Stress-induced mutagenesis, adaptability and adaptedness

Yoav Ram and Lilach Hadany

Corresponding author - [lilach.hadany@gmail.com](mailto:lilach.hadany@gmail.com)

Dept. of Molecular Biology and Ecology of Plants, Tel Aviv University

Tel-Aviv 69978, Israel

Last update: June 23, 2013

# Introduction

Stress-induced mutagenesis (SIM), the phenomenon in which the mutation rate of stressed or maladapted individuals is increased, has been demonstrated in numerous species, both prokaryote and eukaryote (Galhardo, Hastings, and Rosenberg 2007; Sharp and Agrawal 2012; MacLean, Torres-Barceló, and Moxon 2013). More specifically, various stress responses have been shown to regulate increases in the mutation rate by shifting the cell to use error-prone DNA polymerases (Ponder, Fonville, and Rosenberg 2005) and by inhibiting the mismatch repair system (Debora et al. 2010).

There is both experimental, clinical and theoretical evidence that increased mutation rates have an effect on adaptation and that during adaptive evolution constitutive mutators - alleles that constitutively increase the mutation rate – can rise in frequency because of this effect (Sniegowski et al. 2000). However, several works demonstrated that during evolution in a constant environments these mutators suffer from the accumulation of deleterious mutations and are purged from the environment, leading to the idiom of "the rise and fall of the mutator allele" (Taddei et al. 1997; Denamur and Matic 2006). Therefore, it seems that the mutation rate must balance two traits – adaptability and adaptedness, was coined by Leigh (1970).

SIM is considered by many to have a meaningful impact on *adaptability* or *evolvability* - the capacity of individuals and populations to evolve and adapt (Tenaillon, Denamur, and Matic 2004; Rosenberg et al. 2012), but there is currently no theoretical treatment of this effect. The effect of SIM on adaptedness – the capacity of populations to remain adapted to current environmental conditions – did receive theoretical treatment by Agrawal (2002), who showed numerically that in asexual populations SIM doesn't affect the population mean fitness in constant environments. In a more recent work, we have showed that this if rare beneficial mutations can occur than SIM is slightly advantageous in a constant environment (Ram and Hadany 2012).

Here, we analyze a population genetic model of an asexual population. We derive analytical expressions that demonstrate that SIM increases both the mean fitness and the adaptation rate of asexual populations. We compare SIM to CM and non-mutators (NM) to show that SIM is a more efficient mutational strategy, allowing individuals and populations to increase their *adaptability* without compromising their *adaptedness*.

# Model

We consider a population of *N* haploid asexual individuals. The number of new deleterious mutations at replication is Poisson distributed with an average of *U* mutations per genome. The effect of deleterious mutations on fitness are multiplicative (i.e., independent), such that the fitness of an individual with *x* deleterious mutations is *(1-s)x*, where *s* is the selection coefficient, and we assume that selection is stronger than mutation - *s>U*. We consider three mutational strategies: normal mutagenesis (NM), where there is no increase in mutation rates; constitutive mutagenesis (CM) where all individuals increase their mutation rate by *τ*, the mutation rate fold increase; and stress-induced mutagenesis (SIM) where only stressed or maladapted individuals increase their mutation rate by *τ*.

We analyzed two models, one of evolution toward a mutation-selection balance (MSB) in a constant environment, the other of adaptive evolution of a complex trait in a rugged fitness landscape. We developed analytic approximations for the adaptation rate and mean fitness of populations and used stochastic simulations to verify our approximations.

## Mutation-selection balance

Denote the frequency, fitness and deleterious mutation rate of individuals with *x* deleterious mutations by , and , and the population mean fitness by . Let us assume that a small fraction of the mutations are back-mutations or beneficial mutations which reduce the number of deleterious mutations carried by an individual. The frequency of individuals with *x* deleterious mutations in the next generation can therefore be described by (denoting ):

.

