

Population heterogeneity in mutation rate increases mean fitness and the frequency of higher order mutants

Helen K. Alexander[†], Stephanie I. Mayer[†], Sebastian Bonhoeffer

March 22, 2016

[†] joint first authors

Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zürich,
CH-8092 Zürich, Switzerland

Running title: Mutation rate heterogeneity

Keywords: mutator, transient mutagenesis, cancer, pathogen, standing genetic variation

Corresponding author: Helen K. Alexander, Universitätstrasse 16, CH-8092 Zürich, Switzerland; +41 44 6336033; helen.kirsty.alexander@alumni.ethz.ch

Abstract

Mutation rate is a crucial evolutionary parameter that has typically been treated as a constant in population genetic analyses. However, mutation rate is likely to vary among co-existing individuals within a population, due to genetic polymorphisms, heterogeneous environmental influences, and random physiological fluctuations. We explore the consequences of such mutation rate heterogeneity in a model allowing an arbitrary distribution of mutation rate among individuals, either with or without inheritance. We find that variation of mutation rate about the mean results in a higher probability of producing zero or many simultaneous mutations on a genome. Moreover, it increases the frequency of higher order mutants even under ongoing mutation and selection. We gain a quantitative understanding of how this frequency depends on moments of the mutation rate distribution and selection coefficients. In particular, in a two-locus model, heterogeneity leads to a relative increase in double mutant frequency proportional to the squared coefficient of variation of the mutation rate. Relative effect sizes increase with the number of loci. Finally, this clustering of deleterious mutations into fewer individuals results in a higher population mean fitness. Our results imply that mutation rate heterogeneity allows a population to maintain a higher level of adaptedness to its current environment, while simultaneously harboring greater genetic diversity in the standing variation, which could be crucial for future adaptation to a new environment. Our results also have implications for interpreting mutation rate estimates and mutant frequencies in data.

1 Introduction

The mutation rate is a key evolutionary parameter that affects the level of genetic diversity in a population. Genetic diversity in turn affects both the population's current mean fitness and its capacity to adapt to changes in the environment. Most theoretical work to date has assumed that the mutation rate takes on a fixed value in all members of the population. Nonetheless, the mutation rate, like any other trait, can be expected to vary among individuals, due to genetic, environmental, and stochastic effects. The recognition that mutation rate can vary within a population is implicit in the long-standing study of mutation rate evolution, and more recently in considerations of transient or stress-induced mutagenesis, especially in bacteria. However, a comprehensive conceptual understanding of how mutation rate heterogeneity within a population affects the level of standing genetic variation is lacking.

The existence of rare individuals with high mutation rate could be particularly important when a combination of several mutations is relevant for adaptation [62, 8, 23, 22]. Given that the mutation rate is typically low, higher order mutants are generally rare, yet they can be crucial for adaptation to complex new environments. For example, when multiple drugs are applied in combination – a common treatment approach for cancer [1] and several major infectious diseases, including tuberculosis, malaria, and HIV [32] – resistance in the targeted pathogens/cells generally requires multiple mutations. The prevalence of such mutations in the “pre-existing” or standing genetic variation before drug treatment starts, when they are generally expected to be deleterious, is predicted to be crucial to the emergence of resistance during treatment [67, 41]. Multiple mutations are also involved in the initiation and progression of many cancers [39].

There is clearly a genetic contribution to mutation rate via genes involved in replication, proofreading, and repair of the genetic material. This can result in variation of mutation rate even among closely related individuals. Laboratory investigations have

27 identified “mutators” and “antimutators” (having higher, resp. lower mutation rate than
28 the wild type), attributable to one or few specific genetic changes, in a variety of organ-
29 isms. Effect sizes range up to hundreds- to thousands-fold variation in eukaryotic cells,
30 bacteria and DNA viruses, and up to around five-fold in retro- and RNA viruses (details
31 in Supplementary Text I.1).

32 Though these studies indicate the scope for variation, the abundance of such variants
33 in natural populations is less clear. Mutators are expected to arise frequently *de novo*
34 due to the large target size for mutations causing defects in replication or repair genes
35 [21, 19]. Theoretically, under constant conditions, alleles that alter mutation rate can be
36 expected at mutation-selection balance in the long term [35, 19, 50, 20]. Moreover, by
37 hitchhiking with beneficial alleles they generate during phases of adaptation, mutators
38 may rise to higher frequency in the short term [75, 45, 20]. In experimental populations
39 of bacteria, mutators have indeed been observed to spontaneously arise and persist [73]
40 and be enriched through selective sweeps [55], and some parameters of these processes can
41 be estimated [9]. Surveys of clinical and other natural isolates in several bacterial species
42 indicate that strains exhibiting a range of mutation rates also exist outside the laboratory
43 [42, 56, 63, 7, 18, 68, 59, 64, 3, 78]. In RNA virus populations, mutators appear rapidly
44 in laboratory settings [74, 13] and are expected to be present in heterogeneous natural
45 populations [74, 52], but we are not aware of any surveys of natural isolates. Cancerous
46 tumors, which are characteristically genetically unstable and highly heterogeneous [44,
47 31, 4], are also anticipated to be polymorphic in genes affecting mutation rate. It has
48 been hypothesized that a mutator phenotype arises early in carcinogenesis, and moreover
49 increases the chances of successive mutations affecting genomic stability, leading to further
50 non-uniform increases in mutation rate [49, 46, 47, 48]. However, in no case does there
51 appear to be a study quantifying mutation rate in a representative sample of co-existing
52 individuals from a single population (within one infected patient or tumor).

53 Many environmental factors – including temperature, pH, oxygenation, UV radiation,
54 and chemicals – have also been implicated in modulating mutagenesis in bacteria, viruses,
55 and cancerous cells (details in Suppl. Text I.2). Viral mutation rate could also be affected
56 by its host cell’s type, physiological state, and antiviral defenses. However, few quanti-
57 tative estimates relating environmental variables to mutation rate are available. Some
58 antibiotics appear to increase bacterial mutation rates by 2- to around 100-fold [30, 40],
59 while certain antiretrovirals increase HIV-1 mutation rate by roughly five-fold [53]. While
60 it is clear that the relevant environmental factors may be heterogeneously distributed in
61 a population’s habitat, inducing different mutation rates in coexisting individuals, the
62 precise distribution will be highly context-dependent.

63 Finally, mutation rate may vary randomly and non-systematically in a population,
64 due to stochastic effects on individuals’ physiological states [9, 22]. For example, the SOS
65 response, which is associated with production of error-prone polymerases in bacteria [76],
66 exhibited a distribution of induction levels in wild type *E. coli* K12, including 0.3% of
67 the population at least 20-fold above the average level at a given time [57]. Even consti-
68 tutively expressed replication/repair genes are subject to random errors in transcription
69 and translation that affect the protein’s fidelity [62, 8, 58]. Rough calculations suggested
70 that bacterial populations contain resulting “transient mutators” at a total frequency of
71 around 5×10^{-4} , with mutation rates expected to be enhanced to similar degrees as in
72 genetic mutators [62]. Fluctuations in low copy number proteins, particularly upon cell
73 division, could also yield temporary reduction in repair capacity [22], and imbalanced
74 concentrations of protein subunits could produce polymerases missing the proofreading
75 subunit [2]. Thus, even isogenic populations in uniform macroenvironments seem likely
76 to contain individuals with differing propensities to generate mutations, although the few
77 tests to date have yielded mixed results [27, 37].

78 Taken together, this evidence suggests that mutation rate variation within populations

79 is probably common, though there are few direct quantitative estimates. DNA-based or-
80 ganisms appear to have the capacity to vary mutation rate over a few orders of magnitude,
81 while RNA-based viruses appear to tolerate only modest (up to around five-fold) changes
82 in their already high baseline mutation rates [25, 52]. The frequency of mutators in a
83 population could vary widely depending on the source of mutation rate variation and
84 the selective conditions. Furthermore, a broad spectrum arises in the extent to which
85 mutation rate is potentially correlated between parent and offspring. At one extreme, if
86 mutation rate is entirely genetically controlled, the offspring will inherit its parent's muta-
87 tion rate. At the other extreme, erroneously translated polymerases or other intracellular
88 components will have limited if any inter-generational effects before degrading and/or be-
89 ing diluted by new production. If mutation rate is primarily determined by the external
90 environment, parent-offspring correlation could vary over a broad range, depending on the
91 extent to which they share a common environment. If spatial variation in the environment
92 is fine-grained (relative to the typical offspring dispersal distance), correlation will be low,
93 while if variation is coarse-grained, parent and offspring are likely to experience the same
94 environment and thus mutation rate.

95 A large body of theoretical work on evolution of mutation rate takes into account the
96 existence of heritable mutation rate variants (reviewed by [72]). Though the majority
97 of this literature assumes the mutation rate is constitutive, evolution of stress-induced
98 mutagenesis has also been considered [6, 65, 66]. A key factor considered to drive evolution
99 of mutation rate is indirect selection through linkage to other loci that affect fitness, but
100 the focus of these studies is on the dynamics of the mutator allele itself. Far fewer studies
101 have considered how the existence of mutation rate variability in the population, regardless
102 of its source, affects mutational dynamics at other loci [62, 29, 11, 2, 33, 50]. Moreover,
103 these existing models are mostly designed for particular populations and mechanisms of
104 variation, and allow only two possible values of mutation rate.

105 In the present study, we develop a more general theoretical framework to understand
106 the effects of population heterogeneity in mutation rate on the appearance of new genetic
107 variants at one or more loci, and, in conjunction with fitness, the long-term frequency at
108 which these variants are present. That is, we address not only the production of mutants
109 in a single round of replication, but also the temporal dynamics of deleterious mutants
110 under ongoing production and selection. We consider haploid, asexually reproducing in-
111 dividuals, which is a reasonable first approach for many disease-causing microbial and
112 cellular populations of interest, including bacteria (neglecting horizontal gene transfer in
113 some species), viruses (neglecting complementation and in some cases recombination),
114 and cancerous cells (neglecting dominance effects). Our approach allows an arbitrary
115 distribution of mutation rate among individuals, and considers how moments of this dis-
116 tribution and the degree to which mutation rate is inherited affect the population-level
117 frequency of mutants at focal fitness-determining loci. We do not make any assumption as
118 to the biological mechanism underlying this heterogeneity, in particular whether it is an
119 adaptive/regulated response or an unavoidable byproduct of random processes or external
120 environmental factors.

121 We find that variability of the mutation rate about the mean has no effect on single
122 point mutants, but boosts the frequency of higher order mutants, with increasingly large
123 relative effects. Through analytical approximations we gain a quantitative understand-
124 ing, elucidating in particular for a two-locus model that the increase in double mutant
125 frequency is proportional to the variance in mutation rate and depends on the fitness of all
126 mutants. Inheritance of mutation rate strengthens the effect of population heterogeneity,
127 especially when stepwise accumulation of multiple mutations is an important pathway.
128 Finally, we show that a population maintaining a range of mutation rates (whether or
129 not these are inherited) achieves a higher mean fitness than a population in which all
130 individuals have an identical mutation rate.

2 Methods

We model a haploid, asexually reproducing population with non-overlapping generations, and extend classic population genetic models to incorporate a mutation rate that varies among co-existing population members. We focus on genotype dynamics at one or more fitness-determining loci (each of which may consist of one or several base pairs), and assume throughout that mutation rate neither depends on the genotype at the focal loci, nor has any direct fitness effect itself.

