Luria–Delbrück, revisited: the classic experiment does not rule out Lamarckian evolution

Caroline M Holmes1,2, Mahan Ghafari1, Anzar Abbas3, Varun Saravanan3 and Ilya Nemenman1,2,4

1 Department of Physics, Emory University, Atlanta, GA 30322, United States of America
2 Department of Biology, Emory University, Atlanta, GA 30322, United States of America
3 Neuroscience Program, Emory University, Atlanta, GA 30322, United States of America
4 Initiative for Theory and Modeling of Living Systems, Emory University, Atlanta, GA 30322, United States of America

E-mail: ilya.nemenman@emory.edu

Keywords: Lamarck, Darwin, Luria–Delbrück fluctuation test, Bayesian model selection, maximum likelihood

Abstract

We re-examined data from the classic Luria–Delbrück fluctuation experiment, which is often credited with establishing a Darwinian basis for evolution. We argue that, for the Lamarckian model of evolution to be ruled out by the experiment, the experiment must favor pure Darwinian evolution over both the Lamarckian model and a model that allows both Darwinian and Lamarckian mechanisms (as would happen for bacteria with CRISPR-Cas immunity). Analysis of the combined model was not performed in the original 1943 paper. The Luria–Delbrück paper also did not consider the possibility of neither model fitting the experiment. Using Bayesian model selection, we find that the Luria–Delbrück experiment, indeed, favors the Darwinian evolution over purely Lamarckian. However, our analysis does not rule out the combined model, and hence cannot rule out Lamarckian contributions to the evolutionary dynamics.

1. Introduction

From the dawn of evolutionary biology, two general mechanisms, Darwinian and Lamarckian, have been routinely considered as alternative models of evolutionary processes. The Darwinian hypothesis posits that adaptive traits arise continuously over time through spontaneous mutation, and that evolution proceeds through natural selection on this already existing variation. In contrast, the Lamarckian hypothesis proposes that adaptive mutations arise in response to environmental pressures. The nobel prize winning fluctuation test by Salvador Luria and Max Delbrück [1] is credited with settling this debate, at least in the context of evolution of phage-resistant bacterial cells.

Luria and Delbrück realized that the two hypotheses would lead to different variances (even with the same means) of the number of bacteria with any single adaptive mutation. Specific to the case of bacteria exposed to a bacteriophage, this would result in different distributions of the number of surviving bacteria, see figure 1. In the Darwinian scenario, there is a possibility of a phage-resistance mutation arising in generations prior to that subjected to the phage. If this mutation happens many generations earlier, there will be a large number of resistant progeny who will survive (a ‘jackpot’ event). However, there will be no survivors if the mutation does not exist in the population at the moment the phage is introduced. If the same experiment were repeated many times, the variance of the number of survivors would be large. In contrast, in the Lamarckian scenario, the distribution of the number of survivors is Poisson. Indeed, each occurring mutation (and hence each survivor) happens with a small probability, independent of the others. This would result in the usual square-root scaling of the standard deviation of the number of survivors, a much smaller spread than in the Darwinian case.

To test this experimentally, Luria and Delbrück let the cells grow for a few generations, exposed them to a phage, plated the culture, and then counted the number of emergent colonies, each started by a single resistant, surviving bacterium. They found that the distribution of the number of survivors, as measured by the number of colonies grown after plating, was too heavy-tailed to be consistent with the Poisson distribution. They concluded then that the bacteria must evolve using the Darwinian mechanism. They could not derive an analytical form of the distribution of survivors in the Darwinian model, so that their data analysis was semi-quantitative at best. In particular, they could only establish that the Darwinian model fits
the data better than the Lamarckian/Poissonian one, but they could not quantify how good the fit is.

Potentially even more importantly, the original paper contrasted only two scenarios: pure Lamarckian and pure Darwinian ones. However, it is possible that both processes have a role in bacterial evolution, as is abundantly clear now in the epoch of epigenetics and, especially, CRISPR-Cas bacterial immunity, which is essentially Lamarckian [2–6]. In addition, stress can increase the (undirected) mutation rates, so that more mutations arise at the time of the challenge and then get selected, which will also appear Lamarckian: Luria–Delebrück experiments cannot distinguish between induced directed and undirected mutations. Thus ruling out a significant Lamarckian contribution to evolution through either of these or any other mechanism would require us to show not only that the Darwinian model explains the data better than the Lamarckian one, but also that the Darwininan model is more likely than the Combination model, which allows for both types of evolutionary processes. Evolution could also proceed through an entirely different mechanism, so that neither of the proposed models explain the data. Distinguishing between these possibilities requires evaluating whether a specific model fits the data well, rather than which of the models fits the data better.