This can be written as a matrix equation by multiplying the frequencies vector f by the mutation-selection matrix *M*:

The MSB distribution of *x* *f\**[[1]](#footnote-1) fulfills ():

*M* is a positive matrix, and therefore by the *Perron-Frobenius Theorem* (Otto and Day 2007, p. 709) is the largest eigenvalue of *M* and *f\** is its unique non-negative eigenvector with .

Without beneficial mutations, *δ=1* and *β=0*, the above equation simplifies to:

.

So *M* is a triangle matrix and its largest eigenvalue is the largest main diagonal element: . If and then the frequencies vector has been shown to be , that is, the number of deleterious mutations per individual is Poisson distributed with an average *U/s* (Haigh 1978).

With beneficial mutations this eigenvalue problem is harder to solve analytically. By neglecting elements outside the main three diagonals we have shown before (Ram and Hadany 2012) that:

.

However, this framework allows to easily calculate the population mean fitness numerically for finite *n*-by-*n* matrices by defining *n* such that – see Python code in Supplementary File X.

## Adaptation on a rugged fitness landscape

Consider a population in which the wildtype genotype is *ab* and its fitness is 1. Mutations at these loci change *a* to *A* and *b* to *B* at reproduction with probability *µ* (without back-mutations). As before, new deleterious mutations occur with rate *U*.

After the population has reached a MSB it undergoes adaptive evolution to an environmental change which altered that fitness of the *AB* genotype from *(1-s)2*  *1+sH* (*H* is the double mutant relative advantage), making it the optimal genotype (Figure 1). We study four different mutational strategies: NM, CM, SIM-I with which all genotypes except *ab* and *AB* hypermutate, and SIM-II with which all genotypes except *AB* hypermutate.

We developed two models. The first is analytic (Figure 1a), in which we assume that: (i) genotypes with deleterious alleles (except *Ab* and *aB*) do not contribute to the adaptation process and (ii) that the number of deleterious mutations per individual at the MSB is Poisson distributed. The second model is a stochastic Wright-Fisher simulation with selection, mutation and drift (Figure 1b), in which we: (i) let individuals with accumulated deleterious mutations contribute to adaptation, and (ii) let the MSB evolve from a mutation-free population.



**Figure 1 – Adaptive landscape illustration.** Nodes represent genotypes: the alleles *a* or *A* and *b* or *B*, and in panel b also the number of deleterious alleles across the genome. Mutagenesis is induced in stressed genotypes (fitness *<1* or *<1+sH*, depending on the scenario), indicated by ellipses, while adapted genotypes (fitness *>1* or *>1+sH*, depending on the scenario) do not hypermutate. Solid arrows represent site-specific mutations at the *a/A* and *b/B* loci, which occur with probability *µ*. Dashed arrows represent deleterious mutations across the genome which are Poisson distributed with an average *U*. Node colour represents fitness (see colourbar), from pale brown for the fittest genotype (*1+sH*, where *s=0.05* is the selection coefficient and *H=2* is the double mutant advantage) to dark brown for genotypes with accumulated deleterious mutations (*(1-s)x*, where *x* is the number of deleterious mutations). (a) In the analytic model genotypes with deleterious mutations are considered "evolutionary dead-ends" (RIP) and do not contribute to adaptation. (b) In the stochastic model individuals can accumulate up to 25 deleterious mutations (the figure only shows as much as three). Multiple mutations can occur concurrently but are not shown for simplicity of the illustration.

### Constraints on the parameter space

At the MSB, the frequency of wildtype (*ab*) individuals is , the frequency of single mutants (*Ab* and *aB* combined) is and the frequency of double mutants (*AB*) is . Therefore, there are several constraints on the parameter range:

1. The above MSB approximations are only valid when *U/s<1* or *U<s*.
2. If *N*(*µ/s)2>1* then there are double mutants in the population at the MSB and therefore adaptation will be rapid and will not require new mutations.
3. If *Nµ/s<1* then there are no single mutants at the MSB and double mutants must be generated by a double mutation in a wildtype individual. Therefore, increasing the mutation rate of individuals with fitness below 1 will have a much smaller effect than if single mutants were abundant.
4. If we assume that individuals that accumulated deleterious mutations are "evolutionary dead-ends" (figure 1a) and cannot be the origin of adaptation, then the fraction of such individuals must be small - *U/s<1*. This is equal to the first constraint.

