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 27, 2013

# Introduction

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

Stress-induced mutagenesis (SIM), the increase of the mutation rate in stressed or maladapted individuals, 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). Various stress responses regulate the mutation rate by shifting cells to error-prone DNA polymerases (Ponder, Fonville, and Rosenberg 2005) and by inhibiting the mismatch repair system (Debora et al. 2010).

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

Here, we analyzed a population genetic model of an asexual population. We develop general analytical expressions that demonstrate that SIM increases both the mean fitness and the adaptation rate. We compare SIM to constitutive mutagenesis (CM) and normal mutagenesis (NM) to show that SIM is a more efficient mutational strategy and that is allows 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 *ω=(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 analyze two models, one of evolution towards a mutation-selection balance (MSB) in a constant environment, the other of adaptive evolution of a complex trait in a rugged fitness landscape. We develop analytic approximations for the mean fitness and adaptation rate 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 . Assume that a small fraction of all mutations are 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 be described by (denoting ):

.

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

At the MSB, *f\** fulfills (a star \* denotes equilibrium quantities):

Because *M* is a positive matrix and 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 (constant uniform mutation rate) then the frequencies vector is , that is, the number of deleterious mutations per individual is Poisson distributed with average *U/s* (Haigh 1978).

With beneficial mutations this eigenvalue problem is harder to solve analytically. By neglecting elements outside the main three diagonals of *M* 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 the *mean\_fitness* function in the 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:

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

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 – 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:

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

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

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

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

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

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:

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

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:

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

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:

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

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:

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

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

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

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:

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

The adaptation rate is the inverse of the expected adaptation time

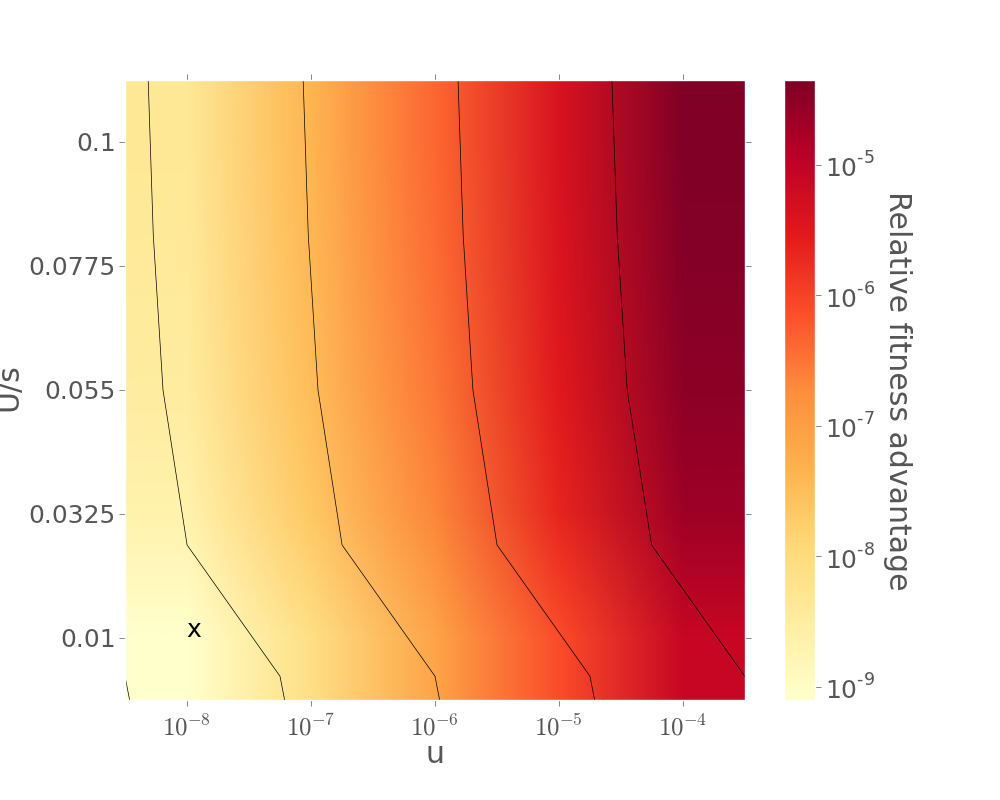
.