Data availability: R code used to generate numerical results is available upon request.

2.1 Occurrence of mutations on a genome in one generation

We consider n loci of interest on the genome, which can be either non-mutant or mutant. We assume that each individual has a given mutation rate u (per locus, per generation) that is uniform across loci; that is, each non-mutant locus in the individual mutates independently with probability u . We neglect back mutations. The number of new mutations that arise thus follows a binomial distribution; in particular, if n loci are non-mutant, then the probability of j mutations occurring simultaneously (i.e. in one generation) is:

$$p_{n,j}(u) = \binom{n}{j} u^j (1-u)^{n-j} \quad (1)$$

In the limit as $n \rightarrow \infty$ and $u \rightarrow 0$ such that $nu \equiv \lambda$, we obtain an “infinite-locus” model in which every mutation occurs at a unique site. Then the number of new mutations that arise in an individual with mutation rate λ (per genome, per generation) follows a Poisson distribution; that is, the probability of j simultaneous mutations is:

$$p_j(\lambda) = e^{-\lambda} \lambda^j / j! \quad (2)$$

151 We note that viruses can have complex, multi-step intracellular replication cycles, which
152 imply that one cycle of cell infection cannot be equated to one genome replication, and
153 therefore a Poisson-distributed number of mutations per genome is not necessarily ex-
154 pected after a single infection cycle [26, 71]. Our present model does not address these
155 complexities.

156 The key novelty in our model is to consider a mutation rate (u or λ) that varies among
157 individuals in any given generation, and can thus be taken as a random variable (denoted
158 U or Λ respectively) in the population as a whole. However, we make the important
159 assumption that the *distribution* of mutation rate in the population does not change over
160 generations.

161 **2.2 Inherited versus non-inherited mutation rate**

162 Once we consider dynamics over more than one generation, we must define the extent to
163 which mutation rate is inherited or correlated from parent to offspring. As described in
164 the Introduction, this correlation could vary over a broad spectrum. Mathematically, we
165 will deal with the two extremes.

166 In the case of no inheritance, each individual in each generation independently draws its
167 mutation rate from the population distribution. Thus one must average the probability or
168 proportion of individuals in the new generation mutating from genotype i to j , conditioned
169 on mutation rate, over the distribution of mutation rate, which is arbitrary but fixed over
170 generations.

171 In the case of perfect inheritance, each individual takes on exactly the same mutation
172 rate as its parent. Thus, the population can be divided into subpopulations character-
173 ized by distinct mutation rates, with no “migration” among subpopulations. Assuming
174 that the subpopulations do not interact, we can describe the population dynamics in each
175 subpopulation separately using a standard model with fixed mutation rate, before finally
176 taking a weighted average of the quantity of interest over subpopulations. If population

177 size regulation acts on the population as a whole, the subpopulations are not truly inde-
178 pendent in their population dynamics. In this situation, indirect selection on mutation
179 rate (due to linkage with focal loci) arises, and if mutations are always deleterious, the
180 lowest mutation rate will be favored [72]. Nonetheless, in line with our aforementioned
181 assumption that the mutation rate distribution does not change over time, we will neglect
182 this predicted evolution of mutation rate, in that we impose selection within each subpop-
183 ulation independently. Since we will consider selection coefficients at the focal loci that
184 are much larger than any subpopulation’s mutation rate, genotype frequency dynamics at
185 the focal loci are expected to occur on a faster timescale than the evolution of mutation
186 rate (cf. [72, 20]), so our approach should provide a reasonable approximation on this
187 faster timescale.

188 **2.3 Genotype frequency dynamics under mutation and se-** 189 **lection**

190 We now derive deterministic recursions describing the change in frequency of each geno-
191 type from one generation to the next, incorporating both mutation and selection. As
192 a basis, our model adopts a standard formulation in discrete population genetic analy-
193 ses. Genotypes are defined by the presence/absence of mutations at the focal (fitness-
194 determining) loci. We denote the frequency of genotype i at generation t by $x_i(t)$ and
195 its relative fitness by w_i . Without loss of generality we take the wild type (carrying no
196 mutations) to have relative fitness 1. We assume that mutations are deleterious, thus
197 $w_i = 1 - s_i$ where the selection coefficients satisfy $0 < s_i \leq 1$ for all types i other than
198 the wild type. (Some results below are also valid for the neutral case, $s_i = 0$, but certain
199 approximations will break down when s_i is too small.) The population mean fitness at
200 time t is given by $\bar{w}(t) := \sum_i w_i x_i(t)$. Census occurs immediately after mutation, before
201 selection (i.e. relative fitness determines the total reproductive output of parents, offspring
202 mutate independently of one another, then they are counted). Generally, then, for any

203 collection of types i and proportions p_{ij} of type i offspring from a parent of type j , one
204 can write a set of recursions:

$$x_i(t+1) = \sum_{\forall j} \frac{w_j p_{ij} x_j(t)}{\bar{w}(t)}$$

205 Note that these equations describe the change in genotype frequencies even in a popula-
206 tion with changing total size, or absolute fitnesses that change over time, as long as the
207 *relative* fitnesses of the types are constant [16, p. 278]. The incorporation of mutation rate
208 heterogeneity, according to the considerations of sections 2.1 and 2.2, essentially lies in
209 the structuring of the population and in the specification of mutation probabilities $\{p_{ij}\}$.

210 **2.3.1 Finite loci**

211 Considering n biallelic loci yields 2^n types, identified by binary notation indicating absence
212 (0) or presence (1) of a mutation at each locus. For finite n , we closely follow the exposition
213 and basic results for fixed mutation rate given by Bürger [10, Ch. III.1.1]. The above
214 recursions can be rewritten as a matrix equation:

$$x(t+1) = \frac{1}{\bar{w}(t)} Mx(t) \tag{3}$$

215 where $x(t)$ collects the frequencies of each genotype at time t into a vector, and M is the
216 $2^n \times 2^n$ “mutation-selection matrix” (independent of time) where $M_{ij} = w_j p_{ji}$.

217 Given a mutation-selection matrix M and an initial frequency vector $x(0)$, the popu-
218 lation mean fitness can be written as

$$\bar{w}(t) = \sum_j (Mx(t))_j$$

219 and the solution of the recursion is then given by

$$x(t) = \frac{M^t x(0)}{\sum_j (M^t x(0))_j} \tag{4}$$

220 The equilibrium frequency solutions (denoted x^*) are given by the eigenvectors of M ,
 221 normalized so the entries add up to one, and the population mean fitness at equilibrium
 222 (\bar{w}^*) is given by the corresponding eigenvalues. Since we neglect back mutation, M will
 223 always be triangular, and thus the eigenvalues are simply the diagonal entries. The reduc-
 224 tion in fitness compared to a homogeneous wild type population due to the production of
 225 deleterious mutants, i.e. $1 - \bar{w}^*$, is known as the “mutational load” [10, p. 105].

226 Given the binomial mutation model (Section 2.1), the mutation probabilities between
 227 types take the form:

$$p_{ij} = \binom{n_i}{n_j} U^{n_i - n_j} (1 - U)^{n_j}$$

228 where U is the per-locus mutation rate and n_i (resp. n_j) is the number of non-mutant loci
 229 in type i (resp. j). We will write the mutation-selection matrix as $M(U)$ to emphasize its
 230 dependence on U . For instance for one locus,

$$M(U) = \begin{pmatrix} 1 - U & 0 \\ U & 1 - s \end{pmatrix} \quad (5)$$

231 while for two loci,

$$M(U) = \begin{pmatrix} (1 - U)^2 & 0 & 0 & 0 \\ U(1 - U) & (1 - U)(1 - s_{01}) & 0 & 0 \\ U(1 - U) & 0 & (1 - U)(1 - s_{10}) & 0 \\ U^2 & U(1 - s_{01}) & U(1 - s_{10}) & (1 - s_{11}) \end{pmatrix} \quad (6)$$

232 We now incorporate mutation rate heterogeneity at the population level. Although we
 233 phrase the following exposition in probabilistic terms, in the present deterministic model
 234 we ultimately treat the probability of a type j parent producing a type i offspring as the
 235 exact proportion of such events. As well as neglecting demographic stochasticity, which
 236 is standard in this modeling approach, we neglect sampling effects from the underlying
 237 mutation rate distribution, which is reasonable if the population is large.

238 **Non-inherited mutation rate:** Each individual (i.e. each offspring) independently
239 draws its per-locus mutation rate from the distribution of U , and conditioned on the mu-
240 tation rate, the number of non-mutant loci that mutate is binomially distributed (Section
241 2.1). Applying the Law of Total Expectation, the expected overall contribution of type j
242 parents to type i offspring is thus given by $\mathbb{E}_U[M_{ij}(U)]x_j/\bar{w}$. This yields the recursion:

$$x(t+1) = \frac{1}{\bar{w}(t)} \mathbb{E}_U[M(U)]x(t) \quad (7)$$

243 where the expectation over U is applied entry-wise to the matrix $M(U)$.

244 **Perfectly inherited mutation rate:** The population is divided into d disconnected
245 subpopulations, where the k^{th} subpopulation is at frequency q_k (with $\sum_{k=1}^d q_k = 1$) and
246 is characterized by mutation rate u_k that is fixed for all individuals within the subpop-
247 ulation. Equivalently, the mutation rate distribution in the entire population is given
248 by the probability mass q_k at value u_k . Neglecting long-term mutation rate evolution as
249 explained in Section 2.2, we independently solve the recursion in each subpopulation:

$$x^{(k)}(t+1) = \frac{1}{\bar{w}^{(k)}(t)} M(u_k)x^{(k)}(t) \quad (8)$$

250 where $x^{(k)}$ is the genotype frequency vector in the k^{th} subpopulation and $\bar{w}^{(k)}(t)$ is the
251 mean relative fitness calculated only within the subpopulation. The population-wide fre-
252 quencies are finally obtained by averaging over subpopulations, i.e. taking the expectation
253 of the fixed-rate results over the distribution of U :

$$x(t) = \sum_{k=1}^d q_k x^{(k)}(t) \quad (9)$$

254

2.3.2 Infinite loci

255

In the infinite-locus limit, we adopt the model of [38], which assumes that fitness is fully

256

determined by the number of mutations carried. Let $x_i(t)$ be the frequency of i -point

257

mutants in generation t ; w_i be the relative fitness of i -point mutants; and λ be the mean

258

number of mutations per genome per generation. The only assumption about the fitness

259

values thus far is that all mutants are less fit than the wild type ($w_i < w_0 = 1$ for $i \neq 0$).

260

Then the recursions describing the dynamics of mutant frequencies over time are [38]:

$$x_i(t+1) = \sum_{j=0}^i \frac{w_{i-j}}{\bar{w}(t)} x_{i-j}(t) \frac{e^{-\lambda} \lambda^j}{j!} \quad (10)$$

261

where $\bar{w}(t) := \sum_{i=0}^{\infty} w_i x_i(t)$ is again the population mean fitness.

