ELSEVIER

Mathematical Biosciences 162 (1999) 1–32

Mathematical Biosciences
an international journal

www.elsevier.com/locate/mathbio

Review

Progress of a half century in the study of the Luria–Delbrück distribution

Qi Zheng *

Division of Biometry and Risk Assessment, National Center for Toxicological Research, HFT-20, 3900 NCTR Drive, Jefferson, AR 72079-9502, USA

Received 19 April 1999; received in revised form 20 July 1999; accepted 18 August 1999

Abstract

The Luria–Delbrück mutation model has been mathematically formulated in a number of ways. This review article examines four most important formulations, focusing on important practical issues closely linked with the distribution of the number of mutants. These issues include the probability generating functions, moments (cumulants), computational methods and asymptotics. This review emphasizes basic principles which not only help to unify existing results but also allow for a few useful extensions. In addition, the review offers a historical perspective and some new explanations of divergent moments. © 1999 Published by Elsevier Science Inc. All rights reserved.

Keywords: Luria–Delbrück model; Luria–Delbrück distribution; Estimation of mutation rate; Poisson-stopped-sum distribution; Filtered Poisson process

Contents

* Tel.: +1-870 543 7176; fax: +1-870 543 7662.
E-mail address: qzheng@nctr.fda.gov (Q. Zheng)

PII: S0025-5564(99)00045-0

1. Introduction

The Luria–Delbrück mutation model had its origin in a series of classic experiments pioneered by Luria and Delbrück [1]. These experiments aimed at settling a fundamental issue from bacteriology: whether phage-resistant bacteria arose from spontaneity (random mutation) or from adaptation (directed mutation). The Luria–Delbrück model not only played a predominant role in helping settle this fundamental issue, but also came into general use as a tool for estimating mutation rates [2]. A number of mathematical formulations of the Luria–Delbrück model came into existence as a result of attempts to improve estimation of mutation rates. Due to its vital importance in estimating mutation rates, the distribution of the number of mutants induced by the Luria–Delbrück model has been the focus of research. We shall call the distribution of the number of mutants determined by a particular formulation of the Luria–Delbrück model as a Luria–Delbrück distribution. Although half a century has elapsed since the Luria–Delbrück model was proposed, our knowledge about the Luria–Delbrück distribution remains fragmentary, and even incoherent in some aspects. This situation is due largely to a rather singular historical development of the field.

When Luria and Delbrück [1] first proposed the model that was to become their namesake, they used a specific mathematical formulation under which both the normal cells and the mutant cells grew deterministically but mutations occurred randomly. Luria and Delbrück found the probability of zero mutations, and from first principles derived the mean and the variance of the distribution of the number of mutants. These results were sufficient to

implement the two methods which Luria and Delbrück suggested for estimating mutation rates. The first method was based on the observed proportion of cultures containing no mutants; the second method relied on the mean number of mutants as determined by their model. These two methods were later known as the P_0 method and the method of means [3]. Aiming at finding the distribution function of the number of mutants, Lea and Coulson [4] tackled a different formulation under which the normal cells grew deterministically but the mutants grew stochastically. There can be little question that the far-reaching work of Lea and Coulson is a hallmark in the mathematical theory of the Luria–Delbrück model. However, it is also clear that Lea and Coulson arrived at an approximate probability generating function (p.g.f.) not so much by design as by accident. Before the paper of Lea and Coulson appeared in print, Coulson must have compared the approximate p.g.f. with an exact p.g.f. communicated to him by Kendall, for Coulson then cautioned that, under some circumstances, the approximate p.g.f. could produce results that "are seriously in error". The exact p.g.f. which eluded Lea and Coulson did not see the light until some three years later, when Armitage [5] presented it to the Royal Statistical Society in 1951. The exact p.g.f. presented by Armitage originated from Bartlett. Armitage gave no derivation details; he probably expected that the void would be filled by either Kendall or Bartlett, both of whom were among the scheduled discussants of Armitage's paper. In his discussion on Armitage's paper, Bartlett presented a p.g.f. derived from yet another formulation under which both the normal cells and the mutant cells grew stochastically. Bartlett offered no derivation details at the time, but pointed out that a limiting form of his p.g.f. coincided with the approximate p.g.f. of Lea and Coulson. On the other hand, Kendall was unexpectedly prevented from joining in the discussion and was later invited to contribute a note on what he originally intended for the discussion. However, having learned Bartlett's results, Kendall deemed that what he had prepared for the discussion "would require drastic revision" and instead he addressed a different issue [6]. Meanwhile, Kendall [7] also proposed formulations that allowed the cellular cycle of mutants to have an arbitrary continuous distribution. Kendall's work was a tremendous theoretical contribution, but for practical purposes, as Mandelbrot [8] put it, 'because of its generality, it lacked explicitness'.

The long-awaited derivations of the exact p.g.f.s did not appear until 1955. In his classic text, Bartlett [9, pp. 115–118; 10, pp. 132–135] offered elegant derivations of the two p.g.f.s, one for the Lea–Coulson formulation and the other for the formulation proposed three years earlier by himself. (To some Bartlett's derivations might appear too concise to be easily accessible.) Some 10 years later Bailey [11, pp. 125–129] put forth a substantially more lucid derivation of the exact p.g.f. for the Lea–Coulson formulation. Another decade later Crump and Hoel [12] employed the filtered Poisson process theory to

simplify the derivation. Unfortunately, little attention was paid to these two illuminating derivations of the exact p.g.f. for the Lea–Coulson formulation. Such unwitting neglect caused research efforts in the ensuing decades to focus on the approximate p.g.f. of Lea and Coulson.

To enhance the applicability of the Luria–Delbrück model, Koch [13] explored ways of extending the Luria–Delbrück model to allow for differential growth between normal and mutant cells. Koch derived the mean and the variance under the Luria–Delbrück formulation, and attempted to generalize an algorithm of Lea and Coulson for computing the probability function. Li et al. [2] applied the Luria–Delbrück model for differential growth to experimental data. In 1988, the work of Cairns et al. [14] immediately stirred not only controversy about the utility of the Luria–Delbrück model [15], but also a phenomenal resurgence of interest in many mathematical facets of the Luria–Delbrück model. Stewart et al. [16] proposed a method for writing down p.g.f.s and an algorithm for computing probability functions from the p.g.f.s; their approach was applicable to almost any case where normal cells were assumed to grow deterministically. Ma et al. [17] and Sarkar et al. [18] soon suggested a simpler and more efficient algorithm to compute the probability function. Ma et al. [17] also kindled interest in the asymptotics of the distribution. As a result, a considerable amount of ingenuities was devoted to the asymptotics by Kemp [19], Pakes [20], Goldie [21] and Prodinger [22].

Such an uneven development of a half century has produced an overwhelming number of results. To organize a great majority of these results into a coherent and accessible theory is the main goal of the present review. Fortunately, the basic principles of the subject are few and simple. For four most important formulations of the Luria–Delbrück model, we shall present these basic principles to elucidate existing results, and occasionally to elicit some minor extensions (when this can be done concisely).

2. Formulation of a mutation

It is helpful at the beginning to briefly review possible ways of modeling a mutation. Kendall [23] proposed three formulations for modeling a mutation. In Kendall's original terminology, the three formulations are

(A) grey$\rightarrow$ black,

(B) grey$\rightarrow$ grey + black,

(C) grey$\rightarrow$ black + black.

In our context, 'grey' stands for a normal cell and 'black' a mutant cell. Mandelbrot [8] paraphrased formulation A in a vivid manner: "a bacterium that mutates may be considered by its non-mutant brethren as having died". In other words, under Formulation A, a mutation entails the loss of a normal cell.

From a biological point of view, Formulation A may not be the best choice, because most mutations are believed to occur at the time of cellular division. As observed by Crump and Hoel [12], many earlier authors adopted Formulation A and made a tacit assumption that "the occurrence of a mutation does not decrease the rate at which mutations occur in subsequent time intervals". Lea and Coulson [4] and Armitage [5] were among such authors. In other words, they ignored the 'death' of a normal cell caused by a mutation. Clearly this practice amounts to adopting Formulation B. As a consequence, some of the results which these authors thought were approximate are in effect exact, if we adopt Formulation B to interpret these results. (Mandelbrot [8] strictly followed Formulation A and hence his results would probably require slight changes when interpreted with Formulation B.) We shall assume Formulation B throughout this review.

3. The Luria–Delbrück formulation

Luria and Delbrück [1] gave this earliest mathematical formulation. Because all subsequent formulations are simple variations of this formulation, we list its underlying assumptions for future reference.

1. The process starts at time $t = 0$ with one normal cell and no mutants.
2. Normal cells grow deterministically at a constant rate, say, β_1. Therefore, the number of normal cells at time t is
$$N(t) = e^{\beta_1 t}. \tag{1}$$
3. Mutants grow deterministically at a constant rate, say, β_2. If a mutant is generated by a normal cell at time $s > 0$, then the clone spawned by this mutant will be of size $e^{\beta_2(t-s)}$ for any $t \geqslant s$.
4. Mutations occur randomly at a rate proportional to $N(t)$. If μ denotes the per cell per unit time mutation rate, then mutations occur in accordance with a Poisson process having an intensity function
$$v(t) = \mu e^{\beta_1 t}. \tag{2}$$
Consequently, the expected number of mutations occurring in the time interval $[0, t)$ is
$$m(t) = \int_0^t v(s)\,\mathrm{d}s = \frac{\mu}{\beta_1}(e^{\beta_1 t} - 1). \tag{3}$$

Lea and Coulson [4] and Armitage [5] believed that Eqs. (1)–(3) held only approximately because they assumed that the number of normal cells $N(t)$ was $e^{\beta_1 t}$ less the number of mutations occurred by time t. Because we adopt Kendall's Formulation B, these equations are exact, as explained in Section 2. For large t, $m(t)$ is roughly the same as

$$\theta(t) = \frac{\mu}{\beta_1} e^{\beta_1 t}. \tag{4}$$

Historically $\theta(t)$ was often equated with $m(t)$. In fact, these two quantities are linked by

$$m(t) = \theta(t)\phi(t), \tag{5}$$

where

$$\phi(t) = 1 - e^{-\beta_1 t} = 1 - N(t)^{-1}. \tag{6}$$

For simplicity, $m(t)$, $\theta(t)$ and $\phi(t)$ will be abbreviated as m, θ and ϕ, when their dependency on time is clear from context. Except when stated otherwise, $X(t)$ denotes the number of mutants existing at time t, and $p_n(t)$ the probability of having n mutants at time t.

Assumptions 1–4 imply that $X(t)$ can be expressed as

$$X(t) = \begin{cases} 0, & M(t) = 0, \\ \sum_{i=1}^{M(t)} \exp\{\beta_2(t - \tau_i)\}, & M(t) \geqslant 1. \end{cases} \tag{7}$$

Here τ_i are the epochs at which mutations occur, and $M(t)$ stands for the mutation process which is a Poisson process with intensity function $v(t)$ given in Eq. (2). Crump and Hole [12] were the first ones who brought out the connection of this formulation to the shot noise process (which falls into the category of filtered Poisson process).

