

Research Article

The ABCs of Experimental Evolution

Zaid Abdo,^{1,2,3,4} Mathew Stein,^{5,6} Andrzej Wojtowicz,^{2,3,7} and Kim M. Pepin^{8,9}

¹ Departments of Statistics, University of Idaho, Moscow, ID 83844, USA

² Department of Mathematics, University of Idaho, Moscow, ID 83844, USA

³ Program of Bioinformatics and Computational Biology (BCB), University of Idaho, Moscow, ID 83844, USA

⁴ USDA-ARS South Atlantic Area, 950 College Station Road, Athens, GA 30605-2720, USA

⁵ Department of Computer Engineering, University of Idaho, Moscow, ID 83844, USA

⁶ Micron Technology Inc., Boise, ID 83716, USA

⁷ Department of Animal Sciences, Washington State University, Pullman, WA 99164, USA

⁸ Fogarty International Center, National Institutes of Health, 31 Center Drive, MSC 2220, Bethesda, MD 20892-2220, USA

⁹ Department of Biology, Colorado State University, Campus Delivery 1878, Fort Collins, CO 80523, USA

Correspondence should be addressed to Zaid Abdo; zaid.abdo@gmail.com

Received 15 January 2013; Accepted 14 February 2013

Academic Editors: F. Castiglione and A. Qiao

Copyright © 2013 Zaid Abdo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Microbial evolution is complex and is influenced by many sources of variation. Experimental evolution is no exception, although it is more controlled, easily replicated, and typically devoid of interactions between species. Mathematical modeling of the evolutionary process can help in understanding the underlying mechanisms that drive outcome of such experiments. These models can be complex and parameter rich, limiting their feasibility for statistical inference. In this paper, we introduce the use of Approximate Bayesian Computation (ABC) as a tool for statistical inference in the study of experimental evolution. ABC is a fast and simple method for fitting complex models to data. We utilize this method, coupled with a mechanistic model of experimental evolution, to study the evolution process of bacteriophage ϕ X174 under benign selection pressure. Our results highlight three mutation-selection scenarios that could explain this process: high mutation/low selection pressure, low mutation/high selection pressure, and low mutation/low selection pressure, with posterior support of 19%, 9.5%, and 71.5% for each of these scenarios, respectively. Sequence data support the first candidate. Though surprising, this scenario was not improbable based on our analysis.

1. Introduction

Mathematical modeling and statistical inference can aid in understanding the underlying evolutionary mechanisms that drive the outcome of evolution experiments (see, e.g., [1–3]). However, modeling and inference attempts are challenged by the complexity of the evolutionary process. This complexity stems from factors such as the stochastic nature of the experimental evolution process, the possibility of clonal interference (interactions between evolving genotypes), and experimental error. To address some of these challenges, we introduce a model of evolution that mimics the experimental evolution process in serial batch cultures. This model accounts for mutation, selection, drift and clonal interference. It also accounts for variability in the observed data resulting from experimental and sampling errors.

Fitting the aforementioned model to estimate evolutionary parameters of interest is complicated by the need to integrate over all possible experimental evolution trajectories. This is a computationally expensive process similar to integrating over all possible genealogies to estimate evolutionary parameters in coalescent-based population genetics [4, 5]. We borrow methodology from the coalescent literature and propose an Approximate Bayesian Computation (ABC) framework to fit this model [5]. Using this approach, we provide an approximation for the posterior distributions of the beneficial mutation probability and the mean selection for a phage population evolving under a novel, low selection pressure.

In what follows, we first introduce the mathematical model. We then introduce the ABC algorithm and methodology used to approximate the posterior distributions of

the parameters of interest. We then follow with the phage example and cap with a discussion and some conclusions.

2. Materials and Methods

2.1. Mathematical Model. One can think of the experimental evolution process of asexual populations in batch culture as proceeding in cycles. Each cycle consists of a number of generations and involves population growth, mutation, and drift, ending with a population bottleneck [1, 3]. Data are collected by sampling from the population at the end of each cycle. In this section, we first introduce a mathematical model that describes within-cycle evolution. We follow that with a mathematical description of bottlenecking and sampling.