262

Similarly to the finite locus case, we extend this model to a distribution of mutation

263

rate under the two inheritance assumptions: If mutation rate is not inherited, then

$$x_i(t+1) = \sum_{j=0}^i \frac{w_{i-j}}{\bar{w}(t)} x_{i-j}(t) \frac{\mathbb{E}_{\Lambda}[e^{-\Lambda} \Lambda^j]}{j!} \quad (11)$$

264

while if mutation rate is perfectly inherited, with rate λ_k in the k^{th} subpopulation at

265

frequency q_k , then the mutant frequencies $\{x_i^{(k)}(t)\}$ in the k^{th} subpopulation are given by

266

Equation 10 with $\lambda \equiv \lambda_k$, and the overall mutant frequencies in the population are given

267

by

$$x_i(t) = \sum_k q_k x_i^{(k)}(t) \quad (12)$$

268

For most of our results, we will deal with a special case of the model in which $w_i =$

269

$(1-s)^i$ [34], meaning that each mutation has an equal effect (cost s) and there is no

270

epistasis.

3 Results

We are interested in the effect of mutation rate heterogeneity on the production and maintenance of mutations in a population’s standing genetic variation. We therefore focus on comparing a “heterogeneous” population, where the mutation rate (per locus, U , or per genome, Λ) has a given distribution, to a baseline “homogeneous” population with mutation rate fixed to the mean of this distribution (denoted $\langle U \rangle$ or $\langle \Lambda \rangle$, respectively). We first consider the probability distribution of the number of mutations occurring “simultaneously”, i.e. on a single genome in one generation, which is independent of their fitness effects. We then consider the temporal dynamics of genotypes over multiple generations of mutation and selection against deleterious mutations, and derive the mutation-selection balance attained when mutation rate varies among individuals.

3.1 Probability of simultaneous mutations

The probability $p_{n,j}(u)$ of j simultaneous mutations among n loci available to mutate, given a mutation rate of u per locus, is given by Equation 1. Averaging over the distribution of mutation rate, the overall probability of j simultaneous mutations is then $p_{n,j} = \mathbb{E}_U[p_{n,j}(U)]$. These probabilities can also be interpreted as the expected frequencies of j -point mutants produced (before selection) by a purely wild type starting population.

To consider the effect of a mutation rate that varies about its mean, we apply Jensen’s Inequality to the functions $p_{n,j}(U)$. If g is any real convex function of a random variable U (i.e. $g''(U) > 0$), then Jensen’s Inequality states that $\langle g(U) \rangle \geq g(\langle U \rangle)$, with equality if and only if g is linear or U takes on a fixed value [14, p. 27]. Thus, we can determine whether variability in mutation rate increases or decreases $p_{n,j} \equiv \langle p_{n,j}(U) \rangle$, relative to $p_{n,j}(\langle U \rangle)$ in the homogeneous case, by analyzing the second derivative of $p_{n,j}(U)$ for each n and j .

If we consider a single locus ($n = 1$), it is clear that the functions $p_{1,j}(U)$ are linear.

296 Thus, the overall probability of mutation at a single locus is fully determined by the mean
297 mutation rate and independent of the extent of variability: specifically $p_{1,0} = 1 - \langle U \rangle$ and
298 $p_{1,1} = \langle U \rangle$. On the other hand, if we consider multiple loci ($n \geq 2$), the functions $p_{n,j}(U)$
299 are non-linear, and in general $\langle p_{n,j}(U) \rangle \neq p_{n,j}(\langle U \rangle)$.

300 Since $p_{n,0}(U) = (1 - U)^n$ and $p_{n,n}(U) = U^n$ are clearly convex for $n \geq 2$, we can
301 conclude that the probabilities of either all or none of the loci mutating are increased by
302 variability in U . Logically, the probability of at least some intermediate numbers of muta-
303 tions must be reduced. We find (Suppl. Text II.1) that $p_{n,j}$ will generally be increased by
304 heterogeneity for the smallest and largest values of j , and decreased in some intermediate
305 range of j , with the exact switching points depending on n and on the particular distri-
306 bution of U . For realistic ranges of mutation rate in most organisms, heterogeneity will
307 increase the chance of zero or of two or more simultaneous mutations and decrease the
308 chance of a single mutation occurring, even with many loci under consideration. For cer-
309 tain RNA viruses with high mutation rates, and possibly cellular populations containing
310 strong mutators, the switching points may be shifted upward (Figure 1).

311 Summary statistics of the probability distribution behave as intuitively expected: the
312 mean number of mutations occurring simultaneously on a single genome is unaffected
313 by variability of the mutation rate about its mean, but the variance in the number of
314 mutations is increased by variance in the mutation rate distribution (exact expressions in
315 Suppl. Text II.1.2).

316 We further analyze the magnitude of the effect of heterogeneity in the case of all n
317 loci mutating simultaneously. Rewriting this probability (Suppl. Text II.1.3):

$$p_{n,n} = \langle U^n \rangle = \langle U \rangle^n + \sum_{i=0}^{n-1} \binom{n}{i} c_{n-i} \langle U \rangle^i \quad (13)$$

318 where $c_i = \langle (U - \langle U \rangle)^i \rangle$ is the i^{th} central moment of the mutation rate distribution. (In
319 particular, $c_0 = 1$, $c_1 = 0$, and c_2 is the variance.) Thus the probability of n simultaneous

320 mutations depends on the first n central moments. For instance, the “boost” in triple
321 mutations increases with the variance, and is larger when the distribution is right-skewed
322 than when it is left-skewed. Note that even-numbered central moments must be positive,
323 while odd-numbered central moments may be positive or negative; however, according to
324 Jensen’s Inequality, any negative terms must be outweighed by the positive terms. It can
325 also be shown (Suppl. Text II.1.3) that the *relative* effect of heterogeneity in mutation
326 rate on the probability of simultaneous mutation at all loci increases with the number of
327 loci under consideration (Fig. 2).

328 **3.2 Deterministic mutant frequency dynamics under muta-** 329 **tion and selection**

330 We now consider genotype frequencies over multiple generations of mutation and selection,
331 with particular attention to the equilibrium (mutation-selection balance). This determin-
332 istic approach neglects demographic stochasticity in a finite population, and importantly
333 in our extension, also neglects sampling effects from the mutation rate distribution. We
334 compare the well-known results for fixed mutation rate to our novel results for heteroge-
335 neous mutation rate. Details of the mathematical results are provided in Supplementary
336 Text II.2.

337 **3.2.1 One locus**

338 Heterogeneity in the mutation rate again turns out to have a negligible effect on mutant
339 frequency dynamics at a single focal locus. In particular, the classic mutation-selection
340 balance is simply replaced by $x_1^* = \frac{\langle U \rangle}{s}$ where $\langle U \rangle$ is the mean mutation rate in the
341 population and s is the cost of the mutation. Population mean fitness at equilibrium is
342 correspondingly given by $\bar{w}^* = 1 - \langle U \rangle$. These solutions are valid regardless of whether
343 mutation rate is inherited.

344 The full temporal dynamics of the mutant frequency are more involved, but still

345 amenable to analytical solution. In the non-inherited case, the temporal solution ex-
346 actly coincides with the homogeneous case. Mathematically, this is because the relevant
347 mutation-selection matrix is identical: due to the linearity of $M(U)$ in U (Equation 5), we
348 have $\langle M(U) \rangle = M(\langle U \rangle)$. In the inherited case, the solutions are not exactly equivalent,
349 but can be shown to coincide up to first order in the maximum mutation rate. Mathemat-
350 ically, the mutant frequency in each subpopulation at time t is nonlinear in u_k , and this
351 nonlinear expression must be averaged over the distribution of mutation rate; however,
352 higher order terms make a negligible contribution (vanishing at equilibrium).

353 **3.2.2 Multiple loci**

354 When multiple loci are involved, the mutation-selection matrix M becomes non-linear in
355 U . Specifically, in the n -locus model terms up to order U^n appear in M , and we expect the
356 highest-order mutants to have frequency with leading order U^n . Thus, the first n moments
357 of the mutation rate distribution will generally play a non-negligible role in genotype fre-
358 quency dynamics. We conduct a detailed mathematical analysis of the two-locus case,
359 while a brief consideration of the infinite-locus limit confirms our key qualitative conclu-
360 sions and suggests how results will extend to more loci. Below we focus on key results
361 and their intuitive interpretation, while detailed expressions for genotype frequencies over
362 time in all model cases are provided in Supplementary Text II.2. Throughout, V denotes
363 the variance of the mutation rate distribution and $c^2 := V/\langle U \rangle^2$ denotes the squared
364 coefficient of variation. To distinguish models under comparison, the short form ‘het’
365 will indicate a heterogeneous mutation rate characterized by a distribution, and ‘hom’ a
366 homogeneous mutation rate fixed to the mean of this distribution. Further, $H = 0$ will
367 indicate that mutation rate is non-inherited and $H = 1$ will indicate perfect inheritance.

368 **Genotype frequencies in the two-locus model** For reference, with fixed muta-
369 tion rate $U \equiv u$, the equilibrium frequencies are approximated to order u^2 by:

$$x_{00}^* \approx 1 - \left(\frac{1}{s_{01}} + \frac{1}{s_{10}} \right) u - \left(\frac{1}{s_{01}s_{11}} + \frac{1}{s_{10}s_{11}} - \frac{1}{s_{11}} - \frac{2}{s_{01}s_{10}} \right) u^2 \quad (14a)$$

$$x_{01}^* \approx \frac{u}{s_{01}} - \frac{u^2}{s_{01}s_{10}} \quad (14b)$$

$$x_{10}^* \approx \frac{u}{s_{10}} - \frac{u^2}{s_{01}s_{10}} \quad (14c)$$

$$x_{11}^* \approx \left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 1 \right) \frac{u^2}{s_{11}} \quad (14d)$$

370 and the population mean relative fitness is given by $\bar{w}^* = (1 - u)^2$.

371 When heterogeneity is introduced, we use the recursions given in Equations 7 (no
 372 inheritance) or 8-9 (perfect inheritance) to solve for the genotype frequencies. Figure 3 il-
 373 lustrates examples of the resulting temporal dynamics of double mutant frequency, $x_{11}(t)$.
 374 Our analytical approximations (Suppl. Text II.2.2) typically show excellent agreement to
 375 results obtained by numerical iteration of the recursions. In all cases, the double mutant
 376 frequency is elevated by variability of the mutation rate about the mean. This increase
 377 is larger (after the first generation) when mutation rate is perfectly inherited. However,
 378 while the non-inherited case falls substantially short for weak selection against single
 379 mutants, the two cases become similar for strong selection against single mutants. Our
 380 analytical solutions indicate that these observations are general: double mutant frequency
 381 is increased by an amount proportional to variance in mutation rate, and depends on the
 382 selection coefficients in a way that will be clarified below. The equilibrium solutions are
 383 summarized in Table 1 for reference. As a note of caution, the error in these approxima-
 384 tions scales with $\max(\langle U \rangle^3, \langle U \rangle V, V^2)$ in the case of non-inheritance and with $\max_k q_k u_k^3$
 385 in the case of perfect inheritance. Further numerical testing (not shown) indicates that
 386 the approximations can break down for extreme mutation rate distributions, particularly
 387 when at least one selection coefficient is small.

388 Heterogeneity also affects the frequencies of the wild type and single-point mutants.
 389 Generally, heterogeneity in the mutation rate has the effect of clustering mutations at

390 mutation-selection balance, i.e. increasing the frequency of the double mutant at the ex-
391 pense of single mutants, similar to the qualitative effect on the production of new muta-
392 tions in each generation. More specifically, if mutation rate is not inherited, heterogeneity
393 decreases the equilibrium frequency of single mutants and increases that of the wild type.
394 If mutation rate is perfectly inherited, heterogeneity still decreases the equilibrium fre-
395 quency of single mutants, but interestingly, the effect on the wild type can take either
396 direction. Namely, x_{00}^* is decreased when epistasis is sufficiently strongly positive, in which
397 case the additional double mutants appear to exert sufficient competition on the wild type
398 to outweigh the increased chance of mutation-free reproduction in each generation.

