Analysis of Morphological Evolution in a Long-term Experiment with *Escherichia coli*

Fangshu Cui, Bo Yuan
Department of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai 200240, China
cuifangshu001@163.com

***Abstract*—Great attentions are still paid to the morphological evolution, such as the waiting time to the morphological stability in constant environment, the contributions of different evolutionary forces to the morphological evolution and so on, despite considerable progress. To investigate these issues, some biologists seek to carry out evolution experiments owing to the incompleteness and uncontrollability of the fossil record and the natural populations. We analyze the morphology (cell size) evolution observed from a long-term evolution experiment with *Escherichia coli* by Lenski et al. and explore these questions more rigorously. We adopt a population genetics model, the Wright-Fisher model, to describe this morphological evolution and calculate the estimates of the waiting time until the ultimate stasis (near stasis) in morphology (cell size) in the long-term experiment by simulations. These calculations have been verified to be in good accordance with the experimental data, which demonstrates the effectiveness of our model. We have shown how the per-locus mutation rate, the average selection advantage per mutation and the population size devote to the morphology (cell size) evolution. Our results indicate that the selective advantage plays a powerful effect on this morphological evolution. By comparison, the mutation rate and population size have a weaker influence.**

Keywords-morphological evolution; evolution experiment; morphological stability; Wright-Fisher model; evolutionary forces

I. INTRODUCTION

The development of evolutionary biology since 1858 is one of the great intellectual achievements of science, written by the philosopher Kim Sterelny [1]. To date, great progress has been made in evolutionary biology. Evolutionary biologists are still attracted by the morphologies that influence the procreative success of the individual organisms, such as how fast populations vary in these morphologies, and whether the rates of variation are constant or variable, how long it will take to reach the morphological stability in constant environment, how the mutation rate, selective advantage and population size affect the morphological evolution, whether mutation or selection is the dominant force in the morphological evolution and so on. In order to examine these questions, evolutionary biologists have relied on the data from the fossil record and the studies of natural populations [2-6]. However, a complete and explicit answer is hard to obtain due to the incompleteness and irreproducibility of the fossil data and the complexity and uncontrollability of the natural populations.

Since 1988, Lenski et al. have embarked on a long-term evolution experiment with *Escherichia coli* in the laboratory. The *Escherichia coli* has many advantages for evolution experiments. For example, it propagates quickly, which allows experiments to run for many generations, and it can be stored in suspended animation and later resurrected, which provides information on the dynamics of the evolutionary process and the extent of evolutionary change by means of the direct comparison of ancestral and evolved types. In virtue of the control of many variables in a laboratory setting, many questions about evolution can be explored with greater rigor. This long-term experiment has shown more complicated and extensive evolutionary dynamics than research of responses to selection that depend either on quantitative variance in a population or on a single allele of major effect. In this experiment, we have perceived that morphology (cell size) of *Escherichia coli* evolves rapidly after the introduction of the study organism into the constant experimental environment, followed by eventual stasis (or near stasis) [7].

With the purpose of a deep and stringent understanding of this morphological evolution, we resort to population genetics, which is the study of the genetic composition of populations and changes under the influence of various factors. By developing mathematical models, we can shed light on various evolutionary processes in a quantitatively accurate way. In this paper, a population genetics model, the Wright-Fisher model, is employed to depict this morphological evolution in the long-term experiment. This is helpful for us to explain how the evolutionary factors work in the progression of cell size. By extensive simulations of our model, we compute the estimates of the waiting time until the morphology (cell size) of *Escherichia coli* is static (or near static) in the long-term experiment, which are in good agreement with the experimental data. Moreover, we analyze and interpret this morphology (cell size) evolution in terms of the per-locus mutation rate, the average selective advantage per mutation and the population size, which provides an essential comprehension of how the different evolutionary forces conduce to this evolutionary process. Our results clarifies that this progression of cell size is strongly influenced by the average selective advantage, whereas the mutation rate and population size play a smaller effect.

