

# **Rates of fitness decline and rebound suggest pervasive epistasis**

L. Perfeito<sup>1\*</sup>, A. Sousa<sup>1\*</sup>, T. Bataillon<sup>2\*\*</sup> and I. Gordo<sup>1\*\*</sup>

**1** Instituto Gulbenkian de Ciência, Oeiras, Portugal

**2** Bioinformatics Research Center, Aarhus University, 8000C Aarhus, Denmark

\*equal contribution

\*\* correspondence: [tbata@birc.au.dk](mailto:tbata@birc.au.dk) , [igordo@igc.gulbenkian.pt](mailto:igordo@igc.gulbenkian.pt)

Contact information: +351 214407915

Running title: Pervasive epistasis in fitness decline and rebound

Keywords: Mutation accumulation, Fitness recovery, Fisher geometrical model, Epistasis, Rate of adaptation

## ABSTRACT

Unraveling the factors that determine the rate of adaptation is a major question in evolutionary biology. One key parameter is the effect of a new mutation on fitness, which invariably depends on the environment and genetic background. The fate of a mutation also depends on population size, which determines the amount of drift it will experience. Here, we manipulate both population size and genotype composition and follow adaptation of 23 distinct *Escherichia coli* genotypes. These have previously accumulated mutations under intense genetic drift and encompass a substantial fitness variation. A simple rule is uncovered: the net fitness change is negatively correlated with the fitness of the genotype in which new mutations appear - a signature of epistasis. We find that Fisher's geometrical model can account for the observed patterns of fitness change and infer the parameters of this model that best fit the data, using Approximate Bayesian Computation. We estimate a genomic mutation rate of 0.01 per generation for fitness altering mutations, albeit with a large confidence interval, a mean fitness effect of mutations of -0.01 and an effective number of traits 9 in *mutS*<sup>-</sup> *E. coli*. This framework can be extended to confront a broader range of models with data and test different classes of fitness landscape models.

## INTRODUCTION

The distribution of effects of new mutations and how their interactions cause changes in fitness are fundamental quantities in our understanding of evolution. It is generally acknowledged that the genetic background has a strong influence on the phenotypic effect of mutations (De Visser et al. 2011). If the phenotype is relevant for fitness, natural selection will act differently on the mutation depending on its genomic background. Throughout this paper, a mutation is said to be epistatic if its fitness effect, and thus ultimately its fate in the population, depends on the genetic background where it appears.

Numerous population genetics models have been built assuming that mutations affect fitness

1 independently, i.e. that epistasis is absent. In the most classical experiment designed to determine  
2 the distribution of fitness effects of mutations, epistasis is typically ignored. In this type of  
3 experiment, populations are propagated at the lowest possible population size, minimizing the effect  
4 of natural selection and allowing deleterious mutations to accumulate close to the rate at which they  
5 appear (Bateman 1959; Mukai 1964). Interestingly, the results of these mutation accumulation  
6 (MA) experiments, have been continuously re-analyzed using different methods (Bataillon 2000;  
7 Halligan and Keightley 2009), but usually under the same model. The rate and mean effect of  
8 mutations are assumed constant and independent of the current fitness of the genotype hit by a new  
9 mutation. However, empirical work suggests that epistasis for fitness might be pervasive, with  
10 positive or negative mean depending on the organism studied (Sanjuán and Elena 2006) . Many  
11 studies are focused on deleterious mutations, as these are the most common. However recent  
12 experiments, focusing on discrete sets of beneficial mutations underlying fitness changes observed  
13 in experimental populations, support the notion that beneficial mutations may generally interact in  
14 an antagonistic way in microorganisms (Chou et al. 2011; Khan et al. 2011; Rokyta et al. 2011).  
15 Other studies show mean levels of epistasis ranging from antagonistic to nearly neutral in bacteria  
16 (Trindade et al. 2009; Ward et al. 2009).

17       Understanding the type of epistasis underlying fitness has important consequences for a  
18 number of evolutionary questions. For example, it is expected that sexual reproduction will be  
19 favored if most deleterious mutations interact synergistically (Kondrashov 1988). One important  
20 type of epistasis is sign epistasis, in which a mutation is deleterious or beneficial depending on the  
21 background. If it is common, then the number of paths that can be taken during adaptation might be  
22 severely constrained (Weinreich et al. 2005). The pattern of epistasis is often thought to be specific  
23 to a given species. However, recurrent observations suggest an interesting empirical rule: genotypes  
24 with lower fitness in a given environment adapt faster to that environment. This is observed across  
25 very diverse organisms: phage (Silander et al. 2007), *Escherichia coli* (Moore and Woods 2006;  
26 Khan et al. 2011), *Saccharomyces cerevisiae* (Murphy and Zeyl 2012), *Aspergillus nidulans*

(Gifford et al. 2011) and *Caenorhabditis elegans* (Estes and Lynch 2003). All of these studies are focused on the fitness effects of beneficial mutations in diverse backgrounds. An open question is whether or not the same rule applies to other types of mutations.

Many experiments, particularly those aimed at studying adaptation, can only detect the effects of mutations that survive genetic drift. Since the probability of escaping stochastic loss depends itself on the selection coefficient of mutations, estimating how fitness effects depend on genetic background requires taking into account the effect of natural selection. Three main factors affect the efficiency of selection and thus the probability of empirically detecting a beneficial mutation: population size, spatial structure and recombination. In small or fragmented populations, selection will be relatively inefficient both at purging deleterious and at fixing beneficial mutations. In very large populations, deleterious mutations will have little chance of fixing, while beneficial mutations enjoy a fixation probability directly proportional to their effect on fitness (Haldane 1927), assuming weak selection. In the absence of recombination, the efficiency of selection is also reduced because mutations are locked together in the same genotype. For example, if a beneficial mutation appears in a background that already harbors a deleterious mutation, it will only have a good chance of fixing if its effect outweighs the deleterious effect of the first mutation (Charlesworth et al. 1993; Bachtrog and Gordo 2004). In the same manner, if two beneficial mutations appear in a population on different genetic backgrounds, they will compete for fixation and one will be lost (Gerrish and Lenski 1998; Gerrish et al. 2007). This interference also occurs in the presence of small amounts of recombination (Barton 1995).

Here we analyze the results of two laboratory experiments, where population size is manipulated in order to modulate the efficiency of natural selection. Specifically we re-analyze the results of a MA experiment using a strain of *E. coli* that exhibits a mutator phenotype (*mutS*<sup>-</sup>). In this experiment *E. coli* evolved under drastic periodic bottlenecks consisting of a single cell. We also complement this MA study with a re-adaptation experiment. In this experiment several independent genotypes, stemming from the MA experiment, are allowed to evolve under large

1 effective population size. We explore how fitness changes (decreases in small populations and  
2 increases in large populations) depend on the genetic background. We show that the fitness change  
3 is strongly correlated with initial fitness in both experiments, suggesting that epistasis is very  
4 common.

