the organism (Heath, 2001). When these proto-oncogenes are mutated, cells reproduce uncontrollably and tumorous behavior emerges to benefit individual cells and defy the organism.

As tissue cells are damaged or intentionally sloughed, reproduction is required (Tannock and Hill, 1998). Cellular reproduction within the tissue of an organism is somatic and semi-conservative (Fairbanks and Andersen, 1999). This process is imperfect; both parent and offspring cell genomes experience random mutations. As the population is replenished, mutations may accumulate to cause genetic dysfunction. Mutational damage to genomic areas controlling reproduction may result in tumorous behavior.

We define multicelluar existence as a population of cells that co-exist using reproductive control. We define tumurous existence as a population of cells that reproduce to maximize population size and resource consumption. In this paper we present a genetic cellular automata abstraction capable of modeling multicellular and tumorous existence.

Cellular automata (CA) are a logical choice for modeling both population types. Just as a biological cell reacts to a local environment using genetic expression, cellular automata react to a local neighborhood using a set of rules. Because multicellular and tumorous existence are the direct result of genetic functionality, we believe these CA rules should reflect the genetic mechanisms involved.

The next section discusses other CA models used to describe tumor growth. The following section introduces genetic CA (GCA) and highlights important features of the design. Next, using results from other published CA models, we show that GCA can evolve similar tumorous behavior. Then we use biopsy-based metrics to classify multicellular and tumorous existence to illustrate the validity of our model. Finally, we discuss results and future work.

Previous Work

CA's have been applied to model avascular tumorous behavior (Moreira and Deutsch, 2002).

A model produced by Qi *et al.* (1993) uses a two dimensional CA to illustrate critical aspects of Gompertzian population growth curves of cancerous cells. Probability-based rules for proliferation and death are parameterized by experimental data. Nutrient resources are dependent on the size of the population–as the population increases less nutrients are available. Simulated population growth curves are then compared to Gompertzian growth curves for validation.

Work by Kansal *et al.* (2000) uses a three dimensional CA to model brain tumor growth. Rules are parameterized by experimental data such that layers of proliferative, non-proliferative, or necrotic cells form based on location relative to the center of the tumor. These layers result from pressure and nutrient constraints on the tumor growth. Rules are deterministic and based on a single set of experimental data. The model is validated via experimentally derived *in vitro* growth curves and Gompertzian growth curves. *Of course*, the layering effects defined by the model's rules are exhibited.

More recently, a two dimensional CA model presented by Dormann and Deutsch (2002) uses a hybrid lattice-gas approach to account for varying channels for migration and chemical diffusion. Rules are also probability-based and derived from experimentally derived differential equations. Cellular rules include behaviors such as mitosis (reproduction), apoptosis (suicide) and necrosis (death). Nutrients are

explicitly accounted for using a lattice-gas diffusion process, although the diffusion process is assisted by expanding the local neighbordhood of each cell to three layers. The proliferative, non-proliferative, and necrotic layering effect is also used to validate this model.

Although these models are insightful to specific data sets, they do not shed light on the underlying general cause of tumorigenesis. All of the models mentioned above simulate a contrived starting population of tumor cells. Tumor development is not contrived, but is the result of genetic dysfunction (Tannock and Hill, 1998). Normal multicellular populations evolve this behavior contrary to the organism. A more complete tumorigenesis model should encompass both multicellular and tumorous existence.

Genetic Cellular Automata Design and Implementation

Cellular Automata Framework

Our model is similar to other CA tumor models in several regards. We model space as a 100x100 lattice of boxes; cell enitities may exist within these boxes. Each cell has a discrete state defined by its binary genome; cell rules are probabilistically determined using this binary genome. Using these rules, cells react to a local Moore neighborhood of eight boxes. Also, time is observed as discrete cycles and all cells on the lattice are updated synchronously.

GCA are atypical to CA tumor models. Nutrients diffuse across the lattice using a stochastic averaging technique. Signals from a cell's environment and interacting genes are interpreted by genes that control the rules of the cell. As a result, biological behaviors, such as reproduction and apoptosis, are controlled by genetic rules. Upon reproduction, cells copy their genome with some mutation rate; the resulting parent and offspring are dissimilar. More than one cell may exist within each box of the lattice. These atypical properties are described below in further detail.

