Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 225 (2003) 275–283
Journal of Theoretical Biology
www.elsevier.com/locate/jtbi

The rescaling method for quantifying the turnover of cell populations ☆

Sergei S. Pilyugin[a,*], Vitaly V. Ganusov[b], Kaja Murali-Krishna[c,†], Rafi Ahmed[c], Rustom Antia[b]

[a] Department of Mathematics, University of Florida, Gainesville, FL 32611-8105, USA
[b] Department of Biology, Emory University, Atlanta, GA, 30322, USA
[c] Emory Vaccine Center, Emory University, Atlanta, GA, 30322, USA

Received 15 January 2003; received in revised form 28 April 2003; accepted 16 June 2003

Abstract

The dynamic nature of immune responses requires the development of appropriate experimental and theoretical tools to quantitatively estimate the division and death rates which determine the turnover of immune cells. A number of papers have used experimental data from BrdU and D-glucose labels together with a simple random birth–death model to quantify the turnover of immune cells focusing on HIV/SIV infections [Mohri et al. 279 (1998) 1223–1227, Hellerstein et al. 5 (1999) 83–89, Bonhoeffer et al. 164 (2000) 5049–5054, Mohri et al. 87 (2001) 1277–1287]. We show how uncertainties in the assumptions of the random birth–death model may lead to substantial errors in the parameters estimated. We then show how more accurate estimates can be obtained from the more recent CFSE data which allow to track the number of divisions each cell has undergone. Specifically, we: (i) describe a general stage-structured model of cell division where the probabilities of division and death are functions of time since the previous division; (ii) develop a rescaling method to identify invariant parameters (i.e. the ones that are independent of the specific functions describing division and death); (iii) show how these invariant parameters can be estimated, and (iv) illustrate this technique by applying it to CFSE data taken from the literature.

Keywords: Quantification of cell turnover; CFSE dye dilution experiment; Stage structured population model

1. Introduction

Understanding the quantitative aspects of cell turnover is a long standing theoretical problem. Specifically, we need a reliable analytical tool for estimating the rates of cell division and cell death that govern the rate of change in the total cell population. In immunology, for example, we would like to understand the mechanism of homeostatic regulation of immune memory which results in a nearly constant cell population. Does such a population consist of quiescent cells or there is a balanced turnover of cells? If turnover occurs, how do cells progress through the cell cycle so that division and death processes balance each other?

Accurate quantification of the dynamics of antigen-specific B- and T-lymphocytes in vivo has become possible recently (Slifka and Ahmed, 1998; Murali-Krishna et al., 1999). In addition, the development of CFSE dye dilution experiments has allowed for accurate tracking of the number of divisions that a given cell has undergone following transfer in vivo (Weston and Parish, 1990; Lyons and Parish, 1994; Lyons, 2000). When cells are stained with CFSE, this fluoresecent dye is included into the cell cytoplasm. In the process of cell division, the CFSE dye is diluted approximately equally between the two daughter cells. For in vivo transfer experiments, the CFSE assay allows to accurately track up to 5–10 successive generations of cells. Consequently,

☆ Abbreviations : CFSE, carboxyl fluorescein succimidyl ester; BrdU, bromodeoxyuridine; FLM, fraction of labeled mitoses; SM, Smith-Martin; CI, confidence interval

*Corresponding author. Tel.: 352-392-0281; fax: 352-392-8357.
E-mail addresses: pilyugin@math.ufl.edu (S.S. Pilyugin), vganuso@emory.edu (V.V. Ganusov), mkaja@u.washington.edu (K. Murali-Krishna), ra@microbio.emory.edu (R. Ahmed), rantia@biology.emory.edu (R. Antia).
† Present address: Department of Immunology, University of Washington, Seattle, WA, 98195, USA.

0022-5193/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0022-5193(03)00245-5

the CFSE technique provides a significantly more accurate description of cell turnover compared to the earlier techniques for measuring cell turnover such as BrdU, thymidine and D-glucose labelling (Mohri et al., 1998; Hellerstein et al., 1999; Bonhoeffer et al., 2000; Mohri et al., 2001; Debacq et al., 2002). There has been a widespread interest in using CFSE data to quantify the birth and death processes underlying the dynamics of immune responses.

A common way to study the quantitative aspects of cell turnover is to formulate a specific model for cell division and death and fit this model to the data. This approach has two serious shortcomings. First, in the absence of a biologically validated model it may be difficult to distinguish between the parameters estimated by using different models (see Section 2). Second, the data may be insufficient to unambiguously determine parameters of the model, that is, several parameter combinations may fit the data equally well.

To formulate a biologically reasonable specific model of cell division, one must analytically describe the mechanisms that govern cell division and cell death. For example, a very elegant quantitative model of cell turnover was formulated by Smith and Martin (1973) in their study of the fraction of labeled mitoses (FLM) curves in the cell culture. In the Smith–Martin (or simply SM) model, the progression of cells through the cell cycle involves a stochastic recruitment of cells from an A-state (approximately corresponding to G1 phase of the cell cycle) into a dividing B-phase (approximately equivalent to the S, G2, and M phases of the cycle). The dividing phase has a fixed duration Δ. The recruitment of cells from the A-state into the B-phase occurs at the fixed rate λ (the waiting time in the A-state has exponential distribution with the parameter λ). The two parameters λ and Δ provide a complete description of cell division (Cain and Chao, 1997a,b). Although the SM model could be viewed as a reasonable first approximation for the process of cell division, alternative models consistent with the FLM data have also been proposed (Castor, 1980; Brooks et al., 1980; Grasman, 1990).

In contrast with cell division, our understanding of processes that regulate cell death during the cell cycle is much poorer[1]. The SM model postulates that cell death occurs at constant rates d_A and d_B in the A-state and the B-phase, respectively. Therefore, the SM model provides a simple yet biologically reasonable description of cell death. Alternatively, cell death events may be restricted to a set of discrete checkpoints within the A-state and the B-phase. Therefore, it is crucial to understand the consequences of making an incorrect assumption about the death process.

