nature ecology & evolution

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41559-018-0758-1

In the format provided by the authors and unedited.

The dynamics of adaptive genetic diversity during the early stages of clonal evolution

Jamie R. Blundell[1,2,3]*, Katja Schwartz[4], Danielle Francois[2], Daniel S. Fisher[1], Gavin Sherlock[4] and Sasha F. Levy[2,4,5,6,7]*

[1]Department of Applied Physics, Stanford University, Stanford, CA, USA. [2]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY, USA. [3]CRUK Cambridge Center Early Detection Programme, Department of Oncology, University of Cambridge, Cambridge, UK. [4]Department of Genetics, Stanford University, Stanford, CA, USA. [5]Joint Initiative for Metrology in Biology, Stanford, CA, USA. [6]National Institute of Standards and Technology, Gaithersburg, MD, USA. [7]SLAC National Accelerator Laboratory, Menlo Park, CA, USA. *e-mail: jrb75@cam.ac.uk; sflevy@stanford.edu

SUPPLEMENTARY INFORMATION

The dynamics of adaptive genetic diversity during the early stages of clonal evolution

Jamie R. Blundell[1,3,5,6], Katja Schwartz[2], Danielle Francois[5], Daniel S. Fisher[1], Gavin Sherlock[2], Sasha F. Levy[2,3,4,5]

1. Department of Applied Physics, Stanford University, Stanford, CA 94305, USA
2. Department of Genetics, Stanford University, Stanford, CA 94305, USA
3. Joint Initiative for Metrology in Biology, Stanford, California 94305-4245, USA
4. National Institute of Standards and Technology, Gaithersburg, MD 20899-1070, USA
5. Laufer Centre for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY 11794-5215, USA
6. CRUK Cambridge Center Early Detection Programme, Department of Oncology, University of Cambridge, UK

November 7, 2018

Contents

1 N-lim experiment and DFE

1.1 Media used for N-lim condition

The barcoded yeast library from [1], containing ∼500,000 barcodes, was evolved by serial batch culture under nitrogen limitation in 100 ml of 5x Delft media [2] with 0.04% ammonium sulfate and 4% dextrose. Cells were grown in 500 ml Delong flasks (Bellco) at 30°C and 223 RPM for 48 hours between each bottleneck. Bottlenecks were performed by adding 400 μl of the evolution to fresh media. Cell counts were performed at each bottleneck to estimate the generation time. Contamination checks for bacteria or other non-yeast microbes were performed regularly. Barcode sequencing and counting was performed, as described [1].

1.2 Inferring the DFE in N-lim

Here we outline how we infer the DFE in N-lim and compare its features with those measured previously in C-lim.

To construct the DFE in N-lim we followed the procedure outlined in detail in [1]. Because adaptation is slower in N-lim we used trajectories for ∼200 generations with read depths given in the following table:

Generation	0	8	16	48	72	80	96	104	112	120	128	136	144	160	168	176	184	192	200
Depth N1 ($\times 10^6$)	222	36	30	36	30	33	39	39	36	35	32	24	23	23	34	39	20	39	–
Depth N2 ($\times 10^6$)	222	–	30	38	34	50	34	29	–	39	–	37	–	29	–	37	–	25	47

Table 1: Read depths across time points for the two replicate evolutions in nitrogen limitation (N-lim). Only these time points are used in the analysis.

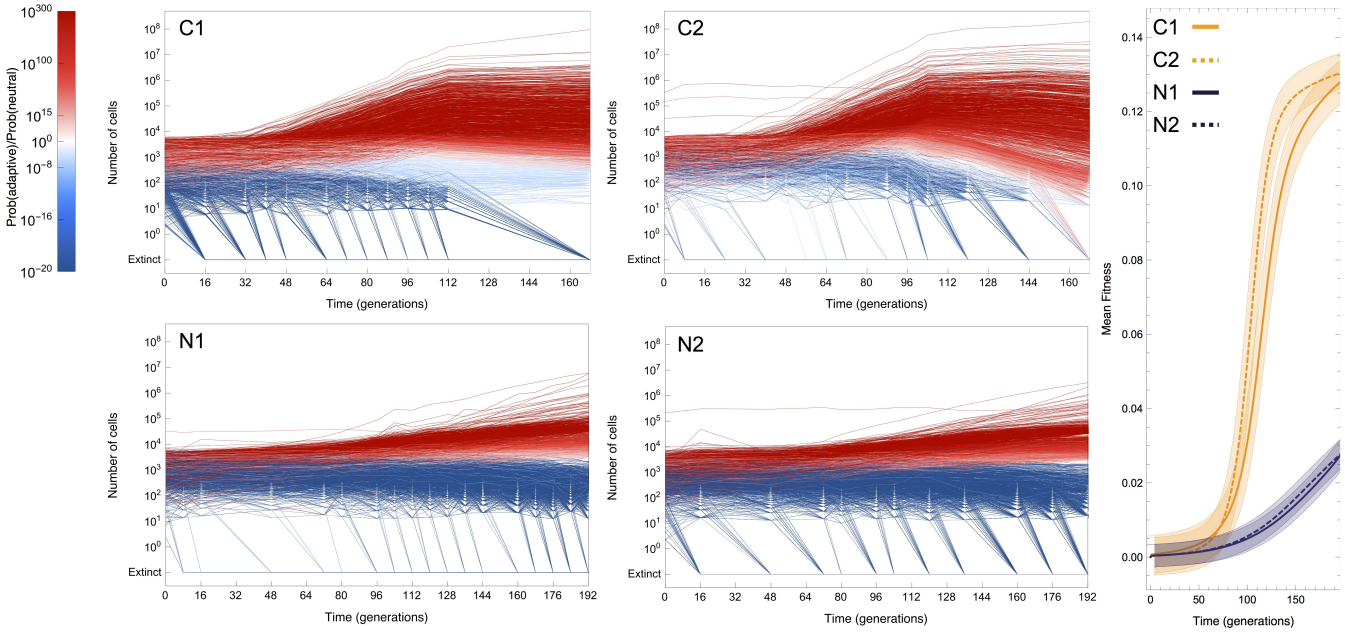
To identify lineages with beneficial mutations, using the same null model as outlined in [1], each barcode trajectory was assigned a posterior probability of harboring an established beneficial mutation with fitness effect s and establishment time τ based on its abundance change between subsequent time points. Lineage trajectories in both N-lim replicates (N1 and N2) are shown in Figure 1 alongside the original C-lim trajectories (C1 and C2) (previously published in [1]) for comparison. The more rapid adaptation observed in C-lim relative to N-lim can be quantified by plotting the mean-fitness trajectory (mean fitness of cells in the population relative to the ancestor), shown in the right-hand panel of Figure 1. These mean fitness trajectories are inferred using the fitness estimates and interpolated trajectories for each adaptive lineage identified in each environment.

To infer the DFE in N-lim, we first filtered out adaptive lineages that were identified as adaptive in both replicates (purple points in Figure 2) as these were likely pre-existing mutations that arose prior to the beginning of the experimental evolutions [1]. Figure 2 shows all beneficial mutations identified as adaptive colored according to whether they are pre-existing (purple) or not (green) in both N-lim replicates and, for comparison, the previously published C-lim replicates.

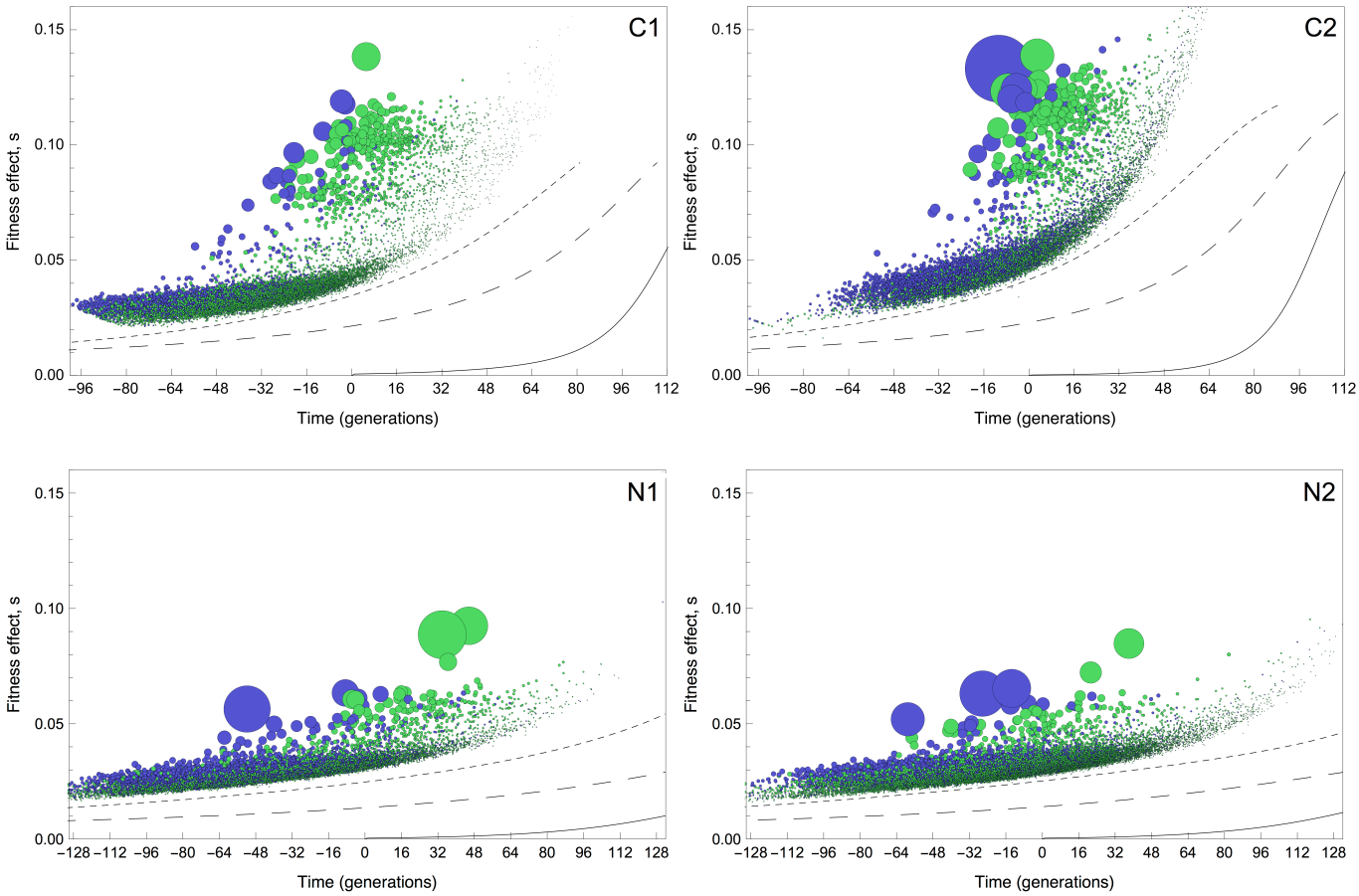
Next we used the estimates of s inferred for each adaptive lineage together with the deterministic approximation outlined in [1] (that relates the mutation rate density $\mu(s)$ to fitness effects in the range $[s, s+ds]$ to the measured fraction $f(ds,t)$ of cells in the population expanding at rates between $[s, s+ds]$ at time t) to estimate the rate of mutation to each fitness interval via:

$$f(ds,t) = \frac{\mu(s)ds}{s} e^{st} \tag{1}$$

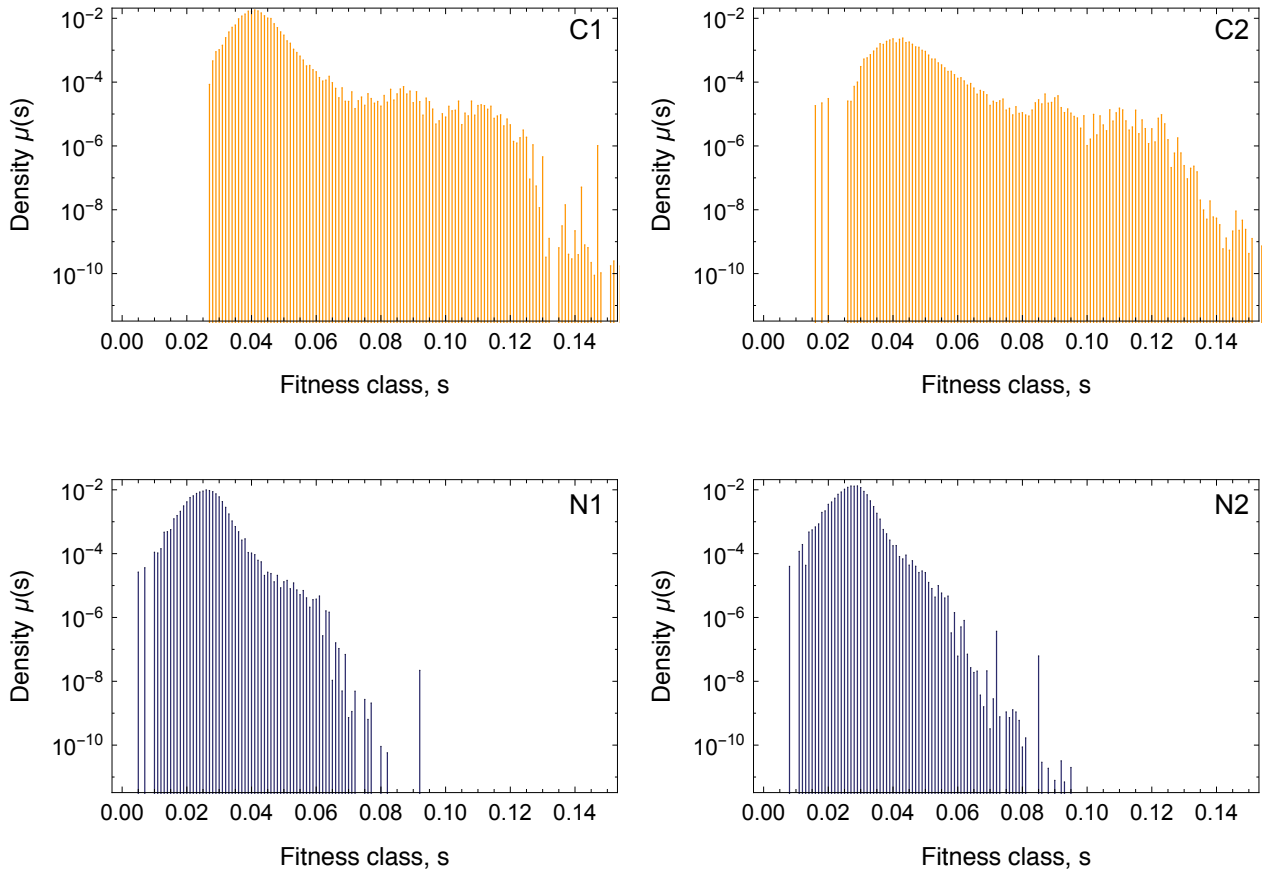
which is independent of the variance in offspring number [1]. The result of this procedure gives one an estimate for the mutation rate density as a function of fitness effect as shown here in Figure 3 and also in the main text Figure 2. The lower two panels are inferences for the N-lim replicates while the upper two panels show previously published DFE estimates for the two C-lim replicates. It should be noted that the y-axes is log-scaled to enable visualizing the shape of the high-fitness nose of the DFE which is critical in determining the dynamics. In short, if the nose is log convex down a predominant fitness effect "s" exists which drives the dynamics whereas if it is log convex up, one is dominated by the largest clone to have occurred [3].



Supplementary Figure 1: (Left) Lineage trajectories from each of the four replicates (C1, C2 performed in C-lim and N1, N2 performed in N-lim) for which lineage tracking was performed colored by probability of containing an adaptive mutation. (Right) The mean fitness of cells in each experiment over time relative to the wild-type. Replicates within the same environment are consistent with one another although do show some variation, particularly at later times, due to the stochastic occurrence time of mutations [1]. Shaded region indicates approximate error on the mean fitness inference.



Supplementary Figure 2: Expansion of lineages can be used to estimate the fitness effect s (y-axis) and establishment time, τ, (x-axis) of all detectable single beneficial mutations that enter the population. Each circle is an independently occurring mutation, the size indicates its abundance in the population at generation 88. Color indicates whether it entered prior to the separation of the replicates (purple, "pre-existing") or after the separation (green, "not pre-existing"). The top panel is from previously published data in [1] and is shown here for comparison.



Supplementary Figure 3: The distribution of fitness effects of all detected single beneficial mutations arising in each of the four replicates. Rates to each fitness effect are inferred by counting the number of cells in a given fitness range over time and using a deterministic approximation to infer $\mu(s)$ (see [1]).

2 Distribution of single-mutant sizes over time and effective U_b

The arguments in this section follow similar arguments presented in the supplemental methods of [1] and [3]. Consider a mutation at a given site in the genome i with fitness advantage s, which occurs at a constant rate $R_i = NU_i$ from a pool of N ancestral feeding cells. The total number of single-mutant cells with this mutation is the convolution of the single mutant size distribution (see [1]) with the distribution of times at which the mutations enter (uniform). This yields a distribution of rescaled clone sizes $\nu = n/\tilde{n} = n/((c/s)e^{st})$:

$$\rho(\nu)d\nu = \frac{1}{\Gamma(R_i)}\frac{e^{-\nu}}{\nu^{1-R_i}}d\nu \tag{2}$$

The rate of mutating a given site is $\mu \approx 3 \times 10^{-10}$ per bp per generation [1] hence $R_i \ll 1$ and the distribution of sizes of unique SNPs is approximately $NUe^{-\nu}/\nu$ with the largest single-mutants being those that arose immediately reaching sizes $n \sim (c/s)e^{st}$ and the smallest being those that arose in the previous generation and that are of size ~ 1.

In general there are a range of possible s with different mutation rates $\mu(s)$ hence the distribution of single-mutant abundances is obtained by summing all of these:

$$\rho(n) = \frac{N}{n}\int_0^\infty \frac{\mu(s)e^{-n/\tilde{n}(s)}}{(1/\tilde{n}(s))}ds \tag{3}$$

Thus the exact form of the single-mutant distribution will depend on the shape of $\mu(s)$ and in general will not retain the characteristic $1/n$ form even at low frequencies unless $\mu(s)$ is sharply peaked over a narrow range.

