

PAPER

Temporal order of mutations influences cancer initiation dynamics

To cite this article: Hamid Teimouri and Anatoly B Kolomeisky 2021 *Phys. Biol.* **18** 056002

View the [article online](#) for updates and enhancements.

You may also like

- [NRPB Chilton Seminar: Professor Bryn Bridges—'Genetic Effects of Radiation: Are the Paradigms Changing?' \(Chilton, November 2001\)](#)
- [Lineage dynamics and mutation–selection balance in non-adapting asexual populations](#)
Sophie Péniisson, Paul D Sniegowski, Alexandre Colato et al.
- [Reply to 'Comments on Hereditary Effects of Radiation'](#)
K Sankaranarayanan and N E Gentner



 **EDINBURGH INSTRUMENTS**

WORLD LEADING MOLECULAR SPECTROSCOPY SOLUTIONS

edinst.com

Physical Biology



PAPER

Temporal order of mutations influences cancer initiation dynamics

RECEIVED
3 March 2021

REVISED
2 June 2021

ACCEPTED FOR PUBLICATION
15 June 2021

PUBLISHED
2 July 2021

Hamid Teimouri^{1,2}  and Anatoly B Kolomeisky^{1,2,3,4,*} 

¹ Department of Chemistry, Rice University, Houston, Texas, United States of America

² Center for Theoretical Biological Physics, Rice University, Houston, Texas, United States of America

³ Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas, United States of America

⁴ Department of Physics and Astronomy, Rice University, Houston, Texas, United States of America

* Author to whom any correspondence should be addressed.

E-mail: tolya@rice.edu

Keywords: cancer initiation, stochastic processes, mutation fixation

Supplementary material for this article is available [online](#)

Abstract

Cancer is a set of genetic diseases that are driven by mutations. It was recently discovered that the temporal order of genetic mutations affects the cancer evolution and even the nature of the disease itself. The mechanistic origin of these observations, however, remain not well understood. Here we present a theoretical model for cancer initiation dynamics that allows us to quantify the impact of the temporal order of mutations. In our approach, the cancer initiation process is viewed as a set of stochastic transitions between discrete states defined by the different numbers of mutated cells. Using a first-passage analysis, probabilities and times before the cancer initiation are explicitly evaluated for two alternative sequences of two mutations. It is found that the probability of cancer initiation is determined only by the first mutation, while the dynamics depends on both mutations. In addition, it is shown that the acquisition of a mutation with higher fitness before mutation with lower fitness increases the probability of the tumor formation but delays the cancer initiation. Theoretical results are explained using effective free-energy landscapes.

It is widely accepted that cancers are caused by genetic alterations in normal tissue cells [1–4]. These changes lead to abnormal growth of those cells that can no longer support the required functions of the tissue. The mechanistic origin of the increased competitiveness of cancer cells have been the focus of significant research efforts that identified the key features of tumors [1, 2, 4, 5]. Traditionally, the cancer phenotype has been considered as a sum of complementary properties that each mutation drives separately [4]. These views, however, have been recently challenged by experimental observations suggesting that the order of mutations is also an important factor in cancer development [6–8]. But the microscopic picture behind the effect of the order of mutations in cancer remains largely unexplored.

There are now several examples of how the different order of mutations lead to dramatically different outcomes in the cancer development [6, 7, 9–13]. In the mouse model of adrenocortical tumors, it was observed that when the mutation in the oncogene Ras

preceded the mutation in p53 gene this led to highly malignant tumors with strong metastatic properties. But the reverse order of mutations (p53 first and then Ras) produced only benign tumors [9]. Another example comes from the studies on chronic myeloproliferative neoplasms, which are myeloid tumors that contain on average between 5 to 10 somatic mutations [7]. Two driver mutations, *TET2* and *JAK2*, were identified as the most common in these cancers. Blood cells analysis showed that patients with *TET2*-first mutations were a decade older than the patients with *JAK2*-first mutations. Moreover, it was observed that the order of mutation acquisition leads to different cancer types and probabilities of survival [7].

These observations stimulated discussions on how to explain the mutations order effect [6, 8]. Three possible phenomenological mechanisms have been proposed [6]. One of them suggests that the first mutation might change the accessibility of specific genomic regions and this prevents the second

mutation from activating or repressing this region. If the order of mutations is reversed the blocking does not happen. The second idea is that initially one mutation can induce rapid cell growth and differentiation, while starting from another mutation lead to much slower growth and less differentiation. The third mechanism argues that different starting mutations might lead to different cellular microenvironment, influencing this way the disease evolution. These mechanisms, however, do not explain the microscopic origin of these phenomena. Theoretical models play an important role in uncovering the mechanisms of cancer progression [1, 14–16]. They provide a critical quantitative link between the underlying microscopic processes and the appearance of tumors. Significant efforts have been made in modeling the dynamics of mutation acquisition and how it is governed by relevant genetic parameters such as the rate of mutations, the size of the population of cells and the rate of mutations proliferation [14, 17–21].

It is worthwhile to note that similar problems have been encountered in various aspects of population genetics [22, 23]. There is a significant progress in our understanding of mutational landscapes and how they are affecting the outcomes of cellular processes. However, to the best of our knowledge, explicit calculations have not been performed to elucidate how the temporal order of mutations influence the probability and dynamics of tumor formation and growth.

In this paper, we present a theoretical model of cancer initiation that allows us to quantitatively explain the effect of different order of mutations. In our theoretical approach, the cancer appearance is viewed as a fixation of two different mutations. This is a so-called two-hit model which assumes that the first type of mutations occupies the whole tissue before the second type of mutations starts and eventually fills the system. Using the method of first-passage processes, we explicitly evaluate the dynamics of cancer initiation. It is found that the order of mutations strongly influences both probabilities and times before the cancer appears. We show that if the first mutation has a higher fitness than the second mutation, this increases the overall probability of the cancer initiation, but, surprisingly, delays the formation of the tumor. These observations are quantified using effective free-energy landscapes for the underlying stochastic processes.