# Results

## Mutation-selection balance

Without beneficial mutations ( and with constitutive mutagenesis (CM), the mean fitness equals (Kimura and Maruyama 1966). Therefore, CM 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 normal mutagenesis (NM), because the least loaded individuals (*x*=0), with fitness *ω0*=1, also have the lowest mutation rate, *U*, and therefore the population mean fitness (the largest eigenvalue of the mutation-selection matrix *M*) is .

With beneficial mutations (), this is still a good approximation (because ), 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 if SIM increases the mutation rate in individuals with at least one deleterious mutation then it increases the population mean fitness. We explore the magnitude of this increase for specific parameter ranges in Fig. X.

**Figure 2 – Mean fitness with stress-induced mutagenesis.** The figure shows the relative fitness advantage of SIM with *τ*=10 in comparison to NM. The x-axis is the rate of beneficial mutations per genome per generation; the y-axis is the ratio of the deleterious mutation rate to the selection coefficient (which is also the average number of deleterious mutations per individual at the MSB). The x marks the parameter set *U*=0.0004, *u*=*U*/5000, *s*=0.05 in which the fitness advantage of SIM is ~5⋅10-9.

## 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*:

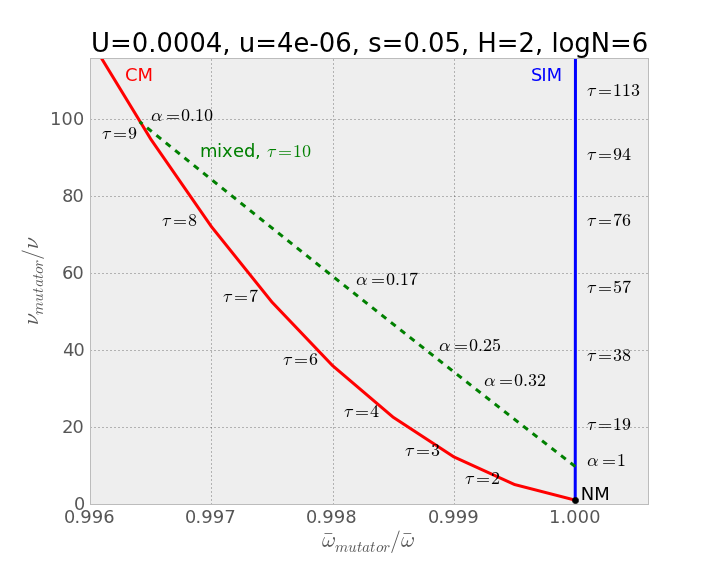
The right-hand side approximations are for small values of *τU*.

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

Figure X shows the mean fitness and adaptation rate of different mutational strategies. Every rate of adaptation *ν* (horizontal line through the figure) can be achieved using both CM and SIM. The highest mean fitness will always be attained with SIM (which even has a slight advantage over NM which cannot be seen in this figure). If the mutation rate fold increase *τ* required by SIM is too high (*i.e.*, *τU>s*) the same adaptation rate can be achieved via a mixed strategy (dashed green line) which combines SIM and CM. For example, a 50-fold increase in adaptation rate can be achieved with CM with *τ=7*, SIM with *τ=50* or a mixed strategy with *τ1=2* and *τ2=10*. The mutational burden will decrease the mean fitness 0.9972-fold with CM and 0.996-fold with the mixed strategy, but will not have an effect with SIM.