2.1.1. Within-Cycle Evolution. In modeling within-cycle evolution, we start by assuming that each cycle, k , is formed of l generations. In proceeding from generation $t - 1$ to generation t within a cycle, we assume that the population grows. This growth occurs such that the number of progeny of the i th subpopulation within the k th cycle at generation t , N'_{ikt} , follows a Poisson distribution with mean $2^{W_0(1+S_{ik(t-1)})}N_{ik(t-1)}$, where W_0 is the absolute fitness of the wild type (or the minimum fitness of all subpopulations existing at the start of an experiment), $S_{ik(t-1)} = (W_{ik(t-1)} - W_0)/W_0$ is the selection coefficient ($W_{ik(t-1)}$ is the absolute fitness of subpopulation i at generation $t - 1$ during cycle k), and $N_{ik(t-1)}$ is the size of the i th subpopulation at generation $t - 1$. The resulting total population size at generation t is

$$N_{kt} = \sum_{i=1}^{M_{kt}} N'_{ikt}, \quad (1)$$

with M_{kt} being the total number of subpopulations during generation t within the corresponding cycle k .

We assume that mutation occurs with a fixed probability μ during each generation within a cycle. Given μ and N'_{ikt} , we model the number of new mutants, within the i th subpopulation, at generation t and within cycle k , X_{ikt} , using a binomial distribution as follows:

$$P(X_{ikt} | \mu, N'_{ikt}) = \binom{N'_{ikt}}{X_{ikt}} \mu^{X_{ikt}} (1 - \mu)^{N'_{ikt} - X_{ikt}}. \quad (2)$$

Original population sizes are adjusted according to the number chosen to mutate as follows:

$$N_{ikt} = N'_{ikt} - X_{ikt}. \quad (3)$$

A new mutant is then assigned a certain selection coefficient, S_{jkt} , chosen from some distribution with probability $P(S_{jkt} | S_{ik(t-1)}, \Theta)$, where Θ is a matrix of parameters. These new mutants are then assigned to that new subpopulation j and are referred to as $X_{(i \rightarrow j)kt}$. Mutants with the same parent that have the same selection coefficient form a new distinct mutant subpopulation.

This process is assumed to repeat itself until a bottleneck is encountered.

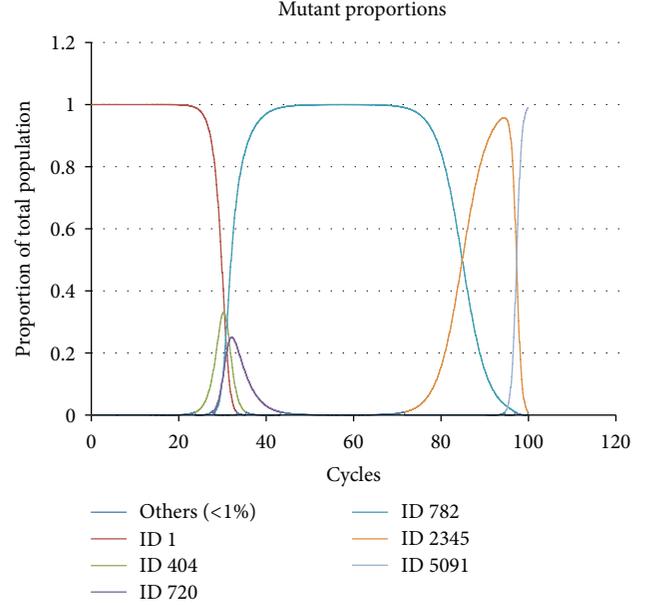


FIGURE 1: This figure is based on one simulation run using mutation probability per genome equal to 10^{-7} , transfer time (number of generations allowed within one cycle) equal to 10, and total number of cycles equal to 100. The simulation was run with initial population size and bottleneck of size 1,000,000. Selection coefficients, assigned to each new mutant, were sampled from an exponential distribution with mean equal to 0.1. The figure shows that simulated phage genotypes with identity numbers ID 404 and ID 720 compete with one another and with ID 782. Simulated genotype ID 782 fixes after the 50th cycle and drives both the ID 404 and ID 720 to extinction. Such clonal interference is also evident between simulated genotypes ID 782, ID 2345, and ID 5091.