Note on clonal interference. The above expression assumes no competition between lineages. This is a good approximation at early times, but for times after the single-mutant class has reached a significant fraction of the entire population clonal interference cannot be ignored. In the simplest model however, clonal interference — which results from the total population size of cells having to remain $\approx N$ — is captured by considering a genotype's fitness advantage over the (increasing) mean-fitness of the population, $\bar{x}(t)$. Thus, while the previous expression will still accurately capture relative frequencies of mutants, the absolute numbers of the different genotypes will be modified by a factor

$$\exp\left(-\int_0^t \bar{x}(u)du\right) \tag{4}$$

Consider now an entire class of mutations with a given fitness effect range $s, s+ds$ rather than a specific mutation. The mutation rates to this range of fitness effects is given by $\mu(s)ds$. Provided the population size $N\mu(s)ds \gg 1$, (true for the majority of the range we consider, but see below for when this does not hold) single mutants occur in large numbers for a given fitness range (since $N\mu(s)ds = R \gg 1$) and therefore, as a class, behave quasi-deterministically. The fraction of single-mutant cells in a given fitness range $f(s)ds$ is simply

$$f(s)ds = \mu(s)ds\frac{e^{st}-1}{s} \tag{5}$$

Since the timescale over which mutations with fitness effects s will contribute is $1/s$ and the number establishing per generation is $N\mu(s)ds$, the number of mutations contributing significantly to the genetic diversity in each fitness class is of order $N\mu(s)ds$. The small lineage size ensures that the majority of single-mutants will occur inside independent lineages, hence the number of lineages that expand at a rate between $s, s+ds$ is also $N\mu(s)ds$.

Effective fitness effects and effective mutation rates. For times $st > 1$, the differences in fitness effects between mutations enable some lineages to expand more relative to others. The range of fitness effects dominating all adaptive cells are those for whom $f(s)ds = \mu(s)ds(e^{st}-1)/s$ is maximized, which corresponds to fitness classes in the region of $\tilde{s}$, where

$$-\frac{d\log\mu}{ds}\bigg|_{\tilde{s}} = t \tag{6}$$

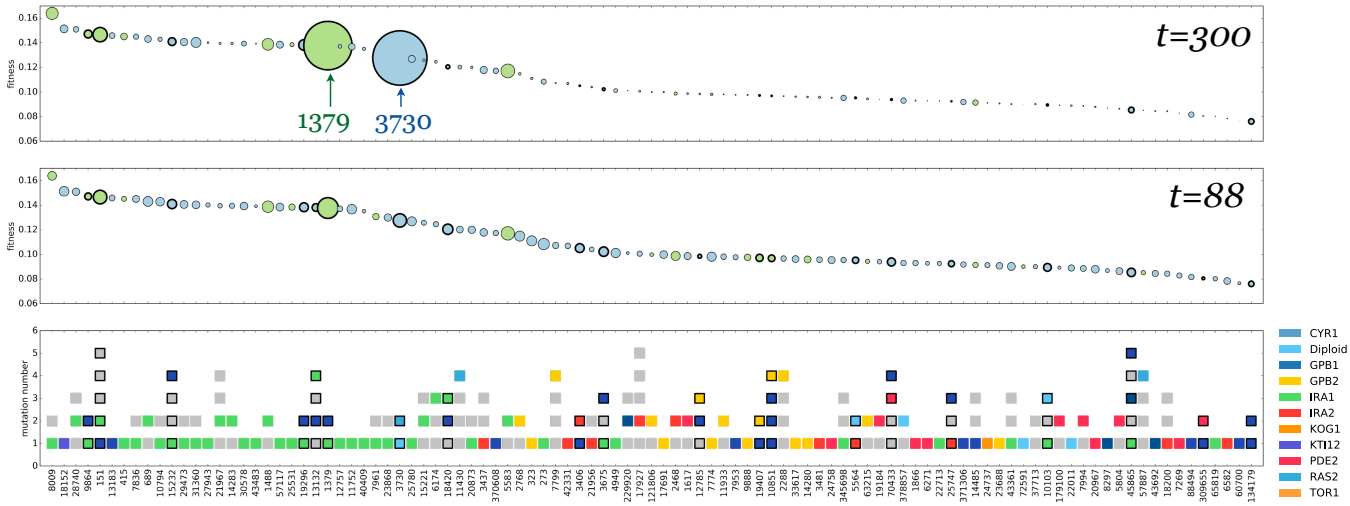
the region having a width $\delta\tilde{s}$ is given by $1/\sqrt{\partial_s^2 \log\mu}$ evaluated at $\tilde{s}$. Mutations in the fitness range $[\tilde{s}(t), \tilde{s}(t)+\delta\tilde{s}(t)]$ dominate the population at t and hence they can be considered an effective s. Similarly, the rate of mutation to these fitness effects $\tilde{U}_b = \mu(\tilde{s}(t))\delta\tilde{s}(t)$ can be considered an effective mutation rate. Together these can be used to determine the number of expanding unique single-mutants contributing to the diversity at time t:

$$\#\text{ single mutants} \sim N\mu(\tilde{s}(t))\delta\tilde{s}(t) \tag{7}$$

Stochastic occurrence of highly fit single mutants. The fitness of the predominant fitness class $\tilde{s}$ typically increases over time because the size of each fitness class has an exponential weighting e^{st}. Since $\mu(s)$ in both environments we analyze decreases with increasing s, this results in a smaller $\tilde{U}_b$ at later times. This has important consequences for stochasticity. The single mutant genetic diversity is determined by an exponential expansion of types, but where the number of types is ever decreasing (see "top-hat" DFE example). For late enough times, the predominant fitness class $\tilde{s}(t)$ with associated width $\delta\tilde{s}(t)$ has an associated mutation rate $\tilde{U}_b(t) = \mu(\tilde{s}(t))\delta\tilde{s}(t)$, which can be small enough that $N\tilde{U}_b(t) \lesssim 1$. At this point, the number of mutations contributing to the adaptive single-mutant population becomes small since the population is dominated by the small number of high fitness clones. This results in lower genetic diversity and increased stochasticity. At times $\tilde{t} \approx (1/\tilde{s}) \ln\left(\tilde{s}/\tilde{U}_b\right)$ single-mutants comprise a significant fraction of all cells. Around this time they begin to interfere with one another, slowing their exponential expansion.

3 Sequencing of clones, remeasuring fitness and "coloring" the DFE

3.1 C-lim



Supplementary Figure 4: C-lim sequenced clones. The top 100 clones ranked (left to right) in order of decreasing re-measured fitness relative to the ancestor in C-lim. Each column is a clone (picked from the population at generation 88). The top panel shows the extrapolated abundance – indicated by disk area – at generation 300, the middle panel the measured abundance – indicated by disk area – at time of sampling (generation 88) and the bottom panel shows the barcode ID, and number of mutations identified in each clone (squares) with colors indicating which gene the mutation landed in. Gray indicates "other" mutations that appear only once, a substantial fraction of which are likely neutral. Multiple mutants are denoted by bold outlines and sometimes also by a dark blue square. The dark blue square indicates an adaptive mutation that did not fall into one of the commonly mutated genes listed in the key on the right-hand side.

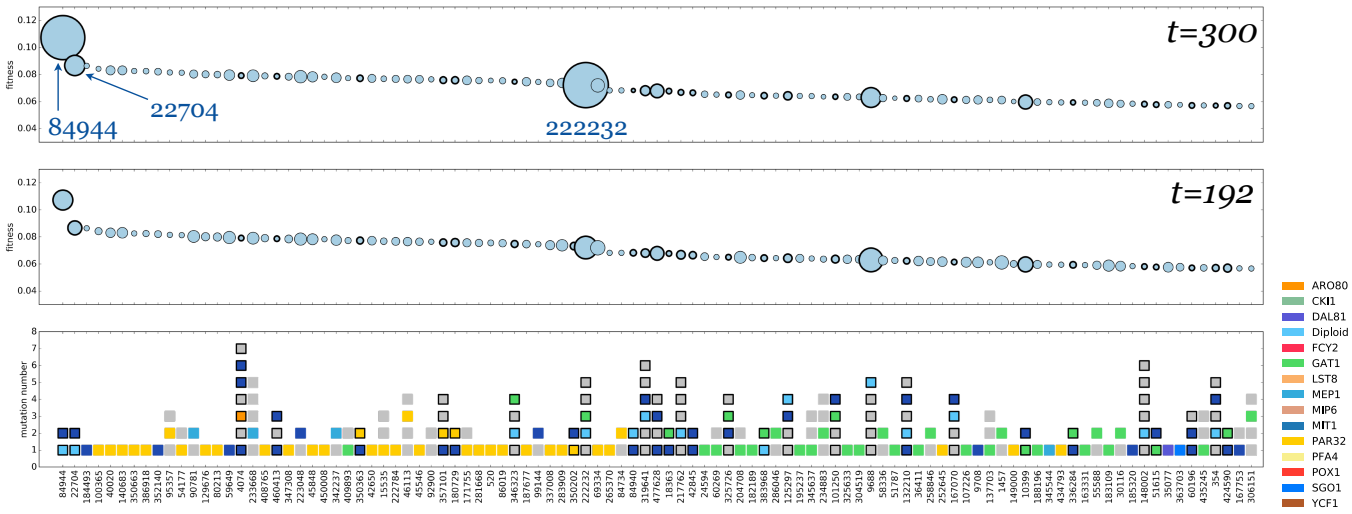
In total 475 clones were picked from the C-lim evolution (369 clones from C1, blue disks in 4, 106 clones from C2, green disks in 4). A large proportion of these clones have been described previously in [5] and we refer readers to this reference for clone picking and sequencing details. Figure 4 shows the top 100 clones ranked by re-measured fitness. Anomalously large double mutant clones are highlighted by barcode numbering on the plot.

Adaptive mutations. We used the following criteria to determine whether mutations were adaptive "driver" mutations (those causing the non-neutral fitness) or neutral passenger mutations.

1. If a gene was mutated multiple times in clones with distinct barcodes, mutations in that gene were designated as `adaptive`. A conservative estimate of the probability that two independently occurring neutral mutations in a given gene would be identified as adaptive via such an approach is small (~ 0.005), as the number of genes mutated twice in independent lineages is 26 and there are $\sim 5,000$ possible genes. This estimate is conservative because (i) it does not use functional information and the vast majority of these mutations indicate they are functional (e.g. missense, frameshift, upstream indel) and (ii) because there is often only one mutation observed in that clone, making it more likely to be the variant causing the fitness difference.

2. If a mutation in a gene was only observed once, but that clone was clearly non-neutral (mean re-measured fitness>0.01), and no other mutations were identified in the clone, then that gene was labeled as adaptive `adaptive`.

Multiple-mutants. A clone that contains two or more mutations that are identified as "adaptive" via the above criteria was classified as a multiple mutant. In replicate C1 the largest clone (barcode 3730) is a confirmed double-mutant with mutations Dip+*RAS2* . C2 is more difficult to assess as the number of clones sampled was smaller and thus a larger fraction of the high abundance barcodes at late times were not sampled at generation 88 (see Figure 24). The largest lineage in C2 — 1379 — is indeed identified as a double-mutant composed of an *IRA1* (mutated 47 times across both replicates) and YIL169C (mutated



Supplementary Figure 5: N-lim sequenced clones.The top 100 clones ranked (left to right) in order of decreasing re-measured fitness relative to the ancestor in N-lim. Each column is a clone (picked from the population at generation 192). The top panel shows the extrapolated abundance – indicated by disk area – at generation 300, the middle panel the measured abundance – indicated by disk area – at time of sampling (generation 192) and the bottom panel shows the barcode ID, and number of mutations identified in each clone (squares) with colors indicating which gene the mutation landed in. Gray indicates "other" mutations that appear only once, a substantial fraction of which are likely neutral. Multiple mutants are denoted by bold outlines and sometimes also by a dark blue square. The dark blue square indicates an adaptive mutation that did not fall into one of the commonly mutated genes listed in the key on the right-hand side.

6 times across both replicates). However a note of caution is warranted here: 5 of the 6 mutations in *YIL169C* are at exactly the same position. While this can occur by chance at the population sizes we use, it is more likely due to a neutral pre-existing mutation that arose prior to barcoding and was barcoded multiple times. This is especially likely in this case since the same SNP is sampled into both replicates (C1 and C2) yet is not adaptive in C1. Hence in C2 there is little evidence for early anomalously fit double-mutants. This lack of early double-mutants is consistent with the diploid trajectory in C2 (Figure 4C of main text, green data points) where the diploid trajectory is outcompeted for a long time and is rescued late, indicating a later double-mutant in this replicate.

3.2 N-lim

We isolated clones from generation 192 of N1, and remeasured the fitness of the 310 clones within that pool, whose trajectories indicated they were adaptive as previously described [5]. We whole genome sequenced all clones from this pool to a mean coverage of 30x. Variants were called using the same pipeline as outlined in [5]. Details of the top 100 clones ranked by re-measured fitness is shown in Figure 5.

From the whole genome sequencing data, we identified SNPs, small indels, larger deletions and insertions, Ty transposition events, and CNVs, including aneuploidy and segmental aneuploidy and annotated the genes within which those mutations fell. Among these genes, there were 39 that had more mutations than would be expected by chance including: *MEP1*, *GAT1*, *PAR32*, *FCY2*, *DAL81* and *MIT1*. In addition, we saw mutations in *MEP2* and *MEP3*. The *MEP* genes encode the ammonium permeases, suggesting that the mutations that we observe are likely to be gain of function, increasing the cell's ability to scavenge ammonium, which is the limiting nutrient. We also observe mutations in *GAT1*, which has been previously identified as mutation in nitrogen limited chemostats [6], and encodes a GATA transcription factor that activates the nitrogen catabolite repression regulon.

Adaptive mutations. Adaptive mutations were identified in the same way as in the C-lim replicates (see above).

Multiple-mutants. In the sequenced N-lim replicate, 28 of the top 100 clones ranked by fitness are double-mutants (defined as above for C-lim). Of these, two lineages go onto dominate the population (BC #84944, containing a Dip+*MEP1* double-

mutant and BC #222232, containing a Dip+*GAT1* double-mutant) both indicating Dip+GoF structure since the gene variants are beneficial as heterozygotes.

3.3 Coloring the DFE by gene / mutation type

Combining the whole genome sequencing of adaptive clones with the lineage tracking data enables one to assign which mutations in which genes contribute to various regions of the DFE. This is shown for both N-lim and C-lim environments in Figure 2 of the main text. To "color" the beneficial mutation rate spectrum, by gene like this, we divided up the fitness range ($[0, 0.15]$ in C-lim and $[0, 0.12]$ in N-lim) into bins of width $\delta x = 0.002$. For each of the replicates with large numbers of sequenced clones (C1 and N1) we then assigned each adaptive barcode to a bin if its fitness estimate (from the maximum likelihood approach using the early time trajectory in Section 1) fell within that fitness bin. For each adaptive barcode in each fitness bin, we then asked whether a clone from this barcode was whole genome sequenced and if it was sequenced, which genes were mutated.

The contribution of a given barcode lineage to the bin rate was determined by estimating the rate using the inferred fitness effect and establishment time via

$$\mu \sim \frac{1}{N} \exp\left(-s\tau\right) \tag{8}$$

We scaled the rate of each gene in a given bin up by a factor such that the rate to all genes in a given bin added up to the total rate inferred for that bin. For example if 2 barcoded lineages each with a different mutated gene contributed to a bin, but the estimates of the rates to each were in the ratio 2:1, and the total rate to the bin was 10^{-6}, then the first gene would be colored to a height of 0.67×10^{-6} and the second to 0.33×10^{-6}. Bins in which no clones were sampled for sequencing were marked as unknown. This method will work well for bins in which many unique lineages that were found to be mutated, and will work less well as the number genes contributing to a bin decreases, since in the latter case there is a good chance we have under-sampled. This effect however does not affect any subsequent results, only the coloration of the DFEs in Figure 2 of the main text. Only single-mutants, which were not pre-existing (e.g. were adaptive replicate C1 (N1) and not in C2 (N2)), and which contained mutations in verified adaptive genes were counted. Verified adaptive genes are defined in section 3.1.

Multiple mutants, i.e. lineages which contain clones that upon sequencing had 2 or more mutations in genes that were independently verified as adaptive (using the above criteria) were colored differently (dark blue) so that the (small) contribution of multiple mutant lineages to the DFE could be assessed.

3.4 Loss-of-Function vs. Gain-of-Function mutations

In the simulations described in section 5, we classify high-fitness mutations as either Loss-of-Function (LoF) or Gain-of-Function (GoF). This classification is inspired by observations that, in both environments, a large fraction of adaptive mutations disrupt gene function (putative LoF mutations) while a smaller fraction are likely to have modified the gene's function (putative GoF mutations). **It should be stressed, however, that our classification of mutations as being either LoF or GoF are not definitive, and none of our conclusions depend on assigning individual clones as harboring LoF or GoF mutations. However multiple lines of evidence do point to these two broad classes of mutation being important to the diversity dynamics**. Specifically:

- **Putative LoF mutations typically occur in negative regulators**. We defined clones as harboring LoF mutations if the adaptive mutation was identified as (i) nonsense, (ii) frameshift or (iii) a large deletion or (iv) missense but in a gene in which at least one other clone harbored a mutation of type (i)-(iii) and was also identified as adaptive. Examples in C-lim of genes which are adaptive under a LoF mutation include *IRA1, IRA2, GPB2, PDE2* [5], examples in N-lim include *PAR32, GAT1, MIT1, FCY2.*
- **Putative GoF mutations often occur in positive regulators**. We defined clones as harboring GoF mutations if the adaptive mutation was identified as (i) missense and in which no other type of mutation was found to be adaptive in that gene or, in the case of N-lim, occasionally a Ty transposition if further biology could justify it as GoF (below). In C-lim, for example, mutations in positive regulators of the *RAS* pathway are exclusively missense mutations in genes such as *RAS2, CYR1, TFS1* strongly indicating they are GoF mutations. In N-lim, mutations in *MEP1, MEP2,*

and MEP3 genes were identified as putative GoF. The MEP genes encode ammonium permeases, suggesting that the mutations that we observe are likely to be gain of function, increasing the cell's ability to scavenge ammonium, which is the limiting nutrient. Mutations in the MEP genes are either very specific missense mutations (sometimes hitting the same exact nucleotide) or Ty insertions downstream of the gene.

- **Putative LoF mutations are more common than GoF mutations**. Based on the mutations in Figure 4 of [5] these classes occur in the ratio of 66:16 or roughly 4:1. Making a quantitative assessment in N-lim is more challenging as definitely assigning LoF or GoF to a given mutation is harder. Thus we elected to use the same ratio 4:1 for N-lim simulations too.
- **Putative LoF mutations partially or totally recessive**. In C-lim, as predicted by our theory, the LoF mutations are almost never seen in diploids. There were however three clones which were sampled which did have LoF mutation in diploids. An *IRA1* homozygote (barcode ID 10103, re-measured fitness 9.6%), an *IRA2* homozygote (barcode ID 5564, re-measured fitness 10.0%) and an *IRA2* heterozygote (barcode ID 3577, re-measured fitness 2.8%). This suggests that high-fitness effect LoF mutations are recessive since the homozygotes retain their high fitness while the heterozygote does not.
- **Putative GoF mutations at least partially dominant**. In contrast to the paucity of LoF mutations in diploids, GoF mutations are observed in diploids and are exclusively seen in heterozygotes. Missense heterozygote mutations in *ACF2* (barcode ID 2039, re-measured fitness 8.3%), *RAS2* (barcode ID 3730, re-measured fitness 13.7%), *PSE1* (barcode ID 9689, re-measured fitness 5.7%), *VPS3* (barcode ID 15337, re-measured fitness 6.9%) and a Chromosomal amplification of Chr11 (barcode ID 3577, re-measured fitness 8.6%) are all clearly supplying a fitness benefit above and beyond the fitness effect of being a diploid (4%±1%).

