

A Branching Process Model of Telomere Shortening

Peter Olofsson

Department of Statistics
Rice University, P.O.Box 1892, Houston, TX 77251

Abstract A branching process model of a certain cell population is considered. The population is divided into types according to the number of remaining chromosome end units. This leads to a reducible multi-type Bellman-Harris branching process which turns out to exhibit polynomial growth dynamics: let $M_{k,j}^{(a)}(t)$ be the expected number of j -type cells of age less than a at time t starting from a k -type ancestor cell. Then, as $t \rightarrow \infty$, $M_{k,j}^{(a)}(t) \sim Ct^{k-j}$ where the constant C depends on a, j and k and can be given explicitly. The proof is fairly short and simple, using elementary results from renewal theory and a Tauberian theorem for Laplace-Stieltjes transforms.

Keywords: Branching process, population dynamics, cell population, telomeres.

1 Introduction

Shortening of telomeres is one of the supposed mechanisms of cellular aging and death. Each time a cell divides it loses pieces of its chromosome ends. These ends are called telomeres and consist of repeated sequences of nucleotides, telomere units. The nature of this process is that when a cell divides, one of its daughter cells will have a maintained chromosome length, the other will have shorter length. When a critical number of telomere units is lost, the cell stops dividing. There is an extensive biological literature on the subject; a good reference for the basic facts and an intuitively appealing explanation of the deletion process is Levy et al. [6]. This also seems to be

the first paper to use mathematical models for this problem.

Our purpose is to describe this process as a branching process and explore its asymptotic behavior. This has been done previously in Arino, Kimmel and Webb [1] and Olofsson and Kimmel [8], where it is also discussed how the models fit to data. These references deal with the case when only one telomere unit is lost in each division round. In this paper we explore a more general model which allows for a random number of deleted telomere units, arbitrary lifelength distributions that may also depend on the current telomere length and that counts cells of different ages. This model is also treated with a differential equation approach in Arino, Sanchez and Webb [2], a reference that we will compare some of our results with.

We describe the cell population as a multi-type Bellman-Harris process with a finite type space, the type of a cell being the remaining number of telomere units. The fact that chromosomes may shorten but not lengthen means that the process is reducible, for some general theory on such processes see Mode [7]. The special structure of our process makes it ideal to analyze using some basic results from renewal theory together with a Tauberian theorem for Laplace-Stieltjes transforms. The main features of the process is that there is polynomial growth and that the growth rates are different for different types.

2 The Model

Let the type of a cell be the number of remaining telomere units. The cell population evolves as follows: a cell of type j has a random lifelength with distribution function G_j . According to the description in the previous section, at the moment of death, it splits into two cells, one of type j and one of some lower type. A cell of type 0 does not die or divide. We denote the probability that a cell of type j gets a type i offspring by $p_{j,i}$ where it is assumed for simplicity that $p_{j,i} > 0$ for all $j < i$, a condition which can easily be relaxed. There is no mortality of cells and hence we have

$$p_{j,j} = 1, \quad \sum_{i=0}^{j-1} p_{j,i} = 1.$$

The model is thus essentially the same as in Arino, Sanchez and Webb [2]. Comparing to their model, we note that they assume lifelengths to

be bounded and continuous (lifelength distributions are defined by hazard rates). We do not need these restrictions; the only assumptions we make on the distribution functions G_j is that $G_j(0) = 0$ for all j and that their means $\mu_j = \int_0^\infty tG_j(dt)$ are finite (although most of our results are trivially true also for infinite μ_j 's). For some of our results we also assume the G_j 's to be non-arithmetic, i.e. not supported by any lattice $\{n\delta, n \in \mathbb{Z}_+\}$. Finally, we describe the process in terms of numbers of cells instead of densities and explore the asymptotics of the expected number of cells of certain types and ages as time goes to infinity.

