Published in final edited form as:

J Theor Biol. 2012 December 7; 314: 84–94. doi:10.1016/j.jtbi.2012.08.015.

Accumulation of neutral mutations in growing cell colonies with competition

Ron Sorace¹ and Natalia L. Komarova^{1,2,*}

¹Department of Mathematics, University of California Irvine, Irvine CA 92697

²Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine CA 92697

Abstract

Neutral mutations play an important role in many biological processes including cancer initiation and progression, the generation of drug resistance in bacterial and viral diseases as well as cancers, and the development of organs in multicellular organisms. In this paper we study how neutral mutants are accumulated in nonlinearly-growing colonies of cells subject to growth constraints such as crowding or lack of resources. We investigate different types of growth control which range from “division-controlled” to “death-controlled” growth (and various mixtures of both). In division-controlled growth, the burden of handling overcrowding lies with the process of cell-divisions, the divisions slow down as the carrying capacity is approached. In death-controlled growth, it is death rate that increases to slow down expansion. We show that division-controlled growth minimizes the number of accumulated mutations, and death-controlled growth corresponds to the maximum number of mutants. We check that these results hold in both deterministic and stochastic settings. We further develop a general (deterministic) theory of neutral mutations and achieve an analytical understanding of the mutant accumulation in colonies of a given size in the absence of back-mutations. The long-term dynamics of mutants in the presence of back-mutations is also addressed. In particular, with equal forward-and back-mutation rates, if division-controlled and a death-controlled types are competing for space and nutrients, cells obeying division-controlled growth will dominate the population.

1 Introduction

Estimating the number of mutants accumulated in a colony of a given size is a mathematical problem which has important biological and medical applications. Spontaneous mutations provide the underlying variability on which selection acts to drive evolution. If sensitive bacteria are exposed to penicillin, eventually the culture becomes resistant to the drug [1]. Similarly, drug resistance in cancers such as Chronic Myeloid Leukemia (CML) is generated by mutations in the oncogene which prevent the binding of small molecule inhibitors [2]. The accumulation of mutations in the reproductive compartments of *C. elegans* presents an evolutionary optimization problem with the goal to minimize damage to the eggs [3].

The task of recovering the number of mutations in a colony, given the size and the mutation rate, can be viewed as an “inverse” problem that can be solved given the growth law of the

© 2012 Elsevier Ltd. All rights reserved

*Corresponding author, komarova@uci.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

cells. The corresponding distribution of the number of mutants is a variant of the well-studied Luria-Delbruck distribution [1]. In the last 60 or so years, much effort has been devoted to studying the Luria-Delbruck distribution and related problems, see for example [4, 5, 6, 7]. An excellent review of work prior to 1999 is provided by [8]. In the last decade, there have been many impressive results in the area of fluctuation analysis [9, 10, 11, 12, 13]. Novel algorithms for mutation rate estimators have been proposed, including [14, 15, 16, 17, 18, 19]; see also the review by [20]. Other important results include [21], where a different form of a branching process was considered, and [22], where an analysis of Luria-Delbruck type was used to extract cellular kinetics parameters, and not just the mutation rate. In his recent paper, [23] extends his earlier results [24], and studies neutral mutations in the absence of cell competition. He finds that under some natural assumptions concerning the life-time distributions and the offspring distributions of cells, the number of mutants in a larger colony converges to a certain stable random variable with index 1.

Most of the mathematical effort in the context of Luria-Delbruck analysis and related problems has been devoted to studying exponentially-growing colonies. On the other hand, when studying real biological systems, it is important to note that a linear birth-death process is only an approximation of the realistic growth law, and it only holds for some systems, and for some time. As a colony expands, cells may exhaust space or nutrients, and growth may slow or stop entirely. Competition among cells starts to play an important role, and the dynamics change. Recently, mutant dynamics in nonlinearly growing colonies have attracted the attention of researchers. In [25], the cells' reproduction probability was considered to be dependent on the population size. Under the regime of small mutation rates, a colony of “residents” (equivalent to the non-mutant cells) was assumed to exist at its carrying capacity, and the dynamics of mutants starting from small numbers was studied. In [26], the number of mutants for a specific class of nonlinear growth laws was calculated.

In the present paper we concentrate on the following question, how does competition/crowding affect the accumulation of mutations in a growing colony? In principle, as the effects of crowding start to be felt in a growing colony, its growth can be regulated in two distinct, and opposite, ways: (i) the division rate of cells may decrease, or (ii) the death rate of cells may increase. Both mechanisms can be present in the same system to varying degrees. Let us consider the two extreme scenarios (see the schematic in figure 1). If it is the ability of cells to divide that decreases with the colony size, we will call such regulation mechanism *division-controlled growth*. In the case where the death rate increases to reduce the net proliferation of the colony, we will refer to such regulation mechanism as *death-controlled growth*. Note that in biologically realistic systems, growth can be regulated by a mixture of these two mechanisms. Therefore, here we will study the whole spectrum of growth-control mechanisms, which lies between the two extrema.

Mathematically, this problems translates into studying nonlinear birth-death processes. The simplest example of such a process is the logistic (or Verhulst) growth law. This law states that the net growth of the population decreases linearly with the total population size. In the usual deterministic formulation of the logistic growth, the exact mechanism of the net growth rate reduction is unspecified. It however becomes important in more complex scenarios, when for instance we add stochastic effects, or consider the generation of mutations. In [25] it was the reproduction probability that was down-regulated by overcrowding, and in [26], both reproduction and death were equally affected. In the experimental study by [27], it is shown that both division- and death-controlled growth are observed the natural history of mammary tumors. In this paper we present a general analysis of the whole spectrum of growth controls, and show that the number of mutants produced by a growing colony depends on the exact type of regulation.

We also go beyond the logistic process and develop a general deterministic theory of neutral mutations. We consider how different mechanisms of growth regulation affect the accumulation of mutations for a general class of nonlinear growth-laws. A stochastic extension of this theory is subject of future work.

This paper is organized as follows. In section 2 we consider two particular types of birth-death processes, the linear process and the generalized logistic (or Verhulst) process. Deterministic and stochastic dynamics are formulated for these processes. In section 3 we investigate how the process of mutant accumulation differs in the linear and nonlinear model. Further, in the logistic systems, we explore how it depends on whether the growth is division-controlled or death-controlled. We consider both deterministic and stochastic formulations of the problem and show that the results hold in both settings. Section 4 develops a general deterministic theory of neutral mutations for any growth laws. It further generalizes our findings to any deterministic growth laws and shows that division-controlled systems have the smallest number of mutants. Both forward- and back-mutations are studied. Discussion is presented in section 5, where we provide a biological example illustrating our theory.

2 Birth-death processes with mutations, and different types of growth control

In this paper we address the simplest scenario tracking the generation of only one kind (or class) of mutants. The mutants are assumed to be neutral, that is, they are neither advantageous nor disadvantageous compared to the wild-type cells. The mutation rate is assumed to be constant. We would like to understand if different types of competition (that is, different growth laws) affect the process of mutant generation.

We begin by considering a stochastically growing colony of cells with death and mutation processes. Denote by i the number of wild-type cells and by j the number of mutant cells. The variable u is reserved for the rate at which mutants are produced, and the back-mutation rate is denoted by u_1 . Note that $0 \leq u, u_1 \leq 1$. We will study a class of continuous-time birth-death processes where the following events can occur during an infinitesimally short time-interval Δt .

- With probability $\lambda_{ij}^w (1 - u) \Delta t$, a wild-type cell divides faithfully, that is, it gives rise to another wild-type cell.
- With probability $\lambda_{ij}^w u \Delta t$, a wild-type cell divides with a mutation giving rise to a mutant cell.
- With probability $\mu_{ij}^w \Delta t$, a wild-type cell dies.
- With probability $\lambda_{ij}^m (1 - u_1) \Delta t$, a mutant cell divides giving rise to another mutant cell.
- With probability $\lambda_{ij}^m u_1 \Delta t$, a mutant cell divides with a mutation giving rise to a wild-type cell.
- With probability $\mu_{ij}^m \Delta t$, a mutant cell dies.
- With probability $1 - (\lambda_{ij}^w + \lambda_{ij}^m + \mu_{ij}^w + \mu_{ij}^m) \Delta t$, no change occurs in the system.

