



Mutation–selection networks of cancer initiation: tumor suppressor genes and chromosomal instability

Natalia L. Komarova^{a,b,*}, Anirvan Sengupta^c, Martin A. Nowak^a

^a*Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

^b*Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK*

^c*Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA*

Received 10 May 2002; received in revised form 26 February 2003; accepted 5 March 2003

Abstract

In this paper, we derive analytic solutions of stochastic mutation–selection networks that describe early events of cancer formation. A main assumption is that cancer is initiated in tissue compartments, where only a relatively small number of cells are at risk of mutating into cells that escape from homeostatic regulation. In this case, the evolutionary dynamics can be approximated by a low-dimensional stochastic process with a linear Kolmogorov forward equation that can be solved analytically. Most of the time, the cell population is homogeneous with respect to relevant mutations. Occasionally, such homogeneous states are connected by ‘stochastic tunnels’. We give a precise analysis of the existence of tunnels and calculate the rate of tunneling. Finally, we calculate the conditions for chromosomal instability (CIN) to precede inactivation of the first tumor suppressor gene. In this case, CIN is an early event and a driving force of cancer progression. The techniques developed in this paper can be used to study arbitrarily complex mutation–selection networks of the somatic evolution of cancer.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Genetic instability; Colon cancer; Stochastic modeling

1. Introduction

Cancer is an evolutionary disease. Multicellular organisms have intricate signal transduction pathways and gene regulation networks to ensure that cells cooperate and only divide when needed. In the process of cancer formation, individual cells alter such networks and revert to an earlier, more primitive evolutionary program of ‘selfish’ reproduction.

Cancers are caused by somatic or inherited mutations in proto-oncogenes, tumor suppressor genes or DNA repair genes. Tumor suppressor genes participate in growth regulatory or differentiation pathways of cells. Loss of their function changes the phenotype of the cell toward becoming a cancer cell. Typically, both copies of a tumor suppressor gene need to be inactivated. Examples of tumor suppressor genes are p53, which is mutated in more than 50% of human cancers, RB1,

the retinoblastoma susceptibility gene, BRCA1 and BRCA2, which are involved in familial breast cancer (Knudson, 2001). Sporadic colorectal cancer seems to start with the inactivation of a tumor suppressor gene called APC (Vogelstein et al., 1988; Nishisho et al., 1991).

There has been a long standing debate whether the somatic evolution of cancer requires cancer cells to have genetic instability (Loeb, 1998; Lengauer et al., 1998; Schutte and Fishel, 1999; Strauss, 1998). Two forms of genetic instability have been found in colorectal cancers. In approximately 87% of colorectal cancers there are large numbers of gross chromosomal changes, which include loss of heterozygosity (LOH), imbalance of chromosome numbers (aneuploidy), chromosomal translocations and gene amplifications (Lengauer et al., 1997; Abdel-Rahman et al., 2001). It is believed that these modifications are the consequence of chromosomal instability (CIN).

The molecular basis for CIN is just beginning to be explored. A large number of gene alterations can give rise to CIN in yeast. The genes include those involved in

*Corresponding author. Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA. Fax: +1-609-951-4438.

E-mail address: natalia@ias.edu (N.L. Komarova).

chromosome condensation, sister chromatid cohesion, kineotcho structure and function, microtubule formation and dynamics, as well as checkpoints that monitor the progress of the cell cycle. To date, the only genes implicated in aneuploidy in human cancer cells are those of the latter class. Heterozygous mutations in the mitotic spindle checkpoint gene hBUB1 were detected in a small fraction of colorectal cancers with the CIN phenotype (Cahill et al., 1998). These mutations in hBUB1 were shown to be dominant negative (Taylor and McKeon, 1997). These results also confirmed cell fusion studies that indicate that the CIN phenotype has a dominant quality, and that it might only require a single mutational “hit” to induce CIN (Lengauer et al., 1997). A current estimate is that our genome contains of the order of 100 genes with the property that mutations in these genes lead to CIN.

In the remaining 13% of colorectal cancers there is no evidence of gross chromosomal changes; there are two copies of each chromosome and there are no signs of chromosome translocations or gene amplifications. In these cells, however, there are mutations in genes that are involved in mismatch repair. This mismatch repair deficiency leads to an increased point mutation rate during cell division. The genomes of such cancer cells display microsatellite instability (MIN) (Markowitz, 2000; Perucho, 1996).

While MIN appears to be a specific property of colon cancers and is rare or absent in most other cancers, CIN emerges more and more as a common feature of almost all cancers. The crucial question is whether CIN is a late stage consequence of cancer evolution or an early event. If CIN was an early event then it can be seen as a driving force of tumor progression. There is indirect evidence that CIN arises early in colon cancer formation. Aneuploidy can be found in early neoplastic lesions, such as benign adenomas of the colon (Bardi et al., 1997; Bomme et al., 1998). Allelic imbalance is detected in the smallest lesions that can be investigated (Shih et al., 2001). In this paper, we will study the question if CIN is likely to emerge before inactivation of the first tumor suppressor gene on the route to cancer.

Our model is motivated by colon cancer, because important genetic details have been discovered for this disease in the last two decades. Colon cancer arises in small compartments of cells called crypts. Crypts contain 1000–4000 cells. At the bottom of the crypt there are 4–6 colon stem cells that divide and replenish the whole crypt (Bach et al., 2000). These cells are believed to be at risk of becoming cancer cells. Differentiated cells that derive from these stem cells divide a few times and travel to the top of the crypt within 3–6 days (Lipkin et al., 1963; Shorter et al., 1964). They undergo apoptosis at the top of the crypt. Sporadic colorectal cancer is thought to arise with the inactivation of the APC gene (Vogelstein et al., 1988; Nishisho

et al., 1991; Lamlum et al., 2000). If both copies of APC are inactivated then the cells fail to undergo apoptosis, and the crypt is taken over by mutated cells resulting in a ‘displastic crypt’. Polyps and colon cancer develop from displastic crypts requiring further mutational events. A crucial question is whether CIN arises before inactivation of APC.

Mathematical modeling of cancer progression has a long and impressive history. We can only cite some examples here. Stochastic multi-stage models for explaining the age-dependent incidence rates of human cancers were first developed by Nordling (1953), Armitage and Doll (1954, 1957), Fisher (1958) and Cook et al. (1969). Bell (1976) models carcinogenesis as an escape from mitotic inhibitors. Moolgavkar and Knudson (1981) provide a stochastic model for Knudson’s ‘two hit hypothesis’; they extend their method to an arbitrary number of hits in their recent paper (Luebeck and Moolgavkar, 2002). Goldie and Coldman (1979, 1985) describe models of drug resistance to chemotherapy. Wheldon (1988) offers an excellent perspective of mathematical models in cancer research up to 1988. Chaplain (1995) studies the dynamics of angiogenesis. Byrne and Chaplain (1996) model cell adhesion and tumor growth. Owen and Sherratt (1999) study immune responses against tumors. Herrero-Jimenez et al. (2000) combine a model for carcinogenesis with an epidemiological model of colon cancer. Bar-Or et al. (2000) study oscillation in a gene regulation network containing p53. Yatabe et al. (2001) develop a mathematical model to analyse turnover rates of colon stem cells using methylation patterns. Wodarz and Krakauer (2001) study genetic instability and angiogenesis. Plotkin and Nowak (2002) explore a stochastic process of apoptosis and genetic instability.

In the following we will study stochastic dynamics of mutation–selection networks that deal with the emergence of CIN and inactivation of a tumor suppressor gene. In Section 2, we analyse a simple, preliminary model that has only 2 types of cells, “A” and “B”. We investigate the mutation of a particular gene (not necessarily a tumor suppressor gene). In “A” cells, the gene is not mutated, in “B” cells, one copy of the gene is mutated. We assume that mutation happens with rate u and is irreversible. “A” cells have a reproductive rate 1, while B cells have a reproductive rate r which can be greater, smaller or equal to 1. If the system has a constant (effective) population size, N , then we have a one-dimensional stochastic process; a particular state is given by the number of cells of type “A”. Analytic time-dependent solutions are very complicated if $r \neq 1$. If, however, the mutation rate u is much smaller than a certain value (roughly given by $1/N$), then for most times the system will contain only “A” or only “B” cells; we will say that the system spends most of the time in *homogeneous* states. The probability that the system is in

the all “B” state at time t is given by $1 - \exp(-Nu\rho t)$, where ρ is the probability that a single B cell will be fixed in a population of $N - 1$ “A” cells.

In Section 3, we study a mutation–selection network containing three types of cells. This network will appear in various forms in the more complex mutation–selection networks of the following sections (and subsequent papers). “A” cells mutate into “B” cells at rate u , “B” cells mutate to “C” cells at rate u_1 . “A” cells have reproductive rate 1, “B” cells have reproductive rate r , and “C” cells have reproductive rate r_1 . This model describes, for example, the inactivation of a tumor suppressor gene. In this case, “A” cells have 2 intact copies of the gene, “B” cells have 1 intact copy, while “C” cells have both copies inactivated. The reproductive rates are $r = 1$ and $r_1 > 1$; hence inactivating both copies of a tumor suppressor gene confers a selective advantage. The mutation rate u describes the probability of inactivating one of the 2 copies of the tumor suppressor gene during a cell division. The mutation rate u_1 describes the probability of inactivating the other copy: this could be caused by an independent point mutation event or by loss of heterozygosity (LOH). Again we provide an analytic description of the simplified stochastic process assuming that at any one time the system contains either only cells of type “A” or only cells of type “B” or only cells of type “C”. As before, this approximation holds whenever mutation rates are small. An interesting observations are stochastic tunnels: for certain choices of parameter values the system proceeds from all “A” to all “C” without visiting all “B”.

In Section 4, we apply the analytic machinery developed in the previous sections to investigate the question if CIN arises before inactivation of the first tumor suppressor gene. The network contains 6 types of cells. As before there are cells that have 2,1 or 0 functioning copies of the tumor suppressor gene. Moreover, all cell types can exist with or without CIN. The main assumption is that CIN enhances the rate of LOH. Hence, in CIN cells the 2nd copy of a tumor suppressor gene is lost rapidly by LOH. It turns out there are 8 cases that can be classified into 5 mutation–selection networks (depending on which tunnels are open or not). Surprisingly, in the end, there are simple conditions for CIN to precede inactivation of the tumor suppressor gene and hence to initiate cancer formation. Our model is certainly motivated by colon cancer, but should, more generally, apply to the interaction between inactivation of a tumor suppressor gene and CIN in any cancer.

Section 5 contains a summary of our findings.

2. Two types of cells

Let us first assume that there are two types of cells in a population, which we will call type “A” and type “B”.

Cells can reproduce, mutate and die. The probability that a cell of type “A” reproduces faithfully is $1 - u$; with probability u it will mutate to type “B”. Cells of type “B” always reproduce faithfully. We will assume that the total number of cells is constant and equal to N . Let the cells of type “A” have reproductive rate 1 and the cells of type “B”—reproductive rate r .

We will use the following convenient short-hand representation of these processes:



Here the reproductive rate of each type is given in parentheses and the mutation rate is marked above the arrow. We will refer to such diagrams as *mutation–selection networks*.

One can envisage the following birth–death process. At each time step, one cell reproduces, and one cell dies. We set the length of each time step to be $1/N$, so that during a unit time interval, N cells are chosen for reproduction and N cells die. We assume that all cells have an equal chance to die (this is equal to $1/N$). On the other hand, reproduction happens differentially depending on the type, and the relative probability of being chosen for reproduction is given by 1 and r for the cells of types “A” and “B” respectively. Obviously, in this setting the total number of cells is preserved.

