Clonal aspects of plant cell proliferation and their applications to animal cells and bacteria

R. W. Korn

Department of Biology, Bellarmine University, Louisville, KY, USA

Received 19 June 2007; revision accepted 13 October 2007

Abstract. *Objectives:* Extensive mathematical studies have been made on cell clone development but little has been advanced in the mathematics of small clone formation and virtually no actual data of small clone size has been collected. *Materials and methods:* Small clone sizes in leaf marginal cells of the aquatic plant *Elodea* and aleurone spot sizes in the grain of *Zea* were counted for later statistical analyses of mean, variance and probability distribution frequencies. *Results:* Simple mathematical models were developed and their calculated results are comparable to data collected on actual plant clones. The parameters in these models were original cell size (s_0), growth rate (T), duration of growth (t) and cell division inequality (i). *Conclusions:* Given T and t , the critical parameter is s_0 . Plant tissue is ideal material to collect data on clone development because growth rate is uniform across a tissue and cells remain in place, so clone size can be measured, unlike microbes and animal cells that have neither feature. In the light of the results, traditional methods for calculating cell cycle duration and mutation rate are questioned. The applications of these plant features to studies on animal cell populations are discussed.

INTRODUCTION

A fundamental feature of microbes and multicellular organisms is cell proliferation, which requires a combination of cell growth and cell division. Each proliferating cell gives rise to a number of descendent cells that at any time constitutes a clone. In the past, it has been rare for a clone to have developmental significance (Korn 2007) but experimentally it plays a critical role in determining DNA mutation rate. The classic study of Luria & Delbruck (1943) established the time, measured in generations, when mutations occur in cultures of *Escherichia coli*. They found the size of mutant clones varied enormously, including jackpot cases of many mutant cells, a feature that could only be explained by mutation occurring at any time, or generation, before a selective agent was introduced. Later, Luria (1951) found r^- mutant viruses in single lysis appear to be in clones, starting at different times, as numbers followed the log–log distribution expected of a geometric series. In both studies, the primary interest was determining the generations when mutations occur and not the actual size of a mutant clone, although they claimed that this could be resolved.

Correspondence: R. W. Korn, Department of Biology, Bellarmine University, 2001 Newburg Road, Louisville, KY 40205, USA. Tel.: 502-452-8200; Fax: 502-452-8096; E-mail: rkorn@bellarmine.edu

At the other end of clonal history is the status of the initial cell. Cell division is usually unequal, which leads to asynchronous proliferation (Prescott 1955). Rarely does cell proliferation occur by synchronous cell division. In plants, the basis for non-synchronous growth is unequal cell division. Here, the average difference in size of a pair of daughter cells is 8% for elongating cells, or a 4% deviation from the average (Korn 1993) and 11%, or 5.5% off equality, for two-dimensionally growing cells (Korn 2001). It was also found that the larger of two daughter cells inevitably divides before the smaller daughter, 19/19 cases observed, suggesting some compensating mechanism exists that keeps cell size within a tissue-specific range (Korn 1993), an idea first suggested by Fanes (1977) in yeast. In plants, four variables are important in determining clone size, c , during development, (1) cell division inequality, i , (2) initial cell size cell, s_0 , (3) growth rate or cell cycle time, T , and (4) duration of growth, t .

Plant tissues are sheets or blocks of cells cemented together thus growth rate, T , is uniform throughout a tissue, unlike more complex cases reported in microbes and animal tissues where cells in a clone may have different growth rates (Rubinow 1968), experience death (Hughes 1955) or are temporarily delayed (Powell 1955). These three features are unlikely in plant tissue because tissues are rarely sheared from variable growth rate or cells found to be dead. Both Kelly & Rahn (1932) using bacteria and Prescott (1959) in the protozoan *Tetrahymena* have shown that cell cycle duration, T , follows what Rubinow (1968) saw as a modified binomial, or gamma, distribution, similar to what Powell (1955) saw as the more general Pearson type III distribution.

In this study, clonal data were collected from two plant tissues, leaf marginal cells in the aquatic plant *Elodea* for distance between consecutive barb cells, and the aleurone layer in maize for spot size, both measured in cell number. The intent here is to inspect the relationship between unequal cell division, initial cell size, cell cycle duration and clonal age in the production of a clone in plants and to compare these results with those from animal studies.

MATERIALS AND METHODS

Elodea densa Caspary was purchased from Pets Palace, Louisville, KY, USA, and kernels of *Zea mays* L. of line 5071 with three doses of the *Ac* allele were obtained from the Maize Genetics Stock Center in Urbana, IL, USA. Grains were soaked in warm water for 2 h, pericarp was peeled off, and cell number of spots counted under $\times 50$ magnification with a Wild dissecting microscope (Fig. 1d). Computer programs were written in Visual Basic 4.0 (Microsoft, Puget Sound, WA, USA) and graphs were prepared with Sigma Plot (SPSS Inc., Chicago, IL, USA).

RESULTS

Microscope observations

Barb cells are somewhat evenly spaced along the edge of an *Elodea* leaf (Fig. 1a). Each barb cell and basipetal unspecialized cells appear to constitute a clone (Fig. 1b). This can be seen by inspecting the embryonic region where all stages of a clone are found, first a bulbous pre-barb cell, next a bulbous cell dividing, followed by both of these daughter cells again dividing, and so on (Fig. 1c). The number of cells in a clone ranges from two to eight, including the terminal barb cell, with an average of 4.22 cells, and a variance of 1.77 (Table 1, 21 I and *Elodea* data). Average length of 15 non-barb daughter cells at maturity is 78.7 μm and a variance of 198.3 with a difference

Table 1. Probability distributions for clones at different ages (iterations) derived from Equations (1) and (2) with a cell cycle (10 iterations) in order to show half cycle values and by a Monte Carlo computer model

[illegible]

Figure 1. Plant tissues. (a) *Elodea* leaf with marginal, spaced barb cells. (b) Marginal clone bracketed of *Elodea* leaf; barb cell at tip of clone with three unspecialized cells below. (c) Meristematic cells of clone developing from left to right (arrow). Numbers refer to number of cells in a developing clone. (d) Maize kernel with aleurone spots and added reference square for doing the fluctuation test (Table 4, column 3). (e) Aleurone spots with cells visible, clones of one cell, aa, two cells, bb, and nine cells, cc.

between sisters of $4.42 \mu\text{m}$, or $2.21/39.35$, 5.6% from the average. Only clones in the top two-thirds of the leaf were scored as those near the bottom and at the end of leaf development may come from a slowing rate of growth.

