Synchronous Growth of Cells and the Generation Time Distribution

By J. D. HARVEY
Department of Cell Biology, University of Auckland,
Private Bag, Auckland, New Zealand

(Accepted for publication 2 October 1971)

SUMMARY

Various features of the growth curve of a synchronous culture of bacteria are related to the generation time distribution of the cells, and the parent-progeny generation time correlation coefficients. It is suggested that the examination of synchronous growth enables the parameters of the generation time distribution, and the correlation coefficients, to be extracted relatively simply and their dependence upon environment to be investigated. Methods of estimating the degree of synchronization, and some corrections for the effects of imperfect synchronization, are suggested.

INTRODUCTION

The generation time distributions of several different organisms in a state of balanced growth have been investigated by direct observation of individual organisms (Kubitschek, 1962a, 1966). More recently, techniques for obtaining synchronous growth of organisms have been improved and, in the case of bacteria, synchronous cultures can now be produced simply by methods which minimize environmental shock. This paper deals with the mathematical description of the growth of synchronous cultures in terms of the generation time distribution and the correlations between parent and progeny cells. It is suggested that the observation of synchronous growth permits the deduction of various features of the generation time distribution simply and rapidly, together with the correlation coefficient between parent and daughter cell generation times. In the subsequent paper (Harvey, 1972), the suggested method is applied to observations of the growth of a colony of bacteria synchronized by the method of Helmstetter & Cummings (1964).

In the first part of this paper expressions for the observed concentrations of organisms in a synchronous culture are derived under the assumption of perfect synchrony at the start of the observations. The derivative of the concentration as a function of time is expressed as a sum of functions localized along the time axis, whose properties are related to those of the generation time distribution of individual organisms and to the correlations between the generation times of these organisms. Finally, the effects of incomplete synchrony on these relations are considered, together with possible methods for correcting for such effects.
THEORY

Perfectly synchronous cultures

Consider a culture of organisms in a state of balanced growth, synchronized in some way such that all are new, first-generation organisms at time, \( T = 0 \). Let \( t_j \) be a random variable representing the lifetime of a \( j \)'th generation organism and define

\[
T_j = t_1 + t_2 + \ldots + t_j,
\]

where \( T_j \) is a random variable representing the time of division of the \( j \)'th generation organism with distribution \( G_j(T_j) \), where

probability of \( j \)'th generation organism dividing before time \( T = \int_0^T G_j(t)\,dt \).

(2)

It may be noted that \( T_1 = t_1 \) and the distribution \( G_j(t) \) is the generation time distribution of the organism in the culture, provided that the initial organisms were collected without bias with respect to parent generation times. Assuming that, upon division, each \( j \)'th generation organism gives rise to two new \((j+1)\)'th generation organisms the number of first generation organisms at any time \( T \) is given by

\[
N_0(T) = N_0\left(1 - \int_0^T G_1(t)\,dt\right),
\]

where \( N_0 \) is the initial number of first-generation organisms in the culture. The total number of \( j \)'th generation organisms \((j > 1)\) at any time \( T \) is given by

\[
N_j(T) = 2^{j-1} N_0\left(1 - \int_0^T G_j(t)\,dt\right),
\]

thus the total number of organisms in the culture at any time will be given by

\[
N(T) = \sum_{j=1}^\infty N_j(T)
= N_0\left(1 + \sum_{j=1}^\infty 2^{j-1} \int_0^T G_j(t)\,dt\right).
\]

(5)

If the lifetimes of different generations of organisms are assumed to be independent, then the distribution of the division time of a \( j \)'th generation organism is given by the \( j \)-fold convolution of the generation time distribution, since \( T_j \) is then the sum of \( j \) independent random variables each with the same distribution (see, for example, Cramer, 1946). Denoting the generation time distribution by \( f(t) \), equation (5) becomes

\[
N(T) = N_0\left(1 + \int_0^T f(t)\,dt + 2 \int_0^T f(t) \int_t^T f(t-t')\,dt'\,dt + \ldots\right).
\]

(6)

In general the different terms in equation (5) all contribute to changes in the number of organisms which are peaked at different times, and, if the generation time distribution has a small standard deviation, lead to successive "steps" in a plot of the number of organisms against time in a synchronous culture. Differentiating equation (6) gives

\[
\frac{dN(T)}{dT} = N_0\left(f(T) + 2 \int_0^T f(t)f(T-t)\,dt + \ldots\right),
\]

\[
= N_0 \sum_{j=1}^\infty F_j(T),
\]