Nutrient Diffusion

All possible forms of eukaryotic cell diet (i.e. glucose, oxygen, etc.) are reduced to one generic nutrient. In biological tissue populations, nutrients are typically delivered by nearby blood vessels and diffuse through adjacent cells. We represent these nutrient sources on the outer edge of the lattice. Nutrient source boxes produce enough nutrients to sustain a lattice population with an average density of 5 cells per box.

Diffusion is emulated by stochastically updating Moore neighborhoods throughout the lattice. The update involves feeding cells within the box and then mixing the nutrient concentration evenly in the neighborhood of boxes. The rate of diffusion may be adjusted by varying the number of updates over the lattice.

This nutrient diffusion process allows our model to simulate nutrient deficiencies and circumvent using a static diffusion representation specific to a particular experimental data set.

Genes Produce Genetic Signals

Cell rules are based on genetic signals. Similar to biological genes, genetic functionality is dependent on the fidelity of a gene; a mutated gene is less likely to be expressed properly. Each gene is implemented as a 64-digit binary string divided into three biologically-based (Fairbanks and Andersen, 1999) regions shown in Figure 1.

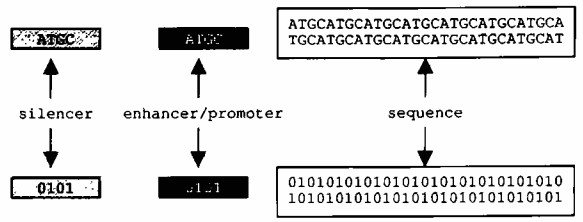

Figure 1: Schematic of a bit string (bottom) used to represent biological genes (top). non-mutated bits = 0; mutated bits = 1.

- Silencer: Negative control. When damaged the gene is over-expressed.
- Enhancer-Promoter: Positive control. When damaged the gene is under-expressed.
- Sequence: Coding region. When damaged the gene product expressed does not function.

Eq. 1 shows genetic functionality (F_g) is related to the number of non-mutated bits of the gene.

$$F_g = \frac{1}{F_s} * F_{ep} * F_s \tag{1}$$

The functionality of the silencer (F_s), enhancer/promoter (F_{ep}), and sequence (F_s) are the percentages of non-mutated bits within each region.

Finally, the genetic signal or probability (P_g) of a gene's proper expression is

$$P_g = Sig_g(c_g, s_g, F_g) = \frac{1}{1 + exp(-s_g(F_g - c_g))}, \tag{2}$$

where Sig_g is a sigmoid function parameterized by slope (s_g) and centering (c_g) for realistic behavior.

Genetic Networks

The actual regulatory genetic networks that control eukaryotic cell behavior are currently incomplete for many multicellular organisms. Integrating the complete genetic networks involved would also be computationally unfeasible. Abstractions are necessary.

Using a coarse-grained approach we focus on basic environment-to-gene and gene-to-gene interactions by collapsing major genetic pathways into genetic classes. For simplicity, we refer to these genetic classes as genes and represent these genes using the bit string design presented in the previous section. Figure 2 shows the collapsed genetic networks used for each cell while Table 1 summarizes these relationships.

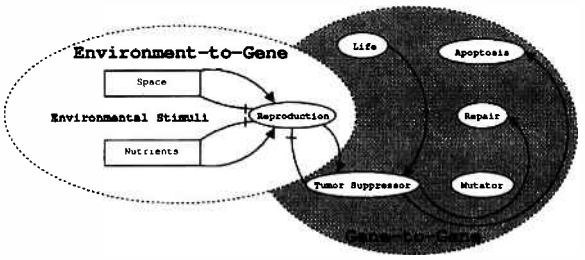

Figure 2: Collapsed genetic network indicating relationships between environment and genes. Environment-to-Gene interactions are shown to the left; Gene-to-Gene interactions are shown to the right. Environmental signals are labeled in rectangles; genetic signals are labeled in ovals. Arrows indicate gene activation; crossed connections indicate inhibition.