Duration of growth from the time, a barb cell is determined until it ceases in the region of elongation of the leaf appears to be constant as initial barb cells arise at the same place in the embryonic region, in the region of eight cells (position 8) from the base of the leaf embryonic region, and clones cease division when passing from the embryonic region, some 33 cells from the base (position 33). Number of cell cycles to make a clone is $\log(33/8)/\log 2$, or 2.04 generations ($8 \rightarrow 16 \rightarrow 32$). This number, based on 10 cases, bodes well as an independent calculation uses the average of 4.22 cells per clone that would require $\log 4.22/\log 2$, or 2.08 cell cycles.

Furthermore, growth rate during this time appears to be constant as tissues neither pucker nor tear from regional differences in growth rate, and cell sizes seem to be uniform throughout the leaf, indicating cell cycle duration is not innately variable. The size or age of a cell moving into position 8 in the embryonic region when differentiation begins varies from S to $2S$ so that the final size of the clone after duration equal to 2.04 generations of proliferation varies accordingly. If differentiation of a barb cell had to occur at a specified time t_x , namely age, or size s_x , in the cell cycle and also at position 8, then most cells would pass through position 8 without differentiating. Hence, to have all cells at position 8 differentiate, neither age nor size is important for determining differentiation. With duration of growth, t , and growth rate, T , constant during clone formation, the only parameter remaining is the size, s_0 , of a cell at the time it specializes as a bulbous barb cell. The parameter s_0 includes that of unequal division, l , as the former depends directly on the latter.

Red spots on the kernels of maize were first identified by Rhoades (1938) and later were shown to be from transposition, or gene jumping (McClintock 1948) within the aleurone tissue, a single-celled layer where all cells can be directly observed (Federhoff 1984). Red spots noticeably vary in size (Fig. 1d,e) and are composed of red vacuolated cells arranged somewhat in orthogonal fashion. The number of deeply red cells was counted, excluding marginal pink cells whose colouration was due to leakage from true red cells (Walbot, personal communication), numbers ranging from the most common value of one to as many as 455, the latter determined by twice-repeated counting and then averaging. The most common size was one-celled with larger spots decreasing in frequency (Figs 1e and 2c) according to an expected geometric distribution. Placing of some sizes within geometric distribution groups, that is, clones of one cell, clones of two cells, clones of four cells, and so on, was difficult, as a clone of six cells, would it belong to the group with the peak at four or those with the peak at eight? To overcome this difficulty $\log_2$ of cell number per clone was plotted against the log of the frequency of all spots with the value for cell number or higher; a linear relationship was found (Fig. 2d) similar to that described by Luria (1951). This was plotted for the number of mutant phages in single lysis. Clearly, number of cells per spot followed an extensive geometric progression with all values at and between doubling steps. Transpositions have been postulated to occur during chromosome duplication (Federhoff 1984) and occur at about the same time in the cell cycle, that is, in S phase. Spot size was then not due to variable size of the initial cell, s_p , at the time of change as in *Elodea* but only to the duration of proliferation time, t . That the expected number of one-celled clones was found indicated that phenotypic lag as noted in animal cells (Rudd *et al.* 1990), did not occur and all mutations were visible immediately.

Mathematical model of clone development

Most investigations on clone development assume variation occurs during development, by variable growth rate, cell death and duration of growth, but there are advantages to determine how much influence the state of an initial cell alone may have on the size of the subsequent clone. For example, a cell may differentiate to begin clone formation at any time in its history, namely, anywhere in the cell cycle, as was found in *Elodea*. Thus, cell size is a function of passing from S to $2S$ with frequencies from $2Y$ to Y , that is, a doubly truncated negative exponential distribution (Fig. 2a, a-a, blue lines) (Korn 2001). The average of a truncated distribution is $1.414S$ and its variance is $1.06S$. Cells growing at rate r over S iterations of growth will double in size and hence the number of cells after i iterations is $2^{i/S}$. For example, if an initial cell is originally 10 units in size and growth rate r is 0.0717 (size doubles over 10 iterations, $1.0717^{10} = 2$), then after 20 and 25 iterations of growth there will be $2^{20/10}$, or 4, and $2^{25/10}$, or 5.6, cells, respectively. This is only a first approximation as other factors enter into the problem.

Figure 2. Data graphs. (a) Cell size distributions starting with doubly truncated distribution, a-a, blue; shown as vertical bar graph, b₁-b₆, green; values converted into normal distributions, c₁-c₆, black, which are summed to give the final cell size distribution, d, olive; and compared to Monte Carlo data, e, red. (b) *Elodea* data compared to data generated by Equation (2). (c) Maize clones sizes and expected from computer model. (d) Linear correlation between log size of aleurone clone in cells against log of number of such clones or greater in size.

Cells divide unequally to have daughter cells on an average $2a$ different in size from each other. Each new cohort of daughter cells by age and size (Fig. 2a, b₁, green line) forms a normal distribution for cells of various sizes (Fig. 2a, c₁, black lines). Symmetrical normal distribution is generated as for every daughter a units smaller than average its sister cell is larger by a , and smaller deviations from equality are expected to be more frequent than larger deviations. As noted for *Elodea*, the average mature cell is $78.7 \mu\text{m}$ with a size difference between sister cells of $4.42 \mu\text{m}$ for a $2.21/39.35$, or 5.6% deviation from expected size from equal division. As cells of each cohort grow exponentially, the cohort separates from others exponentially (Fig. 2a green lines) but remains a normal distribution (Fig. 2a, c₂-c₆, black lines). For purposes considered later continuous normal distribution of daughter cell sizes can be transformed into discrete binomial distribution.

