SUPPORTING INFORMATION

I. THE NEUTRAL MUTATION RATE

To predict the patterns of fitness and mutation accumulation in the LTEE, our population genetic model utilizes the key approximation,

$$U\rho(s) \approx U_n\delta(s) + U_b\rho_b(s)\,, \tag{S1.1}$$

which partitions the DFE into a set of strongly beneficial "driver" mutations and a collection of nearly neutral "passengers" (Good and Desai, 2014; Schiffels et al., 2011). By assumption, the driver mutations set the important evolutionary timescales in the system (e.g., the rate of adaptation and the coalescence timescale), while the passenger mutations constitute a perturbative correction. In recent theoretical work, we have shown that this approximation is accurate when the mutation rate is small compared to the relevant fitness differences in the population (Good and Desai, 2014), which is expected to be the case for the non-mutator lines in the LTEE. Note that as defined above, the passenger portion of the DFE is comprised not only of truly neutral mutations ($|s| \lesssim 1/N$), but also those mutations that approach the neutral substitution rate by hitchhiking with the beneficial drivers. As such, U_n has an implicit dependence on both N and $U_b\rho_b(s)$.

Since U_n is an effective parameter, it cannot be measured directly. Instead, it must be self-consistently inferred from the data along with the other model parameters. In principle, this is straightforward: the passenger mutations do not influence the fitness trajectory by construction, so the neutral mutation rate can be estimated from a regression of the mutation trajectory residuals, $\overline{M}_{\rm obs}(t) - \overline{M}_b(t)$. Nevertheless, we will find it useful to place some crude bounds on U_n, both to limit the range of the regression and to allow for back-of-the-envelope arguments that do not require the full precision of our computational inference scheme. To establish these bounds, we will make use of the fact that the neutral mutation rate cannot exceed the total per genome mutation rate or the observed mutation accumulation rate.

Unlike U_n, the total genomic mutation rate is a directly measurable quantity, although these measurements can be confounded by dependence on the genetic background (Matic et al., 1997), genomic heterogeneity (Lang and Murray, 2008), and the role of complex mutational events. In previous work, Wielgoss et al. (2011) estimated that the total point mutation rate in the LTEE strain is approximately $U_{\rm point} \sim 4 \times 10^{-4}$ based on the number of synonymous mutations that have accumulated over the course of the experiment. A mutation accumulation study in a different strain of *E. coli* found that small indels occur approximately an order of magnitude less frequently than point mutations (Lee et al., 2012), or $U_{\rm indel} \approx 0.1 U_{\rm point}$. This is comparable to observations in other organisms such as yeast (Zhu et al., 2014). In comparison, much less is known about the mutation rate for larger indels and other chromosomal rearrangements, such as those arising from insertion sequence (IS) elements. This uncertainty is particularly problematic for the LTEE, since these complex mutations constitute a substantial fraction of the observed mutation trajectory. A recent mutation accumulation study in a different strain of *E. coli* estimated that the total rate of IS events in their background is approximately $U_{IS} \sim 3 \times 10^{-4}$. In the LTEE lines, ~ 100 rearrangements were observed across all twelve populations by generation 40,000 (Raeside et al., 2014), which suggests that the mutation rate to neutral IS mutations is bounded by $U_{IS,n} \leq 2 \times 10^{-4}$. However, this is likely to be an overestimate, since some of these mutations are probably beneficial. For example, in the single population where timecourse information is available (Ara$-$1), the accumulation of rearrangements slows significantly after 10,000 generations. If we restrict our attention to the mutations that have accumulated after generation 10,000, the bound on $U_{IS,n}$ drops to 6×10^{-5}, which is more in line with the rate of small indels. Note, however, that all of these estimates are based on extremely limited data, and are highly susceptible to statistical fluctuations and other ascertainment biases. In the absence of more precise estimates, we decided to employ the combined bound,

$$U_n \leq U_{\rm point} + U_{\rm indel} + U_{\rm IS,n} \lesssim 7 \times 10^{-4}\,, \tag{S1.2}$$

with the hope that any underestimation of the rate of complex mutational events is balanced by an overestimation of the neutral fraction of point mutations.

