JOURNAL OF THE ROYAL SOCIETY Interface

Selective pressures for and against genetic instability in cancer: a time-dependent problem

Natalia L Komarova, Alexander V Sadovsky and Frederic Y.M Wan

J. R. Soc. Interface 2008 **5**, 105-121
doi: 10.1098/rsif.2007.1054

JOURNAL OF THE ROYAL SOCIETY Interface

J. R. Soc. Interface (2008) **5**, 105–121
doi:10.1098/rsif.2007.1054
Published online 19 June 2007

Selective pressures for and against genetic instability in cancer: a time-dependent problem

Natalia L. Komarova[1,2,*], Alexander V. Sadovsky[2,3] and Frederic Y. M. Wan[1,4]

[1]*Department of Mathematics,* [2]*Institute for Genomics and Bioinformatics,*
[3]*California Institute for Telecommunications and Information Technology, and*
[4]*Center for Complex Biological Systems, University of California, Irvine, CA 92697, USA*

Genetic instability in cancer is a two-edge sword. It can both increase the rate of cancer progression (by increasing the probability of cancerous mutations) and decrease the rate of cancer growth (by imposing a large death toll on dividing cells). Two of the many selective pressures acting upon a tumour, the need for variability and the need to minimize deleterious mutations, affect the tumour's 'choice' of a stable or unstable 'strategy'. As cancer progresses, the balance of the two pressures will change. In this paper, we examine how the optimal strategy of cancerous cells is shaped by the changing selective pressures. We consider the two most common patterns in multistage carcinogenesis: the activation of an oncogene (a one-step process) and an inactivation of a tumour-suppressor gene (a two-step process). For these, we formulate an optimal control problem for the mutation rate in cancer cells. We then develop a method to find optimal time-dependent strategies. It turns out that for a wide range of parameters, the most successful strategy is to start with a high rate of mutations and then switch to stability. This agrees with the growing biological evidence that genetic instability, prevalent in early cancers, turns into stability later on in the progression. We also identify parameter regimes where it is advantageous to keep stable (or unstable) constantly throughout the growth.

Keywords: multistage carcinogenesis; chromosomal instability; somatic evolution; telomeres; optimization, bang-bang control; nonlinear control

1. INTRODUCTION

Genetic instability is found in a great majority of cancers, in very small lesions as well as in relatively advanced tumours. Two main types of genetic instability have been described (Lengauer *et al.* 1998; Sen 2000). One type, termed microsatellite instability (or MSI; Kinzler & Vogelstein 1996; Perucho 1996), is caused by a deficiency in the mismatch repair system and is characterized by an elevated rate of errors in the so-called microsatellites (which are stretches of DNA in which a short motif, usually one to five nucleotides long, is repeated several times). The other broad type of instability is chromosomal instability or CIN. It is characterized by gross chromosomal abnormalities in cancerous cells such as losses and gains of chromosomes, translocations, etc. The causes for MSI are relatively well understood: it is triggered by an inactivation of one of the mismatch repair genes such as hMSH1 and hMLH1. The origins of CIN are still unknown. Several gene candidates have been found whose inactivation causes chromosomal aberrations of cells (Cahill *et al.* 1998; Bardelli *et al.* 2001; Nasmyth 2002; Yarden *et al.* 2002; Rajagopalan *et al.* 2004); another possible reason for CIN may be the telomere crisis (Maser & DePinho 2002, 2004; Bailey & Murnane 2006).

The question whether genetic instability is a driving force or a consequence of cancer is as controversial today as it was three decades ago (Loeb *et al.* 1974; Breivik & Gaudernack 1999; Tomlinson & Bodmer 1999; Li *et al.* 2000; Shih *et al.* 2001; Marx 2002; Breivik & Gaudernack 2004). Analytical and computational approaches have been used to define the timing and other parameters of genetic instability as well as to study its role in carcinogenesis (Breivik & Gaudernack 1999; Breivik 2001, 2005; Komarova *et al.* 2002, 2003; Nowak *et al.* 2002; Little & Wright 2003; Michor *et al.* 2003, 2005; Nowak *et al.* 2006). A common motif of many of these papers is a microevolutionary nature of carcinogenesis.

All types of genetic instability are characterized by an increased rate of change of the cell's genome. This produces at least two effects. One is a possibly increased probability for a cell to experience an advantageous, malignant mutation which can increase the cell's fitness and lead to further growth. The other is an increased chance of deleterious, unwanted changes in the cell's genome which can reduce the cell's fitness or lead to the cell's death. These considerations suggest that, in principle, genetic instability can both increase the rate of cancer progression (by increasing the probability of

*Author for correspondence (komarova@math.uci.edu).

Received 26 March 2007
Accepted 24 May 2007

cancerous mutations) and decrease the rate of cancer growth (by imposing a large death toll on dividing cells). The question central to this paper is whether instability helps cancer, in the sense of Darwinian microevolution inside one organism.

The microevolutionary forces that act on cancer cells during multistage carcinogenesis can be modelled by taking into account these effects (Wodarz & Komarova 2005). In previous work, it was possible to recast the problem by formulating the question from the viewpoint of 'selfish' cancerous cells (Komarova 2004; Komarova & Wodarz 2004). What is the optimal level of instability which makes the cancer progress in the fastest way? The mathematical problem is finding the most efficient (from the point of view of cancer) rate at which genetic changes occur in cells. It was shown (Komarova & Wodarz 2004) that 'too much' instability is detrimental for the cells due to an increased death rate. 'Too little' instability also slows down the progress because the basic rate at which cancerous mutations are acquired is low. An optimal level of genetic instability has been identified which maximizes the rate of progression. This was quantified in terms of the probability of chromosomal loss per cell division and compared with available *in vitro* experimental measurements of this rate. The mathematical result turned out robust (it depended only logarithmically on parameter values), and its order of magnitude was consistent with the data (Lengauer *et al.* 1997).

In this paper, we take this model a step further and study the temporal change of the level of instability. As cancer progresses, the microevolutionary pressures inevitably change. What might have been a good strategy at the beginning of the growth may be detrimental for the colony later on. This time dependence finds experimental support: a recent paper by Chin *et al.* (2004) argues that the level of genetic instability in breast cancers first increases, reaches a peak and then decreases as the cancer passes through telomere crisis. A paper by Rudolph *et al.* (2001) reports data on intestinal carcinoma in mice and humans which is consistent with a similar model: telomere dysfunction promotes chromosomal instability which drives carcinogenesis at early stages; and telomerase activation restores stability to allow further tumour progression. The mechanism of telomerase activation and subsequent prevention of chromosomal instability is also described in the papers by Samper *et al.* (2001) and Artandi & DePinho (2000). It is shown that short telomeres can make mice resistant to skin cancer due to an increased cell death rate (Gonzalez-Suarez *et al.* 2000) which also suggests that telomerase activation and reduction in the level of chromosomal instability may be a necessary step for cancer to develop.

The idea that instability may be beneficial for cancer at an early stage and can become a liability later on is developed in the present paper. We formulate the time-dependent optimization problem to investigate the way to maximize cancerous growth. To model growth and mutations, we employ ordinary differential equations (ODE) similar to quasispecies equations (Eigen & Schuster 1979), widely used in modelling (micro-) evolutionary processes. Using methods of optimal control theory, we find strategies most advantageous for the tumour's growth. The degree of instability (the rate of mutations) appears as an unknown function of time, sought to maximize the growth of the mutants.

Mathematical theory of optimal control has been used in many areas of biosciences (Sontag 2004; Lenhart & Workman 2007). In biomedical applications, control theory has usually been employed to design treatment strategies by methods of optimization (Swan 1990; Kirschner *et al.* 1997; de Pillis *et al.* in press). In this paper, we apply optimal control theory to studying cancer in a very different way. We solve an optimization problem for the dynamics of cancerous growth in order to understand why cancer behaves the way it does. This approach is similar in spirit to the work of Iwasa & Levin (1995) that analysed the optimal timing of life strategies of breeding and migrating organisms. In a sense, we study the 'ecology' of cancer, based on our current knowledge of carcinogenesis, to see that the observed behaviour of tumours is essentially a consequence of the process of optimization.

Our main findings are as follows.

— For a wide range of parameters, the most successful strategy is to keep a high rate of mutations at first and then switch to stability. This explains much of the biological data (Rudolph *et al.* 2001; Chin *et al.* 2004).
— The time of the switching depends, to a small degree, on the 'target' tumour size. It is independent of the basic mutation rate or of the maximum rate of change caused by the instability (as long as the latter is much greater than the former, which is the biologically relevant scenario). The time of the switching is sensitive to the rate of growth of the mutants and is a decaying function of this parameter.
— It turns out that, depending on the concavity of the functional form chosen to express the death rate as a function of the mutation rate, the corresponding optimal strategies are qualitatively different. If the death rate is a linear or concave function of the mutation rate, then the optimal strategy is an abrupt (discontinuous) change from maximum instability to maximum stability. If the function is convex, the transition is more gradual.
— For some parameter regimes, the optimal strategy is to remain maximally unstable throughout the growth. This occurs, for example, if the magnitude of mutation-related death rate is small, while the gain in the mutation rate due to instability is large. On the other hand, a very large death rate and a small gain in mutation rate make instability disadvantageous at all times, and the best strategy then is to remain stable.

The rest of the paper is organized as follows. In §2, we formulate the biologically based mathematical model which describes cells' growth and mutations. Section 3 introduces the formalism of optimal control theory and summarizes the maximum principle for the optimal strategy. Section 4 is a detailed study of a subset of biologically relevant parameters. Optimal strategies are found for different choices of functional form of the death rate of cells. Section 5 considers the entire parameter space of the system and identifies in which cases instability is advantageous. Section 6

contains conclusions and discussion; it also provides suggestions for several experiments that may confirm or refute our theory and guide us in further research of the functional role of genetic instability in cancer.

2. THE MODELS

We model a birth and death process with mutations in a hypothetical setting where the mutation rate can be set to arbitrary (biologically admissible) values at each instant of time. Cells go through a sequence of mutations until an advantageous phenotype is achieved. Until this happens, the colony is subject to a regulatory process which keeps its size constant. The colony must escape this regulation in order to initiate the first wave of clonal expansion; it must 'overcome selection barriers in the race to tumorigenesis' (Cahill *et al.* 1999). Once advantageous mutants are produced, they have the ability to overcome the regulation and spread.

The mutation rate can have two effects. On the one hand, a large mutation rate leads to a high death toll in the population thus reducing the fitness of the cells. On the other hand, an increased mutation rate can lead to a faster production of the advantageous mutants, thus accelerating the net growth of the colony.

2.1. A one-step system

Let us suppose that a colony of cells is currently at a constant population size near a selection barrier. The growth is stalled and the cellular population remains near the 'carrying capacity' which is defined by the available space, nutrients and the cells' ability to divide and die. This barrier can be overcome by the offspring of a mutant whose properties are different. For instance, the mutant cells could have an activated oncogene and show an increased division rate or a decreased death rate. We assume that such transformed cells are created by means of one molecular event, genetic or epigenetic.

Let us denote by x_1 the population of cells that have not undergone cancerous mutations and by x_2 the mutated type. The probability for a cell to acquire an inactivating mutation of a particular gene upon a cell division is denoted by $\bar{\mu}$; this quantity is called the 'basic mutation rate'. The probability p is an additional transformation rate resulting from genetic instability. This quantity measures the degree of genetic instability in cells. It is low in stable cells (cells without CIN), but it can be highly elevated in chromosomally unstable cells. Effectively, if $p \ll \bar{\mu}$, then there is no genetic instability; $p \gg \bar{\mu}$ means a genetically unstable cell population. Both probabilities $\bar{\mu}$ and p are measured per gene per cell division.

With these notations in mind, we can present the processes of growth and mutations described above by means of a mutation diagram, figure 1*a*. This mutation diagram is of the same type as used in Nowak *et al.* (2002) and Komarova *et al.* (2003). The probability p is the parameter of optimization in our problem. We will try to find a strategy (i.e. the function $p(t)$) that maximizes the growth of cancer.

