Epigenetically Inherited Mutation Rates Predicted to Maximize Adaptation Rates

Gabriela Lobinska[1], Yitzhak Pilpel[1*], Yoav Ram[2,3*]

1 Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100 Israel

2 School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv 6997801 Israel

3 School of Computer Science, IDC Herzliya, Herzliya 4610101, Israel

* Correspondence: pilpel@weizmann.ac.il, yoav@yoavram.com

ABSTRACT

The mutation rate is an important determinant of evolutionary dynamics. Because the mutation rate determines the rate of appearance of beneficial and deleterious mutations, it is subject to second-order selection. The mutation rate varies between and within species and populations, increases under stress, and is genetically controlled by mutator alleles. The mutation rate may also vary among genetically identical individuals: empirical evidence from bacteria suggests that the mutation rate may be affected by translation errors and expression noise in various proteins (1). Importantly, this non-genetic variation may be heritable via transgenerational epigenetic inheritance. Here we investigate how the inheritance mode of the mutation rate affects the rate of adaptive evolution on rugged fitness landscapes. We model an asexual population with two mutation rate phenotypes, non-mutator and mutator. An offspring may switch from its parental phenotype to the other phenotype. The rate of switching between the mutation rate phenotypes is allowed to span a range of values. Thus, the mutation rate can be interpreted as a genetically inherited trait when the switching rate is low, as an epigenetically inherited trait when the switching rate is intermediate, or as a randomly determined trait when the switching rate is high. We find that epigenetically inherited mutation rates result in the highest rates of adaptation on rugged fitness landscapes for most realistic parameter sets. This is because an intermediate switching rate can maintain the association between a mutator phenotype and pre-existing mutations, which facilitates the crossing of fitness valleys. Our results provide a rationale for the evolution of epigenetic inheritance of the mutation rate, suggesting that it could have been selected because it facilitates adaptive evolution.

INTRODUCTION

Variation in mutation rates. Mutation rates vary widely between species, between and within populations and between environmental conditions. The lowest known mutation rate was found in *Paramecium tetraurelia,* whose mutation rate is

about 10^{-11} per bp per generation (2). The mutation rate in humans, as well as in several other eukaryotes such as *Drosophila melanogaster* or *Caenorhabditis elegans*, is about $10^{-8} - 10^{-9}$ per bp per generation (3–6), and can reach 10^{-5} per bp per generation in single-strand RNA viruses (7).

The mutation rate also varies between and within populations (8–10). This can be due to an increased mutation rate induced by mutator alleles, such as mutants in mismatch repair genes (e.g. MSH2 in *Escherichia coli*), which can exhibit a mutation rate up to 225-fold higher than the wildtype mutation rate (11). High variation in mutation rate within populations has been observed in isogenic populations of *E. coli* (9,10).

The mutation rate can be regulated by environmental or cellular conditions (12). *Stress-induced mutagenesis* is the increase in mutation rate as a response to stress or maladaptation (13–16). It has been extensively studied in *E. coli* (17), and has also been described in other organisms, including yeast (18), flies (19), and human cancer cells (15,20). It has been demonstrated, using mathematical and computational models, that stress-induced mutagenesis facilitates complex adaptive evolution; increases the population mean fitness at the mutation-selection balance; and can evolve in both constant and changing environments, even in the presence of rare horizontal gene transfer (20–23).

Evolution of the mutation rate. Mutators—individuals with an above-average mutation rate—become burdened with a higher mutational load compared to their non-mutator counterparts (24,25). Hence, we expect a reduction of the mutation rate after the population is well-adapted to its environmental conditions, assuming heritability of the mutator phenotype (26–28), such that the mutation rate will evolve towards some minimum. However, the evolution of the mutation rate may also be affected by genetic drift (29,30), biological and physical constraints (31–33), the cost of replication fidelity (32,34), and second-order selection due to hitchhiking with beneficial mutations (28,35–39). Additionally, the topology of the fitness landscape might constrain the evolution of mutators (40). Mutators may also be favoured in situations where beneficial mutations are frequent, such as frequently changing environments that require constant generation of adaptations (41–44). Frequent generation of mutator alleles could also lead to a substantial equilibrium frequency of these alleles, even under strong selection (45).

A number of theoretical studies are supported by empirical evidence. Experimental validation for the hypothesis that weak mutator alleles could hitch-hike with beneficial mutations in a substantial proportion of cases (35) comes from the *Long-Term Evolutionary Experiment* (38), in which 3 of 12 *E. coli* lineages evolved a 10-100-fold increase in mutation rate after 10,000 generations due to fixation of mutator alleles. However, the mutation rate was later reduced by anti-mutator alleles, perhaps because the opportunities for adaptive evolution dwindled (38). Populations evolving from a single ancestor have been found to exhibit variation in the mutation rate, suggesting that the generation of mutator alleles of various intensities may be frequent (9).

Epigenetic inheritance. The mutator phenotype is commonly thought to be genetically inherited, the result of, for example, an inactivating mutation in DNA mismatch repair genes. Epigenetic inheritance of the mutation rate, including epigenetic inheritance or stochastically determined mutation rates, have received little attention. Ninio (46) has suggested that *transient mutators* could arise, for example, as a result of a translation error in DNA repair proteins. Such an error

could lead to a very strong mutator phenotype, but its inheritance via cytoplasmic inheritance would be limited to a few generations, because it is not encoded in the genotype. In this scenario, individuals can increase their mutation rate for a few generations, thus generating beneficial mutations without a long-term commitment to an elevated mutation rate and the accumulation of deleterious mutations it entails. More recently, Bonhoeffer et al. (47) have shown that stochastic fluctuations in the mutation rate increase the population mean fitness by producing a subpopulation with below-average mutation rate when the population is well-adapted, and a subpopulation with above-average mutation rate when there is potential for further adaptation.

There are several lines of empirical evidence for non-genetic inheritance of the mutation rate. Expression noise in genes involved in mutation rate regulation may cause it to fluctuate stochastically. Such noise has been demonstrated for many gene promoters (48) and has been shown to scale with the abundance of the corresponding protein (49). Translation errors have been shown to occur more often in genes involved in DNA repair and replication (49 and work in progress). This corresponds to the mechanism proposed by Ninio (46). Moreover, a mechanism for epigenetic inheritance of the mutation rate in *E. coli* has been described by Uphoff *et al.* (1): Ada is a DNA repair protein that is induced under high pH stress conditions. Ada has a low copy number—each cell carries, on average, a single Ada molecule. Therefore, during cell division a substantial proportion of daughter cells will inherit no Ada molecules and therefore experience an elevated mutation rate. Furthermore, *Ada* is auto regulated, which intensifies the difference in Ada concentration between cells that inherited zero Ada molecules and cells that have inherited at least one Ada molecule. Thus, cells that initially have, for example, two Ada molecules, will produce daughter cells with a lower mutation rate compared to cells that initially have a single Ada molecule. Thus, stochastic fluctuations in Ada copy number can have transgenerational effects on the mutation rate (1).

The inheritance mode of the mutation rate can therefore be described as a spectrum: at one end, mutator alleles that arise as mutations in genes involved in DNA repair or replication, with little to no effect of stochastic factors ; on the other end of the spectrum, stochastic fluctuations in concentrations of proteins that affect the mutation rate are of such magnitude that there is no correlation between parent and offspring mutation rate; and in the middle of the spectrum, epigenetic modes of inheritance such as cytoplasmic inheritance of mistranslated proteins or transgenerational epigenetic inheritance of fluctuations in protein concentrations.

Here, we investigate the evolution of the mutation rate along this spectrum of inheritance modes. We have developed a Wright-Fisher model with explicit inheritance of two mutation rate phenotypes: non-mutator and mutator. The phenotype switching rate determines the probability that an offspring will switch its mutation rate phenotype rather than inherit it from its parent. We find that intermediate switching rates, corresponding to epigenetic inheritance of the mutation rate, lead to the highest rates of adaptation on theoretical and empirical rugged fitness landscapes. We suggest this is due to the increased association of mutators with the genetic variation they produce, which accelerates the rate of fitness-valley crossing.

MODELS

We first describe our general model, followed by three specific models, each focusing on a different fitness landscape.

General evolutionary model

We consider an asexual haploid population with non-overlapping generations. We model the effects of mutation, phenotypic switching, natural selection, and genetic drift using a Wright-Fisher model (51). Individuals are fully characterized by their **mutation rate phenotype**, either non-mutator (*m*) or mutator (*M*), and their **genotype**, which determines their fitness. The genotype is specified differently in each fitness landscape (see below). The combination of these two characteristics is termed *pheno-genotype* (52), and the frequency of individuals with mutator phenotype *z* and genotype *g* is denoted by f_{zg}. In all cases we assume no recombination and, therefore, complete linkage.