Since the event of no mutants existing at time t is equivalent to the event of no mutations ever happening by time t, it follows from Assumption 4 that

$$p_0(t) = e^{-m(t)}. \tag{8}$$

By virtue of Campbell's theorem or Eq. (5.11) of Snyder and Miller [24], the nth cumulant of $X(t)$ is seen to be

$$\begin{aligned} \kappa_n(t) &= \int_0^t \mu e^{\beta_1 s} e^{n\beta_2(t-s)}\, ds, \\ &= \begin{cases} \frac{\mu}{\beta_1 - n\beta_2}(e^{\beta_1 t} - e^{n\beta_2 t}) & (\beta_1 \neq n\beta_2), \\ \mu t e^{n\beta_2 t} & (\beta_1 = n\beta_2). \end{cases} \end{aligned} \tag{9}$$

Eq. (9) in its present form appears to be new, but the special case $\beta_1 = \beta_2$ is well-known – it was solved first by Armitage [5, p. 9] from first principles, and later by Bailey [11, p. 131] and Crump and Hoel [12, p. 243] by manipulating the cumulant-generating function of $X(t)$. From Eq. (9) it is seen that the mean and the variance of $X(t)$ for the case $\beta_1 = \beta_2$ are

$$E[X(t)] = \mu t e^{\beta_1 t}, \tag{10}$$

$$\mathrm{Var}[X(t)] = \frac{\mu}{\beta_1} e^{\beta_1 t}(e^{\beta_1 t} - 1). \tag{11}$$

These two identities were first derived by Luria and Delbrück (cf. Eqs. (6) and (10) in [1]) by setting $\beta_1 = 1$. Koch [13, p. 137] was the first to tackle the case of differential growth – the case where $\beta_1 \neq \beta_2$. By assuming $\beta_1 = 1$ and denoting β_2 by b, Koch found that

$$E[X(t)] = \frac{\mu N(t)(1 - e^{-(1-b)t})}{1 - b} \quad (b \neq 1) \tag{12}$$

and

$$\mathrm{Var}[X(t)] = \begin{cases} \mu N(t)(e^{(2b-1)t} - 1)/(2b - 1) & (b \neq 0.5), \\ \mu t N(t) & (b = 0.5). \end{cases} \tag{13}$$

(The case $b = 0.5$ was added by this author).

All the formulae so far were derived under Assumption 1. If the process starts at $t = 0$ with $n_0 > 1$ normal cells, it suffices to replace μ in each formula with $n_0\mu$. The reason is as follows. The only effect on $X(t)$ of having n_0 normal cells at $t = 0$ is to increase the chance of a mutation occurring in the interval $[t, t + \Delta t)$ from $\mu e^{\beta_1 t}\Delta t + o(\Delta t)$ to $\mu n_0 e^{\beta_1 t}\Delta t + o(\Delta t)$.

Finally, it follows from Eq. (5.10) in [24] that the characteristic function of $X(t)$ is

$$\Phi(\omega, t) = E[e^{i\omega X(t)}] = \exp\left(\mu \int_0^t \left[\exp(i\omega e^{\beta_2(t-\tau)}) - 1\right] e^{\beta_1 \tau}\, d\tau\right). \tag{14}$$

(The characteristic function given by Crump and Hoel [12, p. 243] was in effect a series expansion of $\log \Phi(\omega, t)$ for the case $\beta_1 = \beta_2$.) In principle, $\Phi(\omega, t)$ can be numerically inverted to obtain the probability distribution function (see, e.g., [25, p. 153]). However, a convenient closed form expression for the probability distribution function seems elusive. Luria and Delbrück [1] remarked that "the calculation of the distribution function involves considerable mathematical difficulties". In a sense, this assertion has remained true to the present day. Bailey [11, p. 130] has shed profound insight into the nature of this issue: "when a mutation occurs $X(t)$ will jump from X to $X + 1$. The distribution of the number of mutants thus involves both continuous and discrete elements. From the point of view of mathematical rigor there are certain analytical difficulties here". A similar observation was made by Crump and Hoel [12, p. 243]. A convenient way of circumventing such analytical difficulties is to discretize the growth function for the mutants.

4. The discretized Luria–Delbrück formulation

4.1. Theoretical considerations

Although this formulation can probably be traced back to the so-called 'second method' of Lea and Coulson [4], Armitage [5] was the first one who rigorously articulated it. This formulation overcame the analytical difficulties just noted by making a slight change in Assumption 3. Instead of adopting the continuous exponential growth function $e^{\beta_2 t}$, we discretize it by an approximating step function

$$h(t) = i \quad \text{for} \quad i \leqslant e^{\beta_2 t} < i+1 \text{ and } i = 1, 2, \ldots. \tag{15}$$

Thus $X(t)$ is expressible as

$$X(t) = \begin{cases} 0, & M(t) = 0, \\ \sum_{i=1}^{M(t)} h(t-\tau_i), & M(t) \geqslant 1, \end{cases} \tag{16}$$

where $M(t)$ is the same Poisson process as in Eq. (7).

For convenience, we now define an integer-valued function K of time t by

$$K(t) = \lfloor e^{\beta_2 t} \rfloor, \tag{17}$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. The dependency of K on time will often be suppressed for simplicity. Furthermore, denote the ratio of the two cellular birth rates by

$$\rho = \frac{\beta_1}{\beta_2}. \tag{18}$$

Clearly, $\{X(t) : t \geqslant 0\}$ is a filtered Poisson process. By virtue of the theory of filtered Poisson process (e.g., the lemma in [26]), we can write the p.g.f. of $X(t)$ as

$$\begin{aligned} G(z;t) = E[z^{X(t)}] &= \exp\left(\int_0^t (z^{h(t-s)} - 1)\mu e^{\beta_1 s}\, ds \right) \\ &= \exp\left(\sum_{j=1}^{K-1} \int_{t-\beta_2^{-1}\log(j+1)}^{t-\beta_2^{-1}\log(j)} (z^j - 1)\mu e^{\beta_1 s}\, ds \right. \\ &\qquad \left. + \int_0^{t-\beta_2^{-1}\log(K)} (z^K - 1)\mu e^{\beta_1 s}\, ds \right) \\ &= e^{-m(t)} \exp\left\{ \theta(t) \left(\sum_{j=1}^{K-1} \left(\frac{1}{j^\rho} - \frac{1}{(j+1)^\rho} \right) z^j + \left(\frac{1}{K^\rho} - \frac{1}{e^{\beta_1 t}} \right) z^K \right) \right\}, \end{aligned} \tag{19}$$

which has hitherto not appeared. In place of Eq. (19), Stewart et al. [16] arrived at an approximate p.g.f.

$$\widetilde{G}(z;t) = \mathrm{e}^{-\theta(t)} \exp\left(\theta(t) \sum_{j=1}^{\infty} \left\{ \frac{1}{j^{\rho}} - \frac{1}{(j+1)^{\rho}} \right\} z^{j}\right). \tag{20}$$

In the case $\beta_1 = \beta_2$, this approximate p.g.f. reduces to the famous p.g.f. of Lea and Coulson [4] which we shall discuss in the next section:

$$\widehat{G}(z;t) = \exp\left(\theta(t) \left\{ \sum_{j=1}^{\infty} \frac{z^{j}}{j(j+1)} - 1 \right\}\right). \tag{21}$$

In contrast, the exact p.g.f. in Eq. (19) reduces to

$$\overline{G}(z;t) = \mathrm{e}^{-m(t)} \exp\left\{\theta(t) \left(\frac{z}{1\cdot 2} + \frac{z^2}{2\cdot 3} + \cdots + \frac{z^{K-1}}{(K-1)K} + \frac{z^{K}(\mathrm{e}^{\beta_1 t} - K)}{K\mathrm{e}^{\beta_1 t}} \right)\right\}, \tag{22}$$

which was first obtained by Armitage [5, Eq. (30b)]. A derivation of Eq. (22) based on the theory of filtered Poisson process was first offered by Crump and Hoel [12, Eq. (2.19)].

It is well-known that the distribution determined by the approximate p.g.f. $\widehat{G}$ possesses only divergent moments. From a historical standpoint, we can scarely overemphasize the often-overlooked fact that such divergent moments are merely a consequence of some sort of approximation. All moments determined by the exact p.g.f. G in Eq. (19) are finite. By repeatedly differentiating $\log G(\mathrm{e}^{\psi};t)$ with respect to ψ, we find the nth cumulant to be

$$\kappa_n(t) = \theta(t) \left\{ \sum_{j=1}^{K-1} \left(\frac{1}{j^{\rho}} - \frac{1}{(j+1)^{\rho}} \right) j^{n} + \left(\frac{1}{K^{\rho}} - \frac{1}{\mathrm{e}^{\beta_1 t}} \right) K^{n} \right\}, \tag{23}$$

which appears to be new. When $\beta_1 = \beta_2$, Eq. (23) simplifies to

$$\kappa_n(t) = \theta(t) \left(\frac{1}{2} + \frac{2^{n-1}}{3} + \cdots + \frac{(K-1)^{n-1}}{K} + K^{n-1}(1 - K\mathrm{e}^{-\beta_1 t}) \right), \tag{24}$$

which was due to Armitage [5]. In particular, the mean and the variance in the case $\beta_1 = \beta_2$ come out to be

$$E[X(t)] = \theta(t) \left(1 + \frac{1}{2} + \cdots + \frac{1}{K} - K\mathrm{e}^{-\beta_1 t} \right), \tag{25}$$

$$\mathrm{Var}[X(t)] = \theta(t) \left(\frac{1}{2} + \frac{2}{3} + \cdots + \frac{K-1}{K} + K(1 - K\mathrm{e}^{-\beta_1 t}) \right). \tag{26}$$

Observe that for large t, $K\mathrm{e}^{-\beta_1 t} \approx 1$ and $1 + 1/2 + \cdots + 1/k \approx \log(K) \approx \beta_1 t$. Therefore,

$$E[X(t)] \approx \mu t e^{\beta_1 t}, \tag{27}$$

$$\mathrm{Var}[X(t)] \approx \frac{\mu}{\beta_1} e^{2\beta_1 t}. \tag{28}$$

So asymptotically the mean and the variance are the same as those given by the Luria–Delbrück formulation (cf. Eqs. (10) and (11)).

4.2. Algorithmic considerations

Stewart et al. [16] devised the following algorithm for computing the probability function of the number of mutants.

Lemma 1 (Ref. [16]). *Let* $\lambda_1, \lambda_2, \ldots \geqslant 0$ *be arbitrary real numbers and* $\lambda = \sum_{j=1}^{\infty} \lambda_j < \infty$. *Let* S *be a discrete random variable having p.g.f.*

$$G(z) = e^{-\lambda} \exp\left(\sum_{j=1}^{\infty} \lambda_j z^j \right). \tag{29}$$

Then $\Pr(S = j)$ *for* $j = 0, 1, \ldots, m$ *can be computed in the following four steps.*

1. *Set* $\Pr(S = 0) = e^{-\lambda}$.
2. *Compute* $Q(n, 1) = \frac{\lambda_1^n}{n!}$ *for* $n = 0, 1, \ldots, m$.
3. *Compute* $Q(n, k) = \sum_{j=0}^{\lfloor n/k \rfloor} \frac{\lambda_k^j}{j!} Q(n - jk, k - 1)$ *inductively for* $0 \leqslant n \leqslant m$ *and* $1 \leqslant k \leqslant m$.
4. *Set* $\Pr(S = j) = e^{-\lambda} Q(j, j)$ *for* $j = 1, \ldots, m$.

