Published in final edited form as:
*Theor Popul Biol*. 2010 August ; 78(1): 54–66. doi:10.1016/j.tpb.2010.05.001.

# Evolutionary dynamics of tumor progression with random fitness values

**Rick Durrett**[1,*], **Jasmine Foo**[2,†], **Kevin Leder**[2,‡], **John Mayberry**[1,§,¶], and **Franziska Michor**[2,‖]

[1] Department of Mathematics, Cornell University, Ithaca, NY 14853

[2] Computational Biology Program, Memorial Sloan-Kettering, Cancer Center, New York, NY 10065

## Abstract

Most human tumors result from the accumulation of multiple genetic and epigenetic alterations in a single cell. Mutations that confer a fitness advantage to the cell are known as driver mutations and are causally related to tumorigenesis. Other mutations, however, do not change the phenotype of the cell or even decrease cellular fitness. While much experimental effort is being devoted to the identification of the functional effects of individual mutations, mathematical modeling of tumor progression generally considers constant fitness increments as mutations are accumulated. In this paper we study a mathematical model of tumor progression with random fitness increments. We analyze a multi-type branching process in which cells accumulate mutations whose fitness effects are chosen from a distribution. We determine the effect of the fitness distribution on the growth kinetics of the tumor. This work contributes to a quantitative understanding of the accumulation of mutations leading to cancer.

### Keywords

cancer evolution; branching process; fitness distribution; beneficial fitness effects; mutational landscape

## 1 Introduction

Tumors result from an evolutionary process occurring within a tissue (Nowell, 1976). From an evolutionary point of view, tumors can be considered as collections of cells that accumulate genetic and epigenetic alterations. The phenotypic changes that these alterations confer to cells are subjected to the selection pressures within the tissue and lead to adaptations such as the evolution of more aggressive cell types, the emergence of resistance, induction of angiogenesis, evasion of the immune system, and colonization of distant organs with metastatic growth. Advantageous heritable alterations can cause a rapid expansion of the cell clone harboring such

[§]Corresponding Author. jm858@cornell.edu, Tel.: +1 607 255 8262, Fax: +1 607 255 7149.
[*]Partially supported by NSF grant DMS 0704996 from the probability program.
[†]Partially supported by NIH grant R01CA138234.
[‡]Partially supported by NIH grant U54CA143798.
[¶]Partially supported by NSF RTG grant DMS 0739164.
[‖]Partially supported by NIH grants R01CA138234 and U54CA143798, a Leon Levy Foundation Young Investigator Award, and a Gerstner Young Investigator Award.

changes, since these cells are capable of outcompeting cells that have not evolved similar adaptations. The investigation of the dynamics of cell growth, the speed of accumulating mutations, and the distribution of different cell types at various timepoints during tumorigenesis is important for an understanding of the natural history of tumors. Further, such knowledge aids in the prognosis of newly diagnosed tumors, since the presence of cell clones with aggressive phenotypes lead to less optimistic predictions for tumor progression. Finally, a knowledge of the composition of tumors allows for the choice of optimum therapeutic interventions, as tumors harboring pre-existing resistant clones should be treated differently than drug-sensitive cell populations.

Mathematical models have led to many important insights into the dynamics of tumor progression and the evolution of resistance (Goldie and Coldman, 1983 and 1984; Bodmer and Tomlinson, 1995; Coldman and Murray, 2000; Knudson, 2001; Maley and Forrest, 2001; Michor et al., 2004; Iwasa et al., 2005; Komarova and Wodarz, 2005; Michor et al., 2006; Michor and Iwasa, 2006; Frank 2007; Wodarz and Komarova, 2007; Schweinsberg, 2008; Durrett, Schmidt, and Schweinsberg, 2009). These mathematical models generally fall into one of two classes: (i) constant population size models, and (ii) models describing exponentially growing populations. Many theoretical investigations of exponentially growing populations employ multi-type branching process models (e.g., Iwasa et al., 2006; Haeno et al., 2007; Durrett and Moseley, 2009), while others use population genetic models for homogeneously mixing exponentially growing populations (e.g., Beerenwinkel et al., 2007; Durrett and Mayberry, 2010). In this paper, we focus on branching process models. In these models, cells with $i \geq 0$ mutations are denoted as type-$i$ cells, and $Z_i(t)$ specifies the number of type-$i$ cells at time $t$. Type-$i$ cells die at rate $b_i$, give birth to one new type-$i$ cell at rate $a_i$, and give birth to one new type-$(i + 1)$ cell at rate $u_{i+1}$. Some authors (e.g., Haeno et al., 2007) consider an alternate version of our model in which mutations occur with probability $\mu_{i+1}$ during birth events which occur at rate $\alpha_i$, but the two versions are equivalent provided $u_{i+1} = \alpha_i\mu_{i+1}$ and $a_i = \alpha_i (1 - \mu_{i+1})$. This relationship between the parameters must be kept in mind when comparing results across papers.

One biologically unrealistic aspect of this model as presented in the literature is that all type-$i$ cells are assumed to have the same birth and death rates. This assumption describes situations during tumorigenesis in which the order of mutations is predetermined, i.e. the genetic changes can only be accumulated in a particular sequence and all other combinations of mutations lead to lethality. Furthermore, in this interpretation of the model, there cannot be any variability in phenotype among cells with the same number of mutations. In many situations arising in biology there is marked heterogeneity in phenotype even if genetically, the cells are identical (Elowitz et al., 2002; Becskei et al., 2005; Kaern et al., 2005; Feinerman et al., 2008). This variability may be driven by stochasticity in gene expression or in post-transcriptional or post-translational modifications. In this paper, we modify the branching process model so that mutations alter cell birth rates by a random amount.

An important consideration for this endeavor is the choice of the mutational fitness distribution. The exponential distribution has become the preferred candidate in theoretical studies of the genetics of adaptation. The first theoretical justification of this choice was given by Gillespie (1983, 1984), who argued that if the number of possible alleles is large and the current allele is close to the top of the rank ordering in fitness values, then extreme value theory should provide insight into the distribution of the fitness values of mutations. For many distributions including the normal, Gamma, and lognormal distributions, the maximum of $n$ independent draws, when properly scaled, converges to the Gumbel or double exponential distribution, $\Lambda(x) = \exp(-e^{-x})$. In the biological literature, it is generally noted that this class of distributions only excludes exotic distributions like the Cauchy distribution, which has no moments. However, in reality, it eliminates all distributions with $P(X > x) \sim Cx^{-\alpha}$. For distributions in

the domain of attraction of the Gumbel distribution, and if $Y_1 > Y_2 \cdots > Y_k$ are the $k$ largest observations in a sample of size $n$, then there is a sequence of constants $b_n$ so that the spacings $Z_i = i(Y_i - Y_{i+1})/b_n$ converge to independent exponentials with mean 1, see e.g., Weissman (1978). Following up on Gillespie's work, Orr (2003) added the observation that in this setting, the distribution of the fitness increases due to beneficial mutations has the same distribution as $Z_1$ independent of the rank $i$ of the wild type cell.

To infer the distribution of fitness effects of newly emerged beneficial mutations, several experimental studies were performed; for examples, see Imhoff and Schlotterer (2001), Sanjuan et al. (2004), and Kassen and Bataillon (2006). The data from these experiments is generally consistent with an exponential distribution of fitness effects. However, there is an experimental caveat that cannot be neglected (Rozen et al., 2002): if only those mutations are considered that reach 100% frequency in the population, then the exponential distribution is multiplied by the fixation probability. By this operation, a distribution with a mode at a positive value develops. In a study of a quasi-empirical model of RNA evolution in which fitness was based on secondary structures, Cowperthwaite et al. (2005) found that fitnesses of randomly selected genotypes appeared to follow a Gumbel-type distribution. They also discovered that the fitness distribution of beneficial mutations appeared exponential only when the vast majority of small-effect mutations were ignored. Furthermore, it was determined that the distribution of beneficial mutations depends on the fitness of the parental genotype (Cowperthwaite et al., 2005; MacLean and Buckling, 2009). However, since the exceptions to this conclusion arise when the fitness of the wild type cell is low, these findings do not contradict the picture based on extreme value theory.

In contrast to the evidence above, recent work of Rokyta et al. (2008) has shown that in two sets of beneficial mutations arising in the bacteriophage ID11 and in the phage $\varphi 6$ – for which the mutations were identified by sequencing – beneficial fitness effects are not exponential. Using a statistical method developed by Biesal et al. (2007), they tested the null hypothesis that the fitness distribution has an exponential tail. They found that the null hypothesis could be rejected in favor of a distribution with a right truncated tail. Their data also violated the common assumption that small-effect mutations greatly outnumber those of large effect, as they were consistent with a uniform distribution of beneficial effects. A possible explanation for the bounded fitness distribution may be found in the culture conditions utilized in the experiments: they evolved ID11 on *E. coli* at an elevated temperature (37° C instead of 33° C). There may be a limited number of mutations that will enable ID11 to survive in increased temperatures. The latter situation may be similar to scenarios arising during tumorigenesis, where, in order to develop resistance to a drug or to progress to a more aggressive stage, the conformation of a particular protein must be changed or a certain regulatory network must be disrupted. If there is a finite, but large, number of possible beneficial mutations, then it is convenient to use a continuous distribution as an approximation.

In this paper, we consider both bounded distributions and unbounded distributions for the fitness advance and derive asymptotic results for the number of type-$k$ individuals at time $t$. We determine the effects of the fitness distribution on the growth kinetics of the population, and investigate the rates of expansion for both bounded and unbounded fitness distributions. This model provides a framework to investigate the accumulation of mutations with random fitness effects.

The remainder of this section is dedicated to statements and discussion of our main results. Proofs of these results can be found in Sections 2–5.

### 1.1 Bounded distributions

The model we consider is a multi-type branching process in which type-$i$ cells have accumulated $i \geq 0$ advantageous mutations. All cells in the population die at rate $b_0$. The initial population consists entirely of type-0 cells that give birth at rate $a_0 > b_0$ to new type-0 cells and produce type-1 cells at rate $u_1$. We assume that the population of type-0 cells starts at a sufficiently large population $V_0$ so that we can approximate its size by $Z_0(t) = V_0 e^{\lambda_0 t}$, where $\lambda_0 = a_0 - b_0$. When a type-0 cell produces a type-1 cell, the new cell gives birth to type-1 cells at rate $a_0 + x$, where $x \geq 0$ is drawn according to a probability distribution $v$ and produces type-2 cells at rate $u_2$. In general, a type-$k$ cell with birth rate $a$ produces a new type-$(k + 1)$ cell at rate $u_k$ and the new type-$(k + 1)$ cell assumes an increased birth rate $a + x$ where $x \geq 0$ is drawn according to $v$. We let $Z_k(t)$ denote the total number of type-$k$ cells in the population at time $t$. When we refer to the $k$th generation of mutants, we mean the set of all type-$k$ cells.

We begin by considering situations in which the distribution of the increase in the birth rate is concentrated on $[0, b]$. In particular, suppose that $v$ has density $g$ with support in $[0, b]$ and assume that $g$ satisfies:

$$g \text{ is continuous at } b,\ g(b)>0,\ g(x) \leq G \text{ for } x \in [0, b] \tag{*}$$

Our first result describes the mean number of first generation mutants at time $t$, $EZ_1(t)$.

**Theorem 1—**If (*) holds, then

$$EZ_1(t) \sim \frac{V_0 u_1 g(b)}{bt} e^{(\lambda_0 + b)t}$$

where a(t) ~ b(t) means a(t)/b(t) → 1.

The next result shows that the actual growth rate of type-1 cells is slower than the mean. Here, and in what follows, we use ⇒ to indicate convergence in distribution.

**Theorem 2—**If (*) holds and $p = b/\lambda_0$, then for $\theta \geq 0$,

$$E\exp(-\theta t^{1+p} e^{-(\lambda_0 + b)t} Z_1(t)) \to \exp(-V_0 u_1 \theta^{\lambda_0/(\lambda_0 + b)} c_1(\lambda_0, b)), \tag{1.1}$$

where $c_1(\lambda_0, b)$ is a constant whose value will be given in (3.8). In particular, we have

$$t^{1+p} e^{-(\lambda_0 + b)t} Z_1 t) \Rightarrow V_1,$$

where $V_1$ has Laplace transform given by the righthand side of (1.1).