Mutation. The pheno-genotype frequencies after mutation are given by

$$f'_{zg} = \sum_j u_{j\to g}\, f_{zj} \quad , \tag{1}$$

where *j* is an index over all possible genotypes, *z* is the mutation rate phenotype, either *m* or *M* for non-mutator and mutator, respectively, and $u_{j\to \boldsymbol{g}}$ is the mutation transition probability from genotype *j* to genotype *g*. The specific transition probabilities $u_{j\to \boldsymbol{g}}$ in each fitness landscape are given below.

Phenotype switching. In every generation an individual may switch its mutation rate phenotype with probability *γ*. We assume this *switching rate γ* is symmetric from non-mutator *m* to mutator *M* and vice versa (see **Figure 1A**). Thus, the pheno-genotype frequencies after phenotype switching are given by

$$\begin{aligned} f''_{mg} &= (1-\gamma)\cdot f'_{mg} + \gamma \cdot f'_{Mg} \\ f''_{Mg} &= \gamma \cdot f'_{mg} + (1-\gamma)\cdot f'_{Mg} \ . \end{aligned} \tag{2}$$

Selection. In accordance with previous work on the evolution of modifier alleles, we assume that the mutator phenotype is neutral (27,45,53), that is, it does not affect fitness. Hence, we can focus on the fitness w_g of genotype *g*, rather than the fitness of the pheno-genotypes. The specific fitness values are determined by the fitness landscape, see below. The pheno-genotype frequencies after selection is

$$f'''_{zg} = \frac{w_g}{\bar{w}} \cdot f''_{zg} \quad , \tag{3}$$

where $\bar{w}$ is the population mean fitness, $\bar{w} = \sum_g w_g \left(f''_{mg} + f''_{Mg}\right)$.

Genetic drift. We model the effect of random genetic drift by drawing the number of individuals of each pheno-genotype from a multinomial distribution parameterized by the population size *N* and the pheno-genotype frequencies after mutation, phenotype switching, and selection, f'''_{zg}.

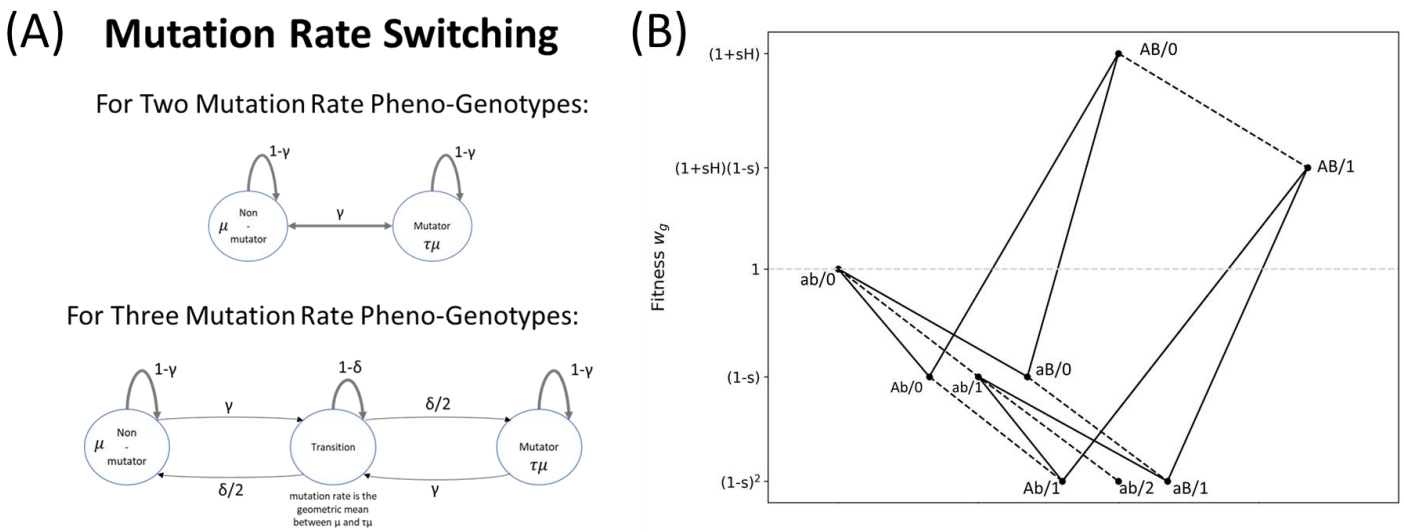


Figure 1. (A) Mutation rate switching is modelled by a Markov chain. (B)Two-peak fitness landscape. Fitness after the environmental changes is shown on the y-axis. Solid and dashed lines denote mutations in major and background loci, respectively. See Table 2 for summary of model parameters. Figure adapted from (20).

Numerical analysis. The effects of random drift can sometimes be neglected (e.g., very large population size). In these cases, pheno-genotype frequencies are only affected by mutation, phenotype switching, and selection. We refer to this version of the model as the *deterministic model* and solve it numerically by iterating eqs. (1-3) until the population reaches an equilibrium (i.e., the population mean fitness does not change from one generation to the next). The equilibrium frequencies can also be obtained by numerically solving for the eigenvalues of the mutation-selection transition matrix, and normalizing the eigenvector corresponding to the leading eigenvalue. We observed that the values obtained by solving the eigenvalue problem corresponded closely to the values obtained from a numerical iteration of the model. The full model with drift is referred to as the *stochastic model*. It is simulated by iteration of eqs. (1-3) and random sampling at each generation using a multinomial distribution. We then set up 500 runs of the full model, and stop the simulation when 475 runs terminate. The results are fitted to a geometric distribution and we report the probability of appearance (or adaptation) as the best fit.

Adaptive evolution. We first consider populations at a mutation-selection balance around a local fitness peak. That is, at the mutation-selection balance the population is considered to be well-adapted to its environment and the frequencies of the pheno-genotypes do not change from generation to generation ($f_{zg}''' = f_{zg}$) due to a balance between mutation and phenotype switching, which generate genetic and phenotypic variation, and selection and/or drift, which eliminate variation.

We assume that selective sweeps (54–57) occur sequentially rather than concurrently, that is, one adaptive genotype goes to fixation before another appears and there is no clonal interference. This assumption is supported by some empirical studies, and corresponds to a regime where selection is strong and mutation is weak (SSWM) (58). In this SSWM regime, mutations are rare and the advantage they confer is large. Hence, frequencies of pheno-genotypes with more than one mutation can be neglected at the mutation-selection balance (59).

We consider two types of adaptative evolution: the crossing of a single fitness valley, and the crossing of multiple fitness valleys until reaching the global fitness peak. In the first case, we approximate the probability of appearance and probability of fixation of an adaptive genotype. From these probabilities, we estimate the adaptation rate of the population and compare it to results of simulations of both the deterministic and the stochastic models. In the case of adaptation to the global fitness peak, we simulate the stochastic model and allow the population to evolve for 50,000 generations, during which we record the population mean fitness and the frequency of non-mutators and mutators in the population.

In the following, we describe three fitness landscapes. In each landscape we define how genotypes are represented and how the pheno-genotype frequencies are affected by mutation (eq. 1) and selection (eq. 3).

Two-Peak fitness landscapes

Genotype. Here the genotype contains two major loci, in which beneficial mutations may occur, and a large number of background loci, in which only deleterious mutations may occur. Thus, the genotype is denoted by the alleles at the two bi-allelic major loci and by the number of deleterious mutations accumulated in the background loci. For example, the wildtype genotype is *ab/0*: it carries alleles *a* and *b* at the major loci and zero mutant alleles in the background loci.

Selection. The fitness of the wildtype genotype $w_{ab/0}$ is set to 1. We assume that all mutant alleles in the background loci are deleterious with a selection coefficient *s,* such that the multiplicative fitness effect of *k* mutant alleles is $(1-s)^k$. Similarly, single mutations at the major loci are deleterious with the same selection coefficient *s*. Initially, the double mutant *AB* is maladaptive with fitness effect $(1-s)^2$. However, after an environmental change occurs, *AB* is adaptive, with fitness effect *(1+sH)*, where *H* is the adaptation coefficient. Hence, after the environmental change there is *sign epistasis* between the two major loci, and the population must cross a *fitness valley* to adapt (60,61). **Table 1** and **Figure 1B** summarize the fitness values of the different genotypes.