3 Branching Process Description

The cell population thus evolves according to a multi-type Bellman-Harris branching process. If the ancestor cell is of type k , the type space is $\{1, 2, \dots, k\}$. Let $Z_{k,j}^{(a)}(t)$ be the number of j -type cells of age less than a at time t starting from an ancestor of type k and let E_k denote expectation if the ancestor's type is k . We will apply the usual branching process technique of decomposing on the first generation. Let τ be the lifelength of the ancestor and let the types of the two daughter cells be k and σ , where $P(\sigma = i) = p_{k,i}, i = 0, 1, \dots, k-1$. This yields

$$Z_{k,j}^{(a)}(t) = Z_{k,j}^{(a)}(t - \tau) + Z_{\sigma,j}^{(a)}(t - \tau)$$

the expectation of which is

$$\begin{aligned} M_{k,j}^{(a)}(t) &= E_k[Z_{k,j}^{(a)}(t)] = E_k[Z_{k,j}^{(a)}(t - \tau)] + E_k[Z_{\sigma,j}^{(a)}(t - \tau)] \\ &= \int_0^t M_{k,j}^{(a)}(t - u)G_k(du) + \sum_{i=0}^{k-1} \int_0^t M_{i,j}^{(a)}(t - u)p_{k,i}G_k(du) \\ &= \sum_{i=j}^k \int_0^t M_{i,j}^{(a)}(t - u)p_{k,i}G_k(du). \end{aligned} \tag{3.1}$$

since $p_{k,k} = 1$.

The special structure of this process makes it possible to analyze using tools from basic renewal theory. Indeed, a j -type cell initiates a renewal process of j -type cells, each of which will initiate a further renewal processes

of type $j - 1$ and so on. The results we use are standard and may be found in Asmussen [3], p.107 and p.120. We state them for the sake of completeness. Let $N(t)$ be the expected number of renewals in a (pure or delayed with a proper delay distribution) renewal process with interarrival distribution G and finite interarrival mean μ . Let further $A(t)$ be the age of the last renewal. Then the elementary renewal theorem states that

$$\frac{N(t)}{t} \rightarrow \frac{1}{\mu}, \quad \text{as } t \rightarrow \infty.$$

If G is assumed to be non-arithmetic we also have Blackwell's renewal theorem

$$N(t) - N(t - a) \rightarrow \frac{a}{\mu}, \quad \text{as } t \rightarrow \infty$$

and a theorem about the asymptotic distribution of $A(t)$:

$$P(A(t) \leq x) \rightarrow \frac{1}{\mu} \int_0^x (1 - G(y)) dy, \quad \text{as } t \rightarrow \infty. \quad (3.2)$$

These three results are used in the proof of the main result in this paper, Theorem 4.1. For a quick illustration how they may apply, consider the process starting from an ancestor of type 1. Then $M_{1,1}^{(a)}(t)$ is equal to $P(A(t) \leq a)$, $A(t)$ as above and (3.2) applies immediately.

If we instead consider $M_{1,0}^{(a)}(t)$, note that the 0-type cells are immortal and hence at time t the number of 0-types younger than a will equal the number of renewals in the 1-type process between $t - a$ and t . Hence, in the above notation, $M_{1,0}^{(a)}(t) = N(t) - N(t - a)$ the asymptotics of which follow from Blackwell's renewal theorem.

4 Main Results

When investigating the asymptotics of $M_{k,j}^{(a)}(t)$, it turns out that the immortality of the 0-types forces us to distinguish between the cases $j > 0$ and $j = 0$. Adopting the standard convention of writing $a_t \sim b_t$ whenever $\lim a_t/b_t = 1$, the main theorem is

Theorem 4.1 Let $M_{k,j}^{(a)}(t)$ be the expected number of j -types of age at most a at time t in a process started from one ancestor of type k . Assume that the life-length distributions G_j are non-arithmetic with means $\mu_j < \infty$. Then, for $k > j > 0$,

$$M_{k,j}^{(a)}(t) \sim \frac{C_{k,j}}{\mu_j} \cdot \int_0^a (1 - G_j(y)) dy \cdot \frac{t^{(k-j)}}{(k-j)!}$$

as $t \rightarrow \infty$ and for $k > 0$

$$M_{k,0}^{(a)}(t) \sim C_{k,0} \cdot a \cdot \frac{t^{k-1}}{(k-1)!}$$

as $t \rightarrow \infty$. The constant $C_{k,j}$ is

$$C_{k,j} = \prod_{i=j+1}^k \frac{p_{i,i-1}}{\mu_i}.$$

Before turning to the proof, we state two further results that are needed. The key ingredient is the following Tauberian theorem which can be found in Feller [4].

Theorem 4.2 Let H be a non-decreasing function such that the Laplace-Stieltjes transform $\widehat{H}(s) = \int_0^\infty e^{-su} H(du) < \infty$ for $s > 0$. Then

$$\widehat{H}(s) \sim \frac{A}{s^r} \quad \text{as } s \rightarrow 0$$

if and only if

$$H(t) \sim \frac{At^r}{\Gamma(r+1)} \quad \text{as } t \rightarrow \infty.$$

where A is a constant.

