Supporting Information

Salverda et al. 10.1073/pnas.1712999114

SI Materials and Methods

Media. LB medium is 10 g/L trypticase peptone, 5 g/L yeast extract, and 10 g/L NaCl. LB tetracycline medium is LB medium containing 15 mg/L tetracycline. Mueller Hinton (Merck) and Mueller Hinton II (BD) medium were prepared according to the manufacturers' instructions. SOC medium is 20 g/L bacto tryptone and 5 g/L yeast extract with 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, and 20 mM glucose. Solid media contained 16 g/L agar.

Recombinant Methodology. Plasmids were prepared from overnight LB tetracycline cultures and purified using a GenElute Plasmid MiniPrep Kit (Sigma). TEM-1 (861 bp ORF) was amplified from pBR322 using the high-fidelity *Pfu* polymerase (Stratagene), sense primer P1 (GGGGGGTCATGAGTATTCAACATTTCCGTGTCG; BspHI site underlined, this enzyme cuts 1 bp upstream of TEM-1's start codon), and antisense primer P2 (CCGAGCTCTTGGTCTGACAGTTACCAATGC; SacI site underlined, this enzyme cuts 17 bp downstream of TEM-1's stop codon) using the following temperature cycle: 30′ at 95 °C, 30′ at 61 °C, and 90′ at 72 °C for 30 cycles, followed by 72 °C for 10 min. The resulting amplicon was digested with BspHI and SacI (New England BioLabs). Plasmid pACSE3 was digested with the same restriction enzymes and dephosphorylated with Calf Intestinal Phosphatase (New England BioLabs). Digested amplicons and vectors were purified using Sigma's GenElute PCR Clean-up Kit, ligated using T4 DNA Ligase (New England BioLabs), and transformed into DH5αE by electroporation. Note that in this protocol, TEM-1's promoter region is not part of the amplicon.

Creation of TEM-1 Mutant Libraries. Mutant TEM alleles were generated by introducing random mutations using the GeneMorph II Mutagenesis Kit (Stratagene) using general procedures described in ref. 1. Primers used for error-prone PCR were P3 (TCATCCGGCTCGTATAATGTGGA) and P4 (ACTCTCTTCCGGGCGCTATCAT), which flank the multiple cloning site of pACSE3. We developed an alternative PCR protocol that uses 0.5 μL of primers P3 and P4, 0.5 μL dNTP, 2.5 μL pfu buffer (Fermentas), 0.2 μL Mutazyme II, 565 ng template plasmid, and MQ to 25 μL. Sequencing of unselected amplicons revealed that the mutation rate of this protocol was ∼1.6 mutations per kilobase. The resulting amplicons were digested with BspHI and SacI, ligated into pACSE3, and electroporated into DH5αE. After recovery for 90 min in SOC medium at 37 °C, the cells were diluted in 500 mL LB tetracycline. An aliquot was taken out directly after mixing and plated onto LB tetracycline agar to determine the library size. The remainder of the culture was incubated at 37 °C overnight to amplify the library. Effective library sizes varied between $\sim 10^5$ and 10^6 transformants. Aliquots of the amplified libraries were stored in 10% glycerol at −80 °C.

Expression and Selection of TEM Alleles. Expression of the TEM-1 allele in pACSE3 is under control of the *pTAC* promoter that is tightly regulated by the lac repressor, encoded by the *lacI* gene on pACSE3. Expression was induced by adding IPTG to a final concentration of 50 μM, which was shown to mimic natural expression of TEM-1 (2). For selection, a series of bottles containing 50 mL Mueller Hinton medium (Merck) was inoculated with CTX (stock solution in 0.1 M $NaPO_4$, pH 7.0; Sigma) concentrations in twofold increments, ranging from 0.0625 mg/L, the MIC of wild-type TEM-1, to 4,096 mg/L. Each bottle was inoculated with a cell number approximately equal to 10 times the library size from the overnight-amplified cultures. Cultures were incubated for 48 h at 37 °C. The culture that grew at the highest concentration of CTX was plated on LB tetracycline agar. The next day, a single colony was selected and grown overnight in LB tetracycline. Plasmid from the overnight culture was isolated using GenElute Plasmid MiniPrep Kit (Sigma) and sequenced using BigDye (Perkin-Elmer) or DYEnamic (AP Biotech) Terminator Cycle Sequence kits.