Let us denote the number of cells of type “A” as a , and the number of cells of type “B” as b , so that $a + b = N$. The probability that a cell of type “A” reproduces is proportional to its frequency and the reproductive rate, and is given by $a/(a + rb)$. Similarly, the probability that a cell of type “B” reproduces is $rb/(a + rb)$. Thus the probability that the new cell is of type “A” or type “B” is given, respectively, by

$$P_{+A} = (1 - u) \frac{a}{a + rb}, \quad P_{+B} = u \frac{a}{a + rb} + \frac{rb}{a + rb}.$$

Cells of both types have a probability to die proportional to their abundance, i.e. the probability that a cell of type “A” (or “B”) dies is given, respectively, by

$$P_{-A} = \frac{a}{N}, \quad P_{-B} = \frac{b}{N}.$$

We will refer to an event consisting of one replication and one cell death by an *elementary event*.

The resulting population dynamics is a Markov process with states $b = 0, 1, \dots, N$, and time steps of length $1/N$. The probability that an elementary event results in an increase of the number of cells of type “B”, is equal to $P_{+B}P_{-A}$, and the probability that the number of cells of type “B” decreases is equal to $P_{-B}P_{+A}$. If P_{ij} is the probability to go to state $b = j$ from state $b = i$,

then the transition matrix is given by

$$P_{ij} = \begin{cases} \frac{u(N-i) + ri}{\mathcal{N}_i} \frac{N-i}{N}, & j = i + 1, \\ \frac{(1-u)(N-i)}{\mathcal{N}_i} \frac{i}{N} & j = i - 1, \\ 1 - P_{i,i+1} - P_{i,i-1}, & j = i, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where $0 \leq i, j \leq N$, and we introduced the notation

$$\mathcal{N}_i = N - i(1 - r).$$

The corresponding Markov process is a biased random walk with one absorbing state, $b = N$. Let us set the initial condition to be $b = 0$ (all cells are of type ‘‘A’’) and study the dynamics of absorption into the state $b = N$.

2.1. Absorption time

The most straightforward method would be to calculate the time of absorption directly. If we denote the number of elementary events until absorption starting from state i as t_i , we have

$$t_i = N + \sum_{m=0}^{N-1} P_{im} t_m, \quad 0 \leq i \leq N - 1. \quad (3)$$

The absorption time is then given by $T_{abs} \equiv t_0$. The factor N appears in the equation above because of our definition of a unit time: $T = 1$ corresponds to N elementary events. Solving system (3) directly is cumbersome, so we will use some approximations.

There are two processes that go on in the system: mutation and selection. If the characteristic time scales of the two processes are vastly different, our task of finding the absorption time simplifies greatly. Let us assume that u is very small, so that once a mutant of type ‘‘B’’ is produced, it typically has time to get fixated or die out before a new mutation occurs. In other words, once a mutant is produced, it is safe to assume that during its lifetime no other mutations occur. In this case of rare mutations, the inverse time to absorption is roughly u times the probability to get absorbed in the state $b = N$ from the state $b = N$ assuming $u = 0$.

For $u = 0$, the system has two absorbing states, $b = 0$ and N . Let us denote the probability to get absorbed in $b = N$ starting from the state $b = i$ as π_i . Then we have approximately

$$\frac{1}{T_{abs}} = Nu\pi_1, \quad (4)$$

where the quantity π_1 is given by the system

$$\pi_i = P_{iN} + \sum_{m=1}^{N-1} P_{im}\pi_m, \quad (5)$$

note that we set $u = 0$ in the expression for P . System (5) can be rewritten as

$$-\pi_{i-1} + (r + 1)\pi_i - r\pi_{i+1} = 0, \quad 1 < i < N - 1,$$

where we canceled the common multiplier in the terms of the matrix $I - P$ in the same row. The boundary conditions are

$$\begin{aligned} (r + 1)\pi_1 - r\pi_2 &= 0, \\ -\pi_{N-2} + (r + 1)\pi_{N-1} &= r. \end{aligned}$$

We can look for a solution in the form $\pi_i = \alpha^i$. The quadratic equation for α gives the roots $\alpha = 1/r$ and $\alpha = 1$. Substituting $\pi_i = Ar^{-i} + B$ into the boundary conditions we obtain the solution

$$\pi_i = \frac{r^{N-i}(1 - r^i)}{1 - r^N}. \quad (6)$$

Let us reserve the notation ρ for the quantity π_1 :

$$\rho \equiv \pi_1 = \frac{r^{N-1}(1 - r)}{1 - r^N}. \quad (7)$$

We have from Eq. (4):

$$\frac{1}{T_{abs}} = Nu\rho. \quad (8)$$

The same result is obtained if we solve system (3) explicitly and then take the first term in the Taylor expansion of t_0 in u .

In order for approximation (4) to be valid, we need to make sure that the time-scale related to mutation $((Nu)^{-1})$ is much longer than the time-scale of the fixation/extinction processes. Only the fraction ρ of all mutants will successfully reach fixation, whereas the rest will be quickly driven to extinction. In order for each mutant lineage to be treated independently, we need to require that time it takes to produce a *successful* mutant, $(\rho Nu)^{-1}$, is much larger than the typical time-scale of fixation, μ_0^{-1} . The value μ_0^{-1} is calculated in Appendix A. We have the general expression,

$$u \ll \frac{\mu_0}{\rho N}.$$

In the case of neutral mutations, $\rho = 1/N$, $\mu_0 = 1/N$ and we arrive to the familiar condition

$$u \ll \frac{1}{N} \quad \text{if } |1 - r| \ll \frac{1}{N} \quad (9)$$

In the case where the mutation is positively or negatively selected, we have

$$u \ll (r^{N-1}N)^{-1} \quad \text{if } r < 1, \quad \frac{1}{N} \ll |1 - r| \ll 1, \quad (10)$$

$$u \ll r/N \quad \text{if } r > 1, \quad \frac{1}{N} \ll |1 - r| \ll 1. \quad (11)$$

2.2. The approximation of ‘‘almost absorbing’’ states

We will call a state of the system *homogeneous*, or pure, if all the N cells are of the same type. In the

two-species model, these are the states $b = 0$ and N . States containing more than one type of cells ($0 < b < N$) will be referred to as *heterogeneous*, or mixed states.

Since the mutation rate is very low relative to the absorption processes in the system (conditions (9)–(11)), the probability of finding the system in a heterogeneous state is very low. More precisely, the probability of finding the state with b cells of species “B” is of the order u for $1 < b < N$. The system spends most of the time in the states $b = 0$ and $b = N$. This allows us to make a further approximation of “almost absorbing” states.

Let us use the capital letters A and B for the probability to find the system in the state $b = 0$ and $b = N$, respectively. Strictly speaking, the state $b = 0$ is not absorbing, but it is long lived. We have approximately, $A + B = 1$. Let us define the following “coarse-grained”, continuous time stochastic process: the system jumps between two states, $A = 0$ and $A = 1$, with the following probabilities:

$$P(A = 0, t + \Delta t | A = 1, t) = u\rho\Delta t,$$

$$P(A = 0, t + \Delta t | A = 0, t) = 1,$$

$$P(A = 1, t + \Delta t | A = 1, t) = 1 - u\rho\Delta t,$$

$$P(A = 1, t + \Delta t | A = 0, t) = 0.$$

The Kolmogorov forward equations for this simple system can be written down, which describe the dynamics of the two-species model, (1):

$$\dot{A} = -uN\rho A \quad A(0) = 1, \tag{12}$$

$$\dot{B} = uN\rho A, \quad B(0) = 0, \tag{13}$$

where A is the probability to find the entire system in state “A”, B is the probability to find the entire system in state “B” and ρ is given by Eq. (6). Eqs. (12) and (13) lead to the solution $A(t) = \exp(-uN\rho t)$ and $B = 1 - \exp(-uN\rho t)$.

A short-hand notation for coarse-grained differential equations (12) and (13) is as follows:

$$A \xrightarrow{uN\rho} B.$$

We will use this notation also for describing the Kolmogorov forward equation of more complex mutation–selection networks.

3. Three types of cells

Now let us suppose that there are three types of cells: type “A”, type “B” and type “C”, and the mutation–selection network that governs the dynamics is as follows:

$$A_{(1)} \xrightarrow{u} B_{(r)} \xrightarrow{u_1} C_{(r_1)}. \tag{14}$$

The reproductive rates are, respectively, 1, r and r_1 . As before, the reproductive rates must be interpreted as *relative* probabilities to be chosen for reproduction,

rather than parameters defining the time-scale. We assume that type “A” can mutate into type “B” with probability u , and type “B” can mutate to type “C” with probability u_1 . There are no other mutation processes in the system.

Let us specify the states of the system by the variables a , b and c , the number of cells of species “A”, “B” and “C”, respectively. They satisfy the constraint $a + b + c = N$. We can characterize a state as a vector (b, c) . In this notation, the state we start with is $(0, 0)$, which is all “A”. The final state, which is the state of interest, is $(0, N)$, or all “C”. The question we will study is again, the time of absorption in the state $c = N$.

We are interested in the case where type “C” has a large selective advantage, i.e. $r_1 \gg (1, r)$, so that once there is one cell of type “C”, this type will invade instantaneously with probability one. Under this assumption we can use a trick which allows us to view the dynamics as a one-dimensional process. Namely, let us consider the following reduced Markov process with the independent stochastic variable b : the states $b = i$ with $0 \leq i \leq N$ correspond to $a = N - i$, $b = i$, $c = 0$, and the state $b = N + 1$ contains all states with $c > 1$. The state $b = N + 1$ is absorbing, because we assume that once a mutant of type “C” appears, then cells “C” invade, so the system cannot go back to a state with $c = 0$. The transition probabilities are given by

$$P_{ij} = \begin{cases} \frac{u(N-i) + (1-u_1)ri}{\mathcal{N}_i} \frac{N-i}{N}, & j = i + 1, \\ \frac{(1-u)(N-i)}{\mathcal{N}_i} \frac{i}{N}, & j = i - 1, \\ \frac{u_1ri}{\mathcal{N}_i}, & j = N + 1, \\ 1 - P_{i,i+1} - P_{i,i-1} - P_{i,N+1}, & j = i, \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

for $0 \leq i \leq N$, $P_{N+1,N+1} = 1$ and $P_{N+1,j} = 0$ for all $j \neq N + 1$. In some special cases, the absorption time can be found from Eq. (3), see Appendix B. However, a direct solution is not possible in the general case, and we will use the so-called *Green function method* to calculate the absorption time. This method is formally introduced in Appendix C. The calculations for the three-type system are performed in Appendix D. Here we give some intuitive considerations and present the results.

3.1. A two-step process and stochastic tunneling

Let us start from the all “A” state. If we are in the regime of homogeneous states, (conditions (9)–(11)), we can consider the lineages of each mutant of type “B” separately. Once a mutant of type “B” is created, it can either get extinct, or get fixated. A mutant of type “C” can be created before or after type “B” reaches fixation. This gives rise to two possible scenarios.

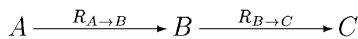
We will call a *genuine two-step process* such a sequence of steps where starting from (0, 0), after some time the system finds itself in the state (N, 0) and then gets absorbed in the state (0, N). In other words, starting from the all “A” state, the system gets to the state where the entire population consists of cells of type “B” and finally reaches fixation in the all “C” state.

We will use the term *tunneling* for such processes where the system goes from (0, 0) to (0, N) without ever visiting state (N, 0). This means that from the all “A” state the system gets absorbed in the all “C” state, skipping the intermediate fixation of type “B”.