Before we go on, we would like to comment further on the biological meaning of the quantity p. The mechanism

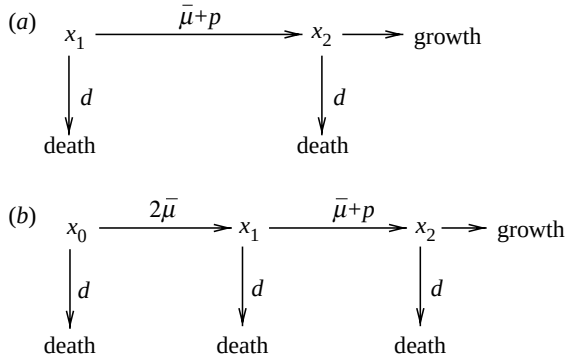

Figure 1. Mutation diagrams for (*a*) the one-step process and (*b*) the two-step process.

of mutations resulting from genetic instability, which is quantified by p, can be the same as or different from that of the basic mutations. For instance, one can assume that the transformation is a small-scale change in the DNA sequence, and the genetic instability is characterized by a deficiency in the nucleotide repair system (Benhamou & Sarasin 2000), which results in an increase (by an additive term p) of the basic mutation rate.

In other scenarios, the molecular mechanism of genetic instability can be different from that of the basic mutations. This is typical for the initiating events of familial colorectal cancers (familial adenomatous polyposis, FAP). Each gene is normally present in two copies (or alleles), a paternal and a maternal one. FAP patients are born with one of the copies of the adenomatosis polyposis coli (APC) gene inactivated in all cells. Then, an inactivation of the second copy eventually causes early lesions and further disease progression. The second copy of the APC gene can become inactivated by a point mutation or by a loss-of-chromosome event. The rate of chromosomal loss is often greatly elevated in colon cancers as a result of genetic (chromosomal) instability or CIN (Lengauer *et al.* 1998). In this situation, $\bar{\mu}$ is the basic point mutation rate and p is the rate of chromosomal loss.

The rate of chromosome loss, p, is a quantity which depends on many factors. No single gene responsible for chromosomal instability has been found; instead, a lot (of the order of hundreds) of genes have been shown to participate in various ways in the process of chromosome duplication, segregation, etc. A defect in any of those genes can change the resulting probability of chromosomal loss. Apart from that, the telomeres (regions of highly repetitive DNA at the end of a linear chromosome that function as disposable buffers) have been shown to play a role in genomic stability. Therefore, a change in any of these factors may be responsible for a change in the level of chromosomal instability and, therefore, can control the value of p.[1]

The dynamics of the cells is modelled here as follows. Cells reproduce and die, and the rate of renewal is taken to be 1 for the type x_1. In the absence of dangerous

[1]The inactivation of the TSG itself may actually be responsible for the change in p directly. It has been suggested (Fodde *et al.* 2001) that the inactivation of the APC plays a role in triggering genetic instability in colon cancer. This scenario remains controversial and is not included in the paper.

mutants, x_2, the total number of cells, x_1, is assumed to obey the well-known logistic growth law, that is, near the equilibrium the population stays constant, with a positive growth rate for the number of cells below the equilibrium, and a negative growth rate above the equilibrium. The mutants x_2 expand at the rate $a>1$.

Mutation diagram in figure 1*a* translates into the following ODEs describing the rate of change of the two cell populations,

$$x_1' = (1-\bar{\mu}-p-d(p))x_1 - \phi x_1, \tag{2.1}$$

$$x_2' = (p+\bar{\mu})x_1 + a(1-d(p))x_2 - \phi x_2, \tag{2.2}$$

where $(\cdot)' = \mathrm{d}(\cdot)/\mathrm{d}t$, $\phi = (1-d(p))x_1/N$, and $x_1(0) = N$, $x_2(0) = 0$. The term ϕ is similar to logistic growth in the absence of mutants, and accounts for the homeostatic control present in a system of x_1-cells. x_2-cells break out of regulation and enter a phase of exponential growth. The term with x_1 in the equation for x_2 is added to represent a partial, non-symmetric, homeostatic control that may play some role at the beginning of the growth of x_2 cells. Later on that term is simply a correction to the growth rate of the x_2 cells. This way of modelling the dynamics is not unique, and in fact the ϕ-term may be removed from the equation for x_2. A more detailed discussion of the robustness of the model is presented in the context of the two-step model, §§2.2 and 4.3.

2.2. A two-step system

In this section, we investigate another model where an inactivation of a tumour suppressor gene (TSG) leads to a clonal expansion. This is a two-step molecular process, whereby the two alleles of the gene are inactivated one at a time. The inactivation of just one allele does not result in any phenotypic changes. The inactivation of the second allele leads to the cells' unchecked growth. These processes can be summarized by a mutation diagram, figure 1*b*. There, x_0 is the population of $\mathrm{TSG}^{+/+}$ cells (i.e. cell with both copies of the TSG intact), x_1 is the population of $\mathrm{TSG}^{+/-}$ cells, where one of the copies of the TSG has been mutated, and x_2 is the population of $\mathrm{TSG}^{-/-}$ cells, where the remaining copy of the TSG has been lost.

The process described by the mutation diagram in figure 1*b* contains two steps: an inactivation of one, and then the other allele of the TSG. Note that the probabilities at which the two inactivation events occur are not equal. In the diagram of figure 1*b*, $\bar{\mu}$ is the basic mutation rate by which an allele can be inactivated. Since there are two alleles of each gene, and either of them can be inactivated first, the total probability of the first inactivation event is $2\bar{\mu}$. The second inactivation event can happen by another mutation (probability $\bar{\mu}$). As in the case of colorectal cancer and the APC gene, there is a different mechanism by which the second copy of the TSG can be turned off. This is a 'loss-of-chromosome' event, which is known to be responsible for the inactivation of a large percentage of TSGs in cancers (Kinzler & Vogelstein 2002). As a result of this event, the whole chromosome corresponding to the TSG in question becomes lost (or, more commonly, is replaced by a copy of the other chromosome where the TSG is mutated). This is a gross chromosomal change, whose probability, p, is our optimization parameter.

Before we go on, we would like to address the question of the asymmetry between the first and the second inactivation events. In principle, the first allele can also be inactivated by a loss-of-chromosome event. However, the fitness of a cell with a missing chromosome and one active copy of the TSG is very low. Such cells will quickly die out and will make no difference for the present analysis. On the other hand, a cell with one chromosome missing and both copies of the TSG inactivated has a selective advantage, because we assume that the inactivation of the TSG leads to an increase in the cell's growth rate. Such cells are produced by the sequence of events depicted in figure 1.

A system of ODEs that describes all these processes is as follows:

$$x_0' = (1-2\bar{\mu}-d(p))x_0 - \phi x_0, \tag{2.3}$$

$$x_1' = 2\bar{\mu}x_0 + (1-\bar{\mu}-p-d(p))x_1 - \phi x_1, \tag{2.4}$$

$$x_2' = (p+\bar{\mu})x_1 + a(1-d(p))x_2, \tag{2.5}$$

where

$$\phi = \frac{(1-d(p))(x_0+x_1)}{N}. \tag{2.6}$$

Cells reproduce and die, and the rate of renewal is taken as 1 for types x_0 and x_1. In the absence of dangerous mutants, x_2, the total number of cells x_0 and x_1 stays constant, i.e. the sum of equations (2.3)–(2.5) with $x_2=0$. The mutants x_2 expand at rate $a>1$.

This formulation for the two-step process is not unique. We will refer to system (2.3)–(2.5) and (2.6) as Model I. One possible modification is to replace the expression for ϕ, equation (2.6), with

$$\phi = 1 - \frac{d(p)(x_0+x_1)}{N}; \tag{2.7}$$

system (2.3)–(2.5) and (2.7) will be referred to as Model II. Also, we may include the ϕ-term in the last equation, that is, replace equation (2.5) with

$$x_2' = (p+\bar{\mu})x_1 + a(1-d(p))x_2 - \phi x_2. \tag{2.8}$$

We will name systems (2.3), (2.4), (2.8), (2.6) and (2.3), (2.4), (2.8), (2.7) Model I* and Model II*, respectively. Such changes in the model equations will lead to quantitative changes in the outcome, but, as we will demonstrate below, the results remain qualitatively robust with respect to such modifications.

Note that our formulation is similar to the well-studied quasispecies model (Eigen & Schuster 1979) in that the competition among cells is captured by means of an additive term (ϕ). This type of a description is widely used in cancer modelling (Wodarz & Komarova 2005) and other areas of mathematical biology. If we suppose $a=0$ (i.e. that the double-mutants, x_2, do not reproduce), then the first two equations, (2.3) and (2.4), are quasispecies equations. The inclusion of the non-zero growth term for the mutants that are not subject

Table 1. Variables, model parameters and their definitions.

notation	biological interpretation
$x_0(t)$, $x_1(t)$	pre-cancerous cell populations
$x_2(t)$	malignant cell population which escapes homeostatic control
$\bar{\mu}$	basic mutation rate of genetically stable cells (probability of mutation per cell division)
$p(t)$	the additional mutation rate resulting from genetic instability
$p_{\min}$, $p_{\max}$	the range of $p(t)$
u_{m}	$p_{\max}-p_{\min}$
N	the population size in the absence of malignant cells
M	the target population size of a growing tumour colony
σ	M/N
T	the time it takes the malignant colony to reach size M
μ	$\bar{\mu}+p_{\min}$
$u(t)$	$(p-p_{\min})/(p_{\max}-p_{\min})$, the scaled rate of genetic instability
A	the exponential growth rate of malignant cells
$d(u)$	the death rate of cells
α	the exponent in the definition of the death rate as a function of the rate of instability, formula (2.9)
d_{m}	the magnitude of the death rate, formula (2.9)

to competition reflects the escape of the biological system from the homeostatic control.

2.3. The death rate parameterization

The death rate, $d(p)$, is a function of the mutation rate, p. If p is small then chromosome losses do not happen, and if p is large a cell often loses chromosomes which results in an increased death rate. Therefore, in general, the function $d(p)$ will be a monotonically increasing function of p.

Here we present an example of a parameterization of the death rate as a function of p. It is convenient to introduce a normalized rate of chromosome loss, u, and express the death rate in terms of this parameter

$$d(u) = d_{\mathrm{m}}(1-(1-u)^{\alpha}), \quad u = \frac{p-p_{\min}}{p_{\max}-p_{\min}}, \quad p_{\min} \le p \le p_{\max}, \quad \alpha > 0. \tag{2.9}$$

The motivation for this particular dependency is as follows. Let us suppose that a cell dies if it loses one of α essential chromosome copies (out of the total of 2×23 copies in a human). Then, if we set $p_{\min}=0$ and $p_{\max}=1$, the death rate (2.9) can be written in a form

$$d(u) = d_{\mathrm{m}} \times [\text{Probability of cell death by chromosome loss}].$$

The constant d_{m} defines the magnitude of the death rate, and is taken to be in the interval $0 \le d_{\mathrm{m}} \le 1$. In this paper, we use general values for $p_{\min}$ and $p_{\max}$ such that $0 \le p_{\min} < p_{\max} \le 1$; these quantities define a biologically relevant range of the mutation rate, u. We allow α, the exponent in equation (2.9), to be a real positive number. In particular, we investigate the influence of the concavity of this function on the optimal solution (cases $\alpha<1$ and $\alpha>1$). The special case $\alpha=1$ yields a control problem where the controls enter linearly, a case much studied in the optimal control literature and rich with analytical results.

The variables and parameters used in our model, as well as their scaled versions employed in the next sections, are summarized in table 1.

2.4. Formulation of the optimization problem

In this paper, we adopt the theoretical framework where it is possible to set the rate of genetic instability to an arbitrary (but meaningful) value at each moment of time. Mathematically, the above framework means specifying the instability rate, $p(t)$, as a function of time. Every choice of such a function determines a growth process of the tumour. We shall seek the choice of $p(t)$ that allows the cancerous population to reach a given size, M, in the shortest possible time. In the terminology of optimization and control theory, the population size M is called the target, the possible values of genetic instability rate p are called 'admissible controls', and each choice of the function $p(t)$ is called a strategy. A strategy steering the system to the target faster than any other strategy is said to be an 'optimal strategy' or 'optimal control'. In this terminology, we seek a strategy for controlling the system to reach the target as soon as possible. A meaningful qualitative comparison between two strategies is now possible: the 'better', or 'more advantageous', strategy is the one allowing the system to reach the target sooner. Thus, an 'advantageous strategy' is advantageous for the cancer in the sense of Darwinian microevolution in an individual organism.