Unlike Luria and Delbrück in 1943, we have powerful computers and new statistical methods at our disposal. Distributions that cannot be derived analytically can be estimated numerically, and model comparisons can be done for models with different numbers of parameters. In this paper, we perform the quantitative analysis missing in the Luria–Delbrück paper and use their original data to evaluate and compare the performance of three models: Darwinian (D), Lamarckian (L), and Combination (C) models. The comparison is somewhat complicated by the fact that both the L and D models are special cases of the C model, so that C is guaranteed to fit not worse than either L or D. We use Bayesian Model Selection [7, 8], which automatically penalizes for more complex models (such as C) to solve the problem. We conclude that, while the L model is certainly inconsistent with the data, D and C explain the data about equally well when this penalty for complexity is accounted for. Thus the Lamarckian contribution to evolution cannot be ruled out by the 1943 Luria and Delbrück data. Further, while D and C fit equally well, neither provide a quantitatively good fit to one of the two primary experimental data sets of Luria and Delbrück’s paper, suggesting that the classic experiment may have been influenced by factors or processes not considered.
Even though we will show that our analysis of the Luria–Delbrück data does not rule out Lamarckian contributions to evolution, our goal is not to challenge the well-established knowledge that this particular system (T1 phage interacting with Escherichia coli) is largely purely Darwinian [9–13]. Instead our goal is to limit ourselves to the original 1943 data, even though additional experiments have been performed many times since then, and to ask: does this data actually tell us what every textbook says, and namely that the 1943 experiment has ruled out Lamarckian evolution in favor of Darwinian?

2. Methods

2.1. Models and notational preliminaries

There have been many theoretical attempts, with varying degrees of success, to find closed-form analytical expressions of the distribution of mutants under the Darwinian—but not the combined—scenario (the Luria–Delbrück distribution). These have followed several different modeling assumptions [14–21], with good reviews in [22, 23]. For example, models with constant and synchronous division times [19, 20] or many different distributions of generation times, in some cases allowing for different growth rates for the wild-type and the mutant populations [24–26], have been proposed to find the distribution of survivors. In this work, we do not advance these analytical treatments. However, we present one such analysis, mainly to introduce notation and to illustrate complications of using analytical expressions for our statistical analysis.

We follow Haldane’s modeling hypotheses [19] and assume that (i) normal cells and mutants have the same fitness until the phage is introduced, (ii) all cells undergo synchronous divisions, (iii) no cell dies before the introduction of the phage, (iv) mutations occur only during divisions, with each daughter becoming a mutant independently (D case), or only when the phage is introduced (L case), and (v) backwards mutations are negligible. With these assumptions, the D and the C models are able to produce very good fits to the experimental data (see below), which suggests that relaxation of these assumptions and design of more biologically realistic models is unnecessary in the context of these experiments.

For the subsequent analyses, let \( N_0 \) be the initial number of wild-type, phage-sensitive bacteria, and \( g \) be the number of generations before the phage is introduced, so that the total number of bacteria after the final round of divisions is \( N = 2^g N_0 \), and the total that have ever lived is \( 2N - N_0 \). We use \( \theta_\text{D} \) to denote the probability of an adaptive (Darwinian) mutation during a division, and \( \theta_\text{L} \) to denote the probability of an adaptive Lamarckian mutation when the phage is introduced. With this, and discounting the probability of another mutation in the already resistant progeny, the mean number of adaptive Darwinian mutations at generation \( g \) is \( m_\text{D} = \theta_\text{D}(2N - N_0) \), and the mean number of adaptive Lamarckian mutations is \( m_\text{L} = \theta_\text{L} N_0 \). The number of survivors in the L model is Poisson-distributed with the parameter \( m_\text{L} \):

\[
P_L(k|\theta_\text{L}, N_0) = \frac{e^{-m_\text{L}} m_\text{L}^k}{k!} \quad (1)
\]