2.1.2. The Bottleneck. A “bottleneck” is defined to be sampling to initiate a new cycle, and it occurs at the end of each evolution cycle after l generations. The dilution factor, D_b , used in a bottleneck is usually much larger than that for sampling (defined below). The resulting total population size after dilution is $N_{(k+1)0} = D_b N_{kl}$. We use multinomial sampling to model the bottleneck as described in (4):

$$\begin{aligned} P(\{N_{i(k+1)0}\} | \{q_{ik}\}, N_{(k+1)0}, M_{kl}) \\ = \frac{N_{(k+1)0}!}{\prod_{i=1}^{M_{kl}} N_{i(k+1)0}!} \prod_{i=1}^{M_{kl}} q_{ikl}^{N_{i(k+1)0}}, \end{aligned} \quad (4)$$

where $q_{ikl} = N_{ikl}/N_{kl}$. M_{kl} is the total number of mutant categories at the end of cycle k . The model assumes that the population is allowed to mutate just before a bottleneck occurs.

Clonal interference is accounted for in this part of the model by (1) allowing for possible existence of multiple genotypes at any point in time depending on how high the mutation probability is, (2) allowing for differential growth of the different genotypes depending on their selection advantage, and (3) enforcing a restricted bottleneck where new populations are initiated with subpopulation sizes proportional to their fitness. An example of the effect of clonal interference as captured by the model is presented in Figure 1.

Parameters used to generate this figure are described in its legend. The figure clearly shows that simulated phage genotypes with identity numbers ID 404 and ID 720 compete with one another and with ID 782. Simulated genotype ID 782 dominates (fixes) after the 50th cycle and drives both the ID 404 and ID 720 to extinction. Such interaction is also evident between simulated genotypes ID 782, ID 2345, and ID 5091. Clonal interference increases the time to fixation of a beneficial mutation, which can allow for other more fit mutations to arise and take over (such as the mutation in association with genotype ID 782 in this example). This allows for a better adapted population to some selective pressure. Occurrence of clonal interference depends on the mutation rate and on the selective pressure imposed [2, 6]. This effect has been observed in many cases of experimental evolution [6, 7].

2.1.3. Sampling. We define “sampling” to be the process of obtaining observable data. Here, we also assume that sampling involves a multinomial sampling process similar to that introduced in (4) and replacing the bottleneck dilution factor with a sampling dilution factor, D_s . This results in a total sample size $Y_k = D_s N_{kl}$ at the end of cycle k . Equation (5) introduces this change:

$$P(\{Y_{ik}\} | \{q_{ikl}\}, Y_k, M_{kl}) = \frac{Y_k!}{\prod_{i=1}^{M_{kl}} Y_{ik}!} \prod_{i=1}^{M_{kl}} q_{ikl}^{Y_{ik}}, \quad (5)$$

where Y_{ik} represents the size of the obtained sample of mutant i at the end of cycle k and q_{ikl} is as defined above.

Observed data from experimental evolution studies can be assumed to be function of mutant sample sizes, selection coefficients, and the number of their types. For example, in the phage evolution data [6] reintroduced below, we observe average population fitness, \bar{W}_k , at the end of cycle k , which is presented in (6):

$$\bar{W}_k = \frac{\ln(N_{kl}) - \ln(N_{k0})}{l \cdot \ln(2)}, \quad (6)$$

where N_{kl} can be thought to equal Y_k/D_s and an error term ϵ can be added to reflect sampling error. Accordingly, (6) becomes

$$\bar{W}_k = \frac{\ln(Y_k) - \ln(D_s) - \ln(N_{k0})}{l \cdot \ln(2)} + \epsilon. \quad (7)$$

ϵ is an error resulting due to inaccuracies in observing the data (a measurement error), which we assume to be a truncated normal distributed with mean equal to zero and variance equal to σ^2 . The point of truncation is such that the mean fitness \bar{W}_k is always positive.

2.2. Approximate Bayesian Computation (ABC). Our inability to observe data during the growth period within a cycle forms a major limitation for parameter estimation using the above model. This is due to the fact that many evolution trajectories within a cycle can result in the same measured fitness

at the end of that cycle. This stochastic effect can be removed by integrating over all possible trajectories within all observed cycles of an experiment, which is analytically impossible and computationally expensive if done numerically. We propose an Approximate Bayesian Computation (ABC) framework [5] to overcome this limitation.

