J. Math. Biology (1986) 23: 231-246

Journal of
Mathematical
Biology

Cell growth and division: a deterministic/probabilistic model of the cell cycle

John J. Tyson and Kenneth B. Hannsgen

Departments of Biology and Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

Abstract. A model of the cell cycle, incorporating a deterministic cell-size monitor and a probabilistic component, is investigated. Steady-state distributions for cell size and generation time are calculated and shown to be globally asymptotically stable. These distributions are used to calculate various statistical quantities, which are then compared to known experimental data. Finally, the results are compared to distributions calculated from a Monte-Carlo simulation of the model.

Key words. Cell division cycle — size distribution — generation time distribution

1. Introduction

The fundamental purpose of the cell division cycle is to make two cells from one. This involves (i) accurately replicating the cell's DNA and apportioning sister chromatids to daughter nuclei at mitosis, and (ii) approximately doubling all other bulk constituents of the cell (e.g. mitochondria, ribosomes,. plasmalemma) and dividing them more-or-less evenly between daughter cells at cell division. Great care is taken by every cell to ensure that each daughter gets an exact copy of the genetic material, but less attention seems to be paid by individual cells to the amount of other materials inherited by their daughters. This "sloppiness" is evidenced by a broad distribution of cell size at birth (coefficient of variation $\simeq 10\%$). Nonetheless, there is some coordination between overall cell growth and division, because there does exist a characteristic size distribution (cells do not become arbitrarily large or small) and because this size distribution is stable (after a perturbation the characteristic size distribution is rapidly re-established). Furthermore, there is evidence that cell size plays an important role in setting the time of cell division (Mitchison 1977; John 1981; Nurse and Streiblova 1984). For example, cell size at division is typically less variable than cell age at division (CV for size $\simeq 10\%$ compared to CV for age $\simeq 20\%$), and there is a strong negative correlation between size at birth and generation time (i.e. time from birth to division).

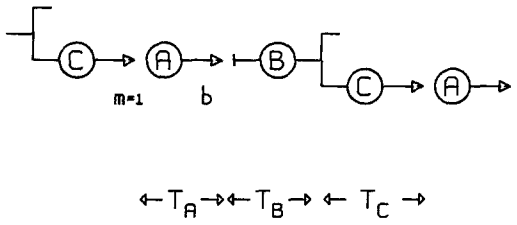

Fig. 1. A deterministic/probabilistic model of cell cycle. After birth, cells must grow to a critical size ($m=1$) before entering the indeterministic A-phase, from which they exit with constant probability per unit time (p). Cells divide a fixed time period (T_B) after leaving the A-state

In this paper we examine a model of the cell cycle that incorporates both a deterministic cell-size monitor and a probabilistic, variability-generating component. The model is illustrated in Fig. 1. When a cell reaches the critical size $m=1$, it enters the A-phase of the cell cycle (Smith and Martin 1973), from which it exits with constant probability per unit time, p. Variable residence times in A-phase, T_A, generate variability in all cell cycle parameters. On leaving A-phase, the cell enters B-phase and divides after a fixed time interval, T_B. If cell size at birth is less than 1, then the cell finds itself in C-phase, during which it must grow until it achieves the critical size $m=1$. If a cell is born with $m>1$, it enters directly into A-phase.

This model has been suggested by a number of authors (Shilo et al. 1976; Fantes 1977; Shields et al. 1978; Nurse 1980) and has been aptly called the "tandem" model by Lord and Wheals (1981). In a recent publication (Tyson and Hannsgen 1985) we have analyzed the tandem model in some detail, concentrating on the distributions of cell size and generation time defined on samples of cells born in a narrow time window $(t, t+\Delta t)$, which we shall refer to as a "contemporaneous" sample of cells. In this paper we reformulate the tandem model in terms of distributions defined on samples of cells belonging to distinct "generations".

Because we assume that cells leave A-phase with constant probability per unit time (p), the residence time in A-phase (T_A) must be an exponentially-distributed random variable, i.e.

$$\text{Prob}\{T_A \geq t\} \equiv \int_t^\infty f_A(s)\, ds = e^{-pt}. \tag{1}$$

We shall also assume that individual cells grow exponentially with specific growth rate k. Let $\tau = \ln 2/k$ denote the mass-doubling time of individual cells. During steady state growth, τ is also the mass-doubling time of the cell population as a whole. We must assume that $\tau > T_B$; otherwise, the cell population would be doubling in mass faster than any cell can possibly divide so there could be no stable size distribution. Under these assumptions it is obvious that the minimum size of a cell at birth is

$$\sigma = e^{kT_B}/2 = e^{-k(\tau - T_B)}, \quad \text{with } \tfrac{1}{2} < \sigma < 1. \tag{2}$$

2. Size distribution functions

There are many different ways to express the distribution of cell size in an asynchronous culture of growing cells. Do we want to know size at birth, size at

division, size at onset of DNA synthesis? Do we want to choose for our sample all cells at time t, e.g. all cells born in narrow time window $(t, t+\Delta t)$, or all newborn cells belonging to a certain generation n? The specification of the sample over which the distribution is defined is a critical and often neglected factor. For instance, the distribution of mass at division in a sample of dividing cells at time t is generally different from the distribution of mass at division in a sample of newborn cells at time t, since in any narrow time window one is more likely to sample rapidly dividing cells than slowly dividing cells.

Painter and Marr (1968) have suggested a useful notation for size distributions. They let the Greek letters ψ, λ, and ϕ denote, respectively, the distributions of size at birth, size at present, and size at division; and they let the subscripts b, e, and m denote the sample on which the distribution is defined, whether newborn cells ("babies"), extant cells, or dividing cells ("mothers"). For example, $\psi_e(x, t)\,dx$ is the probability that, choosing a cell at random from a sample of all cells alive at time t, the chosen cell has a birth size in the interval $(x, x+dx)$. Similarly, $\psi_b(x, t)\,dx$ is the probability that a cell has birth size between x and $x+dx$, if the cell is chosen from a sample of all cells born in a narrow time window $(t, t+\Delta t)$. We must introduce the artifice of a narrow time window when speaking of samples of newborn cells or samples of dividing cells because the size of such samples tends to zero as $\Delta t \to 0$.