For the D model, there are multiple ways to have a certain number of resistant bacteria, \( k \), in the population of size \( N \) before introducing the phage. For instance, there are four ways to have 5 resistant bacteria (i.e. \( k = 5 \)): (i) One mutation occurs 2 generations before the phage introduction (where the total living population is \( N/4 \) at that generation), resulting in 4 resistant progenies in the last generation, and one more mutation at the last generation, making a total of 5 resistant cells before the phage introduction. This is the most likely scenario with probability \( P_\text{D}^{(1)} = (1 - \theta_\text{D})^{2(N - N_0) - 8} \theta_\text{D}^{(N/4)} \left(\frac{N - 4}{1}\right) \), where \( (2N - N_0) - 8 \) is the total number of cells that have ever lived in the entire experiment, so that \( (2N - N_0) - 8 \) is the total number that have ever lived without mutating. The \( \theta_\text{D} \) factor indicates that a total of two mutations have occurred in the population. The first choose factor denotes the number of independent mutational opportunities 2 generations before the phage introduction, and the second one denotes the number of mutational opportunities in the last generation. (ii) Two mutations occur 1 generation before the phage introduction and one more mutation in the last generation. This is less likely than (i) with probability \( P_\text{D}^{(ii)} = (1 - \theta_\text{D})^{2(N - N_0) - 6} \theta_\text{D}^{(N/2)} \left(\frac{N - 2}{3}\right) \). (iii) One mutation occurs 1 generation before the phage introduction and 3 more mutations in the last generation. This is less likely than (ii) with probability \( P_\text{D}^{(iii)} = (1 - \theta_\text{D})^{2(N - N_0) - 4} \theta_\text{D}^{(N)} \). In general, for an arbitrary number of resistant cells, \( k \), let \( \Pi_k \) denote the set of sequences \((a_0, a_1, \ldots)\) that satisfy \( k = \sum_{i=0}^{\infty} a_i 2^i \), such that \( a_i \in \mathbb{Z}_{\geq 0} \). This condition captures all the possible sequences \( \{a_i\} \in \Pi_k \) that produce \( k \) number of resistant cells. For instance, in case (i) the corresponding sequence is \( \{a_i\} = \{a_0, a_1\} = \{0, 1\} \), and in case (ii) it is \( \{a_0 = 2, a_1 = 1\} \). Then, following Haldane’s approach [19], we can write \( P_D(k) \), the probability of finding \( k \) resistant cells given the Darwinian model of evolution, as

\[
P_D(k|\theta_\text{D}, N_0) = \sum_{\{a_i\} \in \Pi_k} (1 - \theta_\text{D})^{2(N - N_0) - \sum_{i=0}^{\infty} a_i 2^{i+1} - 1} \theta_\text{D}^{\sum_{i=0}^{\infty} a_i 2^{i}} \times \prod_{i=1}^{\infty} \left( \frac{N}{2} - \sum_{r=i+1}^{\infty} a_r 2^{r-1} \right), \quad (2)
\]
where \( x = \sum_{i=1}^{n} a_i \) and the probability \( P_D(k|\theta_D, N_0) \) is summed over all the possible sequences \( \{a_i\} \in \Pi_k \) that produce the number \( k \); in the case of \( P_{DL} = 5 \) mentioned earlier, \( P_{DL}(S|\theta_K, N_0) = P_i^k + P_s^i + P_5^m + P_5^{im} + P_5^{im} \).

For the Combination model, both the L and the D processes contribute to generating survivors. Thus we write the distribution of the number of survivors in this case as a convolution

\[
P_C(k|\theta_L, \theta_D, N_0) = \sum_{k=0}^{\infty} P_D(k|\theta_D, N_0) P_L(k - k'|\theta_L, N_0).
\]

(3)

Further analytical progress on the problem is hindered by additional complications. First, in actual experiments, the initial number of bacteria in the culture \( N_0 \) is random and unknown. We view it as Poisson-distributed around the mean that one expects to have, denoted as \( \Pi(N_0|\hat{N}_0) \). This gives:

\[
P_{DL/LC}(k|\theta_D, \theta_L, \hat{N}_0) = \sum_{N_0=0}^{\infty} P_{DL/LC}(k|N_0) \Pi(N_0|\hat{N}_0).
\]