Proof. First, inserting $z = 0$ into Eq. (29) gives $\Pr(S = 0) = e^{-\lambda}$. Next, by re-writing the p.g.f. in Eq. (29) as $G(z) = \prod_{j=1}^{\infty} \exp\{\lambda_j (z^j - 1)\}$, we get a new representation $S = \sum_{j=1}^{\infty} Y_j$, where Y_j / j $(j = 1, 2, \ldots)$ are independent Poisson random variables having mean λ_j. For $n \geqslant 0$ and $k \geqslant 1$, define

$$\begin{aligned} P(n, k) &= \Pr(Y_1 + \cdots + Y_k = n), \\ Q(n, k) &= e^{\lambda} \Pr(Y_1 + \cdots + Y_k = n, Y_{k+1} = \cdots = 0) \\ &= \exp\left(\sum_{j=0}^{k} \lambda_j \right) P(n, k). \end{aligned} \tag{30}$$

Because $\lambda = \sum_{j=0}^{\infty} \lambda_j$, it is clear that

$$Q(n, 1) = e^{\lambda} \Pr(Y_1 = n, Y_2 = \cdots = 0) = \frac{\lambda_1^n}{n!}. \tag{31}$$

Since Y_k are independent and $\Pr(Y_k = jk) = e^{-\lambda_k} \lambda_k^j / j!$,

$$P(n,k) = \sum_{j=0}^{\lfloor n/k \rfloor} \Pr(Y_1 + \cdots + Y_{k-1} = n - jk, Y_k = jk)$$
$$= \sum_{j=0}^{\lfloor n/k \rfloor} \mathrm{e}^{-\lambda_k} \frac{\lambda_k^j}{j!} P(n - jk, k-1). \tag{32}$$

From Eqs. (30) and (32), it follows that

$$Q(n,k) = \exp\left(\sum_{j=1}^{k} \lambda_j\right) \sum_{j=0}^{\lfloor n/k \rfloor} \mathrm{e}^{-\lambda_k} \frac{\lambda_k^j}{j!} P(n - jk, k-1)$$
$$= \exp\left(\sum_{j=1}^{k} \lambda_j\right) \sum_{j=0}^{\lfloor n/k \rfloor} \mathrm{e}^{-\lambda_k} \frac{\lambda_k^j}{j!} \exp\left(-\sum_{j=0}^{k-1} \lambda_j\right) Q(n - jk, k-1)$$
$$= \sum_{j=0}^{\lfloor n/k \rfloor} \frac{\lambda_k^j}{j!} Q(n - jk, k-1). \tag{33}$$

Finally, because $S = n$ implies $Y_j = 0$ for all $j \geqslant n+1$, we deduce that

$$\Pr(S = n) = \mathrm{e}^{-\lambda} Q(n,n). \qquad \square \tag{34}$$

It is worth noting that the above algorithm applies to the Poisson-stopped-sum distribution in general. The p.g.f. of a Poisson-stopped-sum distribution is usually given in the form $G(z) = \exp\{\Lambda(\sum_{j=0}^{\infty} p_j z^j - 1)\}$. By defining $\lambda_j = \Lambda p_j$ for $j = 1, 2, \ldots$ and $\lambda = \Lambda(1 - p_0)$, we readily recast the p.g.f. in the form of Eq. (29). This observation seems to have escaped notice in the literature.

To compute the distribution of the number of mutants as determined by the p.g.f. in Eq. (19), we can invoke Lemma 1 by setting

$$\lambda_j = \begin{cases} \theta(t)(j^{-\rho} - (j+1)^{-\rho}), & j = 1, \ldots, K-1, \\ \theta(t)(K^{-\rho} - \mathrm{e}^{-\beta_1 t}), & j = K, \\ 0, & j = K+1, \ldots, \end{cases} \tag{35}$$

and $\lambda = \sum_{j=1}^{\infty} \lambda_j = m(t)$.

A conceptually simpler and computationally more efficient method was proposed by Ma et al. [17]. This method was in effect a rediscovery of a long-known algorithm for the Poisson-stopped-sum distribution [27, pp. 352–353]. The algorithm relies on a simple result which we present for easy reference.

Lemma 2. *If the series on both sides of*

$$\sum_{j=0}^{\infty} p_j z^j = \exp\left(\sum_{j=0}^{\infty} q_j z^j\right) \tag{36}$$

converge for $|z| < \delta$ *for some* $\delta > 0$, *then*

$$p_0 = \exp(q_0), \tag{37}$$

$$p_n = n^{-1}\sum_{j=1}^{n} jq_j p_{n-j} = n^{-1}\sum_{j=0}^{n-1}(n-j)q_{n-j}p_j \quad (n \geqslant 1). \tag{38}$$

Proof. Setting $z = 0$ in Eq. (36) yields Eq. (37); differentiating Eq. (36) with respect to z and then equating coefficients of power of z gives Eq. (38). □

By setting $q_0 = -m(t)$ and identifying q_j $(j = 1, 2\ldots)$ with the λ_j given in Eq. (35), we can also use Lemma 2 to compute the probability distribution induced by the p.g.f. given in Eq. (19).

As the following lemma suggests, the usefulness of the above two lemmas rests on the fact that the Poisson-stopped-sum distribution appears surprisingly often in practice.

Lemma 3. *Let $X(t)$ be a filtered Poisson process constructed by*

$$X(t) = \sum_{i=1}^{N(t)} Y_i(\tau_i, t),$$

where (i) $\{N(t) : t \geqslant 0\}$ *is a Poisson process having piecewise continuous intensity* $v(\cdot)$, (ii) τ_i *are the times of occurrence of events of* $N(t)$, (iii) $\{Y_i(s,t) : s \geqslant 0, t \geqslant 0\}$ $(i = 1, 2, \ldots)$ *are a sequence of independent and statistically identical stochastic processes taking non-negative integer values, and are independent of* $\{N(t); t \geqslant 0\}$. *Then for any given* $t > 0$, $X(t)$ *follows a Poisson-stopped-sum distribution. Specifically, if* $g(z; s, t) = E[z^{Y_1(s,t)}]$ *is the p.g.f. of* $Y_1(s,t)$, *then the p.g.f. of* $X(t)$ *is expressible as*

$$G(z;t) = E[z^{X(t)}] = \exp\{m(t)(h(z;t) - 1)\}, \tag{39}$$

where $m(t) = \int_0^t v(s)\,\mathrm{d}s$ *and where* $h(z;t)$ *is the p.g.f. of some random variable taking non-negative integer values.*

Proof. Applying the lemma in [26], we have

$$G(z;t) = \exp\left\{\int_0^t (g(z;s,t) - 1)v(s)\,\mathrm{d}s\right\}. \tag{40}$$

On the other hand, because $Y_1(s,t)$ takes only non-negative integer values, we can write $g(z;s,t) = \sum_{j=0}^{\infty} p_j(s,t)z^j$, where $p_j(s,t) = P[Y_1(s,t) = j]$. Hence,

$$\begin{aligned}\int_0^t g(z;s;t)v(s)\,\mathrm{d}s &= \int_0^t \left(\sum_{j=0}^{\infty} p_j(s,t)z^j\right)v(s)\,\mathrm{d}s \\ &= \sum_{j=0}^{\infty}\left(\int_0^t p_j(s,t)v(s)\,\mathrm{d}s\right)z^j = m(t)\sum_{j=0}^{\infty}\pi_j(t)z^j,\end{aligned} \tag{41}$$

where the second equality is justified by the fact that $|p_j(s,t)| \leqslant 1$ and where

$$\pi_j(t) = \frac{\int_0^t p_j(s,t)v(s)\,\mathrm{d}s}{\int_0^t v(s)\,\mathrm{d}s}. \tag{42}$$

It is easy to verify that $\pi_j(t) \geqslant 0$ and $\sum_{j\geqslant 0}\pi_j(t) = 1$. Therefore,

$$h(z;t) = \sum_{j=0}^{\infty}\pi_j(t)z^j \tag{43}$$

is the p.g.f. of the discrete distribution $\{\pi_j(t) : j = 0,1,2,\ldots\}$. Inserting Eqs. (41) and (43) into Eq. (40) completes the proof. □.

Although $X(t)$ in Lemma 3 is constructed as a random sum of non-identical random variables, Lemma 3 reveals that it is expressible as a random sum of some independent and identically distributed random variables. These identical random variables can be conveniently considered as 'average' random variables. Some earlier investigators of the Luria–Delbrück model searched for such average random variables from first principles. The advantage of Lemma 3 arises when such average random variables are difficult to know from first principles.

5. The Lea–Coulson formulation

5.1. A mathematical sketch

This formulation originated from the far-reaching paper by Lea and Coulson [4]. It differs from the previous two formulations in that mutant cell growth is described by a stochastic birth process in place of a deterministic growth function. Specifically, if a mutant is generated by a normal cell at time $s > 0$, then at any given time $t \geqslant s$ the size of the clone spawned by that mutant will have the same distribution as $Y(t-s)$, where $\{Y(\tau) : \tau \geqslant 0\}$ is a Yule process having birth rate β_2 and satisfying $Y(0) = 1$. Let $\{Y_i(\tau) : \tau \geqslant 0\}$ $(i = 1,2,\ldots)$ be a sequence of independent copies of such a Yule process. Then the total number of mutants at any given time $t > 0$ is

$$X(t) = \begin{cases} 0, & M(t) = 0, \\ \sum_{i=1}^{M(t)} Y_i(t-\tau_i), & M(t) \geqslant 1, \end{cases} \tag{44}$$

where $M(t)$ signifies the same Poisson process as in Eq. (7).

