

## Fluctuation Analysis: The Probability Distribution of the Number of Mutants Under Different Conditions

Frank M. Stewart,\* David M. Gordon<sup>†</sup> and Bruce R. Levin<sup>†</sup>

\*Department of Mathematics, Brown University, Providence, Rhode Island 02912 and <sup>†</sup>Department of Zoology, University of Massachusetts, Amherst, Massachusetts 01003

Manuscript received March 16, 1989

Accepted for publication September 16, 1989

### ABSTRACT

In the 47 years since fluctuation analysis was introduced by Luria and Delbrück, it has been widely used to calculate mutation rates. Up to now, in spite of the importance of such calculations, the probability distribution of the number of mutants that will appear in a fluctuation experiment has been known only under the restrictive, and possibly unrealistic, assumptions: (1) that the mutation rate is exactly proportional to the growth rate and (2) that all mutants grow at a rate that is a constant multiple of the growth rate of the original cells. In this paper, we approach the distribution of the number of mutants from a new point of view that will enable researchers to calculate the distribution to be expected using assumptions that they believe to be closer to biological reality. The new idea is to classify mutations according to the number of observable mutants that derive from the mutation when the culture is selectively plated. This approach also simplifies the calculations in situations where two, or many, kinds of mutation may occur in a single culture.

**F**LUCTUATION analysis was introduced by LURIA and DELBRÜCK (1943) in a classic paper. They grew a number of independent cultures of *Escherichia coli* from the same bacteriophage-sensitive strain. The variation among the numbers of phage resistant colonies that appeared, when the different cultures were plated on medium containing the bacteriophage T1, was inconsistent with the hypothesis that phage resistant mutants arose only after the bacteria were exposed to these viruses. Rather, the numbers were consistent with the hypothesis that the mutations occurred at random in the fluctuation test cultures, before the bacteria were ever exposed to the phage, and that the mutants increased in numbers as the cultures grew. In the same paper LURIA and DELBRÜCK also showed how their data might be used to estimate the rate at which the bacteria mutated to the phage resistant type.

In the 47 years since LURIA and DELBRÜCK's paper was published, fluctuation analysis has been widely used. For a review see KENDAL and FROST (1988). Some authors, ARMITAGE (1952, 1953), MANDELROT (1974), FU, LI and CHU (1982), and LI *et al.* (1985), have studied the statistical distribution of the number of mutants that will appear in such fluctuation experiments, but as far as we are aware, only two, LEA and COULSON (1949) and KOCH (1982), have given effective procedures for calculating the probability distribution itself. KOCH generalizes LEA and COULSON's formula to cover both phenotypic lag, where late mutants may fail to produce colonies, and the situation where the mutants grow at a rate different from the

original cells. However, there are other biologically reasonable assumptions that lead to probability distributions that cannot be calculated from the existing theory. Generating functions, as used by ARMITAGE and by MANDELROT, are a powerful theoretical tool. For example, they can be used to decide whether two processes lead to identical distributions, or to calculate low order moments. However, all moments of the LURIA-DELBRÜCK distribution are infinite. If one wishes to compare the results of a fluctuation test with the predictions of some theory, one needs to know the actual distribution that would obtain if the theory were correct.

The need to know these distributions was dramatically illustrated by the recent report of CAIRNS, OVERBAUGH and MILLER (1988). These authors presented evidence which they interpreted as support for the hypothesis that, in addition to producing mutations at random, bacteria have mechanisms for generating mutations that are "directed" by their needs. One line of evidence they presented is the departure of their experimental distributions from the distributions predicted by the formulas of LEA and COULSON.

This paper explains a theoretical framework that makes it possible, at least in principle, to calculate the probability distribution of the number of mutants in any fluctuation test whatever, provided that it meets one simple condition. That condition is that the actual occurrence of a mutation at a particular moment does not affect the probability that a mutation, of the same or a different kind, will occur during any given period. In other words, (1) the cells do not communicate with

each other, by chemical or any other means, and (2) it does not require two separate mutation events to produce the phenotype that will allow the mutants to be counted. On the other hand, there is no restriction on the number of different kinds of mutation or on how the different mutants grow.

We also illustrate how to make the calculation in some representative cases.

The body of the paper consists of five sections. The first treats the number of colonies that will be observed, as a random variable and explains how its probability distribution can be calculated. The section DERIVATION OF THE FORMULAS contains the mathematical arguments that justify the formulas given in the section THE FORMULAS. IN RELATION TO EARLIER WORK, we show how our formula encompasses formulas given by LEA and COULSON, ARMITAGE, and KOCH. The section COMPUTATION AND EXAMPLES returns to biological hypotheses and the distributions that they lead to. It consists of a description of the general procedure for making calculations with the formulas from THE FORMULAS and examples illustrating the procedure. In particular, there are graphs showing the distribution curves (i) when two different mutations yield clones of different fitness, (ii) when the mutation rate varies with the nutrition available to the cells, and (iii) when a randomly obtained fraction of the mutants fail to be counted because the plating process is less than 100% efficient. Although the examples deal with bacterial mutation, we hope that the methods will be useful to workers in other fields, such as somatic cell genetics. In the DISCUSSION we conclude with a brief discussion of the implications of these results for the interpretation of fluctuation test data and, in particular, of the use of probability distributions as evidence for directed mutations.

We will not feel offended if non-mathematicians choose to skip the technical material in DERIVATION OF FORMULAS and RELATION TO EARLIER WORK. (D.M.G. and B.R.L. merely skimmed them.) The formulas in THE FORMULAS are all that is needed to make the calculations that are described in the section on COMPUTATION AND EXAMPLES.

### THE FORMULAS

Imagine a growing culture of bacteria in which any cell may mutate at any time. The formulas to be derived are general and do not depend on any particular assumptions about the rates of mutation or growth. For simplicity, we use the language of differentials and write as if equations that are accurate to the first order were literally correct.

Only two assumptions are needed.

- (i) The probability that a mutation event will take place in the short interval between times  $t$  and  $t +$

$dt$  is independent of what may happen at other times and can be calculated as  $\phi(t) dt$ . Note that  $\phi(t)$  is the over all rate at which events take place. It is the product of the mutation rate per cell per hour and the number of cells present at the time. Fortunately, the actual formula for  $\phi(t)$  is seldom needed. For many situations the probability distribution of the number of mutants can be calculated from the values of a few parameters.

- (ii) If a mutation event takes place at time  $t$  then the eventual number of observed mutants in the resulting clone is a non-negative, integer valued random variable,  $N_t$ , with a probability distribution,

$$p(n; t) = \text{pr}\{N_t = n\} \quad (1)$$

that depends on when the mutation occurs. The situation where the growth of the mutants is assumed to be deterministic is just a special case.

Strictly speaking, assumption (i) can only be an approximation. It is a good approximation when the total number of mutants is small in relation to the total number of cells in the culture.

Assumption (ii) makes it possible to take account of two important facts about mutations. In the first place, there can be many mutant states, alleles of the same or different genes, that, in one sense, code for the same phenotype, *e.g.*, resistance to a bacteriophage or an antibiotic. However, some of these mutations may produce clones that grow slowly while others yield clones with little or no selective disadvantage. Such differences must be taken into account when one attempts to find the distribution of the number of mutant colonies that will be observed. In the second place, the number of colonies that will appear may depend on chance as well as on the particular kind of mutation and the time when it occurs.