Here, the quantities λ_{ij}^w and λ_{ij}^m characterize the intensity of divisions of the wild-type and mutants cells, respectively. Similarly, μ_{ij}^w and μ_{ij}^m describe deaths of the wild-type and

mutants cells. These quantities, in general, are functions of the numbers of wild-type and mutants cells.

While this paper studies the accumulation of mutations for the general birth-death process described, we also consider two specific models of growth which appear in the literature, linear and generalized logistic.

The linear birth-death process

In this classical example, the intensities of divisions and deaths are linear functions of their variables,

$$\lambda_{ij}^w = li, \quad \lambda_{ij}^m = lj, \quad \mu_{ij}^w = di, \quad \mu_{ij}^m = dj,$$

where l and d are constant division and death rates of cells (assumed to be the same for both types of cells). The dynamics of the average numbers of cells are given by the ODEs,

$$\dot{x} = lu_1y + (l(1 - u) - d)x, \quad (1)$$

$$\dot{y} = lxx + (l(1 - u_1) - d)y, \quad (2)$$

where $x = \langle i \rangle$ and $y = \langle j \rangle$ are the expected numbers of wild-type and mutant cells, respectively. The cells in the population on average undergo an unlimited exponential growth. Denoting $z = x + y$ and adding equations (1) and (2), we obtain

$$\dot{z} = (l - d)z.$$

The quantity $l - d$ defines the time-scale of the process and can be taken to be $l - d = 1$. The one independent parameter d measures the relative amount of death in the system. This linear process (in the absence of back-mutations) was considered in [28] in detail. The linear birth-death process exhibits no control of growth, and the population expands exponentially as long as the division rate is larger than the death rate.

Logistic growth

This growth law is also known as the Verhulst law. In the most general formulation, the probability of cell divisions and deaths is affected by the availability of space or other resources. As the resources become sparse, the division rates decrease and the death rates increase linearly with the total number of cells,

$$\lambda_{ij}^w = [c - a(i+j)]^+ i, \quad \lambda_{ij}^m = [c - a(i+j)]^+ j, \quad \mu_{ij}^w = (d+b(i+j))i, \quad \mu_{ij}^m = (d+b(i+j))j,$$

where the positive part of a function is denoted by $[a]^+ = \max\{a; 0\}$. Here we assumed that the division and death rates of the mutants are the same as those for the wild-type cells. A deterministic description of the type of equations (1–2) is also possible in this case, except that the pair of ODEs is no longer the exact description of the expected numbers of cells. As previously, we denote $x = \langle i \rangle$ and $y = \langle j \rangle$ and use the replacement $\langle i^2 \rangle \rightarrow x^2$, $\langle j^2 \rangle \rightarrow y^2$ and $\langle ij \rangle \rightarrow xy$ to obtain

$$\dot{x} = (c - a(x+y))yu_1 + (c - a(x+y))x(1 - u) - (d+b(x+y))x, \quad (3)$$

$$\dot{y} = (c - a(x+y))xu + (c - a(x+y))y(1 - u_1) - (d+b(x+y))y. \quad (4)$$

Adding these two equations yields the logistic equation describing the dynamics of the total population,

$$\dot{z} = (c - d)z \left(1 - \frac{a+b}{c-d}z \right).$$

Note that the quantity $(c - d)$ simply sets the time-scale of the process. For our purposes it can be set to one. Let us require

$$c - d = 1, \quad a + b = \frac{1}{K}, \quad (5)$$

where K is the so-called effective carrying capacity of the system. Note that fixing $c - d$ and K is mathematically equivalent to non-dimensionalizing the variables t and (x, y) respectively. Now the equation for the total population takes its standard form,

$$\dot{z} = z(1 - z/K), \quad (6)$$

with a well-known solution representing the logistic, or sigmoidal, growth (see figure 2(a)):

$$z(t) = \frac{z_0 e^t K}{K + z_0(e^t - 1)}, \quad (7)$$

where $z_0 = x_0 + y_0$ is the initial number of cells.

Let us fix the value K , the effective carrying capacity of the system. Because of condition (5), there are only two independent parameters in this system, d and a . The parameter d represents the amount of death happening in the linear system far from the carrying capacity. As $d \rightarrow \infty$, there is more and more turnover in the system, and in the limit, the deaths balance the divisions. The parameter a measures how much the overcrowding is offset with divisions as compared to deaths. If $a = 1/K$ and $b = 0$, then the death rate does not depend on the population size, and as the size of the system increases, divisions slow. This scenario corresponds to *division-controlled growth*, figure 1. At the opposite extreme, if $a = 0$ and $b = 1/K$, the division rate does not depend on the population size, and the death rate increases as the system grows. This scenario corresponds to *death-controlled growth*, figure 1.

3 Mutant accumulation in the linear and logistic models

In this section we assume that $u_1 = 0$, that is, there are no back-mutations in the system. The consequences of back-mutations are discussed in the next section.

3.1 Deterministic description

We would like to investigate the number of mutants that accumulate in the system by the time the colony reaches a specific size N .

Linear model—The linear ODEs, equations (1–2), can be solved easily, such that $z(t) = z_0 e^{(1-d)t}$, $x(t) = x_0 e^{(1-u-d)t}$, and $y(t) = z(t) - x(t)$. The time when the colony size reaches a given value, N , can be obtained from the equation $z(t_N) = N$, to yield

$$t_N = \ln \frac{N}{z_0}, \quad (8)$$

where we assumed that $1 - d = 1$. Then, the number of mutants corresponding to $t = t_N$ is given by $y(t_N) = N - x(t_N)$, such that

$$y(t_N) = N - x_0 \left(\frac{N}{z_0} \right)^{1-(1+d)u} \quad (9)$$

This quantity is an increasing function of the death rate, as shown in figure 2(b) (the dashed line) and is intuitively clear. Since we fixed the net growth rate of the colony, as the death rate grows (increasing d), the division rates grow as well, which leads to a faster accumulation of mutations. In other words, colonies with a higher turnover rate will have a higher expected number of mutants.

Logistic model—Consider system (3–4) with $u_1 = 0$. In the long run, the system with unidirectional mutations will be dominated by mutants as follows from the stability of the fixed point $(\bar{x}, \bar{y}) = (0, K)$ (see appendix A for a detailed analysis of the solutions). Instead of looking at the steady states, we will focus on measuring the number of mutants at the time when the population reaches a certain size N , see figure 1(b). It is convenient to introduce the parameter α which relates the effective carrying capacity with the value N ,

$$N = \alpha K.$$

We are interested in studying properties of the quantity $y(t_N)$, where t_N is again the time it takes for the colony to reach size N : $z(t_N) = N$.

In figure 2(b) we show the expected number of mutants at size N , $y(t_N)$, as a function of parameter d , the scaled linear death parameter, for different values of a , and for different values of α (see the solid curves). For each value of α , we have $a = 1/K$ for the bottom curve (division-controlled growth), and $a = 0$ for the top curve (death-controlled growth). The intermediate laws are given by the curves between these two extrema, and parameter a decreases in the direction of the arrows. For comparison, the expected numbers of mutants at the same colony size N from the linear birth-death process is also presented (the dashed line). We observe the following trends.

- (a) All logistic growth models give a larger number of mutants compared to the linear model.
- (b) As the death rate increases, the number of mutants increases. For very fast turnover systems (high d) the number of mutants approaches N .
- (c) As a decreases, the number of mutations increases. This means that different growth-control mechanisms yield different mutation loads. If it is the death rate that changes with overcrowding (death-controlled growth), more mutants are generated than with the constant-death, decreasing-division-rate models (division-controlled growth).
- (d) The closer the system to its carrying capacity, the more mutations are accumulated. This corresponds to increasing α in our models.