Let $G(z;t) = E[Z^{X(t)}]$ be the p.g.f. of $X(t)$. Since the p.g.f. of a Yule process having birth rate β_2 is known to be [11, p. 87]

$$y(z;t) = \frac{e^{-\beta_2 t} z}{1 - (1 - e^{-\beta_2 t}) z}, \tag{45}$$

it follows from the filtered Poisson process theory (e.g. the lemma in [26]) that

$$\begin{aligned} \log G(z;t) &= \int_0^t \left(\frac{e^{-\beta_2 (t-s)} z}{1 - (1 - e^{-\beta_2 (t-s)}) z} - 1 \right) \mu e^{\beta_1 s} \, ds \\ &= \mu (z-1) \int_0^t \frac{e^{\beta_1 s}}{1 - (1 - e^{-\beta_2 (t-s)}) z} \, ds \\ &= \mu (z-1) \sum_{k=0}^{\infty} \left\{ \int_0^t (1 - e^{-\beta_2 (t-s)})^k e^{\beta_1 s} \, ds \right\} z^k \\ &= -m(t) + \sum_{k=1}^{\infty} q_k z^k, \end{aligned} \tag{46}$$

where

$$q_k = \mu e^{-\beta_2 t} \int_0^t [1 - e^{-\beta_2 (t-s)}]^{k-1} e^{(\beta_1 + \beta_2) s} \, ds, \tag{47}$$

or, using binomial expansion,

$$q_k = \mu \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \frac{e^{\beta_1 t} - e^{-(j+1)\beta_2 t}}{\beta_1 + (j+1)\beta_2}. \tag{48}$$

That is,

$$G(z;t) = \exp\left(-m(t) + \sum_{k=1}^{\infty} q_k z^k \right). \tag{49}$$

(The dependency of q_k on t is suppressed for simplicity.) Expression (49), first obtained by Stewart et al. [16, Eq. (29)], can be used to compute the probability function by means of Lemma 2. It has been this author's experience that Eq. (47) (with numerical integration) is preferable to Eq. (48) in computing the q_k when k is large. (With large k, Eq. (48) requires adding many quantities of large magnitudes and alternating signs, and the accumulation of rounding errors might swamp the correct answer.) Fig. 1 depicts two probability functions computed using Lemma 2.

Now consider the mean and the variance of $X(t)$. Let $K(\psi, t) = \log\{E[e^{\psi X(t)}]\}$ be the cumulant-generating function of $X(t)$. Applying the random variable technique (see [10, p. 135; 11, pp. 125–129]) readily yields

$$\frac{\partial K}{\partial t} = \beta_2 (e^{\psi} - 1) \frac{\partial K}{\partial \psi} + \mu e^{\beta_1 t} (e^{\psi} - 1). \tag{50}$$

Let $\kappa_j(t)$ $(j \geqslant 1)$ denote the jth cumulant of $X(t)$. By inserting $K(\psi, t) = \sum_{j \geqslant 1} \kappa_j(t) \psi^j / j!$ into Eq. (50) and then equating the coefficients of ψ and ψ^2 on both sides, we obtain

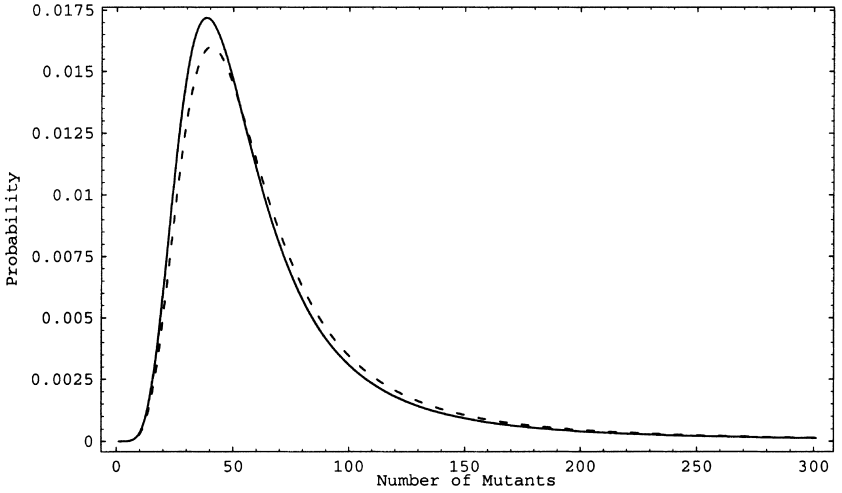

Fig. 1. A comparison of the probability functions between the Lea–Coulson formulation (dashed line) and the discretized Luria–Delbrück formulation (solid line). Parameter values are: $\beta_1 = 3.0$, $\beta_2 = 2.5$, $\mu = 10^{-7}$ and $t = 6.7$.

$$\begin{aligned} \kappa_1'(t) &= \beta_2\kappa_1(t) + \mu e^{\beta_1 t}, \\ \kappa_2'(t) &= \beta_2\kappa_1(t) + 2\beta_2\kappa_2(t) + \mu e^{\beta_1 t}. \end{aligned} \tag{51}$$

Solving Eq. (51) subject to $\kappa_1(0) = \kappa_2(0) = 0$ gives

$$E[X(t)] = \begin{cases} \frac{\mu(e^{\beta_1 t} - e^{\beta_2 t})}{\beta_1 - \beta_2} & (\beta_1 \neq \beta_2), \\ \mu t e^{\beta_1 t} & (\beta_1 = \beta_2) \end{cases} \tag{52}$$

and

$$\mathrm{Var}[X(t)] = \begin{cases} \frac{\mu e^{\beta_2 t}(\beta_1(1 + e^{(\beta_1 - \beta_2)t} - 2e^{\beta_2 t}) + 2\beta_2(e^{\beta_2 t} - 1))}{(\beta_1 - \beta_2)(\beta_1 - 2\beta_2)} & (\beta_1 \neq \beta_2,\ \beta_1 \neq 2\beta_2), \\ \frac{\mu}{\beta_2} e^{\beta_2 t}(1 - e^{\beta_2 t} + 2\beta_2 t e^{\beta_2 t}) & (\beta_1 = 2\beta_2), \\ 2\frac{\mu}{\beta_1} e^{\beta_1 t}(e^{\beta_1 t} - 1) - \mu t e^{\beta_1 t} & (\beta_1 = \beta_2). \end{cases} \tag{53}$$

Eqs. (52) and (53) were obtained for the special case $\beta_1 = \beta_2$ by Bartlett [9, p. 118], Bailey [11, p. 128], and Crump and Hoel [12, p. 244]; the general expressions have hitherto not appeared.

A comparison of Eqs. (10) and (11) with Eqs. (52) and (53) indicates that the means under both the Luria–Delbrück formulation and the Lea–Coulson formulation are the same, but the variances differ. In the special case $\beta_1 = \beta_2$, the variance determined by the Lea–Coulson formulation is roughly twice the variance determined by the Luria–Delbrück formulation. This relation can be

deducted as follows. From Eq. (11) we get $\mathrm{Var}[X(t)] = N(t)m(t)$; from the last case of Eq. (53) we have $\mathrm{Var}[X(t)] \approx 2N(t)m(t)$. Coulson was alerted to this discrepancy by Kendall (see the appendix in [4]), presumably after the untimely death of Lea in June 1947. Coulson pointed out that Eq. (30) in Lea and Coulson [4] "was copied from Luria and Delbrück, is not quite valid". It is now easy to see that the formula for variance by Luria and Delbrück was indeed valid for the intended Luria–Delbrück formulation, but it was not applicable to the Lea–Coulson formulation.

5.2. The Lea–Coulson probability generating function

We shall devote the next four subsections to the special case $\beta_1 = \beta_2$. For convenience, we denote the common value of β_1 and β_2 by β. In the present subsection we limit our attention to the p.g.f. discovered by Lea and Coulson [4].

To begin with, the partial differential equation (PDE) for the p.g.f. of $X(t)$ is

$$\frac{\partial G}{\partial t} = \beta z(z-1)\frac{\partial G}{\partial z} + \mu \mathrm{e}^{\beta t}(z-1)G, \tag{54}$$

with initial condition

$$G(z;0) = 1. \tag{55}$$

Lea and Coulson [4] noted that the initial condition (55) can be replaced with

$$G(0;t) = p_0(t) = \mathrm{e}^{-m(t)}. \tag{56}$$

By use of the time scale transform

$$\theta = \frac{\mu}{\beta}\mathrm{e}^{\beta t}, \tag{57}$$

we can transform Eq. (54) into

$$\frac{\partial F}{\partial \theta} = \frac{1}{\theta} z(z-1)\frac{\partial F}{\partial z} + (z-1)F. \tag{58}$$

(There are misprints in the renditions of Eq. (58) by Kemp [19, Eq. (6)] and Sarkar et al. [18, Eq. (16)].) Thus G and F are related by $F(z;\theta) = G(z;\beta^{-1}\log(\beta\theta/\mu))$. Lea and Coulson [2] discovered a particular solution to Eq. (58) in the form

$$F(z;\theta) = (1-z)^{\theta(1-z)/z}. \tag{59}$$

This particular solution $F(z;\theta)$ was often recast in the convenient form

$$F(z;\theta) = \exp\left[\theta(f(z)-1)\right] \tag{60}$$

with

$$f(z) = 1 + \left(\frac{1}{z} - 1\right)\log(1-z)$$
$$= \sum_{i=1}^{\infty}\left(\frac{1}{i} - \frac{1}{i+1}\right)z^i. \tag{61}$$

The transform in Eq. (57) was due to Lea and Coulson [4]. Lea and Coulson applied this transform to the PDE for $e^{\theta}G$ and arrived at an equation analogous to Eq. (58). Bartlett [10] was the first to derive Eq. (58) using the transform of Lea and Coulson. Ma et al. [17] gave a derivation of Eq. (58) in a manner resembling more closely the quaint style of Lea and Coulson. We shall call Eq. (58) the Lea–Coulson equation, and the particular solution given in Eq. (59) the Lea–Coulson p.g.f., on the basis of their profound historical impact. It is worthy of remark that the Lea–Coulson p.g.f. is not exactly the p.g.f. under the Lea–Coulson formulation, even though it satisfies the Lea–Coulson equation. This assertion can be appreciated from two perspectives. First, the Lea–Coulson p.g.f. satisfies $F(0;\theta) = e^{-\theta}$. Using the original time scale, we express this condition as

$$G(0;t) = p_0(t) = e^{-\theta(t)}. \tag{62}$$

A comparison of Eq. (62) with Eq. (56) reveals that the error of the Lea–Coulson p.g.f. springs from an incorrect initial condition. Second, the Lea–Coulson p.g.f. satisfies $F(z;0) = 1$ which, when expressed in the original time scale, is equivalent to

$$G(z;-\infty) = 1. \tag{63}$$

Since the intensity function given in Eq. (2) is defined for $t \in (-\infty, \infty)$, a comparison of Eq. (63) with Eq. (55) suggests that the Lea–Coulson p.g.f. can be considered as the p.g.f. of a similar process starting at $-\infty$. A moment's thought will then convince us that it is not at all surprising that we find divergent moments at a finite time.

From Eq. (46) we find easily that the exact p.g.f. for the present case is

$$G(z;\theta,\phi) = \exp\left\{\theta\left(\frac{1}{z} - 1\right)\log(1-\phi z)\right\}, \tag{64}$$

which is usually written as

$$G(z;\theta,\phi) = (1-\phi z)^{\theta(1-z)/z}. \tag{65}$$

Because $\phi \approx 1$ for large t, the Lea–Coulson p.g.f. was often considered as a limiting form of the exact p.g.f. Some remarks are in order. First, if $t \to \infty$, then $\theta \to \infty$. Hence the approximate distribution cannot be interpreted as an asymptotic distribution for large t in the usual sense. Bartlett suggested that the size of the normal cell population $N(t)$ would stabilize for large t [10, p. 137]

and hence $\theta(t)$ was approximately a constant. This biologically plausible assumption cannot be inferred from the model itself – $N(t)$ increases exponentially with time according to the model. Next, since the exact p.g.f. is almost as amenable as the Lea–Coulson p.g.f., use of the exact p.g.f. should be encouraged. Even the extra parameter ϕ appearing in Eq. (65) is not a major concern, because it can be inferred from the number of normal cells $N(t)$, which is usually known in practice. Finally, as indicated in Section 4.1, the Lea–Coulson p.g.f. gives rise to infinite moments. This side effect was once a source of confusion and was deemed 'awkward' by some [14].