5 Not accounting for epistasis fails to uncover the salient features of the fitness trajectories observed  
6 in each experiment. Therefore we should consider models where epistasis is pervasive. One widely  
7 used phenotype-based model of evolution is Fisher's geometrical model, hereafter referred to as  
8 FGM (Fisher 1930). This model has been used to interpret patterns of adaptation in large  
9 populations and to derive theoretical distributions of fitness effects of mutations and epistasis  
10 (Martin et al. 2007). It has also been invoked to help understanding the evolutionary trajectory of  
11 small populations (Tenaillon et al. 2007). So far, it has rarely been used to explain the fitness decay  
12 typically observed during evolution of very small populations, in particular of MA experiments (but  
13 see (Poon and Otto 2000; Martin and Lenormand 2006). Under FGM, one can derive the  
14 distribution of fitness effects,  $s$ , of a single mutation depending on the current background and  
15 pairwise epistasis  $\langle \epsilon \rangle$  between random mutations (Martin et al. 2007). The first is expected to  
16 follow a shifted gamma distribution with a given mean,  $E(s)$ , and the latter is predicted to be a  
17 normal distribution with mean  $\langle \epsilon \rangle = \text{zero}$ . Both  $E(s)$  and  $\langle \epsilon \rangle$ , do not depend on the genetic  
18 background, although the variances do in the simplest version of FGM. Other versions of the FGM  
19 where  $\langle \epsilon \rangle$  is non-zero have also been proposed (Gros et al. 2009). The simplest model pictures  
20 evolution as a walk in a geometric space of  $n$  phenotypic traits. It assumes that, within a given  
21 constant environment, a single value for each of the  $n$  of traits is optimal, with fitness decreasing  
22 smoothly for phenotypes away from that optimum. When a population is located at the optimum,  
23 any random mutation, irrespective of its size or direction, will cause a deviation from the optimum  
24 and so will necessarily be deleterious. FGM incorporates naturally a dependence of the rate of  
25 evolution on the initial phenotype. Importantly this model can account for several empirical patterns  
26 (Tenaillon et al. 2007; Gordo and Campos 2012) including the fitness effects of mutations

(Schoustra et al. 2009; Bataillon et al. 2011; Sousa et al. 2012; Trindade et al. 2012) and pairwise epistasis (Martin et al. 2007). Here, we use FGM to analyze the results of the measured fitness change, under both small and large population sizes. We introduce a statistical framework that allows combining information from both experiments to estimate the parameters underlying this model. We show that patterns of fitness change in both experiments can be quantitatively captured under FGM.

## MATERIALS AND METHODS

### *Bacterial strains and growth conditions*

The strains and growth conditions used in this study were as described in (Trindade et al. 2010). Briefly, for the mutation accumulation (MA) experiment and its analysis, the two strains used were *E. coli* K12 MG1655 srl::Tn10 *mutS* and K12 MG1655 srl::Tn10  $\Delta$ *ara mutS*. The first strain was the ancestral for the MA experiment, and the second was used as the reference strain for the competition assays in the fitness measurements (see below). The deletion in the arabinose operon is a phenotypic marker that allows distinguishing the two strains in competition, since these mutants give rise to red colonies when plated in Tetrazolium agar (TA) (Lenski et al. 1991).

For the recovery experiment, 23 clones derived from the MA lines were studied. These clones cover a range of initial fitnesses from 0.77 to 0.99 and include 16 clones from the 50<sup>th</sup> bottleneck, 6 clones from the 20<sup>th</sup> bottleneck and one clone from the 70<sup>th</sup> bottleneck.

To allow the clones to re-adapt, evolution under large population sizes was performed. In the first day, each experimental line was started with 10 ml of LB medium seeded with  $\sim 1.5 \times 10^4$  individuals and incubated at 37°C with agitation for 24h. This period allowed for  $\sim 20$  generations after which appropriate dilutions were made from each culture and another  $\sim 1.5 \times 10^4$  individuals were used to seed the next passage. This procedure was repeated for  $\sim 240$  generations after which fitness from all the lines was measured. One of the clones was propagated in 3 independent replicates and 5 were

propagated in 2 replicates. The remaining clones were propagated once, yielding in total 30 adapted populations.

#### *Mutation rate and Fitness estimation*

Since it is possible for the mutation rate to evolve, we looked for a phenotypic signature of a change in mutation rate, by a fluctuation assay. We observed that the mutator phenotype present in the ancestral clone was maintained in the 34 lines tested, these comprised the 23 lines used as the ancestors of the recovery experiment and 11 lines after the recovery experiment (Figure S1).

The fitness assays of the MA lines are described in (Trindade et al. 2010). Fitness of the recovery lines (after re-adaptation for 120 and 240 generations) was assayed in the same way as the MA lines. Briefly, competitor and reference strains were mixed in a proportion of 1:1 and used to seed 10 ml of LB. The competition tube was then grown for 24 hours at 37°C with agitation. Appropriate dilutions, of both initial and final bacterial populations, were plated in Tetrazolium agar (TA) to access the ratio between competitor and reference strains.

Selective coefficients were estimated as the difference in growth between the test strain and the reference strain, normalized by the number of generations elapsed for the reference strain (Chevin 2011). For each competition replicate, the difference in growth between the test strain ( $a$ ) and the reference strain ( $b$ ) is given by:

$$\Delta r_{ab} = \ln(N_{fa}/N_{ia}) - \ln(N_{fb}/N_{ib}),$$

where  $N_{fa}$  and  $N_{fb}$  are the number of test and reference bacteria, respectively, after the competition, and  $N_{ia}$  and  $N_{ib}$  are the number of test and reference bacteria, respectively, before the competition.

The relative growth rate estimated this way refers to a 24-hour growth cycle and must be normalized by the generation time of the reference in order to be comparable with the population genetics model we use. The number of generations ( $G$ ) elapsed for the reference was estimated for each competition replicate, assuming no death, as follows:

$$G = \log_2(N_{fb}/N_{ib}).$$

1 Fitness of strain  $a$  relative to the reference  $b$  ( $W_a$ ) is then given by:

2  $W_a = 1 + S_{ab}$ , where  $S_{ab} = \Delta r_{ab} / G$ ,

3 The fitness of the ancestral of the MA was measured in the same way and all fitness values were  
4 then standardized such that the fitness of this ancestral was 1.

5

6 *Simulations of the evolution process under FGM*

7 In order to simulate datasets under FGM that are directly comparable with our observed empirical  
8 data, we performed individual-based Monte Carlo simulations.

9 As the MA experiment involves bottlenecks of one cell every day, in the simulations of MA lines  
10 we assumed discrete generations and a demography that reflects this periodic fluctuation in  
11 population size. Accordingly, we crashed the populations to a size  $N=1$ , every 23 generations. After  
12 these crashes the population expanded, doubling in size every generation, until the next crash  
13 occurred. Therefore the harmonic mean of the population size was 23, which was taken as the  
14 effective population size  $N_e$  (Wahl et al. 2002)

15 In our MA experiment, the founding clone was highly fit, so we expect the vast majority of new  
16 mutations to be deleterious. We started the simulations of the MA experiment with a unique  
17 genotype residing at the phenotypic multivariate optimum, i.e. with a maximum fitness of 1. Given  
18 these conditions we simulated the mutation-selection-drift process under FGM. FGM assumes that  
19  $n$  phenotypic traits underlie fitness differences among individuals and each trait is under stabilizing  
20 (optimizing) selection. Without loss of generality we assumed that the optimum phenotypic value  
21 was zero for each trait. A special case of FGM was assumed by considering a Gaussian fitness  
22 landscape, whereby each individual is represented by a vector of  $n$  phenotypic trait values  
23  $\mathbf{x}=(x_1, x_2, \dots, x_n)$  and has fitness  $W=\exp(-\sum_i x_i^2)$ . Mutations are fully pleiotropic and occur each  
24 generation. The FGM we considered here is an isotropic version of the ones in Martin and  
25 Lenormand (2006) and (Waxman and Welch 2005).