Theorem 2 is similar to Theorem 3 in Durrett and Moseley (2009) which assumes a deterministic fitness distribution so that all type-1 cells have growth rate $\lambda_1 = \lambda_0 + b$. There, the asymptotic growth rate of the first generation is $\exp(\lambda_1 t)$. In contrast, the continuous fitness distribution we consider here has the effect of slowing down the growth rate of the first generation by the polynomial factor $t^{1+p}$. To explain this difference, we note that the calculation of the mean given in Section 3 shows that the dominant contribution to $Z_1(t)$ comes from growth

rates $x = b - O(1/t)$. However, mutations with this growth rate are unlikely until the number of type-0 cells is $O(t)$, i.e., roughly at time $r_1 = (1/\lambda_0) \log t$. Thus at time $t$, the number of type-1 cells will be roughly $\exp((\lambda_0+b)(t - r_1)) = \exp((\lambda_0 +b)t)/t^{1+p}$.

To prove Theorem 2, we look at mutations as a point process in $[0, t] \times [0, b]$: there is a point at $(s, x)$ if there was a mutant with birth rate $a_0 + x$ at time $s$. This allows us to derive the following explicit expression for the Laplace transform of $Z_1(t)$:

$$E(e^{-\theta Z_1(t)}) = \exp\left(-u_1 \int_0^b dx\, g(x) \int_0^t ds\, V_0 e^{\lambda_0 s}(1 - \tilde{\varphi}_{x,t-s}(\theta))\right)$$

where $\tilde{\varphi}_{x,r}(\theta) = Ee^{-\theta \tilde{Z}_r^x}$ and $\tilde{Z}_r^x$ is a continuous-time branching process with birth rate $a_0+x$, death rate $b_0$, and initial population $\tilde{Z}_0^x = 1$. In Figure 1, we compare the exact Laplace transform of $t^{1+p} \exp(-(\lambda_0+b)t)Z_1(t)$ with the results of simulations and the limiting Laplace transform from Theorem 2, illustrating the convergence as $t \to \infty$.

Notice that the Laplace transform of $V_1$ has the form $\exp(C\,\theta^\alpha)$ where $\alpha = \lambda_0/(\lambda_0 + b)$ which implies that $P(V_1 > v) \sim v^{-\alpha}$ as $v \to \infty$ (see, for example, the argument in Section 3 of Durrett and Moseley (2009)). To gain some insight into how this limit comes about, we give a second proof of the convergence that tells us the limit is the sum of points in a nonhomogeneous Poisson process. Each point in the limiting process represents the contribution of a different mutant lineage to $Z_1(t)$. More precisely, we define a three dimensional point process $\mathcal{M}(t)$ on $[0, t] \times [0, b] \times (0, \infty)$ by the following rule: there is a point at $(s, x, v)$ if there was a type-1 mutant with birth rate $a_0 + x$ at time $s$ and the number of its type-1 descendants at time $t$, $Z_1^{s,x}(t)$, has $e^{-(\lambda_0+x)(t-s)}Z_1^{s,x}(t) \to v$ as $t \to \infty$ with $v > 0$. We define $F : [0, \infty)^3 \to [0, \infty)$ by

$$F(s, x, v) = vt^{1+p}e^{-(\lambda_0+b)t}e^{(\lambda_0+x)(t-s)}$$

i.e. $F$ maps a point in $\mathcal{M}(t)$ onto its contribution to $V_1 = \lim_{t\to\infty} t^{1+p}e^{-(\lambda_0+b)t}Z_1(t)$.

**Theorem 3—**As $t \to \infty$, $F(\mathcal{M}(t)) \Rightarrow \Lambda$ where $\Lambda$ is a Poisson process on $(0, \infty)$ with mean measure $\mu(z, \infty) = A_1(\lambda_0, b)u_1V_0z^{-\lambda_0/(\lambda_0+b)}$ and $A_1(\lambda_0, b)$ is a constant whose value is given in (3.9). In particular, $V_1 = \lim_{t\to\infty} t^{1+p}e^{-(\lambda_0+b)t}Z_1(t)$ is the sum of the points in $\Lambda$.

A similar result can be obtained for deterministic fitness distributions, see the Corollary to Theorem 3 in Durrett and Moseley (2009). However, the new result shows that the point process limit is not an artifact of assuming that all first generation mutants have the same growth rate. Even when the fitness advances are random, different mutant lines contribute to the limit. This result is consistent with observations of Maley et al. (2006) and Shah et al. (2009) that tumors contain cells with different mutational haplotypes. Theorem 3 also gives quantitative predictions about the relative contribution of different mutations to the total population. These implications will be explored further in a follow-up paper currently in progress.

With the behavior of the type-1 individuals analyzed, we are ready to proceed to the study of type-$k$ individuals. The computation of the mean is straightforward.

**Theorem 4—**If (*) holds, then

$$EZ_k(t) \sim \frac{V_0 \cdot u_1 \cdots u_k \cdot g(b)^k}{t^k b^k k!} e^{(\lambda_0+kb)t}$$

As in the $k = 1$ case, the mean involves a polynomial correction to the exponential growth and again, does not give the correct growth rate for the number of type-$k$ cells. To state the correct limit theorem describing the growth rate of $Z_k(t)$, we will define $p_k$ and $u_{1,k}$ by

$$k+p_k=\sum_{j=0}^{k-1}\frac{\lambda_0+kb}{\lambda_0+jb} \text{ and } u_{1,k}=\prod_{j=1}^{k}u_j^{\lambda_0/(\lambda_0+(j-1)b)}$$

for all $k \geq 1$.

**Theorem 5—**If (*) holds, then for $\theta \geq 0$

$$E\exp(-\theta t^{k+p_k}e^{-(\lambda_0+kb)t}Z_k(t)) \to \exp(-c_k(\lambda_0,b)V_0u_{1,k}\theta^{\lambda_0/(\lambda_0+kb)}) \tag{1.2}$$

where $c_k(\lambda_0, b)$ is a constant whose value will be given in (4.9). In particular, we have

$$t^{k+p_k}e^{-(\lambda_0+kb)t}Z_k(t) \Rightarrow V_k,$$

where $V_k$ has Laplace transform given by the righthand side of (1.2).

If we let $Z_k^{s,x,v}(t)$ be the number of type-$k$ descendants at time $t$ of the 1 mutant at $(s, x, v) \in \mathcal{M}(t)$ where $\mathcal{M}(t)$ is the three dimensional point process described in the paragraph preceding Theorem 3, then $Z_k^{s,x,v}$ is the same as a process in which the initial type (here type-1 cells) behaves like $ve^{(\lambda_0+x)(t-s)}$ instead of $Z_0(t) = V_0e^{\lambda_0 t}$. Therefore, Theorem 5 can be proved by induction. To explain the form of the result we consider the case $k = 2$. Breaking things down according to the times and the sizes of the mutational changes, we have

$$EZ_2(t)=\int_0^b dx_1\, g(x_1)\int_0^b dx_2\, g(x_2)\int_0^t ds_1\int_{s_1}^t ds_2\, V_0e^{\lambda_0 s_1}u_1e^{(\lambda_0+x_1)(s_2-s_1)}u_2e^{(\lambda_0+x_1+x_2)(t-s_2)}$$

As in the result for $Z_1(t)$ the dominant contribution comes from $x_1$, $x_2 = b - O(1/t)$ and as in the discussion preceding the statement of Theorem 2, the time of the first mutation to $b - O(1/t)$ is $\approx r_1 = (\log t)/\lambda_0$. The descendants of this mutation grow at exponential rate $\lambda_0 + b - O(1/t)$, so the time of the first mutation to $2b - O(1/t)$ is $\approx r_2 = r_1 + (\log t)/(\lambda_0 + b)$. Noticing that

$$\exp((\lambda_0+2b)(t-r_1-r_2))=\exp((\lambda_0+2b)t)t^{-(\lambda_0+2b)/\lambda_0-(\lambda_0+2b)/(\lambda_0+b)}$$

tells us what to guess for the polynomial term: $t^{-(2+p_2)}$ where

$$2+p_2=\frac{\lambda_0+2b}{\lambda_0}+\frac{\lambda_0+2b}{\lambda_0+b}$$

In Figure 2, we compare the asymptotic Laplace transform from Theorem 5 with the results of simulations in the case $k = 2$. To explain the slow convergence to the limit, we note that if we take account of the mutation rates $u_1$, $u_2$ in the heuristic from the previous paragraph (which becomes important when $u_1$, $u_2$ are small), then the first time we see a type-1 cell with growth rate $b - O(1/t)$ will not occur until time $\lambda_0^{-1}\log(t/u_1)$ when the type-0 cells reach $O(t/u_1)$ and so the first type-2 cell with growth rate $2b - O(1/t)$ will not be born until time $r=\lambda_0^{-1}\log(t/u_1)+(\lambda_0+b)^{-1}\log(t/u_2)$ when the descendants of the type-1 cells with growth rate $b - O(1/t)$ reach size $O(t/u_2)$. When $u_1 = u_2 = 10^{-3}$, $\lambda_0 = .1$, and $b = .01$, $r \approx 223$. The mutations created at this point will need some time to grow and become dominant in the population. It would be interesting to compare simulations at time 300, but we have not been able to do this due to the large number of different growth rates in generation 1.

### 1.2 Unbounded distributions

In this section, we consider situations in which the fitness distribution is unbounded. We will suppose that the fitness distribution $v$ has tail

$$v(x,\infty) \sim K\, x^{\beta}\exp(-\gamma x^{\alpha}) \tag{1.3}$$

as $x \to \infty$ for some $\alpha$, $\gamma$, $K > 0$, and $\beta \in \mathbb{R}$. Our assumption (1.3) on the tail of $v$ is satisfied by a number of natural distributions including the gamma($\beta$ +1, $\gamma$) distribution which has $\alpha = 1$ (and includes the exponential distribution as the special case $\beta = 0$) and the normal distribution which has $\alpha = 2$, $\beta = -1$.

To analyze this situation, we will again take a Poisson process viewpoint and look at the contribution from a mutation at time $s$ with increased growth rate $x$. A mutation that increases the growth rate by $x$ at time $s$ will, if it does not die out, grow to $e^{(\lambda_0 + x)(t-s)}\,\zeta$ at time $t$ where $\zeta$ has an exponential distribution. The growth rate $(\lambda_0 + x)(t - s) \geq z$ when

$$x \geq \frac{z}{t-s} - \lambda_0.$$

Therefore,

$$\begin{aligned}
\mu(z,\infty) &\equiv E(\#\,\text{mutations with}\,(\lambda_0+x)(t-s)\geq z)\\
&=V_0u_1\int_0^t e^{\lambda_0 s}v(z/(t-s)-\lambda_0,\infty)\,ds\\
&=KV_0u_1\int_0^t e^{\lambda_0 s}q(z/(t-s)-\lambda_0)\left(\tfrac{z}{t-s}-\lambda_0\right)^{\beta}\exp\left(-\gamma\left(\tfrac{z}{t-s}-\lambda_0\right)^{\alpha}\right)ds\\
&=KV_0u_1\int_0^t q(z/(t-s)-\lambda_0)\left(\tfrac{z}{t-s}-\lambda_0\right)^{\beta}\exp(\varphi(s,z))\,ds
\end{aligned}$$

where

$$q(x)=\frac{\nu(x,\infty)}{Kx^{\beta}\exp(-\gamma x^{\alpha})}\to 1 \tag{1.4}$$

as $x \to \infty$ and

$$\varphi(s,z)=\lambda_0 s-\gamma\left(\frac{z}{t-s}-\lambda_0\right)^{\alpha}. \tag{1.5}$$

The size of this integral can be found by maximizing the exponent $\varphi$ over $s$ for fixed $z$. Since

$$\frac{\partial\varphi}{\partial s}(s,z)=\lambda_0-\alpha\gamma\left(\frac{z}{t-s}-\lambda_0\right)^{\alpha-1}\frac{z}{(t-s)^2} \tag{1.6}$$

and

$$\frac{\partial^2\varphi}{\partial s^2}(s,z)=-\alpha(\alpha-1)\gamma\left(\frac{z}{t-s}-\lambda_0\right)^{\alpha-2}\frac{z^2}{(t-s)^4}-\alpha\gamma\left(\frac{z}{t-s}-\lambda_0\right)^{\alpha-1}\frac{2z}{(t-s)^3} \tag{1.7}$$

we can see that $\partial^2\varphi/\partial s^2(s,z) < 0$ when $\alpha z > \lambda_0(t-s)$ so that for all $z$ in this range, $\varphi(s,z)$ is concave as a function of $s$ and achieves its maximum at a unique value $s_z$.

