



Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbiModeling growth and telomere dynamics in *Saccharomyces cerevisiae*Peter Olofsson^{a,*}, Alison A. Bertuch^b^a Trinity University, Mathematics Department, One Trinity Place, San Antonio, TX 78212, USA^b Baylor College of Medicine, Department of Pediatrics, One Baylor Plaza, Houston, TX 77030, USA

ARTICLE INFO

Article history:

Received 15 September 2009

Received in revised form

12 November 2009

Accepted 2 December 2009

MSC:

60G99

60K99

62P10

92D25

Keywords:

Telomere

Branching process

Yeast

Fibonacci sequence

ABSTRACT

A general branching process is proposed to model a population of cells of the yeast *Saccharomyces cerevisiae* following loss of telomerase. Previously published experimental data indicate that a population of telomerase-deficient cells regain exponential growth after a period of slowing due to critical telomere shortening. The explanation for this phenomenon is that some cells engage telomerase-independent pathways to maintain telomeres that allow them to become “survivors.” Our model takes into account random variation in individual cell cycle times, telomere length, finite replicative lifespan of mother cells, and survivorship. We identify and estimate crucial parameters such as the probability of an individual cell becoming a survivor, and compare our model predictions to experimental data.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Telomeres are specialized structures found at the ends of linear chromosomes, which consist of protein bound, tandem repeats of short DNA sequences called telomeric repeats (Denchi, 2009). Telomeres function to prevent the factors that detect and repair chromosome breaks from acting upon natural chromosome ends. When telomere function is lost, either through the loss of telomeric repeats or certain proteins that associate with them, telomeres become substrates of the DNA double strand break repair machinery, which results in telomere–telomere fusions and genomic instability.

During semi-conservative DNA replication of linear chromosomes, there is loss of terminal telomeric DNA due to the so-called end replication problem (Baird, 2008). To offset this loss, certain cell types, such as germ cells, express telomerase, which catalyzes the addition of telomeric repeats onto chromosome termini. Most human somatic cells, however, express insufficient or undetectable levels of telomerase and as a consequence, telomeres shorten in these cells with each cell cycle. Eventually, one or more telomeres reach a critical length that triggers an irreversible halt in cellular proliferation known as cellular senescence. Rarely, cells

bypass the induction of senescence, but go on to die due to the eventual loss of telomeric function and secondary genomic instability. A rare population of cells, however, may bypass this crisis by reactivating telomerase or engaging a recombination-based mechanism of telomere maintenance, allowing for continued proliferation. These processes of telomere dysfunction followed by telomere stabilization are thought to represent important steps in the development of cancer, and thus are of substantial interest.

The budding yeast, *Saccharomyces cerevisiae*, is a genetically tractable model organism for the study of telomere biology. Telomerase-deficient strains can be readily generated and telomere loss rates of 3–5 base pairs have been determined in the absence of telomerase (Lundblad and Szostak, 1989; Singer and Gottschling, 1994). As in higher eukaryotes, telomerase deficiency in yeast results in eventual cellular senescence, and rare cells that evade this fate can be isolated (Lundblad and Blackburn, 1993). Two major classes of these telomerase-independent/post-senescence “survivors” have been characterized, which can be distinguished in part by their telomere structures and growth properties. Chromosome ends in type I survivors have extensively amplified subtelomeric Y' elements and terminate with short tracts of telomeric repeats, whereas the chromosome ends in type II survivors exhibit a much lesser degree of Y' amplification and terminate with heterogeneous and long telomeric repeat tracts (Lundblad and Blackburn, 1993; Teng

* Corresponding author.

E-mail address: polofss@trinity.edu (P. Olofsson).

and Zakian, 1999). Type II telomeres are akin to those that are observed in cancer cells that do not activate telomerase, but rather use an alternative lengthening of telomeres (ALT) pathway for telomere maintenance (Bryan et al., 1997). Although type I survivors arise more frequently than type II survivors, they exhibit severely impaired growth (Teng and Zakian, 1999). Type II survivors, in contrast, grow comparable to wild-type cells, and consequently predominate with serial propagation of telomerase-deficient strains in liquid culture. Although many of the proteins that are required for or influence one or both of the pathways have been identified, the molecular events that drive the telomere recombination choice and the rate they occur are poorly defined.

Previous mathematical models of the process of telomere loss include Levy et al. (1992), Arino et al. (1995, 1997), Olofsson and Kimmel (1999, 2005), Rubelj and Vondracek (1999), Tan (1999), Olofsson (2000), Op Den Buijs et al. (2004), Dyson et al. (2007), and Portugal et al. (2008). In the present paper we develop a model of cellular proliferation in response to telomere dynamics in yeast that takes into account the facts that mothers and daughters are distinguishable individuals, that the proliferative potential of a cell is limited by factors other than telomere length, and, in particular, that population growth can be restored after a period of slowing due to the bypass mechanisms mentioned above.

2. The branching process model

We propose a general branching process (Haccou et al., 2005; Jagers and Nerman, 1984) where individuals reproduce by budding. The times between consecutive budding events (i.e., the cell cycle times) are assumed to be independent random variables with the same distribution. This is where the general branching process comes into play: a mother cell produces daughter cells at different times during her life which is relevant to budding yeast and different from, for example, the binary fission of bacterial reproduction that can be modeled by simpler branching processes. As we are also keeping track of loss of telomeric DNA, we are in fact using a multitype process (Jagers, 1992) where type accounts for telomere length, however, specific results for multitype processes will not be needed. An individual cell contains 16 chromosomes so it has 32 telomeres. In cells that express telomerase, the number of telomeric repeats present varies to a certain extent from end to end resulting in a distribution of telomere lengths around a strain-specific point (Shampay and Blackburn, 1988). In the absence of telomerase, the rate of loss of telomeric DNA follows a probability distribution over a range from 3 to 5 base pairs per end per cell division (Lundblad and Szostak, 1989; Singer and Gottschling, 1994). We therefore define a “telomere unit” to be 4 base pairs and, for the time being, assume that one telomere unit is lost per division. We let the type of a cell be the number of remaining telomere units.