II. LONG-TERM EVOLUTION EXPERIMENT

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

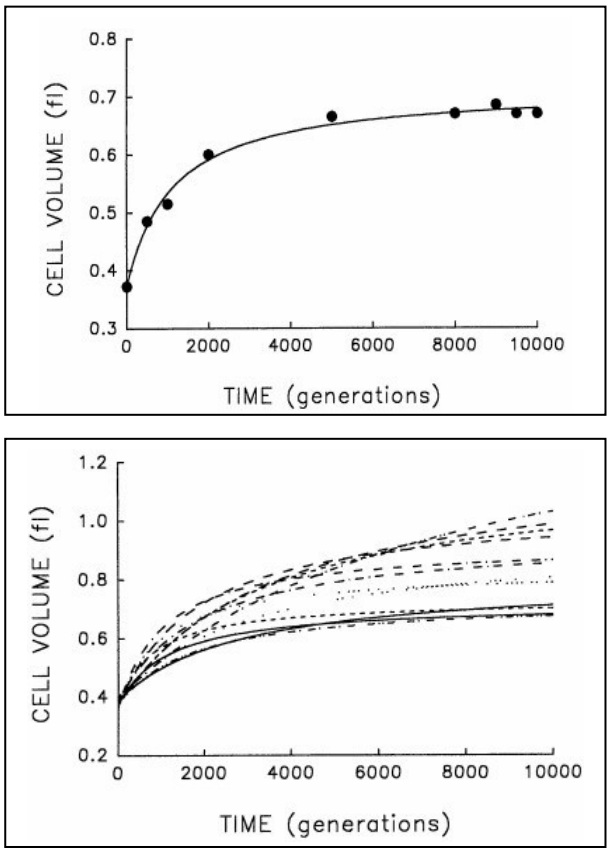

Figure 1. (Top) Trajectory for average cell volume in one population of Escherichia coli during 10, 000 generations of experimental evolution. (Bottom) Trajectories for average cell volume in 12 replicate populations of Escherichia coli during 10, 000 generations (Lenski and Travisano, 1994) [7].

The experiment includes 12 replicate populations of *Escherichia coli*, which have been propagated in identical environment since 1988 [8]. All the populations were founded from the same ancestor, excepting only the neutral marker applied to population identification. Thus, there was essentially no genetic variation either within or between populations, and each population depended totally on new mutations for its subsequent evolution.

The *Escherichia coli* was grown in a glucose-limited minimal salts medium, which was supplemented with glucose at a concentration of 25 mg per L. Each culture was 10 ml, which allowed the bacteria to reach about 5×10^8 cells when the glucose was depleted. The cultures were incubated in 50-ml Erlenmeyer flasks in a shaking incubator at 37℃ and 120 rpm. There was a basic rhythm in this experiment that each population was serially transferred by diluting 0.1 ml into 9.9 ml of fresh medium every day. In the course of 24-h cycle, the populations grew until they had exhausted available resources in the first eight hours or thereabouts, and then was in stationary phase. This course allowed $\log_2 100 \approx 6.64$ generations of growth for each population. Samples from each population were stored in a glycerol-based suspension at -80℃ at intervals, firstly every 100 generations and later 500 generations. Hence, the samples can be recovered, for estimating mean fitness of derived populations relative to their ancestor, or to restart the population if needed and so on.

There were no active protophage or plasmids to mediate horizontal gene transfer in this experiment; the *Escherichia coli* strain was strictly asexual. Particularly, the populations in this experiment lacked any mechanism for genetic exchange, and consequently mutations provided the unique source of genetic variation. The experimentalists did not artificially choose individual organisms on the basis of any special trait. Any mutation that conferred some competitive advantage in exploiting the constant experimental environment would have been favored by natural selection.

In this experiment, the evolutionary dynamics of morphology (cell size) in the evolving populations has been rendered. Cell volume increases quickly for the first about 2000 generations but is almost static for the last several thousand generations (see Figure 1). The procedure for measuring average cell volumes has been described in detail elsewhere [8-10]. And materials and methods are also represented in these literatures, which involve the founding strain, experimental conditions, contamination checks, etc.

III. Model and Methods

For the purpose of mathematical modeling of morphology (cell size) evolution in the long-term experiment, we consider one population of *Escherichia coli* with a constant size N. We assume that each organism (cell) has l susceptible genes that are involved in cell size, and mutations occur independently at rate u per locus. The parameter s ($s>0$) is regarded as the average fitness advantage per mutation. The organism with k mutations out of l sensitive loci is called k-fold mutant, and let $x_k(t)$ be the number of k-fold mutants at time (generation) t, and $r_k(t)=x_k(t)/N$ be the relative frequency. In the long-term experiment, a rational interpretation for the ultimate stasis (near stasis) in cell size is that the organisms have used up all ways to increase their cell sizes to become much better adapted to the constant environment. Here, we assume that this will happen if any m out of l susceptible loci are mutated in a single organism, i.e. $r_m(t)=1/N$. Hence, the relative fitness of a k-fold mutant is

$$q_k=\frac{(1+s)^k}{\sum_{j=0}^{m}(1+s)^j r_j}, \tag{1}$$

where $k=0,1,\cdots$.