ABC is a fast and simple method for fitting complex models to data [4]. This approach requires (1) designing a simulation based on the proposed model, (2) performing large number of simulations based on a set of proposed parameter values, and (3) comparing each of the resulting data sets to the observed data. Comparisons are made using a distance measure. If the simulated data is equivalent to, or close enough to, the real, data then the parameter values that were used to simulate these data are kept. This results in a set of candidate values that are used to construct the approximate posterior distributions for the parameters of interest. It avoids the direct evaluation of a likelihood function, which is computationally expensive. We first describe the implementation of the simulation and then provide details of the algorithm we propose for inference using ABC.

2.2.1. Simulation. We used C++ to develop a simulation based on the proposed model. Both a command line and a GUI version are available from the authors upon request. The GUI format was implemented in C++ using Visual Studio 8 and was linked to *Excel* to provide a graphical output of the competing subpopulations, which was used to generate Figure 1. A user’s manual is also available upon request from the authors.

In implementing this simulation, we assumed that the selection coefficients, S_{ikt} , follow an exponential distribution with mean λ [2, 8, 9]. Implementation of this simulation was simplified by discretizing this distribution to improve computational efficiency.

2.2.2. ABC. We utilized a rejection algorithm [5] in implementing our ABC strategy to sample candidate parameter values as follows.

- (1) Propose B sets of values for the parameters of interest using some prespecified prior distributions.
- (2) Simulate experimental evolution data using the simulation introduced above and utilizing the proposed parameter values from Step 1.
- (3) Compare the simulated data, \bar{W}_{sim} , to the observed experimental data, \bar{W}_{obs} , using a Euclidean distance measure $d(\cdot)$. If $d(\bar{W}_{sim}, \bar{W}_{obs}) \leq \delta$ then keep the proposed parameter values, where δ is some cutoff distance. Parameters resulting in $d(\bar{W}_{sim}, \bar{W}_{obs}) > \delta$ are “rejected.” δ is chosen such that at most 0.1% of the total number of proposed parameter sets are kept.

We used this methodology for statistical inference and estimation in the following section.

3. Example Application Using Bacteriophage ϕ X174

3.1. Data. To demonstrate the utility of our method, we used data from a single replicate of a phage evolution experiment under low selection pressure, referred to as benign line 2 in [6]. Description of these data along with details of the generating experiment is presented in [6]. In short, ϕ X174 phage was evolved in an environment containing a CaCl_2 concentration slightly lower than that used to preadapt the ancestral population and was allowed to reach a fitness optimum. CaCl_2 is a resource important for phage attachment and facilitates injection of the phage genome into its host [6]. Pepin and Wichman [6] refer to this condition where CaCl_2 was slightly reduced as “benign” because it only slightly reduced population fitness. Under these conditions, the authors observed a high rate of beneficial mutation and occurrence of clonal interference.

3.2. Goal. We aim to estimate the mutation probability μ and the mean selection coefficient λ based on this data set. We benchmark by comparing our results against their findings.

3.3. Setup. In addition to μ and λ , we also estimate W_0 , the absolute fitness of the wild type, and σ^2 , the variance component due to experimental error. This is done to alleviate uncertainty about fitness of the initial population and to account for experimental error. To construct the approximate posterior distributions for these parameters, we follow the above introduced ABC algorithm.

Step 1. We generated $B = 200,000$ parameter values as follows. Beneficial mutation probabilities were proposed using a uniform distribution on the exponent of the mutation parameter with bounds between -10 and -1 (translating to mutation probability per genome between 10^{-10} and 10^{-1}). We chose a uniform distribution on the exponent to improve efficiency in sampling values close to zero. Mean selection parameter values were sampled from a uniform distribution with bounds 0 and 0.5 . The absolute fitness of the initial population was sampled using a uniform distribution with bounds between 0 and 8 , and variances due to measurement error were sampled from a uniform distribution with bounds between 0 and 2 . Limits of these distributions were chosen after conducting a number of calibrating runs and form the prior distributions of these parameters of interest. In the original data, the experimenters used a variable bottleneck dilution factor in an effort to maintain a fixed population size following a bottleneck. Accordingly, we assumed a fixed bottleneck size of 3.6×10^6 phage at the start of each cycle which represents the average bottleneck size over the course of the experiment.

