



HAL
open science

Population persistence under high mutation rate: from evolutionary 1 rescue to lethal mutagenesis

Yoann Anciaux, Amaury Lambert, Ophélie Ronce, Lionel Roques, Guillaume Martin

► To cite this version:

Yoann Anciaux, Amaury Lambert, Ophélie Ronce, Lionel Roques, Guillaume Martin. Population persistence under high mutation rate: from evolutionary 1 rescue to lethal mutagenesis. *Evolution - International Journal of Organic Evolution*, 2019, 73 (8), pp.1517-1532. 10.1111/evo.13771 . hal-02370259

HAL Id: hal-02370259

<https://hal.science/hal-02370259v1>

Submitted on 22 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Population persistence under high mutation rate: from evolutionary
2 rescue to lethal mutagenesis

3 Yoann Anciaux^{1, 2} Amaury Lambert^{3, 4} Ophélie Ronce¹ Lionel Roques⁵ Guillaume Martin¹

4

5 ¹ ISEM - Institut des Sciences de l'Evolution de Montpellier, France

6 ² BiRC - Bioinformatics Research Center, Aarhus Danemark

7 ³ CIRB - Centre interdisciplinaire de recherche en biologie, Paris, France

8 ⁴ LPSM UMR 8001 - Laboratoire de Probabilités, Statistique et Modélisation, Paris, France

9 ⁵ INRA - Biostatistics and Spatial Processes (BioSP), Avignon, France

10

11 Contacts : anciaux.yoann@gmail.com , guillaume.martin@umontpellier.fr

12

13 **Keywords** : Evolutionary rescue, Fisher's Geometric Model, Fitness landscape, Lethal mutagenesis,
14 Extinction dynamics, Distribution of extinction times

15

16 Abstract

17 Populations may genetically adapt to severe stress that would otherwise cause their
18 extirpation. Recent theoretical work, combining stochastic demography with Fisher’s geometric
19 model of adaptation, has shown how evolutionary rescue becomes unlikely beyond some critical
20 intensity of stress. Increasing mutation rates may however allow adaptation to more intense stress,
21 raising concerns about the effectiveness of treatments against pathogens. This previous work
22 assumes that populations are rescued by the rise of a single resistance mutation. However, even
23 in asexual organisms, rescue can also stem from the accumulation of multiple mutations in a single
24 genome. Here, we extend previous work to study the rescue process in an asexual population
25 where the mutation rate is sufficiently high so that such events may be common. We predict both
26 the ultimate extinction probability of the population and the distribution of extinction times. We
27 compare the accuracy of different approximations covering a large range of mutation rates.
28 Moderate increase in mutation rates favors evolutionary rescue. However, larger increase leads to
29 extinction by the accumulation of a large mutation load, a process called lethal mutagenesis. We
30 discuss how these results could help design “evolution-proof” anti-pathogen treatments that even
31 highly mutable strains could not overcome.

32

33 Introduction

34 Evolutionary rescue (ER) happens when a population confronted with severe stress avoids
35 extinction by genetic adaptation. Understanding and predicting when and how evolutionary rescue
36 occurs is critical in fields as diverse as conservation biology, invasion biology, emergence of new
37 diseases and the management of resistance to treatment in pests and pathogens (see reviews in
38 Gonzalez *et al.* 2013; Carlson *et al.* 2014; Alexander *et al.* 2014; Bell 2017). In all these situations,
39 genetic variation, be it present before the onset of stress, or generated *de novo* after, is a key
40 ingredient for evolutionary rescue, as expected theoretically (e.g. Gomulkiewicz and Holt 1995)
41 and observed experimentally (e.g Ramsayer *et al.* 2013). Because mutation affects both standing
42 and *de novo* genetic variation, it comes as no surprise that a number of evolutionary rescue models,

43 combining stochastic evolution and demography, have predicted that higher mutation rates are
44 associated with higher probability of evolutionary rescue (Orr and Unckless 2008, 2014; Martin et
45 al. 2013; Anciaux et al. 2018). Few evolutionary rescue experiments have manipulated the
46 mutation rate to test these predictions (reviewed in Bell 2017). For instance, Couce *et al.* (2015)
47 found that two different mutator strains of bacteria with elevated rates of mutations evolved more
48 than 100-fold resistance to antibiotic concentrations that caused the demise of control strains.
49 Mutator alleles are indeed often found in antibiotic resistant strains causing serious health issues
50 (Eliopoulos and Blázquez 2003), raising concern about pathogens escaping our control by evolving
51 higher mutation rates (for theoretical predictions see Taddei *et al.* 1997; Greenspoon and Mideo
52 2017).

53 Most mathematical models of evolutionary rescue assume that the population is rescued
54 from extinction by the spread of a single mutant of large effect (Orr and Unckless 2008, 2014;
55 Martin et al. 2013; Anciaux et al. 2018) and do not describe highly polymorphic populations where
56 several mutations of smaller effects can combine to allow population growth (see however the
57 work of Uecker and Hermisson (2016) and Uecker (2017) where sexual reproduction allows
58 production of such rescue genotypes). The latter situation seems in particular to be common in
59 the evolution of herbicide resistance, especially when the mutational target for resistance is large
60 (Kreiner et al. 2018). Even in asexual organisms, when the mutation rate is high, evolutionary
61 rescue may commonly result from the cumulative effect of multiple mutations accumulating
62 stochastically over time in a given lineage. Such a mutation regime is particularly relevant in highly
63 mutable viruses, mutator strains of bacterial (e.g. Springman *et al.* 2010) or cancer cells (e.g. Loeb
64 2001). Our aim here is to provide theoretical predictions for evolutionary rescue in such a regime
65 with high mutation rates in asexual organisms, complementing existing theory on the subject.

66 Several complications arise when modelling evolutionary rescue in highly polymorphic
67 populations with high mutation rates. First, the dynamics of allelic frequencies at different loci
68 interact in asexuals. For example, the selective sweep of a given beneficial mutation is hindered by
69 the co-segregation of other beneficial mutations (clonal interference, Gerrish and Lenski 1998). A
70 theoretical study by Wilson *et al.* (2017) recently showed that, when evolutionary rescue is likely,
71 it should most often be driven by soft selective sweeps, where multiple resistance mutations

72 spread through the population simultaneously. Wilson *et al.* (2017) still assumed that each of these
73 lineages carried a single mutation, each with the same effect on the population growth rate. When
74 the mutational target is large, different lineages contributing to rescue are however likely to carry
75 mutations with different fitness effects. Modelling the distribution of mutation effects (as in Martin
76 *et al.* 2013; Anciaux *et al.* 2018) then becomes critical. Finally, when the mutation rate is high,
77 multiple mutations may also accumulate in each lineage, either facilitating evolutionary rescue or
78 impeding it, through their *cumulative* effect. Modelling both beneficial and deleterious mutations
79 and, critically, the epistatic interactions between them, also becomes necessary.

80 Previous evolutionary rescue theory predicts that higher mutation rate allows populations
81 to withstand higher levels of stress (e.g. Anciaux *et al.* 2018). Yet, there are reasons to expect this
82 prediction not to hold above some critical mutation rate: increased mutation rates also build-up
83 detrimental mutation loads, thus depressing mean fitness despite ongoing adaptation. Indeed,
84 some previous ER models, including both beneficial and deleterious additive effects on growth
85 rates, have found that ER was most likely at intermediate mutation rates (Loverdo and Lloyd-Smith
86 2013; Greenspoon and Mideo 2017). Artificially increasing the mutation rate has even been
87 proposed as a means to weaken or even eliminate pathogen populations, by a process called lethal
88 mutagenesis (Loeb *et al.* 1999). Models of lethal mutagenesis predict that extinction of the target
89 population could be observed under biologically realistic sets of parameters (Bull *et al.* 2007;
90 Martin and Gandon 2010; Wylie and Shakhnovich 2012). In these models, mean fitness dynamics
91 and extinction stem from the deterministic effects of selection and mutation. Alternatively,
92 Matuszewski *et al.* (2017) discuss the continuity between these models and models of mutational
93 meltdown, where extinction is driven by the interaction of genetic drift and deleterious mutation.
94 Lethal mutagenesis has been investigated empirically for treatment against viruses (Springman *et al.*
95 2010; Arias *et al.* 2014), bacteria (Bull and Wilke 2008) or cancer cells (Liu *et al.* 2015). In
96 particular, the combination of antiviral treatments with mutagenic agents is investigated as a
97 strategy to fight fast evolving viruses, such as influenza (Bank *et al.* 2016). It seems important to
98 improve our ability to predict whether and when such mutagenic agents will increase treatment
99 efficacy or, conversely, facilitate the evolution of resistance.

100 The population genetics of adaptation behind the rescue process, in isolated asexual
101 populations, roughly fall into two alternative regimes: rescue may stem (i) from single mutations
102 of large effect (strong selection weak mutation ‘SSWM’ regime) or (ii) from multiple mutations of
103 small effects (weak selection strong mutation ‘WSSM’ regime) (reviewed in Alexander *et al.* 2014).
104 The dichotomy between SSWM and WSSM entails a somewhat simplistic view of adaptation
105 regimes, at the two extremes of all possible mutation rates. The SSWM regime of adaptation has
106 been extensively investigated in population genetics via “origin-fixation” models describing the
107 average behavior of stochastic evolutionary dynamics (McCandlish and Stoltzfus 2014) whereas
108 the WSSM regime has been widely analyzed via deterministic models of quantitative genetics
109 (Lande 1976, 1980). Corresponding evolutionary rescue models further include a coupling of
110 adaptation and demographic dynamics, and naturally fall into the same two regimes (discussed in
111 Anciaux *et al.* 2018). The SSWM regime of evolutionary rescue is characterized by the fact that the
112 first resistant lineage to establish (and thus cause rescue) is only one mutational step from the
113 predominant, sensitive ‘wild-type’ lineage (e.g. Feder *et al.* 2016). Models that describe highly
114 polymorphic dynamics (WSSM regimes, e.g. the quantitative genetic model in Gomulkiewicz and
115 Holt 1995) often use the infinitesimal model assumptions (many unlinked polymorphic loci), which
116 does not apply to asexual populations. In the WSSM regime, the exact stochastic evolutionary
117 dynamics become quickly intractable, and have often been studied by simulation (e.g. Boulding
118 and Hay 2001). Further the latter models often consider initial standing genetic variance as given
119 and pay little attention to the effect of mutation rates in maintaining this variance. They often
120 ignore *de novo* mutations after the onset of stress, on the argument of short timescales being most
121 critical for evolutionary rescue (e.g. Gomulkiewicz *et al.* 2010).

122 To make analytical progress in our understanding of the effect of mutation rates on the
123 process of evolutionary rescue, we build on two recent theoretical developments (Martin and
124 Roques 2016; Anciaux *et al.* 2018). Anciaux *et al.* (2018) developed a model of evolutionary rescue
125 in the SSWM regime using Fisher’s (1930) geometric model (hereafter “FGM”) to model the
126 distribution of mutation effects on fitness. As in Anciaux *et al.* (2018), we study evolutionary rescue
127 using the FGM. This model assumes a single peak phenotype-fitness landscape, where fitness
128 depends on the position, in phenotype space, of a given genotype relative to an optimum. In the

129 context of ER, stress may affect this landscape in various ways (height, width or position of the
130 peak). In this model, the distribution of mutation effects (both beneficial and deleterious) depend
131 on the context, both genotypic (epistasis) and environmental (e.g. effect of stress). This context-
132 dependence is a key feature of the FGM; it is absent from previous ER models studying the effect
133 of high mutation rates, because they assume additive mutation effects on fitness (e.g. Loverdo and
134 Lloyd-Smith 2013; Greenspoon and Mideo 2017). The variation in the distribution of mutation
135 effects implied by the FGM is qualitatively, and sometimes even quantitatively, consistent with a
136 wealth of empirical observations (reviewed in Tenaillon 2014). Under the assumptions of the FGM,
137 rescue mutants become very rare as the intensity of stress increases, because they require very
138 large mutational steps. As a consequence, Anciaux *et al.* (2018) predict that there is a narrow
139 window of stress levels where the probability of rescue shifts from being very likely to very unlikely.
140 They also predict that this critical level of stress, beyond which adaptation is unlikely, is increased
141 by higher mutation rates (Anciaux et al. 2018). Yet, predictions of this model apply to the SSWM
142 regime and may not hold for higher mutation rates.