Genotype *g*	Fitness w_g	
	before environmental change	**after environmental change**
ab/k	$(1-s)^k$	$(1-s)^k$
Ab/k and *aB/k*	$(1-s)^{k+1}$	$(1-s)^{k+1}$
AB/k	$(1-s)^{k+2}$	$(1+sH)(1-s)^k$

Table 1. Fitness values in the two-peak fitness landscape. *A*, *B* and *a*, *b* are the mutant and wildtype alleles of the major loci, respectively; *k* is the number of mutant alleles in the background loci; *s* is the selection coefficient against mutant alleles; *H* is the adaptation coefficient. See Table 2 for summary of model parameters.

Mutation. The number of mutations per generation that occur at the background loci is Poisson distributed with expected value *U* or *τU* in individuals with non-mutator or mutator phenotype, respectively. Therefore, a genotype with *k*

deleterious mutations will mutate to have *k+l* deleterious (l is assumed to be positive due to neglecting back mutations) mutations with probability $e^{-U}U^{l}/l!$ or $e^{-\tau U}(\tau U)^{l}/l!$ in individuals with non-mutator or mutator phenotype, respectively. Mutations occur at the major loci with probability $\mu = U/n$ or $\mu = \tau U/n$, according to the phenotype, where n is the total number of loci in the genome and μ is the per-locus mutation rate. Therefore, genotype *ab* mutates to *Ab* or *aB* with probability $\mu(1-\mu)$ and to *AB* with probability μ^2, and genotypes *Ab* and *aB* mutate to genotype *AB* with probability μ. We neglect the effects of back mutations. See *Appendix A* for a formal description of the mutation transition probabilities.

Adaptation. The population is initially at a mutation-selection balance around the wildtype genotype *ab*/0, which is the fittest genotype. We assume the population size *N* to be large enough so that single mutants (*Ab* and *aB*) are present even in a population of non-mutators, but small enough so that double mutants (*AB*) are absent even in a population of mutators. Thus, the adaptive genotype must appear due to de-novo mutation rather than from standing variation. Taken together, we assume $e^{U/s}\frac{s}{\mu} < N < e^{\frac{\tau U}{s}}\left(\frac{s}{\tau\mu}\right)^{2}$ (see *Appendix B*).

Table 2 summarizes the model parameters with biologically relevant values.

Symbol	Name	Estimate	References
N	Population size	$10^5 - 10^8$	(62)
U	Genome-wide mutation rate	0.0004–0.003	(63)
n	Number of loci	4000-5000	(64)
μ	Per-locus mutation rate	U/n	
τ	Fold-Increase in mutator mutation rate	10-1000	(37,45,65)
s	Selection coefficient against deleterious mutations	0.001–0.03	(66)
H	Adaptation coefficient	1–10	(66)
γ	Mutator phenotype switching rate	$10^{-6} - 0.5$	(45,46)
δ	Switching rate from transition mutation rate phenotype	0-1	

Table 2. Summary of model parameters. Adapted from (20).

Gradual transition between mutator and non-mutator phenotypes

We mostly consider a model with two mutation rate phenotypes: a mutator and a non-mutator. In biological systems, the transition between a mutator and a non-mutator phenotype would likely be gradual rather than abrupt, and potentially occur over several generations (i.e., cell divisions).

We therefore extend our model to consider an intermediate, transient mutation rate set as the geometric average of the non-mutator and the mutator mutation rates. We call this additional mutation rate pheno-genotype *t*, with f_{tg} the frequency of an individual with genotype *g* and the transient mutation rate phenotype *t*. We introduce a parameter δ that modulates the transition between mutator and non-mutator; the transient mutation rate phenotype (*t*) is inherited with probability $1-\delta$, and switches to the mutator or non-mutator phenotype with probability $\delta/2$ (see **Figure 1A**). Thus, the pheno-genotype frequencies after phenotype switching are given by

$$\begin{aligned} f''_{mg} &= (1-\gamma)\cdot f'_{mg} + \delta/2 \cdot f'_{tg} \\ f''_{tg} &= \gamma \cdot f'_{mg} + (1-\delta)\cdot f'_{tg} + \gamma \cdot f'_{Mg} \\ f''_{Mg} &= (1-\gamma)\cdot f'_{Mg} + \delta/2 \cdot f'_{tg}\,. \end{aligned} \tag{4}$$

NK fitness landscape

Genotype. The NK landscape is commonly used in the study of epistatic interactions (67). This model allows to generate mathematically fitness landscapes by assigning fitness to each possible genotype. It has two parameters: *n* for the number of bi-allelic loci in the genotype and *k* for the number of loci each locus interacts with. The main advantage of the NK landscape is its biological interpretation and its tunable ruggedness (68). Here, the genotype consists of *n=6* bi-allelic loci, which results in $2^6 = 64$ possible genotypes.

Selection. To construct an NK landscape, we first generate all possible substrings of length *k* and assign to each of them a random fitness value between 0 and 1, sampled from a continuous uniform distribution. The fitness of each genotype w_g is then computed as the sum of the fitness effects of the substrings it contains. A position will thus influence the fitness according to the number of substrings that contain it, which increases with *k*. Hence, the ruggedness of a landscape increases with *k* (69). We consider three NK landscapes with low, intermediate, and high ruggedness corresponding respectively to *k* = 1, *k* = 3, and *k* = 5. See **Figs. A11** and **A12** and **Table A2** for properties of the constructed landscapes.

Mutation. A mutation switches the identity of a random bi-allelic locus from one allele to the other. The mutation transition probability $u_{j\to g}$ between two genotypes *j* and *g* that differ in *d* loci is $\varphi(d)$, where the φ is the probability mass function of a binomial distribution $Bin(n, \mu)$ for non-mutators or $Bin(n, \tau\mu)$ for mutators, μ is the per-locus mutation rate, and τ is the mutation rate fold-increase in mutators.

Adaptation. The initial population is at a mutation-selection balance around a genotype that is a local fitness maximum (see *Appendix C*). In the set of NK landscapes that we generated for the least rugged landscape, *k* = 1, there exists only

one local maximum other than the global maximum. This local maximum genotype was at a distance of *d*=4 single-locus mutations from the global maximum genotype. For landscapes of intermediate and high ruggedness with *k=3* and *k=5*, respectively, we also set the initial genotype to be a local maximum at a distance of *d=4* from the global maximum.

Empirical fitness landscape

We examine the insights gained on two-peak and NK landscapes with an empirical fitness landscape (67): the *Aspergillus niger* fitness landscape (69). We chose this landscape because: (i) it is complete with fitness measurements of all 256 genotypes, so we can avoid the interpolation of fitness values of missing genotypes; and (ii) fitness was measured as growth rate relative to the wildtype, as opposed to other studies that quantify some proxy phenotype for fitness such as fluorescence or DNA binding (70–72).

Genotype. The genotype consists of eight bi-allelic loci, each with either the wildtype or the mutant allele. de Visser *et al.* (69) engineered 256 genotypes to bear all possible combinations of wildtype/mutant alleles in these eight loci. The loci are: *fwnA1* (fawn-colored conidiospores); five auxotrophic markers, *argH12* (arginine deficiency), *pyrA5* (pyrimidine deficiency), *leuA1* (leucine deficiency), *pheA1* (phenyl-alanine deficiency), and *lysD25* (lysine deficiency); and two resistances, *oliC2* (oligomycin resistance) and *crnB12* (chlorate resistance).

Selection. de Visser *et al.* (69) estimated fitness by measuring the growth rate of each strain relative to the growth rate of the wildtype strain (i.e., with eight wildtype alleles). Out of 256 genotypes, 70 genotypes were lethal (fitness is zero). The landscape is rugged: it contains 15 local maxima (including the global maximum). See **Figs. A9** and **A10** for the distribution of fitness values.

Mutation. As in the NK landscape, a mutation switches a random bi-allelic locus from one allele to the other. The mutation transition probability $u_{j\to g}$ between two genotypes *j* and *g* that differ in *d* loci is again $\varphi(d)$, where the φ is the probability mass function of a binomial distribution $Bin(n, \mu)$ for non-mutators or $Bin(n, \tau\mu)$ for mutators, μ is the per-locus mutation rate, τ is the mutation rate fold-increase in mutators, and $n = 8$ is the number of loci.

Initialization. The initial population is at a mutation-selection balance around a local maximum genotype (see *Appendix C*) chosen to be *d=4* mutations away from the global maximum.

RESULTS

Adaptation on two-peak fitness landscapes

To analyze the adaptation rate, we focus on four pheno-genotypes (their frequencies appear in parentheses): non-mutators with wildtype genotype *ab/0* (m_0); non-mutators with a single major locus mutant allele, *aB/0* or *Ab/0* (m_1); mutators with wildtype genotypes (M_0); and mutators with a single major locus mutant allele (M_1). Note that a single mutant allele is always deleterious in the two-peak landscape. We approximate the mutation-selection balance (MSB) frequencies of these pheno-genotypes in *Appendix D* and compare these approximations to numerical results of the deterministic and stochastic models. We find a good agreement between all three results.