In Section 2, we discuss some limitations of the random birth–death model commonly used in the analysis of BrdU and CFSE data (Mohri et al., 1998; Bonhoeffer et al., 2000; Mohri et al., 2001; Revy et al., 2001). We illustrate how different assumptions on the timing of death in the cell cycle (as described above) can drastically alter our estimates for the parameters describing cell division and cell death. One potential way to solve this problem might be to estimate the parameters of the SM model by fitting the SM model to the CFSE data. This approach has two limitations:

1. It assumes that the SM model is correct. As we discussed above, while the progression through the cell division cycle is biologically reasonable we do not know if the assumption that the death rates are constant during the A and B phases is correct.
2. The data on the total cell numbers and CFSE distributions is insufficient to estimate all four parameters (λ, Δ, and death rates d_A, d_B).

In this paper, we present a different approach to the problem of estimation of parameters describing cell division and death from the CFSE data. We propose a general model of the cell cycle which makes no specific assumptions about the timing of cell division and death except to say that the probability of division and death are functions of the time since the previous division. To analyze the general model, we develop a method of rescaling (Section 3) and show how this method allows us to identify and estimate kinetic parameters of the cell cycle that are *independent* of the specific mechanisms of cell division and cell death. In Section 4 we explain, with an example, how our method can be used. We will make available online a simple Mathematica program implementing our method for other datasets[2]. Finally, we go over the limitations of the current techniques in the Discussion section. All mathematical derivations are presented in Appendix A.

2. Illustrating the problem

To illustrate the consequences of making different specific assumptions about the death process, we consider two very simple models of the cell cycle. Both models are derived as the limiting cases of the Smith–Martin model with an infinitesimally short B-phase.[3] In both models, we let λ equal the rate of recruitment of

[1] The problem of understanding cell death in comparison with cell division is long standing (Monod, 1949) and stems in part from the fact that cells which have undergone division can be visualized, while dead cells rapidly disappear being eliminated by apoptosis.

[2] The Mathematica notebook implementing the rescaling method can be downloaded from http://www.math.ufl.edu/~pilyugin/Rescaling_method.nb

[3] We present the Smith–Martin model and the mathematical derivation of both limiting cases in the appendix.

cells from the A-state and let $T = 1/\lambda$ be the average division time (that is, the mean duration of the cell cycle). We let $x_n(t)$ denote the number of cells that have divided n times by time t. In the first model, we let $d = d_A = d_B$ so that the death rate is constant throughout the cell cycle. The equations for $x_n(t)$ are

$$\frac{dx_n(t)}{dt} = 2\lambda x_{n-1}(t) - (\lambda + d)x_n(t), \;\; x_0(0) = x_0. \tag{1}$$

This model is analogous to the widely used random birth-death model (Veiga-Fernandes et al., 2000; Bonhoeffer et al., 2000; Revy et al., 2001; Mohri et al., 2001) where the probability rates of both cell division and death are constant. The mathematical solution to Eq. (1) can be written as

$$x_n(t) = \underbrace{\frac{(2\lambda t)^n}{n!}e^{-2\lambda t}}_{\text{distribution}} \underbrace{\left[x_0 e^{(\lambda - d)t}\right]}_{\text{total \#}}. \tag{2}$$

The first model therefore predicts that the total number of cells changes exponentially at the rate $(\lambda - d)$. The distribution in the number of cells that have undergone n divisions by time t is Poisson, with the mean number of divisions that increases at the rate 2λ over time, independently of the death rate. In the second model, we let $d_A = 0$ so that there is no cell mortality in the A-state and assume that a certain fraction f of cells dies in the B-phase. Equations of the second model are

$$\frac{dx_n(t)}{dt} = 2\lambda(1-f)x_{n-1}(t) - \lambda x_n(t), \;\; x_0(0) = x_0. \tag{3}$$

The mathematical solution to Eq. (3) is

$$x_n(t) = \underbrace{\frac{(2\lambda(1-f)t)^n}{n!}e^{-2\lambda(1-f)t}}_{\text{distribution}} \underbrace{\left[x_0 e^{\lambda(1-2f)t}\right]}_{\text{total \#}}. \tag{4}$$

Thus in the second model, the total number of cells grows exponentially at the rate $\lambda(1-2f)$. The distribution in the number of cells that have undergone n divisions by time t is again Poisson, but the mean number of divisions now increases at the rate $2\lambda(1-f)$ over time.

We see that in both models the distribution in the number of cells that have undergone n divisions is Poisson. In the first model, the rate of increase in the mean number of divisions depends on λ whereas in the second model it depends on both λ and f. Consequently, if we were to estimate the mean generation time T using Eq. (1) or (3), we would obtain different estimates depending on the underlying model. For example, in case of a stable population ($\lambda = d$ in Eq. (1) and $f = 0.5$ in Eq. (3)) we would obtain a twofold discrepancy in the estimate for T. Such discrepancies will be even more pronounced for contracting populations and less pronounced for expanding populations. On the other hand, the estimates obtained from both models would be equally "significant" because both models produce equally good fits to the data.

This example illustrates that in the absence of further knowledge on the nature of the death process, we cannot estimate T with much confidence using only the data on the total number of cells and the number of divisions they have undergone (such as given by the CFSE data, for example). Given the uncertainties we have in the formulation of a specific model, we would like to define quantities that describe the cell turnover independently of the underlying model. In the following section, we do this by formulating a very general model of the cell cycle and determining which quantities we can estimate from the CFSE data.