**Figure 3 - Adaptedness and adaptability.** The figure shows the population mean fitness and adaptation rate of a mutator compared to non-mutators (NM) on the x-axis and y-axis, respectively. Constitutive mutators (CM; in red) increase the mutation rate of all individuals; Stress-induced mutators (SIM; in blue) increase the mutation rate of stressed or maladapted individuals; Mixed strategies (in dashed green) increase the mutation rate of *α* of the individuals when stressed and of *1-α* of the individuals constitutively. *τ* is the mutation rate fold increase in hypermutating individuals. SIM breaks the adaptability-adaptedness trade-off of CM, increasing the adaptation rate without reducing the population mean fitness. Parameter values: population size *N=106*, deleterious mutation rate *U=­0.0004*, site-specific mutation rate *µ=U/5000*, selection coefficient *s=0.05*, double mutant advantage *H=2*. SIM-I/SIM-II

# Discussion

We developed theoretical models to analyze the relative effect of stress-induced mutagenesis on both adaptability – the capacity of populations to adapt to new conditions – and on adaptedness – the ability of populations to remain adapted under stable conditions (Leigh 1970). We showed that SIM breaks the trade-off between adaptability and adaptedness, allowing rapid adaptation without compromising the population mean fitness.

One of the features included in our model was that of a mixed strategy – a mutational strategy that combines CM and SIM. There are two examples of such as strategy. First, if individuals' sensing of stress in not perfect (which is probably the case in any biological realistic scenario) then one can expect that there would be errors in the induction of mutagenesis: induction of mutagenesis without stress, as well as failure to induce mutagenesis under stress. Second, a mutator allele can increase the mutation rate without stress and increase it even more under stress (Torres-Barceló et al. 2013).

The traits which were analyzed here – mean fitness and adaptation rate – are population-level traits. Even though SIM creates the most efficient balance between these traits, it does not mean necessarily that it will evolve, because individual-level selection can sometimes work in different directions than population-level selection. Moreover, even if SIM does evolve, it is not clear that is the result of selection on adaptability and adaptedness (2nd order selection (Tenaillon et al. 2001)) or a result of other factors, such as pleiotropic effects of mutator alleles (Torres-Barceló et al. 2013), the cost of DNA replication fidelity (Dawson 1998), and the effect of drift on fidelity of rarely expressed enzymes, aka the "drift barrier hypothesis" (Sung et al. 2012). However, in a previous work we demonstrated that indirect selection on generation of beneficial mutations can lead to the evolution of stress-induced mutagenesis, at least in asexual population (Ram and Hadany 2012).

Complex traits, coded by multiple genes, present an open evolutionary question, first described by Sewall Wright in 1931 (Wright 1931): if different alleles are separately deleterious but jointly advantageous, how can a population evolve from one co-adapted gene complex to a better one? Wright suggested the "shifting-balance theory of evolution" (Wright 1988). His solution is valid (Crow, Engels, and Denniston 1990; Wade and Goodnight 1991; Peck, Ellner, and Gould 2000) but is possibly limited to specific parameter ranges (Moore and Tonsor 1994; Gavrilets 1996; Coyne, Barton, and Turelli 2000; Whitlock and Phillips 2000). As a result, other mechanisms to solve the problem were proposed (Whitlock 1995; Whitlock 1997; Hadany 2003; Hadany and Beker 2003; Weissman et al. 2009) [24–28]. In this work we analyzed the effect of SIM on complex adaptation (Figure X). Our results suggest that SIM is another mechanism that can help resolve this century old problem.

Our work provides a formal theoretical basis to the conjecture that SIM facilitates adaptation and increases the evolvability of populations. The next step would be to experimentally verify our results. This can be done, for example, with *E. coli*, in which one can interfere with the regulation of hypermutation by stress (Cirz and Romesberg 2007). If an experimental population evolves under conditions such as those described in our model, it will be possible to measure the adaptation time with and without SIM and compare it to our analytical approximations. In addition, it will be possible to measure the mean fitness of populations at a mutation-selection balance to determine the relative advantage and disadvantage of the different mutational strategies.

## Conclusions

Stress-induced mutagenesis has been implicated as a driver of adaptive evolution for several decades (Cairns, Overbaugh, and Miller 1988). We provide the first theoretical treatment of this concept. We show that stress-induced mutagenesis indeed increases the rate of adaptation. Furthermore, we show that in contrast to constitutive mutagenesis it does not jeopardize the fitness of populations under stable conditions. Because mutation is such a fundamental factor in every biological process, these results have an important implication on many fields in the life sciences, including epidemiology, ecology and evolutionary biology