Supplemental Information

1. MEAN WAITING TIMES FOR ADAPTATION AND CLEARANCE

We considered the effect of an environmental change on a single site, drawn randomly from genomes in a population (illustrated in Fig. 2 of the main text). This site was then reassigned with either a negative coefficient $s = 0.1$ or with a positive coefficient of s = 0.1 (other scenarios, where robustness interacts with the magnitude of selection are discussed in the sections below). We then estimated the mean waiting times for adaptation as a function of robustness. We defined mean waiting times for adaptation as the number of generations until the newly reassigned adaptive allele reaches a frequency threshold of 0.99, at the verge of fixation. We do not employ the term fixation, as large populations coupled with high mutation rates will rarely allow complete fixation of an allele in a population. As described in the main text, our results show that adaptation will decrease waiting times for adaptation of novel beneficial alleles and increase waiting times for purging of novel deleterious alleles. Here we explored how this effect changes depending on the population size of the virus: constant population size of $N = 10^6$, constant population size of $N = 100$, and a population undergoing a bottleneck of $N = 100$ and undergoing exponential growth till $N = 10^6$ (Fig. S1). For the latter type of exponential growth we further explored the effect of the intrinsic growth rate (denoted logistic r), affecting how fast the population size increases (see Methods of main text). Notably, our results show that exponentially growing viruses are affected by robustness in the same manner as viruses experiencing a constantly large population size. On the other hand, viruses at a small population size of $N = 100$ display a different pattern. Under these conditions, it appears that genetic drift causes an increase in the waiting times for adaptation, and substantially reduced waiting times for clearance. Thus, for smaller population sizes, the negative effects of purging are less pronounced overall, whereas the positive effects of robustness on adaptation are even more pronounced. Indeed, these observations are consistent with the hypothesis that selection in small population sizes favors robustness [7][8].

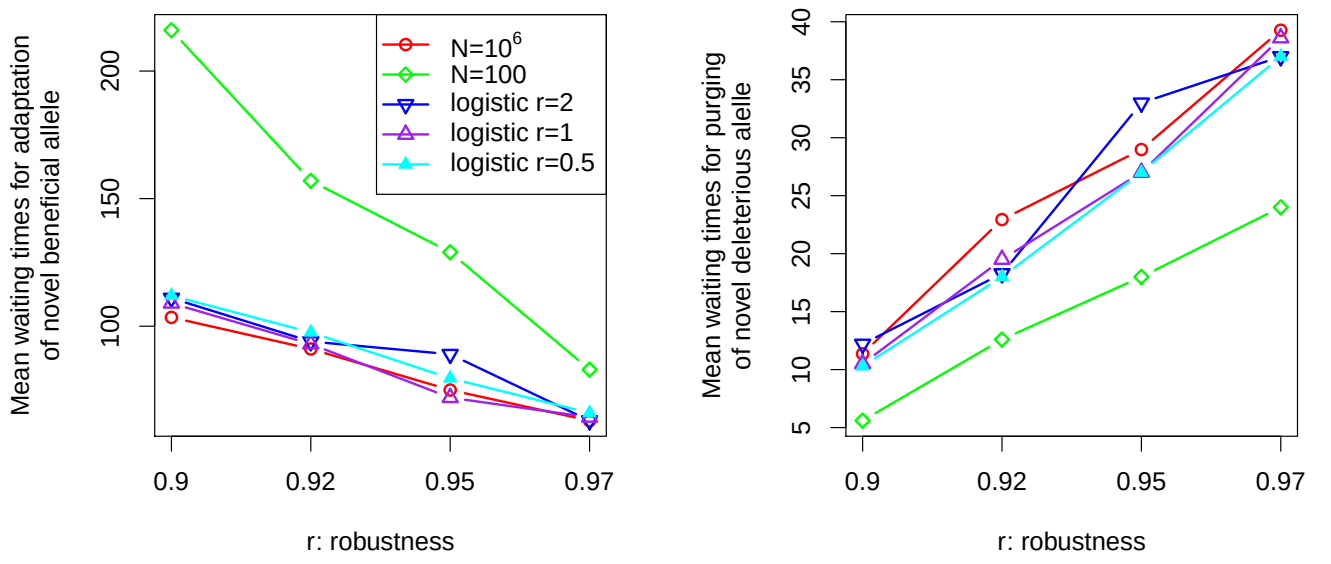

FIG. S1: Mean waiting times in generations are shown for adaptation of a novel beneficial allele (left panel), and purging of a deleterious allele (right panel), following an environmental change. The red circles show a constant $N = 10^6$, green diamonds show a constant $N = 100$. Triangles (blue, violet, and turquoise) show a population with a bottleneck of $N = 100$ at the environmental change, followed by logistic growth with parameter logistic r that defines the rate of growth (see Methods in main text). Robustness is shown to both reduce the time for adaptation and increase the time to purging of deleterious mutations. Related to Figure 2.