These features become important when we outline the epistasis model. Since self-diploidization is a common mutation the distinction between LoF and GoF is likely to be important when considering double-mutants. Indeed, the high fitness double-mutants we observe are typically Dip+GoF mutations and are heterozygote for the GoF mutation since it occurred second. In the epistasis model outlined below we use this biological insight to reason that the order of acquiring mutations is important and profoundly changes the dynamics of genetic diversity. If LoF mutations were able to occur at the same rate on the diploid background as they do on the haploid, for example, the additive model predicts they would dominate the population of double-mutants.

4 Staircase model

Consider the staircase model in which beneficial mutations of (additive) fitness effect s enter at rate U into a population of N cells. In the limit of large numbers of sites, all mutations that enter the population will uniquely label a different clone (infinite sites limit). The frequencies of these clones through time are what we are interested in. Much of this can be understood by considering the median time at which new mutations arise on the background of an expanding subpopulation.

Rank-frequency distribution for arbitrary R and α. Consider a fitness class growing as $n_0 \exp(r_1 t)$ which feeds mutations at rate U that grow have a (higher) fitness r_2. The cumulative number of cell divisions the feeding class has undergone is $C(t) = (n_0/r_1)\,(e^{r_1} - 1)$ and since the probability of a de novo mutation occurring and establishing is $\sim U r_2$, the cumulative probability of the first mutation occurring by time t is

$$C(t) = 1 - \exp\left(-(n_0 U)(r_2/r_1)\,(e^{r_1} - 1)\right) = 1 - \exp\left(-(R/\alpha)\left(e^{r_1 t} - 1\right)\right) \tag{9}$$

where $R = n_0 U$ and $\alpha = r_1/r_2$. For single-mutants $n_0 = N$ while for double-mutants $n_0 = NU/s$. This generalises to mutant class q as $n_0 = Ns/(s/U)^q$ as long as $n_0 \gg 1$. If $n_0 \ll 1$, one has to account for the establishment time of the class. The above expression can be cast as a cumulative distribution over rescaled size, $\nu = n/((1/r_2)e^{r_2 t})$, of the new mutation rather than time of occurrence by using the fact that $n = (\nu/r_2)e^{r_2 t}$. This yields

$$C(\nu) = 1 - \exp\left(-(R/\alpha)\left(\nu^{-\alpha} - 1\right)\right) \tag{10}$$

The typical size of the first mutation can then be determined by asking for the characteristic ν over which the exponent decays. Equating the exponent to 1 gives:

$$\tilde{\nu} = (1 + \alpha/R)^{-1/\alpha} \tag{11}$$

Extending this to ask when the kth mutation will typically enter simply yields

$$\tilde{\nu_k} = (1 + k\alpha/R)^{-1/\alpha} \tag{12}$$

which is the result quoted in the main text.

Feeding rate and effective initial size. What is R in the context of the staircase model? Provided the feeding class is in the deterministic regime i.e. for class q, one finds $R = (Ns)/(s/U)^q$ i.e. for single-mutants $R = NU \gg 1$ while for double-mutants $R = NU^2/s \ll 1$. Generally if $Ns/(s/U)^q \gg 1$ then class q will be well approximated as deterministic and its initial feeding rate is simply $Ns/(s/U)^q$, while if $Ns/(s/U)^q \ll 1$ the initial feeding rate would need to account for the establishment time of the class. Since $R \gg 1$ for single mutants the single-mutant class behaves deterministically and the above formula for the rank-frequency can be approximated as

$$\tilde{\nu_k} = (1 + k\alpha/R)^{-1/\alpha} \sim \exp(-k/R) \qquad \text{for small} \qquad k\alpha/R \tag{13}$$

For double-mutants (or any subsequent mutant class) since $R \ll 1$, they behave stochastically the median scales as

$$\tilde{\nu_k} = (1 + k\alpha/R)^{-1/\alpha} \sim k^{-1/\alpha} \qquad \text{for large} \qquad k\alpha/R \tag{14}$$

In addition to the power-law scaling, the double-mutants exhibit large fluctuations (dominated by the fluctuations in when the first second-mutant occurs).

Number of clones contributing to a class. The above formula enables one to evaluate how many clones contribute significantly to the expansion of the class as a whole. In the $R \gg 1$ limit the only characterstic scale in the exponent is R and thus, on the order of R clones contribute equally to the expansion. For $R \ll 1$ the distribution is a power-law and the number contributing depends on the exponent of that power law. Generally $\nu_k \sim k^{1/\alpha}$ meaning the sum of all the relative frequencies is determined by $\zeta(1/\alpha)$ where ζ is the Riemann Zeta function. Table 2 enumerates how many clones contribute 50%, 90% or 99% of the total for that class, for a number of different α.

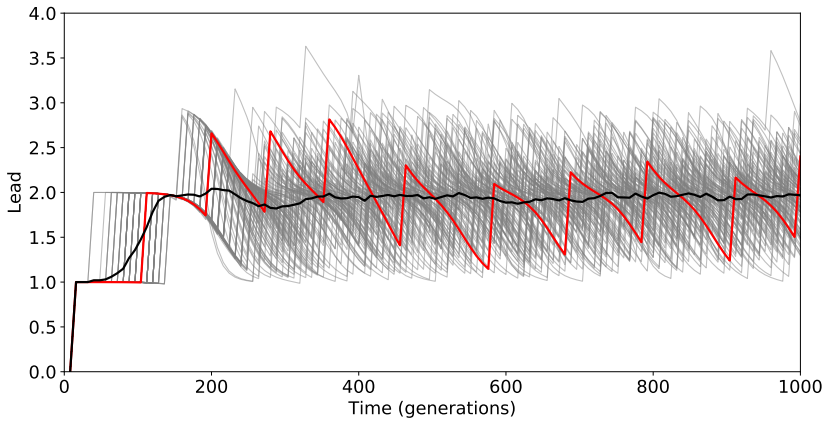
Luria-Delbruck limit. In the limit of $R \ll 1$ and $\alpha = 1$ (double mutants grow at same rate as single-mutants) one recovers the Luria-Delbruck limit in which $\nu_k \sim 1/k$. In this case the total number of clones contributing to the class is not bounded, and in principle continues to grow with time (though in practice is curtailed by the feeding population eventually being

α	# for >50%	# for >90%	# for >99%
0	1	1	1
1/4	1	1	3
1/2	1	6	61
2/3	2	59	5861
1	—	—	—

Table 2: Number of clones that comprise either 50%, 90% or 99% of the total population for that class as a function of the ratio of the fitnesses between those classes α.

outcompeted and thus shrinking). In the absence of this effect, the late clones contribute just as much to the total as the early clones since the cumulative of the $1/k$ distribution gives a $\ln(k)$ dependence i.e. that the total number in the class scales with the log of the number of mutations to have occured. Since the total number scales as $e^{r_1 t}$ there is a linear (i.e. t) term for the total, meaning each generation contributes equally to the total.

Dynamics of the lead.



Supplementary Figure 6: The lead, q, measuring the difference in fitness (in units of s) between the fittest clone in the population relative to the mean over time for 100 simulations of the staircase model (grey lines). A typical instance of the lead trajectory (red) and the mean of all trajectories (black) are highlighted.

Entropy in staircase model. One measure of the genetic diversity in the population is the Shannon entropy

$$S = -\sum_i p_i \ln(p_i) \tag{15}$$

our goal is to quantitatively understand the statistical and dynamical behaviour of the entropy. Considering Figure 3D of the main text there are a number of key features which can be quantitatively understood:

(1) **The timing of the maximum diversity.** The large peak in diversity is caused by the large number of expanding single mutants that enter and expand almost contemporaneously. We can quantitatively understand this by realising that those single-mutants destined to establish enter at a constant rate $NU_b s$ and expand exponentially at rate s. Thus, the density of single-mutant clones at frequency f is $\rho(f) \sim NU_b s(df/f)\exp(-f/\tilde{f})$ where $\tilde{f} = (1/Ns)e^{st}$. The total number of single-mutant clones that contribute significantly to the total number of single-mutant cells is $\sim NU_b \gg 1$. Once the single-mutants reach a total population size on the order of $\sim N$ the wild-type population declines rapidly and the fitness advantage of new mutations is diminished, causing both the rate of establishing new mutations and their fitness advantage to decline and eventually reach zero. This therefore sets the maximum size the first mutants will get to and thus, the timing and magnitude of the maximum in diversity. Using that the single-mutant class as a whole

expands as $n_1 \approx (NU_b/s)e^{st}$ we find that the maximum in genetic diversity should occur when

$$\frac{NU_b}{s}e^{st} \approx N \qquad \text{or} \qquad t = (1/s)\ln(s/U_b) \tag{16}$$

which agrees quantitatively with the simulations (Figure 3D, main text).

(2) **The timing of the diversity crash.** Timing of the diversity crash comes from when the handful of double-mutants outcompete single-mutants. Setting $n_2(t) \approx n_1(t)$ gives:

$$\left(\frac{2NU_b^2}{s}\right)^2 \frac{e^{2st}}{2s} = \left(\frac{NU_b}{s}\right)\frac{e^{st}}{s} \tag{17}$$

$$t \approx \frac{1}{s}\ln\left(\frac{(s/U_b)^3}{2Ns}\right) \tag{18}$$

which is also in quantitative agreement with the simulations.

(3) **The maximum value of entropy before the crash.** The maximum value the entropy can attain can be understood simply for the fact that there will be $\sim NU_b$ single-mutants all of roughly equal size contributing to the expansion of the single mutant class. Hence their contribution to the entropy once single-mutants constitute the majority of the population will be

$$S_{max} \approx \ln(NU_b). \tag{19}$$

This is also in quantitative agreement with simulations.

(4) **The average level of diversity (i.e. entropy) in the steady-state**

In the steady-state the clone sizes are power-law distributed as $\rho(f) \sim f^{\gamma}$ where $(2q-1)/q < \gamma < (2q+1)/(q+1)$ and where q is the lead i.e. the first fitness class for which $Ns(U/s)^k$ which in our case is $q \sim 2$.

If the feeding mutant is growing as $n \sim n_0 \exp((q-1)st)$ and feeds new mutations that grow at rate qs, then one can calculate the entropy by considering the typical time at which mutations enter. The ith mutant will typically occur when

$$Uqs\int_0^{t_i} n(t)dt \sim i \tag{20}$$

and therefore the size it reaches by time t will be

$$n_i'(t) \approx \frac{e^{qs(t-t_i)}}{qs} \tag{21}$$

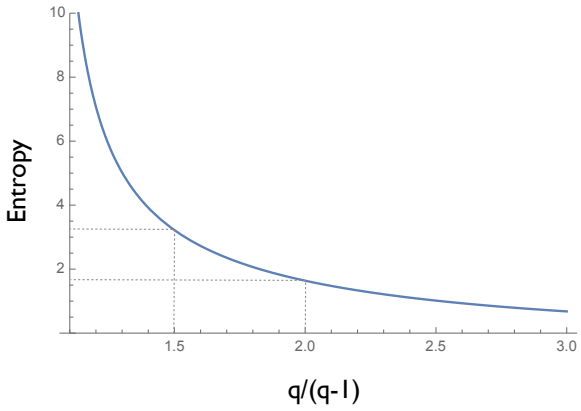
Substituting in for t_i we find that the size of the ith mutant to occur, relative to the first one to occur should follow

$$n_i \sim i^{-q/(q-1)} \tag{22}$$

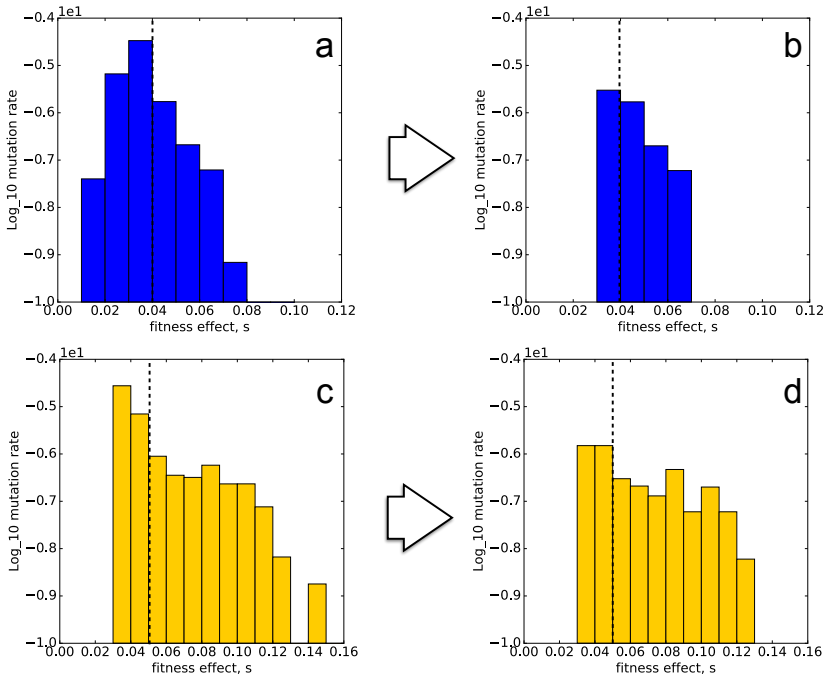
The entropy $S = \sum_i p_i \ln(1/p_i)$ of this distribution is then be calculated as

$$S = \ln(\zeta(q/q-1)) + \frac{q}{q-1}\frac{\zeta'(q/q-1)}{\zeta(q/q-1)} \tag{23}$$

where ζ is the Riemann zeta function. This expression is plotted in Figure 7. For leads $2 < q < 3$ (those applicable to the steady state for the parameters in our experiment) this produces an entropy level of $\approx 2-3$ which agrees with the mean levels observed in Figure 3D of the main text. Double mutants are formed when the lead $q = 2$, which is consistently smaller than the lead of mutants formed in the steady state. Because of this, the entropy expected for double mutants, as a class, is lower than expected for the further multiple-mutants (Figure S6).



Supplementary Figure 7: Entropy predicted by the size distribution of 'typical' mutants i.e. without stochastic fluctuations, as a function of the relative lead $q/(q-1)$. Double mutants (for whom $q/(q-1) = 2$ have a lower entropy ($S \approx 2$)than multiple mutants ($S \approx 3$) since they are formed when the lead is smaller (Figure S6)



Supplementary Figure 8: mDFEs. (a) The inferred mDFE in N-lim from all adaptive lineages and (b) the mDFE resulting from excluding known multiple mutant lineages (from sequencing of clones) and revising the rate to Dip mutations (those below dashed line). (c) and (d) the equivalent plots for the C-lim environment. Note the logarithmic scale on the y-axis.

5 Simulated lineage dynamics with inferred mDFEs under 3 models

Here we perform similar simulations using more realistic mDFEs: those directly inferred in both N-lim and C-lim. We show that the insights from simple mDFEs hold for these more realistic mDFEs: single-mutants drive an expansion in genetic diversity that subsequently crashes due to the emergence of double-mutants.

To simulate the genotype and lineage dynamics with measured single-mutant mDFEs, we started with the mDFEs inferred in each environment (outlined in Section 1) and made the following changes:

1. Lineages that contained multiple-mutants verified via sequencing of clones (see Section 3) were removed from the mDFE
2. The de-novo rate of diploidization (fitness classes below the dashed line in Figure 8) used was 3×10^{-6} per generation. In our previous work [5], we discovered that the large class of mutants in the 4% range in C-lim and 3% range in N-lim were diploids. Based on the high initial fraction ($\sim 1\%$) of the population which harbored a diploid mutation, it was also clear that many of these must have been induced during the transformation of barcodes rather than arising de-novo. To estimate the rate of de-novo diploidization we therefore asked what de-novo rate was most consistent with the measured diploid trajectories (section 6.1). We found that a de-novo rate of $\sim 10^{-6}$ was consistent with the rescue time for diploid trajectories. This rate is broadly consistent with the measured the fraction of diploids present in the population in the same experimental conditions but which did not undergo a barcode transformation. After ~ 120 generations of growth an estimated fraction of $\sim 10\% \pm 3\%$ of cells were diploid (Lucas Herrissant, *unpublished data*). Using the fact that this class will behave deterministically, the rate can be estimated using

$$f(t) \approx (U/s)e^{st} \tag{24}$$

resulting in an estimated rate between $3 \times 10^{-6} - 3 \times 10^{-5}$ per generation. Uncertainties in this number should be stressed too: the exact number of diploids depends on their growth advantage in the growth conditions prior to the serial batch transfer, which are not known.

The results of these two adjustments are shown in Figure 8 (a → b for N-lim and c → d for C-lim). The mutations below the dashed line are all diploidizations (Dip). In both C-lim and N-lim loss-of-function (LoF) mutations occur at a higher rate than gain-of-function (GoF) mutations. As such for our caricature mDFEs we assumed for simplicity that each fitness class above the dashed line has both LoF and GoF mutations which occur with relative rates of 4:1. This is broadly consistent with the number of LoF to GoF mutations seen in each environment (see Figure 4 of [5]).

Simulation details. Simulations were performed using a custom written python code (available on request) that simulated the fate of mutations and lineages and plotted the associated diversity measures (entropy) through time. The simulation closely follows the experimental procedure, briefly:

- A population of $N = 5 \times 10^8$ "cells" are barcoded with $L = 500,000$ tags. The initial population of cells that gets barcoded contains 99% ancestor and 1% Dip. The distribution of initial abundances is Poisson distributed with a mean of 1000.
- Each lineage contains a dynamically updated number of genotypes (unique sets of independently occurring mutations), each of which is initially either the ancestor (WT) or the ancestor + a diploidization (WT+Dip).
- Genotypes give rise to further genotypes via acquiring mutations drawn from the mDFE. The acquisition of new mutations occurs stochastically and in two steps (i) a mutation is generated (ii) a stochastic variable determines whether it establishes, and if it does, the new genotype created starts at establishment size (1/fitness advantage over the mean).
- Each generation genotypes increase in abundance according to their fitness advantage over the "mean fitness" (calculated each generation). Genetic drift (arising from small number fluctuations) is modelled implicitly in the establishment process and after this (i.e. at frequencies large enough that drift is irrelevant) lineages are modeled deterministically. Initially for example a (WT+Dip) genotype would have a fitness advantage of 4% in C-lim or 3% in N-lim since this is the fitness effect of Dip in each environment.
- The abundance of each lineage is simply the sum of the abundances of all genotypes within that lineage (which can be multiple). Genotype and lineage abundances are recorded every 8 generations.