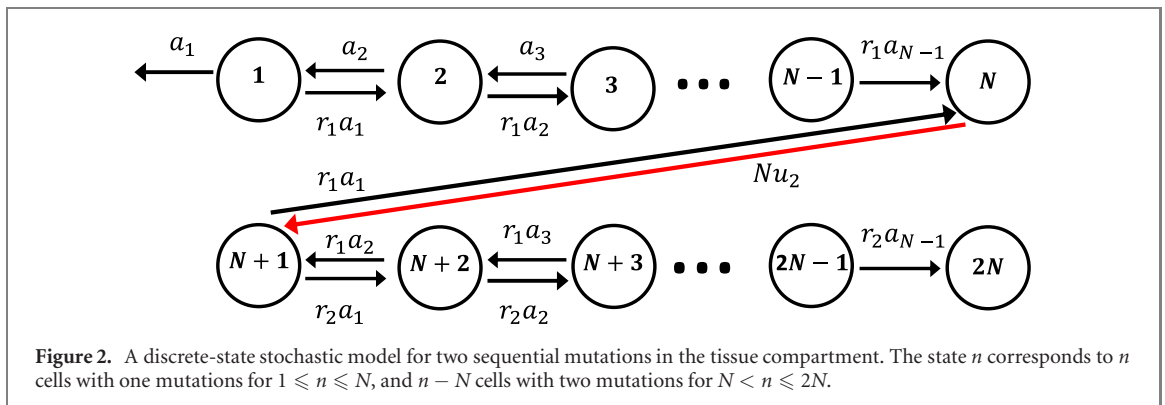
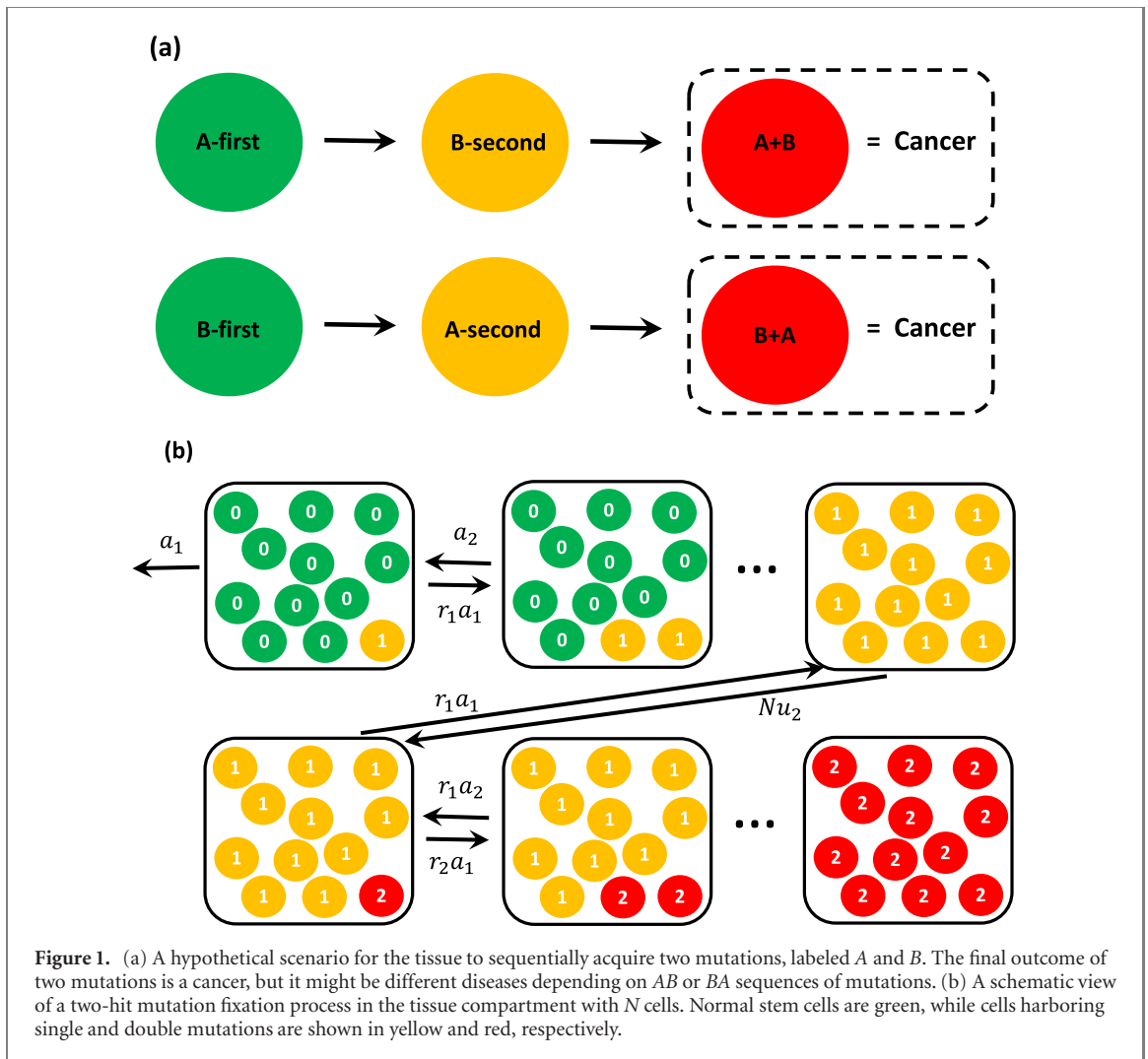
1. Theoretical method

In our theoretical approach, the cancer initiation is viewed as an event that starts when two different mutations take over the healthy tissue as illustrated in figure 1. More specifically, we assume that originally in the well-mixed tissue compartment there are N normal stem cells and the total number of cells is always fixed. At some time, which we view as $t = 0$,

one stem cell gets a first mutation with a probability u_1 . It is realistic to assume that the mutation rate is very low ($Nu_1 \ll 1$) [14, 16, 24, 25], and the further transformations in the system are taking place only via stem cells divisions and removals to keep the total number of cells fixed. Normal cells divide with a speed b , while the cells with the mutation divide with a rate r_1b . The parameter r_1 is known as a fitness parameter, and it is equal to the ratio of the division rates for the mutated cell over the normal cell. It reflects the overall physiological influence of the mutation on cellular metabolism: if $r_1 < 1$ the mutation is disadvantageous, $r_1 = 1$ corresponds to a neutral effect, while for $r_1 > 1$ the mutation is advantageous. The overall number of mutated cells in the fixed population of stem cells will be fluctuating, and there is a time when all cells would become mutated. This is known as a mutation fixation, and the system cannot have normal cells after that. At this moment, we assume that the second mutation will appear with a probability u_2 , which is also very small ($Nu_2 \ll 1$). The cells with one mutation divide with the rate r_1b , while the cells with two mutations divide with the rate r_2b . The fitness parameter r_2 reflects the accumulative effects of both mutations present in the cell.