During balanced, or steady-state, growth these size distributions do not depend on t, so $\phi_b(x)$, etc., can represent the steady-state probability densities. Furthermore, Painter and Marr suggest that the subscripts be dropped from the most natural combinations:

$$\psi(x) \equiv \psi_b(x) = \text{steady-state probability density for size at birth in a sample of newborn cells;}$$

$$\lambda(x) \equiv \lambda_e(x) = \text{steady-state probability density for present size in a sample of extant cells;}$$

$$\phi(x) \equiv \phi_m(x) = \text{steady-state probability density for size at division in a sample of dividing cells.}$$

The distributions used by Painter and Marr (1968) are all defined on samples chosen from cells alive during a brief time span. It was with such samples that we were concerned previously. In this paper we want to consider samples of cells of a certain generation, and we suggest that the subscript n be used to denote such samples. Thus, $\psi_n(x)$ will denote the probability density of size at birth in a sample of cells of generation n, and $\phi_n(x)$ the probability density of size at division in a sample of cells of generation n. (There is no analog to λ, the present size distribution, for generation-based samples.) In the state of balanced growth these distributions no longer depend on generation number, so we suggest that the subscript $_*$ denote the steady-state distributions:

$$\psi_n(x) \to \psi_*(x), \qquad \phi_n(x) \to \phi_*(x) \quad \text{as } n \to \infty.$$

One of our goals is to determine the distribution of generation times in the tandem model, but on what sample should this distribution be defined? Painter and Marr (1968) let $f(T) \equiv f_b(T) =$ steady-state probability density for generation

time on a sample of cells newborn in some time window. But, as Powell (1956, 1964) has emphasized, if there is a correlation between generation times of mother and daughter cells, then the time-window convention produces a biased sample. The bias comes about because a sample of cells newborn in a narrow time window is biased toward the daughters of rapidly dividing cells, and if mother and daughter generation times are negatively correlated, this sample will be biased toward cells with long generation times. Powell (1964) suggested that an unbiased sample would be all newborn cells of generation n where n is large enough so that steady-state conditions can be expected, i.e. $f_*(T)$.

On the basis of experimental observations we asserted above that a population of cells will approach, from arbitrary initial conditions, a steady-state of balanced exponential growth, characterized by time-independent distribution functions for cell size and generation time. For distribution functions based on contemporaneous samples of cells, the global asymptotic stability of the steady-state size distributions can be difficult to prove even in favorable cases (see, e.g. Diekmann et al. 1984; Heijmans 1984; Hannsgen and Tyson 1985). However, for distribution functions based on samples of cells of the same generation, Lasota and Mackey (1984) have shown that, under quite general conditions, the evolution of the birth-size probability density obeys a simple recursion relation from which global asymptotic stability of the size distribution can often be proved by verifying a few simple inequalities.

In the next sections we apply the Lasota–Mackey theory to the tandem model. Since many of our results do not follow directly as trivial special cases of equations in Lasota and Mackey (1984), we present the entire theory in a form suitable to the tandem model.

3. The recursion relation

Suppose we are given the distribution of size at birth in generation 0, $\psi_0(x)$, and we wish to determine $\psi_n(x)$, $n = 1, 2, 3, \ldots$. Consider a cell, in generation n, of given birth-size, $x_n = y$, and let x_{n+1} be the birth-size of its daughters (assuming division of a mother cell into two equally sized daughters). Recall that individual cells are assumed to grow exponentially with specific growth rate k and that $\sigma = \exp\{-k(\tau - T_B)\}$ = minimum size at birth. Now, provided $x \geq \sigma$,

$$\text{Prob}\{x_{n+1} \geq x | x_n = y\} = \begin{cases} \text{Prob}\{T_A \geq k^{-1} \ln(x/\sigma)\} & (\sigma \leq y \leq 1) \\ \text{Prob}\{T_A \geq k^{-1} \ln(x/\sigma y)\} & (1 \leq y \leq x/\sigma) \\ 1 & (x/\sigma \leq y < \infty). \end{cases}$$

Since by assumption (1) T_A is exponentially distributed, we have

$$\text{Prob}\{x_{n+1} \geq x\} = (x/\sigma)^{-p/k} \int_\sigma^1 \psi_n(y)\, dy + (x/\sigma)^{-p/k} \int_1^{x/\sigma} y^{p/k} \psi_n(y)\, dy + \int_{x/\sigma}^\infty \psi_n(y)\, dy.$$

Differentiating this equation with respect to x, we obtain

$$\psi_{n+1}(x) = \int_{\sigma}^{\infty} K(x, y)\psi_n(y)\, dy, \qquad x \geq \sigma \tag{3}$$

where

$$K(x, y) = \begin{cases} (p/k\sigma)(x/\sigma)^{-1-(p/k)}, & \sigma \leq y \leq 1, \\ (p/k\sigma)(x/\sigma)^{-1-(p/k)} y^{p/k}, & 1 \leq y \leq x/\sigma, \\ 0, & y > x/\sigma. \end{cases} \tag{4}$$

Notice that

$$\int_{\sigma}^{\infty} K(x, y)\, dx = 1 \tag{5}$$

for all y, as it must be in order that

$$\int_{\sigma}^{\infty} \psi_n(x)\, dx = 1 \text{ guarantees that } \int_{\sigma}^{\infty} \psi_{n+1}(x)\, dx = 1. \tag{6}$$

Equation (4) can be derived from Eq. (20) in Lasota and Mackey (1984) after identifying the function $q(x)$ in Lasota and Mackey's equation with p/kx if $x > 2\sigma \max\{1, y\}$ and 0 otherwise.

