

# Global Epistasis Makes Adaptation Predictable Despite Sequence-Level Stochasticity

Sergey Kryazhimskiy, Daniel P. Rice,  
Elizabeth Jerison, Michael M. Desai

## Supplemental Online Materials

### Contents

<b>Materials and methods</b>	<b>1</b>
Experimental procedures . . . . .	1
Strains . . . . .	1
Evolution and fitness assays . . . . .	2
Full-genome sequencing . . . . .	5
Knock-out experiments . . . . .	5
Analysis . . . . .	6
Identification of mutations from Illumina reads . . . . .	6
Analysis of covariance of fitness increment with initial fitness . . . . .	9
Analysis of genetic data . . . . .	20
Correlation between fitness increment and the number of mutations . . . . .	22
Interpretation of data in terms of Fisher’s Geometric Model . . . . .	23
<b>References</b>	<b>26</b>
<b>Supplementary Figures</b>	<b>27</b>
<b>Supplementary Tables</b>	<b>37</b>

### Materials and methods

#### Experimental procedures

##### Strains

All strains used in this study are derived from strain DBY15084 (1), a haploid MATa strain with the W303 background and the following genotype: *ade2-1*, *CAN1*, *his3-11*,

*leu2-3,112*, *trp1-1*, *bar1Δ::ADE2*, *hmlαΔ::LEU2*. The ancestor of the evolution experiment, strain ySAK0001 (DivAnc for short), is derived from strain DBY15105 (also known as yGIL432), which has two additional genetic modifications on the DBY15084 background: *gpa1Δ::GPA1-G1406T-NatMX* and *ura3::P<sub>FUS1</sub>-yEVenus* (1,2). Specifically, DivAnc is a clone picked from population BYS1-B05 that was founded with DBY15105 and evolved at the small bottleneck size for 625 generations of the long-term evolution experiment described in Ref. (2).

“Founders” are clones picked from populations founded with DivAnc and evolved in the Diversification phase. Founders were selected via a procedure described in section Diversification and selection of Founders, and subsequently used to found populations for the Adaptation phase. All Founders have shorthand ids “Sddd” or “Lddd”, where “ddd” is a three-digit number, e.g., L013. In addition, Founders whose descendants were sequenced were assigned strain ids ySAK0057 through ySAK0072 (see Table S1). “Evolved clones” are clones picked from populations evolved for 500 generations in the Adaptation phase and are denoted by *x-y-z*, where *x* is the Founder strain, *y* is the population id, and *z* is the clone id, e.g., L013-8-1. Evolved clones L013-8-1, L041-8-2, L096a-4-1, and L096b-6-1 used for knock-out experiments (see section Knock-out experiments) were assigned strain ids ySAK0073 through ySAK0076, respectively (see Table S1).

Two strains, DBY15108 (RmRef for short) and SAK0077 (DivAncCit for short), were used as reference strains for flow cytometry-based fitness assays. Strain DBY15108 has two additional genetic modifications on the DBY15084 background: *gpa1Δ::GPA1<sup>WT</sup>-NatMX* and *ura3Δ::ymCherry* (1,2). Strain ySAK0077 was constructed from DivAnc by amplifying *HIS3-ymCitrineM233I* construct from genomic DNA of strain yJHK111 (courtesy of Melanie Müller, John Koschwanez, and Andrew Murray, Department of Molecular and Cellular Biology, Harvard University) using primers oGW137 and oGW138 (Table S8) and integrating it into the *his3* locus.

Complete list of strains used in this study is given in Table S1. Isolation of clones and transformations were carried out using standard yeast genetics protocols (3).

## Evolution and fitness assays

All evolution experiments and fitness assays were carried out in 128  $\mu$ l of rich laboratory medium YPD (1% yeast extract (BD, VWR catalog #90000-722), 2% peptone (BD, VWR catalog #90000-368), and 2% dextrose (BD, VWR catalog #90000-904)) in round-bottom polystyrene 96-well plates (Corning, VWR catalog #29445-154). Plates were maintained at +30°C without shaking. Serial transfers were performed using the Biomek FX liquid handling robot. Before each transfer, cultures were thoroughly mixed by either pipetting 20  $\mu$ l up and down 20 times or by shaking the plates on an orbital plate shaker (Heidolph Titramax 100) at 1000 rpm for at least 2 minutes. To monitor the rate of cross-contamination events during evolution, all 96-well plates had a pattern of blank wells. No cross-contamination events were observed.

**Diversification and selection of Founders** The purpose of the Diversification experiment was to create a set of clones (referred to as Founders) that would all have DivAnc as their most recent common ancestor, be different from each other by a small number of mutations, and span a range of fitnesses both below and above the fitness of the DivAnc.

To this end, we inoculated six 96-well plates with saturated culture of DivAnc to start 72 populations per plate for a total of 432 populations. We propagated these populations in batch culture in YPD. Half of populations were propagated with a small bottleneck size ( $1 : 2^{15}$  dilution every 36 h) and half were propagated with a large bottleneck size ( $1 : 2^5$  dilution every 12 h) for a total of 240 generations. Populations were frozen in 20% glycerol (v/v) every 60 generations and stored at  $-80^{\circ}\text{C}$ .

From each of the evolved populations we isolated between 0 and 4 clones (3.4 on average) for a total of 1487 clones and measured their fitnesses relative to RmRef at the large bottleneck size without replication. The median fitness of clones evolved under the small bottleneck (S-clones) size was  $-0.45\%$  (5- and 95-percentiles of the distribution of fitnesses  $[-3.2\%, 0.49\%]$ ), while the median fitness of clones evolved under the large bottleneck size (L-clones) was  $0.14\%$  (5- and 95-percentiles  $[-2.08\%, 2.11\%]$ ). From this collection of clones, we selected 128 S-clones, each descended from a distinct population, that spanned fitnesses between  $-4.9\%$  and  $1.0\%$  (median  $-0.41\%$ ) and 128 L-clones, each descended from a distinct population, that spanned fitnesses between  $-3.1\%$  and  $3.4\%$  (median  $0.14\%$ ) and measured their fitnesses again with two-fold replication.

We made the final selection of 32 S-clones and 32 L-clones such that they jointly spanned the range of fitnesses between  $-2.3\%$  and  $3.0\%$  (median  $-0.67\%$ ) relative to RmRef. These 64 clones were the original Founders which were subsequently used to found populations for the Adaptation phase (see section Adaptation). However, after sequencing full genomes of descendants of 13 original Founders, we discovered that the frozen stocks of two of the Founders (originally named L096 and L102) were in fact mixtures of two genotypes each (see section Adaptation for possible explanations). Two genotypes isolated from the L096 stock were named L096a and L096b and two genotypes isolated from the L102 stock were named L102 and L102a. Thus, our library of Founders consists of 66 strains.

**Adaptation** Because we were interested in how the genotype influences the rate of adaptation and the degree of convergence, it was essential to exclude the possibility that replicate populations founded with the same Founder strain shared the same standing genetic variation (SGV). SGV can be present in the frozen stock of the Founder strain due to three reasons: (1) Cross-contamination events (which are most likely when colonies are transferred from an agar plate to a 96-well plate); (2) Non-isogenic origin of a colony (e.g., when two genetically distinct cells stick to each other (4)); (3) New mutations that arise in the culture during clonal expansion in the colony on an agar plate and subsequently in liquid before the culture is frozen. To ensure that replicate populations founded by the same Founder do not share any SGV, each population was started with a clone isolated from the Founder stock. In this way, we founded 10 replicate populations from each of the original Founders resulting in 640 populations distributed across 8 96-well plates.

Replicate populations descended from the same Founder occupied a row in a 96-well plate during the subsequent evolution experiment. Thus, each replicate population had an id number between 1 and 12. As mentioned in section Diversification and selection of Founders, the stocks of original Founders L096 and L102 at the time of inoculation consisted of two distinct genotypes each. Thus, we changed the ids of populations as follows: L096-2, -4, -5, -7, -10, -12 became L096a-2, -4, -5, -7, -10, -12; populations L096-1, -6, -8, -11 became L096a-1, -6, -8, -11; populations L102-2, -3, -5, -6 became L102a-2, -3, -5, -6; populations L102-1, -7, -9, -10, -11, -12 retained their original ids.

In the Adaptation phase, populations were propagated with a large bottleneck size ( $1 : 2^5$  dilution every 12 h) for a total of 500 generations. Populations were frozen in 20% glycerol (v/v) every 50 generations and stored at  $-80^{\circ}\text{C}$ .

**Fitness assays** We measured the fitnesses of our strains and populations using flow cytometry-based competition assays (5). We competed our strains either against the mCherry-labelled RmRef strain or against the ymCitrine-labelled DivAncCit strain (see section Strains and Table S1). Before each fitness assay, 96-well plates with fresh YPD were inoculated from frozen stocks of query and reference strains and acclimated separately for between 24 h and 48 h by serial passaging at  $1 : 2^5$  dilution every 24 h. Thereafter each query strain was mixed with the reference strain in proportion  $1 : 1$  by volume. Mixed cultures were propagated with  $1 : 2^5$  dilution every 12 h for 72 hours, and relative frequencies of query and reference strains were measured using flow cytometry 24 and 72 hours after mixing, which corresponds to 10 and 30 generations of competition. Fitness of the query strain relative to the reference was then calculated as

$$F = \frac{1}{t_2 - t_1} \log \left( \frac{n(t_2)}{n_r(t_2)} \bigg/ \frac{n(t_1)}{n_r(t_1)} \right),$$

where  $n(t)$  and  $n_r(t)$  are the cell counts for the evolved and the reference strains at generation  $t$  after mixing, respectively. In our measurement,  $t_1 = 10$  and  $t_2 = 30$  generations. A detailed fitness assay protocol is available at <http://sergeykryazhimskiy.webs.com/protocols>.

The fitness of RmRef was +1.1% relative to DivAncCit. To rule out non-transitivity, fitnesses of several strains were measured against both references and good correlation was found (Figure S8). Thus, all fitness values are reported with respect to DivAncCit except where especially noted otherwise.

**Replay experiments** We selected four pairs of Founders such that both Founders within a pair had very similar fitnesses but showed large differences in their rate of adaptation during the first 250 generations of evolution in the Adaptation phase (Table S3). We isolated 20 clones from the stocks of each of these Founders and used these clones to found replicate populations, for a total of 160 populations. We evolved these populations for 250 generations under conditions identical to the Adaptation phase and measured the fitnesses of these populations at generations 0 and 250 in duplicate. In 9 out of 160 populations, our fitness measurements were inaccurate (standard deviation among replicate

measurements exceeded 3%), and we excluded such populations from further analysis. The replay experiment confirmed a systematic difference in adaptability among Founders in one out of four pairs (Pair 3, see Table S3).

## Full-genome sequencing

We isolated 1 or 2 clones from generation 500 of the Adaptation phase from each replicate population descended from Founders L003, L013, L034, L041, L048, L094, L096a, L096b, L098, L102, L102a, L125, S002, S028, S121 for a total of 167 clones. Although 167 clones were sequenced, with some populations represented by a pair of clones, all analyses presented in the text were restricted to 128 clones, each from a unique population.

Genomic DNA for subsequent full-genome sequencing was extracted from DivAnc and from each of the 167 clones using PureLink®*Pro* 96 Genomic Purification Kit (Life Technologies catalog # K1821-04A) and quantified using the Qubit platform. Multiplexed sequencing library for the Illumina platform was prepared using the Nextera kit (Illumina catalog # FC-121-1031 and # FC-121-1012) and a modified version of the Illumina-recommended protocol available at <http://sergeykryazhimskiy.webs.com/protocols>. Briefly, the tagmentation reaction was carried out in a 2.5 µl volume, with reaction components added at the Illumina-recommended stoichiometry. In order to attach the standard Illumina barcodes and grafting primers and to amplify the library, 12 cycles of PCR were performed directly on the tagmentation reaction product, using the Kappa Library Amplification kit (Kappa catalog # KK2612). PCR products were purified using AmPure beads (Beckman Coulter catalog # A63881). Purified bar-coded libraries were quantified and quality controlled on the Qubit and Bioanalyzer platforms and pooled in equimolar concentrations. Paired-end 100 bp or 150 bp sequencing was performed on the pooled libraries at the Harvard FAS core facility on the Illumina HiSeq 2000 or 2500 machines. A total of three sequencing runs were performed. Because sequencing runs #2 and #3 were performed on the same pooled library, the reads from these two runs were pooled for subsequent analysis.

## Knock-out experiments

We selected genes *GAT2*, *WHI2*, and *SFL1* for the knock-out experiment after a small-scale pilot which showed that knock outs of these three genes gave a selective advantage under our conditions. In addition, we knocked out gene *ACE2*, which is the most frequently hit gene in our experiment (see Table S6). However, we were unable to accurately measure the fitness of the *ace2*Δ knock-out strains due to cell aggregation phenotype that *ace2*Δ confers. Thus, the *ace2*Δ strains were excluded from analysis. We also knocked out gene *HO* as a negative control because this gene already carries two loss-of-function mutations in the W303 background (6). These five genes were knocked out in a total of 18 backgrounds: 13 Founders, DivAnc and four evolved clones (see Table S1).

To knock out these genes we amplified the *KanMX* cassette with the gene-specific flanking regions from the appropriate yeast haploid knock-out collection strains (Table S8). Transformation was carried out using a standard yeast genetics protocol (3).