Starting from the MSB, we assume an environmental change provides an adaptive opportunity. In particular, under the new environmental conditions, the double mutant *AB* has the highest fitness. The rate of adaptation is calculated from the probability of appearance and the probability of fixation of this adaptive double mutant. These probabilities are derived from the frequencies m_0, m_1, M_0, and M_1. We neglect the frequency of a double mutant at the MSB (i.e., we disregard terms of order μ^2) due to our assumption on the population size and the mutation rate (see above). Hence, the probability q_A of appearance of a double mutant without background mutant alleles, *AB/0*, as a result of a mutation in a single mutant (either *aB/0* or *Ab/0*) in a population that currently does not have any double mutants is

$$q_A = \mu\, e^{-U} m_1 + \tau\mu\, e^{-\tau U} M_1. \tag{5}$$

Here, μm_1 and $\tau\mu M_1$ are the probabilities that a major locus mutation occurs in a single mutant genotype with either the non-mutator or the mutator phenotype, respectively, and e^{-U} and $e^{-\tau U}$ are the probabilities that no background loci mutations occur, again with either the non-mutator or mutator phenotype. Following Eshel (73), the fixation probability ρ_F of a double mutant once it appears in a large population only depends on the selection coefficient *s* and the adaptation coefficient *H*, $\rho_F = \frac{2sH}{1+sH} \approx 2sH$, where the latter approximation applies when *sH* is much smaller than 1.

The probability that in the next generation no double mutant appears and fixes in a population of size *N* is $(1 - q_A\rho_F)^N$. Assuming the number of generations until the appearance of a double mutant that goes to fixation is geometrically distributed, the adaptation rate ν is

$$\nu = 1 - (1 - q_A\rho_F)^N \approx N q_A \rho_F, \tag{6}$$

where the approximation holds when $N q_A \rho_F$ is small. Note that the switching rate γ affects the adaptation rate only indirectly due to its role in the MSB frequencies m_0, m_1, M_0, and M_1. **Figure 2** compares results from: (i) simulations of the stochastic model (i.e., with background loci and genetic drift); (ii) the approximation in Eq. 6 using MSB frequencies from our analytic approximation (*Appendix D*, Eqs. A2-A16); and (iii) the approximation in Eq. 6 using numerically determined MSB frequencies (without drift, Eqs. 1-3). We find good agreement between the two approximations, (ii) and (iii), with $R^2 = 0.96$, as well as between (i) and (ii) ($R^2 = 0.93$) and (i) and (iii) ($R^2 = 0.96$). Importantly, the highest adaptation rate is reached with intermediate switching rates, corresponding to epigenetic inheritance of the mutation rate. In **Table 3**, we record the proportion of mutator, non-mutator, wildtype, and mutant pheno-genotypes for three switching rates: one corresponding to genetic inheritance ($\gamma = 10^{-6}$), one corresponding to epigenetic inheritance ($\gamma = 0.00344$), and one corresponding to random phenotypes ($\gamma = 0.5$). We also compute the association *C* between the mutant genotype and the mutator pheno-genotype. The right column in **Table 3** depicts, for each switching rate, the proportion of mutants that are mutators. This proportion is shown her to be maximized for the intermediate, epigenetic, switching rate. Namely that this regime we get the strongest coupling between being a mutator and being a mutant.

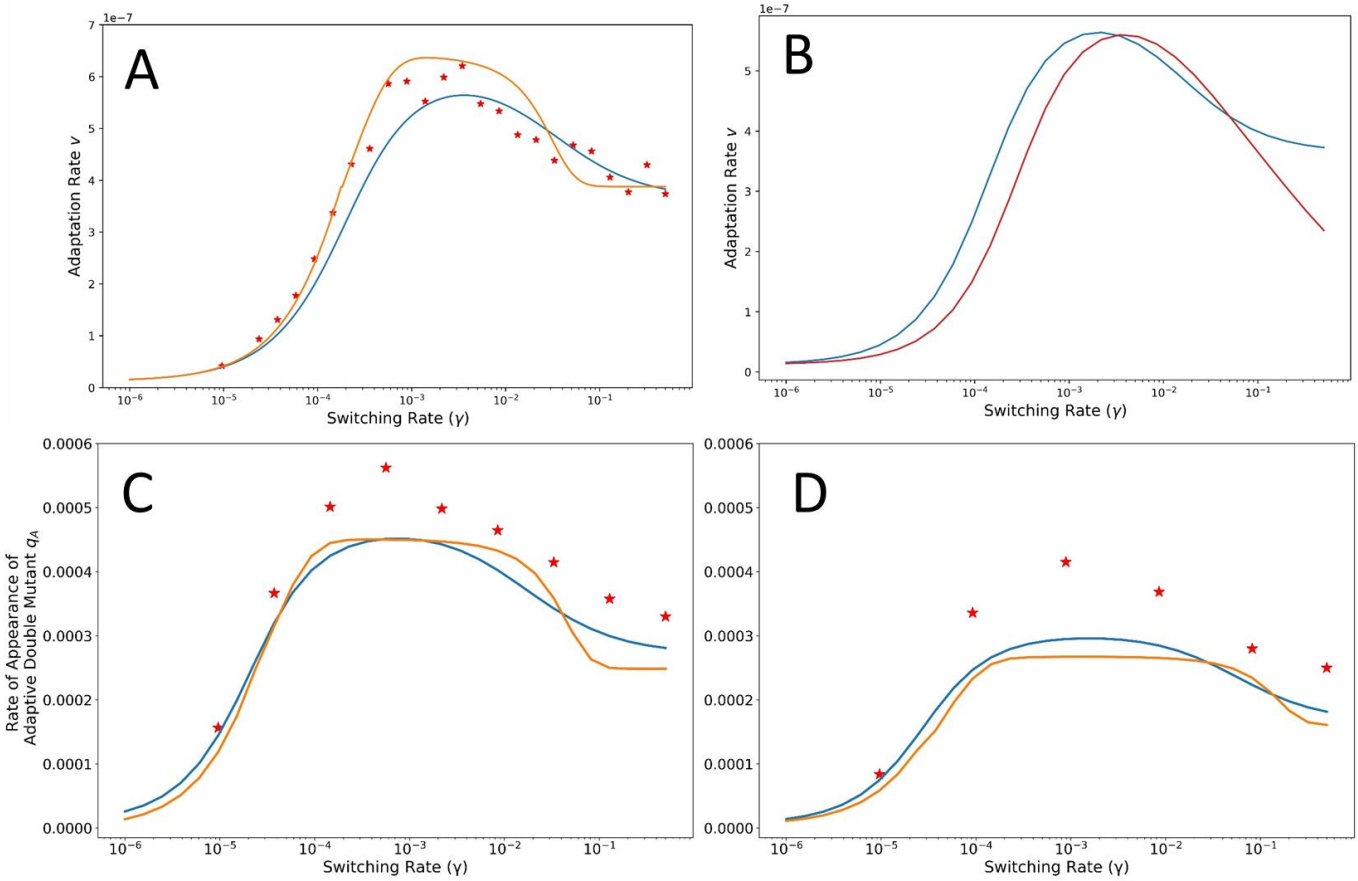


Figure 2. Intermediate switching rates maximize adaptation rates on two-peak fitness landscapes, NK and empirical fitness landscapes. (A,C,D) Stars show the adaptation rate (for NK and empirical landscapes: appearance rate) from simulations, i.e. the stochastic model, and genetic drift. Lines show the analytic approximation (Eq. 6) using MSB frequencies from an analytic approximation (orange; Appendix D, Eqs. A2-A16) or numerical analysis (blue; Eqs. 1-3 without drift). (B) We compare the rate of adaptation from numerical analysis for the standard model and the model with gradual transition between mutator and non-mutator phenotypes. Coefficients of determination for every combination of orange, blue lines and stars are shown in Table A3 and are all greater than 0.9. Parameters: (A) Two-peak landscape: mutation rate, μ = 0.00004; mutation rate increase, τ = 10; selection coefficient, s = 0.03, H=5, N = 10^7. (B) Two-peak landscape, with transition mutation rate: mutation rate, μ = 0.00004; mutation rate increase, τ = 10; selection coefficient, s = 0.03, H=5, δ=0.5, N = 10^7. (C) NK landscape (k=3): mutation rate, μ = 0.00004; mutation rate increase, τ = 10; selection coefficient, s = 0.03, N = 10^6. (D) Aspergillus niger landscape, $\mu = 10^{-6}$; $\tau = 10$, s = 0.15, N = 100,000.