5.1 Single-mutant model

In the single-mutant-model simulations, any cell that has acquired a mutation can no longer mutate and thus only single-mutant genotypes exist and competition between these determines the entire lineage and genotype dynamics. In both environments (C-lim, Figure 9; N-lim, 10) this model produces lineage and diversity dynamics that are not in agreement with observations. The single-mutant model predicts too many lineages that remain at intermediate-high frequencies at the later time points (3rd row, "Muller" plot) as compared with the experimental lineage data. This produces elevated levels of diversity (as measured via the entropy of all adaptive lineages, 4th row) again above observations (Figure 3 of main text).

As expected most of the adaptation in the single-mutant model is driven by the high-fitness LoF and GoF mutations (light and dark green lines in 2nd row of both Figure 9 and 10)

There are a number of reasons this model is not accurately capturing the correct lineage and diversity dynamics. First, is that the whole genome sequencing of clones gives direct evidence that very fit double mutants are the most abundant genotypes in the population at late times in both C-lim and N-lim. Second, the adaptive lineage dynamics do not match the observed ones: competition between single mutants alone is not enough to explain the crash in genetic diversity observed across both environemtns.

5.2 Additive model

In the additive simulations, cells can continue to stochastically acquire multiple adaptive mutations (waves of different colors in row 2 of Figures 11 and 12 (color key on top of plot). New mutations are drawn from the same unchanging mDFE (i.e. the mDFE is independent of genotype) and fitness effects of mutations combine additively. This model would be appropriate if epistasis were rare, and cells were able to acquire many different LoF and GoF mutations that do not interact.

For the C-lim simulations (Figure 11) this model also produces lineage and diversity dynamics that are not in agreement with observations. First, the lineage dynamics (Figure 11) predicted by such a model are qualitatively different than those observed, whereby at intermediate-late times lineages are predicted to undergo large-scale fluctuations over long timescales

(∼50 generations) which we never observe in the experimental data (see columns b, d and e in Figure 11 as examples). These lineage dynamics produce trajectories of the adaptive lineage entropy which are in disagreement with observations: diversity either does not crash as dramatically and remains high even out to 300 generations as the competition between a few fit lineages continues, or it crashed too dramatically due to a very early LoF+LoF mutant (columns a and c in Figure 11 show examples). While these look qualitatively similar to the crashed observed in the data (Figure 1 of main text, and entropy plots in Figure 3 of main text) they are typically sharper: the entropy falling off more rapidly between 150-200 generations than is observed in experiment.

Another key difference between the predictions of the additive model and the observations comes from comparing the types of clones that should dominate the population at intermediate and late times. The additive model predicts that the dominant double-mutant clones over generations 150 - 300 in the 50 simulations performed are LoF+LoF (18/50), Dip+LoF (14/50), early triple mutants (11/50) GoF+LoF (4/50) or Dip+GoF (3/50). This is at odds with the sequencing of clones where we observe Dip+GoF double-mutants to be the most dominant. Furthermore, this additive model predicts that diploid trajectories (see main text Figure 4 and supplementary section 6.1) will undergo a double-dip whereby the rescue, driven by Dip+LoF clones will be outcompeted by LoF+LoF clones, which also disagrees with our data.

For N-lim simulations (Figure 12) using the additive model produces lineage dynamics which are qualitatively similar to those observed. The entropy adaptive lineage trajectories predicted by this model (see 4th row of Figure 12) are also in reasonable agreement with observed trajectories. However model predictions are in disagreement with observations in two important ways. First, the dominant intermediate/late time genotype (the one driving the diversity crash) is predicted to be Dip+LoF, which is not observed in our data. Second, the more modest fitness advantage of the haploid LoF and GoF mutations in N-lim coupled with the the expansion of Dip+LoF mutants, predict that the diploid trajectories in N-lim (see Figure 4 main text) should sweep to fixation almost every time with no dip. This is again in disagreement with data.

5.3 Epsitasis model

The whole genome sequencing of hundred of adaptive clones (see Section 3) and the Diploid trajectories (see Section 6.1 and main text Figure 4) produced a number of somewhat surprising results including:

- In both environments no example of a double-mutant composed of two fit single mutant SNPs was found e.g. (LoF+LoF) or (LoF+GoF) or (GoF+LoF) or (GoF + GoF).
- The double mutants observed (and that dominate the population at late times) were either Dip+GoF (and thus heterozygote for the GoF mutation since it occurred after diploidization) or (LoF+Dip) (and thus homozygote for the LoF mutation since it preceded diploidization).
- Diploid abundance in all replicate evolutions undergoes an expansion-contraction-expansion dynamics.
- Cells do not undergo further genome duplications (e.g. to become tetraploid).
- The fitnesses (those from the original lineage tracking data and the fitness re-measurements) of double mutants (WT+Dip+GoF) and (WT+LoF+Dip) combine approximately additively

Together, these observations suggested a simple model for epistasis in which, while there are 3 classes of single mutants (Dip, LoF, GoF) there are only three possible double mutants as shown Figure 4A of the main text. The key features of the epistasis model we use are:

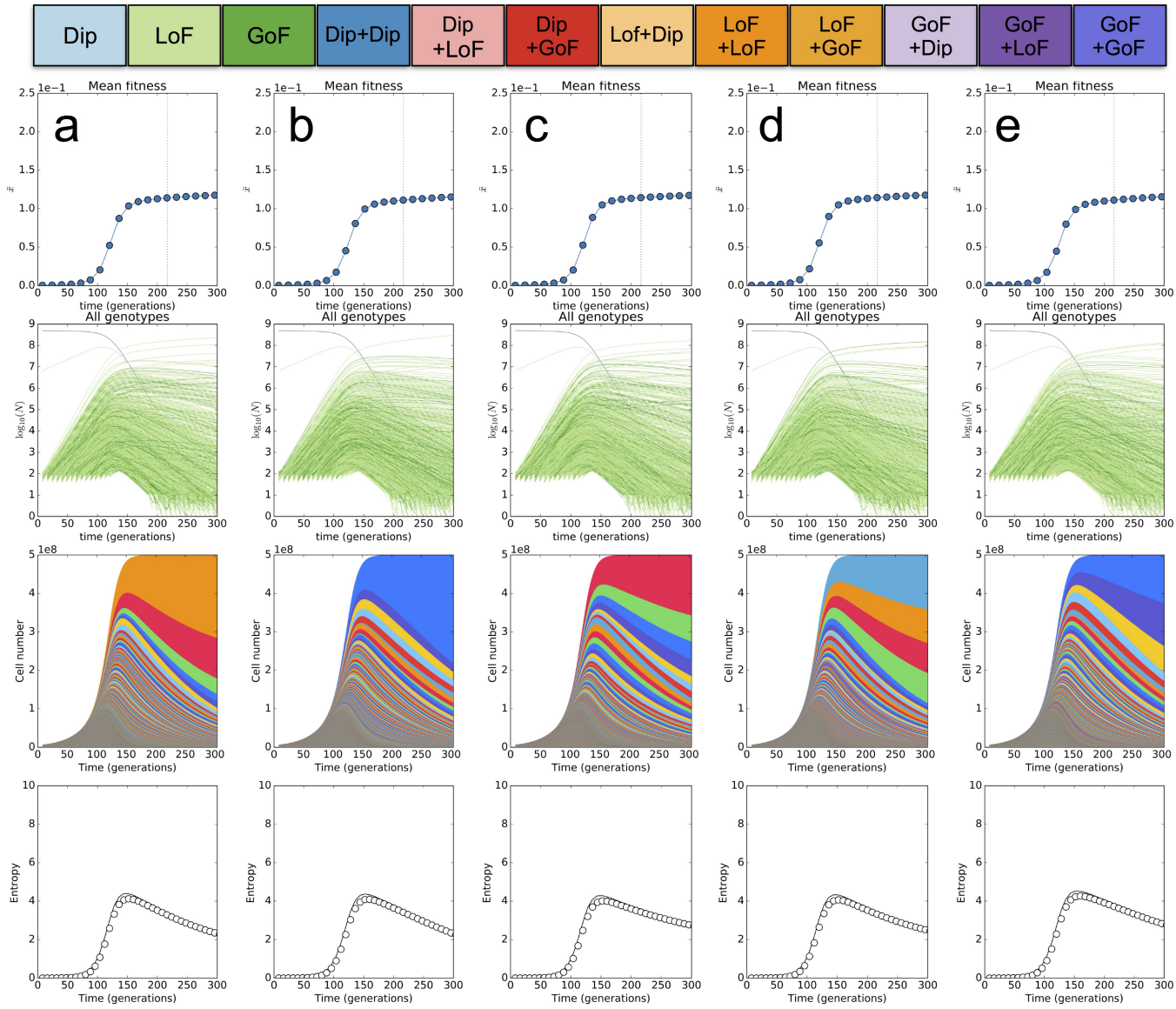
- `if` first mutation is Dip `then` second mutations can only be GoF drawn from rates to GoF mutations in the mDFE for that environment ($\sim 10^{-8}$ per generation).
- `if` first mutation is LoF `then` second mutations can only be Dip which occurs at rate of 10^{-6}/gen independent of LoF mutation
- `if` first mutation is GoF `then` second mutations can only be Dip which occurs at rate of 10^{-6}/gen independent of GoF mutation

The predictions for this model from 5 replicate simulations are shown in Figure 13 (C-lim) and 14 N-lim. For C-lim simulations the epistasis model produces lineage dynamics and diversity trajectories (adaptive lineage entropy trajectories) in close agreement with observations. While there remains considerable stochasticity, the dominance of a handful of double-mutant genotypes (red trajectories, 2nd row Figure 13) and corresponding dominance of a handful of lineages (3rd row) is

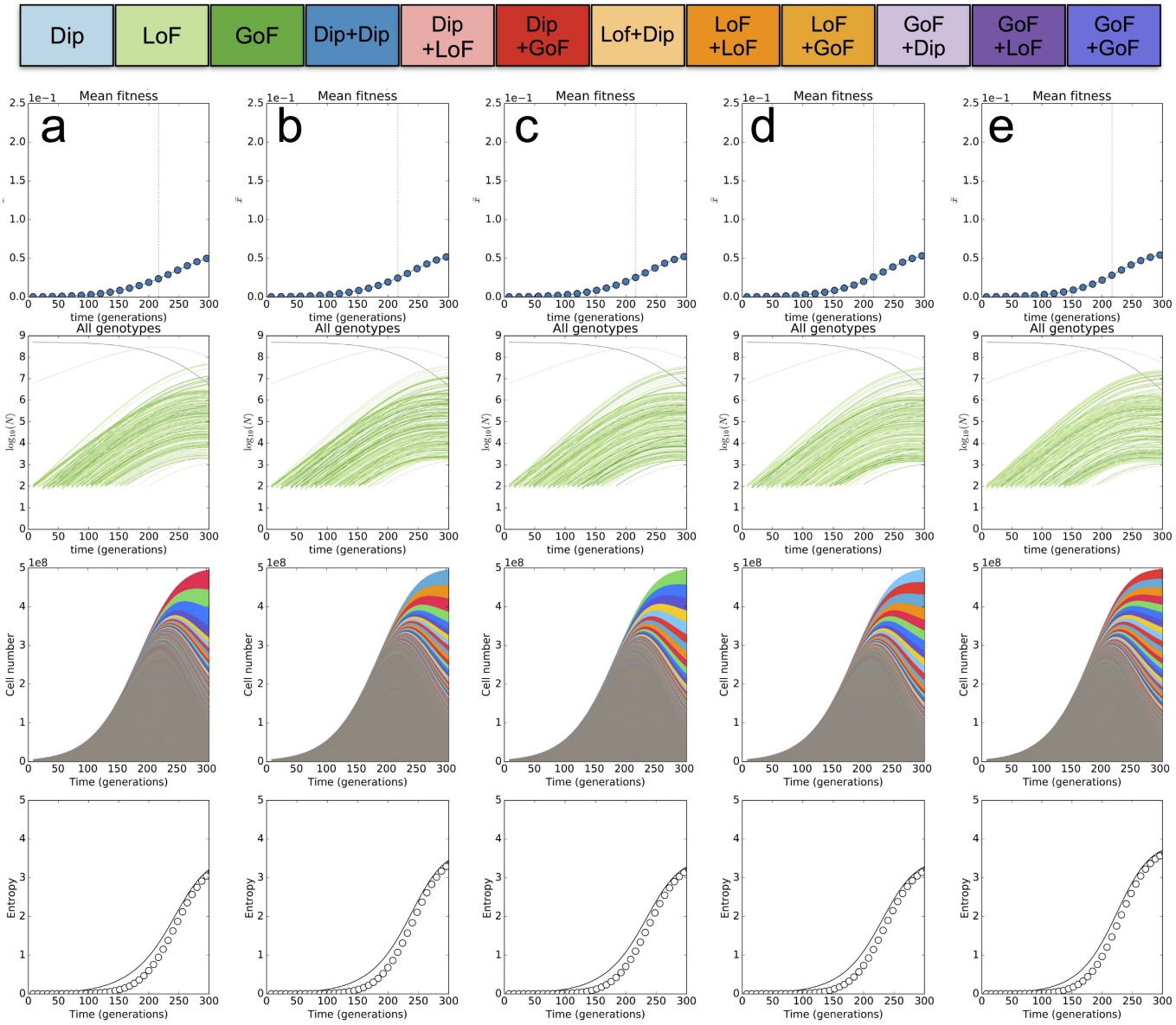
consistent with data. Importantly, this model also predicts that overly dominant double-mutants are more likely to be caused by (Dip+GoF) and that these are most often the cause of the diversity crash.

For N-lim simulations, lineage trajectories alone show less clear evidence that the epistasis model agrees more with observations compared to the additive model. In fact, the epistasis model produces similar lineage trajectories and slightly slower entropy expansion than is observed. However the advantage of the epistasis model is that, while it produces similar lineage dynamics, it predicts that the specific genotypes that dominate the population will be (Dip+GoF) rather than (Dip + LoF) predicted by the multiple-mutant model.

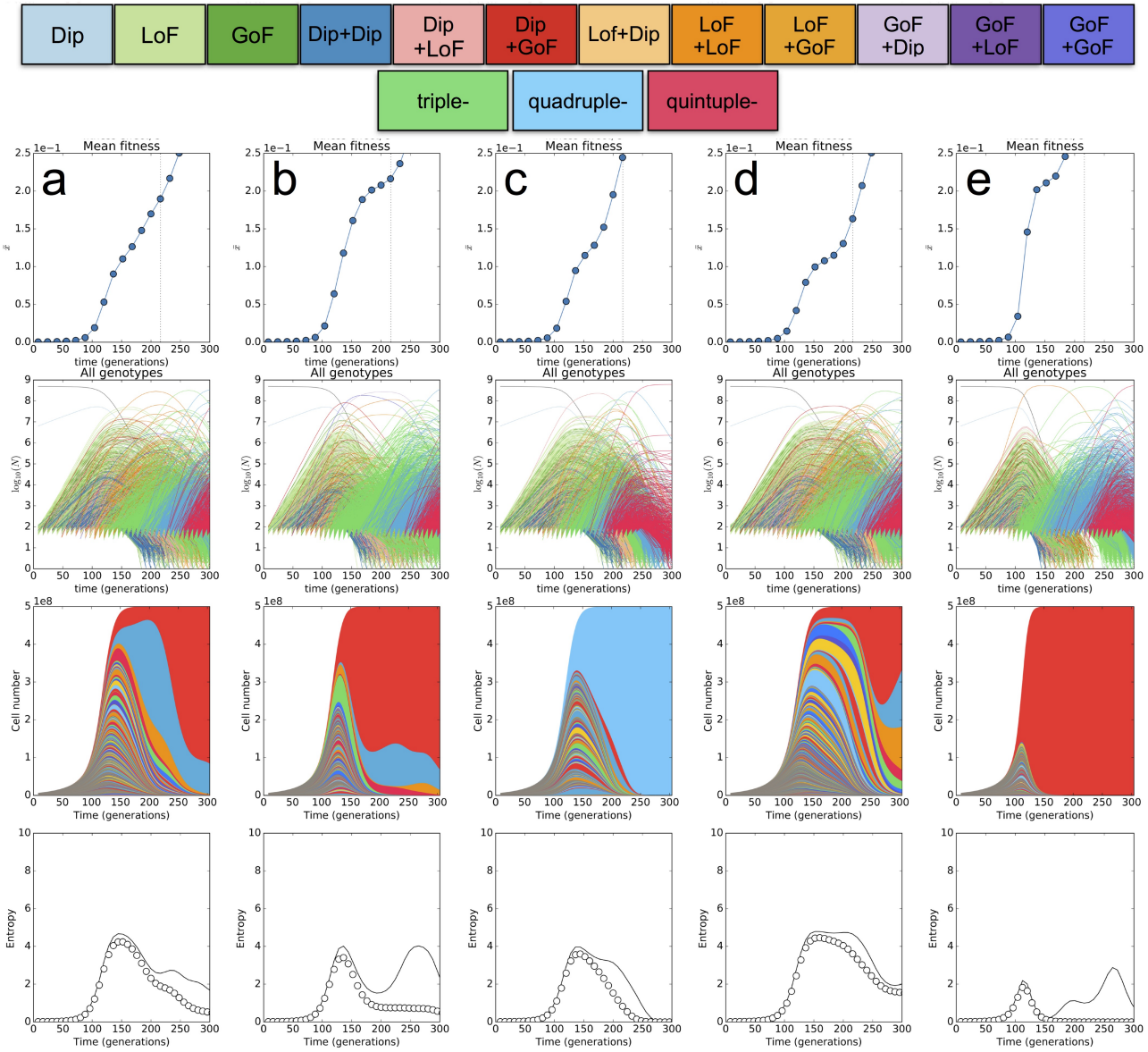
The possibility of further LoF and GoF mutations. This epistasis model shown in Figures 9 - 14 does not permit the the acquisition of more than two mutations. This is clearly unnatural since it is inevitable that further mutations enter at later times and expand. However this simple model appears sufficient to explain the dynamics we observe in the early stages of the experimental evolutions we performed. This suggests that single-mutants with highly fit LoF or GoF mutations have (i) fewer high fitness mutations available to them or (ii) that the fitness effects of new mutations is significantly reduced or (iii) a combination of these. Studying the emergence of multiple-mutants beyond double-mutants, however, and their impact on genetic diversity will likely require re-barcoding techniques as discussed and referenced in the main text as well as longer term whole-population sequencing efforts as in [7, 8].