Because all our strains carry the *NatMX* cassette, we selected the transformants on double-drug plates, G418 488 mg/l (Sigma #A1720-5g) and CloNAT 102 mg/l (Gold-bio #N-500-1), to prevent the integration of the *KanMX* cassette into the *NatMX* locus. We picked between 1 and 10 transformant colonies and confirmed the integration of the *KanMX* cassette into the correct locus using the yeast knock-out collection protocol (see [http://www-sequence.stanford.edu/group/yeast\\_deletion\\_project/protocols.html](http://www-sequence.stanford.edu/group/yeast_deletion_project/protocols.html) and Table S8). Specifically, we retained only those transformants for which both negative control PCRs (with primers *AB* and *CD*) were negative and at least one of the positive control PCRs (with primer pairs *A-KanB* and *KanC-D*) was positive. Thus, we obtained between 1 and 9 (mean 4.4) confirmed transformants for each gene in each genetic background, for a total of 399 strains. Fitnesses of these strains were measured using the fitness assay protocol described above. 30 knock-out strains whose fitnesses were much lower than the fitnesses of the other identical transformants (which we took as an indication of a transformation artifacts in these outlier clones) were excluded from further analysis (Table S1).

## Analysis

### Identification of mutations from Illumina reads

To call mutations, we first used bowtie2 (7) to align de-multiplexed Illumina reads to a W303 reference genome (8). Next, for each barcode, we used GATK software (9) with liberal stringency settings to narrow down the set of variable sites and obtain read depth, reference and alternate allele counts at each such site. We discarded sites at which GATK called multiple alternate alleles because they are certain to be due to sequencing or alignment errors. Most of the variable sites at this stage were still caused by errors, and we observed that the probability that a site is variable varies dramatically across the genome indicating that errors occur at different sites with different probabilities. We modeled the probability distribution of  $\Theta$ , the probability for an error to occur at a given site, by a Beta distribution, i.e.,  $\Theta \sim \text{Beta}(\alpha, \beta)$ . We fit the distribution across all libraries using the MATLAB package *fastfit* and obtained  $\hat{\alpha} = 1 \times 10^{-4}$ ,  $\hat{\beta} = 0.271$  for sequencing run #1 and  $\hat{\alpha} = 1 \times 10^{-4}$ ,  $\hat{\beta} = 0.434$  for the pooled sequencing runs #2 and #3. Next, we used the Bayesian method for calling mutation described below to calculate the posterior probability that each candidate mutation is an error or is a real mutation that occurs in a particular subset of clones (e.g., in both clones that came from the same population, in all clones descended from the same Founder, etc.). We exploited here the hierarchical design of our experiment and the prior knowledge that convergent evolution in yeast in YPD is rare (8). We next discarded all candidate mutations for which the error hypothesis had the highest posterior probability. Having obtained this almost final list of mutations, we discarded those mutations that had a high ( $> 5\%$ ) frequency outside of the clones in which they were called. Finally, we identified “complex” mutations, i.e., mutations at nearby sites that were called in the same clone. Specifically, we clustered mutations that occurred in the same clone using hierarchical clustering (MATLAB functions *linkage*

and *cluster*) with complete linkage, distance as the clustering criterion, and taking 300 bp as the threshold distance (see Figure S9). Mutations within the same cluster were considered to have arisen from the same complex mutational event and were assigned the same Unique Mutation Id. The final list of mutations is presented in Table S4. We assigned mutation types by mapping the positions of mutations on the annotated W303 genome (8). We assumed that mutations within 500 bp upstream of an ORF affect its expression level and we assign them to category “promoter”.

**Bayesian framework for calling mutations** To identify real mutations among sequencing and alignment errors from the output of GATK, we developed the following Bayesian inference method. This method exploits our hierarchical experimental design and the fact that in our experiment the probability of two mutations occurring in the same site in independent populations is extremely small (8). This method allows us to reliably call mutations in data with about 15x average coverage.

Since our data comes from a multiplexed sequencing run, the input into our method is the list of sites that are polymorphic in each barcoded library, together with the respective reference ( $R$ ) and the alternate ( $A$ ) alleles at each of these sites. To identify which of these polymorphic sites are a result of errors versus real mutations, consider first a single site. Let  $a_i$  be the number of reads supporting the alternate allele  $A$  in library  $i$  at this site, and let  $k_i$  be the number of reads at the site that support either  $A$  or  $R$  (all remaining reads, if any, are discarded). Let  $m_i$  be the indicator variable for the event that a real mutation is present at the site in library  $i$ , i.e.,

$$m_i = \begin{cases} 1, & \text{if mutation present in library } i, \\ 0, & \text{otherwise.} \end{cases}$$

We call  $\mathbf{m} = (m_1, m_2, \dots, m_N) \in \{0, 1\}^N$ , where  $N$  is the number of libraries, a configuration of data for the focal site and the given alternate allele. In other words,  $\mathbf{m}$  represents the event that a real mutation is present in a particular subset of clones and is not present in the remaining clones.

We are interested in the posterior probability  $P_{\text{post}}(\mathbf{m}|\mathbf{a}, \mathbf{k})$  of any configuration  $\mathbf{m}$ , given the coverage vector  $\mathbf{k} = (k_1, \dots, k_N)$  and the vector of the number of reads supporting  $A$ ,  $\mathbf{a} = (a_1, \dots, a_N)$ , in each library. By Bayes’ theorem,

$$P_{\text{post}}(\mathbf{m}|\mathbf{a}, \mathbf{k}) = \frac{P_{\text{prior}}(\mathbf{m}) L(\mathbf{a}|\mathbf{m}, \mathbf{k})}{\sum_{\mathbf{m}} P_{\text{prior}}(\mathbf{m}) L(\mathbf{a}|\mathbf{m}, \mathbf{k})}, \quad (\text{S1})$$

where  $L(\mathbf{a}|\mathbf{m}, \mathbf{k})$  is the likelihood of observing a particular number of reads supporting  $A$  in each library given the configuration of data and the coverage vector, and  $P_{\text{prior}}(\mathbf{m})$  is the prior probability of configuration  $\mathbf{m}$ .

Let  $\Theta$  be the probability that allele  $A$  was created at the focal site by error. If  $\Theta$  has

density  $f_{\Theta}(\theta)$ , we can rewrite expression (S1) as

$$P_{\text{post}}(\mathbf{m}|\mathbf{a}, \mathbf{k}) = \frac{P_{\text{prior}}(\mathbf{m}) \int_0^1 L(\mathbf{a}|\mathbf{m}, \mathbf{k}, \theta) f_{\Theta}(\theta) d\theta}{\sum_{\mathbf{m}} P_{\text{prior}}(\mathbf{m}) \int_0^1 L(\mathbf{a}|\mathbf{m}, \mathbf{k}, \theta) f_{\Theta}(\theta) d\theta}. \quad (\text{S2})$$

To obtain the likelihood  $L(\mathbf{a}|\mathbf{m}, \mathbf{k}, \theta)$  of data  $\mathbf{a}$  for a given configuration vector  $\mathbf{m}$ , coverage vector  $\mathbf{k}$  and error probability  $\theta$ , we consider all libraries to be independent and model the number of alternate alleles in each library as a binomial random variable. Specifically, if  $m_i = 0$  in library  $i$ , the alternate alleles occur in this library with the error probability  $\theta$ ; if  $m_i = 1$ , the alternate alleles occur with probability  $1 - \theta$ . Therefore,

$$\begin{aligned} L(\mathbf{a}|\mathbf{m}, \mathbf{k}, \theta) &= \prod_{i=1}^N P(a_i|m_i, \theta, k_i) \\ &= \prod_{i=1}^N \binom{k_i}{a_i} \theta^{m_i(k_i - a_i) + (1 - m_i)a_i} (1 - \theta)^{m_i a_i + (1 - m_i)(k_i - a_i)} \\ &= \theta^{\sum_i [m_i(k_i - a_i) + (1 - m_i)a_i]} (1 - \theta)^{\sum_i [m_i a_i + (1 - m_i)(k_i - a_i)]} \prod_{i=1}^N \binom{k_i}{a_i}. \end{aligned} \quad (\text{S3})$$

Finally, we assume that  $f_{\Theta}$  is a beta distribution with parameters  $\hat{\alpha}$  and  $\hat{\beta}$ , which we fit to the empirical distribution of error rates across the genome over all libraries, i.e.,

$$f_{\Theta}(\theta) = \frac{\theta^{\hat{\alpha}-1} (1 - \theta)^{\hat{\beta}-1}}{B(\hat{\alpha}, \hat{\beta})}, \quad (\text{S4})$$

where  $B(\hat{\alpha}, \hat{\beta})$  is a beta function. Substituting (S3) and (S4) into (S2) and simplifying yields

$$\begin{aligned} P_{\text{post}}(\mathbf{m}|\mathbf{a}, \mathbf{k}) &= \frac{P_{\text{prior}}(\mathbf{m}) \int_0^1 \theta^{\alpha(\mathbf{m}, \mathbf{k})-1} (1 - \theta)^{\beta(\mathbf{m}, \mathbf{k})-1} d\theta}{\sum_{\mathbf{m}} P_{\text{prior}}(\mathbf{m}) \int_0^1 \theta^{\alpha(\mathbf{m}, \mathbf{k})-1} (1 - \theta)^{\beta(\mathbf{m}, \mathbf{k})-1} d\theta} \\ &= \frac{P_{\text{prior}}(\mathbf{m}) B(\alpha(\mathbf{m}, \mathbf{k}), \beta(\mathbf{m}, \mathbf{k}))}{\sum_{\mathbf{m}} P_{\text{prior}}(\mathbf{m}) B(\alpha(\mathbf{m}, \mathbf{k}), \beta(\mathbf{m}, \mathbf{k}))}, \end{aligned} \quad (\text{S5})$$

where

$$\begin{aligned} \alpha(\mathbf{m}, \mathbf{k}) &= \hat{\alpha} + \sum_{i=1}^N [m_i(k_i - a_i) + (1 - m_i)a_i], \\ \beta(\mathbf{m}, \mathbf{k}) &= \hat{\beta} + \sum_{i=1}^N [m_i a_i + (1 - m_i)(k_i - a_i)]. \end{aligned}$$



We used the structure of our experiment to specify the prior probability of each configuration  $\mathbf{m}$ . We set the probability that multiple clones share the same mutation by descent to be  $10^{-6}$ . In other words,  $P_{\text{prior}}(\mathbf{m}) = 10^{-6}$  if  $\mathbf{m}$  corresponds to one of the four configuration types: mutation is present in a single clone, mutation is present in multiple clones from the same population, mutation is present in all descendants from a single Founder, and mutation is present in all clones. To account for the possibility of rare convergent mutations, we assign  $P_{\text{prior}}(\mathbf{m}) = 10^{-12}$  to all configurations  $\mathbf{m}$  that correspond to the events that a mutation is present in two clones from different populations. We assign a zero prior probability to all other configurations, except for the configuration  $\mathbf{m}_0 = (0, 0, \dots, 0)$  which corresponds to a sequencing error. Finally, we set the prior probability of a sequencing error,  $P_{\text{prior}}(\mathbf{m}_0)$ , so that  $\sum_{\mathbf{m}} P_{\text{prior}}(\mathbf{m}) = 1$ .

We use expression (S5) to compute the posterior probability of each configuration  $\mathbf{m}$  that has a non-zero prior probability. We obtain the list of real mutations by discarding all candidate mutations for which the error configuration  $\mathbf{m}_0$  has the highest posterior probability compared to other configurations.

### Analysis of covariance of fitness increment with initial fitness

For each of the time point  $t = 0, 250$ , and  $500$  generations, we measured the fitnesses of all 640 populations. We denote by  $X_{ijk}^{(t)}$  the fitness measurement  $k$  in population  $j$  descended from Founder  $i$  at time  $t$ , where  $i = 1, 2, \dots, n_F = 66$ ,  $j = 1, 2, \dots, n_{\text{pop}}(i)$ , and  $k = 1, 2, \dots, n_{\text{rep}}(i, j)$ . Since for most Founders  $n_{\text{pop}}(i) = 10$  and for most populations  $n_{\text{rep}}(i, j) = 2$ , our design is close to being balanced, but in the following we will consider a general unbalanced design. Since all populations descended from Founder  $i$  are isogenic at time  $t = 0$ ,

$$x_i = \frac{1}{\sum_{j=1}^{n_F(i)} n_{\text{rep}}(i, j)} \sum_{j=1}^{n_F(i)} \sum_{k=1}^{n_{\text{rep}}(i, j)} X_{ijk}^{(0)}$$

is the estimate of the fitness of Founder  $i$ . We denote measurement  $k$  of the fitness increment of population  $j$  descended from Founder  $i$  at time  $t$  as

$$Y_{ijk}^{(t)} = X_{ijk}^{(t)} - x_i.$$

We will omit superscript  $t$  because we will treat the fitness increments at generations 250 and 500 separately.

In order to partition the variance in fitness increments  $Y$  into four components, measurement error, intrinsic evolutionary stochasticity, Founder fitness, and Founder genotype, we consider the following family of six nested statistical models.

**Model 1A. Measurement error.** This is the simplest null model, which supposes that all observed fitness increments are due to measurement error, i.e.,

$$Y_{ijk} = \alpha + \varepsilon_{ijk}, \tag{S6}$$

where  $\alpha$  is the expected fitness increment of a population after evolution and  $\varepsilon_{ijk}$  are independent random variables distributed normally with mean zero and variance  $\sigma_n^2$  representing measurement error.

**Model 1B. Founder fitness and Measurement error.** This model supposes that the fitness increment of a population is correlated with the fitness of its Founder, i.e.,

$$Y_{ijk} = \alpha + \beta x_i + \varepsilon_{ijk}, \quad (\text{S7})$$

where  $\alpha + \beta x_i$  is the expected fitness increment of a population descended from Founder  $i$ . We assume here and below that  $x_i$  are error-free, i.e., these are not random variables.

**Model 2A. Evolutionary stochasticity and Measurement error.** This model supposes that, in addition to Measurement error, different population achieve different fitness increments due to the intrinsic stochasticity of the evolutionary process, but that there are no systematic differences among populations, i.e.,

$$Y_{ijk} = \alpha + A_{ij} + \varepsilon_{ijk}, \quad (\text{S8})$$

where  $\alpha + A_{ij}$  is the expected fitness increment of population  $j$  descended from Founder  $i$ ; all  $A_{ij}$  are assumed to be independent and distributed normally with mean zero and variance  $\sigma_c^2$ .