When $\alpha = 1$, it is easy to set (1.6) to 0 and solve for $s_z$. This in turn leads to an asymptotic formula for $\mu(z,\infty)$ and allows us to derive the following limit theorem for $Z_1(t)$.

**Theorem 6—**Suppose $\alpha = 1$ and let $c_0 = \lambda_0/4\gamma$. Then $t^{-2}\log Z_1(t) \to c_0$ and

$$\frac{1}{t}\left[\log Z_1(t)-c_0t^2\left(1+\frac{(2\beta+1)\log t}{\lambda_0 t}\right)\right]\Rightarrow y^*$$

where y* is the rightmost point in the point process with intensity given by

$$(2c_0)^{\beta}(\pi/\lambda_0)^{1/2}KV_0u_1\exp(\gamma\lambda_0-\lambda_0 y/2c_0). \tag{1.8}$$

When $\alpha \neq 1$, solving for $s_z$ becomes more difficult, but we are still able to prove the following limit theorem for $Z_1(t)$.

**Theorem 7—**For any integer $\alpha > 1$, there exist explicitly calculable constants $c_k = c_k(\alpha,\gamma)$, $0 \le k < \alpha$, and $\kappa = \kappa(\beta,\alpha,\gamma)$ so that $t^{-(\alpha+1)/\alpha}\log Z_1(t) \to c_0$ and

$$\frac{1}{t^{1/\alpha}}\left[\log Z_1(t) - c_0 t^{(\alpha+1)/\alpha}\left(1+\sum_{1\le k<\alpha} c_k t^{-k/\alpha}+\kappa\frac{\log t}{t}\right)\right] \Rightarrow y^*$$

where y* is the rightmost particle in a point process with explicitly calculable intensity.

The complicated form of the result is due to the fact that the fluctuations are only of order $t^{1/\alpha}$ so we have to be very precise in locating the maximum. The explicit formulas for the constants and the intensity of the point process are given in (5.11) and (5.12). With more work this result could be proved for general $\alpha > 1$, but we have not tried to do this or prove Conjecture 1 below because the super-exponential growth rates in the unbounded case are too fast to be realistic.

We conclude this section with two comments. First, the proof of Theorem 7 shows that in contrast to the bounded case, in the unbounded case, most type-1 individuals are descendants of a single mutant. Second, the proof shows that the distribution of the mutant with the largest growth rate is born at time $s \sim t/(\alpha + 1)$ (see Remark 1 at the end of Section 5) and has growth rate $z = O(t^{(\alpha+1)/\alpha})$. The intuition behind this is that since the type-0 cells have growth rate $e^{\lambda 0 s}$ and the distribution of the increase in fitness has tail $\approx e^{-\gamma x^\alpha}$, the largest advance $x$ attained by time $t$ should occur when $s = O(t)$ and satisfy

$$e^{C\lambda_0 t} e^{-\gamma x^\alpha} = O(1) \quad \text{or} \quad x = O(t^{1/\alpha}).$$

The growth rate of its family is then $(\lambda_0 + x)(t - s) = O(t^{(\alpha+1)/\alpha})$.

Since the type-1 cells grow at exponential rate $c_1 t^{(\alpha+1)/\alpha}$, if we apply this same reasoning to type-2 mutants, then the largest additional fitness advance $x$ attained by type-2 individuals should satisfy

$$e^{c_1 t(\alpha+1)/\alpha} e^{-\gamma x^\alpha} = O(1) \quad \text{or} \quad x = O(t^{1/\alpha+1/\alpha^2}).$$

and the growth rate of its family will be $O(t^{1+1/\alpha+1/\alpha^2})$. Extrapolating from the first two generations, we make the following

**Conjecture 1—**Let $q(k) = \sum_{j=0}^{k} \alpha^{-j}$. As t → ∞,

$$\frac{1}{t^{q(k)}}\log Z_k(t) \to c_k$$

Note that in the case of the exponential distribution, $q(k) = k + 1$.

The rest of the paper is organized as follows. Sections 2–5 are devoted to proofs of our main results. After some preliminary notation and definitions in Section 2, Theorems 1–3 are proved in Section 3, Theorems 4–5 in Section 4, and Theorems 6–7 in Section 5. We conclude with a discussion of our results in Section 6.

## 2 Preliminaries

This section contains some preliminary notation and definitions which we will need for the proofs of our main results. We denote by $\mathcal{N}(t)$ the points in a two-dimensional Poisson process on $[0, t] \times [0, \infty)$ with mean measure

$$V_0 e^{\lambda_0 s} ds \nu(dx),$$

where in Sections 3–4, $\nu(dx) = g(x)dx$ with $g$ satisfying (*) and in Section 5, $\nu$ has tail satisfying (1.3). In other words, we have a point at $(s, x)$ if there was a mutant with birth rate $a_0 + x$ at time $s$. Define a collection of independent birth/death branching processes $Z_1^{s,x}(t)$ indexed by $(s, x) \in \mathcal{N}(t)$ with $Z_1^{s,x}(s)=1$, individual birth rate $a_0 + x$, and death rate $b$. $Z_1^{s,x}(t)$ is the contribution of the mutation at $(s, x)$ and

$$Z_1(t)= \sum_{(s,x)\in\mathcal{N}(t)} Z_1^{s,x}(t).$$

It is well known that

$$e^{-(\lambda_0+x)(t-s)} Z_1^{s,x}(t) \to \frac{b}{a_0+x}\delta_0 + \frac{\lambda_0+x}{a_0+x}\zeta,$$

where $\zeta \sim \exp((\lambda_0 + x)/(a_0 + x))$ (see, for example, equation (1) in Durrett and Moseley (2009)). In several results, we shall make use of the three-dimensional Poisson process $\mathcal{M}(t)$ on $[0, t] \times [0, \infty) \times (0, \infty)$ with intensity

$$V_0 e^{\lambda_0 s} \nu(dx) \left(\frac{\lambda_0+x}{a_0+x}\right)^2 e^{-v(\lambda_0+x)/(a_0+x)} dv.$$

In words, $(s, x, v) \in \mathcal{M}(t)$ if there was a mutant with birth rate $a_0 + x$ at time $s$ and the number of its descendants at time $t$, $Z_1^{s,x}(t)$, has $Z_1^{s,x}(t) \sim v e^{(\lambda_0+x)(t-s)}$. It is also convenient to define the mapping $z$: $[0, \infty) \times [0, t] \to [0, \infty)$ which maps a point $(s, x) \in \mathcal{N}(t)$ to the growth rate of the induced branching process if it survives: $z(s, x) = (\lambda_0 + x)(t - s)$ and let

$$\mu(A)=E|\{(s, x) \in \mathcal{N}(t): z(s, x) \in A\}|$$

for $A \subset [0, \infty)$.

We shall use $C$ to denote a generic constant whose value may change from line to line. We write $f(t) \sim g(t)$ if $f(t)/g(t) \to 1$ as $t \to \infty$ and $f(t) = o(g(t))$ is $f(t)/g(t) \to 0$. $f(t) \gg (\ll) g(t)$ means that $f(t)/g(t) \to \infty$ (resp. 0) as $t \to \infty$ and $f(t) = O(g(t))$ means $|f(t)| \le Cg(t)$ for all $t > 0$. We also shall use the notation $f(t) \simeq g(t)$ if $f(t) - g(t) \to 0$ as $t \to \infty$.

## 3 Bounded distributions, $Z_1$

In this section, we prove Theorems 1–3.

### Proof of Theorem 1

Mutations to type-1 cells occur at rate $V_0u_1e^{\lambda 0s}$ so

$$\begin{aligned} EZ_1(t) &= u_1\int_0^t\int_0^b e^{(t-s)(\lambda_0+x)}g(x)\,dx V_0e^{\lambda_0 s}\,ds \\ &= u_1V_0e^{\lambda_0 t}\int_0^b dx\,g(x)\int_0^t e^{(t-s)x}\,ds \\ &= u_1V_0e^{\lambda_0 t}\int_0^b dx\,g(x)\tfrac{e^{tx}-1}{x}. \end{aligned} \tag{3.1}$$

We begin by showing that the contribution from $x \in [0, b - (1 + k)(\log t)/t]$ can be ignored for any $k \in [0, \infty)$. The Mean Value theorem implies that

$$\frac{e^{tx}-1}{x} \le te^{tx}. \tag{3.2}$$

Using this and the fact that $\int_c^d te^{tx}dx \le e^{td}$ for any $c < d$, we can see that

$$t^k e^{-bt}\int_0^{b-(1+k)(\log t)/t} dx\,g(x)\frac{e^{tx}-1}{x}\cdot \le Gt^k e^{-(1+k)\log t} \to 0 \tag{3.3}$$

To handle the other piece of the integral, we take $k = 1$ and note that

$$\int_{b-(2\log t)/t}^b dx\,g(x)\frac{e^{tx}-1}{x} \sim \frac{g(b)}{b}e^{bt}\int_{b-2\log t/t}^b e^{t(x-b)}dx.$$

After changing variables $y = (b - x)t$, $dx = -dy/t$, the last integral

$$= \frac{1}{t}\int_0^{2\log t} e^{-y}\,dy \sim 1/t,$$

which proves the result.

The above proof tells us that the dominant contribution to the type-1 cells comes from mutations with fitness increase $x \ge b_t = b - 2\log t/t$. To describe the times at which the dominant contributions occur, let $S(t) = (2/b)\log\log t$. Then the contribution to the mean from $x \in [b_t, b]$ and $s \ge S(t)$ is by (3.1)

$$\begin{aligned} &\le Gu_1V_0e^{(\lambda_0+b)t}\tfrac{2(\log t)}{t}\int_{S(t)}^\infty e^{-sb_t}\,ds \\ &\le Gu_1V_0e^{(\lambda_0+b)t}\tfrac{2(\log t)}{tb_t}e^{-b_tS(t)}. \end{aligned}$$

Since $b_t S(t) \geq (3/2) \log\log t$ for all $t$ sufficiently large, this quantity is $o(t^{-1}e^{(\lambda_0+b)t})$. In words, the dominant contribution to the mean comes from points close to $(0, b)$ or more precisely from $[0, (2/b) \log\log t] \times [b - (2 \log t)/t, b]$.

## Proof of Theorem 2

It suffices to prove (1.1). The computation in (3.3) with $k = 2 + p$ implies that the contribution from mutations with $x \leq b_t = b - (3 + p)(\log t)/t$ can be ignored. Therefore, we have

$$E\exp(-\theta Z_1(t)e^{-t(\lambda_0+b)}t^{1+p}) \simeq E(\exp(-\theta Z_1(t)e^{-t(\lambda_0+b)}t^{1+p});A_t)$$

where $A_t = \{(s, x) \in \mathcal{N}(t): x > b_t\}$. Lemma 2 of Durrett and Moseley (2009) implies that

$$E(e^{-\theta Z_1(t)};A_t)=\exp\left(-u_1\int_{b_t}^{b}dx\, g(x)\int_0^t ds\, V_0 e^{\lambda_0 s}(1-\tilde{\varphi}_{x,t-s}(\theta))\right)$$

where $\tilde{\varphi}_{x,r}(\theta)=Ee^{-\theta\tilde{Z}_r^x}$ and $\tilde{Z}_r^x$ is a birth/death branching process with birth rate $a_0 + x$, death rate $b_0$, and initial population $\tilde{Z}_0^x=1$. Using

$$e^{-(\lambda_0+b)t}=e^{-(\lambda_0+x)(t-s)}e^{-(\lambda_0+x)s}e^{-(b-x)t} \tag{3.4}$$

we have

$$E(\exp(-\theta Z_1(t)e^{-t(\lambda_0+b)}t^{1+p});A_t)=\exp\left(-u_1V_0\int_{b_t}^{b}dx\, g(x)\int_0^t ds\, e^{\lambda_0 s}\{1-\tilde{\varphi}_{x,t-s}(\theta e^{-(\lambda_0+x)(t-s)}e^{-(\lambda_0+x)s}e^{-(b-x)t}t^{1+p})\}\right)$$