26 We assumed that the number of new mutations in each individual was Poisson distributed with rate



$U$  per individual per generation. In addition, each mutation is fully pleiotropic and changes all traits ( $x_i$ ) and thus ultimately fitness. Each mutation therefore causes a change in each value of  $x_i$ , which is normally distributed with mean 0 and variance  $\sigma^2$ . Individuals are then selected to the next generation according to their fitness. In this model, the mean fitness effect of a mutation “hitting” a genotype initially residing at the multivariate phenotypic optimum is  $E(s) = -n\sigma^2$ .

For the recovery experiment, we assumed a constant effective population size, and three different initial starting fitnesses, covering the range of variation observed in the experiment.  $N_e$  was approximated by the harmonic mean of the population size  $N$  in the actual recovery experiments. We then let the population evolve for 240 generations and record the values of fitness increase after 120 and 240 generations, for each starting fitness. The slope of a linear regression of the fitness change, at each of these time points, as a function of initial fitness was then measured. As these simulations, involving a large  $N_e$ , are computationally very intensive, we did not follow all of the possible starting fitnesses considered in the actual experiment.

For the simulations in Figures 3, the same model was used, except only one mutation per run was recorded. The probability of each mutation surviving drift was estimated by  $(1-e^{-2s})/(1-e^{-2N_es})$  (Haldane 1927). The  $N_e$  used was 23.

Lastly, under the Gaussian FGM assumed here, the increase in fitness when close to the optimum is expected to be linear with initial fitness (Martin and Lenormand, 2008), and we have checked this assumption with a small subset of simulations performed with many starting fitnesses.

### *Estimation of parameters by Approximate Bayesian Computation (ABC)*

Our goal was to estimate jointly the genomic mutation rate for fitness ( $U$ ) and the parameters underlying FGM (i.e.  $n$  the number of independent phenotypic traits under stabilizing selection and  $\sigma$  the scaled mean phenotypic effect of a mutation) from empirical data (collectively termed  $\mathbf{D}$ ). Hereafter we “collect” the 3 parameters in a vector  $\boldsymbol{\theta} = (U, n, \sigma)$ .

Empirical data available to us were of two distinct kinds: fitness measurement of MA lines at

1 regular time intervals and fitness measurement during the fitness recovery experiment.

2 To estimate  $\theta$ , we note that the likelihood of our data  $\mathbf{D}$  cannot be obtained in a closed form under

3 FGM given the complications introduced jointly by demographics and selection. However, this

4 likelihood can be defined implicitly because we can simulate datasets fully comparable to  $\mathbf{D}$  under

5 FGM. We therefore used an approximate Bayesian framework where the likelihood of the full data

6  $\mathbf{D}$  was approximated by replacing the likelihood of  $\mathbf{D}$  by the likelihood of a vector of summary

7 statistics,  $\mathbf{S}_D$ , that summarizes our data  $\mathbf{D}$ . Note that  $\mathbf{S}_D$  comprises 39 statistics summarizing both

8 the MA experiment and the fitness recovery experiment (see Table 1).

9 Rejection sampling was then used to approximate the posterior distribution of  $\theta$  given  $\mathbf{S}_D$ . In brief,

10 to conduct rejection sampling, we simulated  $M$  “composite” datasets comprising data from both MA

11 and fitness recovery lines under FGM. Each simulation generating an MA + recovery dataset was

12 “seeded” by first drawing parameters from a joint prior distribution  $\pi(\theta)$ .

13 We chose the following priors for our parameters:  $\log_{10}(U)$  is uniform in  $[-3.5, -0.5]$  and

14 encompasses very broadly the range of genome wide mutation rates previously reported in *E. coli*;

15 we chose a very flat exponential prior for  $\sigma$  (see Fig. 4); and a uniform discrete distribution on  $[1,$

16  $30]$  that covers broadly the range of previously estimated values for this parameter (Martin and

17 Lenormand 2006; Bataillon et al. 2011; Sousa et al. 2012; Trindade et al. 2012)

18 We implemented rejection sampling by retaining the fraction  $tol$  of parameters that seeded

19 simulations that yielded the set of summary statistics closest to the vector of observed summary

20 statistics  $\mathbf{S}_D$ . In practice, we used  $M=10^6$  simulations and conducted rejection sampling by keeping

21 the closest 1000 simulations ( $tol=0.001$ ). The set of parameters yielding these 1000 accepted

22 datasets was used to approximate  $\Pi^*(\theta)$  the joint posterior distribution of the parameters.

23 We used a Monte Carlo approach to choose a tolerance level  $tol$  and study the statistical

24 performance of our ABC estimators by rejection sampling. One dataset was drawn at random from

25 the pool of  $10^6$  simulated datasets, rejection sampling was then performed using three tolerance

26 levels on the remaining datasets to obtain ABC parameter estimates. Prediction error was calculated

by adding mean (squared) bias and variance as  $MSE = \sum_i (\theta_i' - \theta_i^P)^2 / \text{Var}(\theta^P)$ , where  $\theta_i'$  ( $i=1..50$ ) are the ABC estimates and  $\theta_i^P$  are the true values of the parameters use for simulating data. Preliminary results based on 50 simulated datasets exploring three different tolerance levels ( $tol=0.1\%$ ,  $1\%$ ,  $10\%$ ) suggested that the lowest tolerance ( $tol = 0.1\%$ ) yielded the best prediction errors for the ABC estimates (not shown). We cannot exclude that ABC estimators with even lower MSE can be devised by using alternative weighting and/or alternative sets of summary statistics but this is beyond the scope of the present study. Moreover we showed via simulation that at  $tol=0.1\%$  and with the set of summary statistics we used, ABC estimates do not err systematically (Fig. S3). All rejection sampling, posterior approximations and Monte Carlo simulations for evaluating the performance of the ABC estimator were conducted with the statistical software *R* and using the package *abc* (Csilléry et al. 2012).

## RESULTS

### Initial fitness as a predictor of rates of fitness change through time

#### *(a) Initial fitness predicts the decrease in fitness in populations experiencing intense genetic drift*

We re-analyzed the data of a MA experiment in *E. coli* previously reported (Trindade et al. 2010) to test the extent to which initial fitness of genotypes can influence subsequent changes in fitness due to mutation accumulation. In this experiment, 50 lines of a mutator strain were propagated by daily transfer. Each transfer consisted of a bottleneck down to a single cell (see methods in (Trindade et al. 2010) and mean fitness was measured every 10 bottlenecks, i.e., approximately every 230 generations. During the course of this experiment, mean fitness was reduced by 10% on average, and some of the lines reached a fitness level as low as 0.78 (relative to the ancestor). The decline in mean fitness and increase in variance between MA lines through time allowed the estimation of the deleterious mutation rate ( $U_d$ ) and the mean effect of deleterious