2. ANALYTICAL FORMULA FOR MEAN POPULATION FITNESS

2.1. Aims and specifications

Our goal is to obtain an analytic expression for the difference between the fitnesses of a robust vs a brittle population following an environment change.

2.2. Model description

As per the main text, we assume a haploid population of N individuals, with a genome of length $L = 10000$ bp. Mutations were modeled assuming two allelic states in each site with an equal rate of mutation between the two states of $\mu = 10^{-4}$ mutations/replication/base. Each population is assigned a distribution of mutation fitness effects (DMFE), from which a selection coefficient is drawn for one of the alleles at each position (labeled the alternative or mutant allele), whereas the fitness of the other allele (labeled the wild-type, WT, allele) is set at 1. Log mean fitness was calculated assuming multiplicative effects across sites (see below). We assume that the DMFE follows a mixture of a log-normal distribution of negative selection coefficients ($s < 0$) with probability $(1 - q)$, and a point mass at $s = 0$ with probability q. Robustness is defined by the mean fitness of a mutant allele, $r = \bar{\omega}$ (as defined in the main text).

We assume two populations, a robust one (with robustness r_r) and a brittle one (with robustness r_{nr}) evolving under mutation-selection balance in an environment labeled 1. Each robust and brittle population is characterized by a respective proportion q_r and q_{nr} of sites with $s = 0$, and the remaining proportion of sites with mean $\bar{s}_r < 0$ and $\bar{s}_{nr} < 0$, respectively.We then perform an environmental switch (see main text) to a new environment, labeled environment 2. This consists in assigning different positive or negative selection coefficients to randomly drawn positions. The same environmental switch is applied to both populations. Frequencies of alleles are then evolved via a Wright-Fisher process. Our aim is to derive an expression for difference in log mean fitness following the environmental change from time $t = 0$ until time $t = T$.

2.3. Analytical derivation of log mean fitness over time

We will now obtain an explicit formula for the log mean fitness over time. Under the assumption of no epistasis and no linkage disequilibrium, the log mean fitness of a population is given by

$$\log \bar{\omega}(t) = \sum_{i=1}^{L} \log \bar{\omega}_i(t) = \sum_{i=1}^{L} \log((1 + s_i)x_i(t) + 1(1 - x_i(t))) = \sum_{i=1}^{L} \log(1 + s_i x_i(t)) \approx \sum_{i=1}^{L} s_i x_i(t) \tag{1}$$

We henceforth focus on the fitness of one site, which can be estimated by using a diffusion approximation to the finite Wright-Fisher model, by use of the following expression (see Ref. [1], Ch. 4)

$$\bar{\omega}_i(x_0, s_{i,2}; t) = \int_0^1 (1 + s_{i,2})x' P(x'|x_0, s_{i,2}; t)\rho(x_0|s_{i,1})dx' + \int_0^1 (1 - x')P(x'|x_0, s_{i,2}; t)\rho(x_0|s_{i,1})dx' \tag{2}$$

where $P(x'|x_0, s_{i,2}; t)$ is the probability of observing the allele at frequency x' at time t conditional upon its initial frequency, x_0, and selection coefficient, $s_{i,2}$, and $\rho(x_0|s_{i,1})$ is the distribution of initial frequencies of the allele (see Ref. [1], Eq. 5.72). Implicit formula for the former can be obtained by solving the Fokker Planck equation, with drift and diffusion terms derived by the constraints of the Wright-Fisher process [1], in terms of eigenvalues and eigenfunctions of the transition probability matrix. However, this is of very limited practical use. Therefore we adopt a mean field deterministic approach, which, in the limit of very large populations, is expected to approximate well the stochastic expression in Eq. (2).

We are interested in log mean fitness over time (between time 0 and time T), and our aim is hence to determine

$$\log(\bar{\omega}_i)_T := \int_0^T \log \bar{\omega}_i(t)dt \approx \int_0^T s_i x_i(t)dt \tag{3}$$

Given an initial allele frequency x_0, the deterministic allele frequency at time t is given by the following formula (derived from Ref. [2], using $w = 1 + s$)

$$x(t) = \frac{x_0 w^t}{w^t x_0 + (1 - x_0)} \tag{4}$$

We use this deterministic equation instead of integrating over stochastic paths for mathematical convenience. For strong selection, and intermediate allele frequencies, it should approximate the stochastic path quite accurately. As we will show, it leads to quite accurate predictions in this case.