Following our previous work [26], because the total number of cells in the tissue compartment should remain constant (this is known as a homeostatic equilibrium), we assume that the system follows a Moran process [14, 27, 28]. This means that after division of any randomly chosen cell (with arbitrary number of mutations) the number of cells in the compartment temporarily increases by one, and then one of the randomly chosen cells should be instantly removed to keep the number of cell constant and equal to N : see figure 1 (bottom).

Introducing two mutations in the tissue compartment generates three types of cells (see figure 1(b)). The normal wild-type cells are labeled as type 0 (green circles), the cells with one mutation are labeled as type 1 (yellow circles), and the cells harboring two mutations are labeled as type 2 (red circles). At any time, the tissue might only have the cells of type 0 and 1, or only the cells of type 1 and 2 (figure 1(b)). The dynamic changes in the system can be viewed as stochastic transitions between $2N$ discrete states, as shown in figure 2, and these states are specified by the number of cells with one or two mutations. We define a state n ($1 \leq n \leq N$) as a state that has n cells with one mutation and $N - n$ wild-type cells without mutations. If $N < n \leq 2N$ then the state n describes a situation with $n - N$ cells with two mutations and $2N - n$ cells with one mutations: see figure 2. The state $n = N$ corresponds to the fixation of the first mutation (all cells have one mutation), while the state $n = 2N$ corresponds to the fixation of both mutations (all cells have two mutations). It can be shown that for the states on the first branch ($1 \leq n \leq N$), the forward transition rate from the state n to the state



$n + 1$ is given by $r_1 a_n$ and the backward transition rate from the state n to the state $n - 1$ is equal to a_n where assuming that divisions and removals follow the Moran process [27] one obtains [26]

$$a_n = b \frac{n(N - n)}{N + 1}. \quad (1)$$

For the states on the second branch ($N < n \leq 2N$) it can be shown that the forward transition rate from the state n to the state $n + 1$ is given by $r_2 a_{n-N}$ and the reverse transition from the state n to the state

$n - 1$ is given by $r_1 a_{n-N}$: see figure 2. The transition rate from the state N to the state $N + 1$, which is given by the appearance of the second mutation, is equal to Nu_2 . Transitions between all neighboring states are reversible except two situations. When the system goes from the state $N - 1$ to N it corresponds to the elimination of all wild-type cells and it cannot be reversed. Similarly, the transition from the state $2N - 1$ to $2N$ corresponds to the elimination of all cells with only one mutation. In addition, the single

mutated cell can be eliminated from the system with the rate a_1 from the state 1 (figure 2).

The dynamics of cancer initiation with two sequential mutations can now be fully analyzed using the method of first-passage processes [26]. This is because in our discrete-state stochastic model (figure 2) the cancer initiation corresponds to the events that start in the state 1 and reach the state $2N$ for the first time. We define a function $F_n(t)$ as the corresponding first-passage probability density function to start from any site n at $t = 0$ and to reach the final fixation state $2N$ at time t . The temporal evolution of these functions are governed by the following backward master equations,

$$\begin{aligned} \frac{dF_n(t)}{dt} = & r_1 a_n F_{n+1}(t) + a_n F_{n-1}(t) \\ & - a_n(1 + r_1) F_n(t), \end{aligned} \quad (2)$$

for $1 \leq n < N$, and

$$\begin{aligned} \frac{dF_n(t)}{dt} = & r_2 a_{n-N} F_{n+1}(t) + r_1 a_{n-N} F_{n-1}(t) \\ & - a_{n-N}(r_2 + r_1) F_n(t), \end{aligned} \quad (3)$$

for $N < n < 2N$, while

$$\frac{dF_N(t)}{dt} = Nu_2 F_{N+1}(t) - Nu_2 F_N(t) \quad (4)$$

for $n = N$. In addition we have the boundary condition, $F_{2N}(t) = \delta(t)$, which physically means that if the system starts in this state the cancer initiation process is immediately accomplished.

In the SI (<https://stacks.iop.org/PB/18/056002/mmedia>), we show how to solve explicitly equations (2)–(4), and that allows us to obtain a comprehensive description of cancer initiation dynamics. One can define a probability to reach the fixation state starting from the state n , $\Pi_n = \int_0^\infty F_n(t) dt$, and the calculations presented in the SI show that

$$\Pi_n = \frac{1 - 1/r_1^n}{1 - 1/r_1^{2N}}, \quad (5)$$

for $1 \leq n < N$, while $\Pi_n = 1$ for $N \leq n < 2N$. This formula is a well known result in population genetics: see, e.g. reference [22]. It can be easily understood by considering the properties of the discrete-state stochastic model in figure 2. If the system is found in one of the states $1 \leq n < N$ then there is always a non-zero chance that the mutated cells will be eliminated from the system (exiting eventually to the left from the state 1 with the rate a_1). However, because the transition from the state $N - 1$ to the state N is irreversible, for the states $N \leq n < 2N$ the system can never eliminate the mutated cells and with a probability one it will reach the fixation state. But there is another surprising observation from equation (5). It suggests that the probability of cancer initiation by two mutations (starting from the state $n = 1$) depends only

on the properties of the first mutation. This is again the consequence of the irreversibility in the transition from the first branch and transient fixation of the first mutation before the system can proceed further to the final fixation of both mutations. Since the probability of cancer initiation is directly related to the cancer lifetime risk [14, 29], this immediately shows that the order of mutations does matter. The results for the fixation probability also simplify in several limiting cases. When $r_1 \rightarrow 1$, we obtain $\Pi_n = n/2N$, while for the tissues with very large number of cells and advantageous first mutation ($N \gg 1$ and $r_1 > 1$) we have $\Pi_n = 1 - 1/r_1^n$.