To find the optimal strategy, we consider the quantity, T, which is the solution of the equation

$$x_2(T) = M,$$

where x_2 is the solution of system (2.1) and (2.2) or (2.3)–(2.5). The growth time, T, depends on all the parameters of the system, including the time-dependent mutation rate, p. The optimal strategy is the one that minimizes the value of T.

In the simplest case, we restrict the class of admissible controls, $p(t)$, to constant functions. Then, the result of the optimization problem is a single value, p_{opt}, which will depend on the parameters of the system. A similar problem was solved in Komarova & Wodarz (2004). However, better growth times can be achieved if we allow p to be a function of time. It seems intuitive and is

evident from experimental results that higher initial and lower subsequent values of p will facilitate the growth.

3. MATHEMATICAL APPARATUS

In this section, we develop a mathematical framework for the one-step process. Similar calculations lead to the corresponding formulation for the two-step problem to be presented in appendix A.

3.1. Statement of the one-step problem

3.1.1. Equations of state. Let us define the following parameter combinations: $\sigma = M/N$, $u_{\rm m} = p_{\max} - p_{\min}$. Introducing the scaled quantities $x_1^* = x_1/N$, $x_2^* = x_1/M$, and dropping the asterisks for simplicity, we can rewrite system (2.1) and (2.2) as

$$x_1' = -(\mu + u_{\rm m}u)x_1 + [1-d(u)](1-x_1)x_1 \equiv g_1(x_1, x_2, u), \tag{3.1}$$

$$x_2' = \frac{1}{\sigma}(\mu + u_{\rm m}u)x_1 + [1-d(u)](a-x_1)x_2 \equiv g_2(x_1, x_2, u). \tag{3.2}$$

The death rate, $d(u)$, is given by equation (2.9). Recall that u is the normalized gross chromosomal change rate with $0 \le u \le 1$. As we show below, the three cases $\alpha > 1$, $\alpha = 1$ and $\alpha < 1$ may have to be treated separately.

3.1.2. Boundary conditions. The two ODEs (3.1) and (3.2) are subject to the following three auxiliary conditions:

$$x_1(0) = 1, \qquad x_2(0) = 0, \qquad x_2(T) = 1, \tag{3.3}$$

where T is the time when the dangerous mutant cell population reaches the target size.

3.1.3. Problem. Choose the control function $u(t)$ to minimize the time T needed to reach the target population size of the dangerous mutants subject to the inequality constraint,

$$0 \le u \le 1, \tag{3.4}$$

on the control $u(t)$ and non-negativity constraints on the cell populations,

$$x_1 \ge 0, \quad x_2 \ge 0. \tag{3.5}$$

3.1.4. Restatement of the optimal control problem. To apply the usual Hamiltonian system approach, we recast the problem described in §2.4 by choosing the control function $u(t)$ to minimize the performance index

$$J = \int_0^T 1 \mathrm{d}t, \tag{3.6}$$

subject to the equations of state (3.1) and (3.2), the boundary conditions (3.3) and the inequality constraints (3.4) and (3.5) with $T > 0$ as a part of the solution.

3.2. The maximum principle

3.2.1. The Hamiltonian and adjoint variables. Optimal control problems can be effectively analysed through the Pontryagin maximum principle and its associated Hamiltonian formalism (Pontryagin *et al.* 1962; Bryson & Ho 1969; Wan 1995). In this section, we develop components of the Hamiltonian formalism for our system. The *Hamiltonian* for our problem is

$$H = 1 + \lambda_1(t)g_1 + \lambda_2(t)g_2, \tag{3.7}$$

where λ_1 and λ_2 are the two continuous and piecewise differentiable *adjoint* (or costate) *variables* for the problem chosen to satisfy two *adjoint ODEs*,

$$\lambda_1' = -\left(\lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1}\right), \tag{3.8}$$

$$\lambda_2' = -\left(\lambda_1 \frac{\partial g_1}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2}\right), \tag{3.9}$$

and (for the given auxiliary conditions on the state variable x_1 and x_2) one *transversality condition*

$$\lambda_1(T) = 0. \tag{3.10}$$

Note that (3.1), (3.2), (3.8) and (3.9) form a Hamiltonian system for the Hamiltonian given in (3.7). (More generally, the Hamiltonian should be taken in the form

$$H = \lambda_0 + \lambda_1(t)g_1 + \lambda_2(t)g_2,$$

for some non-negative constant λ_0. But with the terminal time definitely having a role in the optimal control problem, the additional (constant) adjoint variable λ_0 does not vanish. We can rescale the Hamiltonian and simplify it to (3.7).)

With the set of the admissible controls specified by (3.4), an optimal strategy $u(t)$ is continuous or has finite jump discontinuities in $[0, T]$. This follows from the way how $u(t)$ appears in the state and adjoint ODE, (3.1), (3.2), (3.8) and (3.9), that the state and adjoint variables are continuous in $(0, T)$, including instances of jump discontinuities in the optimal control.

3.2.2. Formulation of the maximum principle. The optimal solution of our minimum terminal time problem requires the optimal control function $\bar{u}(t)$ to satisfy the following necessary conditions, known as the maximum principle (Pontryagin *et al.* 1962; Gelfand & Fomin 1963; Wan 1995).

(i) Four continuous and piecewise differentiable functions $\{\bar{x}_1, \bar{x}_2, \bar{\lambda}_1, \bar{\lambda}_2\}$ exist and satisfy the four differential equations (3.1), (3.2), (3.8) and (3.9), and four auxiliary conditions in (3.3) and (3.10) for the admissible control $\bar{u}(t)$.

(ii) The minimum terminal time T obtained with $u = \bar{u}(t)$, satisfies a free end condition

$$[H]_{t=T} = [1 + \bar{\lambda}_2\bar{g}_2]|_{t=T} = 0, \tag{3.11}$$

with

$$\bar{g}_2 = g_2(\bar{x}_1, \bar{x}_2, \bar{\lambda}_1, \bar{\lambda}_2, \bar{u}(t)),$$

where the adjoint boundary condition (3.10) has been used to simplify the expression for H.

(iii) For all t in $[0, T]$, the Hamiltonian achieves its minimum for $u = \bar{u}(t)$, i.e.

$$H(\bar{x}_1(t), \bar{x}_2(t), \bar{\lambda}_1(t), \bar{\lambda}_2(t), \bar{u}(t)) = \inf_v \left[H(\bar{x}_1(t), \bar{x}_2(t), \bar{\lambda}_1(t), \bar{\lambda}_2(t), v)\right], \quad (3.12)$$

for all v in the set of admissible controls restricted by (3.4).

(iv) If there should be a change in the control $\bar{u}$ at the instance T_s that involves a finite jump discontinuity in the value of the control, optimality requires that the Hamiltonian be continuous at T_s (Bryson & Ho 1969; Wan 1995)

$$[H]_{t=T_s-}^{t=T_s+} = 0. \quad (3.13)$$

Given the admissible controls as specified by (3.4), the optimal control $\bar{u}(t)$ can have finite jump discontinuities in $(0, T)$. It follows from the way $u(t)$ appears in the state and adjoint ODE that the state and adjoint variables are continuous in $(0, T)$, including instances of a control jump discontinuity.

3.2.3. The interior control. Suppose the optimal control strategy $\bar{u}(t)$ satisfies $0 \le \bar{u}(t) \le 1$ and the equation

$$\left[\frac{\partial H}{\partial u}\right]_{u=\bar{u}} = 0, \quad (3.14)$$

with

$$\begin{aligned}\frac{\partial H}{\partial u} &= \lambda_1\{-[u_m + d^{\cdot}(u)]x_1 + d^{\cdot}x_1^2\} \\ &\quad + \frac{\lambda_2}{\sigma}\{u_m x_1 - \sigma d^{\cdot} x_2(a - x_1)\} \\ &= u_m x_1\left(\frac{\lambda_2}{\sigma} - \lambda_1\right) - \{\lambda_1 x_1(1 - x_1) \\ &\quad + \lambda_2 x_2(a - x_1)\}d^{\cdot}. \end{aligned} \quad (3.15)$$

Then we call $(x_1, x_2, \lambda_1, \lambda_2, \bar{u})$ an *interior solution* for which the optimal control $\bar{u} = u_{\text{int}}(t)$ is an extremum (or, more correctly, an extremal) of the Hamiltonian. The *stationary* condition (3.14) can be written as

$$\{\lambda_1 x_1(1 - x_1) + \lambda_2 x_2(a - x_1)\}d^{\cdot} = \frac{u_m}{\sigma} x_1\{\lambda_2 - \lambda_1\sigma\}. \quad (3.16)$$

Using the expression for the death rate (2.9) with $d^{\cdot} = \alpha(1-u)^{\alpha-1}$, we obtain from condition (3.17) the following formula for the interior solution,

$$\begin{aligned}u(t) &= u_{\text{int}}(t) \\ &= 1 - \left(\frac{u_m x_1(\lambda_2 - \sigma\lambda_1)}{\alpha\sigma\{\lambda_1 x_1(1 - x_1) + \lambda_2 x_2(a - x_1)\}}\right)^{\frac{1}{\alpha-1}}. \end{aligned} \quad (3.17)$$

It can be shown (Wan *et al.* in preparation) that the interior solution above violates the inequality constraints (3.2) in some part of the solution domain for some range of system parameter values. Consequently, some combination of the *upper corner solution* ($u(t) = 1$), the *lower corner solution* ($u(t) = 0$) and the interior solution has to be considered for the optimal solution. Whenever a corner control is applicable, the adjoint variables (and the corresponding adjoint ODE and auxiliary conditions) may or may not play a role in the solution process since the control variable $u(t)$ is completely specified (and *not* determined by the stationary condition (3.14)).

3.2.4. A vanishing Hamiltonian, $H(t) = 0$. For an autonomous control problem, the Hamiltonian is constant for the optimal solution $\bar{u}(t)$ (Wan 1995; Wan *et al.* in preparation). The free-end condition (3.11) and the continuity condition (3.13) then require $H(t) = 0$. This result (together with the formula for the interior solution, equation (3.17)) will be used to find an optimal control. It can also be used to check how far a candidate control function is from the actual optimal control.

4. THE OPTIMAL RATE OF INSTABILITY

In this section, we consider the special case, $d_m = 1$, see equation (2.9). We will examine the corresponding problem in some detail and then turn to the analysis of the general problem in the following section.

The case $d_m = 1$ corresponds to fixing the death rate (relative to the cell division rate) at its highest possible value. The other parameter in expression (2.9) remains unspecified, such that $\alpha = 1$ corresponds to the linear dependence of the death rate on the mutation rate u, $\alpha > 1$ gives a concave dependence and $\alpha < 1$ a convex dependence. It turns out that the concavity of the function $d(u)$ is critical to the qualitative shape of the optimal control function. We will examine the cases $\alpha \ge 1$ and $0 < \alpha < 1$ separately.

4.1. Non-convex death rates, $\alpha \ge 1$

4.1.1. The case $\alpha = 1$. In this special case, the control variable, u, enters system (2.1) and (2.2) linearly. It follows from the Pontryagin maximum principle (Boltyanskii *et al.* 1962) that the optimal control in this case is *bang-bang* (Wan 1995; Wan *et al.* in preparation), i.e. piecewise constant switching between the values, $u_{\min} = 0$ and $u_{\max} = 1$. In that case, every such control $u(t)$ is completely determined by its initial value, $u(0)$, and the *switching times* $s^1, s^2, \ldots \in (0, T)$ at which $u(t)$ experiences a *switching*, i.e. a discontinuous change of value.