Changing variables $s = r_x + r$ where $r_x=\frac{1}{\lambda_0+x}\log(t^{1+p})$ on the inside integral, $y = (b - x)t$, $dy/t = -dx$ on the outside, and continuing to write $x$ as short hand for $b - y/t$, the above

$$=\exp\left(-u_1V_0\int_0^{(3+p)\log t}\frac{dy}{t}g(x)t^{(1+p)\lambda_0/(\lambda_0+x)}\int_{-r_x}^{t-r_x}dr\, e^{\lambda_0 r}\{1-\tilde{\varphi}_{x,t-r-r_x}(\theta e^{-(\lambda_0+x)(t-r-r_x)}e^{-(\lambda_0+x)r}e^{-y})\}\right) \tag{3.5}$$

Formula (20) in Durrett and Moseley (2009) implies that as $u \to \infty$,

$$1-\tilde{\varphi}_{x,u}(\theta e^{-(\lambda_0+x)u}) \to \frac{\lambda_0+x}{a_0+x}\cdot\frac{\theta}{\theta+\frac{\lambda_0+x}{a_0+x}} \tag{3.6}$$

and therefore, letting $t \to \infty$ and using $(1 + p)\, \lambda_0/(\lambda_0 + b) = 1$, we can see that the expression in (3.5)

$$\to \exp\left(-u_1 V_0 g(b)\int_0^\infty dy \frac{\lambda_0+b}{a_0+b}\int_{-\infty}^\infty dr\, e^{\lambda_0 r}\frac{\theta e^{-(\lambda_0+b)r}e^{-y}}{\theta e^{-(\lambda_0+b)r}e^{-y}+\frac{\lambda_0+b}{a_0+b}}\right)$$

Changing variables $r=\frac{1}{\lambda_0+b}\{q+\log[\theta e^{-y}(a_0+b)/(\lambda_0+b)]\}$, $dr = dq/(\lambda_0 + b)$ gives

$$=\exp\left(-u_1 V_0 g(b)\theta^{\lambda_0/(\lambda_0+b)}\left(\frac{\lambda_0+b}{a_0+b}\right)^{b/(\lambda_0+b)}\int_0^\infty dy\, e^{-y\lambda_0/(\lambda_0+b)}\int_{-\infty}^\infty \frac{dq}{\lambda_0+b}e^{q\lambda_0/(\lambda_0+b)}\frac{e^{-q}}{e^{-q}+1}\right)$$

To simplify the first integral we note that

$$\int_0^\infty dy\, e^{-y\lambda_0/(\lambda_0+b)}=\frac{\lambda_0+b}{\lambda_0}$$

For the second integral, we prove

**Lemma 1—**If $0 < c < 1$

$$\int_{-\infty}^\infty dq\, e^{qc}\frac{e^{-q}}{e^{-q}+1}=\Gamma(c)\Gamma(1-c) \tag{3.7}$$

**Proof:** We can rewrite the integral as

$$\int_{-\infty}^\infty dq\, e^{qc}\int_0^\infty dx\, e^{-x}e^{-q}\exp(-e^{-q}x)$$

so that after interchanging the order of integration and changing variables $w = e^{-q}x$, $dw = -dqe^{-q}x$ so that $w/x = e^{-q}$, $dw/x = -dqe^{-q}$, we have

$$=\int_0^\infty dx\int_0^\infty \frac{dw}{x}(w/x)^{-c}e^{-x}e^{-w}=\int_0^\infty dx\, x^{-1+c}e^{-x}\int_0^\infty dw\, w^{-c}e^{-w}$$

which is $= \Gamma(c)\Gamma(1 - c)$.

Taking $c = \lambda_0/(\lambda_0 + b)$ and letting

$$c_1(\lambda_0,b)=g(b)\frac{\lambda_0+b}{\lambda_0}\cdot\frac{1}{\lambda_0+b}\left(\frac{a_0+b}{\lambda_0+b}\right)^{-b/(\lambda_0+b)}\Gamma(\lambda_0/(\lambda_0+b))\Gamma(1-\lambda_0/(\lambda_0+b)) \tag{3.8}$$

we have proved Theorem 2.

Recall that we have assumed $Z_0(t) = V_0 e^{\lambda_0 t}$ is deterministic. This assumption can be relaxed to obtain the following generalization of Theorem 2 which is used in Section 4.

**Lemma 2—**Suppose that $Z_0(t)$ is a stochastic process with $Z_0(t) \sim e^{\lambda_0 t} V_0$ for some constant $V_0$ as $t \to \infty$. Then the conclusions of Theorem 2 remain valid.

To see why this is true, we can use a variant of Lemma 2 from Durrett and Moseley (2009) to conclude that

$$E\left(e^{-\theta Z_1(t)} | \mathcal{F}_t^0\right) = \exp\left(-u_1 \int_0^b dx\, g(x) \int_0^t ds Z_0(s) \left(1 - \tilde{\varphi}_{x,t-s}(\theta)\right)\right),$$

where $\mathcal{F}_t^0$ is the $\sigma$-field generated by $Z_0(s)$ for $s \le t$. Therefore,

$$E\left(e^{-\theta Z_1(t)}\right) = E\exp\left(-u_1 \int_0^b dx\, g(x) \int_0^t ds Z_0(s) \left(1 - \tilde{\varphi}_{x,t-s}(\theta)\right)\right),$$

Given $\varepsilon > 0$, we can choose $t_\varepsilon > 0$ so that

$$\left| \frac{Z_0(t)}{V_0 \exp(\lambda_0 t)} - 1 \right| < \varepsilon$$

for all $t > t_\varepsilon$. Since the contribution from $t \le t_\varepsilon$ will not affect the limit and the term inside the expectation is bounded, the rest of the proof can be completed in the same manner as the proof of Theorem 2.

We conclude this section with the

## Proof of Theorem 3

Let $\mathcal{M}(t)$ be the three dimensional Poisson process defined in Section 2. Recall that

$$F(s, x, v) = v t^{1+p} e^{-(\lambda_0 + b)t} e^{(\lambda_0 + x)(t-s)}$$

i.e. $F$ maps a point in $\mathcal{M}(t)$ to its contribution to the limit $t^{1+p} e^{-(\lambda_0+b)t} Z_1(t)$. Using (3.4), we see that in order to have $F(s, x, v) > z$ we need

$$v > z t^{-(1+p)} e^{(b-x)t} e^{(\lambda_0 + x)s}$$

Therefore, the expected number of mutations that contribute more than $z$ to the limit is

$$u_1 V_0 \int_0^b dx\, g(x) \int_0^t ds\, e^{\lambda_0 s} \frac{\lambda_0 + x}{a_0 + x} \exp\left(-\frac{\lambda_0 + x}{a_0 + x} \cdot z t^{-(1+p)} e^{(b-x)t} e^{(\lambda_0 + x)s}\right)$$

The exponential term can be simplified by making the change of variables

$$s=\frac{1}{\lambda_0+x}\log\left(\frac{r}{zt^{-(1+p)}e^{(b-x)t\frac{\lambda_0+x}{a_0+x}}}\right),$$

$ds = dr/r(\lambda_0 + x)$ yielding the equivalent expression

$$u_1V_0\int_0^b dx\, g(x)z^{-\lambda_0/(\lambda_0+x)}\left(\frac{\lambda_0+x}{a_0+x}\right)^{x/(\lambda_0+x)}\cdot t^{(1+p)\lambda_0/(\lambda_0+x)}e^{-(b-x)t\lambda_0/(\lambda_0+x)}\int_{\alpha(x,t)}^{\beta(x,t)}\frac{dr}{\lambda_0+x}r^{-x/(\lambda_0+x)}e^{-r}$$

where $\alpha(x, t) = zt^{-(1+p)}e^{(b-x)t}(\lambda_0 + x)/(a_0 + x)$ and $\beta(x, t) = \alpha(x, t)\, e^{(\lambda_0+x)t}$. As in the previous proof, the main contribution comes from $x \in [b_t, b]$ so when we change variables $y = (b - x)t$, $dx = -dy/t$, replace the $x$'s by $b$'s and use $1 = (1 + p)\lambda_0/(\lambda_0 + b)$ we convert the above into

$$g(b)z^{-\lambda_0/(\lambda_0+b)}\frac{u_1V_0}{\lambda_0+b}\left(\frac{\lambda_0+b}{a_0+b}\right)^{b/(\lambda_0+b)}\int_0^\infty dy\, e^{-y\lambda_0/(\lambda_0+b)}\int_0^\infty r^{-b/(\lambda_0+b)}e^{-r}dr$$

Performing the integrals gives the result with

$$A_1(\lambda_0,b)=g(b)\frac{1}{\lambda_0}\left(\frac{\lambda_0+b}{a_0+b}\right)^{b/(\lambda_0+b)}\Gamma(\lambda_0/(\lambda_0+b)) \tag{3.9}$$

## 4 Bounded distributions, $Z_k$

We now move on to the proofs of Theorems 4 and 5. Recall that we have defined $p_k$ by the relation

$$k+p_k=\sum_{j=0}^{k-1}\frac{\lambda_0+kb}{\lambda_0+jb}.$$

### Proof of Theorem 4

Breaking things down according to the times and the sizes of the mutational changes we have

$$\begin{aligned}EZ_k(t)&=\int_0^b dx_1\, g(x_1)\cdots\int_0^b dx_k\, g(x_k)\int_0^t ds_1\cdots\int_{s_{k-1}}^t ds_k\, V_0e^{\lambda_0 s_1}u_1e^{(\lambda_0+x_1)(s_2-s_1)}\cdots u_ke^{(\lambda_0+x_1+\cdots+x_k)(t-s_k)}\\&=\int_0^b dx_1\, g(x_1)\cdots\int_0^b dx_k\, g(x_k)\int_0^t ds_1\cdots\int_{s_{k-1}}^t ds_k\, V_0u_1\cdots u_ke^{\lambda_0 t}e^{x_1(t-s_1)}\cdots e^{x_k(t-s_k)}.\end{aligned} \tag{4.1}$$

The first step is to show

**Lemma 3—**Let $b_t = b - (2k + 1)(\log t)/t$. The contribution to $EZ_k(t)$ from points $(x_1,\ldots x_k)$ with some $x_i \le b_t$ is $o(t^{-2k}e^{(\lambda_0+kb)t})$.

**Proof:** (3.2) implies that

$$\int_{s_{j-1}}^{t} ds_j\, e^{(x_j+\cdots+x_k)(t-s_j)} = \frac{e^{(x_j+\cdots+x_k)(t-s_{j-1})}-1}{x_j+\cdots+x_k} \le t e^{(x_j+\cdots+x_k)(t-s_{j-1})}.$$

Applying this and working backwards in the above expression for $EZ_k(t)$, we get

$$EZ_k(t) \le t^k V_0 u_1 \cdots u_k \int_0^b dx_1\, g(x_1) \cdots \int_0^b dx_k\, g(x_k) e^{(\lambda_0+x_1+\cdots+x_k)t}$$

and the desired result follows.

With the Lemma established, when we work backwards

$$\int_{s_{j-1}}^{t} ds_j\, e^{(x_j+\cdots+x_k)(t-s_j)} = \frac{e^{(x_j+\cdots+x_k)(t-s_{j-1})}-1}{x_j+\cdots+x_k} \sim \frac{e^{(x_j+\cdots+x_k)(t-s_{j-1})}}{(k-j+1)b}$$

From this and induction, we see that the contribution from points $\{x_1, \ldots x_k\}$ with $x_i \in [b_t, b]$ for all $i$ is

$$\sim \frac{V_0 u_1 \cdots u_k}{b^k k!} g(b)^k \int_{b_t}^b dx_1 \cdots \int_{b_t}^b dx_k\, e^{(\lambda_0+x_1+\cdots+x_k)t}$$

Changing variables $y_i = t(b - x_i)$ the above

$$\sim \frac{V_0 u_1 \cdots u_k g(b)^k}{b^k t^k k!} e^{(\lambda_0+kb)t}$$

which proves the desired result.

In the proof of the last result, we showed that the dominant contribution comes from mutations with $x_i > b_t$. To prove our limit theorem we will also need a result regarding the times at which the mutations to the dominant types occur.

**Lemma 4—**Let $\alpha_k = \frac{2k+1}{kb}$. The contribution to $EZ_k\{t)$ from points with $s_1 \ge \alpha_k \log t$ is $o(t^{-2k} e^{(\lambda_0+kb)t})$.