We will say that a mutation at time  $t$  is "of type  $k$ " if its clone produces  $k$  colonies when the culture is plated, *i.e.*, if  $N_t = k$ . The "type" is a property of the individual mutation, not of the mutants. Since the mutants are being screened for some property, they are, *ipso facto*, of the same phenotype from the point of view of the fluctuation experiment. By contrast, the type of the mutation depends not only on the genetic change, but also on the time when the mutation takes place and, possibly, on chance as well.

The number of mutations of type  $k$  that occur in one culture during a fluctuation experiment is a random variable,  $M_k$ . As we will show in DERIVATION OF FORMULAS, the  $M_k$  are stochastically independent and each has a Poisson distribution:

$$\text{pr}\{M_k = j\} = e^{-\lambda_k} \frac{\lambda_k^j}{j!} \quad (2)$$

where the Poisson parameter,  $\lambda_k$ , is given by

$$\lambda_k = \int_0^T p(k; t)\phi(t) dt \quad (3)$$

and  $T$  is the time after which no observable mutations will occur.

Experimental conditions always put some upper limit on the possible values of  $k$ , but this is usually very large. The mathematical theory is equally valid whether we consider a finite or an infinite set of random variables and the notations are a bit simpler in the infinite case. Accordingly, we will often write the upper limit of a sum as  $\infty$ , but the reader is free to interpret this as standing for some large finite number whose actual value has little or no effect on the calculations.

Since each mutation of type  $k$  gives rise to  $k$  observed mutants, the total number,  $M$ , of mutants in a culture is the random variable

$$M = \sum_{k=1}^{\infty} k \cdot M_k \quad (4)$$

The values of the  $\lambda$ 's are determined by the biology and the experimental conditions. Once they are known, the probability distribution of  $M$  can be calculated efficiently by an inductive algorithm. In **RELATION TO EARLIER WORK AND COMPUTATION AND EXAMPLES** we show examples of how to calculate the  $\lambda$ 's under various assumptions.

One feature of this approach is that it enables one to calculate the distribution if the mutation rate changes with the physiological condition of the cells. It is no longer necessary to assume that the rate of mutation is constant per cell per generation. It is possible, for example, that the mutation rate per generation changes as the bacteria begin to exhaust their supply of nutrients. This way of approaching the question can also allow for the possibility that, due to sampling methods or experimental conditions, only a fraction of the mutant cells produce colonies that can be counted.

The algorithm for calculating the distribution of  $M$  depends on introducing the partial sums of the series, (4), and using conditional probabilities. Introduce a sequence of new random variables  $M_1^*, M_2^*, \dots$ . For each value of  $k$  let the random variable  $M_k^*$  be the partial sum:

$$M_k^* = \sum_{j=1}^k j \cdot M_j, \quad (5)$$

let

$$P(m, k) = \text{pr}\{M_k^* = m\} \quad (6)$$

be the probability that  $M_k^*$  assumes the value  $m$ , and

let

$$Q(m, k) = \exp\left(\sum_{i=1}^k \lambda_i\right) P(m, k). \quad (7)$$

As we show in the next section, the probability we are seeking turns out to be

$$\text{pr}\{M = m\} = e^{-\Lambda} Q(m, m) \quad (8)$$

where  $\Lambda = \sum_{i=1}^{\infty} \lambda_i$  is the average number of mutations to be expected in one culture. We do not attach any biological meaning to the  $Q(m, k)$ . They merely provide a tool for calculating  $P(m, m) = \text{pr}\{M = m\}$ .

We will also show that the  $Q$ 's can be calculated by induction, using the formula

$$Q(m, k) = \sum_{j=0}^{\lfloor m/k \rfloor} \frac{\lambda_k^j}{j!} Q(m - k \cdot j, k - 1) \quad (9)$$

where  $\lfloor m/k \rfloor$  stands for the largest integer that is no greater than  $m/k$ . Since  $M^*$  and  $M_1$  are one and the same,  $P(m, 1)$  is just the probability function of a Poisson random variable with mean value  $\lambda_1$  and the inductive computation begins with

$$Q(m, 1) = \frac{\lambda_1^m}{m!}. \quad (10)$$

#### DERIVATION OF THE FORMULAS

The assumptions (i) and (ii) imply that we are dealing with a Markov process. But it is a Markov process of a very simple kind and the necessary formulas can be derived from first principles, without reference to the general theory of Markov processes.

We will say that the system is in state  $E_k(j)$  if exactly  $j$  mutations of type  $k$  have already taken place. The conditions (i) and (ii) say that the rate of transition from state  $E_k(j)$  to state  $E_k(j + 1)$  is the function  $\phi_k(t) = p(k; t)\phi(t)$  and that transitions for different values of  $k$  take place independently of one another.

This means that, if we let  $f(j; k, t)$  be the probability that, at time  $t$ , the system is in state  $E_k(j)$ , then

$$\frac{df(j; k, t)}{dt} = \phi_k(t)[f(j - 1; k, t) - f(j; k, t)] \quad (11)$$

where, of course,  $f(j; k, t) = 0$  when  $j < 0$ . Since the system is certain to be in state  $E_k(0)$  at time  $t = 0$ , the differential equations (11) are complemented by the initial conditions

$$f(j; k, 0) = \begin{cases} 1, & \text{if } j = 0; \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

Integration of the differential equations from 0 to  $T$ , yields

$$f(j; k, T) = e^{-\lambda_k} \frac{\lambda_k^j}{j!} \quad (13)$$

where

$$\lambda_k = \int_0^T \phi_k(t) dt. \tag{14}$$

In other words, the probability,  $\text{pr}\{M_k = j\}$ , that  $j$  mutations of type  $k$  take place during the course of the experiment is given by Equation 2. Since the processes are quite independent of each other, the random variables  $M_1, M_2, \dots$ , are stochastically independent.

Since each mutation of type  $k$  produces  $k$  mutants,  $M_k^*$  cannot assume the value  $m$  unless  $M_k$  has a value  $j$  with  $j \cdot k \leq m$  and, in that case,  $M_k^* = m$  if and only if the mutations of types  $1, 2, \dots, k - 1$  combine to produce  $m - j \cdot k$  mutants. Because the  $M_j$ 's are independent and since  $\text{pr}\{M_k = j\} = e^{-\lambda_k} \lambda_k^j / j!$ , the probability that these two events occur is

$$\begin{aligned} \text{pr}\{M_k = j\} \cdot \text{pr}\{M_{k-1}^* = m - j \cdot k\} \\ = e^{-\lambda_k} \frac{\lambda_k^j}{j!} P(m - j \cdot k, k - 1). \end{aligned} \tag{15}$$

Summing over the possible values of  $j$  yields

$$\begin{aligned} P(m, k) = \text{pr}\{M_k^* = m\} \\ = \sum_{j=0}^{\lfloor m/k \rfloor} e^{-\lambda_k} \frac{\lambda_k^j}{j!} P(m - j \cdot k, k - 1). \end{aligned} \tag{16}$$