(4)

Finally, in some of the Luria–Delbrück experiments, they plated only a fraction \( r \) the entire culture subjected to the phage. This introduced additional randomness in counting the number of survivors after the plating, \( k_p \) which we again model as a Poisson distribution with the mean \( r k \), \( \Pi(k_p|rk) \) [27–29], resulting in the overall distribution of survivors:

\[
P_{DL/LC}(k_p|\theta_D, \theta_L, \hat{N}_0) = \sum_{k=0}^{\infty} \Pi(k_p|rk) P_{DL/LC}(k).
\]

(5)

Equation (5) illustrates the main complication of using analytical results for statistical inference studies: it is not computationally efficient, involving multiple nested (and infinite) sums. Alternative approaches (e.g. [14]) represent the distribution as recursive expressions, through the inverse Fourier transform of the characteristic functional, or through low-order moments. These expressions are also not easy to calculate, or are hard to compare to the experimental probability distribution in the inference step due to additional complications, such as dilution or the Lamarckian contribution in the C model. One can try to develop an efficient algorithm for evaluating the probability of the number of survivors for the C model, similar to the ones that have already been developed for the D case [30, 31], but this is not a trivial task. Instead, since our focus is on the inference and not on the analytics, we chose a simpler approach: sampling the distribution using Monte-Carlo techniques.

2.2. Computational models

Our simulations assume that each culture begins with a Poisson-distributed number of bacteria, with a mean number of 135, as in the original paper. The bacteria were modeled as dividing in discrete generations for a total of \( g = 21 \) generations. Both of the numbers are easily inferable from the original paper using the known growth rate and the final cell density numbers. Cells divide synchronously, and each of the daughters can gain a resistance mutation at division with the probability \( \theta_D \), which is nonzero in C and D models. Daughters of resistant bacteria are themselves resistant. Non-resistant cells in the final generation are subjected to a bacteriophage, which induces Lamarckian mutations with probability \( \theta_L \), nonzero in C and L models. We note again that this total number of Lamarckian-mutated cells is Poisson-distributed with the mean \( \theta_L \) times the number of the remaining wild type bacteria.

To speed up simulations of the Darwinian process, we note that the total number of cells that have ever lived is \( N_F = 2N_02^g - N_0 \). Thus the total number of Darwinian mutation attempts is Poisson distributed, with mean \( N_0\theta_D \). We generate the number of these mutations with a single Poisson draw and then distribute them randomly over the multi-generational tree of cells, marking every offspring of a mutated cell as mutated. We then correct for overestimating the probability of mutations due to the fact that the number of mutation attempts in each generation decreases if there are already mutated cells there. For this, we remove original mutations (and unmark their progenies) at random with the probability equal to the ratio of mutated cells in the generation when the mutation appeared to the total number of cells in this generation. Note that since mutations are rare, such unmarking is not very common in practice, making this approach substantially faster than simulating mutations one generation at a time.

To estimate \( P_C(k_p|\theta_D, \theta_L) \), we estimate this probability on a \( 41 \times 41 \) grid of values of \( \theta_D \) and \( \theta_L \). For each pair of values of these parameters, we perform \( n = 3000000 \) simulation runs (see below for the explanation of this choice) starting with a Poisson-distributed number of initial bacteria, then perform simulations as described above, and finally perform a simulated Poisson plating of a fraction of the culture if the actual experiment we analyze had such plating. We measure the number of surviving bacterial cultures \( k_p \) in each simulation run and estimate \( P_C(k_p|\theta_D, \theta_L) \) as a normalized frequency of occurrence of this \( k_p \) across runs, \( f_C(k_p|\theta_D, \theta_L) \). The Darwinian case is evaluated as \( P_C(k_p|\theta_D = 0, \theta_L) \), and the Lamarckian case as \( P_C(k_p|\theta_D = 0, \theta_L) \).