**Proof:** Replace the $X_i$'s in the exponents by $b$'s, we can see from (4.1) that the expected contribution from points with $s_1 \ge \alpha_k \log t$ is

$$\begin{aligned} &\le b^k G^k V_0 u_1 \cdots u_k \int_{\alpha_k \log t}^t ds_1 \int_{s_1}^t ds_2 \cdots \int_{s_{k-1}}^t ds_k e^{\lambda_0 t} e^{b(t-s_1)} \cdots e^{b(t-s_k)} \\ &\le C e^{\lambda_0 t} \int_{\alpha_k \log t}^t e^{kb(t-s_1)} ds_1 \\ &\le C e^{(\lambda_0+kb)t} t^{-\alpha_k kb} \end{aligned}$$

and the desired result follows.

Recall that

$$k+p_k=\sum_{j=0}^{k-1}\frac{\lambda_0+kb}{\lambda_0+jb}.$$

For the induction used in the next proof, we will also need the corresponding quantity with $\lambda_0$ replaced by $\lambda_0 + x$ and $k$ by $k - 1$

$$k-1+p_{k-1}(x)=\sum_{j=0}^{k-2}\frac{\lambda_0+x+(k-1)b}{\lambda_0+x+jb}$$

which means

$$p_{k-1}(x)=\sum_{j=0}^{k-2}\frac{(k-1-j)b}{\lambda_0+x+jb}$$

The limit will depend on the mutation rates through

$$u_{1,k}=\prod_{j=1}^{k}u_j^{\lambda_0/(\lambda_0+(j-1)b)}$$

Again we will need the corresponding quantity with $k - 1$ terms

$$u_{2,k}(x)=\prod_{j=1}^{k-1}u_{j+1}^{(\lambda_0+x)/(\lambda_0+x+(j-1)b)}.$$

We shall write $u_{2,k} = u_{2,k}(b)$ and note that

$$u_{1,k}=u_1u_{2,k}^{\lambda_0/(\lambda_0+b)} \tag{4.2}$$

**Proof of Theorem 5—**We shall prove the result under the more general assumption that $Z_0(t) \sim V_0e^{\lambda 0t}$ for some constant $V_0$. The result then holds for $k = 1$ by Lemma 2. We shall prove the general result by induction on $k$. To this end, suppose the result holds for $k -1$. Let $Z_k^{s,x,v}(t)$ be the type-$k$ descendants at time $t$ of the 1 mutant at $(s,x,v) \in \mathcal{M}(t)$. Since $Z_1^{s,x}(t) \sim ve^{(\lambda_0+x)(t-s)}$ compared to $Z_0(t) \sim V_0e^{\lambda 0t}$, it follows from the induction hypothesis that

$$E\exp(-\theta(t-s)^{k-1+p_{k-1}(x)}e^{-(\lambda_0+x+(k-1)b)(t-s)}Z_k^{s,x,v}(t)) \to \exp(-c_{k-1}(\lambda_0+x,b)vu_{2,k}(x)\theta^{(\lambda_0+x)/(\lambda_0+x+(k-1)b)}) \tag{4.3}$$

Integrating over the contributions from the three-dimensional point process we have

$$E\exp(-\theta Z_k(t))=\exp\left(-\int_0^b dx\, g(x)\int_0^t ds\, u_1 V_0 e^{\lambda_0 s}\int_0^\infty dv\left(\frac{\lambda_0+x}{a_0+x}\right)^2\exp\left(-\frac{\lambda_0+x}{a_0+x}v\right)(1-\varphi_{x,v,t-s}^{k-1}(\theta))\right)$$

where $\varphi_{x,v,t-s}^{k-1}(\theta)=E\exp(-\theta Z_k^{0,x,v}(t-s))$. To prove the desired result we need to replace $\theta$ by $\theta t^{k+pk}e^{-(\lambda_0+kb)t}$. Doing this with (4.3) in mind we have

$$E\exp(-\theta t^{k+p_k}e^{-(\lambda_0+kb)t}Z_k(t))$$
$$=\exp\left(-\int_0^b dx\, g(x)\int_0^t ds\, u_1 V_0 e^{\lambda_0 s}\int_0^\infty dv\left(\frac{\lambda_0+x}{a_0+x}\right)^2\exp\left(-\frac{\lambda_0+x}{a_0+x}v\right)\{1-\varphi_{x,v,t-s}^{k-1}(\theta t^{k+p_k}e^{-(\lambda_0+x+(k-1)b)(t-s)}e^{-(b-x)t}e^{-(\lambda_0+x+b(k-1))s})\}\right) \tag{4.4}$$

By Lemmas 3 and 4, we can restrict attention to $x \in [b_t, b]$ and $s \le \alpha_k \log t$. The first restriction implies that all of the $x$'s except the one in $(b-x)$ can be set equal to $b$ and the second that we can replace $t$ by $t-s$. Since $(k+p_k)-(k-1+p_{k-1}(b)) = (\lambda_0+kb)/\lambda_0$, the term in the exponential on the righthand side of (4.4) is

$$\simeq -\int_{b_t}^b dx\, g(x)\int_0^{\alpha_k \log t} ds\, u_1 V_0 e^{\lambda_0 s}\int_0^\infty dv\left(\frac{\lambda_0+b}{a_0+b}\right)^2\exp\left(-\frac{\lambda_0+b}{a_0+b}v\right)(1$$
$$-\varphi_{x,v,t-s}^{k-1}(\theta(t-s)^{k-1+p_{k-1}(b)}e^{-(\lambda_0+kb)(t-s)}t^{(\lambda_0+kb)/\lambda_0}e^{-(b-x)t}e^{-(\lambda_0+kb)s}))$$

Changing variables $s = R(t) + r$ where $R(t) = (1/\lambda_0)(\log t)$, and $y = (b-x)t$, $dy = -tdx$ the above becomes

$$=-g(b)\int_0^{(2k+1)\log t} dy\int_0^\infty dy\left(\frac{\lambda_0+b}{a_0+b}\right)^2\exp\left(-\frac{\lambda_0+b}{a_0+b}v\right)\int_{-R(t)}^{\alpha_k\log t-R(t)} dr\, u_1 V_0 e^{\lambda_0 r}(1-\varphi_{x,v,t-s}^{k-1}(\theta(t-s)^{k-1+p_{k-1}(b)}e^{-(\lambda_0+kb)(t-s)}e^{-y}e^{-(\lambda_0+kb)r}))$$

Using (4.3) now we have that the $1-\varphi$ term converges to

$$1-\exp(-c_{k-1}(\lambda_0+b,b)vu_{2,k}[\theta e^{-y}]^{(\lambda_0+b)/(\lambda_0+kb)}e^{-(\lambda_0+b)r})$$

To simplify this expression, we let

$$r=\frac{1}{\lambda_0+b}(q+Q(v,y)) \text{ where } Q(v,y)=\log\{c_{k-1}(\lambda_0+b,b)vu_{2,k}[\theta e^{-y}]^{(\lambda_0+b)/(\lambda_0+kb)}\}$$

$dr = dq/(\lambda_0 + b)$. Plugging this into $e^{\lambda 0 r}$ results in

$$e^{q\lambda_0/(\lambda_0+b)}(c_{k-1}(\lambda_0+b,b)vu_{2,k})^{\lambda_0/(\lambda_0+b)}\theta^{\lambda_0/(\lambda_0+kb)}e^{-y\lambda_0/(\lambda_0+kb)}$$

so we conclude that the term in the exponential on the righthand side of (4.4) converges to

$$-g(b)\frac{c_{k-1}(\lambda_0+b,b)^{\lambda_0/(\lambda_0+b)}}{\lambda_0+b}V_0u_1u_{2,k}^{\lambda_0/(\lambda_0+b)}\theta^{\lambda_0/(\lambda_0+kb)}\int_0^\infty dv\left(\frac{\lambda_0+b}{a_0+b}\right)^2 v^{\lambda_0/(\lambda_0+b)}\exp\left(-\frac{\lambda_0+b}{a_0+b}v\right)\int_0^\infty dy\, e^{-y\lambda_0/(\lambda_0+kb)}\int_{-\infty}^\infty \frac{dq}{\lambda_0+b}e^{q\lambda_0/(\lambda_0+b)}(1-\exp(- \tag{4.5}$$

To obtain a cleaner expression for the constants, we begin by noting that the change of variables $w = v(\lambda_0 + b)/(a_0 + b)$, $dw = dv(\lambda_0 + b)/(a_0 + b)$ implies that

$$\int_0^\infty dv\left(\frac{\lambda_0+b}{a_0+b}\right)^2 v^{\lambda_0/(\lambda_0+b)}\exp\left(-\frac{\lambda_0+b}{a_0+b}v\right)=\left(\frac{a_0+b}{\lambda_0+b}\right)^{-1+\lambda_0/(\lambda_0+b)}\Gamma(1+\lambda_0/(\lambda_0+b)) \tag{4.6}$$

Furthermore, we also have

$$\int_0^\infty dy\, e^{-y\lambda_0/(\lambda_0+kb)}=\frac{\lambda_0+kb}{\lambda_0} \tag{4.7}$$

Finally, to evaluate the third integral in (4.5), we make the change of variables $x = e^{-q}$, $dx = -e^{-q}\, dq$, or $dq = -dx/x$ to show that it is

$$=\int_0^\infty dx\, x^{-1-\lambda_0/(\lambda_0+b)}(1-e^{-x})\, dx.$$

Integrating by parts $f(x) = 1-e^{-x}$, $g'(x) = x^{-1-\lambda_0/(\lambda_0+b)}$, $f'(x) = e^{-x}$, $g(x) = x^{-\lambda_0/(\lambda_0+b)}(\lambda_0+ b)/\lambda_0$ shows that the previous expression is

$$\frac{\lambda_0+b}{\lambda_0}\Gamma(1-\lambda_0/(\lambda_0+b)) \tag{4.8}$$

Combining (4.6) – (4.8) and using (4.2), we conclude that the expression in (4.5) is

$$=c_{k-1}(\lambda_0+b,b)^{\lambda_0/(\lambda_0+b)}\cdot g(b)\frac{\lambda_0+kb}{\lambda_0}\cdot V_0u_{1,k}\theta^{\lambda_0/(\lambda_0+kb)}\cdot\frac{1}{\lambda_0}\left(\frac{a_0+b}{\lambda_0+b}\right)^{-1+\lambda_0/(\lambda_0+b)}\Gamma(1+\lambda_0/(\lambda_0+b))\Gamma(1-\lambda_0/(\lambda_0+b)) \tag{4.9}$$

Setting $c_k(\lambda_0,b)$ equal to the quantity in (4.9) divided by $V_0u_{1,k}\theta^{\lambda_0/(\lambda_0+kb)}$ we have proved the result.

To work out an explicit formula for the constant and to compare with Durrett and Moseley (2009), it is useful to let $\lambda_j = \lambda_0 + jb$, $a_j = a_0 + jb$ and

$$c_{h,j}=\frac{1}{\lambda_{j-1}}\left(\frac{a_j}{\lambda_j}\right)^{-1+\lambda_{j-1}/\lambda_j}\Gamma(1+\lambda_{j-1}/\lambda_j)\Gamma(1-\lambda_{j-1}/\lambda_j)$$

From this we see that

$$\begin{aligned}c_k(\lambda_0,b)&=c_{k-1}(\lambda_1,b)^{\lambda_0/\lambda_1}g(b)^{\frac{\lambda_k}{\lambda_0}}c_{h,1}\\&=c_{k-2}(\lambda_2,b)^{\lambda_0/\lambda_2}\cdot\left(g(b)^{\frac{\lambda_{k-1}}{\lambda_0}}c_{h,2}\right)^{\lambda_0/\lambda_1}\cdot g(b)^{\frac{\lambda_k}{\lambda_0}}c_{h,1}\end{aligned}$$

and hence

$$c_k(\lambda_0,b)=\prod_{j=1}^{k}\left(g(b)\frac{\lambda_{k-j+1}}{\lambda_0}c_{h,j}\right)^{\lambda_0/\lambda_{j-1}}$$

In Durrett and Moseley (2009) if we let $\mathcal{F}_{k-1}$ be the $\sigma$-field generated by $Z_j(t)$ for $j \le k$ and all $t \ge 0$ then

$$E(e^{-\theta V_k}|\mathcal{F}_{k-1})=\exp(-u_kV_{k-1}c_{h,k}\theta^{\lambda_{k-1}/\lambda_k})$$

Iterating we have

$$\begin{aligned}E(e^{-\theta V_k}|\mathcal{F}_{k-2})&=E(\exp(-u_kV_{k-1}c_{h,k}\theta^{\lambda_{k-1}/\lambda_k})|\mathcal{F}_{k-2})\\&=\exp\left(-u_{k-1}u_k^{\lambda_{k-2}/\lambda_{k-1}}V_{k-2}c_{h,k-1}c_{h,k}^{\lambda_{k-2}/\lambda_{k-1}}\theta^{\lambda_{k-2}/\lambda_k}\right)\end{aligned}$$

and hence

$$E(e^{-\theta V_k}|V_0)=\exp(-c_{\theta,k}V_0u_{1,k}\theta^{\lambda_0/\lambda_k})$$

where $c_{\theta,k}=\prod_{j=1}^{k}c_{h,j}^{\lambda_0/\lambda_{j-1}}$.