These constraints are summarized by:

|  |  |
| --- | --- |
|  | (1) |
|  | (2) |

Table 1 summarizes the model parameters with estimated values for *Escherichia coli*. Taking the conservative estimations, the population size *N* must be between105and107, but the constraint can also be met for other combinations of the parameter values.

Table 1 – Model parameters and estimated values for bacteria

|  |  |  |  |
| --- | --- | --- | --- |
| Symbol | Name | Estimate | References |
| *s* | Selection coefficient | 0.001-0.03 | (Kibota and Lynch 1996; Gordo, Perfeito, and Sousa 2011) |
| *H* | Double mutant advantage | 1-10 | (Gordo, Perfeito, and Sousa 2011) |
| *U* | Genomic deleterious mutation rate | 0.0004-0.003 | (Drake et al. 1998; Wielgoss et al. 2011) |
| *µ* | Site-specific beneficial mutation rate | U/5000 | (Gordo, Perfeito, and Sousa 2011) |
| *τ* | Fold-increase in mutation rate | 1-100 | (Bjedov et al. 2003; Hall and Henderson-Begg 2006) |
| *N* | Population size | 104-1010 |  |

### Appearance of a double mutant

Because there are no double mutants (*AB*) at the time of the environmental change, double mutants can appear either via a double mutation in a wildtype individual *ab*, or via a single mutation in a single mutant *Ab* or *aB* (Figure 1a). We assume that at the MSB the number of deleterious mutations per individual follows a Poisson distribution (Haigh 1978). Therefore, the frequencies of mutation-free wildtype *ab* and single mutants *aB* and *Ab* are and . The probability *q* that a random individual in the next generation is a double mutant, given there are no double mutants and neglecting individuals with deleterious mutations is:

|  |  |
| --- | --- |
|  | (3) |

With SIM-I the mutation rate of single mutants is increased *τ*-fold and the appearance probability is:

|  |  |
| --- | --- |
|  | (4) |

Note that stress-induction increases the transition from single mutants to other types, but does not significantly change the MSB frequency of single mutants, because this frequency is mainly determined by the mutation rate of the wildtype which does not hypermutate.

The above expressions can be simplified by using first-order approximations. Starting with Eq. (3) for populations without SIM-I:

The last step assumes that *2s* is much larger than *s2* and *sU* is much larger than *2µ.* Rearranging the last expression gives us

|  |  |
| --- | --- |
|  | (5) |

For a population with SIM-I the first-order approximation is based on the full expression in Eq. (4):

The last approximation assumes that *Us* is smaller than *U* and that *τU* is much larger than *µ/s.* Now,

The last approximation assumed that *2τ>s* and 2*τ2>1*, because *τ>1* and probably even *τ≥10*. Rearranging the last expression gives us the first order approximation for populations with SIM:

|  |  |
| --- | --- |
|  | (6) |

Note that by setting and because , is consistent with .

### Fixation probability with stress-induced mutagenesis

Assuming that the advantage of the double mutant is considerable (for example, *H>1)* and that the population size is large (constraint (2) ensures that), a double mutant has two possible fates after its appearance: fixation or extinction. Following Eshel (1981), the fixation probability ρ of the double mutant is:

|  |  |
| --- | --- |
|  | (7) |

where *α* is the fitness of the double mutant relative to the population mean fitness and assuming that fitness is measured by the number of progeny which is Poisson distributed:

|  |  |
| --- | --- |
|  | (8) |

Here, we only take the fraction of progeny that do not have deleterious mutations -. This factor cannot be ignored because there is variation in mutation rates in the population.

At this stage, double mutants are still very rate, so we can use the population mean fitness at the MSB. Without stress-induced mutagenesis, this evaluates to (see ‎3.1). We find that:

|  |  |
| --- | --- |
|  | (9) |

Assuming *sH* is small we can simplify this to:

|  |  |
| --- | --- |
|  | (10) |

which is a classic result of population genetics theory (Eshel 1981).

