Modeling Mutator Advantage in Stress-Induced Mutagenesis

Kenneth Hu
8.592 Final Project
(Dated: May 12, 2011)

Stress-Induced Mutagenesis (SIM) is the increase in mutation frequency observed in a number of bacterial populations following periods of stress. A number of hypotheses have been proposed to explain this often dramatic increase in mutation frequency. One such theory proposes that these bacteria induce a molecular response to stress which results in error-prone DNA replication and repair. It is hypothesized that such a pathway has evolved in natural bacterial populations which face a constantly fluctuating environment. However, it is not immediately clear that a global increase in mutation rate is necessarily beneficial, and if so, for what kinds of conditions. In this paper, we attempt to simulate a simple scenario of a population of *Escherichia coli* consisting of normal and SIM-capable bacteria placed in a stressful environment. The simulation shows that indeed a mutator genotype is evolutionarily advantageous in a stressful environment. In addition, we show that being able to switch off hypermutation by SIM mechanisms in our model provides an even better fitness relative to normal.

I. INTRODUCTION

Darwinian evolution has classically been based around a steady spontaneous mutation rate in populations which generates the genetic diversity that natural selection acts on. Mutation creates the raw material that natural selection refines into adaptations. Obviously, the rate of mutation is then of great importance when considering the rate of evolution. Thus the rate of mutation itself can be a valuable parameter to adjust in a population as it adapts to a variety of environments. This selection for certain rates of mutagenesis or second-order selection has been theorized to explain the experimental observation of stress-induced mutagenesis (SIM).

It is immediately obvious that for well-adapted populations growing at near-optimal rates, the lower the mutation rate, the better. A fundamental feature of biological systems is that deleterious and even lethal mutations are much more likely to occur than beneficial. Simply put, it is much easier to break a biological system, than to improve it. Thus increasing the rate of mutation would result in more of a genetic burden than any possible improvement. As such, there exists a vast array of molecular mechanisms to proofread and ensure genetic stability. Mutations still occur and this spontaneous mutation rate has been measured at around $5X10^{-10}$ mutations per base pair per generation in *Escherichia coli.*[1]

However, in natural settings, there are a variety of bacterial environments that are far from ideal. These volatile environments are constantly fluctuating in terms of nutrients, pH, temperature, and other factors. In these situations, the bacterial population will find itself under sometimes intense selection for beneficial mutations that are adaptive for the new environment. In these situations, it has been experimentally verified that mutator phenotypes are evolutionarily advantageous.[2] For example, Gibson et al. showed that a *mutT E. coli* strain, which displays a 1000 fold increase in mutation rates, outcompetes wild-type mut^+ organisms.[3] In addition, strains with loss of function mutations in the methyl-directed mismatch repair system (MRS) display mutation rates 100-fold higher and have been isolated from natural sources.[4]

Intuitively, it seems to make sense that under extremely stressful situations, an increased mutation rate would be helpful. A higher mutation rate should create much more genetic diversity for selection to act on. However, there is also the tradeoff that a higher mutation rate also means a higher rate of accumulating deleterious or even lethal mutations. In these situations, a constitutive mutator faces a delicate balance between a higher likelihood of a beneficial mutation as well as a higher likelihood of becoming burdened with fitness-reducing mutations. Thus, it is interesting to see if a constitutive mutator phenotype is truly advantageous in such a situation, and if so, under what situational parameters.This we will investigate using the simulation described below.

II. METHODS

A. Experimental Basis

In this paper, our simulation will be motivated by starvation experiments performed by Mao et al. in which they used carbon starvation as the selective pressure.[2] They utilized a Lac^- *E. coli* strain which was able to revert to a Lac^+ phenotype by making a simple addition of G to a G-G-G-G-G-G sequence. The bacteria were grown in minimal lactose media. The rate of reversion was measured by simply counting growing colonies.

B. Simulation Parameters

In order to understand the evolutionary dynamics of SIM, we will assess the ability of a mutator subpopulation to fix or take over the total bacterial population. We start with two alleles at the mutator gene with m^+ denoting a mutator allele and m^- denoting a non-mutator. We will

begin the simulation with equal frequencies of these two subpopulations such that the frequencies at generation 0 of each genotype (f_{m+}, f_{m-}) are 0.5.