## 5 Proofs for unbounded distributions

In this Section, we prove Theorems 6 and 7. The first step is to show that unlike in the case of bounded mutational advances, for unbounded distributions, the main contribution to the limit is given by the descendants of a single mutation. We will later show that the largest growth rate will come from $z = O(t^{(\alpha+1)/\alpha})$ so the next result is enough. Recall that $z(s, x) = (\lambda_0 + x)(t - s)$ is the growth rate of the family descended from a mutant at $(s, x)$.

### Lemma 5

For any $\bar{z}$, $t > 0$, we have

$$E\left(\sum_{(s,x):z(s,x)\le\bar{z}} Z_1^{s,x}(t)\right) \le V_0 u_1 t e^{\lambda_0 t+\bar{z}}$$

**Proof—**$z(s, y) \le \bar{z}$ if and only if we have a mutation at time $s$ which increases fitness by $y \le \bar{z}/(t-s) - \lambda_0$ and hence, the expected number of individuals produced by mutations with growth rates $\le \bar{z}$ is

$$V_0 u_1 \int_0^t \int_0^{\frac{\bar{z}}{t-s}-\lambda_0} e^{\lambda_0 s} \cdot e^{z(s,y)} \nu(dy)\, ds \le V_0 u_1 t e^{\lambda_0 t+\bar{z}}$$

since $\int_0^\infty \nu(dy) = 1$.

To motivate the proof of the general result, we begin with the case when $\alpha = 1$. Recall that the mean number of mutations with growth rate larger than $z$ has

$$\mu(z,\infty) = K V_0 u_1 \int_0^t q(z/(t-s) - \lambda_0)\left(\frac{z}{t-s} - \lambda_0\right)^\beta \exp(\varphi(s,z))\, ds$$

where $q$, $\varphi$ are as in (1.4), (1.5).

### Proof of Theorem 6

Since

$$Z_1(t) = \sum_{(s,x)\in\mathcal{N}(t)} Z_1^{s,x}(t) = \sum_{(s,x):z(s,x)\le z} Z_1^{s,x}(t) + \sum_{(s,x):z(s,x)>z} Z_1^{s,x}(t)$$

for any $z > 0$, we have

$$\frac{1}{t}\log Z_1(t) \sim \frac{1}{t}\left[\log\left(\sum_{(s,x):z(s,x)\le z} Z_1^{s,x}(t)\right) \vee \log\left(\sum_{(s,x):z(s,x)>z} Z_1^{s,x}(t)\right)\right]$$

as $t \to \infty$. Lemma 5 tells us that if there is a mutation with growth rate $z = O(t^2)$, then the contribution from mutations with growth rates smaller than $z - \varepsilon$ can be ignored so it suffices to describe the distribution of the largest growth rates. We will show that if

$$z_t = c_0 t^2 \left(1 + \frac{(2\beta+1)\log t}{\lambda_0 t} + f(t)\right)$$

then

$$\mu(z_t, \infty) \to \begin{cases} 4c_0^{\beta}(\pi/\lambda_0)^{1/2} V_0 u_1 \exp(\gamma\lambda_0 - 2\lambda_0 x/2c_0) & \text{if} \quad tf(t) \to \frac{x}{c_0},\ x \geq 0 \\ 0 & \text{if} \quad tf(t) \to \infty \end{cases} \tag{5.1}$$

so that the largest growth rate is $O(t^2)$ and comes from the rightmost particle in the point process with intensity given by (1.8).

To prove (5.1), we first need to locate the maximum of $\varphi$. Let $z > \lambda_0 t$ so that there exists a unique maximum $s_z$. Solving $\varphi_s(s, z) = 0$ and using the expression for $\varphi_s$ in (1.6) yields

$$s_z = t - a_0 z^{1/2}$$

where $a_0 = (\gamma/\lambda_0)^{1/2} = (4c_0)^{-1/2}$ which leads to the expression

$$\begin{aligned} \varphi(s_z, z) &= \lambda_0 t - \lambda_0(t - s_z) - \gamma\left(\tfrac{z}{t-s} - \lambda_0\right) \\ &= \lambda_0 t - \lambda_0 a_0 z^{1/2} - \gamma z^{1/2}/a_0 + \gamma\lambda_0 \\ &= \lambda_0(t - 2a_0 z^{1/2}) + \gamma\lambda_0. \end{aligned} \tag{5.2}$$

If we take

$$z_t = c_0 t^2 \left(1 + \frac{(2\beta+1)\log t}{\lambda_0 t} + f(t)\right) = \left(\frac{t}{2a_0}\right)^2 \left(1 + \frac{(2\beta+1)\log t}{\lambda_0 t} + f(t)\right)$$

in (5.2) and use $(1 + y)^{1/2} = 1 + y/2 + O(y^2)$, we obtain

$$\varphi(s_{z_t}, z_t) = -\frac{(2\beta+1)\log t}{2} - \lambda_0 t f(t) + \gamma\lambda_0 + o(1) \tag{5.3}$$

as $t \to \infty$. Furthermore, (1.7) implies that

$$\varphi_{ss}(s_{z_t}, z_t) = -\frac{2\gamma z_t}{(t-s_{z_t})^3} = -\frac{2\gamma}{a_0^3 z_t^{1/2}} = -\frac{a}{t} + o(1)$$
$$\varphi_{sss}(s_{z_t}, z_t) = -\frac{6\gamma z_t}{(t-s_{z_t})^4} = -\frac{6\gamma}{a_0^4 z_t} = -\frac{24\gamma}{a_0^4 t^2} + o(1)$$

as $t \to \infty$ with $a = 4\gamma/a_0^2$. Since $\varphi_s(s_z, z) = 0$, taking a Taylor expansion around $s_z$ yields

$$\varphi(s, z_t) = \varphi(s_{z_t}, z_t) - \frac{a}{2t}(s - s_{z_t})^2 + g(s, z_t) \tag{5.4}$$

where $|g(s, z)| \le C|s - s_z|^3/t^2$ for all $s$. Also note that letting

$$\psi(s, z) = q(z/(t-s) - \lambda_0)\left(\frac{z}{t-s} - \lambda_0\right)^\beta$$

and recalling (1.4), we have

$$\begin{aligned}\psi(s_{z_t}, z_t) &= q(z_t/(t - s_{z_t}) - \lambda_0)\left(\tfrac{z}{t-s_{z_t}} - \lambda_0\right)^\beta \\ &= c_t z_t^{\beta/2}/a_0^\beta + o(z_t^{\beta/2}) \\ &= c_t (2c_0)^\beta t^\beta + o(t^\beta)\end{aligned}$$

where $c_t \to 1$ as $t \to \infty$ so that

$$\psi(s, z_t) = c_t (2c_0)^\beta t^\beta + g_2(s, z_t)$$

where $|g_2(s, z)||s - s_z|^{-1} t^{-\beta} = o(1)$.

Write

$$\int_0^t \psi(s, z_t) e^{\varphi(s, z_t)} ds = \int_A \psi(s, z_t) e^{\varphi(s, z_t)} ds + \int_{A^c} \psi(s, z_t) e^{\varphi(s, z_t)} ds$$

where $A = \{s: |s - s_{z_t}| \le C(t \log t)^{1/2}\} \cap [0, t]$. Since concavity implies that for $s \in A^c$ and $C$ sufficiently large, we have

$$\exp(\varphi(s, z_t)) \le \frac{1}{t^{2+\beta}} \exp(\varphi(s_{z_t}, z_t))$$

the contribution of the second integral is negligible. After the change of variables $s = s_{z_t} + (t/a)^{1/2} r$, when $t$ is large, the first integral becomes

$$\int_A \psi(s,z_t)e^{\varphi(s,z_t)}ds=(c_t(2c_0)^\beta t^\beta+o(1))e^{\varphi(s_{z_t},z_t)}\int_{-C(\log t)^{1/2}}^{C(\log t)^{1/2}} e^{g(s,z_t)}e^{-r^2/2}(t/a)^{1/2}dr.$$

and therefore since $|g(s, z_t)| \le C(t \log t)^{3/2}/t^2$ when $s \in A$, we have

$$\mu(z_t,\infty)=KV_0u_1\int_0^t \psi(s,z_t)e^{\varphi(s,z_t)}ds \sim KV_0u_1bt^{\beta+1/2}e^{\varphi(s_{z_t},z_t)} \tag{5.5}$$

where $b=(2c_0)^\beta\sqrt{2\pi/a}=(2c_0)^\beta(\pi/\lambda_0)^{1/2}$. Since

$$\varphi(s_{z_t},z_t)=-\frac{(2\beta+1)\lambda_0\log t}{2}-\lambda_0 tf(t)+\gamma\lambda_0$$

we can conclude that

$$\mu(z_t,\infty)\to\begin{cases} KV_0u_1b\exp(\gamma\lambda_0-2\lambda_0a_0^2x)=V_0u_1b\exp(\gamma\lambda_0-2\lambda_0x/2c_0) & \text{if } tf(t)\to\frac{x}{c_0}\\ 0 & \text{if } tf(t)\to\infty\end{cases}$$

which proves (5.1).

When $\alpha \neq 1$, we no longer have an explicit formula for the maximum value $s_z$ which complicates the process of identifying the largest growth rate. We shall assume for convenience that $\alpha > 1$ is an integer.

## Proof of Theorem 7

As in the proof of Theorem 6, it suffices to describe the distribution for the largest growth rates. Let $z > \lambda_0 t$ so the maximum $s_z$ exists. To find a useful expression for the value of $\varphi(s_z, z)$, we write

$$\varphi(s,z)=\lambda_0 t-\lambda_0(t-s)-\gamma\left(\frac{z}{t-s}-\lambda_0\right)^\alpha.$$

Using the definition of $s_z$ as the solution to $\varphi_s(s_z, z) = 0$ yields the condition that

$$(t-s_z)^{\alpha+1}=\frac{\alpha\gamma}{\lambda_0}z^\alpha\left(1-\lambda_0\frac{t-s_z}{z}\right)^{\alpha-1}$$

i.e.,

$$t-s_z=\left(\frac{\alpha\gamma}{\lambda_0}\right)^{1/(\alpha+1)}z^{\alpha/(\alpha+1)}\left(1-\lambda_0\frac{t-s_z}{z}\right)^{(\alpha-1)/(\alpha+1)}.$$

If we substitute the right side of this equation back in for $t - s_z$ in the parenthesis, then writing $a_0 = (\alpha\gamma/\lambda_0)^{1/(\alpha+1)}$ we have

$$t - s_z = a_0 z^{\alpha/(\alpha+1)} \left(1 - \lambda_0 a_0 z^{-1/(\alpha+1)} \left(1 - \tfrac{\lambda_0 (t-s_z)}{z}\right)^{\frac{\alpha-1}{\alpha+1}}\right)^{\frac{\alpha-1}{\alpha+1}}$$
$$= a_0 z^{\alpha/(\alpha+1)} \left(1 - \lambda_0 a_0 z^{-1/(\alpha+1)} \left(1 - \lambda_0 a_0 z^{-1/(\alpha+1)} \left(1 - \tfrac{\lambda_0 (t-s_z)}{z}\right)^{\frac{\alpha-1}{\alpha+1}}\right)^{\frac{\alpha-1}{\alpha+1}}\right)^{\frac{\alpha-1}{\alpha+1}}$$

