Stress-induced mutagenesis and the evolution of complex adaptations

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: 12 June 2013

# Summary (max 200)

Regulation of mutagenesis by stress responses has been evidenced in numerous species, both prokaryote and eukaryote, and has been suggested to contribute to adaptation and evolvability. However, the theoretical basis for this contribution is lacking.

Here we analyze a population genetic model of a rugged fitness landscape. We derive analytical expressions that show that stress-induced mutagenesis increases the adaptation rate and present the results of stochastic simulations that validate our analysis. Our results suggest that stress-induced mutagenesis indeed facilitates adaptation and promotes evolvability.

Furthermore, we show that stress-induced mutagenesis helps to resolve the problem of adaptive peak shifts, first described by Sewall Wright in 1931, by accelerating the evolution of complex traits.

# Keywords (3-6)

population genetics; mathematical modeling; evolvability; stress-induced variation; adaptive peak shifts; fitness landscape

# Short title

Stress-induced mutation & complex adaptations

# Introduction

Stress-induced mutagenesis (SIM), the phenomenon in which stressed or maladapted individuals increase their mutation rate, has been demonstrated in numerous species, both prokaryote and eukaryote [1–3]. SIM is considered by many to have a meaningful impact on *evolvability* - the capacity of individuals and populations to adapt [4,5]. In a previous work we showed that SIM is favored by natural selection and that it increases the mean fitness of populations due to the increased generation of beneficial mutations in maladapted individuals [6]. Here, we focus instead on the effect of SIM on the evolution of complex adaptations.

Here, we analyze a population genetic model of an asexual population undergoing an adaptive peak shift. We derive analytical expressions that demonstrate that SIM increases the population adaptation rate and show the results of stochastic simulations that validate our analytic expressions.

# Material and methods

## Analytical model

We consider two loci with alleles *A/a* and *B/b* and a population at a mutation-selection balance (MSB). The genotype *ab* is the wildtype with a fitness value of *1*, single mutants (*Ab* and *aB*) have a fitness value of *1-s*, with *s* as the selection coefficient, and double mutants (*AB)* have a fitness value of *(1-s)2*.

Mutation from *a* to *A* and from *b* to *B* occurs with a probability *µ* at reproduction (we neglect back-mutation) - *µ* is therefore the site-specific beneficial mutation rate. In addition, the number of new deleterious mutations that occur across the genome at reproduction follows a Poisson distribution with an average *U* - the genomic deleterious mutation rate.

Individuals with SIM and fitness lower than 1 increase both mutation rates *τ*-fold.

The model assumptions impose two general constraints on the parameters:

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

The first constrain is required for the MSB. The second constraint requires that the population is large enough for single mutants to have an impact on the adaptation process, but small enough so that double mutants are not present at the MSB.

Table 1 – Model parameters and estimated values for bacteria

|  |  |  |  |
| --- | --- | --- | --- |
| Symbol | Name | Estimate | References |
| *s* | Selection coefficient | 0.001-0.03 | [7,8] |
| *H* | Double mutant advantage | 1-10 | [8] |
| *U* | Genomic deleterious mutation rate | 0.0004-0.003 | [9,10] |
| *µ* | Site-specific beneficial mutation rate | U/5000 | [8] |
| *τ* | Fold-increase in mutation rate | 1-100 | [11,12] |
| *N* | Population size | 104-1010 |  |

Table 1 summarizes the model parameters and estimated values for *Escherischia 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. We assume both of these constraints hold thourghout the Results section.

We are interested in the capacity of the population to adapt to an environmental change after which the double mutant *AB* is the optimal genotype. Figure 1a presents an illustration of the analytical model.