For the problem stated in §§3.1 and A.1, the optimal control is even simpler. It is possible to show (Wan *et al.* in preparation) that the optimal control starts with $u_{\max} = 1$ for a period $0 \le t < T_s$ and then switches to $u_{\min} = 0$ for the rest of the growth process, $T_s < t < T$; there is only one switching in the optimal control function $\bar{u}(t)$. Biologically, in order to maximize the growth, it is reasonable to take u as large as possible, namely, $u = 1$ at the start, such that some pool of mutants is created quickly, and then to switch to $u = 0$ later on, to take advantage of the exponential growth of $x_2(t)$. Therefore, the optimal solution has the form

$$\bar{u}(t) = \theta(T_s - t), \quad (4.1)$$

where θ is the Heaviside function and T_s is some switching time, $0 < T_s < T$, to be determined as a part of the solution process.

4.1.2. The case $\alpha > 1$. Next, we turn to strictly concave death rates. For nonlinear control problems, the optimal control is generally not bang-bang. One normally attempts to solve equations (3.1), (3.2), (3.8) and (3.9) with boundary conditions (3.3), (3.8), and the conditions (3.11) and (3.13) at the terminal time and at the switch points, T_s, by $u(t)$, respectively, with the control variable expressed in terms of the state and adjoint variables by (3.17). For our problem, the solution obtained by this method, however, violates the constraint $0 \le u(t) \le 1$, and thus is not applicable, at least, initially. Therefore, an optimal solution is to start with a corner control. It can be shown (Wan *et al.* in preparation) that optimal controls are again bang-bang, as in the case $\alpha = 1$, starting with $u = 1$ and switching to $u = 0$. In short, the nonlinear case $\alpha > 1$ of the control problem is characterized by a bang-bang solution of the form (4.1) just like the linear case $\alpha = 1$.

Formula (4.1) shows that an optimal control is completely determined by the value T_s of the switching time. In theory, this value is determined by $H(t_s) = 0$ which involves the solution for the adjoint variables. It is simpler to find T_s using the following approach. Varying T_s over a suitable interval $[r_1, r_2]$, compute for each T_s value the corresponding terminal time, T, by solving the initial-value problem (3.1)–(3.3). This process yields a function $T(T_s)$, $r_1 \le T_s \le r_2$. The value T_s at which $T(T_s)$ attains a minimum specifies an optimal control.

Figure 2 illustrates this procedure for a particular set of parameters and shows that the function $T(T_s)$ has one minimum, given by the value $T_s = \bar{T}_s \approx 1.04$. This value of T_s minimizes the time to target, and the control function, $\bar{u}(t)$ with $T_s = \bar{T}_s$, equation (4.1), is the optimal control for the set of parameter values shown in the legend of figure 2.

We have also investigated how the switching time depends on various parameters of the system. The following are the results.

— T_s is a monotone, very gradually increasing, concave function of σ for $\sigma > 1$ (recall that $\sigma = M/N$ defines the target colony size relative to the normal colony size, N), see figure 3*a*. The relative switch time, T_s/T, is a decreasing function of σ (not shown).
— T_s is a monotone, decreasing, convex function of the parameter a, the growth rate of mutants, see figure 3*b*.
— The value of T_s is a decreasing function of μ. As μ decreases below the value u_m, the function T_s reaches saturation and does not change much with the value of μ.
— T_s is an increasing function of u_m, such that for $u_m \to 0$, $T_s \to 0$. However, as u_m becomes greater than μ, T_s reaches saturation. Therefore, for the biologically relevant regime of $u_m \gg \mu$, T_s is essentially independent of u_m.

We can see that as long as $u_m \gg \mu$ (a biologically relevant parameter regime), the switching time of the optimal control is effectively independent of the mutation rates μ and u_m. The switching time also does not depend strongly on σ (it is a slowly increasing function of σ). Hence, the switching time essentially depends only on the growth rate of the mutant cells, a. The faster the mutants grow, the sooner the switch happens.

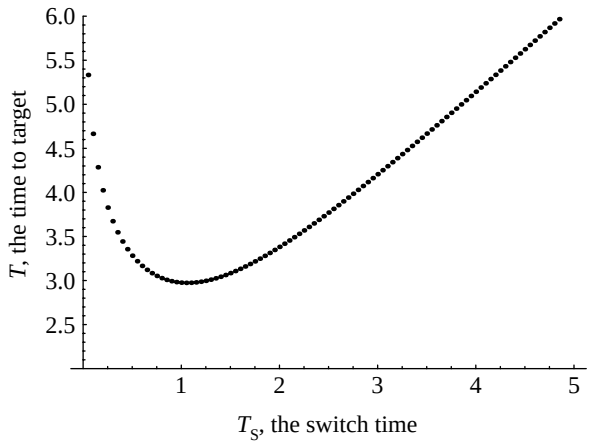

Figure 2. The minimum time to target, T, for different values of the switch time, T_s. The parameters are $a = 2$, $\sigma = 10$, $\alpha = 2$, $u_m = 1$, $\mu = 10^{-7}$.

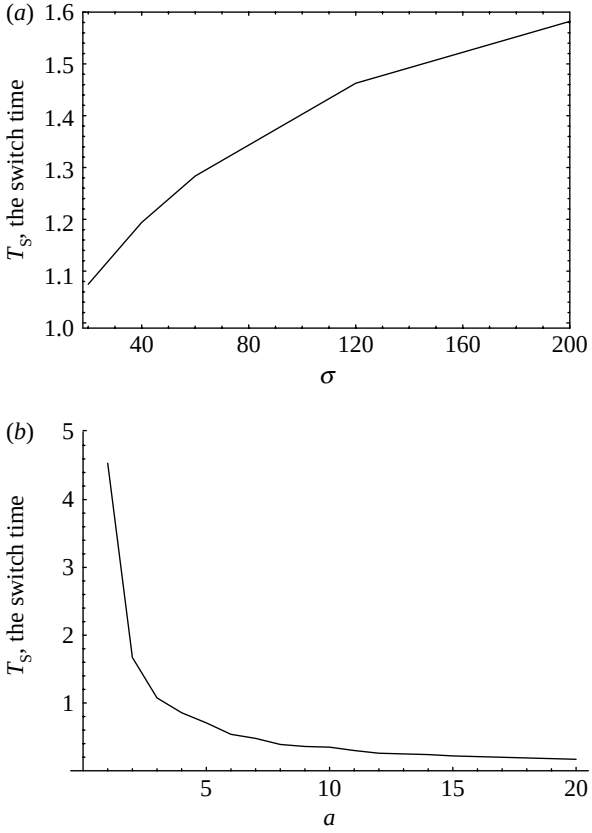

Figure 3. The switch time, T_s, as a function of parameters (*a*) σ and (*b*) a. The other parameters are $a = 2$ and $u_m = 1$ in (*a*), $\sigma = 10$ and $u_m = 10^{-1}$ in (*b*), $\alpha = 1.5$, $\mu = 10^{-5}$.

4.2. Convex death rates, $0 < \alpha < 1$

We now turn to the remaining case of $0 < \alpha < 1$. The death rate $d(u)$ is then a convex function of u. Optimal controls are not bang-bang in this case, as shown in Wan *et al.* (in preparation). This can also be seen from the following numerical experiment. Let us first assume that $u(t)$ is given by formula (4.1) and find the value T_s which minimizes the time it takes for $x_2(t)$ to grow to size 1. The corresponding time, T, tells us how well a bang-bang control does. Now, let us expand the class of

possible control functions to a two-parametric family,

$$u_1(t) = 1 - \frac{1}{2}\left(1 + \tanh\frac{t - T_s}{w}\right), \qquad (4.2)$$

where T_s is the characteristic time of the 'transition' from high to low values of u, and w is the width of this smooth transition. In fact, it is enough to leave only one free parameter, w, and fix T_s to the 'best' switching time obtained for the bang-bang control.

Now, let us vary the parameter w and calculate the time T needed for the colony of mutants to grow to size M. We will obtain a function $T(w)$. The value w corresponding to the minimum T gives the best performance of family (4.2). If the best width is $w=0$, then we can conclude that the sharp, bang-bang-type transition cannot be improved by smoothing it out. However, as figure 4 illustrates for a particular set of parameters, the best control for family (4.2) corresponds to non-zero w. That is, a smooth transition can do better than the best of the bang-bang family. This means that the optimal control is not bang-bang.[2]

The next question is finding the actual optimal control for $\alpha<1$. Solving the boundary value problem with the interior solution again does not work. In fact, one can prove that (unless α is very small) the interior solution, (3.17), fails at and near the terminal time (Wan *et al.* in preparation).

In order to find the optimal control, we have designed the following method.

(i) Start from any control, $0\le u_0(t)\le 1$, e.g. $u_0(t)=1$.
(ii) Solve the initial value problem (3.1)–(3.3) for $0\le t\le T_1$ such that $x_2(T_1)>1$.
(iii) Find the solution, $t=T_0$, of the equation $x_2(t)=1$. This is the zeroth approximation to the best terminal time.
(iv) Solve the boundary value problem (3.8)–(3.11) on $0\le t\le T_0$ with the functions $x_1(t)$, $x_2(t)$ known from the previous step.
(v) Use the obtained state and adjoint variables to calculate the function $u(t)$ by formula (3.17). Call this function $\bar{u}_1(t)$.
(vi) Take $u_1(t)=\bar{u}_1(t)\theta(\bar{u}_1(t))$ (again, θ is the Heaviside step-function). That is, replace all negative values of the control by zero. This gives us the next approximation to the optimal control, $u_1(t)$.
(vii) Go to step (ii) and repeat all the operations.

Numerical evidence suggests that this method converges to a fixed point which is indeed the optimal control. To verify the optimality, we can evaluate the value of the Hamiltonian, equation (3.7), at each step. According to the maximum principle for our problem, the Hamiltonian must vanish for all t, see §3.2.4. A sample run of the algorithm is presented in figure 5*a*, where we plot the logarithm of the value $||H_i|| = \int_0^{T_i} |H(u_i, t)|\mathrm{d}t$ for consecutive iterations. We can see that the Hamiltonian converges to $H(t)=0$, and therefore the corresponding limiting function $u_\infty(t)$ is the optimal control for the problem. Figure 5*b* shows the values of the time to target after each iteration. Figure 6 shows some examples of the optimal control functions, $u_i(t)$ for large i, found by this method for different α values.

[2]If we perform the same operations in the $\alpha\ge 1$ case, we obtain that $w=0$ gives the best result. This of course does not prove that the optimal controls are bang-bang, but it is consistent with the conclusions of §4.1.2.

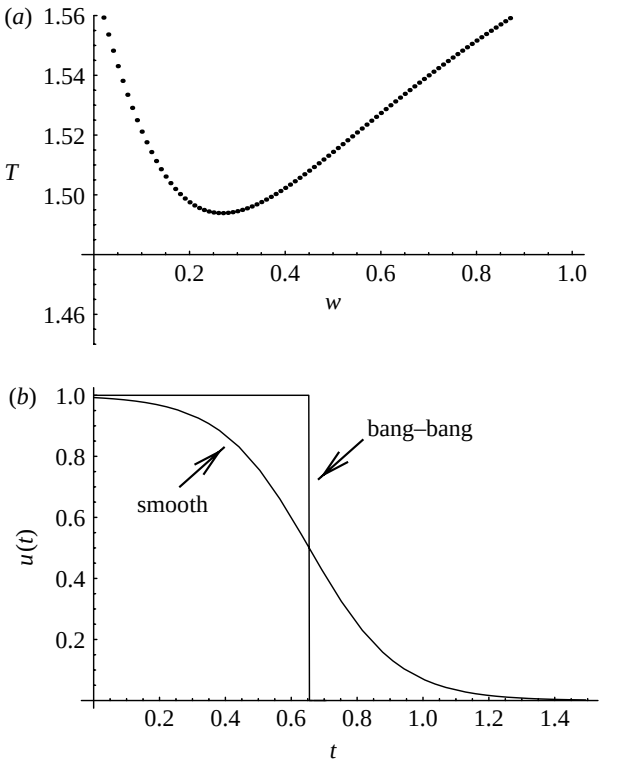

Figure 4. Does bang-bang work in the $\alpha<1$ case? (*a*) The time to target, T, as a function of parameter w in formula (4.2). The other parameters are $a=2$, $\sigma=10$, $\alpha=0.5$, $u_m=1$, $\mu=10^{-1}$. (*b*) The control function, $u_1(t)$, corresponding to the best value of w. Also, the best bang-bang control for these parameters is shown. The time to target for the bang-bang control is $T\approx 1.57$. For the function $u_1(t)$, it is $T\approx 1.49$.