II. THE "RUNNING OUT OF MUTATIONS" MODEL

In the main text, we analyzed the patterns of fitness and mutation accumulation in a simple "running out of mutations" model of macroscopic epistasis. Here, we present our model in more detail, and show how the continuum analysis in

the main text emerges from a model which is fundamentally based on a finite number of sites. In the most general form of this model, we consider a collection of L_b sites with fitness effects $\{s_i\}_{i=1}^{L_b}$ and target sizes $\{\mu_i\}_{i=1}^{L_b}$, from which we can define a *joint* distribution of target sizes and fitness effects,

$$f_0(\mu, s) = \frac{1}{L_b} \sum_{i=1}^{L_b} \delta(\mu - \mu_i)\delta(s - s_i)\,. \tag{S2.1}$$

Here, $f_0(\mu, s)$ can be interpreted as the probability density that a randomly drawn site has a target size μ and fitness effect s. Similarly, the marginal distribution $f_0(s) = \int f_0(\mu, s)\, d\mu$ can be interpreted as the probability density that a randomly drawn site has fitness effect s, independent of the target size. Note that $f_0(s)$ differs from the traditional DFE, $\rho_0(s)$, which is the probability density that a randomly drawn *mutation* has fitness effect s. Since mutations will be biased towards sites with larger target sizes, the DFE corresponds to the weighted integral,

$$U_b \rho_0(s) = L_b \int \mu f_0(\mu, s)\, d\mu\ . \tag{S2.2}$$

So far, our discussion has been purely notational. The content of this model comes from the assumption that, once a given site has mutated, further mutations at that site are no longer beneficial. For example, after a loss-of-function mutation in a particular pathway, further loss-of-function mutations in the same pathway are expected to have little to no effect ($s \approx 0$), while the reversion will restore the original function of the gene ($s < 0$). In the large population limit, these mutated sites will behave as if they are effectively removed from the beneficial portion of the DFE. Based on this intuition, we define a collection of indicator variables $\{I_i(t)\}_{i=1}^{L_b}$, where $I_i(t) = 1$ if a mutation at site i has fixed by time t. We can use these indicator variables to define a time-dependent version of $f_0(\mu, s)$,

$$f(\mu, s, t) = \frac{1}{L_b} \sum_{i=1}^{L_b} [1 - I_i(t)]\delta(\mu - \mu_i)\delta(s - s_i)\,, \tag{S2.3}$$

which satisfies the initial condition $f(\mu, s, 0) = f_0(\mu, s)$. In the SSWM limit, $p_{\rm fix}(s) \approx 2s$ is independent of $f(\mu, s, t)$, so that the latter evolves as

$$\partial_t \langle f(\mu, s, t)\rangle = -2N\mu s \langle f(\mu, s, t)\rangle\,, \tag{S2.4}$$

or

$$\langle f(\mu, s, t)\rangle = f_0(\mu, s) e^{-2N\mu s t}\,, \tag{S2.5}$$

where $\langle f(\mu, s, t)\rangle$ denotes the expectation value over $\{I_i(t)\}_{i=1}^{L_b}$. It is similarly straightforward to show that the average fitness and mutation trajectories, $\overline{X}(t) = \sum s_i \langle I_i(t)\rangle$ and $\overline{M}(t) = \sum \langle I_i(t)\rangle$, evolve as

$$\partial_t \overline{X}(t) = \int 2NL_b\mu s^2 e^{-2N\mu s t} f_0(\mu, s)\, d\mu\, ds\,, \tag{S2.6a}$$

$$\partial_t \overline{M}_b(t) = \int 2NL_b\mu s e^{-2N\mu s t} f_0(\mu, s)\, d\mu\, ds\,. \tag{S2.6b}$$

Thus, our model corresponds to the strong-selection ($Ns \to \infty$) limit of the weak-mutation model analyzed by McCandlish et al. (2014). We recover Eqs. (5), (6), and (7) in the main text by demanding that all sites have the same target size $\mu = U_b/L_b$, so that

$$f_0(\mu, s) = \delta\left(\mu - \frac{U_b}{L_b}\right) \rho_0(s)\,. \tag{S2.7}$$

Note that in the present framework, our pseudo-continuous DFEs exist purely for notational convenience, since the set of available mutations is always bounded in practice. However, in certain cases it will be computationally convenient to consider the large L_b limit of these equations, in which sums over discrete numbers of sites are replaced with integrals over a continuous distribution. We have constructed our notation in such a way that this limit requires no modification of our equations, with the implicit caveat that we can only consider distributions for which this limit is well-defined. This excludes pathological cases like heavy-tailed DFEs, which lead to singular dynamics in the $L_b \gg 1$ limit.