The fixation probability with SIM-I equals that of NM and CM because the mutation rate of the wildtype *ab* equals that of the double mutant *AB*. However, with SIM-II the mutation rate of *ab* is higher than that of *AB* and therefore *AB* has an additional advantage due to the accumulation of deleterious mutations in the stressed population (we can assume the population reached a MSB after the environmental change because the convergence to MSB is much quicker then adaptation):

### Adaptation rate

From the probability *q* that a random newborn is a double mutant we can derive the probability that some double mutants appear in the next generation: . The constraint (2) guarantees that *Nq* is very small and therefore this probability can be approximated by *Nq*.

Once a double mutant appears it has a probability *ρ* to go to fixation.

The time for adaptation *T* can be approximated by the waiting time for a double mutant who will go to fixation, *Tw*. This is true as long as fixation is much faster than mutation (guaranteed by *µ2<2* which is a weaker constraint than that given by Eq. (1)). *Tw* follows a geometric distribution with probability *Nqρ* and therefore the expected time for adaptation can be approximated by:

|  |  |
| --- | --- |
|  | (11) |

The adaptation rate is the inverse of the expected adaptation time

.

# Results

## Mutation-selection balance

Without beneficial mutations and with a mutation rate constant in time and uniform across the population, the mean fitness equals (Kimura and Maruyama 1966). Constitutive mutagenesis causes the population mean fitness to exponential decay as a function of the mutation rate fold-increase *τ*. Stress-induced mutagenesis, as was numerically shown by Agrawal (Agrawal 2002), does not change the population mean fitness with respect to NM, because the fittest individuals, with fitness *1*, also have the lowest mutation rate, *U*, and therefore the largest eigenvalue of the mutation-selection matrix is .

With beneficial mutations, this is still a good approximation, but the actual value of the population mean fitness is slightly higher than . As we have shown before (Ram and Hadany 2012), CM decreases the population mean fitness with respect to NM, but SIM can increase the population mean fitness with respect to NM. A sufficient condition is that the mutation rate of individuals with below average fitness is increased. Because we assume that *,* then , so setting *π=1* ensures that the condition is met and therefore SIM increases the population mean fitness. We explore the magnitude of this increase for specific parameter ranges in Fig. X.

## Adaptation on a rugged fitness landscape

We define two adaptive evolution scenarios: in scenario I, the appearance of a new carbon source which can only be utilized by the double mutant *AB* appears, making the *AB* genotype the optimal one without affecting the fitness values of other genotypes, and therefore in this scenario SIM does not induce hypermutation in neither *AB* nor *ab*. In the second scenario, an antibiotic drug renders all genotypes less fit, and the double mutant *AB*, which confers resistance to the drug, is the optimal genotype. In this scenario SIM induces hypermutation in all genotypes except for *AB*. Therefore, when analyzing SIM we differ between SIM-I and SIM-II.

The adaptation rate *ν* is the inverse of the expected adaptation time which can be approximated, similar to, by *1/Nρq*:

The right-hand side approximations are for small values of *τU*.

There are several interpretations we can make from these expressions. First, adaptation with CM is faster than with NM because of faster appearance of double mutants and an unchanged fixation probability. Second, adaptation with SIM-I is also faster than with NM, but not as fast as with CM because only single mutants *aB* and *Ab* hypermutate. Third, adaptation with SIM-II is the fastest, because the appearance of double mutants is the same as with CM but their fixation is more probable because of the difference in mutation rate between double mutants and the rest of the population. Note that this advantage increases linearly with *τ*.

These approximations, and therefore these results, all depend on the constraints above, of which the most important is *τU<s*. Deleterious mutation rates in microbes are generally on the order of 10-4-10-2 and selection coefficients are between 10-1 and 10-2 (see Table 1), so the limit on *τ* can be anything between 1 and 1,000. More generally, we can expect that is selection is strong and mutation rates are low, then CM and SIM will have a big advantage over NM.

## Adaptability and adaptedness

1. A star *\** denotes equilibrium values [↑](#footnote-ref-1)