399 **Understanding the roles of mutation rate inheritance and simultaneous**

400 **mutations:** Our results can be understood by considering which of the underlying
401 mutational pathways to the double mutant are affected by heterogeneity in the mutation
402 rate. In the absence of inheritance, the mutation rate experienced by multiple loci on
403 a genome only shows an association over one generation. Any boost in multiple mutant
404 frequency due to mutation rate heterogeneity must thus be achieved through a boost in
405 simultaneous mutations. In contrast, effects of heterogeneity can act across generations
406 when the mutation rate is inherited. Then multiple mutants can be boosted not only
407 by simultaneous mutations, but also by stepwise accumulation of mutations over several
408 generations.

409 To confirm this reasoning we consider a variant of the mutation model, where simulta-
410 neous double mutations are not allowed. Then higher-order terms in U no longer appear
411 in the mutation-selection matrix (Equation S10 in Suppl. Text II.2.2.5), indicating that
412 in the absence of inheritance, only the mean mutation rate matters. This is in line with
413 our insight that all effects of non-inherited variation in mutation rate must act via the
414 eliminated simultaneous mutations. On the other hand, when mutation rate is perfectly
415 inherited, variance can still have an effect, albeit reduced, via stepwise accumulation of

416 mutations. Mathematically, multiplication of the mutation-selection matrix (i.e. iteration
417 over generations) gives rise to higher-order terms in U before the result is averaged over
418 the mutation rate distribution.

419 We are now in a position to interpret the increases in double mutant frequency due to
420 mutation rate heterogeneity. We denote the absolute increase by

$$\Delta_{\text{abs}}(t) := x_{11}(t; \text{het}) - x_{11}(t; \text{hom})$$

421 and the relative increase by

$$\Delta_{\text{rel}}(t) := (x_{11}(t; \text{het}) - x_{11}(t; \text{hom})) / x_{11}(t; \text{hom})$$

422 It turns out that $\Delta_{\text{abs}}(t)$ is proportional to the variance (V) and $\Delta_{\text{rel}}(t)$ is proportional
423 to the squared coefficient of variation ($c^2 = V / \langle U \rangle^2$) at all times t (Suppl. Text II.2.3.1).

424 With perfect inheritance of mutation rates, mutation rate heterogeneity affects all
425 pathways equally. Thus selection plays the same role as in the homogeneous case and the
426 double mutant frequency is simply scaled up by a constant factor:

$$\Delta_{\text{rel}}(t; H = 1) = c^2$$

427 independent of t and $\{s_i\}$. Plotting the equilibrium frequency x_{11}^* as a function of single
428 mutant cost, this effect manifests as parallel curves in the homogeneous and heteroge-
429 neous cases (Fig. 4). When single mutants are sufficiently fit, stepwise accumulation of
430 mutations is the main pathway, and blocking simultaneous mutations has little effect on
431 the double mutant frequency; however, when single mutants are very costly, blocking si-
432 multaneous mutations has a drastic effect. The precise ranking of x_{11}^* all model cases, as
433 determined by $\{s_i\}$ and c^2 , is given in Supplementary Text II.2.3.2. Importantly, since
434 the perfectly inherited case yields the maximal increase compared to the homogeneous
435 case, we can conclude that c^2 in fact provides an upper bound on the relative increase in
436 double mutant frequency that can be achieved by mutation rate heterogeneity across all

437 choices of selection coefficients, times (since starting from a wild type population), and
438 inheritance assumptions.

439 In the non-inherited case, the precise increase in double mutant frequency depends
440 on the selection coefficients. (We focus here on the equilibrium; transient dynamics are
441 analyzed in Suppl. Text II.2.3.1.) The absolute increase

$$\Delta_{\text{abs}}^*(H = 0) = V/s_{11}$$

442 is due entirely to the increased influx of simultaneous double mutations, which are filtered
443 by selection according to coefficient s_{11} . The selection coefficients of the single mutants do
444 not appear, because stepwise accumulation of mutations cannot be affected by variance
445 in mutation rate in this case. However, the relative increase

$$\Delta_{\text{rel}}(H = 0) = c^2 / \left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 1 \right)$$

446 increases with s_{01} and s_{10} , because the stronger the selection against single mutants, the
447 fewer the double mutants can be produced via these intermediates, and the more relatively
448 important simultaneous double mutation becomes. In Figure 4, the heterogeneous non-
449 inherited case thus approaches the homogeneous case as $s_{01} = s_{10} \rightarrow 0$, and approaches the
450 heterogeneous perfectly inherited case as $s_{01} = s_{10} \rightarrow 1$, where all double mutants must
451 be generated directly by simultaneous mutation from the wild type. When simultaneous
452 mutation is blocked, the heterogeneous non-inherited case drops to align perfectly with
453 the homogeneous case over the full range of single mutant costs.

454 **Genotype frequencies in the infinite-locus model:** The two-locus results are
455 supported and extended by considering an infinite-locus model. For simplicity, we consider
456 the special case in which each individual mutation has an equal fitness effect (cost s) and
457 there is no epistasis. In this case, for fixed per-genome mutation rate λ , the equilibrium
458 frequency of i -point mutants is given by [34]:

$$x_i^* = e^{-\lambda/s} (\lambda/s)^i / i! \quad (15)$$

459 The equilibrium frequencies thus take the same form as the probabilities of simultaneous
460 mutation, namely a Poisson distribution, but now with parameter λ/s instead of λ .

461 With a heterogeneous mutation rate, in the case of perfect inheritance, one must sim-
462 ply average the fixed-rate solution over the distribution of Λ . Thus an analysis of the
463 effect of heterogeneity, applying Jensen's Inequality, proceeds exactly as in the simulta-
464 neous mutation analysis: Heterogeneity in mutation rate has the effect of increasing the
465 frequency of zero- or possibly few-point mutants, decreasing the frequency of intermediate
466 mutants, and increasing the frequency of many-point mutants. Moreover, decreasing s has
467 the same effect as increasing the mutation rate, namely shifting upwards the values of i
468 at which the directional effect of heterogeneity switches sign. In the case of non-inherited
469 mutation rate, analytical progress does not appear straightforward, but we intuitively
470 expect the same qualitative effect. We confirm this pattern by numerical iteration of the
471 recursions for both cases (Equations 10-12). The switching points in the directional effect
472 appear as predicted in both cases, but are shifted upward when mutation rate is non-
473 inherited (Fig. 5). Furthermore, among the higher mutant classes where heterogeneity
474 boosts mutant frequency, the relative magnitude of this effect increases with number of
475 mutations, as also seen in the simultaneous mutation results.

476 Thus, in qualitative agreement with the two-locus results, heterogeneity has the effect
477 of clustering mutations at mutation-selection balance. Furthermore, heterogeneity again
478 appears to have a larger impact on the mutant frequency distribution when the mutation
479 rate is inherited. More specifically, there is a larger increase in the frequency of the
480 lowest- and highest-order mutants (and correspondingly larger decrease in the frequency
481 of intermediates). Again, the relative importance of inheritance varies with the strength of
482 selection: the non-inherited case appears more similar to the homogeneous case when s is

483 smaller, and gradually becomes more similar to the perfectly-inherited case as s increases
484 (Fig. S1).

485 Clearly, increasing the variance of the mutation rate distribution generally increases
486 the effects of heterogeneity (Fig. S1). However, the differences in mutant frequencies are
487 no longer directly proportional to variance, since all higher moments of the mutation rate
488 distribution also affect the results in the infinite-locus case.

489 **Frequency of mutant alleles and average number of mutations per genome:**

490 We have seen that heterogeneity in mutation rate clusters mutations at multiple loci. We
491 now ask whether it has any effect on the frequency of the mutant allele at any given
492 locus ℓ (f_ℓ^*) or on the average number of mutations carried per genome ($m^* = \sum_\ell f_\ell^*$) at
493 equilibrium.

494 In the two-locus model, it turns out that both these quantities differ between the het-
495 erogeneous and homogeneous cases by an amount proportional to the variance in mutation
496 rate. However, the magnitude and direction of the effect depends on epistasis and muta-
497 tion rate inheritance (details in Suppl. Text II.2.2.4). With perfect inheritance, variance
498 increases the frequency of any given mutant allele if and only if epistasis (ϵ) is positive.
499 Intuitively, since variance in the mutation rate increases double mutant frequency at the
500 expense of single mutants, the overall frequency of the mutant allele at any individual
501 locus will be elevated precisely when the double mutant is fitter than expected from the
502 singles. Without inheritance, the frequency of the mutant allele is increased by variance
503 when the double mutant is fitter than the single mutant, which can be translated into a
504 more stringent condition on epistasis than in the perfect-inheritance case ($\epsilon > \epsilon_c > 0$).
505 Similar effects arise for m^* .

506 In the infinite-locus model, every mutation occurs at a unique site, so it no longer
507 makes sense to consider f_ℓ^* , but we can still consider m^* . We analyze only the special
508 case of the model where every mutation has equal cost and there is no epistasis. Then

509 consistent with the two-locus finding, in the case of perfect inheritance m^* is not affected
510 by heterogeneity in the mutation rate (Suppl. Text II.2.4.2). In the case of no inheritance,
511 extrapolating from the two-locus model, we expect that m^* is reduced by heterogeneity
512 in the absence of epistasis. Although we lack an analytical result, this indeed appears
513 numerically to be the case (confirmed for the cases illustrated in Figures 5 and S1).

514 **Mutational load:** In the two-locus model, the population mean fitness at equilibrium
515 is given by:

$$\bar{w}^* = \mathbb{E}_U [(1 - U)^2] = (1 - \langle U \rangle)^2 + V \quad (16)$$

516 regardless of whether mutation rate is inherited (Suppl. Text II.2.2). Thus, mutational
517 load is decreased by an amount equal to the variance of mutation rate in the population.
518 This effect can be explained by the clustering of mutations into fewer individuals.

519 In the limiting infinite-locus model in which fitness is determined by number of muta-
520 tions (with otherwise arbitrary fitness costs), the population mean fitness at equilibrium
521 is given by

$$\bar{w}^* = 1 - \langle e^{-\Lambda} \rangle \quad (17)$$

522 again regardless of inheritance. An application of Jensen's Inequality demonstrates that
523 \bar{w}^* is likewise enhanced, i.e. mutational load is reduced, by variability in the mutation
524 rate.