4. The steady state size distribution

Any steady state size distribution, $\psi_*(x)$, satisfies

$$\psi_*(x) = \int_{\sigma}^{\infty} K(x, y)\psi_*(y)\, dy, \qquad \int_{\sigma}^{\infty} \psi_*(x)\, dx = 1. \tag{7}$$

If we look for solutions of the form $\psi_*(x) = x^{-r}$ (r real, $r > 1$), we find that the exponent r must satisfy the characteristic equation

$$p - k(r-1) = p\sigma^{(r-1)}, \tag{8}$$

or

$$1 - s = e^{-as}, \tag{8a}$$

with $s = k(r-1)/p$ and $a = p(\tau - T_B)$. In order that $\psi_*(x) = x^{-r}$ be integrable, i.e. in order that $r > 1$, we must insist that $k/p < -\ln \sigma$, i.e. that $p(\tau - T_B) > 1$. Properly normalized, our steady state size distribution is

$$\psi_*(x) = [(r-1)/\sigma](x/\sigma)^{-r}, \qquad x \geq \sigma \tag{9}$$

or

$$\psi_*(x) = (ps/k\sigma)(x/\sigma)^{-1-(ps/k)}, \qquad x \geq \sigma. \tag{9a}$$

To prove uniqueness and global asymptotic stability of the steady state solution (9), we must first restate the machinery developed by Lasota and Mackey (1984).

Let D be the set of all real-valued functions in $L^1(=L^1[\sigma, \infty))$ which satisfy

$$\psi(x) \geq 0 \quad \text{and} \quad \int_\sigma^\infty \psi(x)\, dx = 1. \tag{10}$$

Let $K(x, y)$ be a real-valued function defined on $x, y \in [\sigma, \infty)$

$$K(x, y) \geq 0 \quad \text{and} \quad \int_\sigma^\infty K(x, y)\, dx = 1 \quad \text{for all } y. \tag{11}$$

Let the operator $P: L^1 \to L^1$ be defined by

$$P\psi(x) = \int_\sigma^\infty K(x, y)\psi(y)\, dy. \tag{12}$$

A function $h \in L^1$ will be called a non-trivial *lower bound function* for P if, for every $\psi \in D$, there exists an integer $n_0(\psi)$ such that

$$P^n\psi(x) \geq h(x), \; n \geq n_0(\psi), \tag{13}$$

and if

$$h(x) \geq 0 \quad \text{and} \quad \int_0^\infty h(x)\, dx > 0. \tag{14}$$

Theorem 1 (Lasota and Yorke 1982; Lasota and Mackey 1984). *If, for kernel $K(x, y)$ satisfying* (11), *there exists a nontrivial lower bound function for the operator* (12), *then Eq.* (7) *has a unique solution $\psi_* \in D$. Moreover, for any $\psi \in D$,*

$$\lim_{n\to\infty} \int_\sigma^\infty |P^n\psi(x) - \psi_*(x)|\, dx = 0. \tag{15}$$

Theorem 2 (Lasota and Mackey 1984). *If $K(x, y)$ satisfies* (11), *and if*

$$\int_\sigma^\infty xK(x, y)\, dx \leq \gamma y + \delta, \qquad y \geq 0 \tag{16}$$

for some nonnegative constants γ and δ, $\gamma < 1$, and if

$$\int_0^\infty \inf_{y\in[0,\,\alpha]} K(x, y)\, dx > 0 \tag{17}$$

for some $\alpha > \delta/(1-\gamma)$, then there exists a nontrivial lower bound function for the operator (12).

(Note: Lasota and Mackey require a stronger condition than our (17); a simple modification of their proof justifies our version.)

Thus, to prove uniqueness and global asymptotic stability of (9), we must verify (16) and (17). Assuming for a moment that $p > k$, we find that Eq. (16) holds with $\gamma = \sigma p/(p-k)$ and $\delta = (1-\sigma)\sigma p/(p-k)$. In order that $\gamma < 1$, we must insist not merely that $p > k$ but further that

$$1 - \sigma > k/p. \tag{18}$$

Then, taking $\alpha = \{(1-\sigma)\sigma p/[p(1-\sigma)-k]\}+\varepsilon$ $(\varepsilon>0)$, we find that

$$\inf_{y\in[0,\alpha]} K(x,y) = \begin{cases} 0, & \text{if } x<\alpha\sigma \\ (p/k\sigma)(x/\sigma)^{-1-(p/k)}, & \text{if } x>\alpha\sigma \end{cases}$$

so (17) is satisfied. Thus, (18) is a sufficient condition for uniqueness and global asymptotic stability of $\psi_*(x)$. It is interesting to note that (18) is identical to the condition $(r>2)$ that $\psi_*(x)$ given by (9) have finite first moment, i.e. that there be a finite average birth size. For $\frac{1}{2}<\sigma<1$ and

$$1-\sigma<k/p<-\ln\sigma, \tag{19}$$

we have $1<r<2$, so $\psi_*(x)$ given by (9) is integrable but $x\psi_*(x)$ is not. We suspect that $\psi_*(x)$ is still unique and globally asymptotically stable in case (19), but we cannot prove so using the Lasota-Mackey theorems. Furthermore, we suspect that there does not exist an integrable, stable steady-state solution to (3) when $k/p>-\ln\sigma$, but we have not tried to prove this assertion.

5. The generation time distribution

Now that we know $\psi_*(x)$, the steady-state distribution of birth size on a sample of cells belonging to the same generation, we can calculate the steady-state distribution of residence times in C-phase (again) on a sample of cells belonging to the same generation, which we denote $f_C(t)$. This distribution will have three contributions:

1) the probability that a cell, chosen at random from a sample of cells of the same generation, has $T_C=0$ is just the probability that a cell has birth mass >1, i.e.

$$\int_1^\infty \psi_*(x)\,dx = \sigma^{(ps/k)} = 1-s. \tag{20}$$

2) the probability that a cell, chosen at random from a sample of cells of the same generation, has $T_C>\tau-T_B$ is identically zero, since the maximum value of T_C is given by $1=\sigma\exp(kT_C^{\max})$.