Repeating this $\alpha$ times and then using the approximation $(1 - x)^n = 1 - nx + O(x^2)$ with $n = (\alpha - 1)/(\alpha + 1)$ yields

$$t - s_z = z^{\alpha/(\alpha+1)} \left( \sum_{j=0}^{\alpha} a_j z^{-j/(\alpha+1)} + O(z^{-1}) \right) \tag{5.6}$$

where

$$a_j = a_0 \left( \frac{\lambda_0 a_0 (\alpha - 1)}{\alpha + 1} \right)^j$$

for $j \geq 1$. The error term is $O(z^{-1})$ because

$$0 < (1 - \lambda_0 (t - s)/z) \leq 1$$

for all $z > \lambda_0 t$ and $s \leq t$. Factoring out $a_0$ in (5.6) and using $(1 + x)^{-1} = \Sigma(-x)^j$ when $|x| < 1$, we have that

$$\frac{z}{t-s} - \lambda_0 = a_0^{-1} z^{1/(\alpha+1)} \left(1 - \sum_{i_1=1}^{\alpha} a_0^{-1} a_{i_1} z^{-i_1/(\alpha+1)} + \sum_{i_1,i_2=1}^{\alpha} a_0^{-2} a_{i_1} z^{-(i_1+i_2)/(\alpha+1)} - \cdots + (-1)^{\alpha} \sum_{i_1,\ldots,i_\alpha=1}^{\alpha} a_0^{-\alpha} \prod_{j=1}^{\alpha} a_{i_j} z^{-\sum_{j=1}^{\alpha} i_j/(\alpha+1)} + O(z^{-1})\right) - \lambda_0 z^{1/(\alpha+1)} z^{-1}$$
$$= z^{1/(\alpha+1)} \left( \sum_{j=0}^{\alpha} b_j z^{-j/(\alpha+1)} + O(z^{-1}) \right) \tag{5.7}$$

for large $z$ where the $b_j$ are given by

$$b_0=1/a_0$$
$$b_1=-a_1/a_0^2-\lambda_0$$
$$b_2=-(a_2-a_1^2)/a_0^3$$
$$b_3=-(a_4-2a_1a_3-a_2^2-3a_1^2a_2+a_1^4)/a_0^4$$

and in general,

$$b_i=\sum_{k=1}^{\alpha}\sum_{\substack{i_1,\dots,i_k=1\\ i_1+\cdots+i_k=i}}^{\alpha}(-a_0)^{-(k+1)}\prod_{j=1}^{k}a_{i_j}.$$

(5.7) implies that

$$\begin{aligned}-\gamma\left(\frac{z}{t-s}-\lambda_0\right)^{\alpha}&=-\gamma z^{\alpha/(\alpha+1)}(b_0^{\alpha}\\&+\alpha b_0^{\alpha-1}b_1z^{-1/(\alpha+1)}\\&+\left(\alpha b_0^{\alpha-1}b_2+\binom{\alpha}{2}b_0^{\alpha-2}b_1^2\right)z^{-2/(\alpha+1)}\\&+\cdots+(\alpha b_0^{\alpha-1}b_{\alpha}\\&+\cdots+b_1^{\alpha})z^{\alpha/(\alpha+1)}+O(z^{-1}))\end{aligned}$$

and therefore,

$$\begin{aligned}\varphi(s_z,z)&=\lambda_0t+\lambda_0(t-s)-\gamma\left(\tfrac{z}{t-s}-\lambda_0\right)^{\alpha}\\&=\lambda_0t+\sum_{j=0}^{\alpha}d_jz^{\frac{\alpha-j}{\alpha+1}}+O(z^{-1/(\alpha+1)})\end{aligned}\tag{5.8}$$

where the $d_j$ can be calculated explicitly, for example:

$$d_0=-\lambda_0a_0-\gamma b_0^{\alpha}$$
$$d_1=-\lambda_0a_1-\gamma\alpha b_0^{\alpha-1}c_1$$
$$d_2=-\lambda_0a_2-\gamma\left(\alpha b_0^{\alpha-1}b_2+\binom{\alpha}{2}b_0^{\alpha-2}b_1^2\right)$$
$$d_3=-\lambda a_3-\gamma\left(\alpha b_0^{\alpha-1}b_3+\binom{\alpha}{2}b_0^{\alpha-2}b_1b_2+\binom{\alpha}{3}b_0^{\alpha-3}b_1^3\right).$$

To figure out the distribution of the growth rate for the largest mutant, we let $c_0=(-\lambda_0/d_0)^{(\alpha+1)/\alpha}$ and then search for $\kappa_j$, $j=1,\dots,\alpha-1$ and $\kappa$ so that plugging

$$z_t = c_0 t^{(\alpha+1)/\alpha}\left(1+\sum_{j=1}^{\alpha-1}\kappa_j t^{-j/\alpha}+\frac{\kappa \log t}{t}+f(t)\right)$$

into (5.8) yields

$$\varphi(s_{z_t}, z_t) = k_1 - k_2 t f(t) - k_3 \log t \tag{5.9}$$

for some constants $k_1$, $k_2$, $k_3$. Substituting $z_t$ into (5.8) and writing $\kappa_0 = 1$, $\kappa_\alpha = x/c_0$ to ease the notation we obtain

$$\varphi(s_{z_t}, z_t) = \lambda_0 t + \sum_{j=0}^{\alpha} d_j\left(-\frac{\lambda_0 t}{d_0}\right)^{(\alpha-j)/\alpha}\left(\sum_{j=0}^{\alpha}\kappa_j t^{-j/\alpha}+\kappa t^{-1}\log t\right)^{(\alpha-j)/(\alpha+1)} + O(t^{-1/\alpha}).$$

Since $\lambda_0 t + d_0(-\lambda_0 t/d_0) = 0$, the first order terms in this expansion is $t^{(\alpha-1)/\alpha}$ and after using the Taylor series expansion

$$(1+x)^p = 1 + px + p(p-1)x^2/2 + \cdots + p(p-1)\cdots(p-\alpha+1)x^\alpha/\alpha! + O(x^{\alpha+1})$$

we obtain

$$\varphi(s_{z_0}, z_0) = \sum_{j=1}^{\alpha}\rho_j t^{(\alpha-j)/\alpha} + \rho \log t + O(t^{-1/\alpha}\log t) \tag{5.10}$$

where

$$\rho = d_0\left(-\tfrac{\lambda_0}{d_0}\right)\left(\tfrac{\alpha}{\alpha+1}\right)\kappa = -\tfrac{\alpha\lambda_0}{\alpha+1}\kappa$$

$$\rho_1 = d_0\left(-\tfrac{\lambda_0}{d_0}\right)\left(\tfrac{\alpha}{\alpha+1}\right)c_1 + d_1\left(-\tfrac{\lambda_0}{d_0}\right)^{(\alpha-1)/\alpha}$$

$$\rho_2 = d_0\left(-\tfrac{\lambda_0}{d_0}\right)\left[\tfrac{\alpha}{\alpha+1}c_2 + \tfrac{\alpha}{\alpha+1}\left(\tfrac{\alpha}{\alpha+1}-1\right)c_1^2\right] + d_1\left(-\tfrac{\lambda_0}{d_0}\right)^{(\alpha-1)/\alpha}\left(\tfrac{\alpha-1}{\alpha}\right)c_1 + d_2\left(-\tfrac{\lambda_0}{d_0}\right)^{(\alpha-2)/\alpha}$$

and in general

$$\rho_j = \sum_{i=0}^{j} d_i\left(-\frac{\lambda_0}{d_0}\right)^{(\alpha-i)/\alpha}\sum_{k=1}^{j-i}\prod_{\ell=1}^{k}\left(\frac{\alpha-i}{\alpha+1}-\ell+1\right)\kappa_{i_\ell}$$

$j = 1, 1, \ldots, \alpha$ where for each $i$ and $k$, in the inner product, $i_1, \ldots, i_k$ are always chosen to satisfy $i_1 + i_2 + \cdots + i_k = j - i$. Since $\rho_j$ depends only on $\kappa_i$, $i \le j$, then after noting that the coefficient

of $\kappa_j$ in $\rho_j$ is $-\alpha\lambda_0/(\alpha+1)$, we can use forward substitution to solve the system $\rho_j = 0$, $j = 1, 2, \ldots, \alpha - 1$ for $\kappa_j$ to obtain the recursive formulas

$$c_j \equiv \kappa_j = -\frac{\alpha+1}{\alpha\lambda_0}\left(\rho_j - \frac{-\alpha\lambda_0}{\alpha+1}\kappa_j\right) \tag{5.11}$$

for $i = 1, 2, \ldots, \alpha - 1$. Setting $\rho = -k_3$ yields

$$\kappa = \frac{(\alpha+1)k_3}{\alpha\lambda_0}$$

and for this choice of $c_j$, $\kappa$, we obtain (5.9) with

$$k_2 = -\frac{\alpha}{\alpha+1}\frac{d_0}{c_0}\left(-\frac{\lambda_0}{d_0}\right) = \frac{\alpha\lambda_0}{(\alpha+1)c_0}$$

and $k_1 = -(\rho_\alpha - k_2 x)$. Since

$$\begin{aligned}\left(\frac{z_t}{t - s_{z_t}} - \lambda_0\right)^\beta &= z_t^{\beta/(\alpha+1)}/a_0^\beta + o(z_t^{\beta/(\alpha+1)}) \\ &= \left(\frac{c_0^{1/(\alpha+1)}}{a_0}\right)^\beta t^{\beta/\alpha} + o(z_t^{\beta/(\alpha+1)})\end{aligned}$$

choosing $k_3 = (2\beta/\alpha + 1)/2$ replaces (5.3) in the proof of Theorem 6.

Now substituting (5.6) and (5.7) in (1.7) yields

$$\begin{aligned}\varphi_{ss}(s_z, z) &= -\alpha(\alpha-1)\gamma z^{\frac{\alpha-2}{\alpha+1}}\left(\sum_{j=0}^{\alpha} b_j z^{-j/(\alpha+1)} + O(z^{-1})\right)^{\alpha-2} \\ &\times \frac{z^2}{z^{4\alpha/(\alpha+1)}\left(\sum_{j=0}^{\alpha} a_j z^{-j/(\alpha+1)} + O(z^{-1})\right)^4} \\ &- \alpha\gamma z^{\frac{\alpha-1}{\alpha+1}}\left(\sum_{j=0}^{\alpha} b_j z^{-j/(\alpha+1)} + O(z^{-1})\right)^{\alpha-1} \\ &\times \frac{2z}{z^{3\alpha/(\alpha+1)}\left(\sum_{j=0}^{\alpha} a_j z^{-j/(\alpha+1)} + O(z^{-1})\right)^3} \\ &= [-\alpha(\alpha-1)\gamma b_0^{\alpha-2}/a_0^4 - \alpha\gamma b_0^{\alpha-1}/a_0^3] z^{-\alpha/(\alpha+1)} + o(z^{-\alpha/(\alpha+1)}) \\ &= -\frac{\alpha^2\gamma}{a_0^{\alpha+2}} z^{-\alpha/(\alpha+1)} + o(z^{-\alpha/(\alpha+1)})\end{aligned}$$

where in the second to last line we have used the fact that $b_0 = a_0^{-1}$. When $z = z_t$, this becomes

$$\varphi_{ss}(s_{z_t}, z_t) = -\frac{a}{t} + o(t^{-1})$$

where

$$a = \frac{\alpha^2 \gamma}{a_0^{\alpha+2} c_0^{\alpha/(\alpha+1)}}.$$

Since $\varphi_s(s_z, z) = 0$ and a calculation similar to the one above shows that $\varphi_{sss}(s_{z_t}, z_t) = O(t^{-2})$, we have

$$\varphi(s, z_t) = \varphi(s_{z_t}, z_t) - \frac{a}{2}(s - s_{z_t})^2 + g(s_{z_t}, z_t)$$

where $|g(s,z)| \le C|s - s_z|^3/t^2$ for all $s$. This replaces (5.4) from the $\alpha = 1$ proof and the rest of the proof is the same. Note that the intensity for the limiting point process is given by

$$KV_0 u_1 \left( \frac{c_0^{1/(\alpha+1)}}{a_0} \right)^{\beta} \sqrt{2\pi/a} \exp(k_1 - k_2 x). \tag{5.12}$$

**Remark 1—**From (5.6), we have

$$t - s_{z_t} \sim a_0 (c_0 t^{(\alpha+1)/\alpha})^{(\alpha+1)/\alpha} = \frac{\alpha t}{\alpha+1}$$

which tells us that the time at which the mutant with largest growth rate is born is ~ t/(α + 1).

## 6 Discussion

In this paper, we have analyzed a multi-type branching process model of tumor progression in which mutations increase the birth rates of cells by a random amount. We studied both bounded and unbounded distributions for the random fitness advances and calculated the asymptotic rate of expansion for the *k*th generation of mutants.