A $\gamma = 10^{-6}$	Non-Mutator	Mutator	% Mutator
Wildtype	0.996	0.0028	*0.3%*
Mutant	$5.1 \cdot 10^{-7}$	$1.4 \cdot 10^{-8}$	*2.6%*

B $\gamma = 0.00344$	Non-Mutator	Mutator	% Mutator
Wildtype	0.519	0.474	*48%*
Mutant	$3.9 \cdot 10^{-7}$	$2.3 \cdot 10^{-6}$	***85%***

C $\gamma = 0.5$	Non-Mutator	Mutator	% Mutator
Wildtype	0.497	0.496	*49.9%*
Mutant	$1.3 \cdot 10^{-6}$	$1.4 \cdot 10^{-6}$	*51.7%*

Table 3. Pheno-genotype frequencies at the mutation-selection balance explain why intermediate switching rates maximize the adaptation rate. The association between mutator phenotypes and mutant genotypes is higher for intermediate switching rates (value in red in B). Here, background loci mutation rate, U = 0.00004; mutator mutation rate increase, τ = 10; selection coefficient, s = 0.03; switching rate, (A) γ = 10-6 (B) γ = 0.00344, and (C) γ = 0.5.

In Eq. 5, the frequency M_1 of mutators with a single major locus mutant allele has a large effect on the appearance probability q_A. Thus, to determine the role of the switching rate γ in the adaptation rate, we must determine how it affects M_1. This is because M_1 describes the association of standing genetic variation (mutant allele in a major locus) with an increased mutation rate. We define this association as:

$$C = \frac{M_1}{m_1 + M_1} = \frac{M_1}{p_S} \tag{7}$$

M_1 is maximized when the switching rate is intermediate, namely (see Eqs. A17 and A18, see *Appendix E*):

$$\frac{(\tau-1)U}{2} < \gamma < \frac{\log \tau}{s}. \tag{8}$$

This is because M_1 is approximately the ratio between a logistic function with rate $(\tau - 1)U/2$, describing the proportion of mutants at MSB, and a logistic function with rate $\log(\tau)/s$, describing the decrease in association between the mutator phenotype and the mutant genotype (see *Appendix D*). We find that in the 'strong selection–weak mutation' regime, intermediate switching rates in the range of $10^{-3} - 10^{-2}$ maximize the rate of adaptation (**Figure 3**; values outside the SSWM regime, $\tau U < s$, are beyond the scope of this study.)

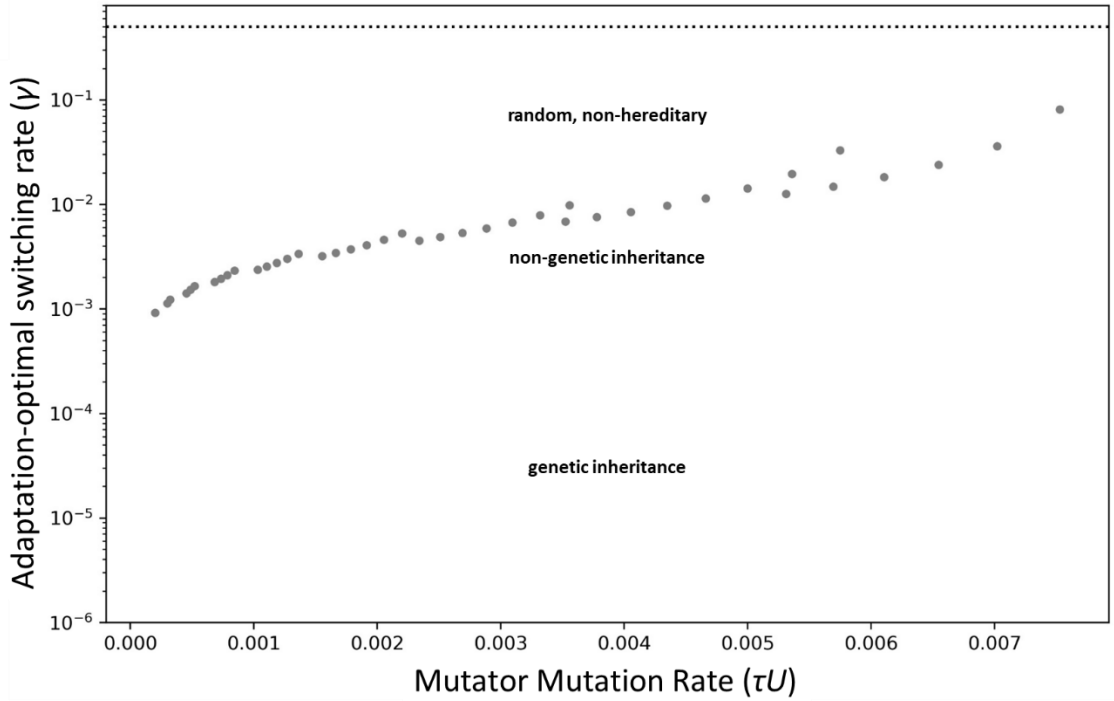


Figure 3. Adaptation-optimal switching rates. The adaptation-optimal switching rate increases with the mutator mutation rate, τU, until reaching $\gamma = 0.5$ when $\tau U = s$. Here, selection coefficient, s = 0.01. Details on the optimization can found in Appendix G.

Gradual transition between mutator and non-mutator phenotypes

When considering a third, transient phenotype between the non-mutator and mutator phenotypes, we also observe that the adaptation rate is maximised for intermediate switching rates (**Figure 2**). The curve describing the adaptation rate as a function of the switching rate is slightly shifted to the right; this can be explained by the fact that, because every transition between the mutator and non-mutator phenotypes requires a minimum of two generations, the switching rate is effectively lower.

The ratio between the maximum adaptation rate and the adaptation rate achieved for $\gamma = 0.5$ is higher than in the two mutation rate phenotypes model, which supports our conclusion that intermediate switching rates result in the highest rates of adaptation. In Appendix F, we show the equilibrium frequencies of the mutator, non-mutator, and transient mutation rate phenotypes as a function of the switching rate and provide an intuitive explanation for the larger difference between the adaptation rates for random and epigenetic switching rates compared to the two mutation rate phenotypes model.

Adaptation on NK landscapes and empirical landscapes

Crossing a single fitness valley

In the NK and empirical landscapes, the initial genotype is defined as the wildtype while all other genotypes are defined as mutants. To reach a mutation-selection balance (MSB) around the wildtype, it was temporarily set to have the global maximum fitness (*Appendix C*), and so the frequency of mutants was low compared to that of the wildtype. After reaching the MSB, the fitness values of the mutants were reset to their landscape values.

The rate of appearance of the double mutant was approximated using Eq. 5 with MSB frequencies m_1 and M_1 determined numerically (from Eqs. 1-3) or with our analytic approximation, as in the two-peak fitness landscape (*Appendix D*, Eqs. A2-

A16). Here, m_1 and M_1 correspond to the frequencies of non-mutators and mutators with genotypes one mutation away from the wildtype. Our approximation thus neglects the frequencies of genotypes more than one mutation away from the wildtype. **Figure 2** shows a comparison of these approximations with results of simulations of the stochastic model (with drift). Here we focus on the appearance rate, because the switching rate does not affect the fixation probability, which is the other factor determining the adaptation rate.

We find that the highest rate of double mutant appearance occurs for an intermediate switching rate. Similar to the two-peak landscape, we conclude that when crossing fitness valleys in multi-dimensional fitness landscapes such as NK landscapes and empirical landscapes, intermediate switching rates maximize the adaptation rate. This conclusion is true regardless of the extent of ruggedness (see **Figure A12**).

Crossing multiple fitness valleys

Next, we sought to confirm that the advantage of intermediate switching rates extends beyond a single valley crossing. Again, we initialized the population at a MSB around a local maximum of the fitness landscape, which we mark as the wildtype (*Appendix C*). We then allowed the population to evolve for 50,000 generations. At each generation, we recorded the proportion of simulations (out of 500) in which the genotype with the highest fitness (the global maximum) became dominant in the population. In the NK landscape with k = 5, we use the proportion of simulations in which a genotype fitter than the wildtype fixed in the population, because none of the simulation reached the global maximum, probably due to the high ruggedness of the landscape.