Now, using Equation 7 to write this in terms of the  $Q$ 's, we get

$$\begin{aligned} Q(m, k) &= \exp\left(\sum_{i=1}^k \lambda_i\right) P(m, k) \\ &= \exp\left(\sum_{i=1}^k \lambda_i\right) \sum_{j=0}^{\lfloor m/k \rfloor} e^{-\lambda_k} \frac{\lambda_k^j}{j!} \\ &\quad \cdot \exp\left(-\sum_{i=1}^{k-1} \lambda_i\right) Q(m - j \cdot k, k - 1) \\ &= \sum_{j=0}^{\lfloor m/k \rfloor} \frac{\lambda_k^j}{j!} Q(m - j \cdot k, k - 1). \end{aligned} \tag{17}$$

Finally, observe that  $M$  has the value  $m$  if and only if there are no mutations of type greater than  $m$  and the mutations of types  $1, 2, \dots, m$  produce precisely those  $m$  mutants. These two events are independent. The probability of the former is  $\prod_{j=m+1}^{\infty} e^{-\lambda_j}$  and the probability of the latter is  $P(m, m) = \exp(-\sum_{i=1}^m \lambda_i) Q(m, m)$ . Hence the probability that  $M$  has the value  $m$  is

$$\begin{aligned} \text{pr}\{M = m\} \\ &= \left(\prod_{j=m+1}^{\infty} e^{-\lambda_j}\right) P(m, m) \\ &= \exp\left(-\sum_{j=m+1}^{\infty} \lambda_j\right) \exp\left(-\sum_{i=0}^m \lambda_i\right) Q(m, m) \\ &= \exp\left(-\sum_{j=1}^{\infty} \lambda_j\right) Q(m, m) = e^{-\Lambda} Q(m, m). \end{aligned} \tag{18}$$

For future reference, note that since  $\sum_{k=1}^{\infty} p(k; t) = 1$ ,

$$\Lambda = \sum_{k=1}^{\infty} \lambda_k = \sum_{k=1}^{\infty} \int_0^T p(k; t) \phi(t) dt = \int_0^T \phi(t) dt. \tag{19}$$

RELATION TO EARLIER WORK

The first step in applying the theory developed in the previous paragraphs is to calculate the  $\lambda$ 's. Below, we will do so under the different assumptions made by LURIA and DELBRÜCK, LEA and COULSON, ARMITAGE, and KOCH.

None of these authors calculate the  $\lambda$ 's, as such, explicitly, but LEA and COULSON (1949, p. 268) and ARMITAGE (1952, p. 13), give probability generating functions for their distributions. The probability generating function for KOCH's distribution will be derived by making minor changes in LEA and COULSON's argument. Since the probability generating function for  $\text{pr}\{M = m\}$  has a simple expression in terms of the  $\lambda$ 's, the generating functions provide a convenient way to compare our method with theirs.

The probability generating function is easy to calculate because there is no need to group together events that yield the same value of  $M$ . Let  $E(j_1, j_2, \dots, j_k, \dots)$  be the event that there are  $j_1$  mutations of type 1,  $j_2$  mutations of type 2,  $\dots$ . The probability of this event is

$$\text{pr}\{E(j_1, j_2, \dots, j_k, \dots)\} = \prod_{k=1}^{\infty} e^{-\lambda_k} \frac{\lambda_k^{j_k}}{j_k!} \tag{20}$$

You may consider all possible sequences,  $(j_1, j_2, \dots)$ , but it is clear that  $E(j_1, j_2, \dots)$ 's with an infinite number of nonzero  $j$ 's have probability zero and may be ignored. When  $E(j_1, j_2, \dots)$  occurs,  $M$  has the value  $\sum_{k=1}^{\infty} k \cdot j_k$ . Hence the probability generating function of  $M$  is

$$\begin{aligned} P(z) &= \sum_{m=0}^{\infty} \text{pr}\{M = m\} \cdot z^m \\ &= \sum_{j_1, j_2, \dots=0}^{\infty} \left(\prod_{k=1}^{\infty} e^{-\lambda_k} \frac{\lambda_k^{j_k}}{j_k!}\right) \cdot z \sum_{k=1}^{\infty} k \cdot j_k \\ &= \sum_{j_1=0}^{\infty} \dots \sum_{j_k=0}^{\infty} \dots \prod_{k=1}^{\infty} e^{-\lambda_k} \prod_{k=1}^{\infty} \frac{(\lambda_k z^k)^{j_k}}{j_k!} \\ &= e^{-\sum_{k=1}^{\infty} \lambda_k} \prod_{k=1}^{\infty} \sum_{j_k=0}^{\infty} \frac{(\lambda_k z^k)^{j_k}}{j_k!} \\ &= e^{-\Lambda} \exp\left(\sum_{k=1}^{\infty} \lambda_k z^k\right). \end{aligned} \tag{21}$$

In calculating the  $\lambda$ 's we will use the following variables and boundary values. Let

- $t$  be the time,
- $n$  be the population size,

$T$  be the time when mutation ceases or, more precisely, after which no mutations will produce observable colonies,

$n_0$  be the initial population size, and let  $n_T$  be the final population size.

**The LURIA-DELBRÜCK distribution:** Consider first the original formulation of the problem by LURIA and DELBRÜCK. They assume that the cell populations grow exponentially at a constant rate and that the rate at which mutations occur is proportional to the number of cells present. LEA and COULSON base their second method on what amounts to the same assumptions.

Let  $\kappa$  be the growth rate, and  $\mu$  be the mutation rate per cell per time unit. Then  $dn = \kappa n dt$  and  $\phi(t) = \mu n$ . A clone that arises from a mutation at a time when there are  $n$  nonmutants will grow to a size  $n_T/n$ . Since the cells are discrete entities, assume that the true number is the greatest integer less than or equal to this value. Accordingly,

$$p(k; t) = \begin{cases} 1, & \text{when } k \leq n_T/n < k + 1; \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

Hence

$$\begin{aligned} \lambda_k &= \int_0^T p(k; t) \mu n dt = \frac{\mu}{\kappa} \int_{t=0}^{t=T} p(k, t) dn \\ &= \frac{\mu}{\kappa} \int_{n_T/(k+1)}^{n_T/k} 1 dn = \frac{\mu}{\kappa} \frac{n_T}{k(k+1)} \end{aligned} \quad (23)$$

when  $k + 1 \leq n_T/n_0$  and  $\lambda_k = 0$  when  $k \geq n_T/n_0$ .

LEA and COULSON assume that  $n_0$  is small enough so that one may take  $\Lambda = \int_0^T \mu n dt = (\mu/\kappa)(n_T - n_0)$  to be  $\mu n_T/\kappa$ . They denote this by  $m$  so, when the  $\lambda$ 's are given by (21), our Equation 14 coincides with their Equation 15.

**LEA and COULSON's first method:** LEA and COULSON also calculate the distribution under the assumption that the growth of a mutant clone is a random process. Specifically, they assume that the probability that a given cell will divide in a short period of time,  $dt$ , is  $\kappa dt$  where  $\kappa$  is equal to the growth rate of the nonmutant cells. Since the average size of such clones that arise from a mutation at time  $t$  is the same as the corresponding clone size in the deterministic model, it is to be expected that the two methods will give similar results, but it is not obvious that the results will be identical.