Another important characteristics of the cancer initiation process is the cancer initiation time. It is defined as a time for the tissue to reach the two-mutations fixation for the first time. In our theoretical framework, it also corresponds to the mean first-passage time to reach the state $2N$ starting from the state n . It is generally defined as $T_n = \frac{\int_0^\infty t F_n(t) dt}{\Pi_n}$. As shown explicitly in the SI, the cancer initiation time starting from the state $n = 1$ can be written as

$$\begin{aligned} T_1 = & \frac{N + 1}{b} \sum_{n=1}^{N-1} \frac{1}{n(N-n)} \left(\frac{r_1^n - 1}{r_1 - 1} \right) \left(\frac{r_1^{N-n} - 1}{r_1^N - 1} \right) \\ & + \frac{1}{Nu_2} \left[\frac{1 - (r_1/r_2)^N}{1 - r_1/r_2} \right] \\ & + \frac{N + 1}{br_2} \sum_{n=1}^{N-1} \frac{1}{n(N-n)} \left(\frac{1 - (r_1/r_2)^{N-n}}{1 - r_1/r_2} \right). \end{aligned} \quad (6)$$

There are three contributions to the average fixation time of two mutations in equation (6), i.e. $T_1 = T_{11} + T_{\text{trans}} + T_{12}$. The first term T_{11} describes the time for the system to reach the state N , which corresponds to the fixation of the first mutation. The second term T_{trans} describes the effective rate of acquiring the second mutation in cells that are fully fixed by the first mutation. It can be shown that

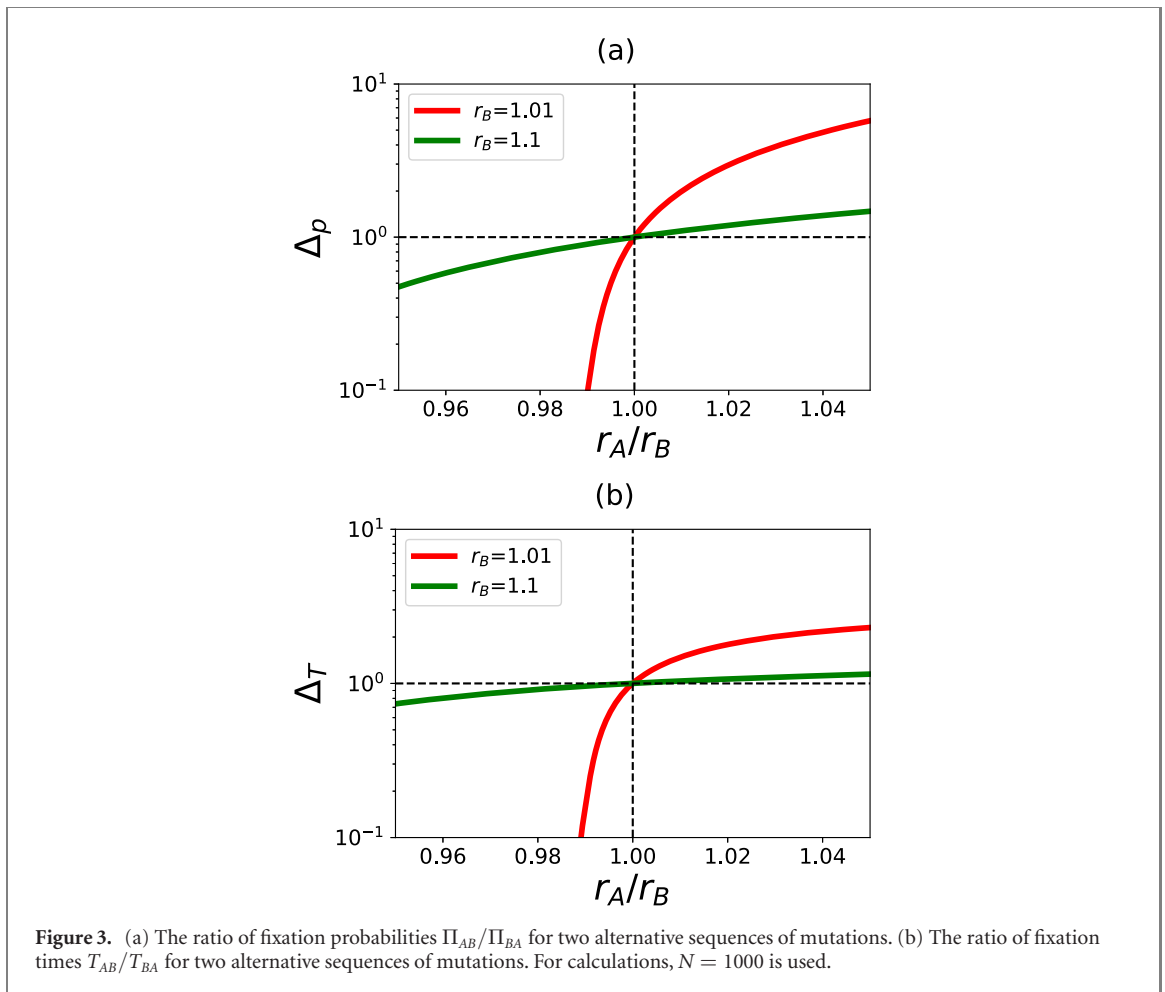
$$T_{\text{trans}} = \frac{1}{Nu_2} \frac{1}{\Pi_{12}}, \quad \Pi_{12} = \frac{1 - r_1/r_2}{1 - (r_1/r_2)^N}, \quad (7)$$

where Π_{12} is the fixation probability of the second mutation starting from the state $N + 1$. Thus, this contribution to the overall fixation time reflects the possibility of reverse transitions from the state $N + 1$ back to the state N . The third term, T_{12} describes the time to reach the final fixation starting from the state $N + 1$.