3) the probability that a cell, chosen at random from a sample of cells of the same generation, has T_C greater than some chosen time, $0<t<T_C^{\max}$, is

$$\begin{aligned} \text{Prob}\{t<T_C<T_C^{\max}\} &= \text{Prob}\{\sigma<\text{birth size}<e^{-kt}\} \\ &= \int_\sigma^{\exp(-kt)} \psi_*(x)\,dx \\ &= 1-(1-s)\,e^{pst}. \end{aligned} \tag{21}$$

Differentiating (21) with respect to t, we find that

$$f_C(t) = ps(1-s)\,e^{pst}, \quad \text{for } 0<t<\tau-T_B.$$

Collecting these three facts, we can write

$$\begin{aligned} f_C(t) &= (1-s)2\delta(t)+ps(1-s)e^{pst}, \quad 0\leqslant t\leqslant\tau-T_B, \\ &= 0, \qquad\qquad\qquad\qquad\quad t>\tau-T_B, \end{aligned} \tag{22}$$

where $\delta(t)$ is the Dirac delta function, which satisfies $\delta(t)=0$ if $t\neq 0$, and $\int_0^\infty \delta(t)\,dt=\frac{1}{2}$.

Now we are in a position to calculate $f_*(T)$, the steady-state generation-time distribution on a sample of cells belonging to the same generation. Since $T-T_B=T_A+T_C=$ sum of two independent random variables,

$$f_*(T)=\int_0^{T-T_B} f_A(T-T_B-t)f_C(t)\,dt.$$

Using (1) and (22), we find that

$$\begin{aligned} f_*(T)&=0, \quad \text{if } 0\leqslant T<T_B,\\ &=C[1+s\exp\{p(1+s)(T-T_B)\}]p\exp\{-p(T-T_B)\}, \quad \text{if } T_B\leqslant T\leqslant\tau\\ &=C[1+s\exp\{p(1+s)(\tau-T_B)\}]p\exp\{-p(T-T_B)\}, \quad \text{if } T\geqslant\tau, \end{aligned} \tag{23}$$

where $C=(1-s)/(1+s)$.

The mean generation time is given by

$$\langle T\rangle=\int_0^\infty tf_*(t)\,dt=\langle T_A\rangle+T_B+\langle T_C\rangle=\tau. \tag{24}$$

That is, the mean generation time is *identical* to the mass-doubling time of the population. At first sight, this result seems wrong, since Painter and Marr (1967) emphasize that the mean generation time is always larger than the mass-doubling time. However, we must recall that the Painter-Marr proof is for a generation-time distribution based on a sample of cells newborn *at a particular instant of time*, whereas our result (24) refers to the average of a generation-time distribution based on a sample of cells *belonging to a particular generation.*

In the Appendix we prove that $\langle T\rangle=\tau$ whenever cells grow exponentially and divide exactly in half. Otherwise, such a simple relation is not generally true.

6. Alpha and beta curves

Since cell kineticists often plot generation-time data as survivorship functions, we calculate the function $\alpha(t)=\int_t^\infty f_*(T)\,dT$,

$$\begin{aligned} \alpha(t)&=1, \quad \text{if } 0<t<T_B,\\ &=1-C[\exp\{ps(t-T_B)\}-\exp\{-p(t-T_B)\}], \quad \text{if } T_B\leqslant t\leqslant\tau,\\ &=C\exp\{-p(t-T_B)\}[1+(s/1-s)e^a], \quad \text{if } t\geqslant\tau. \end{aligned} \tag{25}$$

Equation (25) can be derived from Eq. (8) in Lasota and Mackey (1984), but the derivation is quite tedious.

Notice that, for $t>\tau$, $\ln\alpha(t)=\ln(\text{constant})-p(t-T_B)$, which gives a straight line when $\ln\alpha$ is plotted against t (an "alpha-curve"). That this must be so follows from the fact that $\max(T_C)=\tau-T_B$: all cells with $T>\tau$ must be in the A-state $\tau-T_B$ time units after birth; these cells leave the A-state with constant probability per unit time so $\alpha(t)$ must be a decaying exponential for $t>\tau$. For $T_B\leqslant t\leqslant\tau$, there is a shoulder which connects the "exponential tail" with the

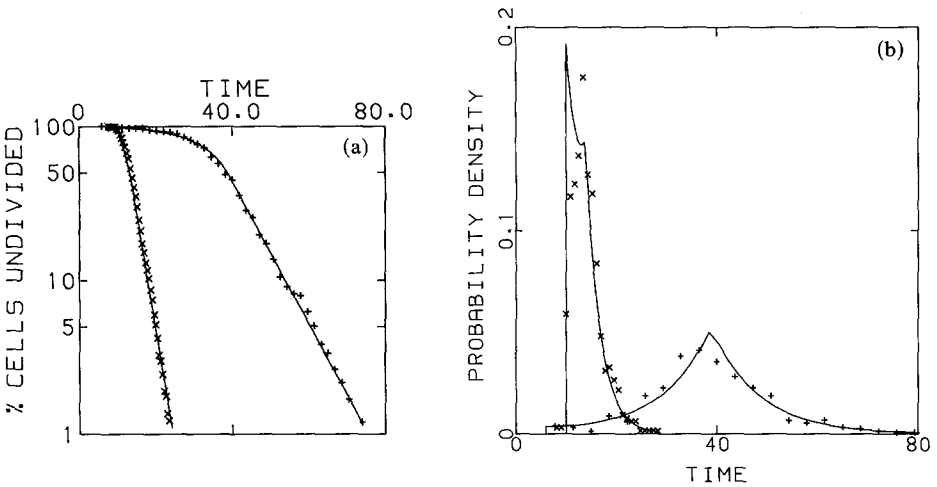

Fig. 2a, b. Generation-time distributions for (×) *Balbc 3T3* cells (Shields 1980) and for (+) *Staphylococcus albus* (Shields 1978). The time unit is hours for the murine cell data and minutes for the bacterial cell data. The continuous curves are calculated from the model with the following parameter values: for *Balbc 3T3*, $p = 0.35\ \mathrm{h}^{-1}$, $\tau = 13.8$ h, and $T_B = 10$ h; for *S. albus*, $p = 0.106\ \mathrm{min}^{-1}$, $\tau = 38.8$ min, $T_B = 6$ min. (**a**) α-curves (cumulative distributions plotted semilogarithmically) calculated from the data and from Eq. (25). (**b**) Histograms of cell generation times constructed from the data, and probability distribution functions calculated from Eq. (23)

point $\ln \alpha = 0$ at $t = T_B$. If $p(\tau - T_B) \gg 1$, then, from Eq. (8a), $s \simeq 1$ and

$$\begin{aligned} \alpha(t) &\simeq 1 - e^{-a} \cosh[p(t - T_B)], \quad T_B \leqslant t \leqslant \tau, \\ &\simeq 0.5 \exp[-p(t - \tau)], \qquad t \geqslant \tau. \end{aligned} \tag{26}$$

Figure 2a illustrates the excellent fit of Eq. (25) to experimental alpha-curves for bacterial and murine cell populations. These are not "best fits" in any statistical sense, but merely casual fits obtained by estimating p from the slope of the exponential tail, T_B from the minimum observed generation time, and τ from the median generation time (cf. Eq. (26)). Using these values of p, T_B and τ we can sketch the expected generation-time distribution $f_*(T)$ according to Eq. (23) and compare the expected distribution with the observed histograms, as in Fig. 2b. Presumably, one could improve the fit by standard statistical procedures, such as the principle of maximum likelihood, but we have not attempted to do this.

