

Simulating Tumor Growth Models and Sampling and its effects on Heterogeneity

Student:

Mr. Zafarali Ahmed

zafarali [dot] ahmed [at] mail [dot] mcgill [dot] ca

Supervisor:

Dr. Simon Gravel

Submission Date:

20th September 2015

Location:

McGill University

Abstract

Significance: Tumor heterogeneity is the name given to the phenomenon of subclonal populations within a tumor having different genotypic composition. Understanding how tumor heterogeneity is distributed in space in a tumor during the time of therapeutic decision is of great importance as it changes the effectiveness of a targeted treatment. Quantification of tumor heterogeneity also proves to be a challenge as it is inherently affected by tumor sampling bias which causes the decision to be made under situations where we have incomplete information.

We use Cellular Potts Model (CPM) to simulate tumors in their natural environments and allow them to grow under three models: the traditional linear progression model, the Cancer Stem Cell (CSC) model and a baseline model. We compare the spatial patterns of these models and then proceed to quantify their genetic landscapes based on different sampling strategies that one might use in a clinical setting.

With the advancement of modern technologies like deep sequencing, we are able to now obtain single-cell resolution of the genetic landscape which can be used to validate models presented in this poster.

Through this we obtain a clearer picture into the biology of how tumor heterogeneity varies with model and sampling methodology. This can be used to make better informed decisions regarding therapy and forms the basis for further work.

Methods: We use Cellular Potts Model (CPM) to simulate tumors in their natural environments and allow them to grow under three models: long standing traditional linear progression model, the Cancer Stem Cell (CSC) model, and a baseline model. We emulate sampling and sequencing of their genomes to obtain the relationship between spatial expanse and the genetic landscape of the tumor using a custom built library. We calculate the expected proportion of pairwise differences and the number of segregating sites as popular statistics to quantify the heterogeneity. **Results:** Main results and observations **Limitations:** Due to memory and time constraints the entire simulation is limited to around 25,000 cells. In reality, a tiny incision of a tumor can contain an excess of 10^6 cells. **Conclusions:** Conclusions of the paper

Keywords: cancer, tumor heterogeneity, targeted medicine, computational oncology, cellular potts model.

1 Introduction

Cancer is the autonomous and uncontrolled growth of cells that forms a feed forward system to reinforce the tumors existence [2]. Cancer originates from accumulated mutations in tumor suppressor genes, proto-oncogenes and DNA repair genes. These mutations hinder a cell's ability to regulate its cell cycles, causing it to divide uncontrollably. *** MIGHT BE UNNECESSARY: However, for cancerous cells to keep growing and overcome the limitations of efficient diffusion of nutrients, tumors have two solutions: angiogenesis and metastasis. Angiogenesis permits a tumor to grow blood vessels and bring in nutrients. Metastasis is the migration of cancer cells to other parts of the body often via the permeable blood vessels formed during angiogenesis.**** It is often after metastasis that cancer becomes difficult to contain and treat despite surgical resection [1].

1.1 Background

1.1.1 Tumor Heterogeneity

Not all cancer cells have the same phenotypic characteristics. Higher mutation rates and weaker DNA repair mechanisms lead to a wide variety of phenotypes amongst the cells in the tumor. Cancer cells consist of subclonal populations that have different growth rates, metastasizing abilities and response to therapeutic drugs [4]. This phenomenon of differing subclonal phenotypes within a tumor is known as tumor heterogeneity. If certain subclones have a mutation that gives them resistance to a therapy, these cancer cells will survive and eventually take over and lead to a therapeutically resilient population. This proves to be a challenge for later treatment even in the context of surgery. Tumor heterogeneity develops during the mitotic events during cancer division in the form of random mutations, chromosomal rearrangements and replication errors. From the mutations encoded within the genomes of cancer cells in the tumor we can reconstruct the method of progression [5]. However, there is noise in this data presented from neutral mutations which do not provide an advantage nor a disadvantage to the cell. In small populations a neutral mutation might be prevalent due to genetic drift. ***TALK ABOUT DAVID NELSON HERE?***

1.1.2 Cancer Stem Cells (CSCs)

The traditional view of tumor progression is a linear evolutionary model having a single cell origin. Initially, a normal cell develops a growth advantage allowing it to compete favourably against neighbouring normal cells. At each mitotic event, the cancer cells accumulate different mutations that leads to competition and eventual selection of a few of the most advantageous mutations [6].