143 We extend our previous analysis of evolutionary rescue over Fisher's geometric model
144 (Anciaux et al. 2018) to the more complex and more polymorphic WSSM regime. To do so, we use
145 the approach in Martin and Roques (2016) to model the non-equilibrium dynamics of fitness
146 distributions, in large asexual populations. The form of fitness epistasis assumed may have
147 particular impact on the results because lineages accumulate multiple mutations at different sites
148 over the rescue process. We use the FGM here, which implies a particular form of epistasis that
149 has proven consistent with several observed patterns in fitness epistasis among mutations (Martin
150 et al. 2007; Perfeito et al. 2014; Blanquart and Bataillon 2016). Moreover, Martin and Roques
151 (2016) showed that, under the FGM, while the fitness dynamics are more complex at higher
152 mutation rates, they are also more predictable and less prone to stochastic fluctuations, even in
153 relatively small populations. To model evolutionary rescue, we still need to describe the
154 demographic stochasticity associated with the extinction process. In the WSSM regime, we thus
155 use a combination of two analytically tractable theories: a deterministic approximation to the
156 dynamics of mean fitness (Martin and Roques 2016) and a diffusion approximation to the
157 stochastic dynamics of population sizes (from Bansaye and Simatos 2015).

158 Beyond a derivation of the probability of ultimate rescue or extinction, this approach
159 further allows tracking the rescue process over time. As stated in Gomulkiewicz *et al.* (2017),
160 tracking of transient dynamics (population size dynamics, distributions of extinction times) are of
161 high interest for applications of evolutionary rescue theory, yet are not available from existing
162 predictions, which focused mainly on ultimate outcomes. Gomulkiewicz *et al.* (2017) studied the
163 distribution of extinction times for populations doomed to extinction, mostly in the absence of
164 mutation (i.e. with a fixed arbitrary set of competing asexual genotypes at the onset of stress). We
165 extend this analysis to include frequent *de novo* mutation, rescue events involving several
166 mutational steps, a particular form of epistasis and variable mutation effects depending on stress
167 intensity, and an explicit description of the dynamics of mutation load. Our approach captures the
168 continuum from evolutionary rescue to lethal mutagenesis, as mutation rate increases.
169 Interestingly, some parameter ranges prove to greatly limit evolutionary rescue at all mutation
170 rates, i.e. in spite of the possible mutator genotypes.

171

172 Methods

173 **I. General framework**

174 The present work focuses on an asexual population with independent lineages (e.g. a
175 population of asexual microbes without horizontal gene transfer), facing an abrupt and stressful
176 environmental change (e.g. an antimicrobial treatment). The population initially consists of N_0
177 individuals with either a single or multiple genotypes and is initially adapted to a non-stressful
178 environment, where its mean growth rate is positive. At the onset of stress, the population is
179 shifted from the non-stressful environment to a stressful environment, where its mean growth rate
180 becomes negative (definition of stress here). In such an environment, in the absence of evolution,
181 the population is doomed to extinction. Evolutionary rescue occurs if at least one resistant lineage
182 (with a positive growth rate in the new environment) establishes, in spite of demographic
183 stochasticity. These resistant mutant lineages can either already be present in the population or
184 arise *de novo* after the onset of stress. It is thus crucial to determine how the number and growth
185 rates of such mutants depend on the new environmental conditions and on the parental genotypes

186 already present in the population. We do so using the FGM detailed below. Note that the main
187 notations used here are summarized in Table 1.

188

189 *Fitness landscape*

190 In the FGM, a given phenotype is a vector in a phenotypic space of n dimensions that
191 determine fitness (here the growth rate r). The phenotype of an individual with genotype i , is
192 characterized by a vector $\mathbf{z}_i \in \mathbb{R}^n$ of the breeding values (heritable components) for the n traits,
193 and its growth rate is r_i . In a given environment, fitness decays as a quadratic function of the
194 phenotypic distance to a single phenotypic optimum, where the growth rate r_{max} is maximal at a
195 given absolute level (height of the fitness peak). We assume that each environment is associated
196 with a single optimum and fitness peak. In the scenario investigated here, in the non-stressful
197 environment, the population is close to the ‘ancestral’ optimum \mathbf{z}_A . When the environment
198 changes, it is assumed to determine a new optimum \mathbf{z}_* . Without loss of generality, the height of
199 the peak may also differ between the ancestral and new environments. However, we do require
200 that the n dimensions that determine fitness remain the same (in nature and number) across
201 environments. In the new environment, the growth rate of an individual with genotype i is given
202 by:

$$r(\mathbf{z}_i, \mathbf{z}_*) = r_{max} - \frac{\|\mathbf{z}_i - \mathbf{z}_*\|^2}{2}. \quad [1]$$

203 This is an isotropic version of the FGM (all directions are equivalent for selection and mutation)
204 where phenotypes are scaled by selective strength.

205

206 *Stochastic demographic dynamics*

207 We confine our analysis to finite haploid asexual populations. Individuals have independent
208 evolutionary and demographic fate (frequency or density dependence are ignored). Each genotype
209 i has a growth rate r_i and a reproductive variance σ_i ($r_i = r(\mathbf{z}_i, \mathbf{z}_*)$) in a given environment with

210 optimum \mathbf{z}_*), which define its stochastic demographic parameters in the context of a Feller
 211 diffusion approximation (Feller 1951), as in e.g. Martin et al. (2013), Gomulkiewicz *et al.* (2017) or
 212 Anciaux *et al.* (2018). For simplicity, we further assume here that the average stochastic variance
 213 in reproduction is constant over time: $\bar{\sigma}(t) = 1/K_t \sum_{i=1}^{K_t} \sigma_i = \bar{\sigma}$, where the average is taken over
 214 the K_t genotypes present at time t . This can for example be accurate whenever the σ_i are roughly
 215 constant across genotypes i (discussed in Martin *et al.* 2013 and Anciaux *et al.* 2018).

216

Notation	Description	Formula
N_t, N_0	N_t : population size at time t after the onset of the stress. N_0 : initial population size at the onset of the stress.	
U	Mutation rate <i>per individual per unit time</i> .	
n	Number of traits under stabilizing selection, or phenotypic dimensionality.	
λ	Variance of mutational effects: variance of the phenotypic effects of mutations, per trait, in a trait space scaled by the strength of stabilizing selection	$\mathbf{dz} \sim N(\mathbf{0}, \lambda \mathbf{I}_n)$
μ	Composite parameter : square root of mutational variance per trait	$\mu = \sqrt{U\lambda}$
\mathbf{z}	n -dimensional vector $\mathbf{z} \in \mathbb{R}^n$ of (breeding values for) phenotype	
$\mathbf{z}_A, \mathbf{z}_*$	\mathbf{z}_* : optimal phenotype in the new environment \mathbf{z}_A : average phenotype of the ancestral population (before the onset of stress)	
r, σ	Growth rate (r) and reproductive variance (σ) of a given genotype, in the new environment.	Eq.[1]
$\bar{r}_t, \bar{\sigma}$	Mean growth rate (\bar{r}_t) and mean reproductive variance ($\bar{\sigma}$) of the population, in the new environment.	Eqs. [2] and [3]
$\langle \bar{r}_t \rangle$	'ensemble mean' growth rate : expected mean growth rate \bar{r}_t across stochastic replicates	Eqs. [2] and [3]
r_{max}	Maximum possible growth rate in the new environment.	$r(\mathbf{z}_*) = r_{max}$
r_D	Rate of decay of the ancestral phenotype \mathbf{z}_A in the new environment.	$r(\mathbf{z}_A) = r_{max} - \ \mathbf{z}_A\ ^2/2 = -r_D$
y_D	Rate of decay scaled by r_{max} .	$y_D = r_D/r_{max}$
ϵ	Ratio between the mutational load and the maximal growth rate	$\epsilon = n \mu / (2 r_{max})$
$\omega_{DN}, \omega_{DN+SV}$	ω_{DN} : rate of rescue from <i>de novo</i> mutations scaled by N_0 ω_{DN+SV} : rate of rescue from both <i>de novo</i> and standing variance scaled by N_0	Eqs.[5],[6]
P_{ext}, P_R	P_{ext} : Extinction probability P_R : Rescue probability	P_{ext} see Eq.[4] $P_R = 1 - P_{ext} = 1 - \exp(-N_0 \omega_R)$

U_c	Critical mutation rate beyond which the WSSM regime is valid.	$U_c = n^2 \lambda/4$
U_{max}	Mutation rate beyond which certain extinction is enforced by lethal mutagenesis	$U_{max} = 4 r_{max}^2/(n^2 \lambda)$
U_*	Mutation rate at which the rescue probability rises to 1/2	Eq.[7]
$r_D^*(p)$	Rate of decay for which the maximal rescue probability for any U or λ is equal to p	Eq.[8]

217 ***Table 1:*** Notations

218

219

220 II. Evolutionary dynamics.

221 In this section, we describe the model of evolutionary dynamics over the fitness landscape
222 (FGM of the previous section), which is embedded into the ER model. In the following, *de novo*
223 mutations (appearing after the onset of stress) are denoted “DN” and mutations from standing
224 genetic variance (mutants already present before the onset of stress) are denoted “SV”.
225 Correspondingly, evolutionary rescue dynamics from an isogenic population, adapting only from
226 *de novo* mutations, are labelled “DN” and dynamics from a polymorphic population, adapting from
227 both *de novo* mutations and standing genetic variance, are labelled “DN + SV”.

228

229 1. Evolutionary dynamics from an isogenic population.

230 The population is maladapted in the new stressful environment and its growth rate is $-r_D$,
231 corresponding to a decay rate $r_D > 0$. Mutations arise following a Poisson process with rate U per
232 unit time per capita. For a given parent phenotype, each mutation creates a random perturbation
233 $d\mathbf{z}$ on phenotype, which is unbiased and follows an isotropic multivariate Gaussian distribution,
234 $d\mathbf{z} \sim \mathcal{N}(0, \lambda \mathbf{I}_n)$ where \mathbf{I}_n is the identity matrix in n dimensions and λ is the variance of mutational
235 effects on traits, standardized by the strength of selection. Mutation effects are additive on
236 phenotype, but not on fitness because $r(\cdot)$ is nonlinear (epistasis on fitness and not on
237 phenotype).

238 In the WSSM regime, the mean growth rate of the population shows limited stochastic
 239 variation among replicates, even in reasonably small populations. We thus approximate the
 240 evolutionary process by a deterministic fitness trajectory, derived in the WSSM regime under the
 241 FGM (Martin and Roques 2016). This seemingly rough approximation can be justified *a priori*: most
 242 of the ER process is determined by the speed at which the population adapts at the very onset of
 243 stress. This early trajectory takes place when the population is still large and the adaptive process
 244 proves to be relatively deterministic, especially over this short timescale, provided that the
 245 mutation rate is high enough compared to the mean fitness effect of random mutations (WSSM:
 246 $U \gg U_c = n^2\lambda/4$, with $n\lambda/2$ the mean fitness effect of random mutations). Both analytical
 247 arguments and simulations detailed in Martin and Roques 2016, showed that mean fitness
 248 trajectories are indeed close to the deterministic prediction (with limited variation among
 249 replicates), provided that $U \gg U_c$ (WSSM) and $NU \gg 1$ (large mutational input). Here, we use the
 250 deterministic fitness trajectory corresponding to these conditions to approximate the growth rate
 251 trajectory of all replicate populations under stress ($\bar{r}_t \approx \langle \bar{r}_t \rangle$, with $\langle . \rangle$ the expectation over
 252 replicates).