For increased clarity, we will first describe a simplified model and successively extend it to incorporate all the features we need, leading up to the description in Proposition 2.2. The population starts from a single telomerase-deficient cell of type n and upon completion of the cell cycle, this cell has produced one daughter cell. As all telomeres have shortened by one unit in the preceding round of DNA replication and are randomly allocated to mother and daughter cell, it is reasonable to assume that both mother and daughter, after division, have type $n-1$. For simplicity, we refer to the two cells as the first generation. When a cell reaches type 0 (the critical telomere length mentioned in Section 1), it stops dividing and becomes senescent, unless one of the alternative mechanisms for telomere maintenance described in Section 1 is established. We will address this situation later. Senescent cells

remain in the population but do not further reproduce. As the ultimate fate of a senescent cell is likely to be death, we could let each senescent cell remain in the population for a time following some probability distribution. However, as the mean in this distribution is likely far to exceed the duration of the experiment, senescent cells can for practical purposes be considered simply nondividing. Again, it is not difficult to adjust the model but any possible gain would hardly outweigh the increased complexity of the model.

We assume that cell cycle times are independent random variables with the common cumulative distribution function (cdf) F . The main quantity of interest is the number of cells in the population at time t , which we denote by Z_t . The population starts from a single cell of type n at time $t = 0$. In this basic model, the expected value of Z_t is

$$E[Z_t] = 1 - F(t) + \sum_{k=1}^{n-1} 2^k (F^{*k}(t) - F^{*(k+1)}(t)) + 2^n F^{*n}(t)$$

where F^{*k} denotes k -fold convolution of F , that is, the cdf of the sum of k cell cycle times. The factor 2^k is the expected number of cells in the k th generation (recall our use of “generation” mentioned above). Now note that each cell in the k th generation is present in the population if the sum of k cell cycle times is less than t but the sum of $k+1$ cell cycle times is greater than t . As the probability of this event is $F^{*k}(t) - F^{*(k+1)}(t)$, the expected number of cells from the k th generation that are present at t equals

$$2^k (F^{*k}(t) - F^{*(k+1)}(t))$$

and, noting that cells of type 0 do not further reproduce, summing over k gives the expression for $E[Z_t]$. Note that $E[Z_t] \rightarrow 2^n$ as $t \rightarrow \infty$ so 2^n is the final number of cells.

Additional factors influence the replicative capacity of budding yeast cells, such that a given cell will undergo a finite number of cell divisions even when telomerase is expressed and telomere length is maintained (D’Mello and Jazwinski, 1991). We denote this maximum number of divisions by n_0 , refer to it as the replicative lifespan, and assume that it is constant. The expression for $E[Z_t]$ is now

$$E[Z_t] = 1 - F(t) + \sum_{k=1}^{n-1} m(k) (F^{*k}(t) - F^{*(k+1)}(t)) + m(n) F^{*n}(t)$$

where $m(k)$ is the expected number of cells in the k th generation. As long as $k \leq n_0$, we still have $m(k) = 2^k$ and for $k > n_0$ we describe a recursive scheme that enables us to compute $m(k)$. To that end, in any given generation, let k_j be the number of cells that are able to reproduce j times for $j = 0, 1, \dots, n_0$. Each cell with $j > 1$ produces a daughter cell that is able to reproduce n_0 times and is then itself able to reproduce another $j-1$ times. Cells with $j = 0$ remain unchanged. For $n_0 \leq k \leq n$, the transition from generation $k-1$ to generation k is therefore as follows:

$$\begin{cases} \text{Generation } k-1: & (k_0, k_1, \dots, k_{n_0-1}, k_{n_0}) \\ \text{Generation } k: & \left(k_0 + k_1, k_2, \dots, k_{n_0}, \sum_{i=1}^{n_0} k_i \right) \end{cases}$$

and after relabeling in generation k to $(k_0, k_1, \dots, k_{n_0})$ again, we have

$$m(k) = \sum_{j=0}^{n_0} k_j$$

The initial configuration in generation 0 is $(0, 0, \dots, 0, 1)$ since there is one cell that is able to divide n_0 times. For $k > n$, all cells are senescent so $m(k)$ stays constant and we have $m(k) = m(n)$ for $k \geq n$. It is interesting to note that the sequence $\{k_1, \dots, k_{n_0}\}$ (excluding k_0) constitutes an “ n_0 -nacci” sequence, that is, a

generalized Fibonacci sequence where each new number is obtained by adding the previous n_0 numbers, starting from $(0, 0, \dots, 0, 1)$. Hence, it is possible to obtain explicit expressions for the k_j (Flores, 1967). We state this observation as a proposition that is interesting in its own right; for our computational purposes, however, the recursive formula above is sufficient.