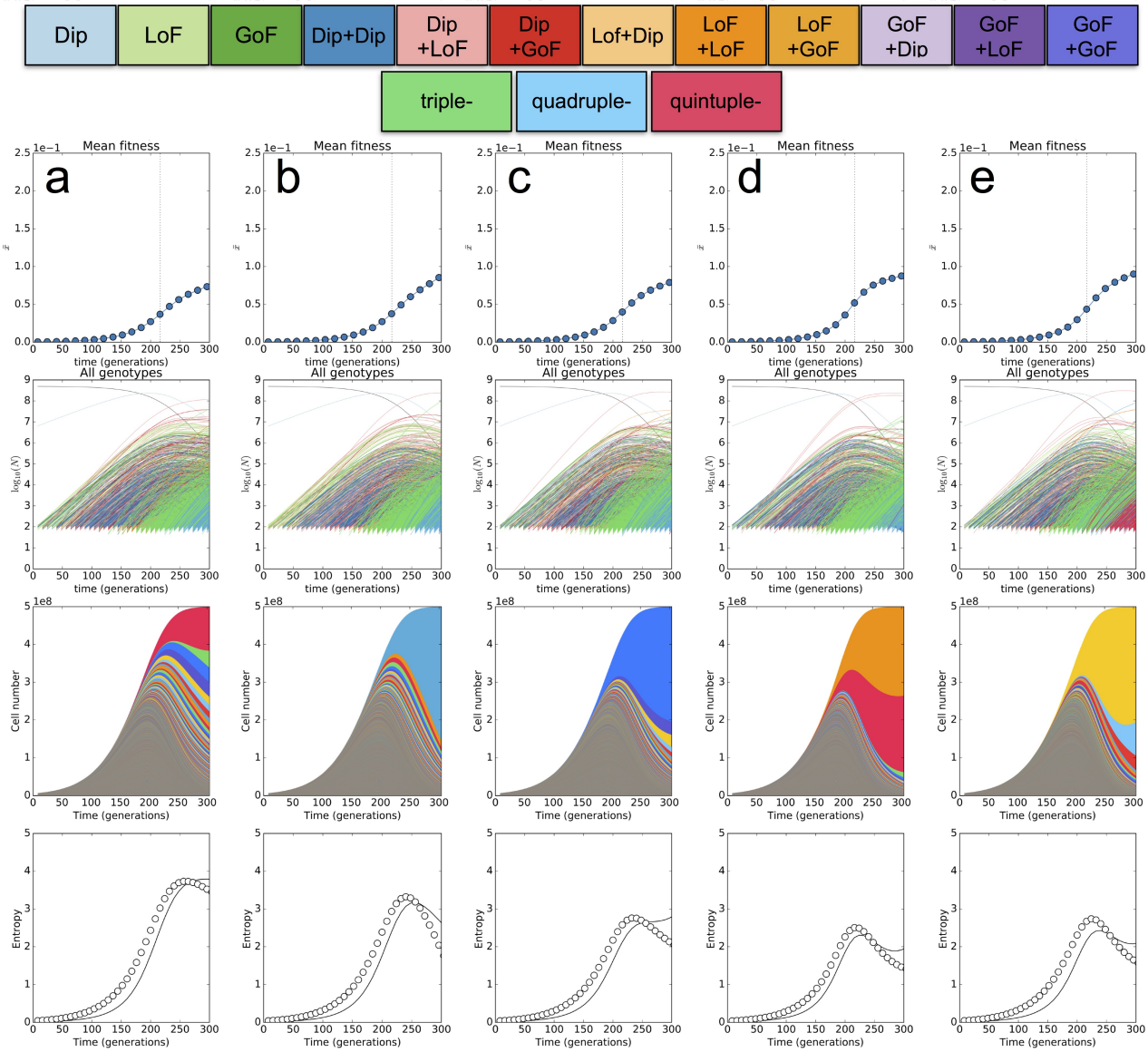
Supplementary Figure 9: C-lim Single-mutants. (a–e) 5 randomly sampled simulated replicates of the dynamics in C-lim under a single-mutant only model. Each column is a replicate and shows the mean fitness of cells relative to the ancestor over time (1st row), the abundance of every genotype in the population through time on a log-scale (2nd row, genotypes colored according to color-key on top of figure), the abundance of each adaptive lineage (different colors are here chosen arbitrarily for visualization purposes) ranked top to bottom by lineage size at generation 300 (third row) and the entropy of all adaptive lineages through time (4th row) as measured via all genotypes (solid line) and via adaptive lineages (data points).



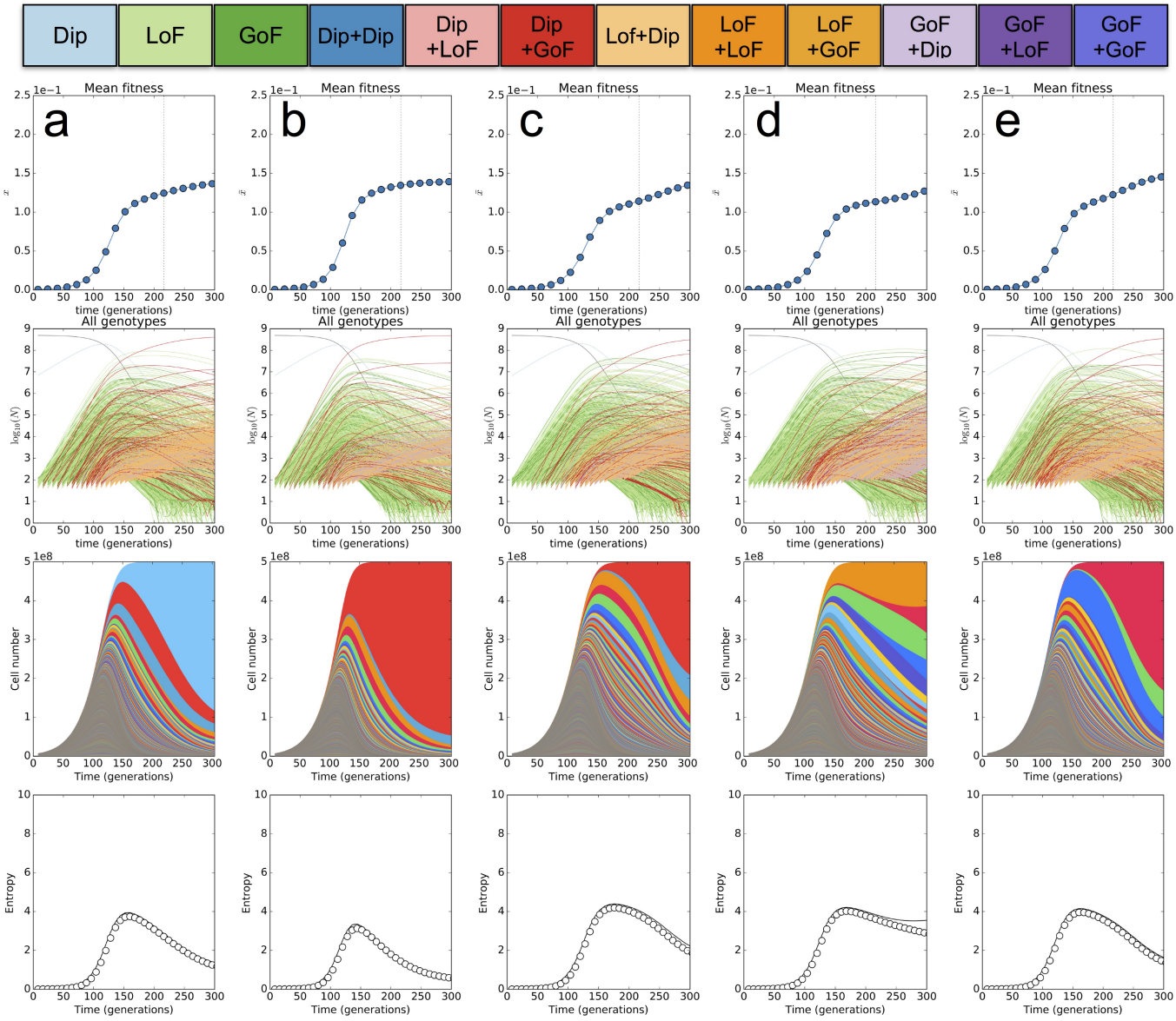
Supplementary Figure 10: N-lim Single-mutants. (a–e) 5 randomly sampled simulated replicates of the dynamics in N-lim under a single-mutant only model. Each column is a replicate and shows the mean fitness of cells relative to the ancestor over time (1st row), the abundance of every genotype in the population through time on a log-scale (2nd row, genotypes colored according to color-key on top of figure), the abundance of each adaptive lineage (different colors are here chosen arbitrarily for visualization purposes) ranked top to bottom by lineage size at generation 300 (third row) and the entropy of all adaptive lineages through time (4th row) as measured via all genotypes (solid line) and via adaptive lineages (data points).



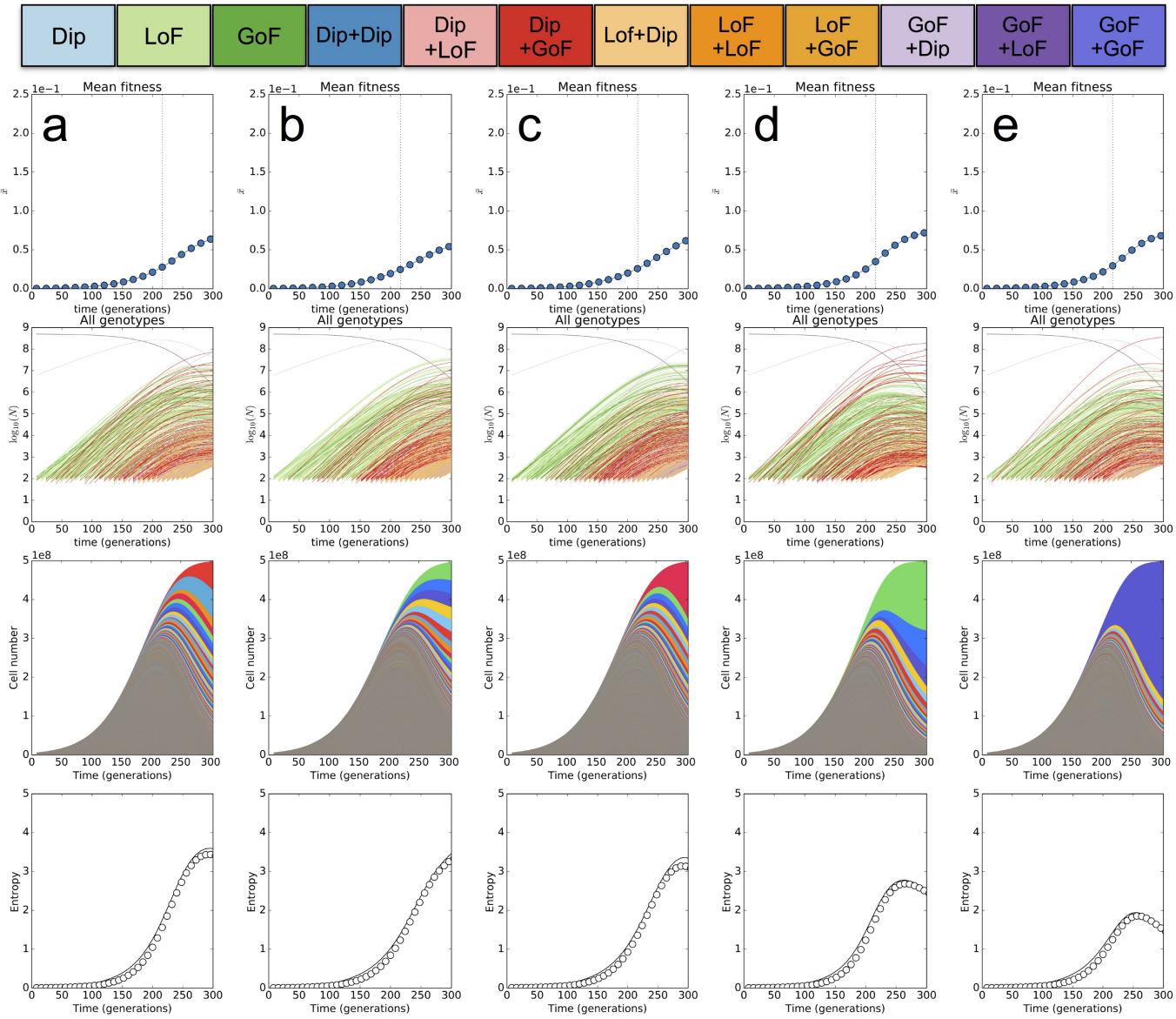
Supplementary Figure 11: C-lim Multiple-mutants. (a–e) 5 randomly sampled simulated replicates of the dynamics in C-lim under the additive model. Each column is a replicate simulation and shows the mean fitness of cells relative to the ancestor over time (1st row), the abundance of every genotype in the population through time on a log-scale (2nd row, genotypes colored according to color-key on top of figure), the abundance of each adaptive lineage (different colors are here chosen arbitrarily for visualization purposes) ranked top to bottom by lineage size at generation 300 (third row) and the entropy of all adaptive lineages through time (4th row) as measured via all genotypes (solid line) and via adaptive lineages (data points).



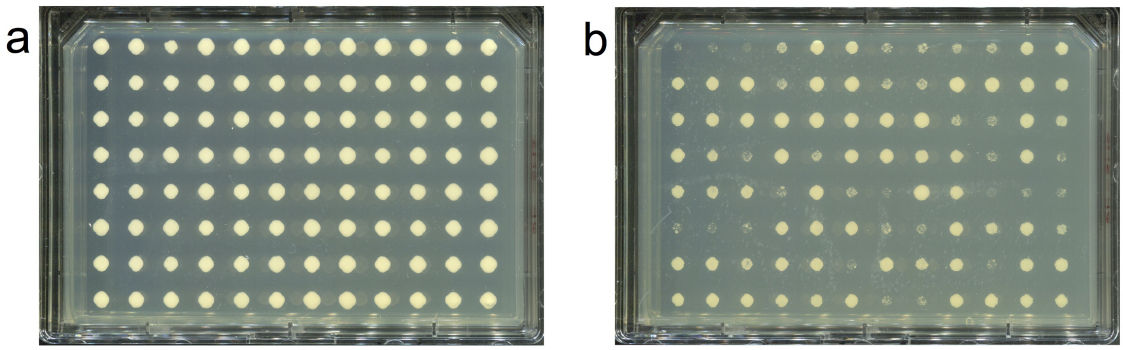
Supplementary Figure 12: N-lim Multiple-mutants. (a–e) 5 randomly sampled simulated replicates of the dynamics in N-lim under the multiple-mutant no epistasis model. Each column is a replicate and shows the mean fitness of cells relative to the ancestor over time (1st row), the abundance of every genotype in the population through time on a log-scale (2nd row, grey is the ancestor, blue are single mutants), the abundance of each adaptive lineage (different colors) ranked top to bottom by lineage size at generation 300 (third row) and the entropy of all adaptive lineages through time (4th row) as measured via all genotypes (solid line) and via adaptive lineages (data points).



Supplementary Figure 13: C-lim Epistasis model. (a–e) 5 randomly sampled simulated replicates of the dynamics in C-lim under the epistasis model. Each column is a replicate and shows the mean fitness of cells relative to the ancestor over time (1st row), the abundance of every genotype in the population through time on a log-scale (2nd row, grey is the ancestor, blue are single mutants), the abundance of each adaptive lineage (different colors) ranked top to bottom by lineage size at generation 300 (third row) and the entropy of all adaptive lineages through time (4th row) as measured via all genotypes (solid line) and via adaptive lineages (data points).



Supplementary Figure 14: N-lim Epistasis model. (a–e) 5 randomly sampled simulated replicates of the dynamics in N-lim under the epistasis model. Each column is a replicate and shows the mean fitness of cells relative to the ancestor over time (1st row), the abundance of every genotype in the population through time on a log-scale (2nd row, grey is the ancestor, blue are single mutants), the abundance of each adaptive lineage (different colors) ranked top to bottom by lineage size at generation 300 (third row) and the entropy of all adaptive lineages through time (4th row) as measured via all genotypes (solid line) and via adaptive lineages (data points).



Supplementary Figure 15: Benomyl assay for quantifying diploid abundances. Example plates showing the growth of 96 randomly sampled colonies from the population on a control plate (YPD) (a) and on a plate containing benomyl (YPD+Benomyl) (b) where the growth of diploids is inhibited.

6 Diploid trajectories

6.1 Measuring diploid abundance through time

We suspended cells (from either the carbon or nitrogen limited evolution) from a frozen stock into 1mL of water. We used a coulter counter (Beckman Coulter: 6605700) to quantify cell concentration, then used 5-8 glass beads to spread $\sim$ 400 cells onto Nunc plates (ThermoFisher Cat #: 267060) containing YPD agar (10g yeast extract Fisher: 212750, 20g peptone Fisher: 211677, 20g dextrose Sigma Aldrich: G8270-5KG, 20g agar Fisher: 214010, bring up to 1L with DO water) supplemented with 2x G418 (8mL of 50mg/mL stock into 1L YPD) and allowed them to grow for two days at 30°C. Then we scanned plates at 600 dpi resolution and inputted the resulting images into a custom written MATLAB code (available on request), which infers coordinates for individual colonies based on the scanned image. We then used a ROTOR STINGER robot (Singer Instruments) to pick colonies into 96 well plates (ThermoFisher: 265301) containing YPD with 2x G418. These plates were placed in a 30°C incubator and grown overnight. We next used the ROTOR STINGER robot to replicate the 96 well plates onto a square plate with Benomyl media, and control plate with DMSO (Sigma Aldrich: 472301) media. The Benomyl media was made from a 10mg/ml stock of Benomyl (Sigma Adrich Cat #: 381586-5G) in DMSO by adding 2 mL of stock Benomyl drop-wise while stirring to one liter of YPD that had been allowed to cool to about 55°C. The final concentration of the Benomyl media was 10 μg/ml. The cells replicated onto Benomyl and DMSO plates were grown up for 2 days at 30°C. After 48 hours growth, each plate was scanned. Haploids are able to grow better on Benomyl media while diploids grow more slowly. For each colony that grew on the control plate, we checked the corresponding positions on Benomyl plate and counted how many grew poorly (see Figure 15). Diploid frequencies were then calculated as the fraction of those colonies that grew poorly.

6.2 Predicted diploid trajectories under the 3 models

To track diploid trajectories, one does not need lineage information. Thus, diploid trajectory simulations were implemented independently since they are far more efficient without the large number of lineage labels that are required in the simulations outlined in Section 5.

The frequency trajectory of diploids in the population though time is then calculated using the following stochastic simulation:

- The initial abundance of diploids is set at 1%, based on direct measurement from the colony growth assay.
- The initial abundance of fitter LoF and GoF mutations is generated stochastically by realizing that their effective starting number will be roughly Poisson distributed with mean $n_0 = NU_{LoF}$ or $n_0 = NU_{GoF}$ which then grow deterministically through time as $n_0 e^{st}/s$. (parameter values in the table below). Drift for the single mutants can be safely ignored since n_0 for each of the three classes is a large parameter i.e $NU_{LoF} \gg 1$ and $NU_{GoF} \gg 1$.
- These three expanding classes can each acquire further mutations whose mutation rates and fitness effects are outlined in the tables below. These double mutants change in frequency between time points by generating a Poisson sample the

new expected number of cells and normalizing by N in order to include the stochasticity of drift: which is crucial for the double-mutants.