Distribution frequency of cell sizes (s) is a joint distribution from two primary distributions, a truncated exponential distribution (Fig. 2a, a-a, blue lines) and a set of symmetrical binomial distributions (Fig. 2a, c₁-c₆, black lines) to give Equation (1) calculated earlier (Korn 2001) to provide the following cumbersome distribution,

$$f(s) = \sum ap_g q_g^{s_g-1} \binom{n-j}{x} p_{cd}^x q_{cd}^{n-j-x}$$

n is given
 $s_{cd} = x - \text{Int}(n/2)$
 if n is odd then $j = 1$
 if n is even then $j = 0$
 $s = s_g + s_{cd}$

(1)

where s is the final consideration of cell size, s_g is cell size by perfect equality of daughter cell sizes and s_{cd} is the deviation from unequal division. The first expression on the left, $ap_g q_g^{s_g-1}$, is the doubly truncated distribution of frequency of cell sizes, $f(s_g)$ (Fig. 2a, a), and the second term on the left, is the binomial probability distribution of daughter cell sizes deviant from average, $f(s_{cd})$ (Fig. 2a, c), and $f(s)$ is the frequency of final cell size (Fig. 2a, d). For example, when s_g ranges from five to ten units and deviations from unequal division can be as great as three units, the frequency of daughter cells 6.0 units in size is the sum of $(f(5_g) + f(1_d)) + ((f(6_g) + f(0_d) + (f(7_g) + f(-1_d)) + ((f(8_g) + (f(-2_d)) + (f(9_g) + f(-3_d))$, or $f(s) = (f(s_g) + f(s_{md}))$ where s equals s_g plus s_d . The data generated by Equation (1) provide a similar distribution to that found for cell sizes of *Elodea* clone cells (Fig. 2a, e, red lines). When (a) values of $ap_g q_g^{s_g-1}$ are transformed into those for a geometric distribution (b) the average remains the same as the binomial distribution is symmetrical and (c) the maximum value of s is defined, then Equation (1) can be given as a binomial distribution.

$$f(s) = \binom{h-d}{s-d} p^{s-d} q^{(h-d)-(s-d)}$$

$d = S - n/2$
 $h = 2S + n/2 - d$

(2)

With the mean size still $1.414S$ and maximum size is $2S + \text{int}(n/2)$ the value of p in the binomial distribution is the fraction of mean over the maximum value (Meyer 1965). Given the fixed number of trials, h , and p is known, a new binomial distribution can be determine with d subtracted from h to bring the smallest value on the abscissa down to zero. When values of Equation (2) are displaced to the right to give a mean of $64 \mu\text{m}$ then they fit the *Elodea* data by a χ^2 analysis, with a P -value of >0.25 (Fig. 2b). Hence, Equation (2) is a binomial expression displaced to the right.

Variance for clone size can be worked out to be $0.272e^{(i/(S+a))-2}$. The value 0.272, variance for clone size, after one-half of a cell cycle of unequal divisions, is reasoned as follows. The doubly truncated distribution beginning with $2Y$ at X and ending with Y at $2X$ (Fig. 2a, a, blue lines) has a cumulative value of 1.0. The younger half of the distribution, from X to $1.5X$ is $-e^{-1.5r} + e^{-1r}$, by integration totals 0.585786437626905 , and the older half is $-e^{-2r} + e^{-1.5r}$, or integrates as 0.414213562373095 , with r the decay constant of $\ln 2$ or 0.69. Given a population of initial cells of different sizes, according to Equation (2) and Fig. 2(a), blue lines, then 41.4% of cells divide after half a cycle and all have divided after one cycle. After half a cycle, this 41.4% is the number of two-celled clones and thus 58.6% of clones are still one-celled. Average clone size after one-half of a cell cycle is then $2 \times 0.414 + 1 \times 0.586$, or 1.41 cells. Variance is then $0.414 \times (2-1.414)^2 + 0.586 \times (1-1.414)^2$ or 0.272 after one-half of a cell cycle. Because the number of cycles is $i/(S+a)$, variance is $e^{(i/(S+a))-2}$ where a is a cell cycle correction factor to be explained shortly. The value of -2 is included as a correction factor because the value $0.272e^{(i/S)}$

is not added for the first two cell cycles of proliferation. As variance arises independently at each half of the cell cycle after two cell cycles, it is added at each half cycle after this time (Table 1, column 3).

Each cell will then proliferate into a clone over time, t , or iterations of growth, i , with an initially small cell, one recently birthed, forming a relatively small clone or a large initial cell, one ready to divide, becomes a relatively large clone. A clone of c cells is then

$$c = \text{Int}(s_0(1 + r)^i/h) \quad (3)$$

where h is the number of iterations per cell cycle. For example, a cell 12 units in size (s_0) with a growth rate of 0.0717/iteration (r), over 100 iterations (i) with 10 iterations per cycle (h) becomes

$$\text{Int}(12(1.0717)^{100}/10) = 1220 \text{ cells}$$

and the frequency of clones composed of c cells is then

$$f(c) = f(s:n) \quad (4)$$

where $f(s:n)$ is the frequency of any cell that become a clone of c cells.

To match this mathematical procedure with the *Elodea* data, sizes of the initial cells of populations follows the distribution in Equation (2) and was found to grow for 21 iterations by Equations (3) and (4) to form clones with the average, variance and distribution of various number of cells (Table 1).

To test these results a Monte Carlo computer program was written that includes keeping track of sizes of initial cells and all subsequent cells from 5 to 24 units in size; growth rate is 0.07177/iteration, unequal division value is average of 19 random numbers in order to give a deviation of 5.6% from average as in *Elodea*, and cell division occurs after cells reach at least 20 units. These computer data are also included in Table 1 and compare well with the mathematical calculations.