Step 2. We utilized the proposed parameter values to simulate data. Simulation runs were restricted in time to two hours each and were conducted using the Beowulf cluster at the IBEST computational core of the University of Idaho. Incomplete simulations included high mutation rate in association

with high selection parameters resulting in distant outcomes compared to the observed data.

Step 3. We chose δ such that only 0.1% of the proposed parameters were kept resulting in a sample of size 200 of acceptable parameter values that were used to construct the approximate posterior distributions.

We used *R* [10] for all analyses and sampling.

4. Results and Discussion

Figure 2 introduces the approximate posterior densities of the beneficial mutation probability (Figure 2(a)) and mean selection (Figure 2(b)) obtained using our method. It also shows the approximate posterior densities of the initial fitness of the experimental population (Figure 2(c)) and the variation due to sampling error (Figure 2(d)). Given the skewed nature of these distributions, we used the posterior median, rather than the posterior mean, as an estimator of these parameters. We also provide the 95% credibility intervals (CI) for these estimates.

Figure 2(a) shows the posterior density of the logarithm, base 10, of the probability of beneficial mutations. Based on this distribution, our estimate of this parameter was equal to 5.07×10^{-9} (CI = $[1.16 \times 10^{-10}, 2.95 \times 10^{-2}]$) per genome. This wide variation in the estimate of the mutation probability might be attributed to the observed bimodality of this distribution. This bimodality indicates that there are two mutation scenarios that can explain the pattern of experimental evolution observed for this data set. The first of these scenarios involves low probability of beneficial mutation, and the second involves high probability of such mutation. In what follows, we focus on the second scenario and we revisit the first below in conjunction with the associated mean selection. With a chance of about 19% of having beneficial mutations with probability greater than 10^{-5} , we cannot rule the second scenario out as candidate explanation for the pattern observed in these data. The median beneficial mutation probability for this scenario, excluding the first one, is estimated at 7.7×10^{-4} indicating that a total of about 12 beneficial mutations could have occurred through this experiment. This further indicates that about 26% of all possible mutations could be beneficial (assuming that the size of ϕ X174's genome is ~ 5000 bp and that the mutation rate per genome is about 0.003, [11]). Not surprisingly, this high mutation probability is associated with low selection coefficients, median equal to 0.015, reflecting the low selective pressure of this part of the experiment. This scenario is consistent with the 4, out of 9, (45%) beneficial mutations observed by Pepin and Wichman [6] for this lineage under what they refer to as benign experimental conditions.

Our estimate of the mean selection was 0.033 (CI = $[0.0014, 0.2103]$), based on the posterior distribution presented in Figure 2(b). The associated CI is also wide, yet the median estimate is consistent with the benign nature of the experiment. That said, here too, we can identify two plausible scenarios: one with low selection and one with high selection. The probability of having mean selection of more than 0.15