3. General stage-structured model of cell cycle

3.1. Formulation

We postulate a set of relatively general biological assumptions about the nature of cell proliferation. We define cell *division* as an event when one mother cell leaves its generation and at the same time two identical daughter cells enter the next generation[4]. We define cell *death* as an event when one cell leaves its generation and no daugther cells are produced. Using this terminology, we assume that (i) The cycle of a given cell is terminated either by cell division or by cell death. (ii) Cell division and cell death are independent random events whose probabilities of occurrence depend only on the time since the cell was born. In particular, such probabilities are independent of a particular cell and a given generation. (iii) The probability that cell division and cell death occur simultaneously is negligibly small. (iv) The system is closed, that is, new cells enter the population only through division and cells leave the circulating pool only through death.

We let $x_n(t,s)$ denote the density of cells at time t that entered the n-th generation at time $t-s$. We refer to s as the age of cells inside the generation. We let $\lambda(s)$ denote the probability rate of cell division at age s and $d(s)$ denote the probability rate of cell death at age s inside the generation. The dynamics of the cell density within the n-th generation is described by the partial differential equation

$$\frac{\partial x_n(t,s)}{\partial t} + \frac{\partial x_n(t,s)}{\partial s} = -(\lambda(s) + d(s))x_n(t,s), \;\; n \geqslant 0. \tag{5}$$

The number of cells in the n-th generation that divide anywhere between the times t and $t + dt$ is given by

$$\left(\int_0^\infty \lambda(s)x_n(t,s)\, ds\right) dt,$$

[4] We refer to the cells which have undergone n divisions as cells in the n-th generation.

and therefore twice the number of cells enter the $(n+1)$-th generation between t and $t+\mathrm{d}t$. Consequently, the dynamics of two consecutive generations are coupled through the boundary condition

$$x_n(t,0) = 2\int_0^\infty \lambda(s)x_{n-1}(t,s)\,\mathrm{d}s,\ \ n\geqslant 1. \tag{6}$$

The set of Eqs. (5) and (6) constitutes the general stage-structured model of the cell cycle. Since the system is closed, no cells enter the 0-th generation, that is, $x_0(t,0)=0$.

The dynamics of the cell density is governed by a set of linear partial differential equations (5) where the boundary condition (6) describes the rate at which cells enter the n-th generation. Therefore, rescaling this rate by a factor of $a\geqslant 0$ will result in the identical rescaling of the cell density $x_n(t,s)$. Equivalently, the dynamics of the rescaled cell densities $x_n(t,s,a)=a^n x_n(t,s)$ must satisfy the equations

$$\frac{\partial x_n(t,s,a)}{\partial t} + \frac{\partial x_n(t,s,a)}{\partial s} = -(\lambda(s)+d(s))x_n(t,s,a), \tag{7}$$

$$x_n(t,0,a) = 2a\int_0^\infty \lambda(s)x_{n-1}(t,s,a)\,\mathrm{d}s. \tag{8}$$

The set of Eqs. (7) and (8) constitutes the *rescaled* model of the cell cycle.

The biological meaning of the rescaling is intuitively simple: Suppose that each mother cell produces not 2 but $2a$ daughter cells, then the cell density in each subsequent generation will be rescaled by a factor of a relative to the cell density in the preceding generation. In particular, the cell density in th n-th generation will be rescaled by a factor of a^n. Importantly, such rescaling does not affect the distribution of division or death events within a given generation, but only the rate of transfer of cells from one generation to the next.

In the next section, we show that for any nonnegative value of a, the rescaled model (7)–(8) describes an exponentially growing or decaying population. We also show how one can estimate parameters describing cell division and death from the relationship between the net proliferation rate $r(a)$ of the rescaled population and the value of a.

3.2. Method of rescaling

A typical set of experimental data (e.g., obtained from CFSE experiments) is presented as a time series for each generation of cells. We let $X_n(t)=\int_0^\infty x_n(t,s)\,\mathrm{d}s$ denote the total number of cells in the n-th generation at time t and consider the rescaled time series $X_n(t,a)=a^nX_n(t)$ with $a\geqslant 0$. In the previous section, we explained that the time series $X_n(t)$ corresponds to the original model (5)–(6) if and only if the rescaled time series $X_n(t,a)$ corresponds to the rescaled model (7)–(8). The total population sizes are given by

$$X(t)=\sum_{n=0}^\infty X_n(t),\quad X(t,a)=\sum_{n=0}^\infty a^nX_n(t),$$

respectively. Since Eqs. (5) and (6) are a special case of Eqs. (7) and (8) with $a=1$, we analyze the more general case $a\geqslant 0$.

According to Eqs. (7) and (8), the dynamics of the rescaled total cell density $x(t,s,a)=\sum_{n=0}^\infty a^nx_n(t,s)$ satisfies the classical von Foerster equation (Foerster, 1959),

$$\frac{\partial x(t,s,a)}{\partial t} + \frac{\partial x(t,s,a)}{\partial s} = -(\lambda(s)+d(s))x(t,s,a), \tag{9}$$

$$x(t,0,a) = 2a\int_0^\infty \lambda(s)x(t,s,a)\,\mathrm{d}s. \tag{10}$$

This model is also equivalent to the celebrated Lotka renewal equation (Sharpe and Lotka, 1911) widely used in demographical applications (Keyfitz, 1968, 1985). Model (9)–(10) is linear, and therefore it predicts an asymptotically exponential net proliferation of $X(t,a)$ at a rate $r(a)$ which will vary as we vary the value a (Bellman and Cooke, 1963). The underlying relationship between $r(a)$ and a is given by the characteristic equation of the rescaled model

$$1 = 2a\int_0^\infty \lambda(s)\mathrm{e}^{-\Lambda(s)-D(s)}\mathrm{e}^{-r(a)s}\,\mathrm{d}s, \tag{11}$$

where $\Lambda(s)=\int_0^s\lambda(z)\,\mathrm{d}z$ and $D(s)=\int_0^s d(z)\,\mathrm{d}z$ (Sharpe and Lotka, 1911) (We present the derivation of Eq. (11) in the appendix). Note that $r(a)$ is a strictly increasing function of a for $a\geqslant 0$.