253 Provided $U \gg U_c$ we thus approximate the trajectory of the mean growth rate of all replicate
 254 populations by its deterministic trajectory for the WSSM (Martin and Roques 2016):

$$\bar{r}_t^{\text{DN}} \approx \langle \bar{r}_t^{\text{DN}} \rangle \approx r_{\text{max}} - \underbrace{\frac{(r_D + r_{\text{max}})(\text{sech}(\mu t))^2}{2}}_{\text{convergence to the optimum}} - \underbrace{\frac{\bar{n}}{2} \mu \tanh(\mu t)}_{\text{Mutation load build-up}}, \quad [2]$$

255 where $\text{sech}(z) = 2/(\exp(z) + \exp(-z))$ is the hyperbolic secant, $\tanh(z) = (\exp(z) -$
 256 $\exp(-z))/(\exp(z) + \exp(-z))$ is the hyperbolic tangent and $\mu = \sqrt{U\lambda}$ is a composite parameter
 257 of the mutational parameters. In Eq.[2] the first term stems from the population nearing the
 258 phenotypic optimum, while the second term stems from the build-up of the mutation load as the
 259 genetic variance accumulates in the initially clonal population. Recall that r_D is the decay rate of
 260 the isogenic population and r_{max} is the maximum fitness that can be reached in the new
 261 environment (with $r_D + r_{\text{max}}$ the fitness distance between the parent genotype's fitness and the

262 top of the fitness peak). The mean growth rate in Eq.[2] reaches a plateau at infinite time of $r_{max} -$
 263 $n \mu/2$ corresponding to the maximal growth rate minus the mutational load.

264

265 *2. Evolutionary dynamics from an initially polymorphic population (at mutation-selection balance)*

266 The evolutionary dynamics of rescue from a polymorphic population is obtained by a similar
 267 approximation. We assume that the population is initially at mutation-selection balance in the non-
 268 stressful environment, with an arbitrary positive mean growth rate. The phenotypic distribution at
 269 the onset of stress is centered on a mean phenotype \bar{z} , whose growth rate in the new environment
 270 is used to characterize the harshness of the stress imposed. For consistency with the isogenic
 271 population model above, we thus denote this growth rate $-r_D = r(\bar{z}, \mathbf{z}_*)$, where $r_D > 0$ is the
 272 decay rate of the central genotype as in the previous section. Resistant genotypes may already be
 273 present in the population at the onset of stress (“SV”) or appear by *de novo* mutation (“DN”), or
 274 arise as combinations of these (multiple step rescue mutants). Using the same reasoning as in the
 275 previous subsection, we approximate the mean growth rate of all replicate populations by the
 276 deterministic trajectory for the WSSM, i.e. whenever $U \gg U_c$ (Martin and Roques 2016):

$$\bar{r}_t^{DN+SV} \approx \langle \bar{r}_t^{DN+SV} \rangle \approx r_{max} - \underbrace{\frac{\exp(-2 \mu t)(r_D + r_{max})}{\text{convergence to the optimum}}}_{\text{Mutation load}} - \frac{\tilde{n} \tilde{\mu}}{2} . \quad [3]$$

277 In a polymorphic population at mutation-selection balance, the presence of a mutational load
 278 implies that the mean growth rate of the population in Eq.[3] is lower than the mean growth rate
 279 of an isogenic population in the same environmental conditions, with the same central genotype
 280 ($\bar{r}_0^{DN+SV} = -r_D - n \mu/2$). Contrary to Eq.[2], the mutation load is stable (the phenotypic variance
 281 is already at equilibrium) while the convergence to the optimum is faster due to the presence of
 282 standing variance. The mean growth rate in Eq.[3] ultimately reaches the same plateau as Eq. [2]
 283 (the maximal growth rate minus the mutational load).

284 Note that in both Eqs.[2] and [3] the two mutational parameters U and λ play entirely symmetric
 285 roles through the composite parameter $\mu = \sqrt{U \lambda}$. This is because the WSSM regime amounts to

286 a diffusive approximation of the effect of mutation in fitness space (Appendix E.III of Martin and
 287 Roques 2016). Like any diffusive approximation, only the mean and variance of the mutation kernel
 288 contribute to the dynamics: in this case $U \lambda$ is the variance induced by mutation during one time
 289 unit.

290

291 III. Evolutionary rescue probability

292 As stated in introduction, we neglect evolutionary stochasticity (i.e. the variance over
 293 replicates of the mean fitness dynamic, induced by drift and mutation) but not the demographic
 294 stochasticity. The latter is approximated by an inhomogeneous Feller diffusion (see Appendix I
 295 section I) with parameters \bar{r}_t (the expected mean growth rate) and $\bar{\sigma}$ (the reproductive variance
 296 averaged across segregating genotypes, which we assume is stable over time): \bar{r}_t may vary
 297 stochastically as genotype frequencies change under the effects of drift, selection and mutation.
 298 However, as explained in the previous subsection, we approximate each replicate's fitness
 299 trajectory by its deterministic expectation under the WSSM regime: $\bar{r}_t \approx \langle \bar{r}_t^{\text{DN}} \rangle$ or $\langle \bar{r}_t^{\text{DN+SV}} \rangle$, using
 300 the relevant cases from Eqs.[2] or [3] (the validity of this approximation is tested against
 301 simulations in the Results section). Therefore, the model approximately reduces to a Feller
 302 diffusion with constant $\bar{\sigma}$ and time-inhomogeneous deterministic growth rate \bar{r}_t . Under these
 303 hypotheses, the demographic dynamics follow the stochastic differential equation: $dN_t =$
 304 $\bar{r}_t N_t dt + \sqrt{\bar{\sigma} N_t} dB_t$ where B_t is a Weiner process (see Appendix I section I for more details). We
 305 can then use the results from Theorem 1 of Appendix II (see also Bansaye and Simatos 2015) on
 306 inhomogeneous Feller diffusions to derive the probability that the population is extinct before
 307 time t :

$$P_{\text{ext}}(t) = \exp\left(-\frac{2N_0}{\bar{\sigma} \int_0^t \exp(-\rho_u) du}\right), \quad [4]$$

$$\rho_u = \int_0^u \bar{r}_v dv ,$$

308 where \bar{r}_t is given by Eqs.[2] or [3]. This expression can be evaluated numerically at any time $t > 0$.
309 ER happens whenever evolution (change in \bar{r}_t) allows extinction to be avoided. Hence, the general
310 form of the rescue probability is readily obtained as the complementary probability of the infinite
311 time limit of Eq.[4], namely the probability of never getting extinct: $P_R = 1 - P_E$ with $P_E =$
312 $P_{\text{ext}}(\infty)$, the extinction probability after infinite time. Depending on the scenarios, this rescue
313 probability can be computed explicitly, or approximated via Laplace approximations to the integral
314 $\int_0^\infty \exp(-\rho_u) du$, in some parameter ranges (see Results and Appendix I).

315

316 **IV. Individual-based stochastic simulations:** The analytical predictions are tested against exact
317 stochastic simulations of the population size and genetic composition of populations across
318 discrete, non-overlapping generations (**Supplementary Figures 8** shows examples of simulated
319 dynamics of the mean fitness and the population size). Rescue probability was estimated by
320 running 100 replicate simulations until either extinction or rescue occurred. A population was
321 considered rescued when it reached a population size N_t and mean growth rate \bar{r}_t such that its
322 ultimate extinction probability, if it were monomorphic, would lie below 10^{-12} ($\exp(-2 N_t \bar{r}_t) <$
323 10^{-12}). The simulation algorithm is described in Anciaux *et al.* (2018) and also detailed in Appendix
324 I section VIII. Briefly, the number of offspring is Poisson distributed every generation with
325 parameter e^{r_i} for genotype i , mutations occur according to a Poisson process with constant rate
326 U per capita per generation. The phenotypic effects of the mutations are drawn from a multivariate
327 normal distribution, with multiple mutations having additive effects on phenotype. Fitnesses are
328 computed according to the FGM using Eq.[1]. With such a Poisson offspring distribution, the
329 reproductive variance of genotype i is $\sigma_i = 1 + r_i \approx 1$, assuming small growth rates $r_i \ll 1$, in per-
330 generation time units. So this particular demographic model satisfies our assumption of constant
331 $\bar{\sigma}$ in spite of changes in the genotypic composition of the population: here $\bar{\sigma} = 1/K_t \sum_{i=1}^{K_t} \sigma_i \approx 1$.
332 Note also that the analytical derivations (relying on a Feller diffusion (1951) approximation)
333 approximately cover other demographic models (e.g. birth death models, see Martin *et al.* 2013;
334 Gomulkiewicz *et al.* 2017), as long as they also satisfy $\bar{\sigma}(t) \approx \bar{\sigma}$ constant.

335 For *de novo* rescue, the initial population consisted of a single genotype optimal in the environment
 336 before stress. When considering contribution from standing variance, the initial population was at
 337 mutation selection balance in the former environment (before the onset of stress). More precisely,
 338 10 replicate initial equilibrium populations were generated and 100 replicate simulations were run
 339 from each of these populations. The overall rescue probability is then the average, over the 10
 340 equilibrium populations, of the rescue probabilities estimated by the proportion of rescued
 341 populations over the 100 simulations.

342

343 **V. Appendices and supplementary information:** Analytical derivations and supplemental figures are
 344 described in Appendix I. Appendix II contains the derivations for the inhomogeneous Feller
 345 diffusion. Supplementary file 1 provides the details of the analytical derivations, the code
 346 producing the figures and the simulation code as a Mathematica© .cdf file (MATHEMATICA v. 11.3
 347 Wolfram Research 2018) which can be open using the free “CDF player” available on the Wolfram
 348 website. Supplementary file 2 provides the Matlab© (MATLAB 2015a, The MathWorks, Natick,
 349 2015) source code for the curve fitting procedure used for Eq.[8].

350

351 Results

352 **Rescue probability: general form.** In spite of their difference in the population genetics underlying
 353 ER, the two scenarios with or without standing variance yield similar expressions for the probability
 354 of ER (as shown in Eqs.(A5) and (A11-A12) in Appendix I):

$$P_R = 1 - \exp(-N_0 \omega) \quad \text{with} \quad \omega = \frac{2 \mu}{\int_0^\infty h(\tau)^{n/2} \exp(-n/(2 \epsilon) f(\tau)) d\tau}$$

$$\text{and} \quad \begin{cases} DN : f(\tau) = \tau - (1 + y_D) \tanh(\tau) , & h(\tau) = \cosh(\tau) \\ DN + SV : f(\tau) = \tau - \frac{1}{2}(1 + y_D)(1 - \exp(-2 \tau)) , & h(\tau) = \exp(\tau) \end{cases} , \quad [5]$$

355 where the particular form of the functions $f(\cdot)$ and $h(\cdot)$ depend on the chosen scenario. Here,
356 $\cosh(z) = (\exp(z) + \exp(-z))/2$ is the hyperbolic cosine and we introduce two scaled
357 parameters: $\epsilon = n \mu / (2 r_{max})$ and $y_D = r_D / r_{max}$ (recall that $\mu = \sqrt{U \lambda}$ depends on both mutation
358 rate and effect). The parameter y_D describes how fast the initial clone decays, compared to how
359 fast the optimal genotype grows, it gives a scaled measure of the harshness of the stress imposed
360 (see also Anciaux *et al.* 2018). The parameter ϵ is the ratio of the mutation load (at mutation-
361 selection balance) and the maximal absolute growth rate that can be reached in the stress. Certain
362 extinction by lethal mutagenesis occurs whenever the load is equal to or larger than the maximal
363 growth rate (i.e. $\epsilon \geq 1$). A small value of ϵ means that we are far from this certain extinction
364 regime.

365 The quantity ω in Eq.[5] is akin to a ‘per lineage rate of rescue’. Analytical expressions for this rate
366 are either unavailable (DN) or complicated ($DN + SV$, Eq.(A12) in Appendix I).

367

368 **Weak selection, intermediate mutation approximation:** To get a more direct insight into the impact
369 of each parameter, we sought an approximate expression for ω , detailed in Appendix I section III
370 and IV. This approximation applies with an intermediate mutation rate, where ER mainly depends
371 on the early adaptation of the population to stress, and not on the ultimate mutation load (lethal
372 mutagenesis). More precisely, it requires that the WSSM approximation be accurate while ϵ
373 remains small, which implies a small λ and intermediate U : $n^2 \lambda / 4 \ll U \ll r_{max}$. This is why we
374 call this range a ‘weak selection, intermediate mutation’ approximation (WSIM in **Figure 1** and **2**).

