Ordinary Differential Equations System for Multistep transformation to Cancer Revisited

Miao Liu (U21118674)

Chenxin Sun (U28161804)

Yanlin Chen (U)

MA579 Project

Supervisor: Professor Samuel Isaacson

May 4, 2021

Introduction

The progression of normal cells (N) to primary tumor cells (T) involves several stochastic mutations. to simplify the model to a feasible mathematical model, only four crucial mutations were considered: angiogenesis (A), decreased apoptosis (D), genetic instability (G), and increased replication rate (R), which are heritable variations that can be recombined. Cells with all four mutations are turned into primary tumor cells. The multistep progression of primary tumor cells to metastatic cells (M) was cut down to a one-step process. To simplify the system of ordinary differential equations (ODEs), the following assumptions were made:

1. The population of normal cells is constant, 10^8 [1].
2. The size of the tumor cannot grow past 10^6 cells unless 10% of the population has acquired the mutation A [2].
3. The lethal tumor limit is 10^{13} cells [3].
4. Numbers of genes involved in single, double, and triple mutations are 100, 10, and 1, respectively [4].

Through the simulation, we are interested in whether the specific order of mutations is important and how that affects the kinetics of tumor development. First, the simulations of single, double, triple mutations were run. According to different numbers of cells with various types of mutations, the fastest path to cancer was hereby concluded. In this paper, formulation of the ODE with multistep cancer transformation is discussed, and the fastest path to form tumors is re-examined, with further discussion of biological interpretations.