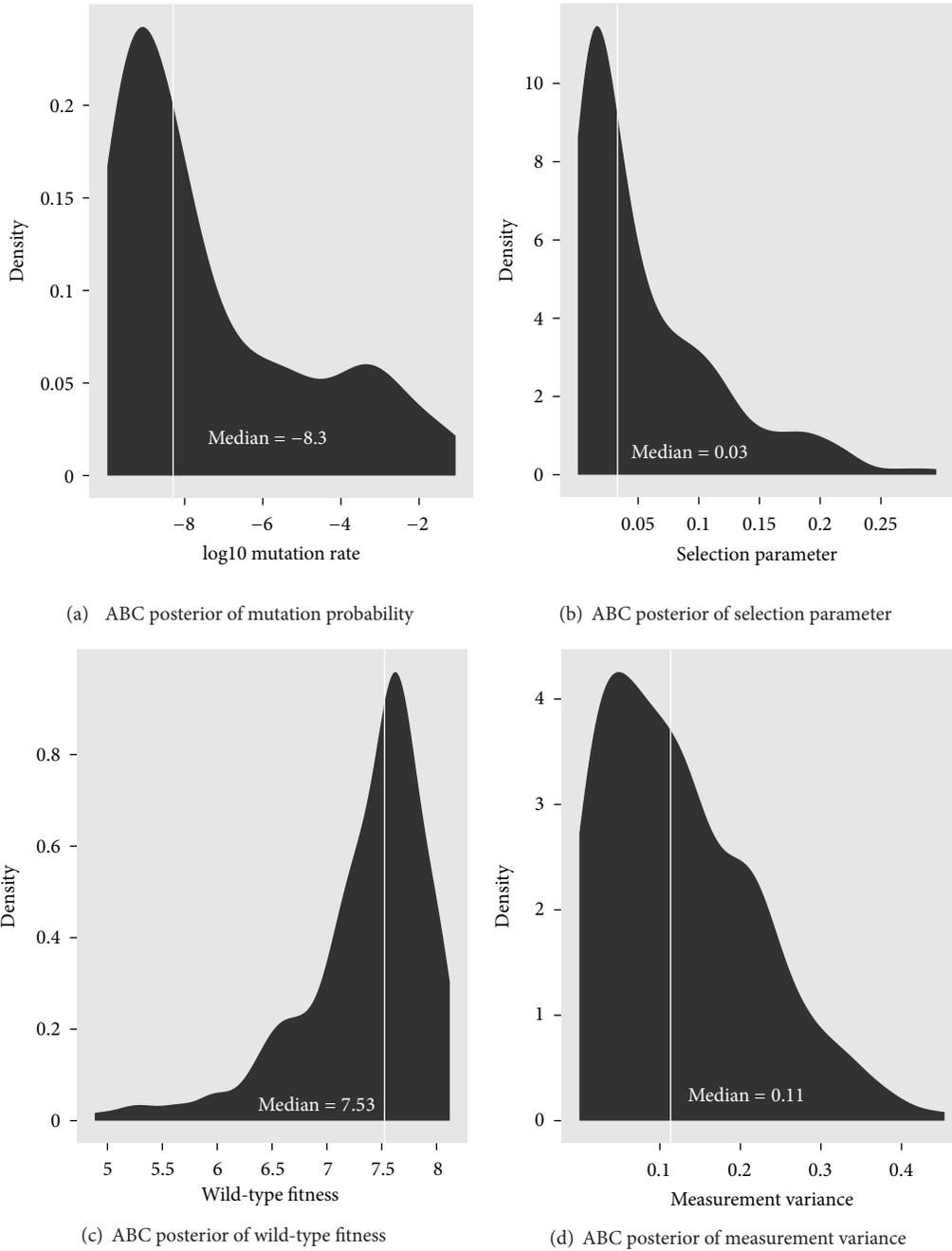


FIGURE 2: Resulting ABC posterior distributions of parameters associated with the phage data.

is about 9.5%. This high selection scenario is associated with very low mutation probability, median 2.24×10^{-10} . Hence, based on this scenario one expects mutations to be rare yet when they occur these mutations are likely to have a profound effect resulting in a drastic increase in fitness relative to the wild type. Although possible, it has low probability of occurring. This scenario is not supported by the associated sequence data.

Figure 3 depicts the Spearman correlation between all variables. This figure indicates a high negative correlation between μ (μ) and S with Spearman correlation coefficient

equal to -0.66 . The associated correlation plot indicates a nonlinear relationship between these two variables with both scenarios, presented above, clearly observed. In concordance with these scenarios, this figure shows that some low mutation probabilities are associated with high selection coefficients and that some high mutation probabilities are associated with low selection coefficients. The figure also highlights the last, most dominant, scenario that associates low mutation probabilities with low selection coefficients. This scenario excludes the two presented above and occurs with probability of about 71.5%. This scenario was what we