3.2. Transition rates

3.2.1. Genuine two-step process

In the case of a genuine two-step process, the dynamics can be represented by the diagram



with

$$R_{A \rightarrow B} = Nu\rho, \quad R_{B \rightarrow C} = Nu_1.$$

Here we used the short-hand notation

$$\rho = \frac{r^{N-1}(1-r)}{1-r^N},$$

and assumed that $r_1^{N-1}(1-r_1)/(1-r_1^N) \approx 1$. The corresponding differential equations are

$$\begin{aligned} \dot{A} &= -Nu\rho A, & A(0) &= 1, \\ \dot{B} &= Nu\rho A - Nu_1 B, & B(0) &= 0. \end{aligned}$$

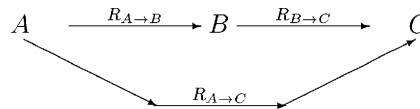
The exact conditions when the genuine two-step process is a good approximation can be given. In the case where “B” is neutral, i.e. $|1-r| \ll 1/N$, it requires the inequality $1/N \gg \sqrt{u_1}$ to hold. In the case where “B” is positively or negatively selected, $|1-r| \gg \max\{1/N, \sqrt{u_1}\}$, it requires $u_1 \ll p_c$, where we introduced

$$p_c = \begin{cases} (1-r)^2 r^{N-2}, & r < 1, \\ \frac{r-1}{rN \log \left[\frac{N(r-1)}{r} \right]}, & r > 1. \end{cases} \quad (16)$$

3.2.2. Tunneling

Having a high mutation rate, u_1 will increase the probability of tunneling. Also, in the case of a large population size, N , the fixation of type “B” becomes less probable thus making tunneling a more likely scenario. Finally, if type “B” is greatly disadvantageous, we also expect the system to tunnel from “A” to “B”.

In the general case, we have the following diagram:



The two-step rates, $R_{A \rightarrow B}$ and $R_{B \rightarrow C}$ are as before, and the tunneling rate, $R_{A \rightarrow C}$, is different depending on whether the intermediate mutant, “B”, is positively or negatively selected. We have three cases:

- *Type “B” negatively selected:* If $r < 1$ and $|1-r| \gg \sqrt{u_1}$, then we have tunneling from “A” to “C” with the rate $R_{A \rightarrow C} = \frac{Nuru_1}{1-r}$.
- *Type “B” positively selected:* If $r > 1$ and $|1-r| \gg \sqrt{u_1}$, then we have tunneling from “A” to “C” with the rate $R_{A \rightarrow C} = uu_1 N^2 \log \left[\frac{N(r-1)}{r} \right]$.
- *Type “B” neutral:* If $|1-r| \ll \sqrt{u_1}$, then we have tunneling from “A” to “C” with the rate $R_{A \rightarrow C} = Nu\sqrt{u_1}$.

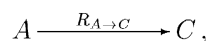
3.2.3. When is tunneling the leading process?

It can be shown that the tunneling process is faster than the two-step process if

$$\sqrt{u_1} \gg 1/N. \quad (17)$$

This condition can be derived in the following intuitive way. As with the condition for homogeneity, (9), we need to require that time it takes to produce a *successful* mutant is much larger than the typical time-scale of fixation. For condition (9), only the fraction $\rho = 1/N$ of mutants was successful. Now, because of a large advantage of the type “C”, all mutants will get fixated. This leads immediately to condition (17).

When condition (17) is satisfied and tunneling is the leading process, the dynamics can be represented as a one-step process,



and the corresponding differential equation is

$$\dot{A} = -R_{A \rightarrow C} A, \quad A(0) = 1.$$

Fig. 1 summarizes all the conditions.

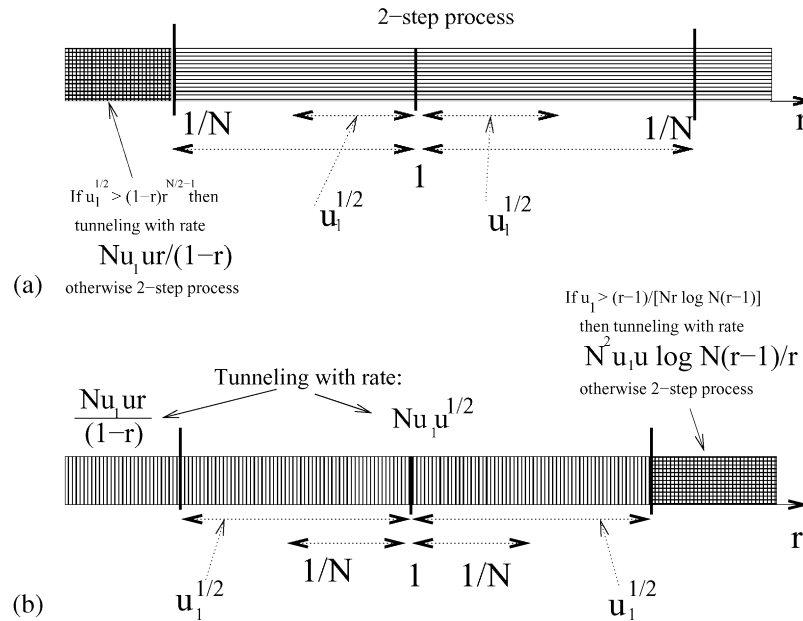


Fig. 1. Conditions for tunneling. (a) Corresponds to the case $1/N \gg \sqrt{u_1}$, (b) to $\sqrt{u_1} \gg 1/N$. Depending on r (the horizontal axis), there may be a genuine two-step process or tunneling with different rates. Horizontal shading means a two-step process, vertical shading means tunneling.

4. A tumor suppressor gene and chromosomal instability

The methods developed above can be applied to describe the processes that may lead to colon cancer.

4.1. Summary of the cellular processes leading to the formation of displastic crypts

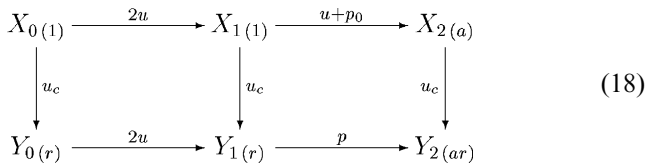
Below we introduce notations for our mathematical model and describe the mutation–selection processes which are believed to be important in the first step leading to cancer—the formation of displastic crypts.

- X_0 are cells in which both copies of the tumor suppressor gene are functional; they have reproductive rate 1.
- X_1 are cells with one functional and one inactivated copy of the tumor suppressor gene. They result from mutations of cells X_0 . The probability of such a mutation is $2u$; the factor 2 comes from the fact that either of the two copies of the tumor suppressor gene can mutate, each with probability u . The reproductive rate of X_1 cells is the same as that of X_0 cells and equals 1.
- X_2 are cells with both copies of the tumor suppressor gene inactivated. This can happen if an independent mutation occurs in the second copy of the tumor suppressor gene; the probability of this is u . Also, loss of heterozygosity (LOH) might occur, where the chromosome containing the second copy of the tumor suppressor gene is deleted. The probability of such an event is given by p_0 . Therefore, the mutation

rate from X_1 to X_2 is $u + p_0$. The quantity p_0 may be smaller, of the order of, or larger than u . The resulting cells have a large selective advantage, a , compared to the types X_0 and X_1 .

- Y_0 are cells with chromosomal instability (CIN). The mutation rate from X_0 to Y_0 is u_c . Cells with CIN can have a selective disadvantage compared with types X_0 and X_1 , because some LOH events will lead to non-viable genomic modifications, in which case we can assume that their reproductive rate is $r < 1$. On the other hand, there may be circumstances under which CIN cells have an advantage (i.e. the presence of carcinogens), so r can also be greater than 1.
- Y_1 are CIN cells with one functional and one inactivated copy of the tumor suppressor gene. The mutation rate from Y_0 to Y_1 is $2u$, like in the case of mutation from X_0 to X_1 . In addition, cells of the type Y_1 can appear as a result of a mutation of X_1 cells, with mutation rate u_c . Y_1 cells have the reproductive rate r , the same as Y_0 cells. Cells with CIN have a large chance of losing a chromosome (LOH). We assume that the rate at which they lose the chromosome with the remaining tumor suppressor gene is $p \gg p_0$. The resulting cell type is Y_2 .
- Y_2 are CIN cells with both copies of the tumor suppressor gene inactivated. They result from cells Y_1 by means of chromosome deletion. In addition, such cells can originate from cells of type X_2 with the mutation rate u_c . These cells combine the disadvantage of having CIN with the advantage of not having the tumor suppressor gene, so their reproductive rate is taken to be $ar > 1$.

The following mutation–selection network summarizes the processes which lead to the formation of displastic crypts with or without a CIN mutation:



Here the mutation rate is marked by the arrows and the reproductive rate of each type is given in brackets.

The mutation rate, u_c , can be further expressed in terms of n_c , the total number of genes that, if inactivated, lead to chromosomal instability. We have

$$u_c = 2n_c u. \tag{19}$$

Remark. As before, we assume that for all types of cells, reproduction and death happen with the same rate (say, N divisions and N deaths every two days), and consequently the generation turnover always happens with the same pace. This is an assumption that can be easily removed; we prefer to use it because it simplifies calculations.

4.2. Different pathways of inactivating a tumor suppressor gene

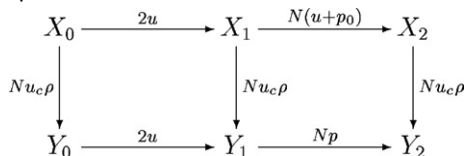
Now let us describe the dynamics specified by mutation–selection diagram (18). We will use the methods developed in the previous sections to reduce the stochastic process to a Markov process with a few almost absorbing states. In order to identify which of the homogeneous states are long lived (almost absorbing) we need to check what parameter range we are in. In principle, there can be three tunnels in the system:

- from X_0 to X_2 avoiding X_1 ,
- from Y_0 to Y_2 avoiding Y_1 and
- from X_1 to Y_2 avoiding Y_1 .

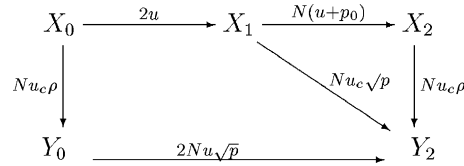
Let us write down the differential equations corresponding to processes in diagram (18).