Model Description

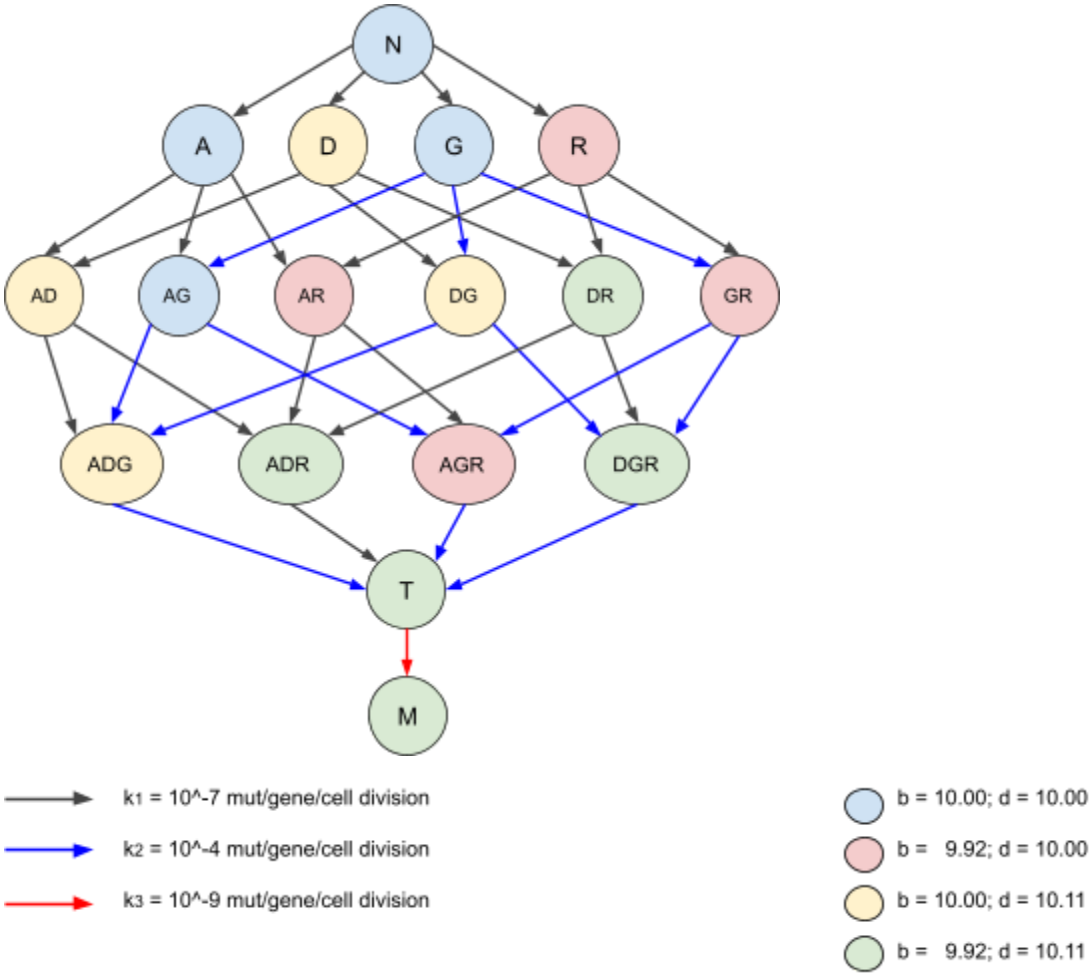

Figure 1 - State diagram of the model. Normal cells acquire one of the four mutations (angiogenesis (A), decreased apoptosis (D), genetic instability (G), and increased replication rate (R)) at the rate of k_1. The mutation rate increases to k_2 after G is acquired. Similarly, the cell division rate without R is every 10 days; the death rate without D is every 10 days. The division rate increases to every 9.92 days after mutation R; the death rate decreases to every 10.11 days after mutation D. Cells with all four mutations are primary tumor cells (T). The transition from tumor cells (T) to metastatic cells (M) happens at a constant rate of k_3. Cells with multiple mutations are named alphabetically.

The ordinary differential equation (ODE) model is suitable to understand the kinetics of cancer progression. 17 ODEs were constructed, corresponding to cells with 17 types of gene mutations. Cells can gain one out of four mutations at one time along the pathway (Figure 1).

The increment of cells in each state corresponds to the replication of current-state cells and the influx from the previous state cells, while the total decrement of cells is the sum of apoptosis and efflux of cells proceeding to the next state. All ODEs are bounded by two logistic terms: the angiogenesis cap and lethal tumor cells burden, respectively.

Parameter	*Description*	*Value*
k_1	Mutation rate without G	10^{-7} mutation/gene/cell division
k_2	Mutation rate with G	10^{-4} mutation/gene/cell division
k_3	Metastasis rate	10^{-9} /cell division
b	Cell birth rate without R	/10 days
b_R	Cell birth rate with R	/9.92 days
d	Cell death rate without D	/10 days
d_D	Cell death rate with D	/10.11 days

Table 1 - Parameters used in ODEs

The ODEs were reformatted into the vector equation:

$$\frac{dy}{dt} = (\mathrm{diag}(\mathrm{diag}(y^T \cdot k) \cdot b)^T \cdot M + (\mathrm{diag}((b-d)^T \cdot y))^T) \cdot S$$
$$\cdot(1 - a(y)P_{nm}/10^6) \cdot (1 - P_{nm}/10^{13}) + m_m$$

Notice that the equation above is slightly different than the version proposed by Spencer et al.[5]. y is the row vector that corresponds to the number of cells in each state. We partition the state from figure 1 as left to right, top to bottom, i.e., $y[1]$ (here we choose index starting from 1) is the number of normal cells, $y[2]$ is the number of cells with A mutation, $y[3]$ is the number

of cells with D mutation, etc. *k* is the rate row vector that corresponds to mutation rate in for each state, i.e., k_i (mutation/genes/cell division) represents the rate with $y[i]$ in y. $\mathrm{diag}(\cdot)$ is the matrix operator that gives diagonal entries of a matrix to form a column vector. b is the replication rate row vectors of cells in each state, and d is the death rate row vector of cells in each state. Note that the notation of those two vectors might be confusing since b_i is in the form of either $1/b$ or $1/b_r$ and similar for d. M is a 17 by 17 upper triangular matrix where M_{ij} corresponds to the number of genes to transform from state i to state *j* for $i \neq j$. The diagonal entry of M is computed as

$$M_{ii} = -\sum_{i \neq j} M_{ij},$$

where the main diagonal of the M corresponds to the number of genes leaving the state i.

Matrix S is another 17 by 17 matrix that helps to apply appropriate cell population cap for each state and is calculated by

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

P_{nm} is the number of non-normal(N) and non-metastatic(M) cells, which is calculated by

$$P_{nm} = \sum_{i=2}^{16} y_i.$$

The terms before S represent the rate of net increment and decrement of cells in each state, the term after S represents the cap each state is enforced. The first cap is angiogenesis with logistic terms of 10^6, proposed by Folkman[6], that only exists when the population of angiogenesis is less or equals to 10% of the total non-normal, non-metastatic cells, which captured by function $a(y)$. The second cap addresses the burden of lethal tumor cells in P_{nm} with logistic terms of 10^{13} suggested by Friberg and Mattson[7].