Again, it is interesting to consider limiting cases. For $r_1 = r_2 = 1$, i.e. for two successive neutral mutations and $N \rightarrow \infty$, we obtain,

$$T_1 \simeq \frac{N}{b} + \frac{1}{bu_2} + \frac{N}{b} \ln N \simeq \frac{N}{b} \left[1 + \frac{1}{Nu_2} + \ln N \right]. \quad (8)$$

This results suggests that in this limit the system spends most of the time on the second branch of



discrete states. Also, for large N equation (6) can be simplified into (see the SI):

$$T_1 \simeq \frac{2(1 + r_1^N) \ln N + \text{Ei}(\ln r_1) + 2r_1^N \text{Ei}(-\ln r_1)}{b(r_1 - 1)(r_1^N - 1)} + \frac{1}{Nu_2} \left[\frac{1 - (r_1/r_2)^N}{1 - r_1/r_2} \right] + \frac{2 \ln N + \left(\frac{r_1}{r_2}\right)^N \text{Ei}\left(-\ln \frac{r_1}{r_2}\right) + \text{Ei}\left(\ln \frac{r_1}{r_2}\right)}{br_2(1 - r_1/r_2)} \quad (9)$$

,where $\text{Ei}(x)$ is the exponential integral defined as $\text{Ei}(x) \equiv -\int_{-x}^{\infty} \frac{e^{-z}}{z} dz$, and γ is the Euler–Mascheroni constant. For $r_1 < r_2$, terms multiplied by $(r_1/r_2)^N$ vanish and consequently the fixation time is further simplified,

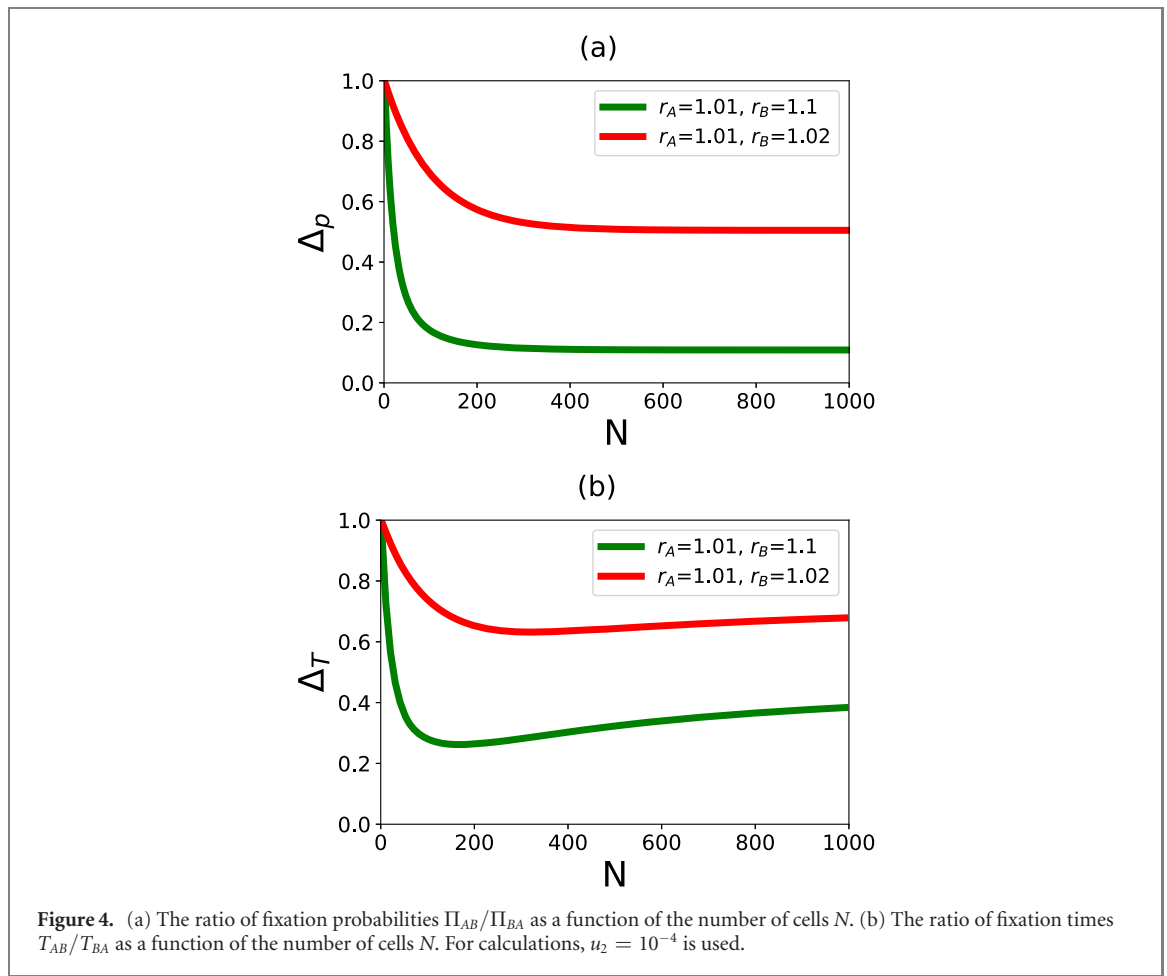
$$T_1 \simeq \frac{2(1 + r_1^N) \ln N + \text{Ei}(\ln r_1) + 2r_1^N \text{Ei}(-\ln r_1)}{b(r_1 - 1)(r_1^N - 1)} + \frac{1}{Nu_2} \left[\frac{1}{1 - r_1/r_2} \right] + \frac{1}{br_2(1 - r_1/r_2)} \left[2 \ln N + \text{Ei}\left(\ln \frac{r_1}{r_2}\right) \right]. \quad (10)$$

Now we can analyze the effect of the mutations order by considering two specific mutations A and B with the fitness parameters given by r_A and r_B , respectively. There are two different two-hit mutations in

this system, labeled as AB and BA . For the events of type AB the mutation A is the first one, while the mutation B is the second one. Then we have $r_1 = r_A$ while $r_2 = \varepsilon r_A r_B$. Here, the combined fitness parameter r_2 reflects the presence of both mutations and the parameter ε describes the cooperativity between the mutations. When $\varepsilon > 1$ the presence of both mutations is more advantageous and this is a positive cooperativity, while for $\varepsilon < 1$ we have a negative cooperativity when the mutations counterbalance each other. To simplify our calculations, we assume $\varepsilon = 1$, which corresponds to a neutral cooperation, i.e. when the effect of both mutations is independent of each other. But note that the cooperativity effects can be explicitly considered in our theoretical approach. The situation is opposite for events of type BA . Here we have $r_1 = r_B$ and $r_2 = r_A r_B$. These two mutational sequences have different probabilities of cancer initiation defined as Π_{AB} and Π_{BA} , respectively. To quantify the difference we define a ratio,

$$\Delta_p = \frac{\Pi_{AB}}{\Pi_{BA}} = \left(\frac{1 - 1/r_A}{1 - 1/r_B} \right) \left(\frac{1 - 1/r_B^N}{1 - 1/r_A^N} \right). \quad (11)$$