		Input Signals				
		Envioronment		*Gene*		
		Nutrients	Space	Life	Repro.	Tumor Sup.
Genes	Life					
	Reproction	+/-	+/-			-
	Tumor Suppressor			+	+	
	Apoptosis					+
	Repair					+
	Mutator					

Table 1: Genetic Network Summary. '+' means positive control (activation); '-' means negative control (inhibition).

Cellular Automata Rules

Each cell's rules are dictated by the genetic relationships described in the previous section. Rule behavior is probabilistic and based on the summation of environmental and genetic input signals of relevant genes shown in Table 2.

For example, a cell's overall mutation rate is based on a mutation rate (10^{-6}) that does not cause premature mutational meltdown. Repair signals lower the mutation rate (Fairbanks and Andersen, 1999) whereas mutator signals increase the mutation rate (Kuby, 1997). If the repair gene becomes damaged the mutation rate may increase whereas if the mutator gene becomes damaged the mutation rate may decrease.

With rules based on multicellular genetics, we capture general multicellular behavior. Instead of cells reproducing to fill the lattice to a maximum density (10 cells per box),

		Genes					
		Life	Repro.	Tumor Sup.	Apop.	Repair	Mutator
Rules	Reproduction		X				
	Mutation Rate					X	X
	Genetic Death	X					
	Suicide				X		

Table 2: Genetic Rules. 'X' indicates a direct relationship between cell rules and genes.

the cells are genetically designed to reproduce to an acceptable density (3 cells per box). As individuals accumulate mutations, genetic rules may ignore proper environmental signalling–just as genetic dysfunction leads to tumorous existence in organisms.

Algorithmic Description

Initially a normal cell (with no mutations) is placed randomly on the lattice. The cell reproduces itself and clonal reproduction continues until an appropriate density of cells is present within each box. Cell densities are kept relatively homogenous as a natural consequence of the GCA rules. After the lattice has reached population homeostasis, random cell removal (.0005 chance a cell will be removed) encourages reproduction.

A Comparison to Previous Work

Several of the CA models mentioned above verify their work using comparisons to experimentally derived features. Some of these tumorous features are discussed and evidence is presented to show that our model evolves similar behavior without directly designing the model to do so.

Gompertzian Growth

Typically unicellular populations grow exponentially in the presence of excess nutrients. As mentioned above, exponential growth is prevented in multicellular organisms through proto-oncogenes; cells are genetically designed to reproduce when appropriate to the organism. A population of tumorous cells typically displays exponential growth until space or nutrient limitations are met (Tannock and Hill, 1998).

Models presented by Qi *et al.* and Kansal use Gompertzian growth to illustrate validity. Gompertzian growth curves are used to describe growing populations in the presence of resource limitations. Typical Gompertzian curves are initially exponential and dampen to a stationary phase with little growth (Qi et al., 1993) (Kansal et al., 2000).

$$V = V_0 = \frac{A}{B} - exp(-Bt)) \tag{3}$$

The above equation utilizes several parameters to fit the growth curve to some experimentally derived data set. Figure 4 shows growth curves derived using this equation with the parameters A = 0.997, B = .0787, V_0 = 0.000103.