525 4 Discussion

526 4.1 Significance

527 The critical role of mutations in producing the raw material for evolution has long been
528 recognized by biologists and mathematically analyzed by population geneticists. In par-

529 ticular, the maintenance of standing genetic variation is known to be shaped by both
530 mutation rate and selection coefficients. While the distribution of mutational fitness ef-
531 fects has been intensively studied, the distribution of mutation rates among members of
532 a population has received far less attention. Heterogeneity in mutation rate could stem
533 from a wide range of sources, including genetic differences, environmental influences, and
534 random physiological fluctuations. In this study, we showed that such heterogeneity has
535 potentially far-reaching evolutionary consequences by affecting the chance of multi-point
536 mutants appearing, the frequency of mutant genotypes harbored in the standing genetic
537 variation, and the population mean fitness. Our general modeling approach allowed an
538 arbitrary distribution of mutation rate, clearly separated the role of variation about the
539 mean from effects on the mean itself, and compared inherited versus non-inherited forms
540 of variation. Our first key finding is that mutation rate heterogeneity results in an over-
541 representation of higher-order mutants and an under-representation of intermediate mu-
542 tants. We gained a quantitative understanding of these effects as a function of moments
543 of the mutation rate distribution and selection coefficients of the mutants. Our second key
544 finding is that, due to this clustering of mutations into fewer individuals, the population
545 as a whole has a reduced mutational load.

546 The qualitative effect that variability in mutation rate among individuals should result
547 in an over-representation of multiple-point mutants has previously been recognized [62,
548 23, 22, 27], but our rigorous analysis of the quantitative conditions for this effect to
549 occur is novel. Moreover, while previous calculations of the contribution of mutators
550 have implicitly only considered simultaneous mutations [62, 11, 23], we also analyzed the
551 effects on accumulated mutant frequency when both *de novo* mutation and selection act
552 over time. We thus showed that heterogeneity in mutation rate increases not only the
553 probability of multiple mutations arising simultaneously on a genome (Section 3.1), but
554 also the frequency of multiple mutants at mutation-selection balance (Section 3.2). In

555 Supplementary Text III, we additionally show numerically that heterogeneity reduces the
556 waiting time for the first appearance of double mutants in a stochastic branching process
557 model.

558 Interestingly, this skewing of the mutant frequencies leads to the appearance of posi-
559 tive linkage disequilibrium. If detected empirically, the standard interpretation of such an
560 observation would be the existence of epistasis. However, our results show that an alterna-
561 tive explanation is the existence of population heterogeneity in mutation rate. Drake and
562 colleagues observed that multiple mutants often appear to be over-represented (relative
563 to a Poisson distribution) in mutant counts from experiments, and indeed interpreted this
564 finding as a sign of heterogeneity in the mutation rate [23, 22]. However, it was unclear
565 whether selection could be ruled out in all the reported experiments, and thus whether
566 epistatic fitness effects might also have boosted double mutant frequency. An interesting
567 direction for future work would be to disentangle the effects of mutation rate heterogeneity
568 and epistasis, with a rigorous statistical analysis of the existing experimental findings.

569 We derived expressions for mutant frequency that both indicate the key parameters at
570 play and allow a quantitative estimation of their effects. The dynamics of mutants at n
571 focal loci are driven by the first n moments of the mutation rate distribution. Thus if we
572 examine only one locus, the population mean mutation rate is sufficient to predict mutant
573 dynamics, but to predict the joint dynamics at two loci, variance must be considered, and
574 so on. Specifically, for two loci, we found that the frequency of double mutants is boosted
575 (compared to a population with mutation rate fixed to the same mean) by an absolute
576 amount proportional to the variance, or a relative amount proportional to (and at most
577 equal to) the squared coefficient of variation. Variance likewise appeared to determine the
578 reduction in waiting time for double mutants in the stochastic model (Suppl. Text III.2
579 and Fig. S4). The quantitative effects can be substantial. For example, parameterizing
580 from data on mutation rates to rifampicin resistance in bacteria [43], the presence of hy-

581 permutator individuals with 200-fold increased mutation rate at 1% frequency is predicted
582 to increase the equilibrium frequency of double mutants by ~ 9 -fold through effects on the
583 mean, but up to ~ 45 -fold further through effects on the variance at fixed mean (Fig. 3 and
584 analytical solutions in Section 3.2). On the other hand, more evenly distributed mutation
585 rates result in smaller effects; for instance, using the best-fitting per base pair mutation
586 rate distribution inferred in a population of budding yeast [37], variance increases double
587 mutant frequency by maximally 1.4-fold relative to a homogeneous population with the
588 mean mutation rate.

589 Moreover, we predict that mutation rate heterogeneity becomes increasingly signifi-
590 cant with more loci under consideration, in that the relative boost in frequency is larger
591 for higher-order mutants. This suggests that “mutators” in a population could make a
592 much greater contribution to adaptation than suggested by analyses focusing on single
593 loci (see also [62]). Furthermore, one should proceed with caution in applying mutation
594 rate estimates obtained by assaying a phenotype that can be conferred by a single point
595 mutation. If mutation rate is heterogeneous in the population, such an assay will in fact
596 estimate the mean rate, a reasonable result in itself. However a naive extrapolation as-
597 suming this rate to be fixed would underestimate the chance that such a population will
598 harbor multiple mutants. This is particularly concerning for analyses of whether multi-
599 drug resistant pathogens or cancerous cells are likely to “pre-exist” before a patient starts
600 a drug treatment [67, 41, 12]. Progression to cancer via accumulation of multiple muta-
601 tions may also occur faster than expected in a cellular population, even when controlling
602 for changes in mean mutation rate. (The importance of generally elevating mutation rate
603 in cancerous cells has previously been pointed out; [49, 46, 47, 48].)

604 Our analysis also clarified the role of parent-offspring correlation of mutation rate,
605 by directly comparing cases where mutation rate is either perfectly inherited or drawn
606 independently at random by each individual, such that the population-level distribution

607 is the same in both cases. (Realistically, different underlying mechanisms are likely to give
608 rise to both different levels of correlation and different distributions of mutation rate, but
609 we construct this hypothetical scenario to isolate the effect of the former.) Realistically,
610 parent-offspring correlation falls on a broad spectrum between the two extremes analyzed
611 here, and thus the effects of mutation rate variation are likely to be intermediate. The more
612 strongly mutation rate is correlated through a lineage, the greater the effect of variation
613 about the population mean, via the boosted rate of stepwise accumulation of mutations.
614 When mutation rates are uncorrelated, variation about the mean can only have an effect
615 through increasing the chance of simultaneous mutations, since non-random association
616 of mutation rate at multiple loci on a genome only lasts for one generation. A previous
617 study concluded that simultaneous mutations make a negligible contribution to multi-
618 locus adaptation [50], but their results supposed that single mutants are neutral or only
619 slightly deleterious. We considered a wider range of effects and found that if intermediates
620 are highly deleterious, simultaneous mutations play a crucial role in generating multiple
621 mutants directly. In this case, the extent to which mutation rate is inherited makes
622 little difference, so long as the population as a whole contains individuals with elevated
623 mutation rate (Section 3.2 and Suppl. Text III). The case of low-fitness intermediates
624 would be highly relevant for instance when a multi-drug combination therapy is applied
625 to a pathogen population, and each mutation confers resistance to a single drug but no
626 protection against the other drugs.

627 Taken together, our results suggest that populations with heterogeneous mutation
628 rate have a greater capacity to adapt. Firstly, they generate multi-point mutants faster,
629 which applies to beneficial *de novo* mutations as well as deleterious ones. In particular,
630 crossing fitness valleys is expected to be accelerated (see also [2, 50]). Secondly, in the
631 long term they harbor costly multi-point mutants at a higher frequency in the standing
632 genetic variation. This diversity is expected to be an important source for adaptation if

633 the environment changes and previously deleterious genotypes become favorable [5].

634 Finally, the clustering of multiple mutations into fewer individuals implies that a popu-
635 lation with heterogeneous mutation rate not only can explore genotype space more widely,
636 but simultaneously maintains higher population mean fitness. Such a possibility has pre-
637 viously been suggested verbally [23, 13] and uncovered in particular models assuming
638 mutation rate can take on two possible values, with no parent-offspring correlation [29, 2].
639 We find that this effect is much more general, holding for any distribution of mutation
640 rate, and, perhaps surprisingly, independently of the degree of parent-offspring correlation.
641 Furthermore, it does not require that mutation rate is specifically elevated in individu-
642 als with low fitness, which was previously found to yield a similar effect of simultaneous
643 “adaptedness and adaptability” in a model of stress-induced mutagenesis [66]. Thus, be-
644 sides concentrating high mutation rates into limited time periods or limited parts of the
645 genome [72], concentrating high mutation rates into particular individuals in a population
646 (whether heritably or not) adds another possible solution to the conundrum of balancing
647 gains in adaptability against fitness losses through deleterious mutations

648 4.2 Model limitations and extensions

649 We point out a few noteworthy assumptions of our modeling approach. Firstly, we de-
650 scribed an asexually reproducing haploid population. We expect similar but weaker qual-
651 itative effects to hold in sexual populations. Recombination can bring together mutations
652 generated in different lineages, thus reducing the importance of rare hypermutators in ac-
653 celerating the first appearance of multi-point mutants. In the longer term, recombination
654 would counteract positive linkage disequilibrium and thereby reduce the excess multiple
655 mutants generated by mutation rate heterogeneity.

656 Secondly, we neglected any fitness effects directly associated with mutation rate, in-
657 cluding any dependence of mutation rate on the alleles carried at the focal loci. How-
658 ever, particularly for viruses, there is a correlation between replicative fitness and mu-

659 tation rate, since both are influenced by the speed of replication of the genetic material
660 [28, 15]. Furthermore, environmental heterogeneities that may affect mutation rate (such
661 as chemotherapeutic drugs) also are likely to affect fitness. The level of “stress” experi-
662 enced by an individual and hence its mutation rate could then depend on its genotype:
663 for example, bacteria appear to down-regulate the SOS stress response (associated with
664 mutagenesis) as they adapt to a stressful environment [77].

665 Finally, we assumed that the mutation rate distribution was fixed from generation to
666 generation. Nonetheless, temporal environmental changes or evolution of the mutation
667 rate could yield changes in the population’s distribution of mutation rate over time. Ana-
668 lyzing the joint effects of inter-individual variation and population-level temporal changes
669 in mutation rate would be an interesting direction for future work.

670 **4.3 Empirical Outlook**

671 Given the evolutionary significance we predict for mutation rate heterogeneity, it would
672 clearly be of interest to gain a better empirical understanding of this heterogeneity. There
673 are two approaches to this question. Firstly, mutation rate as a function of some genetic
674 or environmental variable can be quantified using standard techniques to estimate (mean)
675 mutation rate in a population under each fixed condition. Combining this functional
676 relationship with a measurement or model of how the relevant variable is distributed
677 in a natural population would suggest in turn how mutation rate is distributed. The
678 latter point could be achieved for instance by expanding on recent advances in isolating
679 single virus-infected host cells [13] or single cells from cancerous tumors [79, 4, 60] and
680 quantifying mutation rate in their descendant lineages.

681 While it is well established that various alleles and environmental factors can affect
682 mutation rate, there is still work to be done to quantify these relationships in a way that is
683 useful for parameterizing population genetic models and thereby predicting evolutionary
684 consequences of mutation rate variation. We echo other authors [69, 51] in emphasizing

685 the need to infer mutation rate per generation, as opposed to reporting only the mean
686 frequency of mutants counted at the end of culture growth. While the former is a more
687 informative and reproducible measure [69], the latter remains in frequent use in experi-
688 mental studies of stress-induced mutagenesis and characterizations of natural isolates.