In such a birth process the probability that a mutation occurring at time  $t$  will produce a clone with exactly  $k$  members is [FELLER (1968, p. 450, Equation 3.5)]

$$p(k; t) = e^{-\kappa(T-t)}(1 - e^{-\kappa(T-t)})^{k-1} \quad (24)$$

so

$$\begin{aligned} \lambda_k &= \int_0^T \mu n_0 e^{\kappa t} e^{-\kappa(T-t)}(1 - e^{-\kappa(T-t)})^{k-1} dt \\ &= \mu n_0 e^{\kappa T} \int_0^T e^{-2\kappa(T-t)}(1 - e^{-\kappa(T-t)})^{k-1} dt \quad (25) \\ &= \mu n_T \int_{e^{-\kappa T}}^1 u(1 - u)^{k-1} \frac{du}{\kappa} \approx \frac{\Lambda}{k(k+1)}. \end{aligned}$$

Since  $e^{-\kappa T} = n_0/n_T$  the approximation process, setting  $n_0 = 0$ , is exactly that used by LEA and COULSON.

**Phenotypic lag:** ARMITAGE (1952, p. 13), introduces the concept of *effective mutants* in terms of the delay that is sometimes observed before the mutated genotype is expressed as the phenotype that can be detected, but his ideas and formulas are valid in a much wider context. In his terminology, the "true size" of a mutant clone is the number  $k$  that satisfies  $k \leq n_T/n < k + 1$ . In his notation,  $\pi_{kj}$  is the probability that a clone of true size  $k$  will be counted as a clone of size  $j$ . In our notation, his assumption is  $p(j; t) = \pi_{kj}$  when  $k \leq n_T/n < k + 1$ . Thus

$$\lambda_j = \Lambda \sum_{k=1}^{\infty} \frac{\pi_{kj}}{k(k+1)}. \quad (26)$$

The exponent that appears in braces in ARMITAGE's Equation 33 is really a double sum that is the same as that in our (21) provided the  $\lambda$ 's are given by (26).

**Selection against the mutants:** To allow for a growth differential between the original cell line and the mutants, only small changes are needed in Equations 22 and 23. If the original cells have a growth rate of  $\kappa$  and the mutants have a growth rate of  $w\kappa$  then, under a continuous growth model, there would be  $(n_T/n)^w$  mutants in a clone that arose when there were  $n$  nonmutants. Thus, (22) must be replaced by

$$p(k; s) = \begin{cases} 1, & \text{if } k \leq (n_T/n)^w < k + 1, \\ 0, & \text{otherwise,} \end{cases} \quad (27)$$

so the  $k$  and  $k + 1$  in (23) must be replaced by  $k^{1/w}$  and  $(k + 1)^{1/w}$ . This gives

$$\begin{aligned} \lambda_k &= \int_0^T p(k; t) \mu n dt = \frac{\mu}{\kappa} \int_{n_T/(k+1)^{1/w}}^{n_T/k^{1/w}} 1 dn \\ &= \frac{\mu}{\kappa} \left( \frac{n_T}{k^{1/w}} - \frac{n_T}{(k+1)^{1/w}} \right) \\ &= \Lambda(k^{-1/w} - (k+1)^{-1/w}). \end{aligned} \quad (28)$$

This formula does not agree with those given by KOCH because LEA and COULSON's two methods no longer give identical results.

In Equation 25,  $\kappa$  occurs first in the factor,  $\mu n_0 e^{\kappa t}$ , that relates to the number of nonmutants present. The two later occurrences of  $\kappa$  come from equation (24) and relate to the number of mutants. To take account of a difference in growth rate, the later oc-

currences must be replaced by  $w\kappa$ . With the replacements, (25) becomes

$$\begin{aligned} \lambda_k &= \mu n_0 e^{\kappa T} \int_0^T e^{-(1+w)\kappa(T-t)} (1 - e^{-w\kappa(T-t)})^{k-1} dt \\ &= \frac{\Lambda}{w} \int_{e^{-w\kappa T}}^1 u^{1/w} (1-u)^{k-1} du \\ &\approx \frac{\Lambda}{w} B(k, 1 + 1/w) \end{aligned} \tag{29}$$

where  $B$  is the beta function,  $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$ .

To verify that this formula (with  $n_0 = 0$ , as before) will give the same probabilities as KOCH's method, start with KOCH's Equation 12 and follow LEA and COULSON's (1949, p. 268), method for finding the generating function for their  $q$ 's. Since KOCH's  $b$  is our  $w$ , their differential equation for  $\psi$  now becomes

$$w(x - x^2)\psi'(x) + \psi(x) = x. \tag{30}$$

Since LEA and COULSON's  $q_0$  is 1,  $\psi(0) = 0$  and, according to Equation 21, we seek a solution of (30) in the form  $\psi(x) = \sum_{k=1}^{\infty} c_k x^k$  where  $\lambda_k = \Lambda c_k$ . The only such solution has  $c_1 = 1/(1 + w)$  and

$$\begin{aligned} c_k &= \frac{k-1}{k + \frac{1}{w}} c_{k-1} \\ &= \frac{1}{w} \frac{1 \cdot 2 \dots (k-1)}{\left(1 + \frac{1}{w}\right)\left(2 + \frac{1}{w}\right) \dots \left(k + \frac{1}{w}\right)} \\ &= \frac{1}{w} \frac{\Gamma(k)\Gamma\left(1 + \frac{1}{w}\right)}{\Gamma\left(k + 1 + \frac{1}{w}\right)} \end{aligned} \tag{31}$$

which agrees with (29).

KOCH's treatment of phenotypic lag is simple and practical. There is no difficulty in incorporating it in our equations.

### COMPUTATION AND EXAMPLES

From a mathematical point of view, the algorithm for calculating the probability distributions is completely described by Equations 8, 9 and 10, but a practical way to make the calculations is best described by a short program segment. The following paragraphs explain the reasons for organizing the calculations in a particular way and the pseudo-code provides a concise description of the algorithm.

Suppose you wish to calculate, for values of  $m$  up to some  $m_{\max}$ , the probability that  $m$  mutants will be observed, *i.e.*, the probability,  $\text{pr}\{M = m\}$ , that the random variable  $M$  will assume the value  $m$ . You will need to calculate the Poisson parameters,  $\lambda_k$ , for  $k = 1$  up to  $k = m_{\max}$ , but the values you will want to use

for the  $\lambda$ 's will depend on your assumptions about the biology of the situation so it is best to isolate their calculation from that of the  $Q(m, k)$ . Assume, then, that  $\lambda_1, \lambda_2, \dots, \lambda_{m_{\max}}$  are known. In the four following subsections, we will illustrate ways to calculate the  $\lambda$ 's for different experimental conditions.

If you examine the inductive formula, (9), you will see that the value of  $Q(m, k)$  depends only on the values of  $Q(m', k - 1)$  with  $m' \leq m$ . If you calculate, not up from  $m = 0$ , but down from  $m = m_{\max}$ , you can make the calculation in a single one dimensional array, making the program simpler and the computation more efficient. Think of  $Q$ , then, as an array  $Q[1], Q[2], \dots, Q[m_{\max}]$  and, when it is helpful, denote the value of  $Q[m]$  at the  $k$ th stage of the computation as  $Q_k[m]$ .