An important consequence of unequal cell division is cell cycle duration. For example, if two daughter cells from a mother cell of size $2S$ deviate from a 0.5 : 0.5 ratio by a , then their sizes are $S - a$ and $S + a$. Given exponential growth with rate r , then the iterations of growth, i_{S-a} and i_{S+a} for the daughters to reach size $2S$ are

$$i_{S-a} = (\log 2S - \log (S - a))/\log(1 + r) \quad i_{S+a} = (\log 2S - \log(S + a))/\log(1 + r) \quad (5)$$

and since $\log (S - a) + \log (S + a) \neq 2 \log(S)$ the short cycle is shorter than the long cycle is longer from the average to further uncouple cell age and size. For example, if a mother cell of 20 units divides into daughters with sizes 10.56 and 9.45 units, the average difference of 5.6% from each other as earlier referenced for *Elodea*, the daughters will have cycle durations in iterations of growth of

$$(\log 20 - \log 10.56)/\log 1.0717 = 9.21 \quad (\log 20 - \log 9.44)/\log 1.0717 = 10.83.$$

which provide an average of 10.04 iterations per cell cycle, 0.04 more than from an equal division. This deviation from the average, 0.04, is the correction factor, a , for calculating average cell number which is now $2^{i/(10+a)}$, and the variance cited earlier becomes $0.272e^{(i/(S+a))-2}$ (see Table 1, column 2). Prescott (1959) qualitatively stated that with unequal division the 'generation time

can be more easily lengthened than it can be shortened'. This lengthening of generation time can be seen in Table 1, column 4, from the computer model for number of cells after 10, 20 and 30 iterations that would give 1, 2 and 3 cell generations or 2.0, 4.0 and 8.0 cells but instead give less, 1.99, 3.96 and 7.89 cells.

It was stated earlier that unequal cell divisions will generate a normal distribution of daughter cell sizes (s_{cd}).

$$f(s_d) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(m-s_d)^2/2\sigma^2} \quad (6)$$

and the cycle duration (cd) from Equation (5) is

$$cd = \log(2S) - \log(s_d)/\log(1+r).$$

Solving for s_d

$$s_d = \frac{2S}{(1+r)^{cd}} \quad (7)$$

Although the distribution of cell cycle durations is a reciprocal expression of initial cell size, it is a lognormal distribution, a normal distribution stretched out to the right, as cd is an exponent (Equation (7)) such that logarithmic values of cd's generate normal distribution (Fig. 3c). A summary of mean and variance from the two parameters of interest, initial cell size and duration of growth, is given in Table 2.

Table 2. Statistics for different stages and types of clones

Conditions	Cell number/clone		cv
	Mean	Variance	
1. After half a cell cycle ($S/2$)			
A. Initial cells of same size	1.0	0	0
B. Initial cells of different sizes	1.414	0.272	0.36
2. After many iterations of growth (I)			
A. Initial cells of same size	$i/(S+a)$	$0.272e^{(i/(S+a))-2}$	23
B. Initial cells of different sizes.	$i/(S+a)$	$0.272e^{(i/(S+a))-2}$	23

Mathematical model of mixed clones

Any single large clone developed above can be divided into smaller subclones and these can be identified either by cellular differentiation as in the case of *Elodea* or mutational transposition in maize, or typical mutations in general. The most important statistic is variance, as Luria & Delbruck (1943) used it to determine when mutations would arise during cell culture growth. A further mathematical model can be developed here to provide the mean and variance for all subclones collectively within a population when knowing the size (s) and frequency ($f(s)$) of all subclones. For example, the histories of many of sixteen-celled clones coming from initial cells include subclones in a ratio of 16 of one-celled subclones to 8 two-celled subclones to 4

Figure 3. Sixteen-celled clones. (a). Determinate clone of five levels (L) with a total of 31 cells (2^{L-1}). (b). An indeterminate clone with total levels of four to seven. Cell cycle duration as a function of initial cell size. (c). Logarithmic transformation of cell cycle duration to give a normal distribution indicating cell cycle duration is a lognormal distribution. Influence of sampling size. (d). Calculated frequency of cultures without mutant cells (heavy line) and mutation rate (light line). The average number of mutant cells per sample is held constant for different sample fractions.

subclones of four cells to 2 subclones of eight cells to 1 clone of sixteen cells for a total of 31 possible clones (Fig. 3a). Here there is a final level L , as $1 + \log(s_{\max})/\log(2)$, with $s_{\max}$ as 16 so L is 5.0. The final level is composed of 2^{L-1} cells, each level is composed of 2^{L-1} cells, and the total cells during the history of a population is $2^L - 1$. The frequency of each clone size is $2^{L-1}/(2^L - 1)$. Individual level reference (1) can be replaced with L since

$$L \cdot 2^{L-1} = \sum 2^{L-l} \cdot 2^{l-1}$$

The mean number of cells for all clones during the history of a proliferative hierarchy is

$$m = \sum s \cdot f(s) = \sum_{l=1}^L 2^{L-l} \cdot \frac{2^{l-1}}{2^L - 1} = \frac{L \cdot 2^L - 1}{2^L - 1} \quad (8)$$

The variance is $\sum s^2 \cdot f(s) - (\sum s \cdot f(s))^2$ (Meyer, 1965) and here becomes

$$\begin{aligned}
v &= \sum_{l=1}^L (2^{L-l})^2 \cdot \frac{2^{l-1}}{2^L - 1} - \left(\sum_{l=1}^L 2^{L-l} \cdot \frac{2^{l-1}}{2^L - 1} \right)^2 \\
&= \frac{\sum_{l=1}^L 2^{L-l} 2^{L-l} \cdot 2^{l-1}}{2^L - 1} - \left(\frac{2^{L-1}}{2^L - 1} \right)^2 \\
&= \frac{\sum_{l=1}^L 2^{L-l} \cdot 2^{L-1} \cdot 2}{2^{L-l} \cdot 2^{l-1}} - m^2 \\
&= 2^{L-1} - m^2
\end{aligned} \tag{9}$$

Values for the mean and variance by the long-hand method give the same results as by Equations (8) and (9). When final cell number (n) is not in the geometric series, then L is $1 + \log(n)/\log(2)$ where n is the largest clone counted in cell number (31 in Fig. 3b). When using L in this case for Equations (8) and (9), it still gives values for the mean and variance of within 8% of those by the correct long-hand method (Fig. 3b).

A Monte Carlo computer model was constructed for mixed clones with an initial cell proliferating into 1220 cells. Growth rate was adjusted for doubling according to Equation (6) or slightly more than 10.0 iterations per cell cycle, an average of 5.6% difference between daughter cell size (average of 19 random numbers), and a mutation rate of 1.00E-01 for both daughter cells. The data profile comes close to the actual data from red spots of maize aleurone with peaks at 1, 2, 4, 8 and near 16 cells (Fig. 2c). The sample size here is 149 clones that is small and gives different mean and variance than derived by Equations (8) and (9). This suggests that the data set collected by Luria & Delbruck (1943) is undersized.