1 mutations ( $E(s_d)$ ), using a variant of the Bateman-Mukai (BM) method, which accounts for the  
2 effects of natural selection between bottlenecks (Gordo and Dionisio 2005; Trindade et al. 2010).  
3 The inferred mutation rate was  $U=0.005$  deleterious mutations per genome per generation with a  
4 mean effect of  $E(s_d)=-0.03$ . This inference assumes that all mutations are deleterious with identical  
5 fitness effects, and that no epistasis occurs. There are two ways in which we can look for evidence  
6 of epistasis in an MA experiment. One method that has been widely used is to examine the decrease  
7 in fitness, averaged over all MA lines, through time. If there is no directional epistasis ( $\langle \epsilon \rangle = 0$ ), then  
8 the decrease in the logarithm of fitness with time is expected to be linear. Deviations from linearity  
9 indicate some level of mean epistasis, either positive or negative, depending on the curvature found  
10 (De Visser et al. 2011) . For this MA, the regression coefficient of the logarithm of fitness with time  
11 was 0.98 (Akaike information Criterion – AIC – of -36.3). A quadratic model applied to this data  
12 did not improve the fit (AIC= -36.5). Hence, using this method, we have no evidence that the mean  
13 epistatic effect was different from zero. However we note that this approach merely regresses the  
14 mean fitness of all lines. Independent lines accumulated different numbers of mutations along time;  
15 hence the mean may not be very sensitive. As a complement, we examined the change in fitness  
16 after a period of 10 bottlenecks as a function of the initial fitness before these 10 bottlenecks. If  
17 epistatic interactions do not exist, no correlation between the fitness of a line and its decrease in  
18 fitness in subsequent bottlenecks is to be expected. Contrary to this expectation, we observed that  
19 lines with a higher fitness were more likely to show a decrease in fitness than lines with lower  
20 initial fitness. Figure 1 shows that the change in fitness and initial fitness are correlated (Spearman  
21 correlation coefficient of -0.34,  $p < 10^{-7}$ ). We conclude that the change in fitness depends strongly on  
22 the initial fitness. This points to the occurrence of widespread epistasis among spontaneous  
23 mutations affecting fitness in *E. coli*. These findings prompted us to re-analyze the data under a  
24 model that explicitly accounts for the possibility of widespread epistasis among mutations thus  
25 yielding novel estimates of the rate and effects of mutations in this MA experiment (see below).

26

1 **(b) Initial fitness predicts the increase in fitness in large populations.**

2 Fitness changes observed in the course of MA experiments are generally negative (see  
3 (Halligan and Keightley 2009) for a review). We sought to perform a recovery experiment where  
4 the change in fitness is expected to be positive. We also wanted to test if the rate of fitness increase  
5 depends on the genetic background. We did so by using a sub-set consisting of 23 MA strains  
6 covering a wide range of relative fitness (from 0.77 to 0.99) to start our fitness recovery experiment.  
7 Among these strains, 16 were obtained from the 50<sup>th</sup> bottleneck, 6 were from the 20<sup>th</sup> bottleneck and  
8 the 23<sup>rd</sup> was obtained by propagating one of the MA lines for another 20 bottlenecks.

9 We adapted these clones to the medium where the competitions were performed, during 240  
10 generations, under serial dilution thereby imposing an effective population size ( $N_e$ ) of  $3 \times 10^5$ . We  
11 then measured fitness in the same conditions used for measuring fitness of the MA lines. Figure 2  
12 shows the change in mean population fitness as a function of the fitness of the founding clone.  
13 Clearly, the rate of adaptation depends on the genetic background (Spearman correlation coefficient  
14 of -0.76,  $p < 10^{-5}$ ). In fact, the lower the initial fitness of the clone, the higher its fitness increase.  
15 This relationship was observed for the clones sampled from the MA at the 20<sup>th</sup> and at the 50<sup>th</sup>  
16 bottleneck, as shown in Figure S2. In this experiment, 63% of the variance in fitness could be  
17 predicted by the initial fitness, and only two lines exhibited no increase in fitness during the 240  
18 generations (Fig. 2).

19 Although not impossible, back-mutation is an unlikely explanation for these results, given  
20 the estimates of point mutation rates in *E. coli* (Drake 1991) (Wielgoss et al. 2011). Compensatory  
21 mutations happening at other sites throughout the genome are, on the other hand, expected to be far  
22 more common (Poon et al. 2005). However, studies addressing their rate and distribution remain  
23 rare, with most studies examining compensation in the context of antibiotic resistance (Schrage et al.  
24 1997; Levin et al. 2000; Sousa et al. 2012).

# 1 Fisher geometrical model predicts the patterns observed in MA and in the recovery 2 experiment

3

## 4 (a) *The patterns of fitness change expected under FGM*

5       The data of both the MA and the recovery experiment show a clear dependence of fitness  
6 change on initial fitness ( $W_0$ ) as illustrated in Figures 1 and 2. This pattern has been very rarely  
7 reported in previous MA experiments (see Bataillon 2000; Halligan and Keightley 2009 for a  
8 review). We note however that under FGM, a fitness landscape model that has received a lot of  
9 theoretical attention, the distribution of effects of incoming mutations on a genotype, changes  
10 drastically as a function of its current fitness relative to the optimum. In particular, under FGM, we  
11 expect that in well-adapted populations new mutations will be mostly deleterious, while in poorly  
12 adapted populations a sizeable fraction of the mutations will be beneficial (compensatory).

13       We compared the predictions of FGM to the experimental observations, taking into account  
14 the demographic parameters used and the parameters that best fit the data (see the next section for  
15 the ABC approach we used to estimate them). We modeled the extreme fluctuations in population  
16 size in the MA, and allowed selection to occur during the growth of a line before it is bottlenecked  
17 (see Methods). When simulating FGM under these conditions, we observed that even with the  
18 drastic bottleneck procedure, the distribution of mutations accumulated is skewed by selection for  
19 beneficial and small effect deleterious mutations (see Fig. 3A). To evaluate the effect of selection  
20 occurring during the MA we simulated MA datasets with and without selection. In these  
21 simulations, mutations drawn from FGM accumulated with or without selection and we calculated  
22 the slopes of the linear regression of the change in fitness with starting fitness. Without selection the  
23 distribution of slopes was centered on zero with a small range of negative or positive values, while  
24 for MA data simulated with selection an average negative slope was observed (see Fig 3B). Using  
25 the appropriate demography for each experiment, we conclude that FGM can reproduce the patterns  
26 seen in the experimental data, both in the MA (Fig. 1) and fitness recovery (Fig. 2).

The linear relation of change in fitness with starting fitness appears to capture well the pattern of the recovery experiment. Under FGM, and assuming a Gaussian fitness landscape, Martin and Lenormand (2008) predicted a linear relation between the distance to the peak ( $s_0$ ) and the mean effect of beneficial mutations fixed during the first step of adaptation (i.e.  $E(s_b|fix)=4s_0/(4+n)$ ). This prediction is valid when the starting genotype is close to the optimum and ignoring clonal interference. In order to compare this prediction with our data, we took the clones with highest starting fitness ( $W_i>0.85$ ). These are the closest to the peak, where clonal interference is least expected to occur. If we assume that the fitness increase observed (DW) results from the first step of adaptation, we can calculate the slope of the linear regression of DW on  $s_0 \sim 1-W_i$ , and get a rough estimate of  $n$ . Under these assumptions  $n$  is estimated to be around 3.