Let  $c[j]$  or  $c_k[j]$  be the coefficient,  $\lambda_k^j/j!$ , that comes from the Poisson distribution. Equation 9 now assumes the form  $Q_k[m] = Q_{k-1}[m] + \sum_{j=1}^{\lfloor m/k \rfloor} c_k[j] Q_{k-1}[m - jk]$ .

The calculation can be described in pseudo-code as follows. Key words are written in bold face, variables and expressions in *italic*, and informal descriptions in roman type.

```

declare
  k, m, j, mmax: integer;
  λ, Q, c: array of real;
  procedure set_Poisson (P: array of real, k: integer)
    For appropriate values of j, set  $P[j] := (\lambda_k)^j/j!$ .
begin calculation
  set_Poisson (Q, 1);
  for k starting from 2 to mmax by steps of 1
  begin loop 1
    set_Poisson (c, k);
    for m starting from mmax to k by steps of -1
    begin loop 2
      for j starting from 1 to m/k by steps of 1
      begin loop 3
         $Q[m] := Q[m] + c[j] * Q[m - j * k];$ 
      end loops 1, 2 and 3;
    end loops 1, 2 and 3;
  end calculation;
  
```

Source code for a complete program is available, either in hard copy or electronically via BITNET.<sup>1</sup>

<sup>1</sup> Hard copy can be obtained from the Zoology Department, University of Massachusetts, Amherst, Massachusetts 01003. There are three ways someone with a network connection can request the file.

1. For BITNET nodes which support interactive messages, send the command: "get mutants c" to LISTSERV@BROWNVN. For example, on an IBM VM/CMS system, issue:

```
tell listserv at brownvm get mutants c
```

2. From any BITNET node, send mail or a NOTE file to LISTSERV@BROWNVN. Include as the first line of text the command:

```
get mutants c
```

3. From any Internet node, send mail to

```
listserv@brownvm.brown.edu
```

and include as the first line of text the command:

```
get mutants c.
```

The program is written in *C* and contains modules both for calculating the  $\lambda_k$  under several different hypotheses, and for simulations that will produce random samples drawn from the resulting distributions.

The expected number of mutants,  $\Lambda$ , affects only the absolute probabilities of the different values of  $M$ . The relative frequencies of cultures with different numbers of mutants are independent of  $\Lambda$ . If  $\Lambda$  is not known on theoretical grounds, the only harm in using  $\sum_{k=1}^{m_{\max}} \lambda_k$ , instead, will be to underestimate the probability of "jackpots" with more than  $m_{\max}$  mutants.

We now turn to the question of finding the  $\lambda$ 's and give examples of three very different kinds of argument that may be needed.

**The LURIA-DELBRÜCK distribution:** The distribution implied by the original assumptions of LURIA and DELBRÜCK was discussed from a different point of view in *The LURIA-DELBRÜCK distribution* in the section on RELATION TO EARLIER WORK. The following treatment is very close to that of LEA and COULSON.

Consider the case where the rate of mutation per cell division is a constant,  $\mu$ . Let  $n_T$  be the ultimate number of cells in a culture. During the time that the population grows from  $n_T/(k + 1)$  to  $n_T/k$  cells, there will be

$$\frac{n_T}{k} - \frac{n_T}{k + 1} = \frac{n_T}{k(k + 1)} \quad (32)$$

cell divisions. Hence the expected number of mutations in that period is  $\mu n_T/k(k + 1)$ . From the time when such a mutation takes place to the time when the culture stops growing, the overall population will increase somewhere between  $k$ - and  $(k + 1)$ -fold. Since the mutant clone will not consist of a fractional number of cells, it is natural to assume that such a mutation will produce a clone of  $k$  mutants. In other words, the expected number of mutations of type  $k$ , is  $\lambda_k = \mu n_T/k(k + 1)$ .

**Mutation rate depending on the nutritional state of the cells:** Although the rate of mutation per generation may be constant when the cells are growing exponentially, it is conceivable that this, per generation, mutation rate could change when the rate of cell division decreases as the bacteria run out of resources. Theoretically, the change could be either an increase or a decrease in the rate per cell per generation. For purposes of illustration, we consider the case where the mutation rate combines a rate per cell per generation and a rate per cell per hour. The net effect is that the mutation rate declines as the resource concentration declines, but not as fast as the growth rate declines.

To be specific, let  $r$  stand for some measure of the concentration of resources, let  $t$ ,  $n$ ,  $n_0$ , and  $n_T$  be defined as in the section, THE LURIA-DELBRÜCK DIS-

TRIBUTION and assume that:

$$Vr = e(n_T - n), \quad dn = \frac{Pr}{Q + r} n dt, \quad (33)$$

$$\text{and } \phi(t) dt = \mu dn + \nu n dt,$$

where  $V$ ,  $e$ ,  $P$ ,  $Q$ ,  $\mu$ , and  $\nu$ , are some constants. Reasonable values for them will be suggested in the section, *Mixing and matching*.

The last of the Equations 33 says that the rate at which mutations occur has two components. There are (i) mutations that occur at a rate proportional to the rate at which the cells are dividing and (ii) mutations that occur at a steady rate independent of the growth rate.

Of course, the continuous approximation ceases to be meaningful when the resource is nearly exhausted so the parameter  $T$  cannot be related to the ultimate population size,  $n_T$ . What happens, in reality, at this late stage in the growth of a culture is not at all clear, but a plausible model can be based on the assumption that mutation ceases at a time,  $T_1$ .

When we replace  $r$  by its value in terms of  $n$  we get  $n dt = \{QV/Pe(n_T - n) + 1/P\} dn$  so

$$\phi(t) dt = \left( \mu + \frac{\nu}{P} + \frac{QV\nu}{Pe} \frac{1}{n_T - n} \right) dn. \quad (34)$$

To simplify the notation, let  $A = \mu + \nu/P$  and let  $B = QV\nu/Pe$ .

To calculate  $\lambda_k$  we must integrate  $\phi(t) dt$  over the time interval when each mutation will produce a clone of type  $k$ . According to the assumptions made in *THE LURIA-DELBRÜCK DISTRIBUTION*, this is just the time when  $n$  lies between  $n_T/(k + 1)$  and  $n_T/k$  so, using (34),

$$\begin{aligned} \lambda_k &= \int_{n_T/(k+1)}^{n_T/k} \left( A + \frac{B}{n_T - n} \right) dn \\ &= A \left( \frac{n_T}{k} - \frac{n_T}{k + 1} \right) - B \ln \left( \frac{n_T - n_T/k}{n_T - n_T/(k + 1)} \right) \\ &= \frac{An_T}{k(k + 1)} - B \ln \left( 1 - \frac{1}{k^2} \right) \end{aligned} \quad (35)$$

for  $k > 1$ .