Proposition 2.1. Consider the vector $(k_0, k_1, \dots, k_{n_0})$ in the k th generation of the branching process above for $k \leq n$. Let F_j denote the j th n_0 -nacci number and set $F_0 = F_1 = \dots = F_{n_0-1} = 0$, $F_{n_0} = 1$. Then $(k_0, k_1, \dots, k_{n_0})$ equals

$$\left(\sum_{j=0}^k F_j, F_{k+1}, \dots, F_{k+n_0} \right)$$

To account for the observed phenomenon of restored exponential growth, we assume that cells that have reached type 0 have the possibility to turn into “survivors” or become senescent. The expression for $m(k)$ remains the same for $k \leq n$, but for $k > n$ it changes since cells of type 0 may now escape senescence and keep reproducing. We assume that a cell of type 0 becomes a survivor with probability p and that the survivor status is inherited by all of its daughter cells. Thus, each survivor starts a population of survivors where telomere length is generally maintained and we assume the only limiting factor is the replicative lifespan n_0 . The nonsurvivors turn senescent.

To arrive at an expression for the expected number of cells in generation k for $k > n$, we first need to figure out how many cells in generation n that are eligible to become survivors, and for each of those, how many more times it is able to reproduce. There is a total of $m(n)$ cells, but many of those cells have ceased reproducing after reaching the threshold imposed by the maximum number of offspring n_0 . For now, we assume that the only cells in generation n that are eligible to become survivors are those that have reached type 0 and are still able to reproduce at least once, an assumption that will later be relaxed. To account for those cells, we use the recursive scheme above and in generation n arrive at the vector $(k_0, k_1, \dots, k_{n_0})$. We then have $k_1 + \dots + k_{n_0}$ cells that are eligible to become survivors. The cells that do not become survivors become senescent, joining the k_0 cells that are no longer able to reproduce. Thus, in generation n there is an expected number of $k_0 + (1-p)(k_1 + \dots + k_{n_0})$ senescent cells and an expected number of $p(k_1 + \dots + k_{n_0})$ survivors. Since the survivors differ in the number of possible daughter cells, they do not contribute equally to the population and we need to deal with each k_j separately.

Let us start with the k_1 cells that may reproduce once more. Each such cell produces one daughter cell that starts a population now capable of maintaining telomere length, limited only by the reproductive lifespan n_0 . The mother cell turns senescent. Thus, in generation $n+j$, the expected size of the population stemming from the original survivor is $1+m(j-1)$. The k_2 cells that may reproduce twice more contribute an expected number of $1+m(0)=2$ cells in generation $n+1$ and an expected number of $1+m(1)+m(0)$ cells in generation $n+2$. After that, the mother turns senescent and, in generation $n+j$, the expected size stemming from an initial survivor of this kind is $1+m(j-1)+m(j-2)$. Generally, a survivor that can reproduce i more times contributes an expected number of $m_{i,j}$ cells to generation $n+j$ where

$$m_{i,j} = 1 + \sum_{l=1}^{\min(i,j)} m(j-l)$$

Further, let $s(n)$ be the number of cells in generation n that are eligible to become survivors ($s(n) = k_1 + \dots + k_{n_0}$ in the notation

above) and let $d(n)$ be the number of cells in generation n that are already senescent from before ($d(n) = k_0$ in the notation above). We are now ready for the final expression.

Proposition 2.2. Let $M(k)$ denote the expected number of cells in generation k in a branching process where

- cell cycle times are independent with cdf F ;
- the initial telomere length is n ;
- the replicative lifespan is n_0 ;
- the probability of becoming a survivor is p .

Then

$$M(k) = \begin{cases} 2^k & \text{for } k \leq n_0 \\ m(k) & \text{for } n_0 < k \leq n \\ d(n) + (1-p)s(n) + p \sum_{i=1}^{n_0} m_{i,k-n} & \text{for } k > n \end{cases}$$

and the expected number of cells in the population at time t is

$$E[Z_t] = 1 - F(t) + \sum_{k=1}^{\infty} M(k)(F^{*k}(t) - F^{*(k+1)}(t))$$

Note that all senescent cells are included in the expression for $M(k)$ even though they have ceased to reproduce. However, the expression for $E[Z_t]$ is still valid by the telescoping properties of the terms in the sum; once a cell is senescent, it stays in the population forever which is mathematically equivalent to saying that it “replaces itself” at time intervals with cdf F .

3. Data

To fit and calibrate our model, we use data from Bertuch and Lundblad (2004). In this study, telomerase-deficient haploid cells (containing one set (1n) of chromosomes) were generated from telomerase-proficient diploid (2n) cells, in which one of the two copies of the gene that encodes the catalytic subunit of telomerase (*EST2*) was replaced by a selectable marker (*est2Δ*). Upon meiosis of a parental diploid cell, four daughters would be formed—two that would inherit the wild-type *EST2* copy and two that would inherit the deleted *est2Δ* copy, which would be telomerase-proficient and telomerase-deficient, respectively. The parental diploid strain also contained one wild-type and one deleted copy of a gene that encodes an enzyme that degrades dysfunctional telomeres, *EXO1*. These similarly segregated 2:2 to the offspring haploids. The individual cells were separated from one another on solid media and allowed to grow into colonies for 4 days. After 4 days, the individual colonies were resuspended in their entirety in 5 milliliters (ml) media and allowed to grow for 1 day. Cells counts were performed and the cultures subsequently diluted into fresh media at a concentration of 10^5 cells/ml. The cultures were grown for 22 h after which time cell counts were obtained. The dilution, growth, and cell counting protocol was repeated every 22 h for a total of 9 days. The genotypes of the cultures (*EST2* vs. *est2Δ* and *EXO1* vs. *exo1Δ*) were ultimately determined.