First, we assume that $\sqrt{u+p_0} \ll 1/N$. This means that there is no tunnel from X_0 to X_2 , but the other two tunnels may or may not be open. Depending on whether the tunneling phenomenon or a genuine two-step process takes place, we can have the following seven cases:

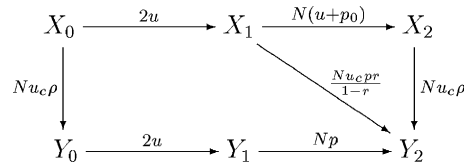
- (ia) $\sqrt{p} \ll 1/N$ and $|1-r| \ll 1/N$: no tunnels,



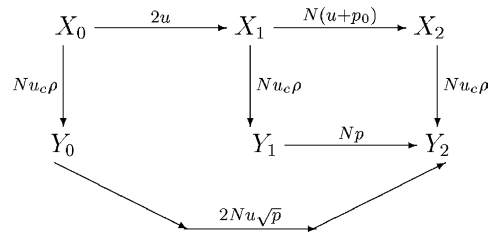
- (ib) $\sqrt{p} \ll 1/N$ and $|1-r| \gg 1/N$ and $p \ll p_c$, where p_c is defined in Eq. (16): tunnels, see the diagram in (ia);
- (iia) $\sqrt{p} \gg 1/N$ and $|1-r| \ll \sqrt{p}$: tunnels from Y_0 to Y_2 with rate $2Nu\sqrt{p}$ and from X_1 to Y_2 with rate $Nu_c\sqrt{p}$,



- (iib) $r < 1$ and $\sqrt{p} \gg 1/N$ and $|1-r| \gg \sqrt{p}$: tunnels from Y_0 to Y_2 with rate $2Nu\sqrt{p}$ and from X_1 to Y_2 with rate $Nu_c p r / (1-r)$, similar to the diagram in (iia).
- (iic) $r > 1$ and $\sqrt{p} \gg 1/N$ and $|1-r| \gg \sqrt{p}$ and $p \gg p^+$: tunnels from Y_0 to Y_2 with rate $2Nu\sqrt{p}$ and from X_1 to Y_2 with rate $Nu_c p N \log[N(r-1)/r]$; similar to the diagram in (iia).
- (iii) $r < 1$ and $p \gg p_c$ and $\sqrt{p} \ll 1/N \ll |1-r|$: tunnel from X_1 to Y_2 with rate $Nu_c p r / (1-r)$, but not from Y_0 to Y_2 ,



- (iv) $r > 1$ and $(r-1) \gg \sqrt{p} \gg 1/N$ and $p \ll p_c$: tunnel from Y_0 to Y_2 with rate $2Nu\sqrt{p}$, but no tunnel from X_1 to Y_2 ,



Note that the seven limiting cases above exhaust all the logical possibilities. This can be shown by using the following two inequalities:

$$\frac{1}{N} \geq r^{N/2-1}(1-r), \quad r < 1, \tag{20}$$

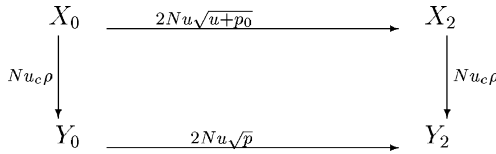
$$\frac{1}{N} \leq \sqrt{\frac{r-1}{rN \log \frac{N(r-1)}{r}}}, \quad r > 1. \tag{21}$$

Inequality (20) can be proved if we note that $\max_r \{r^{N/2-1}(1-r)\}$ corresponds to $r = (N-2)/N$ and equals to $2(N-2)^{N/2-1}N^{-N/2}$. Let us show that this is smaller than $1/N$. We have $2(N-2)^{N/2-1}N^{-N/2} \leq 1/N \leftrightarrow [(N-2)/N]^{N/2-1} \leq 1/2$. The latter inequality is true for $N \geq 4$, which proves Eq. (20) for these values of N .

Inequality (21) is equivalent to $\log[N(r-1)/r] \leq N(r-1)/r$, which always holds for $N(r-1)/r \geq 1$.

Next, we consider the case where $\sqrt{u+p_0} \gg 1/N$, i.e. the tunnel from X_0 to X_2 is open. This leads to only one additional case (since $p \gg p_0$ we automatically have the tunnel from Y_0 to Y_2).

(v) The diagram in this case is:

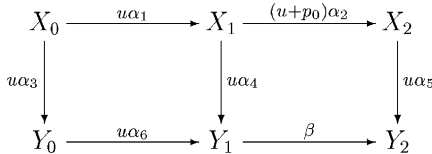


The coarse-grained differential equations corresponding to each of the cases follow from these diagrams in a straightforward manner. These linear equations can be solved analytically. However, the exact expressions for $Y_2(t)$ are enormous, and some approximations must be made to allow further analysis.

4.3. Conditions for Y_2 to dominate over X_2

Let us determine, for each of the five types of diagrams above, what the relative values of X_2 and Y_2 are. Our goal is to find out whether X_2 is larger or smaller than Y_2 for the time-scale comparable with the duration of human life.

Let us first consider the following diagram (case (i) of the previous section):



where $u \ll 1$. The corresponding differential equations are as follows:

$$\dot{X}_0 = -uX_0(\alpha_1 + \alpha_3), \tag{22}$$

$$\dot{X}_1 = u\alpha_1 X_0 - X_1(u\alpha_4 + (u+p_0)\alpha_2), \tag{23}$$

$$\dot{X}_2 = (u+p_0)X_1\alpha_2 - uX_2\alpha_5, \tag{24}$$

$$\dot{Y}_0 = uX_0\alpha_3 - uY_0\alpha_6, \tag{25}$$

$$\dot{Y}_1 = uX_1\alpha_4 + uY_0\alpha_6 - \beta Y_1, \tag{26}$$

$$\dot{Y}_2 = \beta Y_1 + uX_2\alpha_5. \tag{27}$$

For very long times, i.e. for $t \gg 1/u\alpha_i$, $1 \leq i \leq 6$, $Y_2 \approx 1$ and $X_2 \ll 1$. In the very short time range ($t \ll 1/\beta$), we have $X_2 \propto t^2$ and $Y_2 \propto t^3$, i.e. $X_2 > Y_2$. However, we are interested in the intermediate time range such that

$$\frac{1}{\beta} \ll t \ll \frac{1}{u\alpha_i}, \quad t \ll \frac{1}{(u+p_0)\alpha_2}. \tag{28}$$

In this case, we can solve the differential equations approximately, keeping only the largest contribution to each term. We have

$$X_0 = 1 + O(ut), \quad X_1 = \alpha_1 ut + O((ut)^2),$$

$$X_2 = \alpha_1 \alpha_2 \frac{u(u+p_0)t^2}{2} + O((ut)^3), \quad Y_0 = u\alpha_3 t + O(ut),$$

$$Y_1 = (\alpha_1 \alpha_4 + \alpha_3 \alpha_6) \frac{u^2 t}{\beta} + O((ut)^2),$$

$$Y_2 = \frac{(\alpha_1 \alpha_4 + \alpha_3 \alpha_6)(ut)^2}{2} + O((ut)^3).$$

Comparing X_2 and Y_2 , we obtain the following condition for Y_2 to win over X_2 in the intermediate time-scale (28):

$$\alpha_1 \alpha_4 + \alpha_3 \alpha_6 > \alpha_1 \alpha_2 \left(1 + \frac{p_0}{u}\right). \tag{29}$$

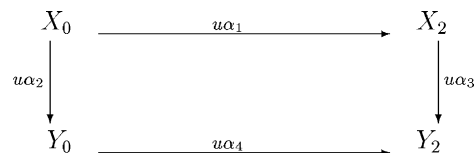
Similarly, we obtain conditions for Y_2 to win for diagrams of types (ii)–(iv) of the previous section:

$$\alpha_1 \alpha_7 + \alpha_3 \alpha_8 > \alpha_1 \alpha_2 \left(1 + \frac{p_0}{u}\right), \tag{30}$$

$$\alpha_1 \alpha_7 + \alpha_3 \alpha_8 > \alpha_1 \alpha_2 \left(1 + \frac{p_0}{u}\right), \tag{31}$$

$$\alpha_3 \alpha_8 + \alpha_1 \alpha_4 > \alpha_1 \alpha_2 \left(1 + \frac{p_0}{u}\right). \tag{32}$$

Note that in the regime defined by inequalities (28), both functions $X_1(t)$ and $Y_2(t)$ are quadratic in time. This means that in order to determine which of them is larger, we need to compare the coefficients. In other cases, X_2 and Y_2 may turn out to grow as different powers of t , in which case the condition $Y_2 > X_2$ will be time dependent. Let us consider the case corresponding to the tunnel from X_0 to X_2 . We have



The solutions are $X_2 = \alpha_1 ut + O((ut)^2)$, $Y_2 = (\alpha_1 \alpha_3 + \alpha_2 \alpha_4)(ut)^2/2 + O((ut)^3)$; note that X_2 is now linear in time. The condition for Y_2 to be larger than X_2 is now $(\alpha_1 \alpha_3 + \alpha_2 \alpha_4) \frac{tu}{2} > \alpha_1$.

In general, the condition for CIN to win can be written down by using the following rule. We need to consider all paths leading from X_0 to X_2 and from X_0 to Y_2 . Then we identify fast and slow leaps in the diagram. This depends on the rates, u, u_c, p and p_0 , and on the time scale of interest. In regime (28), leaps with rates multiplying u, u_c and p_0 are slow, and leaps corresponding to rate p are fast. For each path, we calculate the number of slow leaps—this number defines the power of t . We will only need the paths with the smallest number of slow steps. For each such path, we multiply the corresponding slow rates together; fast leaps do not count. Finally, we add all the paths together and divide by the factorial of the number of slow leaps.

To illustrate this let us consider the regime where p_0 is sufficiently large so that for the time-scale of interest, we have

$$p_0 \gg u, \quad ut \ll 1 \ll pt, \quad u_c t \ll 1 \ll p_0 t. \tag{33}$$

The leap with p_0 is now a fast step. Let us apply the rule described above. We have one slow leap (with or without tunneling) from X_0 to $X_2^{(1)}$ and two slow leaps (with or without tunneling) to reach $Y_2^{(1)}$. Therefore, $X_2(t) \propto ut$ and $Y_2(t) \propto uu_c t^2$. This means that if $u_c \sim u$ then X_2 will win. However, if $u_c \gg u$, there is a possibility that CIN may win. The corresponding calculations will depend on the particular parameter regime. For instance, if we are in regime (32), we have $X_2 = \alpha_1 ut$ and $Y_2 = (u^2 \alpha_2 \alpha_8 + u^2 \alpha_1 \alpha_4 + u^2 \alpha_1 \alpha_5) t^2/2$ so that CIN wins if

$$(\alpha_2 \alpha_8 + \alpha_1 \alpha_4 + \alpha_1 \alpha_5) \frac{ut}{2} > \alpha_1.$$

4.4. Conditions for CIN to inactivate the first tumor suppressor gene

We can combine the results of Sections 4.2 and 4.3 to obtain conditions for CIN to initiate colon cancer. The detailed calculations are presented in Appendix E; here we only show the result. The conditions are different depending on whether the CIN mutation is negatively selected, neutral or positively selected. We will say that CIN is *neutral* if the following condition holds:

$$|1 - r| \ll \max \left\{ \frac{1}{N}, \sqrt{p} \right\}.$$

Similarly, we will say that CIN is *positively selected* if

$$r > 1 \quad \text{and} \quad |1 - r| \gg \max \left\{ \frac{1}{N}, \sqrt{p} \right\},$$

and that CIN is *negatively selected* if

$$r < 1 \quad \text{and} \quad |1 - r| \gg \max \left\{ \frac{1}{N}, \sqrt{p} \right\}$$

Let us first assume that we are in regime (28). We have three cases:

$$\text{CIN is neutral : } n_c > \frac{1 + p_0/u}{2} \min \left\{ \frac{1}{2p}, \frac{1}{\sqrt{p}(N\rho + 1)} \right\},$$

CIN is negatively selected :

$$n_c > \frac{1 + p_0/u}{2} \min \left\{ \frac{1}{2p}, \frac{1 - r}{pr} \right\},$$

CIN is positively selected :

$$n_c > \frac{1 + p_0/u}{2} \min \left\{ \frac{1}{2p}, \left(pN \log \left[\frac{N(r - 1)}{r} \right] + N\rho \sqrt{p} \right)^{-1} \right\},$$

If on the other hand, p_0 is large enough to be considered as a fast rate (condition (33)) we have the same conditions except the factor $1 + p_0/u$ must be replaced by the factor $2/(Nut)$.

All these cases can be summarized as follows. The condition for CIN to appear before the second copy of the tumor suppressor gene is inactivated is

$$n_c > S \min \left(\frac{1}{2\rho}, C \right). \tag{34}$$

Here, the quantity C is defined as follows:

$$C = \begin{cases} 1/[(N\rho + 1)\sqrt{p}], & \text{CIN is neutral,} \\ (1 - r)/(rp), & \text{CIN is negatively} \\ & \text{selected,} \\ 1/\{pN \\ \log[N(r - 1)/r] \\ + N\rho \sqrt{p}\}, & \text{CIN is positively} \\ & \text{selected.} \end{cases}$$

The quantity S depends on whether p_0 defines a slow leap or a fast leap:

$$S = \begin{cases} (1 + p_0/u)/2, & p_0 t \ll 1, \\ 1/(Nut), & p_0 t \gg 1. \end{cases}$$

The case of weak selection: The conditions above depend among other parameters on the effective population size, N , and on the selective advantage or disadvantage of CIN cells. If, however, N is small then these dependencies are not relevant and the order of magnitude of the critical number of CIN genes is determined by the ratio p_0/u , where p_0 is the rate of LOH in a non-CIN cell and u is the mutation rate. Both quantities are expressed as probabilities per gene per cell division. If n_c exceeds p_0/u , then it is likely that CIN causes inactivation of the first tumor suppressor gene

and therefore acts as early driving force of tumor progression.