Because Equation 34 is not a good approximation near the end of the growth period, Equation 35 cannot be used to calculate  $\lambda_1$ . Instead, we must integrate from the time, call it  $T_2$ , at which  $n = n_T/2$ , to the time  $T_1$ . This  $T_1$  may be a time when the population size,  $n_1 = n(T_1)$ , is somewhat less than its theoretical limit,  $n_T$ . Alternatively, one may assume that random mutation continues during stationary phase or even after the cells are on the plates. In that case,  $T_1$  need not be actual clock time, but may simply be a parameter that is adjusted to give a reasonable approximation for the total amount of mutation anticipated.

A straight forward, but tedious, computation—in-

tegrating the first two equations in (33) from the time,  $T_2$ , when  $n(T_2) = n_T/2$  to the time,  $T_1$ —shows that the ratio  $\rho = n_1/n_T$  is determined by the transcendental equation

$$\left(1 + \frac{en_T}{QV}\right) \ln(\rho) - \ln(1 - \rho) = \frac{en_T}{QV} [P(T_1 - T_2) - \ln 2]. \tag{36}$$

If follows, more easily, that

$$\lambda_1 = An_T \left(\rho - \frac{1}{2}\right) - B \ln 2(1 - \rho) \tag{37}$$

and  $\Lambda = An_1 - B \ln(1 - \rho)$ .

For most purposes, it is unnecessary to solve (36). Once the cells enter stationary phase,  $\rho$  ceases to relate to actual population sizes and enters into the equations merely as a way to relate the time to the expected number of mutations. Then  $\ln(\rho)$  is so close to 0 that it can be ignored and the equations (37) can be replaced by the simpler

$$\lambda_1 = A \frac{n_T}{2} + B(r - \ln 2) \quad \text{and} \quad \Lambda = An_T + Br \tag{38}$$

where  $r = -\ln(1 - \rho) = \frac{en_T}{QV} [PT - \ln 2]$  and  $T$  is the time interval from  $T_2$  until the time when mutation ceases.

**The role of plating efficiency:** It is obvious that if only a random fraction of the mutants are actually observed then some of the clones will have no representatives in the observed sample. In other words the mutations will turn out to be of type 0. Now  $\Lambda = \sum_{k=1}^{\infty} \lambda_k$  is the average number of mutations of strictly positive type and hence  $\Lambda$  will be somewhat less than the average number of mutations. Not so obvious is how these random losses may affect the nature of the distribution. Here is a case where it is not easy to give simple closed form expressions for the  $\lambda_k$  but where the methods described above can still be applied.

Suppose you can calculate the  $\lambda$ 's under the assumption that all mutants will be observed. For example, assume that they follow the LURIA-DELBRÜCK distribution described in The *LURIA-DELBRÜCK distribution*. Use primes to denote quantities thought of in terms of this initial distribution. Thus we will write

$$\lambda'_k = \frac{\Lambda'}{k'(k' + 1)} \tag{39}$$

and say that a mutation that produces a clone with  $k'$  mutants of which only  $k$  produce observable colonies is "of type'  $k'$ " and "of type  $k$ ."

If the probability that a mutant will be observed is  $p = 1 - q$  then of the mutants of type'  $k'$ , a fraction,  $\binom{k'}{k} p^k q^{k'-k}$  of them, will be type  $k$ 's. All the types'  $k'$  with  $k' \geq k$  will contribute to type  $k$ . It is not difficult to write a program to calculate

$$\lambda_k = \sum_{k'=k}^{\infty} \binom{k'}{k} p^k q^{k'-k} \lambda'_k \tag{40}$$

to any desired degree of accuracy. However, for large values of  $k$  the convergence may be very slow. Care must be taken to ensure that enough terms are used and that the floating point arithmetic provides sufficient range and accuracy. For example, the naive program used to prepare Figure 3 would not run on an IBM mainframe because of floating point overflow, so the calculations were made on a 512K Macintosh.

**Mixing and matching:** A lagniappe associated with the algorithm for calculating the distribution is that situations where there are two or more different kinds of mutation can be handled without increasing the complexity of the calculation. For example, in some experiments one may expect that some common mutation will produce cells that are at a great selective disadvantage, while the bacteria from another mutation are at a slight disadvantage, and yet other mutants are at no selective disadvantage at all. Another, perhaps common, situation is that observable mutants may arise not only from mutations in the culture vessels, but also from mutations that occur after plating (B. R. LEVIN, D. M. GORDON and F. M. STEWART, unpublished data).

For simplicity, consider the case where there are two kinds of mutation. With minor changes of notation, the same argument is equally applicable when several different kinds of mutation must be considered. Symbols with a single prime, like  $\lambda'_i$ , will refer to mutations and mutants of the first kind and symbols with a double prime will refer to mutations and mutants of the second kind. For example 'a mutation of type"  $k$ ' will mean a mutation of the second kind whose clone produces  $k$  observable colonies. The word 'type' and the symbols ' $\lambda$ ', ' $\Lambda$ ', and ' $a$ ' without a prime will refer to the over all situation.

Two equivalent ways of looking at the situation lead to formulas that differ in appearance, but not in content. The easier to describe simply looks at the individual mutations.

The number of mutations of type'  $k$  is a Poisson random variable with mean  $\lambda'_k$  and the number of mutations of type"  $k$  is a Poisson random variable with mean  $\lambda''_k$ . The overall number of mutations of type  $k$  is just the sum of the type'  $k$  mutations and the type"  $k$  mutations so it, too, is Poisson and it has mean  $\lambda_k = \lambda'_k + \lambda''_k$ .



FIGURE 1.—The distribution when mutants grow more slowly than nonmutants. The solid lines show the LURIA-DELBRÜCK distribution with  $\Lambda = 5$  and  $\Lambda = 10$ . The other curves are for mutants with  $\Lambda = 10$ , but with different assumptions about the growth rates: (-----) all mutants grow at a rate 55% that of the nonmutants; (-----) half the mutants have a growth rate 40% that of the nonmutants and half have a growth rate of 70% that of the nonmutants; (- - -) 90% of the mutants grow half as fast as the nonmutants and the other 10% grow at the same rate as the nonmutants.