**Model 2B. Founder fitness, Evolutionary stochasticity and Measurement error.** This model supposes that the fitness increment of a population is correlated with the fitness of its Founder, and that different populations descended from the same Founder achieve different fitness increments due to the evolutionary stochasticity, but that there are no systematic differences among such populations, i.e.,

$$Y_{ijk} = \alpha + \beta x_i + A_{ij} + \varepsilon_{ijk}. \quad (\text{S9})$$

**Model 3A. Founder, Evolutionary stochasticity and Measurement error.** This model supposes that, in addition to Evolutionary stochasticity, population descended from different Founders achieve systematically different fitness increments, i.e.,

$$Y_{ijk} = \alpha + B_i + A_{ij} + \varepsilon_{ijk}, \quad (\text{S10})$$

where  $\alpha + B_i$  is the expected fitness increment of population descended from Founder  $i$ ;  $B_i$  are assumed to be independent and distributed normally with mean zero and variance  $\sigma_g^2$ . This model is analogous to that considered by Travisano et al (10).

**Model 3B. Founder fitness, Founder genotype, Evolutionary stochasticity and Measurement error.** This model supposes that, in addition to Evolutionary stochasticity, the fitness increment of a population is determined by the fitness of its Founder as well as by the Founder genotype so that descendants of different Founders with the same fitness achieve systematically different fitness increments, i.e.,

$$Y_{ijk} = \alpha + \beta x_i + B_i + A_{ij} + \varepsilon_{ijk}, \quad (\text{S11})$$

where  $\alpha + \beta x_i + B_i$  is the expected fitness increment of descendants of Founder  $i$ .

Our goal is to (a) identify which of the above models gives the best fit to the data, and (b) determine how much variance in final fitness is explained by different factors. We fit the models using the maximum likelihood approach and compare the fits using the likelihood ratio test.

In the next sections, we derive the expressions for the likelihood functions for each of the models, after introducing useful notations and recalling some elementary properties of the normal distribution.

**Notations** Our data has three levels of nestedness, such that each data point is associated at the bottom level with a particular measurement, at the middle level it is associated with a particular population, and at the top level it is associated with a particular Founder. To reflect this nesting, we will denote vectors that correspond to the full data, i.e., vectors of size  $N = \sum_{i=1}^{n_F} \sum_{j=1}^{n_{\text{pop}}(i)} n_{\text{rep}}(i, j)$ , by three subscript dots, e.g.,

$$x_{\dots} = \left( x_{111}, x_{112}, \dots, x_{n_F n_{\text{pop}}(n_F) n_{\text{rep}}(n_F, n_{\text{pop}}(n_F))} \right)^T,$$

where the dot in the first position (closest to the  $x$  symbol) corresponds to the top level of nesting, the dot in the second position (middle) corresponds to the middle level of nesting, and the dot in the third position (farthest from the  $x$  symbol) corresponds to the bottom level of nesting. Continuing with this logic, symbols with one or two subscript dots will denote lower-dimensional vectors that correspond to one or two out of three levels of nesting. Thus,  $x_{\cdot}$  (dot in the first position) denotes a vector of size  $n_F$ , i.e., a vector of values corresponding to Founders;  $x_{i\cdot}$  (dot in the second position) denotes a vector of size  $n_{\text{pop}}(i)$  that corresponds to all populations descended from Founder  $i$ ; and  $x_{ij\cdot}$  (dot in the third position) denotes a vector of size  $n_{\text{pop}}(i, j)$  which corresponds all measurements of population  $j$  descended from Founder  $i$ . Analogously,  $x_{i\cdot\cdot}$  (dot in the second and third positions) denotes a vector of size  $\sum_{j=1}^{n_{\text{pop}}(i)} n_{\text{rep}}(i, j)$ , i.e., a vector of values corresponding to all measurements in all populations descended from Founder  $i$ . We have

$$\begin{aligned} x_{\cdot} &= (x_1, x_2, \dots, x_{n_F})^T, \\ x_{i\cdot} &= (x_{i1}, x_{i2}, \dots, x_{i n_{\text{pop}}(i)})^T, \\ x_{ij\cdot} &= (x_{ij1}, x_{ij2}, \dots, x_{ij n_{\text{rep}}(i, j)})^T, \\ x_{i\cdot\cdot} &= (x_{i11}, x_{i12}, \dots, x_{i n_{\text{pop}}(i) n_{\text{rep}}(i, n_{\text{pop}}(i))})^T. \end{aligned}$$

We introduce the following notations for the empirical means.

$$\begin{aligned}
M[x...] &= \frac{1}{N} \sum_{i=1}^{n_F} \sum_{j=1}^{n_{\text{pop}}(i)} \sum_{k=1}^{n_{\text{rep}}(i,j)} x_{ijk} \quad \text{is the grand mean,} \\
M_i[x...] &= \frac{1}{\sum_{j=1}^{n_{\text{pop}}(i)} n_{\text{rep}}(i,j)} \sum_{j=1}^{n_{\text{pop}}(i)} \sum_{k=1}^{n_{\text{rep}}(i,j)} x_{ijk} \quad \text{is the mean for Founder } i, \\
M_{ij}[x...] &= \frac{1}{n_{\text{rep}}(i,j)} \sum_{k=1}^{n_{\text{rep}}(i,j)} x_{ijk} \\
&\quad \text{is the mean for population } j \text{ descended from Founder } i, \\
M.[x...] &= (M_1[x...], M_2[x...], \dots, M_{n_F}[x...])^T \\
&\quad \text{is a vector of Founder means,} \\
M..[x...] &= (M_{11}[x...], M_{12}[x...], \dots, M_{n_F n_{\text{pop}}(n_F)}[x...])^T \\
&\quad \text{is a vector of population means,} \\
M[M.[x...]] &= \frac{1}{n_F} \sum_{i=1}^{n_F} M_i[x...] \quad \text{is the mean of Founder means,} \\
M[M..[x...]] &= \frac{1}{\sum_{i=1}^{n_F} n_{\text{pop}}(i)} \sum_{i=1}^{n_F} \sum_{j=1}^{n_{\text{pop}}(i)} M_{ij}[x...] \quad \text{is the mean of population means.}
\end{aligned}$$

Note that the grand mean, the mean Founder means, and the mean of population means are all different in an unbalanced design, but identical in a balanced design. The notations for the empirical variances are as follows.

$$\begin{aligned}
V[x...] &= M[(x... - M[x...])^2] \quad \text{is the grand variance,} \\
V_i[x...] &= M_i[(x... - M_i[x...])^2] \quad \text{is the variance within Founder } i, \\
V_{ij}[x...] &= M_{ij}[(x... - M_{ij}[x...])^2] \\
&\quad \text{is the variance within population } j \text{ descended from Founder } i, \\
V.[x...] &= (V_1[x...], V_2[x...], \dots, V_{n_F}[x...])^T \\
&\quad \text{is a vector of within-Founder variances} \\
V..[x...] &= (V_{11}[x...], V_{12}[x...], \dots, V_{n_F n_{\text{pop}}(n_F)}[x...])^T \\
&\quad \text{is a vector of within-population variances} \\
M[V.[x...]] &= \frac{1}{n_F} \sum_{i=1}^{n_F} V_i[x...] \quad \text{is the mean within-Founder variance} \\
M[V..[x...]] &= \frac{1}{\sum_{i=1}^{n_F} n_{\text{pop}}(i)} \sum_{i=1}^{n_F} \sum_{j=1}^{n_{\text{pop}}(i)} V_{ij}[x...] \quad \text{is the mean within-population variance.}
\end{aligned}$$

We denote the empirical covariance as

$$C[x., y.] = M[(x. - M[x.]) (y. - M[y.])].$$

Where there is no ambiguity, we will omit the dots next to the symbol denoting the variable and write  $M[x]$ ,  $V[x]$ ,  $M_i[x]$ ,  $V_i[x]$ , etc.

**Useful properties of the normal distribution** We denote the normal probability density function with mean  $m$  and variance  $\sigma^2$  taken at  $x$  by  $P_N(x; m, \sigma^2)$ . The property of the normal distribution that we will need is

$$\prod_{k=1}^n P_N(x; \mu_k, \sigma_k^2) = A P_N(x; \tilde{\mu}, \tilde{\sigma}^2), \quad (\text{S12})$$

where

$$\begin{aligned} A &= (2\pi)^{-\frac{n-1}{2}} \left( \frac{\tilde{\sigma}^2}{\prod_k \sigma_k^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left( \sum_k \frac{\mu_k^2}{\sigma_k^2} - \frac{\tilde{\mu}^2}{\tilde{\sigma}^2} \right) \right\}, \\ \tilde{\sigma}^2 &= \left( \sum_k \frac{1}{\sigma_k^2} \right)^{-1}, \\ \tilde{\mu} &= \tilde{\sigma}^2 \sum_k \frac{\mu_k}{\sigma_k^2}. \end{aligned}$$

In the special case when  $\sigma_k^2 = \sigma$ ,  $k = 1, \dots, n$ , equation (S12) becomes

$$\prod_k P_N(x; \mu_k, \sigma^2) = A P_N(x; M[\mu.], \sigma^2/n),$$

where

$$A = (2\pi\sigma^2)^{-\frac{n-1}{2}} n^{-\frac{1}{2}} \exp \left\{ -\frac{V[\mu.]}{2\sigma^2/n} \right\}.$$

In the special case when  $n = 2$ , we have

$$P_N(y; \mu_1, \sigma_1^2) P_N(y; \mu_2, \sigma_2^2) = P_N(\mu_1; \mu_2, \sigma_1^2 + \sigma_2^2) P_N(y; \mu_3, \sigma_3^2), \quad (\text{S13})$$

where

$$\begin{aligned} \sigma_3 &= \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \\ \mu_3 &= \sigma_3 \left( \frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2} \right). \end{aligned}$$

It follows from equation (S13) that

$$\int_{-\infty}^{\infty} P_N(y; \mu_1, \sigma_1^2) P_N(y; \mu_2, \sigma_2^2) dy = P_N(\mu_1; \mu_2, \sigma_1^2 + \sigma_2^2). \quad (\text{S14})$$

**Model 1A. Measurement error** This model, described by equation (S6), has two parameters,  $\alpha$  and  $\sigma_n^2$ . The likelihood of the data is

$$L_{1A}(Y; \alpha, \sigma_n^2) = \prod_{i=1}^{n_{FG}} \prod_{j=1}^{n_{pop}(i)} \prod_{k=1}^{n_{rep}(i,j)} P_N(Y_{ijk}; \alpha, \sigma_n^2). \quad (S15)$$

Under the balanced design, the maximum likelihood parameter estimates are

$$\begin{aligned} \hat{\alpha} &= M[Y], \\ \hat{\sigma}_n^2 &= V[Y]. \end{aligned}$$

**Model 1B. Fitness and Measurement error** This model, described by equation (S7), has three parameters,  $\alpha$ ,  $\beta$ , and  $\sigma_n^2$ . The likelihood of data is given by

$$L_{1B}(Y; \alpha, \beta, \sigma_n^2) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{pop}(i)} \prod_{k=1}^{n_{rep}(i,j)} P_N(Y_{ijk}; \alpha + \beta x_i, \sigma_n^2), \quad (S16)$$

If the design is balanced, we can obtain the expressions for the ML parameter values as follows. We minimize function

$$\ell_{1B} = -2 \log L_{1B} = \sum_{ijk} \frac{(Y_{ijk} - \alpha - \beta x_i)^2}{\sigma_n^2} + N \log \sigma_n^2 + N \log(2\pi)$$

by setting its partial derivatives to zero

$$\begin{aligned} \frac{\partial \ell_{1B}}{\partial \alpha} &= -\frac{2}{\sigma_n^2} \sum_{ijk} (Y_{ijk} - \alpha - \beta x_i) = 0, \\ \frac{\partial \ell_{1B}}{\partial \beta} &= -\frac{2}{\sigma_n^2} \sum_{ijk} (Y_{ijk} - \alpha - \beta x_i) x_i = 0, \\ \frac{\partial \ell_{1B}}{\partial \sigma_n^2} &= -\frac{1}{\sigma_n^4} \sum_{ijk} (Y_{ijk} - \alpha - \beta x_i)^2 + \frac{N}{\sigma_n^2} = 0, \end{aligned}$$

where we used the shorthand

$$\sum_{ijk} \equiv \sum_{i=1}^{n_F} \sum_{j=1}^{n_{pop}} \sum_{k=1}^{n_{rep}}.$$

Thus, we obtain the following equations for ML parameter values  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{\sigma}_n^2$ .

$$\begin{aligned} 0 &= M[Y] - \hat{\alpha} - \hat{\beta} M[x], \\ 0 &= n_{pop} n_{rep} \sum_i M_i[Y] x_i - N \hat{\alpha} M[x] - N \hat{\beta} M[x^2] = 0, \\ 0 &= -\sum_{ijk} (Y_{ijk} - \hat{\alpha} - \hat{\beta} x_i)^2 + N \hat{\sigma}_n^2 = 0, \end{aligned}$$

which, together with the fact that, under the balanced design,  $M[Y] = M[M.[Y]]$ , lead to

$$\begin{aligned}\hat{\beta} &= \frac{C[M.[Y], x]}{M[x^2] - (M[x])^2}, \\ \hat{\alpha} &= M[Y] - \hat{\beta}M[x], \\ \hat{\sigma}_n^2 &= \frac{1}{N} \sum_{ijk} \left( Y_{ijk} - \hat{\alpha} - \hat{\beta}x_i \right)^2.\end{aligned}$$