The mutation rate was set at 1.00E-02, relatively high, in order to follow development of early arising, large subclones, also, mutation can occur in both daughter cells. Equations (8) and (9) are deterministic, and hence mutation can occur only when there are at least 100 cells in a clone, an undesirable condition if jackpot cases are to occur. Consequently, a stochastic feature was added so a mutation could occur at any generation, although at low frequency. A run of a single clone gives iteration (age), cell number, mutations, mutant cells and mutant clone sizes (Table 3). The added stochastic feature provides a good example of a jackpot clone with 90 mutant cells in a population of only 1220 cells. Mutation rate can be determined directly. There were 18 mutations in this run, of the final 1220 cells, for a total of $1220 \times 2 - 2$, or 2438 DNA helices less the total of 391 mutant cells times 2.0, for a total of 1710 cells. The mutation rate is then $18/1710$, or 1.05E-02, close to the known rate of 1.00E-02. These 18 clones averaged 2.86 mutant cells with a variance of 891.

The Luria & Delbruck's (1943) fluctuation test results can be compared to the computer model and data from maize aleurone spots (Table 4, square in Fig. 1d). When rates are adjusted to offer some computer trials with no mutation, $p_o > 0$, similar profiles are produced to those of maize and bacteria. The data set of Luria & Delbruck (1943) of the number of mutant cells is based on a 1/8 sample of cultures and, because bacteria move around, it is also from a mixture of clones. To simulate these two conditions, that is mixing and then sampling, the computer model is run 1000 times and a random sampling of 1/8 of the population generated gave a mean of 4.2 mutant cells, a variance of 24.4 and a mutation rate of 2.8E-02 (Table 4). When the computer population size was adjusted to present an average of 4.2 mutant cells without sampling, the variance was 47.5 and mutation rate was 0.95E-02, both values different from those from the 1/8 dilution. Two effects of dilution are decreased mutation rate and increased variance (Fig. 3d). Clearly, the 1/8 sampling procedure decreases the calculated mutation rate by a factor close to 3.0.

Table 3. History of one population of mixed mutant clones (from computer model)

Time (iterations)	Number of cells	New mutations	Number of mutant cells	Size of clones in cells (in cell number)
5	1	0	0	
10	2	0	0	
15	3	0	0	
20	4	0	0	
25	6	0	0	
30	9	0	0	
35	13	0	0	
40	19	1	1	1
45	27	0	2	2
50	38	0	2	2
55	54	1	5	4, 1
60	76	1	5	5, 2, 1
65	108	0	8	8, 2, 2
70	152	1	12	11, 4, 2, 1
75	216	1	18	16, 5, 4, 2, 1
80	305	2	41	22, 8, 5, 2, 1
85	431	0	61	32, 11, 8, 4, 2, 2, 2
90	610	4	89	45, 16, 11, 5, 4, 2, 2, 1, 1, 1, 1
95	863	3	134	62, 22, 16, 8, 5, 4, 4, 2(4), 1(3)
100	1220	4	381	90 (jackpot), 32, 22, 11, 8, 5, 2(7), 1(5)

Table 4. Mutant subclones frequencies of computer program, maize aleurone and data from Lucia and Delbruck (1943)

Number	Number of special cells		
	Model mutant cells	Maize aleurone square of 269 cells	Luria and Delbruck's data
1	4	1	1
2	4	0	0
3	2	3	3
4	3	0	0
5	1	4	0
6	0	1	5
7	0	0	0
8	1	3	5
9	7	0	0
10	2	15	6
11	1	168 jackpot	107 jackpot
12	4	0	0
13	2	31	0
14	2	0	0
15	6	0	1
16	2	0	0
17	1	4	0
18	2	1	64 jackpot
19	0	1	0
20	3	0	33 jackpot
mean	2.3	12.0	11.35
variance	99.7	1476.5	694
mutation rate	3.9E-02*	4.75E-03	1.1E-08

*Calculated by Lederberg's method while the model uses 1.0E-02.

Besides the fluctuation test to calculate mutation rate, the continuous growth method (Novick & Szilard 1950) also gives excellent results. This method determines the increase in frequency of mutant cells over a known number of generations, or $\mu = f/d$ where μ is the mutation rate, f is the frequency of mutant cells and d is the number of cell cycles during the experiment (Araten *et al.* 2005). This equation can be simplified for plants as

$$\mu = f/(N_T/2) \quad (10)$$

where f is the frequency of one-celled clones and $N_T/2$ is half the final size of the population, or the number of new cells. In the computer model above (Table 3), the final number of cells (100th iteration) of 1220 has 18 mutant clones and at earlier cell generation (90th iteration), there are 610 cells with 11 mutant clones, indicating a mutation rate of $(18-11)/(1220-610)$, or 0.011 close to the value of 0.010 used in the model. In the corn kernel, there is an average of 5.0 one-celled clones from 10 samples of a 1-mm² area (square in Fig. 1d) of about 2700 cells for an average transpositional rate of $5/(2700/2)$, or 0.0037 per cell.

DISCUSSION

The use of cell size as the parameter to study the kinetics of cell proliferation was first evaluated in protozoa (Popoff 1909) and only later in plants (Korn 1993, 2001). Kendall (1948) considered that a cell passes through a series of probabilistic steps, the sum of which follows a Pearson type III distribution, a modified exponential distribution, for duration of the cell cycle. Stuart & Merkel (1948) invoked a maturation process while Rubinow (1968) considered a variety of maturation rates and Scherbaum & Rasch (1957) emphasized age as the critical parameter. Of these hypotheses, Rubinow (1968) found his variable maturation hypothesis fitted Prescott's empirical data best. Stuart and Merkle's hypothesis predicts a symmetrical distribution of ages that clearly is not supported by the asymmetric data sets from many cases examined (Kelly & Rahn 1932; Powell 1955; Prescott 1959; Korn 1993, 2001). Smith & Martin (1973) proposed a two-phase model with one expressing an exponential distribution of duration and so of cycle age. It was shown there that age is not a good measure as unequal division produces daughter cells of different sizes but are the same age. In addition, that the larger daughter cell divides before the smaller daughter cell (Korn 2001) indicates that size and not age is the better measure of cell advancement towards division. Alexandersson (2001) recently used the size-dependent cell cycle feature in her model.