All of these observations can be proved analytically. First we note that system (3–4) can be solved exactly. Use of solution (7) obtains:

$$x(t) = x_0 e^{t(1-(1+d)u)} \left(\frac{K}{K+z_0(e^t-1)} \right)^{1-auK}, \quad y(t) = z(t) - x(t). \quad (10)$$

The expression for the number of mutants at time t_N can be obtained in a closed form. We have

$$t_N = \ln \frac{N(K-z_0)}{z_0(K-N)},$$

and the number of mutants at time t_N for the logistic growth is

$$y(t_N) = N - x_0 \left(\frac{N}{z_0} \right)^{1-(1+d)u} \left(\frac{K-z_0}{K-N} \right)^{(aK-(1+d)u)}, \quad (11)$$

to be compared with formulas (8) and (9) for the linear growth. Using the facts that $N/z_0 > 1$ and $(K-z_0)/(K-N) > 1$ we can prove that $y(t_N)$ is an increasing function of d , and it tends to N as $d \rightarrow \infty$ (observation (b) above). Further, the derivative of $y(t_N)$ with respect to a is negative, proving observation (c). Next, as α increases, K decreases, and $y(t_N)$ decreases, proving statement (d). Finally, observation (a) corresponds to comparing logistically growing colonies (positive values of α), with linearly growing colonies ($\alpha = 0$). It follows from the previous argument that the number of mutants corresponding to $\alpha = 0$ (the linear case) is the lowest.

3.2 Stochastic description

The deterministic result clearly shows that to minimize the number of mutants produced by the logistically growing colony, the overcrowding must be controlled by reducing the division rate, not by increasing the death rate. In [28] we demonstrated that the deterministic system may not be a good approximation for the fully-stochastic, fixed-size problem. The question remains if the conclusions of the previous section remain true in the fully-stochastic system. To test the hypothesis that division-driven systems minimize the accumulation of mutations, we studied the corresponding stochastic process where the number of mutants is assessed at the time when the system reaches size $N < K$.

One approach to this problem would be to study a continuous-time process. However, since it is the average number of mutants at size N , and not the actual dynamics of mutant accumulation, that we are concerned with, a simplification can be implemented which reduces the problem to a simpler, discrete-time process. Both the discrete-time and the continuous-time processes would result in exactly the same probability distribution of the number of mutants, when the colony reaches a given size N . Here we will use a method generalizing the approach of [28] for logistic birth-death processes.

Stochastic setup—The states of the system belong to a two-dimensional simplex, $i+j = N$, $i, j = 0$, figure 3(a). We assume that all states (i, j) such that $i+j = N$ are absorbing. This is equivalent to stopping the process as soon as the colony reaches size N . The transition probabilities can be characterized as up-, down-, left- and right-probabilities, as shown in figure 3(a). These are related in a straightforward manner to the rates defined in Section 2:

$$P_{\rightarrow}^{ij} = R_{ij}^{-1} \lambda_{ij}^w (1-u), \quad P_{\leftarrow}^{ij} = R_{ij}^{-1} \mu_{ij}^w, \quad P_{\uparrow}^{ij} = R_{ij}^{-1} (\lambda_{ij}^w u + \lambda_{ij}^m), \quad P_{\downarrow}^{ij} = R_{ij}^{-1} \mu_{ij}^m, \quad (12)$$

for $i, j > 0$, $i+j < N$, with the normalization factor

$$R_{ij} = \lambda_{ij}^w + \lambda_{ij}^m + \mu_{ij}^w + \mu_{ij}^m.$$

The rest of the elements in the transition matrix are identically zero. Obviously, the state $(0, 0)$ is absorbing. In this model, the condition $P_{\rightarrow}^{ij} + P_{\leftarrow}^{ij} + P_{\uparrow}^{ij} + P_{\downarrow}^{ij} = 1$ means that the clock shifts one unit only if a birth or a death occurs.

Our goal is to calculate the mean number of mutants found in a colony of size N , starting from one wild-type cell. This quantity, which we denote by $E(\text{mut})$, is given by $E(\text{mut}) \equiv E(\text{mut}; 1, 0)$, where $E(\text{mut}; i, j)$ stands for the expected number of mutants starting from i wild-type and j mutant cells, given that colony reaches size N . Note that the mean value, $E(\text{mut}; i, j)$, is evaluated in the context of conditional probability, that the colony reaches size N , and does not go extinct before reaching size N . To study this quantity, we need to study the appropriate unconditional mean, as well as the probability of non-extinction.

Derivation of algebraic equations for the means—Let us denote as h_{ij}^k the probability to be absorbed at the state $(N-k, k)$ starting from (i, j) , figure 3(a). We have the following recursive equations for the quantity h_{ij}^k , the probability to have k mutants at size N starting from state (i, j) :

$$h_{ij}^k = P_{\rightarrow}^{ij} h_{i+1,j}^k + P_{\leftarrow}^{ij} h_{i-1,j}^k + P_{\uparrow}^{ij} h_{i,j+1}^k + P_{\downarrow}^{ij} h_{i,j-1}^k, \quad 0 \leq i+j \leq N-1. \quad (13)$$

Inserting the values for the transition probabilities from equations (12), we get

$$R_{ij} h_{ij}^k = \lambda_{ij}^w (1-u) h_{i+1,j}^k + \mu_{ij}^w h_{i-1,j}^k + (\lambda_{ij}^w u + \lambda_{ij}^m) h_{i,j+1}^k + \mu_{ij}^m h_{i,j-1}^k, \quad (14)$$

for $0 \leq i+j \leq N-1$. The boundary conditions are

$$h_{00}^k = 0, \quad h_{m,N-m}^k = \delta_{k,N-m},$$

where $\delta_{k,N-m}$ is the Kronecker symbol. The first of these conditions states that the colony cannot reach size N from $(0, 0)$, another absorbing state. The second condition means that starting from absorbing state $(N-k, k)$, the colony obviously ends up with k mutants.

Since we are interested in the mean numbers of mutants, let us introduce the variable

$$z_{ij} = \sum_{k=0}^N h_{ij}^k k. \quad (15)$$

This quantity has the meaning of the expected number of mutants starting at state (i, j) , as the system either reaches size N or goes extinct. To derive an equation for z_{ij} , we simply multiply equations (14) by k and sum over $0 \leq k \leq N$. This yields a set of equations for z_{ij} with $0 \leq i+j \leq N$,

$$R_{ij} z_{ij} = \lambda_{ij}^w (1-u) z_{i+1,j} + \mu_{ij}^w z_{i-1,j} + (\lambda_{ij}^w u + \lambda_{ij}^m) z_{i,j+1} + \mu_{ij}^m z_{i,j-1}, \quad (16)$$

$$z_{0,j} = N P_{0,j}^{\text{non-ext}}, \quad z_{N-j,j} = j, \quad 0 \leq j \leq N. \quad (17)$$

Here we denoted by $P_{0,j}^{non-ext}$ the probability that starting from point $(0, j)$, the colony reaches size N . The first of the two boundary conditions means that, starting from j mutant cells, the system will either terminate at state $(0, 0)$ or, with probability $P_{0,j}^{non-ext}$ will have N mutants. Equations (16–17) can be solved directly as an algebraic system of equations for sufficiently small N , and for larger system sizes, a boundary layer approximation can be used (see [28, 29]).

Note that by using a similar method, one can derive similar equations for higher (say, l th) moments of the number of mutants. In this case, we would have k^l instead of k on the right hand side of equation (15), and the corresponding equations would be obtained by multiplying equation (14) by k^l .