#### ***(a) Estimating parameters of FGM from experimental data***

Given the capacity of FGM to reproduce the pattern present in the data (Fig. 1 and 2), we next sought to estimate jointly the parameters of this model ( $n$  and  $\sigma$ , see Methods for details) as well as the genomic mutation rate ( $U$ ) from our experimental data. To do so, we used an ABC framework where we combined empirical data from both experiments using several summary statistics (Table 1). These statistics describe the fitness distribution along bottlenecks of the MA experiment and the fitness change in two time points of the recovery experiment (see Methods). We then confronted these empirically observed summary statistics with those predicted under FGM given a value of  $U$ ,  $n$  and  $\sigma$ . ABC approaches have been very popular when analyzing patterns of nucleotide diversity under various demographic models (Robert et al. 2011). However, to our knowledge, they have never been used systematically to analyze these types of experiments. Previous studies relied on Monte-Carlo simulations of datasets which were used to “fit” data from MA experiments with EMS mutagenesis (Davies 1999) or to estimate fitness effect of two interfering beneficial mutations from patterns of adaptation over time in experimental populations (Hegreness 2006).

1 Since demography and selection are key factors to reproduce the empirical data, we cannot  
2 derive an explicit likelihood function in a closed form for each experiment. Instead, we adopted an  
3 ABC approach to obtain approximate posterior distributions for the genome-wide mutation rate and  
4 parameters of the FGM model. Our approach also allowed us to integrate data from both  
5 experiments (but see also fig S4 for posterior distribution of the parameters inferred separately from  
6 each experiment). We first evaluated the performance of our ABC estimator when considering data  
7 generated under the FGM model itself. Using sets of simulated datasets under FGM that mimic our  
8 data, we show that our ABC framework yields accurate estimates of genomic mutation rates and the  
9 FGM landscape parameters (Fig. S3).

10 Figure 4 shows the approximate posterior distributions obtained by our ABC method. Our estimates  
11 (median and 95% central posterior intervals) are  $U = 0.01$  [0.003 - 0.19],  $n = 9$  [4 - 20] and  $\sigma = 0.034$   
12 [0.01, 0.065]. The latter parameter estimates imply that the mean effect of single mutations on  
13 fitness is  $E(s) = -n\sigma^2 = -9 \cdot 0.034^2 = -0.01$ . This estimate of  $U$  is close to the one obtained using the  
14 BM method, and note that the posterior suggest substantial sampling variance. The difficulties in  
15 obtaining accurate estimates of  $U$  are well known for different methods (Keightley 2004). Our new  
16 estimate of the mean effect of mutations on fitness is smaller than previous BM estimate. This is not  
17 surprising given that it is well known that the BM method overestimates the mean selection  
18 coefficient of mutations when the variance of these effects is significant (Keightley and Eyre-  
19 Walker 1999; Halligan and Keightley 2009; Trindade et al. 2010).

## 21 DISCUSSION

22  
23 We reanalyzed jointly a traditional MA experiment and a fitness recovery experiment using  
24 genotypes resulting from the MA. The patterns of fitness degeneration in MA and subsequent  
25 recovery are not new and have been reported in other organisms (Estes and Lynch 2003). What is  
26 new here is the fact that our analysis of the data also provides strong evidence for the pervasive role



1 of epistasis among mutations underlying fitness and lends support to the idea that fitness effects of  
2 mutations vary substantially, i.e. the variance in mutational fitness effects is large. Our argument  
3 relies first on the observation that patterns of temporal fitness change in both experiments are  
4 strongly dependent on initial fitness levels. This qualitative pattern of strong genetic background-  
5 dependence in the data can hardly be explained without invoking the existence of epistasis. Second,  
6 we note that the fitness effects of new mutations theoretically expected under FGM, exhibit the  
7 salient features we empirically observed (Fig. 1 and 2). This motivates the use of FGM as a  
8 plausible framework for model-based statistical inference in experimental evolution. FGM has been  
9 recently used to derive quantitative predictions for the marginal distribution of fitness effects of  
10 single mutations and/or the expected patterns of epistasis among pairs of mutations (Martin et al.  
11 2007; Martin and Lenormand 2008; Bataillon et al. 2011). These predictions have been confronted  
12 with a few datasets. In particular the distribution of pairwise epistasis between deleterious  
13 mutations, in *E. coli*, has been found to follow a Normal distribution with mean zero, following the  
14 expectation of this simple model. FGM currently stands as an attractive alternative to statistical  
15 heuristics based on extreme value theory (Rokyta et al. 2008; Bataillon et al. 2011). Here we go one  
16 step further and demonstrate that parameters underlying FGM can also be estimated jointly with  
17 genome wide mutation rate from data on temporal fitness changes in either MA or adaptation  
18 experiments.

19 Most of the adaptation detected in the recovery lines is likely due to compensatory  
20 mutations. Compensatory mutations are epistatic by definition: they are only beneficial in a specific  
21 background as they allow a mutant strain to revert to the ancestral phenotype but do not occur at the  
22 same nucleotide (Nagaev et al. 2001). They contribute to adaptation by increasing the number of  
23 possible beneficial mutations as fitness decreases. In addition to strictly compensatory mutations, it  
24 is possible that the ancestral background was not fully adapted to the medium where the experiment  
25 took place. In that case, there can also be new mutations, which are beneficial irrespective of the  
26 background. It is possible that both changes in the beneficial mutation rate ( $U_b$ ) and in the mean

1 effect of beneficial mutations ( $s_b$ ) are contributing to the pattern observed in Figure 2. As was the  
2 case of the MA, our re-adaptation experiment supports the view that there is widespread epistasis  
3 between spontaneous mutations in *E. coli*. Empirical evidence that the distribution of fitness effects  
4 of compensatory mutations depends on the genetic background has been obtained in several studies  
5 examining the cost of antibiotic resistance (MacLean et al. 2010; Gordo et al. 2011). Some of us  
6 have recently used FGM in this context and have found that it could provide an accurate  
7 quantitative description of experimental data on the effects of these mutations (Sousa et al. 2012;  
8 Trindade et al. 2012). However, FGM has not been used to jointly study adaptation of different  
9 genotypes, which differ in mutations scattered throughout the whole genome. This is an important  
10 test, because it contributes to the establishment of general patterns.

11 Using FGM, one can derive how the rate and effects of both beneficial and deleterious  
12 mutations change along the evolutionary process, i.e. as we walk along the single peaked fitness  
13 landscape parameterized by the experimental data. Figure 3 illustrates our quantitative predictions,  
14 which can be challenged further by performing jointly several MA experiments and many replicate  
15 adaptation experiments, starting from distinct initial fitnesses (see for instance Gifford et al. 2011), a  
16 task that is beyond the scope of this paper. One salient qualitative feature of the distribution of  
17 spontaneous mutations under FGM is that while the mean effect does not depend on the genetic  
18 background, the variance does. In particular, as fitness decreases, the variance of the distribution of  
19 new mutations increases, with more strongly deleterious and strongly beneficial mutations. Because  
20 of that, the mean effect of mutations that survive genetic drift is predicted to depend strongly on  
21 initial fitness (Figure 3).

22 Our estimate of the genomic mutation rate towards fitness altering mutations in a *mutS*<sup>-</sup> *E.*  
23 *coli* is 0.01 per generation. Considering that this mutator has a 60-fold increase in mutation rate  
24 over the wild type, this estimate is similar to previous ones (Kibota and Lynch 1996; Trindade et al.  
25 2010) but note that the posterior interval we report  $U$  [0.003 - 0.19] is very broad.