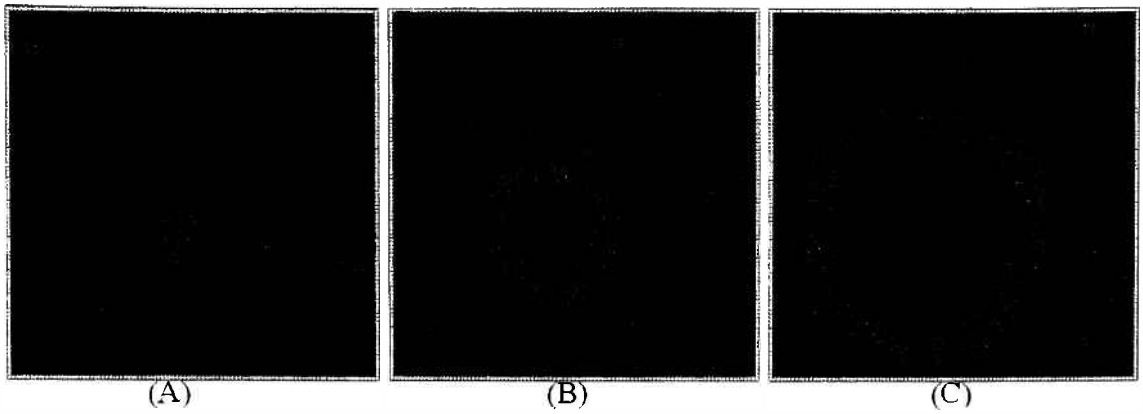

Figure 3: Cells dividing are light gray. (A) Shows initial tumor growth. (B) Proliferative (light gray) and non-proliferative (dark gray) layers are visible. (C) Necrotic center is colored black. Note random background growth due to random cell removal.

After a tumorous population of cells emerges, nutrient and space limitations inhibit growth. The result is a population growth curve that is implicitly Gompertzian (Figure 4).

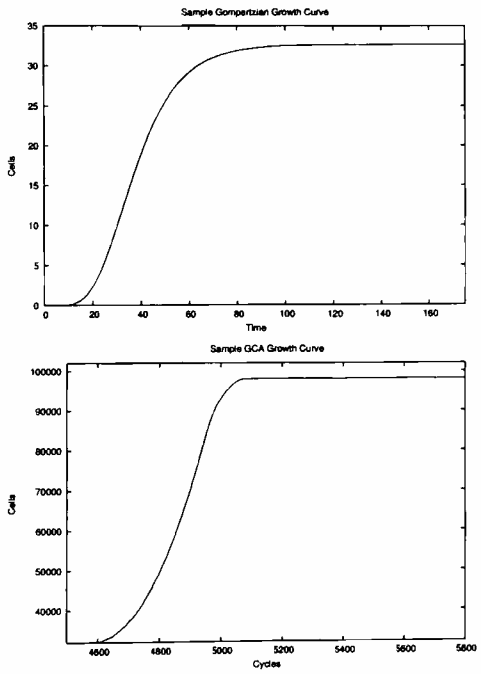

Figure 4: Comparison of Gompertzian growth derived from experimental data (Steel, 1977) (top) and a sample GCA growth curve (bottom).

Tumor Layering

As a result of nutrient and space limitations (Tannock and Hill, 1998), biological tumor masses often manifest three distinct layers:

1. Proliferative: Tumorous cells on the exterior of the tumor actively reproducing.
2. Non-Proliferative: Quiescent tumor cells without enough nutrients and space to reproduce.
3. Necrotic: Dead tumor cells.

Models by Kansal *et al.* (2000), Dormann and Deutsch (2002) use this layering phenomena to confirm their models. Such behavior emerges from our GCA model due to nutrients unable to support the dense tumor cell population. As a result, Figure 3 shows cells at the center of the tumor mass starve and may become non-proliferative or necrotic.

Tumor data (Kansal et al., 2000) also suggests that the outer proliferative layer is fixed in depth. If the density of the tumor is homogenous, nutrients would only be able to penetrate a fixed depth into the tumor mass. Our data support this feature: The outer circumference of the proliferative layer is a steady distance from the non-proliferative layer of the tumor.

A Clinical Approach to Classifying Multicellular and Tumorous Growth

To diagnostically confirm cancer, physicians perform clinical biopsies (AJCC, 2002). A sample of suspect tissue is removed and used for labratory testing. These tests are used to classify the tissue as being normal or tumorous.

We validate our GCA's ability to express both multicellular and tumorous existance using metrics which parallel those used by physicians for cancer diagnosis.