Another statistic that is commonly reported for cell populations is

$$\beta(t) = \text{probability that } |T_1 - T_2| \geqslant t, \quad \text{where } T_1 \text{ and } T_2 \text{ are the generation times of sister cells.}$$

For the tandem model, sister cells (which are by assumption identical in size at birth) have identical values of T_C and differ only in T_A. Since T_A is exponentially distributed, we have $\beta(t) = e^{-pt}$. Thus, a plot of $\ln \beta$ *versus* t ("beta-curve") will be a straight line through the origin with the same slope as the tail of the alpha-curve for the cell population. This situation is commonly observed (Minor and Smith 1974; Shields and Smith 1977). Indeed, for the bacterial cell population whose alpha-curve is reported in Fig. 2, Shields (1978) has found that the

beta-curve is a straight line through the origin with slope $= -0.1\ \text{min}^{-1}$, exactly as we would expect from the tandem model. However, it should be recognized that beta-curves are not always parallel to the tail of the corresponding alpha-curve. For instance, for the murine cell population whose alpha-curve is given in Fig. 2, Shields (1980) has found that the beta-curve is a straight line through the origin with slope $= -0.6\ \text{h}^{-1}$, which is twice the expected value.

7. Correlation coefficients

Our model can be tested further by evaluating the correlation in generation times between sister cells and between mother-daughter pairs. For sister cells the correlation coefficient is (Green 1980)

$$r_{\text{sis}} = \text{var}(T_C)/[\text{var}(T_C) + \text{var}(T_A)]. \tag{27}$$

Since $\text{var}(T_A) = p^{-2}$ and $\text{var}(T_C) = p^{-2}[1 - 2s^{-1}(1-s)(a-1)]$, we have

$$r_{sis} = [2(1-s)(a-1) - s]/[2(1-s)(a-1) - 2s]. \tag{28}$$

Since $s = s(a)$, r_{sis} is a function of a only, and it is plotted in Fig. 3. Though the model predicts that $0 < r_{sis} < 0.5$, sisiter cell correlation coefficients are often greater than 0.5. (Powell 1956; Schaechter et al. 1962; Kubitschek 1962; Minor and Smith 1974; Shields 1980).

Mother and daughter generation times (T^m and T^d) are negatively correlated since, as long as $T_A^m + T_B < \tau$, $T^d = \tau - T_A^m + T_A^d$. Provided that $T_A^m + T_B < \tau$, the correlation coefficient is

$$r_{md} = -\text{var}(T_A)/\{2\,\text{var}(T_A)[\text{var}(T_A) + \text{var}(T_C)]\}^{1/2}, \tag{29}$$

which reduces to $-0.5\{s/[1 - a(1-s)]\}^{1/2}$. Now, $T_A^m < \tau - T_B$ for a fraction of mother-daughter pairs $= 1 - e^{-a}$. The rest of the pairs, for which $T_A^m > \tau - T_B$, are uncorrelated. Thus, roughly speaking,

$$r_{md} \simeq -0.5(1 - e^{-a})\{s/[1 - a(1-s)]\}^{1/2}, \tag{30}$$

which is plotted in Fig. 3. The model predicts that $r_{md} \simeq -0.5$ for all values of a, which is, in general too negative (Powell 1956; Schaechter et al. 1962; Kubitschek 1962; Shields and Smith 1977).

That the predicted correlation coefficients for sister cells and for mother-daughter pairs are too negative is a common failing of deterministic models of the cell cycle (Koch and Schaechter 1962). This does not invalidate the notion

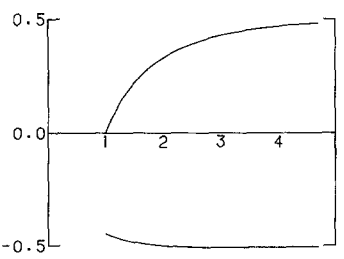

Fig. 3. Product-moment correlation coefficients. The *upper curve* is the sibling correlation coefficient calculated from Eq. (28), and the *lower curve* is the mother-daughter correlation coefficient calculation from Eq. (30)

of size control, however, since small heterogeneities in growth rate would tend to make both correlation coefficients more positive (Koch 1980).

8. Monte-Carlo simulations

A Monte-Carlo simulation of the tandem model can be implemented as follows. Choose an arbitrary initial mass distribution for a suitable large number, N, of cells. Calculate T_C for each cell, assuming exponential growth ($T_C = 0$ if birth mass ≥ 1). Next, generate N random numbers, T_A's, from an exponential distribution. Compute the division mass of each cell from its birth mass and its generation time ($T_A + T_B + T_C$), assuming exponential growth. Finally, imagine that each cell divides exactly in half, and retain only one daughter. This generates a new sample of N cells of known initial masses. The process can be iterated until the mass distribution and generation-time distribution settle down to their asymptotic forms. The results of such simulations are illustrated in Fig. 4, and it is seen that the analytic expressions for $\psi_*(T)$ and $\alpha(t)$ agree quite well with the Monte-Carlo results.