In demographic applications, the functions $\Lambda(s)$ and $D(s)$ can be readily obtained from the census data (Keyfitz, 1985) which provides the age distribution of individuals within one generation. In our application, the absence of such information (i.e. the age distribution of cells within a single generation) is a major obstacle. Fortunately, the CFSE data provide the age distribution of cells according to their division numbers. The novelty and the advantage of the rescaling method is that it allows one to approximate the characteristic equation (11) only using the distribution of cells by division numbers. To approximate the characteristic equation from a given experimental time series $X_n(t)$, we generate a family of rescaled time series $X_n(t,a)$, calculate the change in the total population size $X(t,a)$ with time, and evaluate the exponential proliferation rate $r(a)$ for each value of a. As a result, we can obtain the function $r(a)$ by manipulating a single time series of CFSE data.

3.3. Estimation of intrinsic parameters

With no additional assumptions on model (5)–(6), we can estimate several intrinsic kinetic parameters of the cell cycle. Estimation of intrinsic parameters can be

performed without specific knowledge of functions $\lambda(s)$ and $d(s)$. Such parameters include:

1. the fraction of cells that die in one generation (δ);
2. the mean generation time of surviving cells (τ).

Suppose that a cohort of N_0 cells enters a given generation (simultaneously) at $t = 0$. Eq. (9) implies that the fraction of cells that have not divided and remain alive by time s is equal to $N(s) = N(0)\exp(-\Lambda(s) - D(s))$. Consequently, the number of cells in this cohort that will eventually divide is given by

$$N_1 = \int_0^\infty \lambda(s)N(s)\,ds = N_0 \int_0^\infty \lambda(s)e^{-\Lambda(s)-D(s)}\,ds.$$

Assuming that in the absence of death all cells eventually divide (that is, $\int_0^\infty \lambda(s)e^{-\Lambda(s)}\,ds = 1$), we find that the fraction of cells that die in one generation is given by $\delta = 1 - N_1/N_0$. Consequently,

$$\delta = 1 - \int_0^\infty \lambda(s)e^{-\Lambda(s)-D(s)}\,ds. \tag{12}$$

Similarly, the fraction of cells that survive through one generation (and therefore divide) is given by

$$1 - \delta = \int_0^\infty \lambda(s)e^{-\Lambda(s)-D(s)}\,ds.$$

Let a^* be the real root of r, that is, $r(a^*) = 0$. Eqs. (11) and (12) imply that

$$\delta = 1 - \frac{1}{2a^*}, \quad 1 - \delta = \frac{1}{2a^*}. \tag{13}$$

If not all cells die during one generation (i.e. $\delta < 1$), we can define the mean generation time for surviving cells as

$$\tau = \frac{1}{1-\delta}\int_0^\infty s\lambda(s)e^{-\Lambda(s)-D(s)}\,ds. \tag{14}$$

In the Appendix, we show that

$$\tau = \frac{1}{a^* r'(a^*)}. \tag{15}$$

According to Eqs. (13) and (15), the quantities δ and τ can be estimated from a given experimental time series by first finding a^* as the a-intercept of the graph $r = r(a)$ and then finding the slope to this graph at $a = a^*$.

In addition to the mean generation time for surviving cells τ, we can estimate the variance of τ, that is, the variance of generation times for surviving cells. Such variance is mathematically expressed as

$$\sigma_\tau^2 = \frac{1}{1-\delta}\int_0^\infty s^2\lambda(s)e^{-\Lambda(s)-D(s)}\,ds - \tau^2.$$

In the appendix, we show that σ_τ^2 is given by

$$\sigma_\tau^2 = \tau^2\left(1 + (a^*)^2 r''(a^*)\tau\right). \tag{16}$$

4. Illustration of the rescaling method

In this section, we analyze the time series describing the dynamics of CFSE labeled P-14 transgenic naive CD8 T cells after adoptive transfer into irradiated hosts (Murali-Krishna and Ahmed, 2000) (Fig. 1). The original data show that during the first week following transfer, the cells grow exponentially ($r \approx 0.28\ \text{day}^{-1}$) and the mean number of divisions they have undergone increases approximately linearly. The CFSE data is best represented as the table of the number of cells having undergone n divisions at time t (Table 1).

To estimate the parameters δ and τ we rescale the data by a, where a varies from 0 upwards. We do so by multiplying the number of cells in the n-th generation by a^n; for example, if $a = 2$ then the number of cells in the n-th generation is multiplied by 2^n. Using Table 1 for day 3 and $a = 2$, the rescaled numbers will be $(1.77\ 10^4)\cdot 2^0$, $(6.10\ 10^4)\cdot 2^1$, $(6.58\ 10^4)\cdot 2^2$, $(1.28\ 10^4)\cdot 2^3$ and $0.0\cdot 2^4$. After rescaling the entire Table 1 with a given a, we sum the cells in all generations and estimate the exponential growth rate $r(a)$ for the corresponding value of a (see Fig. 1). Plotting $r(a)$ as a function of a, we estimate a^* as the a-intercept and $r'(a^*)$ as the slope of $r(a)$ at a^*. Then we use the resulting values and Eqs. (13) and (15) to estimate δ and τ, respectively.

We applied the rescaling method to the data on the dynamics of Tg CD8 T cells during proliferation in lymphopenic hosts to estimate the relationship between a and $r(a)$. For a list of values a (from $a = 0.2$ to 1), we fitted the logarithmic time series $\ln(X(t,a))$ with a linear function and used the resulting slope as the estimate for $r(a)$ (Fig. 1). The estimated graph of a vs. $r(a)$ is shown by a bold solid line in Fig. 2. The thin dashed lines represent the 67% confidence intervals for $r(a)$ which were calculated as the standard errors for $r(a)$ for each value of a.