Again, we find that populations with an intermediate switching rate have the highest adaptation rate (**Figure 4**). We hypothesized that this advantage is due to the association between the mutator phenotype and standing genetic variation, which increases the probability to generate double mutants. To test this hypothesis, we artificially removed this association: at each generation, each individual was assigned a mutation rate phenotype at random, independent of its parent phenotype and its genotype, such that the frequency of the phenotypes is equal to their frequency at that generation from a simulation of the stochastic model. With an intermediate switching rate $\gamma = 3.4 \cdot 10^{-3}$, and with $N = 10^6$, $\mu = 10^{-6}$, and either $\tau = 10$ (for the NK landscape) or $\tau = 100$ (for the empirical landscape), the mutator frequency at MSB is almost 50%, close to the mutator frequency at MSB for high switching rate, $\gamma = 0.5$. Therefore, after reassigning mutation rate phenotypes at random among all individuals in the population, we expected the population to evolve similarly to a population with a high switching rate and without phenotype reassignment. Indeed, **Figure 4** shows that populations with intermediate switching rates and random phenotype reassignment (dotted red lines) evolve slower than corresponding populations with intermediate switching rate and without random phenotype assignment, and similarly to populations with high switching rates (solid orange lines). Thus, we conclude that fast adaptation with intermediate switching rates is due to the association between the mutator phenotype and the mutant genotype.

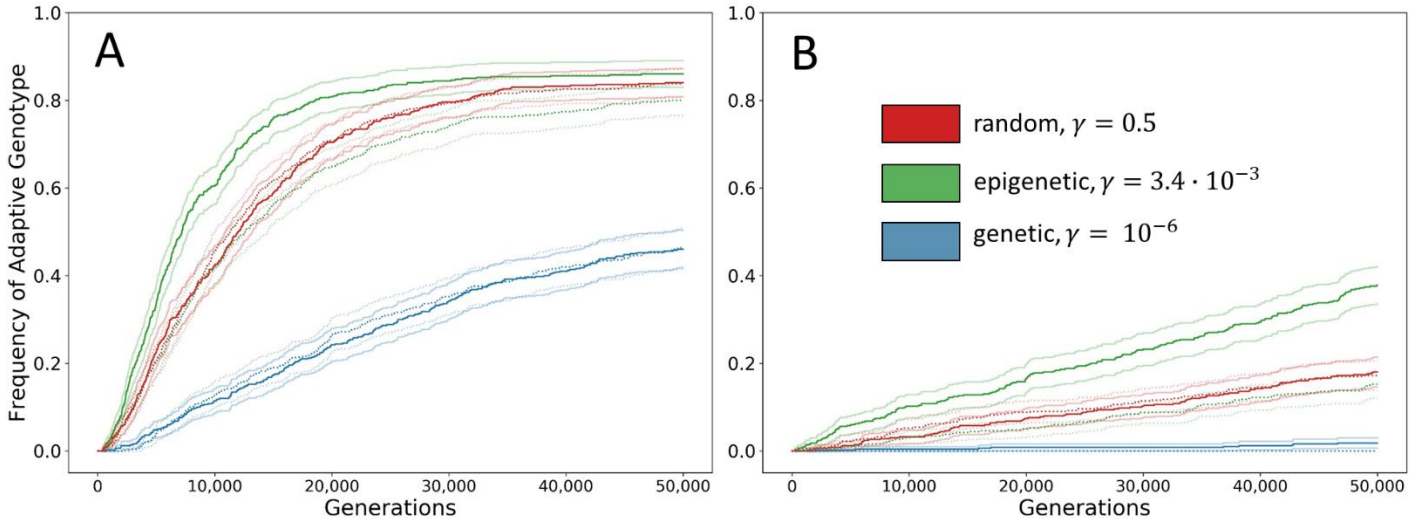


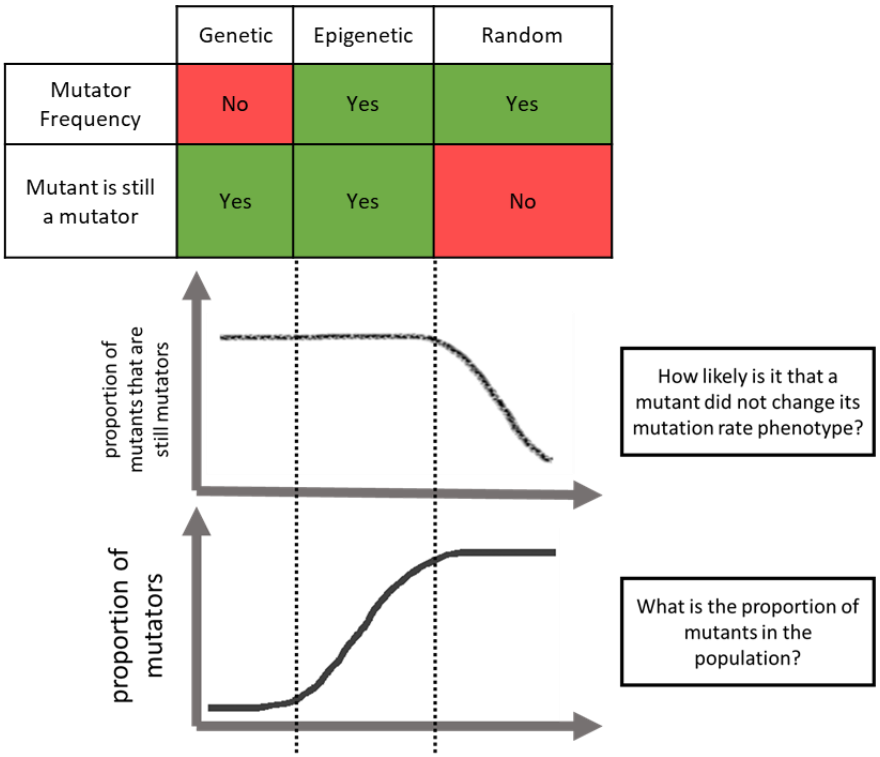
Figure 4. Intermediate switching rates maximize the adaptation rates on NK and empirical fitness landscapes. Solid lines show frequency of the adaptive genotype (global maximum) from simulations (stochastic model with drift). We observe that intermediate switching rate lead to the highest frequencies of adaptive genotype, and hypothesized that this is due to association between the mutator phenotype and pre-existing genetic variation. In order to test this hypothesis, we reran the simulation with removing the association between the mutator and its genotype – at each generation, each individual was re-assigned a random mutator rate phenotype. The mutator frequency was conserved and equal to the mutator frequency in the original simulation at any given generation. Dotted lines are results of simulations where mutator phenotypes were randomly reassigned. Transparent lines show the corresponding confidence intervals. (A) NK landscape with k=3; $N = 10^6$; $\mu = 10^{-6}$; $\tau = 10$. (B) Empirical fitness landscape; $N = 10^6$; $\mu = 10^{-6}$; $\tau = 100$.

DISCUSSION

Interpretation of results

Through a combination of analytical, numerical, and computational analysis, we have demonstrated that the adaptation rate on rugged fitness landscapes is maximized by intermediate rates of mutation-rate phenotype switching, roughly $10^{-3} < \gamma < 10^{-2}$ (albeit assuming strong selection and weak mutation, i.e., $s > \tau U$). We call such intermediate switching rates "epigenetic inheritance of the mutation rate".

With a low switching rate at $\gamma = 10^{-6}$, which corresponds to genetic inheritance of the mutation rate, mutators are strongly associated with their mutant alleles, both deleterious and beneficial ones. Therefore, selection efficiently acts against mutators due to the deleterious mutations they generate, and their frequency will be low (**Table 3A**). On the other side of the spectrum, $\gamma = 0.5$ and mutators reach a high frequency; however, their association to the mutant alleles they generate is quickly broken by phenotype switching, and the generation of double mutants is not maximized by the high frequency of mutators (**Table 3C**). In contrast, intermediate switching rates strike a good balance between a high frequency of mutators in the population and a strong association between the mutator and the mutant alleles it generates (**Table 3B,** for schematic explanation **see Figure 5**). This allows rapid generation of double mutants, which are crucial for adaptation on rugged fitness landscapes. Overall, intermediate switching rates are within the range $(\tau - 1)U/2 < \gamma < \log(\tau)/\, s$ (*Appendix D*).