Nodule Formation

Tumorous nodules are the result of abnormally high cellular density (Tannock and Hill, 1998). By the definition of

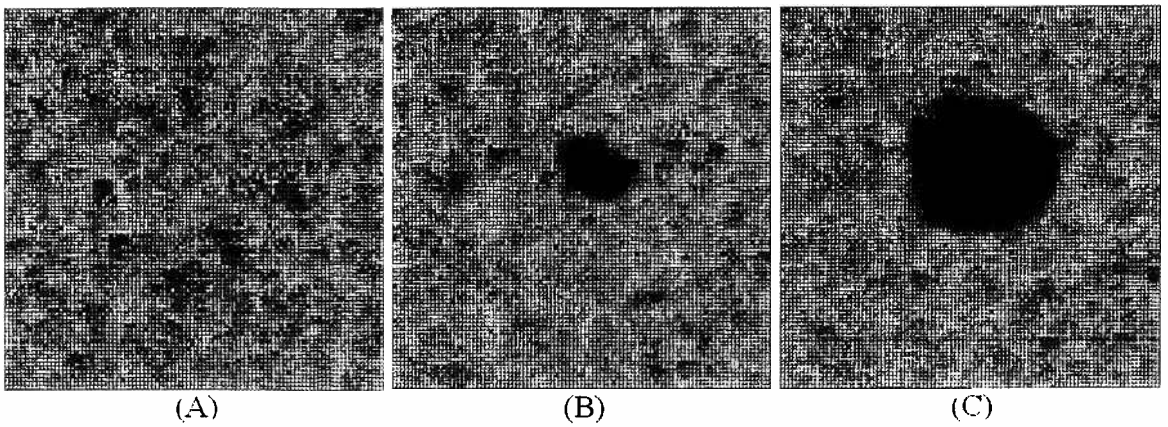

Figure 5: Grayscale indicates cell density (light = lower density, dark = higher density). (A) Homoestatic growth. (B) Initial nodule growth. (C) Expanding nodule growth.

cancer, this high cellular density is the result of abnormal reproduction. Tumorous nodules are often used as the initial indication of a tumorous population.

Figure 5(A) shows a homeostatic population of cells that exist at a moderate density throughout the lattice. Figures 5(B) and 5(C) show a tumorous population emerging (in darker coloring) via exponential growth to take over the lattice. The tumorous population is surrounded by a homeostatic population of cells–this contrast in density results in visible nodules.

Genetic Instability

Chromosomal annealing tests are used to observe major genomic inconsistencies of tumorous cells (Tannock and Hill, 1998). The DNA from a normally behaving cell of an organism may be compared to the genome of a tumorous cell to observe a ratio of similarity. With increased mutation and reproductive rates, tumor cells often genetically diverge from the organism's genome. We simulate chromosomal annealing by analyzing the average number of mutations present for cells at each box of the lattice. This test takes into account the magnitude of mutations.

Unlike actual biological genomic experiments, the GCA model accounts for **all** mutations that have acrued in the population since all cell genomes are accessible for analysis. Results shown in Figure 6 are based on the entire population, not samples that are used to statistically infer the entire population.

Growth Curves

After removing a biopsy tissue sample, cells may be grown in a culture for examination (AJCC, 2002). Features of Gompertzian growth are used above to validate tumorous growth curves. We use these characteristics to classify both

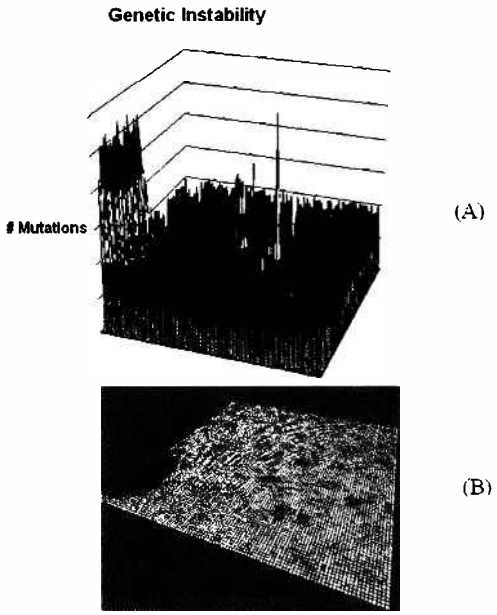

Figure 6: (A) Plot of relative mutations. (B) Corresponding grayscale density lattice.

normal and tumorous cells. Figure 7 shows growth curves resulting from several GCA runs.