Derivation of Eq. 10

In this section, we show how one can obtain a logarithmic fitness trajectory from a finite sites model with an ancestral DFE,

$$\rho_0(s) \propto \begin{cases} s^{-2} e^{-s/\sigma} & \text{if } s > \epsilon\sigma, \\ 0 & \text{else,} \end{cases} \tag{S2.8}$$

where ϵ is a small parameter. We will make use of the asymptotic expansion

$$I_p(\epsilon) \equiv \int_\epsilon^\infty \xi^{-p} e^{-\xi}\, d\xi \sim \begin{cases} \frac{\epsilon^{1-p}}{p-1} & \text{if } p > 1, \\ \log\left(\frac{1}{\epsilon}\right) & \text{if } p = 1, \\ \Gamma(1-p) & \text{if } p < 1. \end{cases} \tag{S2.9}$$

The normalizing constant for the DFE is therefore given by

$$C = \frac{\sigma}{\int_\epsilon^\infty \xi^{-2} e^{-\xi}\, d\xi} \approx \epsilon\sigma\,, \tag{S2.10}$$

and the mean and mean-squared fitness effects are

$$\langle s \rangle = \epsilon\sigma \int_\epsilon^\infty \xi^{-1} e^{-\xi}\, d\xi \approx \epsilon\sigma \log\left(\frac{1}{\epsilon}\right)\,, \tag{S2.11}$$

$$\langle s^2 \rangle = \epsilon\sigma^2 \int_\epsilon^\infty e^{-\xi}\, d\xi \approx \epsilon\sigma^2\,. \tag{S2.12}$$

Substituting Eq. (5) into Eq. (6), we find that the fitness trajectory is given by

$$\begin{aligned} \partial_t \overline{X}(t) &= \int 2NU_b s^2 e^{-2NUst/L} \rho_0(s) = 2NU_b \epsilon\sigma^2 \int_\epsilon^\infty \exp\left[-\left(\frac{2NU_b\sigma t}{L} + 1\right)\xi\right] d\xi\,, \\ &\approx \left(\frac{2NU_b\epsilon\sigma^2}{1 + \frac{2NU_b\sigma t}{L}}\right) e^{-\frac{2NU_b\sigma\epsilon t}{L}}\,, \end{aligned} \tag{S2.13}$$

which yields a logarithmic fitness trajectory

$$\overline{X}(t) = (L\sigma\epsilon) \log\left(1 + \frac{2NU_b\epsilon\sigma^2 t}{L\sigma\epsilon}\right)\,, \tag{S2.14}$$

provided that $t \ll L/(2NU_b\sigma\epsilon)$. When $t \approx L/(2NU_b\sigma\epsilon)$, we have $\overline{X}(t) \approx \overline{X}(\infty)$, so we can also write this condition in the form $\overline{X}(t) \ll \overline{X}(\infty)$. Even after fitting $X_c = L\sigma\epsilon$ and $v_0 = 2NU_b\epsilon\sigma^2$, there is still sufficient freedom to choose the parameters so that this condition holds for any finite time $t_{\max}$ or fitness $\overline{X}(t_{\max})$.

Heterogeneous target sizes

It is also useful to investigate the consequences of the running out of mutations model when we relax the uniform target size assumption. This becomes unwieldy in our original DFE notation, but simplifies considerably if we change variables from the target size μ to the substitution rate $R = 2N\mu s$. From the change of variables theorem, this induces induces a related joint distribution,

$$g_0(R, s) = \frac{f_0(R/2Ns, s)}{2Ns}\,, \tag{S2.15}$$

which allows us to rewrite the fitness and mutation trajectories in the form

$$\partial_t \overline{X}(t) = L_b \int sRe^{-Rt} g_0(R, s)\, dR\, ds = L_b \int h(R) Re^{-Rt} g_0(R)\, dR\,, \tag{S2.16a}$$

$$\partial_t \overline{M}_b(t) = L_b \int Re^{-Rt} g_0(R, s)\, dR\, ds = L_b \int Re^{-Rt} g_0(R)\, dR\,. \tag{S2.16b}$$