Figure 3(a) shows the difference in the cancer initiation probabilities for different mutations sequences as a function of the relative fitness parameters. One



can see that when the first mutation is more advantageous ($r_A/r_B > 1$) the AB fixation probability is higher than the BA fixation probability. The effect is stronger when the fitness parameters are closer to being equal to one. This observation can be easily explained by noting again that the overall fixation probability depends only on the nature of the first mutation. Thus, if the first mutation has a larger fitness parameter for one sequence it will lead to a higher probability of cancer initiation.

Our theoretical framework also allows us to evaluate the dynamics of cancer initiation. We can explicitly calculate the fixation times T_{AB} and T_{BA} for AB and BA mutation sequences, respectively. To the best of our knowledge, these quantities have never been explicitly evaluated for two-hit mutations systems. To quantify the difference in fixation times, the following function is defined,

$$\Delta_T = \frac{T_{AB}}{T_{BA}}, \quad (12)$$

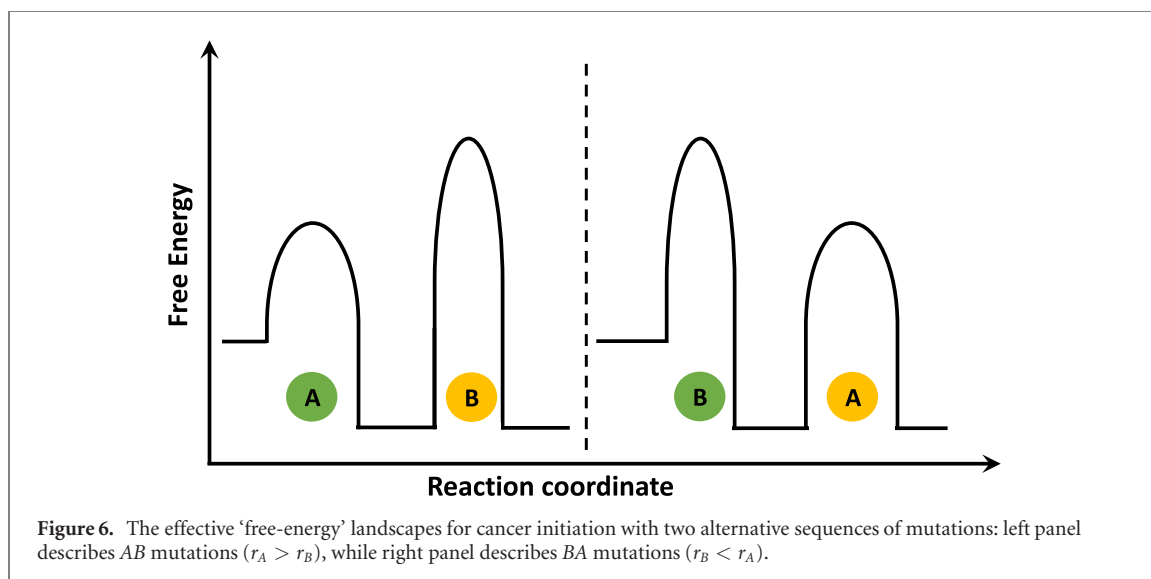
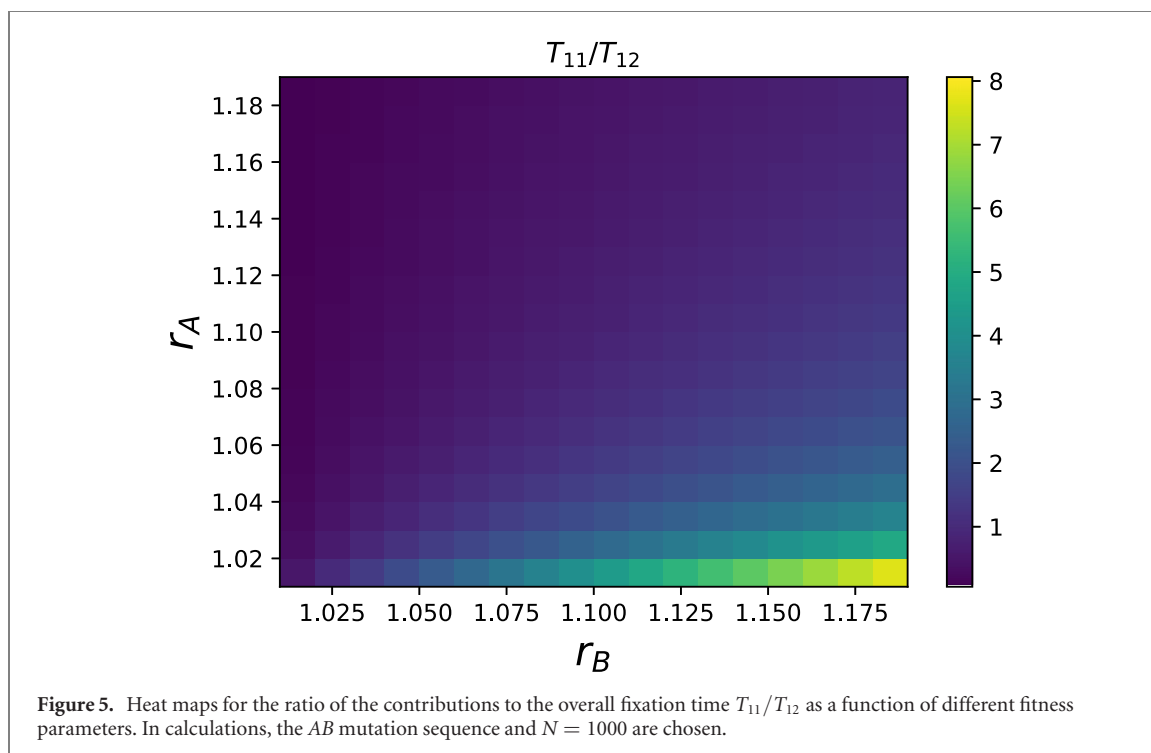
and the results are presented in figure 3(b). It shows that there is a different cancer initiation dynamics depending on the mutations order, and the effect is stronger when fitness parameters are close to unity. However, the unexpected result is that if the first mutation is more advantageous ($r_A/r_B > 1$) it takes longer to initiate the cancer. This contrasts with the higher probability of the overall fixation for this

situation, i.e. something that is more probable takes longer to achieve, which is opposite to naive expectations.