The observed stochastic nature that tumorous populations emerge at different simulated time points is notable. This result is encouraging because tumorous growths occur stochastically throughout the lifetime of organisms.

Conclusion

We have designed a CA model with these unique features:

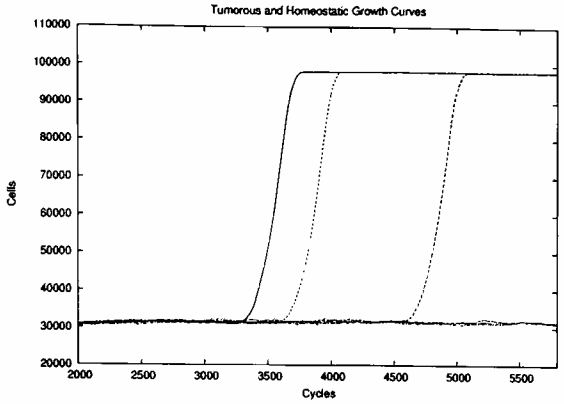

Figure 7: Growth curves generated using six GCA runs. Three runs remain homeostatic at 30,000 cells. Three runs stochastically grow in a Gompertzian manner to 100,000 cells.

- Genetically based rules
- Tumorous behavior evolves stochastically
- Non-tumorous and tumorous states are modeled

We validate our GCA model by presenting features used to validate previous CA models. We also present data based on metrics used to classify biospy tissue samples to verify that GCA can exhibit both multicellular and tumorous existence. Although this model is abstract, it reflects the basic genetic mechanisms that cause tumor growth.

GCA models may be extended to provide a framework for the addition of established genetic networks. Such genetic simulations may help the understanding of genetic dysfunction pathways and possibly provide treatment strategies for tumorigenesis. Also, unlike biological analyses, GCA models allow for the observation of the entire population–instead of drawing statistical relevance from a population sample. The GCA model presented evolves behaviors produced by other CA models that are specific to experimentally derived data sets. The behaviors evolved are the result of generalized genetic mechanisms that are known to result in tumorous behavior.

It is well known that cancer is an evolutionary disease. The task of treating a disease that evolves to the host's defenses and the physician's medications is a difficult one. It is our hope that this model will be used to simulate useful experiments and treatment strategies.

Future Work

We hope to expand the genetic networks represented by our GCA to involve established pathways derived from experimental data. Also, a three dimensional GCA is being considered that would allow a more realistic comparison to experimentally derived *in vitro* and *in vivo* data sets since these data are produced in three dimensions. Finally, we hope to use this model to generate hypotheses about genetic dysfunction pathways that lead to tumorigenesis.

Acknowledgements

This work is funded by NIH Grant #P20 RR16448-01. Experiments were simulated on the University of Idaho Beowulf Cluster Supercomputer funded by Intel Corporation and NSF Grant #EPS80935.

References

AJCC (2002). *AJCC Cancer Staging Manual.* American Cancer Society.

Fairbanks, D. J. and Andersen, R. W. (1999). *Genetics: The Continuity of Life.* Brooks/Cole Publishing Company.

Heath, J. K. (2001). *Principles of Cell Proliferation.* Iowa State University Press.

Kansal, A., Torquoato, S., Harsh IV, G., Chiocca, E., and Deisboeck, T. (2000). Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. *Journal of Theoretical Biology*, 203:367–382.

Kuby, J. (1997). *Immunology.* W.H. Freeman and Company.

Moreira, J. and Deutsch, A. (2002). Cellular automaton models of tumor development: A critical review. *Advances in Complex Systems*, 5:247–267.

Qi, A.-S., Zheng, X., Du, C.-Y., and Bao-Sheng, A. (1993). A cellular automaton model of cancerous growth. *Journal of Theoretical Biology*, 161:1–12.

Steel, G. G. (1977). *Growth Kinetics of Tumors.*

Tannock, I. F. and Hill, R. P. (1998). *The Basic Science of Oncology.* Mcgraw-Hill.