The value a^* at which the rate of exponential increase is zero is 0.549. Thus we estimated $a^* = 0.549$ with 67% confidence intervals $(0.546, 0.551)$ (these values are a-intercepts of 67% CI for $r(a)$, i.e., values at which the

Table 1
The dynamics of P-14 Tg CD8 T cells after adoptive transfer into irradiated hosts (Murali-Krishna and Ahmed, 2000). The numbers below equal the number of cells per spleen divided by 10^4

	t (days)			
$x_n(t)$	**0.5**	**1.25**	**3**	**8**
0	7.38	7.07	1.77	0.0
1	0.0	0.64	6.10	0.29
2	0.0	0.0	6.58	5.71
3	0.0	0.0	1.28	19.97
4	0.0	0.0	0.0	18.83
5	0.0	0.0	0.0	7.99
6+	0.0	0.0	0.0	4.00

n stands for generation.

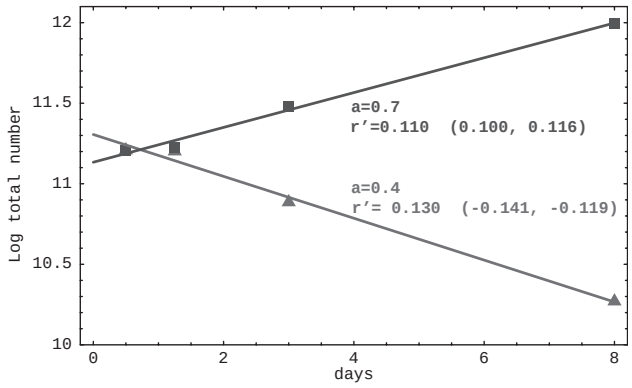

Fig. 1. The change in the total population size of the rescaled population (with $a = 0.4$ and 0.7) with time using the data in Table 1. The regression lines for the total size increase are shown (r' is the mean slope; the 67% confidence intervals for the slope are in parantheses).

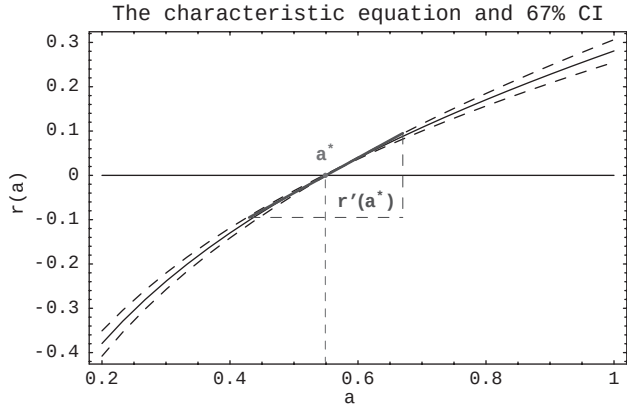

Fig. 2. The plot of a vs. $r(a)$ obtained from the Tg cells data (see Table 1). The bold solid line is the graph $r = r(a)$ obtained from the data. The 67% confidence intervals for $r(a)$ are represented by the thin dashed lines. From this graph, we estimated the average fraction of surviving cells as $1 - \delta = 0.91$ and the mean generation time of surviving cells as $\tau = 2.33$ days.

dashed curves in Fig. 2 intersect the a-axis). Consequently, we estimate that $\delta = 0.089$ with the same small error ($CI = (0.085, 0.092)$). Thus, approximately 9% of the cells die on average during a single cell cycle in the given experiment.

We estimated the slope of the function $r(a)$ at $a^* = 0.549$ using a five-point difference scheme (from $a^* - 2\Delta a$ to $a^* + 2\Delta a$ with a step $\Delta a = 0.02$; other small steps gave approximately the same result) and found that $r'(a^*) = 0.782$. To obtain the confidence intervals for $r'(a^*)$, we again used the five-point difference scheme but now for 67% CI of $r(a)$ to calculate the slopes $r'_+(a^*)$ and $r'_-(a^*)$. We found the 67% confidence intervals for $r'(a^*)$ to be: $(0.733, 0.831)$. Finally, we estimated the mean generation time of surviving cells $\tau = 1/a^* r'(a^*)$ as $\tau = 2.33$ days. The 67% confidence intervals for τ were $(2.18, 2.50)$.

Based on the CFSE time series presented in Table 1, we conclude that following transfer of CD8 T cells into irradiated hosts approximately 9% die per cell cycle, and the average division time for surviving cells is approximately 2.3 days.

5. Discussion

We have shown how the existing experimental data on the dynamics of cell populations and the number of divisions they have undergone can give us a quantitative picture of the underlying processes of cell division and cell death. We have argued that the problem is harder than previously thought (Gett and Hodgkin, 2000; Veiga-Fernandes et al., 2000; Revy et al., 2001). The reason is that making accurate estimates of parameters requires having a biologically reasonable quantitative model of both cell division and cell death. While the features of cell division can be quantitatively described (at least to a first approximation) much less is known about cell death. We have shown that choosing different, biologically plausible scenarios for cell death can lead to large inaccuracies in the estimation of the parameters of *both* cell division and cell death.

As the next step, we have investigated a general class of stage structured models, where the division and death of a cell are random variables that depend only on the time since cell division. For this class of models, we have introduced the method of rescaling, identified parameters that are independent of the particular mechanisms of cell division and cell death, and shown how they can be estimated from the CFSE data. These parameters are the fraction of cells which die per cell cycle (δ) and the average time that surviving cells take to complete the cell cycle (τ). We should, however, note that these are not the only invariant parameters, but have been chosen because they have a clear biological interpretation. Other parameters such as the variance of generation times for surviving cells σ_τ^2 can also be obtained from the characteristic equation $r = r(a)$.