689 The second and more challenging direction is to quantify how mutation rate varies from
690 individual to individual within a population even under fixed experimental conditions,
691 including the detection of non-inherited variation. This requires statistical analysis of
692 the distribution of number of mutations per individual to test for deviations from the
693 expectation under a model with uniform mutation rate [23, 22, 27, 37]. Recently developed
694 individual-based methods provide new ways to count mutations, including fluorescent
695 labeling of nascent mutant foci in *E. coli* [27] and isolation and whole-genome sequencing
696 of mother and daughter cells in budding yeast [37]. Nonetheless, an inherent challenge
697 is that mutations are rare events, and at least double mutants are required to detect
698 non-uniformity. Besides examining the whole genome or using mutator strains [27, 37],
699 technical advances could make it feasible to examine more individuals in a population.
700 In parallel, a statistical power analysis (determining the sample size required to detect
701 mutation rate differences of a given magnitude) would be valuable both for experimental
702 design and for retrospective interpretation of existing results.

703 Acknowledgements

704 SB acknowledges funding by the ETH Zurich, by the European Research Council under
705 the 7th Framework Programme of the European Commission (PBDR:Grant Agreement
706 Number 268540), and by the Swiss National Science Foundation (Grant Nr. 155866).
707 The authors thank members of the Theoretical Biology group at ETH Zurich for in-
708 sightful comments on this work, particularly Dominique Cadosch and Antoine Frenoy for
709 discussion of stress-induced mutagenesis, and Antoine Frenoy for comments on a draft of

710 this manuscript.

References

711

712 [1] B. Al-Lazikani, U. Banerji, and P. Workman. Combinatorial drug therapy for cancer
713 in the post-genomic era. *Nature Biotechnology*, 30:1–13, 2012.

714 [2] K. Aoki and M. Furusawa. Promotion of evolution by intracellular coexistence of
715 mutator and normal DNA polymerases. *Journal of Theoretical Biology*, 209:213–222,
716 2001.

717 [3] M.-R. Baquero, A. I. Nilsson, M. C. Turrientes, D. Sandvang, J. C. Galán, J. L.
718 Martínez, N. Frimodt-Møller, F. Baquero, and D. I. Andersson. Polymorphic muta-
719 tion frequencies in *Escherichia coli*: emergence of weak mutators in clinical isolates.
720 *Journal of Bacteriology*, 186:5538–5542, 2004.

721 [4] L. J. Barber, M. N. Davies, and M. Gerlinger. Dissecting cancer evolution at the
722 macro-heterogeneity and micro-heterogeneity scale. *Current Opinion in Genetics &
723 Development*, 30:1–6, 2015.

724 [5] R. D. H. Barrett and D. Schluter. Adaptation from standing genetic variation. *Trends
725 in Ecology and Evolution*, 23:38–44, 2008.

726 [6] I. Bjedov, O. Tenaillon, B. Gérard, V. Souza, E. Denamur, M. Radman, F. Taddei,
727 and I. Matic. Stress-induced mutagenesis in bacteria. *Science*, 300:1404–1409, 2003.

728 [7] B. Björkholm, M. Sjölund, P. G. Falk, O. G. Berg, L. Engstrand, and D. I. Andersson.
729 Mutation frequency and biological cost of antibiotic resistance in *Helicobacter pylori*.
730 *PNAS*, 98:14607–14612, 2001.

731 [8] L. Boe. Translational errors as the cause of mutations in *Escherichia coli*. *Molecular
732 & General Genetics*, 231:469–471, 1992.

733 [9] L. Boe, M. Danielsen, S. Knudsen, J. B. Petersen, J. Maymann, and P. R. Jensen.

- 734 The frequency of mutators in populations of *Escherichia coli*. *Mutation Research*,
735 448:47–55, 2000.
- 736 [10] R. Bürger. *The Mathematical Theory of Selection, Recombination and Mutation*.
737 Wiley, 1st edition, 2000.
- 738 [11] J. Cairns. Mutation and cancer: the antecedents to our studies of adaptive mutation.
739 *Genetics*, 148:1433–1440, 1998.
- 740 [12] C. Colijn, T. Cohen, A. Ganesh, and M. Murray. Spontaneous emergence of multiple
741 drug resistance in tuberculosis before and during therapy. *PLoS One*, 6:e18327, 2011.
- 742 [13] M. Combe, R. Garijo, R. Geller, J. M. Cuevas, and R. Sanjuán. Single-cell analysis
743 of RNA virus infection identifies multiple genetically diverse viral genomes within
744 single infectious units. *Cell Host & Microbe*, 18:424–432, 2015.
- 745 [14] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley &
746 Sons, Inc., Hoboken, NJ, 2nd edition, 2006. Accessed online.
- 747 [15] M. J. Dapp, R. H. Heineman, and L. M. Mansky. Interrelationship between HIV-1
748 fitness and mutation rate. *Journal of Molecular Biology*, 425:41–53, 2013.
- 749 [16] T. Day. Modelling the ecological context of evolutionary change: déjà vu or something
750 new? In K. Cuddington and B. E. Beisner, editors, *Ecological Paradigms Lost: Routes*
751 *of Theory Change*, chapter 13, pages 273–309. Academic Press (Elsevier), Burlington,
752 MA, USA, 2005.
- 753 [17] R. del Campo, M.-I. Morosini, E. Gómez-G. de la Pedrosa, A. Fenoll, C. Muñoz-
754 Almagro, L. Máiz, F. Baquero, R. Cantón, and the Spanish Pneumococcal Infection
755 Study Network. Population structure, antimicrobial resistance, and mutation fre-
756 quencies of *Streptococcus pneumoniae* isolates from cystic fibrosis patients. *Journal*
757 *of Clinical Microbiology*, 43:2207–2214, 2005.

- 758 [18] E. Denamur, S. Bonacorsi, A. Giraud, P. Duriez, F. Hilali, C. Amorin, E. Bingen,
759 A. Andremont, B. Picard, F. Taddei, and I. Matic. High frequency of mutator
760 strains among human uropathogenic *Escherichia coli* isolates. *Journal of Bacteriol-*
761 *ogy*, 184:605–609, 2002.
- 762 [19] E. Denamur and I. Matic. Evolution of mutation rates in bacteria. *Molecular Micro-*
763 *biology*, 60:820–827, 2006.
- 764 [20] M. M. Desai and D. S. Fisher. The balance between mutators and nonmutators in
765 asexual populations. *Genetics*, 188:997–1014, 2011.
- 766 [21] J. W. Drake. General antimutators are improbable. *Journal of Molecular Biology*,
767 229:8–13, 1993.
- 768 [22] J. W. Drake. Too many mutants with multiple mutations. *Critical Reviews in*
769 *Biochemistry and Molecular Biology*, 42:247–258, 2007.
- 770 [23] J. W. Drake, A. Bebenek, G. E. Kissling, and S. Peddada. Clusters of mutations
771 from transient hypermutability. *PNAS*, 102:12849–12854, 2005.
- 772 [24] J. W. Drake, B. Charlesworth, D. Charlesworth, and J. F. Crow. Rates of sponta-
773 neous mutation. *Genetics*, 148:1667–1686, 1998.
- 774 [25] J. W. Drake and J. J. Holland. Mutation rates among RNA viruses. *PNAS*, 96:13910–
775 13913, 1999.
- 776 [26] S. Duffy, L. A. Shackelton, and E. C. Holmes. Rates of evolutionary change in viruses:
777 patterns and determinants. *Nature Reviews Genetics*, 9:267–276, 2008.
- 778 [27] M. Elez, A. W. Murray, L.-J. Bi, X.-E. Zhang, I. Matic, and M. Radman. Seeing
779 mutations in living cells. *Current Biology*, 20:1432–1437, 2010.

- 780 [28] V. Furió, A. Moya, and R. Sanjuán. The cost of replication fidelity in an RNA virus.
781 *PNAS*, 102:10233–10237, 2005.
- 782 [29] M. Furusawa and H. Doi. Promotion of evolution: disparity in the frequency of
783 strand-specific misreading between the lagging and leading DNA strands enhances
784 disproportionate accumulation of mutations. *Journal of Theoretical Biology*, 157:127–
785 133, 1992.
- 786 [30] S. H. Gillespie, S. Basu, A. L. Dickens, D. M. O’Sullivan, and T. D. McHugh. Effect of
787 subinhibitory concentrations of ciprofloxacin on *Mycobacterium fortuitum* mutation
788 rates. *Journal of Antimicrobial Chemotherapy*, 56:344–348, 2005.
- 789 [31] R. J. Gillies, D. Verduzco, and R. A. Gatenby. Evolutionary dynamics of carcinogen-
790 esis and why targeted therapy does not work. *Nature Reviews Cancer*, 12:487–493,
791 2012.
- 792 [32] D. E. Goldberg, R. F. Siliciano, and W. R. Jacobs Jr. Outwitting evolution: fighting
793 drug-resistant TB, malaria, and HIV. *Cell*, 148:1271–1283, 2012.
- 794 [33] C. Gonzalez, L. Hadany, R. G. Ponder, M. Price, P. J. Hastings, and S. M. Rosen-
795 berg. Mutability and importance of a hypermutable cell subpopulation that produces
796 stress-induced mutants in *Escherichia coli*. *PLoS Genetics*, 4:e1000208, 2008.
- 797 [34] J. Haigh. The accumulation of deleterious genes in a population – Muller’s Ratchet.
798 *Theoretical Population Biology*, 14:251–267, 1978.
- 799 [35] T. Johnson. The approach to mutation-selection balance in an infinite asexual pop-
800 ulation, and the evolution of mutation rates. *Proceedings of the Royal Society B*,
801 266:2389–2397, 1999.
- 802 [36] J. M. Kang, N. M. Iovine, and M. J. Blaser. A paradigm for direct stress-induced
803 mutation in prokaryotes. *The FASEB Journal*, 20:2476–2485, 2006.

- 804 [37] S. R. Kennedy, E. M. Schultz, T. M. Chappell, B. Kohn, G. M. Knowels, and A. J.
805 Herr. Volatility of mutator phenotypes at single cell resolution. *PLoS Genetics*,
806 11:e1005151, 2015.
- 807 [38] M. Kimura and T. Maruyama. The mutational load with epistatic gene interactions
808 in fitness. *Genetics*, 54:1337–1351, 1966.
- 809 [39] A.G. Knudson. Two genetic hits (more or less) to cancer. *Nature Reviews Cancer*,
810 1:157–162, 2001.
- 811 [40] M. A. Kohanski, M. A. DePristo, and J. J. Collins. Sublethal antibiotic treat-
812 ment leads to multidrug resistance via radical-induced mutagenesis. *Molecular Cell*,
813 37:311–320, 2010.
- 814 [41] N. L. Komarova and D. Wodarz. Drug resistance in cancer: principles of emergence
815 and prevention. *PNAS*, 102:9714–9719, 2005.
- 816 [42] J. E. LeClerc, B. Li, W. L. Payne, and T. A. Cebula. High mutation frequencies
817 among *Escherichia coli* and *Salmonella* pathogens. *Science*, 274:1208–1211, 1996.
- 818 [43] H. Lee, E. Popodi, H. Tang, and P. L. Foster. Rate and molecular spectrum of
819 spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-
820 genome sequencing. *PNAS*, 109:E2774–E2783, 2012.
- 821 [44] C. Lengauer, K. W. Kinzler, and B. Vogelstein. Genetic instabilities in human can-
822 cers. *Nature*, 396:643–649, 1998.
- 823 [45] K. R. Loeb and L. A. Loeb. Significance of multiple mutations in cancer. *Carcino-*
824 *genesis*, 21:379–385, 2000.
- 825 [46] L. A. Loeb. Mutator phenotype may be required for multistage carcinogenesis. *Can-*
826 *cer Research*, 51:3075–3079, 1991.

