Theoretical Population Biology 77 (2010) 42–48

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Evolution of resistance and progression to disease during clonal expansion of cancer

Richard Durrett *, Stephen Moseley

Department of Mathematics, Cornell University, Ithaca, NY 14853, United States
Field of Applied Mathematics, Cornell University, Ithaca, NY 14853, United States

ARTICLE INFO

Article history:
Received 15 June 2009
Available online 5 November 2009

Keywords:
Cancer models
Waiting times
Moran model
Branching process

ABSTRACT

Inspired by previous work of Iwasa et al. (2006) and Haeno et al. (2007), we consider an exponentially growing population of cancerous cells that will evolve resistance to treatment after one mutation or display a disease phenotype after two or more mutations. We prove results about the distribution of the first time when k mutations have accumulated in some cell, and about the growth of the number of type-k cells. We show that our results can be used to derive the previous results about a tumor grown to a fixed size.

1. Introduction

The mathematical investigation of cancer began in the 1950s, when Nordling (1953), Armitage and Doll (1954, 1957), and Fisher (1959) set out to explain the age-dependent incidence curves of human cancers. For a nice survey see Frank (2007). Armitage and Doll (1954) noticed that log–log plots of cancer incidence data are linear for a large number of cancer types; for example, colorectal cancer incidence has a slope of 5.18 in men and 4.97 in women. The authors used this observation to argue that cancer is a multi-stage process and results from the accumulation of multiple genetic alterations in a single cell. The math underlying this hypothesis was very simple. Suppose X_i are independent and have an exponential distribution with rates u_i (i.e., the density function is $u_i e^{-u_i t}$ and the mean is $1/u_i$). Noting that the sum $X_1 + \cdots + X_k$ has a density function that is asymptotically

$$u_1 \cdots u_k \frac{t^{k-1}}{(k-1)!} \quad \text{as } t \to 0, \tag{1}$$

the authors inferred that the slope of the age–incidence curve was the number of stages minus 1, making colon cancer a six-stage process.

Later on, Knudson (1971) performed a statistical analysis of retinoblastoma, a childhood eye cancer. His study showed that familial cases of retinoblastoma have an earlier age of onset than the sporadic cases that emerge in families without a history of the disease. Based on age–incidence curves in the two groups, he hypothesized that two mutagenic events or "hits" are necessary to cause cancer in the sporadic case, but in individuals with the inherited form of the disease, a single hit is sufficient since one mutation is already present at birth. This study led to the concept of a tumor suppressor gene, i.e., a gene which contributes to tumorigenesis if inactivated in both alleles. See Knudson (2001) for a survey.

Knudson's research led to an explosion of papers on the multi-stage theory of carcinogenesis too numerous to list here. Most studies, like the ones cited in the last two paragraphs, merely fit curves to data on age-specific incidence without considering a population genetic model for the cell population. Iwasa et al. (2004, 2005) were the first to study waiting times in this way. They used a Moran model for a population of a fixed size N in which type-i cells are those with $i \geq 0$ mutations, and type i mutates to type $i+1$ at rate u_{i+1}. Let τ_k be the first time at which there is a type-k cell. They considered a variety of scenarios based on the relative fitnesses of mutants. In the neutral case, i.e., if the mutation does not alter the fitness or growth rate of the cell, they showed:

Theorem 1. *In a population of N cells, τ_2 is approximately exponentially distributed with rate $Nu_1u_2^{1/2}$, provided $1/\sqrt{u_2} \ll N \ll 1/u_1$.*

They called this result "stochastic tunneling" because the 2's arise before the 1's reach fixation. Durrett et al. (2009), see also Schweinsberg (2008), generalized this result to cover τ_k.

In many cases, such as leukemia and polyps in colon cancer, the cell population does not have constant size. For these reasons, Iwasa et al. (2006) considered the time to develop one mutation in an exponentially growing population and Haeno et al. (2007) extended the analysis to waiting for two mutations. Their model is a multi-type branching process in which type-i cells are those with $i \geq 0$ mutations. Type-i cells give birth at rate a_i and die at rate b_i,

* Corresponding address: Department of Mathematics, Cornell University, 523 Malott Hall, Ithaca, NY 14853, United States.
E-mail address: rtd1@cornell.edu (R. Durrett).