Due to the linearity of model (9)–(10), the total population size $X(t)$ should reach an exponential phase after a short transient (Bellman and Cooke, 1963). The rescaling method is based on the assumption that the population is in the exponential phase, and therefore it can be applied only to the data in which the total population size changes exponentially (i.e., increases, declines, or stays constant). Using the CFSE data we can calculate the exponential rate of change of the total population size $r(a)$ if each cell divides not in 2 but in $2a$ daughter cells for each value of a by rescaling the CFSE data. The value of a at which the rescaled population size does not change (a^*), the slope and curvature of the experimental curve $r(a)$ at this point ($r'(a^*)$ and $r''(a^*)$) allow us to estimate δ, τ, and σ_τ^2.

The rescaling method allows us to estimate the invariant parameters without making any specific assumptions about how and when cells divide and die. Once a more detailed mechanistic model of cell division and death is validated then the parameters of such a model could be estimated. However, our present lack of information about cell death precludes formulating such a model.

5.1. Limitations of the rescaling method

The rescaling method is considerably more general than existing methods to analyze cell turnover. Nevertheless, it has several important limitations:

1. The CFSE assay can only detect up to a maximum of 5–10 divisions (Lyons, 2000) (after this, the CFSE dye becomes so dilute that it cannot be detected above the background) and this causes truncation errors. The truncation errors result from excluding cell counts with higher division numbers from the total population. It is therefore important that the CFSE measurements are obtained prior to such truncations having occurred.
2. In formulating the general model of cell turnover and the rescaling method we assume that the cell turnover is independent of both time and number of divisions. This may not be always valid, for example, as in the case of programmed immune responses (Kaech and Ahmed, 2001; van Stipdonk et al., 2001).
3. The rescaling method evaluates the function $r(a)$ assuming that the total population size changes exponentially and disregards the transient effects.
4. The rescaling method cannot be used to analyse the data obtained with such commonly used labels of cell division as BrdU or *D*-glucose because these labels do not allow the quantification of the number of divisions that cells have undergone. Thus it is easier to point out the limitations of the earlier models for the analysis of BrdU data than to suggest ways in which the analysis can be improved.
5. The rescaling method described can only be applied to a homogenous population of cells. Some of the problems associated with measurement of turnover of heterogenous populations of cells using BrdU labeling have been elegantly considered by Asquith et al. (2002).

5.2. Existing analysis of CFSE data

Prior to this work, the CFSE data has been interpreted using two different approaches. For completeness, we describe them here. One approach is based on a purely statistical description of the data. Such description includes the dynamics of the mean number of divisions for surviving cells, the variance in the number of divisions for surviving cells, and the changes in the total number of precursors (obtained by dividing the number of cells undergone n divisions by 2^n) (Gett and Hodgkin, 2000). The model of cell division and death used by Gett and Hodgkin (2000) may serve as an alternative to the model presented in this paper. Gett and Hodgkin analysed the CFSE data assuming that once a cell commits to division, its progeny continues to divide in equal intervals of time. They concluded that the rate at which cells commit to division follows a normal curve. These two models can be distinguished by estimating the variance in the number of divisions for surviving cells. The general model of the cell cycle predicts a continuous linear increase in the variance while the model of Gett and Hodgkin predicts that after a transient increase this variance will remain nearly constant. Another approach is to fit the random birth-death model to the CFSE data (Veiga-Fernandes et al., 2000; Revy et al., 2001). This approach includes the estimation of the division and death rates (assuming that they are constant). We have demonstrated that the parameter estimation performed in this fashion largely depends on the specific underlying model of cell division and death. Discrepancies in the underlying model often lead to large inaccuracies in the resulting estimates. In particular, we note that broad inferences such as that of Revy et al. (2001) suggesting that "relative sizes of the CFSE peaks do not depend on the death rate and can be used to unambiguously determine the mean division time", are not generally valid.

We have emphasized our lack of a quantitative understanding of cell death and explained why the CFSE data alone cannot provide further insights into the regulation of cell death during the cell cycle. We hope that this theory will be instrumental in designing direct experimental studies that would elucidate such mechanisms of regulation. For instance, in vitro FLM studies of immune cell populations in expansion and contraction phases might give us a better understanding of where cell death occurs in the cell cycle. Such studies might allow us to generate a biologically valid explicit model for cell death and stimulate the development of new quantitative methods for estimating specific parameters of the model from the data.

Acknowledgements

We thank Rob DeBoer for many useful discussions. We are grateful to the anonymous referees for their valuable comments. The research of R.A. and V.G. was supported by grant NIH R01 AI 49334.

Appendix A. Mathematical derivations

A.1. Two limiting cases of Smith–Martin equations

The original SM model (Smith and Martin, 1973) is formulated as a set of differential equations

$$\frac{dA_n(t)}{dt} = 2b_{n-1}(t,\Delta) - (\lambda + d_A)A_n(t), \tag{A.1}$$

$$\frac{\partial b_n(t,s)}{\partial t} + \frac{\partial b_n(t,s)}{\partial s} = -d_b b_n(t,s),\ 0<s<\Delta, b_n(t,0) = \lambda A_n(t), \tag{A.2}$$

where n is the number of divisions that a given cell has undergone by time t. $A_n(t)$ is the number of cells in the A-state and $b_n(t,s)$ is the fraction of cells that entered the B-phase at time $t-s$. The amount of time that cells spend in the A-state is exponentially distributed with parameter λ. All cells spend a constant amount of time (Δ) in the B-phase. The mean generation time (i.e. the mean duration of the cell cycle) is given by $T = \frac{1}{\lambda} + \Delta$. All cells divide when they reach the end of the B-phase and their daughter cells immediately re-enter the A-state of the next generation. The total number of cells in the B-phase is given by $B_n(t) = \int_0^{\Delta} b_n(t,s)\,ds$. The death rates in both states are constant and denoted by d_A and d_B. One can reduce the SM model by expressing $b_n(t,s)$ in terms of $A_n(t)$ as

$$b_n(t,s) = \lambda e^{-d_B s} A_n(t-s),\ 0<s<\Delta, \tag{A.3}$$

and replace Eqs. (A.1) and (A.2) by

$$\frac{dA_n(t)}{dt} = 2\lambda e^{-d_B\Delta} A_{n-1}(t-\Delta) - (\lambda + d_A)A_n(t), \tag{A.4}$$

$$B_n(t) = \lambda \int_0^{\Delta} e^{-d_B s} A_n(t-s)\,ds. \tag{A.5}$$

In Eq. (A.4), the factor $e^{-d_B\Delta}$ is the fraction of cells that survive in the B-phase. The mean generation time defined as the average duration of the cell cycle is $T = 1/\lambda + \Delta$.

