Appendix I : Mathematical derivations and approximations

2 In the following, we use a supplementary notation to simplify the mathematical derivations: $\theta = n/2$.

4

1

5 I. General WSSM approximation to ER.

6 In the present model, we approximate the stochastic dynamics of each lineage by a Feller diffusion 7 (Feller 1951) or continuous branching CB-process (Lambert 2008), with parameters $\{r_i, \sigma_i\}$ for 8 lineage *i*. We ignore any density or frequency dependence and assume that all lineages have similar 9 stochastic reproductive variance ($\sigma_i \approx \sigma$). The resulting total population size N_t (cumulating all lineage 10 that co-segregate) can then also be approximated by a CB process, which follows the stochastic 11 differential equation:

$$dN_t = \bar{r}_t N_t dt + \sqrt{\sigma N_t} dB_t , \qquad (A1)$$

12 where B_t is a Weiner process, $\bar{r}_t = 1/N_t \sum_{i=1}^{N_t} r_i$ is the mean growth rate of all lineages present at 13 time t and σ is the common stochastic reproductive variance of all lineages. In the WSSM regime (i.e. 14 when $U \gg U_c = \theta^2 \lambda$), we can ignore the evolutionary stochasticity introduced by mutation and drift 15 as a first approximation. Then, the mean growth rate $\bar{r}_t \approx \langle \bar{r}_t \rangle$ (expectation over replicates denoted 16 by $\langle . \rangle$) is approximately deterministic and given by the WSSM results in (Martin and Roques 2016) for 17 the FGM. The probability of such a time inhomogeneous CB process to be extinct by time t is given by 18 (Bansaye and Simatos 2015):

$$P_{\rm ext}(t) = \exp\left(-\frac{2N_0}{\psi_t}\right), \qquad (A2)$$

19 where $\psi_t = \sigma \int_0^t \exp(-\rho(v)) dv$ and $\rho(u) = \int_0^u \langle \bar{r}_v \rangle dv$. The probability of ER is the complementary 20 probability of not being extinct over infinite time, namely $P_R = 1 - P_{\text{ext}}(\infty)$. For consistency with the 21 main text, $\sigma \approx 1$ in the following.

22 II. Explicit expression for a population initially clonal

Define r_D the decay rate of the initial clone, and r_{max} the maximal attainable growth rate (that of a genotype optimal in the stressful environment). In this case, the mean fitness trajectory, relative to the optimal genotype ($\langle \bar{m}_t \rangle = \langle \bar{r}_t \rangle - r_{max}$), under the WSSM approximation, for an initially clonal population is (eq. (12) in (Martin and Roques 2016)

$$\langle \overline{m}_t \rangle = m_0 \left(\operatorname{sech}(\mu t) \right)^2 - \theta \,\mu \tanh(\mu t) ,$$
 (A3)

27 where $\mu = \sqrt{U \lambda}$ and $\theta = n/2$, sech(.) and tanh(.) are the hyperbolic secant and tangent functions

and m_0 is the fitness difference between the original clone and the optimal genotype. Now, we require the absolute mean fitness trajectory (mean growth rate), which is obtained by noting that, by definition, $\langle \overline{m}_t \rangle = \langle \overline{r}_t \rangle - r_{max}$ and $m_0 = -r_D - r_{max}$. Denoting $y_D = r_D / r_{max}$, $\epsilon = \theta \mu / r_{max}$ we obtain:

$$\langle \bar{r}_t \rangle = \frac{\theta \,\mu}{\epsilon} \left(1 - (y_D + 1) \operatorname{sech}(\mu \, t)^2 - \epsilon \, \operatorname{tanh}(\mu \, t) \right).$$
 (A4)

32 The integral $\rho_t = \int_0^t \langle \bar{r}_v \rangle \, dv$ of this growth rate over time can be expressed in compact form by using