Solving the integral of Eq. (4) over time , and approximating $\log \omega = \log(1+s) \approx s$ yields

$$\int_0^T sx(t)dt = \frac{(w-1)}{\log(w)} \log((w^T-1)x_0+1) \approx \log((w^T-1)x_0+1) \tag{5}$$

We next apply our mean field approach (i.e., assuming allele frequencies are given by their expectation before the environmental change) by assuming

$$x_i(t) = qx_i(t, x_0^{(n)}) + (1-q)x_i(t, x_0^{(d)}) \equiv qx_i^{(n)}(t) + (1-q)x_i^{(d)}(t) \tag{6}$$

where $x_0^{(n)}$ and $x_0^{(d)}$ represent mean equilibrium frequencies for neutral and deleterious alleles, respectively, in environment 1. We assume that N_+ and N_- sites, in both the robust and non-robust population, are reassigned positive (s^+) and negative (s^-) coefficients in the novel environment. $M = L - N_+ - N_-$ sites retain their coefficient from environment 1. Per the Haldane-Muller principle, log mean fitness of the $(1-q)M$ deleterious sites is given approximately by $-\mu$, determined by mutation-selection equilibrium frequencies. On the other hand log mean fitness of the qM neutral sites will be 0 ($s = 0$). We hence obtain:

$$\begin{aligned}
\log(\bar{\omega}(t)) &\approx \sum_{i=1}^{L} s_i x_i = \sum_{i=1}^{N_+} s_i^+ x_i + \sum_{i=1}^{N_-} s_i^- x_i - (1-q)M\mu && (7)\\
&= \sum_{i=1}^{N_+} qs_i^+ x_i^{(n)}(t) + \sum_{i=1}^{N_+}(1-q)s_i^+ x_i^{(d)}(t) + \sum_{i=1}^{N_-} qs_i^- x_i^{(n)}(t) + \sum_{i=1}^{N_-}(1-q)s_i^- x_i^{(d)}(t)\\
&- (1-q)M\mu && (8)\\
&= \sum_{i=1}^{N_+} qs_i^+ x_i^{(n)}(t) + \sum_{i=1}^{N_+}(1-q)s_i^+ x_i^{(d)}(t) + \sum_{i=1}^{N_-} qs_i^- x_i^{(n)}(t) + \sum_{i=1}^{N_-}(1-q)s_i^- x_i^{(d)}(t)\\
&- (1-q)M\mu && (9)
\end{aligned}$$

Integrating the previous in time, we obtain

$$\begin{aligned}
\log(\bar{\omega})_T &\approx \sum_{i=1}^{N_+} q\log(((s_i^+ + 1)^T - 1)x_0^{(n)} + 1) + \sum_{i=1}^{N_+}(1-q)\log(((s_i^+ + 1)^T - 1)x_0^{(d)} + 1)\\
&+ \sum_{i=1}^{N_-} q\log(((s_i^- + 1)^T - 1)x_0^{(n)} + 1) + \sum_{i=1}^{N_-}(1-q)\log(((s_i^- + 1)^T - 1)x_0^{(d)} + 1)\\
&- M(1-q)\mu T
\end{aligned} \tag{10}$$

We are interested in evaluating the difference between the log mean fitness of a robust vs a brittle population. We assume below that neutral mutations will segregate in both populations at the same mean frequency before the environmental switch, while deleterious mutations will slightly differ in mean frequencies (reflected in the notations below). We thus write:

$$\begin{aligned}
\Delta\log(\bar{\omega})_T &\approx \sum_{i=1}^{N_+} q_r\log(((s_i^+ + 1)^T - 1)x_0^{(n)} + 1) + \sum_{i=1}^{N_+}(1-q_r)\log(((s_i^+ + 1)^T - 1)x_{0_r}^{(d)} + 1)\\
&+ \sum_{i=1}^{N_-} q_r\log(((s_i^- + 1)^T - 1)x_0^{(n)} + 1) + \sum_{i=1}^{N_-}(1-q_r)\log(((s_i^- + 1)^T - 1)x_{0_r}^{(d)} + 1)\\
&- \sum_{i=1}^{N_+} q_{nr}\log(((s_i^+ + 1)^T - 1)x_0^{(n)} + 1) - \sum_{i=1}^{N_+}(1-q_{nr})\log(((s_i^+ + 1)^T - 1)x_{0_{nr}}^{(d)} + 1)\\
&- \sum_{i=1}^{N_-} q_{nr}\log(((s_i^- + 1)^T - 1)x_0^{(n)} + 1) - \sum_{i=1}^{N_-}(1-q_{nr})\log(((s_i^- + 1)^T - 1)x_{0_{nr}}^{(d)} + 1)\\
&+ M\mu T(q_r - q_{nr})
\end{aligned} \tag{11}$$