Our simulation will incorporate one main beneficial gene. This gene is analogous to the Lac gene in the Mao et al. experiment. This single gene, when mutated from starting allele b^- to allele b^+ will provide a fitness advantage of b. Our model will also incorporate 3 deleterious genes. This choice was partly motivated by simplicity as well as previous simulations done in Tenaillon et al.[5] This choice of deleterious genes is somewhat arbitrary, but can be adjusted in future simulations to better represent bacterial genomes. Our bacterial population starts in state d^0 denoting no deleterious genes have been mutated into their harmful alleles. Then d^n denotes that n deleterious genes have been mutated into their harmful allele. We will treat all 3 genes as being identical, thus it does not matter which is mutated, only how many. The fitness penalty will then be represented as $-0.05 * n$. Finally, to model lethal mutations, we begin with state l^- and can mutate to the lethal l^+ allele state. Bacteria with the l^+ allele are automatically assigned a fitness $s = -1$.

In terms of fitness, the population starts with $s = 1$. This applies for both mutator and non-mutator. Here we ignore possible pleiotropic effects which we will discuss later. Thus mutators and non-mutators with the same beneficial/deleterious alleles have the same fitness. So for example, a m^- and m^+ with the beneficial allele and two deleterious alleles has fitness $s = 1 + b - 0.1$. All bacteria with the l^+ allele are automatically placed in a "lethal" state which is given $s = -1$.

Mutation rates describe the flow between these different allele states. In this simulation, we will use mutation rates estimated from Kibota and Lynch (1996) and also used in Tenaillon et al. (1999).[6][5] μ_d, the rate of mutation to a deleterious allele was set at 10^{-4}.In addition, there is a chance of reverting the deleterious mutation denoted by $\mu_{dr} = 10^{-8}$. Lethal mutations ocurred at a rate $\mu_l = 10^{-5}$. The rate of mutation to the beneficial allele was set at $\mu_b = 10^{-8}$. For the mutator strains with the m^+ allele, all mutation rates were increased by a factor of m which we will vary in the course of the simulation.

In summary, we have a total of 9 states for both mutators and non-mutators. We did not take into account any mutation from m^+ to m^- state or vice-versa. The individual can have either 0,1,2,3 deleterious mutations, have the beneficial mutation or not, or be in lethal state. In addition, we now have mutation rates representing rates of flow from each state to another. In the course of the simulation, we keep track of these 18 possible genotypes in a frequency array. We only record the frequency of the genotype in the population.

C. Algorithm

For the simulation, we make use of a similar setup done in Tenaillon et al. (1999).[5] The first step is the growth or replication step. In this step, we increment all genotypes according to their fitness and renormalize their frequencies. So for genotype g with fitness s_g, the frequency of this genotype, f'_g, can be represented after this step as:

$$f'_g = f_g \left(\frac{1 + s_g}{1 + \sum_{allgenotypesi} f_i s_i} \right) \tag{1}$$

After the growth step, the simulation enters the mutation step, which causes shifts of frequencies amongst the genotypes. This is the key differential step between non-mutators and mutators. We can draw a network of states with the associated flows which we will not show here in the interests of space. This allows us to create a "mutation matrix" which can be multiplied by our frequency array to give the new frequencies. The following equation describes the changes in genotype frequencies after mutation:

$$f''_g = f'_g \left(1 - \sum_{i \neq g} \mu_{g \to i} \right) + \sum_{i \neq g} f'_i \mu_{i \to g} \tag{2}$$

After the mutation step, all frequencies should still be normalized. For our simulation, we deal with a finite fixed population size. In order to select the next generation of individuals that proceed to the next generation, we used a Poisson sampling process. Using the frequencies at this point, we sample from a Poisson distribution with mean equal to $f''_g N$ where $N = 10^6$ is our population size. Finally, with the new numbers of individuals for each genotype, we normalize again to the new total and this frequency array moves onto the next iteration.

D. Data Collected

The first set of data we are interested in is simply the progression of the mutator frequency. This mutator frequency, f_{m+}, is simply the sum of the frequency of all genotypes that have allele m^+. We run the simulation for 1500 steps and keep track of this frequency, then plot versus generation. The graphs presented display 20 independent runs of the simulation.

The second set of data we are interested in is the fixation frequency. This is the frequency of the mutator subpopulation taking over the entire bacterial population, $f_{m+} = 1$. We ran the simulation 100 times and keep track of how many times we observe fixation.