33 the change of variable $\tau = \mu t$, yielding

$$\rho_{t} = \frac{\theta}{\epsilon} f(\tau) - \theta \log(h(\tau))$$

$$h(\tau) = \cosh(\tau) \qquad (A5)$$

$$f(\tau) = \tau - (1 + y_{D}) \tanh(\tau)$$

Finally, the same change of variable ($dt = \mu d\tau$) can also be used to express the indefinite integral that

determines extinction probabilities, yielding the rate of ER per lineage present at the onset of stress:

$$\omega^{DN} = -\frac{\log P_{\text{ext}}(\infty)}{N_0} = \frac{2}{\int_0^\infty \exp(-\rho_u) \, du} \qquad (A6)$$
$$\omega^{DN} = 2 \, \mu / (\int_0^\infty h(\tau)^\theta \, \exp(-\frac{\theta}{\epsilon} \, f(\tau)) \, d\tau)$$

36

37

38 III. Laplace approximations for small ϵ , for a population initially clonal

Eq.(A6) is fully analytic but yields no explicit expression, which would be useful to get a more intuitive grasp of how each parameter affects the ER probability. When the mutation rate and effects are small enough that the load $\theta \mu$ is small relative to the maximal growth rate r_{max} (so that we are well away from lethal mutagenesis), then $\epsilon \ll 1$. This means that $\theta/\epsilon \gg 1$ (with $\theta \ge 1/2$) and the integral in Eq.(A6) is amenable to the Laplace Approximation (as $h(\tau)$ is monotonic over $\tau \in \mathbb{R}^+$). This approximation can be formulated as follows :

$$\int_{0}^{\infty} h(\tau)^{\theta} \exp(-\frac{\theta}{\epsilon} f(\tau)) d\tau \underset{\epsilon \to 0}{\approx} \sqrt{\frac{2 \pi \epsilon}{\theta f''(\tau_0)}} h(\tau_0)^{\theta} \exp(-\frac{\theta}{\epsilon} f(\tau_0))$$
(A7)

- 45 where $\tau_0 = \cosh^{-1}(\sqrt{1+y_D})$ is the unique minimum of f(.) over $\tau \in \mathbb{R}^+$. The Laplace approximation
- 46 in Eq.(A7) then yields a fully explicit expression for the rate of ER per lineage present at the onset of
- 47 stress defined in Eq.(A6). Rewriting in terms of the original parameters ($\theta/\epsilon = r_{max}/\mu$), we get

$$\omega^{DN} \underset{\epsilon \to 0}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} \gamma(y_D)\right)$$
$$\gamma(y_D) = \sqrt{y_D(1+y_D)} - \cosh^{-1}\left(\sqrt{1+y_D}\right) + \epsilon \left(\left(1+\frac{1}{2\theta}\right)\frac{\log(1+y_D)}{2} - \frac{1}{\theta}\frac{\log(y_D)}{4}\right).$$
(A8)

- 48 The accuracy of this expression is illustrated in **Supplementary Figure 1**.
- 49



51 **Supplementary Figure 1:** P_R as a function of y_D for three values of the dimensionality (given in legend), 52 as computed from the 'exact' Eq.(A6) (solid colored lines) vs. Laplace Approximation Eq.(A8) (dashed 53 purple lines). Parameters are $r_{max} = 1$, $\lambda = 5$. 10^{-4} , U = 0.25, $N_0 = 10^7$.

55 In fact, unless we consider narrow range of stress variation, as in Supplementary Figure 1, it appears 56 that, in the limit $\epsilon \to 0$, P_R shows limited dependency on dimensionality, provided that it remain 57 limited (n varies by five-fold above). This is confirmed by further simplifying Eq.(A8) to produce a rough 58 but reasonably accurate approximation, in a range that is a priori of most biological relevance, say for 59 $y_D > 0.1$, i.e. not-too mild a stress. In this range both $\log(1 + y_D)/2$ and $\log(y_D)/4$ are of similar or smaller order than $g(y_D) = \sqrt{y_D(1+y_D)} - \cosh^{-1}(\sqrt{1+y_D})$. Therefore, as $\epsilon \to 0$ (as we assume 60 here), the right hand factor in $\gamma(y_D)$, proportional to ϵ , becomes negligible, relative to the left hand 61 62 term, and $\gamma(y_D) \approx g(y_D)$. Eq. (A8) thus simplifies to the expression given in Eq.(6) for *de novo* rescue:

$$\omega_{y_D \ge 0.1}^{DN} \approx_{\substack{\epsilon \to 0 \\ y_D \ge 0.1}} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} g(y_D)\right).$$

$$g(y_D) = \sqrt{y_D(1+y_D)} - \cosh^{-1}\left(\sqrt{1+y_D}\right)$$
(A9)