In the bounded setting, we found that there are only two parameters of the distribution that affect the limiting growth rate of the *k*th generation (see Theorems 1, 2, 4, and 5): the upper bound for the support of the distribution and the value of its density at the upper bound. This is a rather intuitive result since one would expect that in the long run, the *k*th generation will be dominated by mutants with the maximum possible fitness. In addition, we found that there is a polynomial correction to the exponential growth of the *k*th generation. This correction is not present in the case where the fitness advances are deterministic. We have discussed this point in further detail in Section 1.1 and after the proof of Theorem 5 in Section 4. Finally, we showed that the limiting population is descended from several different mutations (see Theorem 3).

In the unbounded setting, we assumed that the distribution of the fitness advance has tail

$$\nu(x, \infty) \sim K x^{\beta} e^{-\gamma x^{\alpha}} \tag{6.1}$$

where $\alpha, \beta, \gamma$, and $K$ are parameters. We found that the population of cells with a single mutation grows asymptotically at a super-exponential rate $\exp(t^{(\alpha+1)/\alpha})$ (see Theorems 6 and 7) and at large times, most of the first generation is derived from a single mutation (see Lemma 5 and the preceding paragraph). The super-exponential growth rate suggests that distributions of the form (6.1), which includes the exponential distribution that is often used to model fitness advances of organisms under selective pressure, is not a good choice for modeling the mutational advances in the progression to cancer where there is very little evidence for populations growing at a super-exponential rate.

These conclusions provide several interesting contributions to the existing literature on evolutionary models of cancer progression. First, our model generalizes previous multi-type branching models of tumor progression by allowing for random fitness advances as mutations are accumulated and provides a mathematical framework for further investigations into the role played by the fitness distribution of mutational advances in driving tumorigenesis. Second, we have discovered that bounded distributions lead to exponential growth whereas unbounded distributions lead to super-exponential growth. This dichotomy might provide a new method for testing whether a tumor population has evolved with an unbounded distribution of mutational advances. Third, we observe that in the case of bounded distributions, the growth rate of the tumor is somewhat 'robust' with respect to the mutational fitness distribution and depends only on its upper endpoint. Finally, our calculations of the growth rates for the $k$th generation of mutants serve as a groundwork for studying the evolution and role of heterogeneity in tumorigenesis. These implications will be explored further in future work.

## References


Becskei A, Kaufmann BB, van Oudenaarden A. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. Nature Genetics 2005;9:937–944. [PubMed: 16086016]

Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler KW, Velculescu VE, Vogelstein B, Nowak MA. Genetic progression and the waiting time to cancer. PLoS Computational Biology 2007;3 paper e225.

Beisel CJ, Rokyta DR, Wichman HA, Joyce P. Testing the extreme value domain of attraction for distributions of beneficial fitness effects. Genetics 2007;176:2441–2449. [PubMed: 17565958]

Bodmer W, Tomlinson I. Failure of programmed cell death and differentiation as causes of tumors: some simple mathematical models. Proc Natl Acad Sci USA 1995;92:11130–11134. [PubMed: 7479951]

Coldman AJ, Murray JM. Optimal control for a stochastic model of cancer chemotherapy. Mathematical Biosciences 2000;168:187–200. [PubMed: 11121565]

Cowperthwaite MC, Bull JJ, Meyers LA. Distributions of beneficial fitness effects in RNA. Gentics 2005;170:1449–1457.

Durrett R, Mayberry J. Traveling waves of selective sweeps. Ann Appl Prob. 2009 to appear.

Durrett R, Moseley S. Evolution of resistance and progression to disease during clonal expansion of cancer. Theor Pop Biol. 2009 to appear.

Durrett R, Schmidt D, Schweinsberg J. A waiting time problem arising from the study of multi-stage carcinogenesis. Ann Appl Prob 2009;19:676–718.

Elowitz MB, et al. Stochastic gene expression in a single cell. Science 2002;297:1183–1186. [PubMed: 12183631]

Feinerman O, et al. Variability and robustness in T cell activation from regulated heterogeneity in protein levels. Science 2008;321:1081. [PubMed: 18719282]

Frank SA. Dynamics of Cancer: Incidence, Inheritance and Evolution. Princeton Series in Evolutionary Biology. 2007

Gillespie JH. A simple stochastic gene substitution model. Theor Pop Biol 1983;23:202–215. [PubMed: 6612632]

Gillespie JH. Molecular evolution over the mutational landscape. Evolution 1984;38:1116–1129.

Goldie JH, Coldman AJ. Quantitative model for multiple levels of drug resistance in clinical tumors. Cancer Treatment Reports 1983;67:923–931. [PubMed: 6627236]

Goldie JH, Coldman AJ. The genetic origin of drug resistance in neoplasms: implications for systemic therapy. Cancer Research 1984;44:3643–3653. [PubMed: 6744284]

Haeno H, Iwasa Y, Michor F. The evolution of two mutations during clonal expansion. Genetics 2007;177:2209–2221. [PubMed: 18073428]

Iwasa Y, Michor F, Komorova NL, Nowak MA. Population genetics of tumor suppressor genes. J Theor Biol 2005;233:15–23. [PubMed: 15615616]

Iwasa Y, Nowak MA, Michor F. Evolution of resistance during clonal expansion. Genetics 2006;172:2557–2566. [PubMed: 16636113]

Kassen R, Bataillon T. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. Nature Genetics 2006;38:484–488. [PubMed: 16550173]

Kaern M, et al. Stochasticity in gene expression: from theories to phenotypes. Nature Reviews Genetics 2005;6:451.

Knudson AD. Two genetic hits (more or less) to cancer. Nature Reviews Cancer 2001;1:157–162.

Komarova NL, Wodarz D. Drug resistance in cancer: principles of emergence and prevention. Proc Natl Acad Sci USA 2005;102:9714–9719. [PubMed: 15980154]

Maley CC, et al. Genetic clonal diveresity predicts progression to esophageal adenocarcinoma. Nature Genetics 2006;38:468–473. [PubMed: 16565718]

Maley CC, Forrest. Exploring the relationship between neutral and selective mutations in cancer. Artif Life 2001;6:325–345. [PubMed: 11348585]

Michor F, Iwasa Y, Nowak MA. Dynamics of cancer progression. Nature Reviews Cancer 2004;4:197–205.

Michor F, Nowak MA, Iwasa Y. Stochastic dynamics of metastasis formation. J Theor Biol 2006;240:521–530. [PubMed: 16343545]

Michor F, Iwasa Y. Dynamics of metastasis suppressor gene inactivation. J Theor Biol 2006;241:676–689. [PubMed: 16497335]

Nowak MA, Michor F, Iwasa Y. Genetic instability and clonal expansion. J Theor Biol 2006;241:26–32. [PubMed: 16405914]

Nowell PC. The cloncal evolution of tumor cell populations. Science 1976;194:23–28. [PubMed: 959840]

Orr HA. The distribution of fitness effects among beneficial mutations. Genetics 2003;163:1519–1526. [PubMed: 12702694]

Otto SP, Jones CD. Detecting the undetected: Estimating the total number of loci underlying a quantitative trait. Genetics 2002;156:2093–2107. [PubMed: 11102398]

Rokyta DR, Beisel CJ, Joyce P, Ferris MT, Burch CL, Wichman HA. Beneficial fitness effects are not exponential in two viruses. J Mol Evol 2008;67:368–376. [PubMed: 18779988]

Rozen DE, de Visser JAGM, Gerrish PJ. Fitness effects of fixed beneficial mutations in microbial populations. Curret Biology 2002;12:1040–1045.

Sanjuán R, Moya A, Elena SF. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc Natl Acad Sci, USA 2004;101:8396–8401. [PubMed: 15159545]

Schweinsberg J. The waiting time for *m* mutations. Electron J Probab 2008;13:1442–1478.

Shah SP, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature 2009;461:809–813. [PubMed: 19812674]

Weissman I. Estimation of parameters and large quantiles based on the *k* largest observations. j Amer Stat Assoc 1978;73:812–815.

Wodarz D, Komarova NL. Can loss of apoptosis protect against cancer? Trends Genet 2007;23:232–237. [PubMed: 17382429]

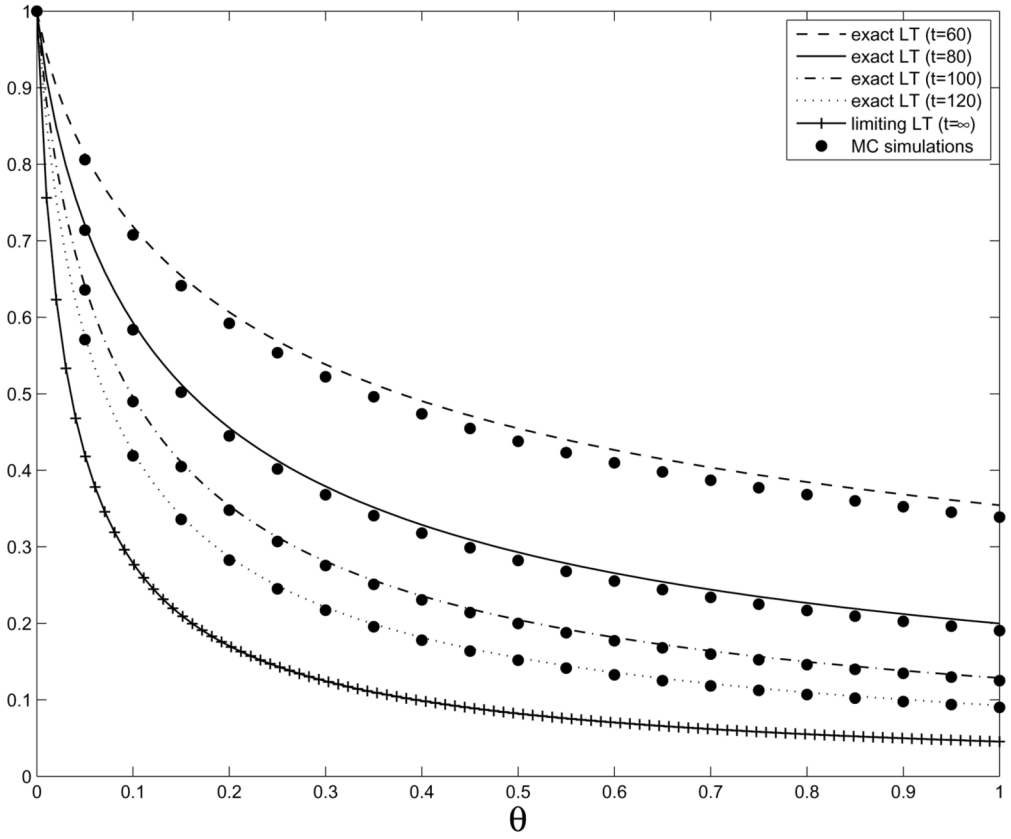


**Figure 1.**
Plot of the exact Laplace transform (LT) for $t^{(1+p)}\, e^{-(\lambda_0+b)t}\, Z_1(t)$ at times $t$ = 60, 80,100,120, the approximations from Monte Carlo (MC) simulations at the corresponding times, and the asymptotic Laplace transform from Theorem 2. Parameter values: $a_0 = 0.2$, $b_0 = 0.1$, $b = 0.01$, and $u_1 = 10^{-3}$. $g$ is uniform on [0, .01].

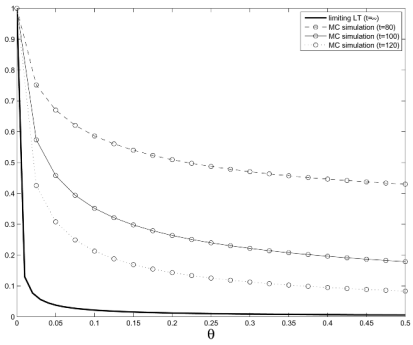


**Figure 2.**

Plot of the approximations to the Laplace transform of $t^{2+p_2}e^{-(\lambda_0+2b)t}\,Z_2(t)$ from Monte Carlo (MC) simulations at times $t$ = 80,100,120 along with the asymptotic Laplace transform from Theorem 5. Parameter values: $a_0 = 0.2$, $b_0 = 0.1$, $b = 0.01$, and $u_1 = u_2 = 10^{-3}$. $g$ is uniform on [0, 0.01].