III. RESULTS FOR THE CONSTITUTIVE MUTATOR

Progression of mutator frequency for 20 runs of the simulation are shown below for different parameters. We varied m or the strength of the mutator, and b, the benefit to fitness of the beneficial allele. The first striking obser-

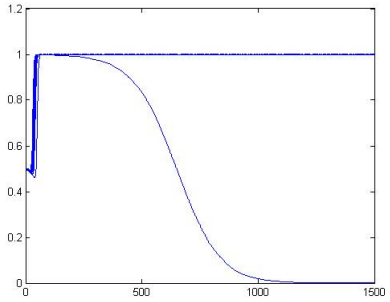

FIG. 1: Mutator frequency progression for m=100, b=1

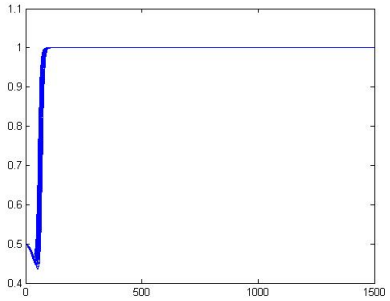

FIG. 2: Mutator frequency progression for m=100, b=0.5

vation is that even down to a fitness benefit of $b = 0.05$ in Fig. 5 which is of the same level as the deleterious effects, a large fraction of the time the mutator allele fixes. As b increases, we can see the mutators begin to fix almost all the time. This immediately suggests that the constitutive mutator phenotype is evolutionarily advantageous in certain situations. The situation itself is quite crucial in determining the advantageousness of the mutator. As we can see in Fig. 6, too low of a benefit makes the mutator genotype disadvantageous. This data displays the inherent balancing act that mutator strains face: they have to balance the positive effects of a beneficial mutation with the more likely negative effects from deleterious and lethal mutations. As the reward increases, the expected benefit for the mutator also increases.

The shape of these curves also suggests a narrative for the progression of the mutator genotype. As the benefit decreases, so does the average time it takes for the mutator to fix. The graphs are progressively pushed out further as b decreases. In addition, we note that many curves display an initial dip in mutator frequency, before rising up and increasing to 1. This initial decrease represents the initial penalties the mutator population incurs

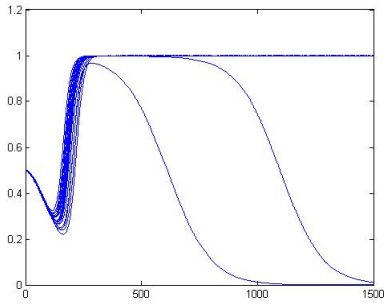

FIG. 3: Mutator frequency progression for m=100, b=0.15

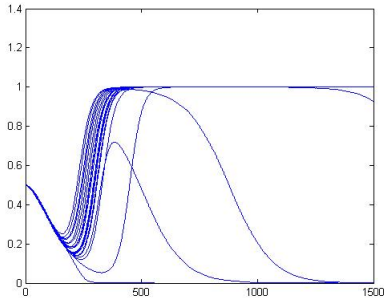

FIG. 4: Mutator frequency progression for m=100, b=0.1

in terms of deleterious and lethal mutations. These mutations are much more likely than the beneficial one and generally appear earlier. The average fitness of the mutators is thus initially lower. However, as time progresses, some individuals acquire the beneficial mutation. Once this step has occured, the mutator individuals carrying these beneficial mutations start to leverage their higher fitness and rapidly take over. This idea of the mutator allele "hitchhiking" along with beneficial mutations is one of the major explanations for the advantage of mutators in stressful environments. One can think of these select individuals carrying the beneficial mutations as "rescuing" the mutator allele from its initial descent in fitness. However, this crucial step is stochastic as we can see in many cases; this "rescue" fails, perhaps coming too late or a victim of genetic drift. In these cases, the intitial burden of deleterious and lethal mutations drags down the mutator genotype to loss. The frequency of loss clearly increases as b decreases. In addition, note that the dip widens and bottoms out at progressively lower values as b decreases. This shows that as the benefit decreases, the individuals carrying the beneficial mutation do not proliferate as quickly and thus the rescue phase is more drawn out.