The Monte-Carlo simulation generates the asymptotic birth-mass distribution function as well as the generation-time distribution function. In Table 1, we compare the mean, variance and skewness of $\psi_*(x)$ with the mean, variance and skewness of the set of random numbers representing the birth masses of 1,000 cells of a Monte-Carlo simulation in the 20th generation. The mean, variance and skewness of $\psi_*(x)$ are

$$\text{mean} = \sigma(r-1)(r-2)^{-1}, \tag{31}$$

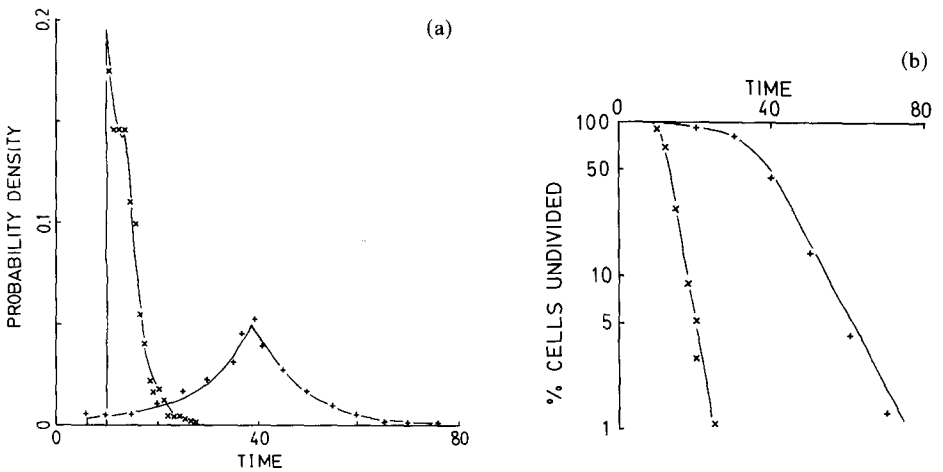

Fig. 4a, b. Generation-time distributions from Monte-Carlo simulations of the tandem model. The fundamental parameters, p, k and T_B, were the same used to model (×) the murine cell data and (+) the bacterial cell data in Fig. 2. In the Monte-Carlo simulation, $N = 1000$ cells, and the generation-time distribution was determined in the 25th generation after starting the population with all cells of unit mass. (**a**) Histograms of simulated cell generation times, and probability distribution functions calculated from Eq. (23). (**b**) Alpha-curves calculated from the same simulations and from Eq. (25)

Table 1. Mean, variance and skewness of the birth-mass distribution function, according to the analytical result and Monte-Carlo simulation[a]

No.	p	k	T_B	r	Mean	Variance	Skewness
1	0.580	0.0559	8.0	10.33	0.88 (0.88)	0.011 (0.012)	2.89 (2.88)
2	0.390	0.0408	13.0	6.93	1.02 (1.01)	0.045 (0.045)	3.85 (3.73)
3	0.180	0.0246	21.0	4.07	1.24 (1.22)	0.47 (0.38)	68.7 (5.9)
4	0.140	0.0224	23.0	2.29	3.73 (1.7↑)	∞ (1.8)	∞ (5.5)

[a] The Monte-Carlo simulations were performed as described in the text, with $N = 1000$. The mean, variance and skewness were calculated in the 20th generation after starting from a population of 1000 cells of unit mass. The Monte-Carlo results are given in parentheses after the theoretical values predicted by Eqs. (31)-(33). Simulations 1 and 2 had reached the stationary mass distribution by the 20th generation, but simulation 4 was not yet at a stationary state (the mean was still increasing at generation 20). The values chosen for p, k and T_B are those used by Tyson and Hannsgen (1981) to model some experimental cell populations.

$$\text{variance} \equiv \mu_2 = \sigma^2(r-1)(r-2)^{-2}(r-3)^{-1}, \tag{32}$$

$$\text{skewness} \equiv \mu_3/\mu_2^{3/2} = 2r(r-4)^{-1}[(r-3)(r-1)^{-1}]^{1/2}, \tag{33}$$

where $\mu_n = n$th moment of $\psi_*(x)$ about the mean. The results in Table 1 show that the analytic formulae agree excellently with the Monte-Carlo results as long as r is not too small. For r small, the higher moments will not agree because, e.g. if $r \leqslant 4$, $\mu_n = \infty$ for $n = 3, 4, \ldots$, yet any simulation will necessarily give a finite value for all these moments.

9. Theoretical and experimental distributions

In this paper we have been concerned with age and size distributions defined on samples of cells belonging to the same generation (from some distant ancestor). In a previous publication we looked at distributions defined on samples of cells that were all born contemporaneously (i.e. in some narrow time window). Neither of these sampling conventions, which are convenient from a theoretical point-of-view, conform very closely to experimentally measured distributions, for which the sampling convention is rarely defined explicitly but seems in most cases to consist of all cells in a time-lapse record for which adequate measurements can be made with some correction, perhaps, for truncation of the record. It would behoove experimentalists to pay more attention to the sample of cells on which their histograms are constructed. Two sampling conventions, due to Powell, seem particularly clear and appropriate for both theoretical and experimental purposes. First, one could commence a time-lapse record with just a few cells (preferably one) in the field of view and then use as one's sample all the progeny of these cells in, say, the fifth or sixth generation. This would give an adequate approximation to the distributions $\psi_*(x)$ and $f_*(T)$ discussed in this paper. Alternatively, one could use a time-lapse record with many hundreds of cells, choose a frame in the middle of the film, and define as one's sample all cells alive in this frame. This is a sample of "extant" cells, which can be used to define the distributions

$\psi_e(x)$ and $f_e(T)$. In particular (Powell, 1964; Painter and Marr, 1968),

$$f_e(T) = 2(1 - e^{-kT})f(T),$$

so $f(T)$, the steady-state probability density for generation time defined on a contemporaneous sample of newborn cells, can easily be derived from $f_e(T)$ and k, the specific growth rate of the cell population. Unfortunately, we know of no comparable expression relating $\psi_e(x)$ to $\psi(x)$, the birth-size distribution defined on a contemporaneous sample of newborn cells.