The following simple lemma on Laplace-Stieltjes transforms is also used in the proof.

Lemma 4.3 *Let G be a distribution function of a non-negative, non-degenerate random variable with finite mean, $\mu = \int_0^\infty tG(dt) < \infty$ and Laplace-Stieltjes transform $\widehat{G}(s) = \int_0^\infty e^{-st}G(dt)$. Then*

$$\frac{\widehat{G}(s)}{1 - \widehat{G}(s)} \sim \frac{1}{s\mu}$$

as $s \rightarrow 0+$.

Proof. Consider a renewal process with inter-arrival times distributed according to G and delay distribution G . Let $N(t)$ be the number of renewals up to time t . Then, by the elementary renewal theorem, $N(t) \sim t/\mu$ as $t \rightarrow \infty$. By Theorem 4.2, $\widehat{N}(s) \sim 1/s\mu$ as $s \rightarrow 0+$. But, since

$$N(t) = \sum_{n=1}^{\infty} G^{*n}(t),$$

and transforms turn convolutions into products, we also get

$$\begin{aligned} \widehat{N}(s) &= \int_0^\infty e^{-st} N(dt) = \sum_{n=1}^{\infty} \int_0^\infty e^{-st} G^{*n}(dt) \\ &= \sum_{n=1}^{\infty} \widehat{G}(s)^n = \frac{\widehat{G}(s)}{1 - \widehat{G}(s)} \end{aligned}$$

which proves the lemma. ■

Proof of Theorem 4.1. For any k, j , by Equation (3.1),

$$M_{k,j}^{(a)}(t) = \sum_{i=j}^k \int_0^t M_{i,j}^{(a)}(t-u) p_{k,i} G_k(du)$$

which has Laplace-Stieltjes transform

$$\widehat{M}_{k,j}^{(a)}(s) = \sum_{i=j}^k p_{k,i} \widehat{M}_{i,j}^{(a)}(s) \widehat{G}_k(s).$$

Now, $p_{k,k} = 1$, and hence

$$\widehat{M}_{k,j}^{(a)}(s) = \widehat{M}_{k,j}^{(a)}(s) \widehat{G}_k(s) + \sum_{i=j}^{k-1} p_{k,i} \widehat{M}_{i,j}^{(a)}(s) \widehat{G}_k(s)$$

i.e.

$$\widehat{M}_{k,j}^{(a)}(s) = \frac{\widehat{G}_k(s)}{1 - \widehat{G}_k(s)} \sum_{i=j}^{k-1} p_{k,i} \widehat{M}_{i,j}^{(a)}(s). \quad (4.1)$$

First consider the case $j > 0$. By Theorem 4.2 we are done if we can prove that

$$\widehat{M}_{k,j}^{(a)}(s) \sim \frac{1}{s^{k-j}} \frac{C_{k,j}}{\mu_j} \int_0^a (1 - G_j(y)) dy \quad (4.2)$$

as $s \rightarrow 0+$. We do this by induction on k . First let $k = j$ and note that

$$M_{j,j}^{(a)}(t) = P(A(t) \leq a)$$

where $A(t)$ is the age of the last renewal before time t . By (3.2)

$$M_{j,j}^{(a)}(t) = P(A(t) \leq a) \rightarrow \frac{1}{\mu_j} \int_0^a (1 - G_j(y)) dy$$

as $t \rightarrow \infty$. Hence by Theorem 4.2,

$$\widehat{M}_{j,j}^{(a)}(s) \rightarrow \frac{1}{\mu_j} \int_0^a (1 - G_j(y)) dy \quad (4.3)$$

as $s \rightarrow 0+$. Next let $k = j + 1$. By Equation (4.1) we have

$$\widehat{M}_{j+1,j}^{(a)}(s) = \frac{\widehat{G}_{j+1}(s)}{1 - \widehat{G}_{j+1}(s)} \cdot p_{j+1,j} \cdot \widehat{M}_{j,j}^{(a)}(s)$$

so that, by Lemma (4.3) and Equation (4.3), we obtain

$$\widehat{M}_{j+1,j}^{(a)}(s) \sim \frac{1}{s\mu_{j+1}} \cdot \frac{1}{\mu_j} \cdot p_{j+1,j} \cdot \int_0^a (1 - G_j(y)) dy.$$