(7)
where

$$F_j(T) = 2^{j-1} G_j(T).$$  \hspace{1cm} (8)

Each term in equation (7) is the convolution of the generation time distribution with the previous term, and if the Laplace transform of the generation time distribution is denoted by $\phi(s)$ then the Laplace transform of the $j$'th term is given by $2^{j-1} \phi(s)^j$. From the properties of the distributions of the sum of $j$ independent random variables, it follows that

$$\int_0^\infty F_j(T) dT = 2^{j-1},$$  \hspace{1cm} (9)

$$2^{1-j} \int_0^\infty TF_j(T) dT = j\tau,$$  \hspace{1cm} (10)

$$2^{1-j} \int_0^\infty (T-j\tau)^2 F_j(T) dT = j\sigma^2,$$  \hspace{1cm} (11)

$$2^{1-j} \int_0^\infty (T-j\tau)^3 F_j(T) dT = j\mu_3,$$  \hspace{1cm} (12)

where $\tau$, $\sigma$ and $\mu_3$ are the mean, standard deviation, and skewness (third central moment) of the generation time distribution. More complicated formulae apply to the higher moments.

The equations (9) to (12) show that the means of the different functions $F_j(t)$ occur at integral multiples of the generation time, while the functions gradually spread out (become less localized along the $t$ axis) until the stepwise change in $N(T)$ goes over to an exponential growth curve for large values of $T$. From the point of view of practical applications, however, only the region of time around the first few steps in $N(T)$ is of interest. If the generation time distribution has a small coefficient of variation ($= \sigma/\tau$), then both $F_1(T)$ and $F_2(T)$ will be negligible around $T = 1.5\tau$, and the generation time distribution can be inferred directly from the derivative of $N(T)$. In practice, the tails of the functions $F_1(T)$ and $F_2(T)$ interfere slightly, leading to difficulties in estimating the higher moments of the distribution, or the exact form of the tail region.

The effect of correlated generation times on synchronous cultures

Correlations between the generation times of sister cells and those of parent and progeny cells have been observed by several previous workers (see, for example, Kubitschek, 1966). This indicates that the assumption of complete independence of the generation times, made in the previous section, is hardly justified; indeed this assumption leads to a standard deviation of $F_2(T)$ of $\sqrt{2}\sigma$, where $\sigma$ is the standard deviation of the generation time distribution. This result disagrees with the observed data on synchronous growth (Harvey, 1972).

To investigate the effect of correlations between sister cell generation times, consider the effect of separating all second-generation pairs of sister cells into two new cultures (in an imaginary experiment) as the first-generation organisms divide. Each of these cultures will behave in exactly the same way as time progresses, and mixing the two cultures together will not affect the way in which the concentration varies with time. No dependence of sister cell generation times will produce an observable effect on the concentrations observed during synchronous growth. Conversely, synchronous cultures do not provide any information on the correlations between sister cell generation times. This is a consequence of the very large population of organisms in the culture. Growth is an example of a binary branching process, and, if growth from a single individual is considered, the correlations between sister cell generation times can be important. It can be shown for example (Harris, 1963)
that if sister cell lifetimes are correlated, then the expected number of organisms (derived
from a single individual) at any subsequent time is the same as if the lifetimes were un-
correlated, but the variance of this number is greater than, or equal to, that observed if the
lifetimes were uncorrelated. This can be reconciled with the earlier statement on synchronous
cultures by considering a very large number \( N \) of initial individual organisms, in which
case the number at any subsequent time becomes \( N \) times the mean (or expected) number
from a single individual, independent of the higher moments of the distribution of this
number. In the case of batch cultures, Powell (1956) has shown that the growth rate is
independent of sister cell generation time correlations.

In contrast to the effects of correlations between the generation times of sister cells,
correlations between the generation times of organisms of different generations can have
a marked effect on the observed growth of synchronous cultures. Consider the effect of
correlations between the generation times of mother and daughter cells. Introducing the
joint normalized distribution function \( H(t_1, t_2) \) of the generation times of the mother \( (t_1) \)
and daughter \( (t_2) \), the marginal distributions of mother and daughter generation times are
both the generation time distribution, i.e.

\[
\int_0^\infty H(t_1, t_2) dt_1 = f(t_2); \quad \int_0^\infty H(t_1, t_2) dt_2 = f(t_1),
\]

and the conditional distribution of the daughter cell generation times is given by

\[
f_c(t_2|t_1) = H(t_1, t_2)/f(t).
\]