The exact p.g.f. given in Eq. (65) first appeared in print in Armitage [5, Eq. (30a)] (communicated to him by Bartlett). This p.g.f. might be the same p.g.f. communicated to Coulson by Kendall (see the appendix in [4]). Derivation details were given first by Bartlett [10] and then by Bailey [11]. Bailey [11] worked directly on Eq. (54); Bartlett tackled the transformed equation, the Lea–Coulson equation. Crump and Hoel [12] were the first to give a derivation based on the filtered Poisson process theory. (Their p.g.f. presented in Eq. (2.17) contains a misprint.)

If the process starts with n_0 normal cells at time $t = 0$, Eqs. (64) and (65) for the p.g.f. are still valid, so long as the mutation rate μ is replaced with $n_0\mu$. The reasons are the same as explained in Section 3.

5.3. The distribution induced by the equal growth case

We say a random variable X has an $\mathrm{LD}(\theta, \phi)$ distribution, if the p.g.f. of X, denoted by $G(z; \theta, \phi) = E[z^X]$, is given by Eq. (64). It is evident from Eqs. (4) and (6) that $G(z; \theta, \phi)$ is an legitimate p.g.f. for $\theta > 0$ and $0 \leqslant \phi < 1$. Moreover, $\phi = 1$ is also admissible because $\mathrm{LD}(\theta, 1)$ represents the distribution induced by the Lea–Coulson p.g.f. An $\mathrm{LD}(\theta, \phi)$ distribution is infinitely divisible. Moreover, if X_i $(i = 1, 2, \ldots, n)$ are independent $\mathrm{LD}(\theta_i, \phi)$ random variables, then $\sum_{i=1}^{n} X_i$ is an $\mathrm{LD}(\sum_{i=1}^{n} \theta_i, \phi)$ random variable.

Let $m = \theta\phi$. Rewriting Eq. (64) as

$$G(z; \theta, \phi) = \exp\left(-m + \sum_{i=1}^{\infty} \theta \left(\frac{\phi^i}{i} - \frac{\phi^{i+1}}{i+1} \right) z^i \right) \tag{66}$$

and applying Lemma 2, we get

$$\begin{aligned} p_0 &= \mathrm{e}^{-m}, \\ p_n &= \frac{\theta}{n} \sum_{j=0}^{n-1} \phi^{n-j} \left(1 - \frac{(n-j)\phi}{n-j+1} \right) p_j \quad (n \geqslant 1) \\ &= \frac{\theta}{n} \sum_{j=1}^{n} \phi^j \left(1 - \frac{j\phi}{j+1} \right) p_{n-j} \quad (n \geqslant 1). \end{aligned} \tag{67}$$

This recurrence relation was first noticed for the special case $\phi = 1$ by Ma et al. [17], and for the general case by Sarkar et al. [18].

By differentiating Eq. (64) with respect to z repeatedly, we find, if $X \sim \mathrm{LD}(\theta, \phi)$, then

$$E[X] = -\theta \log(1-\phi), \tag{68}$$

$$\mathrm{Var}[X] = \theta\left[\frac{2\phi}{1-\phi} + \log(1-\phi)\right], \tag{69}$$

$$\alpha_3(X) = \frac{\theta[3\phi^2 - (1-\phi)^2 \log(1-\phi)]}{(1-\phi)^2(\mathrm{Var}[X])^{3/2}}, \tag{70}$$

$$\alpha_4(X) = 3 + \frac{2\phi(1+\phi+2\phi^2) + (1-\phi)^3 \log(1-\phi)}{\theta(1-\phi)[2\phi + (1-\phi)\log(1-\phi)]^2}. \tag{71}$$

Eqs. (70) and (71) are expressions for skewness and kurtosis. The derivation of these two identities entails an inordinate amount of tedious algebra; they were derived with the help of *Mathematica* [28]. In the numerical example of Armitage [29, p. 179], $N(t) = 3.1 \times 10^8$ and the parameter θ was estimated to be 8.30. Therefore, $\phi = 1 - 3.2 \times 10^{-9} \approx 1$, and $\mathrm{LD}(8.30, 1)$ should be a good approximation to the exact distribution. But we must rely on the exact distribution to find the moments. Using Eqs. (68)–(71), we get

$$E[X] = 162.28, \qquad \mathrm{Var}[X] = 3.15 \times 10^9,$$
$$\alpha_3(X) = 6482.14, \qquad \alpha_4(X) = 7.47 \times 10^7.$$

We now consider the limiting behavior of an $\mathrm{LD}(\theta, \phi)$ distribution when the parameter ϕ approaches zero. In view of Eqs. (5) and (64), it follows from the elementary relation $\lim_{\phi \to 0} \log(1-\phi z)/\phi = -z$ that

$$\lim_{\phi \to 0} \log G(z; \theta, \phi) = m\left(\frac{1}{z} - 1\right) \lim_{\phi \to 0} \frac{\log(1-\phi z)}{\phi} = m(z-1). \tag{72}$$

That is, the limiting distribution of an $\mathrm{LD}(\theta, \phi)$ is a Poisson distribution with mean $m = \theta\phi$. This new finding is noteworthy in that the Luria–Delbrück model was originally intended to detect deviations from the Poisson distribution.

Of tremendous importance is the limiting behavior of the individual probability $p_n = \Pr[X = n]$. A number of authors have investigated the asymptotic behavior of p_n for the case $X \sim \mathrm{LD}(\theta, 1)$. Ma et al. [17] proved that $p_n \approx c/n^2$ for some constant c and numerically verified that $c = 1$ in the case $\theta = 1$. Appealing to the theory of convolution powers of subexponential laws, Pakes [20] was the first to refine the finding of Ma et al. by showing that

$$p_n \approx \frac{\theta}{n^2}. \tag{73}$$

Using an elementary approach, Kemp [19] also proved result (73) and deduced from it what was previously conjectured by Cairns et al. [14]:

$$\sum_{j \geqslant n} p_j \approx \frac{\theta}{n}. \tag{74}$$

Goldie [21] established the validity of Eq. (73) by use of properties of infinitely divisible sequences. The latest proof of Eq. (73) was given by Prodinger [22] by means of singularity analysis of generating functions. Among the four proofs of Eq. (73), only the one by Kemp is self-contained in the sense that it does not rely on any advanced results from a highly specialized branch of mathematics. Small wonder that Kemp's proof is the longest of the four. We can render Kemp's proof considerably shorter and arguably more lucid by some modifications. The refined proof comes about as follows.

In the notation of Eqs. (60) and (61) we have

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\{zF(z;\theta)\} = \theta\big(2f'(z) + zf''(z)\big)F(z;\theta) + \theta^2 z f'(z)^2 F(z;\theta). \tag{75}$$

Simple algebraic work gives

$$2f'(z) + zf''(z) = \sum_{n=0}^{\infty} z^n,$$

$$zf'(z)^2 = \sum_{n=0}^{\infty} T_n z^{n+1}$$

with

$$T_n = \sum_{i=0}^{n} \frac{1}{(\mathrm{i}+2)(n-\mathrm{i}+2)}.$$

Substituting these relations into Eq. (75) and equating coefficients of z^n, we arrive at

$$n(n+1)p_n = \theta(p_0 + \cdots + p_n) + \theta^2(p_0 T_{n-1} + \cdots + p_{n-1}T_0). \tag{76}$$

Now denote the nth harmonic number $1 + 1/2 + \cdots + 1/n$ by h_n and observe that

$$T_n = \frac{1}{n+4}\sum_{i=0}^{n}\left(\frac{1}{\mathrm{i}+2} + \frac{1}{n-\mathrm{i}+2}\right)$$

$$= \frac{2(h_{n+2}-1)}{n+4} = \mathrm{O}(n^{-1}\log(n)).$$

So $T_n \to 0$ as $n \to \infty$. Since $\{p_n\}$ is a proper probability sequence ($F(1;\theta) \equiv 1$), $\lim_{n\to\infty}(p_0 + \cdots + p_n) = 1$. Therefore, from elementary calculus we know that $\lim_{n\to\infty} T_n = 0$ implies

$$\lim_{n\to\infty}(p_0 T_{n-1} + \cdots + p_{n-1}T_0) = 0.$$

(It is also a special case of [30, Theorem 2.5.5, p. 47].) Therefore, it follows from Eq. (76) that

$$\lim_{n\to\infty} n(n+1)p_n = \theta, \tag{77}$$

which is equivalent to $n^2 p_n \to \theta$. The proof is thus complete.

With the help of a computer algebra system and the theory of singularity analysis, Prodinger [22] was able to refine Eqs. (73) and (74):

$$\begin{aligned} p_n &= \frac{\theta}{n^2} + 2\theta^2 \frac{\log(n)}{n^3} + \frac{(2\gamma-3)\theta^2 - \theta}{n^3} + \mathrm{O}\left(\frac{\log^2(n)}{n^4}\right), \\ \sum_{j\geqslant n} p_j &= \frac{\theta}{n} + \theta^2 \frac{\log(n)}{n^2} + \frac{(\gamma-1)\theta^2 - \theta}{n^2} + \mathrm{O}\left(\frac{\log^2(n)}{n^3}\right), \end{aligned} \tag{78}$$

where $\gamma = 0.577216\ldots$ denotes Euler's constant.

Pakes [20] was the only one who found the asymptotic behavior of p_n for the $\mathrm{LD}(\theta,\phi)$ distribution with $0 < \phi < 1$. Pakes gave

$$p_n \sim \frac{\phi^n n^{\theta(1-\phi)-1}}{\Gamma(\theta(1-\phi))} \tag{79}$$

and thus suggested that the asymptotic behavior of p_n be made more tangible by linking it with that of a negative binomial distribution. If Z has a negative binomial distribution $\mathrm{NB}(\theta(1-\phi), 1-\phi)$, then it can be easily verified using Stirling's formula [31, Eq. (1.4.25)] that for large n

$$\Pr[Z=n] \approx (1-\phi)^{\theta(1-\phi)} \frac{\phi^n n^{\theta(1-\phi)-1}}{\Gamma(\theta(1-\phi))}.$$

Thus, the tail behavior of an $\mathrm{LD}(\theta,\phi)$ distribution for $\phi \in (0,1)$ is proportional to that of a negative binomial distribution, and hence is quite different from that of an $\mathrm{LD}(\theta,1)$ distribution. In other words, the conjecture of Cairns et al. [14] does not hold in terms of the exact distribution.