The next parameter of interest was the mutator strength itself or m. We varied its value from 10 to 1000 to represent the natural range of mutator strengths observed in natural isolates. Comparing figures 7 and 8 to the earlier figure 2, we can see that at high b ($b = 0.5$), a higher mutation rate is more advantageous and allows

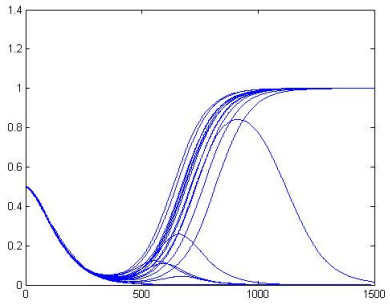

FIG. 5: Mutator frequency progression for m=100, b=0.05

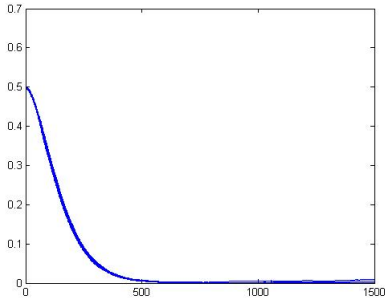

FIG. 6: Mutator frequency progression for m=100, b=0.025

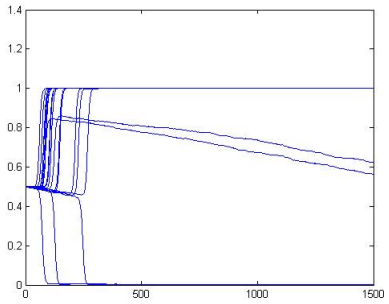

FIG. 7: Mutator frequency progression for m=10, b=0.5

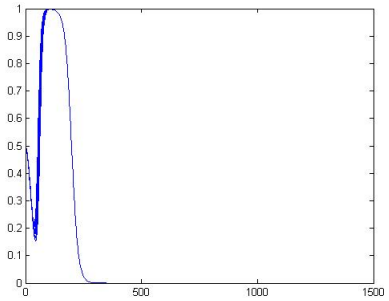

FIG. 8: Mutator frequency progression for m=1000, b=0.5

for faster fixation of the mutator allele. Meanwhile a low mutation rate can result in a failure to rescue the mutator allele in time. This makes sense since having a higher mutation frequency would allow greater likelihood of picking up the beneficial mutation early on and thus rescue.

Here we see the effects of mutator strength in the regime of low b ($b = 0.1$). Comparing figures 9 and 10 to figure 4, we can see several notable features. Clearly a high mutation rate at low b is disadvantageous. This can be explained by the rapid generation of deleterious and lethal mutants in the mutator population before any appreciable benefit to fitness. With too low a benefit, mutating at higher rates is simply not worth it from a fitness perspective. Meanwhile, a low mutation rate avoids this problem, and while it does lag behind $m = 100$, it fixes the mutator allele most of the time. Thus at low benefit regime, a high mutator strength is quite disadvantageous since it accelerates the accumulation of genetic burden on the mutator before the weak benefit can rescue.

We have varied the mutator strength and observed varying behavior for different b. We would now like to examine this constitutive mutator on a more quantitative level. To do so, we will measure the fixation frequencies for varying parameters. We first observed fixation probabilities for varying b at $m = 100$. The graph is shown in Fig. 11. We observe a sharp rise in fixation frequency at low benefits (from $b = 0.05$ to $b = 0.1$). However beyond $b = 0.1$, we see the increase start to level off. We observe a slight increase in frequency as b goes to 1.5, but the slope starts to greatly level off. This suggests that at higher benefits, the actual rate of mutation is the limiting factor. The chance of getting a beneficial mutation is governed solely by the mutation frequency and thus the mutator strength m. Thus for a given mutator strengthbm, we will see a leveling off of fixation frequency at some point.

Naturally, an interesting question to ask is how fixation frequency depends on m. Does it increase monotonically with m or is there a maximum at some m? In order to study this, we plotted fixation frequency versus strength of the mutator m for $b = 0.15$. The result is shown in Fig.12. We observe that fixation frequency increases slightly then peaks at around $m = 100$ and then sharply falls from there. There does indeed appear to be a maximum at around $m = 100$. We can explain this behavior by intuition. At low mutator strengths, one does not acquire a beneficial mutation as frequently and as such, the likelihood of rescue decreases. At very high mutator frequencies, we start to accumulate too many deleterious and lethal mutations too quickly for the beneficial allele to rescue. Thus we seek a balance between too little and too much mutation which is achieved at the maximum. Further runs are needed to examine the shape of this graph as we vary b. Will the maximum shift to the right as b increases? This seems to be the most intuitive result, but needs further testing which we do not perform in this paper.