Note that, although in our analysis we utilized for calculations $N = 10^3$, a larger number of stem cells is more realistic ($N \simeq 10^5 - 10^9$) [30, 31]. This choice of parameters for N , however, does not have any qualitative effects on our conclusions, as long as N is quite large. Thus, for simplicity we employ $N = 10^3$ since the results of calculations effectively do not change for larger values of N . One could also see this in figure 4, which is also explained below.

Another relevant parameter in the cancer initiation is the number of stem cells. Figure 4 shows how varying N influences the differences in the cancer initiation probabilities and times. The differences are relatively small for low number of cells, but they start to increase and eventually saturate at $N \gg 1$. The differences are stronger for the fitness parameters that are closer to unity. This can be easily explained for the fixation probabilities because $\Pi_1 = 1 - 1/r_1$ in the limit of $N \rightarrow \infty$. In this limit, from equation (10) the fixation time can be approximated as $T_1 \sim \ln N$ with the coefficient that depends only the fitness parameters r_1 and r_2 , and this leads to the constant ratio of the fixation times for different mutation sequences.

To understand the microscopic origin of the surprising observation that if the first mutation is more



advantageous it leads simultaneously to higher probability of cancer initiation and to slower cancer initiation dynamics, we plot in figure 5 the relative contributions of the first and third terms in the fixation times, T_{11}/T_{12} for different fitness parameters. These times correspond to the system being found on the first branch when only wild-type and cells with one mutation are present (T_{11} , $1 \leq n < N$); and for being on the second branch when only cells with one or two mutations are present (T_{12} , $N + 1 \leq n < 2N$). For convenience, let us consider the AB mutation sequence. When the first mutation has a higher fitness parameter than the second mutation (figure 5, upper left corner), it is found that $T_{11}/T_{12} < 1$, i.e. the system spends most of the time on the second branch of discrete states. But if the second mutation is more

advantageous (figure 5, lower right corner) the situation is reversed, and the system spends most of the time on the first branch of discrete states.

Based on these observations, we can construct effective qualitative ‘free-energy’ landscapes for cancer initiation dynamics driven by stochastic transitions. A similar but more quantitative energy landscape was recently utilized to describe the attractor states in the cancer gene network state space as distinct biological functional states (normal, cancer and apoptosis states) [32]. One can associate longer times with higher effective barriers and shorter times with smaller effective barriers. The idea here is that it takes longer for a stochastic ‘particle’ to overcome larger kinetic barriers. This leads to the schematic picture

shown in figure 6 for alternative sequences of mutations. The first barrier corresponds to the states on the first branch, and the second barrier corresponds to the states on the second branch. The deep region between two barriers reflect the irreversible transition between two branches of states. Now we can understand better the difference in cancer initiation for two sequences of mutations. The probability of cancer initiation is determined only by the first barrier, and then it is clear that if the first mutation is more advantageous ($r_A > r_B$, left panel on figure 6) the cancer lifetime risk is higher. But in the opposite case ($r_B < r_A$, right panel in figure 6), the cancer lifetime risk is lower. At the same, the cancer initiation dynamics depends on both barriers, although the highest barrier dominates the behavior as a rate-limiting step. It is reasonable to assume that the highest barrier will lead to the slowest dynamics, and this is described by the situation on the left panel of figure 6 (*AB* sequence). This is because the second barrier (distance between the minimum and the maximum in the effective free-energy landscape) is higher. For the case on the right panel in figure 6 (sequence *BA*), while the first transition is relatively slow it is not as slow as for the *AB* sequence (barrier is lower). These arguments explain the effect of temporal order of mutations in cancer initiation.

One should also note that such anti-correlations between the probabilities of events and dynamics have been observed in various stochastic processes. The most striking example is chemical reactions. While the probabilities of reactions is determined by the free-energy differences between the products and reagents, there are multiple chemical reactions that are not happening due to high kinetic barriers despite the fact the reactions are very probable from the free-energy difference point of view.

2. Summary and discussion

In conclusion, we developed a theoretical model to quantitatively explain the effect of temporal order of mutations in cancer initiation. The appearance of tumor is associated with fixation of several mutations in the tissue compartment that originally had a fixed number of normal stem cells. Our idea is that the cancer initiation process can be viewed as a sequence of stochastic transitions between discrete states that are defined by different numbers of cells with one or two mutations. This allows us to evaluate properties of cancer initiation and compare different mutations sequences. Specifically, we analyzed in detail the cancer initiation dynamics after the fixation of two different sequential mutations. It is shown that the probability of cancer initiation is fully determined by the fixation of the first mutation. This also suggests that the probability of the tumor formation is higher if the first mutation is more advantageous than the second one. These theoretical results explain recent experimental observations that emphasize the

special role of initial truncating mutations in the human cancers [8]. We also found that cancer initiation times depend on both mutations, and the fastest dynamics is observed if the second mutation is more advantageous. These theoretical predictions are able to explain striking observations that the reverse order of mutations delays the formation of tumor by many years or might not even lead to cancer at all [6]. One of the most surprising observations is the apparent anti-correlation between the probability of cancer initiation and the speed of tumor formation, which contrasts with naive expectations. To explain this result, we utilized the analogy of effective free-energy landscapes for cancer initiation dynamics that allowed us to clarify the mechanistic origin of this phenomenon. The analogy with chemical reactions is also suggested as a possible way to understand the underlying microscopic processes.