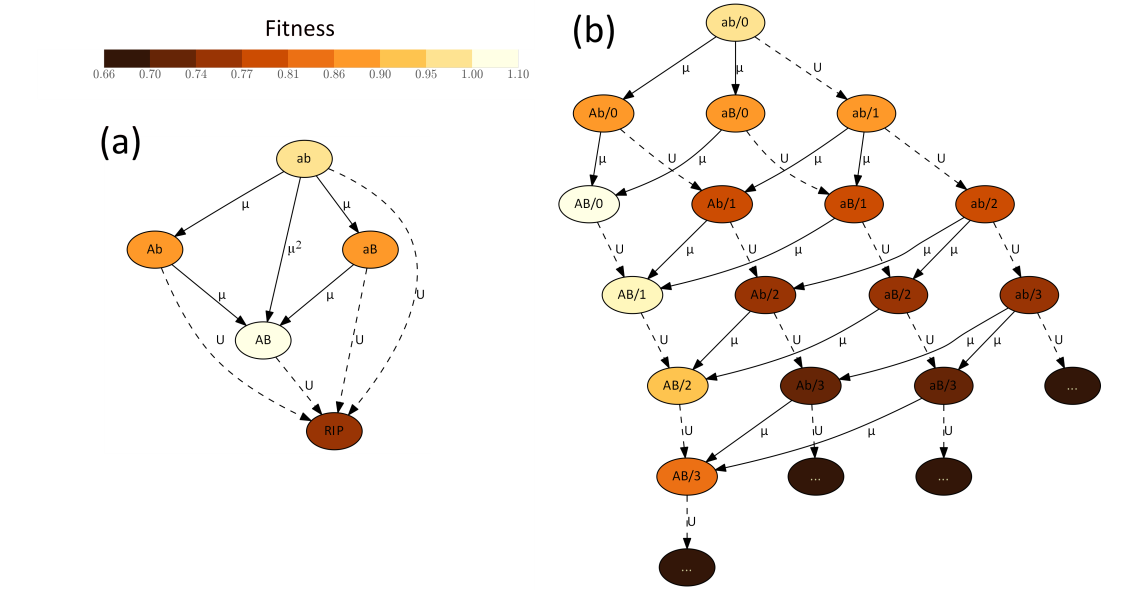


Figure 1 – Adaptive landscape illustration. Each node represent a specific genotype. Node labels specify the alleles at the *A/*a and *B/*b loci and the number of additional deleterious alleles after the forward slash (only in (b) and only as much as 3 deleterious mutations are shown to keep the figure simple). Arrows define the direction of mutation (solid for site-specific beneficial mutations with rate µ, dashed for genomic deleterious mutations with rate *U*). Node colour indicates the fitness of a genotype (see colourbar), from pale brown for optimal fitness (*1+sH*, with *s* as the selection coefficient and *H* as the relative advantage of *AB*) to dark brown for lower fitness (*(1-s)m* with *m* as the number of deleterious mutations). Parameters used: *s=0.01*, *H=2*. (a) In the analytical model individuals with deleterious mutations are considered evolutionary dead-ends (*RIP*) anddo not contribute to adaptation.(b) In the stochastic model we make no such assumption. In this panel only single mutations are shown for simplicity.

## Stochastic model

To validate our analytic approximations, we developed a Wright-Fisher simulation with mutation, selection and random genetic drift. The main differences between the analytical and the stochastic models are: (i) The simulations incorporate genetic drift by randomly sampling each generation from the previous one using a multinomial distribution. (ii) Individuals with deleterious mutations are not "evolutionary dead-ends" - individuals are allowed to accumulate up to 25 deleterious mutations (figure 1b). (iii) Simulations start with an *ab* mutation-free population. After reaching MSB, the environment is changed so that *AB* is advantageous. Therefore, the stochastic model assumes nothing about the distribution of deleterious mutations at the MSB.

# Results

The adaptation process is divided to two distinct processes: (i) the appearance of the double mutant *AB*; and (ii) its subsequent fixation. The appearance of the double mutant mainly depends on mutation and therefore takes much longer than fixation, which depends on the interplay between selection and genetic drift. In the following sections we provide analytic expressions for the appearance and fixation probabilities of the double mutant and the expected total waiting time for adaptation. In addition, we compare these analytic approximations to simulation results.

The complete derivations of all the expressions are given in the *Electronic Supporting Material*.

## Appearance of a double mutant

In a population at an MSB without double mutants *AB*, the probability *q* that a random newborn is a double mutant and neglecting individuals with deleterious mutations is (model parameters are summarized in Table 1):

|  |  |
| --- | --- |
|  |  |

However, if mutation is stress-induced, then the mutation rate of single mutants is increased *τ*-fold and the appearance probability is:

|  |  |
| --- | --- |
|  |  |

Deriving with respect to *τ*, we find that as long as , increasing *τ* increases the appearance probability of the double mutant.

## Fixation of a double mutation

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. We derive the fixation probability *ρ* of the double mutant following Eshel [13]. Assuming that fitness is measured by the number of progeny and is Poisson distributed we find:

which is a classic result in population genetics [13].