What we have depicted in this section is a simulation of a constitutive mutator bacterial population. This popu-

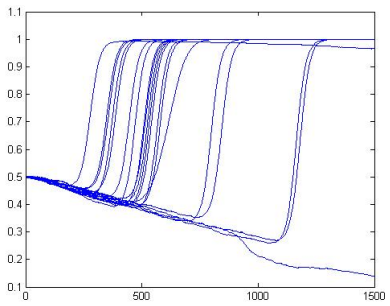

FIG. 9: Mutator frequency progression for m=10, b=0.1

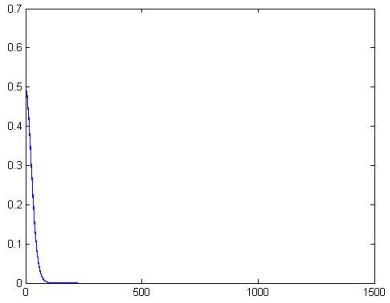

FIG. 10: Mutator frequency progression for m=1000, b=0.1

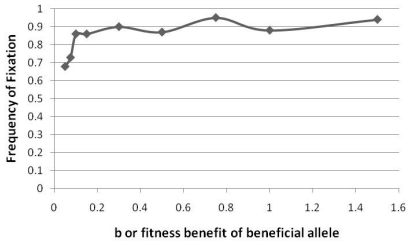

FIG. 11: Fixation Frequencies for Varying b with fixed $m = 100$

lation has a high overall mutation rate for the duration of the simulation, even when it has mutated the beneficial gene. It is quite clear that in the short run, this strategy can be advantageous in the right situations. However, it is questionable how viable this strategy may be in the long run. Once we start looking at the long run, it is clear that these constitutive mutators will keep acquiring deleterious mutations, even when they are well-adapted, slowly weighing down the population's average fitness.There is however an alternative strategy: inducing hypermutation only under stressful conditions.

IV. INCORPORATING STRESS-INDUCED MUTATION

As it so happens, many bacteria possess the capability to control the rate at which new mutations are

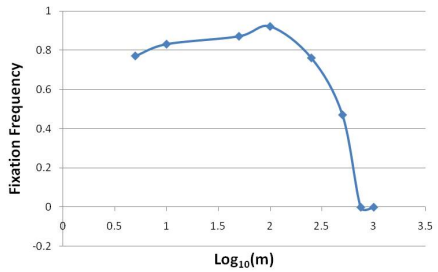

FIG. 12: Fixation Frequencies for Varying m at $b = 0.15$

introduced into their genome. Experimental evidence has shown multiple species of bacteria greatly increase their mutation rates when exposed to a variety of environmental stresses and then downregulate their rate of mutation once the stressor is absent.[7] These stresses include oxidative stress, starvation, UV irradiation, and antibiotics. This regulated increase in mutation rate is accomplished through several molecular pathways. The stress response induces a switch from the high fidelity polymerase normally used, to PolIV, a highly error-prone DNA polymerase.[8]

There is an ongoing debate over the evolutionary cause behind this observed effect. The pleiotropic hypothesis claims that this switch to error-prone mechanisms is simply a side-effect of the stress. There is evidence that by skimping on proofreading, the cell can save some significant amount of energy. One example is the MRS system for proofreading which requires specialized enzymes which cost energy to synthesize. During stress, the bacterium may simply place its priorities elsewhere, thus explaining the higher mutation rate. It is simply a cost-cutting measure, rather than a direct response to stress.

Meanwhile, the second-order selection hypothesis claims that this response has been selected for because of its ability to control mutation rate. By turning on hypermutation in times of stress, then strategically returning mutation rates to normal levels once the stress is gone, bacterial populations can more quickly adapt to stresses and then reduce the long-term burden to their average fitness. Here, we will attempt to model such a SIM subpopulation's dynamics and show that this strategy is evolutionarily favored under the right conditions, and even more so than the constitutive mutator strategy.