For small values of T (as explored in the main text, $T = 20$), and assuming that mean frequencies of deleterious mutations $x_0^{(d)} \ll 1$, we assume that all terms of the form $\log(((s_i+1)^T-1)x_0^{(d)}+1) \approx 0$. Thus, the previous equation can be rewritten as:

$$\Delta \log(\bar{\omega})_T \;\approx\; \Delta q \left(\sum_{i=1}^{N_+} \log(((s_i^+ + 1)^T - 1)x_0^{(n)} + 1) + \sum_{i=1}^{N_-} \log(((s_i^- + 1)^T - 1)x_0^{(n)} + 1) + M\mu T \right) \tag{12}$$

We conclude that the difference in log mean fitness values is dominated by $\Delta q = (q_r - q_{nr})$, by the mean frequency of neutral alleles in the first environment $x_0^{(n)}$, by the time T between environmental changes, and by the magnitude of adaptive versus deleterious changes following the environment change. The "standard" genetic load component $M\mu T$, reflecting the load of alleles at equilibrium frequencies, will likely play a smaller role in this context of short time-scales.

3. EXTRAPOLATION OF FITNESS FROM EXPERIMENTAL DATA

3.1. Aims and specifications

Our theoretical approach allows us to predict under what conditions a brittle versus robust population would have higher fitness following a theoretical environmental change, composed of changes in fitness value of different mutations. The experimental datasets published by [18] and [17] allow us to extrapolate the difference in log fitness of a brittle versus a robust population following a host change as measured in these studies. In Table S1 and in the text below, we give in detail the calculations performed to reach these extrapolated fitness values, together with associated calculations of 95% confidence intervals (CIs).

Family	Host species	Mean $\sum s^+$	CI of $\sum s^+$	Mean $\sum s_-$	CI of $\sum s_-$	Δ log fitness	CI fitness
Solanaceae	*Nicotiana tabacum*	0.57	[-5.12,6.27]	-162	[-182.58,-141.43]	-756.39	[-913.54,-599.25]
Solanaceae	*Nicotiana benthamiana*	1.41	[-1.19,4.02]	-272	[-318,-226]	-1110	[-1414,-806]
Solanaceae	*Datura stramonium*	197	[181,213]	-999	[-1056,-941]	11	[-381,404]
Solanaceae	*Capsicum annuum*	6.8	[2.98,10.61]	-811	[-901,-720]	-2476	[-3067,-1886]
Solanaceae	*Solanum lycopersicum*	257.35	[231.6,283.1]	-3628	[-3638,-3618]	-1135	[-1327,-944]
Amaranthaceae	*Gomphrena globosa*	0	[0,0]	-2288	[-2351,-2224]	-5950	[-6336,-5564]
Amaranthaceae	*Spinacea oleracea*	0	[0,0]	-2547	[-2611,-2484]	-6108	[-6470,-5746]
Asteraceae	*Helianthus annuum*	0	[0,0]	-2235	[-2297,-2173]	-5906	[-6257,-5555]
–	*Salmonella typhimurium*	4.48	NA	-749	NA	-429	NA

TABLE S1: Extrapolated difference in log fitness values between a robust and brittle population after a host shift (first eight rows from [17], last row from [18]), and 95% confidence intervals. Related to Figure 4.

3.2. Calculations

Mean $\sum s_i$ was calculated as described in the Methods section of the main text, based on the six replicas used to estimate s in [17]. In order to calculate the confidence intervals, we assumed independent and identically distributed random variables (*iid*), in line with the inherent assumption throughout the manuscript. This allowed us to calculate the variance and standard deviation of $\log(((s+1)^T-1)x_0+1)$, with $T = 10$ and $x_0 = 0.5$. Variance of the sum Δ log fitness $\approx \sum_i \log(((s_i+1)^T-1)x_0+1)$ and $\sum s_i$ was then calculated assuming covariances of zero. We further assumed a normal distribution over $\sum s_i$ and Δ log fitness, so that bounds of the 95% CI were calculated as $\bar{x} \pm 1.96 \cdot \sigma/\sqrt{n}$, where x is the variable at hand ($\sum s_i$ or Δ log fitness), n is the number of replicas used to estimate s (normally 6), and σ is the standard deviation as described above.