0040-5809/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2009.10.008

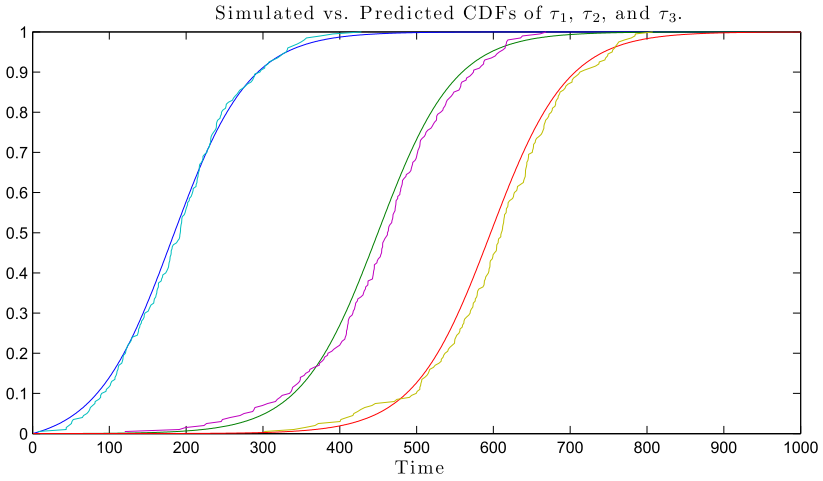

Fig. 1. Results of 200 runs of the system with $a_0 = 1.02$, $a_1 = 1.04$, $a_2 = 1.06 b_i = 1.0$, $u_i = 10^{-5}$. The smooth curves are the limit results for τ_i when $i = 1, 2, 3$.

where $\lambda_i = a_i - b_i > 0$. The previous papers consider a number of different possibilities, but here will restrict our attention to the case in which $i \to \lambda_i$ is increasing.

We suppose that during their lifetimes, type-i cells mutate at rate u_{i+1}, becoming type-$i+1$ cells. This is slightly different than the previous approach of having mutations with probability u_{i+1} at birth, which translates into a mutation rate of $a_i u_{i+1}$, and this must be kept in mind when comparing results. In applications, the mutation rates are small compared to birth and death rates, so the reduction of the birth rate of type-i cells to $a_i(1 - u_{i+1})$ is an insignificant difference.

1.1. Growth of type-0 cells

The number of type-0 cells, $Z_0(t)$, is a branching process, so if $Z_0(0) = 1$, $EZ_0(t) = e^{\lambda_0 t}$ and $e^{-\lambda_0 t} Z_0(t)$ is a nonnegative martingale. Well-known results imply that $e^{-\lambda_0 t} Z_0(t) \to W_0$ as $t \to \infty$. A closed-form formula for the generating function $Ex^{Z_0(t)}$ is known; see (15). To find the Laplace transform of W_0, we let $x = \exp(-\theta e^{-\lambda_0 t})$ in the closed-form solution and look at the limit as $t \to \infty$ to conclude that

$$Ee^{-\theta W_0} = \frac{b_0}{a_0} + \left(1 - \frac{b_0}{a_0}\right) \frac{1 - b_0/a_0}{1 - b_0/a_0 + \theta}.$$

From this we see that, if δ_0 is a pointmass at 0, and $\lambda_0 = a_0 - b_0$,

$$W_0 =_d \frac{b_0}{a_0}\delta_0 + \frac{\lambda_0}{a_0} \text{ exponential}(\lambda_0/a_0) \tag{2}$$

where the exponential(r) distribution has density re^{-rt} and mean $1/r$.