Here, $g_0(R) = \int g_0(s,R)\,ds$ is the marginal distribution of R and $h(R) = g_0(R)^{-1}\int s g_0(R,s)\,ds$ is the conditional mean of s given R. Thus, it is easy to see that for a *fixed* $h(R)$, the relationship between $\overline{X}(t)$ and $\overline{M}_b(t)$ is completely determined. For example, we recover Eq. (12) in the main text when $h(R) \propto R$, even if the distribution of target sizes is not completely uniform. However, it is also clear that if we are allowed to tune $h(R)$ and $g_0(R)$ independently, then it is possible to fit both $\overline{X}(t)$ and $\overline{M}(t)$ simultaneously with the inverse Laplace transforms

$$g_0(R) \propto \frac{\mathcal{L}^{-1}\left\{\partial_t \overline{M}_b\right\}}{L_b R}\,, \tag{S2.17a}$$

$$h(R) \propto \frac{\mathcal{L}^{-1}\{\partial_t \overline{X}\}}{\mathcal{L}^{-1}\{\partial_t \overline{M}_b\}}\,, \tag{S2.17b}$$

subject to the same technical restrictions on $\overline{X}(t)$ and $\overline{M}_b(t)$ that we encountered in the text. See also related results by (McCandlish et al., 2014), who study similar questions while relaxing the strong selection requirement.

For example, we can reproduce both the fitness and mutation trajectories of the global diminishing returns model in Eq. (22) by choosing

$$g_0(R) \propto \begin{cases} R^{-3/2} e^{-R/\tilde{R}} & \text{if } R > \epsilon\tilde{R}\,, \\ 0 & \text{else,} \end{cases} \tag{S2.18}$$

$$h(R) = \tilde{s}\left(\frac{R}{\tilde{R}}\right)^{1/2}\,, \tag{S2.19}$$

where

$$\tilde{R} = \frac{2NU_b\langle s\rangle\langle s\rangle_f}{X_c}\,, \quad c = \langle s\rangle_f\,, \quad L_b = \frac{2X_c}{\langle s\rangle_f\sqrt{\epsilon}}\,, \tag{S2.20}$$

and $\epsilon \ll 1$ is a small parameter chosen to maintain normalization. We can achieve this by choosing a joint distribution for R and s of the form

$$g_0(R,s) \propto \begin{cases} \delta\left[s - c\left(\frac{R}{\tilde{R}}\right)^{1/2}\right] R^{-3/2} e^{-R/\tilde{R}} & \text{if } R > \epsilon\tilde{R}, \\ 0 & \text{else.} \end{cases} \tag{S2.21}$$

Switching back to μ and s, we have

$$\begin{aligned} f_0(\mu,s) &\propto \begin{cases} (2Ns)\cdot\delta\left[s - c\left(\frac{2Ns\mu}{\tilde{R}}\right)^{1/2}\right](2Ns\mu)^{-3/2}e^{-2Ns\mu/\tilde{R}} & \text{if } 2Ns\mu > \epsilon\tilde{R}, \\ 0 & \text{else,} \end{cases} \\ &\propto \begin{cases} (2Ns)^{-1/2}\cdot\delta\left[\mu - \frac{\tilde{R}s}{2Nc^2}\right]\mu^{-3/2}e^{-2Ns\mu/\tilde{R}} & \text{if } 2Ns\mu > \epsilon\tilde{R}, \\ 0 & \text{else,} \end{cases} \\ &\propto \begin{cases} \delta\left[\mu - \frac{U_b\langle s\rangle s}{X_c c}\right] s^{-2}\exp\left[-\left(\frac{s}{c}\right)^2\right] & \text{if } s > c\sqrt{\epsilon}, \\ 0 & \text{else,} \end{cases} \end{aligned} \tag{S2.22}$$

where the normalization factor is

$$C = \frac{1}{\int_{c\sqrt{\epsilon}}^{\infty} s^{-2}e^{-(s/c)^2}\,ds} = \frac{2c}{\int_{\epsilon}^{\infty}\xi^{-3/2}e^{-\xi}\,d\xi} = c\epsilon^{1/2}\,. \tag{S2.23}$$