2.3. Quality of fit

In the original Luria and Delbrück publication [1], no definitive quantitative tests were done to determine the quality of fit of either of the model to the data. We can use the estimated values of \( P_C(k_p|\theta_D, \theta_L) \) for this task. Namely, Luria and Delbrück have provided us not with frequencies of individual values of \( k_p \), but with frequencies of occurrence of \( k_p \) within bins of \( x \in (0, 1, 2, 3, 4, 5, 6, 10, 11, 20, 21–50, 51–100, 101–200, 201–500, 501–1000) \). By summing \( f_C(k_p|\theta_D, \theta_L) \) over \( k_p \in x \), we evaluate \( f_C(x|\theta_D, \theta_L) \), which allows us to write the probability that each experimental set of measurements, \( \{N_x\} \), came from the model:
\[ P(n_x | \theta_D, \theta_L) = C \prod_{x} \frac{n_x}{\sum_{x} n_x} e^{(x | \theta_D, \theta_L)} , \quad (6) \]

where \( C \) is the usual multinomial normalization coefficient. This probability can also be viewed as the likelihood of each parameter combination, and the peak of the probability gives the usual maximum likelihood estimation of the parameters [32]. To guarantee that the estimated value of the likelihood has small statistical errors, we ensured that each of the \((\theta_D, \theta_L)\) combinations has 30 000 000 simulated cultures. Then, at parameter combinations close to the maximum likelihood, each bin \( x \) has at least 10 000 samples. Correspondingly, at these parameter values, the sampling error in each bin is smaller than 1%.

Finally, to evaluate the quality of fit of a model, rather than to find the maximum likelihood parameter values, we calculate empirically the values of \( \log_{10} P(n_x \| \theta_D, \theta_L) \) for each parameter combination, where \( n_x \) are synthetic data generated from the model with \( \theta_D, \theta_L \). Mean and variance of \( \log_{10} P \) gives us the expected range of the likelihood if the model in question fits the data perfectly.

### 2.4. Comparing models

In comparing the L, the D, and the C models, we run into the problem that C is guaranteed to have at least as good of a fit as either D or L since it includes both of them as special cases. Thus in order to compare the models quantitatively, we need to penalize C for the larger number of parameters (two mutation rates) compared to the two simpler models. To perform this comparison, we use Bayesian model selection [7, 8], which automatically penalizes for such model complexity.

Specifically, Bayesian model selection involves calculation of probability of an entire model family \( M = \{ L, D, C \} \) rather than of its maximum likelihood parameter values:

\[ P(M | \{ n_x \}) \propto \int d\theta_D P(\theta_D | \{ n_x \}) P(M) \]  

(7)

where the posterior distribution of \( \theta \) is given by the Bayes formula,

\[ P(\theta_D | \{ n_x \}, M) \propto P(\{ n_x \} | \theta_D, M) P(\theta_D | M) \]  

(8)

and \( P(\{ n_x \} | \theta_D, M) \) comes from equation (6). Finally, \( P(M) \) and \( P(\theta_D | M) \) are the \( a \ priori \) probabilities of the model and the parameter values within the model, which we specify below.

The integral in equation (7) is over as many dimensions as there are parameters in a given model. Thus while more complex models may fit the data better at the maximum likelihood parameter values, a smaller fraction of the volume of the parameter space would provide a good fit to the data, resulting in an overall penalty on the posterior probability of the model. Thus posterior probabilities \( P(M | \{ n_x \}) \) can be compared on equal footing for models with different number of parameters to say which specific model is \( a \ posteriori \) more likely given the observed data. Often the integral in equation (7) is hard to compute, requiring analytical or numerical approximations. However, here we already have evaluated the likelihood of combinations of \((\theta_D, \theta_L)\) over a large grid, so that the integral can be computed by direct summation of the integrand at different grid points.

To finalize computation of posterior likelihoods, we must now define the \( a \ priori \) distributions \( P(M) \) and \( P(\theta_D | M) \). We choose \( P(C) = P(M) = P(L) = 1/3 \), indicating our ignorance about the actual process underlying biological evolution. The choice of \( P(\theta_D | M) \) is tricky, as is often the case in applications of Bayesian statistics. We point out that the experiment was designed so that the number of surviving mutants is almost always 1 or less, for a population with \( \approx 0.25 \times 10^8 \) individuals, which indicated that \( a \ priori \) both \( \theta_D \) and \( \theta_L \) are less than \( 4 \times 10^{-9} \). Further, we assume that, for the combined model, \( P(\theta_D) = P(\theta_L)P(\theta_D) \). Beyond this, we do not choose one specific form of \( P(\theta_D) \), but explore multiple possibilities to ensure that our conclusions are largely independent of the choice of the prior.