Finally, we present some miscellaneous new results inspired by the Lea–Coulson algorithm which we shall discuss in Section 5.4. By expanding the p.g.f. in Eq. (66) in the form

$$G(z;\theta,\phi) = \mathrm{e}^{-m} \prod_{i=1}^{\infty} \left\{ \sum_{k=0}^{\infty} \left(\frac{\phi^i}{i} - \frac{\phi^{i+1}}{i+1} \right)^k z^{ik} \frac{\theta^k}{k!} \right\},$$

we deduce that p_n is expressible as

$$p_n = \mathrm{e}^{-m} \sum_{j=1}^{n} C_{j,n} \frac{\theta^j}{j!} \quad (n \geqslant 1), \tag{80}$$

where $C_{j,n}$ are independent of θ. It is clear that

$$C_{1,n} = \frac{\phi^n}{n} - \frac{\phi^{n+1}}{n+1},$$
$$C_{n,n} = \left(\phi - \frac{\phi^2}{2}\right)^n. \tag{81}$$

Therefore, for large θ, we have

$$p_n \approx \mathrm{e}^{-\theta\phi}\left(1 - \frac{\phi}{2}\right)^n \frac{\phi^n \theta^n}{n!}, \tag{82}$$

and for small θ,

$$p_n \approx \mathrm{e}^{-\theta\phi}\left(\frac{1}{n} - \frac{\phi}{n+1}\right)\phi^n \theta. \tag{83}$$

5.4. The Lea–Coulson algorithm

Lea and Coulson [4] devised the oldest algorithm for computing (approximately) the probability function of the number of mutants. When $\beta_1 = \beta_2$ (which we denote by β), the process $\{X(t) : t \geqslant 0\}$ is defined by

$$\Pr[X(t+\Delta t) - X(t) = 1 | X(t) = n] = (n\beta + \mu \mathrm{e}^{\beta t})\Delta t + \mathrm{o}(\Delta t).$$

Thus $p_i(t) = \Pr[X(t) = i]$ $(i = 0, 1, \ldots)$ satisfy (cf. [11, Eq. (8.18)])

$$\frac{\mathrm{d}p_0(t)}{\mathrm{d}t} = -\mu \mathrm{e}^{\beta t} p_0(t), \tag{84}$$
$$\frac{\mathrm{d}p_n(t)}{\mathrm{d}t} = ((n-1)\beta + \mu \mathrm{e}^{\beta t})p_{n-1}(t) - (n\beta + \mu \mathrm{e}^{\beta t})p_n(t) \quad (n \geqslant 1). \tag{85}$$

Employing the same rescaling of time as in Eq. (57), we transform Eq. (85) to

$$\frac{\mathrm{d}r_n(\theta)}{\mathrm{d}\theta} = \left(1 + \frac{n-1}{\theta}\right) r_{n-1}(\theta) - \left(1 + \frac{n}{\theta}\right) r_n(\theta), \tag{86}$$

where $r_n(\theta) = p_n(\beta^{-1} \log(\beta\theta/\mu))$. Defining $q_n(\theta) = \mathrm{e}^{\theta} r_n(\theta)$, we rewrite Eq. (86) as

$$\frac{\mathrm{d}q_n(\theta)}{\mathrm{d}\theta} + \frac{n}{\theta} q_n(\theta) = \left(1 + \frac{n-1}{\theta}\right) q_{n-1}(\theta), \tag{87}$$

which is Eq. (6) of [4]. If $\phi = 1$, then $m = \theta$ and Eq. (80) can be rewritten as

$$q_n(\theta) = \sum_{j=1}^{n} C_{j,n} \frac{\theta^j}{j!} \quad (n \geqslant 1). \tag{88}$$

In other words, $q_n(\theta)$ is a degree n polynomial function of θ when $m = \theta$. Substituting Eq. (88) into Eq. (87) and equating coefficients of powers of θ on both sides yield a recurrence relation

$$
\begin{aligned}
C_{j,n} &= \frac{j}{j+n} C_{j-1,n-1} + \frac{n-1}{j+n} C_{j,n-1}, \\
C_{1,n} &= \frac{1}{n(n+1)}, \\
C_{n,n} &= \frac{1}{2^n}.
\end{aligned} \tag{89}
$$

From Eqs. (88) and (89), q_n (and hence p_n) can be easily computed for $n \geqslant 1$, the case $q_0(\theta) \equiv 1$ being trivial. On the other hand, if $\phi \neq 1$, it is clear that the last two relations in Eq. (89) should be replaced with Eq. (81). But Eq. (88) no longer holds when $\phi \neq 1$. Thus, for $\phi \approx 1$, the Lea–Coulson algorithm gives approximate results.

Koch [13] made an attempt to generalize the Lea–Coulson algorithm to the case of differential growth. Defining $r_n(\theta) = p_n(1/\beta_1 \log(\beta_1\theta/\mu))$ and $q_n(\theta) = \mathrm{e}^{\theta} r_n(\theta)$, we have

$$
\frac{\mathrm{d}q_n(\theta)}{\mathrm{d}\theta} + \frac{n}{\theta}\rho q_n(\theta) = \left(1 + \frac{n-1}{\theta}\rho\right) q_{n-1}(\theta), \tag{90}
$$

where $\rho = \beta_2/\beta_1$. By inserting Eq. (88) into Eq. (90) and equating coefficients of powers of θ, we obtain Koch's recurrence relation:

$$
\begin{aligned}
C_{j,n} &= \frac{j}{j+n\rho} C_{j-1,n-1} + \frac{(n-1)\rho}{j+n\rho} C_{j,n-1}, \\
C_{n,n} &= \frac{1}{(1+\rho)^n}, \\
C_{1,n} &= \frac{(n-1)\rho}{1+n\rho} C_{1,n-1}.
\end{aligned} \tag{91}
$$

It is worth noting that the use of the recurrence relation (91) relies on the assumption that $q_n(\theta)$ be expressible as a polynomial function of θ. As we just indicated, in the case where $\beta_1 = \beta_2$, this assumption is approximately true when $N(t)$ is large (or $\phi \approx 1$). In the case of differential growth, the condition $\beta_1 \approx \beta_2$ also seems necessary for that assumption to hold approximately (compare Eq. (49) with Eq. (66)). Fig. 2 provides a numerical example to illustrate this point.

6. The Bartlett formulation

Bartlett first proposed this formulation in 1951 while discussing the paper by Armitage (see discussion [5, p. 37]). This fully stochastic formulation of the Luria–Delbrück model is a two-dimensional birth process $\{(X_1(t), X_2(t)): t \geqslant 0\}$, where $X_1(t)$ and $X_2(t)$ represent the population size at time t of the

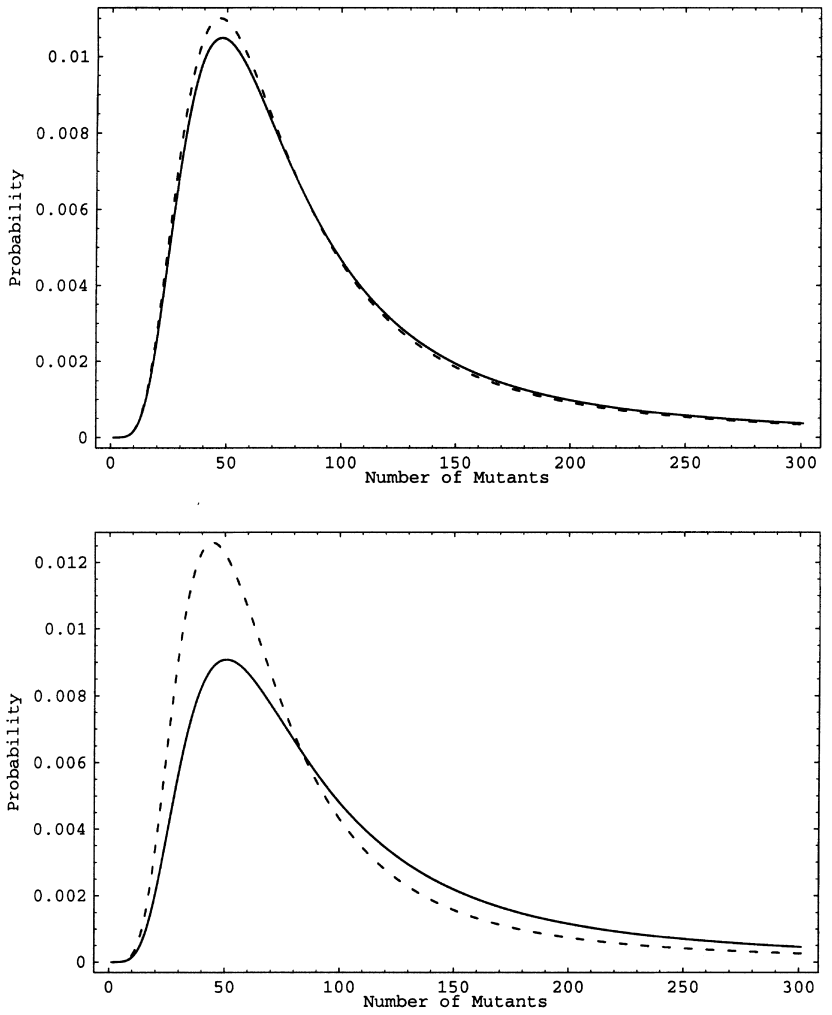

Fig. 2. The Lea–Coulson algorithm (solid line) is compared with the algorithm based on Lemma 2 (dashed line). In the upper panel, $\beta_1 = 3.0,\ \beta_2 = 2.97,\ \mu = 10^{-7}$ and $t = 6.7$. Because β_1/β_2 is close to 1, the two algorithms are roughly in agreement. In the lower panel, β_2 is changed to 2.80, and the discrepancy becomes more pronounced.

normal cells and that of the mutant cells, respectively. The formulation is called fully stochastic because it models the growth of both the normal cells and the mutant cells by stochastic growth processes. In other words, both Assumptions 2 and 3 in Section 3 are relaxed by this formulation. Fig. 3 captures the salient features of this formulation. Note in particular that the occurrence of mutations is modeled with Kendall's Formulation B. The Bartlett formulation can

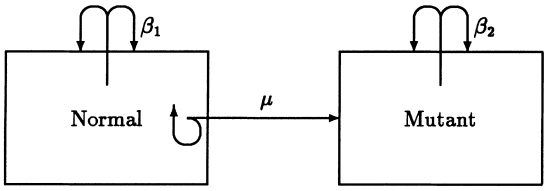

Fig. 3. The Bartlett formulation of the Luria–Delbrück model.

be characterized by the following two transitions that can happen in an infinitesimal time interval of length Δt:

$$\begin{aligned}&\Pr[(X_1,X_2)\to(X_1+1,X_2)]=\beta_1X_1\Delta t+\mathrm{o}(\Delta t),\\&\Pr[(X_1,X_2)\to(X_1,X_2+1)]=(\mu X_1+\beta_2X_2)\Delta t+\mathrm{o}(\Delta t).\end{aligned}\tag{92}$$