The function \( G_2(T) \) becomes

\[
G_2(T) = \int_0^T f(t)f_c(T-t|t) dt
\]

\[
= \int_0^T H(t,T-t) dt,
\]

while the mean and standard deviation of \( F_2(T) \) are now given by

\[
\frac{1}{2} \int_0^\infty TF_2(T) dT = 2\tau,
\]

and

\[
\frac{1}{2} \int_0^\infty (T-2\tau)^2 F_2(T) dT = 2\sigma^2(1+\rho),
\]

where \( \rho \) is the correlation coefficient between mother and daughter cells given by

\[
\sigma^2\rho = \int_0^\infty \int_0^\infty (t_1-\tau)(t_2-\tau)H(t_1, t_2) dt_1 dt_2.
\]

The situation with \( \rho = 1 \) corresponds to the daughter cell generation time always being
equal to that of its parent, leading to a standard deviation of \( F_2(t) \) which is twice that of
\( f(t) \). The case \( \rho = -1 \) corresponds to the situation where the deviation of the parent cell
generation time from the mean is exactly compensated for by that of the daughter cell
generation time, leading to \( t_1 + t_2 = 2\tau \) and a standard deviation of zero in \( F_2(t) \). The higher
moments of \( F_2(t) \) now involve complicated measures of correlation between mother and
daughter cell generation times, but equation (17) indicates that the correlation coefficient can
be simply extracted by the measurement of the variance of the functions \( F_1 \) and \( F_2 \).

To extend the previous treatment to correlations extending over more than one generation,
it is necessary to introduce multidimensional probability distributions of the progeny generation times. Defining the joint distribution function of the first \( j \) generations
\[
H(t_1, t_2, \ldots, t_j) \ dt_1, dt_2, \ldots, dt_j = 1, \tag{19}
\]
the function \( G_j(T) \) becomes
\[
G_j(T) = \int_0^T dt_1 \int_0^{T} dt_2 \ldots \int_0^{T} dt_{j-1} H(t_1, t_2 - t_1, \ldots, t_{j-1} - t_{j-2}, T - t_{j-1}). \tag{20}
\]

It may be verified that this expression for \( G_j(T) \) yields the same normalization and mean value of the functions \( F_j(T) \) as was found in the case of independent generation times (equations (9) and (10)) and leads to the following generalization of equation (11),
\[
2^{1-j} \int_0^\infty (T-j\tau)^2 F_j(T) dT = j\sigma^2 + 2\sigma^2 \sum_{m=1}^j \sum_{n=m+1}^j \rho_{mn}, \tag{21}
\]
where, for all \( k \),
\[
f(t_k) = \int_0^\infty \cdots \int_0^\infty H(t_1, t_2, \ldots, t_j) dt_1, \ldots, dt_{k-1}, dt_{k+1} \ldots dt_j \tag{22}
\]
and
\[
\sigma^2 \rho_{mn} = \int_0^\infty \cdots \int_0^\infty (t_m - \tau)(t_n - \tau)H(t_1, \ldots, t_j) dt_1 \ldots dt_j. \tag{23}
\]

Writing now \( \rho^{(k)} \ldots \rho_{i\uparrow i+k} \), where \( \rho^{(1)} \) is the mother–daughter generation time correlation coefficient and \( \rho^{(2)} \) is the mother–granddaughter generation time correlation coefficient, etc., equation (21) becomes
\[
\sigma_j^2 = 2^{1-j} \int_0^\infty (T-j\tau)^2 F_j(T) dT = j\sigma^2 + 2\sigma^2 \sum_{k=1}^{j-1} (j-k) \rho^{(k)}. \tag{24}
\]

The expressions for the higher moments of the functions \( F_j(T) \) again involve complicated measures of correlations between parent and progeny generation times and are not considered further here. It is to be emphasized that the use of equation (24) to extract the various parameters \( \tau, \sigma \), and the \( \rho^{(k)} \), etc., is only possible if the successive terms in equation (5) are well resolved. The coefficient of variation of the generation rate distribution has been found to be between 0.1 and 0.2 for most cases investigated previously, and it can be expected that at least the first few terms in equation (5) will be reasonably well resolved. This means that the derivative of \( N(T) \) should fall to near zero between division times, or that the top of the first few ‘steps’ in the concentration curve should be relatively flat. Nevertheless, even for quite large values of the standard deviation of the generation time, the form of the first ‘step’ on the plot of concentration against time yields considerable information on the shape of the generation time distribution.