In Section 2, we present two limiting cases of model (A.4)–(A.5) as Δ becomes infinitesimally small. Letting $x_n(t) = A_n(t) + B_n(t)$, we observe that due to Eq. (A.5), in the limiting case $\Delta \to 0$ we also have $A_n(t) \to x_n(t)$ and $A_{n-1}(t-\Delta) \to x_{n-1}(t)$.

In the first limiting case, we let $d = d_A = d_B$ so that the death rate is constant throughout the cell cycle. In the limit $\Delta \to 0$, Eq. (A.4) is replaced by

$$\frac{dx_n(t)}{dt} = 2\lambda x_{n-1}(t) - (\lambda + d)x_n(t),\ x_0(0) = x_0. \tag{A.6}$$

The mathematical solution to Eq. (A.6) can be written as

$$x_n(t) = \frac{(2\lambda t)^n}{n!} e^{-2\lambda t}[x_0 e^{(\lambda - d)t}]. \tag{A.7}$$

In the second limiting case, we let $d_A = 0$ so that there is no cell mortality in the A-state and assume that $e^{-d_B\Delta} \to 1 - f$ as $\Delta \to 0$. Here f is a limiting fraction of cells that die in the B-phase. This model is described by

$$\frac{dx_n(t)}{dt} = 2\lambda(1-f)x_{n-1}(t) - \lambda x_n(t),\ x_0(0) = x_0. \tag{A.8}$$

The mathematical solution to Eq. (3) is

$$x_n(t) = \frac{(2\lambda(1-f)t)^n}{n!} e^{-2\lambda(1-f)t}\left[x_0 e^{\lambda(1-2f)t}\right]. \tag{A.9}$$

A.2. The characteristic equation of the general model

To derive the characteristic equation of Eqs. (9)–(10), we substitute $x(t,s,a) = y(s,a)e^{r(a)t}$ into Eq. (9). Since the partial derivatives of $x(t,s,a)$ are given by

$$\frac{\partial x}{\partial t}(t,s,a) = r(a)y(s,a)e^{r(a)t}, \quad \frac{\partial x}{\partial s}(t,s,a) = \frac{\partial y}{\partial s}(s,a)e^{r(a)t},$$

$y(s,a)$ is a solution of

$$\frac{\partial y}{\partial s}(s,a) = -(\lambda(s) + d(s) + r(a))y(s,a),\ s \geqslant 0.$$

Solving this equation for y, we find that

$$y(s,a) = y(0,a)e^{-\Lambda(s) - D(s)}e^{-r(a)s}, \tag{A.10}$$

where $\Lambda(s) = \int_0^s \lambda(z)\,dz$ and $D(s) = \int_0^s d(z)\,dz$. Substituting Eq. (A.10) into Eq. (10), we obtain the equation

$$e^{r(a)t}y(0,a) = 2ae^{r(a)t}y(0,a)\int_0^{\infty} \lambda(s)e^{-\Lambda(s)-D(s)}e^{-r(a)s}\,ds.$$

Since $e^{r(a)t}y(0,a) \neq 0$, we cancel this quantity on both sides of the above equation to obtain the final form of the characteristic equation

$$1 = 2a\int_0^{\infty} \lambda(s)e^{-\Lambda(s)-D(s)}e^{-r(a)s}\,ds. \tag{A.11}$$

A.3. The expressions for τ and σ_τ^2

To derive the expression for τ, we differentiate Eq. (A.11) with respect to a to obtain

$$0 = 2\int_0^{\infty} \lambda(s)e^{-\Lambda(s)-D(s)}e^{-r(a)s}\,ds - 2ar'(a)\int_0^{\infty} s\lambda(s)e^{-\Lambda(s)-D(s)}e^{-r(a)s}\,ds. \tag{A.12}$$

We substitute $a = a^*$ and $r(a^*) = 0$ into Eq. (A.12) and obtain

$$0 = \frac{1}{a^*} - 2a^*r'(a^*)(1-\delta)\tau.$$

Using the fact that $2a^*(1-\delta) = 1$, we find that

$$\tau = \frac{1}{a^*r'(a^*)}. \tag{A.13}$$

To derive the expression for σ_τ^2, we differentiate Eq. (A.12) with respect to a to obtain

$$0 = -2r'(a)\int_0^\infty s\lambda(s)e^{-\Lambda(s)-D(s)}e^{-r(a)s}\,ds + 2a(r'(a))^2\int_0^\infty s^2\lambda(s)e^{-\Lambda(s)-D(s)}\,e^{-r(a)s}\,ds - 2(r'(a)+ar''(a))\int_0^\infty s\lambda(s)e^{-\Lambda(s)-D(s)}e^{-r(a)s}\,ds. \tag{A.14}$$

We substitute $a = a^*$, $r(a^*) = 0$, and Eq. (A.13) into Eq. (A.14) and find that

$$\frac{1}{1-\delta}\int_0^\infty s^2\lambda(s)e^{-\Lambda(s)-D(s)}\,ds = \tau^2(2+(a^*)^2r''(a^*)\tau).$$

Finally, we compute

$$\sigma_\tau^2 = \frac{1}{1-\delta}\int_0^\infty s^2\lambda(s)e^{-\Lambda(s)-D(s)}\,ds - \tau^2 = \tau^2(1+(a^*)^2r''(a^*)\tau). \tag{A.15}$$

References

Asquith, B., Debacq, C., Macallan, D.C., Willems, L., Bangham, C.R., 2002. Lymphocyte kinetics: the interpretation of labelling data. Trends Immunol. 23 (12), 596–601.