We presume that the progression of cell size evolves according to the Wright-Fisher model [11], which is a widely used stochastic model of evolving populations first introduced by Fisher (1922) and Wright (1931). Within this model, the organisms evolve in discrete, non-overlapping generations. Each organism independently originates from a parent organism of the previous generation with a probability proportional to the fitness of the parent. The organisms are clonally inherited, so that each organism is exactly the same as its parent aside from the additional mutations with probability u per locus. Moreover, the daughter organism inherits the mutations of the mother organism (Figure 2). This process goes on until an organism has accumulated m mutations out of l

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

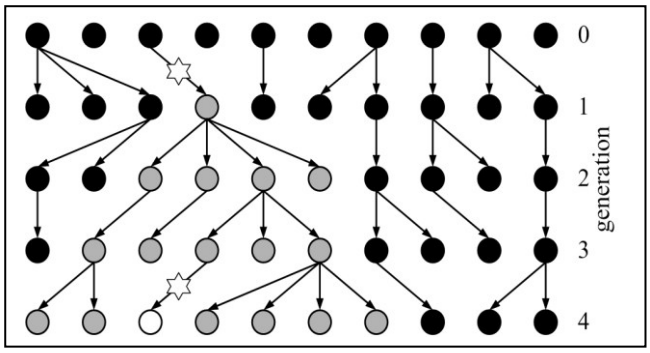

Figure 2. Diagrammatic sketch of the Wright-Fisher model for describing the cell size evolution. Shown is a population with a constant size of $N=10$ organisms, evolving according to the Wright-Fisher model. Organisms independently spring from the parent organisms of the previous generation with the probability proportional to the fitness of the parent. Originally, at generation $t=0$, all organisms are of wild type (black). At generation $t=1$, the first organism with one mutation occurs (gray), which has a selective advantage to produce offspring more probably. At generation $t=4$, the organism with two mutations firstly appears (white), and the waiting time until the accumulation of two mutations is $t_2=4$ generations.

susceptible genes that are involved in cell size, when the cell size of *Escherichia coli* is static (or near static) in the long-term experiment.

Initially, all organisms are of the wild type, and the conformation of the numbers of k-fold mutants at time (generation) $t=0$ is

$$\left(x_0(0),x_1(0),\cdots \qquad \cdots \quad . \right. \tag{2}$$

Neglecting back mutation, the probability that an organism in the next generation will have k mutations is

$$p_k=\sum_{i=0}^{k}\binom{l-i}{k-i}u^{k-i}(1-u)^{l-k}q_i r_i(t)\,, \tag{3}$$

where $k=0,1,\cdots$, and $\sum_{k=0}^{m}p_k=1$.

Therefore, the probability of the conformation $\left(x_0(t+1),x_1(t+1),\cdots \quad 1)\right)$ is offered by the multinomial distribution

$$\frac{N!}{x_0(t)!x_1(t)!\cdots}\qquad \prod_{=0}^{m}p_k^{x_k(t)}\,, \tag{4}$$

where $\sum_{k=0}^{m}x_k(t)=N$.

This Wright-Fisher process for describing the evolution of cell size stops when $x_m(t)>0$, i.e. the m-fold mutant firstly appears in the population of *Escherichia coli* in the long-term experiment.