This simpler approximation is less precise but still provides a good order of magnitude, when ϵ is small enough. **Supplementary Figure 2** illustrates its accuracy and the fact that the ER rate is indeed roughly

65 independent of dimensionality θ in this parameter range : Eq.(A9) captures the order of magnitude of

66 stress at which P_R drop from 1 to 0.





69 **Supplementary Figure 2:** Same as **Supplementary Figure 1** (same parameters, except r_{max} , indicated

70 on the graph), this time as a function of r_D and compared to both Eqs.(A8) (purple dashed lines) and

71 (A9) (black dotdashed lines).

73 IV. Explicit expression for a population initially at mutation-selection balance

74 This time we assume initially mutation-selection balance, then a shift in optimum occurs without any 75 change in U or λ : adaptation is driven by both standing variance and de novo mutations. Under the 76 WSSM approximation (corresponding to our assumption here), the mutation selection balance in an 77 asexual population under the isotropic FGM corresponds to a simple Gaussian distribution of 78 phenotypes, centered on the optimum (in the environment where the equilibrium sets) and a variance 79 μ per trait: $\mathbf{z} \sim N(\mathbf{0}, \mu \mathbf{I}_n)$. Once this optimum is shifted in the new environment, the phenotypic 80 distribution progressively shifts (retaining the same normality and variance) towards the new 81 optimum. This yields the following dynamics for the mean growth rate (based on the corresponding 82 WSSM approximation in eq. (13) in (Martin and Roques 2016), with the same parameterization as for 83 Eq.(A4)) we get

$$\langle \bar{r}_t \rangle = r_{max} (1 - \exp(-2 \,\mu \, t) (1 + y_D) - \epsilon) ,$$
 (A10)

84 which then yields the corresponding expression for ρ_t (same form as Eq.(A5)):

$$\rho_{t} = \frac{\theta}{\epsilon} f(\tau) - \theta \log(h(\tau))$$

$$f(\tau) = \tau - (1 + y_{D}) (1 - \exp(-2\tau))/2$$

$$h(\tau) = \exp(\tau)$$
(A11)

The rate of ER takes a similar form as in the DN scenario (Eq.(A6)), but this time an explicit exact expression is found:

$$\omega^{DN+SV} = -\log P_{\text{ext}}(\infty) / N_0 = \frac{2}{\int_0^\infty \exp(-\rho_u) \, du} = 2 \, \mu / \int_0^\infty h(\tau)^\theta \exp(-\frac{\theta}{\epsilon} f(\tau)) d\tau$$

$$\omega^{DN+SV} = 4 \, \mu \frac{\exp(-\xi) \, \xi^\beta}{\Gamma(\beta) - \Gamma(\beta, \xi)}$$
(A12)

87 Where $\xi = (1 + y_D) \theta/(2\epsilon)$ and $\beta = (1 - \epsilon) \theta/(2\epsilon)$, and $\Gamma(.)$ and $\Gamma(.)$ are respectively the

88 Euler's gamma function and the incomplete gamma function.

89 V. Laplace approximations for small ϵ , for a population initially at mutation-selection balance

Although this is already explicit, a Laplace approximation can be used to get a simpler result, again away from the lethal mutagenesis regime ($\epsilon \ll 1$). The function f(.) has a unique minimum at $\tau_0 = \log(\sqrt{1+y_D})$ and the resulting Laplace approximation for the rate of ER is of the same form as Eq.(A8):

$$\omega^{DN+SV} \underset{\epsilon \to 0}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} \gamma(y_D)\right),$$

$$\gamma(y_D) = \frac{y_D - (1-\epsilon)\log(1+y_D)}{2},$$
(A13)