375 Under this approximation, the *per capita* rate of rescue ω takes a roughly similar form for
376 both scenarios with or without standing variance (detailed in Eqs.(A8) through (A14) in Appendix
377 I):

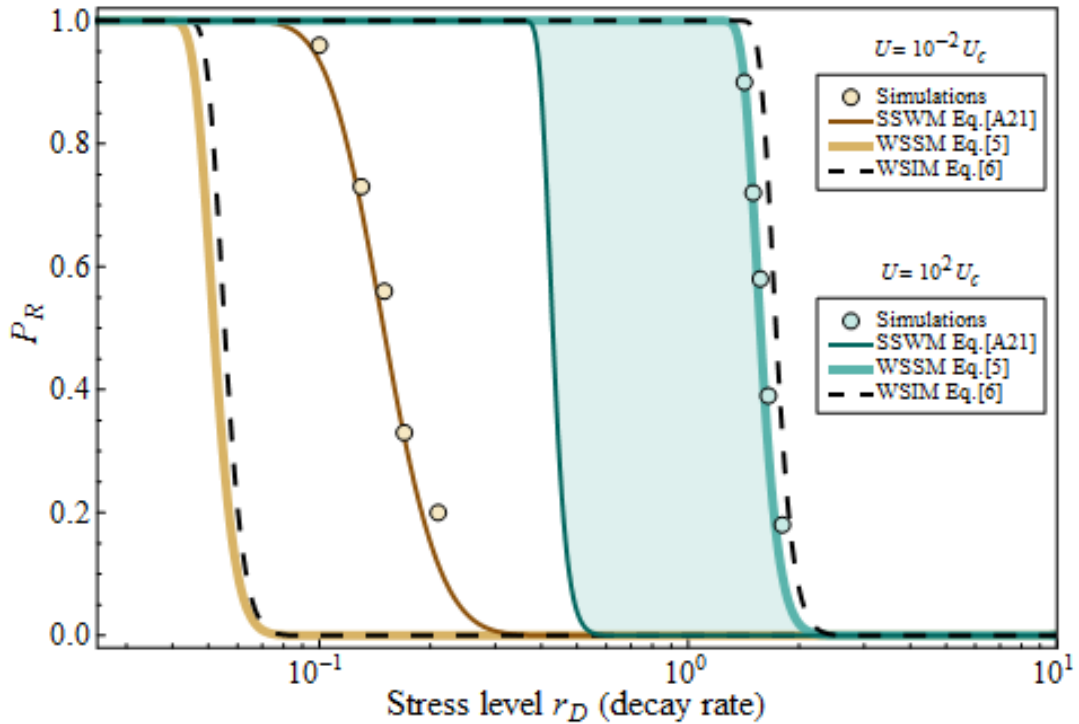
$$\omega_{\epsilon \ll 1} \approx 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} g(y_D)\right)$$

$$\text{with } \begin{cases} DN : g(y) = g_{DN}(y) = \sqrt{y(1+y)} - \cosh^{-1}(\sqrt{1+y}) \\ DN + SV : g(y) = g_{DN+SV}(y) = (y - \log(1+y))/2 \end{cases} \quad [6]$$

378 In both cases, the function $g(\cdot)$ is positive and increases (roughly log-log linearly) with $y_D =$
379 r_D/r_{max} . The accuracy of this approximation is illustrated in **Supplementary Figures 2 and 3** and
380 **Figures 1 and 2**. Note that, whenever the approximation applies, the ER probability is independent
381 of dimensionality (n). This directly stems from the fact that the growth rate trajectories in Eqs.[2]
382 and [3] only depend on dimension via the ‘mutation load’ terms, whose contribution is negligible
383 when far from the lethal mutagenesis threshold ($U \ll r_{max}$).

384

385 **Sharp decay in ER probability with increasing stress levels:** A possible measure of stress intensity in
386 ER is the rate of decay r_D of a population after the environmental change (see also Anciaux *et al.*
387 2018). However, stress might also affect other parameters of the FGM: the height of the fitness
388 peak r_{max} , the mutation rate U or the variance of mutational effects: λ . We detail the two latter
389 effects (which affect ER via the composite parameter $\mu = \sqrt{U\lambda}$), in the next section, and focus
390 here on r_{max} and r_D . In the following we use the *per capita* rate of rescue ω_{DN} from Eq.(5) of
391 Anciaux *et al.* (2018) in the SSWM regime (recalled in Appendix I section VIII, Eqs.(A20)-(A21)), as
392 a comparison to the present results in the WSSM regime. By basic properties of the FGM described
393 in Anciaux *et al.* (2018), increased r_D means both a faster decay (purely demographic effect of
394 stress) and a larger shift in optimum (which affects the whole distribution of fitness effect of
395 mutations), which both decrease the ER probability. On the contrary, increasing r_{max} increases the
396 ER probability through two effects. First, the size of the phenotypic space of resistance increases
397 with r_{max} (as in the SSWM regime, see Anciaux *et al.* 2018). Second a large r_{max} counterbalances
398 a high mutational load $n\mu/2$ as can be seen in Eq.[2] (this latter effect is only captured in the
399 WSSM approximation).



400
 401 **Figure 1:** ER probability against decay rate r_D for a population without standing variance (DN). Dots
 402 show the results of 10^2 simulations; thin plain lines: Eq.(5) from Anciaux et al. 2018 (recalled in
 403 Appendix Eq.(A21)) derived under the SSWM regime; thick plain lines: Eq.[5] derived under the
 404 WSSM regime; dashed lines: the corresponding closed form expression in Eq.[6] derived under the
 405 weak selection intermediate mutation regime (WSIM). The shaded area corresponds to the extra
 406 contribution to ER from multiple mutants, compared to single mutants. All models and simulations
 407 are shown for a high mutation rate (blue) or a low mutation rate (brown), indicated in legend. As
 408 expected, the model from Anciaux et al. (2018) derived in the SSWM regime captures the
 409 simulations at low mutation rates ($U = 10^{-2}U_c$, brown), whereas Eqs.[5] and [6] captures the
 410 simulations at high mutation rates ($U = 10^2U_c$, blue). Other parameters are $N_0 = 10^5$, $n = 4$,
 411 $r_{max} = 1$ and $\lambda = 5 \times 10^{-3}$.

412
 413 **Figure 1** illustrates how the ER probability drops sharply with the decay rate r_D , both in the
 414 SSWM regime (brown curves, $U \ll U_c$, see legend) and in the WSSM regime (blue curves, $U \gg$
 415 U_c). This qualitatively similar behavior is *a priori* due to common geometric constraints imposed by
 416 the FGM. The model of (Anciaux et al. 2018) and the present model, derived under complementary
 417 approximations (SSWM *versus* WSSM), capture the results of simulations in a different and
 418 complementary portion of the range of possible mutation rates: compare the blue ($U = 10^2U_c$)

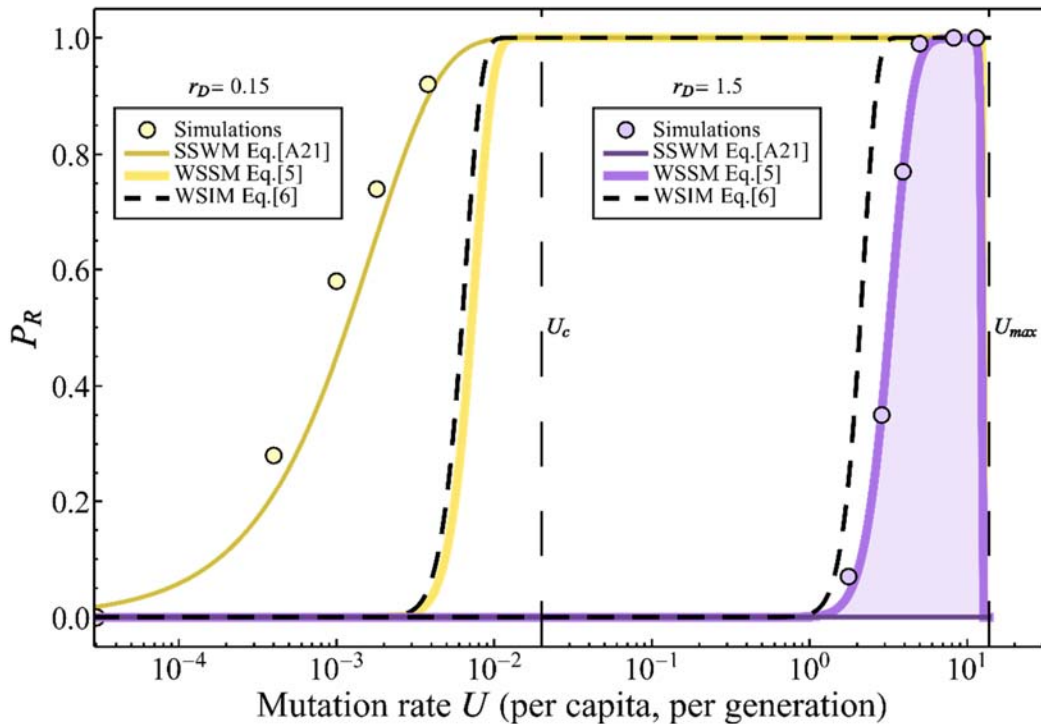
419 vs. brown ($U = 10^{-2}U_c$) curves to the dots in **Figure 1**. Higher mutation rates allows higher stress
420 levels (larger r_D) to be endured, but it is not their only effect, as we now detail.

421

422 **Non-monotonic relationship between ER probability and mutational parameters:** In the following
423 section, we investigate the effect of mutational parameters. Both the mutation rate U and the
424 variance of mutational effects λ affect the system in a similar fashion through the composite
425 parameter $\mu = \sqrt{U\lambda}$. At small t (in Eq.[4]), an increase in μ speeds-up the early adaptive process,
426 thus favoring rescue but also increases the ultimate mutation load, favoring extinction by lethal
427 mutagenesis. These antagonistic effects of μ create a non-monotonic relationship between the
428 rescue probability and mutational parameters. This is illustrated in **Figure 2**, which also shows how
429 Eq.[5] (thick lines) captures this effect: not neglecting the parameter $\epsilon = n\mu/(2r_{max})$ makes the
430 relationship between μ and P_R non-monotonic. For low stress r_D , P_R is maximal and approximately
431 equal to 1 over a range of mutation rates (plateau **Figure 2**, see also **Supplementary Figure 6** for
432 higher stress values). Beyond this range, the rescue probability in Eq.[5] drops to 0 at $U = U_{max} \equiv$
433 $(4r_{max}^2)/(n^2\lambda)$. U_{max} is the mutation rate beyond which certain extinction is enforced by lethal
434 mutagenesis because the mutation load ($\mu n/2$) is larger than the maximal growth rate that can
435 be reached in the stress (r_{max}). Hence, even if ER allowed the population to invade the new
436 environment, it could not generate a stable population once at mutation-selection balance.

437 Note that Eq.[6], which is only valid for intermediate μ , does not capture the decrease in P_R close
438 to $U = U_{max}$. It does however capture the increase in P_R as mutation rate increases, far
439 below U_{max} .