We convert the growth rate data to logarithmic population size data. For example, if the initial count by day 1 is 10^8 cells per ml, the logarithmic population size is $\log(5 \times 10^8) \approx 20.0$. Next, suppose the count by day 2 is again 10^8 . This means that the 10^5 cells (per ml) that were harvested at the end of day 1 grew to 10^8 cells (per ml) by day 2, that is, a 1000-fold increase. We then

estimate the logarithmic population size to be $20.0 + \log(1000) \approx 26.9$, and so on and so forth.

4. Estimation of model parameters

In order to estimate cell cycle parameters, we make the assumption that cell cycle times follow a gamma distribution. This distribution is a flexible two-parameter family that is commonly used to model lifetimes (Oprea and Kepler, 2001; Larsson et al., 2008). Specifically, if the parameters are a and b , the probability density function is

$$f_{a,b}(t) = e^{-bt} b^a \frac{t^{a-1}}{\Gamma(a)}$$

where $\Gamma(a)$ is the gamma function. The model parameters are thus: initial telomere length n , replicative lifespan n_0 , survivor probability p , and the parameters a and b in the gamma distribution. Our first goal is to estimate a and b , which will also give an estimate of the growth rate. For this purpose, we use data for wild-type, telomere-proficient strains, also obtained by Bertuch and Lundblad. Since the data indicate that this population is in steady exponential growth, we use asymptotic results for exponentially growing branching processes to estimate the mean and variance of the population (Haccou et al., 2005; Jagers and Nerman, 1984). At this point, we propose a simplification that disregards the effects of n_0 based on the following argument.

The proliferative lifespan, n_0 , has been estimated to be on average about 25 (Sinclair and Guarente, 1997; Sinclair et al., 1998) and at the beginning of each 22-h time period, cells are distributed according to Proposition 2.1. By day 5, when the first cell counts are performed, there have been about 30 population doublings and by day 8, there have been about 100 doublings; hence, we have k in the range 30–100 (n can be considered infinite as telomere length is maintained in the wild-type population). With n_0 and k in these ranges, Proposition 2.1 shows that the fraction of cells that are not able to keep producing daughter cells for the full 22 h is very small. For example, more than 99% of cells are able to produce at least 18 daughter cells which with overwhelming probability takes far more than 22 h. In fact, almost 50% of cells are newborn and another 25% are second generation. In conclusion, almost all cells are able to produce daughter cells for the full 22-h period, which holds true also for other values of n_0 , varying within realistic limits (Sinclair and Guarente, 1997; Sinclair et al., 1998).

Hence, we can view the population as a binary splitting process, also known as a Bellman–Harris process, where lifetimes are $\Gamma(a, b)$ and each cell produces 2 daughter cells (one of which is actually the mother cell). In such a Bellman–Harris process, the population grows asymptotically as $e^{\alpha t}$ where the growth rate α is known as the *Malthusian parameter* and satisfies the relation

$$2\hat{F}(\alpha) = 1$$

where F is the cdf of the lifetime (cell cycle time) and \hat{F} denotes Laplace transform:

$$\hat{F}(\alpha) = \int_0^\infty e^{-\alpha t} F(dt)$$

For the $\Gamma(a, b)$ distribution we get

$$2\hat{F}(\alpha) = 2 \int_0^\infty e^{-\alpha t} f_{a,b}(t) dt = \frac{2b^a}{(\alpha + b)^a}$$

and setting this expression equal to 1 gives an explicit solution for the Malthusian parameter as

$$\alpha = b(2^{1/a} - 1)$$

Since we have disregarded n_0 we get a slight overestimate of α . This overestimate is negligible for practical purposes; for a more thorough account see the Appendix. Another source of overestimation of α is the fact that the cells in our data set of daily counts are not newborn, but have lived for a while at the time of harvesting. It is possible to use stable population theory (Jagers and Nerman, 1984) to establish that the age of a harvested cell has probability density function given by

$$f(t) = ce^{-\alpha t}(1 - F(t))$$

where F is the cdf of the cell cycle time and c is a normalizing constant. However, such detailed modeling is not likely to yield much improvement considering all the other sources of error and uncertainty in the data sets. It is enough for our purposes to get a rough estimate of α , knowing that it is likely to be a slight overestimate which also has bearing on the estimates of a and b .

The asymptotic mean and variance of the population size Z_t at time t are

$$E[Z_t] \sim e^{\alpha t} \frac{1}{2\alpha\beta} \tag{4.1}$$

and

$$\text{Var}[Z_t] \sim e^{2\alpha t} \frac{4\hat{F}(2\alpha) - 1}{4\alpha^2\beta^2(1 - 2\hat{F}(2\alpha))} \tag{4.2}$$

where

$$\beta = 2 \int_0^\infty te^{-\alpha t} f_{a,b}(t) dt = \frac{2ab^a}{(\alpha + b)^{a+1}}$$

For details see Haccou et al. (2005) or the classic Harris (1963). From the data for the wild-type strain in Bertuch and Lundblad (2004), we get the following 18 daily estimated cell population counts per ancestor cell:

2250, 2300, 3000, 2100, 2530, 1530, 2450, 3650, 2550, 2570, 2570
2700, 3900, 2600, 3950, 1880, 2920, 2600

As each culture is grown for 22 h, these are observed values of Z_t where $t = 22$ h. The sample mean is $\bar{Z} = 2136$, and the sample variance is $s_z^2 = 261952$. Using the approximations provided in (4.1) and (4.2), we solve the equations $E[Z_{22}] = 2669$ and $\text{Var}[Z_{22}] = 409300$ to get the moment estimates $a = 17$ and $b = 9$, rounded to nearest integers. The mean cell cycle time is estimated to $a/b = 1.9 = 22$ h with an approximate 95% confidence interval 1.9 ± 0.2 and the estimate of the Malthusian parameter is $\alpha = 0.37$.