94 From which, as $\epsilon \ll 1$, arises the same expression as Eq.(A9) with a different function g(.):

$$\omega^{DN+SV} \underset{\epsilon \to 0}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} g(y_D)\right).$$

$$g(y_D) = (y_D - \log(1+y_D))/2$$
(A14)

95 The accuracy of this approximation is illustrated in **Supplementary Figure 3** below. Note also how the

96 drop in ER probability occurs at higher stress levels here than with purely *de novo* mutation (in

97 Supplementary Figure 2).





102

98

103 The accuracy shown in the main text between Eq.[5] (Eq.A(12)) and the simulations decrease when the 104 growth rates of the individuals in the population increase (in absolute value). Indeed, when growth 105 rates are too low or high ($|r| \ge 1$), the continuous time approximation used in our model fails to 106 predict accurately the discrete time dynamic of the population size and therefore the probability of 107 rescue given by our simulations. This is illustrated in **Supplementary Figure 4** where we can see that 108 increasing r_{max} leads to an underestimation of the rescue probability predicted by the simulations. 109 Moreover, at higher U the ER probability drop from highly likely to highly unlikely at larger r_D , of order 110 1 or more, for which the discrepancies between continuous time and discrete time dynamics111 mentioned above increase, as shown in **Supplementary Figure 4.**





114 *Supplementary figure 4:* Accuracy of the predictions of ER probabilities for DN+SV scenario (blue:

115 Eq.[5]) for two values of r_{max} and two values of U against simulations (see legend). Note that

116 the green line and the corresponding simulations are the same as in Figure 3a. Other parameters

117 are $N_0 = 10^5$, n = 4, $\lambda = 5.10^{-3}$.

118 VI. Width of the mutation window of ER

119 We now wish to evaluate the mutation rate at which ER switches from very likely to very unlikely. We 120 know that the rate of ER drops sharply at some upper threshold mutation rate, due to lethal 121 mutagenesis effects (i.e. when the mutation load $\mu \theta = r_{max}$). This implies a maximal mutation rate 122 $U_{max} = r_{max}^2/(\theta^2 \lambda)$, above which lethal mutagenesis leads to certain extinction of the population.

Here, we are interested in the parameter range, far below lethal mutagenesis, where the ER probability increases (sharply too) with the mutation rate. We first seek the value μ_* at which ER occurs 50% of the time. This critical μ_* , is the one above which ER becomes likely, so that the rescue probability is $P_R(\mu_*) = 1/2$ and the ER rate is $\omega_* = -\log(1 - P_R(\mu_*))/N_0 = \log(2)/N_0$. In this range of parameter ($\epsilon \ll 1$), we can use the approximate expressions in Eqs.(A9) and (A14) : $\omega(\mu_*) \approx$ $2 \sqrt{r_{max} \mu_*/\pi} \exp(-r_{max} g(y_D)/\mu_*)$. Solving for this equation yields a unique solution:

$$\mu_* \approx \frac{2 g(y_D) r_{\max}}{\mathcal{W}\left(\frac{8 g(y_D)}{\pi} \left(\frac{N_0 r_{max}}{\log(2)}\right)^2\right)} , \qquad (A15)$$

129 where $\mathcal{W}(.)$ is Lambert's function. This rate can be approximated to a simpler form if we note that N_0^2

130 is typically very large compared to $g(y_D)$ and r_{max} , which are of order 1. Therefore the denominator 131 in Eq.(A15) is driven by the asymptotic limit of $\mathcal{W}(.)$ ($\mathcal{W}(v) \approx \log(v/\log(v))$), for large v) and by the

terms in N_0 . Overall, to a reasonably good approximation:

$$\mu_{*} \underset{N_{0} \to \infty}{\approx} \frac{2 g(y_{D}) r_{\max}}{\log\left(\frac{N_{0}^{2}}{\log(N_{0}^{2})}\right)} \quad . \tag{A16}$$