10. Conclusion

We have analyzed a model of the cell cycle that incorporates in a simple way the ideas of size control and random transitions as regulatory elements for progress toward division. We have (i) proved that there exists a globally asymptotically stable distribution of cell size at birth, (ii) given an analytic expression for the birth-size distribution, (iii) derived the generation-time distribution from the known birth-size distribution, and (iv) compared the generation-time distribution with experimental data. Our model is superior to both elementary models from which it is derived, in that it is able to account quite well for generation-time histograms, α-curves and β-curves. However, the model is incorrect in at least three regards: it predicts values for mother-daughter and sister-sister correlations which do not agree with experiment, and it predicts a birth-size distribution with an abrupt discontinuity at $x = \sigma =$ minimum birth size, which is far from the observed smooth, almost Gaussian, distribution of cell size at birth. These problems are traceable to the assumptions of (i) exponential growth of individual cells and (ii) a size-monitoring mechanism that prevents cells from entering A-phase until they reach *precisely* $x = 1$. We are currently investigating models for which cell growth is non-exponential and size-monitoring is imprecise.

Acknowledgments. We thank Robert Shields for graciously providing the experimental data in Fig. 2; he obtained the bacterial data from E. O. Powell. Paul Nurse drew our attention to the cell cycle model discussed in this paper. This work has been supported by grants from the National Science Foundation (MCS-8300559 and MCS-8301104) and the National Institutes of Health (1 RO1 GM27629).

Appendix: mean generation time

In Sect. 5 we noted that, for the tandem model, the steady-state mean generation time (over a sample of cells belonging to the same generation) is identical to the mass-doubling time of the population (assuming that individual cells grow exponentially). It is easy to show that this result holds quite generally for exponential cell growth and exact binary fission. Consider a cell chosen at random from a sample of same-generation cells. Let x be its birth mass, t its generation time, and y the birth mass of its daughters. Then $y = x\,e^{kt}/2$, or $\ln y = \ln x + kt - \ln 2$. Taking the expected value of both sides of this latter equality over a sample of same-generation cells, we have $\langle \ln y \rangle = \langle \ln x \rangle + k\langle t \rangle - \ln 2$. Under steady-state conditions $\langle \ln y \rangle$ must be the same as $\langle \ln x \rangle$, so $\langle t \rangle = \ln 2/k \equiv \tau$.

This proof can be generalized further. Lasota and Mackey (1984) have shown that

$$\phi_n(x) = \int_0^x [-\alpha_t(y, T(x, y))][V(x)]^{-1}\psi_n(y)\,dy, \tag{A1}$$

where

ψ_n = probability density for *birth* size in nth generation,

ϕ_n = probability density for *division* size in nth generation,

$V(x)$ = rate of growth of a cell of size x,

$T(x, y)$ = time needed for a cell to grow from size y to size x,

$\alpha(x, t)$ = probability that a cell has generation time $\geq t$, given that its birth size is x.

$\alpha_t(x, t) = \partial\alpha(x, t)/\partial t$,

If cells divide exactly in half, then $\psi_{n+1}(x) = 2\phi_n(2x)$ and (A1) becomes Eq. (21) in Lasota and Mackey. If cells divide into unequally sized daughters, then (Powell 1964)

$$\psi_{n+1}(x) = \int_0^1 p^{-1}K(p)\phi_n(x/p)\,dp, \tag{A2}$$

where $K(\cdot)$ is a probability density function for the ratio of daughter-size to mother-size. Obviously $K(\cdot)$ must be symmetric around $\frac{1}{2}$. If $K(\cdot)$ is a delta "function" at $\frac{1}{2}$, then $\psi_{n+1}(x) = 2\phi_n(2x)$, as expected. From (A1) and (A2) we have

$$\psi_{n+1}(x) = \int_0^1 \int_0^{x/p} p^{-1}K(p)[-\alpha_t(y, T(x/p, y))][V(x/p)]^{-1}\psi_n(y)\,dy\,dp. \tag{A3}$$

Let $\langle \ln x\rangle_n = \int_0^\infty \ln x\,\psi_n(x)\,dx$. Then, from (A3),

$$\langle \ln x\rangle_{n+1} = \langle \ln m(x, t)\rangle_n + \int_0^1 \ln p\,K(p)\,dp, \tag{A4}$$

where $m(x_0, t)$ is the solution of the initial value problem $dx/dt = V(x)$, $x(0) = x_0$, and

$$\langle \ln m(x, t)\rangle_n = \int_0^\infty \int_0^\infty \ln m(x, t)[-\alpha_t(x, t)]\psi_n(x)\,dx\,dt.$$

For exponential growth, $m(x, t) = x\,e^{kt}$ and

$$\langle \ln m(x, t)\rangle_n = \langle \ln x\rangle_n + k\langle t\rangle_n, \tag{A5}$$

where we have used the fact that the generation-time distribution is $f_n(t) = \int_0^\infty [-\alpha_t(x, t)]\psi_n(x)\,dx$.

In steady state, the expectations are independent of n so (A4) and (A5) imply that

$$k\langle t\rangle = -\int_0^1 \ln p\,K(p)\,dp. \tag{A6}$$

For $K(p)$ a delta "function" at $p = 1/2$, the right hand side of (A6) is simply $\ln 2$, so $\langle t\rangle = \ln 2/k = \tau$ as derived at the start. It is easy to show that, in general,

$$-\int_0^1 \ln p\,K(p)\,dp = \ln 2 + 2\mu_2 + R \tag{A7}$$

where μ_2 is the variance of $K(p)$ and the remainder is given by

$$R = \sum_{n=2}^{\infty} (2n)^{-1}2^{2n}\mu_{2n}$$

with $\mu_i = i$'th central moment of $K(p)$. We can estimate the size of the remainder by assuming that $K(p)$ is a uniform distribution with variance $= \mu_2$. Then

$$R = (36\mu_2^2/5)\sum_{n=0}^{\infty} r_n(12\mu_2)^n$$

with $r_n = 20[(2n+4)(2n+5)]^{-1}$. Obviously,

$$(36\mu_2^2/5) < R < (36\mu_2^2/5)(1-12\mu_2)^{-1}.$$

Thus, $R = 0(\mu_2^2)$, so to first-order in μ_2,

$$k\langle t\rangle = \ln 2 + 2\mu_2. \tag{A8}$$

We have derived (A8) by assuming that $K(p)$ is a narrow uniform distribution. A more reasonable form for $K(p)$ would be a narrow symmetric beta distribution, but one can show by a similar but lengthy argument that $R = O(\mu_2^2)$ for this case as well. For cells like budding yeast, which divide asymmetrically, it may be that $R = O(\mu_2)$ and that μ_2 is not particularly small. In such cases $k\langle t\rangle$ is not necessarily close to ln 2.