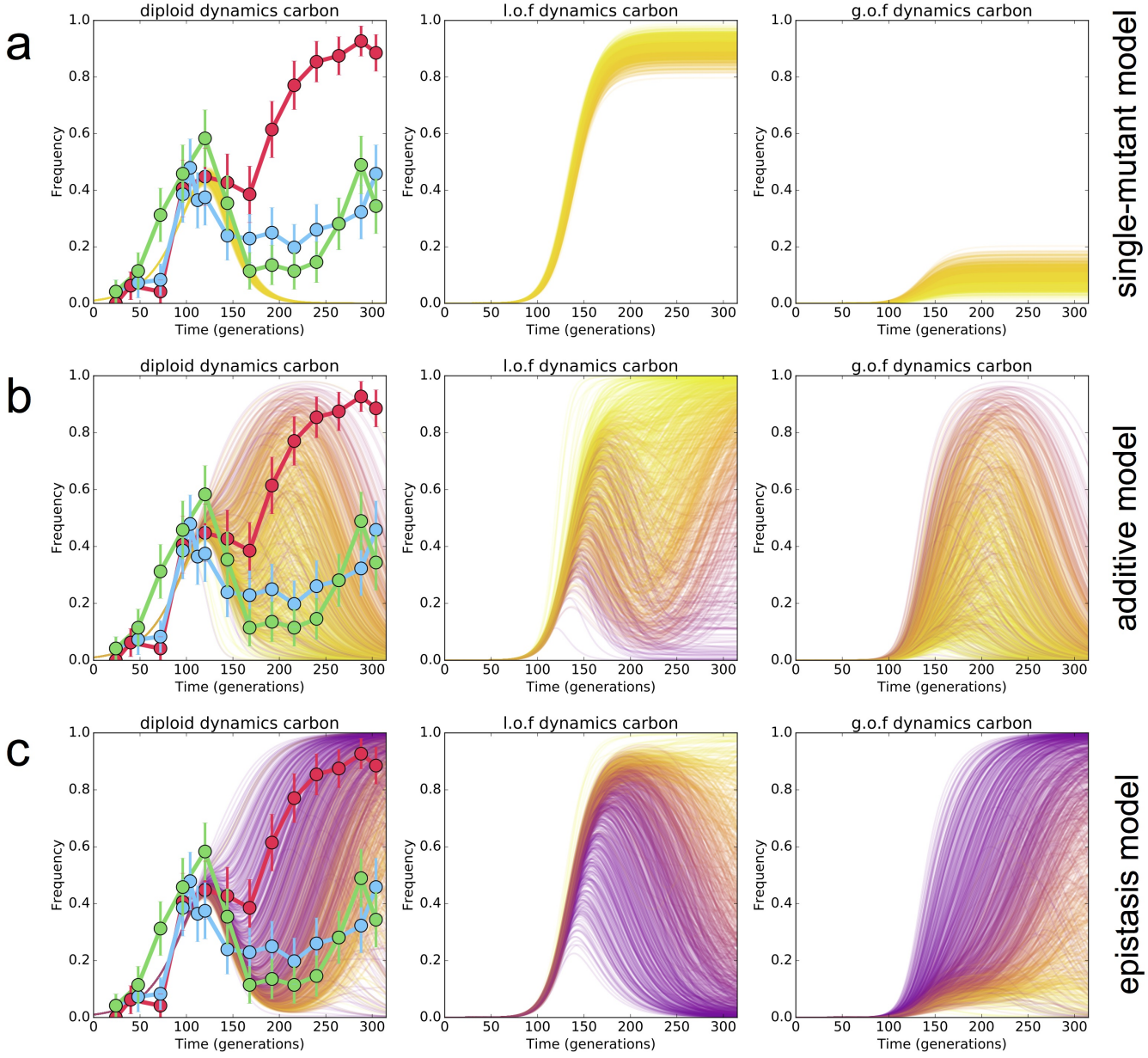
Each model is specified by a number of parameters:

Mutant	Measured fitness	Measured rate	`Single` s & U	`Additive` s & U	`Epistasis` s & U
Dip	0.04	10^{-6}	—	—	—
LoF	0.10	10^{-7}	—	—	—
GoF	0.10	10^{-8}	—	—	—
$Dip + Dip$		—	—	0.08 & 10^{-6}	—
$Dip + LoF$		—	—	0.14 & 7.5×10^{-7}	—
$Dip + GoF$		—	—	0.14 & 7.5×10^{-8}	0.14 & 10^{-8}
$LoF + Dip$		—	—	0.14 & 10^{-6}	0.14 & 10^{-6}
$LoF + LoF$		—	—	0.20 & 7.5×10^{-7}	—
$LoF + GoF$		—	—	0.20 & 7.5×10^{-8}	—
$GoF + Dip$		—	—	0.14 & 10^{-6}	0.14 & 10^{-6}
$GoF + LoF$		—	—	0.20 & 7.5×10^{-7}	—
$GoF + GoF$		—	—	0.20 & 7.5×10^{-8}	—

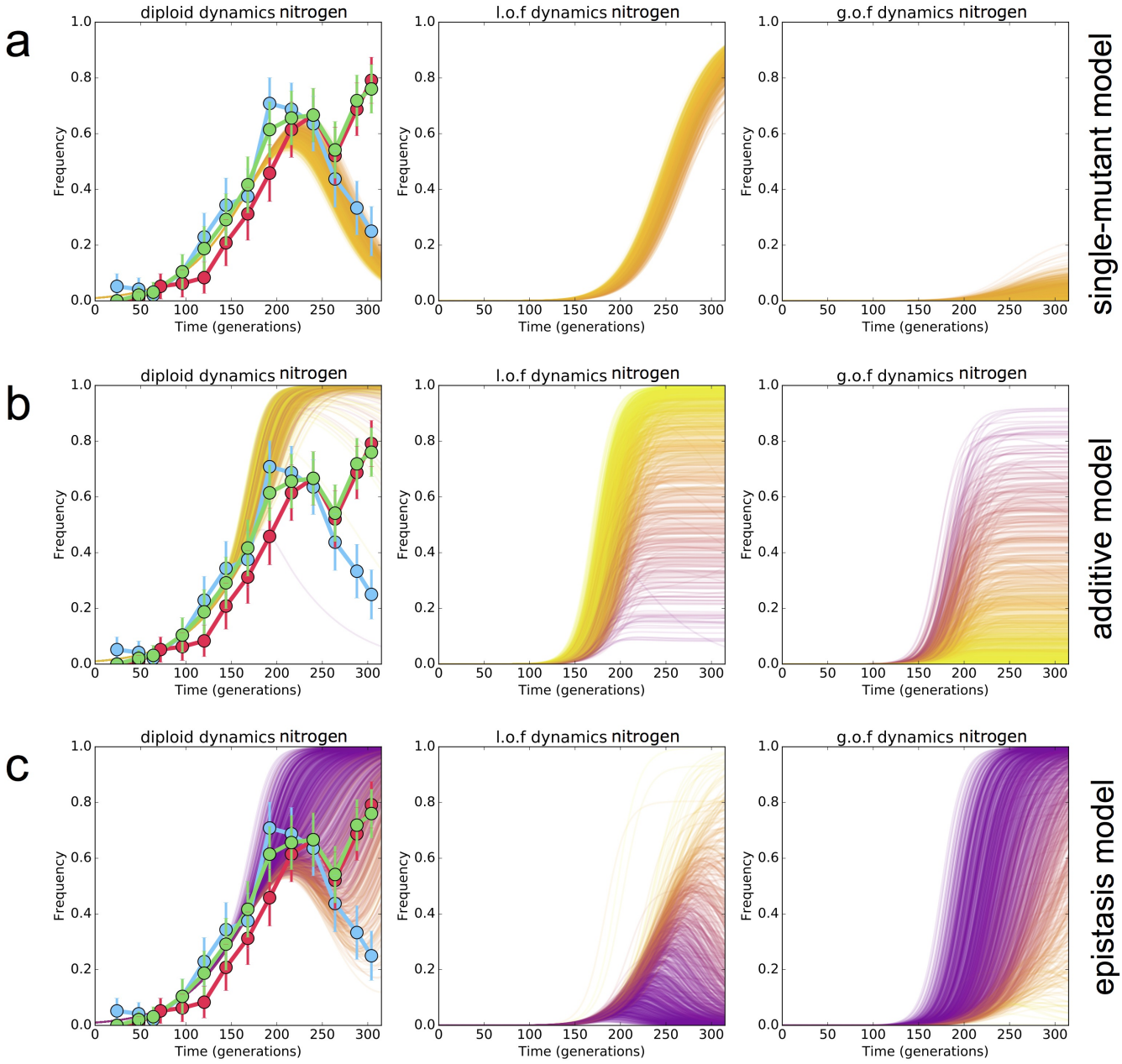
Table 3: Values used for the simulations of diploid trajectories in C-lim.

Mutant	Measured fitness	Measured rate	`Single` s & U	`Additive` s & U	`Epistasis` s & U
Dip	0.03	10^{-6}	—	—	—
LoF	0.06	10^{-7}	—	—	—
GoF	0.06	2.5×10^{-8}	—	—	—
$Dip + Dip$		—	—	0.06 & 10^{-6}	—
$Dip + LoF$		—	—	0.09 & 10^{-7}	—
$Dip + GoF$		—	—	0.09 & 2.5×10^{-8}	0.09 & 10^{-8}
$LoF + Dip$		—	—	0.09 & 10^{-6}	0.09 & 10^{-6}
$LoF + LoF$		—	—	0.12 & 10^{-7}	—
$LoF + GoF$		—	—	0.12 & 2.5×10^{-8}	—
$GoF + Dip$		—	—	0.09 & 10^{-6}	0.09 & 10^{-6}
$GoF + LoF$		—	—	0.12 & 10^{-7}	—
$GoF + GoF$		—	—	0.12 & 2.5×10^{-8}	—

Table 4: Values used for the simulations of diploid trajectories in N-lim.



Supplementary Figure 16: C-lim diploid, LoF and GoF trajectories for 3 models of mutation acquisition. (a) single-mutant model (b) additive model and, (c), epistasis model. The 2nd and 3rd columns show the trajectory of LoF and GoF mutants as classes. Coloring is consistent across the three plots so that a yellow trajectory in one column is the same color across all three columns since it comes from the same simulation.



Supplementary Figure 17: N-lim diploid, LoF and GoF trajectories for 3 models of mutation acquisition. (a) single-mutant model (b) additive model and, (c), epistasis model. The 2nd and 3rd columns show the trajectory of LoF and GoF mutants as classes. Coloring is consistent across the three plots so that a yellow trajectory in one column is the same color across all three columns since it comes from the same simulation.

7 Maximum likelihood procedure for goodness-of-fit testing

To measure the goodness-of-fit of the different models that could explain the diploid frequency trajectories shown in Figure 4 of the main text we define the log-likelihood function, L

$$L = -\sum_i \ln \binom{N}{k_i} + k_i \ln(f_i) + (N - k_i)\ln(1 - f_i) \tag{25}$$

where i runs over all time points in the diploid trajectory, $N = 96$ is the number of colonies tested at each time point, k_i are the number of diploid colonies measured at time point i and f_i is the frequency of diploids predicted by the model at time point i. With this we then perform the following

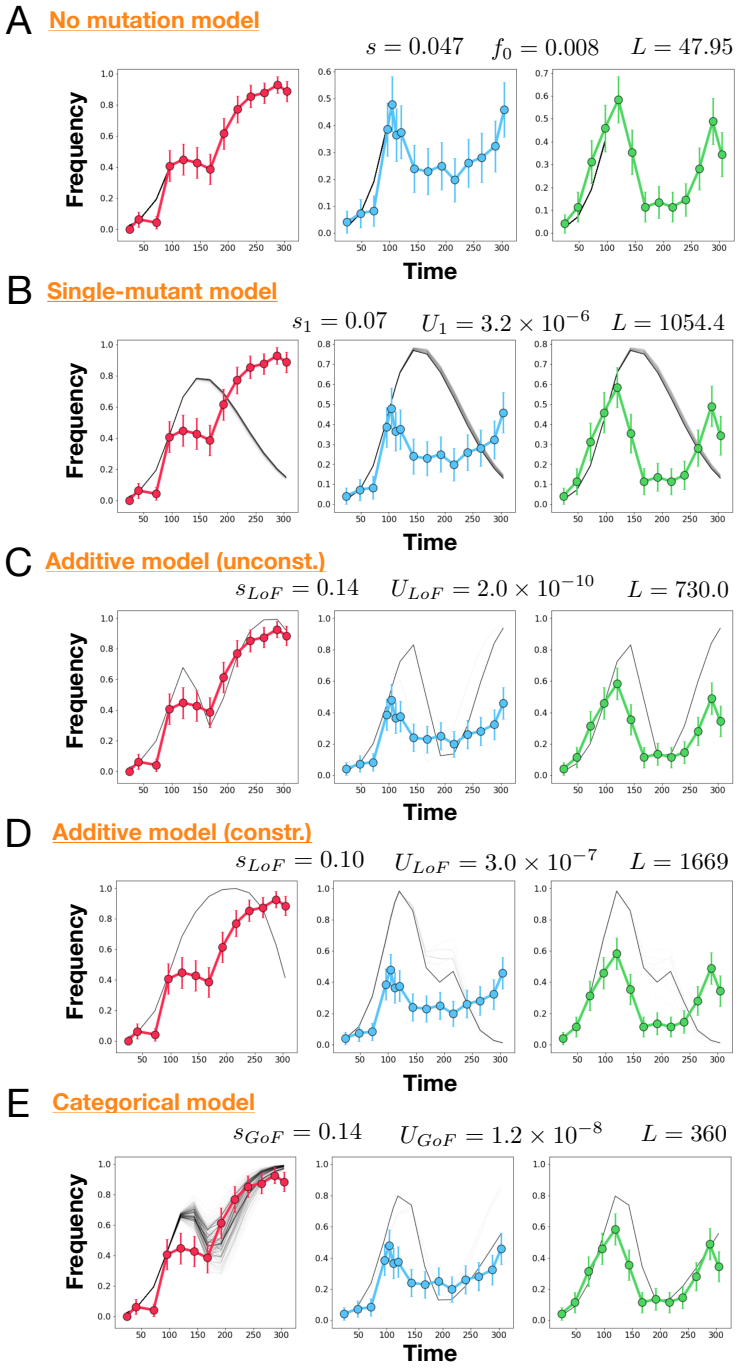
1. **No mutation model.** We determine the best fit values for the fitness effect, s_{Dip}, and initial frequency, f_0 of diploids across all three replicate experiments. The range of values considered for (s_{Dip}, f_0) were $s_{Dip} \in 0.015 - 0.055$ in steps of $\delta s_{Dip} = 0.001$ and $f_0 \in 0.005 - 0.015$ in steps of $\delta f_0 = 0.001$. We restricted analysis to time points $t \leq 104$ to minimize the effect of de-novo mutations (those that occur during the evolution) from influencing best-fits for these parameters. For each parameter pair (s_{Dip}, f_0) calculated the log-likelihood of the 100 simulated trajectories and recoded the log-likelihood of the smallest (i.e. the most likely) for each of the three replicates and added these together to form the total log-likelihood. Under this procedure the maximum likelihood values for (s_{Dip}, f_0) were determined to be $(s_{Dip} = 0.047, f_0 = 0.008)$ with a total log-likelihood of $L = 47.95$. The fitted trajectories compared to the data for these best fit parameters are shown in Figure 18 A.

2. **Single-mutant model.** With the s_{Dip} and f_0 values fixed, we determine the best fit values under a single-mutant model in which there is a rate U_{LoF} to a fitness effect s_{LoF} (representing large fitness effect LoF and GoF mutations). The range of values considered for (s_{LoF}, U_{LoF}) were $s \in 0.04 - 0.14$ in steps of $\delta s_{LoF} = 0.01$ and $U_{LoF} \in 1 \times 10^{-7} - 60 \times 10^{-7}$ in steps of $\delta U_{LoF} = 10^{-7}$. We included all time points in this analysis. Performing the same procedure as in (1) we find the best fit values are $(s_{LoF} = 0.07, U_{LoF} = 3.2 \times 10^{-6})$ with a total log-likelihood of $L = 1054.4$. The best fit trajectories for these parameters are show in Figure 18 B.

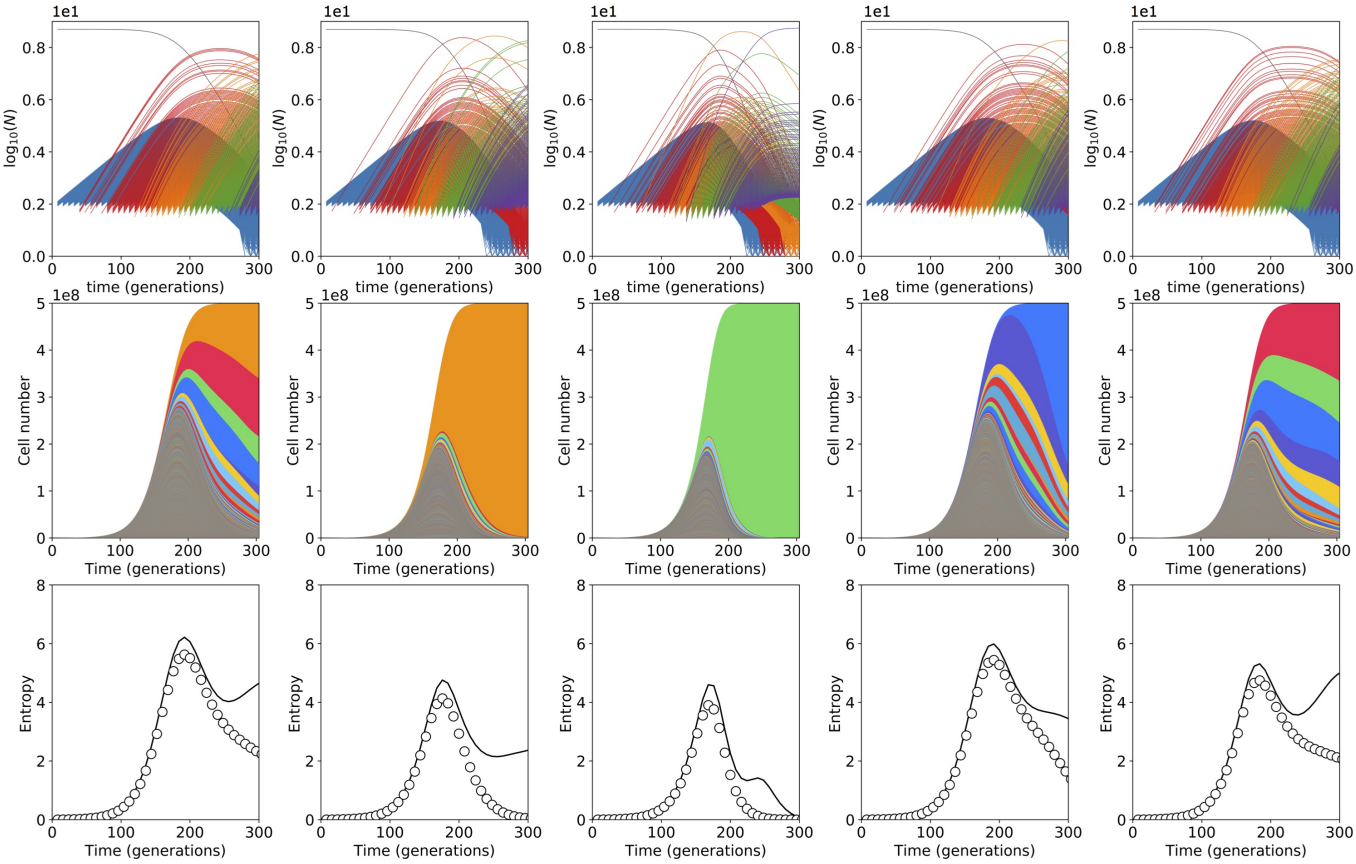
3. **Additive multiple-mutant model (unconstrained)** Here we allow multiple mutants (both additional diploids and additional LoF and GoF mutations) to occur and for their fitnesses to add. In order to perform a two-parameter search only, we fix the mutation rate to diploidization to 10^{-5}, since this is the measured value. We then performed a two-parameter maximum likelihood search on the parameter pair (s_{LoF}, U_{LoF}) (in this model there is no difference between LoF and GoF mutations). The best-fit parameters are found to be $(s_{LoF} = 0.14, U_{LoF} = 2.0 \times 10^{-10})$ with a total log-likelihood value of $L = 730.0$. The best fit trajectories for these parameters are shown in Figure 18 C.