Because SIM can have an effect on the population mean fitness [6], the fixation probability with SIM is:

Comparing the righthand side of the last two equations, and because *s<1* and τ*<*1, this demonstrates that and that SIM increases the fixation probability of the double mutant. This is because maladapted individuals accumulate more deleterious mutations then adapted individuals, producing a wider fitness distribution and higher relative fitness for adapted individuals. However, the derivative of with respect to *τ* is , so increasing *τ* only has a small effect on .

## Adaptation rate

The adaptation time can be approximated by the number of generations before the appearance of a double mutant *AB* that goes to fixation. This number is geometrically distributed with probability *1/Nqρ*, where *N* is the population size and *q* and *ρ* are the probabilities shown in the previous sections. The expected adaptation time without and with SIM is therefore:

We compared these expressions to find an approximate sufficient condition for SIM to decrease the adaptation time:

That is, the mutation rate of maladapted individuals must be larger than that of well-adapted individuals but lower than one mutation per genome per generation. Two estimates of the genomic deleterious mutation rate *U* in *E. coli* are 0.003 and 0.0004, which sets the upper limit on *τ* to be between 333 and 2,500. This is higher than the values of *τ* found in the literature *- 1<τ<100* (see Table 1 for estimated parameter values).

## Simulation results

Figure 2 shows the effect os SIM on complex adaptation. The approximations (in dashed blue), fit the simulation results (in black circles) very well.

The starting point of all the lines is at *τ=1* which represents populations without SIM. Hence, both the approximations and the simulation results agree that SIM increases the adaptation rate, and that this effect increases with *τ*, thee fold increase of mutation rate in stressed individuals.



Figure 2 – Stress-induced mutagenesis reduces the adaptation time. Analytical approximations in dashed blue and simulation results in black circles. (a) Fold-change in waiting time for appearance of a double mutant with SIM compared to without. Note that a 10-fold increase in mutation rate results in a ~10-fold decrease in waiting time. (b) Fold-change in the number of double mutant appearances before fixation. The increase in fixation probability (Eq. 12) is to small to be seen on this scale. (c) The waiting time for adaptation – appearance and fixation of a double mutant – in generations, as a function of the fold-increase in mutation rate in stressed individuals. Note how a minor increase in mutation rate results in a considerable decreae in adaptation time. Parameters used: selection coefficient *s=0.05*, double mutant advantage *H=2*, genomic mutation rate *U=0.0004*, locus specific mutation rate *µ=U/5000*, population size *N=106*.

# Discussion

We studied the effect of SIM on complex adaptation. Our analysis, validated by simulatios, shows that SIM can considerably reduce adaptation time by significantly increasing the appearance and slightly increasing the fixation of double mutants.

An alternative mutational strategy is constitutive mutagenesis, in which individuals increase their mutation rate at all times. Although constitutive mutagenesis increases the appearance of double mutants, its long-term effect is detrimental [14] and it is easily outcompeted by SIM [6].

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 experimentaly verify our results. This can be done, for example, with *E. coli*, in which one can interfere with the regulation of hypermutation by stress [15]. If an experimental population evolves in under specific conditions such as described in our model, it will be possible to measure the adaptation time with and without SIM and compare it to our analytical approximations.

Complex traits, coded by multiple genes, present an open evolutionary question, first described by Sewall Wright in 1931 [16]: *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" [16]. His solution is valid [17–19] but is possibly limited to specific parameter ranges [20–23]. As a result, other mechanisms to solve the problem were provided [24–28]. This work presents SIM as another mechanism that can help resolve this century old problem.

# Acknowledgments

The work was supported in part by the Israel Science Foundation 840/08 (L.H.) and by Marie Curie reintegration grant 2007-224866 (L.H.).

# Data accessability

The simulations source code and the source code used to analyse the data and generate the figures are available as a *git* repository on GitHub at <https://github.com/yoavram/ruggedsim.git>. The codeis written in Python (<http://www.python.org>) using NumPy (<http://www.numpy.org>) and SciPy (<http://www.scipy.org>).