As explained in Section 5.1, the PDE of the joint p.g.f. of $X_1(t)$ and $X_2(t)$ can be readily written as

$$\frac{\partial G}{\partial t}=\{\beta_1z_1(z_1-1)+\mu z_1(z_2-1)\}\frac{\partial G}{\partial z_1}+\beta_2z_2(z_2-1)\frac{\partial G}{\partial z_2},\tag{93}$$

with initial condition

$$G(z_1,z_2;0)=z_1.\tag{94}$$

In a similar vein, the PDE of the cumulant-generating function is given by

$$\frac{\partial K}{\partial t}=\{\beta_1(\mathrm{e}^{\theta_1}-1)+\mu(\mathrm{e}^{\theta_2}-1)\}\frac{\partial K}{\partial\theta_1}+\beta_2(\mathrm{e}^{\theta_2}-1)\frac{\partial K}{\partial\theta_2}.\tag{95}$$

with initial condition

$$K(\theta_1,\theta_2;0)=\theta_1.\tag{96}$$

Let $\kappa_{i,j,}(t)$ $(i+j\geqslant 1)$ be the (i,j)th joint cumulant of $X_1(t)$ and $X_2(t)$. Expanding $K(\theta_1,\theta_2;t)$ in θ_1 and θ_2 as

$$K(\theta_1,\theta_2;t)=\sum_{i+j\geqslant 1}\kappa_{i,j}(t)\frac{\theta_1}{i!}\frac{\theta_2}{j!}$$

and inserting it into Eq. (95), we can equate coefficients of θ_1, θ_2, θ_1^2, $\theta_1\theta_2$ and θ_2^2 to obtain a system of ordinary differential equations (ODEs):

$$\begin{aligned}&\kappa'_{1,0}(t)=\beta_1\kappa_{1,0}(t),\\&\kappa'_{0,1}(t)=\beta_2\kappa_{0,1}(t)+\mu\kappa_{1,0}(t),\\&\kappa'_{2,0}(t)=\beta_1\kappa_{1,0}(t)+2\beta_1\kappa_{2,0}(t),\\&\kappa'_{1,1}(t)=(\beta_1+\beta_2)\kappa_{1,1}(t)+\mu\kappa_{2,0}(t),\\&\kappa'_{0,2}(t)=\beta_2\kappa_{0,1}(t)+2\beta_2\kappa_{0,2}(t)+\mu\kappa_{1,0}(t)+2\mu\kappa_{1,1}(t).\end{aligned}\tag{97}$$

These equations can be solved exactly, using the initial condition that when $t = 0$, $\kappa_{1,0} = 1$, all other cumulants being zero. By solving just the first two equations we can get an expression for $E[X_2(t)] = \kappa_{0,1}(t)$ and find it to be the same as the corresponding expression under the Lea–Coulson formulation as given in Eq. (52). Expressions for the second-order cumulants are a little too cumbersome to be given here. But it is easy to see that if the term $2\mu\kappa_{1,1}(t)$ were deleted from the last equation in Eq. (97), then the resulting equation is equivalent to the equation of $\kappa_2(t)$ given in Eq. (51). It can also be inferred from the fourth equation in Eq. (97) that $\kappa_{1,1}(t) > 0$ for all $t > 0$. Therefore, we conclude that $\kappa_{0,2}(t) > \kappa_2(t)$ for all $t > 0$, in agreement with intuition.

We now derive the p.g.f. of $X_2(t)$ by following the ingenious approach of Bartlett [9, pp. 115–118; 10, pp. 132–136]. Bartlett derived the joint p.g.f. of $X_1(t)$ and $X_2(t)$ first, and then extracted the p.g.f. of $X_2(t)$ from the joint p.g.f. For simplicity we shall start directly from the p.g.f. of $X_2(t)$. (Some ideas in the following derivation are also drawn from Kendall [23] and Puri [32].)

Let T denote the waiting time for the first transition, which necessarily occurs in the first compartment (see Fig. 3). Clearly, T is exponentially distributed with mean $1/(\beta_1 + \mu)$. The first transition is either a division of the initial normal cell or a mutation (the initial normal cell splits into a normal daughter cell and a mutant daughter cell). Furthermore, given that the first transition does occur, the probability of its being a cellular division is $\beta_1/(\beta_1 + \mu)$, and its being a mutation is $\mu/(\beta_1 + \mu)$. If $g(z;t) = E[z^{X_2(t)}]$ denotes the p.g.f. of $X_2(t)$, then

$$E[z^{X_2(t)}|T = \tau] = \begin{cases} 1 & (\tau > t), \\ \frac{\beta_1}{\beta_1+\mu} g(z;t-\tau)^2 + \frac{\mu}{\beta_1+\mu} y(z;t-\tau)g(z;t-\tau) & (\tau \leqslant t), \end{cases} \tag{98}$$

where $y(z;t)$ is the p.g.f. of the Yule process defined in Eq. (45). Therefore, the p.g.f. of $X_2(t)$ is

$$\begin{aligned} g(z;t) &= \int_0^\infty E[z^{X_2(t)}|T = \tau](\beta_1 + \mu)\mathrm{e}^{-(\beta_1+\mu)\tau}\,\mathrm{d}\tau \\ &= \int_t^\infty (\beta_1 + \mu)\mathrm{e}^{-(\beta_1+\mu)\tau}\,\mathrm{d}\tau \\ &\quad + \int_0^t [\beta_1 g(z;t-\tau)^2 + \mu y(z;t-\tau)g(z;t-\tau)]\mathrm{e}^{-(\beta_1+\mu)\tau}\,\mathrm{d}\tau \\ &= \mathrm{e}^{-(\beta_1+\mu)t} + \int_0^t [\beta_1 g(z;s)^2 + \mu y(z;s)g(z;s)]\mathrm{e}^{-(\beta_1+\mu)(t-s)}\,\mathrm{d}s. \end{aligned}$$

The last equality is obtained by applying the change of variable $s = t - \tau$. Multiplying both sides by $\mathrm{e}^{(\beta_1+\mu)t}$ and suppressing the dependency of g and y on z give

$$e^{(\beta_1+\mu)t}g(t) = 1 + \int_0^t [\beta_1 g(s)^2 + \mu y(s)g(s)]e^{(\beta_1+\mu)s}\,ds.$$

Differentiating with respect to t and rearranging, we get

$$g'(t) = \beta_1 g(t)^2 - (\beta_1 + \mu - \mu y(t))g(t). \tag{99}$$

Eq. (99) is subject to the initial condition $g(0) = 1$.

Dividing both sides of Eq. (99) by $-g(t)^2$ and introducing $R(t) = 1/g(t)$, we obtain

$$R'(t) = (\beta_1 + \mu - \mu y(t))R(t) - \beta_1. \tag{100}$$

This equation is subject to $R(0) = 1$. Note that Eq. (100) depends on β_2 through $y(t)$ given in Eq. (45).

Because a mutation is modeled by Kendall's Formulation B, the condition of equal growth between normal cells and mutants is interpreted as $\beta_1 + \mu = \beta_2$. It is under this condition that Eq. (100) admits a simple solution

$$g(t) = \frac{1}{R(t)} = \frac{ze^{-\beta t}}{ze^{-\beta t} + 1 - z - (1-z)(ze^{-\beta t} + 1 - z)^p}. \tag{101}$$

Here we set $\beta = \beta_2$ and $p = \mu/(\beta_1 + \mu)$ for convenience. This p.g.f. first appeared without a derivation in Bartlett's discussion on Armitage's paper [5]. A derivation first appeared in the well-known text of Bartlett [9, p. 112] about three years later.

Differentiating the p.g.f. in Eq. (101) with respect to z, Bartlett also found the mean and the variance:

$$E[X(t)] = e^{\beta t} - e^{\beta(1-p)t}, \tag{102}$$

$$\mathrm{Var}[X(t)] = e^{\beta(1-p)t} - e^{\beta t} + e^{2\beta t}(e^{-\beta p t} - 1)^2 + 2pe^{\beta(2-p)t} - 2pe^{\beta(1-p)t}. \tag{103}$$

Note that the probability of zero mutants, $p_0(t) = \Pr[X_2(t) = 0]$, is clearly independent of the division rate of mutants, β_2. This observation allows us to extract a general expression for $p_0(t)$ from the p.g.f. in Eq. (101), even though the p.g.f. was derived under the assumption $\beta_1 + \mu = \beta_2$. Letting z approach zero in Eq. (101) gives

$$p_0(t) = \frac{\beta_1 + \mu}{\beta_1 + \mu e^{(\beta_1+\mu)t}}. \tag{104}$$

This formula, essential to the P_0 method, is presented for the first time here. If there are n_0 normal cells at $t = 0$, then $p_0(t)^{n_0}$ will be the desired probability. Fig. 4 presents a comparison of a $p_0(t)$ determined by the Bartlett formulation with the corresponding quantities determined by the other three formulations.

Bartlett [10, p. 134] has shown that under appropriate conditions the p.g.f. in Eq. (101) can be approximated by the Lea–Coulson p.g.f. Recall from elementary calculus that (1) if $x \approx 1$, then $x \approx 1 + \log x$ and (2) for $|x| < n$ and n

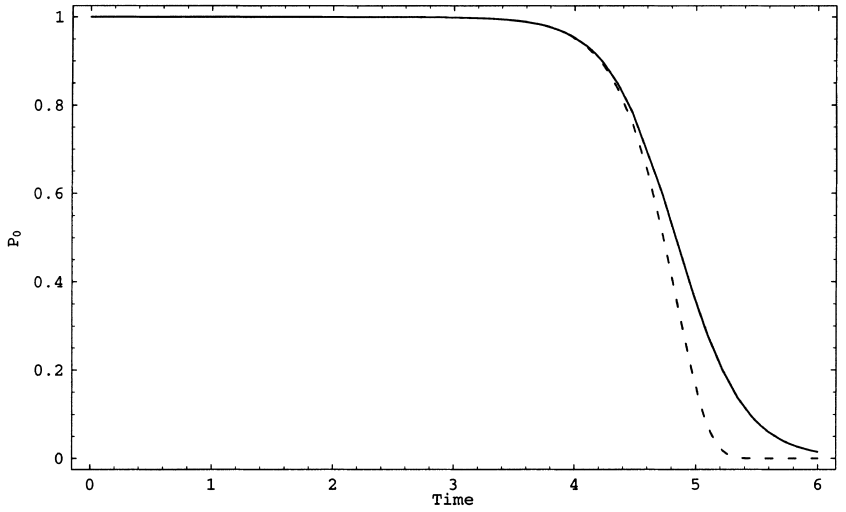

Fig. 4. The P_0 method for estimating mutation rates depends on the quantity $p_0(t)$. This figure depicts $p_0(t)$ with parameter values $\beta_1 = 3.6$ and $\mu = 10^{-7}$. Note that $p_0(t)$ is independent of β_2. The dashed line represents the Luria–Delbrück formulation, the discretized Luria–Delbrück formulation and the Lea–Coulson formulation; $p_0(t)$ is the same among these three formulations. The solid line gives $p_0(t)$ for the Bartlett formulation.

large, $(1 - x/n)^{-n} \approx e^x$. Now assume that the process starts with a large number of normal cells, say n_0. Assume further that p is small and t is large. We then have

$$\begin{aligned} g(t)^{n_0} &= \left\{1 + \frac{1-z}{z} e^{\beta t}\left(1 - [1 - (1 - e^{-\beta t})z]^p\right)\right\}^{-n_0} \\ &\approx \left\{1 - \frac{1-z}{z} e^{\beta t} p \log\left(1 - (1 - e^{-\beta t})z\right)\right\}^{-n_0} \\ &\approx \exp\left\{\Theta\left(\frac{1}{z} - 1\right)\log(1-z)\right\}, \end{aligned}$$

where $\Theta = n_0 p\, e^{\beta t}$.