4. **Additive multiple-mutant model (constrained)** The unconstrained additive model best fit parameter for $U_{LoF} = 2 \times 10^{-10}$ is over three-orders of magnitude smaller than the measured value from our data. Therefore we also explored a maximum likelihood model in which we constrained $3 \times 10^{-7} < U_{LoF} < 10^{-6}$ (within half an order of magnitude of measured values). We then performed a two-parameter sweep over possible (s_{LoF}, U_{LoF}) values, finding the maximum likelihood values to be $(s_{LoF} == 0.10, U_{LoF} = 3 \times 10^{-7})$ with a maximum-likelihood of $L \approx 1669$. This model, with more realistic parameters clearly does much worse than the unconstrained model with unrealistic parameters. We therefore asked whether a model with minimal epistasis would be able to provide a more parsimonious explanation of the data.

5. **Categorial epistasis model**. In this model we constrain $U_{Dip} = 10^{-5}$, $U_{LoF} = 3x10^{-7}$, and $s_{Dip} = 0.047$, $s_{LoF} = 0.11$ as measured. We furthermore, we allow fitness effects to combine additively. We then ask, if Diploids can acquire GoF mutations of fitness effect s_{GoF} at rate U_{GoF} can this model provide a more likely explanation of the data. We find that the maximum likelihood values $(s_{GoF} = 0.14, U_{GoF} = 1.2 \times 10^{-8})$ are both in the range we would expect based on GoF single-mutants, yet this model provides a much more likely explanation of the data since $L \approx 360$.



Supplementary Figure 18: Plots showing the three diploid trajectories in C-lim C1(red data points), C2 (blue data points) and C3 (green data points) compared to the best fit trajectories (black lines) generated by the maximum likelihood models (A) no mutation (B) single-mutant model (C) additive model with no constraint on parameters, (D) additive model in which the mutation rate to LoF mutations is constrained to be the measured rate and (E) the categorical epistasis model. Thie produces the lowest -log-likelihood (best fit) and trajectories that most closely resemble the data.



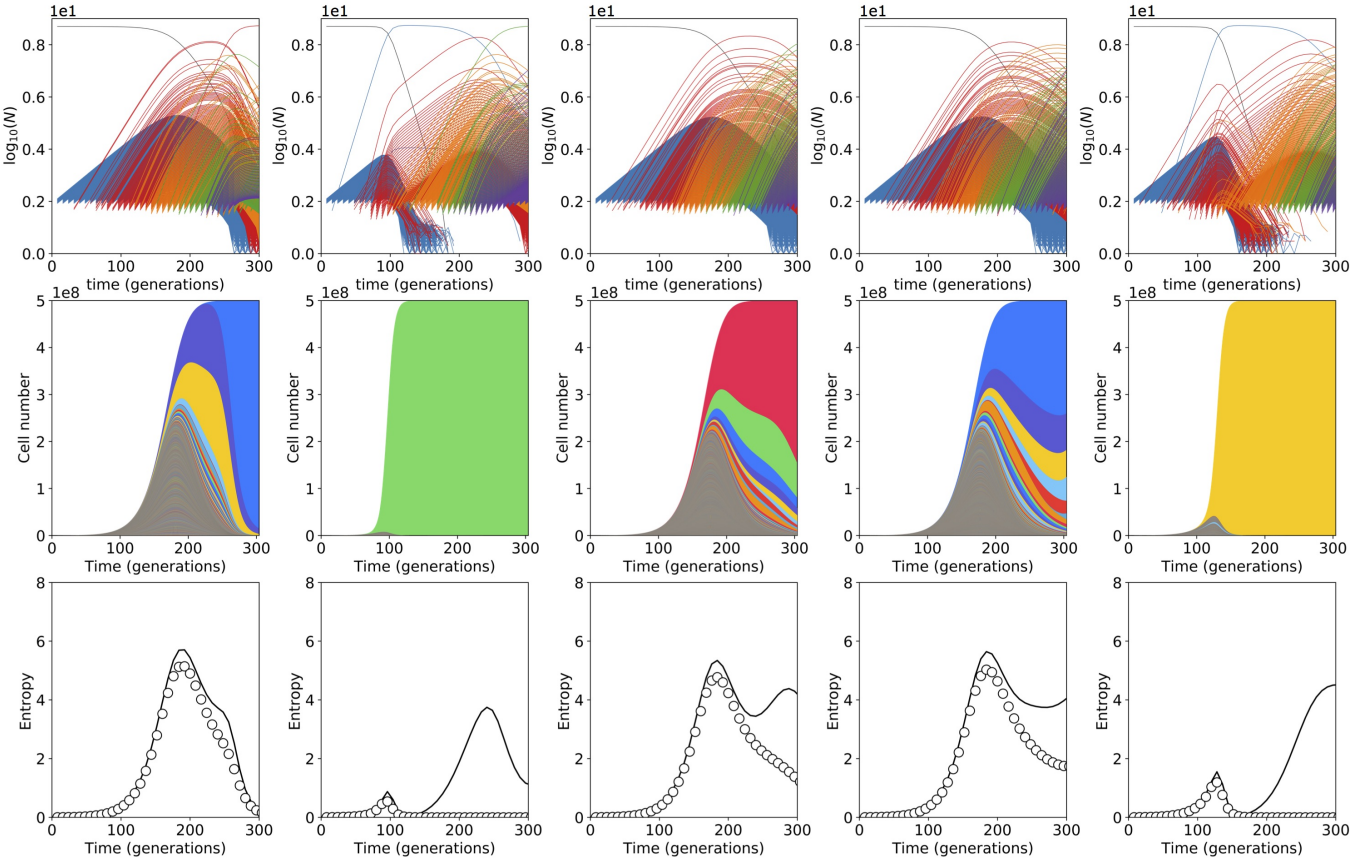
Supplementary Figure 19: Diversity dynamics for mDFE modeling growth-inhibiting drug. Here the mDFE is composed of two delta-functions (at $s_\alpha = 0.05$ with rate $U_\alpha = 10^{-5}$ and $s_\beta = 0.20$ with rate $U_\beta = 10^{-11}$). Top row, the genotype trajectories of all unique genotypes that arise in the simulation (single-mutants blue, double-mutants red, triple-mutants orange, quadruple-mutants purple). Large crashes in diversity (column 2 and 3) are caused by anomalously early double mutants comprised of two small-effect mutations. Bottom row: the entropy trajectory (see Section 9) of all genotypes in the population (data points) compared to the theoretical prediction from using Eq. 35 in the text that assumes single mutants only.

8 Growth-inhibiting drug mDFE and Power-law mDFEs

Here we model how adaptive genetic diversity evolves under a mDFE with a small supply of large-effect mutations, as might occur in the presence of a growth-inhibiting drug. As an example of what happens in such a scenario, we use a model mDFE where small-fitness-effect mutations ($s_\alpha = 0.05$) occur at a relatively high rate ($U_\alpha = 10^{-5}$ per cell division) while large fitness-effect mutations ($s_\beta = 0.20$) occur at a lower rate (in the range $10^{-11} < U_\beta < 10^{-9}$ per cell division). In this context we take "large" fitness effects to mean that $s_\beta > s_\alpha \times (\ln(N s_\alpha)/\ln(s_\alpha/U_\alpha))$. That is, large-effect mutations destined to establish will typically sweep regardless of the fitness of the genotype on which they fall (since the fitness effect is larger than the range of fitnesses expected in the population at any time). In this limit the the dynamics of adaptive genetic diversity depends most sensitively on the supply of large-effect mutations U_β.

Large effect mutants enter typically every $\sim 1/NU_\beta s_\beta$ generations. The early time dynamics of the modest-effect mutations studied here takes place over timescales $\tau_{sw} = (1/s_\alpha)\ln(s_\alpha/U_\alpha)$ called the "sweep time" (~ 170 generationsfor the parameters above and those in Figures 19 - 21). It is the relative magnitudes of these timescales that can be used to assess the impact of the large-effect mutants on the adaptive genetic diversity.

When the supply of large-effect mutations is low enough, large-effect mutations have little impact on the early adaptive diversity dynamics (Figure 19, where $U_\beta = 10^{-11}$). In this case the typical time it takes for a large-effect mutation to

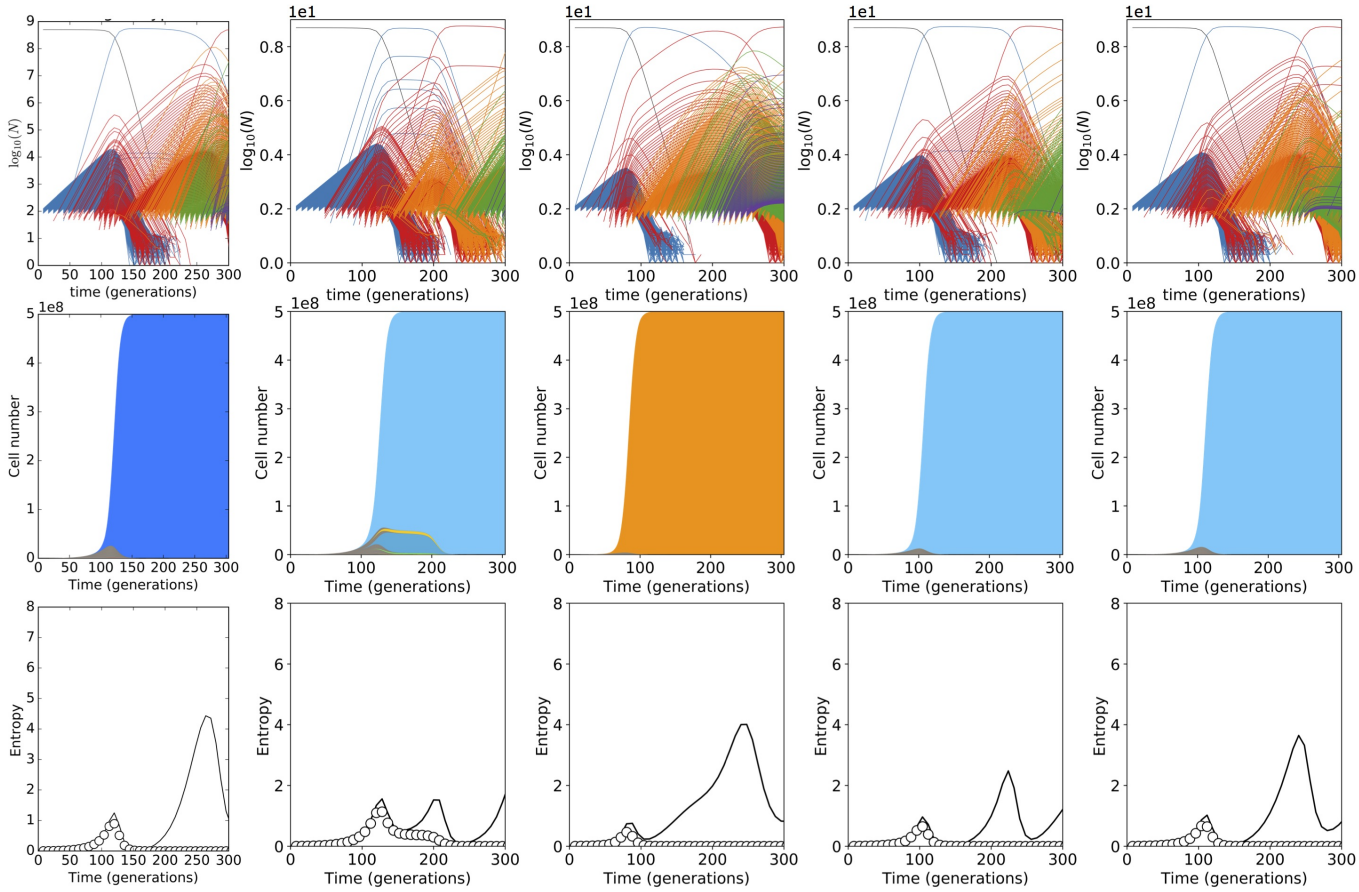


Supplementary Figure 20: Diversity dynamics for mDFE modeling growth-inhibiting drug. Here the mDFE is composed of two delta-functions (at $s_\alpha = 0.05$ with rate $U_\alpha = 10^{-5}$ and $s_\beta = 0.20$ with rate $U_\beta = 10^{-10}$). Top row, the genotype trajectories of all unique genotypes that arise in the simulation (single-mutants blue, double-mutants red, triple-mutants orange, quadruple-mutants purple). Here, because the supply of large-effect single-mutants is higher, large crashes in diversity (e.g. column 1 and 5) can be caused large-effect single-mutants which sweep. Bottom row: the entropy trajectory (see Section 9) of all genotypes in the population (data points) compared to the theoretical prediction from using Eq. 35 in the text that assumes single mutants only. In cases where large-effect single-mutants enter early, they drive a diversity crash before single-mutant diversity has much time to build up.

enter is $\sim 10^3$ generations and hence the probability of large-effect mutation entering during the early time dynamics (<170 generations) is small. Thus, when large diversity crash does occur, it is driven by anomalously early double mutants as can be seen in Figure 19 and as is the case for the empirical mDFEs observed in the experimental evolutions.

As one increases the supply of large-effect mutations (Figure 20, where $U_\beta = 10^{-10}$) the timing of the first large-effect mutation and the sweep time become comparable. For parameters used here, this is predicted to occur at $U_\beta \sim 10^{-10}$, and Figure 20 indeed shows that, in a substantial fraction of simulations, a large-effect mutation enters and causes a single-mutant driven diversity crash. As the supply of large-effect mutants in increased further (Figure 21 where $U_\beta = 10^{-9}$), large-effect single mutants sweep in almost all simulations and the diversity does not have enough time to build up before a large-effect mutant sweeps.

The presence of a growth-inhibiting drug could therefore be expected to alter the early adaptive diversity dynamics since stochastically occurring large-effect single-mutants could drive a crash, which would be quantitatively different from the case where the diversity crash is driven by a small number of fit multiple-mutants. It should be noted however, that in this case, the evolutionary dynamics changes more generally: it is driven by successive sweeps of large-effect mutants and is no longer truly in the clonal-interference multiple-mutation limit studied here.



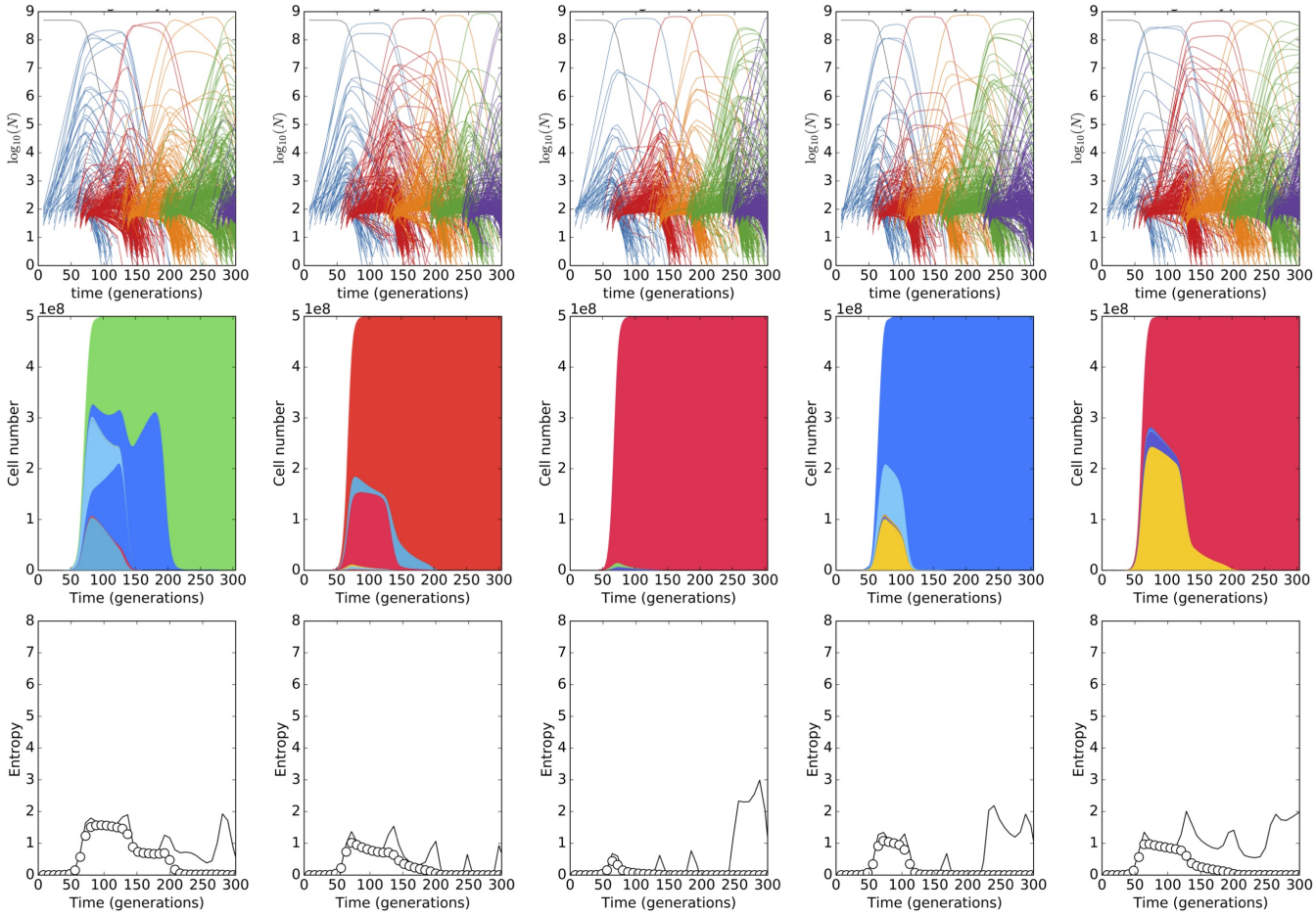
Supplementary Figure 21: Diversity dynamics for mDFE modeling growth-inhibiting drug. Here the mDFE is composed of two delta-functions (at $s_\alpha = 0.05$ with rate $U_\alpha = 10^{-5}$ and $s_\beta = 0.20$ with rate $U_\beta = 10^{-9}$). Top row, the genotype trajectories of all unique genotypes that arise in the simulation (single-mutants blue, double-mutants red, triple-mutants orange, quadruple-mutants purple). With a larger supply of large-effect single-mutants almost all simulations exhibit a diversity crash driven by large-effect single-mutants which sweep. Bottom row: the entropy trajectory (see Section 9) of all genotypes in the population (data points) compared to the theoretical prediction from using Eq. 35 in the text that assumes single mutants only. In all cases a diversity crash caused by large-effect single-mutants occurs before diversity can build up substantially. This is followed by large oscillations in diversity occurring over timescales at which large-effect mutations enter and sweep i.e. every $1/(NU_\beta s_\beta) \sim 100$ generations.