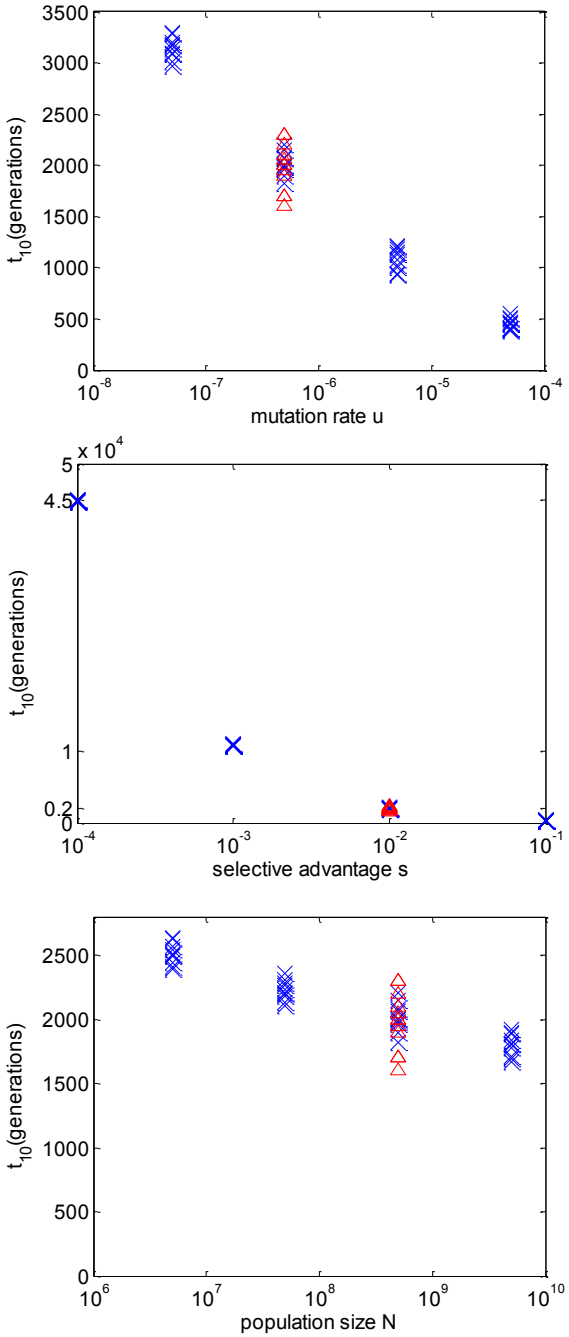

Figure 3. The anticipated time to the occurrence of an organism with $m=10$ mutations out of $l=20$ sensitive loci, t_{10}, which is simulated by the Wright-Fisher process for cell size evolution and plotted versus (Top) the mutation rate u per locus, when $N=5\times10^8$, $s=0.01$; (Middle) the selective advantage s per mutation, when $N=5\times10^8$, $u=5\times10^{-7}$; and (Bottom) the population size N, when $u=5\times10^{-7}$, $s=0.01$. The crosses show the results of 10 independent simulations at each parameter set and the triangles are the experimental data in the long-term experiment with *Escherichia coli*.

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

IV. RESULTS AND DISCUSSION

A. Results

We have applied a population genetics model, i.e. Wright-Fisher model, to study the evolutionary process of morphology (cell size) of *Escherichia coli* in the long-term experiment. In this model, the waiting time to the final stasis (near stasis) in cell size is equal to the time until the first m-fold mutant occurs in the population, i.e. t_m. We assume $l = 20$ susceptible loci, and set $N = 5\times10^8$, $s = 0.01$, $u = 5\times10^{-7}$ from the experimental data [7,12]. Considering $m = 10$, we obtain the estimates of the time t_{10} by simulations, which are well consistent with the data in the long-term experiment (see Figure 1, Figure 3). Thus, the Wright-Fisher model is verified to be effective for the evolutionary process of cell size, and gives a definite tradeoff among the evolutionary forces.

For a broad range of parameters, the anticipated time for developing the first m-fold mutant t_m, for $m = 10$, is depicted in Figure 3. The time t_{10} decreases with increasing population size, with increasing selection strength, and with increasing mutation rate. Furthermore, t_{10} diminishes fastest as the selection strength increases, which demonstrates that the selective advantage has a powerful impact on the genetic progression of cell size. Alternatively, a higher selective advantage would enable a smaller population with a smaller mutation rate to attain the stability (near stability) of cell size more quickly.

B. Discussion

Great progress has been made in the application of population genetics models to investigate the expected time until the accumulation of *m* mutations, since Armitage and Doll and Knudson began to engage in this theoretical work [13, 14]. Iwasa et al. [15] explored a two-stage model for a population of cells, which evolves according to the Moran model. Later, Beerenwinkel et al. [16] studied the anticipated time to reach *m* mutations in terms of the Wright-Fisher model for very large population sizes. Schweinsberg and Durrett and Schweinsberg originated the asymptotic distributions of the waiting time to *m* mutations [17, 18]. In this paper, we use the discrete Wright-Fisher process to describe the genetic progression of morphology (cell size) of *Escherichia coli*. Although the Moran model has the advantage of being mathematically more amenable than the Wright-Fisher model, and both models have the similar performance for large population sizes, the Wright-Fisher process permits for much more efficient computer simulations than the Moran model [11, 16].