387

388 **Figure 5. Intermediate switching rate (corresponding to epigenetic inheritance) maximizes adaptation rates due to the combination of high**
389 **mutator frequencies at mutation-selection balance and retaining the association between a mutator phenotype and single mutants.**

390 Mutation rate inheritance and phenotype switching

391 Switching between a mutator and a non-mutator phenotype is a type of phenotype switching (74). The increased
392 adaptability of populations with intermediate switching rates, compared to populations with low switching rates, is due
393 to the high rate of mutator generation, which balances the strength of selection against mutators at the mutation-
394 selection balance. Generation of a phenotype that is disadvantageous in the short term but potentially advantageous in
395 the long term is known as *bet-hedging* (74) and has been implicated as driving the evolution of phenotype switching
396 mechanisms (74–76). Here, our phenotype of interest is the mutation rate, and the product of the phenotype is the rate
397 of generation of mutations. We find that intermediate switching rates maximize the adaptation rate in rugged fitness
398 landscapes. Likewise, we suggest that intermediate switching rates may be optimal in other systems in which a specific
399 phenotype has a short-term disadvantage, and therefore a high switching rate is required to counter the effect of natural
400 selection, but the same phenotype may also be advantageous if it occurs in consecutive generations of the same lineage,
401 and therefore a not too high switching rate is beneficial in maintaining parent-offspring phenotype correlation.

402 Evolution of epigenetic inheritance

403 Although intermediate switching rates maximize the adaptation rate, it is not clear that an adaptation-optimal switching
404 rate can be selected for when competing against lower or higher switching rates. First, the proportion of mutants increases
405 with the switching rate (**Figure A2**). This means that the mutational load of subpopulations with higher switching rates
406 will be higher than the mutational load of subpopulations with lower switching rates, thus selecting for low switching
407 rates. Second, the advantage of a subpopulation with an adaptation-optimal switching rate depends on the mutation

supply. The adaptation rate of a subpopulation of size N_1 with a given switching rate γ_1 resulting in an appearance rate q_{A1} will be higher than the adaptation rate in a subpopulation of size N_2 with a given switching rate γ_2 resulting in an appearance rate q_{A2} if $\frac{N_2}{N_1} > \frac{q_{A1}}{q_{A2}}$ (Eq. 6). Therefore, selection for an allele that induces an intermediate switching rates will depend on its frequency in the population, similar to the case of mutator alleles (77).

For example, in the two-peak landscape with parameter values from **Figure 2,** the highest adaptation rate (with $\gamma = 3.4 \cdot 10^{-3}$) is 36-fold higher than the lowest adaptation rate (with $\gamma = 10^{-6}$). Hence, a subpopulation with the adaptation-optimal switching rate might be outcompeted by a subpopulation with a low switching rate if its initial frequency is less than 1/36 ~ 3%. Note that before the adaptive genotype appears, the frequency of the subpopulation with a high switching rate may decrease due to its higher mutational load, which further decreases its adaptive advantage. A more detailed treatment of the evolution of the adaptation-optimal switching rate is in progress in order to conclude whether we could expect to observe selection for epigenetic inheritance of the mutation rate in natural populations.

Conclusions

Empirical evidence indicates that the mutation rate may be determined both by genetically inherited factors, such as mutator alleles coding for deficient mismatch repair mechanisms (11,38), and by stochastic factors, such as fluctuations in protein concentrations, which may also be epigenetically inherited for a limited number of generations (1,48,50). We found that the adaptation-optimal switching rates are generally in the epigenetic range and increase with the mutator mutation rate until reaching values that can be interpreted as random determination of the mutator phenotype (**Figure 3**). However, adaptation-optimal switching rates were not found in the range of genetic inheritance. This could explain why stochasticity in mutation rate inheritance is adaptive, as it appears to facilitate rapid adaptation on rugged landscapes. Stochastic mutation rates have been previously suggested to increase the population mean fitness both in well-adapted and adapting populations (47). This stochasticity can be manipulated experimentally (48,78) and hence the interplay between the switching rate and the capacity to adapt could be tested empirically.

REFERENCES

1. Uphoff S, Lord ND, Okumus B, Potvin-Trottier L, Sherratt DJ, Paulsson J. Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. Science (80-). 2016;

2. Sung W, Tucker AE, Doak TG, Choi E, Thomas WK, Lynch M. Extraordinary genome stability in the ciliate Paramecium tetraurelia. Proc Natl Acad Sci U S A. 2012;

3. Roach JC, Glusman G, Smit AFA, Huff CD, Hubley R, Shannon PT, et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science (80-). 2010;

4. Keightley PD, Ness RW, Halligan DL, Haddrill PR. Estimation of the spontaneous mutation rate per nucleotide site in a Drosophila melanogaster full-sib family. Genetics. 2014;

442 5. Denver DR, Dolan PC, Wilhelm LJ, Sung W, Lucas-Lledó JI, Howe DK, et al. A genome-wide view of Caenorhabditis elegans
443 base-substitution mutation processes. Proc Natl Acad Sci U S A. 2009;

444 6. Lynch M. Rate, molecular spectrum, and consequences of human mutation. Proc Natl Acad Sci U S A. 2010;

445 7. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral Mutation Rates. J Virol. 2010;

446 8. Matic I, Radman M, Taddei F, Picard B, Doit C, Bingen E, et al. Highly variable mutation rates in commensal and pathogenic
447 Escherichia coli. Science (80-). 1997;

448 9. Ramiro RS, Durão P, Bank C, Gordo I. Low mutational load and high mutation rate variation in gut commensal bacteria. PLoS
449 Biol. 2020;

450 10. Woo AC, Faure L, Dapa T, Matic I. Heterogeneity of spontaneous DNA replication errors in single isogenic Escherichia coli
451 cells. Sci Adv. 2018;

452 11. Lang GI, Parsons L, Gammie AE. Mutation rates, spectra, and genome-wide distribution of spontaneous mutations in
453 mismatch repair deficient yeast. G3 (Bethesda). 2013;

454 12. Matic I. Molecular mechanisms involved in the regulation of mutation rates in bacteria. Periodicum Biologorum. 2016.

455 13. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. Critical Reviews in
456 Biochemistry and Molecular Biology. 2007.

457 14. Gonzalez C, Hadany L, Ponder RG, Price M, Hastings PJ, Rosenberg SM. Mutability and importance of a hypermutable cell
458 subpopulation that produces stress-induced mutants in Escherichia coli. PLoS Genet. 2008;

459 15. Fitzgerald DM, Hastings PJ, Rosenberg SM. Stress-induced mutagenesis: Implications in cancer and drug resistance. Annu Rev
460 Cancer Biol. 2017;

461 16. Bjedov I, Tenaillon O, Gérard B, Souza V, Denamur E, Radman M, et al. Stress-induced mutagenesis in bacteria. Science (80-).
462 2003;

463 17. Foster PL. Stress-induced mutagenesis in bacteria. Critical Reviews in Biochemistry and Molecular Biology. 2007.

464 18. Heidenreich E. Adaptive mutation in Saccharomyces cerevisiae. Critical Reviews in Biochemistry and Molecular Biology.
465 2007.

466 19. Sharp NP, Agrawal AF. Evidence for elevated mutation rates in low-quality genotypes. Proc Natl Acad Sci U S A. 2012;

467 20. Ram Y, Hadany L. Stress-induced mutagenesis and complex adaptation. Proc R Soc B Biol Sci. 2014;

468 21. Ram Y, Hadany L. The evolution of stress-induced hypermutation in asexual populations. Evolution (N Y). 2012;

469 22. Dellus-Gur E, Ram Y, Hadany L. Errors in mutagenesis and the benefit of cell-to-cell signalling in the evolution of stress-
470 induced mutagenesis. R Soc Open Sci. 2017;

471 23. Ram Y, Hadany L. Evolution of stress-induced mutagenesis in the presence of horizontal gene transfer. Am Nat. 2019;

472 24. Kimura M. On the evolutionary adjustment of spontaneous mutation rates. Genet Res. 1967;

473 25. Kimura M, Maruyama T. THE MUTATIONAL LOAD WITH EPISTATIC GENE INTERACTIONS IN FITNESS. Genetics. 1966;

474 26. Altenberg L, Liberman U, Feldman MW. Unified reduction principle for the evolution of mutation, migration, and
475 recombination. Proc Natl Acad Sci U S A. 2017;

476 27. Liberman U, Feldman MW. Modifiers of mutation rate: A general reduction principle. Theor Popul Biol. 1986;

477 28. Wielgoss S, Barrick JE, Tenaillon O, Wiser MJ, Dittmar WJ, Cruveiller S, et al. Mutation rate dynamics in a bacterial population
478 reflect tension between adaptation and genetic load. Proc Natl Acad Sci U S A. 2013;

479 29. Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, et al. Genetic drift, selection and the evolution of the
480 mutation rate. Nature Reviews Genetics. 2016.