We can rewrite (A8) in terms of the coefficient of variation (CVP) of the distribution of $p \equiv$ daughter-size/mother-size;

$$k\langle t\rangle_{\text{gen}} \simeq \ln 2 + (\text{CVP})^2/2, \tag{A9}$$

where $\langle t\rangle_{\text{gen}}$ is the mean generation time on a sample of same-generation cells. Equation (A9) should be compared to (Painter and Marr 1967)

$$k\langle t\rangle_{\text{con}} \simeq \ln 2 + (\text{CVT})^2/4 \tag{A10}$$

where $\langle t\rangle_{\text{con}}$ is the mean generation time on a sample of contemporaneous newborn cells, and CVT is the coefficient of variation of generation times on such a sample. Typically, CVP $\simeq$ 5–10% whereas CVT $\simeq$ 10–20% so the difference between mean generation time and mass-doubling time ($\tau = \ln 2/k$ for exponential cell growth) is much less for same-generation samples than for contemporaneous samples.

For cell growth laws other than exponential there is no simple relation between $\langle t\rangle_{\text{gen}}$ and τ.

References

Diekmann, O., Heijmans, H. J. A. M., Thieme, H. R.: On the stability of the cell size distribution. J. Math. Biol. **19**, 227–248 (1984)

Fantes, P. A.: Control of cell size and cycle time in *Schizosaccharomyces pombe*. J. Cell Sci. **24**, 51–67 (1977)

Green, P. J.: A 'random transition' in the cell cycle? Nature **285**, 116 (1980)

Hannsgen, K. B., Tyson, J. J.: Stability of the steady-state size distribution in a model of cell growth and division. J. Math. Biol. **22**, 293–301 (1985)

Heijmans, H. J. A. M.: On the stable size distribution of populations reproducing by fission into two unequal parts. Math. Biosci. **72**, 19–50 (1984)

John, P. C. L.: The cell cycle. London: Cambridge University Press 1981

Koch, A. L.: Does the variability of the cell cycle result from one or many chance events? Nature **286**, 80–82 (1980)

Koch, A. L., Schaechter, M.: A model for statistics of the cell division process. J. Gen. Microbiol. **29**, 435–454 (1962)

Kubitschek, H. E.: Normal distribution of cell generation rate. Exp. Cell Res. **26**, 439–450 (1962)

Lasota, A., Yorke, J. A.: Exact dynamical systems and the Frobenius-Perron operator. Trans. Am. Math. Soc. **273**, 375–384 (1982)

Lasota, A., Mackey, M. C.: Globally asymptotic properties of proliferating cell populations. J. Math. Biol. **19**, 43–62 (1984)

Lord, P. G., Wheals, A. E.: Variability in individual cell cycles of *Saccharomyces cerevisiae*. J. Cell Sci. **50**, 361–376 (1981)

Minor, P. D., Smith, J. A.: Explanation of degree of correlation of sibling generation times in animal cells. Nature **248**, 241–243 (1974)

Mitchison, J. M.: The timing of cell cycle events. In: Little, M., Paweletz, N., Petzelt, C., Ponstingl, H., Schroeter, D., Zimmerman, H.-P. (eds.) Mitosis: Facts and questions, pp. 1–13. Berlin Heidelberg New York: Springer 1977

Nurse, P.: Cell cycle control—both deterministic and probabilistic? Nature **286**, 9–10 (1980)

Nurse, P., Streiblova, E.: The microbial cell cycle. Boca Raton, Florida: CRC Press 1984

Painter, P., Marr, A. G.: Inequality of mean interdivision time and doubling time. J. Gen. Microbiol. **48**, 155–159 (1967)

Painter, P. R., Marr, A. G.: Mathematics of microbial populations. Ann. Rev. Microbiol. **22**, 519–548 (1968)

Powell, E. O.: Growth rate and generation time of bacteria, with special reference to continuous culture. J. Gen. Microbiol. **15**, 492–511 (1956)

Powell, E. O.: A note on Koch and Schaechter's hypothesis about growth and fission of bacteria. J. Gen. Microbiol. **37**, 231–249 (1964)

Schaechter, M., Williamson, J. P., Hood, J. R., Koch, A. L.: Growth, cell and nuclear divisions in some bacteria. J. Gen. Microbiol. **29**, 421–434 (1962)

Shields, R.: Further evidence for a random transition in the cell cycle. Nature **273**, 755–758 (1978)

Shields, R.: Transition probability and the regulation of the cell cycle. In: Jiminez de Asua, L., Levi-Montalcini, R., Shields, R., Iacobelli, S. (eds.) Control mechanisms in animal cells, pp. 157–164. New York: Raven Press 1980

Shields, R., Smith, J. A.: Cells regulate their proliferation through alterations in transition probability. J. Cell. Physiol. **91**, 345–355 (1977)

Shields, R., Brooks, R. F., Riddle, P. N., Capellaro, D. F., Delia, D.: Cell size, cell cycle and transition probability in mouse fibroblasts. Cell **15**, 469–474 (1978)

Shilo, B., Shilo, V., Simchen, G.: Cell-cycle initiation in yeast follows first-order kinetics. Nature **264**, 767–770 (1976); see also Nature **267**, 648–649 (1977)

Smith, J. A., Martin, L.: Do cells cycle? Proc. Natl. Acad. Sci. USA **70**, 1263–1270 (1973)

Tyson, J. J., Hannsgen, K. B.: Analysis of a deterministic/probabilistic model of the cell division cycle. In: Rotenberg, M. (ed.) Biomathematics and cell kinetics, pp. 167–176. Amsterdam: Elsevier/North Holland 1981

Tyson, J. J., Hannsgen, K. B.: The distributions of cell size and generation time in a model of the cell cycle incorporating size control and random transitions. J. Theor. Biol. **113**, 29–62 (1985)

Received September 19, 1984/Revised February 26, 1985