**Model 2A. Evolutionary stochasticity and Measurement error** This model, described by equation (S8), has three parameters,  $\alpha$ ,  $\sigma_c^2$ , and  $\sigma_n^2$ . The likelihood of data is given by

$$\begin{aligned}L_{2A}(Y; \alpha, \sigma_c^2, \sigma_n^2) &= \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} \mathbb{P}(Y_{ij\cdot}) \\ &= \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} \int_{-\infty}^{\infty} \mathbb{P}(A = a_{ij}) \mathbb{P}(Y_{ij\cdot} | A = a_{ij}) da_{ij} \\ &= \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} \int_{-\infty}^{\infty} \mathbb{P}(A = a_{ij}) \prod_{k=1}^{n_{\text{rep}}(i,j)} \mathbb{P}(Y_{ijk} | A = a_{ij}) da_{ij} \\ &= \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} \int_{-\infty}^{\infty} P_N(a_{ij}; 0, \sigma_c^2) \prod_{k=1}^{n_{\text{rep}}(i,j)} P_N(Y_{ijk}; \alpha + a_{ij}, \sigma_n^2) da_{ij},\end{aligned}\quad (\text{S17})$$

where we used the fact that replicate measurements of the same population are independent conditional on the fitness of the population. Using equations (S12) and (S14), we obtain

$$\begin{aligned}\prod_{k=1}^{n_{\text{rep}}(i,j)} P_N(Y_{ijk}; \alpha + a_{ij}, \sigma_n^2) &= \prod_{k=1}^{n_{\text{rep}}(i,j)} P_N(a_{ij}; Y_{ijk} - \alpha, \sigma_n^2) \\ &= A_{\text{rep}}(i, j) P_N(a_{ij}; M_{ij}[Y] - \alpha, \sigma_n^2/n_{\text{rep}}(i, j))\end{aligned}$$

and

$$\int_{-\infty}^{\infty} P_N(a_{ij}; 0, \sigma_c^2) P_N(a_{ij}; M_{ij}[Y] - \alpha, \sigma_n^2/n_{\text{rep}}(i, j)) da_{ij} = P_N(M_{ij}[Y]; \alpha, \tilde{\sigma}_{ij}^2),$$

where

$$A_{\text{rep}}(i, j) = (2\pi\sigma_n^2)^{-\frac{n_{\text{rep}}(i,j)-1}{2}} (n_{\text{rep}}(i, j))^{-\frac{1}{2}} \exp \left\{ -\frac{V_{ij}[Y]}{2\sigma_n^2/n_{\text{rep}}(i, j)} \right\}, \quad (\text{S18})$$

$$\tilde{\sigma}_{ij}^2 = \sigma_n^2/n_{\text{rep}}(i, j) + \sigma_c^2. \quad (\text{S19})$$

Thus, equation (S17) becomes

$$L_{2A}(Y; \alpha, \sigma_c^2, \sigma_n^2) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} A_{\text{rep}}(i, j) P_N(M_{ij}[Y]; \alpha, \tilde{\sigma}_{ij}^2). \quad (\text{S20})$$

In case of a balanced design equations (S18), (S19) simplify to

$$\begin{aligned} A_{\text{rep}}(i, j) &= (2\pi\sigma_n^2)^{-\frac{n_{\text{rep}}-1}{2}} n_{\text{rep}}^{-\frac{1}{2}} \exp \left\{ -\frac{V_{ij}[Y]}{2\sigma_n^2/n_{\text{rep}}} \right\}, \\ \tilde{\sigma}_{ij}^2 &\equiv \tilde{\sigma}^2 = \sigma_n^2/n_{\text{rep}} + \sigma_c^2, \end{aligned} \quad (\text{S21})$$

and we can easily obtain the analytical expressions for the maximum likelihood values of parameters  $\alpha$ ,  $\sigma_n^2$  and  $\sigma_c^2$ . As before, we minimize function

$$\ell_{2A} = -2 \ln L_{2A} = \sum_{ij} \left[ n_{\text{rep}} \frac{V_{ij}[Y]}{\sigma_n^2} + (n_{\text{rep}} - 1) \ln(2\pi\sigma_n^2) + \frac{(M_{ij}[Y] - \alpha)^2}{\tilde{\sigma}^2} + \ln(2\pi\tilde{\sigma}^2) \right],$$

where we used shorthand  $\sum_{ij} \equiv \sum_{i=1}^{n_F} \sum_{j=1}^{n_{\text{pop}}}$ . Next, we set the first partial derivatives of  $\ell_2$  to zero, recalling that  $\tilde{\sigma}^2 = \sigma_c^2 + \sigma_n^2/n_{\text{rep}}$ .

$$\begin{aligned} \frac{\partial \ell_2}{\partial \alpha} &= \frac{1}{\tilde{\sigma}^2} \sum_{ij} 2(M_{ij}[Y] - \alpha) = 0, \\ \frac{\partial \ell_2}{\partial \sigma_c^2} &= \frac{1}{\tilde{\sigma}^2} \sum_{ij} \left[ -\frac{(M_{ij}[Y] - \alpha)^2}{\tilde{\sigma}^2} + 1 \right] = 0, \\ \frac{\partial \ell_2}{\partial \sigma_n^2} &= \frac{1}{\sigma_n^2} \sum_{ij} \left[ -n_{\text{rep}} \frac{V_{ij}[Y]}{\sigma_n^2} + (n_{\text{rep}} - 1) \right] + \frac{1}{n_{\text{rep}}} \frac{\partial \ell_2}{\partial \sigma_c^2} = 0. \end{aligned}$$

Taking into account that  $\sigma_c^2 \geq 0$ , these equations yield

$$\begin{aligned} \hat{\alpha} &= M[Y], \\ \hat{\sigma}_n^2 &= \frac{n_{\text{rep}}}{n_{\text{rep}} - 1} M[V_{..}[Y]] \\ \hat{\sigma}_c^2 &= \max \left\{ 0, V[M_{..}[Y]] - \frac{\hat{\sigma}_n^2}{n_{\text{rep}}} \right\}. \end{aligned}$$

**Model 2B. Fitness, Evolutionary stochasticity and Measurement error** This model, described by equation (S9), has four parameters,  $\alpha$ ,  $\beta$ ,  $\sigma_c^2$ , and  $\sigma_n^2$ . We obtain the equation for the likelihood of data under this model analogously to equation (S20).

$$L_{2B}(Y; \alpha, \beta, \sigma_c^2, \sigma_n^2) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} A_{\text{rep}}(i, j) P_N(M_{ij}[Y]; \alpha + \beta x_i, \tilde{\sigma}_{ij}^2), \quad (\text{S22})$$



where  $A_{\text{rep}}(i, j)$  and  $\tilde{\sigma}^2$  are given by equations (S18) and (S19), respectively. Analogously to Model 2A, to obtain equations for the maximum likelihood parameters under the balanced design we minimize function

$$\ell_{2B} = -2 \ln L_{2B} = \sum_{ij} \left[ n_{\text{rep}} \frac{V_{ij}[Y]}{\sigma_n^2} + (n_{\text{rep}} - 1) \ln(2\pi\sigma_n^2) + \frac{(M_{ij}[Y] - \alpha - \beta x_i)^2}{\tilde{\sigma}^2} + \ln(2\pi\tilde{\sigma}^2) \right].$$

We obtain

$$\begin{aligned} 0 &= M[Y] - \alpha - \beta M[x], \\ 0 &= M[M.[Y] \cdot x] - \alpha M[x] - \beta M[x^2], \\ 0 &= \frac{1}{\tilde{\sigma}^2} \sum_{ij} \left[ -\frac{(M_{ij}[Y] - \alpha - \beta x_i)^2}{\tilde{\sigma}^2} + 1 \right], \\ 0 &= \frac{1}{\sigma_n^2} \sum_{ij} \left[ -n_{\text{rep}} \frac{V_{ij}[Y]}{\sigma_n^2} + (n_{\text{rep}} - 1) \right] + \frac{1}{n_{\text{rep}}} \frac{\partial \ell_2}{\partial \sigma_c^2}. \end{aligned}$$

Thus,

$$\begin{aligned} \hat{\beta} &= \frac{C[M.[Y], x]}{V[x]}, \\ \hat{\alpha} &= M[Y] - \hat{\beta} M[x], \\ \hat{\sigma}_n^2 &= \frac{n_{\text{rep}}}{n_{\text{rep}} - 1} M[V..[Y]], \\ \hat{\sigma}_c^2 &= \max \left\{ 0, \frac{1}{n_F n_{\text{pop}}} \sum_{ij} (M_{ij}[Y] - \hat{\alpha} - \hat{\beta} x_i)^2 - \frac{\hat{\sigma}_n^2}{n_{\text{rep}}} \right\}. \end{aligned}$$

**Model 3A. Founder, Evolutionary stochasticity and Measurement error** This model, described by equation (S10), has four parameters,  $\alpha$ ,  $\sigma_g^2$ ,  $\sigma_c^2$ , and  $\sigma_n^2$ . The likelihood of data is given by

$$\begin{aligned} L_{3A}(Y; \alpha, \sigma_g^2, \sigma_c^2, \sigma_n^2) &= \prod_{i=1}^{n_F} \mathbb{P}(Y_{i..}) = \prod_{i=1}^{n_F} \int_{-\infty}^{\infty} db_i \mathbb{P}(B_i = b_i) \mathbb{P}(Y_{i..}|B_i = b_i) \\ &= \prod_{i=1}^{n_F} \int_{-\infty}^{\infty} db_i P_N(b_i; 0, \sigma_g^2) \prod_{j=1}^{n_{\text{pop}}(i)} \mathbb{P}(Y_{ij.}|B_i = b_i) \\ &= \prod_{i=1}^{n_F} \int_{-\infty}^{\infty} db_i P_N(b_i; 0, \sigma_g^2) \prod_{j=1}^{n_{\text{pop}}(i)} \int_{-\infty}^{\infty} da_{ij} \mathbb{P}(A_{ij} = a_{ij}|B_i = b_i) \mathbb{P}(Y_{ij.}|A_{ij} = a_{ij}, B_i = b_i) \end{aligned}$$

$$\begin{aligned}
&= \prod_{i=1}^{n_F} \int_{-\infty}^{\infty} db_i P_N(b_i; 0, \sigma_g^2) \prod_{j=1}^{n_{\text{pop}}(i)} \int_{-\infty}^{\infty} da_{ij} P_N(a_{ij}; 0, \sigma_c^2) \\
&\times \prod_{k=1}^{n_{\text{rep}}(i,j)} P_N(Y_{ijk}; \alpha + b_i + a_{ij}, \sigma_n^2). \tag{S23}
\end{aligned}$$

As for Model 2A, we obtain

$$\int_{-\infty}^{\infty} da_{ij} P(a_{ij}; 0, \sigma_c^2) \prod_{k=1}^{n_{\text{rep}}(i,j)} P_N(Y_{ijk}; \alpha + b_i + a_{ij}, \sigma_n^2) = A_{\text{rep}}(i, j) P_N(M_{ij}[Y]; \alpha + y_i, \tilde{\sigma}_{ij}^2),$$

where  $A_{\text{rep}}(i, j)$  and  $\tilde{\sigma}_{ij}^2$  are given by equations (S18) and (S19), respectively. This implies

$$L_{3A} = \prod_{i=1}^{n_F} \left( \prod_{j=1}^{n_{\text{pop}}(i)} A_{\text{rep}}(i, j) \right) \int_{-\infty}^{\infty} db_i P_N(b_i; 0, \sigma_g^2) \prod_{j=1}^{n_{\text{pop}}(i)} P_N(M_{ij}[Y]; \alpha + b_i, \tilde{\sigma}_{ij}^2).$$

Applying equations (S12) and (S14), we obtain

$$\begin{aligned}
&\prod_{j=1}^{n_{\text{pop}}(i)} P_N(M_{ij}[Y]; \alpha + b_i, \tilde{\sigma}_{ij}^2) = \prod_{j=1}^{n_{\text{pop}}(i)} P_N(b_i; M_{ij}[Y] - \alpha, \tilde{\sigma}_{ij}^2) \\
&= A_{\text{pop}}(i) P_N(b_i; \tilde{M}_i[\text{M..}[Y]] - \alpha, \tilde{\sigma}_i^2),
\end{aligned}$$

and

$$\int_{-\infty}^{\infty} P_N(b_i; 0, \sigma_g^2) P_N(b_i; \tilde{M}_i[\text{M..}[Y]] - \alpha, \tilde{\sigma}_i^2) db_i = P_N(\tilde{M}_i[\text{M..}[Y]]; \alpha, \sigma_g^2 + \tilde{\sigma}_i^2),$$

where

$$A_{\text{pop}}(i) = (2\pi)^{-\frac{n_{\text{pop}}(i)-1}{2}} \left( \frac{\tilde{\sigma}_i^2}{\prod_{j=1}^{n_{\text{pop}}(i)} \tilde{\sigma}_{ij}^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\tilde{V}_i[\text{M..}[Y]]}{2\tilde{\sigma}_i^2} \right\}, \tag{S24}$$

$$\tilde{\sigma}_i^2 = \left( \sum_{j=1}^{n_{\text{pop}}(i)} \frac{1}{\tilde{\sigma}_{ij}^2} \right)^{-1}, \tag{S25}$$

$$\tilde{M}_i[\text{M..}[Y]] = \tilde{\sigma}_i^2 \sum_{j=1}^{n_{\text{pop}}(i)} \frac{M_{ij}[Y]}{\tilde{\sigma}_{ij}^2}, \tag{S26}$$

$$\tilde{V}_i[\text{M..}[Y]] = \tilde{\sigma}_i^2 \sum_{j=1}^{n_{\text{pop}}(i)} \frac{(M_{ij}[Y])^2}{\tilde{\sigma}_{ij}^2} - \tilde{M}_i^2[\text{M..}[Y]]. \tag{S27}$$

Thus, equation (S23) for the likelihood of data becomes

$$L_{3A} = \prod_{i=1}^{n_F} \left( \prod_{j=1}^{n_{\text{pop}}(i)} A_{\text{rep}}(i, j) \right) A_{\text{pop}}(i) P_N \left( \tilde{M}_i [\text{M}.. [Y]] ; \alpha, \sigma_g^2 + \tilde{\sigma}_i^2 \right). \quad (\text{S28})$$