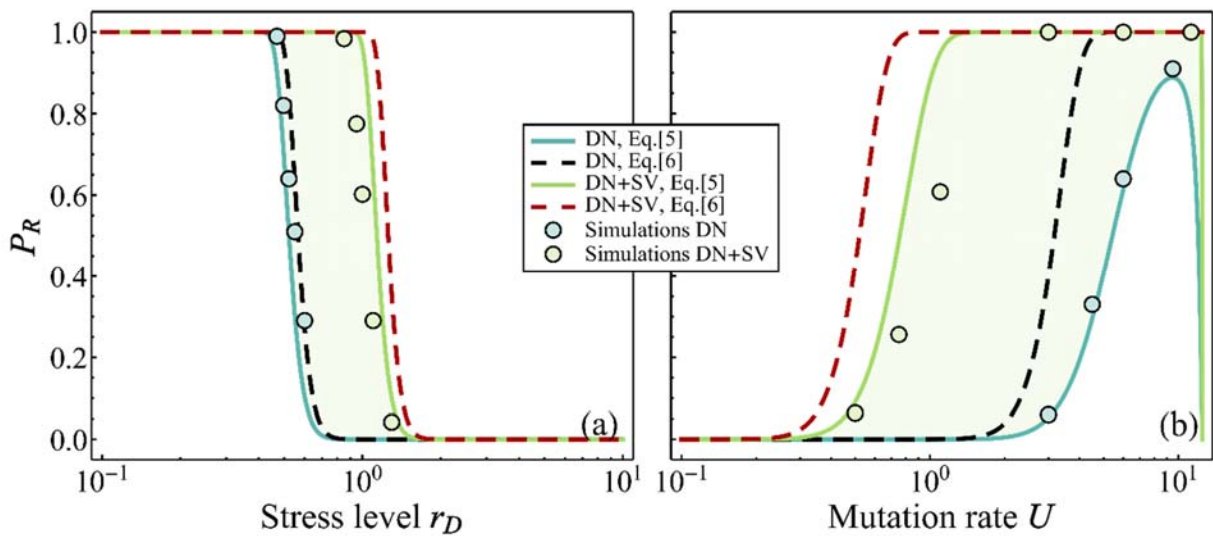
440



441
 442 **Figure 2:** ER probability against mutation rate U for a population without standing variance (DN).
 443 Dots show the results of 10^2 simulations; thin plain lines: Eq.(5) from Anciaux et al. 2018 (recalled
 444 in Appendix Eq.(A21)) derived under the SSWM regime; thick plain lines: Eq.[5] derived under the
 445 WSSM regime; dashed lines: the corresponding closed form expression in Eq.[6] derived under the
 446 weak selection intermediate mutation regime (WSIM). The shaded area corresponds to the extra
 447 contribution to ER from multiple mutants, compared to single mutants. All models and simulations
 448 are shown for a high decay rate (purple) or a low decay rate (yellow), indicated in legend. For low
 449 mutation rates, corresponding to the SSWM regime, simulations show that populations handle only
 450 low stresses, which is captured by the model from Anciaux et al. (2018) derived in the SSWM regime.
 451 Whereas for high mutation rates, corresponding to the WSSM regime, simulations show that
 452 populations handle high stresses, which is captured by the present model derived in the WSSM
 453 regime. Other parameters are $N_0 = 10^5$, $n = 4$, $r_{max} = 0.5$ and $\lambda = 5 \times 10^{-3}$.

454
 455 **Evolutionary dynamics from an initially polymorphic population (at mutation-selection balance):** The
 456 results presented in the previous figures illustrate rescue from *de novo* mutations. In the presence
 457 of additional standing genetic variance, rescue mutants can arise from *de novo* mutants, from pre-
 458 existing genotypes or from a combination of both. **Figure 3** shows the qualitative similarity between
 459 the case with and without standing genetic variance, in their dependence on r_D and U , as observed
 460 in simulations and captured by Eqs.[5] and [6]. Indeed, **Figure 3** confirms that the addition of

461 standing genetic variance does not qualitatively modify the relationship between the rescue
 462 probability and stress intensity (r_D , **Figure 3a**) or mutational parameters (here U , **Figure 3b**). Note
 463 that the accuracy of Eq.[5] is lower for higher r_{max} , where the continuous time approximations
 464 become less accurate to capture discrete time simulations (see **Supplementary Figure 4**). In the
 465 next sections, for the sake of clarity and simplicity, we will mainly discuss the scenario of ER from
 466 *de novo* mutations only, as the qualitative behaviors are similar with an extra contribution from
 467 standing variance.



468 **Figure 3:** ER probabilities from *de novo* mutants only (blue: Eq.[5], black dashed line: Eq.[6]) or from
 469 both *de novo* and pre-existent mutants (green: Eq.[5], red dashed line: Eq.[6]) against the stress
 470 level r_D (a) or the mutation rate U (b). Dots show the results of 10^2 simulations (started from 10
 471 simulated populations at mutation-selection balance for the DN + SV scenario). The shaded area
 472 corresponds to the contribution from the standing genetic variance to the rescue compared to *de novo*
 473 mutation. Other parameters are $N_0 = 10^5$, $r_{max} = 0.5$, $n = 4$ and $\lambda = 5 \times 10^{-3}$. $U = 10 U_c$
 474 in panels (a) and $r_D = 1.8$ in panel (b).
 475

476

477 **Mutation window for ER:** In the previous subsections, we have shown that the ER probability drops
 478 sharply with increasing stress and is maximal over a finite range of mutation rates, which we denote
 479 “mutation window” for ER. The “width” (range of mutation rates) and “height” (maximum of ER
 480 probability over the range) of this window strongly depends on stress.

481 Width of the window: To characterize the mutation window, its upper and lower bounds must be
 482 defined. The lower bound of the window (denoted U_*) corresponds to the mutation rate at which
 483 the rescue probability rises to $1/2$. Thus, this lower bound is only defined if the height of the
 484 window lies above or at $1/2$ (i.e. $\max(P_R) \geq 1/2$). The upper bound is set to the mutation rate
 485 U_{max} beyond which certain extinction is enforced by lethal mutagenesis. The ER probability drops
 486 off very sharply close to U_{max} , so that, approximatively, ER is only likely within the mutation
 487 window $U_* \leq U \leq U_{max}$. These two bounds are derived in Appendix I Eq. (A17):

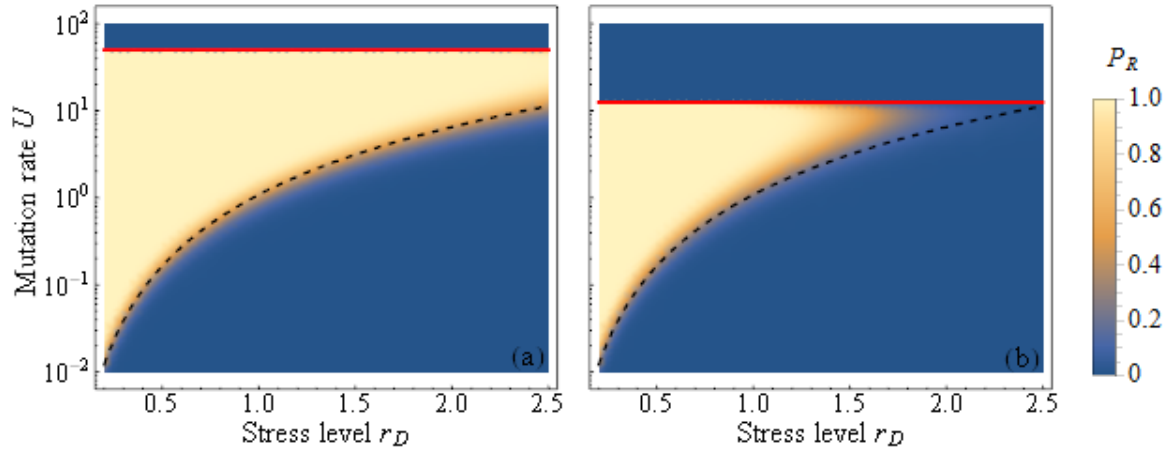
$$U_{max} = \frac{4 r_{max}^2}{n^2 \lambda}$$

$$U_* \approx \frac{1}{\lambda} \left(\frac{2 g(y_D) r_{max}}{\mathcal{W} \left(\frac{8 g(y_D)}{\pi} \left(\frac{N_0 r_{max}}{\log(2)} \right)^2 \right)} \right)^2 \underset{N_0 \rightarrow \infty}{\approx} \frac{4 g(y_D)^2 r_{max}^2}{\lambda \log \left(\frac{N_0^2}{\log(N_0^2)} \right)^2} \quad [7]$$

488 where $\mathcal{W}(\cdot)$ is the Lambert W function, which converges to $\mathcal{W}(z) \approx \log(z/\log(z))$ as z gets large
 489 (yielding the right hand side approximation above). Here, $g(y_D)$ is the function of stress intensity
 490 given in Eq.[6], which describes how stress intensity ($y_D = r_D/r_{max}$) affects ER rates. Depending
 491 on the scenario, one uses $g(y_D) = g_{DN+SV}(y_D)$ or $g(y_D) = g_{DN}(y_D)$ in the presence or absence
 492 of standing genetic variance, respectively (note that the window is always wider in the former
 493 case).

494 **Figure 4** illustrates, for an initially clonal population, the sharpness of the transition from ER being
 495 almost certain to being highly unlikely ($P_R = 1$ to $P_R = 0$) as a function of U and r_D . It also shows
 496 the accuracy of the approximation for U_* in Eq.[7] (dashed black line corresponding to the right
 497 hand side approximation of Eq.[7]), compared to its numerical estimation from Eq.[5] (color
 498 gradient). **Supplementary Figure 5** shows a similar result for a population starting with standing
 499 genetic variance.

500



501

502 **Figure 4:** ER probabilities from de novo mutants only (Eq.[5]) for different values of r_D and U . The
503 color gradient gives the value of P_R (see legend). The red straight line corresponds to $U = U_{max}$
504 (Eq.[7]) and the black dashed line corresponds to $U = U_*$ ($P_R = 1/2$) (right hand side
505 approximation of Eq.[7]). For a given U , the ER probability drops sharply from $P_R = 1$ (light yellow)
506 to $P_R = 0$ (blue), over a short increase in r_D . For a given r_D , P_R rises sharply as U increases about U_*
507 and then drops sharply as U increases around U_{max} . Other parameters are $r_{max} = 1$, $N_0 = 10^4$
508 and $\lambda = 5 \times 10^{-3}$. $n = 4$ in panel (a) and $n = 8$ in panel (b).

509

510 The upper bound U_{max} is independent of initial conditions or stochasticity, as lethal
511 mutagenesis depends on the deterministic equilibrium state of the population, once adapted to
512 the stress. Therefore, U_{max} does not depend on the presence or absence of initial standing
513 variance, the decay rate imposed by the environmental change (γ_D) or the initial population
514 size N_0 . On the contrary, the lower bound U_* depends on these factors as it is determined by the
515 capacity of the population to transiently adapt to the new conditions. It shows, however, little
516 dependence on dimensionality (n), as is also apparent in **Figure 4**, by the accuracy of Eq.[7], which
517 is independent of n . Overall, the width of the mutation window where ER is likely decreases with
518 increasing stress r_D (**Figure 4**) and increases with initial population size N_0 (eq. [7]: U_* decreases
519 with N_0 , and U_{max} is unchanged). The width of the window decreases with dimensionality n and
520 increases with the maximum fitness in the new environment r_{max} , because both parameters affect
521 the upper bound of the window, U_{max} (eq. [7]). Lower dimensionality and higher maximum
522 population growth rate reduce the mutation load, and thus allow persisting under higher mutation
523 rates.

524 Finally, note that we focused on the effect of U here, but similar results could be obtained if λ was
 525 varied (as both parameters affect P_R as a product $\mu = \sqrt{U\lambda}$). This is apparent in Eq.[7] where U
 526 and λ could be exchanged.

527

528 Height of the window and “Mutation proof extinction”: within the mutation window ($U_* \leq U \leq$
 529 U_{max}), the ER probability P_R rises above 50% and then drops back to zero. Yet, for more extreme
 530 stresses, it cannot even reach above 50% for any value of the mutation rate: the height of the
 531 mutation window lies below 1/2). When this height is low, extinction is ‘mutation proof’, in that it
 532 is highly likely whatever the mutation rate(s) U or the variance of mutational effects λ in the
 533 population. To illustrate this, we compute the maximum of the ER probability $\max(P_R)$ when U
 534 varies from U_c to U_{max} , by numerically evaluating Eq.[5] over this range. **Supplementary Figure 6**
 535 shows detailed profiles of ER probabilities against mutation rates (illustrating how $\max(P_R)$ is
 536 found), in the presence or absence of initial standing variance. **Figure 5** shows the maximum P_R
 537 attainable as a function of r_D and N_0 : it drops (transition from yellow to blue areas) with increasing
 538 stress r_D and decreasing population size N_0 . In this example, a large part of the combinations of
 539 the two parameters N_0 and r_D correspond to $\max(P_R)$ lower than 10% (blue area below the lower
 540 white dashed line in **Figure 5**). Therefore, for a given inoculum size N_0 , there is always a threshold
 541 of stress level beyond which ER is nearly impossible, whatever the mutation rates in the population.