We note that a control found by simply maximizing $x_2(t)$ at each t is not optimal. This control is obtained by solving $\partial g_2/\partial u=0$, see equation (3.2). This gives the value of u maximizing the time-derivative of x_2

$$u = 1 + \left(\frac{u_m x_1}{\alpha\sigma(a - x_1)x_2}\right)^{\frac{1}{\alpha-1}}. \qquad (4.3)$$

We have solved the initial value problem obtained from (3.1)–(3.3) with the above expression for $u(t)$. The solution $x_1(t)$, $x_2(t)$ can be inserted in the expression (4.3). The resulting control is compared with the optimal control in figure 7; we can see that the two functions differ from each other. The performance of the optimal control obtained from the maximum principle is found to be better, as expected.

We have also employed sequential quadratic programming (SQP) algorithm (Gill *et al.* 2005) with direct collocation and automatic differentiation (implemented, respectively, in the software packages SNOPT v. 7 (Gill *et al.* 2005), DIRCOL v. 2.1 (Von Stryk 2000) and ADIFOR v. 2.0 (ADIFOR 1994)) to find an accurate approximation to the optimal solution. A sample of such numerical solutions is shown in figure 8*a*. We see that as $\alpha\to 1$ from below, the transition from $u=1$ to $u=0$ becomes sharper and sharper. The optimal control becomes bang-bang control as $\alpha\to 1$ from below.

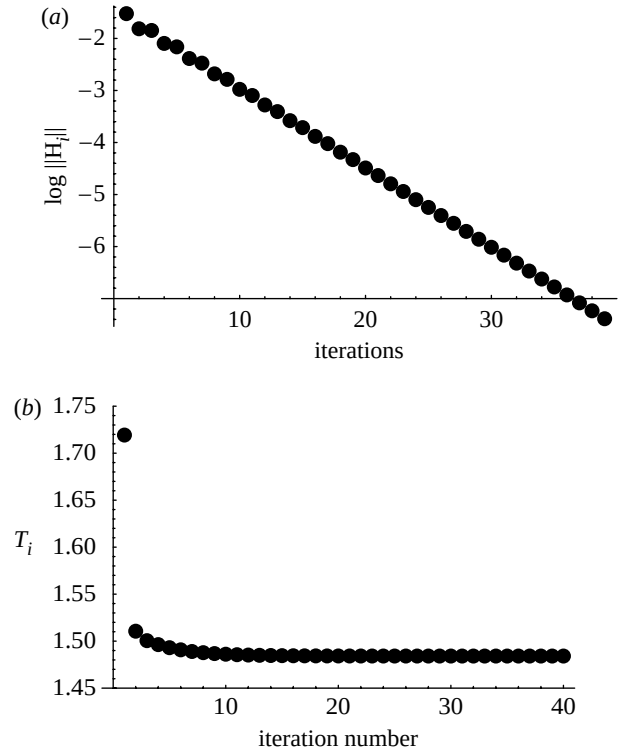

Figure 5. The iterative algorithm to find the optimal control in the $\alpha<1$ case. (a) The value $H_i=\int_0^{T_i}|H(u_i,t)|\mathrm{d}t$ for consecutive iterations. (b) The iterations of the minimum time to target, T. The parameters are $a=2$, $\sigma=2$, $\alpha=0.5$, $u_\mathrm{m}=1$, $\mu=10^{-1}$.

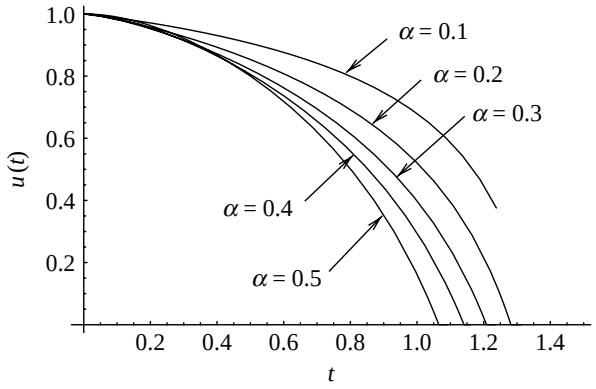

Figure 6. The optimal control found by the iterative method for $\alpha<1$. The optimal functions for five values of α are presented. The optimal time to target is $T=1.238$ for $\alpha=0.1$, 1.330 for $\alpha=0.2$, 1.395 for $\alpha=0.3$, 1.445 for $\alpha=0.4$ and 1.484 for $\alpha=0.5$. The parameters are $a=2$, $\sigma=2$, $u_\mathrm{m}=1$, $\mu=10^{-1}$.

4.3. The two-step problem

Most of the qualitative conclusions obtained for the one-step problem also hold for the two-step problem. Namely, for $\alpha\geq1$ we have a bang-bang optimal control, and for $0<\alpha<1$ we have an interior solution for part of the domain $[0,\ T]$. This can be demonstrated again by comparing the best bang-bang solution with a family of continuous functions $u(t)$ with a finite width. Note that an improvement on the bang-bang control for $0<\alpha<1$ may not be found in the class of functions defined by equation (4.2). In some instances, we have used the function (4.2) multiplied by some small power of t. In all cases, a non-zero width $w>0$ gives a better performance than the bang-bang control.

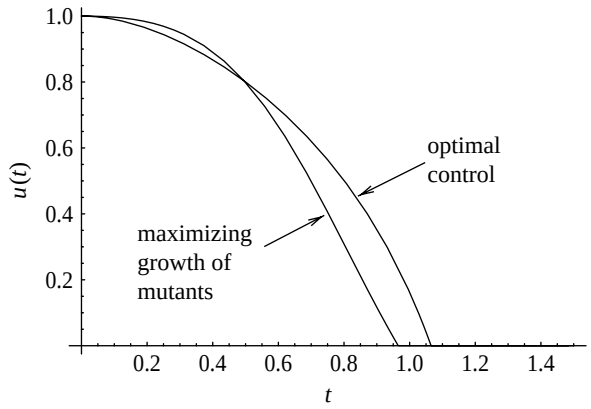

Figure 7. The optimal control found by the iterative method, compared to the function $u(t)$ found from formula (4.3), which maximizes the growth of $x_2(t)$. The parameters are $a=2$, $\alpha=0.5$, $\sigma=2$, $u_\mathrm{m}=1$, $\mu=10^{-1}$.

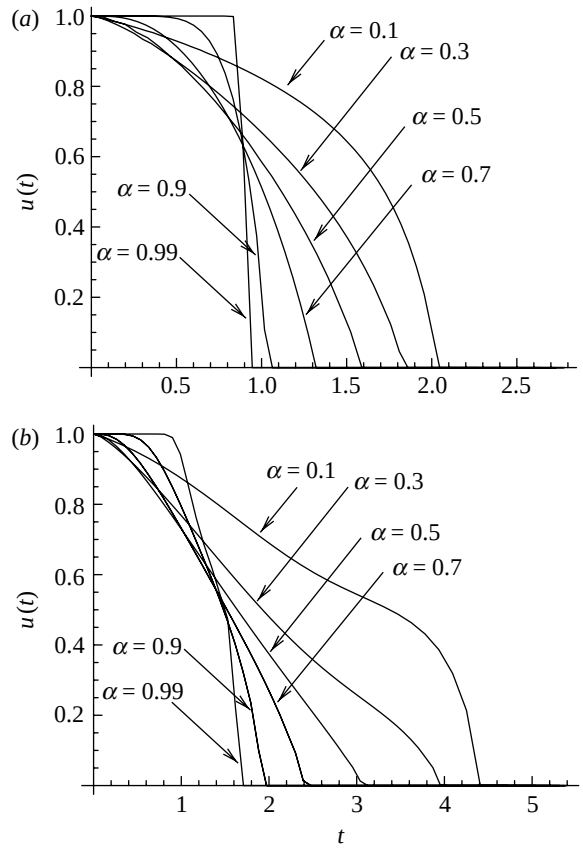

Figure 8. The optimal control found by the SQP method, for different values of $\alpha<1$, for (a) a one-step process and (b) a two-step process. The other parameters are $a=2$, $\sigma=10$, $u_\mathrm{m}=1$, $\mu=10^{-1}$.

The iterative method developed above for approximating the optimal control for the one-step model does not work for the two-step problem. There, we do not observe a convergence of the algorithm to a fixed point. Instead, we used SQP algorithm to find the solution. An example is shown in figure 8*b*, where the optimal control is found numerically for different values of $\alpha<1$. As in the case of a one-step process, the control becomes steeper as $\alpha\to1$.

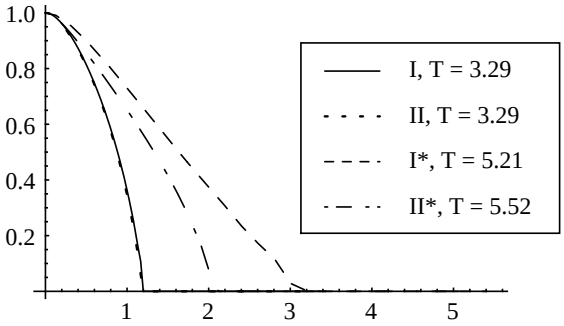

Figure 9. A comparison of the optimal controls obtained for models I, II, I* and II*, with parameters $a=2$, $\sigma=10$, $u_m=1$, $\mu=0.1$, $\alpha=0.5$. The optimal control is plotted as a function of time.

Table 2. Optimal controls (of the bang-bang type) for Models I, II, I* and II*, for parameters $a=2$, $\sigma=10$, $u_m=1$, $\mu=0.1$ and $\alpha=1.5$.

model	I	II	I*	II*
T_S	0.75	0.70	1.36	1.10
T	3.36	3.37	5.40	5.68

Robustness of the model has been checked. We have compared the qualitative shapes of the optimal controls for Models I, II, I* and II*, formulated in §2. For all the models, the optimal control with $\alpha \geq 1$ is of the bang-bang type, and for $\alpha<1$ it is a continuous monotonically decaying function of time, starting at $u=1$ at $t=0$. Figure 9 presents the optimal controls found by SQR algorithm for the four models, for $\alpha=0.5$ (and all other parameters taken the same for all models). The optimal control (of the bang-bang type) for $\alpha=1.5$ is presented for the four cases in table 2. We can see that there are quantitative differences, but the qualitative behaviour is the same.

5. THE OPTIMAL STRATEGY FOR CANCER

In what follows, we drop the restriction $d_m=1$ in formula (2.9) and investigate the two-dimensional parameter subspace of the problem, (u_m, d_m). The magnitude u_m tells us the maximum mutation rate (due to genetic instability) that is possible in the system. The value d_m characterizes the maximum magnitude of the death rate associated with the instability.