If the design is balanced, expressions (S24) and (S28) simplify to

$$A_{\text{pop}}(i) = (2\pi\tilde{\sigma}^2)^{-\frac{n_{\text{pop}}-1}{2}} n_{\text{pop}}^{-\frac{1}{2}} \exp \left\{ -\frac{V_i [\text{M}.. [Y]]}{2\tilde{\sigma}^2/n_{\text{pop}}} \right\},$$

$$L_{3A} = \prod_{i=1}^{n_F} \left( \prod_{j=1}^{n_{\text{pop}}(i)} A_{\text{rep}}(i, j) \right) A_{\text{pop}}(i) P_N (M_i [Y] ; \alpha, \sigma_g^2 + \tilde{\sigma}^2/n_{\text{pop}}),$$

with  $\tilde{\sigma}^2$  given by equation (S21), and we can find analytical expressions for the maximum likelihood parameter values  $\alpha$ ,  $\sigma_g^2$ ,  $\sigma_c^2$ , and  $\sigma_n^2$ . As before, we look for the minimum of function

$$\ell_{3A} = -2 \ln L_{3A} = \sum_{i=1}^{n_F} \left[ \sum_{j=1}^{n_{\text{pop}}} \left( n_{\text{rep}} \frac{V_{ij} [Y]}{\sigma_n^2} + (n_{\text{rep}} - 1) \ln (2\pi\sigma_n^2) \right) \right. \\ \left. + n_{\text{pop}} \frac{V_i [\text{M}.. [Y]]}{\tilde{\sigma}^2} + (n_{\text{pop}} - 1) \ln (2\pi\tilde{\sigma}^2) + \frac{(M_i [Y] - \alpha)^2}{\sigma_g^2 + \tilde{\sigma}^2/n_{\text{pop}}} + \ln (2\pi (\sigma_g^2 + \tilde{\sigma}^2/n_{\text{pop}})) \right].$$

Setting the first partial derivatives of  $\ell_{3A}$  to zero, we obtain

$$\frac{\partial \ell_{3A}}{\partial \alpha} = \frac{1}{\sigma_g^2 + \tilde{\sigma}^2/n_{\text{pop}}} \sum_{i=1}^{n_F} 2 (M_i [Y] - \alpha) = 0,$$

$$\frac{\partial \ell_{3A}}{\partial \sigma_g^2} = \frac{1}{\sigma_g^2 + \tilde{\sigma}^2/n_{\text{pop}}} \sum_{i=1}^{n_F} \left[ -\frac{(M_i [Y] - \alpha)^2}{\sigma_g^2 + \tilde{\sigma}^2/n_{\text{pop}}} + 1 \right] = 0,$$

$$\frac{\partial \ell_{3A}}{\partial \sigma_c^2} = \frac{1}{\tilde{\sigma}^2} \sum_{i=1}^{n_F} \left[ -n_{\text{pop}} \frac{V_i [\text{M}.. [Y]]}{\tilde{\sigma}^2} + (n_{\text{pop}} - 1) \right] + \frac{1}{n_{\text{pop}}} \frac{\partial \ell_{3A}}{\partial \sigma_g^2} = 0,$$

$$\frac{\partial \ell_{3A}}{\partial \sigma_n^2} = \frac{1}{\sigma_n^2} \sum_{i=1}^{n_F} \sum_{j=1}^{n_{\text{pop}}} \left[ -n_{\text{rep}} \frac{V_{ij} [Y]}{\sigma_n^2} + (n_{\text{rep}} - 1) \right] + \frac{1}{n_{\text{rep}}} \frac{\partial \ell_{3A}}{\partial \sigma_c^2} = 0,$$

which yields the following maximum likelihood parameter estimates.

$$\hat{\alpha} = M [Y],$$

$$\hat{\sigma}_n^2 = \frac{n_{\text{rep}}}{n_{\text{rep}} - 1} M [V.. [Y]],$$

$$\hat{\sigma}_c^2 = \max \left\{ 0, \frac{n_{\text{pop}}}{n_{\text{pop}} - 1} M [V. [\text{M}.. [Y]]] - \frac{\hat{\sigma}_n^2}{n_{\text{rep}}} \right\},$$

$$\hat{\sigma}_g^2 = \max \left\{ 0, V [\text{M}.. [Y]] - \frac{\hat{\sigma}_c^2}{n_{\text{pop}}} - \frac{\hat{\sigma}_n^2}{n_{\text{pop}} n_{\text{rep}}} \right\}.$$

**Model 3B. Founder fitness, Founder genotype, Evolutionary stochasticity and Measurement noise** This model, described by equation (S11), has five parameters,  $\alpha$ ,  $\beta$ ,  $\sigma_g^2$ ,  $\sigma_c^2$ , and  $\sigma_n^2$ . We obtain the equation for the likelihood of data under this model analogously to equation (S28).

$$L_{3B}(Y; \alpha, \beta, \sigma_g^2, \sigma_c^2, \sigma_n^2) = \prod_{i=1}^{n_F} \left( \prod_{j=1}^{n_{\text{pop}}(i)} A_{\text{rep}}(i, j) \right) A_{\text{pop}}(i) P_N \left( \tilde{M}_i[\text{M.}[Y]]; \alpha + \beta x_i, \sigma_g^2 + \tilde{\sigma}_i^2 \right), \quad (\text{S29})$$

where  $A_{\text{rep}}(i, j)$ ,  $A_{\text{pop}}(i)$ ,  $\tilde{\sigma}_i^2$ , and  $\tilde{M}_i[\text{M.}[Y]]$  are given by equations (S18), (S24), (S25) and (S26), respectively. The maximum likelihood parameter estimates under the balanced design are given by

$$\begin{aligned} \hat{\beta} &= \frac{C[\text{M.}[Y], x]}{V[x]}, \\ \hat{\alpha} &= M[Y] - \hat{\beta}M[x], \\ \hat{\sigma}_n^2 &= \frac{n_{\text{rep}}}{n_{\text{rep}} - 1} M[V..[Y]], \\ \hat{\sigma}_c^2 &= \max \left\{ 0, \frac{n_{\text{pop}}}{n_{\text{pop}} - 1} M[V. [\text{M.}[Y]]] - \frac{\hat{\sigma}_n^2}{n_{\text{rep}}} \right\}, \\ \hat{\sigma}_g^2 &= \max \left\{ 0, V[\text{M.}[Y]] - \frac{\hat{\sigma}_c^2}{n_{\text{pop}}} - \frac{\hat{\sigma}_n^2}{n_{\text{pop}} n_{\text{rep}}} \right\}. \end{aligned}$$

## Analysis of genetic data

**Expected number of mutations and dN/dS** We calculated the number of synonymous, nonsynonymous and nonsense mutations expected under neutrality as follows. First, each codon  $c$  was assigned the number of synonymous,  $n_{\text{syn}}(c)$ , nonsynonymous,  $n_{\text{non}}(c)$ , and nonsense,  $n_{\text{stop}}(c)$ , single point-mutation neighbors based on data from Table 1 in Ref. (11). Next, for each codon we obtained its usage frequency  $\nu_c$  in *S. cerevisiae* genome from the *Saccharomyces* genome database website ([http://downloads.yeastgenome.org/unpublished\\_data/codon/ysc.orf.cod](http://downloads.yeastgenome.org/unpublished_data/codon/ysc.orf.cod)) on October 4, 2013. Thus, the expected frequencies of synonymous, nonsynonymous, and nonsense mutations are  $f_i = \frac{1}{9} \sum_c \nu_c n_i(c)$ , with  $i = \text{syn, non, stop}$ , respectively. Note that the  $f_{\text{stop}}$  overestimates the frequency of potentially neutral nonsense mutations because many premature stop codons in essential genes are lethal. We compared the observed numbers of mutations with the numbers expected from these frequencies in three ways using the  $\chi^2$  test (see Table S9). First, we found that the observed proportions of the three types of mutations were inconsistent with the neutral expectation (Table S9, column 3). The number nonsynonymous mutations relative to the number of synonymous mutations is slightly higher than expected (dN/dS = 1.06), but this difference is not statistically significant (Table S9, column 4). In contrast, the number of nonsense mutations significantly exceeds the

neutral expectation despite this expectation being an overestimate (Table S9, column 5). We conclude that most observed nonsense mutations are beneficial.

**Convergence and parallelism** We tested whether the observed degree of convergent evolution is stronger than expected by chance at two levels of biological organization, genes and GO Slim biological process terms. We downloaded the mapping of *S. cerevisiae* genes onto the GO Slim terms from the *Saccharomyces* genome database website ([http://downloads.yeastgenome.org/curation/literature/go\\_slim\\_mapping.tab](http://downloads.yeastgenome.org/curation/literature/go_slim_mapping.tab)) on October 23, 2012.

We observed that in our data 818 putatively functional mutations affected 676 genes with 80 genes having mutated in 2 or more lines, and 24 genes having mutated in 3 or more lines. First, we calculated the number of genes that would be affected by mutation if the putatively functional mutations were randomly distributed across all yeast ORFs. We carried out  $10^4$  simulations in which we distributed 818 mutations across 6579 yeast ORFs using the multinomial distribution in which the probability of an ORF receiving a mutation was proportional to  $L_{\text{gene}} + L_{\text{prom}}$ , where  $L_{\text{gene}}$  is the gene length and  $L_{\text{prom}} = 500$  is the length of our operationally defined gene’s promoter region upstream of the gene. The results are shown in Table S5. We found that the total number of genes affected by mutation in our data is significantly smaller than expected by chance and, correspondingly, the number of genes with 2 or more mutations or with 3 or more mutations is significantly higher than expected.

Next, we calculated the convergence and parallelism indices at a given level of biological organization as follows. Let  $c(i, j)$  be the number of categories (genes or GO Slim terms) affected by mutation in both clones  $i$  and  $j$ . We define the convergence index  $CI$  the expected number of mutated categories shared among any two randomly picked clones. We define the parallelism index  $PI$  as the expected number of mutated categories shared among any two randomly picked clones descended from the same Founder. More precisely,

$$CI = \frac{2}{N(N-1)} \sum_i \sum_{j>i} c(i, j),$$

$$PI = \frac{2}{\sum_k N_k(N_k-1)} \sum_i \sum_{\substack{j>i: \\ \phi(i)=\phi(j)}} c(i, j).$$

Here,  $\phi(i)$  is the Founder of clone  $i$ ,  $N = 105$  is the total number of sequenced clones and  $N_k$  is the number of sequenced clones descended from Founder  $k$ .

To calculate the null distribution of  $CI$ , we randomly distributed all putatively functional mutations across all yeast ORFs as described above, for each clone separately. Because we are interested in whether the observed value of  $PI$  is greater than expected, given the overall degree of convergence, we calculate the conditional null distribution of  $PI$  by randomly permuting the Founder labels of sequenced clones. We found that the observed  $CI$  values are significantly greater than expected by chance (genes,  $P < 10^{-4}$ ; GO Slim,  $P < 0.01$ ; Figure S7), whereas the observed  $PI$  values lie well within the null distribution (genes,  $P = 0.053$ ; GO Slim,  $P = 0.596$ ; Figure S7).

We conclude that, while adaptation in our environment leads to significant overall convergence at multiple levels of biological organization, populations descended from the same Founder are no more likely to acquire mutations in the same functional units than populations descended from different Founders.

### Correlation between fitness increment and number of mutations

Here we consider the statistical problem of correlating the observed mean fitness increment  $Y_{ij}$  in population  $j$  descended from Founder  $i$  after 500 generations of adaptation with the number of mutations  $m_{ij}$  in multihit genes discovered in the clone sampled from it. We consider three models.

**Model 1A. Simple regression** In this model, we regress the fitness increment against the number of mutations,

$$Y_{ij} = \alpha + \beta m_{ij} + \varepsilon_{ij},$$

where  $\varepsilon_{ij}$  are independent normally distributed random variables with mean 0 and variance  $\sigma_n^2$ . This model has three parameters,  $\alpha$ ,  $\beta$ , and  $\sigma_n^2$ , and the likelihood of the data is given by

$$L_{1A}(Y..) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} P_N(Y_{ij}; \alpha + \beta m_{ij}, \sigma_n^2),$$

where we used the vector notations introduced in section Analysis of covariance of fitness increment with initial fitness. We compare this model with Model 0 in which  $\beta = 0$  and the likelihood of data is given by

$$L_0(Y..) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} P_N(Y_{ij}; \alpha, \sigma_n^2),$$

using the likelihood ratio test.

**Model 2A. Multiple regression** Next, we carry out the multiple regression analysis where we use the initial fitness  $x_i$  of Founder  $i$  as a covariate,

$$Y_{ij} = \alpha + \beta m_{ij} + \gamma x_i + \varepsilon_{ij}.$$

This model has three parameters,  $\alpha$ ,  $\beta$ , and  $\sigma_n^2$ , and the likelihood of the data is given by

$$L_{2A}(Y..) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} P_N(Y_{ij}; \alpha + \beta m_{ij} + \gamma x_i, \sigma_n^2).$$

We compare this model with Model 1A described above and Model 1B in which  $\beta = 0$  and the likelihood of data is given by

$$L_{1B}(Y..) = \prod_{i=1}^{n_F} \prod_{j=1}^{n_{\text{pop}}(i)} P_N(Y_{ij}; \alpha + \gamma x_i, \sigma_n^2),$$

using the likelihood ratio test.