In our model, the starving bacteria suffer starvation-induced stress and when they mutate the beneficial gene from b^- to b^+, they counter this stress. By mutating this gene, they gain its benefit and are no longer suffering from starvation stress. This results in a reversion to normal mutation rates in these individuals which we represent simply with an altered mutation matrix. Simply put, all outgoing mutations from states with b^+ are no longer multiplied by m for the mutator genotype. We will now investigate what effect this has on the behavior of these SIM mutators.

First we look at some qualitative aspects of the progression of the SIM mutator genotype frequency. The results are shown below. We observe several key features. Notably, the fraction of times that the mutator fixes is higher in our SIM scenarios. Comparing figures 15 and 5 we can see a dramatic increase in the fraction of runs that reach fixation. This increase can be observed at higher and lower m as well. As depicted in Fig.16 versus 10, we see a massive increase in the fixation frequencies for $m = 1000$ at $b = 0.1$. We also observe a similar effect for $m = 10$ as seen in Fig.17. Clearly SIM is even more advantageous than the constitutive mutator strategy within the same situational parameters. This is due to formation of beneficial mutants which then revert their mutation frequencies which prevents the increased accumulation of deleterious mutations afterwards. This results in only beneficial alleles being acquired by certain mutator individuals. They are free to acquire the beneficial allele initially without worrying about deleterious alleles in the longer run.We can observe this increase in fixation frequency across b values in Fig.18.

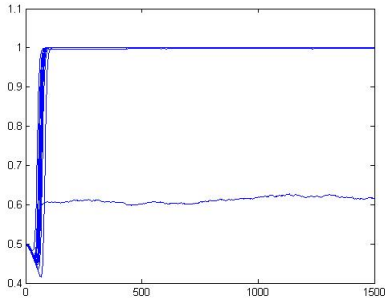

FIG. 13: Mutator frequency progression under SIM for m=100, b=0.5

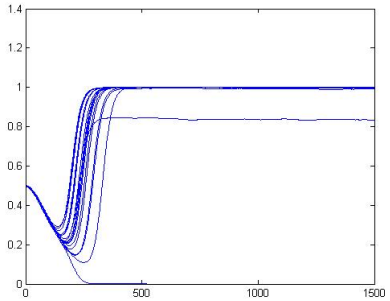

FIG. 14: Mutator frequency progression under SIM for m=100, b=0.1

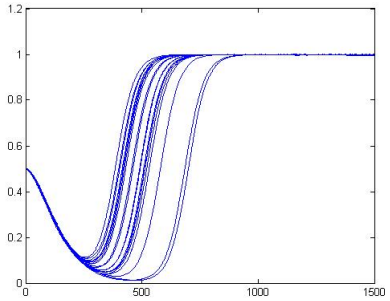

FIG. 15: Mutator frequency progression under SIM for m=100, b=0.05

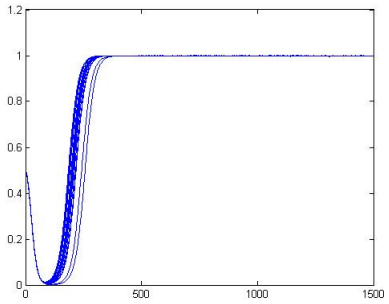

FIG. 16: Mutator frequency progression under SIM for m=1000, b=0.1

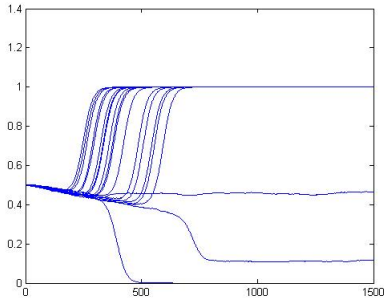

FIG. 17: Mutator frequency progression under SIM for m=10, b=0.1

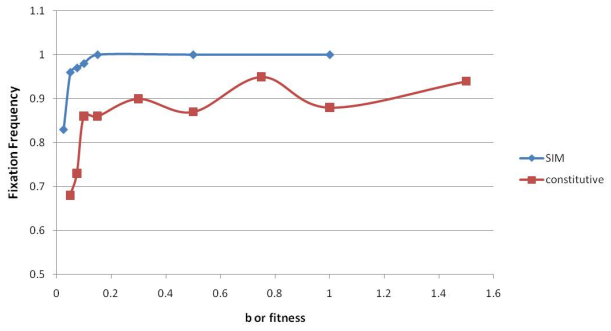

FIG. 18: SIM versus constitutive mutator fixation frequencies

V. CONCLUSIONS

In this paper, we attempt to model and simulate the behavior of both a constitutive mutator and a SIM mu-

tator strain subpopulation. We initially begin with equal amounts of the mutator strain and a non-mutator. The simulations show that both types of mutators are evolutionarily favored in times of environmental stress under certain conditions. Generally, the conditions that favor these mutators are high benefit from beneficial alleles and comparatively low penalties from deleterious alleles. In addition, we show that the mutator strength achieves an optimal value for fixation frequency that is neither too low nor too high.