Probability of non-extinction—Finally, we need to derive a relation between the variable z_{ij} and the quantity of interest, $E(mut; i, j)$. As mentioned before, $E(mut; i, j)$ is the mean number of mutants, given that the colony reaches size N . Therefore, we can present this conditional mean as a re-normalized unconditional mean,

$$E(mut; i, j) = \frac{z_{ij}}{P_{i,j}^{non-ext}}, \quad (18)$$

where $P_{i,j}^{non-ext}$ is the probability that starting from point (i, j) , the colony reaches size N . We have $P_{i,j}^{non-ext} = \sum_{k=0}^N h_{ij}^k$. There is, however, an easier way to calculate the probability of non-extinction in our case. As long as the mutants grow and die with the same rates as the wild-type cells, that is, as long as the quantities $\lambda_{ij}^w + \lambda_{ij}^m$ and $\mu_{ij}^w + \mu_{ij}^m$ only depend on the sum $(i + j)$, we have $P_{i,j}^{non-ext} = P_{i+j}^{non-ext}$, the probability of non-extinction starting from the total of $(i + j)$ cells. The latter is easily calculated,

$$P_{i+j}^{non-ext} = \frac{1 + \sum_{l=1}^{i+j-1} \prod_{k=1}^l (\mu_k^w + \mu_k^m) / (\lambda_k^w + \lambda_k^m)}{1 + \sum_{l=1}^{N-1} \prod_{k=1}^l (\mu_k^w + \mu_k^m) / (\lambda_k^w + \lambda_k^m)}, \quad (19)$$

a generalization of the classical 1D gambler's ruin problem, e.g., [30], [31].

Summary of the results—In order to calculate $E(mut; i, j)$, the mean number of mutants in a colony of size N , starting from i wild-type and j mutant cells, we need to solve system (16–17) for unconditional means $z_{i,j}$, and then re-normalize by equation (18), using the probability of non-extinction given by equation (19). The quantity of our immediate interest is given by $E(mut) = E(mut; 1, 0)$. Note that this method of calculating the mean number of mutants does not rely on performing multiple stochastic runs and taking the average. Instead, we solve an algebraic system of equations which, after renormalization, yields the desired quantity. Please see [28, 29] for a discussion of computational costs and the boundary layer method of approximate solution of systems of type (16–17).

Figure 3(b) demonstrates a numerical solution of the stochastic system in the case of the logistic growth with varying parameters. We plot the quantity $E(mut)$ as a function of parameter d . For comparison, we also present the result of the linear stochastic model (the dashed line). This case corresponds to the result of [12], who showed that in an exponentially growing colony, the number of mutants grows with the death rate.

In parallel with the deterministic figure 2(b), the number of mutants corresponding to logistic growth is represented by solid curves. Again, the different curves correspond to different values of the parameter a , such that $a = 1/K$ for the bottom curve (division-controlled growth), and $a = 0$ for the top curve (death-controlled growth). The intermediate laws are given by the curves between these two extrema, and parameter a decreases in the direction of the arrow. We can see that the results obtained in the deterministic system hold in the stochastic system as well. The number of mutants increases with the turnover rate (the parameter d). The logistically growing colony produces more mutants than a linearly growing colony with similar parameters. Under the logistic growth law, the colony with diminishing division rates (not with increasing death rates) minimizes the number of mutations.

Comparing figures 2(b) and 3(b), one can see that the results for the deterministic and stochastic models are quantitatively somewhat different. This is not surprising. These are two different models, and one cannot say that one is an approximation of the other. Nor can one claim that between the deterministic and the stochastic models one is “right” and the other is “wrong”. They are both alternative descriptions of some biological assumptions, and their behavior is different in many respects. Therefore, it is important to establish that the results that we report hold for *both* models. The fact that our findings for the logistic growth hold both in the deterministic and stochastic framework means that the phenomenon we report possesses a certain degree of robustness, and is not an artifact of a particular modeling choice.

In the next section, we generalize this result (in the deterministic setting) to arbitrary growth laws.

4 General deterministic theory of neutral mutations

Observations (a–d) of the previous section have so far been only discussed in the context of logistic growth models. In this section, we will consider a more general system where the net growth is regulated by the total population size. We will prove some generalizations of the previous findings, and also provide some insights into general dynamics of neutral mutations with and without back mutations. This section only deals with the deterministic setting. An extension of these theories to general nonlinear stochastic models is subject of future work.

4.1 Neutral mutants in colonies with restricted growth

Consider a deterministic system, where the wild-type cells and mutants evolve according to the following equations

$$\dot{x} = G_1(z) u_1 y + G(z) (1 - u) x - D(z) x, \quad (20)$$

$$\dot{y} = G(z) u x + G_1(z) (1 - u_1) y - D_1(z) y. \quad (21)$$

Here the functions $G(z)$, $G_1(z)$ and $D(z)$, $D_1(z)$ are the division and death rates of the wild-type and mutants cells. Let us define the net growth rates as

$$f(z) = G(z) - D(z), \quad f_1(z) = G_1(z) - D_1(z),$$

and assume that these are non-increasing functions of their argument:

$$\frac{df}{dz} \leq 0, \quad \frac{df_1}{dz} \leq 0.$$

This means that the net growth rates of the wild-type cells and the mutants are arbitrary non-increasing functions of the total population size. We will further assume that there are numbers K and K_1 such that

$$f(K)=0, \quad f_1(K_1)=0. \quad (22)$$

In other words, the carrying capacities are defined as zeros of the growth rates.

Note that if $f(z) = f_1(z)$, then adding equations (20) and (21) yields

$$\dot{z} = z f(z), \quad (23)$$

that is, the total population size satisfies a self-consistent equation. If $f(z) \neq f_1(z)$, this is not the case and one needs to solve equations (20–21) as a system.

In general, to conform with our understanding of the biological systems, the functions G , G_1 , D and D_1 must satisfy the following constraints:

- (i) the division rates are non-increasing functions of the population size, $dG/dz \leq 0$ and $dG_1/dz \leq 0$;
- (ii) the linear division rates are given by $G(0) = G_{max}$ and $G_1(0) = G_{max,1}$;
- (iii) the death rates are non-decreasing functions of the population size, $dD/dz \geq 0$ and $dD_1/dz \geq 0$;
- (iv) the linear death rates are given by $D(0) = D_{min}$ and $D_1(0) = D_{min,1}$.

Note that in the presence of the Allee effect, some of these assumptions are violated; we however do not include this effect in the present model. System (20–21) together with the conditions on its various components, is a very general two-type deterministic system of equations which includes (1) back- and forward mutations, (2) the possibility that division and death rates of mutants are different from those of the wild-type cells, (3) the negative regulation of the division rates and/or the positive regulation of the death rates by the total population size; this reflects the effect of growth saturation, without fixing any specific functional forms. Below are some examples of such systems.

Examples—An example already considered here is the logistic growth process where we have

$$G(z) = G_1(z) = c - az, \quad D(z) = D_1(z) = d + bz,$$

and the second of conditions (5) assures equation (22). To formulate a slightly more general model of a logistic process, we note that in its usual formulation, the division and death rates are linear functions of the total population. In a nonlinear version of this process [32], we can set

$$\dot{z} = rz(1 - (z/K)^\gamma).$$

Other growth laws that describe cellular competition for space and nutrients have been used in the literature, for example, a variant of the Gompertz law [33]

$$\dot{z} = rz(1 - \ln z),$$

or exponentially-saturating growth [29]

$$\dot{z} = rz(e^{-\beta z} - e^{-\beta K}).$$

In these equations the constant r can be taken to be $r = 1$. In all the above examples it is assumed $f(z) = f_1(z)$, such that a single equation can be formulated for the total population. If only an equation of type (23) is provided, it is not clear how the burden of overcrowding is controlled by the division and death process regulation. For any growth law of the form (23), we propose the following parametrization:

$$G(z) = d + 1 - A(1 - f), \quad D(z) = d + (1 - A)(1 - f),$$

where the parameter $0 \leq A \leq 1$ corresponds to the system being division-controlled ($A = 1$) or being death-controlled ($A = 0$). Varying A from 1 to 0 we recover the whole range of various mixtures of growth- and death control laws.

General definition of neutrality—Returning to the most general formulation of the population-regulated two-component system formulated above, we ask the question: *For which functions G_1 and D_1 is the mutant neutral with respect to the wild-type, that is, under what conditions is the type y neither advantageous nor disadvantageous compared to the wild-type?* To this end, we find fixed points of system (20–21) in the absence of mutations, $u = u_1 = 0$. There are three cases.