**Model 2B. Analysis of covariance** Third, we carry out the analysis of covariance where we regress the fitness increment against the number of mutations but allow the intercept of this regression to vary from one Founder to another.

$$Y_{ij} = \alpha + A_i + \beta m_{ij} + \varepsilon_{ij}, \quad (\text{S30})$$

where  $A_i$  are independent normally distributed random variables with mean 0 and variance  $\sigma_g^2$ . This model, given by equation (S30), has four parameters,  $\alpha$ ,  $\sigma_g^2$ ,  $\beta$ , and  $\sigma_n^2$ , and we derive the likelihood of data in this model analogously to section Analysis of covariance of fitness increment with initial fitness.

$$L_{2B}(Y..) = \prod_{i=1}^{n_{\text{FG}}} A(i) P_N(M_i[F]; \alpha + \beta M_i[m], \tilde{\sigma}_i^2), \quad (\text{S31})$$

where

$$\begin{aligned} A(i) &= (2\pi\sigma_n^2)^{-\frac{n_{\text{pop}}(i)-1}{2}} (n_{\text{pop}}(i))^{-\frac{1}{2}} \exp \left\{ -\frac{V_i[Y] - 2\beta C_i[Y, m] + \beta^2 V_i[m]}{2\sigma_n^2/n_{\text{pop}}(i)} \right\}, \\ \tilde{\sigma}_i^2 &= \sigma_n^2/n_{\text{pop}}(i) + \sigma_g^2, \end{aligned}$$

and we used the notations for the means and variances introduced in section Analysis of covariance of fitness increment with initial fitness as well as notation

$$C_i[x.., y..] = M_i[(x.. - M_i[x..])(y.. - M_i[y..])]$$

for the covariance within Founder  $i$ .

We compare Model 2B with Model 1A described above and Model 1C in which

$$Y_{ij} = \alpha + A_i + \varepsilon_{ij},$$

and the likelihood of the data is given by

$$L_{1C}(Y..) = \prod_{i=1}^{n_{\text{FG}}} A(i) P_N(M_i[Y]; \alpha, \tilde{\sigma}_i^2),$$

using the likelihood ratio test.

The results are presented in Table S10 and Figure S6.

## Comparison of data with predictions of Fisher's Geometric Model

Fisher's Geometric Model assumes that the genotype of an organism determines  $n$  quantitative traits that in turn determine its fitness in the given environment (12). Each genotype is therefore represented as a point in an  $n$ -dimensional space. It is typically assumed that there is a single point in this space where fitness is maximal, and that fitness

as a function of the underlying traits is a multivariate Gaussian function centered at the optimum (13). Mutations perturb the traits and are represented by vectors in the phenotypic space. It is usually assumed that the distribution of the lengths and directions of mutational vectors does not depend on the particular genotype which they perturb, i.e., there is no epistasis at the level of phenotypes; moreover, the directions of these vectors are assumed to be uniformly distributed. Fisher’s model is typically used to estimate the effective number of phenotypic traits,  $n$ , which is taken as a measure of complexity of the organism (14–16), or to predict the statistical patterns of epistasis among mutations (13, 16). Here we ask whether our data is consistent with the predictions of Fisher’s model.

Fisher’s model makes two major predictions against which we can test our data. First, in Fisher’s model there are fewer beneficial mutations of a particular magnitude available to high-fitness versus low-fitness backgrounds (17). This is true because a mutation of a given magnitude is more likely to move the genotype farther away from the optimum if the genotype is already close to the optimum than if it is far from the optimum. As a consequence, the rate of adaptation of a population adapting on Fisher’s fitness landscape declines as the population approaches the optimal phenotype (14). In other words, Fisher’s model predicts the rule of declining adaptability. Accordingly, our fitness data shown in Figure 1 in the main text are consistent with this prediction of Fisher’s model.

Second, Fisher’s model predicts that the effect of any given mutation on fitness would strongly depend on the genetic (or, more precisely, phenotypic) background in which it occurs (Figure S2), i.e., it predicts extensive epistasis. Importantly, the fitness effect of a mutation is predicted to depend not only on the distance of the background phenotype to the optimum, but on its specific position relative to the optimum (Figure S2B). Therefore, because the fitness of the genetic background is entirely determined by its distance to the optimum, the effect of a given mutation is not expected to depend on fitness alone. Therefore, in a general case, Fisher’s model predicts that (a) different mutations would be beneficial in different genetic backgrounds and (b) the effects of a given mutation on fitness would strongly depend on the genotype but not on fitness of the genetic background in which it occurs. In other words, Fisher’s model is a special case of the idiosyncratic epistasis model discussed in the main text. Consequently, our genetic and knock-out data does not support Fisher’s model.

We note as an aside that it is in principle possible to reconcile our genetic and knock-out data with Fisher’s model, if all sequenced Founders fall on a line that goes through the optimum, on the same side with respect to the optimum. This degenerate case is equivalent to Fisher’s model with  $n = 1$  (Figure S2A). While such situation is formally possible, this implies that there is only one physiological way to achieve a given value of fitness, which appears biologically unlikely.

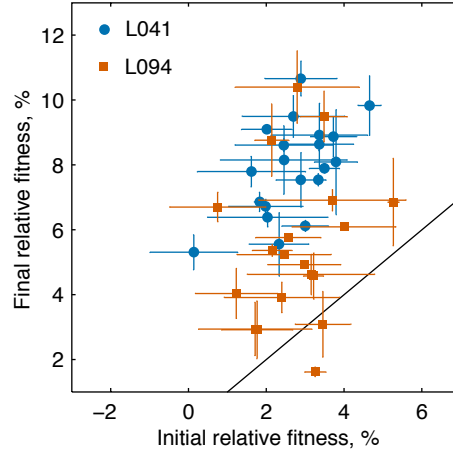


## References

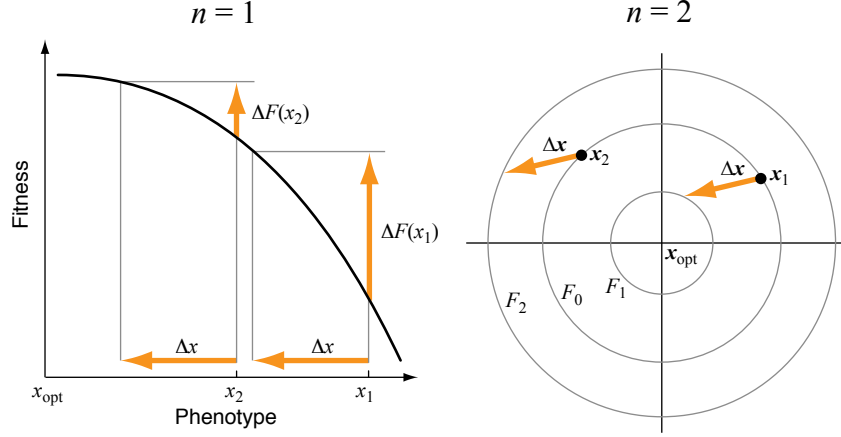
1. Lang GI, Murray AW, Botstein D (2009) The cost of gene expression underlies a fitness trade-off in yeast. *Proc Natl Acad Sci USA* 106: 5755–5760.
2. Lang GI, Botstein D, Desai MM (2011) Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* 188: 647–661.
3. Sherman F, Fink GR, Lawrence CW (1974) *Methods in yeast genetics: laboratory manual*. Cold Spring Harbor Laboratory.
4. Smukalla S, Caldara M, Pochet N, Beauvais A, Guadagnini S, et al. (2008) *FLO1* is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* 135: 726–737.
5. Breslow DK, Cameron DM, Collins SR, Schuldiner M, Stewart-Ornstein J, et al. (2008) A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods* 5: 711–718.
6. Meiron H, Nahon E, Raveh D (1995) Identification of the heterothallic mutation in HO-endonuclease of *S. cerevisiae* using *HO/ho* chimeric genes. *Curr Genet* 28: 367–373.
7. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359.
8. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, et al. (2012) Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* 500: 571–574.
9. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43: 491–498.
10. Travisano M, Mongold JA, Bennett AF, Lenski RE (1995) Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267: 87–90.
11. Zhang J (2005) On the evolution of codon volatility. *Genetics* 169: 495–501.
12. Fisher RA (1999) *The genetical theory of natural selection*. Oxford University Press.
13. Martin G, Elena SF, Lenormand T (2007) Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nature Genet* 39: 555–560.
14. Orr HA (2000) Adaptation and the cost of complexity. *Evolution* 54: 13–20.

15. Tenaillon O, Silander OK, Uzan JP, Chao L (2007) Quantifying organismal complexity using a population genetic approach. *PLoS One* 2: e217.
16. Weinreich DM, Knies JL (2013) Fisher’s geometric model of adaptation meets the functional synthesis: Data on pairwise epistasis for fitness yields insights into the shape and size of phenotype space. *Evolution* 67: 2957–2972.
17. Hartl DL, Taubes CH (1998) Towards a theory of evolutionary adaptation. *Genetica* 102/103: 525–533.
18. Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, et al. (2012) The molecular diversity of adaptive convergence. *Science* 335: 457–461.

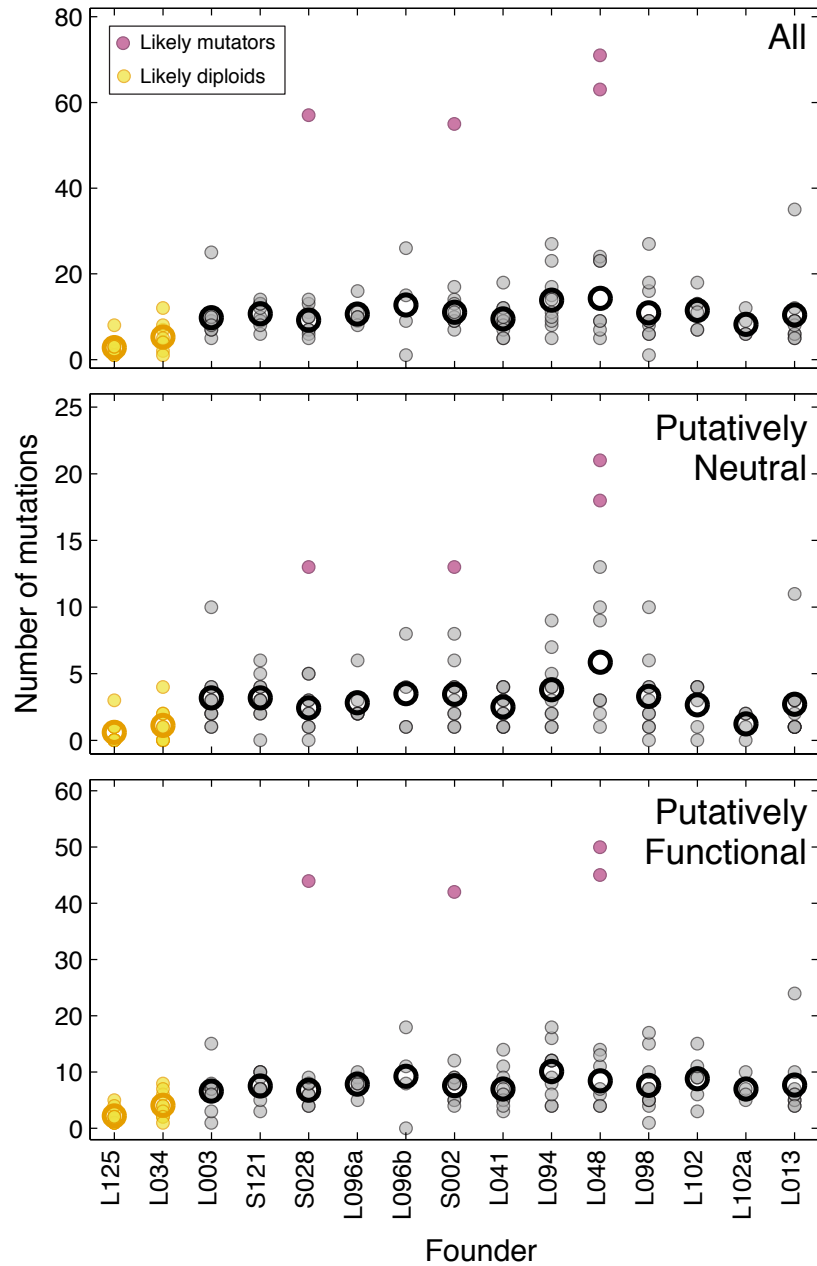
## Supplementary Figures



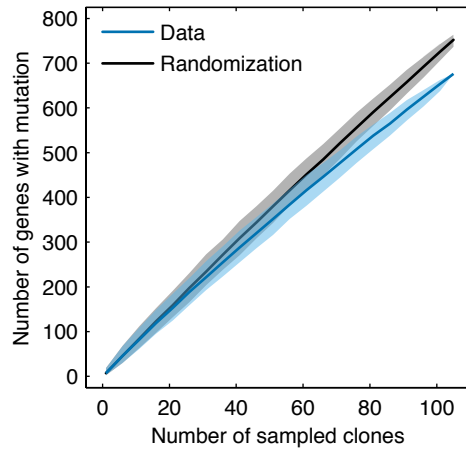
**Figure S1.** Replay experiment. To test whether the apparent difference in adaptability between Founders with similar fitness, L041 and L094, is real, we founded additional 20 replicate populations from each Founder and evolved them for 250 generations in conditions identical to the main experiment. We show the correlation between the initial fitness of each replicate line and its fitness after 250 generations of adaptation. Fitnesses of L041- and L094-descended populations were indistinguishable initially (Nested ANOVA,  $F_{1,37} = 0.03$ ,  $P = 0.86$ ), but significantly different after 250 generations of evolution ( $F_{1,37} = 15.95$ ,  $P < 10^{-3}$ ). See also Table S3. Error bars show  $\pm 1$  SEM.