7. Phenotypic delay

Under both the discretized Luria–Delbrück formulation and the Lea–Coulson formulation, the p.g.f. of the number of mutants at any given time $t > 0$ is expressible in the form (cf. Eqs. (19) and (49))

$$G(z;t) = \exp\left\{m(t)\left(\sum_{j=1}^{\infty} q_j z^j - 1\right)\right\}. \tag{105}$$

Thus the number of mutants has the same distribution as $X_1 + \cdots + X_{M(t)}$ where $M(t)$ is a Poisson process having mean $m(t)$, independent of all X_i, and where $X_1, X_2, \ldots$ are independent and identically distributed random variables satisfying $\Pr[X_1 = j] = q_j$. Because $m(t)$ in Eq. (105) coincides with the expected number of mutations occurring in the time interval $[0, t)$ (cf. Lemma 3), it is intuitively appealing to consider each X_i as the average size of the mutant clone spawned by the ith mutation. For this reason, Armitage [5] suggested modeling phenotypic delay by 'diluting' each X_i in such a way that if $X_1 = n$, then the probability of j $(0 \leqslant j \leqslant n)$ out of the n mutants being phenotypically expressed is $\pi_{n,j}$. Consequently, the number of expressed mutants is distributed as $Y_1 + \cdots + Y_{M(t)}$ where each Y_j is independently and identically distributed with

$$\Pr[Y_1 = j] = \sum_{n=j}^{\infty} q_n \pi_{n,j}. \tag{106}$$

The p.g.f. of Y_1 is

$$\sum_{j=0}^{\infty} \left(\sum_{n=j}^{\infty} q_n \pi_{n,j} \right) z^j = \sum_{n=1}^{\infty} q_n \left(\sum_{j=0}^{n} \pi_{n,j} z^j \right) = \sum_{n=1}^{\infty} q_n H_n(z), \tag{107}$$

where $H_n(z) = \sum_{j=0}^{n} \pi_{n,j} z^j$ are the p.g.f.s of $\pi_{n,j}$ $(j = 0, 1, \ldots, n)$. Hence the p.g.f. of the number of expressed mutants is

$$G^*(z; t) = \exp\left\{ m(t) \left(\sum_{n=1}^{\infty} q_n H_n(z) - 1 \right) \right\}. \tag{108}$$

The above approach apparently does not apply to the Bartlett formulation under which the number of mutants does not follow a Poisson-stopped-sum distribution. Kendall [33] suggested adding an intermediate cell type between normal cells and mutant cells to incorporate phenotypic delay into the Bartlett formulation. Kendall's solution was essentially a three-compartment model; in Kendall's original terminology the three compartments were called (1) normal; (2) mutant-but-not-resistant and (3) resistant. As indicated in Section 6, moments (cumulants) from such models can be computed by solving ODEs. By adapting an approach for computing survival probability in carcinogenesis modeling [34], we can also compute the probability of zero mutations through solving ODEs. Lack of space prevents the inclusion of details.

8. Summary

To sum up, we recapitulate some conclusions drawn by this review.

First, among the four formulations we discussed, the Bartlett formulation is the most general, but also the least studied. Clamoring for solution is the issue

of finding an efficient algorithm for computing the probability function of the number of mutants. Because the other three formulations are all filtered Poisson processes, their properties are much better understood. The Lea–Coulson formulation is obviously preferable among the three formulations. As far as the method of means is concerned, three of the four formulations are equivalent, the exception being the discretized Luria–Delbrück formulation. Similarly, three of the four formulations are equivalent as far as the P_0 method is concerned, the exception being the Bartlett formulation.

Second, both the algorithm of Stewart et al. and that of Ma et al. compute the exact probability function. Both are applicable to the discretized Luria–Delbrück formulation and the Lea–Coulson formulation, but neither applies to the Bartlett formulation. The algorithm of Ma et al. is simpler and more efficient. On the other hand, the Lea–Coulson algorithm is not an exact method and it is applicable only to the Lea–Coulson formulation. In the case of differential growth, the approximation is in general unsatisfactory. The algorithm proposed by Ma et al. is therefore the preferred method.

Finally, all infinite moments result from approximation. During the past half a century research efforts focused on the approximate p.g.f. of Lea and Coulson. This distortion often caused controversy and confusion. The present review attempts to redress the balance. In fact, the exact distribution $\mathrm{LD}(\theta, \phi)$ is as easy to use as the approximate distribution $\mathrm{LD}(\theta, 1)$, because not only is the extra parameter ϕ known in practice, but the two distributions are equally amenable from a computational point of view.

Acknowledgements

This article could not have been written were it not for Dr Samuel Kotz of the George Washington University who kindly directed my attention to the Luria–Delbrück distribution. The enthusiastic reviews of an earlier draft by two anonymous referees have resulted in a vast number of improvements in the presentation.

Note added in proof

I owe sincerely thanks to Dr A.G. Pakes who recently brought my attention to three additional references [35–37].

References

[1] S.E. Luria, M. Delbrück, Mutations of bacteria from virus sensitivity to virus resistance, Genetics 28 (1943) 491.

[2] I.-C. Li, S.C.H. Wu, J. Fu, E.H.Y. Chu, A deterministic approach for the estimation of mutation rates in cultured mammalian cells, Mutation Res. 149 (1985) 127.
[3] W.S. Kendal, P. Frost, Pitfalls and practice of Luria–Delbrück fluctuation analysis: a review, Cancer Res. 48 (1988) 1060.
[4] D.E. Lea, C.A. Coulson, The distribution of the numbers of mutants in bacterial populations, J. Genetics 49 (1949) 264.
[5] P. Armitage, The statistical theory of bacterial populations subject to mutation, J. Royal Statist. Soc. B 14 (1952) 1 (with discussion on pp. 34–40).
[6] D.G. Kendall, On the choice of a mathematical model to represent normal bacterial growth, J. Royal Statist. Soci. B 14 (1952) 41.
[7] D.G. Kendall, Les processus stochastique de croissance en biologie, Ann. Inst. Henri Poincaré 13 (1952) 43.
[8] B. Mandelbrot, A population birth-and-mutation process I: Explicit distributions for the number of mutants in an old culture of bacteria, J. Appl. Prob. 11 (1974) 437.
[9] M.S. Bartlett, An Introduction to Stochastic Processes, Cambridge, London, 1955.
[10] M.S. Bartlett, An Introduction to Stochastic Processes, Cambridge, London, 1978.
[11] N.T.J. Bailey, The Elements of Stochastic Processes with Applications to the Natural Sciences, Wiley, New York, 1964.
[12] K.S. Crump, D.G. Hoel, Mathematical models for estimating mutation rates in cell populations, Biometrika 61 (1974) 237.
[13] A.L. Koch, Mutation and growth rates from Luria–Delbrück fluctuation tests, Mutation Res. 95 (1982) 129.
[14] J. Cairns, J. Overbaugh, S. Miller, The origin of mutants, Nature 335 (1988) 142.
[15] R.E. Lenski, M. Slatkin, F.J. Ayala, Mutation and selection in bacterial populations: alternatives to the hypothesis of directed mutation, Proc. Nat. Acad. Sci. USA 86 (1989) 2775.
[16] F.M. Stewart, D.M. Gordon, B.R. Levin, Fluctuation analysis: the probability distribution of the number of mutants under different conditions, Genetics 124 (1990) 175.
[17] W.T. Ma, G.v.H. Sandri, S. Sarkar, Analysis of the Luria–Delbrück distribution using discrete convolution powers, J. Appl. Prob. 29 (1992) 255.
[18] S. Sarkar, W.T. Ma, G.v.H. Sandri, On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants, Genetica 85 (1992) 173.
[19] A.W. Kemp, Comments on the Luria–Delbrück distribution, J. Appl. Prob. 31 (1994) 822.
[20] A.G. Pakes, Remarks on the Luria–Delbrück distribution, J. Appl. Prob. 30 (1993) 991.
[21] C.M. Goldie, Asymptotics of the Luria–Delbrück distribution, J. Appl. Prob. 32 (1995) 840.
[22] H. Prodinger, Asymptotics of the Luria–Delbrück distribution via singularity analysis, J. Appl. Prob. 33 (1996) 282.
[23] D.G. Kendall, Birth-and-death process and the theory of carcinogenesis, Biometrika 47 (1960) 13.
[24] D.L. Snyder, M.I. Miller, Random Point Processes in Time and Space, Springer, New York, 1991.
[25] K.L. Chung, A Course in Probability Theory, Academic Press, New York, 1974.
[26] Q. Zheng, On a compartmental analysis result, Math. Biosci. 130 (1995) 203.
[27] N.L. Johnson, S. Kotz, A.W. Kemp, Univariate Discrete Distributions, Wiley, New York, 1992.
[28] S. Wolfram, The Mathematica Book, fourth ed., Copublished by Wolfram Media, Champaign, IL and Cambridge University, Cambridge, New York, 1999.
[29] P. Armitage, Statistical concepts in the theory of bacterial mutation, J. Hygiene 51 (1953) 162.
[30] J.J. Hunter, Mathematical Techniques of Applied Probability, Vol. 1, Discrete Time Models: Basic Theory, Academic Press, New York, 1983.
[31] N.N. Lebedev, Special Functions and Their Applications, (revised English editon, translated and edited by R.A. Silverman), Dover, New York, 1972.

[32] P.S. Puri, Interconnected birth and death processes, J. Appl. Prob. 5 (1968) 334.
[33] D.G. Kendall, Stochastic processes and the growth of bacterial colonies, Symposia of the Society for Experimental Biology 7 (1953) 55.
[34] Q. Zheng, A unified approach to a class of stochastic carcinogenesis models, Risk Anal. 17 (1997) 617.
[35] W.Y. Tan, On distribution theories for the number of mutants in cell populations, SIAM J. Appl. Math. 42 (1982) 719.
[36] T.D. Tlsty, B.H. Margolin, K. Lum, Differences in the rates of gene amplification in nontumorigenic and tumorigenic cell lines as measured by Luria–Delbrück fluctuation analysis, Proc. Nat. Acad. Sci. USA 86 (1989) 9441.
[37] M. Kimmel, D.E. Axelrod, Fluctuation test for two-stage mutations: application to gene amplification, Mutation Res. 306 (1994) 45.