Hence (4.2) is true for $k = j + 1$. Now consider an arbitrary $k > j + 1$. Using the relation (4.1), where we note that the $(k - 1)$ -term dominates, and the induction hypothesis we obtain

$$\widehat{M}_{k,j}^{(a)}(s) \sim \frac{1}{s\mu_k} \cdot p_{k,k-1} \cdot \frac{1}{s^{k-j-1}} \cdot \frac{1}{\mu_j} \cdot C_{k-1,j} \int_0^a (1 - G_j(y)) dy$$

$$= \frac{1}{s^{k-j}} \cdot \frac{1}{\mu_j} C_{k,j} \int_0^a (1 - G_j(y)) dy$$

which by (4.2) proves the first assertion of the theorem.

Next, consider the case $j = 0$. This time the theorem follows if we can prove that

$$\widehat{M}_{k,0}^{(a)}(s) \sim \frac{C_{k,0} \cdot a}{s^{k-1}}$$

as $s \rightarrow 0+$. Again use induction on k . First let $k = 1$. Since 0-types are immortal, $M_{1,0}^{(a)}(t)$ will equal the number of renewals between $t - a$ and t in a renewal process with interarrival distribution G_1 . With $N(u) =$ the number of renewals up to time u , Blackwell's renewal theorem yields

$$M_{1,0}^{(a)}(t) = N(t) - N(t - a) \rightarrow \frac{a}{\mu_1}$$

as $t \rightarrow \infty$. By Theorem 4.2, this is also the limit of $\widehat{M}_{1,0}^{(a)}(s)$ as $s \rightarrow 0+$ and since $p_{1,0} = 1$, the assertion is true for $k = 1$. Using the relation (4.1), Lemma (4.3) and induction gives

$$\widehat{M}_{k,0}^{(a)}(s) \sim \frac{1}{s\mu_k} \cdot p_{k,k-1} \cdot \frac{1}{s^{k-2}} \cdot a \cdot C_{k-1,0} = \frac{1}{s^{k-1}} \cdot a \cdot C_{k,0}$$

as $s \rightarrow 0+$ and the proof is complete. ■

Note. If some of the $p_{j,i}$'s are zero, the asymptotics are still polynomial as long as all the $p_{j,j}$ are equal to one. The polynomial powers and the constants in the limit then depend on which $p_{j,i}$'s are zero. If all the $p_{j,j-1}$'s are non-zero, the asymptotics are the same as in the theorem. If a number, r say, of the $p_{j,j-1}$'s are zero but all other $p_{j,i}$'s are non-zero, the polynomial growth rate starting from a k -type will be t^{k-r} and the constant will be the product over those j that have $p_{j,j-1} > 0$ and the $p_{j,j-2}$ for the other j 's. If further $p_{j,i}$'s are zero, the power in the polynomial asymptotics decreases in a similar manner.

Note. The assumption of non-arithmetic G_j 's may be relaxed, keeping in mind that Blackwell's renewal theorem and (3.2) still hold whenever a is a

multiple of the span δ in the supporting lattice $\{n\delta, n \in Z_+\}$.

Note. If the life lengths are continuous, we have that the density of a -aged cells, $m_{k,j}^{(a)}(t)$ say, asymptotically is

$$m_{k,j}^{(a)}(t) \sim \frac{C_{k,j}}{\mu_j} \cdot (1 - G_j(a)) \cdot \frac{t^{(k-j)}}{(k-j)!}$$

for $k > j > 0$ and

$$m_{k,0}^{(a)}(t) \sim C_{k,0} \cdot \frac{t^{k-1}}{(k-1)!}.$$

By choosing $a = 0$ we obtain Theorem 1 in Arino, Sanchez and Webb [2] for the densities of newborn cells (starting from one newborn ancestor cell).

If we are only interested in the number of cells alive at time t , which may indeed be reasonable from the point of view of applications, we need not assume that the G_j be non-arithmetic:

Theorem 4.4 *Let $M_{k,j}(t)$ be the expected number of j -type cells at time t , starting from an ancestor cell of type k . Suppose that $\mu_j < \infty$ for all j . Then*

$$M_{k,j}(t) \sim \frac{C_{k,j}}{\mu_j} \cdot \frac{t^{k-j}}{(k-j)!}$$

as $t \rightarrow \infty$.