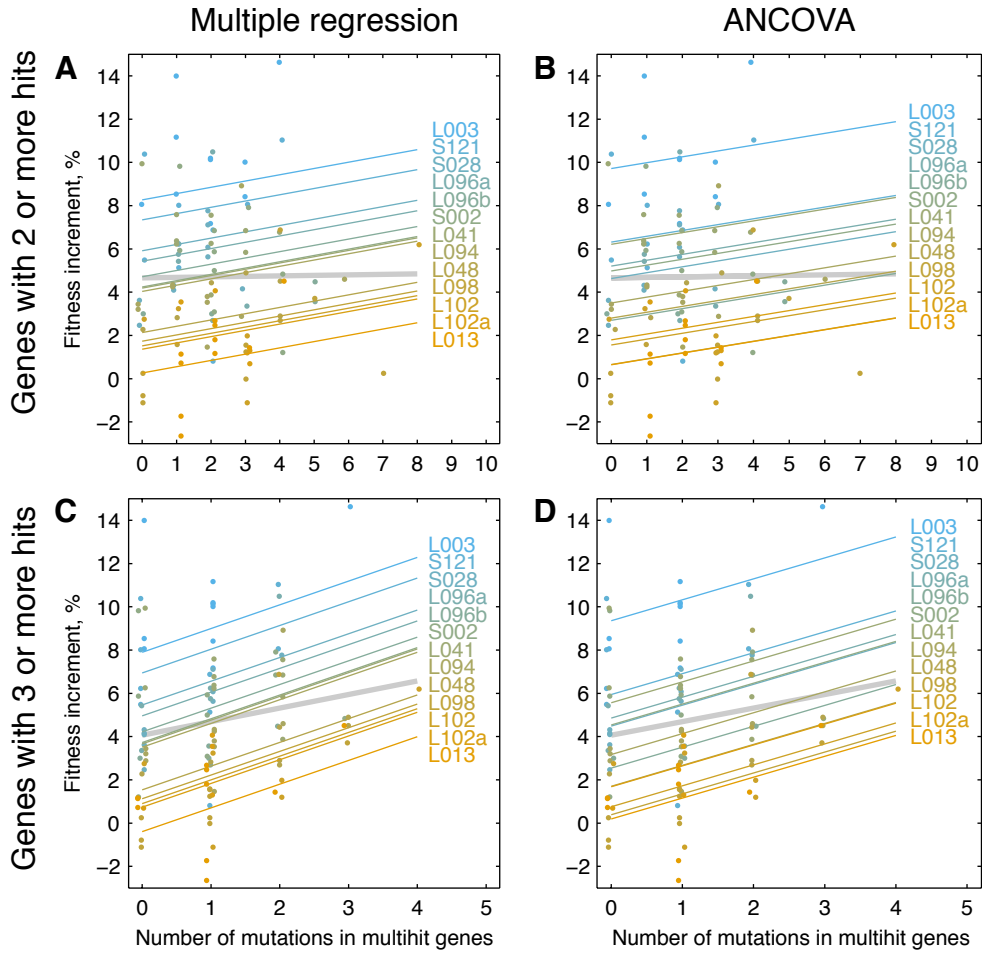
**Figure S2.** Fisher's geometric model in one and two dimensions. **A.** In this schematic, black line shows fitness  $F$  as a function of a single phenotypic trait  $x$ . A particular beneficial mutation moves the phenotype value by the same  $\Delta x$  in the phenotype space towards the optimal value  $x_{\text{opt}}$ , irrespectively of the background (assuming no epistasis at the phenotype level). The resulting fitness effect of this mutation is larger if it occurs in a low-fitness background  $x_1$  than if it occurs in a high-fitness background  $x_2$ , i.e.,  $\Delta F(x_1) > \Delta F(x_2)$ . This leads to global diminishing returns epistasis observed in Figure 4 in the main text. **B.** In this schematic, the coordinate plane represents a two-dimensional phenotype space, and gray circles show fitness isoclines. A particular mutation moves the organism in the phenotype space by the same vector  $\Delta \mathbf{x}$  in the phenotype space irrespectively of the background (assuming no epistasis at the phenotype level), but the sign and size of its effect on fitness strongly depend on the position of the background strain in the phenotype space, not just on its distance to the optimum  $\mathbf{x}_{\text{opt}}$ . In the extreme case shown here, a mutation is beneficial ( $F_1 - F_0 > 0$ ) if it occurs in a strain with phenotype  $\mathbf{x}_1$  and fitness  $F_0$ , but it is deleterious ( $F_2 - F_0 < 0$ ) if it occurs in a strain with phenotype  $\mathbf{x}_2$  and the same fitness  $F_0$ . This implies extensive sign epistasis which we do not observe in our data (Figures 3 and 4 in the main text).



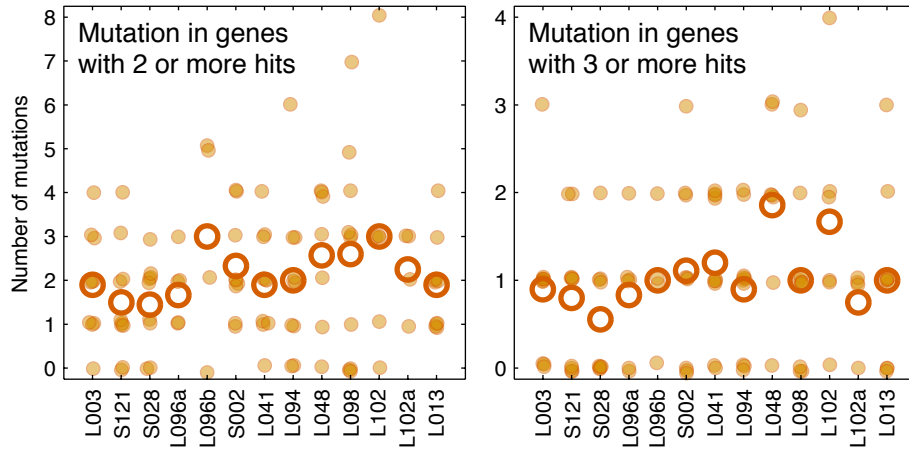
**Figure S3.** Number of mutations in each line descended from each of the 15 sequenced Founders. Larger circles indicate the average number of mutations for populations descended from the same Founder. Clones shown in yellow are likely diploid, limiting our ability to accurately call mutations. Clones shown in purple apparently acquired mutator phenotypes. Top, middle, and bottom panels show all mutations, putatively neutral mutations, and putatively functional mutations, respectively. Founders are ordered from least-fit (left) to most-fit (right).



**Figure S4.** Rarefaction curve (blue) showing the number of unique genes in which we identified any putatively functional mutation as a function of the number of clones sequenced. Black line indicates the expectation if mutations were randomly drawn from all yeast ORFs. Envelopes show 95% confidence intervals. The observed degree of gene-level convergent evolution, while highly significant, is much weaker than observed in *E. coli* by Tenaillon et al. (18)

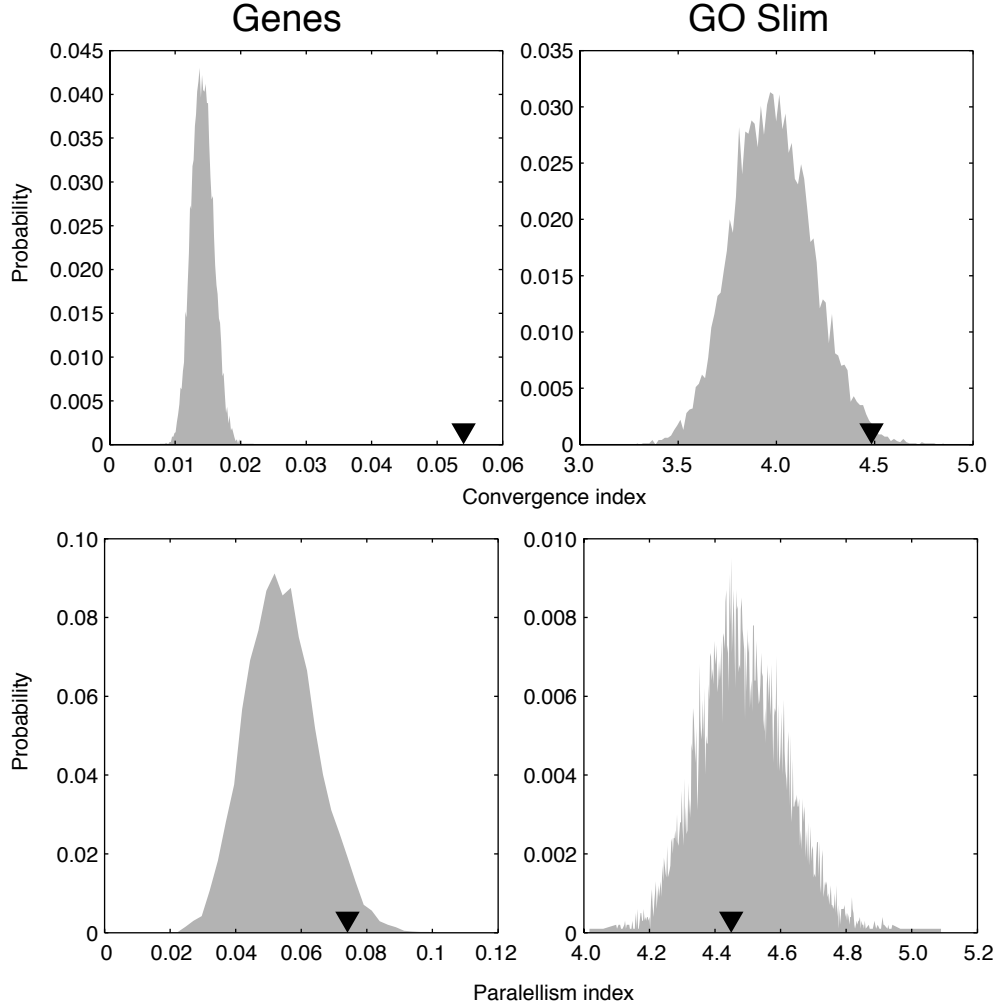


**Figure S5.** Fitness increment of each line after 500 generations of adaptation as a function of the number of mutations within 2 or more hit genes (**A,B**) or 3 or more hit genes (**C,D**) in the clone sampled from that line. Lines descended from each Founder are indicated with the corresponding color according to the legend. Overall there is only a weak, non-significant correlation between the number of mutations in multi-hit genes and the increase in fitness of a given clone (gray lines). However, conditional on the fitness of the Founder (**A,C**) or its identity (**B,D**), such correlation is strong and significant (colored lines). See Table S10 for the best-fit parameters used in this figure; note that the slopes of lines for different Founders are identical by construction (see section Correlation between fitness increment and number of mutations and Table S10). Points are slightly displaced along the  $x$ -axis for clarity.

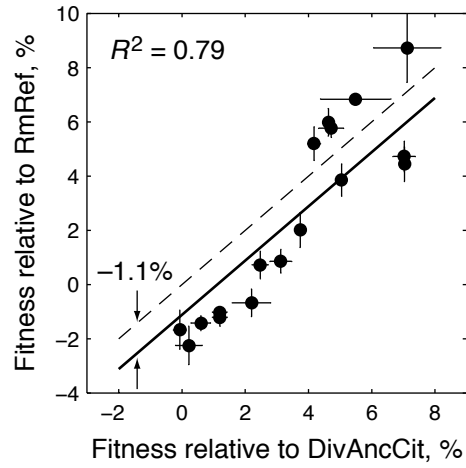


**Figure S6.** Number of multi-hit mutations in each sequenced line, organized by Founder. Founders are ordered from least-fit (left) to most-fit (right). Note that there is no tendency for lines descended from less-fit Founders to acquire more multi-hit mutations than for those descended from more-fit Founders. Points are slightly displaced along the  $x$ -axis for clarity.

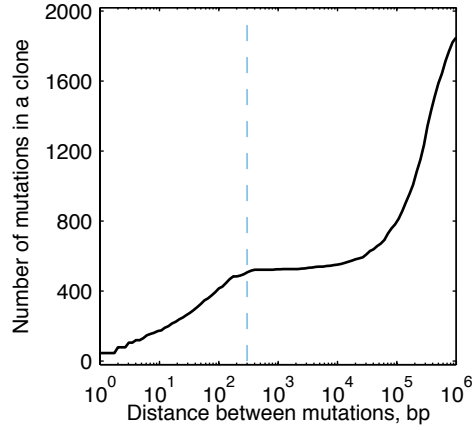




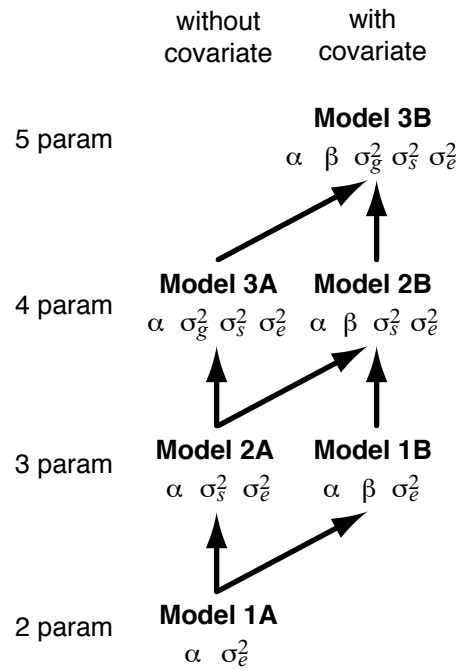
**Figure S7.** Degree of convergence and parallelism at the gene and GO Slim levels as measured by the convergence index  $CI$  (top) and parallelism index  $PI$  (bottom). The null distributions for the  $CI$  and  $PI$  estimated from  $10^4$  randomizations are shown in gray (see Materials and Methods for details). Black triangles show the observed values of  $CI$  and  $PI$  in our data. Note that the  $PI$  null distributions are conditional on the observed degree of overall convergence.



**Figure S8.** Comparison of fitness measurements against two reference strains, DivAncCit ( $x$ -axis) and RmRef ( $y$ -axis), for all strains used in the knock-out experiments, with the exception of strain L096a-4-1 (ySAK0075), whose fitness could not be accurately measured with respect to RmRef. Dashed line shows the diagonal; solid line shows the line with slope 1 with the best fit intercept which indicates that RmRef is about 1% more fit than DivAncCit.