Let us first take $\alpha<1$, fix a value $u_m<1$ and find an optimal control $\bar{u}(t)$ for several values of the magnitude of the death rate, d_m. Figure 10 presents the functions $\bar{u}(t)$ obtained by means of the method described in §4.2 We have also used the SQP algorithm to double-check the results. Figure 10*a* shows several runs for $\alpha=0.1$, and figure 10*b* shows several runs for $\alpha=0.9$. We observe the following trend: as d_m becomes smaller, the optimal controls become larger. In other words, for small values of the death rate, optimal strategies tend to favour large values of $u(t)$ for a longer initial period of time. For example, in the case $\alpha=0.1$ the transition to $u=0$ does not happen for $d_m<0.001$. On the other hand, for large values of d_m, optimal controls decrease

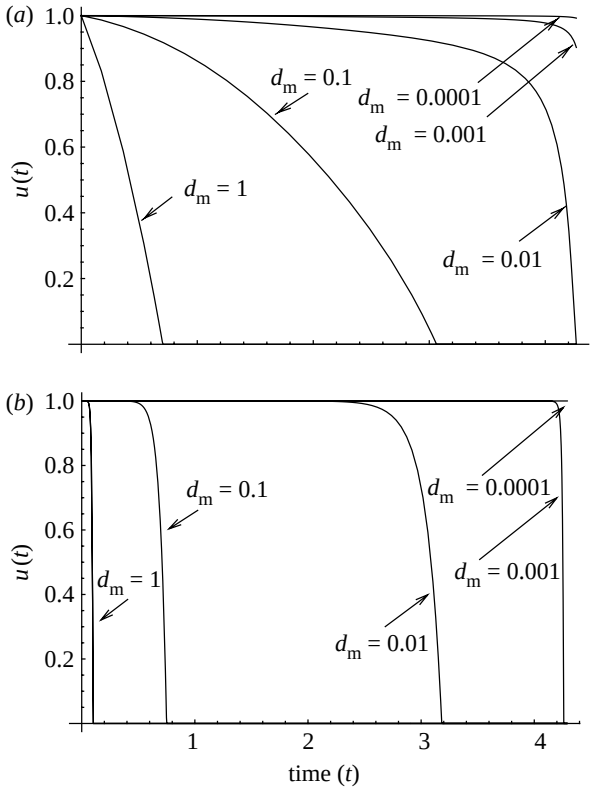

Figure 10. The optimal controls $\bar{u}(t)$ for $\alpha<1$ and various values of d_m. (*a*) $\alpha=0.1$, the optimal times to target are 4.37, 4.35, 4.28, 4.27 and 4.27 for $d=1$; 0.1; 0.01; 0.001 and 0.0001, respectively. (*b*) $\alpha=0.9$, the optimal times to target are 4.36, 4.31, 4.28, 4.27 and 4.27. The other parameters are $a=2$, $u=0.1$, $\sigma=10$ and $u_m=0.01$.

continuously from their maximum value to zero very early in the growth process. Note that increasing d_m to values larger than 1 will lead to an even sharper decrease in $\bar{u}(t)$.

Next, we turn to the values $\alpha \geq 1$. In this case, the results do not depend on the actual value of α; optimal controls were found (Wan *et al.* in preparation) to be bang-bang with exactly one switching. In other words, the result that no feasible interior solution exists for $\alpha \geq 1$ extends to the case $d_m \neq 1$. The function $\bar{u}(t)$ assumes only one or both of the values $u=1$ and $u=0$ (with no more than one switching), and the value of the death rate, $d(\bar{u})$, satisfies $d(0)=0$ and $d(1)=1$ for all $\alpha \geq 1$. This simplifies the problem because the form of optimal controls for all $\alpha \geq 1$ is the same, and it is enough to perform simulations for just one value $\alpha \geq 1$, say, $\alpha=1$.

We therefore take $u(t)$ to be of the bang-bang form, given by equation (4.1), and find the optimal switching time, T_s, for various points in the (u_m, d_m) plane. Figure 11 presents a two-dimensional density plot of the quantities T_s/T, the relative switching time, for various pairs (u_m, d_m). The lighter shades correspond to smaller values of the relative switching time. The corresponding diagram for $\alpha \geq 1$ in the (u_m, d_m) parameter subspace for the two-step model looks qualitatively the same and is not presented here.

As d_m changes, the behaviour of the optimal controls in the $\alpha \geq 1$ case follows the same trend as in the case $\alpha<1$. Namely, for large values of d_m, the period of time

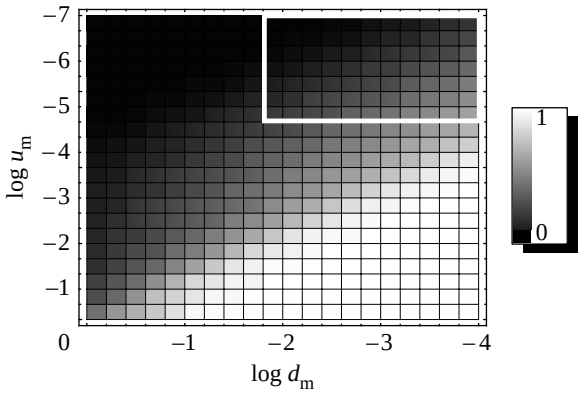

Figure 11. The relative switching time, T_s/T, for the optimal strategy, depending on the parameters u_m and d_m. Black corresponds to $T_s/T=0$ (an immediate switching) and white to $T_s/T=1$ (no switching). The parameters are $a=2$, $\alpha\geq 1$, $\sigma=10$, $\mu=10^{-4}$.

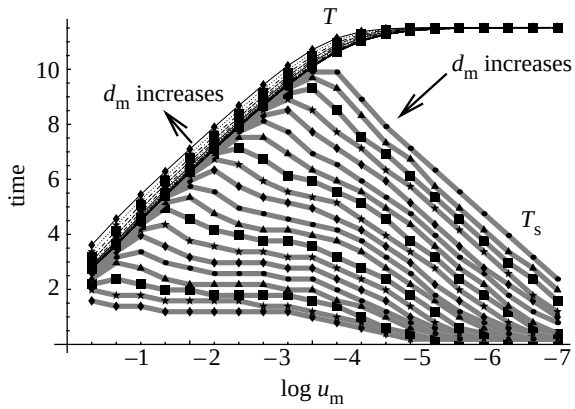

Figure 12. The optimal time to target, T (black lines), and the corresponding switching time, T_s (grey lines), as functions of u_m for different values of d_m. The values of log d_m vary from -4 to 0, the direction of the increase of d_m is indicated. The other parameters are the same as in figure 11.

where the maximum value of u is advantageous becomes shorter. We now present a mathematical explanation together with a biological interpretation of the various regions in the parameter space of figure 11, with $\alpha\geq 1$.

(i) Small values of u_m and large values of d_m correspond to $T_s/T\ll 1$, the regime where genetic instability is never advantageous, except in a very short time-interval at the beginning of the growth (the black cells in figure 11). This is because the small gain in the mutation rate (u_m) is not worth risking the penalty (the large death rate), and the cells are better off without the instability. Mathematically, there is one switching from $u=1$ to $u=0$ in the function $\bar{u}(t)$, but this switching happens so early in the growth that, biologically, the initial, unstable, period of time (with $u=1$) is negligible, and the cells are characterized by low mutation rates at all times.

(ii) The opposite situation arises when instability is advantageous and it does not become a liability for the entire duration of the growth (until the colony reaches size M, see white cells in figure 11). This happens when u_m is large and d_m is small: a small penalty for a large gain in the mutation rate. In this regime, $|1-T_s/T|\ll 1$, that is, for most, or all, of the time the optimal control is $\bar{u}(t)=1$.

(iii) The grey cells in figure 11 correspond to intermediate values of T_s, such that T_s/T and $[1-T_s/T]$ are not too small. In this regime, the optimal strategies are one-switch bang-bang controls with an unstable strategy advantageous at first and then becoming disadvantageous, prior to reaching the target. The switching occurs at some intermediate time, neither at the very start nor close to the end of the growth. This regime characterizes the middle portion of the parameter space.

(iv) Finally, we have the white rectangle in figure 11 which corresponds to small values of d_m and u_m. In such cases, the shape of optimal controls is bang-bang with an intermediate switch, similar to case (iii) above. However, in this regime, changes in $u(t)$ affect the time to target, T, only very slightly. No matter what the shape of $u(t)$ is, the changes in the mutation rate and the death rate are bounded by very small u_m and d_m, respectively, and cannot significantly alter the solutions of the ODEs. Inside the white rectangle in figure 11, the difference between the maximum and the minimum time to target, T, as a function of T_s, is less than 1%. In this (biologically irrelevant) region of the parameter space, genetic instability is unimportant. The results for this region are included for mathematical completeness only.

A more quantitative picture for optimal controls in the $\alpha\geq 1$ case is presented in figure 12. There, we plot the optimal time to target, T (thin black lines), and the corresponding switching time, T_s (thick grey lines), as functions of u_m, for different values of d_m. The lines corresponding to T are monotonically increasing, concave functions of log u_m; the different lines correspond to different values of d_m; the direction in which d_m increases is marked by an arrow. The lines corresponding to T_s have a one-hump shape; again, different lines correspond to different d_m.

For large values of u_m and small values of d_m, the lines for T and T_s coincide. In this case, no switching is observed and instability is advantageous at all times. As u_m decreases, the value T_s tends towards zero. This means that stability throughout the growth is observed. Intermediate values of u_m and d_m correspond to a switching sometime in the course of growth, not too close to $t=0$ or $t=T$.

6. DISCUSSION

6.1. Summary

We have examined optimal strategies for a cancerous colony with respect to the magnitude of the mutation rate, as the colony acquires carcinogenic mutations and enters a phase of a clonal expansion. Biologically, a large mutation rate corresponds to genetic instability

and a small mutation rate to stability of cancerous cells. The two types of carcinogenic mutations that we considered are activation of an oncogene (the one-step model) or inactivation of a TSG (the two-step model).

In order for a cancer to progress, a cell colony first has to generate carcinogenic mutants and then to grow. Genetic instability may expedite the first of these processes and slow down the second. Therefore, this process can be examined as an optimization problem.

Genetic instability is 'blind', i.e. it does not necessarily 'hit' the exact genes necessary for cancer progression. It may cause defects in other genes thus creating deleterious cells. The question is whether the gain in progression speed due to the increased mutation rate would outweigh the losses suffered by the cells as a result of spurious, deadly mutations created by the instability. In order to model this, we introduce two parameters: u_{m}, the maximum rate of mutations; and d_{m}, the magnitude of the death rate. Small values of u_{m} mean a small gain in creating cancerous mutations. Large values of d_{m} mean a large penalty paid by the colony as a result of many mutation-related deaths.

First we examined in detail a subset of the $(u_{\mathrm{m}}, d_{\mathrm{m}})$ parameter space, namely, $d_{\mathrm{m}}=1$. We found that large mutation rates at first and lower mutation rates later on constitute the optimal strategy. The exact shape of the optimal mutation strategy depends on the concavity of the function $d(u)$, the instability-dependent death rate. We distinguish two cases. For non-convex death rates (§4.1), the best performance is achieved if the mutation rate jumps (in a discontinuous, abrupt fashion) from maximum to minimum. For convex death rates (§4.2), the transition in an optimal control is gradual. In both cases, having the highest possible mutation rate is advantageous at first; later on, it pays off to switch to a lower mutation rate.

We also performed numerical simulations to find optimal strategies for all biologically reasonable values of u_{m} and d_{m}. We found three qualitatively different forms of optimal strategies. In one scenario, the instability makes a minimal contribution to creating carcinogenic mutations (small u_{m}), but significantly increases the death rate (large d_{m}). Consequently, it does not pay to be unstable at any stage of the growth in this case. At the other extreme (large u_{m} and small d_{m}), instability is useful and it 'comes cheap'; in other words, the death toll paid by the affected cells is small. Therefore, the colony is better off being unstable at all times. Finally, between these two extremes, optimal controls start at the maximum admissible mutation rate and then drop to the lowest possible value at some time during the dynamics. This corresponds to genetic instability being advantageous at the beginning and becoming a liability later. This explains the growing experimental literature suggesting that tumours switch from genetic instability to stability some time in the course of cancer progression.

The advantage of our approach is that several results can be obtained analytically. These first simple models can be extended in many ways to include more information about the biological reality. For instance, the models have a stochastic version. That is, instead of average quantities, one could use probability distributions. More details about specific mutations can be included if one chooses to focus on a particular case study. Various additional constraints on the strategies can be imposed, reflecting the dependence of the mutation rate on other characteristics of the system, e.g. the system size. Finally, the parameterization of the death rate as a function of the instability can be generalized. However, the general trends found in the simple models are biologically intuitive and experimentally supported. With possible non-principal modifications, they are likely to persist in certain more sophisticated scenarios.

6.2. Does cancer solve an optimization problem?

Not literally, of course. However, by solving this problem, one can obtain valuable information about the growth of cancer. This is similar to the general philosophy of the evolutionary game theory, as, for example, in Maynard Smith (1982). In that paper, different strategies are played against each other to see which one wins. In principle, one could design an 'ideal' strategy which leads to the maximal pay-off in the game. The game can be a model of something that happens in nature, for instance the behaviour of animals in different situations or adaptations of cells in various environments. The ideal (optimal) strategy may not even be realistic (there are many constraints in nature which escape modelling, but can make a strategy impossible). What occurs in reality, however, tends to approximate an optimal strategy. Finding the 'evolutionary stable strategy' or the 'Nash equilibrium' in the system helps us understand the general experimentally observed trend. A plausible explanation for the survival of those animals is that they have won the evolutionary game against other animals that used an inferior strategy.