542 A closed form approximation can be obtained to describe this transition (detailed in
 543 Appendix I section VII): denote $r_D^*(p)$ the value of r_D at which $\max(P_R) = p$ for some $p \in [0,1]$.
 544 We obtain the following simple expression for the threshold of level of stress beyond
 545 which $\max(P_R)$ cannot exceed some level p , independently of μ :

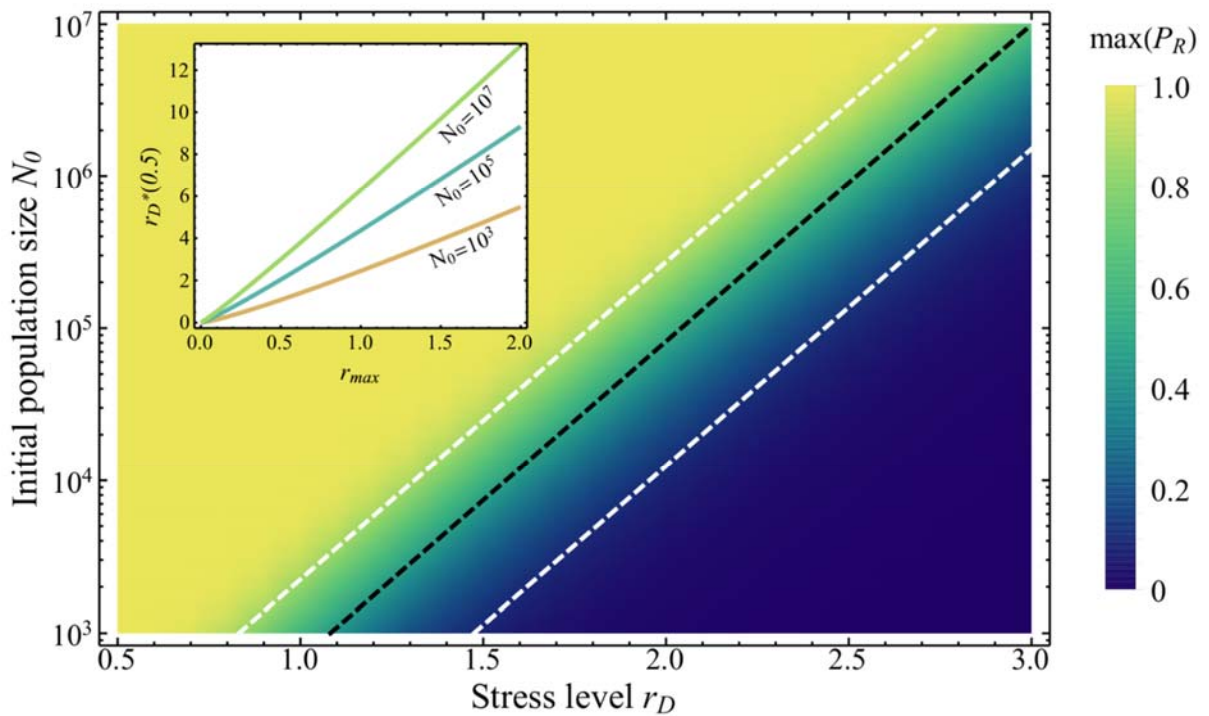
$$r_D^*(p) = \frac{r_{max}}{n \delta} \log \left(\frac{N_0 r_{max}}{n \log(1/(1-p))} \right) \text{ with } \begin{cases} DN : \delta = 0.6 \\ DN + SV : \delta = 0.31 \end{cases}, \quad [8]$$

546 where the values of δ come from a curve-fitting procedure (detailed in Appendix I section VII).
 547 Setting $p \ll 1$ in Eq.[8] thus provides the stress level beyond which ER is very unlikely, whatever
 548 the mutation rate U or the variance of mutational effects λ . This means, in particular, that the

549 evolution of higher mutation rates (via hypermutator strains) or higher variance of mutational
 550 effects (larger λ) would not allow the population to avoid extinction, when confronted to this stress
 551 level. The validity of the heuristic in Eq. [8] is illustrated in **Figure 5**, where we see that the dashed
 552 lines ($r_D^*(p)$ with $p = \{0.1, 0.5, 0.9\}$ see legend) accurately predict the transition from high to low
 553 values of $\max(P_R)$, computed numerically from Eq.[5].

554 This whole argument applies to both DN and $DN + SV$ scenarios (by choosing δ
 555 accordingly in Eq. [8]). Interestingly, we can also see that populations initially at mutation-selection
 556 balance (standing genetic variance) can withstand stresses twice larger than populations only
 557 adapting from *de novo* mutations (initially clonal).

558



559
 560 **Figure 5:** Maximum ER probability reached as U is varied, for different values of r_D and N_0 for a
 561 population with no initial polymorphism. The color gradient gives the value of $\max(P_R)$ this time
 562 (see legend). The black dashed line gives the value of $r_D^*(0.5)$ (Eq.[8] with $\delta = 0.6$) and the white
 563 dashed lines the value of $r_D^*(0.1)$ and $r_D^*(0.9)$. The maximum of the ER probability attainable (for
 564 all possible U) drops sharply over a short range of increasing r_D for a given N_0 , or over a short range

565 of decreasing N_0 for a given r_D . Other parameters are $r_{max} = 0.5$, $n = 4$, and $\lambda = 5 \times 10^{-3}$. The
 566 inset panel shows how $r_D^*(0.5)$ from Eq.[8] varies with r_{max} and N_0 .

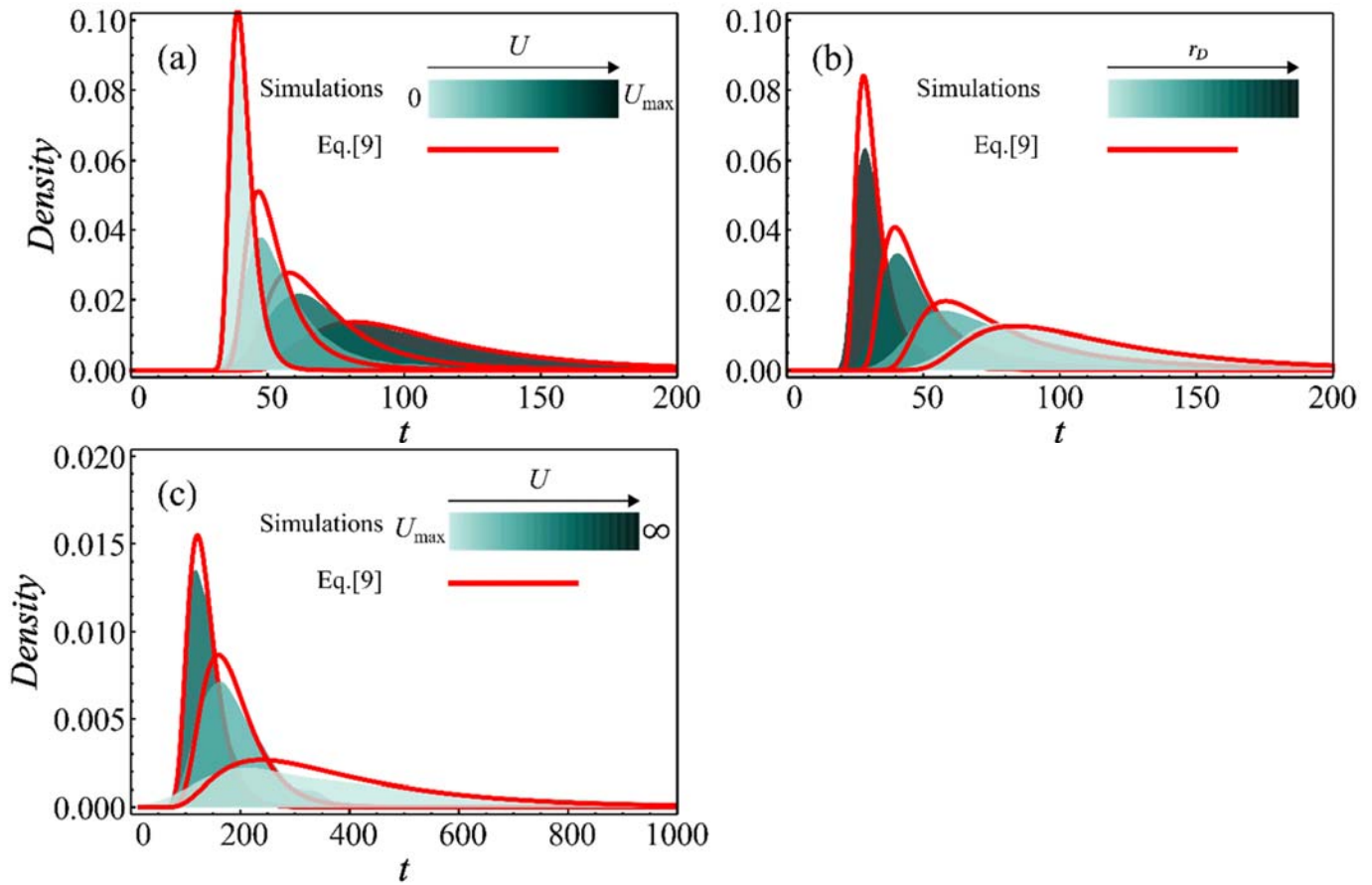
567

568 **Distribution of extinction times:** From Eq.[4] we can derive the probability density $\varphi(t)$ of this
 569 distribution, in either of the two scenarios considered (purely clonal population DN or population
 570 at mutation-selection balance $DN + SV$). We get:

$$\varphi(t) = \frac{\partial_t P_{ext}(t)}{P_{ext}(\infty)} = 2 N_0 \frac{P_{ext}(t)}{P_{ext}(\infty)} \frac{\Psi'(t)}{\Psi(t)^2} , \quad [9]$$

$$\Psi(t) = \int_0^t h(\mu v)^{n/2} \exp\left(-\frac{n}{2\epsilon} f(\mu v)\right) dv$$

571 where the functions $f(\cdot)$ and $h(\cdot)$ depend on the scenario considered and are given explicitly in
 572 Eq.[5]. **Figure 6** illustrates the accuracy of this result and how the distribution of extinction times
 573 varies with stress intensity r_D and mutation rate U . In spite of neglecting evolutionary stochasticity,
 574 Eq.[9] still captures the shape and scale of extinction time distributions, in the WSSM regime.



575

576 **Figure 6:** Density of the extinction probability dynamics for different values of U (panel (a) and (c))
 577 and r_D (panel (b)). The distributions of extinction times from simulations started with an isogenic
 578 population are shown by shaded histograms, with the corresponding theory (Eq.[9]) given by the
 579 plain red lines. The color gradient corresponds to increasing levels of r_D (panel (b)) or of U (lower
 580 range in panel (a) and higher range in panel (c), as indicated on the legend). Each simulated
 581 distributions is drawn from 1000 extinct populations among a sufficient number of replicates to
 582 observe these 1000 extinctions. The mutation rates covered in panel (a) are $U = \{0; 4 U_c; 8 U_c; 14 U_c\}$
 583 and in panel (c) are $U = \{2 U_{max}; 1.5 U_{max}; 1.1 U_{max}\}$ and in both panel
 584 the decay rate is $r_D = 0.28$. The decay rate covered in panel (b) are $\{0.285; 0.33; 0.4; 0.5\}$ and the
 585 mutation rate is $U = 15 U_c$. Other parameters are $N_0 = 10^5$, $n = 4$, $r_{max} = 0.1$ and $\lambda =$
 586 5×10^{-3} .