In terms of the traditional DFE, $\rho_0(s)$, we have

$$\rho_0(s) \propto \int \mu f_0(\mu,s)\,d\mu \propto \begin{cases} s^{-1}e^{-(s/c)^2} & \text{if } s > c\sqrt{\epsilon}, \\ 0 & \text{else,} \end{cases} \tag{S2.24}$$

where the overall mutation rate is given by

$$\begin{aligned} U_b &= L_b\int \mu f_0(\mu,s)\,d\mu\,ds = \frac{L_b\tilde{R}\sqrt{\epsilon}}{4Nc}\int_{\epsilon}^{\infty}\xi^{-1}e^{-\xi}\,d\xi\,, \\ &\approx \frac{v_0}{2Nc^2}\log\left(\frac{1}{\epsilon}\right)\,. \end{aligned} \tag{S2.25}$$

These expressions show that (within the context of the SSWM limit) the scaling of $\overline{X}(t)$ and $\overline{M}(t)$ with NU_b cannot be used to distinguish between the generalized running out of mutations model and the global diminishing returns model in Eq. (22). However, the time-dependent DFE, $\rho(s,t)$, differs between the two models, which implies that they still constitute different models of macroscopic epistasis. In principle, these differences can be elucidated by considering additional observables or by probing the scaling with N and U_b in the clonal interference regime.

III. COMPARISON WITH RECONSTRUCTION DATA

In the main text, we focused on signatures of epistasis in long-term patterns of fitness and mutation accumulation. Since our inferences were conducted at this aggregate level, it is natural to ask how they relate to more traditional measures of epistasis derived from the fitness effects of individual mutations. Actual data in this case is somewhat limited, given the general difficulty of identifying and reconstructing mutations that arose during the course of an evolution experiment. Fortunately, in the case of the LTEE, a small-scale study was recently carried out by Khan et al. (2011), who reconstructed all 2^5 allelic combinations of the first 5 mutations that fixed in the Ara$-$1 population. These reconstructions were used to measure the fitness effects of 5 mutations in 16 different genetic backgrounds, and 3 of the 5 mutations showed signatures of diminishing returns epistasis.

One the one hand, it is tempting to apply our global diminishing returns models directly to the full dataset, e.g., plotting $s(X)/s(0)$ as a function of X for each of the 5 mutations. Wiser et al. (2013) employed a related method to support the global diminishing returns model in Eq. (22). However, as we have argued above, we must be careful when extrapolating from macroscopic models of epistasis to fundamentally microscopic measurements. In particular, we showed in the text that one cannot define a consistent model of microscopic epistasis where $s(X) = s(0)e^{-X/X_c}$, since this would violate the bookkeeping property (Nagel et al., 2012). Thus, in a technical sense, the dependence of $s(X)/s(0)$ on X cannot provide additional evidence for the model in Eq. (22).

However, there is one aspect of the reconstruction data that *can* be used to gauge the support for the class of macroscopic models in Eq. (17). If global diminishing returns is responsible for the decelerating fitness trajectory in Fig. 1, then the fitness effects of fixed mutations should decrease along the line of descent, independent of their effects in other backgrounds. For example, in the SSWM limit, the distribution of fitness effects of fixed mutations (measured in the background in which they arise) must satisfy the scaling relation,

$$\rho_f(s|X) \propto \left(\frac{s}{f(X)}\right) \rho_0 \left(\frac{s}{f(X)}\right) . \tag{S3.1}$$

This distribution becomes more complicated in the clonal interference regime, but the scaling with $f(X)$ is approximately preserved up to logarithmic corrections (Good et al., 2012). In either case, we expect the fitness effects of fixed mutations to decrease roughly proportionally to $f(X)$. However, the data itself shows no such decrease (Fig. S5). Of course, this finding should be treated with a degree of caution, since it is based on a sample of five mutations from $\rho_f(s)$ with considerable sampling noise. Nevertheless, it implies that there is limited evidence in the Khan et al. (2011) data to suggest that global diminishing returns is driving the deceleration in the fitness trajectory. As suggested by Draghi and Plotkin (2013), the stronger diminishing returns signal off the line of descent may reflect ascertainment biases inherent in any set of fixed mutations.

REFERENCES

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.

Draghi, J. and Plotkin, J. B. 2013. Selection biases the prevalence and type of epistasis along adaptive trajectories. *Evolution* 67:3120–3131.

Good, B. H. and Desai, M. M. 2014. Deleterious passengers in adapting populations. *Genetics* 198:1183–1208.

Good, B. H., Rouzine, I. M., Balick, D. J., Hallatschek, O., and Desai, M. M. 2012. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc. Natl. Acad. Sci.* 109:4950–4955.