Within our model, we view the selective advantage *s* as the average fitness increment per mutation. In fact, each mutation has a different impact on the fitness of the organism [12]. Usually, the effects of one mutation on the phenotype of the organism are influenced in a complicated way by the hereditary factors. Epistasis, i.e. interactions between genes, which is the phenomenon where the impacts of one gene are modified by one or several other genes, can play an effect on the accumulation of the mutations. But here the effects of epistasis are shared equally between the mutations.

With numerical simulations of our model, we have effectively calculated estimates of the time until the appearance of the first m-fold mutant, which are in good accordance with the experimental data. Furthermore, we have shown how varying mutation rate, selection advantage and population size affect the waiting time to the stability of cell size in the long-term experiment. Here, the selective advantage has been manifested to play a vigorous effect on this morphological evolution. Relatively, the mutation rate and population size have less influence.

ACKNOWLEDGMENT

This work was supported in part by project No. Z-XT-003 (S.J.).

REFERENCES

[1] K. Sterelny, Philosophy of Evolutionary Thought, In Michael Ruse & Joseph Travis, Evolution: The First Four Billion Years, Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2009, pp. 313.

[2] D. Jablonski, "Micro- and macroevolution scale and hierarchy in evolutionary biology and paleobiology," in D. H. Erwin and S. L.Wing, Eds. Deep time: Paleobiology's perspective. Paleobiology 26(Suppl. to No. 4), 2000, pp. 15–52.

[3] R. B. Huey, G. W. Gilchrist, M. L. Carlson, D. Berrigan, and L. Serra, "Rapid evolution of a geographic cline in size in an introduced fly," Science 287, 2000, pp. 308–309.

[4] C. D. Thomas, E. J. Bodsworth, R. J. Wilson, A. D. Simmons, Z. G. Davies, M. Musche, and L. Conradt, "Ecological and evolutionary processes at expanding range margins," Nature 411, 2001, pp. 577–581.

[5] D. J. Futuyma, "Evolutionary constraint and ecological consequences," Evolution 64,2010, pp. 1865–1884.

[6] A. Matamoro-vidal, C. A. Furness, P. –H. Gouyon, K. J. Wurdack, and B. Albert, "Evolutionary stasis in Euphorbiaceae pollen: selection and constraints," Journal of Evolutionary Biology, Volume 25, Issue 6, pp. 1077–1096, June 2012.

[7] R. E. Lenski and M. Travisano, "Dynamics of adaptation and diversification: a 10,000-generation experiment with bacteria populations," Proc. Natl. Acad. Sci. USA 91, 1994, pp. 6808–6814.

[8] R. E. Lenski, M. R. Rose, S. C. Simpson, and S. C. Tadler, "Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am. Nat. 138, 1991, pp. 1315-1341.

[9] S. F. Elena and R. E. Lenski, "Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation," Nat. Rev. Genet. 4, 2003, pp. 457.

[10] R. E. Lenski, "Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*," Plant Breed. Rev. 24, 2004, pp. 225-265.

[11] W. J. Ewens, Mathematical population genetics, New York: Springer, 2004.

[12] J. E. Barrick, D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski, and J. F. Kim, "Genome evolution and adaptation in a long-term experiment with *Escherichia coli*," Nature 461, 2009, pp. 1243–1247.

[13] P. Armitage and R. Doll, "The age distribution of cancer and a multi-stage theory of carcinogenesis," Brit. J. Cancer, 11, 1954, pp. 161-169.

[14] A. G. Knudson, "Genetic instabilities in human cancers," Proc. Natl. Acad. Sci. U.S.A., 68, 1971, pp. 820-823.

[15] Y. Iwasa, F. Michor, and M. A. Nowak, "Stochastic tunnels in evolutionary dynamics," Genetics, 166, 2004, pp. 1571-1579.

978-1-4673-4398-5/12/$31.00 ©2012 IEEE

[16] N. Beerenwinkel, T. Antal, D. Dingli, A. Traulsen, K. W. Kinzler, V. E. Velculescu, B. Vogelstein, and M. A. Nowak, "Genetic progression and the waiting time to cancer," PLoS Comput. Biol, 3, 2007, pp. 2239-2246.

[17] J. Schweinsberg, "The waiting time for m mutations," Electron. J. Probab., 13, 2008, pp. 1442-1478.

[18] R. Durrett, D. Schmidt, and J. Schweinsberg, "A waiting time problem arising from the study of multi-stage carcinogenesis," Ann. Appl. Probab, 19, 2009, pp. 676-718.

978-1-4673-4398-5/12/$31.00 ©2012 IEEE