Label	Effect on s	Mechanism	Description
Enhancing effect	$s_+ \uparrow$	Chaperones	Chaperones may increase the advantage of a novel allele by allowing accumulation of the novel protein form that may otherwise be sent to degradation.
		MOI	High MOI may increase the advantage of a novel beneficial allele through the accumulation of multiple copies of the protein in the cell
Masking effect	$s_+ \downarrow$	Chaperones	Chaperones may decrease the advantage of a novel beneficial allele by not allowing folding of a protein domain that is advantageous in the new environment; thus, for example, they may preferentially mediate folding of the original WT allele protein over the beneficial mutant allele.
		MOI	High MOI may decrease the advantage of a novel beneficial allele via complementation: for example, a viral capsid protein with enhanced binding may encapsidate a defective genome (e.g., with a defective polymerase). This will cause the the advantageous capsid to infect a cell to no avail, thus masking its potentially advantageous effect.
Deteriorating effect	$s_- \uparrow$	Chaperones	– Unlikely –
		MOI	High MOI may increase negative effects of novel deleterious allele via accumulation of multiple deleterious products and increased aggregation of unfolded proteins.
Buffering effect	$s_- \downarrow$	Chaperones	Chaperones may decrease the fitness cost of a novel deleterious allele by decreasing the probability of degradation of a novel protein fold.
		MOI	High MOI may decrease the negative effects of a novel deleterious allele via complementation: a defective genome may be encapsidated in a functional particle, thus reducing its fitness cost for one generation and increasing its frequency in the population.

TABLE S2: Possible effects of robustness on selection coefficients in a new environment, related to Figure 1.

3.3. Conclusions

For all but one of the nine host-changes (*Datura stramonium* as the exception), the confidence intervals support a negative Δ log fitness which is equal or well below -200 (the minimal fitness difference measured in Fig. 4 of the main text). Thus we can assume that under these host shifts, a brittle viral population would be at an advantage as compared to a more robust one. Calculating CI for [18] was not possible, as the raw data for s measurements was not available. However the mean Δ log fitness was also compatible with the conclusions above.

4. INTERACTIONS BETWEEN ROBUSTNESS AND SELECTION

Few studies have studied the interaction between robustness and how it affects beneficial mutations in a new environment. In Table S2, we explore different hypothetical options whereby robustness may alter the fitness effects of beneficial or deleterious mutations. In the section below, we analyze the effects of high MOI on deleterious and beneficial mutations. While our particular experimental results support masking of beneficial mutations, we envisage that other scenarios may be possible as well.

We define extrinsic robustness as robustness mediated by an external factor. Two such factors come to mind in the context of viruses: (a) cellular chaperones, which aid in regulating the maturation of viral proteins [3] [4], and (2) multiplicity of infection (MOI), as described in the main text. We list in Table S2 several modes whereby these extrinsic factors can affect the selection coefficient of an allele that changes is fitness effect following an environmental change.

5. MODELING A WITHIN-HOST CHANGE

In the simulations described in the main text (as displayed in Fig. 4), the new environment is characterized by an initial low MOI, regardless of whether the founder population was derived from low or high MOI in the previous host. However, it is possible that the population in the new environment is under conditions where mutations are still efficiently buffered. To examine this possibility, we considered the effect of a drug on a population replicating at high or low MOI Figure S2. One possibility is that the drug will immediately reduce the viral population size and, consequently, the MOI. In this case our model predicts a similar situation as that described for the host-to-host transmission (Fig. 3 of main text). However, if MOI remains high, for instance if the drug does not effectively inhibit the virus in all organs, then high MOI may be possible and thus co-infection may affect the magnitude of the selection coefficients in the new environment as described below.

As revealed herein and in previous experiments [19] co-infection may buffer the effects of deleterious mutations via complementation. However, what are the effects of high multiplicity of infection on beneficial alleles in a new environment? In theory, higher multiplicity of infection may either: (a) enhance the fitness effects of adaptive alleles through the potentially increased population frequency of these alleles, or (b) mask the fitness effects of adaptive alleles, for example when a genome bearing the adaptive allele is packaged in a defective capsid (see also Supplemental section 3).

To determine which of these two effects predominates we further analyzed poliovirus populations that were serially passed at low or high multiplicity of infection (Fig. 1C and D, main text). The results of our population genetic estimation of the DMFE indicate that beneficial mutations are masked at high MOI: the number of sites with a beneficial fitness effect, as well as the magnitude of this fitness effect, was smaller in the high MOI condition as compared to the low MOI condition ($p_i 0.005$; paired t-test). On average, the masking effect of the beneficial alleles consisted of a 6% reduction in the adaptive selection co-efficient. Thus, consistent with option (b) described above, it appears that under the high MOI passaging in this experiment, robustness reduces the beneficial fitness effects of advantageous mutations.