Khan, A. I., Din, D. M., Schneider, D., Lenski, R. E., and Cooper, T. F. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332:1193–1196.

Lang, G. I. and Murray, A. W. 2008. Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* 178:67–82.

Lee, H., Popodi, E., Tanga, H., and Foster, P. L. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci* 109:E2774–E2783.

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., and Elion, J. 1997. Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* 277:1833–1834.

McCandlish, D. M., Otwinowski, J., and Plotkin, J. B. 2014. On the role of epistasis in adaptation. *arXiv* .

Nagel, A. C., Joyce, P., Wichman, H. A., and Miller, C. R. 2012. Stickbreaking: A novel fitness landscape model that harbors epistasis and is consistent with commonly observed patterns of adaptive evolution. *Genetics* 190:655–667.

Raeside, C., Gaffé, J., Deatherage, D. E., Tenaillon, O., Briska, A., Ptashkin, R. N., Cruveiller, S., Médigue, C., Barrick, R. E. L. J. E., and Schneider, D. 2014. Large chromosomal rearrangements during a long-term evolution experiment with escherichia coli. *mBio* 5:e01377–14.

Schiffels, S., Szöllösi, G., Mustonen, V., and Lässig, M. 2011. Emergent neutrality in adaptive asexual evolution. *Genetics* 189:1361–1375.

Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R. E., Schneider, D., and Andrews, B. J. 2011. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with escherichia coli. *G3 (Bethesda)* 1:183–186.

Wiser, M. J., Ribeck, N., and Lenski, R. E. 2013. Long-term dynamics of adaptation in asexual populations. *Science* 342:1364–1367.

Zhu, Y., Siegal, M. L., Hall, D. W., and Petrov, D. A. 2014. Precise estimates of mutation rate and spectrum in yeast. *Proc Natl Acad Sci* 111:E2310–E2318.

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.

Draghi, J. and Plotkin, J. B. 2013. Selection biases the prevalence and type of epistasis along adaptive trajectories. *Evolution* 67:3120–3131.

Good, B. H. and Desai, M. M. 2014. Deleterious passengers in adapting populations. *Genetics* 198:1183–1208.

Good, B. H., Rouzine, I. M., Balick, D. J., Hallatschek, O., and Desai, M. M. 2012. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc. Natl. Acad. Sci.* 109:4950–4955.

Khan, A. I., Din, D. M., Schneider, D., Lenski, R. E., and Cooper, T. F. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332:1193–1196.

Lang, G. I. and Murray, A. W. 2008. Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* 178:67–82.

Lee, H., Popodi, E., Tanga, H., and Foster, P. L. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci* 109:E2774–E2783.

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., and Elion, J. 1997. Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* 277:1833–1834.

McCandlish, D. M., Otwinowski, J., and Plotkin, J. B. 2014. On the role of epistasis in adaptation. *arXiv* .

Nagel, A. C., Joyce, P., Wichman, H. A., and Miller, C. R. 2012. Stickbreaking: A novel fitness landscape model that harbors epistasis and is consistent with commonly observed patterns of adaptive evolution. *Genetics* 190:655–667.

Raeside, C., Gaffé, J., Deatherage, D. E., Tenaillon, O., Briska, A., Ptashkin, R. N., Cruveiller, S., Médigue, C., Barrick, R. E. L. J. E., and Schneider, D. 2014. Large chromosomal rearrangements during a long-term evolution experiment with escherichia coli. *mBio* 5:e01377–14.

Schiffels, S., Szöllösi, G., Mustonen, V., and Lässig, M. 2011. Emergent neutrality in adaptive asexual evolution. *Genetics* 189:1361–1375.

Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R. E., Schneider, D., and Andrews, B. J. 2011. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with escherichia coli. *G3 (Bethesda)* 1:183–186.

Wiser, M. J., Ribeck, N., and Lenski, R. E. 2013. Long-term dynamics of adaptation in asexual populations. *Science* 342:1364–1367.

Zhu, Y., Siegal, M. L., Hall, D. W., and Petrov, D. A. 2014. Precise estimates of mutation rate and spectrum in yeast. *Proc Natl Acad Sci* 111:E2310–E2318.

File S2

Estimated likelihoods used for the figures

Available for download as a .zip file at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.172460/-/DC1