However, the mathematical model given in the article did not contribute to the exact same plots. Appropriately, some modifications were made to duplicate their result. The following initial conditions were chosen according to Figure 2 of the original paper, such that:

$y_0 = \left(10^8, 10^4, 10^4, 10^4, 10^4, 10^3, 10^3, 10^3, 10^3, 10^3, 10^3, 10^2, 10^2, 10^2, 10^2, 0, 0\right)$,

meaning at time zero, the numbers of normal cells, cells with a single mutation, double mutations, triple mutations are 10^8, 10^4, 10^3, 10^2, respectively, whereas the numbers of tumor cells and metastatic cells are 0. To match the graph of the original paper, $a(y)$ was set to zero, which means the angiogenesis cap was ignored. The modified equation was then solved by the Runge-Kutta method using package Tsit5 provided by Julia with relative tolerance of adaptive steps 10^{-5} and absolute tolerance of 10^{-4}. After the adjustments, our model shows the qualitatively identical result as proposed by S.L. Spencer (Figure 2).

Result & Discussion

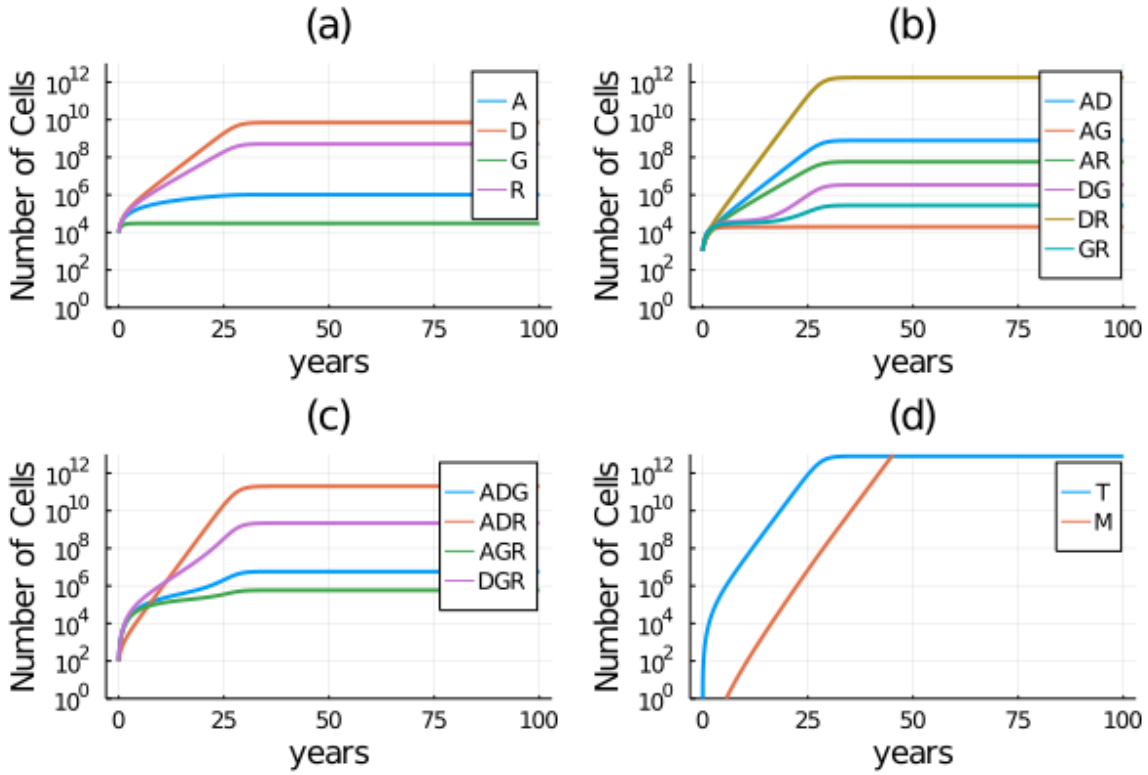

Figure 2 - Dynamic simulation of cell populations with single mutation (a), double mutation (b), triple mutation (c), tumor cells, and metastatic cells (d) vs. time.