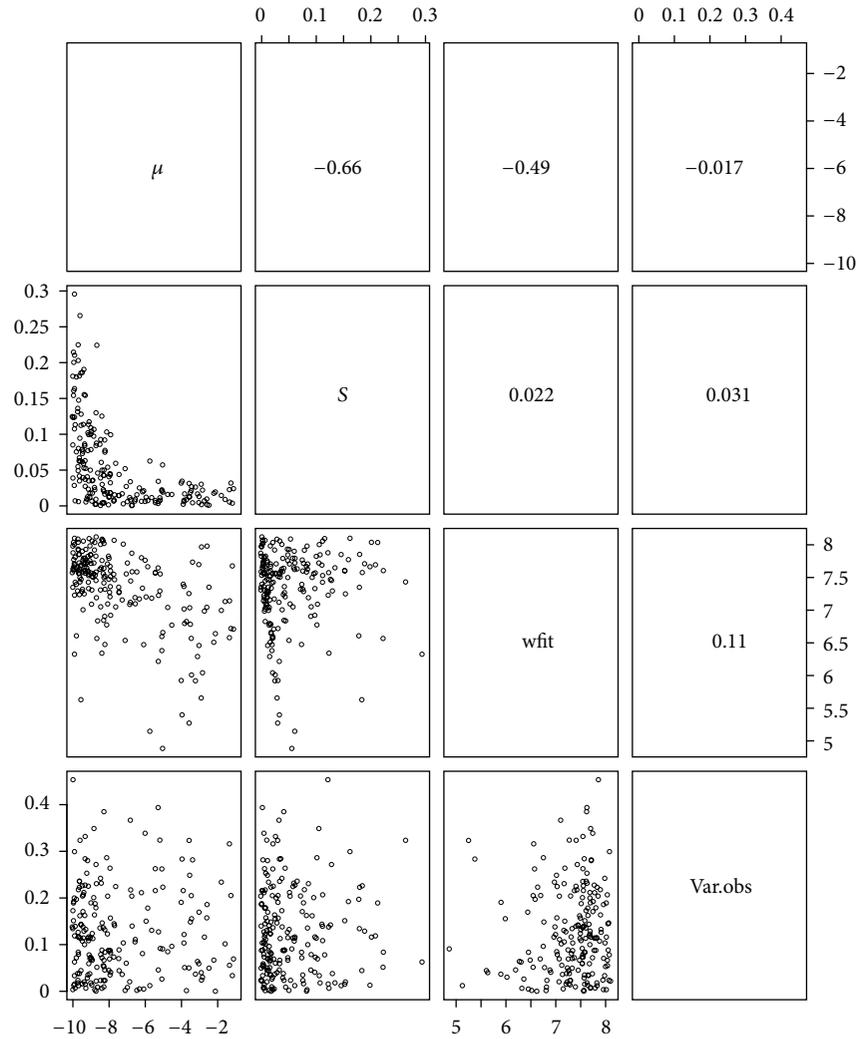


FIGURE 3: Correlation plot between the four different parameters estimated for the phage example, where μ represents the mutation rate, S represents the selection parameter, $wfit$ is the wild-type fitness, and $var.obs$ is the variance of the observation error. Lower triangle shows the correlation graphs while upper triangle introduces the computed Spearman correlation coefficients.

usually expect under the presented experimental conditions. Estimates of the beneficial mutation probability and mean selection parameter were closer to the overall estimate at 2.66×10^{-9} and 0.039, respectively. Though as stated above, the other scenarios were not completely excluded with high probability of occurrence for the first one (high μ and low S) of 19%. This highlights the ability of this method to identify the different candidate explanations of fitness data results prior to sequencing.

The wild-type fitness and the measurement error were introduced as nuisance parameters to account for the possible variation that can result from an error in estimating the initial fitness of the population and due to the estimation of the population fitness during the experiment. These parameters were estimated to be 7.53 (CI = [5.63, 8.07]) and 0.113 (CI = [0.002, 0.339]), based on the posterior distributions presented in Figures 2(c) and 2(d), respectively. Estimate of the wild-type fitness is computed per generation; considering that

we have three generations per cycle of the experiment, this estimate would be 22.59 per cycle and is higher than the 18.0 presented in Pepin and Wichman, [6] yet the latter still lies in the CI interval (CI = [16.89, 24.22]) per cycle. On the other hand, it was quite interesting and surprising that the effect of measurement error was not large, in concordance with the findings of Joyce et al. [3]. Though, incorporating both the uncertainty in our estimate of the wild-type fitness and the measurement error provides a more realistic model for these data. It is worth noting that fitness of the wild type showed negative correlation, Spearman correlation coefficient = -0.49 , with the mutation probability (Figure 3). This is in concordance with our expectation; when starting at a higher fitness within an environment, the expected selection pressure would be lower and less beneficial mutations would be available to improve fitness. Measurement error did not show discernible correlation with any of the other variables (Figure 3).