5. Discussion

We have developed a mathematical machinery to study stochastic dynamics of mutation–selection networks of cancer progression. Our approximations hold in the limit of small mutation rates, u , and small population sizes, N . Roughly we need $u \ll 1/N$. The other extreme of very large population sizes could perhaps be studied with a deterministic model.

We were mostly interested in the question whether CIN emerges before inactivation of the first tumor suppressor gene. If the number of genes that can induce CIN, n_c , exceeds a critical value, n^* , then there is a high probability that inactivation of the tumor suppressor gene happens in CIN cells. The required number of CIN genes, n^* , depends on the effective population size, N , the selective advantage or disadvantage of CIN cells and on the ratio p_0/u (the rate of LOH in normal cells vs the mutation rate). Let t be the characteristic time of human life divided by the time of cell division of the cells that are at risk of becoming cancer cells. If $p_0 t \ll 1$ holds, then we have the following result:

$$n^* = \frac{1}{2} \left(1 + \frac{p_0}{u} \right) \min \left(\frac{1}{2\rho}, C \right), \tag{35}$$

where $\rho = r^{N-1}(1-r)/(1-r^N)$, and $C = 1/([N\rho + 1]\sqrt{p})$ if CIN is neutral, $C = (1-r)/(rp)$ if CIN is negatively selected and $C = 1/\{pN \log[N(r-1)/r] + N\rho\sqrt{p}\}$ if CIN is positively selected. The more general formula which holds for the case $p_0 t \gg 1$ is given by Eq. (34).

We will illustrate our results using parameter estimates consistent with experimental data for colorectal cancers. The mutation rate per gene per cell division, u , is normally assumed to be $u = 10^{-7}$. It can be argued that the rate of LOH in normal cells is not much larger than u . Indeed, if p_0 were much larger than u , then, with or without CIN, the second copy of APC (the tumor suppressor gene in colorectal cancers) should always be inactivated by an LOH event. This is not the case: in about 30% of colon carcinomas, APC is inactivated by 2 mutational events (Miyoshi et al., 1992). We assume that the crucial effect of CIN is to increase the rate of LOH, which implies that the rate of LOH in CIN cells, p , satisfies $p \gg p_0$ (Lengauer et al., 1997; Bardelli et al., 2001).

The typical time-scale of human life can be taken as 70 years. The life cycle of stem cells is thought to be of the order of 1–20 days (see e.g. Bach et al., 2000; Potten et al., 1992). Therefore, the total number of cell divisions per human lifetime is of the order of 25,600 or smaller. Together with $p_0 \sim 10u \approx 10^{-6}$ we obtain $p_0 t < 0.026$, so

the inequality $p_0 t \ll 1$ holds even for these rather conservative estimates, and we can apply formula (35).

Let us first assume that it is the stem cells that are at risk of developing mutations leading to cancer. Then we have $N \approx 4-6$ (Yatabe et al., 2001). The relative reproductive rate of CIN cells, r , can be smaller, of the order of, or larger than 1. Let us consider all the three possibilities. For $N = 4$, we have

$$\frac{p_0}{u} = 1: \quad n^* = \begin{cases} 3, & r = 0.8, \\ 2, & r = 1.0, \\ 2, & r = 1.2, \end{cases}$$

$$\frac{p_0}{u} = 10: \quad n^* = \begin{cases} 16, & r = 0.8, \\ 11, & r = 1.0, \\ 9, & r = 1.2. \end{cases}$$

Note that for $N = 4$ and $p = 10^{-2}$, we have $|1-r| \ll \max\{1/N, \sqrt{p}\}$ for all the three values of r above, so that CIN is neutral.

More generally, we expect N to be somewhere between 1 and 1000. For $N = 100$, CIN is negatively selected for $r = 0.8$, neutral for $r = 1$ and positively selected for $r = 1.2$, and we have

$$\frac{p_0}{u} = 1: \quad n^* = \begin{cases} 25, & r = 0.8, \\ 5, & r = 1.0, \\ 1, & r = 1.2, \end{cases}$$

$$\frac{p_0}{u} = 10: \quad n^* = \begin{cases} 138, & r = 0.8, \\ 28, & r = 1.0, \\ 2, & r = 1.2. \end{cases}$$

It has been argued that the number of CIN genes is of the order of 100. For a wide range of parameter values, the required number of CIN genes, n^* , is small, and the condition $n_c > n^*$ is satisfied, which makes it likely that colon cancer is initiated by CIN mutations.

In general, our model suggests that the likelihood of an early CIN mutation depends on the number, N , of cells at risk, the ratio p_0/u of the LOH rate to the point mutation rate in normal cells and the number, n_c , of CIN genes. Measuring these quantities experimentally will shed light on molecular mechanisms underlying cancer development, specifically, the stem cell development and function, the cellular dynamics of LOH and the genetic causes of CIN.

Acknowledgements

This work was supported by the Packard Foundation, the Ambrose Monell Foundation, the Leon Levy and Shelby White Initiatives Fund, and Jeffrey Epstein. N.L.K. and M.A.N. would like to thank Yoh Iwasa for helpful discussions.

Appendix A. The method of differential equations

Let us denote the probability to be in state $a = i$ at time t as $\varphi_i(t)$. Using the transition matrix for two types, (2), we can write down the Kolmogorov forward equation for φ :

$$\begin{aligned} \frac{\partial \varphi_i}{\partial t} = & (1 - u) \left[\varphi_{i-1} \frac{(i-1)[N - (i-1)]}{rN - (i-1)(r-1)} \right. \\ & \left. - \varphi_i \frac{i(N-i)}{rN - i(r-1)} \right] \\ & + \varphi_{i+1} \frac{[r(N - (i+1)) + (i+1)u](i+1)}{rN - (i+1)(r-1)} \\ & - \varphi_i \frac{[r(N-i) + iu](N-i)}{rN - i(r-1)}. \end{aligned}$$

It is convenient to introduce the variable $\eta = i/N$. Taking the continuous limit and expanding into the Taylor series up to the second order, we obtain the following partial differential equation for $\varphi(\eta, t)$:

$$\frac{\partial \varphi}{\partial t} = \frac{\partial}{\partial \eta} (M\varphi) + \frac{1}{N} \frac{\partial^2}{\partial \eta^2} (V\varphi), \tag{A.1}$$

where

$$\begin{aligned} M &= \frac{\eta(1-r)(1-\eta) - u}{\eta(r-1) - r}, \\ V &= -\frac{1}{2} \frac{\eta[(1-\eta)(1+r) - u(1-2\eta)]}{\eta(r-1) - r}. \end{aligned}$$

When $r = 1 + s/N$ with $s \ll N$, we have the following equation:

$$N \frac{\partial \varphi}{\partial t} = s \frac{\partial}{\partial \eta} (\eta(1-\eta)\varphi) + \frac{\partial^2}{\partial \eta^2} (\eta(1-\eta)\varphi).$$

This equation is studied in Kimura (1994, paper 2). In the case $s \ll 1$ the principal term in the expression for $\varphi(\eta, t)$ is proportional to $e^{-\mu_0 t}$, where

$$\mu_0 = \frac{1}{N} (1 + O(s^2)).$$

This sets the typical time-scale of the process.

We can also study the case $1 \ll |s| \ll N$. In that limit, for $s > 0$, the region of interest is $\eta \ll 1$ (remember that $\eta = 0$ corresponds to the all “B” state). Thus the equation simplifies to

$$N \frac{\partial \varphi}{\partial t} = s \frac{\partial}{\partial \eta} (\eta\varphi) + \frac{\partial^2}{\partial \eta^2} (\eta\varphi).$$

This equation could be solved in terms of Laguerre polynomials, in general:

$$\varphi(\eta, t) = e^{-s\eta} \sum_{n=0}^{\infty} c_n L_n^1(s\eta) e^{-s(1+n)t/N}.$$

The Laguerre polynomials, $L_n^z(x)$ satisfy the differential equation

$$\left\{ x \frac{d^2}{dx^2} + (\alpha + 1 - x) \frac{d}{dx} + n \right\} L_n^z(x) = 0.$$

Note that the leading transient gives $\mu_0 = |s|/N = |1 - r|$ in this limit. One could similarly treat the case of $s < 0$, $|s| \ll 1$, by concentrating near the region $1 - z \ll 1$. In general, $\mu_0 = \frac{1}{N} f(s)$ where $f(s) = 1 + O(s^2)$ for small s but $f(s) \approx |s|$ for large s .

Appendix B. Absorption time for three types of cells, the neutral case

Let us study the model with three types of cells. The absorption time is defined by Eqs. (3) with the transition matrix given by Eq. (15). It turns out that for $r = 1$, and for small mutation rate u , the solution has a simple form. The first term in the Taylor expansion of $1/t_1$ in terms of u can be written as

$$\begin{aligned} \frac{1}{t_1} &= \frac{u \sum_{i=0}^{N-1} \binom{N-1}{i}^2 u_1^i}{\sum_{i=0}^{N-1} \binom{N-1}{i} \binom{N}{i+1} u_1^i} \\ &= \frac{uF(-N+1, -N+1, 1, u_1)}{F(-N+1, -N+1, 2, u_1)}, \end{aligned} \tag{B.1}$$

where $F(\alpha, \beta, \gamma, z)$ is the hypergeometric function. It can be shown that in the limit where $N^2 u_1 \rightarrow \infty$ and $u_1 \rightarrow 0$, we have

$$\frac{1}{t_1} = Nu\sqrt{u_1}. \tag{B.2}$$

To see this, we note that if $N^2 u_1 \ll 1$ then the sums in Eq. (B.1) are dominated by the term with u_1^0 . If, on the other hand, $N^2 u_1 \gg 1$, the largest contribution comes from the term with $i = i_0$ where $1 \ll i_0 \ll N$. To find i_0 , let us consider each term of $F(-N+1, -N+1, 1, u_1)$, $f_i = \binom{i}{N-1}^2 u_1^i$ and use the Stirling formula to approximate $\log f_i$. Then, differentiating $\log f_i$ with respect to i , we find that the maximum is achieved at $i = i_0 = N\sqrt{u_1}$. The same is true for the terms in the expression for $F(-N+1, -N+1, 2, u_1)$. The main contribution to both functions can be found:

$$f_{i_0} = e^{2N\sqrt{u_1}}$$

for $F(-N+1, -N+1, 1, u_1)$ and

$$f_{i_0} = e^{2N\sqrt{u_1} - \log(N\sqrt{u_1})}$$

for $F(-N+1, -N+1, 2, u_1)$. Plugging this into formula (37) we obtain estimate (B.2).

Appendix C. The Green function method

In this section we will introduce the Green function method of calculating transition probabilities. As we

will see later, this method is well adapted for approximations that involve “summing over” short-lived intermediate states to get *effective* transition rates between longer-lived states, allowing us to concentrate on a lower dimensional problem. This approach confirms our results for the two-species process obtained by other methods (Section 2), and allows us to carry out the calculations for a more complicated, three-species processes, Section 3.