### 2.5. Statistical power of the tests

Before analyzing the original Luria–Delbrück data, it is important to understand the statistical power of our approach in discriminating among the models. Anticipating that the L model will be easy to rule out (see Results), we focus on disambiguating just the D and the C models by investigating the relationship between \( N \), the number of cultures, and the expected ratio \( P(D)/P(C) \) for synthetic data that resembles that of the Luria–Delbrück experiment 22 (see Results), which is fitted well either with the D model with \( \theta_D = 2.0 \times 10^{-9} \), or with the C model with \( \theta_D = 1.8 \times 10^{-9} \), and \( \theta_L = 0.4 \times 10^{-9} \).

First, we note that ‘ruling out’ the Lamarckian contribution would require having \( P(D)/P(C) > 1 \). Thus we created synthetic data with a fixed Darwinian mutation rate \( \theta_D = 2.0 \times 10^{-9} \) and zero Lamarckian contribution and explored how strong the evidence in favor of purely Darwinian model would be at different \( N \). For this, we varied \( N \) and investigated the ratio \( P(D)/P(C) \), using a prior \( P(\theta_D) \) that is uniform between 0 and \( 4 \times 10^{-9} \), as we later use for the analysis of the actual experiments. Figure 2(a) illustrates the dependence. Notice that \( P(D)/P(C) \sim 20 \), which would correspond to rejection of the Combined model at about 95% confidence, only for \( N > 10^3 \). This suggests that the Luria–Delbrück experiment was not designed well to achieve this demarcation.

Next we analyzed the number of cultures that would be needed to demonstrate existence of the Lamarckian influence as a function as a function of \( \theta_L \). We defined the demonstration as \( P(D)/P(C) < 1/20 \), and then explored the number of cultures required to
reach this threshold for different Lamarckian rates, while keeping $\theta_D = 1.8 \times 10^{-9}$, see figure 2(b). As expected, when the Lamarckian contribution is higher, the number of cultures decreases. However, crucially, even at $\theta_L = \theta_D$, one would require $N > 200$ cultures to demonstrate the Lamarckian contribution conclusively. This is because at $\theta_L = \theta_D$, the expected number of Lamarckian mutations in a culture (half that of the Darwinian ones) is $< 1$. This again suggests that the Luria–Delbrück experiment should have had larger numbers of cells to be designed optimally for discriminating among these different models.

3. Results

Luria and Delbrück’s paper provided data from multiple experiments, where in each experiment they grew a number of cultures, subjected them to the phage, and counted survivors. Most of the experiments have $O(10)$ cultures, which means that their statistical power for distinguishing different models is very low. We exclude these experiments from our analysis and focus only on experiments No. 22 and 23, which have $N = 100$ and $N = 87$ cultures, respectively. Our previous analysis suggests that even this is likely to be too few cultures for conclusive results, but these are the numbers we have to work with. The experimental protocols differ in that experiment 23 plated the entire culture subjected to the phage, while experiment 22 plated only 1/4 of the culture. We analyze these experiments separately from each other.

3.1. Experiment 22

We evaluated the posterior probability of different parameter combinations, $P(\theta_M | \{n_k\}, M)$, numerically, as described in Methods. The likelihood (posterior probability without the prior term) is illustrated in figure 3. Note that the peak of the likelihood is at $\theta_L^{22} \approx 4.0 \times 10^{-10}$ and $\theta_D^{22} \approx 1.8 \times 10^{-9}$, illustrating that the data suggests that the Combination model is better than either of the pure models in explaining the data, though the pure Darwinian model comes close. The fit of the maximum likelihood Combination model is shown in figure 4. The quality of the best fit is $\theta_L^{22} | \theta_D^{22}, P(n_{k}) \log, 63.7$. This matches surprisingly well with the likelihood expected if the data was indeed generated by the model, $\theta_L^{22} | \theta_D^{22}, P(n_{k}) \log, 62.1 \pm 5.1$. Thus the model fits the data perfectly despite numerous simplifying assumptions, suggesting no need to explore more complex physiological scenarios, such as asynchronous divisions, or different growth rates for mutated and non-mutated bacteria.