8.1 Power-law mDFEs

In this section we explore how the dynamics of adaptive genetic diversity would be expected to evolve when the mDFE has a supply of beneficial mutations spread over roughly two orders-of-magnitude. We consider mDFEs where $\mu(s) \sim s^{-\gamma}$ with a lower limit of $s = 0.005$ (set since this mutations below this effect size would not be able to establish before the end of the observation period ~ 300 generations) and with an upper limit of $s = 0.3$ (set to remain biologically plausible, given the observed fitness effects). The total rate of beneficial mutations is set to be $U_b = 10^{-6}$.

Figure 22 shows the early time dynamics of beneficial clones for $\gamma = 2$. In this case, the dynamics effectively enters a successive sweep regime whereby stochastically occurring large-fitness-effect clones effectively fix before the occurrence of the $n + 1$th mutant clone. In this case, diversity has a limited time to build up and is typically purged causing characteristic humps in the entropy (last row, Figure 22).

Power law mDFEs thus produce qualitatively different adaptive diversity dynamics than we observe in the experimental data, though we note that a full quantitative analysis of the dynamics expected under a power-law mDFE is beyond the scope of the present work [3].



Supplementary Figure 22: Diversity dynamics for a power-law mDFE. In these simulations $\mu(s) \sim s^{-2}$ for $0.005 < s < 0.3$. The total beneficial mutation rate is normalized to $U = 10^{-6}$. As in previous simulations the bottleneck population size is 7×10^7 which grows by a factor of 256 between bottlenecks resulting in an effective population size of $\approx 5 \times 10^8$. Top row, the genotype trajectories of all unique genotypes that arise in the simulation (single-mutants blue, double-mutants red, triple-mutants orange, quadruple-mutants purple). Middle row: Muller plots of adaptive lineages. Bottom row: the entropy trajectory of all genotypes in the population (data points) compared to the theoretical prediction from using Eq. 35 in the text that assumes single mutants only. This is followed by large oscillations in diversity occurring over timescales at which large-effect mutations enter and sweep i.e. every $1/(NU_\beta s_\beta) \sim 100$ generations.

9 Lineage abundances and calculation of adaptive lineage diversity

Here we outline how we produced Figure 1 of the main text as well as trajectories for the genetic diversity (entropy) in Figure 3 of the main text.

Interpolating and extrapolating lineage abundances To plot Figure 1 of the main text we use all of the time points sequenced to best interpolate and extrapolate the abundance of each lineage in the population. To do this we use the first 104 generations of C-lim and the first 192 generations of N-lim data to estimate the fitness of each lineage using the same methods as outlined in [1], which produce fitness estimates for each lineage as well as the time evolution of the population mean fitness. We then "forecast" the abundance for times later than T104 (C-lim) or T192 (N-lim) by:

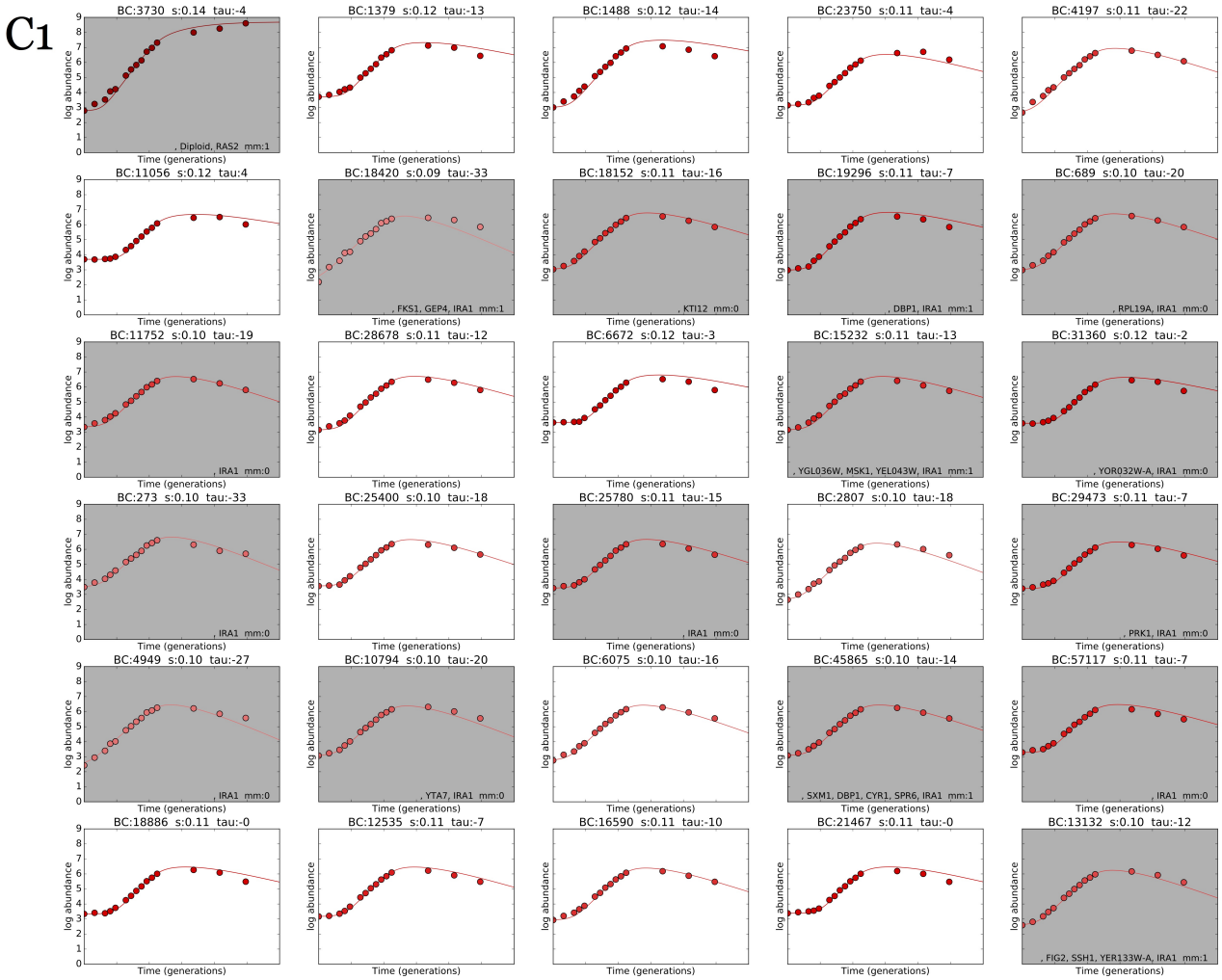
- Using the inferred fitness x_i and the initial mean fitness $\bar{x}(t_0)$ to predict the frequencies at the next time point:

$$f_i(t_0 + \Delta t) = f_i(t_0) \exp\left[(x_i - \bar{x}(t_0))\Delta t\right] \tag{26}$$

- Recalculate the new mean fitness at this later time point:

$$\bar{x}(t_0 + \Delta t) = \boldsymbol{x} \cdot \boldsymbol{f}(t_0 + \Delta t) \tag{27}$$

and iterate each generation until T304. The result of this procedure can be validated against the known abundance of the lineage measured from sequencing data at the later time points. Comparisons for how well this procedure works is shown in Figures 23 and 24.

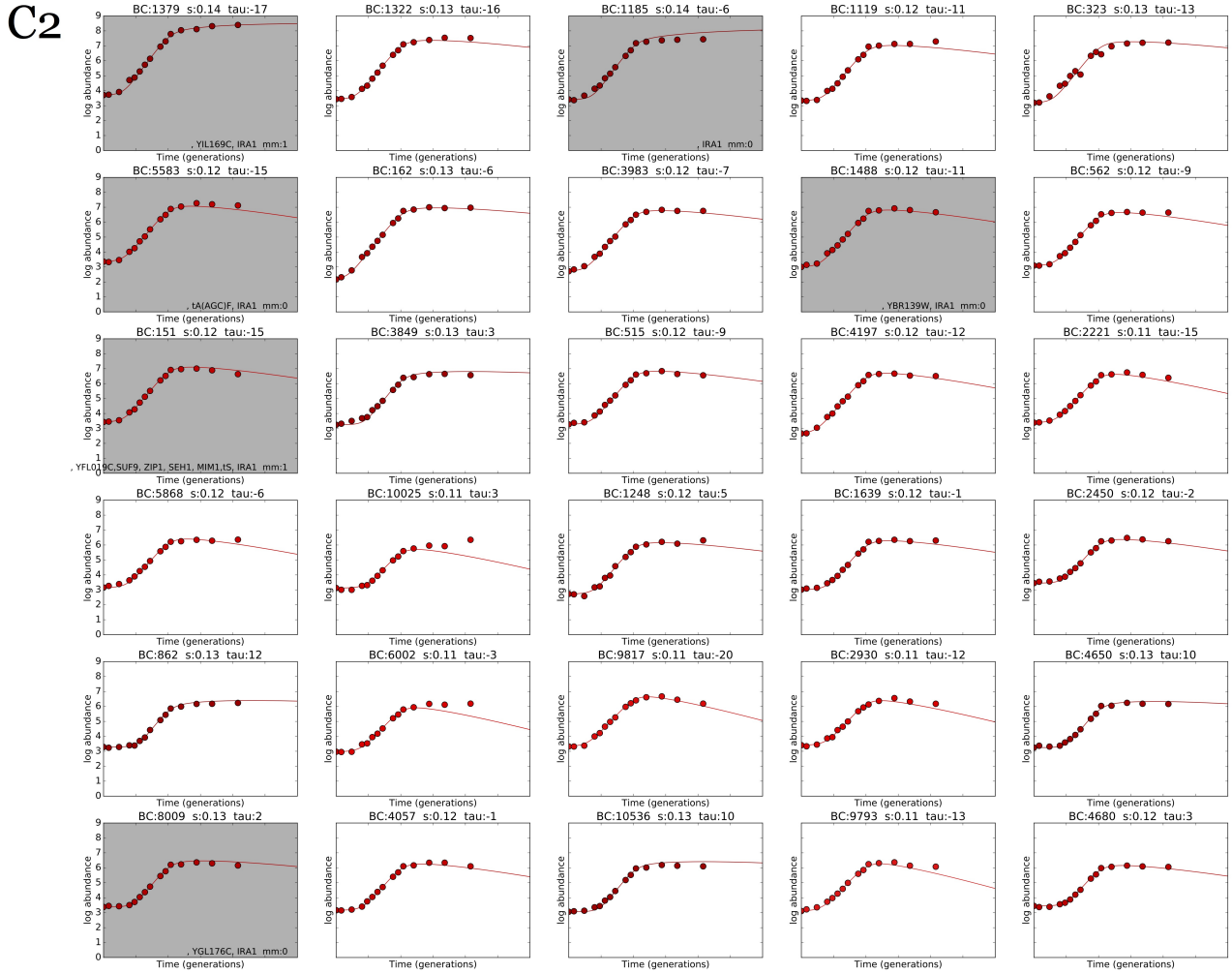


Supplementary Figure 23: Interpolation and extrapolation of lineage abundance trajectories, C-lim replicate 1. The top 30 barcodes ranked by abundance. Each panel shows a measured barcode trajectory (data points) vs. the predicted trajectory from the extrapolation procedure (solid line). Colors indicate fitness level. Shaded panels are those lineages from which a clone was picked and subsequently whole-genome sequenced.

Diversity measured via adaptive lineage entropy To quantify genetic diversity we use the Shannon entropy of adaptive lineage through time defined here as

$$S(t) = \sum_i f_i(t) \ln\left(\frac{1}{f_i(t)}\right) \tag{28}$$

where $f_i(t)$ is the frequency of lineage i at time t and the sum is over all *adaptive* lineages (excluding diploids) identified using the same method as in [1] (using lineage deviation from a neutral expectation). The numerical value of S can be interpreted, very roughly, as there being effectively $\sim e^S$ adaptive lineages present in the population. Note that we exclude diploids



Supplementary Figure 24: Interpolation and extrapolation of lineage abundance trajectories, C-lim replicate 1. The top 30 barcodes ranked by abundance. Each panel shows a measured barcode trajectory (data points) vs. the predicted trajectory from the extrapolation procedure (solid line). Colors indicate fitness level. Shaded panels are those lineages from which a clone was picked and subsequently whole-genome sequenced.

from our definition because most diploid lineages are genetically identical despite adaptive clones arising in a large number of lineages. SNPs by contrast are typically genetically unique.

Other possible measures of diversity. The diversity is really a statement about the distribution of adaptive clones sizes and how this changes through time as shown by the colored lines in the 2nd rows of Figures 9 - 14.

However, if one wants to compare many replicates it becomes difficult unless one chooses a statistic based on the distribution of clone or lineage abundances. The entropy is one such measure. We do not claim that the entropy is the "correct" measure of diversity since there are many alternatives, but it *does* capture the important qualitative distinction between the early vs. late distribution of clone sizes in the populations. Early, there are many small clones of roughly equal size, whereas late there is a very skewed distribution with a small number of very large clones, and the entropy measure captures this difference as discussed in the main text. Crucially, the entropy of adaptive lineages tracks the entropy of adaptive genotypes into the diversity crash.

One consideration for a diversity measure is how much to weight low frequency lineages. Consider the case where one adaptive lineage is large, at frequency p while the $k \gg 1$ other lineages are at frequency $1 - p/k$. In this case the contribution from the low frequency lineages would be $(1-p)\log(k/(1-p))$ which would mean that the low frequency lineages would contribute more to the entropy than the dominant lineage when $p < 1 - 1/\log(k)$. A reasonable alternative measure of

diversity, would be to weight variants not with $\log(p_i)$ (as one does in the entropy measure), but with a different functional form. For example, if we used a diversity measure:

$$D = \sum_i f_i^\gamma \tag{29}$$

where $0 \leq \gamma \leq 1$, the contribution from the low frequency lineages would be enhanced over simple linear weighting, (and in the limit of $\gamma = 0$ it is simply a count of how many are present). Following the same line of argument as above, the contribution of the low frequency lineages would be larger than the single large lineage if $1 - p < k^{1-1/\gamma}$.

Other reasonable choices of the diversity measure could be made. One simple example would be a step function counting how many lineages reach above a threshold size p_0 over time. Because many adaptive lineages reach frequencies above 10^{-5}, but many of them dip below this frequency at late times, for parameter values $10^{-5} < p_0 < 10^{-3}$, this measure would be qualitatively similar to the entropy measure, though has the drawback of having to make a somewhat arbitrary choice for p_0.

References

[1] Sasha F. Levy, Jamie R. Blundell, Sandeep Venkataram, Dmitri A. Petrov, Daniel S. Fisher, and Gavin Sherlock. Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature, 519(7542):181–186, March 2015. ISSN 0028-0836. URL http://www.nature.com/nature/journal/v519/n7542/abs/nature14279.html.

[2] Cornelis Verduyn, Erik Postma, W. Alexander Scheffers, and Johannes P. Van Dijken. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast, 8(7): 501–517, July 1992. ISSN 1097-0061. URL http://onlinelibrary.wiley.com/doi/10.1002/yea.320080703/abstract.

[3] Michael M. Desai and Daniel S. Fisher. Beneficial mutation selection Balance and the Effect of Linkage on Positive Selection. Genetics, 176(3):1759–1798, July 2007. ISSN 0016-6731, 1943-2631. URL http://www.genetics.org/content/176/3/1759.

[4] Philipp W. Messer and Richard A. Neher. Estimating the Strength of Selective Sweeps from Deep Population Diversity Data. Genetics, 191(2):593–605, June 2012. ISSN 0016-6731, 1943-2631. URL http://www.genetics.org/content/191/2/593.

[5] Sandeep Venkataram, Barbara Dunn, Yuping Li, Atish Agarwala, Jessica Chang, Emily R. Ebel, Kerry Geiler-Samerotte, Lucas Herrissant, Jamie R. Blundell, Sasha F. Levy, Daniel S. Fisher, Gavin Sherlock, and Dmitri A. Petrov. Development of a Comprehensive Genotype-to-Fitness Map of Adaptation-Driving Mutations in Yeast. Cell, 166 (6), 1585 - 1596 September 2016. ISSN 0092-8674, 1097-4172. URL /cell/abstract/S0092-8674(16)31010-8.

[6] Jungeui Hong and David Gresham. Molecular Specificity, Convergence and Constraint Shape Adaptive Evolution in Nutrient-Poor Environments. PLOS Genetics, 10(1):e1004041, January 2014. ISSN 1553-7404. URL http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004041.

[7] Gregory I. Lang, Daniel P. Rice, Mark J. Hickman, Erica Sodergren, George M. Weinstock, David Botstein, and Michael M. Desai. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature, 500(7464):571–574, August 2013. ISSN 0028-0836. URL http://www.nature.com/nature/journal/v500/n7464/full/nature12344.html.

[8] Daniel J. Kvitek and Gavin Sherlock. Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment. PLoS Genet, 9(11):e1003972, November 2013. URL http://dx.doi.org/10.1371/journal.pgen.1003972.