5. Conclusions

The main goal of this paper was to introduce and demonstrate the use of Approximate Bayesian Computation (ABC) in experimental evolution studies. ABC is a useful approach that allows for the use of complex models to study different aspects of the evolution process by alleviating the need to fit these models to the data. Instead, this approach relies on comparisons between simulated data and the observed data, either directly or through the use of sufficient summary statistics [12]. We showed the use of this approach in inference by utilizing a, deceptively, simple model of experimental evolution. This model was complicated by the required integration over all possible, unobservable trajectories of evolution within an experiment to estimate experimental parameters of interest.

Utilizing our model, combined with the ABC approach, we studied the experimental evolution process of bacteriophage ϕ X174, under benign selection conditions [6]. Our results highlighted three candidate mutation-selection scenarios that could explain the pattern observed in these data. These scenarios were high mutation with low selection, low mutation with high selection, and low mutation with low selection, with an estimated chance of 19%, 9.5%, and 71.5%, respectively, per scenario. Sequence data obtained by Pepin and Wichman [6] favored the first scenario, which was surprising to the authors though not improbable based on our analysis. In this analysis, we accounted for uncertainty of the estimate of the wild-type fitness at the start of the experiment and for uncertainty due to measurement error.

In our approach, we relied solely on fitness data based on one replicate of experimental evolution. Our model and approach can be improved by utilizing more replicates and by utilizing the observed sequence data to improve and focus inference on the region of the parameter space that best describes the evolutionary pattern. This would allow us to obtain better estimates of the parameters of interest, such as mutation and selection, and can better inform our understanding of the experimental evolution process under study, which is an avenue of future research.

Acknowledgments

The authors thank Holly Wichman, Eva Top, Holger Heuer, Julie Hughes, and Hirokazu Yano for constructive discussions. They also thank the reviewers for their comments. Funding was provided by NIH/NCRR P20RR16448 and NIH-1R01AI084918-01A1. Bioinformatics staff and facilities were supported by NIH/NCRR P20RR16448 and P20RR016454. K. M. Pepin was supported by the RAPIDD Program of the Science and Technology Directorate, U.S. Department of Homeland Security, and the Fogarty International Center, NIH.

References

- [1] L. De Gelder, J. M. Ponciano, Z. Abdo, P. Joyce, L. J. Forney, and E. M. Top, "Combining mathematical models and statistical

methods to understand and predict the dynamics of antibiotic-sensitive mutants in a population of resistant bacteria during experimental evolution," *Genetics*, vol. 168, no. 3, pp. 1131–1144, 2004.

- [2] P. J. Gerrish and R. E. Lenski, "The fate of competing beneficial mutations in an asexual population," *Genetica*, vol. 102–103, pp. 127–144, 1998.
- [3] P. Joyce, Z. Abdo, J. M. Ponciano, L. De Gelder, L. J. Forney, and E. M. Top, "Modeling the impact of periodic bottlenecks, unidirectional mutation, and observational error in experimental evolution," *Journal of Mathematical Biology*, vol. 50, no. 6, pp. 645–662, 2005.
- [4] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré, "Markov chain Monte Carlo without likelihoods," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15324–15328, 2003.
- [5] P. Marjoram and S. Tavaré, "Modern computational approaches for analysing molecular genetic variation data," *Nature Reviews Genetics*, vol. 7, no. 10, pp. 759–770, 2006.
- [6] K. M. Pepin and H. A. Wichman, "Experimental evolution and genome sequencing reveal variation in levels of clonal interference in large populations of bacteriophage ϕ X174," *BMC Evolutionary Biology*, vol. 8, no. 1, article 85, 2008.
- [7] J. M. Hughes, B. K. Lohman, G. E. Deckert et al., "The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation," *MBio*, vol. 3, no. 4, Article ID e00077-12, 2012.
- [8] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory*, The Blackburn Press, Caldwell, NJ, USA, 1970.
- [9] J. H. Gillespie, *Population Genetics: A Course Guide*, The Johns Hopkins University Press, Baltimore, Md, USA, 1998.
- [10] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012.
- [11] J. W. Drake, "A constant rate of spontaneous mutation in DNA-based microbes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 16, pp. 7160–7164, 1991.
- [12] P. Joyce and P. Marjoram, "Approximately sufficient statistics and bayesian computation," *Statistical Applications in Genetics and Molecular Biology*, vol. 7, no. 1, article 26, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