The use of plant tissue for inspecting proliferation kinetics has several simplifications (Korn 1993, 2001) over animal cells and microbes. First, all cells must grow at the same rate otherwise tissue would tear from shear stress, and such rupturing has never been found (Fig. 1a). Second, division rate is uniform across a tissue otherwise some regions would have either larger or smaller cells than are found elsewhere in the organ, a feature also never encountered. Third, the parameter of cell size in a monolayer of cells adherent to each other is easily and precisely determined (Fig. 1b). Fourth, the relative size of daughter cells is also easily determined (Fig. 1b), (Korn 1993). These features are noted best in elongating epidermal cells (Korn 1993) but can be seen in isodiametric cells as well (Korn 2001). Furthermore, clone size judged by cell number is measurable because clones are spatially separated from each other unlike unicellular organisms and animal tissue where cells of various clones intermingle. Finally, valuable data of clone size were collected on *Elodea* barb clones and maize aleurone red spot clones, and these can be

used to check the validity of data derived from equations and computer models. It is rather surprising how little data (Kelly & Rahn 1932; Powell 1955; Prescott 1959) have been collected on which so much theory has been developed (Kendall 1948; Stuart & Merkel 1948; Powell 1955; Painter & Mary 1968; Rubinow 1968).

The source of all variation in plant cell size and clone size is unequal division and this feature has four consequences: (i) uncoupling cell size from age, (ii) instigating asynchronous cell proliferation, (iii) cell cycle durations follow a lognormal distribution and (iv) formation of a skewed distribution of cell sizes in a growing cell population. First, age and size are uncoupled by unequal cell division thus leading to variable generation times. As was shown through Equation (5) unequal division leads to different durations for each daughter cell before they divide and the average duration of two daughters is greater than that from equal division. It is initial size and not progressive ageing that determines cell cycle duration.

The second consequence of unequal cell division is asynchronous cell division. Perfect asynchronous division is theoretically impossible because many paired daughters are about the same size and so will divide at about the same time later, but most divisions are noticeably unequal and over several generations a size distribution is found as for cells of *Elodea* in Fig. 2b and Equation (2), to give mean and variance values of 78.7 μm and 198.3, respectively. The Monte Carlo computer model was run beginning with one cell and proliferation continued until the average cell size and variance were closest to 78.7 μm and 198.3, namely, 81.4 μm and 182.5, respectively. These approximations are reached after 4.1 cell cycles when there were $2^{4.1}$, or 17 cells. The status of a clone can then be completely synchronous, partially synchronous, or nearly asynchronous. By this criterion, cells of barb clones in *Elodea* are partially synchronous because clones are much less than 17 cells. So too are most of the cloned red spots in maize, however, a few are over 17 cells in size thereby attaining asynchrony.

The third consequence of unequal cell division is that it creates a variation of initial cell sizes that generate a lognormal distribution of cell cycle durations. The collection of normal distributions of initial cell sizes by Equation (5) generates a symmetric distribution, skewed to the right, that appears to be a lognormal type as its logarithmic values generate a symmetric distribution (Fig. 3c). Cell cycle duration (cd) is derived from daughter cell size, s_d , as cd equals $\log(2S - \log(s - a)/\log(1 + r))$, an equation that generates lognormal distribution (Fig. 3C). The distribution of cell cycle times was first interpreted by Kendall (1948) as following the general Pearson type III distribution that includes a specific gamma distribution, the latter then adopted later by others (Rubinow 1968; Kepler & Oprea 2001; Zheng 2005). Kepler & Oprea (2001) suggest several trees of cell lineages are possible. For example, cell A can divide into cells B and C which, in turn, divide into four cells D–G. Cell A can divide into cells B and C, but cell B divides into cells D and E and then cell D divides into cells E and F. This sequence is unlikely because cell C would grow for two generations and become so large as to be immediately recognized. In addition, the range of cell cycle duration is not found to be S to $4S$. In general, exponential distribution of cell sizes would create a range of cell sizes so great that smallest cells would hardly have space for a nucleus and chloroplasts and the largest cells would collapse under tissue pressure. The proposed range of cell cycle durations as generated by Equation (5) allows for small extensions below S and is less complex than the assumptions required for the gamma distribution; more importantly, it makes more biological sense as it is derived directly from observed data of sister cell sizes. Feller (1940) warned that growth data from some biological examples that give logistic growth expression also fit a variety of equations with different assumptions. Painter & Mary (1968) also noted that just because a model generates a data set that fits experimental data, the goodness of fit does not validate the model but only invites further studies. Here, a gamma distribution is one such example for range of cell cycle duration,

but its assumptions in requiring many probabilistic steps have no apparent biological basis. The transition probability that gives exponential distribution proposed by Smith & Martin (1973) also does not fit the observed distribution of cycle times. Kimmel *et al.* (1984) proposed that observed variation in cell cycle duration of cultured Chinese hamster cells is determined by unequal cell division resulting in unequal distribution of some RNAs. Equations (6) and (7) developed here for plant cells quantify these findings in animal cells.

Finally, continuous distribution of cell sizes in asynchronous dividing populations is skewed to the right and appears to be described by an asymmetric binomial distribution. The expression

$$\binom{h-j}{s-j}$$

is the number of sequences to get a cell of size s and these sequences are different from initial sizes plus the growth period. For example, cell of size 6.0 can come from an initial cell of sizes 2, 3, 4, 5 and 6 with number of growth periods of 4, 3, 2, 1 and 0. In addition, when p is less than 0.5 the binomial expression $p^s q^{n-s}$ generates the truncated geometric distribution $ap_g q_g^{s-1}$ generated by Equation (1). Asymmetrical distributions of cell sizes also fit a Weibull (Meyer 1965) distribution but its assumption of two failures has no counterpart in actual cell growth and is therefore discounted as a possible means of explaining cell size variation. A case similar to size of plant cells it that for fish length; MacDonald (1987) studied the length of pike over several years. The size of fry follows normal distribution as do other cohorts of fish of different ages. Cohorts of cells, too, follow normal distribution of size that persists over time.