Another compelling model exists based on evidence that certain cancer cells acquire unlimited proliferative potential but producing cancer cells that have limited proliferative potential [8]. The cells with unlimited proliferative potential are known as Cancer Stem Cells (CSCs).

It is expect that these two models will produce vastly different spatial and genetic landscapes.

1.2 Motivation and Significance

Understanding how tumor heterogeneity is distributed in space in a tumor during the time of therapeutic decision is of great importance. Quantification of tumor heterogeneity also proves to be a challenge as it is inherently affected by tumor sampling bias [4][5] which causes the decision to be made under situations where we have incomplete information.

In this paper, we take a look at two theoretical cancer models: the Mutational Response Model and Cancer Stem Cell (CSC) Model, in addition to a simpler model for comparison. We compare the spatial patterns of these models and then proceed to quantify their genetic landscapes based on different sampling strategies that one might use in a clinical setting. With the advancement of modern technologies like deep sequencing, we are able to now obtain single-cell resolution of the genetic landscape which can be used to validate models presented in this paper.

Through this we obtain a clearer picture into the biology of how tumor heterogeneity varies with model and sampling methodology. This can be used to make more informed decisions regarding therapy and forms the basis for further work.

2 Methods

2.1 Cellular Potts Model (CPM)

The Cellular Potts Model (CPM) [3] was used to simulate tissue cultures. The model is an extension of the q-Potts Model in statistical mechanics and is powerful due to its simplicity.

The Cellular Potts Model consists of a $M \times M$ lattice of *spins*. For a more biological interpretation, we can consider a *spin* to represent an area of *cell cytoplasm*. To maintain consistency with established literature, we will continue to use the term *spin*. At any location of the lattice, (x, y) , a spin can take on a value, $\sigma(x, y) = \{0, 1, \dots, N\}$. A *cell*, σ , is defined as a collection of these spins; as such, this restricts same-spin values to be clustered together in the lattice. If $\sigma(x, y) = 0$, then this region is said to be the *extra cellular matrix (ECM)*. To be able to simulate different *types* of cells, for example proliferating and migrating cells, we introduce $\tau(\sigma)$, which returns the *type* of the cell σ . Two cells can be seen in Figure ??.

For state of the lattice, we define a *Hamiltonian*, H , which captures the total energy of the system. Surface tension between cells, restrictions in cell areas and volumes and effects of external potentials are amongst the phenomenon taken into account by H .

The simulation then proceeds via a Monte Carlo method to flip spins on the lattice in an attempt to grow and move cells. Cells will move to minimize the Hamiltonian, H .

For simulations in this paper, we have use CompuCell3D [7] (version 3.7.4) which is a widely adopted software by computational biologists aiming to simulate many different tissue systems based on the Cellular Potts Model.

2.2 The Infinite Genome

To simulate genomes and phenotypes of the cancer cells, we built an extension to the CompuCell3D software. The positional loci of the genome is modelled as a random variable, $L = 10^{15} \times L^*$, $L^* \sim U(0, 1)$, where L^* is picked from the uniform distribution with parameters $a = 0, b = 1$. By the limitation of doubles, the number of loci are limited to 10^{15} . Thus the probability of any mutation

appearing between loci at position l and l' is $P(l' < L < l) = \frac{l'-l}{10^{15}}$. At each mitotic event we pick, $\lambda, \lambda \sim Pois(\lambda)$, loci to mutate. Once a locus is marked as mutated, it cannot switch back to an unmutated state.

2.3 The Models

To simulate background neutral mutations and natural cell shapes we divided our simulation into two distinct phases. In phase one, known in the literature as the *naturalization period*, we grow a single normal cell to divide and accumulate a lattice of area 900×900 within a 1000×1000 lattice. To avoid edge effects, we allow the remaining space to consume cells and allow the central area to extend infinitely. In phase two, we then select a single cell within 10×10 square units of the lattice center to be the cancer cell and then resume the simulation this time turning off growth for normal cells. In all models, cancer cells have a lower dividing threshold and a higher mutation rate. When a single cancer cell reaches the edge of the lattice, we terminate the simulation.