26 Our estimate of the mean selection coefficient of a mutation,  $E(s) = -0.01$ , lies below our

1 previous one (Trindade et al. 2010). This is not surprising because our estimate is based on a model  
2 that intrinsically assumes a distribution of fitness effects with a variance larger than zero and  
3 explicitly accounts for compensatory mutations. This new estimate is close to that of Kibota and  
4 Lynch (1996).

5 Finally we have estimated a relatively small number of traits ( $n \sim 9$ ) under natural selection,  
6 which implies a low level of phenotypic complexity in this environment. This estimate is close to  
7 comparable measures of phenotypic complexity reported in other bacterial experiments (Bataillon et  
8 al. 2011; Trindade et al. 2012) but clearly more work is needed in that area and having a single  
9 common model to estimate  $n$  would help comparisons.

10 We have emphasized that FGM is an attractive model that naturally incorporates a number  
11 of biological important components: epistasis, various degrees of phenotypic complexity underlying  
12 fitness, pleiotropic effects of mutations, etc. We have also shown that FGM can account for patterns  
13 of fitness decline and recovery observed in empirical data. That being said, a number of alternative  
14 fitness models exists. One natural extension of our work is to consider a broader range of FGM  
15 based models featuring different forms of pleiotropy (Lourenço et al. 2011), different patterns of  
16 epistasis (Gros et al. 2009) , and/or different types of fixed, moving or “shaking” fitness optima  
17 (Kopp and Hermisson 2009; Gordo and Campos 2012). Another avenue is to examine radically  
18 different fitness landscape models (such as the mutational landscape model (Gillespie JH 1984),  
19 seascape model (Mustonen and Lässig 2009) or the recently proposed stickbreaking model (Nagel  
20 et al. 2012) and put these in competition with the various versions of FGM for explaining patterns  
21 of adaptation. We have shown that our ABC approach, although it is rather brute-force and relies on  
22 a large number of simulations, yields fairly accurate parameter estimates of the classical FGM (see  
23 Fig. S3) and this framework can be extended to encompass a variety of models. Although model  
24 selection under ABC requires caution (see Robert et al. 2011), the methodology is naturally prone  
25 towards such an analysis. Our approach can also be a fruitful starting point to devise novel  
26 frameworks for the analysis of experimental evolution empirical datasets that borrow strength from

- 1 both temporal patterns of fitness changes and snapshots of the genome at various time points of the
- 2 experiments.
- 3

## ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) /ERC grant agreement no 260421–ECOADAPT. I.G. acknowledges the salary support of LAO/ITQB & FCT. TB acknowledges financial support from The Danish Council for Independent Research (FNU).

## LITERATURE CITED

- Bachtrog, D. and I. Gordo. 2004. Adaptive evolution of asexual populations under Muller's ratchet. *Evolution* 58:1403–1413.
- Barton, N. H. 1995. Linkage and the limits to natural selection. *Genetics* 140:821–841.
- Bataillon, T. 2000. Estimation of spontaneous genome-wide mutation rate parameters: whither beneficial mutations? *Heredity (Edinb)* 84 ( Pt 5):497–501.
- Bataillon, T., T. Zhang and R. Kassen. 2011. Cost of Adaptation and Fitness Effects of Beneficial Mutations in *Pseudomonas fluorescens*. *Genetics* 189:939–949.
- Bateman, A. B. 1959. The viability of near-normal irradiated chromosomes. *International Journal of Radiation Biology* 1:170–180.
- Charlesworth, B., M. T. Morgan and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303.
- Chevin, L.-M. 2011. On measuring selection in experimental evolution. *Biol. Lett.* 7:210–213.
- Chou, H.-H., H.-C. Chiu, N. F. Delaney, D. Segrè and C. J. Marx. 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332:1190–1192.
- Csilléry, K., O. François and M. G. B. Blum. 2012. abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* 3:475–479.
- Davies, E. K. 1999. High Frequency of Cryptic Deleterious Mutations in *Caenorhabditis elegans*. *Science* 285:1748–1751.
- De Visser, J. A. G. M., T. F. Cooper and S. F. Elena. 2011. The causes of epistasis. *Proceedings of the Royal Society B: Biological Sciences* 278:3617–3624.
- Drake, J. W. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. U.S.A.* 88:7160–7164.
- Estes, S. and M. Lynch. 2003. Rapid fitness recovery in mutationally degraded lines of *Caenorhabditis elegans*. *Evolution* 57:1022–1030.
- Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Press: Oxford, UK.
- Gerrish, P. J., A. Colato, A. S. Perelson and P. D. Sniegowski. 2007. Complete genetic linkage can subvert natural selection. *Proc. Natl. Acad. Sci. U.S.A.* 104:6266–6271.
- Gerrish, P. J. and R. E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* 102-103:127–144.
- Gifford, D. R., S. E. Schoustra and R. Kassen. 2011. The length of adaptive walks is insensitive to starting fitness in *Aspergillus nidulans*. *Evolution* 65:3070–3078.
- Gillespie JH. 1984. Molecular evolution over the mutational landscape. *Evolution* 38:1116–1129.
- Gordo, I. and P. R. A. Campos. 2012. Evolution of clonal populations approaching a fitness peak. *Biology letters*, doi: 10.1098/rsbl.2012.0239.

1 Gordo, I. and F. Dionisio. 2005. Nonequilibrium model for estimating parameters of deleterious  
2 mutations. *Phys Rev E Stat Nonlin Soft Matter Phys* 71:031907.

3 Gordo, I., L. Perfeito and A. Sousa. 2011. Fitness effects of mutations in bacteria. *J. Mol. Microbiol.*  
4 *Biotechnol.* 21:20–35.

5 Gros, P.-A., H. Le Nagard and O. Tenaillon. 2009. The Evolution of Epistasis and Its Links With  
6 Genetic Robustness, Complexity and Drift in a Phenotypic Model of Adaptation. *Genetics* 182:277–  
7 293.

8 Haldane, J. B. S. 1927. A Mathematical Theory of Natural and Artificial Selection, Part V: Selection  
9 and Mutation. *Mathematical Proceedings of the Cambridge Philosophical Society* 23:838.

10 Halligan, D. L. and P. D. Keightley. 2009. Spontaneous Mutation Accumulation Studies in  
11 Evolutionary Genetics. *Annual Review of Ecology, Evolution, and Systematics* 40:151–172.

12 Hegreness, M. 2006. An Equivalence Principle for the Incorporation of Favorable Mutations in  
13 Asexual Populations. *Science* 311:1615–1617.

14 Keightley, P. D. 2004. Comparing analysis methods for mutation-accumulation data. *Genetics*  
15 167:551–553.

16 Keightley, P. D. and A. Eyre-Walker. 1999. Terumi Mukai and the riddle of deleterious mutation  
17 rates. *Genetics* 153:515–523.

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

20 Kibota, T. T. and M. Lynch. 1996. Estimate of the genomic mutation rate deleterious to overall  
21 fitness in *E. coli*. *Nature* 381:694–696.

22 Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature*  
23 336:435–440.

24 Kopp, M. and J. Hermisson. 2009. The genetic basis of phenotypic adaptation II: the distribution of  
25 adaptive substitutions in the moving optimum model. *Genetics* 183:1453–1476.

26 Lenski, R. E., M. R. Rose, S. C. Simpson and S. C. Tadler. 1991. Long-Term Experimental  
27 Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *The*  
28 *American Naturalist* 138:1315–1341.