- **Disadvantageous mutants.** If $K_1 < K$, the stable fixed point is given by $(K, 0)$, that is, the mutant is eliminated.
- **Advantageous mutants.** If $K_1 > K$, the stable fixed point is given by $(0, K_1)$, that is, the mutant dominates the system.
- **Neutral mutants.** If $K_1 = K$, there is a family of neutrally-stable fixed points which satisfies $\bar{x} + \bar{y} = K$. In this case, the long-term outcome of the dynamics, that is, the exact proportion of mutants and the wild-type cells, will depend on the initial conditions.

Note that the equality of the carrying capacities is enough to establish neutrality. The mutants may have completely different division and death rates, and for example could be characterized by a higher ($G_{\max,1} > G_{\max}$) or lower ($G_{\max,1} < G_{\max}$) linear division rate compared to the wild-type, but as long as $K_1 = K$, they will have no advantage or a disadvantage in the long run.

In the presence of mutations, the definition for advantageous and disadvantageous types changes slightly resulting in stronger inequalities for K and K_1 . Namely, if K_1 is sufficiently higher than K (more precisely, $K_1 > K + O(u_1 G_1(K_1)/f'(K))$), then the stable solution to the leading term in the mutation rates becomes

$$\left(\bar{x}, \bar{y}\right) = \left(\frac{-G_1'(K_1) K_1 u_1}{f(K_1)}, K_1 + G_1(K_1) K_1 u_1 \left(\frac{1}{f(K_1)} + \frac{1}{K_1 f_1'(K_1)}\right)\right),$$

where the prime denotes differentiation with respect to z . We can see that the presence of (small) mutation rates shifts the stable point away from $(0, K_1)$. Similarly, if

$K > K_1 + O(uG(K)/f_1'(K_1))$, the stable point is an appropriate perturbation of $(K, 0)$. In either case in the presence of mutations, the advantageous type dominates the system, and the disadvantageous type is maintained at a selection-mutation balance.

In the case of neutral mutations, the quantities $uG(K)/f_1'(K_1)$ and $u_1 G_1(K_1)/f'(K)$ set approximate bounds for the difference between K and K_1 for the mutants to have characteristics of neutrality. As long as

$$K_1 \in \left[K - O\left(\frac{uG(K)}{f_1'(K_1)}\right), K + O\left(\frac{u_1 G_1(K_1)}{f'(K)}\right)\right] \quad (24)$$

at the steady state, both mutants and the wild-type cells are present in significant proportions.

4.2 No back mutations: what growth control mechanism minimizes the number of mutants?

Let us consider unidirectional mutations ($u > 0, u_1 = 0$) and assume that $f(z) = f_1(z)$ for all z . The equations describing the accumulation of mutants are

$$\dot{x} = G(x+y)x(1-u) - xD(x+y), \quad (25)$$

$$\dot{y} = G(x+y)xu + G_1(x+y)y - D_1(x+y)y, \quad (26)$$

where the sum of the cell types $z = x + y$ satisfies equation (23). We further assume that the functions G, G_1, D and D_1 satisfy requirements (i)–(iv) of section 4.1. Under these constraints, we can solve the problem of the minimization of the number of mutants in a growing colony of cells which corresponds to a given colony size N .

We ask the question: what type of kinetic laws (division rates, $G(z)$ and $G_1(z)$, and death rates, $D(z)$ and $D_1(z)$), minimize the number of mutants found in a colony of a given size?

Suppose that the function $f(z)$ is fixed, and the solution of equation (23) is given by $x(t)$. Since $y(t) = z(t) - x(t)$, the task of minimizing the number of mutants is equivalent to the task of maximizing the number of wild-type cells. Integrating equation (25), we have

$$\ln(x/x_0) = \int_0^N (f(z) - G(z)u) dt = \int_{x_0+y_0}^N \frac{f(z) - uG(z)}{zf(z)} dz$$

where we made a change of variables $z dt = dz$ and used equation (23) for the derivative of z . This function is maximized if $G(z)$ is chosen to be as small as possible. From the constraints imposed on the functions G and D , this corresponds to the choice $D(z) = D_{min}$ and $G(z) = f(z) + D_{min}$. In other words, the number of mutants is minimized if the death rate does not increase with the population size, and instead, the division rate decreases as a response to

overcrowding. The opposite choice, $G(z) = G_{max}$ and $D(z) = G_{max} - f(z)$, maximizes the number of mutants produced. It is interesting that the minimization problem does not depend on the mutant functions, G_1 and D_1 , as long as $G_1(z) - D_1(z) = f(z)$. The result of the minimization is a direct generalization of the example of the logistic growth explored in section 3.

The proof above holds only in the absence of back-mutations, and under the equality $f_1 = f$. An extension to the case where $u_1 = 0$ but $f_1 \neq f$ is described in Appendix B. It follows that minimizing the function G minimizes the number of mutants accumulated at a given size N . This holds for any growth laws of the wild-type and mutant cells, even if we do not assume neutrality of the mutants (Appendix B).

To summarize, we found that in a two-component system of type (20–21) without back-mutations, under assumptions (i)–(iv) of section 4.1, minimizing the divisions function $G(z)$ is equivalent to minimizing the number of mutations found in a colony of a given size.

4.3 With back-mutations: the number of mutants in a colony of a given size

In the presence of back-mutations, the picture is more complicated. The number of mutants at size N is now defined by a balance of the division rates of both types of cells and the mutation rates, see Appendix C. As an example, figure 4 shows the number of mutants as a function of the parameter d for the logistic growth model without (a) and with (b) back-mutations. The equations we used for the wild-type and mutant populations are as follows:

$$\dot{x} = (c_1 - a_1(x+y))yu_1 + (c - a(x+y))x(1-u) - (d+b(x+y))x, \quad (27)$$

$$\dot{y} = (c - a(x+y))xu + (c_1 - a_1(x+y))y(1-u_1) - (d_1+b_1(x+y))y. \quad (28)$$

Here, parameters (a, b, c, d) characterize the wild-type cells, and parameters (a_1, b_1, c_1, d_1) characterize the mutants. As before, for the wild-type cells we have conditions (5). For mutant cells, c_1 and d_1 are independent variables, and the quantity $c_1 - d_1$ is the linear growth rate of the mutants. Since we are interested in neutral mutants, we set the mutant carrying capacity, $K_1 = K$, such that $b_1 = 1/K_1 - a_1$.

In the absence of back-mutations (figure 4(a)), we observe that division-controlled growth minimizes the number of mutants, and the number of mutants grows with the death rate d . In the presence of back-mutations (figure 4(b)), for the particular choice of parameters presented, the number of mutants is actually smaller for death-controlled growth, and the dependence on d is non-monotonic.

Because of this absence of a clear pattern of behavior for systems with back-mutations, we will employ a different strategy which allows us to get some general results. In the presence of back-mutations, instead of considering the number of mutants at a given colony size, we will concentrate on a different question, namely, that of a long-term mutant dynamics.

4.4 With back-mutations: the long-term dynamics of neutral mutants

Let us determine the long-term dynamics of neutral mutants in the presence of forward- and back- mutations, system (20–21). If we set $K_1 = K$, the stable fixed point of equations (20–21) can be found exactly in full generality. We have

$$\left(\bar{x}, \bar{y}\right) = \left(\frac{Ku_1G_1(K)}{uG(K) + u_1G_1(K)}, \frac{KuG(K)}{uG(K) + u_1G_1(K)}\right), \quad (29)$$

where $\bar{x} + \bar{y} = K$. A straightforward substitution in equations (20–21) yields zeros for both equations. In the particular case where $u_1 = 0$, the fixed point contains only the mutants. If on the other hand $u = 0$ and $u_1 > 0$, the colony is comprised of the wild-type cells.