Three models of cancer were simulated:

1. The 1st Order Model only had cancer cells which have a higher mutation rate.
2. The Mutational Response Model had the growth rate, dV , according to the logistic equation (Eq 1):

$$dV = \frac{1}{1 + \exp \left(- \left[0.01X - \ln \left(\frac{1}{dV_{min}} - 1 \right) \right] \right)} \quad (1)$$

X = Number of mutations in the first 10% of the genome

dV_{min} = The minimum growth rate of the cell

3. In the cancer stem cell model, we have two cancer cell types: cancer stem cell and the regular tumor cell. The starter cancer cell is labelled as a CSC. After division, the daughter cell A inherits the CSC characteristic with probability p_{CSC} or remain with a regular cancer phenotype with probability $1 - p_{CSC}$, while the daughter cell B can lose its CSC phenotype with probability p_{endCSC} .

Exact parameters for each simulation are summarized in table ***DO TABLE****

2.4 Sampling

When the simulation is completed, we need to sample from the tumor to obtain a subset of cells that we will use to quantify diversity. Cells are sampled by ellipses with varying eccentricities $0 \leq \epsilon < 1$. An ellipse with $\epsilon = 0$ resembles circular patch from the tumor while $\epsilon = 0.99$, resembles rectangular slice from the tumor. We scale the area of ellipses, $A_{ellipse}$ according to the number of cells in that sample, N , and the average area of a cell \bar{A} : $A_{ellipse} \propto N\bar{A}$. We can write the major and minor axes, a and b respectively in terms of these parameters:

$$a^2 = \frac{\bar{A}N}{\pi\sqrt{1-\epsilon^2}}; \quad b^2 = \frac{\bar{A}N\sqrt{1-\epsilon^2}}{\pi}$$

We select 20 points, k_1, \dots, k_{20} , within the tumor to sample our ellipses with a random angle, θ . For each point k_i , we select the N closest cells to that point. After this, we subsample a fixed number of cells, M , which ensures that our statistics are not affected due to the increasing N but only due to area $A_{ellipse}$.

2.5 Proportion of Pairwise Differences and Number of Segregating Sites

Two statistics are used to quantify diversity in tumor samples. The expected proportion of pairwise differences, $E(\pi)$, is defined in terms of the number of chromosomes, n , the number of mutated chromosomes at loci l , m_l , and total number of loci, M :

$$E(\pi) = 2 \sum_{l=1}^M \left(\frac{m_l}{n} \right) \left(\frac{n - m_l}{n} \right) \quad (2)$$

The number of segregating sites, S , is defined as the total number of mutated sites in the tumor sample. We scale S using the Harmonic series, $\frac{S}{H_n}$. H_n is the theoretical expected growth in the number of segregating sites given a randomly mating population.

3 Results

4 Discussion

5 Conclustions

References

- [1] R. Demicheli, M. W. Retsky, W. J. M. Hrushesky, M. Baum, and I. D. Gukas. The effects of surgery on tumor growth: a century of investigations. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*, 19(11):1821–8, Nov. 2008.
- [2] Douglas Hanahan and R. A. Weinberg. Hallmarks of Cancer: The Next Generation. *Cell*, 2011.
- [3] F. Graner and J. a. Glazier. Simulation of Biological Cell Sorting Using a Two-Dimensional Extended Potts Model. *Physical Review Letters*, 69(13):2013–2016, 1992.
- [4] G. H. Heppner and B. E. Miller. Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer and Metastasis Review*, 2(1):5–23, 1983.
- [5] K. Naxerova and R. K. Jain. Using tumour phylogenetics to identify the roots of metastasis in humans. *Nature Reviews Clinical Oncology*, 12(5):258–272, 2015.
- [6] P. Nowell. The clonal evolution of tumor cell populations. *Science*, 194(4260):23–28, 1976.
- [7] M. H. Swat, G. L. Thomas, J. M. Belmonte, A. Shirinifard, D. Hmeljak, and J. a. Glazier. *Multi-Scale Modeling of Tissues Using CompuCell3D*, volume 110. Elsevier Inc., 2012.
- [8] M. H. Tomasson. Cancer stem cells: A guide for skeptics, 2009.