In the present paper, we use similar ideas. We solve the optimization problem for cancerous growth and find optimal strategies. Does cancer always use an optimal strategy? Probably not. One obstacle to optimality is that cancer is unable to adjust its level of instability instantaneously throughout the entire population to optimize the growth. However, a cancer which follows the general trend, i.e. a strategy close to an optimal one, will grow faster. These are the colonies that 'succeed' causing a disease. Other pre-cancerous colonies of cells might exist in any organ of an organism, but if they do not use a strategy sufficiently close to optimal, they do not succeed with further growth and are not observed.

6.3. Implications for somatic evolution of cancer

In this paper, we formulate the problem of cancer growth from the point of view of cancerous cells, in order to find the most optimal mutation strategy. Here we discuss our model in the context of somatic evolution of cells.

Selection in this problem takes place on two levels. One level is the level of individual cells, where normal cells are competing for space and nutrients (this is one of the mechanisms of homeostatic control), and cancerous cells escape this control to enter a phase of

exponential growth. The forces of selection are manifested in the nonlinearities of the basic ODEs. The second level of selection is the level of cell colonies. Different colonies are characterized by different functions $u(t)$. They are not assumed to be in direct interaction with one another. The competition in this case manifests itself in whether a given colony will reach a cancerous state quickly and become observable. The mathematical problem we solve is to find the optimal strategy $u(t)$, such that the colony with this strategy will be the first one to 'make it'.

Several other papers have proposed evolutionary models of genetic instability. An important component of many such models is costs and benefits of cell repair, (Breivik & Gaudernack 1999, 2004; Breivik 2001; Komarova & Wodarz 2003; Breivik 2005). It is argued that an evolutionary reason for instability is the fact that cell repair is costly, and under some circumstances it may pay off to avoid repair, which leads to genomic instability. In this paper, we do not focus on the detailed analysis of costs and benefits of cell repair; instead, we concentrate on their effect on the choice of the most advantageous strategy with respect to the rate of chromosome loss. However, the model includes the costs of repair in the following indirect way. Cells with efficient repair have a disadvantage because repair is costly (having to enter cell cycle arrest in order to repair the genome decreases the growth rate of the population). Cells with inefficient repair have a disadvantage of creating deleterious mutations. The difference between the two, that is, the relative disadvantage of instability, is captured in the quantity $d(u)$. The function $d(u)$ reduces the fitness of unstable cells with respect to stable cells. We assume that this quantity is positive, due to the large damage inflicted by losses of 'wrong' chromosomes.

6.4. Suggestions for experiments

The theory developed here would greatly benefit from further experiments that would aim at quantification of the processes of tumour growth and mutation. We propose two types of experiments.

The rate of instability as a function of the tumour stage. Despite many reports that the degree of genetic instability goes down as the tumour progresses (Rudolph *et al.* 2001; Chin *et al.* 2004), a quantification of this phenomenon is still lacking. The rate of chromosomal loss has been measured by Lengauer *et al.* (1997). In a similar way, a thorough study of the 'natural history' of genetic instability can be performed. A systematic study of cells harvested at different stages of tumorigenesis would yield a curve, where the mutation rate is a function of the tumour age. This study can be done in a controlled manner with animal models.

The death rate of cells as a function of the mutation rate. In our mathematical model, we postulate that the death rate of cells is a function of the cells' mutation rate. In the presence of instability, we assume that the death rate is elevated; this can be quantified experimentally. In the last several years, much work has been done to understand the role instability plays in tumours, in particular, in the way it affects the death of cells (Lowe *et al.* 2004; Bartkova *et al.* 2005; Chen *et al.* 2005; Gorgoulis *et al.* 2005; Deng 2006). However, a detailed measurement of the death rate as a function of a mutation rate has not been performed. One way to do this is as follows, similar in spirit to previous work (Marder & Morgan 1993; Nagar *et al.* 2003; Smith *et al.* 2003) on cells treated by radiation. As the amount of radiation increases, the mutation rate increases and so does the cell death. The dependence of the cell death on the mutation rate could serve to improve our model of the function $d(u)$, and help us reason about optimal strategies of tumours. Another way to quantify the dependence of death on the degree of genetic instability is to measure the death rate together with the degree of instability in a series of experiments with cells harvested at different stages of tumour growth.

N.L.K.'s research was supported by the Sloan Fellowship, and NIH grants 1R01AI058153-01A2, 1R01 CA118545-01A1 and R01GM075309. A.S.'s research was supported by the NIH National Research Service Award 5 T15 LM007443 from the National Library of Medicine, awarded to P. Baldi. The research of F.Y.M.W. was supported by NIH grants R01GM067247, R01GM075309 and UC Irvine research grant 445861 (the two NIH R01 grants were awarded through the Joint NSF/NIGMS Initiative to Support Research in the Area of Mathematical Biology). We thank P. Gill for help with obtaining and using SNOPT v. 7.0, M. Fagan for help with obtaining ADIFOR v. 2.0 and O. Von Stryk for help with obtaining DIRCOL v. 2.1. Thanks are also due to A. Vladimirsky for a helpful discussion of computational approaches to finding time-optimal controls.

APPENDIX A. THE TWO-STEP PROBLEM: THE MATHEMATICAL FORMALISM

A.1. Statement of the problem

A.1.1. Equations of state. Let $x_0(t)$ be the population size of normal cells and $x_1(t)$ be the population size of harmless mutants with only one functioning copy of its TSG, both normalized by the initial normal cell population size N. Let $x_2(t)$ be the population size of dangerous TSG mutant cells normalized by its final target population size M. With $p_{\min}=0$, the time evolution of the three cell populations is governed by the following three first-order ODEs:

$$x_0' = -2\mu x_0 + x_0(1-d(u))(1-x_0-x_1) \equiv g_0(x_0,x_1,x_2,u), \tag{A 1}$$

$$x_1' = 2\mu x_0 - (\mu + u_{\mathrm{m}}u)x_1 + x_1(1-d(u))(1-x_0-x_1) \equiv g_1(x_0,x_1,x_2,u), \tag{A 2}$$

$$x_2' = \frac{1}{\sigma}(\mu + u_{\mathrm{m}}u)x_1 + x_2(1-d(u))(a-x_0-x_1) \equiv g_2(x_0,x_1,x_2,u), \tag{A 3}$$

again with $()' = \mathrm{d}()/\mathrm{d}t, \sigma = M/N \gg 1, a > 1, 0 < u_{\mathrm{m}} \leq 1$, $0 \leq \mu \ll 1$, and the death rate (2.9) where the non-negative normalized *gross chromosomal change rate* u is limited by the inequality constraint (3.4). System (A 1)–(A 3) corresponds to system (2.3), (2.4) and (2.8) of §2.2 of the main text.

A.1.2. Boundary conditions. The three equations of state (A 2)–(A 4) are subject to the following four auxiliary conditions:

$$x_0(0)=1, \quad x_1(0)=0, \quad x_2(0)=0, \quad x_2(T)=1, \tag{A 4}$$

where T is the time when the dangerous mutant cell population reaches the target size.

A.1.3. Problem. Choose the control function $u(t)$ to minimize the time T needed to reach the target population size of the dangerous mutants subject to the inequality constraint (3.4), on the control $u(t)$ and the non-negative constraints

$$x_k \geq 0, \qquad k=0,1,2, \tag{A 5}$$

on the cell populations. The problem can be restated as choosing the control function $u(t)$ to minimize the performance index J defined in (3.6) subject to the equations of state (A 1)–(A 3), the boundary conditions (A 4), the inequality constraints (3.4) and (A 5) with $T>0$ as a part of the solution.

A.2. The Hamiltonian and adjoint variables

The *Hamiltonian* for the two-step problem is

$$H = 1+\lambda_0 g_0+\lambda_1 g_1+\lambda_2 g_2, \tag{A 6}$$

where λ_k are the three adjoint variables for the problem described by three adjoint ODEs,

$$\lambda_k' = -\left(\lambda_0\frac{\partial g_0}{\partial x_k}+\lambda_1\frac{\partial g_1}{\partial x_k}+\lambda_2\frac{\partial g_2}{\partial x_k}\right), \quad (k=0,1,2), \tag{A 7}$$

and (for the given auxiliary conditions on the state variable (x_k)) two transversality conditions

$$\lambda_0(T)=\lambda_1(T)=0. \tag{A 8}$$

For the optimal solution of our minimum terminal time problem, it is customary to investigate first the possibility of an *interior* solution, by seeking

(i) a control $u=u_{\mathrm{int}}(t)$ to satisfy the optimality condition (3.14),
(ii) six quantities $\{x_i(t),\lambda_j(t)\}$ to satisfy the six differential equations (A 1)–(A 3) and (A 7) and six auxiliary conditions in (A 4) and (A 8), and
(iii) the terminal time T to satisfy a free end condition which again becomes (3.11) after simplification by the adjoint boundary condition (A 8).

For the state equations (A 1)–(A 3), the three adjoint differential equations (A 7) take the form

$$\lambda_0' = -\lambda_0\{(1-2x_0-x_1)(1-d)-2\mu\} - \lambda_1\{2\mu-x_1(1-d)\}+\lambda_2 x_2(1-d), \tag{A 9}$$

$$\lambda_1' = \lambda_0(1-d)x_0-\lambda_1\{(1-x_0-2x_1)(1-d) -(u_{\mathrm{m}}u+\mu)\}-\frac{\lambda_2}{\sigma}\{(u_{\mathrm{m}}u+\mu)-\sigma x_2(1-d)\}, \tag{A 10}$$

$$\lambda_2' = -\lambda_2(a-x_0-x_1)(1-d). \tag{A 11}$$

Consider the interior control $u(t)$ determined by the optimality condition (3.15), $\partial H/\partial u=0$, with

$$\frac{\partial H}{\partial u}=\frac{u_{\mathrm{m}}x_1}{\sigma}(\lambda_2-\sigma\lambda_1)-\{(\lambda_0x_0+\lambda_1x_1)(1-x_0-x_1) +\lambda_2x_2(a-x_0-x_1)\}d^{\cdot}, \tag{A 12}$$

where $d^{\cdot}=\alpha(1-u)^{\alpha-1}$ for the death rate (2.9). We have from (3.14)

$$d^{\cdot}=\alpha(1-u)^{\alpha-1} = \frac{x_1\{\lambda_2-\lambda_1\sigma\}}{\sigma_{\mathrm{m}}\{(\lambda_0x_0+\lambda_1x_1)(1-x_0-x_1)+\lambda_2x_2(a-x_0-x_1)\}}. \tag{A 13}$$

REFERENCES

ADIFOR 1994 ADIFOR 2.0: automatic differentiation of FORTRAN. A collaborative project between the Mathematics and Computer Science Division at Argonne National Laboratory and the Center for Research on Parallel Computation at Rice University, http://www-unix.mcs.anl.gov/autodiff/ADIFOR/.

Artandi, S. E. & DePinho, R. A. 2000 A critical role for telomeres in suppressing and facilitating carcinogenesis. *Curr. Opin. Genet. Dev.* **10**, 39–46. (doi:10.1016/S0959-437X(99)00047-7)

Bailey, S. M. & Murnane, J. P. 2006 Telomeres, chromosome instability and cancer. *Nucleic Acids Res.* **34**, 2408–2417. (doi:10.1093/nar/gkl303)

Bardelli, A., Cahill, D. P., Lederer, G., Speicher, M. R., Kinzler, K. W., Vogelstein, B. & Lengauer, C. 2001 Carcinogen-specific induction of genetic instability. *Proc. Natl Acad. Sci. USA* **98**, 5770–5775. (doi:10.1073/pnas.081082898)

Bartkova, J. *et al.* 2005 DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864–870. (doi:10.1038/nature03482)

Benhamou, S. & Sarasin, A. 2000 Variability in nucleotide excision repair and cancer risk: a review. *Mutat. Res.* **462**, 149–158. (doi:10.1016/S1383-5742(00)00032-6)

Boltyanskii, V. G., Gamkrelidze, R. V., Mishchenko, E. F. & Pontryagin, L. S. 1962 *The mathematical theory of optimal processes.* New York, NY: Wiley.