29 Levin, B. R., V. Perrot and N. Walker. 2000. Compensatory mutations, antibiotic resistance and the  
30 population genetics of adaptive evolution in bacteria. *Genetics* 154:985–997.

31 Lourenço, J., N. Galtier and S. Glémin. 2011. Complexity, pleiotropy, and the fitness effect of  
32 mutations. *Evolution* 65:1559–1571.

33 MacLean, R. C., A. R. Hall, G. G. Perron and A. Buckling. 2010. The population genetics of  
34 antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature Reviews*  
35 *Genetics* 11:405–414.

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

38 Martin, G. and T. Lenormand. 2008. The Distribution of Beneficial and Fixed Mutation Fitness  
39 Effects Close to an Optimum. *Genetics* 179:907–916.

40 Martin, G. and T. Lenormand. 2006. The fitness effect of mutations across environments: a survey  
41 in light of fitness landscape models. *Evolution* 60:2413–2427.

42 Moore, F. B.-G. and R. Woods. 2006. Tempo and constraint of adaptive evolution in *Escherichia*  
43 *coli* (Enterobacteriaceae, Enterobacteriales). *Biological Journal of the Linnean Society* 88:403–411.

44 Mukai, T. 1964. The genetic structure of natural populations of *Drosophila Melanogaster*. I.  
45 spontaneous mutation rate of polygenes controlling viability. *Genetics* 50:1–19.

46 Murphy, H. A. and C. W. Zeyl. 2012. Prezygotic isolation between *Saccharomyces cerevisiae* and  
47 *Saccharomyces paradoxus* through differences in mating speed and germination timing. *Evolution*  
48 66:1196–1209.

49 Mustonen, V. and M. Lässig. 2009. From fitness landscapes to seascapes: non-equilibrium dynamics  
50 of selection and adaptation. *Trends Genet.* 25:111–119.

51 Nagaev, I., J. Björkman, D. I. Andersson and D. Hughes. 2001. Biological cost and compensatory  
52 evolution in fusidic acid-resistant *Staphylococcus aureus*. *Mol. Microbiol.* 40:433–439.

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

4 Poon, A., B. H. Davis and L. Chao. 2005. The coupon collector and the suppressor mutation:  
5 estimating the number of compensatory mutations by maximum likelihood. *Genetics* 170:1323–  
6 1332.

7 Poon, A. and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of  
8 small populations. *Evolution* 54:1467.

9 Robert, C. P., J.-M. Cornuet, J.-M. Marin and N. S. Pillai. 2011. Lack of confidence in approximate  
10 Bayesian computation model choice. *Proc. Natl. Acad. Sci. U.S.A.* 108:15112–15117.

11 Rokyta, D. R., C. J. Beisel, P. Joyce, M. T. Ferris, C. L. Burch and H. A. Wichman. 2008. Beneficial  
12 Fitness Effects Are Not Exponential for Two Viruses. *Journal of Molecular Evolution* 67:368–376.

13 Rokyta, D. R., P. Joyce, S. B. Caudle, C. Miller, C. J. Beisel and H. A. Wichman. 2011. Epistasis  
14 between beneficial mutations and the phenotype-to-fitness Map for a ssDNA virus. *PLoS Genet.*  
15 7:e1002075.

16 Sanjuán, R. and S. F. Elena. 2006. Epistasis correlates to genomic complexity. *Proc. Natl. Acad. Sci.*  
17 *U.S.A.* 103:14402–14405.

18 Schoustra, S. E., T. Bataillon, D. R. Gifford and R. Kassen. 2009. The properties of adaptive walks  
19 in evolving populations of fungus. *PLoS Biol.* 7:e1000250.

20 Schrag, S. J., V. Perrot and B. R. Levin. 1997. Adaptation to the fitness costs of antibiotic resistance  
21 in *Escherichia coli*. *Proc. Biol. Sci.* 264:1287–1291.

22 Silander, O. K., O. Tenaillon and L. Chao. 2007. Understanding the evolutionary fate of finite  
23 populations: the dynamics of mutational effects. *PLoS Biol.* 5:e94.

24 Sousa, A., S. Magalhães and I. Gordo. 2012. Cost of antibiotic resistance and the geometry of  
25 adaptation. *Mol. Biol. Evol.* 29:1417–1428.

26 Tenaillon, O., O. K. Silander, J.-P. Uzan and L. Chao. 2007. Quantifying Organismal Complexity  
27 using a Population Genetic Approach. *PLoS ONE* 2:e217.

28 Trindade, S., L. Perfeito and I. Gordo. 2010. Rate and effects of spontaneous mutations that affect  
29 fitness in mutator *Escherichia coli*. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 365:1177–1186.

30 Trindade, S., A. Sousa and I. Gordo. 2012. Antibiotic resistance and stress in the light of Fisher's  
31 model. *Evolution* no–no.

32 Trindade, S., A. Sousa, K. B. Xavier, F. Dionisio, M. G. Ferreira and I. Gordo. 2009. Positive  
33 Epistasis Drives the Acquisition of Multidrug Resistance. *PLoS Genetics* 5:e1000578.

34 Wahl, L. M., P. J. Gerrish and I. Saika-Voivod. 2002. Evaluating the impact of population  
35 bottlenecks in experimental evolution. *Genetics* 162:961–971.

36 Ward, H., G. G. Perron and R. C. Maclean. 2009. The cost of multiple drug resistance in  
37 *Pseudomonas aeruginosa*. *J. Evol. Biol.* 22:997–1003.

38 Waxman, D. and J. J. Welch. 2005. Fisher's microscope and Haldane's ellipse. *Am. Nat.* 166:447–  
39 457.

40 Weinreich, D. M., R. A. Watson and L. Chao. 2005. Perspective: Sign epistasis and genetic  
41 constraint on evolutionary trajectories. *Evolution* 59:1165–1174.

42 Wielgoss, S., J. E. Barrick, O. Tenaillon, S. Cruveiller, B. Chane-Woon-Ming, C. Medigue, R. E.  
43 Lenski, D. Schneider and B. J. Andrews. 2011. Mutation Rate Inferred From Synonymous  
44 Substitutions in a Long-Term Evolution Experiment With *Escherichia coli*. *G3* 1:183–186.

45 Genes|Genomes|Genetics 1:183–186.

46

47

48

49

1

2

3

4

5

6



Table 1: Overview of the set of summary statistics used for ABC inference

Summary Statistic	Observed	Experiment
<0.77 <sup>a</sup>	0	Bot 10
[0.77 - 0.81[	0	
[0.81 - 0.85[	1	
[0.85 - 0.89[	1	
[0.89 - 0.93[	3	
[0.93 - 0.97[	20	
>0.97	25	
<0.77	0	Bot 20
[0.77 - 0.81[	0	
[0.81 - 0.85[	4	
[0.85 - 0.89[	2	
[0.89 - 0.93[	12	
[0.93 - 0.97[	24	
>0.97	8	
<0.77	0	Bot 30
[0.77 - 0.81[	0	
[0.81 - 0.85[	2	
[0.85 - 0.89[	10	
[0.89 - 0.93[	10	
[0.93 - 0.97[	21	
>0.97	7	
<0.77	0	Bot 40
[0.77 - 0.81[	2	
[0.81 - 0.85[	5	
[0.85 - 0.89[	8	
[0.89 - 0.93[	16	
[0.93 - 0.97[	13	
>0.97	6	
<0.77	0	Bot 50
[0.77 - 0.81[	4	
[0.81 - 0.85[	10	
[0.85 - 0.89[	9	
[0.89 - 0.93[	15	
[0.93 - 0.97[	11	
>0.97	1	
slope 120 <sup>b</sup>	-0.7986	Fitness Recovery experiment
intercept 120	0.7585	
slope 240	-0.8873	
intercept 240	0.8647	

a: Observed distribution fitness values in the MA experiment. Distributions at each time point (Bottleneck 10, 20,30,40,50) are summarized using counts in 7 classes of fitness values (all counts sum to 50 at each time point).  
b: The observed slope and intercept of the regression line describing the fitness recovery (after 120 or 240 generations) as a function of initial fitness.