The probability that, in t steps, the system makes a transition from state i to state j is $(P^t)_{ij}$, where P is the transition matrix; in the case of the two species model it is given by Eq. (2). This could be alternatively expressed as

$$\begin{aligned} (P^t)_{ij} &= \frac{1}{2\pi i} \oint_C dz z^t (zI - P)_{ij}^{-1} \\ &= \frac{1}{2\pi i} \oint_C dz z^t G_{ij}(z), \end{aligned} \tag{C.1}$$

where the contour C encloses the disk $|z| \leq 1$ and I is the identity matrix in the state space. Thus, knowing the matrix element $G_{ij}(z) = (zI - P)_{ij}^{-1}$ for all z is enough to calculate the transition probability from state i to state j , for any time t . We call $G_{ij}(z)$ the *Green function* in analogy with a similar quantity in quantum mechanics.

Let us break the state space S into two subspaces: S_{dir} (for *direct*) and S_{int} (for *intermediate*). The space S_{dir} contains the states of interest, i.e. the long-lived (or “almost absorbing”) states. The space S_{int} consists of short-lived states. Then the transition matrix P has the block structure:

$$P = \begin{pmatrix} P_{dir} & P_{dir-int} \\ P_{int-dir} & P_{int} \end{pmatrix},$$

P_{dir} corresponds to direct transitions between long-lived states in the sub-space S_{dir} . P_{int} corresponds to transitions among intermediate states, which are members of S_{int} , that we would like to eliminate in our effective picture. $P_{int-dir}$ and $P_{dir-int}$ contain the transitions linking the subspaces S_{int} and S_{dir} .

Let us suppose that α and β are two states in the subspace S_{dir} . An element $G_{\alpha\beta}(z)$ of the Green function can be decomposed as follows:

$$\begin{aligned} G_{\alpha\beta}(z) &= (zI - P)_{\alpha\beta}^{-1} \\ &= [zI - P_{dir} - P_{dir-int}(zI - P_{int})^{-1}P_{int-dir}]_{\alpha\beta}^{-1}. \end{aligned} \tag{C.2}$$

Under certain circumstances, one could make a “long time” approximation, which says that for large t , in expression (C.1) one can replace $G_{\alpha\beta}(z)$ with

$$G_{\alpha\beta}(z) \approx (zI - P^{eff})_{\alpha\beta}^{-1}, \tag{C.3}$$

where

$$P^{eff} = P_{dir} + P_{dir-int}(I - P_{int})^{-1}P_{int-dir}. \tag{C.4}$$

This is equivalent to taking $z = 1$ in $(zI - P_{int})^{-1}$ in decomposition (C.2). Let us discuss the conditions under which this can be done.

From an intuitive standpoint, this procedure is justified as long as the intermediate processes are much faster than the time scale of interest, which is transition time between states in the reduced subspace, S_{dir} . To be more precise, let λ be an eigenvalue of P and λ_{appr} —the eigenvalue of P^{eff} approximating λ . Then if λ_0 is the largest (i.e. closest to 1) eigenvalue of P_{int} , we can write

$$\frac{\lambda - \lambda_{appr}}{1 - \lambda_{appr}} \sim O\left(\frac{1 - \lambda}{1 - \lambda_0}\right).$$

Therefore, if one can show that

$$1 - \lambda \ll 1 - \lambda_0, \tag{C.5}$$

then approximation (C.3) is legitimate. In many calculations, we could check a posteriori, that inequality (C.5) indeed holds for the eigenvalues governing long time behavior, i.e. for λ_{appr} 's which are close to one. In the models to be discussed, this is a consequence of the assumption that the time required to get a relevant mutation is typically much longer than the time required for a the population to become homogeneous.

As an application of this approximation, let us look at the two species model. As was noted in Section 2.2, for small non-zero u , $b = N$ is the only absorbing state, and the state $b = 0$ is long lived. Hence we take the states $b = 0$ and N to span our space of interest, S_{dir} , whereas all the rest of the states, with $0 < b < N$, span S_{int} . The crucial effective transition rate $P_{0,N}^{eff}$ between the two states in S_{dir} is given by expression (C.4),

$$P_{0,N}^{eff} = P_{0,1}(I - P_{int})_{1,N-1}^{-1}P_{N-1,N}. \tag{C.6}$$

From Eq. (2), we have $P_{0,1} = u$. Since u is small, we can set $u = 0$ in the expression for P_{int} in Eq. (C.6). Next, let us rewrite Eq. (5) as $(I - P)\pi = (0, 0, \dots, P_{N-1,N})^T$, where $\pi = (\pi_1, \pi_2, \dots, \pi_{N-1})^T$. From this we can see that $(I - P)_{1,N-1}^{-1}P_{N-1,N} = \pi_1 = \rho$. Substituting this in Eq. (C.7), we obtain

$$P_{0,N}^{eff} = u\rho, \tag{C.7}$$

from which Eq. (8) follows immediately.

Let us go back and show that approximation (C.3) can indeed be used in this problem. We have $1 - \lambda_{appr} = 1/(NT_{abs}) = u\rho$. On the other hand, setting $u = 0$ in the expression for the matrix P_{int} , we can see that the value of $1 - \lambda_0$ depends on r and N and is typically non-zero in the limit when $u \rightarrow 0$; the value $1 - \lambda_0 = \mu_0$ is calculated in Appendix A. Therefore, inequality (C.5) holds for small u .

Appendix D. Analysis of the system with three types

D.1. Long-lived states and effective transition rates

Here we will calculate the effective transition rates between the states by using the Green function method developed in Section C. We will define the state space S_{dir} to consist of $(0, 0)$, $(N, 0)$ and $(0, N)$, i.e. the all “A”, all “B” and all “C” states; S_{int} consists of the rest of the states. The all “C” state is absorbing, so it is long lived by definition. The all “A” state is “almost absorbing” because of the smallness of u , the probability to create a mutant of type “B”. The all “B” state may or may not be long lived.

The effective transition rate from $(0, 0)$ to $(0, N)$ is given by formula (C.4),

$$P_{(0,0)(0,N)}^{eff} = \sum_{b=0,1} P_{(0,0)(1,0)}(I - P)_{(1,0)(b,N-1)}^{-1} \times P_{(b,N-1)(0,N)}. \tag{D.1}$$

In order to evaluate the quantity $(I - P)_{(1,0)(b,N-1)}^{-1}$, let us break S_{int} into two parts. One consists of the states $(b, 0)$, $0 < b < N$, with no “C” cells. We call this subspace S_{int}^0 . The other part consists of states (b, c) with $c > 0$. This is called S_{int}^+ . The transition matrix for the intermediate states, P_{int} , has the block structure

$$P_{int} = \begin{pmatrix} P^0 & P^{0+} \\ P^{+0} & P^+ \end{pmatrix}.$$

The transition rate between S_{int}^0 and S_{int}^+ can be used as a small parameter due to smallness of u_1 . Namely, let us consider the following Taylor expansion (where Q_1 and Q_2 are matrices, and Q_2 is “small”):

$$(I - Q_1 - Q_2)^{-1} = (I - Q_1)^{-1} + (I - Q_1)^{-1} Q_2 (I - Q_1)^{-1} + \dots$$

Setting

$$Q_1 = \begin{pmatrix} P^0 & 0 \\ P^{+0} & P^+ \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 0 & P^{0+} \\ 0 & 0 \end{pmatrix},$$

and letting $0 \leq b \leq 1$, we can evaluate the quantity $(I - P)_{(1,0)(b,N-1)}^{-1}$ up to the first order in P^{0+} :

$$(I - P)_{(1,0)(b,N-1)}^{-1} \approx 0 + [(I - P^0)^{-1} P^{0+} (I - P^+)^{-1}]_{(1,0)(b,N-1)}. \tag{D.2}$$

This approximation is valid provided that, starting from S_{int}^0 , the chance of getting absorbed at $(0, 0)$ is higher than the chance of jumping to S_{int}^+ . This assumption holds in our model because of the smallness of the rate u_1 .

Now, we can combine Eqs. (D.1) and (D.2) to write down the effective rates for the tunneling and genuine two-step processes.

D.1.1. Tunneling

In case of tunneling we have

$$P_{(0,0)(0,N)}^{eff-tunn} = \sum_{i=1}^{N-1} \sum_{b=0}^1 P_{(0,0)(1,0)}(I - P^0)_{(1,0)(i,0)}^{-1} P_{(i,0)(i-1,1)}^{0+} \times (I - P^+)_{(i-1,1)(b,N-1)}^{-1} P_{(b,N-1)(0,N)}. \tag{D.3}$$

The first term on the right-hand side of Eq. (D.3) is $P_{(0,0)(1,0)} = u$, i.e. the probability to create one mutant of type “B”. Each term

$$(I - P^0)_{(1,0)(i,0)}^{-1} P_{(i,0)(i-1,1)}^{0+} \tag{D.4}$$

is the probability that the system produces a mutant of type “C” while at the state $(i, 0)$. Finally, the summation over b ,

$$\sum_{b=0}^1 (I - P^+)_{(i-1,1)(b,N-1)}^{-1} P_{(b,N-1)(0,N)},$$

gives the probability to get absorbed in the all “C” state starting from one cell of type “C”. We will take this to be a number close to 1, for all i .

Let us concentrate on expression (D.4). From definition (15), we have

$$P_{(i,0)(i-1,1)}^{0+} = u_1 r i / \mathcal{N}_i, \tag{D.5}$$

which is the probability to create a mutant of type “C”, given that there are i cells of type “B” in the population. Let us show that

$$\sum_{i=1}^{N-1} (I - P^0)_{(1,0)(i,0)}^{-1} P_{(i,0)(i-1,1)}^{0+} = u_1 r \sum_{i=1}^{N-1} \frac{q_i i}{\mathcal{N}_i}, \tag{D.6}$$

where the vector $\mathbf{q} = (q_1, \dots, q_{N-1})$ satisfies the equation

$$\mathbf{q}(I - P^0) = \mathbf{f} \tag{D.7}$$

with $\mathbf{f} = (1, 0, \dots, 0)$, and P^0 the $(N - 1) \times (N - 1)$ matrix obtained from matrix P , Eq. (15), by taking $1 \leq i, j \leq N - 1$ and setting $u = 0$. Because of Eq. (D.5), it is enough to show that $(I - P^0)_{(1,0)(i,0)}^{-1} = q_i$ for all i . From (D.7) we have $\mathbf{q} = (1, 0, \dots, 0)(I - P^0)^{-1}$, or $q_i = (I - P^0)_{1i}^{-1}$, which is the desired relation. Combining Eqs. (D.3) and (D.6), we obtain the expression for the tunneling rate from A to C :

$$NP_{(0,0)(0,N)}^{eff-tunn} = R_{A \rightarrow C} = Nu u_1 r \sum_{i=1}^{N-1} \frac{q_i i}{\mathcal{N}_i}. \tag{D.8}$$

The multiplier N appears because of our definition of a unit time, which corresponds to N elementary events. In the next subsections we will calculate the vector \mathbf{q} and obtain expressions for $R_{A \rightarrow C}$ in different parameter regimes.

D.1.2. Genuine two-step process

For the genuine two-step process we have formally:

$$\begin{aligned}
 & P_{(0,0)(0,N)}^{eff-2-step} \\
 &= \sum_{b=0}^1 P_{(0,0)(1,0)} \\
 &\times (I - P^0)_{(1,0)(N,0)}^{-1} P_{(N,0)(N-1,1)}^{0+} \\
 &\times (I - P^+)_{(N-1,1)(b,N-1)}^{-1} P_{(b,N-1)(0,N)}. \tag{D.9}
 \end{aligned}$$

that we have to solve Eq. (D.7). There are two cases: $\sqrt{u_1} \ll |1 - r|$ and $\sqrt{u_1} \gg |1 - r|$. They will be considered separately.