While the presented theoretical model provides a description of the effect of mutations order in the cancer initiation, it is crucial to discuss its limitations. It assumes that the second mutation does not start until the first one is fully fixed. This is known as a linear evolution in tumor formation [33]. However, there is a limited amount of experimental data that support the linear evolution in cancers [19, 33]. Most studies favor a branching evolution picture in which different cellular clones with different mutations evolve in parallel. At the same time, the linear evolution is a limiting case of the branched evolution when the mutation rate are very low, and there are arguments suggesting that early stages of cancer initiation can be well described by the linear evolution [33]. Another simplification is the assumption that two mutations would lead to the formation of the tumor, while current data suggest that as many as 10 mutations are needed to drive cancers. In addition, the cancer might start even if not the whole tissue is overtaken by the mutated cells. Some of these effects can be accounted in our approach and it will be a subject of future studies. However, despite these limitations, the model provides a clear physical picture, explaining the effect of the temporal order of mutations during the formation of tumors, that can be tested experimentally.

Acknowledgments

The work was supported by the Welch Foundation (C-1559), by the NSF (CHE-1953453 and MCB-1941106), and by the Center for Theoretical Biological Physics sponsored by the NSF (PHY-2019745).

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

Hamid Teimouri  <https://orcid.org/0000-0002-3319-1187>

Anatoly B Kolomeisky  <https://orcid.org/0000-0001-5677-6690>

References

- [1] Weinberg R A 2013 *The Biology of Cancer* (New York: Garland Science Taylor & Francis Group LLC)
- [2] Vogelstein B, Papadopoulos N, Velculescu V E, Zhou S, Diaz L A and Kinzler K W 2013 Cancer genome landscapes *Science* **339** 1546–58
- [3] Hassler M R and Egger G 2012 Epigenomics of cancer—emerging new concepts *Biochimie* **94** 2219–30
- [4] Hanahan D and Weinberg R A 2011 Hallmarks of cancer: the next generation *Cell* **144** 646–74
- [5] Knudson A G 2001 Two genetic hits (more or less) to cancer *Nat. Rev. Cancer* **1** 157
- [6] Kent D G and Green A R 2017 Order matters: the order of somatic mutations influences cancer evolution *Cold Spring Harbor Perspect. Med.* **7** a027060
- [7] Ortmann C A *et al* 2015 Effect of mutation order on myeloproliferative neoplasms *New Engl. J. Med.* **372** 601–12
- [8] Levine A J, Jenkins N A and Copeland N G 2019 The roles of initiating truncal mutations in human cancers: the order of mutations and tumor cell type matters *Cancer Cell* **35** 10–5
- [9] Herbet M, Salomon A, Feige J-J and Thomas M 2012 Acquisition order of Ras and p53 gene alterations defines distinct adrenocortical tumor phenotypes *PLoS Genet.* **8** e1002700
- [10] Klampfl T *et al* 2013 Somatic mutations of calreticulin in myeloproliferative neoplasms *New Engl. J. Med.* **369** 2379–90
- [11] Nangalia J *et al* 2013 Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2 *New Engl. J. Med.* **369** 2391–405
- [12] Levine R L and Gilliland D G 2007 JAK-2 mutations and their relevance to myeloproliferative disease *Curr. Opin. Hematol.* **14** 43–7
- [13] Delhommeau F *et al* 2009 Mutation in TET2 in myeloid cancers *New Engl. J. Med.* **360** 2289–301
- [14] Nowak M A 2006 *Evolutionary Dynamics: Exploring the Equations of Life* (Cambridge, MA: Harvard University Press)
- [15] Frank S A 2007 *Dynamics of Cancer: Incidence, Inheritance, and Evolution* vol 1 (Princeton, NJ: Princeton University Press)
- [16] Orr H A 2005 The genetic theory of adaptation: a brief history *Nat. Rev. Genet.* **6** 119–27
- [17] Michor F, Iwasa Y and Nowak M A 2004 Dynamics of cancer progression *Nat. Rev. Cancer* **4** 197
- [18] Komarova N L, Sengupta A and Nowak M A 2003 Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability *J. Theor. Biol.* **223** 433–50
- [19] Nowak M A, Michor F and Iwasa Y 2003 The linear process of somatic evolution *Proc. Natl Acad. Sci.* **100** 14966–9
- [20] Haeno H, Maruvka Y E, Iwasa Y and Michor F 2013 Stochastic tunneling of two mutations in a population of cancer cells *PLoS One* **8** e65724
- [21] Ashcroft P, Michor F and Galla T 2015 Stochastic tunneling and metastable states during the somatic evolution of cancer *Genetics* **199** 1213–28
- [22] Durrett R 2008 *Probability Models for DNA Sequence Evolution* (Berlin: Springer)
- [23] Desai M M and Fisher D S 2007 Beneficial mutation–selection balance and the effect of linkage on positive selection *Genetics* **176** 1759–98
- [24] Lynch M 2010 Rate, molecular spectrum, and consequences of human mutation *Proc. Natl Acad. Sci.* **107** 961–8
- [25] Gillespie J H 1984 Molecular evolution over the mutational landscape *Evolution* **38** 1116–29
- [26] Teimouri H, Kochugaeva M P and Kolomeisky A B 2019 Elucidating the correlations between cancer initiation times and lifetime cancer risks *Sci. Rep.* **9** 18940
- [27] Moran P A P *et al* 1962 Oxford: Oxford University Press
- [28] Moran P A P 1958 Random processes in genetics *Math. Proc. Camb. Phil. Soc.* **54** 60–71
- [29] Nowak M A and Waclaw B 2017 Genes, environment, and ‘bad luck’ *Science* **355** 1266–7
- [30] Lee-Six H *et al* 2018 Population dynamics of normal human blood inferred from somatic mutations *Nature* **561** 473–8
- [31] Tomasetti C and Vogelstein B 2015 Variation in cancer risk among tissues can be explained by the number of stem cell divisions *Science* **347** 78–81
- [32] Li C and Wang J 2014 Quantifying the underlying landscape and paths of cancer *J. R. Soc. Interface* **11** 20140774
- [33] Davis A, Gao R and Navin N 2017 Tumor evolution: linear, branching, neutral or punctuated? *Biochim. Biophys. Acta* **1867** 151–61