- 827 [47] L. A. Loeb. A mutator phenotype in cancer. *Cancer Research*, 61:3230–3239, 2001.
- 828 [48] L. A. Loeb, K. R. Loeb, and J. P. Anderson. Multiple mutations and cancer. *PNAS*,
829 100:776–781, 2003.
- 830 [49] L. A. Loeb, C. F. Springgate, and N. Battula. Errors in DNA replication as a basis
831 of malignant changes. *Cancer Research*, 34:2311–2321, 1974.
- 832 [50] M. Lynch and A. Abegg. The rate of establishment of complex adaptations. *Molecular*
833 *Biology and Evolution*, 27:1404–1414, 2010.
- 834 [51] R. C. MacLean, C. Torres-Barceló, and R. Moxon. Evaluating evolutionary models of
835 stress-induced mutagenesis in bacteria. *Nature Reviews Genetics*, 14:221–227, 2013.
- 836 [52] L. M. Mansky and K. S. Cunningham. Virus mutators and antimutators: roles in
837 evolution, pathogenesis and emergence. *Trends in Genetics*, 16:512–517, 2000.
- 838 [53] L. M. Mansky, E. Le Rouzic, S. Benichou, and L. C. Gajary. Influence of reverse
839 transcriptase variants, drugs, and Vpr on Human Immunodeficiency Virus Type 1
840 mutant frequencies. *Journal of Virology*, 77:2071–2080, 2003.
- 841 [54] L. M. Mansky and H. M. Temin. Lower in vivo mutation rate of human immun-
842 odeficiency virus type 1 than that predicted from the fidelity of purified reverse
843 transcriptase. *Journal of Virology*, 69:5087–5094, 1995.
- 844 [55] E. F. Mao, L. Lane, J. Lee, and J. H. Miller. Proliferation of mutators in a cell
845 population. *Journal of Bacteriology*, 179:417–422, 1997.
- 846 [56] I. Matic, M. Radman, F. Taddei, B. Picard, C. Doit, E. Bingen, E. Denamur, and
847 J. Elion. Highly variable mutation rates in commensal and pathogenic *Escherichia*
848 *coli*. *Science*, 277:1833–1834, 1997.

- 849 [57] J. D. McCool, E. Long, J. F. Petrosino, H. A. Sandler, S. M. Rosenberg, and S. J.
850 Sandler. Measurement of SOS expression in individual *Escherichia coli* K-12 cells
851 using fluorescence microscopy. *Molecular Microbiology*, 53:1343–1357, 2004.
- 852 [58] J. H. Miller. Spontaneous mutators in bacteria: insights into pathways of mutagenesis
853 and repair. *Annual Reviews in Microbiology*, 50:625–643, 1996.
- 854 [59] M.-I. Morosini, M.-R. Baquero, J. M. Sánchez-Romero, M.-C. Negri, J.-C. Galán,
855 R. del Campo, J. C. Pérez-Díaz, and F. Baquero. Frequency of mutation to rifampicin
856 resistance in *Streptococcus pneumoniae* clinical strains: *hexA* and *hexB* polymor-
857 phisms do not account for hypermutation. *Antimicrobial Agents and Chemotherapy*,
858 47:1464–1467, 2003.
- 859 [60] N. E. Navin. The first five years of single-cell cancer genomics and beyond. *Genome*
860 *Research*, 25:1499–1507, 2015.
- 861 [61] R. W. Ness, A. D. Morgan, R. B. Vasanthakrishnan, N. Colegrave, and P. D. Keight-
862 ley. Extensive *de novo* mutation rate variation between individuals and across the
863 genome of *Chlamydomonas reinhardtii*. *Genome Research*, 2015. Published in ad-
864 vance online 10 Aug. 2015.
- 865 [62] J. Ninio. Transient mutators: a semiquantitative analysis of the influence of transla-
866 tion and transcription errors on mutation rates. *Genetics*, 129:957–962, 1991.
- 867 [63] A. Oliver, R. Cantón, P. Campo, F. Baquero, and J. Blázquez. High frequency
868 of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science*,
869 288:1251–1253, 2000.
- 870 [64] A.-L. Prunier, B. Malbruny, M. Laurans, J. Brouard, J.-F. Duhamel, and R. Leclercq.
871 High rate of macrolide resistance in *Staphylococcus aureus* strains from patients with

- 872 cystic fibrosis reveals high proportions of hypermutable strains. *Journal of Infectious*
873 *Diseases*, 187:1709–1716, 2003.
- 874 [65] Y. Ram and L. Hadany. The evolution of stress-induced hypermutation in asexual
875 populations. *Evolution*, 66:2315–2328, 2012.
- 876 [66] Y. Ram and L. Hadany. Stress-induced mutagenesis and complex adaptation. *Pro-*
877 *ceedings of the Royal Society B*, 281:20141025, 2014.
- 878 [67] R. Ribeiro and S. Bonhoeffer. Production of resistant HIV mutants during antiretro-
879 viral therapy. *PNAS*, 97:7681–7686, 2000.
- 880 [68] A. R. Richardson, Z. Yu, T. Popovic, and I. Stojiljkovic. Mutator clones of *Neisseria*
881 *meningitidis* in epidemic serogroup A disease. *PNAS*, 99:6103–6107, 2002.
- 882 [69] W. A. Rosche and P. L. Foster. Determining mutation rates in bacterial populations.
883 *Methods*, 20:4–17, 2000.
- 884 [70] S. M. Rosenberg. Spontaneous mutation: real-time in living cells. *Current Biology*,
885 20:R810–R811, 2010.
- 886 [71] R. Sanjuán, M. R. Nebot, N. Chirico, L. M. Mansky, and R. Belshaw. Viral mutation
887 rates. *Journal of Virology*, 84:9733–9748, 2010.
- 888 [72] P. D. Sniegowski, P. J. Gerrish, T. Johnson, and A. Shaver. The evolution of mutation
889 rates: separating causes from consequences. *BioEssays*, 22:1057–1066, 2000.
- 890 [73] P. D. Sniegowski, P. J. Gerrish, and R. E. Lenski. Evolution of high mutation rates
891 in experimental populations of *E. coli*. *Nature*, 387:703–705, 1997.
- 892 [74] P. Suárez, J. Valcárcel, and J. Ortín. Heterogeneity of the mutation rates of Influenza
893 A viruses: isolation of mutator mutants. *Journal of Virology*, 66:2491–2494, 1992.

- 894 [75] F. Taddei, M. Radman, J. Maynard-Smith, B. Toupance, P. H. Gouyon, and
895 B. Godelle. Role of mutator alleles in adaptive evolution. *Nature*, 387:700–702,
896 1997.
- 897 [76] O. Tenaillon, E. Denamur, and I. Matic. Evolutionary significance of stress-induced
898 mutagenesis in bacteria. *Trends in Microbiology*, 12:264–270, 2004.
- 899 [77] C. Torres-Barceló, M. Kojadinovic, R. Moxon, and R. C. MacLean. The SOS response
900 increases bacterial fitness, but not evolvability, under a sublethal dose of antibiotic.
901 *Proceedings of the Royal Society B*, 282:20150885, 2015.
- 902 [78] M. E. Watson Jr, J. L. Burns, and A. L. Smith. Hypermutable *Haemophilus in-*
903 *fluenzae* with mutations in *mutS* are found in cystic fibrosis sputum. *Microbiology*,
904 150:2947–2958, 2004.
- 905 [79] C. Zong, S. Lu, A. R. Chapman, and X. S. Xie. Genome-wide detection of single-
906 nucleotide and copy-number variations of a single human cell. *Science*, 338:1622–1626,
907 2012.

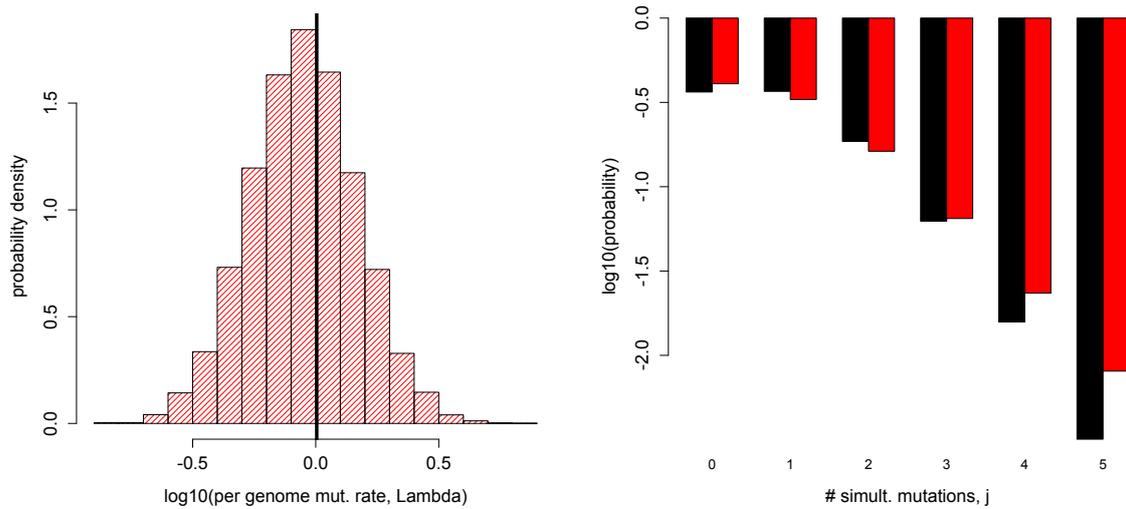


Figure 1: **Directional effect of mutation rate heterogeneity on simultaneous mutation probability.** Left: the distribution of per-genome mutation rate (Λ) obtained by sampling 10000 times from a log-normal distribution, with the thick black vertical line indicating the sample mean. Sample mean ≈ 1.0 (similar to estimates for some RNA viruses [25]), sample variance ≈ 0.29 , range $\approx 0.14 - 6.3$. Right: probability of j simultaneous mutations occurring in the genome under the infinite-locus model, i.e. p_j , if Λ is heterogeneous following the chosen distribution (red) versus fixed to the sample mean (black). The directional effects agree with those predicted analytically (keeping in mind that higher probabilities will be less negative on the log scale).