Bellman, R., Cooke, K.L., 1963. Differential-Difference Equations. Academic Press, New York.

Bonhoeffer, S., Mohri, H., Ho, D., Perelson, A.S., 2000. Quantification of cell turnover kinetics using 5-bromo-2'-deoxyuridine. J. Immunol. 164, 5049–5054.

Brooks, R.F., Bennett, D.C., Smith, J.A., 1980. Mammalian cell cycles need two random transitions. Cell 19 (2), 493–504.

Cain, S.J., Chao, P.S., 1997a. Transition probability cell cycle model. Part 1: balanced growth. J. Theor. Biol. 185, 55–67 DOI:10.1006/jtbi.1996.0289.

Cain, S.J., Chao, P.S., 1997b. Transition probability cell cycle model. Part 2: non-balanced growth. J. Theor. Biol. 185, 69–79 DOI:10.1006/jtbi.1996.0290.

Castor, L.N., 1980. A G1 rate model accounts for cell-cycle kinetics attributed to 'transition probability'. Nature 287, 857–859.

Debacq, C., Asquith, B., Kerkhofs, P., Portetelle, D., Burny, A., Kettmann, R., Willems, L., 2002. Increased cell proliferation, but not reduced cell death, induces lymphocytosis in bovine leukemia virus-infected sheep. Proc. Natl. Acad. Sci. U.S.A 99, 10048–10053.

Foerster, H.v., 1959. Some remarks on changing populations. In: Stohlman Jr., F. (Ed.), The Kinetics of Cellular Proliferation. Grune and Stratton, New York, pp. 328–407.

Gett, A.V., Hodgkin, P.D., 2000. A cellular calculus for signal integration by T cells. Nat. Immunol. 1, 239–244.

Grasman, J., 1990. A deterministic model of the cell cycle. Bull. Math. Biol. 52 (4), 535–547.

Hellerstein, M., Hanley, M.B., Cesar, D., Siler, S., Papageorgopolous, C., Weider, E., Schmidt, D., Hoh, R., Neese, R., Macallan, D., Deeks, S., McCune, J.M., 1999. Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. Nat. Med. 5 (1), 83–89 (see comments).

Kaech, S.M., Ahmed, R., 2001. Memory CD8 T cells differentiation: initial antigen encounter triggers a developmental program in naive cells. Nat. Immunol. 2, 415–421.

Keyfitz, N., 1968. Introduction to the Mathematics of Population. Addison-Wesley, Reading, MA.

Keyfitz, N., 1985. Applied Matehmatical Demography. Springer, New York.

Lyons, A.B., 2000. Analyzing cell division in vivo and in vitro using flow cytometric measurement of CFSE dye dilution. J. Immunol. Methods. 243, 147–154 DOI:10.1016/S0022-1759(00)00231-3.

Lyons, A.B., Parish, C.R., 1994. Determination of lymphocyte division by flow cytometry. J. Immunol. Methods 171, 131–137 DOI:10.1016/0022-1759(94)90326-4.

Mohri, H., Bonhoeffer, S., Monard, S., Perelson, A.S., Ho, D., 1998. Rapid turnover of T lymphocytes in SIV-infected rhesus macaques. Nature 279, 1223–1227.

Mohri, H., Perelson, A.S., Tung, K., Ribeiro, R.M., Ramratnam, B., Markowitz, M., Kost, R., Hurley, A., Weinberger, L., Cesar, D., Hellerstein, M.K., Ho, D., 2001. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. J. Exp. Med. 194 (9), 1277–1287.

Monod, J., 1949. The growth of bacterial cultures. Ann. Rev. Microbiol. 3, 371–394.

Murali-Krishna, K., Ahmed, R., 2000. Naive T cells masquerading as memory cells. J. Immunol. 165 (4), 1733–1737.

Murali-Krishna, K., Lau, L.L., Sambhara, S., Lemonnier, F., Altman, J., Ahmed, R., 1999. Persistence of memory CD8 T cells in MHC class I-deficient mice. Science 286, 1377–1381.

Revy, P., Sospedra, M., Barbour, B., Trautmann, A., 2001. Functional antigen-independent synapses formed between T cells and dendritic cells. Nat. Immunol. 2 (10), 925–931.

Sharpe, F.R., Lotka, A.J., 1911. A problem in age-distribution. Philos. Mag. 21 (6), 435–438.

Slifka, M.K., Ahmed, R., 1998. Long-lived plasma cells: a mechanism for maintaining persistent antibody production. Curr. Opin. Immunol. 10 (3), 252–258.

Smith, J.A., Martin, L., 1973. Do cells cycle? Proc. Natl. Acad. Sci. U.S.A 85, 411–426.

van Stipdonk, M.J., Lemmens, E.E., Schoenberger, S.P., 2001. Naive CTLs require a single brief period of antigenic stimulation for clonal expansion and differentiation. Nat. Immunol. 2, 423–429.

Veiga-Fernandes, H., Walter, U., Bourgeois, C., McLean, A., Rocha, B., 2000. Response of naive and memory CD8+ T cells to antigen stimulation in vivo. Nat. Immunol. 1 (1), 47–53.

Weston, S.A., Parish, C.R., 1990. New fluorescent dyes for lymphocyte migration studies. Analysis by flow cytometry and fluorescence microscopy. J. Immunol. Methods 133, 87–97 DOI:10.1016/0022-1759(90)90322-M.