If we let $\Omega_0^0 = \{Z_0(t) = 0 \text{ for some } t \geq 0\}$ then (14) implies $P(\Omega_0) = b_0/a_0$, i.e., $W_0 = 0$ if and only if the process dies out. Letting $\Omega_\infty^0 = \{Z_0(t) > 0 \text{ for all } t \geq 0\}$ we have

$$(e^{-\lambda_0 t} Z_0(t) | \Omega_\infty^0) \to V_0 = \text{exponential}(\lambda_0/a_0) \tag{3}$$

and hence the Laplace transform

$$Ee^{-\theta V_0} = \frac{\lambda_0}{\lambda_0 + a_0\theta} = \left(1 + c_{\theta,0}\theta\right)^{-1} \tag{4}$$

where $c_{\theta,0} = a_0/\lambda_0$. Here, and in what follows, the c's are constants that only depend on the birth and death rates, and not on the mutational rates.

1.2. Type-1 results

Let τ_1 be the time of occurrence of the first type-1 cell. Since type-1 cells are produced at rate $u_1 Z_0(t)$,

$$P(\tau_1 > t | Z_0(s), s \leq t, \Omega_\infty^0) = \exp\left(-u_1 \int_0^t Z_0(s) ds\right). \tag{5}$$

τ_1 will occur when $\int_0^t Z_0(s)\, ds$ is of order $1/u_1$. A typical choice for $u_1 = 10^{-5}$, so $1/u_1$ is a large number, and we can use the approximation $(Z_0(s)|\Omega_\infty^0) \approx e^{\lambda_0 s} V_0$. Evaluating the integral, taking the expected value, and using (4), we conclude that

$$\begin{aligned} P(\tau_1 > t | \Omega_\infty^0) &\approx E \exp\left(-u_1 V_0 (e^{\lambda_0 t} - 1)/\lambda_0\right) \\ &= \frac{\lambda_0}{\lambda_0 + a_0 u_1 (e^{\lambda_0 t} - 1)/\lambda_0} \\ &= \left(1 + c_{\tau,1} u_1 (e^{\lambda_0 t} - 1)\right)^{-1} \end{aligned} \tag{6}$$

where $c_{\tau,1} = a_0/\lambda_0^2$. The median $t_{1/2}^1$ of the distribution has $\lambda_0^2 = a_0 u_1 (e^{\lambda_0 t_{1/2}^1} - 1)$, so

$$t_{1/2}^1 = \frac{1}{\lambda_0} \log\left(1 + \frac{\lambda_0^2}{a_0 u_1}\right). \tag{7}$$

Fig. 1 shows that (6) agrees well with the values of τ_1 observed in simulations. Parameters are given in the figure caption.

Our next step is to consider the growth of $Z_1(t)$. In Section 3 we show that

$$M_t = e^{-\lambda_1 t} Z_1(t) - \int_0^t u_1 e^{-\lambda_1 s} Z_0(s)\, ds \quad \text{is a martingale}$$

and use this to conclude

Theorem 2. *$e^{-\lambda_1 t} Z_1(t) \to W_1$ a.s. with*

$$EW_1 = u_1/(\lambda_1 - \lambda_0).$$

On Ω_∞^0 we will eventually get a type-1 mutant with an infinite line of descent so $\{W_1 > 0\} \supset \{\Omega_\infty^0\}$.

Let $T_M = \min\{t : Z_0(t) = M\}$. The results of simulations given in Figure 3 of Iwasa et al. (2006) show that when $\log P(W_1 > x | T_M < \infty)$ is plotted versus $\log x$, a straight line results. Since their M is large, this suggests that $(W_1|\Omega_\infty^0)$ has a power law tail. As we will now show, this is only approximately correct. To begin, we consider $Z_i^*(t)$, the number of type-i cells at time t in a system with $Z_0^*(t) = e^{\lambda_0 t} V_0$ for all $t \in (-\infty, \infty)$. Let

$$c_{h,1} = \frac{1}{\lambda_0} \left(\frac{a_1}{\lambda_1}\right)^{\lambda_0/\lambda_1 - 1} \Gamma(1 - \lambda_0/\lambda_1)\Gamma(\lambda_0/\lambda_1 + 1).$$