As a side note, let us point out that the Bellman–Harris approximation is bad for small values of n_0 . For example, if $n_0 = 2$, properties of the regular Fibonacci sequence shows that the distribution in Proposition 2.1 quickly converges to the proportions $(1, \phi, 1)$ where $\phi \approx 0.62$ is the golden section. Thus, in that case, as many as 38% of cells ($1/(2 + \phi) \approx 0.38$) have reached the end of their lifespan and are no longer able to produce daughter cells.

5. Curve fitting

We are now ready to compare our model to data. For that purpose, we use the data for the *est2Δ EXO1* strain where estimated cell counts exist for 9 consecutive days (22-h periods). The first such count is after 5 days (see the description of the experiment above) so we have data for days 5–13, and as each population starts from one single cell, we also have a data point for day 0. Table 1 and Fig. 1 display the seven data sets on a logarithmic scale. It is noticeable how the growth is slower in the beginning. Between days 0 and 5 there have been about 30

Table 1
Logarithmic cell counts of 7 yeast cultures, days 5–13.

Day	5	6	7	8	9	10	11	12	13
Culture 1	21.5	28.1	34.0	37.9	40.9	43.3	47.0	54.2	61.4
Culture 2	21.1	28.5	36.6	39.9	42.8	45.5	49.0	55.7	63.1
Culture 3	21.5	29.2	35.9	40.9	43.6	46.3	48.8	54.9	62.5
Culture 4	21.4	29.2	34.9	40.3	42.8	44.8	49.0	55.8	63.3
Culture 5	21.3	29.4	36.5	42.7	47.2	50.5	53.5	60.9	68.4
Culture 6	21.4	28.7	34.9	39.1	42.4	48.6	56.2	64.2	71.8
Culture 7	20.7	28.3	34.8	39.0	41.8	43.0	48.2	55.3	62.3

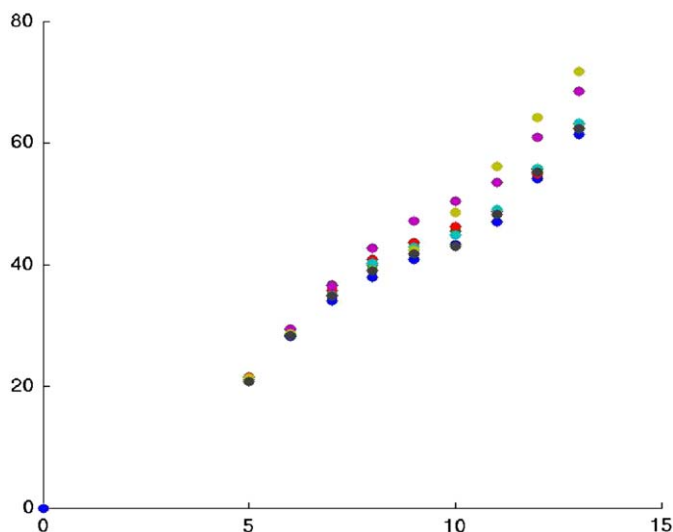


Fig. 1. Cell counts of 7 yeast cultures for 13 days (logarithmic scale).

population doublings ($21/\log 2 \approx 30$), whereas it only takes a little over 3 days for an additional 30 doublings. Possible reasons for this phenomenon are a lag in the time from when a single cell is plated onto rich media and resumes mitotic cell growth, and differences in growth on solid (the first 4 days) as compared to liquid media (the subsequent days). All of these factors are unknown so rather than trying to model the initial 5 days, we relabel day 5 as day 0 and rescale the data so that each data set has one cell on day 0 (by simply dividing each data set by the cell count on its day 0). These relabeled and rescaled data are displayed in Fig. 2 together with the mean (solid line). Note that the line is the logarithm of the average population counts which is why it is not “in the middle” of the dots (the logarithm of the average is always greater than the average of the logarithms, by Jensen’s inequality). Hence, we want our model to give a result such that $\log E[Z_t]$ resembles the solid line in Fig. 2.

Other than the already estimated n_0 , a , and b , our model also contains the parameters n and p . As for n , the initial telomere length at day 0, note that the population by then has passed through about 30 population doublings and 10^5 cells are harvested to start the population which means that rather than one value of n , we have a distribution over many different values. There is also random variation due to the fact that telomere loss might not be constant, and due to the asynchronous cell division. Fig. 2 indicates that telomere shortening starts significantly influencing population growth around days 3–4, when there has been about 30–35 population doublings so we let the average number of telomere units on day 0 be in that range.

Another potential source of variation comes from the way in which cells turn into survivors. It is currently unknown how

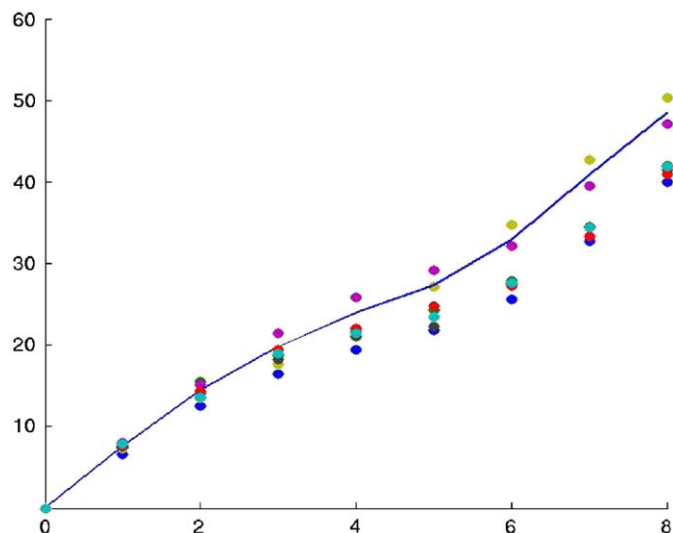


Fig. 2. Data from Fig. 1 rescaled, with mean.