133 Recalling that $\mu = \sqrt{U \lambda}$, the critical mutation rate where the ER probability is 50% is

$$U_* \approx \frac{1}{\lambda} \left(\frac{2 g(y_D) r_{\max}}{\mathcal{W}\left(\frac{8 g(y_D)}{\pi} \left(\frac{N_0 r_{\max}}{\log(2)}\right)^2\right)} \right)^2 \approx \frac{4 g(y_D)^2 r_{\max}^2}{\lambda \log\left(\frac{N_0^2}{\log(N_0^2)}\right)^2}.$$
 (A17)

This means that ER is only likely within a window of mutation rate $U_* \le U \le U_{max}$. This window narrows down as stress increases (increased decay rate and hence $g(y_D)$, see **Supplementary Figure 6**) or as N_0 gets smaller, and its lower bound is roughly independent of dimensionality. The accuracy of this approximation is illustrated in **Figure 4** for the DN scenario ($g(y_D) = g_{DN}(y_D) =$ $\sqrt{y_D(1+y_D)} - \cosh^{-1}(\sqrt{1+y_D})$) and in **Supplementary Figure 5** below for the SV+DN scenario in the presence of standing variance ($g(y_D) = g_{DN+SV}(y_D) = (y_D - \log(1+y_D))/2$).





141

Supplementary Figure 5: same as Figure 4 with standing variance plus de novo mutation.





144

Supplementary Figure 6: Decrease of the width and height of the mutation window with stress. Colored plain lines show Eq.[6] for ER from *de novo* mutations (a) or from pre-existing standing genetic variance and *de novo* mutations (b). Black lines give $max(P_R)$ from Eq.(A18) (for $r_D = 2.2$ in (a) and $r_D = 4$ in (b)). The colors and the parameters are the same as in **Figure 4** for increasing r_D .

149

Supplementary Figure 6 shows that, for sufficiently high stress r_D or low inoculum size N_0 , the ER probability cannot reach above some limited maximum $p = \max_{\mu \in \mathbb{R}^+} (P_R)$. Figure 5 also shows that max (P_R) drops sharply (from max $(P_R) = 1$ to max $(P_R) \ll 1$) as r_D increases or N_0 decreases. The goal of this section is to compute the values of r_D where this transition occurs, namely the stress levels beyond which the population cannot avoid extinction, even if its mutation rate was to changed (e.g. by hyper-mutators).

156 The transition observed in Figure 5 occurs at some pairs of parameters $\{r_D^*(p), N_0^*(p)\}$ for 157 which $\max(P_R) = p$ for some threshold value p. The shape of the transition in Figure 5 suggests that 158 the values of $r_D^*(p)$ and $\log(N_0^*(p))$ are linearly related: so that $\log N_0^*(p) = a + b r_D^*(p)$ for 159 some (a, b) > 0. We also know that $P_R = 1 - \exp(-N_0 \omega)$ where ω depends on r_D, r_{max}, μ, n

(Eq.(5)). At $\{r_D^*, \log N_0^*\}$ we thus have $p = \max(P_R) = 1 - \exp(-N_0^*(p) \omega^*(p))$, where $\omega^*(p)$ is the 160 corresponding ER rate along the transition of height p. This implies that $\log N_0^*(p) = \log(\log(1 - 1))$ 161 $|p|| - \log \omega_*(p) = a + b r_D^*(p)$, so that $\omega_*(p) = \omega(r_D^*(p)) = A \exp(-B r_D^*(p))$ for some (A, B). 162 Since ω_* is independent of N_0 and is a maximum over μ (so independent of μ), it must only depend on 163 164 the remaining parameters $\{r_D, r_{max}, n\}$. The coefficients A, B, which determine its relationship with r_D^* , 165 must thus only depend on n and r_{max} . Therefore, we look for approximations of the form $\omega_* \approx$ $A \exp(-B r_D^*)$, for some functions $A = A(n, r_{max})$ and $B = B(n, r_{max})$ for both DN and DN + SV 166 167 scenarios. A simple fitting procedure, across a range of values of $\{n, r_{max}\}$, carried out with Matlab© 168 Optimization Toolbox (the Matlab© code is available in supplementary file 2) suggests that $A \approx$ 169 r_{max}/n and $B \approx \delta n/r_{max}$, leading to:

$$\max(P_R) = 1 - \exp(-N_0^*(p)\omega^*(p))$$

$$\omega_*(p) = \frac{r_{max}}{n} \exp\left(-\delta \frac{n}{r_{max}} r_D^*(p)\right).$$
(A18)
with
$$\begin{cases}
DN : \delta = 0.6 \\
DN + SV : \delta = 0.31
\end{cases}$$