D.2.1. The case $\sqrt{u_1} \ll |1 - r|$

In this regime, we can take the limit $u_1 \rightarrow 0$ in the transition matrix, so we have

$$\sum_{i=1}^{N-1} (\delta_{im} - P_{im})q_i = \delta_{1m}, \tag{D.13}$$

$$P = \begin{pmatrix}
 1 - \frac{(N-1)(r+1)}{N\mathcal{N}_1} & \frac{r(N-1)}{N\mathcal{N}_1} & 0 & 0 & \dots \\
 \frac{2(N-2)}{N\mathcal{N}_2} & 1 - \frac{2(N-2)(r+1)}{N\mathcal{N}_2} & \frac{2r(N-2)}{N\mathcal{N}_2} & 0 & \dots \\
 0 & \frac{3(N-3)}{N\mathcal{N}_3} & 1 - \frac{3(N-3)(r+1)}{N\mathcal{N}_3} & \frac{3r(N-3)}{N\mathcal{N}_3} & 0 \\
 0 & 0 & \frac{4(N-4)}{N\mathcal{N}_4} & 1 - \frac{4(N-4)(r+1)}{N\mathcal{N}_4} & \frac{4r(N-4)}{N\mathcal{N}_4} \\
 \dots & & & &
 \end{pmatrix}$$

Here, as before, $P_{(0,0)(1,0)} = u$ is the probability to create one mutant of type ‘‘B’’. $(I - P^0)_{(1,0)(N,0)}^{-1}$ is the probability for a mutant of type ‘‘B’’ to get fixated; it is nothing but ρ , given in formula (7). $P_{(N,0)(N-1,1)}^{0+}$ is the probability to create one mutant of type ‘‘C’’ starting from the all ‘‘B’’ state, it is equal to u_1 . Finally, the summation over b is the probability for type ‘‘C’’ to invade and it is approximately equal to one. Therefore, we can express the transition rate from A to B to C as

$$N^2 P_{(0,0)(0,N)}^{eff-2-step} = R_{A \rightarrow B} R_{B \rightarrow C}, \tag{D.10}$$

where

$$R_{A \rightarrow B} = Nu \frac{r^{N-1}(1-r)}{1-r^N}, \quad R_{B \rightarrow C} = Nu_1. \tag{D.11}$$

D.2. Two-step or tunneling?

Let us write down conditions for each of the two possible processes to take place, and the corresponding rates. It is easy to see that if

$$R_{A \rightarrow B} \ll R_{A \rightarrow C}, \tag{D.12}$$

then starting from the state with the dominant type ‘‘A’’, the system will get absorbed into the all ‘‘B’’ state with the rate $R_{A \rightarrow B}$ and then go to the all ‘‘C’’ state with the rate $R_{B \rightarrow C} = Nu_1$. On the other hand, if we have the ‘‘ \gg ’’ sign in condition (D.12), then the system will tunnel to the all ‘‘C’’ state from the all ‘‘A’’ state with the rate $R_{A \rightarrow C}$ without ever settling in the all ‘‘B’’ state.

The rate $R_{A \rightarrow B}$ is given by formula (D.11). In order to find whether or not condition (D.12) is satisfied, we need to calculate the rate $R_{A \rightarrow C}$, expression (D.8). This means

where the matrix P is found from Eq. (2) with $u = 0$. Let us denote

$$\xi_i = \frac{q_i i(N-i)}{N\mathcal{N}_i}.$$

For this variable we obtain the system

$$-\xi_{i-1}r + (r+1)\xi_i - \xi_{i+1} = 0, \quad 1 < i < N-1 \tag{D.14}$$

with boundary conditions

$$(r+1)\xi_1 - \xi_2 = 1, \tag{D.15}$$

$$-\xi_{N-2}r + (r+1)\xi_{N-1} = 0. \tag{D.16}$$

It is convenient to look for a solution in the form $\xi_i = \alpha^i$ which gives $\alpha_1 = r$ and $\alpha_2 = 1$. Substituting $\xi_i = Ar^i + B$ into the boundary conditions, we find

$$q_i = \frac{N\mathcal{N}_i}{ri(N-i)} \frac{r^i - r^N}{1 - r^N}. \tag{D.17}$$

Using definition (D.8) of $R_{A \rightarrow C}$, we obtain

$$R_{A \rightarrow C} = Nu_1 u \sum_{i=1}^{N-1} \frac{N}{N-i} \frac{r^i - r^N}{1 - r^N}. \tag{D.18}$$

We will now analyse different limits of this expression.

Type ‘‘B’’ is neutral, $|1 - r| \ll 1/N$: In this case expression (D.18) becomes

$$R_{A \rightarrow C} = Nu_1(1-r)N$$

and

$$R_{A \rightarrow B} = u,$$

so the condition for tunneling is

$$\sqrt{u_1} \gg \frac{1}{N}.$$

Since in the case under consideration, $\sqrt{u_1} \ll |1 - r| \ll 1/N$, we must have $\sqrt{u_1} \ll 1/N$, i.e. we do not have tunneling in this case.

Type “B” negatively selected, $r < 1, 1 - r \gg 1/N$: In this limit, we have

$$R_{A \rightarrow C} = Nu u_1 \frac{r}{1 - r}$$

and

$$R_{A \rightarrow B} = Nur^{N-1}(1 - r),$$

which gives the condition for tunneling:

$$\sqrt{u_1} \gg (1 - r)r^{N/2-1}.$$

Type “B” positively selected, $r > 1, 1/N \ll r - 1$: Let us rewrite the expression for $R_{A \rightarrow C}$ as

$$\begin{aligned} R_{A \rightarrow C} &\approx Nu u_1 \sum_{i=1}^{N-1} \frac{N}{N-i} (1 - r^{i-N}) \\ &= N^2 u u_1 \left(\sum_{k=1}^{N-1} \frac{1}{k} - \sum_{k=1}^{N-1} \frac{r^{-k}}{k} \right). \end{aligned}$$

The first of the terms in parentheses on the right-hand side is H_N , the harmonic number, which is a discreet equivalent of a logarithm. We have $H_N - \log N \rightarrow \gamma$ for large values of N (here $\gamma \approx 0.577216$ is Euler’s constant). The second term can be evaluated to give $B_{1/r}(N, 0) + \log((r - 1)/r)$, where the incomplete beta function, $B_{1/r}(N, 0)$, only gives a small correction. Putting everything together, we obtain

$$R_{A \rightarrow C} = u u_1 N^2 \left(\gamma + \log \left[\frac{N(r - 1)}{r} \right] \right). \tag{D.19}$$

On the other hand,

$$R_{A \rightarrow B} = Nu \frac{r - 1}{r},$$

so we have tunneling if

$$u_1 \gg \frac{r - 1}{rN \log[N(r - 1)/r]}.$$

D.2.2. The case $\sqrt{u_1} \gg |1 - r|$

Now the rate u_1 is significant, and we cannot neglect the terms with u_1 in the matrix P . Again, we have several sub-cases.

The case $\sqrt{u_1} \gg 1/N$. Let us first assume that the components q_i of \mathbf{q} from Eq. (D.7) decay fast with i , so that only the ones with $i \ll N$ need to be calculated. The resulting expressions for q_i will in fact show that this is a correct assumption.

Setting $u = 0$ and neglecting i compared to N in the matrix P (Eq. (15)), we obtain the following equations for $\xi_i = q_i i$:

$$-r \xi_{i-1} + \xi_i(r + 1 + u_1) - \xi_{i+1} = 0, \quad 1 < i < N.$$

With the substitution $\xi_i = \alpha^i$, we obtain a quadratic equation for α , whose roots under assumption $|1 - r| \ll \sqrt{u_1}$ are $\alpha_{1,2} = 1 \pm \sqrt{u_1}$. Setting $B = 0$

in $\xi_i = A(1 - \sqrt{u_1})^i + B(1 + \sqrt{u_1})^i$, and applying the initial condition for $i = 1$,

$$\xi_1(r + 1 + u_1) - \xi_2 = N,$$

we have

$$q_i = \frac{N(1 - \sqrt{u_1})^{i-1}}{ir}.$$

In order for this calculation to be valid, we need to require $(1 - \sqrt{u_1})^N \ll 1$, which gives the condition $\sqrt{u_1} \gg 1/N$, which is exactly the regime we are considering. Using the formula for the tunneling rate (D.8), we obtain

$$R_{A \rightarrow C} = Nu \sqrt{u_1}. \tag{D.20}$$

The rate $R_{A \rightarrow B}$ can be estimated as

$$R_{A \rightarrow B} = u$$

in the neutral case (when $|1 - r| \ll 1/N$), which gives us $\sqrt{u_1} \gg 1/N$ as the condition for tunneling; this condition of course always holds in this regime.

In the opposite case when $|1 - r| \gg 1/N$, we can also show that tunneling takes place. Let us first assume that $r < 1$. Neglecting r^N compared to 1 in the expression for $R_{A \rightarrow B}$, we see that for tunneling we need

$$u_1 \gg r^{N-1} \frac{1 - r}{N}. \tag{D.21}$$

Using $1 - u_1 \gg \sqrt{u_1}$, we can replace $1 - r$ by $\sqrt{u_1}$ in condition (D.21) to make it even more strict. Similarly, we can replace r^{N-1} by 1. The resulting condition gives $\sqrt{u_1} \gg 1/N$ which is satisfied. Next, let us set $r > 1$ and neglect 1 in comparison with r^N in the expression for $R_{A \rightarrow B}$. We need for tunneling

$$u_1 \gg \frac{r - 1}{rN}.$$

Replacing r by 1 in the denominator and $r - 1$ by $\sqrt{u_1}$ in the numerator, we again obtain $\sqrt{u_1} \gg 1/N$ which is satisfied in the present case. We conclude that tunneling with the rate given by Eq. (D.20) takes place in the regime where $\sqrt{u_1} \gg 1/N$.

The case $\sqrt{u_1} \ll 1/N$: In this regime, it is safe to take the limit $r \rightarrow 1$ first and obtain the corresponding rate $R_{A \rightarrow C}$. This is done in Section B. In the case where $\sqrt{u_1} \ll 1/N$, the sums in the expression for $R_{A \rightarrow C} = 1/t_1$, formula (37), can be replaced by the terms with $i = 0$, which gives the estimate for the tunneling rate,

$$R_{A \rightarrow C} = \frac{u}{N}.$$

On the other hand, we have $R_{A \rightarrow B} = u$, which is much larger, so we conclude that there is no tunneling effect in this case.

Appendix E. Conditions for CIN to initiate colon cancer: derivation

In what follows we assume that $\sqrt{p_0} \ll 1/N$, i.e. there is no tunnel X_0 to X_2 . The relevant time scale satisfies Eq. (28). We have three cases.

CIN is neutral: In this case, scenarios (ia) and (iia) apply. Using formulas (29) and (30) we obtain the following conditions for the probability $Y_2(t)$ to take over:

$$(ia): n_c > (1 + p_0/u)/(4\rho), \quad \sqrt{p} \ll 1/N,$$

$$(iia): n_c > \frac{1 + p_0/u}{2\sqrt{p}(N\rho + 1)}, \quad \sqrt{p} \gg 1/N$$

(we used definition (19) for the parameter u_c). Note that in the latter case, $\sqrt{p}(N\rho + 1) \gg \rho$. Therefore, we can write down both conditions in the following concise way:

$$n_c > (1 + p_0/u) \min \left\{ \frac{1}{4\rho}, \frac{1}{2\sqrt{p}(N\rho + 1)} \right\}.$$

If this is satisfied, then CIN will initiate colon cancer.