481 30. Jain K, Nagar A. Fixation of mutators in asexual populations: The role of genetic drift and epistasis. Evolution (N Y). 2013;

482 31. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A. 1991;

483 32. Furió V, Moya A, Sanjuán R. The cost of replication fidelity in an RNA virus. Proc Natl Acad Sci U S A. 2005;

484 33. Lynch M. The lower bound to the evolution of mutation rates. Genome Biol Evol. 2011;

485 34. Dawson KJ. Evolutionarily stable mutation rates. J Theor Biol. 1998;

486 35. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. Role of mutator alleles in adaptive evolution.
487 Nature. 1997;

488 36. Giraud A, Matic I, Tenaillon O, Clara A, Radman M, Fons M, et al. Costs and benefits of high mutation rates: Adaptive
489 evolution of bacteria in the mouse gut. Science (80-). 2001;

490 37. Tenaillon O, Toupance B, Nagard H Le, Taddei F, Godelle B. Mutators, population size, adaptive landscape and the adaptation
491 of asexual populations of bacteria. Genetics. 1999;

492 38. Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of E. coli. Nature. 1997;

493 39. Loh E, Salk JJ, Loeb LA. Optimization of DNA polymerase mutation rates during bacterial evolution. Proc Natl Acad Sci U S A.
494 2010;

495 40. Clune J, Misevic D, Ofria C, Lenski RE, Elena SF, Sanjuán R. Natural selection fails to optimize mutation rates for long-term
496 adaptation on rugged fitness landscapes. PLoS Comput Biol. 2008;

497 41. Tanaka MM, Bergstrom CT, Levin BR. The evolution of mutator genes in bacterial populations: The roles of environmental
498 change and timing. Genetics. 2003;

499 42. Leigh, EG. Natural Selection and Mutability. Am Nat. 1970;

500 43. Lachmann M, Jablonka E. The inheritance of phenotypes: An adaptation to fluctuating environments. J Theor Biol. 1996;

501 44. Kussell E, Leibler S. Ecology: Phenotypic diversity, population growth, and information in fluctuating environments. Science
502 (80-). 2005;

503 45. Desai MM, Fisher DS. The balance between mutators and nonmutators in asexual populations. Genetics. 2011;

504 46. Ninio J. Transient mutators: A semiquantitative analysis of the influence of translation and transcription errors on mutation
505 rates. Genetics. 1991;

506 47. Bonhoeffer S, Mayer SI, Alexander HK. Population Heterogeneity in Mutation Rate Increases the Frequency of Higher-Order
507 Mutants and Reduces Long-Term Mutational Load. Mol Biol Evol [Internet]. 2016 Nov 11;34(2):419–36. Available from:
508 https://doi.org/10.1093/molbev/msw244

509 48. Sharon E, Van Dijk D, Kalma Y, Keren L, Manor O, Yakhini Z, et al. Probing the effect of promoters on noise in gene expression
510 using thousands of designed sequences. Genome Res. 2014;

511 49. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, et al. Noise in protein expression scales with natural protein
512 abundance. Nat Genet. 2006;

513 50. Mordret E, Dahan O, Asraf O, Rak R, Yehonadav A, Barnabas GD, et al. Systematic Detection of Amino Acid Substitutions in
514 Proteomes Reveals Mechanistic Basis of Ribosome Errors and Selection for Translation Fidelity. Mol Cell. 2019;

515 51. OTTO SP, DAY T. Mathematical Modeling in Biology. In: A Biologist's Guide to Mathematical Modeling in Ecology and
516 Evolution. 2019.

517 52. Feldman MW, Cavalli-Sforza LL. The evolution of continuous variation. II. Complex transmission and assortative mating.
518 Theor Popul Biol. 1977;

519 53. James A, Jain K. Fixation probability of rare nonmutator and evolution of mutation rates. Ecol Evol. 2016;

520 54. Hermisson J, Pennings PS. Soft sweeps: Molecular population genetics of adaptation from standing genetic variation.
521 Genetics. 2005;

522 55. Pennings PS, Hermisson J. Soft sweeps II - Molecular population genetics of adaptation from recurrent mutation or
523 migration. Mol Biol Evol. 2006;

524 56. Pennings PS, Hermisson J. Soft sweeps III: The signature of positive selection from recurrent mutation. PLoS Genet. 2006;

57. Smith JM, Haigh J. The hitch-hiking effect of a favourable gene. Genet Res. 1974;

58. Gillespie JH. A simple stochastic gene substitution model. Theor Popul Biol. 1983;

59. Unckless RL, Orr HA. The population genetics of adaptation: Multiple substitutions on a smooth fitness landscape. Genetics. 2009;

60. Wright S. Evolution of Mendelian populations. Elsevier/ Acad Press. 1931;

61. Wright S. Surfaces of Selective Value Revisited. Am Nat. 1988;

62. Berg OG. Selection intensity for codon bias and the effective population size of Escherichia coli. Genetics. 1996;

63. Wielgoss S, Schneider D, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, et al. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with escherichia coli. G3 Genes, Genomes, Genet. 2011;

64. Serres MH, Gopal S, Nahum LA, Liang P, Gaasterland T, Riley M. A functional update of the Escherichia coli K-12 genome. Genome Biol [Internet]. 2001 [cited 2021 Apr 30];2(9):research0035.1. Available from: /pmc/articles/PMC56896/

65. Hall LMC, Henderson-Begg SK. Hypermutable bacteria isolated from humans – a critical analysis. Microbiology [Internet]. 2006;152(9):2505–14. Available from: https://mic.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.29079-0

66. Gordo I, Perfeito L, Sousa A. Fitness effects of mutations in bacteria. Journal of Molecular Microbiology and Biotechnology. 2012.

67. Obolski U, Ram Y, Hadany L. Key issues review: Evolution on rugged adaptive landscapes. Reports on Progress in Physics. 2018.

68. Kauffman S, Levin S. Towards a general theory of adaptive walks on rugged landscapes. J Theor Biol [Internet]. 1987;128(1):11–45. Available from: http://www.sciencedirect.com/science/article/pii/S0022519387800292

69. De Visser JAGM, Park SC, Krug J. Exploring the effect of sex on empirical fitness landscapes. Am Nat. 2009;

70. Aguilar-Rodríguez J, Payne JL, Wagner A. A thousand empirical adaptive landscapes and their navigability. Nat Ecol &Amp; Evol [Internet]. 2017 Jan 23;1:45. Available from: https://doi.org/10.1038/s41559-016-0045

71. Aguilar-Rodríguez J, Peel L, Stella M, Wagner A, Payne JL. The architecture of an empirical genotype-phenotype map. Evolution (N Y). 2018;

72. Banerjee S, Schenkelberg CD, Jordan TB, Reimertz JM, Crone EE, Crone DE, et al. Mispacking and the Fitness Landscape of the Green Fluorescent Protein Chromophore Milieu. Biochemistry. 2017;

73. Eshel I. On the survival probability of a slightly advantageous mutant gene with a general distribution of progeny size-A branching process model. J Math Biol. 1981;

74. Levien E, Min J, Kondev J, Amir A. Non-genetic variability: survival strategy or nuisance? [Internet]. 2020. Available from: https://arxiv.org/abs/2010.05672

75. Carjaa O, Liberman U, Feldman MW. Evolution in changing environments: Modifiers of mutation, recombination, and migration. Proc Natl Acad Sci U S A. 2014;

76. Stajic D, Bank C, Gordo I. Epigenetic switching outcompetes genetic mutations during adaptation to fluctuating environments. bioRxiv [Internet]. 2021 Mar 12 [cited 2021 Jul 5];2021.03.11.434930. Available from: https://doi.org/10.1101/2021.03.11.434930

77. Chao L, Cox EC. COMPETITION BETWEEN HIGH AND LOW MUTATING STRAINS OF *ESCHERICHIA COLI*. Evolution (N Y) [Internet]. 1983 Jan 1 [cited 2021 Apr 28];37(1):125–34. Available from: http://doi.wiley.com/10.1111/j.1558-5646.1983.tb05521.x

78. You ST, Jhou YT, Kao CF, Leu JY. Experimental evolution reveals a general role for the methyltransferase Hmt1 in noise buffering. PLoS Biol. 2019;

79. Gordo I, Dionisio F. Nonequilibrium model for estimating parameters of deleterious mutations. Phys Rev E - Stat Nonlinear, Soft Matter Phys. 2005;

80. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for

568 scientific computing in Python. Nat Methods [Internet]. 2020 Mar 1 [cited 2021 May 21];17(3):261–72. Available from:
569 https://doi.org/10.1038/s41592-019-0686-2

570 81. Chevin LM. On measuring selection in experimental evolution. Biol Lett [Internet]. 2011 Apr 23 [cited 2021 Apr 29];7(2):210–
571 3. Available from: https://royalsocietypublishing.org/

572