**Figure S9.** Number of pairs of mutations called in the same clone within a certain distance of each other as a function of this distance, prior to identifying complex mutations. Mutations located on different chromosomes are assumed to be infinitely far apart. At above 300 bp (vertical blue line), the number of mutation pairs within a window of a given size grows linearly with the window size indicating that mutations are independently uniformly distributed along the chromosome. The inflection point at around 300 bp indicates that there is an excess of mutations that occur close to each other within the same clone, which suggests that these mutational events are not independent. Therefore, we conservatively treat such genomic changes as the same complex mutational event.



**Figure S10.** Nested models in the analysis of covariance of fitness increment with initial fitness. Arrows point from a less general to a more general model of which the less general model is a special case.

## Supplementary Tables

**Table S1.** Strains used in this work.

Generation	Model	ML parameter values					$\ell$
		$\alpha$	$\beta$	$\sigma_g^2$	$\sigma_c^2$	$\sigma_n^2$	
250	1A	3.32	0	0	0	9.05	6290.2
	1B	4.26	-0.47	0	0	6.25	5828.5
	2A	3.32	0	0	7.51	1.53	5558.6
	2B	4.25	-0.47	0	4.73	1.53	5302.9
	3A	3.28	0	3.06	4.46	1.53	5394.9
	3B	4.25	-0.47	0.27	4.45	1.53	5297.3
500	1A	6.55	0	0	0	12.15	6808.0
	1B	7.88	-0.67	0	0	6.53	6016.3
	2A	6.55	0	0	9.56	2.59	6192.1
	2B	7.88	-0.67	0	3.94	2.59	5727.2
	3A	6.48	0	6.18	3.48	2.59	5839.1
	3B	7.88	-0.67	0.47	3.47	2.59	5711.0

**Table S2.** Maximum likelihood parameter values for the ANCOVA models for the fitness increment. In all comparisons of nested models, the likelihood ratio test rejects the less complex model with  $P \ll 10^{-3}$ , except for the comparison of models 3A and 3B at Generation 250, in which model 3A is rejected with  $P = 0.018$ .

	Founder	Fitness, %	Rate of adaptation, %		Nested ANOVA		
			Original	Replay	<i>F</i>	<i>dfs</i>	<i>P</i>
Pair 1	S104	−0.5	2.7	3.7	4.36	1, 33	0.045
	S075	−1.0	1.5	3.3			
Pair 2	L103	1.3	1.8	3.8	0.7	1, 35	0.280
	S016	1.1	0.9	2.7			
Pair 3	L041	2.3	2.1	5.2	11.69	1, 38	0.002
	L094	2.4	0.6	2.0			
Pair 4	L015	3.3	1.2	1.8	1.23	1, 37	0.275
	L031	3.6	0.1	3.0			

**Table S3.** Results of the replay experiment. Nested ANOVA was used to test whether replicate populations descended from different Founders achieved systematically different final fitnesses after 250 generations of evolution.

**Table S4.** List of mutations.



Genes affected by at least	Observed	Expected	$P$
1 mutation	676	754.0	$< 10^{-4}$
2 mutations	80	59.4	$< 0.005$
3 mutations	24	4.2	$< 10^{-4}$

**Table S5.** Observed and expected number of genes affected by mutation.

Gene (hits)	Pop.	Mut. type	Pos. <sup>1</sup>	Pop.	Mut. type	Pos. <sup>1</sup>
<i>ACE2</i> (12)	L003-11	frameshift	543	L048-3	frameshift	742
	S121-4	frameshift	328	L048-5	T to K	626
	L096a-4	Q to *	143	L048-12	frameshift	325
	S002-2	promoter	−16	L102	frameshift	169
	S002-4	C to S	666	L102a-5	E to *	102
	L094-4	K to *	697	L102a-6	N to K	619
<i>IRA1</i> (8)	L003	S to *	1911	L094-12	I to L	1654
	S028-10	K to *	2287	L048	N to Y	658
	L041-4	frameshift	700	L098	S to *	1855
	L094-4	T to K	684	L013-1	frameshift	324
<i>SFL1</i> (6)	S028-1	R to K	204	L098-9	R to K	117
	L041-10	S to *	213	L013-5	R to K	117
	L048-5	frameshift	429	L013-8	W to *	84
<i>SUN4</i> (6)	L096a-7	G to *	312	L102-11	S to *	337
	S002-8	S to L	119	L102-12	frameshift	71
	L102-7	W to C	190	L013-11	W to *	184
<i>IRA2</i> (6)	L096b-1	G to V	2581	L041-1	F to S	2658
	L096b-8	P to Q	2568	L048-9	A to S	1483
	S002-12	S to *	2504	L098-12	P to Q	1091
<i>HMG1</i> (6)	L003-6	R to K	1025	S028-11	G to V	811
	S121-2	frameshift	776	S002-6	F to I	751
	S121-8	R to I	801	L041-9	Q to K	968
<i>WHI2</i> (5)	S121-4	Q to *	232	L096b-6	Q to *	29
	S121-10	frameshift	74	L102-11	frameshift	74
	S028-1	promoter	−137			
<i>GRR1</i> (5)	L003-11	R to L	418	L096a-10	F to L	952
	S121-7	R to S	680	L094-5	C to Y	465
	L096a-2	L to F	738			
<i>KEL1</i> (4)	S002-7	frameshift	725	L013-1	frameshift	746
	L094-7	T to K	405	L013-11	frameshift	857
<i>FUS3</i> (4)	L003-10	S to *	165	L094-2	A to T	256
	L041-2	Y to *	200	L013-4	E to D	94

Gene (hits)	Pop.	Mut. type	Pos. <sup>1</sup>	Pop.	Mut. type	Pos. <sup>1</sup>
<i>PTR2</i> (4)	S002-8	S to Y	515	L048-6	L to W	129
	S002-12	I to N	398	L102-7	W to L	437
<i>IRC8</i> (4)	S028-9	L to P	122	L048-7	S to *	618
	L096a-7	frameshift	710	L102-10	M to I	1
<i>EGT2</i> (4)	S002-2	frameshift	575	L098-9	E to *	302
	L048-12	S to *	457	L102a-2	frameshift	764
<i>MIH1</i> (4)	S121-7	Q to *	391	L098-12	frameshift	408
	L041-11	S to *	213	L102-10	S to L	221
<i>ECM21</i> (3)	L048-6	R to S	51	L013-1	D to E	1023
	L098-4	frameshift	776			
<i>MSH3</i> (3)	L094-5	K to N	22	L013-7	R to Q	833
	L094-11	E to Q	471			
<i>HSL1</i> (3)	L041-6	E to *	799	L013-12	I to S	320
	L041-9	E to *	479			
<i>CSI1</i> (3)	L048-3	D to V	194	L098-10	G to C	261
	L098-6	V to I	189		G to R	262
<i>SUP35</i> (3)	L096b-1	V to F	505	L048-9	G to C	449
	S002-8	E to V	610			
<i>PIN4</i> (3)	L048-12	P to Q	231	L102	T to I	74
	L098-8	D to E	304			
<i>DIG1</i> (3)	L041-4	frameshift	186	L098	S to P	406
	L094	W to *	355			
<i>CDC28</i> (3)	L003-2	D to V	270	L102-11	R to H	131
	L003-4	K to N	13			
<i>GAT2</i> (3)	L041-1	R to I	491	L041-12	Q to *	280
	L041-2	frameshift	465			
<i>SUR2</i> (3)	L003-1	frameshift	173	S121-1	F to S	238
	L003-8	R to S	177			

**Table S6.** Genes hit 3 or more times. <sup>1</sup>Position shows residue number if mutation happened within the ORF or the nucleotide distance to ATG if it happened outside of the ORF.

**Table S7.** GO term enrichment.

Purpose	Primer id	Primer sequence (5' to 3')
<i>HIS3-ymCitrineM233I</i> amplification	oGW137	TTGGTGAGCGCTAGGAGTC
	oGW138	TATGAAATGCTTTTCTTGTTGTTCTTACG
KO confirmation	KanB	CTGCAGCGAGGAGCCGTAAT
	KanC	TGATTTTGATGACGAGCGTAAT
<i>KanMX</i> amplification for <i>gat2</i> KO	oSK-GAT2-F03	TTTGAGTTTCAGTTTCTTGATATCTTGGT
	oSK-GAT2-R03	CATTCCCCTGTGAATCTAAACAATTTATG
<i>gat2Δ</i> confirmation	oSK-GAT2-A	AATTGTGCAAAAGCTAAACTCAAC
	oSK-GAT2-B1	CCACTACTGTTTTGTTGCCACTAT
	oSK-GAT2-C	GAAGATCTATTGATTTAGCCAACGA
	oSK-GAT2-D	TTTGAAATCTCGAGTATATTGAGGC
<i>KanMX</i> amplification for <i>ace2</i> KO	oSK-ACE2-F01	GGTTATGTCCCTATAAACGATGACTATTG
	oSK-ACE2-R01	GTACACATAAACTTATAACAACCTCTCTCG
<i>ace2Δ</i> confirmation	oSK-ACE2-A	GGAGTACTGTTTGTCTTACTCGCAAT
	oSK-ACE2-B	TCCCTTATATGGTCAAAGTCATCAT
	oSK-ACE2-C	GACACAAGTCCCGTAAAAGAACTA
	oSK-ACE2-D	TGTATTTTCGAGATCAACTTTAGC
<i>KanMX</i> amplification for <i>whi2</i> KO	oSK-WHI2-F01	TAAGGGAAAAGAGAAACAACGTTATTTAGT
	oSK-WHI2-R03	ATGTGCTTTGAGACAGTTCATATAATTTTG
<i>whi2Δ</i> confirmation	oSK-WHI2-A	GCATAGGCATAGTGATAGAGTGTGA
	oSK-WHI2-B	AGGAGGGAAATTAACAATGTAGACC
	oSK-WHI2-C	CAGTACCATTTCTACAGCAAGGAAT
	oSK-WHI2-D	CTATTGTTTTATACCGGATCAATGC
<i>KanMX</i> amplification for <i>sfl1</i> KO	oSK-SFL1-F04	AAAAAGAACATAGTGAACCTTCATCGC
	oSK-SFL1-R03	ATAGTTATAATCACAAGGATCAGGAGGAAA
<i>sfl1Δ</i> confirmation	oSK-SFL1-A	ATTTTTCAAAGAAGTCGTAAGTGGTG
	oSK-SFL1-B	CCACTACTGTTTTGTTGCCACTAT
	oSK-SFL1-C	TTCTTCAAATTTAATAAACTCCCCC
	oSK-SFL1-D	TGGAAATGAAAGTGGAATAACAAT
<i>KanMX</i> amplification for <i>ho</i> KO	oSK-HO-F02	ACGCACTATTCATCATTAATATTTAAAGC
	oSK-HO-R02	TTGTTTTTCTTTCGCGTAAATATTTAGCTT
<i>hoΔ</i> confirmation	oSK-HO-A	TATTAGGTGTGAAACCACGAAAAGT
	oSK-HO-B	ACTGTCATTGGGAATGTCTTATGAT
	oSK-HO-C	GAGTGGTAAAAATCGAGTATGTGCT
	oSK-HO-D	CATGTCTTCTCGTTAAGACTGCAT

**Table S8.** Primers used in this work.

Mutation type	Observed	Expected		
		All	Syn-Non	Syn-Stop
Syn	133	143.8	139.2	152.1
Non	488	498.1	481.8	N/A
Stop	54	33.0	N/A	34.9
$\chi^2$		14.34	0.35	12.82
<i>dfs</i>		2	1	1
<i>P</i>		< 0.001	0.554	< 0.001

**Table S9.** Observed and expected numbers of synonymous (Syn), nonsynonymous (Non), and nonsense (Stop) mutations. Last three columns show the expected numbers for all types of mutations, nonsynonymous relative to synonymous mutations, and nonsense relative to synonymous mutations, respectively. Bottom half of the table shows the results of the corresponding  $\chi^2$  tests.

Mutations in	Model	ML parameter values					$\ell$	$P_1$	$P_2$
		$\alpha$	$\beta$	$\gamma$	$\sigma_g^2$	$\sigma_n^2$			
genes with 2 and more hits	0	4.71	0	0	0	11.04	549.1	N/A	N/A
	1A	4.66	0.02	0	0	11.15	549.1	1.00	N/A
	1B	7.05	0	-0.73	0	5.48	474.6	N/A	N/A
	1C	4.48	0	0	5.43	4.44	491.0	N/A	N/A
	2A	6.52	0.29	-0.75	0	5.34	470.8	$\ll 10^{-6}$	0.051
	2B	3.97	0.24	0	5.76	4.27	487.4	$\ll 10^{-6}$	0.060
genes with 3 and more hits	0	4.71	0	0	0	11.04	549.1	N/A	N/A
	1A	4.06	0.63	0	0	10.82	546.0	0.078	N/A
	1B	7.05	0	-0.73	0	5.48	474.6	N/A	N/A
	1C	4.48	0	0	5.43	4.44	491.0	N/A	N/A
	2A	6.08	1.10	-0.78	0	4.56	454.4	$\ll 10^{-6}$	$7 \times 10^{-6}$
	2B	3.44	1.00	0	6.05	3.60	474.9	$\ll 10^{-6}$	$6 \times 10^{-5}$

**Table S10.** Maximum likelihood parameter values for models of correlation of the fitness increment of a population with the number of mutations in a clone sampled from it.  $\ell = -2 \log L$ , where  $L$  is the likelihood of the data.  $P_1$  and  $P_2$  are the  $P$ -values for the comparison with the two nested models using the likelihood ratio test, number of degrees of freedom is 1 in all comparisons. Model 1A is compared only against model 0; Model 2A is compared against Models 1A and 1B; Model 2B is compared against Models 1A and 1C.