In the presence of both back- and forward-mutations, it is interesting to see which type is more abundant at the equilibrium. From formula (29), we can see that the type that mutates more (has a larger value of the division rate at carrying capacity times the mutation rate) is less abundant, and the type that mutates less dominates. This is consistent with our previous statement that decreasing $G(z)$ minimizes the number of mutations. In the context of the generalized logistic growth, this corresponds to decreasing d and increasing a . In the long run, the neutral type with the smallest number of mutants produced wins the evolutionary competition. In particular, if the mutation rates are equal, and a division-controlled and a death-controlled types are competing for space and nutrients, the cells with the division-controlled growth will be the dominant population.

We note that the balance between the wild-type cells and the mutants, equation (29), is very sensitive toward changes in the carrying capacity of the cells. Formula (29) corresponds to the exact equality $K = K_1$. As we change K_1 in the interval defined by (24), the proportions of wild-type and mutants at the equilibrium change significantly but stay commensurate inside the region of neutrality.

5 Discussion and conclusions

In this paper we considered the process of accumulation of neutral mutations in stochastically and deterministically growing colonies. We showed that the expected number of mutants in a colony of a given size depends on the exact growth law. In particular, for systems where the growth slows as a result of overcrowding or the lack of resources, it is important exactly how the overcrowding is handled, by the reduction in the reproduction rates or by increase in cell deaths. We showed that the number of mutants is minimized by the colonies whose growth is controlled by divisions rather than by those which experience a death-controlled growth. We further developed a general deterministic theory of neutral mutations and found that in a colony with back-mutations, the long-term winner will be the type whose growth is division-controlled.

Our results can be applicable to many biological systems where mutations are accumulated in the presence of restricted growth. For example, from the point of view of tissue design in the context of organ development, it appears that growth control must be handled by a downregulation of divisions rather than an upregulation of apoptosis. This minimizes the number of mutations accumulated in the tissue, which in turn minimizes the risk of cancer.

We expect different mechanisms in neoplastic tissues. According to the theory of carcinogenesis, tumors have phases of growth alternating with phases of dormancy (plateaus). We hypothesize that, generally, a successful tumor is likely to exhibit a death-controlled growth to maximize the number of mutations accumulated. In other words, when the growth plateau is approached, we expect the tumor cells to continue to divide, and their apoptosis rates to increase, to slow down the net growth. In this scenario, the number of mutants accumulated by the tumor is maximized, which increases the chances for the tumor to emerge from the plateau phase and continue its growth. In the absence of further mutations, the tumor may remain dormant.

Specific biological example

A fascinating example of some of these mechanisms is provided in the paper by [27] which examined the escape from senescence in human mammary epithelial cells (HMECs). Early

exponential growth (phase (a), figure 5) of HMECs in culture was observed to slow down, as the population of cells entered a plateau (a senescent phase, phase (b)). Unlike other types of cells, such as human mammary fibroblasts, HMECs were reported to spontaneously exit the plateau phase, and undergo a second phase of growth (phase (c)), which resulted in a second plateau (phase (d)). Interestingly, the characteristics of the two plateau stages (phases (b) and (d)) were very different. In phase (b), the cells were characterized by genetic integrity and a low proliferative index. Very little deaths were observed at this stage. In other words, phase (b) was a truly senescent phase, which suggested that the growth in phase (a) was controlled by a decrease in cell divisions. On the other hand, in phase (d), the cells were in a state similar to a telomere-based crisis [34]. Their high proliferative index was counterbalanced by an increase in cell death. Further, the cells exhibited a high frequency of gross chromosomal abnormalities, which were accumulated rapidly in the second part of growth phase (c), before entering plateau (d).

What is observed in the experiments of [27] can be explained in the context of the theory presented here. The first (senescent) plateau observed is a result of a division-controlled growth, which is an adaptation of a multicellular organism to minimize cancer risk. In the literature, cellular senescence and tight control of genomic stability are believed to be important barriers for the development of cancerous lesions [35]. Cancers that manage to overcome that barrier proceed to high levels of genomic instability [36], phase (c) and (d) in figure 5. The second plateau is a result of a death-controlled growth, which results in a fast accumulation of further mutations and an increase in the lesion's malignancy and its overall invasive potential.

In other words, we hypothesize that the first plateau in [27] is a result of division-controlled growth, which is an evolutionary adaptation protecting the organism against cancer. The second stage of growth characterizes breaking out of this control, and entering a truly malignant stage where death-controlled growth takes place, which leads to a fast accumulation of further mutations.

Broader context and future work

The main result of this paper is the effect that crowding and competition have on the accumulation of mutants. As the population size approaches the carrying capacity, the net growth rate decreases. We show that the number of mutants accumulated in this process depends on how much the division and death rates are affected by crowding/competition.

It is useful to compare our results with the recent paper [26], which states that the expected number of mutants present in a tumor of a given size is independent of the type of the curve assumed for the average growth of the tumor. Our results clearly suggest that the number of mutants at size N , $y(t_N)$, depends on whether the growth is linear or logistic, and even within the logistic framework, it depends on the balance between births and deaths. To resolve this paradox, we note that in [26], only a very specific type of growth laws was used. Namely, both the division and the death rates were reduced by the competition/crowding in exactly the same way. In our notation, it corresponds to $c = d\lambda$ and $a = -b\lambda$ in equations (3–4), where λ is some positive parameter. With scaling (5) this condition reads $a = (1 + d/K)$, and $b = d/K < 0$ (that is, the death rate decreases with crowding). Using this condition in formula (11) we see immediately that the result for the nonlinear model coincides with the result for the linear model, equation (9), that is, both the linear and the nonlinear colony produce exactly the same number of mutants. This regime does not appear in figure 2(b) because we do not consider negative values of b (a decrease of deaths with crowding). From these arguments, we conclude that under the very specific restriction where both division and death rates decrease with crowding in exactly the same way, the logistically growing population will have the same number of mutants as the exponentially growing population.

More generally, however, the two populations will result in different numbers of mutants, as argued in our paper.

The theory presented here concerns neutral mutations. Neutral mutations are arguably the most common ones [37], and they play an important role in evolution [38]. Mathematically, neutral mutations are more difficult to handle compared to advantageous mutations which sweep through the population and come to a rapid fixation. As neutral mutants are produced, their dynamics are dictated by a balance between production, growth and death, and are subject to a random drift. In the present paper we have developed a stochastic approach to handle the random nature of neutral mutant dynamics in colonies with restricted growth. In the example of a stochastic logistic growth that we considered, we saw that division-controlled growth minimized the expected number of mutants produced. However, the general results concerning the division-controlled and death-controlled growth mechanisms were only obtained in a deterministic setting. One important extension of this work would be to create a general stochastic theory of mutant accumulation.

A Dynamics of wild-type and neutral mutant cells in a logistic model

System (3–4) under conditions (5) has two fixed points, $x = y = 0$ which is unstable and $x = 0, y = K$ which is stable. However in the special case where $a = 1/K$ and $d = 0$, the system acquires a neutrally stable family of solutions $x + y = K$. In this case, the long-term dynamics of wild-type cells and mutants are defined by the initial conditions. The reason for this is that in this special case, the death rate of cells ($d + bz$) is zero, and a portion of wild-type cells will remain in the population as $t \rightarrow \infty$.