Informed by these results, we next simulated an environmental change within an infected host. Under this scenario, both the novel deleterious and the novel advantageous alleles are buffered (i.e., shifted towards neutrality) when the within-host change affects the high MOI virus populations, but not the low MOI virus populations. The simulation results encompassing 400 environmental changes show an essentially similar pattern as compared to the transmission low MOI scenario (Figure S2). However, the differences between robust and brittle populations become less pronounced. Essentially, this occurs because deleterious buffering reduces the fitness drop associated with an environment with a high influx of deleterious alleles, and also because beneficial masking reduces the advantage of the robust population when there is a high influx of beneficial alleles (Fig. 4A & B, main text).

Our results on beneficial mutational masking have interesting implications for disease management. While it is common to treat diseases early on, this implies that MOI will be low, allowing beneficial mutations (e.g., antiviral drug resistance mutations) to exert their effects in full. Later on in infection, when MOI is high, the beneficial effect of these same mutations may be masked, which may significantly slow down the rate of adaptation of the viral population to a drug. These issues are further compounded by the population size and the site of replication, calling for future studies on this topic.

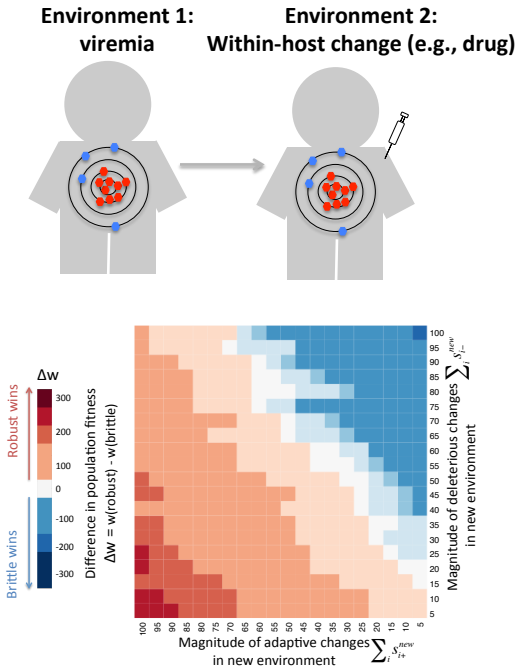

FIG. S2: Modelling of within-host changes, when robustness and selection in the new environment interact. Top: During a within-host change, populations evolving both at high and at low MOI are affected. Bottom: a map of fitness differences between robust and brittle populations over 400 different environmental changes of type within-host, where buffering and masking of deleterious and beneficial alleles, respectively, occurs in the robust populations. Notations are as in Fig. 4 of the main text. Related to Figure 4.

6. SIMULATION OF ENVIRONMENTAL CHANGES OVER TIME WITHOUT RECOMBINATION

6.1. Aims and specifications

The main text simulations and analytic calculations are based on the assumption of no linkage, i.e., infinite recombination. We note that this is clearly a limitation of our model (see discussion in the main text). However, for many viruses, recombination has been found to be extremely prevalent (e.g., [9][10]) . Specifically for Poliovirus, the rate of recombination has been estimated as 1-20% recombination events per genome per infectious cycle [11][13][12].

To further address how linkage affects our results, we test how robustness affects adaptation during the opposite scenario, namely sites that are linked (no recombination). We assume that most viruses will fall somewhere in between zero and infinite recombination. Tracking haplotypes with more than two linked sites requires tracking the mutations from every possible haplotype into every other possible haplotype (2^n haplotypes with n sites in the genome), which would be extremely computationally intensive. A proof of concept is hence shown below for two linked sites. Further research is required to determine the effects of linkage disequilibrium on robustness and adaptation to new environments.

6.2. Model description

Lets assume two binary sites in a genome. Four haplotypes are possible:
$00, 01, 10, 11$.

Fitness is multiplicative and defined as the product of fitness with respect to each locus, assuming no epistasis. Hence, fitness values for the haplotypes are:

$$00: \quad 1$$

$$
\begin{aligned}
01: &\quad 1 \times \omega_2 = 1 \times (1 + s_2) \\
10: &\quad \omega_1 \times 1 = 1 \times (1 + s_1) \\
11: &\quad \omega_1 \times \omega_2 = (1 + s_1)(1 + s_2)
\end{aligned}
$$

where w_i and s_i represent the fitness and selection coefficient at site i, respectively ($i \in 1, 2$). Assuming that in environment 1 $s_1 = s_2 = -0.1$, there are three robustness levels possible: $q = 1$ (both sites are neutral), $q = 0.5$ (one sites is neutral), $q = 0$ (zero sites are neutral).