**Effects of incomplete synchronization**

In the mathematical developments of the previous section, it has been assumed that the culture is perfectly synchronized at \( T = 0 \). In any practical experiment, perfect synchronization will not be obtainable and, if the results are to be used to extract information concerning the generation time distribution it is necessary to estimate the effects of any asynchrony.

Consider a culture partially synchronized, such that, instead of consisting entirely of organisms that had just completed division at \( T = 0 \), new organisms are introduced over
a period of time and, with a specified distribution into a previously uninhabited medium.

Let the number of new organisms introduced between time \( t \) and \( t + dt \) be \( N_0 \int f(t)dt \), where \( f(t) \) is normalized with mean and standard deviation \( \mu \) and \( \sigma \), respectively.

\[
\int_{0}^{\infty} f(t)dt = 1, \tag{25}
\]

\[
\int_{0}^{\infty} tf(t)dt = \mu, \tag{26}
\]

\[
\int_{0}^{\infty} (t-\mu)^2 f(t)dt = \sigma^2, \tag{27}
\]

thus the total number of new first generation organisms is again \( N_0 \).

Each set of new organisms introduced into the culture at time \( t \) will develop according to equation (5), and thus the behaviour of the culture at any subsequent time is given by summing the effects of all of the new organisms to give,

\[
\bar{N}(T) = N_0 \int_{0}^{T} f(t)N(T-t)dt
\]

\[
= N_0 \sum_{j=1}^{\infty} \bar{F}_j(T), \tag{28}
\]

where \( \bar{F}_j(T) \) is the convolution of \( f(t) \) with \( F_j(T) \) defined in equation (8). For a culture which is perfectly synchronized at \( t = \mu \), \( f(t) = \delta(t-\mu) \) (where \( \delta(x) \) is the Dirac delta function) and equation (28) reduces to

\[
\bar{N}(T) = \begin{cases} 
0 & T < \mu \\
N(T-\mu) & T > \mu .
\end{cases} \tag{29}
\]

The means and variances of the functions \( F_j \) are now given by

\[
\int_{0}^{\infty} T \bar{F}_j(T)dT = j\mu + \epsilon, \tag{30}
\]

\[
\int_{0}^{\infty} (T-j\mu-\epsilon)^2 \bar{F}_j(T)dT = \sigma_j^2 + \sigma^2, \tag{31}
\]

assuming no correlation between the time of introduction of the new organisms and their generation times. These equations may be used to correct, or to estimate the error in, the measured values of \( \sigma \) and \( \rho \) provided that some estimate of \( \sigma_0 \) is available. Consider an analysis in which the standard deviation of the first three functions \( \bar{F}_j(T) \) are determined, then from equations (31) and (24),

\[
\sigma_1^2 = \sigma^2 + \sigma_0^2, \tag{32}
\]

\[
\sigma_2^2 = 2\sigma^2(1 + \rho^{(1)}) + \sigma_0^2, \tag{33}
\]

\[
\sigma_3^2 = 3\sigma^2 + 2\sigma^2(2\rho^{(1)} + \rho^{(2)}) + \sigma_0^2. \tag{34}
\]

These three equations yield unique values of \( \sigma, \rho^{(1)} \) and \( \rho^{(2)} \), only if \( \sigma_0^2 \) is either known or neglected. If the correlation coefficient between mother and granddaughter cell generation times is assumed to be purely that spurious correlation due to their mutual correlation with the daughter cell generation time, then \( \rho^{(2)} = [\rho^{(1)}]^2 \) and the equations can be solved for \( \sigma, \sigma_0, \) and \( \rho^{(1)} \); alternatively, \( \sigma_0 \) can be neglected and the third equation used to check for consistency. Similar considerations apply to higher values of \( j \), each new equation in \( \sigma_j^2 \)
introducing one new correlation coefficient (see equation (24)). Correlations between a large number of generations are not to be expected, however, so that with large $j$ values, it is reasonable to neglect the highest correlation coefficient, and hence determine $\sigma_0$, $\sigma$ and all the lower correlation coefficients. Unfortunately the standard deviations of the function $F$ for large values of $j$ are difficult to determine accurately from the experimental data, and, if possible, it is preferable to reduce $\sigma_0$ to a negligible value by design of the experiment.