## List of Figures

**Figure 1. Fitness change as a function of initial fitness during the MA.** The fitness change of each line over 10 bottlenecks (or 230 generations) is shown as a function of that line's fitness before that interval: the difference between the 10<sup>th</sup> and the 20<sup>th</sup> bottlenecks is shown as a function of fitness at the 10<sup>th</sup> bottleneck, the difference between the 20<sup>th</sup> and the 30<sup>th</sup> is shown as a function of fitness at the 20<sup>th</sup>, and so on. Each MA line is represented 4 times, once for each interval. The experimental data (shown as dots) is taken from (Trindade et al. 2010). Fifty simulations of datasets under FGM (shown as crosses) were performed as described in the text, using as parameters  $n=9$ ,  $U=0.01$  and  $\sigma=0.034$ . In both cases, a significant correlation between the change in fitness and initial fitness was obtained. For the experimental data, the slope of the linear regression is -0.32 with CI (-0.43, -0.21) and for the data under FGM the slope is -0.15 with the 95% confidence interval (-0.24, -0.05).

**Figure 2. Fitness change as a function of the initial fitness.** Fitness change after 12 bottlenecks (shown as dots) as a function of the initial fitness. Clones with different starting fitness were adapted for 240 generations under a large effective population size (see Methods for details). The results of FGM simulations with parameters values set to  $n=9$ ,  $U=0.01$  and  $\sigma=0.034$  are shown as crosses. In both cases, a significant correlation between the change in fitness and initial fitness was obtained. For the experimental data, the slope of the linear regression is -0.77 with CI (-0.97, -0.57) and for the data under FGM the slope is -0.84 with the 95% confidence interval (-0.86, -0.82).

**Figure 3. A) Distribution of fitness effect of mutations before and after selection.** The distribution of spontaneous mutations was generated by at least 100000 Monte-Carlo simulations of

FGM as described in the text, using as parameters  $n=9$ ,  $U=0.01$  and  $\sigma=0.034$ . The distribution of mutations that escape genetic drift was estimated as the fraction of spontaneous mutations that escape drift in a population of effective size of 23 individuals, corresponding to the population expansion occurring in the MA (Gordo and Dionisio 2005). **B) Distribution of slopes of the linear regression of change in fitness with initial fitness under FGM** with no selection (black bars, mean = -0.003) and with selection (grey bars, mean = -0.015) in the MA lines. The dashed line shows the location of slope 0. Parameter values used are the same as in panel A.

**Figure 4. Approximate posterior densities for parameters of FGM.** All posterior densities are approximated using a set 1000 accepted simulations (tolerance rate of 0.001, see methods for details). Prior densities are in orange. Marginal posteriors are in blue and regions of highest posterior densities are in grey. **(A)** Prior and approximate marginal posterior for the genome wide mutation rate for fitness ( $U$ ). Note the  $\log_{10}$  scale. **(B)** Prior and approximate marginal posterior for the scaled phenotypic effect of mutations,  $\sigma$ . **(C)** Prior and approximate posterior probability distributions for the number of dimensions,  $n$ . Note that these are discrete probability distributions **(D)** Bivariate approximate posterior densities for  $U$  and  $\sigma$ . A blue circle depicts the location of the region with highest posterior density. The regions comprising 0.5, 0.9, 0.95 and 0.99 of the probability mass of the posterior density are depicted using grey shadings. **(E)** Bivariate approximate posterior densities for  $U$  and  $n$ . Legend as in D. Note that for graphical convenience the joint posterior distribution over  $n$ , and  $U$  was smoothed by “jittering” the values of  $n$  to make them apparently continuous. **(F)** Bivariate approximate posterior densities for  $\sigma$  and  $n$ . Legend as in 3D. Note that for graphical convenience the joint posterior distribution over  $n$ , and  $\sigma$  was smoothed by jittering the values of  $n$  to make them apparently continuous.

List of supporting information figures caption

**Supporting Figure S1 – Fitness changes as a function of initial fitness for two points in the MA experiment.** The blue dots show the data for the 6 clones taken from the 20<sup>th</sup> bottleneck (B20), the red dots show the clones taken from the 50<sup>th</sup> bottleneck (B50) and the green dot corresponds to the clone from the 70<sup>th</sup> bottleneck (B70).

**Supporting Figure S2. Change in fitness after 120 (red symbols) and 240 generations of adaptation (black symbols).**

**Supporting Figure S3. Distribution of ABC estimates for data simulated under FGM and known genomic mutation rates.** We used 500 simulated datasets to estimate the bias and sampling variance of the ABC estimator for the joint parameters  $\theta = (U, n, \sigma)$  of the model. Datasets were simulated using a broad range of parameters values (spanning the prior we used for our data analysis) and to match the data we analyze here (i.e. an MA experiment and a fitness recovery experiment). Bias is visualized on the scatter plots. Perfect unbiased estimation would follow the  $y = x$  line (orange) and the difference between the ideal line and the actual local robust regression (blue line) reveals the bias. The number of dimensions underlying the fitness landscape,  $n$ , is estimated with some imprecision (see scatter above and below the diagonal). Bias in estimation varies with the true number of dimensions:  $n$  is systematically overestimated by about 1 dimension for  $n < 10$ , and underestimated for  $n > 15$ . Estimation of the mean phenotypic effect of a single mutation,  $\sigma$  is fairly precise with bias and variance being minimum in the area where our data lies ( $\sigma < 0.1$ ). Estimation of  $U$  is virtually unbiased in the range of  $U$  values that are likely of empirical interest ( $U=0.005$  to  $U=0.05$ ).

Figure S4. Posterior distributions obtained by fitting separately MA and Fitness Recovery datasets.

1 The ABC inference was carried as described in the methods but using separately the fitness  
2 recovery dataset (4 summary statistics) and the fitness of all 50 MA lines at 5 times points ( $7*5=35$   
3 summary statistics).

4

5 **Figure S4. Marginal Posterior distributions obtained by fitting separately MA and Fitness**  
6 **Recovery datasets.**

7 The ABC inference was carried as described in the methods but using separately the fitness  
8 recovery dataset (4 summary statistics) and the fitness of all 50 MA lines at 5 times points ( $7*5=35$   
9 summary statistics). We report in orange the prior used for the each analysis (an identical prior was  
10 used for each inference) and the marginal posteriors obtained for the genome-wide mutation rate  
11 (**panel A**), the phenotypic effect of a mutation (**panel B**) and the number of dimensions underlying  
12 the fitness landscape (**panel C**). For all the parameters there is a large overlap between the posterior  
13 distributions for both experiments.