According to our simulation result, we successfully identified the individual contribution of the four mutations during the progression of normal cells to cancer cells. For cells with a single mutation (Figure 2(a)), acquiring the delayed apoptosis (D) mutation enables them to proliferate for a longer period, thus outrunning the other three types of mutation. In the following state, when the second mutation is gained (Figure 2(b)), cells with increased replication rate (R) further increase the population of mutants by having a large birth rate and small death rate. As a result, the DR cells plateau at the highest value. However, because of the angiogenesis cap, when the number of tumor precursor cells approaches the 10^6 threshold, they need blood vessels to supply nutrients as well as oxygen, which explains why the angiogenesis (A) mutation is

required for cells to grow past 10^6, at which the ADR and DGR cells intersect (Figure 2(c)). In the last stage, genetic instability (G) is acquired (Figure 2(d)), cells with all four mutations become tumor cells that metastasize at the rate of 10^{-9} per cell division to metastatic cells. Since we only monitored cell populations, mutation G comes to play a role in the final state because it does not have any selective advantages that escalate cell proliferation, such as delayed death and faster division rate. The population of metastatic cells increases almost linearly and sooner becomes the dominant species, surpassing that of primary tumor cells. Another observation that reflects our assumption is that since the non-normal, non-metastatic cells are capped by the lethal tumor burden limit, 10^{13} cells, all the other 15 types of mutants except the metastatic cells cannot grow beyond 10^{13}. Therefore, the plots all plateau at various values lower than 10^{13}.

We hereby concluded the fastest path to cancer (Figure 3), is by acquiring mutation D, followed by R, A, and finally G, as demonstrated in the original paper.

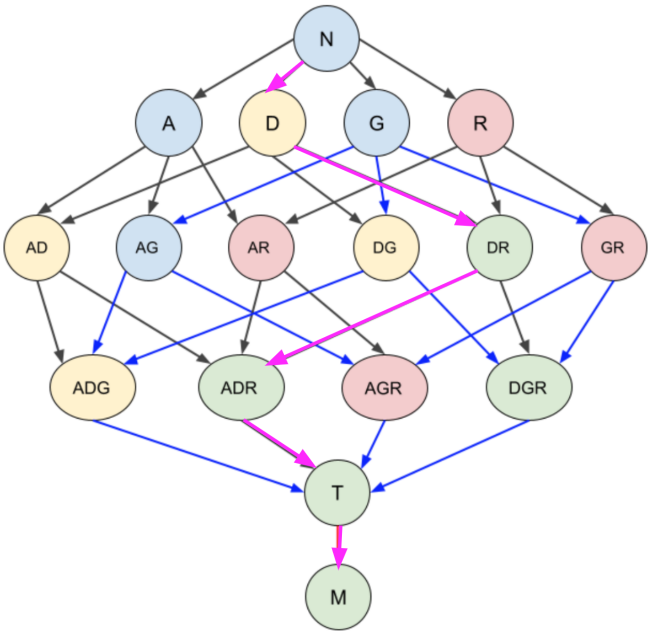

Figure 3 - Fastest path to cancer. The order of acquiring mutations is D, R, A, and G.