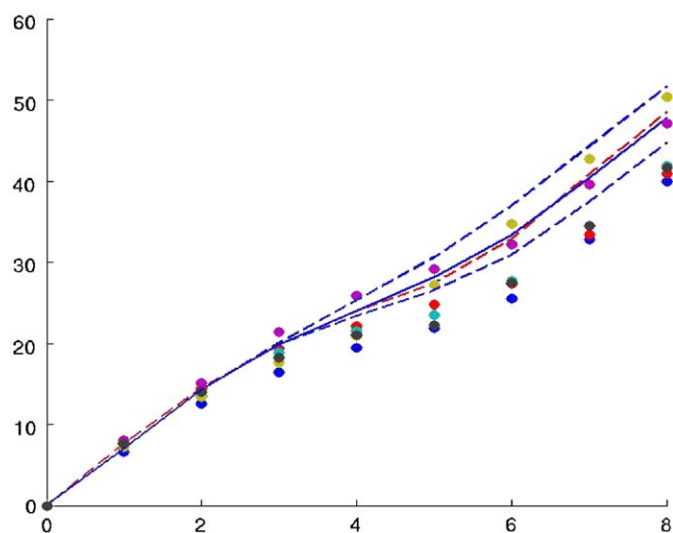


Fig. 3. Cell count data sets with mean (dashed red line) and model predictions for different values of p (blue lines) with uniform n .

survivorship is related to telomere length, but it is certainly possible that a cell may achieve survivor status not only when it is about to become senescent, but also while it still proliferating. Thus, in our model we can let cells become survivors not only when their type is 0, but also at type 1,2,3, etc. Rather than incorporating this directly into the model, we can view a population where cells may become survivors at type j as a population starting with type $n-j$, and average in some suitable way over such populations. This is possible as long as we deal with the expected value $E[Z_t]$; if we were also to compute the variance of Z_t , we lose linearity and with it, many of the suggested simplifications.

Thus, all of the sources of variation mentioned above can be incorporated by letting n have a probability distribution over some suitable range of values and compute the expected value $E[Z_t]$ when n is chosen accordingly. There is much uncertainty and many unknown factors; we shall assume that n has a probability distribution on some range of integers where the mean is in the range 30–35.

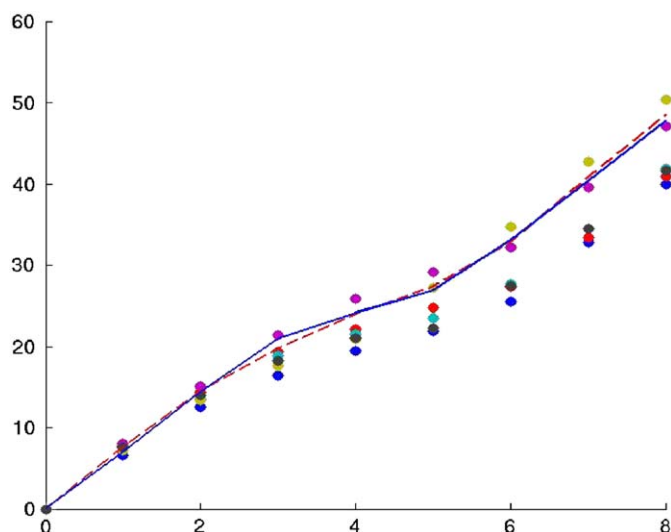


Fig. 4. Data, mean, and model prediction with binomial n .

The final unknown parameter is p , the probability that a cell becomes a survivor, for which we try different values to examine the fit to data. Fig. 3 displays the seven data sets (dots), the mean (dashed red line), and three lines computed from our model where n has a uniform distribution with mean 35 (range 20–50). The solid line, which presents the best fit, has $p = 2 \times 10^{-5}$ whereas the two dashed lines have $p = 10^{-3}$ (upper) and $p = 2 \times 10^{-6}$ (lower), respectively. In Fig. 4, n instead has a binomial distribution with mean 35.

6. Discussion

We used a general branching process to model growth and telomere dynamics in the yeast *S. cerevisiae*. The model accounts for random variability in cell cycle times and shortening of telomeres by “telomere units” of 4 bp. Cells also have the possibility to turn into survivors and we assumed that this happens independently with probability p in cells that are approaching a critical telomere length. The model also takes into account the known fact that cells undergo a finite number of cell divisions even when telomerase is expressed and telomere length is maintained. The general branching process easily deals with this problem since it allows a mother cell to repeatedly give birth to daughter cells until the maximum number is reached.

Cell cycle times were assumed to follow a gamma distribution whose parameters were estimated from data on telomerase-proficient yeast by applying asymptotic branching process results. We then used these estimates together with values of initial cell lifespan and telomere length, taken from the literature, to predict the development of mean population size over time. Different values of the remaining parameter p were tried and the best fit to data occurred for p on the order of 2×10^{-5} . It should be emphasized that our goal was not to get as precise estimates as possible; rather, we aimed at building a stochastic model that incorporates the necessary parameters to describe cellular proliferation in response to telomere attrition and maintenance. Our model agrees with data very well, both qualitatively and quantitatively, but in order to perform a more sophisticated analysis, a new experimental design needs to be implemented to eliminate some of the uncertainties present in the current data.