170 From Eq.(A18) we can see that the height $max(P_R)$ of the mutation window increases with an increase

in the maximal growth rate r_{max} or the population size N_0 and decreases with an increase in the dimensionality n or the harshness of stress r_D (the accuracy of this fitting for different values of r_{max} and n is illustrated in Supplementary figures 7). The figure below provides a comparison between the numerical value of max(P_R) from Eq.[5] and the result of Eq.(A18).

- 175
- 176



Supplementary Figure 7: $max(P_R)$ as a function of r_D for different values of the dimensionality and the maximal growth rate (given in legend), as computed from the 'exact' Eq.[5] vs. fitted approximation Eq.A(18). (a) Population adapting from de novo mutation only. (b) Population adapting from both de novo mutation standing genetic variance. In both cases $N_0 = 10^5$.

182

183 Solving $\max(P_R) = p = 1 - \exp(-N_0 \omega_*(p))$ in terms of $r_D^*(p)$ using Eq. A(18) yields the threshold 184 stress at which $\max(P_R) = p$:

$$r_D^*(p) = \frac{r_{max}}{n\,\delta} \log\left(\frac{N_0 r_{max}}{n\,|\log(1-p)|}\right) \tag{A19}$$

185 where δ is given in Eq. (A 18) depending on the scenario.

186

187 VIII. Analytical and simulation models from Anciaux et al. (2018)

188 Explicit expression for a population initially clonal under the SSWM approximation from Anciaux et al.189 (2018):

- 190 Here we provide the equations from Anciaux et al. (2018) for the rate of rescue per individual for a
- 191 population initially clonal ω_{DN} in the SSWM regime ($U < U_c$). In this model, the distribution of
- fitness effects of mutations is also based on the FGM. Hence, from Eq.(2) of Anciaux et al. (2018), for
- 193 a clonal population with a growth rate $-r_D$ in the new environment, the scaled growth rates of
- single-step mutants $y = r/r_{max}$, have the following probability density function:

$$F_{y}(y) = \exp\left(-\frac{r_{max}(2+y_{D}-y)}{\lambda}\right) \left(\frac{r_{max}}{\lambda}\right)^{n/2} (1-y)^{n/2-1} \frac{{}_{0}F_{1}\left(\frac{n}{2}, \left(\frac{r_{max}}{\lambda}\right)^{2}(1+y_{D})(1-y)\right)}{\Gamma(n/2)}, \qquad (A20)$$

$$y \in]-\infty, 1]$$

- 195 with $y_D = r_D / r_{max} \in [0, +\infty]$ and r_{max} , n, λ defined in table 1. Where ${}_0F_1(.,.)$ is the
- 196 confluent hypergeometric function and $\Gamma(z)$ is the gamma function.
- 197 Following Eq.(5) from Anciaux et al. (2018), the rate of rescue for a population initially clonal is:

$$\omega_{DN} = \frac{U}{r_D} \int_0^1 \pi(y) f_y(y) dy , \qquad (A21)$$

198 with $\pi(y) = 1 - \exp(-2y r_{max}/\sigma)$ the probability of establishment of a resistant genotype with 199 scaled growth rate y > 0 in the new environment (in figures 1 and 2 from the main text $\sigma =$ 200 1).