6.3. Numerical simulations

Populations are first simulated with the WT haplotype (00) frequency set to 1, and evolved for $150,000$ generations to allow mutation-selection-drift equilibrium. Simulations of haplotype frequencies are performed using the Wright-Fisher model. Similar to the main text, binomial sampling is performed at each generation, and haplotype frequencies are weighted by their fitness and by the probability of mutation from one haplotype to another. Following an environmental change, one of the selection coefficients is randomly chosen and is changed to either -0.1 (deleterious) or 0.1 (advantageous). Figure S3A shows mean waiting times for adaptation of the novel advantageous allele, and figure S3B shows mean waiting times for purging of the novel deleterious allele. Adaptation and purging are defined as in the main text when an allele reaches a frequency of 0.99 or 0.01, respectively.

We next measure mean population fitness over time given each of the simplistic environmental changes performed (one selection coefficient becomes advantageous, or one becomes deleterious). Figure S3 depicts the behavior of the population over time.

6.4. Conclusions

Essentially, the results under infinite recombination hold, with the robust population outcompeting the robust one in the case of one allele becoming adaptive, and losing for the case of one allele becoming deleterious.

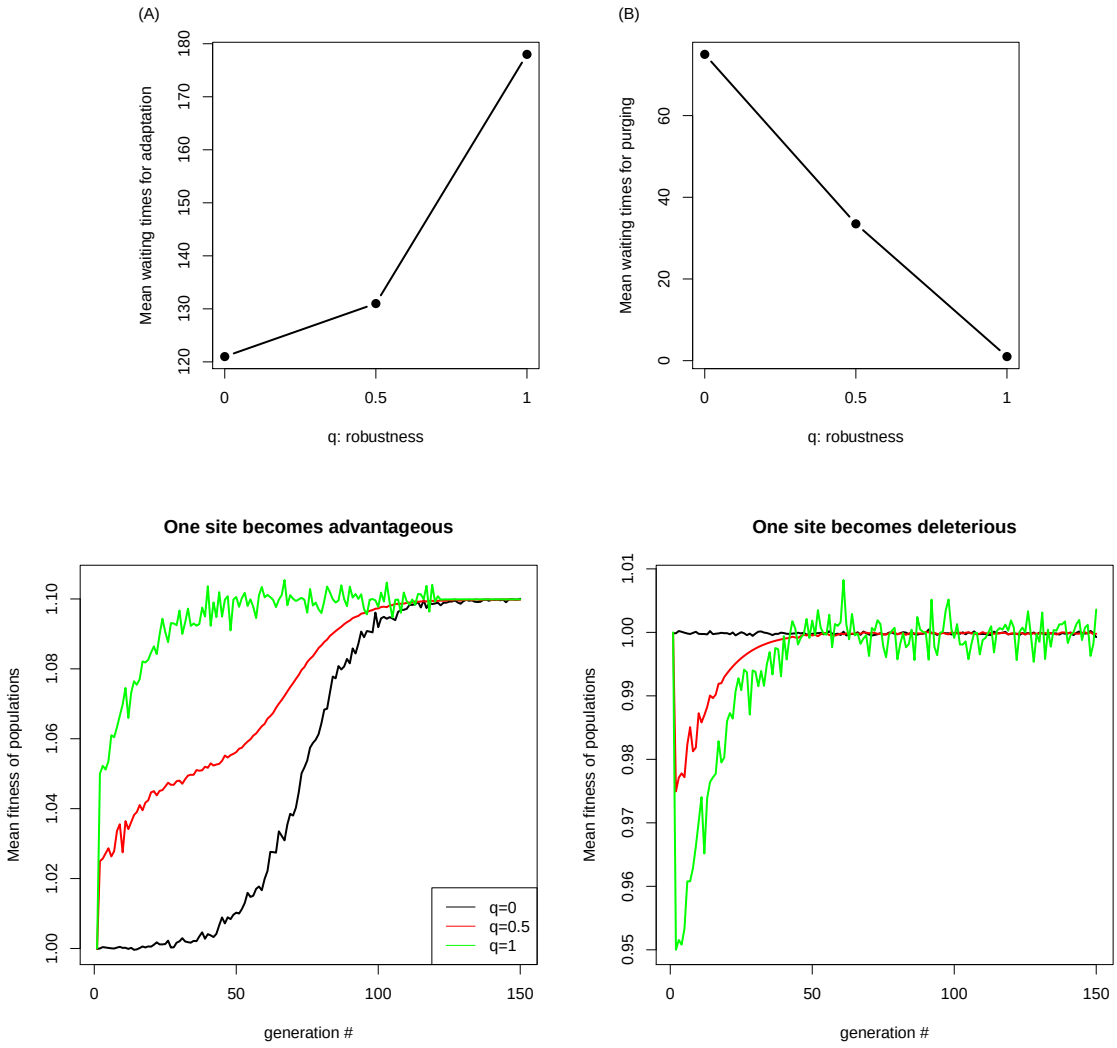

FIG. S3: Results of simulations assuming no recombination. Mean waiting times in generations for (A) adaptation of a novel adaptive allele, and (B) purging of a deleterious allele, following an environmental change. Mean fitness of populations of non-recombining genomes following an environmental change when (C) one site becomes advantageous, (D) one site becomes deleterious. Related to experimental procedures.