We have made several simplifications in the model. Here, we considered the average replicative lifespan of pre- and post-senescent survivors to be equivalent, however, recent work on

survivors indicates that this limit on replicative lifespan is reduced in type II survivors (Chen et al., 2009). We also disregarded the possibility that cell cycle times slow down progressively during senescence progression or replicative aging (Sinclair et al., 1998). In addition, we viewed population counts as actual counts rather than estimates based on sequential sampling. A more thorough analysis could take into account sampling errors when the 10^5 cells are harvested from much larger populations to initiate another day of growth.

The approach taken here allows for future evaluation of the contribution of specific factors on growth dynamics and survivor formation in response to telomerase-deficiency. For example, the effects of an *exo1Δ* deletion, which has been shown to slow senescence progression and favor the formation of type I survivors (Bertuch and Lundblad, 2004; Maringele and Lydall, 2004) or the effects of a *rad50Δ* mutation, which eliminates the formation of type II survivors (Teng et al., 2000), can be mathematically modeled.

Appendix A

In Section 4, we disregarded the effects of n_0 , thus approximating the general branching process by the simpler Bellman–Harris process. In this appendix, we take a closer look at how the two models relate. In a general branching process, the fundamental object is the reproduction process $\xi(t)$ which gives the number of daughter cells a mother cell has up to age t . Let the times between consecutive births (i.e., cell cycle times) be independent random variables with the common cdf F . If we denote the time of the k th birth by T_k , the reproduction process becomes

$$\xi(t) = \sum_{k=1}^{n_0} I\{T_k \leq t\}$$

where I denotes indicator function. The mean of $\xi(t)$ is

$$\mu(t) = \sum_{k=1}^{n_0} P(T_k \leq t) = \sum_{k=1}^{n_0} F^{*k}(t) \quad (\text{A.1})$$

where F^{*k} denotes the k -fold convolution of F with itself, that is, the cdf of T_k . For simplicity, we assume that n_0 is constant, however, a random n_0 is easily dealt with, see Olofsson (2009). See the discussion after (A.5) for justification of this assumption. Note that we can write

$$\mu(t) = \int_0^t \mu(du)$$

where

$$\mu(du) = \sum_{k=1}^{n_0} f^{*k}(u) du$$

if F is a continuous distribution with probability density function f . The Laplace transform of μ is defined as

$$\hat{\mu}(s) = \int_0^\infty e^{-st} \mu(dt)$$

and the asymptotic growth rate of the process is given by the Malthusian parameter α which is defined through the relation

$\hat{\mu}(\alpha) = 1$. Finally, let

$$\beta = \int_0^{\infty} t e^{-\alpha t} \mu(dt)$$

The asymptotic mean and variance of Z_t are then given by

$$E[Z_t] \sim e^{2t} \frac{1}{\alpha\beta} \quad (\text{A.2})$$

and

$$\text{Var}[Z_t] \sim e^{2t} \frac{\text{Var}[\hat{\xi}(\alpha)]}{\alpha^2 \beta^2 (1 - \hat{\mu}(2\alpha))} \quad (\text{A.3})$$

where

$$\hat{\xi}(\alpha) = \int_0^{\infty} e^{-\alpha t} \xi(dt)$$

For technical details and conditions, see Jagers and Nerman (1984).

The Laplace transform of the $\Gamma(a, b)$ distribution is

$$\hat{f}(\alpha) = \int_0^{\infty} e^{-\alpha t} f(t) dt = \left(\frac{b}{\alpha + b} \right)^a$$

As T_k is the sum of k independent $\Gamma(a, b)$ variables, we get $T_k \sim \Gamma(ka, b)$ and as the Laplace transform of the sum of independent random variables equals the product of the individual Laplace transforms, (A.1) gives

$$\hat{\mu}(\alpha) = \sum_{k=1}^{n_0} \left(\frac{b}{\alpha + b} \right)^{ka} = \frac{1 - \left(\frac{b}{\alpha + b} \right)^{a(n_0 + 1)}}{1 - \left(\frac{b}{\alpha + b} \right)^a} - 1 \quad (\text{A.4})$$

In the Bellman–Harris approximation in Section 4 we got

$$\alpha = b(2^{1/a} - 1)$$

Note that this relation implies that

$$\left(\frac{b}{\alpha + b} \right)^a = \frac{1}{2}$$

which, by (A.4) gives

$$\hat{\mu}(\alpha) = 1 - \left(\frac{1}{2} \right)^{n_0} \quad (\text{A.5})$$

As we have $\hat{\mu}(\alpha) < 1$ for any n_0 , the Bellman–Harris approximation overestimates α , which makes intuitive sense since it disregards the restrictions imposed by the finite lifespan of cells. However, for realistic values of n_0 in the range 20–30 (Sinclair and Guarente, 1997; Sinclair et al., 1998) we have $\hat{\mu}(\alpha) \approx 1$, confirming that our approximation is reasonable. Eq. (A.5) also provides some justification for assuming a constant n_0 ; over the range of values that have been experimentally determined (Sinclair and Guarente, 1997; Sinclair et al., 1998) the effect of varying n_0 is negligible.