201

202 Simulation algorithm from Anciaux et al. (2018):

The following algorithm is described in the Methods section of Anciaux et al. (2018) and the explanations are directly extracted from the same section. The Mathematica code is provided in supplementary file I.

while COND	
1. Muta	ation:
•	Draw the number of mutations for each individual from a Poisson distribution with rate ${\cal U}$
•	add mutation effects (when relevant) to the current parent phenotype, by draowing these into a multivariate normal distribution $N(0, \lambda \mathbf{I}_n)$
2. Selec	ction and drift:
•	Genotypes forming the new generation are sampled with replacement from the
	previous one with weight $W_i = e^{r_i}$ for the genotype <i>i</i> .
3. Dem	ographic stochasticity:
•	The size N_{t+1} of population at generation $t + 1$ is drawn as a Poisson number
	$N_{t+1} \sim Poisson(N_t \overline{W})$, with $\overline{W} = \overline{\exp(r)}$ the mean Darwinian fitness ($W =$
	$exp(r)$ and N_t the population size, in the previous generation.
end while	
With COND = if ($\bar{r}_t > 0$, $\exp(-2 N_t \bar{r}_t) < 10^{-12}$) & $N_t > 0$. A population is considered rescued wher	

With COND = if $(\bar{r}_t > 0, \exp(-2N_t \bar{r}_t) < 10^{-12}) \& N_t > 0$. A population is considered rescued when it reaches a population size N_t and mean growth rate \bar{r}_t such that its ultimate extinction probability, if it were monomorphic and stable over time (probability $\exp(-2N_t \bar{r}_t)$), would lie bellow 10^{-12} . This is a conservative criterion: once \bar{r}_t has become positive, we expect it to remain so, yielding further increases in population size and thus further decreasing the probability of future extinction.

Points 2 and 3 from the algorithm correspond to reproduction and have been separated into 2 separate
phases to increase the speed of the algorithm. However, this is exactly equivalent to drawing
independent Poisson reproductive outputs from each individual.

- In the algorithm in supplementary file 1, the effects of mutations are not drawn every generations as in point 1. Instead a large number (10^6) of mutation effects are drawn before the simulations and
- 216 mutations effects are drawn in this "pre-set" every generations to increase the speed of the algorithm.

For rescue from populations at mutation-selection balance, 10 replicate initial equilibrium populations were generated, each by starting from an optimal clone and running the same algorithm with fixed population size ($N_t = 10^6$) until the mean growth rate had visually stabilized to a fixed value (close to its theoretical equilibrium value $\bar{r}_{eq} = r_{max} - U$ (for $U < U_c$) for more than 1000 generations. Then the optimum was shifted by $\sqrt{2(r_D + r_{max})}$ phenotypic units, and 1000 replicate ER simulations were performed (same algorithm as for de novo rescue), from each of the 8 replicate equilibrium populations.

- **Supplementary figure 8** shows the dynamic of the mean fitness and the population size of 4 replicates
- (2 rescues and 2 extinctions) of a simulated population using the above algorithm for a population
- 226 with high mutation rate ($U > U_c$).

227





229 **Supplementary Figure 8:** Dynamics of the mean fitness \bar{r}_t (A) and the population size N_t (B) of 230 replicate populations starting from a clone at $-r_D = -1.5$ at size $N_0 = 10^5$. The blue lines represents 231 rescues and the yellow lines extinctions. Parameters for the simulations are $r_{max} = 1$, $U = 100 U_c$, 232 n = 4 and $\lambda = 5 \times 10^{-3}$.

234 Bibliography:

.

235 Anciaux, Y., L.-M. Chevin, O. Ronce, and G. Martin. 2018. Evolutionary Rescue over a Fitness

236 Landscape. Genetics 209:265–279.

- 237 Bansaye, V., and F. Simatos. 2015. On the scaling limits of Galton-Watson processes in varying
- environments. Electron. J. Probab. 20.
- 239 Feller, W. 1951. Diffusion processes in genetics. University of California Press Berkeley, Calif.
- Lambert, A. 2008. Population Dynamics and Random Genealogies. Stoch. Models 24:45–163.
- 241 Martin, G., and L. Roques. 2016. The Non-stationary Dynamics of Fitness Distributions: Asexual Model
- with Epistasis and Standing Variation. Genetics 204:1541–1558.

243