Next we evaluate the posterior probabilities of all three models by performing the Bayesian integral, equation (7). We use two different priors for $\theta_L$ and $\theta_D$ to verify if our conclusions are prior-independent: uniform between $0$ and $\times − 4 10 9$ and uniform in the logarithmic space between $\times − 1 10 10$ and $\times − 4 10 9$. For the uniform prior, $\approx\approx \approx\approx 2.8 1 , 1 0 6$. In other words, the purely Lamarckian model is ruled out by an enormous margin, as suggested in the original publication. However, in agreement with our estimate of the statistical power of the analysis, the ratio of posterior probabilities of the Darwinian and the Combination models is only 2.8, and this ratio is 2.0 for the logarithmic prior, which is way over
the usual 5% significance threshold for ruling out a hypothesis. In other words,

The darwinian and the Combination models of evolution have nearly the same posterior probabilities after controlling for different number of parameters in the models. Thus contribution of Lamarckian mechanisms to evolution in the Luria–Delbrück Experiment 22 cannot be ruled out.

3.2. Experiment 23
We performed similar analysis for experiment 23 and evaluated the posterior likelihood, figure 5, for each combination of parameters. See figure 6 for the best fit model. Here, however, the posterior is several orders of magnitude smaller than for experiment 22. This is because the experimental data has a tail that is heavier than typical realizations of even the Darwinian model would predict. Indeed, Luria and Delbrück themselves noted this excessively heavy tail. However, as they were only choosing whether the Darwinian or the Lamarckian model fits better, this led further credence to the claim that the Lamarckian model could not describe the data. Now we are able to quantify this: for experiment 23, at the maximum likelihood parameters \((\theta^2_L = 4.0 \times 10^{-10} \text{ and } \theta^2_D = 1.8 \times 10^{-9})\), the quality of fit is \(\log_{10} P(\mathbf{n}_i | \theta^2_D, \theta^2_D) \approx -90.0\). In contrast, for data generated from the model, we get \(\log_{10} P(\text{data}|\theta_D, \theta_L) = -76.9 \pm 3.1\). In fact, by generating \(10^5\) data sets using these parameter combination, we estimate that the probability of generating data
from this model that is as unlikely as the experimental data is \( p < 10^{-4} \). Thus the tail of the distribution of the number of mutants in experiment 23 is so heavy that it cannot be fit well by either of the hypotheses considered. Instead, it is likely that some other dynamics are at play here, such as some form of contamination, or additional non-Darwinian processes. In other words,

Luria–Delbrück Experiment 23 cannot be explained by any of the proposed hypotheses (the Lamarckian, the Darwinian, or the Combination one), and thus cannot be used to rule out one hypothesis over another.

4. Discussion

The classic Luria and Delbrück 1943 experiment [1] is credited with ruling out the Lamarckian model in favor of Darwinism for explaining acquisition of phage resistance in bacteria. However, while heralded as a textbook example of quantitative approaches to biology, the data in the paper was analyzed semi-quantitatively at best. We performed a quantitative analysis of the fits of three models of evolution (Lamarckian, Darwinian, and Combination) to these classic data, experiments 22 and 23. Our analysis was based on a very simplified model of the process: we started each colony with a Poisson-distributed (mean

\[ \begin{align*}
\tau_L & = 0 \\
\tau_D & = 0
\end{align*} \]
wild-type bacteria and allowed them to replicate synchronously for exactly 21 times, with mutations occurring continuously (Darwinian model) or at the last generation (Lamarckian model). Additionally, we did not consider the possibility that multiple mutations might be needed to acquire resistance, or that growth rates of bacteria may be inhomogeneous. Nonetheless, the simple model fits experiment 22 data perfectly, suggesting no need for more complex modeling scenarios.

For experiment 22, by a ratio of $\approx 10^{-10}$, the Lamarckian model is a posteriori less likely than the Darwinian one, agreeing with the original Luria and Delbrück conclusion that the pure Darwinian evolution is a better explanation of the data than the pure Lamarckian evolution. However, the posterior odds of the pure Darwinian model are only $2\sim3$ times higher than those for the Combination model (suitably penalized for model complexity), which has nonzero Darwinian and Lamarckian mutation rates. Even by liberal standards of modern day hypothesis testing, there is insufficient evidence to rule out the Combination model, and, therefore, contribution of Lamarckian processes to bacterial evolution in this experiment. This was in agreement with our analysis of the statistical power of the data: the number of cultures and the mean number of mutations per culture were too small to effectively discriminate between the D and the C model with a small Lamarckian rate. In contrast, for experiment 23, neither of the three considered models could quantitatively explain the data, suggesting that additional processes must be in play beyond simple Lamarckian and Darwinian mutations.