The dynamics of wild-type cells and mutants have two distinct phases, as seen from the analysis of the quantity

$$\ln(x/x_0) = t(1 - (d+1)u) + 1(1 - auK) \ln\left(\frac{K}{K+z_0(e^t - 1)}\right). \quad (30)$$

Initially, the wild-type cells grow exponentially, with the rate

$$1 - z_0/K + u(z_0a - (d+1)).$$

This expression is obtained from the Taylor series of $\ln x$ for small t . For very large values of d , more precisely for values satisfying

$$d > \left(\frac{1}{u} - 1\right) \left(1 - \frac{z_0}{K} \left(\frac{1 - auK}{1 - u}\right)\right),$$

the growth never occurs, and the population of the wild-type cells decays from the start. If the value of d is lower than the preceding bound, the initial stage of growth will be replaced by an exponential decay with the rate

$$u(aK - (d+1)) < 0$$

as seen from the Taylor series of the same expression for large t . The time when the growth starts to slow to give way to decay is given by

$$t = \ln \frac{(K - z_0)(1 - u(d+1))}{uz_0(d+1 - aK)}$$

as seen from the zero of the time-derivative of expression (30). The value of $x = x_m$ at the maximum is given by

$$\frac{x_m}{x_0} = \left(\frac{(K - z_0)(1 - u(d+1))}{uz_0(d+1 - aK)} \right)^{u(aK - (d+1))} \left(\frac{K(1 - u(d+1))}{z_0(1 - auK)} \right)^{1 - auK}.$$

If we assume that $K \gg z_0$ and $du \ll 1$, we can approximate the above expression with

$$\frac{x_m}{x_0} = (K/z_0)^{1 - u(d+1)} [u(d+1 - aK)]^{u(d+1 - aK)} (1 - aKu)^{(1 - aKu)}.$$

Roughly speaking, the value

$$d_c \approx \frac{\ln 2}{u \ln K} - 1$$

correspond to $x_m/x_0 \approx K/2$. Thus, for $d > d_c$, the wild-type cells never reach more than a half of the carrying capacity. For $d < d_c$, the wild-type cells rise to significant numbers before being replaced by the mutants.

These calculations show that as d increases, the wild-type cells grow slower and slower, and the mutant cells increase in abundance for each given colony size.

B Non-equal growth rates of mutants and wild-type cells

Suppose that $f_1 = f$, and that there are no back-mutations in the system. We ask how the mutant number at a given colony size depends on the division rate of the cells. The equations describing the dynamics are given by

$$\dot{x} = fx - Gux \equiv F(x, y), \quad (31)$$

$$\dot{y} = f_1y - Gux \equiv F_1(x, y), \quad (32)$$

$$x(0) = x_0 \quad y_0 = y_0. \quad (33)$$

Let us further assume that the function G depends on a parameter α such that $G = G(z, \alpha)$, and, without loss of generality, $G_\alpha(z, \alpha) > 0$. We would like to investigate how the number of mutants at size N depends on the parameter α .

We have by definition of t_N ,

$$z(t_N) = N,$$

such that

$$dt_N/d\alpha = -z_\alpha(t_N)/\dot{z}(t_N),$$

where the subscript α denotes partial differentiation with respect to α . This is obtained from differentiation of the previous equation with respect to α . We have

$$\frac{dy(t_N)}{d\alpha} = y_\alpha + \dot{y} \frac{dt_N}{d\alpha} = y_\alpha - \frac{\dot{y} z_\alpha}{\dot{z}}.$$

Here the functions x , y , and their α - and time- derivatives are evaluated at time t_N . Since $\dot{z} \geq 0$, we have the sign of $dy(t_N)/d\alpha$ coincides with that of

$$\text{sign}(dy(t_N)/d\alpha) = \text{sign}(y_\alpha \dot{z} - z_\alpha \dot{y}).$$

Rewriting $\dot{z} = \dot{x} + \dot{y}$ and $z_\alpha = x_\alpha + y_\alpha$ and simplifying, we obtain further

$$\text{sign}(dy(t_N)/d\alpha) = \text{sign}(y_\alpha \dot{x} - x_\alpha \dot{y}).$$

Let us denote by H the expression $H = y_\alpha \dot{x} - x_\alpha \dot{y}$ and derive an ODE describing the evolution of H . We note that

$$\dot{H} = \ddot{x} \dot{y}_\alpha + \dot{x} \ddot{y}_\alpha - \ddot{y} x_\alpha - \dot{y} \ddot{x}_\alpha. \quad (34)$$

Differentiating equations (31–32) with respect to time we obtain

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \end{pmatrix} = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}.$$

Similarly, differentiating equations (31–32) with respect to α we obtain

$$\begin{pmatrix} \dot{x}_\alpha \\ \dot{y}_\alpha \end{pmatrix} = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \end{pmatrix} \begin{pmatrix} x_\alpha \\ y_\alpha \end{pmatrix} + u G_\alpha x \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Inserting these expression into equation (34), we obtain the following equation for H :

$$\dot{H} = H \left(\frac{\partial F}{\partial x} + \frac{\partial F_1}{\partial y} \right) + G_\alpha u x (\dot{x} + \dot{y}), \quad (35)$$

$$H(0) = 0, \quad (36)$$

where the initial condition is obtained from differentiating initial condition (33) with respect to α . On the right hand side of equation (35) the coefficient multiplying H is given by

$$\frac{\partial F}{\partial x} + \frac{\partial F_1}{\partial y} = f'x + f'_1y + f + f_1 - Gu.$$

Equations (35–36) comprise an initial value problem for H . Note that equation (35) is linear inhomogeneous, and that the sign of its solution is determined by the sign of the function $G_\alpha ux \dot{z} \geq 0$. We conclude that $H \geq 0$, and therefore

$$\frac{dy(t_N)}{d\alpha} \geq 0. \quad (37)$$

This shows that an increase in the production term $G(z)u$ will lead to an increase in the number of mutants in the colony of a given size. This result holds for growth-laws of mutants different from that of the wild-type cells. One does not need to assume mutant neutrality to obtain inequality (37).

C The effect of back-mutations

Suppose $f_1 = f$ and $u_1 > 0$, that is, the mutants grow identically to the wild-type cells, and both forward- and back-mutations are allowed. We ask how the mutant number at a given colony size depends on the division rate of the cells. The equations describing the dynamics are given by

$$\dot{x} = Gu_1y + fx - Gux, \quad (38)$$

$$\dot{y} = Gux + fy - Gu_1y. \quad (39)$$

As in section B we assume that the function G depends on the parameter α such that $G_\alpha \geq 0$, and ask how the number of mutants $y(t_N)$ depends on α . In this case, $f = f$ holds and, therefore, neither $x(t)$ nor t_N depend on α . We have the following equation for y

$$\dot{y} = uGz + (f - G(u + u_1))y.$$

Differentiating this with respect to α , we obtain the equation for y_α

$$\dot{y}_\alpha = (ux - u_1y)G_\alpha + (f - G(u + u_1))y_\alpha, \quad y_\alpha(0) = 0.$$

This is a non-homogeneous linear equation for y_α , and the sign of the solution depends on the sign of $(ux - u_1y)G_\alpha$. In particular, if $u_1 = 0$, we have $y_\alpha \geq 0$ as shown in section 4.2. If u_0 and $u_1 > 0$, we have $y_\alpha \geq 0$, which also is a consequence of the above result. In general, the sign of y_α depends on the dynamics of x and y , and in particular, on the initial conditions $x(0)$ and $y(0)$. There is no universal condition of biological interest which guarantees a particular sign for y_α in this case.

References

- [1]. Luria SE, Delbruck M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*. 1943; 28:491–511. [PubMed: 17247100]
- [2]. Druker BJ. Imatinib as a paradigm of targeted therapies. *Adv. Cancer Res*. 2004; 91:1–30. [PubMed: 15327887]