Breivik, J. 2001 Don't stop for repairs in a war zone: Darwinian evolution unites genes and environment in cancer development. *Proc. Natl Acad. Sci. USA* **98**, 5379–5381. (doi:10.1073/pnas.101137698)

Breivik, J. 2005 The evolutionary origin of genetic instability in cancer development. *Semin. Cancer Biol.* **15**, 51–60. (doi:10.1016/j.semcancer.2004.09.008)

Breivik, J. & Gaudernack, G. 1999 Genomic instability, DNA methylation, and natural selection in colorectal carcinogenesis. *Semin. Cancer Biol.* **9**, 245–254. (doi:10.1006/scbi.1999.0123)

Breivik, J. & Gaudernack, G. 2004 Resolving the evolutionary paradox of genetic instability: a cost-benefit analysis of DNA repair in changing environments. *FEBS Lett.* **563**, 7–12. (doi:10.1016/S0014-5793(04)00282-0)

Bryson, A. E. & Ho, Y. C. 1969 *Applied optimal control.* Waltham, MA: Ginn and Company.

Cahill, D. P., Lengauer, C., Yu, J., Riggins, G. J., Willson, J. K., Markowitz, S. D., Kinzler, K. W. & Vogelstein, B. 1998 Mutations of mitotic checkpoint genes in human cancers. *Nature* **392**, 300–303. (doi:10.1038/32688)

Cahill, D. P., Kinzler, K. W., Vogelstein, B. & Lengauer, C. 1999 Genetic instability and darwinian selection in tumours. *Trends Cell Biol.* **9**, 57–60. (doi:10.1016/S0962-8924(99)01661-X)

Chen, Z. *et al.* 2005 Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* **436**, 725–730. (doi:10.1038/nature03918)

Chin, K. *et al.* 2004 *In situ* analyses of genome instability in breast cancer. *Nat. Genet.* **36**, 984–988. (doi:10.1038/ng1409)

Deng, C.-X. 2006 BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Res.* **34**, 1416–1426. (doi:10.1093/nar/gkl010)

de Pillis, L. G., Gu, W., Fister, K. R., Head, T., Maples, K., Murugan, A., Neal, T. & Yoshida, K. In press. Chemotherapy for tumors: an analysis of the dynamics and a study of quadratic and linear optimal controls, *Math. Biosci.*

Eigen, M. & Schuster, P. 1979 *The hypercycle: a principle of natural self-organization.* Berlin, Germany: Springer.

Fodde, R. *et al.* 2001 Mutations in the APC tumour suppressor gene cause chromosomal instability. *Nat. Cell Biol.* **3**, 433–438. (doi:10.1038/35070129)

Gelfand, I. M. & Fomin, S. V. 1963 *Calculus of variations.* New Jersey, NJ: Prentice Hall.

Gill, P., Murray, W. & Saunders, M. A. 2005 SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM Rev.* **47**, 99–131. (doi:10.1137/S0036144504446096)

Gonzalez-Suarez, E., Samper, E., Flores, J. M. & Blasco, M. A. 2000 Telomerase-deficient mice with short telomeres are resistant to skin tumorigenesis. *Nat. Genet.* **26**, 114–117. (doi:10.1038/79089)

Gorgoulis, V. G. *et al.* 2005 Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **434**, 907–913. (doi:10.1038/nature03485)

Iwasa, Y. & Levin, S. A. 1995 The timing of life history events. *J. Theor. Biol.* **172**, 33–42. (doi:10.1006/jtbi.1995.0003)

Kinzler, K. W. & Vogelstein, B. 1996 Lessons from hereditary colorectal cancer. *Cell* **87**, 159–170. (doi:10.1016/S0092-8674(00)81333-1)

Kinzler, K. W. & Vogelstein, B. 2002 *The genetic basis of human cancer*, 2nd edn. New York, NY: McGraw-Hill.

Kirschner, D., Lenhart, S. & Serbin, S. 1997 Optimal control of the chemotherapy of HIV. *J. Math. Biol.* **35**, 775–792. (doi:10.1007/s002850050076)

Komarova, N. L. 2004 Does cancer solve an optimization problem? *Cell Cycle* **3**, 840–844.

Komarova, N. L. & Wodarz, D. 2003 Evolutionary dynamics of mutator phenotypes in cancer: implications for chemotherapy. *Cancer Res.* **63**, 6635–6642.

Komarova, N. L. & Wodarz, D. 2004 The optimal rate of chromosome loss for the inactivation of tumor suppressor genes in cancer. *Proc. Natl Acad. Sci. USA* **101**, 7017–7021. (doi:10.1073/pnas.0401943101)

Komarova, N. L., Lengauer, C., Vogelstein, B. & Nowak, M. A. 2002 Dynamics of genetic instability in sporadic and familial colorectal cancer. *Cancer Biol. Ther.* **1**, 685–692.

Komarova, N. L., Sengupta, A. & Nowak, M. A. 2003 Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. *J. Theor. Biol.* **223**, 433–450. (doi:10.1016/S0022-5193(03)00120-6)

Lengauer, C., Kinzler, K. W. & Vogelstein, B. 1997 Genetic instability in colorectal cancers. *Nature* **386**, 623–627. (doi:10.1038/386623a0)

Lengauer, C., Kinzler, K. W. & Vogelstein, B. 1998 Genetic instabilities in human cancers. *Nature* **396**, 643–649. (doi:10.1038/25292)

Lenhart, S. & Workman, J. T. 2007 *Optimal control applied to biological models.* Boca Raton, FL: Chapman and Hall/CRC.

Li, R., Sonik, A., Stindl, R., Rasnick, D. & Duesberg, P. 2000 Aneuploidy vs. gene mutation hypothesis of cancer: recent study claims mutation but is found to support aneuploidy. *Proc. Natl Acad. Sci. USA* **97**, 3236–3241. (doi:10.1073/pnas.040529797)

Little, M. P. & Wright, E. G. 2003 A stochastic carcinogenesis model incorporating genomic instability fitted to colon cancer data. *Math. Biosci.* **183**, 111–134. (doi:10.1016/S0025-5564(03)00040-3)

Loeb, L. A., Springgate, C. F. & Battula, N. 1974 Errors in DNA replication as a basis of malignant changes. *Cancer Res.* **34**, 2311–2321.

Lowe, S. W., Cepero, E. & Evan, G. 2004 Intrinsic tumour suppression. *Nature* **432**, 307–315. (doi:10.1038/nature03098)

Marder, B. A. & Morgan, W. F. 1993 Delayed chromosomal instability induced by DNA damage. *Mol. Cell Biol.* **13**, 6667–6677.

Marx, J. 2002 Debate surges over the origins of genomic defects in cancer. *Science* **297**, 544–546. (doi:10.1126/science.297.5581.544)

Maser, R. S. & DePinho, R. A. 2002 Connecting chromosomes, crisis, and cancer. *Science* **297**, 565–569. (doi:10.1126/science.297.5581.565)

Maser, R. S. & DePinho, R. A. 2004 Telomeres and the DNA damage response: why the fox is guarding the henhouse. *DNA Repair (Amst)* **3**, 979–988. (doi:10.1016/j.dnarep.2004.05.009)

Maynard Smith, J. 1982 *Evolution and the theory of games.* Cambridge, UK: Cambridge University Press.

Michor, F., Iwasa, Y., Komarova, N. L. & Nowak, M. A. 2003 Local regulation of homeostasis favors chromosomal instability. *Curr. Biol.* **13**, 581–584. (doi:10.1016/S0960-9822(03)00172-6)

Michor, F., Iwasa, Y., Vogelstein, B., Lengauer, C. & Nowak, M. A. 2005 Can chromosomal instability initiate tumorigenesis? *Semin. Cancer Biol.* **15**, 43–49. (doi:10.1016/j.semcancer.2004.09.007)

Nagar, S., Smith, L. E. & Morgan, W. F. 2003 Mechanisms of cell death associated with death-inducing factors from genomically unstable cell lines. *Mutagenesis* **18**, 549–560. (doi:10.1093/mutage/geg033)

Nasmyth, K. 2002 Segregating sister genomes: the molecular biology of chromosome separation. *Science* **297**, 559–565. (doi:10.1126/science.1074757)

Nowak, M. A., Komarova, N. L., Sengupta, A., Jallepalli, P. V., Shih, I.-M., Vogelstein, B. & Lengauer, C. 2002 The role of chromosomal instability in tumor initiation. *Proc. Natl Acad. Sci. USA* **99**, 16 226–16 231. (doi:10.1073/pnas.202617399)

Nowak, M. A., Michor, F. & Iwasa, Y. 2006 Genetic instability and clonal expansion. *J. Theor. Biol.* **241**, 26–32. (doi:10.1016/j.jtbi.2005.11.012)

Perucho, M. 1996 Cancer of the microsatellite mutator phenotype. *Biol. Chem.* **377**, 675–684.

Pontryagin, L. S., Boltyanski, V. G., Gamkrelidze, R. V. & Mishchenko, E. F. 1962 *The mathematical theory of optimal control processes.* New York, NY: Interscience Publishers.

Rajagopalan, H., Jallepalli, P. V., Rago, C., Velculescu, V. E., Kinzler, K. W., Vogelstein, B. & Lengauer, C. 2004 Inactivation of hCDC4 can cause chromosomal instability. *Nature* **428**, 77–81. (doi:10.1038/nature02313)

Rudolph, K. L., Millard, M., Bosenberg, M. W. & DePinho, R. A. 2001 Telomere dysfunction and evolution of intestinal carcinoma in mice and humans. *Nat. Genet.* **28**, 155–159. (doi:10.1038/88871)

Samper, E., Flores, J. M. & Blasco, M. A. 2001 Restoration of telomerase activity rescues chromosomal instability and premature aging in Terc-/- mice with short telomeres. *EMBO Rep.* **2**, 800–807. (doi:10.1093/embo-reports/kve174)

Sen, S. 2000 Aneuploidy and cancer. *Curr. Opin. Oncol.* **12**, 82–88. (doi:10.1097/00001622-200001000-00014)

Shih, I. M., Zhou, W., Goodman, S. N., Lengauer, C., Kinzler, K. W. & Vogelstein, B. 2001 Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. *Cancer Res.* **61**, 818–822.

Smith, L. E., Nagar, S., Kim, G. J. & Morgan, W. F. 2003 Radiation-induced genomic instability: radiation quality and dose response. *Health Phys.* **85**, 23–29. (doi:10.1097/00004032-200307000-00006)

Sontag, E. D. 2004 Some new directions in control theory inspired by systems biology. *Syst. Biol. (Stevenage)* **1**, 9–18. (doi:10.1049/sb:20045006)

Swan, G. W. 1990 Role of optimal control theory in cancer chemotherapy. *Math. Biosci.* **101**, 237–284. (doi:10.1016/0025-5564(90)90021-P)

Tomlinson, I. & Bodmer, W. 1999 Selection, the mutation rate and cancer: ensuring that the tail does not wag the dog. *Nat. Med.* **5**, 11–12. (doi:10.1038/4687)

Von Stryk, O. 2000 User's guide for DIRCOL (version 2.1): a direct collocation method for the numerical solution of optimal control problems, Fachgebiet Simulation und Systemoptimierung (SIM), Technische Universität Darmstadt.

Wan, F. Y. M. 1995 *Introduction to the calculus of variations and its applications*, 2nd edn. Boca Raton, FL: Chapman and Hall/CRC.

Wan, F. Y. M., Sadovsky, A. V. & N. L. Komarova. In preparation. Genetic instability in cancer: an optimal control problem.

Wodarz, D. & Komarova, N. L. 2005 *Computational biology of cancer: lecture notes and mathematical modeling*. Singapore: World Scientific.

Yarden, R. I., Pardo-Reoyo, S., Sgagias, M., Cowan, K. H. & Brody, L. C. 2002 BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage. *Nat. Genet.* **30**, 285–289. (doi:10.1038/ng837)