Two parameters can be assigned coefficient of variation values ($cv = 100 \times sd/mean$), difference in size between sister cells and variation in size among all cells. Prescott's (1955) data on the protozoan *Tetrahymena* for the duration of cell cycle from daughter cells selected at the start of the cycle, is single-peaked with synchrony gradually lost. He found the average cell cycle duration was 111.4 min and I calculated the standard deviation of 7.76 from duration of 28 sister pairs he listed to give a coefficient of variation of $100 \times 7.6/111.4$, or 7.2. The average found here for *Elodea* daughter cells is 78.8 μm and standard deviation of size difference from the average of two daughters of 11.9 μm for a cv of $100 \times 5.95/78.8$, or 7.4, close to that for *Tetrahymena* sister cells. It seems that the Prescott data on mean generation time can be explained by unequal cell division alone. The cv for cell size of an asynchronous population is significantly greater. Data derived by Scherbaum & Rasch (1957) also from *Tetrahymena*, indicate the average cell area of a population of cells is 73 000 μm^3 with a variance of 33 300 for a cv value of $100 \times 33\,300^{0.5}/73\,000$, or 22.6. In *Elodea*, the same cells noted above had average size of 78.8 μm and an SD of 198.3^{0.5}, for a cv value of 17.9, again not too unlike that found for *Tetrahymena*. It would seem that cell size may be a valid parameter for inspecting status in protozoans as it does for plant cells. These data are also corroborated by experiments in mammalian cells where it has been demonstrated that entry into S phases is strictly controlled by cell size (Dolznic *et al.* 2004).

Various models have been proposed to calculate mutation rate (Luria & Delbruck 1943; Lederberg 1951; Luria 1951; Kepler & Oprea 2001; Zheng 2005) but all are based on a sampling method that has been shown here to give values noticeably lower than what actually occurs. The reason for this change in mutation rate is found in the equation for mutation rate $\mu = -\log(p_0)/(N - N_0)$ as given by Lederberg (1951) where p_0 is the frequency of culture samples without mutant cells and $N - N_0$ is the number of cells sampled. Close inspection reveals that the net effect of progressively smaller sampling it that p_0 arithmetically increases slower than cell number exponentially decreases (Fig. 3d). Besides the sampling problem, past calculations are based on one mutant target per dividing cell when the number of targets may be two, one

for each new strand of DNA synthesized, so $N - N_o$ becomes $2(N - N_o)$ leading to μ literally halving in value. Calculated mutation rate is significantly lowered from both sampling and number of targets considered.

Most recent studies on cell proliferation and clone formation have been on animal cells, particularly cancer cells. Tumour progression seems to involve sequential mutation and Darwinian selection resulting in a Russian dolls picture of clones within clones that are progressively more capable of proliferation, thus, they become more malignant. The Delbruck and Luria's fluctuation test has been used on cultured mouse lymphoma cells (Rudd *et al.* 1990; Boesen *et al.* 1994). These results appear valid as all cells of cultures were tested, that is, no sampling was necessary, and the problem of phenotypic lag, a problem pointed out by Kendal & Frost (1988), was experimentally resolved. The second, and more direct, method of continuous growth as applied to corn aleurone spots has been used to determine mutation rate in normal human lymphoblastoid cells (Araten *et al.* 2005). In their equation $u = f/d$, u is the mutation rate, d is the number of cell generations over which growth occurs; f is then the frequency of accumulated mutant cells from both proliferation and mutation. Early mutations occur when the population is relatively small and error is great because it is magnified by proliferation. In the more direct plant method (Equation (10)) only mutant cells, by mutation during the last cell cycle when the population is the largest, give a more accurate picture of mutation rate.

Other common studies on animal cells are those of stem cells and the clones they produce. In small intestinal crypts, there are a few proliferating stem cells, sc, many more proliferating transit cells, tc, and many non-proliferating differentiated cells, dc. In humans, a crypt seems to have as few as one stem cell, 150 proliferating cells and several hundred differentiated cells (Potten & Loeffler 1990). Near the base of a crypt one sc gives rise to a daughter tc which, in turn, divides into two granddaughter tcs, and so on. Thus, a file or ribbon of cells is a clone of those derived from a stem cell composed of subclones of 1, 2, 4, 8, 16, etc., cells, the last batch of which are siphoned off into the pool of differentiated of cells. It seems that tcs divide twice as fast as scs (Cairns 2002) so a file of cells would have clones of either 1(sc), 1(tc), 2(tc), 8(tc), 32(tc) and 128(tc) cells or 1(sc), 1(tc), 4(tc), 16(tc) and 64(tc) cells. Cells per clone and their number with 8 levels are 128 (1), 64(2), 32(4), 16(8) and 128(1) to give a mean size of clones by Equation (8) of $8.2^{8-1}/(2^8 - 1)$, or 512/127, or 4.03 cells. Accordingly, Equation (9) gives a variance of 239.8. Tumourgenesis of the intestine seems to involve a particular sequence of mutations in stem cells in becoming locked in stem cells (Pinto & Clevers 2005), the mathematics of which have been worked out by Cairns (2002). It still would be of interest to identify the subclones of transit cells coming from stem cells and their sizes, as suggested in the preceding paragraph, because knowing the normal kinetics of cell proliferation would help to characterize tumour formation.

Clones in higher organisms are more important in evolution than in development. An organism is a clone, all cells of which have the same genotype, so evolutionary competition is between organisms of pure types. Were an organism composed of mixed phenotypes and successful in competition, success is uncertain as to which genotype provided better survival. Hence, for evolutionary importance, organisms must be genetically pure clones. But mutation at any time produces mutant clones of various sizes and in various places. It is highly unlikely that a mutant clone includes both the region of a body where it contributes to its success as well as in the reproductive region so it can be passed on. Most (3/4) of mutations occur in the last two cell generations and if they are resident in reproductive cells, they are inherited and belong to pure clones in the offspring. Hence, it is important as to where mutations occur in a multicellular organism but not as to when in the organismal clone they occur. In microbes where each cell

expresses nearly all traits, almost all mutant clones are evolutionarily tested and inherited, hence adaptation is much faster than in multicellular forms, not because cell division is faster but because all mutants are immediately tested.