In any experiment there will generally be, in addition to the partially synchronized organisms, a ‘background’ of completely unsynchronized organisms. If these cells are assumed to be in a state of exponential growth (as is reasonable if the experiment has not disturbed the growth rate of the organisms), then their effect could be subtracted from the observed number $N_{obs}(T)$ by a process of curve fitting. In this case, the number of cells observed as a function of time would be given by

$$N_{obs}(T) = N(T) + n_2e^{tr},$$

where $N(T)$ is given by equation (5) and $n$ could, in principle, be determined by fitting equation (35) to the observed growth curve in the region where $N(T)$ is almost constant. In most successful synchronization techniques, the derivative of $N_{obs}(T)$ is almost zero for $T < 0.57$, implying that, within the accuracy of measurement, this ‘background’ count is negligible.

**DISCUSSION**

The generation time distribution of single organisms in a state of balanced growth has been a subject of continuing investigation. Theoretical models of growth and cell division yield predictions for the distribution of the generation time, or that of its reciprocal (the generation rate), which can only be tested by accurate experimental observations. The earliest suggestion for the form of the generation time distribution was that of Rahn (1932), and various other forms have been proposed subsequently (Kendall, 1952; Powell, 1956; Koch & Schaechter, 1962; Goodwin, 1970). Kubitschek (1962b, 1966a) has performed a series of experiments to measure the generation time distribution of bacteria, and has collected data on the observed coefficient of variation of the generation rate distribution of many different organisms; he found considerable evidence in support of a normal distribution for the generation rate. This implies a skewed distribution for the generation time, but not all of the experiments performed have supported the hypothesis of a normally distributed generation rate, and it has been claimed (Koch & Schaechter, 1962) that the skewness of the observed generation time distribution has decreased as experimental techniques have advanced. In view of these uncertainties, a new method of observing the generation time distribution would seem desirable.

The most obvious method of obtaining some measure of this distribution is by direct observation and timing of the division cycle of individual cells. These experiments have been performed on a wide range of organisms but are tedious to perform, and have yielded somewhat inconclusive results. In order to obtain a reasonable measure of the distribution, a histogram of observed frequencies must be built up from the observation of a very large number of divisions. Several experimental difficulties appear immediately. It is necessary that the organisms be in a state of balanced growth but this is difficult to achieve if the organisms are constrained to move in only two dimensions on a microscope slide. A further difficulty involves the introduction of bias into the experimental observations, which are of necessity performed over a limited number of generations; this problem is discussed in more detail by Kubitschek (1966a).
The difficulties mentioned previously are avoided by the use of a synchronous culture to measure the parameters of the generation time distribution. It is emphasized that the sample of cells contained in a synchronous culture is so large that the actual concentration of organisms varies smoothly with time, and stochastic fluctuations in the concentration are quite negligible. The determination of the concentration (and hence \( N(T) \)) is a separate problem which will be considered elsewhere (Harvey, 1972). The synchronous culture can easily be maintained in a state of undisturbed growth by standard techniques, leading to easily reproducible results when a reliable method of synchronization is used. Finally, the measurement is relatively simple and fast to perform experimentally, and this makes possible the investigation of the dependence of the parameters of the generation time distribution on environmental factors.

Powell (1964, 1969) has discussed the distinction between 'real' and 'artificial' distributions of the generation time of organisms in a state of multiplicative growth. In brief, the distinction arises because a sample of new organisms collected from an exponentially growing culture contains more cells with short lived parents than cells with long lived parents. If the generation times of parent and daughter cells are dependent, the observed generation time distribution will thus differ from that observed in a sample collected without bias with respect to the parent generation time. This indicates the need for careful assessment of the synchronization technique to decide whether any such bias has been introduced in the collection of the synchronized sample. In the method of Helmstetter & Cummings (1964) for example, in which organisms are eluted from a culture growing on a membrane, it can be assumed that the organisms are in a state of balanced growth, and the distribution observed (i.e. the function \( F_l(T) \)), is then the 'real' distribution, called by Powell 'the population distribution'. The difference between the two distributions is not great, however (Powell, 1969) especially for small coefficients of variation of the generation time distribution and small correlation coefficients between parent and daughter cell generation times.

The use of synchronous cultures for the measurement of generation time distributions has the disadvantage that the tail of the distribution, and hence its high central moments, are not well defined. The greatest disadvantage is that, at present, relatively few cultures can be successfully synchronized without disturbing their metabolism. For those organisms which can be successfully synchronized, however, the method provides an accurate measure of the central regions of the distribution (or the low central moments), and of the correlation coefficient between parent and progeny cells.

The author would like to thank Dr C. T. Tindle and Professor R. E. F. Mathews for helpful discussions, and the N.Z.U.G.C. for financial support.

REFERENCES

Generation times during synchrony