# References

1 Galhardo, R. S., Hastings, P. J. & Rosenberg, S. M. 2007 Mutation as a stress response and the regulation of evolvability. *Crit. Rev. Biochem. Mol. Biol.* **42**, 399–435. (doi:10.1080/10409230701648502)

2 Sharp, N. P. & Agrawal, A. F. 2012 Evidence for elevated mutation rates in low-quality genotypes. *Proc. Natl. Acad. Sci. US. A.* **109**, 6142–6. (doi:10.1073/pnas.1118918109)

3 MacLean, R. C., Torres-Barceló, C. & Moxon, R. 2013 Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nat. Rev. Genet.* **14**, 221–7. (doi:10.1038/nrg3415)

4 Tenaillon, O., Denamur, E. & Matic, I. 2004 Evolutionary significance of stress-induced mutagenesis in bacteria. *Trends. Microbiol.* **12**, 264–70. (doi:10.1016/j.tim.2004.04.002)

5 Rosenberg, S. M., Shee, C., Frisch, R. L. & Hastings, P. J. 2012 Stress-induced mutation via DNA breaks in Escherichia coli: A molecular mechanism with implications for evolution and medicine. *Bioessays.* , 1–8. (doi:10.1002/bies.201200050)

6 Ram, Y. & Hadany, L. 2012 The evolution of stress-induced hypermutation in asexual populations. *Evolution* **66**, 2315–28. (doi:10.1111/j.1558-5646.2012.01576.x)

7 Kibota, T. T. & Lynch, M. 1996 Estimate of the genomic mutation rate deleterious to overall fitness in E. coli. *Nature.* **381**, 694–6. (doi:10.1038/381694a0)

8 Gordo, I., Perfeito, L. & Sousa, A. 2011 Fitness effects of mutations in bacteria. *J. Mol. Microbiol. Biotechnol.* **21**, 20–35. (doi:10.1159/000332747)

9 Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. 1998 Rates of spontaneous mutation. *Genetics.* **148**, 1667–86.

10 Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R. E. & Schneider, D. 2011 Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With Escherichia coli. *G3: Genes, Genomes, Genetics* **1**, 183. (doi:10.1534/g3.111.000406)

11 Bjedov, I., Tenaillon, O., Gérard, B., Souza, V., Denamur, E., Radman, M., Taddei, F. & Matic, I. 2003 Stress-induced mutagenesis in bacteria. *Science (80-. ).* **300**, 1404–9. (doi:10.1126/science.1082240)

12 Hall, L. M. C. & Henderson-Begg, S. K. 2006 Hypermutable bacteria isolated from humans--a critical analysis. *Microbiology* **152**, 2505–14. (doi:10.1099/mic.0.29079-0)

13 Eshel, I. 1981 On the survival probability of a slightly advantageous mutant gene with a general distribution of progeny size - a branching process model. *Journal of Mathematical Biology* **12**, 355–362. (doi:10.1007/BF00276922)

14 Taddei, F., Radman, M., Maynard Smith, J., Toupance, B., Gouyon, P.-H. & Godelle, B. 1997 Role of mutator alleles in adaptive evolution. *Nature.* **387**, 700–2. (doi:10.1038/42696)

15 Cirz, R. T. & Romesberg, F. E. 2007 Controlling mutation: intervening in evolution as a therapeutic strategy. *Crit. Rev. Biochem. Mol. Biol.* **42**, 341–54. (doi:10.1080/10409230701597741)

16 Wright, S. 1988 Surfaces of selective value revisited. *American Naturalist* **131**, 115–123. (doi:10.1086/284777)

17 Crow, J. F., Engels, W. R. & Denniston, C. 1990 Phase Three of Wright’s Shifting-Balance Theory. *Evolution* **44**, 233. (doi:10.2307/2409403)

18 Wade, M. & Goodnight, C. 1991 Wright’s shifting balance theory: an experimental study. *Science (80-. ).* **253**, 1015–1018. (doi:10.1126/science.1887214)

19 Peck, S. L., Ellner, S. P. & Gould, F. 2000 Varying migration and deme size and the feasibility of the shifting balance. *Evolution* **54**, 324–7. (doi:10.1111/j.0014-3820.2000.tb00035.x)

20 Moore, F. B.-G. & Tonsor, S. J. 1994 A Simulation of Wright’s Shifting-Balance Process: Migration and the Three Phases. *Evolution* **48**, 69. (doi:10.2307/2410004)

21 Gavrilets, S. 1996 On phase three of the shifting-balance theory. *Evolution* **50**, 1034–1041. (doi:10.2307/2410644)