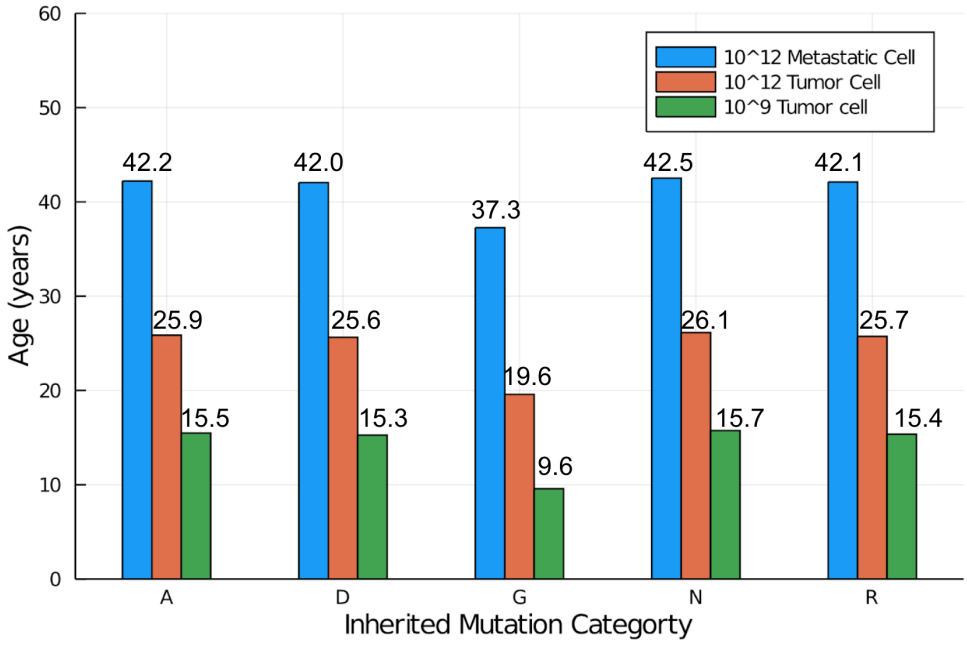

Figure 4 - Cancer onset age vs. inherited mutations. Normal cells (no mutation is inherited) are denoted as N. Inheriting a mutation in G generates a significantly earlier onset.

Although the effect of obtaining mutation G is the least in cancer progression dynamics, it affects the onset of cancer drastically. In this project, we also analyzed the cancer onset ages of different inherited mutations, which are inactivated and the initial cells at time zero are still functioning as normal cells. However, the functional normal cells were set to have a larger mutation rate. Instead of having the rate of 10^{-7} mutation/gene/cell division, the mutation rate from normal cells to cells with a single inherited mutation increased to 10^{-5} mutation/gene/cell division. To calculate the onset age, the diagnosable tumor size was set to three thresholds: 10^{9} primary tumor cells (T), 10^{12} primary tumor cells (T), and 10^{12} metastatic cells (M). The resulting histogram (Figure 4) shows that patients with an inherited single mutation in A, D, and R would not have significantly different onset ages comparing to patients with no inherited mutation (N). The figure is shown by running 5 different simulations with an appropriate change in rate in A,

D, G, R, respectively (rate in group N means no change in rate), along with the new initial condition $y_0 = \left(10^8, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\right)$.

On the other hand, patients who have inherited a mutation in G show a compelling earlier onset of tumors, which is on average 6 years earlier than patients with no mutation. One thing to notice is that our simulation produced the same pattern comparing to the original paper, whereas different values of the onset age. This might be caused by confusion about the initial cell numbers that were presented in the original paper.

Nonetheless, there are several problems with the original paper that arose during our reproduction of the simulation. First, as mentioned in the previous section, the application of the angiogenesis cap remains unclear, although the mechanism of angiogenesis limitation is well explained. Furthermore, they did not clarify how they chose the initial values. One possible reason might be "to account for the fact that some genes function in more than one category, we allow double and triple state transitions but reduced the number of genes involved to the order of 10 and 1, respectively" [5]. Likewise, the initial populations of cells with single, double, and triple mutations decrease by an order of 10.

Last but not least, since the author only capped the growth of primary tumor cells, the growth of metastatic cells is not inhibited, therefore the number of metastatic cells can keep increasing until it becomes exponential. To solve this problem, another cap may be introduced. For example, the maximum number of body cells including normal cells and all 16 possible mutants cannot surpass 10^{13}. Another possible solution is to construct a division-limiting equation according to the Hayflick limit[8], which reduces the cell division to about 50 times before entering a senescence phase.

References

[1] Nowak, M. A., Komarova, N. L., Sengupta, A., Jallepalli, P. V., Shih, I., Vogelstein, B., & Lengauer, C. (2002). The role of chromosomal instability in tumor initiation. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(25), 16226–16231.

[2] Kerbel R. S. (2008). Tumor angiogenesis. The New England journal of medicine, 358(19), 2039–2049.

[3] van Leeuwen, I. M., Zonneveld, C., & Kooijman, S. A. (2003). The embedded tumour: host physiology is important for the evaluation of tumour growth. *British journal of cancer*, *89*(12), 2254–2263.

[4] Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., & Stratton, M. R. (2004). A census of human cancer genes. *Nature reviews. Cancer*, *4*(3), 177–183.

[5] Spencer, S. L., Berryman, M. J., García, J. A., & Abbott, D. (2004). An ordinary differential equation model for the multistep transformation to cancer. *Journal of Theoretical Biology*, *231*(4), 515-524.

[6] Folkman, J. (1990). What is the evidence that tumors are angiogenesis dependent?. *JNCI: Journal of the National Cancer Institute*, *82*(1), 4-7.

[7] Friberg, S., & Mattson, S. (1997). On the growth rates of human malignant tumors: implications for medical decision making. *Journal of surgical oncology*, *65*(4), 284-297.

[8] Hayflick L. (1965). The limited in vitro lifetime of human diploid cell strains. *Experimental cell research*, *37*, 614–636.