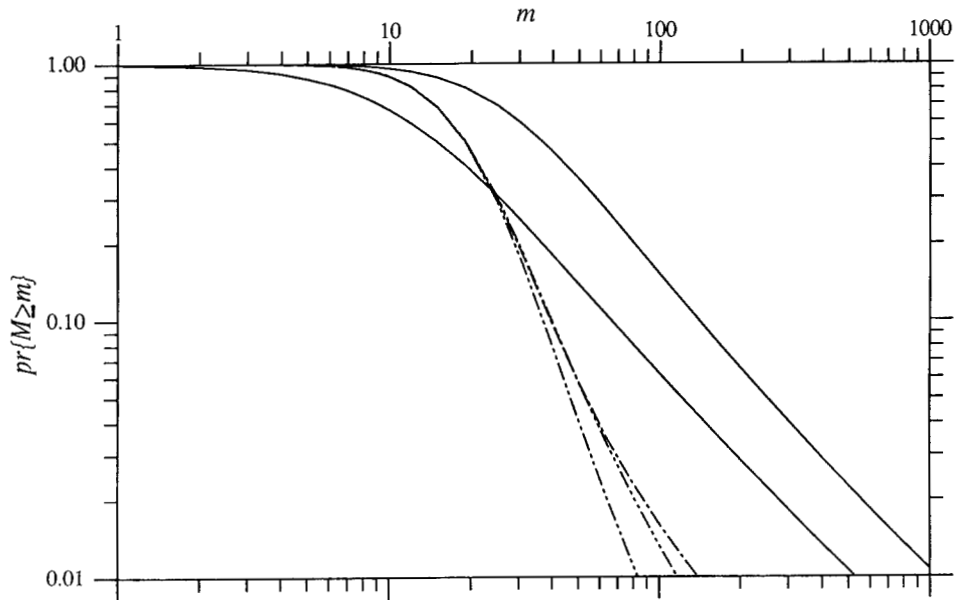
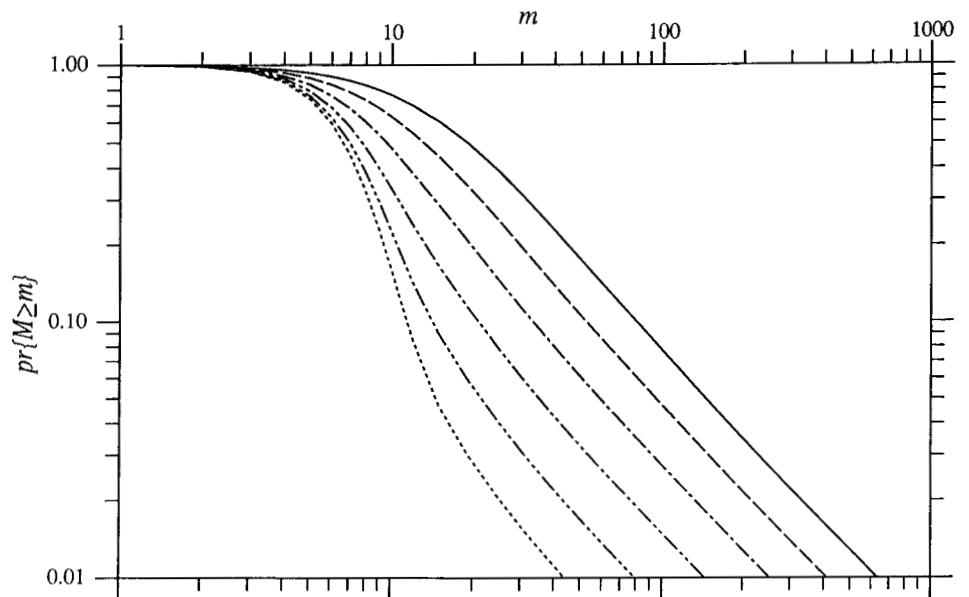


FIGURE 2.—Dependence on  $T$  when the mutation rate is not strictly proportional to growth rate. For all curves  $\Lambda = 6$ . The solid line shows the LURIA-DELBRÜCK distribution. For the other curves,  $\mu = \nu/P$  is chosen to make  $\Lambda = 6$  and  $V$  is 3 ml. Other parameters are as given in the text. (-----)  $T = 1.5$ ; (-----)  $T = 3$ ; (-----)  $T = 6$ ; (-----)  $T = 12$ ; (-----)  $T = 24$ .



To get the over all distribution, you calculate, independently, the  $\lambda_k$ 's for each kind of mutation and add them to get the over all  $\lambda_k$ 's. Then use the combined  $\lambda$ 's to calculate the over all distribution. The  $\lambda$ 's, unlike the distributions, are simply added like vectors.

At times, it may be convenient to think in terms of the total expected number of mutations,  $\Lambda$ , and what fraction of them are of a given kind. Suppose the  $\lambda$ 's are given by  $\lambda_k' = \Lambda' \cdot a_k'$  and  $\lambda_k'' = \Lambda'' \cdot a_k''$  where  $a_k'$  and  $a_k''$  are simple expressions in  $k$ . If a fraction  $p'$  of the mutations are of the first kind and a fraction  $p''$  are of the second kind, then  $\lambda_k = \Lambda \cdot a_k$  where  $a_k = p' a_k' + p'' a_k''$ .

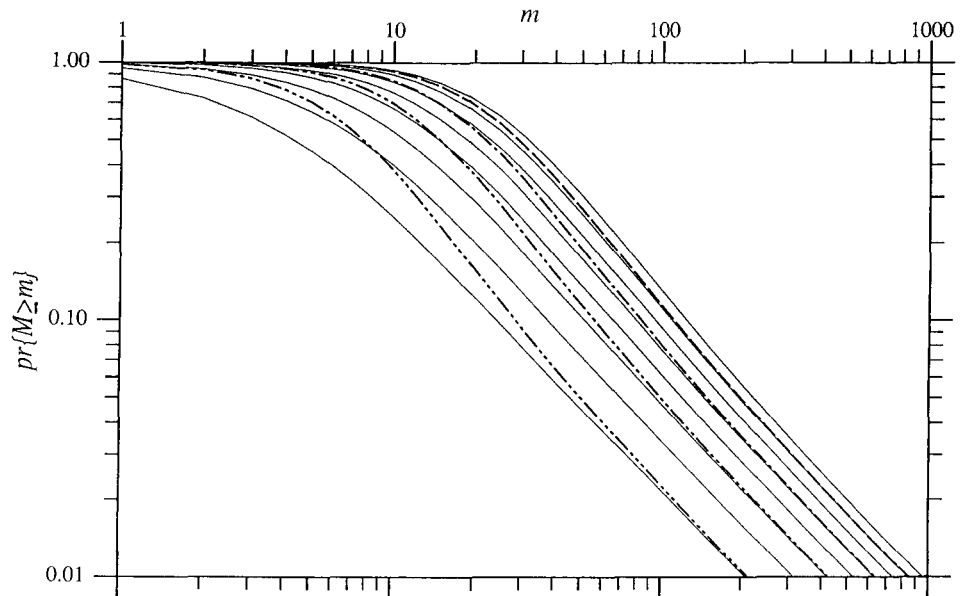
**Results:** The effects of changing the biological assumptions are best shown graphically. The figures show the cumulative probability distribution functions

for several specific sets of hypotheses. They are plotted in a form that seems to be popular for fluctuation analysis. The abscissas are the values of  $\log_{10}(m)$  and the ordinates are  $\log_{10}(\text{pr}\{M \geq m\})$ .

Figure 1 shows what happens when the mutants are at a selective disadvantage relative to the parental strains. Curves are shown both for the case where there is only a single mutant type and for the case where there are two types with different growth rates. As LENSKI, SLATKIN and AYALA (1989), have pointed out, the latter distribution can be approximated by the distribution for a single type whose growth rate is the average of the two, but the graphs show that there is some difference.