22 Coyne, J. A., Barton, N. H. & Turelli, M. 2000 Is Wright’s shifting balance process important in evolution? *Evolution* **54**, 306–317. (doi:10.1111/j.0014-3820.2000.tb00033.x)

23 Whitlock, M. C. & Phillips, P. C. 2000 The exquisite corpse: a shifting view of the shifting balance. *Trends in Ecology & Evolution* **15**, 347–348. (doi:10.1016/S0169-5347(00)01930-3)

24 Whitlock, M. C. 1995 Variance-Induced Peak Shifts. *Evolution* **49**, 252. (doi:10.2307/2410335)

25 Whitlock, M. C. 1997 Founder Effects and Peak Shifts Without Genetic Drift: Adaptive Peak Shifts Occur Easily When Environments Fluctuate Slightly. *Evolution* **51**, 1044. (doi:10.2307/2411033)

26 Hadany, L. 2003 Adaptive peak shifts in a heterogenous environment. *Theoretical Population Biology* **63**, 41–51. (doi:10.1016/S0040-5809(02)00011-4)

27 Hadany, L. & Beker, T. 2003 Fitness-associated recombination on rugged adaptive landscapes. *Journal of Evolutionary Biology* **16**, 862–870. (doi:10.1046/j.1420-9101.2003.00586.x)

28 Weissman, D. B., Desai, M. M., Fisher, D. S. & Feldman, M. W. 2009 The rate at which asexual populations cross fitness valleys. *Theor. Popul. Biol.* **75**, 286–300. (doi:10.1016/j.tpb.2009.02.006)

# Electronic supporting material

## Constraints on the parameter space

At the mutation-selection balance (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 [@Kimura1966]. 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 mutatns 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 Eqs. (1-2).

## 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). At the MSB the number of deleterious mutations per individual follows a Poisson distribution [@Haigh1978]. Therefore, the frequencies of mutation-free wildtype *ab* and single mutants *aB* and *Ab* are and . The probability *q* that a random newborn is a double mutant, given there are no double mutants and neglecting individuals with deleterious mutations is:

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

If mutation is stress-induced, then 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 stress-induced mutation:

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 stress-induced mutation 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 stress-induced mutation:

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

Note that by setting and because , is consistent with .

## Fixation probability with stress-induced mutation

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 [@Eshel1981] 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 [@Kimura1966]. Therefore, and

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

Assuming that is small we get

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

However, as we have shown before [@Ram2012], the mean fitness of a population with stress-induced mutagenesis can be different from because of mutation rate variation. The mean fitness with stress-induced mutagenesis can be calculated by divding the population to the mutation-free fraction which has fitness *1* and the rest of the population with a fraction of . Within the the latter fraction, individuals have at least one deleterious mutation and additional mutations are Poisson distributed with expectation , because these individuals are hypermutating. Therefore the mean fitness of this fraction is . Taken together, the mean fitness of a population with stress-induced mutagenesis is:

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

Pluging in the population mean fitness to the relative fitness of the double mutant we get

Pluging that in the fixation probability gives the final result:

|  |  |
| --- | --- |
|  | (12) |

This fixation probability can be further simplified by first-order approximations:

|  |  |
| --- | --- |
|  | (13) |

## Adaptation time

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 which will go to fixation *Tw*. This is true as long as fixation is much faster than mutation (guaranteed by *µ*2*<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:

|  |  |
| --- | --- |
|  | (14) |

Without stress-induced mutation, we plug in Eqs. (10) and (5) and get:

|  |  |
| --- | --- |
|  | (15) |

With stress-induced mutation we plug in Eqs. (6) and (13) and get:

|  |  |
| --- | --- |
|  | (16) |

Comparing these two expressions, we can write the adaptation rate (the inverse of the expected adaptation time) with stress-induced mutagenesis as a function of the adaptation rate without it:

Now, because the second term is positive, if then we can infer that the rate with stress-induced mutagenesis is faster than without. This condition can ve rewritten:

Using the quadrate formula this translates to:

Because *U* is very small, *1-2U* is well apprxomated by *1*, and *1/U2* is much larger than *2*, the RHS of this inequality can be approximated by 2 and we get the following condition:

Therefore, an approximate sufficient condition for stress-induced mutagenesis to decrease the adaptation time is:

|  |  |
| --- | --- |
|  | (17) |