References

- Arino, O., Kimmel, M., Webb, G.F., 1995. Mathematical modeling of the loss of telomere sequences. *J. Theor. Biol.* 177, 45–57.
 Arino, O., Sanchez, M., Webb, G.F., 1997. Polynomial growth dynamics of telomere loss in a heterogeneous cell population. *Dynam. Contin. Discrete Impuls. Syst.* 3, 262–282.

- Baird, D.M., 2008. Telomere dynamics in human cells. *Biochimie* 90, 116–121.
 Bertuch, A.A., Lundblad, V., 2004. *EXO1* contributes to telomere maintenance in both telomerase-proficient and telomerase-deficient *Saccharomyces cerevisiae*. *Genetics* 166, 1651–1659.
 Bryan, T.M., Dalla-pozza, L., Dunham, M.A., Reddel, R.R., 1997. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat. Med.* 3, 1271–1274.
 Chen, X.F., Meng, F.L., Zhou, J.Q., 2009. Telomere recombination accelerates cellular aging in *Saccharomyces cerevisiae*. *PLoS Genet.* 5, e1000535.
 D'Mello, N.P., Jazwinski, S.M., 1991. Telomere length constancy during aging of *Saccharomyces cerevisiae*. *J. Bacteriol.* 173, 6709–6713.
 Denchi, E.L., 2009. Give me a break: how telomeres suppress the DNA damage response. *DNA Repair (Amst)* 8, 1118–1126.
 Dyson, J., Vिलлелла-бressан, R., Webb, G.F., 2007. Stabilization of telomeres in nonlinear models of proliferating cell lines. *J. Theor. Biol.* 244 (3), 400–408.
 Flores, I., 1967. Direct calculation of k -generalized Fibonacci numbers. *Fibonacci Q.* 5, 259–266.
 Haccou, P., Jagers, P., Vatutin, V., 2005. *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, Cambridge.
 Harris, T.E., 1963. *The Theory of Branching Processes*. Springer, Berlin.
 Jagers, P., Nerman, O., 1984. The growth and composition of branching populations. *Adv. Appl. Probab.* 16, 221–259.
 Jagers, P., 1992. Stabilities and instabilities in population dynamics. *J. Appl. Probab.* 29, 770–780.
 Larsson, S., Ryden, T., Holst, U., Oredsson, S., Johansson, M., 2008. Estimating the total rate of DNA replication using branching processes. *Bull. Math. Biol.* 70 (8), 2177–2194.
 Levy, M.Z., Allsop, R.C., Futcher, A.B., Greider, C.W., Harley, C.B., 1992. Telomere end-replication problem and cell aging. *J. Molec. Biol.* 225, 951–960.
 Lundblad, V., Blackburn, E.H., 1993. An alternative pathway for yeast telomere maintenance rescues *est1* senescence. *Cell* 73, 347–360.
 Lundblad, V., Zostak, J.W., 1989. A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57, 633–643.
 Maringele, L., Lydall, D., 2004. *EXO1* plays a role in generating type I and type II survivors in budding yeast. *Genetics* 166, 1641–1649.
 Olofsson, P., Kimmel, M., 1999. Stochastic models of telomere shortening. *Math. Biosci.* 1, 75–92.
 Olofsson, P., 2000. A branching process model of telomere shortening. *Comm. Stat. Stochastic Models* 16 (1), 167–177.
 Olofsson, P., Kimmel, M., 2005. Mathematical modeling of telomere shortening. In: Haccou, P., Jagers, P., Vatutin, V. (Eds.), *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, Cambridge, pp. 225–231.
 Olofsson, P., 2009. Can telomere shortening explain sigmoidal growth curves? *J. Biological Dynamics*, to appear.
 Oprea, M., Kepler, T.B., 2001. Improved inference of mutation rates: II. Generalization of the Luria–Delbruck distribution for realistic cell-cycle time distributions. *Theor. Popul. Biol.* 59 (1), 49–59.
 Op Den Buijs, J., Van denbosch, P.P.J., Musters, M.W.J.M., Van riel, N.A.W., 2004. Mathematical modeling confirms the length-dependency of telomere shortening. *Mech. Ageing Dev.* 125 (6), 437–444.
 Portugal, R.D., Land, M.G.P., Svaiter, B.F., 2008. A computational model for telomere-dependent cell-replicative aging. *Biosystems* 91, 262–267.
 Rubelj, I., Vondracek, Z., 1999. Stochastic mechanism of cellular aging—abrupt telomere shortening as a model for stochastic nature of cellular aging. *J. Theor. Biol.* 197, 425–438.
 Shampay, J., Blackburn, E.H., 1988. Generation of telomere-length heterogeneity in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 85, 534–538.
 Sinclair, D.A., Guarente, L., 1997. Extrachromosomal rDNA circles—a cause of aging in yeast. *Cell* 91 (7), 1033–1042.
 Sinclair, D.A., Mills, K., Guarente, L., 1998. Aging in *Saccharomyces cerevisiae*. *Annu. Rev. Microbiol.* 52, 533–560.
 Singer, M.S., Gottschling, D.E., 1994. TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science* 266, 404–409.
 Tan, Z., 1999. Intramitotic and intraclonal variation in proliferative potential of human diploid cells: explained by telomere shortening. *J. Theor. Biol.* 198, 259–268.
 Teng, S.C., Zakian, V.A., 1999. Telomere–telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol. Cellular Biol.* 19 (12), 8083–8093.
 Teng, S.C., Chang, J., McCowan, B., Zakian, V.A., 2000. Telomerase-independent lengthening of yeast telomeres occurs by an abrupt Rad50p-dependent, Rif-inhibited recombinational process. *Mol. Cell* 6, 947–952.