Proof. The proof is essentially the same as that of Theorem 4.1. Remove the superscripts in the basic relation (3.1), note that we do not have to treat 0-cells separately and thus obtain

$$M_{j,j-1} \sim p_{j,j-1} \frac{t}{\mu_j}$$

for all $j \geq 1$. Keeping in mind that $p_{1,0} = 1$, the theorem now follows by induction. ■

Note. The theorem is trivially true also in the case $\mu_j = \infty$ in the sense that $t^{-(k-j)}M_{k,j}(t) \rightarrow 0$.

5 Concluding Remarks

Our main theoretical result is Theorem (4.1) which gives the precise polynomial asymptotics of our reducible multi-type Bellman-Harris process. This could be regarded as a minor contribution to the theory of branching processes; it is not a very deep result but still interesting as an example of non-exponential growth in a process motivated by a biological application. It is also a good example of how branching process models are very well suited for problems in cell biology with their conceptual clarity and powerful limit results.

It could also be noted that Theorem (4.4) could equally well have been stated as a result on iterated renewal processes: at each renewal point, a new renewal process is started. If all these renewal processes have the same inter-arrival distribution, G , then we would obtain the asymptotics

$$N_k(t) \sim \frac{t^k}{\mu^k k!}, \quad t \rightarrow \infty,$$

for the k th generation renewals.

More important is that this is a model of a much-studied biological phenomenon, the loss of telomere sequences. Such models have been compared with data in Arino, Kimmel and Webb [1] and Levy *et al.* [5]. Both of these references use the assumption that only one telomere unit is lost in each division. It is clearly reasonable from a biological point of view that this number may be variable and we therefore allow for a random number of telomere units to be lost. Also, we allow the life lengths to depend on the current telomeric state. It is not clear if there are any biological motivations for this; it may or may not be the case that telomeres are also involved in the regulation of the division rate of a cell.

There is one problem with this model that we discuss briefly. The reader should note that, although this has not been pointed out explicitly, we identify the type of the cell with the telomeric state of one single chromosome. This is reasonable if there is one particular chromosome that is responsible for stopping cell division (or if there only is one chromosome such as is the case for *Tetrahymena thermophila*; see Larson *et al.* [5]). If however cell division is stopped whenever any of several chromosomes reaches the critical telomere length, the situation is different. We then consider a minimum of several processes that are not independent, and the dynamics are the same.

For a simple illustration of this, assume a cell has two chromosomes and that the type of the cell is (j, k) where these are the remaining numbers of telomere units on the respective chromosomes. As soon as one of the chromosomes hit 0, the cell stops dividing. Now suppose the cell is of type $(1, 1)$. Then the two type 1 chromosomes each produce two daughter chromosomes, one of type 0 and one of type 1. If chromosomes are randomly allocated to daughter cells, the two daughter cells may thus either be of types $(0, 0)$ and $(1, 1)$ or of types $(0, 1)$ and $(1, 0)$ with equal probabilities. In the second case, both cells will stop dividing, in the first case the $(1, 1)$ -type will go on dividing. Clearly there will only be a finite number of divisions in this model (in fact the total number of cells will be a geometric random variable) and the population no longer exhibits polynomial growth. These issues are not further addressed here but should be of interest for future investigations.

6 References

- [1] Arino, O., Kimmel, M. and Webb, G.F. Mathematical modeling of the loss of telomere sequences. *J. Theor. Biol.* **177** (1995), 45-57.
- [2] Arino, O., Sanchez, E. and Webb, G.F. Polynomial growth dynamics of telomere loss in a heterogeneous cell population. *Dynam. Contin. Discrete Impuls. Systems* **3** (1997), 262-282.
- [3] Asmussen, S. *Applied Probability and Queues*. Wiley (1987), New York.
- [4] Feller, W. *An Introduction to Probability Theory and its Applications, Vol 2*. Wiley (1971), New York.
- [5] Larson, D.D., Spangler, E.A. and Blackburn, E.H. Dynamics of telomere length variation in tetrahymena thermophila. *Cell* **50** (1987), 477-483.
- [6] Levy, M.Z., Allsopp, R.C., Futcher A.B., Greider C.W. and Harley C.B. Telomere end-replication problem and cell Aging. *J. Molec. Biol.* **225** (1992), 951-960
- [7] Mode, C. *Multitype Branching Processes*. American Elsevier (1971), New

York

[8] Olofsson, P. and Kimmel M. Stochastic models of telomere shortening.
Math. Biosci. **1** (1999), 75-92