Figure 2 shows what happens when the mutation rate does not decrease as rapidly as the growth rate. The parameters  $\epsilon$  and  $P$  are measures, respectively, of

FIGURE 3.—Effect of plating efficiency. The thin, unbroken lines show the LURIA-DELBRÜCK distribution curves for  $\Lambda = 2, 3, \dots, 9$ . The other curves show the distribution when  $\Lambda' = 10$ , but each mutant has only a probability  $p$  of being counted: (---)  $p = 0.8$ ,  $\Lambda \approx 0.89$ ,  $p\Lambda' = 8$ ; (- - -)  $p = 0.6$ ,  $\Lambda \approx 0.77$ ,  $p\Lambda' = 6$ ; (- - - -)  $p = 0.4$ ,  $\Lambda \approx 0.61$ ,  $p\Lambda' = 4$ ; (- - - - -)  $p = 0.2$ ,  $\Lambda \approx 0.40$ ,  $p\Lambda' = 2$ .



the amounts of resources used by a cell between two cell divisions and the maximum rate at which the cells can grow when provided with a plentiful supply of nutrients. The following values for them and the other parameters are within the range that may be expected for cells growing in LURIA broth. If we choose a scale where  $r$  will range from 0 to 300 units per ml, then we may take  $e = 5 \times 10^{-7}$  units per cell,  $Q = 50$  units per ml, and  $P = 1.5 \text{ hr}^{-1}$ . The graphs are drawn for a population of  $n_T = 1.0 \times 10^9$  cells, and with the mutation rates,  $\mu$  mutations per cell division and  $\nu/P$  mutations per cell, equal, the exact value being chosen so that the expected number of mutations will be 6. The time between the moment,  $T_2$ , when  $n(T_2) = n_T/2$  and the time when the cells stop growing may be about an hour and three quarters. Figure 2 shows the effect of different assumptions about the time after which no observable mutations will occur.

Figure 3 shows how the LURIA-DELBRÜCK distribution is modified when a randomly obtained fraction of the mutants fail to be counted, *i.e.*, when only a sample of the culture is plated or when the mutants have a plating efficiency below 100%. As in the section, *The role of plating efficiency*, let  $\Lambda'$  be the true mean number of mutations and let  $\Lambda$  be the apparent mean number of mutations, *i.e.*, the expected number of mutant clones that will be represented by at least one colony that actually appears on the plates. Let  $p$  be the probability that any given mutant cell will generate a colony. It can be proved that

$$\Lambda = \frac{-p \ln(p)}{(1-p)} \Lambda'. \quad (41)$$

Figure 3 shows the actual distribution and, for comparison, the LURIA-DELBRÜCK distributions for several values of the mean number of mutants. The upper parts of the true distribution curves are approximated

by the LURIA-DELBRÜCK curves with mean number of mutants equal to  $\Lambda$  while the lower parts are approximated by curves whose mean number of mutants is equal to  $p\Lambda'$ .

#### DISCUSSION

The primary purpose of this study is to provide mathematical tools to extend the applicability and improve the accuracy of fluctuation analysis. However, the main motivation of the two junior authors is to get a better understanding of the meaning and limitations of fluctuation test distributions as a tool for drawing inferences about the processes that give rise to mutations.

The theory and methods derived here are general. They can equally well be applied to situations different from those we used for our illustrations. These "situations," biological processes that affect the distributions of mutants in fluctuation test experiments, are handled by modifying the equations that determine the Poisson parameters,  $\lambda_k$ . The modification can take many forms, some of which have been illustrated in the previous sections. Situations where two or more kinds of mutation process must be considered can be treated with little more effort than is required for the components because, unlike the distributions themselves, the sequences of Poisson parameters,  $\lambda_1, \lambda_2, \dots$ , can be added as if they were vectors.

In all of our illustrations, as well as those of LENSKI, SLATKIN and AYALA (1989), the distributions appear to be more or less like that anticipated from the LURIA-DELBRÜCK assumptions but modified somewhat as if there were an admixture of mutants that had a Poisson distribution. For both mathematical and biological reasons we would anticipate this to be the case in most realistic situations.

The distribution of random, selectively neutral mu-

tations in independent cultures, the LURIA-DELBRÜCK distribution, is extreme. For a LURIA-DELBRÜCK distribution, the series  $\sum \lambda_i = \Lambda \sum 1/k(k+1)$  is very slowly convergent. Even when, to be realistic, one uses only a finite number of terms, the terms with moderately large values of  $k$  have a dominant effect on the distribution. By contrast, among all distributions for non-interacting mutants in clones arising from independent mutations, the Poisson distribution is an extreme at the other end of the range of possibilities. For it, the series consists of the single term  $\lambda_1$ . All of  $\lambda_2, \lambda_3, \dots$ , are zero. Thus any process that tends to increase the relative importance of the  $\lambda_1$  will give the distribution a Poisson-like component.

We have argued that, on mathematical and biological grounds, distributions of mutants similar to those anticipated from the operation of directed mutation would be expected for many (we conjecture most) biologically realistic deviations from the conditions specified by LURIA and DELBRÜCK. The implication is straight forward. By themselves, deviations from the LURIA-DELBRÜCK distribution cannot be used as evidence that a particular process affects the incidence of mutation. More specifically, determining whether there are mechanisms for choosing which mutations occur is an empirical problem that requires direct evidence.

We are most grateful to ZVIA AGUR and to two very careful reviewers for a number of helpful comments and suggestions. This work was supported, in part, by grant GM33782 from the National Institutes of Health. We also express our gratitude to those who finance and maintain BITNET, which greatly facilitated this inter-university collaboration. In particular, we thank Brown University's

Computer and Information Services for making it possible to distribute the program referred to in the section COMPUTATION AND EXAMPLES and PETER DICAMILLO who saw to it that the program is available.

#### LITERATURE CITED

- ARMITAGE, P., 1952 The statistical theory of bacterial populations subject to mutation. *J. R. Statist. Soc. B* **14**: 1-40.
- ARMITAGE, P., 1953 Statistical concepts in the theory of bacterial mutation. *J. Hygiene* **51**, 162-184.
- CAIRNS, J., J. OVERBAUGH, and S. MILLER, 1988 The origin of mutants. *Nature* **335**: 142-145.
- FELLER, W., 1968 *An Introduction to Probability Theory and Its Applications*, Ed. 3. John Wiley & Sons, New York.
- FU, J., I-C. LI and E. H. Y. CHU, 1982 The parameters for quantitative analysis of mutation rates with cultured mammalian somatic cells. *Mutat. Res.* **105**: 363-370.
- KENDAL, W. S., and P. FROST, 1988 Pitfalls and practice of Luria-Delbrück fluctuation analysis: a review, *Cancer Res.* **48**: 1060-1065.
- KOCH, A. L., 1982 Mutation and growth rates from Luria-Delbrück fluctuation tests. *Mutat. Res.* **95**: 129-143.
- LEA, D. E., and C. A. COULSON, 1949 The distribution of the numbers of mutants in bacterial populations. *J. Genetics* **49**: 264-285.
- LENSKI, R. E., M. SLATKIN and F. J. AYALA, 1989 Mutation and selection in bacterial populations: alternatives to the hypothesis of directed mutation, *Proc. Natl. Acad. Sci. USA* **86**: 2775-2778.
- LI, I-C., S-C. H. WU, J. FU and E. H. Y. CHU, 1985 A deterministic approach for the estimation of mutation rates in cultured mammalian cells. *Mutat. Res.* **149**: 127-132.
- LURIA, S., and M. DELBRÜCK, 1943 Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**: 491-511.
- MANDELBROT, B., 1974 A population birth-and-mutation process. I. Explicit distributions for the number of mutants in an old culture of bacteria. *J. Appl. Prob.* **11**: 437-444.

Communicating editor: E. THOMPSON