587

588 **Figure 6b** shows that decreasing the stress level (r_D) increases the mean persistence time
 589 of the population and also increases the variance of this duration. This behavior could be seen as
 590 a mere scaling: even in the absence of evolution the population takes longer time to become
 591 extinct with a smaller r_D . On the contrary, increasing the mutation rate, keeping it below the lethal

592 mutagenesis threshold ($0 \leq U \leq U_{max}$, **Figure 6a**), increases the mean and the variance of the
593 persistence time. This likely stems from subcritical mutations ($0 > r > r_D$, beneficial but not
594 resistant) that can transiently invade the population, thus delaying its extinction. However, beyond
595 the lethal mutagenesis threshold ($U > U_{max}$, **Figure 6c**), the trend is reversed: extinctions (which
596 are always certain then) occur faster at high mutation rates (panel c). Therefore, even in those
597 cases where ER probabilities are uninformative ($P_R = 0$), the distribution of extinction times
598 conveys important information on the underlying adaptive or maladaptive dynamics.

599

600 Discussion

601 We investigated the effect of an abrupt environmental change on the persistence of
602 asexual populations with a large mutational input of genetic variance (WSSM regime), adapting
603 either from *de novo* mutations arising after the environmental change (*DN* scenario) or from both
604 *de novo* and pre-existing mutations (*DN + SV* scenario). In a previous study (Anciaux et al. 2018),
605 we studied evolutionary rescue when considering adaptation over a phenotype-fitness landscape
606 (FGM), which implies pervasive epistasis between multiple mutations and imposes a relationship
607 between the initial decay rate of the population, the proportion and growth rate of resistance
608 alleles, and their selective cost in the ancestral (before the stress) environment. However, we
609 assumed that rescue resulted from rare mutations with strong effects (SSWM). The key
610 contributions of the present model, building on our previous work, are to (i) allow for the
611 cumulative effect of multiple mutations and (ii) provide insights into the distribution of extinction
612 times in the presence of an evolutionary response.

613

614 **Single step (SSWM regime) vs. multiple step (WSSM) regimes in ER:** In spite of its complexity, the ER
615 process in high mutation rate regimes can readily be captured by simple analytic approximations
616 (Eqs.[5] and [6]), which neglect evolutionary stochasticity and only account for demographic
617 stochasticity. Overall, the SSWM and WSSM approximations roughly capture the ER process in
618 complementary domains of the mutation rate spectrum (**Figures 1-3**).

619 This approach shows how multiple mutations allow withstanding higher stress than what
620 the single step approximation (SSWM in Anciaux *et al.* 2018) predicts (**Figures 1 and 3**). However,
621 this is only true for intermediate mutation rates: a further increase in mutation rate ultimately
622 shifts the system to a lethal mutagenesis regime (**Figures 2 to 4**). Indeed, the dependence between
623 the ER probability and the mutation rate is not monotonic. The model shows an optimal mutation
624 rate for the ER probability, at which the maximal ER probability may be less than 1 (depending on
625 the stress, **Figure 5**). Beyond this rate, the ER probability drops down, to some point (U_{max} , Eq. [7])
626 where the mutation load is so large that absolute fitness is negative at mutation selection balance.
627 This non-monotonic dependence reflects the continuum between ER and lethal mutagenesis along
628 a gradient of mutation rates.

629

630 **Similarities to existing ER models:** Some of our key previous findings (Anciaux *et al.* 2018), regarding
631 how ER depends on the parameters of the fitness landscape (FGM here) are still valid in the more
632 polymorphic WSSM regime. The main common features are (i) the sharp decrease of ER
633 probabilities with stress (decay rate r_D), (ii) their log-linear increase with initial population size N_0 ,
634 (iii) the fact that standing variance allows withstanding higher stress, (vi) the limited effect of
635 dimensionality n (for mutation rates far from lethal mutagenesis, eq. [6]). The effect of initial
636 population size ($P_R = 1 - \exp(-N_0 \omega)$ eq.[5]) is exactly the same as in previous models where ER
637 stemmed from single mutants (SSWM regime Orr and Unckless 2008, 2014; Martin *et al.* 2013;
638 Anciaux *et al.* 2018). This is expected of any model ignoring interactions between individuals, be it
639 evolutionary (e.g. sexual reproduction or frequency-dependent selection) or demographic (e.g.
640 density-dependence): each of the N_0 lineages initially present contributes independently to ER
641 (with some rate ω per individual). Decay rate has a broadly similar (but quantitatively different)
642 effect in previous ER models not based on a fitness landscape. The other parameters (r_{max} , n , λ)
643 are not defined outside the FGM. More generally, the key implications of considering the FGM to
644 model ER are detailed more thoroughly in Anciaux *et al.* (2018).

645

646 **Experimental test and parametrization:** To experimentally test the predictions, the assumptions of
647 the WSSM regime *a priori* imply the need to use organisms with relatively high mutation rates,
648 such as viruses, highly mutating strains of bacteria or possibly cancer cells. How large should the
649 mutational parameters be for the WSSM to apply? Empirically mutation effects and rates, at least
650 in microbes, are typically scaled by growth rates. In our model, if this growth rate is the maximal
651 growth rate (that of the wild type adapted to the lab environment), and assuming the lab
652 environment and the stressful environment have the same r_{max} , then, these scaled parameters
653 would correspond to $\bar{s} \equiv n \lambda / (2 r_{max})$ and $u \equiv U / r_{max}$. If r_{max} is approximately equal to a birth
654 rate (i.e. the optimal 'wild-type' has a small death rate in the lab environment) \bar{s} is akin to a mean
655 mutation effect per division, while u is akin to a mutation rate per division. Expressed in these
656 parameters, the WSSM approximation (Eq.[5]) applies whenever $u \gg u_c \equiv \bar{s} n / 2$, lethal
657 mutagenesis occurs when $u = u_{max} \equiv 2 / (n \bar{s})$ and the intermediate mutation weak selection
658 approximation (Eq.[6]) is valid when $u_c \ll u \ll 1$ (see the section weak selection intermediate
659 mutation approximation from the results). Estimates of dimensionality n are typically not very large
660 based on mutation fitness effects analyzed under the FGM (Martin and Lenormand 2006; Perfeito
661 et al. 2014). Therefore, based on this very rough analysis, the WSSM ER model proposed here
662 would apply in strains with mutation rates u higher than mutation effects \bar{s} , measured in the same
663 per division time units.

664 For a proper experimental test of the predictions, the parameters N_0 , U , r_D , r_{max} , λ and n
665 must be measured, in the stressful environment(s) where ER is studied. The methods and
666 challenges in estimating these parameters are discussed in Anciaux *et al.* (2018). Note however
667 that (i) the WSSM results only depend on the product $\mu = \sqrt{U \lambda}$ and that (ii) an error on the
668 dimensionality may not be critical given its relatively small effect (except in setting U_c above which
669 the model applies). The composite parameter μ can be estimated directly from mutation
670 accumulation (MA) experiments as $\mu = \sqrt{U \lambda} = \sqrt{2 \Delta M / n}$ where $\Delta M = U E(s)$ is the change in
671 mean fitness per unit time in the MA experiment ($E(s) = n \lambda / 2$ is the mean selection coefficient
672 of random mutations).

673

674 Finally, if the model is valid, fitting an observed distribution of extinction times (**Figure 6**)
675 might provide estimates of r_D , r_{max} , n and possibly μ (note however that large or low values of
676 $\mu > \mu_c$ could produce similar distributions, see **Figure 6**). This would allow using not only the
677 information from rescued populations but that from extinct ones. Such empirical test would
678 require fine-scale time series of the population size or at least the extinction status over time.

679
680 **Treatment against pathogens, hyper-mutators and lethal mutagenesis:** Our model as well as
681 previous ones all suggest that the effectiveness of a given treatment depends on the mutation rate
682 of the organism. Polymorphism for mutation rate and invasion of hyper-mutator genotypes are
683 thus potentially important issues for treatments against pathogens. However, our results suggest
684 that a sufficiently strong stress could be effective in spite of hyper-mutator evolution. Indeed,
685 because of the lethal mutagenesis effect (**Figure 2**), ER is only possible within a mutation rate
686 window: a hyper-mutator would have to hit this window to be advantageous, and the width of the
687 window narrows with increased stress (**Figure 4**). At sufficiently higher stress levels, ER is unlikely
688 whatever the mutation rate (**Figure 5**), making these strong treatments robust to hyper-mutator
689 evolution. Whether this pattern is confirmed empirically and whether the required treatment
690 levels are then not too harmful for the treated subject remain open questions. Note also that the
691 same line of argument could be used, not for mutation rate evolution, but for the evolution of the
692 variance of mutational effects, as the end result depends on the product $\mu^2 = U \lambda$.

693 Our model covers the continuum from stress induced extinction to extinction induced by
694 lethal mutagenesis. The latter might be an option, especially for organisms for which no “stress
695 treatment” exists, or whose high mutation rates (above U_* in our model) allows them to withstand
696 even strong stresses. Our results indeed confirm that increasing the mutation rate in this context
697 (above U_{max}) will allow to completely eliminate the population. The addition of a stressor might
698 also help in the process, as has been suggested before (Pariante et al. 2001, 2003, 2005): indeed,
699 the ER probability does drop faster with increasing U (as we approach U_{max}) in the presence of a
700 strong stress (with high y_D), see **Supplementary Figure 6**.

701

702 **Treatment duration issues:** One may also wonder how long a treatment must last (be it by stress
703 effect or by lethal mutagenesis) for it to be efficient. From the distribution of extinction times
704 (**Figure 6**), it is possible to predict the duration of the treatment needed to get rid of, say, 99% of
705 the pathogen populations. Strong stresses (dark blue histogram in **Figure 6b**) tend to both decrease
706 the overall ER probability and shrink the distribution of times to extinction. This means they are
707 more efficient overall and require shorter treatment durations, with less risk associated with
708 imperfect treatment compliance. However, when the mutation rate increases towards U_{max}
709 (treatment by lethal mutagenesis), although extinction becomes highly likely ($P_R \rightarrow 0$, **Figure 2**),
710 extinction times get more variable and the mode of the distribution is higher (**Figure 6a**). If the
711 mutation rate increases beyond U_{max} , the opposite pattern is observed: the distribution of
712 extinction times shrinks as U increases further away from U_{max} (**Figure 6c**). Hence, a treatment by
713 lethal mutagenesis, even if guaranteeing total extinction after infinite time, may need to be applied
714 for a long time to significantly decrease ER probability (if the resulting U is close to U_{max}). Thus,
715 observed distributions of extinction times hold valuable information.

716

717 **Limits of the model:** Obviously, the different results and applications described above are limited
718 by the hypotheses of the model.

719 First, the model ignores density (discussed in Anciaux et al. 2018) or frequency-dependent
720 effects. Hence, our model cannot describe “competitive release” effects which have been
721 discussed in the context of antibiotic resistance (Read et al. 2011; Day and Read 2016), whereby
722 higher stresses, by eliminating sensitive genotypes more rapidly, can “release” limiting resources
723 for resistant ones, and ultimately increase ER probability. However, we do not expect *a priori* the
724 present model to show such a decrease in ER probability at higher stress levels. In the FGM in
725 particular, given the very sharp decay in P_R with stress level r_D , it is highly likely that density-
726 dependence (hence room for competitive release) will only mitigate the result and that ER
727 probability will still decrease at higher stress. However, this must be studied quantitatively in a
728 dedicated model, which would account for density-dependence.

729 The approximation of the demographic dynamics by a Feller diffusion also imposes that
730 demographic variations remain smooth, which may be inaccurate, e.g. for some viruses showing
731 occasional large burst events. The Feller diffusion can also fail to predict discrete time demography
732 (used here in the simulations) at high growth rates per generation ($r \geq 1$).

733 Second, in the demographic dynamics, we also make the approximation that σ is roughly
734 constant across genotypes and thus constant over time. This approximation can be accurate if the
735 mutant growth rates are typically small compared to their reproductive variance $r_i \ll \sigma_i$, which is
736 valid as long as $b_i - d_i \ll b_i + d_i$ (with b_i (resp. d_i) the birth rate (resp. the death rate) of each
737 genotype; discussed in Martin *et al.* 2013). However if mutations affect substantially but
738 dissimilarly the birth and death rates this approximation might fail. This could be handled by
739 considering the deterministic time dynamics of the mean reproductive variance $\bar{\sigma}(t)$ in addition
740 to that of the mean growth rate $\bar{r}(t)$, in the inhomogeneous Feller diffusion.s

741 Third our results assume a sharp change in the environment, which does not reflect all
742 forms of stresses, including in the context of sudden treatment (e.g. antibiotics can have complex
743 pharmacokinetic patterns over time Regoes *et al.* 2004).