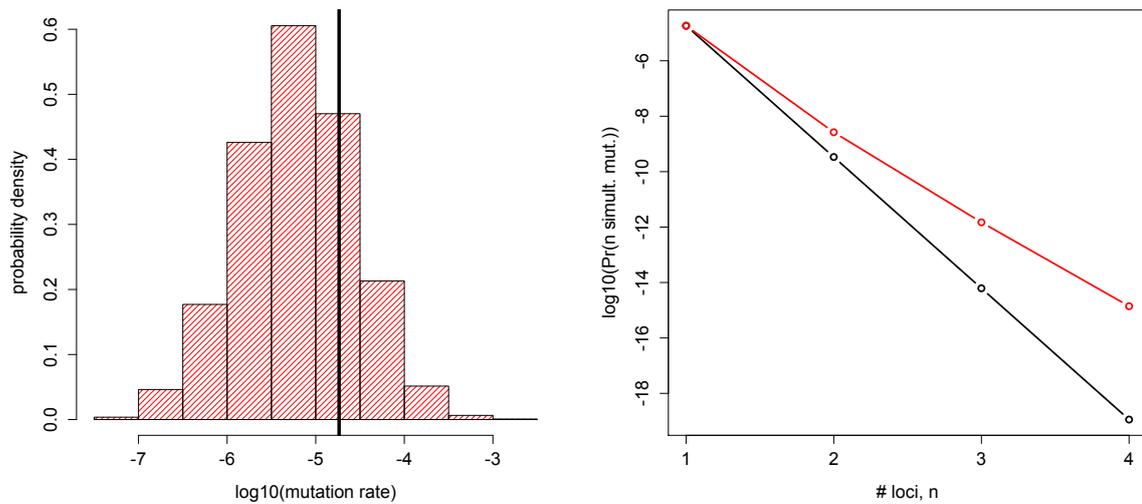
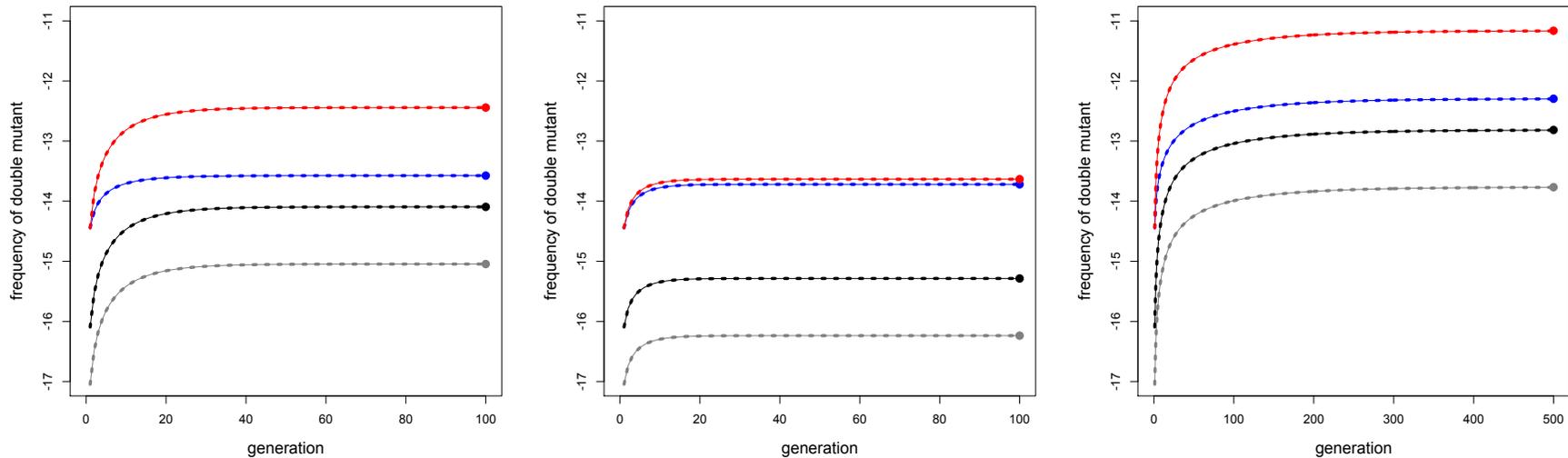


Figure 2: **Relative effect of mutation rate heterogeneity increases with the number of mutating loci.** Left: the distribution of per-locus mutation rate (U) obtained by sampling 10000 times from a log-normal distribution, with the thick black vertical line indicating the sample mean. Sample mean $\approx 1.8 \times 10^{-5}$ (close to the base substitution rate estimated for HIV [54]), sample variance $\approx 2.3 \times 10^{-9}$, range $\approx 3.2 \times 10^{-8} - 1.4 \times 10^{-3}$. Right: probability of all n loci under consideration mutating simultaneously, i.e. $p_{n,n}$, as a function of n , if U is heterogeneous following the chosen distribution (red) versus fixed to the sample mean (black).



45

Figure 3: **Temporal dynamics of the double mutant with approach to equilibrium**, assuming the population is initially composed entirely of the wild type. Mutation rate takes on either of two values: $U = u_\ell = 3 \times 10^{-9}$ with probability 0.99 or $U = u_h = 6 \times 10^{-7}$ with probability 0.01, thus $\langle U \rangle \approx 9.0 \times 10^{-9}$ and $V \approx 3.5 \times 10^{-15}$. These values are reasonable for loci with large target sizes in a bacterial population containing a strong hypermutator at 1% frequency, parameterizing approximately from [43]: u_ℓ is the wild type *E. coli* mutation rate to rifampicin resistance estimated from a fluctuation assay and u_h/u_ℓ is the fold-increase in mutation rate in the mismatch repair defective MutL⁻ strain. The selection coefficients vary across panels: left, $s_{01} = s_{10} = 0.1$ and $s_{11} = 0.19$; center, $s_{01} = s_{10} = 0.9$ and $s_{11} = 0.19$; right, $s_{01} = s_{10} = 0.1$ and $s_{11} = 0.01$. Black indicates the homogeneous case with mutation rate fixed to $\langle U \rangle$; blue is the heterogeneous case with no inheritance; and red is the heterogeneous case with perfect inheritance. Grey additionally shows the result when $U \equiv u_\ell$. Thus comparing black to grey indicates the effect of changing the mean mutation rate by adding a hypermutator, while comparing blue/red to black isolates the effect of increasing variance with fixed mean. The solid line in each case indicates the result of numerically iterating the recursions, while the dashed line indicates the analytical approximation for the temporal dynamics, and the large point at the end of each curve indicates the analytical approximation of the equilibrium.

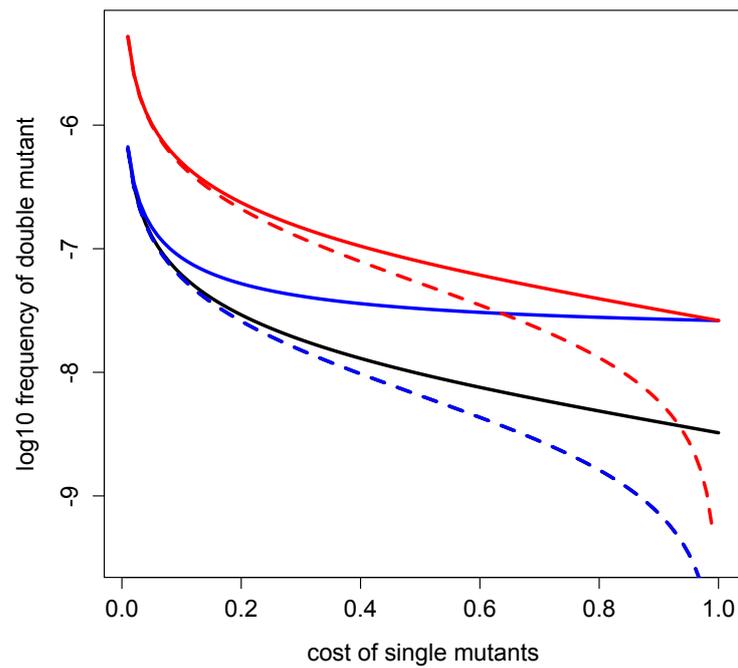


Figure 4: **Equilibrium frequency of the double mutant as a function of single mutant costs.** The analytical approximations for the equilibrium double mutant frequency (Table 1) are plotted for the various model cases: homogeneous – black; heterogeneous, no inheritance – blue; heterogeneous, perfect inheritance – red. Solid lines indicate the result when simultaneous mutations are allowed; dashed lines when simultaneous mutations are blocked. We take $\langle U \rangle = 1.8 \times 10^{-5}$ and $V = 2.3 \times 10^{-9}$ as in Figure 2.

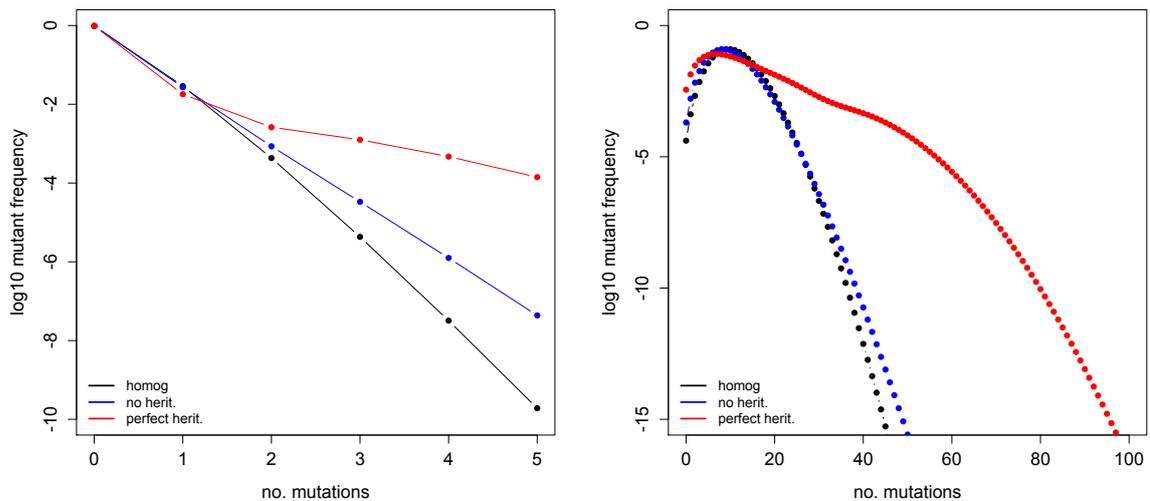


Figure 5: **Equilibrium mutant frequencies under the infinite-locus model.** Specifically, the base 10 log of the genotype frequency at the numerically determined equilibrium is plotted as a function of number of mutations carried, in each model case (homogeneous – black; heterogeneous with no inheritance – blue, or perfect inheritance – red). The selection coefficient per mutation is $s = 0.1$. The results are illustrated for two example mutation rate distributions. Left: Λ takes on two values, 0.0015 with probability 0.99 or 0.15 with probability 0.01 (mean ~ 0.003 is bacteria-like; [24]). Right: Λ is given by 1000 draws from a log-normal distribution. Sample mean is 1.01 (RNA- or retrovirus-like; [24]), sample variance is 0.29, and range is 0.19-4.2.

	With simultaneous mutation	Without simultaneous mutation
Homog. mut. rate	$\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 1\right) \frac{\langle U \rangle^2}{s_{11}}$	$\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 2\right) \frac{\langle U \rangle^2}{s_{11}}$
Heterog. mut. rate, no herit.	$\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 1\right) \frac{\langle U \rangle^2}{s_{11}} + \frac{V}{s_{11}}$	$\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 2\right) \frac{\langle U \rangle^2}{s_{11}}$
Heterog. mut. rate, perfect herit.	$\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 1\right) \frac{\langle U \rangle^2 + V}{s_{11}}$	$\left(\frac{1}{s_{01}} + \frac{1}{s_{10}} - 2\right) \frac{\langle U \rangle^2 + V}{s_{11}}$

Table 1: Approximate equilibrium frequency of double mutants (x_{11}^*) in various cases of the two-locus model. In the homogeneous case, we suppose the mutation rate u is fixed to $\langle U \rangle$ for comparison with the heterogeneous cases, where the equilibrium is defined by both the mean $\langle U \rangle$ and variance V of the mutation rate distribution.