REFERENCES

- Alexandersson M (2001) On the existence of the stable birth-type distribution in a general branching process cell cycle model with unequal division. *J. Appl. Probab.* **38**, 685–695.
- Araten DJ, Golde DW, Zhang RH, Thaler HT, Gargiulo L, Notaro R, Luzzatto L (2005) A quantitative measurement of the human somatic mutation rate. *Cancer Res.* **65**, 8111–8117.
- Boesen JJ, Niericker MJ, Dieteren N, Simons JW (1994) How variable is a spontaneous mutation rate in cultured mammalian cells? *Mutat. Res.* **307**, 121–129.
- Cairns J (2002) Somatic stem cells and the kinetics of mutagenesis and arcinogenesis. *Proc. Natl. Acad. Sci. USA* **99**, 0567–0570.
- Dolznic H, Grebien F, Sauer T, Beug H, Mullner EW (2004) Evidence for a size-sensing mechanism in animal cells. *Nat. Cell Biol.* **6**, 899–905.
- Fanes PA (1977) Control of cell size and cycle time in *Schizosaccharomyces pombe*. *J. Cell Sci.* **24**, 51–67.
- Federhoff NV (1984) Transposable genetic elements in maize. *Sci. Am.* **250**, 84–98.
- Feller W (1940) On the logistic law of growth and its empirical verification in biology. *Acta Biotheor.* **5**, 51–66.
- Hughes WH (1955) The inheritance of differences in growth rate in *Escherichia coli*. *J. Gen. Microbial.* **12**, 265–268.
- Kelly CD, Rahn O (1932) The growth rate of individual bacterial cells. *J. Bacteriol.* **23**, 147–153.
- Kendal WS, Frost P (1988) Pitfalls and paractice of Luria-Delbruck fluctuation analysis: a review. *Cancer Res.* **48**, 1060–1065.
- Kendall DG (1948) On the role of variable generation time in the development of a stochastic birth process. *Biometrika* **35**, 316–330.
- Kepler TB, Oprea M (2001) Improved inference of mutation rates. II. Generalization of the Luria-Delbruck distribution for realistic cell-cycle time distributions. *Theor. Popul. Biol.* **59**, 41–48.
- Kimmel M, Darzynkiewicz Z, Ovide A, Tragano F (1984) Analysis of a cell model based on unequal division of metabolic constituents to daughter cell during cytokinesis. *J. Theor. Biol.* **110**, 637–664.
- Korn R. (1993) The geometry of elongating plant cells. *Bull. Math. Biol.* **55**, 345–364.
- Korn R. (2001) The geometry of proliferating dicot cells. *Cell Prolif.* **34**, 43–54.
- Korn R. (2007) Watermelon stripes: a case for the clonal mosaic model in plants. *J. Theor. Biol.* **247**, 859–861.
- Lederberg J (1951) Inheritance, variation and adaptation. In: Werkman, CH Wilson, PW, eds. *Bacterial Physiology*, pp. 67–100. New York: Academic Press.
- Luria SE (1951) The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction. *Cold Spring Harb. Symp. Quant. Biol.* **16**, 463–470.
- Luria SE, Delbruck M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511.
- MacDonald PDM (1987) Analysis of length-frequency distribution. In: Summerfelt RC, Hall GE, eds. *Age and Growth of Fish*, pp. 371–384. Ames, IA: Iowa State University.
- McClintock B (1948) Mutable loci in maize. *Carnegie Inst. Wash. Year Book* **47**, 155–169.
- Meyer PL (1965) *Introductory Probability and Statistical Applications*. Reading, MA: Addison-Wesley.
- Novick A, Szilard L (1950) Experiments with the chemostat on spontaneous mutations of bacteria. *Proc. Natl. Acad. Sci. USA* **36**, 708–719.
- Painter PR, Mary AG (1968) Mathematics of microbial populations. *Annu. Rev. Microbial.* **22**, 519–548.
- Pinto D, Clevers H (2005) Wnt, stem cells and cancer in the intestine. *Biol. Cell* **97**, 185–196.
- Popoff M (1909) Experimentelle Zellstudien. *Arch. Zellf.* **3**, 124–180.
- Potten CS, Loeffler M (1990) Stem cells: attributes, cycles, spirals, pitfalls and uncertainties Lessons for and from the crypt. *Development* **110**, 1001–1020.
- Powell EO (1955) Some features of the generation times of individual bacteria. *Biometrika* **24**, 16–44.
- Prescott DM (1955) Relations between cell growth and cell division. *Exp. Cell Res.* **9**, 328–337.
- Prescott DM (1959) Variations in the individual generation times of *Tetrahymena geleii* HS. *Exp. Cell Res.* **16**, 279–284.
- Rhoades MM (1938) Effect of the *Dt* gene on the mutability of the *a₁* allele in maize. *Genetics* **23**, 37–397.
- Rubinow SI (1968) A maturity-time representation for cell populations. *Biophys. J.* **8**, 1055–1073.

- Rudd CJ, Daston DS, Caspary WJ (1990) Spontaneous mutation rates in mammalian cells: effect of differential growth rates and phenotypic lag. *Genetics* **126**, 435–442.
- Scherbaum O, Rasch G (1957) Cell size distribution and single cell growth in *Tetrahymena pyriformis* GL. *Acta Pathol. Microbial. Scand.* **41**, 161–182.
- Smith JA, Martin L (1973) Do cells cycle? *Proc. Nat. Acad. Sci. USA* **70**, 1263–1267.
- Stuart RN, Merkel TC (1948) *Calculation of Treatment Schedules for Cancer Chemotherapy, Part II*. O. Lawrence Radiation Laboratory Report, UCRL140505, University of California, Berkeley, CA.
- Zheng Q (2005) Update on estimation of mutation rates using data from fluctuation experiments. *Genetics* **171**, 861–864.