CIN is negatively selected: In this case, scenarios (ib), (iib) and (iii) apply. The corresponding conditions (obtained from formulas (29), (30) and (31), respectively) are as follows:

$$(ib): n_c > \frac{1 + p_0/u}{4\rho}, \quad \sqrt{p} \ll (1 - r)r^{N/2-1},$$

$$(iib): n_c > \frac{1 + p_0/u}{2} \left(\frac{pr}{1 - r} + N\rho\sqrt{p} \right)^{-1}, \quad \sqrt{p} \gg 1/N,$$

$$(iii): n_c > \frac{1 + p_0/u}{2} \left(\frac{pr}{1 - r} + \rho \right)^{-1}, \\ (1 - r)r^{N/2-1} \ll \sqrt{p} \ll 1/N.$$

Note that because of inequality (20), these cases exhaust all the possibilities to arrange $\sqrt{p}, 1/N$ and $(1 - r)r^{N/2-1}$.

In case (iib), let us show that

$$\frac{pr}{1 - r} \gg N\rho\sqrt{p}. \tag{E.1}$$

This inequality is the same as

$$\frac{\sqrt{p}}{N} \gg r^{N-2}(1 - r)^2$$

(we used the fact that $r^N \ll 1$ in the definition of ρ). The latter inequality can be proved by writing

$$\frac{\sqrt{p}}{N} \gg \frac{1}{N^2} > r^{N-2}(1 - r)^2,$$

where the first of the inequalities follows from $\sqrt{p} \gg \frac{1}{N}$, and the second one from inequality (20). This shows that Eq. (E.1) is true.

Next, let us note that in case (iii),

$$\frac{pr}{1 - r} \gg \rho. \tag{E.2}$$

This follows from $\sqrt{p} \gg r^{N-2}(1 - r)^2$. Therefore, conditions for Y_2 to win can be summarized as

$$(ib): n_c > \frac{1 + p_0/u}{4\rho}, \quad \sqrt{p} \ll (1 - r)r^{N/2-1},$$

$$(iib), (iii): n_c > \frac{1 + p_0/u}{2} \left(\frac{pr}{1 - r} \right)^{-1}, \quad \sqrt{p} \gg (1 - r)r^{N/2-1}.$$

Moreover, we have $\sqrt{p} \gg (1 - r)r^{N/2-1} \leftrightarrow pr/(1 - r) \ll \rho$, and similarly with the “ \gg ” sign. Therefore, we finally obtain

$$n_c > \frac{1 + p_0/u}{2} \min \left\{ \frac{1}{\rho}, \frac{1 - r}{pr} \right\}.$$

CIN is positively selected: In this case, scenarios (ic), (iic) and (iv) apply. The corresponding conditions (obtained from formulas (29), (30) and (32), respectively) are as follows:

$$(ic): n_c > \frac{1 + p_0/u}{4\rho}, \quad \sqrt{p} \ll 1/N,$$

$$(iic): n_c > \frac{1 + p_0/u}{2} (pN \log \kappa + N\rho\sqrt{p})^{-1}, \\ \times \sqrt{p} \gg \sqrt{(r - 1)/(rN \log \kappa)},$$

$$(iv): n_c > \frac{1 + p_0/u}{2} (\rho[1 + N\sqrt{p}])^{-1}, \\ \times 1/N \ll \sqrt{p} \ll \sqrt{(r - 1)/(rN \log \kappa)},$$

where we introduced the notation

$$\kappa = \frac{N(r - 1)}{r}.$$

Note that because of inequality (21), the three cases above exhaust all the possibilities to arrange $\sqrt{p}, 1/N$ and $\sqrt{(r - 1)/(rN \log \kappa)}$.

In case (iv), it is easy to see that the contribution from $N\sqrt{p}$ can be neglected in comparison with one, so the condition for CIN to win can be replaced by $n_c > 1/2(N\rho\sqrt{p})^{-1}$. On the other hand, we can show in this case

$$pN \log \kappa \ll N\rho\sqrt{p}. \tag{E.3}$$

This inequality is the same as

$$\sqrt{p}N \ll \frac{\kappa}{\log \kappa}$$

(we used the fact that $\rho \approx (1 - r)/r$ in this limit). The latter inequality can be proved by writing

$$\sqrt{p}N \ll \sqrt{\frac{\kappa}{\log \kappa}} < \frac{\kappa}{\log \kappa}$$

(here $\kappa > 1$). Therefore, the condition in case (iv) is also equivalent to

$$n_c > \frac{1 + p_0/u}{2} (pN \log \kappa + N\rho\sqrt{p})^{-1},$$

which is the same condition as in (iic). To summarize so far, we have

$$(ic): n_c > \frac{1 + p_0/u}{4\rho}, \quad \sqrt{p} \ll 1/N,$$

$$(iic), (iv): n_c > \frac{1 + p_0/u}{2} (pN \log \kappa + N\rho\sqrt{p})^{-1}, \\ \sqrt{p} \gg 1/N.$$

This can be rewritten as the following concise condition:

$$n_c > \frac{1 + p_0/u}{2} \min \left\{ \frac{1}{2\rho}, (pN \log \kappa + N\rho\sqrt{p})^{-1} \right\}.$$

Large p_0 : Finally, we consider the case where the relevant time scale satisfies (33). We have the same seven cases as before plus case (v) of Section 4.2. It is easy to see that the argument is exactly the same for the first seven cases, except the time dependence of X_2 is now linear. This means that in all the conditions, the factor $1 + u/p_0$ should be simply replaced by a factor $2/(Nut)$. Finally, one can show that case (v) does not change the resulting inequality.

References

- Abdel-Rahman, W.M., et al., 2001. Spectral karyotyping suggests additional subsets of colorectal cancers characterized by pattern of chromosome rearrangement. *Proc. Natl Amer. Soc.* 98, 2538–2543.
- Armitage, P., Doll, R., 1954. The age distribution of cancer and a multistage theory of carcinogenesis. *Br. J. Cancer* 8, 1–12.
- Armitage, P., Doll, R., 1957. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Br. J. Cancer* 11, 161–169.
- Bach, S.P., Renehan, A.G., Potten, C.S., 2000. Stem cells: the intestinal stem cell as a paradigm. *Carcinogenesis* 21, 469–476.
- Bardi, G., et al., 1997. Cytogenetic comparisons of synchronous carcinomas and polyps in patients with colorectal cancer. *Br. J. Cancer* 76, 765–769.
- Bardelli, A., et al., 2001. Carcinogen-specific induction of genetic instability. *Proc. Natl Amer. Soc.* 98 (10), 5770–5775.
- Bar-Or, R.L., et al., 2000. Generation of oscillations by the p53-Mdm2 feedback loop: a theoretical and experimental study. *Proc. Natl Amer. Soc.* 97, 11250–11255.
- Bell, G.L., 1976. Models of carcinogenesis as an escape from mitotic inhibitors. *Science* 192, 569–572.
- Bomme, L., et al., 1998. Allelic imbalance and cytogenetic deletion of 1p in colorectal adenomas: a target region identified between DIS199 and DIS234. *Genes Chromosomes Cancer* 21, 185–194.
- Byrne, H.M., Chaplain, M.A., 1996. Growth of necrotic tumors in the presence and absence of inhibitors. *Math. Biosci.* 35, 187–216.
- Cahill, D.P., et al., 1998. Mutations of mitotic checkpoint genes in human cancers. *Nature* 392, 300–303.
- Chaplain, M.A.J., 1995. The mathematical modelling of tumour angiogenesis and invasion. *Acta Biotheoret.* 43, 387–402.
- Cook, P.J., Doll, R., Fellingham, S.A., 1969. A mathematical model for the age distribution of cancer in man. *Int. J. Cancer* 4, 93–112.
- Fisher, J.C., 1958. Multiple mutation theory of carcinogenesis. *Nature* 181, 651–652.
- Goldie, J.H., Coldman, A.J., 1979. A mathematical model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer Treatment Rep.* 63, 1727–1733.
- Goldie, J.H., Coldman, A.J., 1985. Genetic instability in the development of drug resistance. *Semin. Oncol.* 12, 222–230.
- Herrero-Jimenez, P., et al., 2000. Population risk and physiological rate parameters for colon cancer. The union of an explicit model for carcinogenesis with the public health records of the United States. *Mutat. Res.* 447, 73–116.
- Kimura, M., 1994. *Population Genetics, Molecular Evolution, and the Neutral Theory: Selected Papers*. University of Chicago Press, Chicago and London.
- Knudson, A., 2001. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* 1 (2), 157–162.
- Lamlum, H., et al., 2000. APC mutations are sufficient for the growth of early colorectal adenomas. *Proc. Natl Amer. Soc.* 97, 2225–2228.
- Lengauer, C., Kinzler, K.W., Vogelstein, B., 1997. Genetic instability in colorectal cancers. *Nature* 386, 623–627.
- Lengauer, C., Kinzler, K.W., Vogelstein, B., 1998. Genetic instabilities in human cancers. *Nature* 396, 643–649.
- Lipkin, M., Bell, B., Shelrock, P., 1963. Cell proliferation kinetics in the gastrointestinal tract of man. *J. Clin. Invest.* 42, 767.
- Loeb, L.A., 1998. Cancer cells exhibit a mutator phenotype. *Foundations in Cancer Research*, Academic Press, New York, pp. 25–56.
- Luebeck, E.G., Moolgavkar, S.H., 2002. Multistage carcinogenesis and the incidence of colorectal cancer. *Proc. Natl Amer. Soc.* 99 (23), 15095–15100.
- Markowitz, S., 2000. DNA repair defects inactivate tumor suppressor genes and induce hereditary and sporadic colon cancers. *J. Clin. Oncol.* 18, 75s–80s.
- Miyoshi, Y., et al., 1992. Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum. Mol. Genetics* 1, 229–233.
- Moolgavkar, S.H., Knudson, A.G., 1981. Mutation and cancer: a model for human carcinogenesis. *J. Natl. Cancer Inst.* 66, 1037–1052.
- Nishisho, I., et al., 1991. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* 253, 665–669.
- Nordling, C.O., 1953. A new theory on the cancer inducing mechanism. *Br. J. Cancer* 7, 68–72.
- Owen, M.R., Sherratt, J.A., 1999. Mathematical modelling of macrophage dynamics in tumours. *Math. Models Methods Appl. Sci.* 9, 513–539.
- Perucho, M., 1996. Cancer of the microsatellite mutator phenotype. *Biol. Chem.* 377, 675–684.
- Plotkin, J.B., Nowak, M.A., 2002. The different effects of apoptosis and DNA repair on carcinogenesis. *J. Theor. Biol.* 214, 453–467.
- Potten, C.S., et al., 1992. Measurement of *in vivo* proliferation in human colorectal mucosa using bromodeoxyuridine. *Gut* 33, 71–78.
- Schmutte, C., Fishel, R., 1999. Genomic instability: first step to carcinogenesis. *Anticancer Res.* 19, 4665–4696.
- Shih, I.M., et al., 2001. Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. *Cancer Res.* 61, 818–822.
- Shorter, R.G., Moertel, C.G., Titus, J.L., J, R.R., 1964. Cell kinetic of in the jejunum and rectum of man. *Am. J. Dig. Dis.* 760.
- Strauss, B.S., 1998. Hypermutability in carcinogenesis. *Genetics* 148, 1619–1626.
- Taylor, S.S., McKeon, F., 1997. Kinetochores localization of murine Bubl is required for normal mitotic timing and checkpoint response to spindle damage. *Cell* 89, 727–735.
- Vogelstein, B., et al., 1988. Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* 319, 525–532.
- Wheldon, T.E., 1988. *Mathematical Models in Cancer Research*. IOP Publishing Ltd., Bristol.
- Wodarz, D., Krakauer, D., 2001. Genetic instability and the evolution of angiogenic tumour cell lines. *Oncol. Rep.* 8, 1195–1201.
- Yatabe, Y., Tavaré, S., Shibata, D., 2001. Investigating stem cells in human colon by using methylation patterns. *Proc. Natl Amer. Soc.* 98, 10839–10844.