744

745 **Conclusion:** Recently, lethal mutagenesis treatments have received renewed interest for their
746 ability to provide alternative treatments against viruses resistant to other type of drugs (Jiang *et*
747 *al.* 2016; Escibano-Romero *et al.* 2017). The present model attempts to provide a framework to
748 predict the efficacy of such treatments. The strategy used here of coupling deterministic
749 evolutionary trajectories with stochastic demographic dynamics could in principle be applied to
750 other models, to cope with a large input of mutations affecting the ER process. We hope that future
751 experimental tests will evaluate its accuracy and potential to tackle various pressing applied issues.

752

753 Bibliography

- 754 Alexander, H. K., G. Martin, O. Y. Martin, and S. Bonhoeffer. 2014. Evolutionary rescue: linking theory for
755 conservation and medicine. *Evol. Appl.* 7:1161–1179.
- 756 Anciaux, Y., L.-M. Chevin, O. Ronce, and G. Martin. 2018. Evolutionary Rescue over a Fitness Landscape.
757 *Genetics* 209:265–279.
- 758 Arias, A., L. Thorne, and I. Goodfellow. 2014. Favipiravir elicits antiviral mutagenesis during virus
759 replication in vivo. *eLife* 3.
- 760 Bank, C., N. Renzette, P. Liu, S. Matuszewski, H. Shim, M. Foll, D. N. A. Bolon, K. B. Zeldovich, T. F. Kowalik,
761 R. W. Finberg, J. P. Wang, and J. D. Jensen. 2016. An experimental evaluation of drug-induced
762 mutational meltdown as an antiviral treatment strategy. *Evolution* 70:2470–2484.
- 763 Bansaye, V., and F. Simatos. 2015. On the scaling limits of Galton-Watson processes in varying
764 environments. *Electron. J. Probab.* 20.
- 765 Bell, G. 2017. Evolutionary Rescue. *Annu. Rev. Ecol. Evol. Syst.* 48:605–627.
- 766 Blanquart, F., and T. Bataillon. 2016. Epistasis and the Structure of Fitness Landscapes: Are Experimental
767 Fitness Landscapes Compatible with Fisher’s Geometric Model? *Genetics* 203:847–862.
- 768 Bull, J. J., R. Sanjuán, and C. O. Wilke. 2007. Theory of Lethal Mutagenesis for Viruses. *J. Virol.* 81:2930–
769 2939.
- 770 Bull, J. J., and C. O. Wilke. 2008. Lethal Mutagenesis of Bacteria. *Genetics* 180:1061–1070.
- 771 Carlson, S. M., C. J. Cunningham, and P. A. H. Westley. 2014. Evolutionary rescue in a changing world.
772 *Trends Ecol. Evol.* 29:521–530.
- 773 Couce, A., A. Rodríguez-Rojas, and J. Blázquez. 2015. Bypass of genetic constraints during mutator
774 evolution to antibiotic resistance. *Proc R Soc B* 282:20142698.
- 775 Day, T., and A. F. Read. 2016. Does High-Dose Antimicrobial Chemotherapy Prevent the Evolution of
776 Resistance? *PLOS Comput. Biol.* 12:e1004689.

777 Eliopoulos, G. M., and J. Blázquez. 2003. Hypermutation as a Factor Contributing to the Acquisition of
778 Antimicrobial Resistance. *Clin. Infect. Dis.* 37:1201–1209.

779 Escribano-Romero, E., N. J. de Oya, E. Domingo, and J. C. Saiz. 2017. Extinction of West Nile Virus by
780 Favipiravir through Lethal Mutagenesis. *Antimicrob. Agents Chemother.* 61:e01400-17.

781 Feder, A. F., S.-Y. Rhee, S. P. Holmes, R. W. Shafer, D. A. Petrov, and P. S. Pennings. 2016. More effective
782 drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1. *Elife* 5:e10670.

783 Feller, W. 1951. *Diffusion processes in genetics*. University of California Press Berkeley, Calif.

784 Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Oxford University Press.

785 Gerrish, P. J., and R. E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population.
786 *Genetica* 102:127.

787 Gomulkiewicz, R., and R. D. Holt. 1995. When does Evolution by Natural Selection Prevent Extinction?
788 *Evolution* 49:201.

789 Gomulkiewicz, R., R. D. Holt, M. Barfield, and S. L. Nuismer. 2010. Genetics, adaptation, and invasion in
790 harsh environments. *Evol. Appl.* 3:97–108.

791 Gomulkiewicz, R., S. M. Krone, and C. H. Remien. 2017. Evolution and the duration of a doomed
792 population. *Evol. Appl.* n/a-n/a.

793 Gonzalez, A., O. Ronce, R. Ferriere, and M. E. Hochberg. 2013. Evolutionary rescue: an emerging focus at
794 the intersection between ecology and evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 368:20120404–
795 20120404.

796 Greenspoon, P. B., and N. Mideo. 2017. Evolutionary rescue of a parasite population by mutation rate
797 evolution. *Theor. Popul. Biol.* 117:64–75.

798 Holt, R., M. Barfield, and R. Gomulkiewicz. 2004. Temporal Variation Can Facilitate Niche Evolution in
799 Harsh Sink Environments. *Am. Nat.* 164:187–200.

800 Holt, R. D., R. Gomulkiewicz, and M. Barfield. 2003. The phenomenology of niche evolution via
801 quantitative traits in a “black-hole” sink. *Proc. R. Soc. B Biol. Sci.* 270:215–224.

802 Jiang, Y.-C., H. Feng, Y.-C. Lin, and X.-R. Guo. 2016. New strategies against drug resistance to herpes
803 simplex virus. *Int. J. Oral Sci.* 8:1–6.

804 Kreiner, J. M., J. R. Stinchcombe, and S. I. Wright. 2018. Population Genomics of Herbicide Resistance:
805 Adaptation via Evolutionary Rescue. *Annu. Rev. Plant Biol.* 69:611–635.

806 Lande, R. 1976. Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution* 30:314.

807 Lande, R. 1980. The Genetic Covariance Between Characters Maintained by Pleiotropic Mutations.
808 *Genetics* 94:203–215.

809 Liu, L. L., F. Li, W. Pao, and F. Michor. 2015. Dose-Dependent Mutation Rates Determine Optimum
810 Erlotinib Dosing Strategies for EGFR Mutant Non-Small Cell Lung Cancer Patients. *PLoS ONE* 10.

811 Loeb, L. A. 2001. A Mutator Phenotype in Cancer. *Cancer Res.* 61:3230–3239.

812 Loeb, L. A., J. M. Essigmann, F. Kazazi, J. Zhang, K. D. Rose, and J. I. Mullins. 1999. Lethal mutagenesis of
813 HIV with mutagenic nucleoside analogs. *Proc. Natl. Acad. Sci.* 96:1492–1497.

814 Loverdo, C., and J. O. Lloyd-Smith. 2013. Evolutionary Invasion and Escape in the Presence of Deleterious
815 Mutations. *PLoS ONE* 8:e68179.

816 Martin, G., R. Aguilée, J. Ramsayer, O. Kaltz, and O. Ronce. 2013. The probability of evolutionary rescue:
817 towards a quantitative comparison between theory and evolution experiments. *Phil Trans R Soc B*
818 368:20120088.

819 Martin, G., S. F. Elena, and T. Lenormand. 2007. Distributions of epistasis in microbes fit predictions from a
820 fitness landscape model. *Nat. Genet.* 39:555–560.

821 Martin, G., and S. Gandon. 2010. Lethal mutagenesis and evolutionary epidemiology. *Philos. Trans. R. Soc.*
822 *Lond. B Biol. Sci.* 365:1953–1963.

823 Martin, G., and T. Lenormand. 2006. The Fitness Effect of Mutations Across Environments: A Survey in
824 Light of Fitness Landscape Models. *Evolution* 60:2413–2427.

825 Martin, G., and L. Roques. 2016. The Non-stationary Dynamics of Fitness Distributions: Asexual Model with
826 Epistasis and Standing Variation. *Genetics* 204:1541–1558.

827 Matuszewski, S., L. Ormond, C. Bank, and J. D. Jensen. 2017. Two sides of the same coin: A population
828 genetics perspective on lethal mutagenesis and mutational meltdown. *Virus Evol.* 3.

829 McCandlish, D. M., and A. Stoltzfus. 2014. Modeling Evolution Using the Probability of Fixation: History
830 and Implications. *Q. Rev. Biol.* 89:225–252.

831 Orr, H. A., and R. L. Unckless. 2008. Population Extinction and the Genetics of Adaptation. *Am. Nat.*
832 172:160–169.

833 Orr, H. A., and R. L. Unckless. 2014. The Population Genetics of Evolutionary Rescue. *PLOS Genet.*
834 10:e1004551.

835 Pariente, N., A. Airaksinen, and E. Domingo. 2003. Mutagenesis versus Inhibition in the Efficiency of
836 Extinction of Foot-and-Mouth Disease Virus. *J. Virol.* 77:7131–7138.

837 Pariente, N., S. Sierra, and A. Airaksinen. 2005. Action of mutagenic agents and antiviral inhibitors on foot-
838 and-mouth disease virus. *Virus Res.* 107:183–193.

839 Pariente, N., S. Sierra, P. R. Lowenstein, and E. Domingo. 2001. Efficient Virus Extinction by Combinations
840 of a Mutagen and Antiviral Inhibitors. *J. Virol.* 75:9723–9730.

841 Perfeito, L., A. Sousa, T. Bataillon, and I. Gordo. 2014. Rates of Fitness Decline and Rebound Suggest
842 Pervasive Epistasis. *Evolution* 68:150–162.

843 Ramsayer, J., O. Kaltz, and M. E. Hochberg. 2013. Evolutionary rescue in populations of *Pseudomonas*
844 *fluorescens* across an antibiotic gradient. *Evol. Appl.* 6:608–616.

845 Read, A. F., T. Day, and S. Huijben. 2011. The evolution of drug resistance and the curious orthodoxy of
846 aggressive chemotherapy. *Proc. Natl. Acad. Sci.* 108:10871–10877.

847 Regoes, R. R., C. Wiuff, R. M. Zappala, K. N. Garner, F. Baquero, and B. R. Levin. 2004. Pharmacodynamic
848 Functions: a Multiparameter Approach to the Design of Antibiotic Treatment Regimens.
849 *Antimicrob. Agents Chemother.* 48:3670–3676.

850 Springman, R., T. Keller, I. J. Molineux, and J. J. Bull. 2010. Evolution at a High Imposed Mutation Rate:
851 Adaptation Obscures the Load in Phage T7. *Genetics* 184:221–232.

852 Taddei, F., M. Radman, J. Maynard-Smith, B. Toupance, P. H. Gouyon, and B. Godelle. 1997. Role of
853 mutator alleles in adaptive evolution. *Nature* 387:700–702.

854 Tenaillon, O. 2014. The Utility of Fisher’s Geometric Model in Evolutionary Genetics. *Annu. Rev. Ecol. Evol.*
855 *Syst.* 45:179–201.

856 Uecker, H. 2017. Evolutionary rescue in randomly mating, selfing, and clonal populations. *Evolution*
857 71:845–858.

858 Uecker, H., and J. Hermisson. 2016. The Role of Recombination in Evolutionary Rescue. *Genetics* 202:721–
859 732.

860 Wilson, B. A., P. S. Pennings, and D. A. Petrov. 2017. Soft Selective Sweeps in Evolutionary Rescue.
861 *Genetics* 205:1573–1586.

862 Wylie, C. S., and E. I. Shakhnovich. 2012. Mutation Induced Extinction in Finite Populations: Lethal
863 Mutagenesis and Lethal Isolation. *PLOS Comput. Biol.* 8:e1002609.

864