- [3]. Denver DR, Dolan PC, Wilhelm LJ, Sung W, Lucas-Lledo JI, Howe DK, Lewis SC, Okamoto K, Thomas WK, Lynch M, Baer CF. A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:16310–16314. [PubMed: 19805298]
- [4]. Lea D, Coulson C. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* 1949; 49:264.
- [5]. Armitage P. The statistical theory of bacterial populations subject to mutation. *J. Roy. Stat. Soc. B.* 1952; 14:1.
- [6]. Crump K, Hoel D. Mathematical models for estimating mutation rates in cell populations. *Biometrika.* 1974; 61:237.
- [7]. Ma W, Sandri G, Sarkar S. Analysis of the Luria-Delbrck distribution using discrete convolution powers. *J. Appl. Prob.* 1992; 29:255.
- [8]. Zheng Q. Progress of a half century in the study of the Luria-Delbrck distribution. *Math Biosci.* 1999; 162:1–32. [PubMed: 10616278]
- [9]. Kepler TB, Oprea M. Improved inference of mutation rates: I. An integral representation for the Luria-Delbrck distribution. *Theor Popul Biol.* 2001; 59(1):41–48. [PubMed: 11243927]
- [10]. Oprea M, Kepler TB. Improved inference of mutation rates: II. Generalization of the Luria-Delbrck distribution for realistic cell-cycle time distributions. *Theor Popul Biol.* 2001; 59(1): 49–59. [PubMed: 11243928]
- [11]. Dewanji A, Luebeck EG, Moolgavkar SH. A generalized Luria-Delbrck model. *Math Biosci.* 2005; 197(2):140–152. [PubMed: 16137718]
- [12]. Iwasa Y, Nowak M, Michor F. Evolution of resistance during clonal expansion. *Genetics.* 2006; 172(4):2557. [PubMed: 16636113]
- [13]. Haeno H, Iwasa Y, Michor F. The evolution of two mutations during clonal expansion. *Genetics.* 2007; 177(4):2209. [PubMed: 18073428]
- [14]. Zheng Q. Statistical and algorithmic methods for fluctuation analysis with salvador as an implementation. *Mathematical biosciences.* 2002; 176(2):237–252. [PubMed: 11916511]
- [15]. Zheng Q. New algorithms for luria-delbrück fluctuation analysis. *Mathematical biosciences.* 2005; 196(2):198–214. [PubMed: 15950991]
- [16]. Hall B, Ma C, Liang P, Singh K. Fluctuation analysis calculator: a web tool for the determination of mutation rate using luria-delbrück fluctuation analysis. *Bioinformatics.* 2009; 25(12):1564–1565. [PubMed: 19369502]
- [17]. Lang G, Murray A. Estimating the per-base-pair mutation rate in the yeast *saccharomyces cerevisiae*. *Genetics.* 2008; 178(1):67. [PubMed: 18202359]
- [18]. Niccum B, Poteau R, Hamman G, Varada J, Dshalalow J, Sinden R. On an unbiased and consistent estimator for mutation rates. *Journal of Theoretical Biology.*
- [19]. Wu X, Strome E, Meng Q, Hastings P, Plon S, Kimmel M. A robust estimator of mutation rates. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis.* 2009; 661(1–2): 101–109.
- [20]. Foster P. Methods for determining spontaneous mutation rates. *Methods in enzymology.* 2006; 409:195–213. [PubMed: 16793403]
- [21]. Xiong X, Boyett J, Webster R, Stech J. A stochastic model for estimation of mutation rates in multiple-replication proliferation processes. *Journal of mathematical biology.* 2009; 59(2):175–191. [PubMed: 18846374]
- [22]. Jean L, Suchorolski M, Jeon J, Luebeck E. Multiscale estimation of cell kinetics. *Computational and Mathematical Methods in Medicine.* 2010; 11(3):239–254. [PubMed: 20582763]
- [23]. Angerer W. Proliferation model dependence in fluctuation analysis: the neutral case. *Journal of mathematical biology.* 2010; 61(1):55–93. [PubMed: 19707763]
- [24]. Angerer W. An explicit representation of the luria-delbrück distribution. *Journal of Mathematical Biology.* 2001; 42(2):145–174. [PubMed: 11261316]
- [25]. Klebaner F, Sagitov S, Vatutin V, Haccou P, Jagers P. Stochasticity in the adaptive dynamics of evolution: the bare bones. *Journal of Biological Dynamics.* 2011; 5(2):147–162. [PubMed: 22873437]

- [26]. Tomasetti C. On the probability of random genetic mutations for various types of tumor growth. *Bulletin of Mathematical Biology*. 2012;1–17. [PubMed: 21431450]
- [27]. Romanov SR, Kozakiewicz BK, Holst CR, Stampfer MR, Haupt LM, Tlsty TD, Tlsty TD. Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. *Nature*. 2001; 409:633–637. [PubMed: 11214324]
- [28]. Komarova NL, Wu L, Baldi P. The fixed-size Luria-Delbruck model with a nonzero death rate. *Math Biosci*. 2007; 210:253–290. [PubMed: 17583754]
- [29]. Sorace, R. Ph.D. thesis. University of California; Irvine: 2011. Accumulation of mutations in stochastically growing colonies: Theory and applications.
- [30]. Feller, W. An introduction to probability theory and its applications. Vol. Vol. 1. John Wiley & Sons; 1968.
- [31]. Karlin, S.; Taylor, HM. A First Course in Stochastic Processes. Academic Press; New York: 1975.
- [32]. Richards F. A flexible growth function for empirical use. *J. Exp. Bot.* 1959; 10:290–300.
- [33]. Kozusko F, Bajzer Z. Combining Gompertzian growth and cell population dynamics. *Math Biosci*. 2003; 185:153–167. [PubMed: 12941534]
- [34]. Chin K, de Solorzano CO, Knowles D, Jones A, Chou W, Rodriguez EG, Kuo WL, Ljung BM, Chew K, Myambo K, Miranda M, Krig S, Garbe J, Stampfer M, Yaswen P, Gray JW, Lockett SJ. In situ analyses of genome instability in breast cancer. *Nat. Genet.* 2004; 36:984–988. [PubMed: 15300252]
- [35]. Cahill DP, Kinzler KW, Vogelstein B, Lengauer C. Genetic instability and darwinian selection in tumours. *Trends Cell Biol.* 1999; 9:57–60. [PubMed: 10087619]
- [36]. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. *Nature*. 1997; 386:623–627. [PubMed: 9121588]
- [37]. Kimura M. Evolutionary rate at the molecular level. *Nature*. 1968; 217:624–626. [PubMed: 5637732]
- [38]. Ohta T, Gillespie JH. Development of Neutral and Nearly Neutral Theories. *Theor Popul Biol.* 1996; 49:128–142. [PubMed: 8813019]

Figure 1.

Population-regulated growth. The two extreme cases are illustrated. A growing colony of cells which gradually fills its carrying capacity, K , is shown schematically. The cell divisions are represented as in-coming arrows, and deaths are out-going arrows. At the initial stages of growth (the top chart), the division rate exceeds the death rate. As the colony grows (the two charts at the bottom), the net growth rate must decrease. In the case of division-controlled growth, the division rate decreases and the death rate stays constant. In the case of death-controlled growth, the division rate stays constant, and the death rate increases.

Figure 2.

Deterministic studies of mutant accumulation. (a) The logistic growth function $x(t)$ with $x_0 = 1$ and different values of K . The number of mutants is assessed when the population reaches a defined size, $N < K$ ($N = 100$ is marked by a dashed line in the figure). (b) The expected number of mutants, for logistic growth models (solid lines) and the linear birth-death process (dashed line), as predicted by the deterministic system. Two sets of logistic growth curves correspond to parameters $\alpha = 0.99$ (further from the carrying capacity) and $\alpha = 0.9999$ (closer to the carrying capacity). The different logistic growth curves in each set correspond to 10 different values of parameter a from 0 (the top curve) to $1/K$ (the bottom curve). The other parameters are $N = 100$, $u = 0.05$, $x_0 = 1$, $y_0 = 0$.

Figure 3.

Stochastic studies of mutation accumulation. (a) A schematic showing the two-dimensional simplex and the notation for transition probabilities. (b) The expected number of mutants, for logistic growth models (solid lines) and the linear birth-death process (dashed line) obtained from the stochastic system. The different logistic growth curves correspond to 10 different values of parameter a , from 0 (the top curve) to $1/K$ (the bottom curve). The other parameters are $N = 100$, $\alpha = 0.99$, $u = 0.05$.

Figure 4.

The expected number of mutants, for logistic growth models with back-mutations, as predicted by the deterministic system. (a) No back-mutations, $u_1 = 0$; (b) with back-mutations, $u_1 = 2u$. The curves marked “Death-controlled growth” correspond to parameter $a = 0$, and the curves marked as “Division-controlled growth” to $a = 1/K$. The other parameters are $N = 100$, $\alpha = 0.99$, $u = 0.005$, $x_0 = 1$, $y_0 = 0$, $a_1 = a$, and $d_1 = c_1 = 1.5$.

Figure 5.
A schematic representing the phases of growth of HMEC cells, reported in [27].