[1] W. J. Ewens, *Mathematical Population Genetics*, 2nd Ed., Springer, Berlin (2004).

[2] R. Nielsen and M. Slatkin, *An Introduction to Population Genetics: Theory and Applications*, 1st Ed., Sinauer Associates, Inc. (2013).

[3] R. Geller, M. Vignuzzi, R. Andino, J. Frydman, *Evolutionary constraints on chaperone-mediated folding provide an antiviral approach refractory to development of drug resistance.*, Genes Dev 21(2):195-205 (2007).

[4] R. Geller, R. Andino, J. Frydman, *Hsp90 Inhibitors Exhibit Resistance-Free Antiviral Activity against Respiratory Syncytial Virus*, PlOS ONE 8(2): e56762 (2013).

[5] J. Hermisson and P.S. Pennings, *Soft sweeps: molecular population genetics of adaptation from standing genetic variation.* Genetics 169, 2335-2352 (2005).

[6] J.A. Draghi, T.L. Parsons, G.P. Wagner, and J.B. Plotkin *Mutational robustness can facilitate adaptation*, Nature 463, 353-355 (2010)

[7] D.C. Krakauer and J.B. Plotkin *Redundancy, antiredundancy, and the robustness of genomes*, Proc Natl Acad Sci U S A 99, 1405-1409 (2002)

[8] E. Rajon and J. Masel *Evolution of molecular error rates and the consequences for evolvability*, Proc Natl Acad Sci U S A 18;108(3):1082-7 (2011)

[9] N. Tromas, M.P. Zwart, M. Poulain, and S.F. Elena*Estimation of the in vivo recombination rate for a plant RNA virus*, J

Gen Virol, 95(Pt 3):724-32 (2014)
[10] J. Zhuang, A.E. Jetzt, G. Sun, H. Yu, G. Klarmann, Y. Ron, B.D. Preston, and J.P. Dougherty *Human Immunodeficiency Virus Type 1 Recombination: Rate, Fidelity, and Putative Hot Spots* J Virol. 76(22):11273-82. (2002)
[11] K. Kirkegaard, D. Baltimore D *The mechanism of RNA recombination in poliovirus*, Cell 47: 433443 (1986)
[12] R. Duggal, A. Cuconati, M. Gromeier, E. Wimmer *Genetic Recombination of Poliovirus in a Cell-Free System* Proc Natl Acad Sci U S 94: 1378613791 (1997)
[13] C. Runckel, O. Westesson, R. Andino, J.L. DeRisi *Identification and Manipulation of the Molecular Determinants Influencing Poliovirus Recombination* PLoS Pathog. 9(2):e1003164 (2013)
[14] A. Acevedo, L. Brodsky, R. Andino *Mutational and fitness landscapes of an RNA virus revealed through population sequencing* Nature 505: 686-690 (2014)
[15] K.T. Thai, M.R. Henn, M.C. Zody., V. Tricou, N.M. Nguyet, P. Charlebois, N.J. Lennon, L. Green, P.J. de Vries, T.T. T.T., et al. *High-resolution analysis of intrahost genetic diversity in dengue virus serotype 1 infection identifies mixed infections* J Virol. 86, 835-843. (2012)
[16] B.T. Grenfell, O.G. Pybus, J.R. Gog, J.L. Wood, J.M. Daly, J.A. Mumford, E.C. *Holmes Unifying the epidemiological and evolutionary dynamics of pathogens* Science 303, 327-332. (2004)
[17] J. Lalic, J.M. Cuevas, S.F. Elena *Effect of host species on the distribution of mutational fitness effects for an RNA virus.* PLoS Genet 7, e1002378. (2011)
[18] P.F. Vale, M. Choisy, R. Froissart, R. Sanjuan, S. Gandon *THE distribution of mutational fitness effects of phage phiX174 on different hosts* Evolution 66, 3495-3507. (2012)
[19] R. Froissart, C.O. Wilke, R. Montville, S.K. Remold, L. Chao, and P.E. Turner *Co-infection weakens selection against epistatic mutations in RNA viruses* Genetics 168, 9-19. (2004)