In summary, while subsequent experiments have certainly established the Darwinian nature of mutations in the T1- E. coli system, our analysis shows that the classic Luria–Delbrück 1943 data cannot be used to rule out Lamarckian contributions to bacterial evolution in favor of Darwinism, potentially necessitating changes to the exposition in many biology textbooks.

Acknowledgments

This work was partially supported by NSF Grant No. PoLS-1410978, James S McDonnell Foundation Grant No. 220020321, NIH NINDS R01 NS084844, Woodruff Scholars Program and Laney Graduate School at Emory University.

References

[1] Luria S E and Delbrück M 1943 Mutations of bacteria from virus sensitivity to virus resistance Genetics 28 491
[2] Koonin E V and Wolf Y I 2009 Biol. Direct 4 12
[3] Jablonka E and Lamb M J 2002 Ann. New York Acad. Sci. 981 82–96
[4] Jablonka E and Raz G 2009 Q. Rev. Biol. 84 131–76
[5] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero D A and Horvath P 2007 Science 315 1709–12
[6] Koonin E V and Wolf Y I 2016 Biol. Direct 11 9
[7] MacKay D J 1992 Neural Comput. 4 415–47
[8] MacKay D J 2003 Information Theory, Inference, and Learning Algorithms (Cambridge: Cambridge University Press)
[9] Lederberg J and Lederberg E M 1952 Replica plating and indirect selection of bacterial mutants J. Bacteriol. 63 399
[10] Newcombe H B et al 1949 Nature 164 150–0
[11] Taylor A L 1963 Proc. Natl Acad. Sci. 50 1043–51
[12] Hayes W et al 1964 The Genetics of Bacteria and Their Viruses (Studies in Basic Genetics and Molecular Biology) (New York: Wiley)
[13] Hanke K and Braun V 1978 Functional interaction of the tonA/tonB receptor system in Escherichia coli J. Bacteriol. 135 190–7
[14] Lea D and Coulson C A 1949 J. Genet. 49 264–85
[15] Armitage P 1952 The statistical theory of bacterial populations subject to mutation J. R. Stat. Soc. B 14 1–40
[16] Mandelbrot B 1974 J. Appl. Probab. 11 437–44
[17] Bartlett M S 1978 An Introduction to Stochastic Processes: with Special Reference to Methods and Applications (Cambridge: Cambridge University Press)
[18] Kendall D 1932 Stochastic processes of growth in biology Ann. Inst. Henri Poincare Inst. 13 43–108
[19] Sarkar S 1991 Haldane’s solution of the Luria–Delbrück distribution Genetics 127 257
[20] Zheng Q 2007 Math. Biosci. 209 500–13
[21] Angerer W P 2001 J. Math. Biol. 42 145–74
[22] Zheng Q 1999 Math. Biosci. 162 1–32
[23] Ycart B 2013 PLos One 8 e80958
[24] Koch A L 1982 Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis vol 95 (New Y ork: Elsevier) pp 129–41
[25] Jones M, Thomas S and Rogers A 1994 Luria–Delbrück fluctuation experiments: design and analysis Genetics 136 1209–16
[26] Jaeger G and Sarkar S 1995 Genetics 96 217–23
[27] Stewart F M, Gordon D M and Levin B R 1990 Fluctuation analysis: the probability distribution of the number of mutants under different conditions Genetics 124 175–85
[28] Stewart F 1991 Genetica 84 51–5
[29] Montgomery–Smith S, Le A, Smith G, Billstein S, Oveys H, Pischcko D and Yates A 2016 (arXiv:1608.04175)
[30] Ma W T, Sandri G Vh and Sarkar S 1992 Analysis of the Luria–Delbrück distribution using discrete convolution powers J. Appl. Probab. 29 255–67
[31] Sarkar S, Ma W and Sandri G V H 1992 Genetica 85 173–9
[32] Nelson P 2015 Physical Models of Living Systems (San Francisco, CA: Freeman)