The course of mutator fixation displays an initial dip in frequency due to the faster accumulation of negative mutations. This is a stochastic phase where the mutator genotype may or may not make a beneficial mutation in time to overcome this initial burden and rescue the mutator genotype. The mutator genotype then hitchikes with the succesful beneficial allele into fixation. Otherwise, it will be overcome by the negative mutations and be lost.

Our SIM mutators meanwhile display even more advantage then the constitutive mutators under similar conditions. The evidence does suggest that SIM can be a evolutionarily advantageous strategy and supports the second-order selection hypothesis. However, it does not disprove the pleiotropic hypothesis. Indeed both are still equally likely and could act in conjunction to select for SIM.

Of course, our model clearly fails to capture the true complexity of the bacterial genome. There are definitely many more deleterious and beneficial mutations than our model accounts for. In addition, environmental stresses are rarely easily modeled as a constant fitness effect. In short, much of our model is greatly simplified from the vast complexity of natural biological systems. However, it does capture the key components and these complexities can be worked into our simulation by simply adjusting parameters. In addition, the effects of varying N, the population size, remain yet to be explored. Intuitively, a larger population should increase mutator fixation frequency given the larger expected number of mutants with the beneficial allele.

The idea of hitchhiking mutator alleles is only applicable to asexual populations. With recombination, the mutator allele would be uncoupled from the beneficial allele too frequently. However, this does not mean that studying mutators is limited to bacteria. Loeb et al. have shown that tumorogenesis, which requires multiple "hits" to the genome, cannot progress at reasonable timescales without some increased level of mutation. They propose that early in tumor progression, the cancer cells acquire a mutator genotype and consequently are able to mutate tumor-supressor genes quickly.[9] This also believed to underlie development of drug resistance in cancer.

Finally, understanding SIM and mutator genotypes in bacteria presents an exciting possibility that bacterial evolution may be even faster than once thought. These bacteria are no longer constrained by an inherent spontaneous mutation rate and can dynamically control their mutation rates. This is a fascinating topic that shows that the mutation rate, originally one of the fixed parameters in evolutionary dynamics, is itself a target of natural selection.

[1] Drake J.W. A constant rate of spontaneous mutation in DNA-based microbes. (1991) PNAS **88**(16): 7160-4.

[2] Mao E.F., Lane L., Lee J., Miller J.H. Proliferation of mutators in a cell population. (1997) J. Bacteriol. **179**(2): 417-22.

[3] Gibson T.C., Scheppe M.L., Cox E.C. Fitness of an Escherichia coli mutator gene. (1970) Science **169**(946): 686-8.

[4] LeClerc J.E., Li B., Payne W.L., Cebula T.A., High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. (1996). Science **274**: 1208-1211.

[5] Tenaillon O., Toupance B., Nagard H.L., Taddei F., Godelle B., Mutators, Population Size, Adaptive Landscape and the Adaptation of Asexual Populations of Bacteria. (1999) Genetics **152**: 485-493.

[6] Kibota T.T., Lynch M., Estimate of the genomic mutation rate deleterious to overall fitness in E. coli. (1996) Nature **381**(6584): 694-6.

[7] Bjedov I., Tenaillon O., Gerard B., Souza V., Denamur E., Radman M., Taddei F., Matic I. Stress-induced mutagenesis in bacteria. (2003) Science **300**(5624): 1404-9.

[8] Layton J.C., Foster P.L. Error-prone DNA polymerase IV is controlled by the stress-response sigma factor, RpoS, in *Escherichia coli*. (2003) Mol. Microbiol. **50**: 549-561.

[9] Loeb L.A. Mutator phenotype may be required for multistage carcinogenesis. (1991) Cancer Res. **51**: 3075-3079.