MIC Assays. The MIC was determined from 150-μL cultures at a titer of 10^5 cells per milliliter in Mueller Hinton II medium. A 150-μL solution of twofold serial dilutions of antibiotic in MH II was added to these cultures as well as IPTG to render a final concentration of 50 μM. Cultures were grown for 20 h at 37 °C, growth was determined by visual inspection, and MIC was defined as the lowest concentration of antibiotic that completely prevents visible growth. To exclude phenotypic variation in bacteria and vector, before each MIC assay, plasmid was isolated from all selected clones, after which the TEM allele was amplified, placed in a naïve vector background, and retransformed into isogenic *E. coli* DH5αE cells. At least three replicate assays were performed to estimate the MIC of each genotype.

Bulk Competition and Deep Sequencing. Competitions were performed in 5 mL Mueller Hinton supplemented with 50 μM IPTG, and CTX concentrations ranging from 4 to 4,096 mg/L with twofold increments. The competitions were started with exponential phase cells expressing the eight final-round TEM alleles at equal frequencies, and a cumulative cellular density similar to that of the UD lines during passaging (10^5 cells per milliliter), and incubated at 37 °C for 48 h. At numerous time points we took 50- or 250-μL samples of the competitions and stored them at −20 °C until further processing. For each replicate, we then analyzed time points 0, 4, 8, 24, and 48 h for the highest CTX concentration at which we visually observed growth (1,024 mg/L for six replicates, 2,048 mg/L for two replicates). We then PCR-amplified the TEM gene using the P3 and P4 primers, by directly adding a small volume of the competition samples to the PCR mix (1 μL sample in a 50-μL PCR). The Nextera XT DNA library preparation kit (Illumina Inc.) was used to prepare libraries from the purified PCR products, and these were sequenced by Illumina HiSeq 2500 with 100-bp paired-end reads. Library preparation and sequencing were performed by the Cologne Center for Genomics. The data were analyzed with CLC Genomics Workbench version 8.0.2 (Qiagen Aarhus A/S) by trimming of reads (ambiguous limit of two bases, Phred score >30, and discarding reads shorter than 15 bases), mapping onto the pACTEM reference sequence with default mapping settings, and the export of the alignment (as per-position, per-nucleotide coverage information in *.tsv format). Each evolved line has 2–7 unique single-nucleotide mutations (including synonymous mutations), so at the relevant sites the relative frequency of the expected nucleotide for that allele was calculated. The mean relative frequency of the unique nucleotides (f) was used for estimating the selection rate constant (r) of each allele relative to the seven other alleles over a time interval t as $\ln(r) = (\ln(Z_t) - \ln(Z_0))/t$, where $Z = f/(1-f)$.

For testing whether initial cellular density affected selection of beneficiary types, we used a similar setup but with the following modifications. HiSeq 4000 2 × 75-bp paired-end reads was used, and more stringent criteria were used for trimming (Phred score >35.2 and discarding reads shorter than 40 bases) and alignment (length and similarity fractions set to 0.9). Because the median rate at which nucleotides did not match the time sequence

in the alignment was very low [$(7.40 \pm 0.9) \times 10^{-4}$] and SNP-calling algorithms cannot handle high coverage mappings (here also >500,000) efficiently, we again did not perform a SNP calling but only considered those changes with a frequency higher than 0.01. Note that other than those mutations reported, no other mutations were found to be close to this threshold. The relatively low complexity of the selected populations allowed haplotypes to be inferred heuristically, and they were checked manually by considering whether mutations were present on the same paired-end reads.

Statistical Analyses. We fitted a linear ($MIC_t = ct$) and an exponential model ($MIC_t = c(1-k^t)/(1-k)$) to the MIC doubling data, where the constant c is the rate of increase, and k causes attenuation of the MIC increase over time for the exponential model. The models were fitted to the data by least squares, using nonlinear regression (SPSS 22; IBM), and the SEM of parameter estimates was determined by bootstrapping. The results were verified by grid searches over large parameter spaces in R 3.3.0.

1. Salverda MLM, et al. (2011) Initial mutations direct alternative pathways of protein evolution. *PLoS Genet* 7:e1001321.

2. Barlow M, Hall BG (2002) Predicting evolutionary potential: *In vitro* evolution accurately reproduces natural evolution of the tem beta-lactamase. *Genetics* 160:823–832.

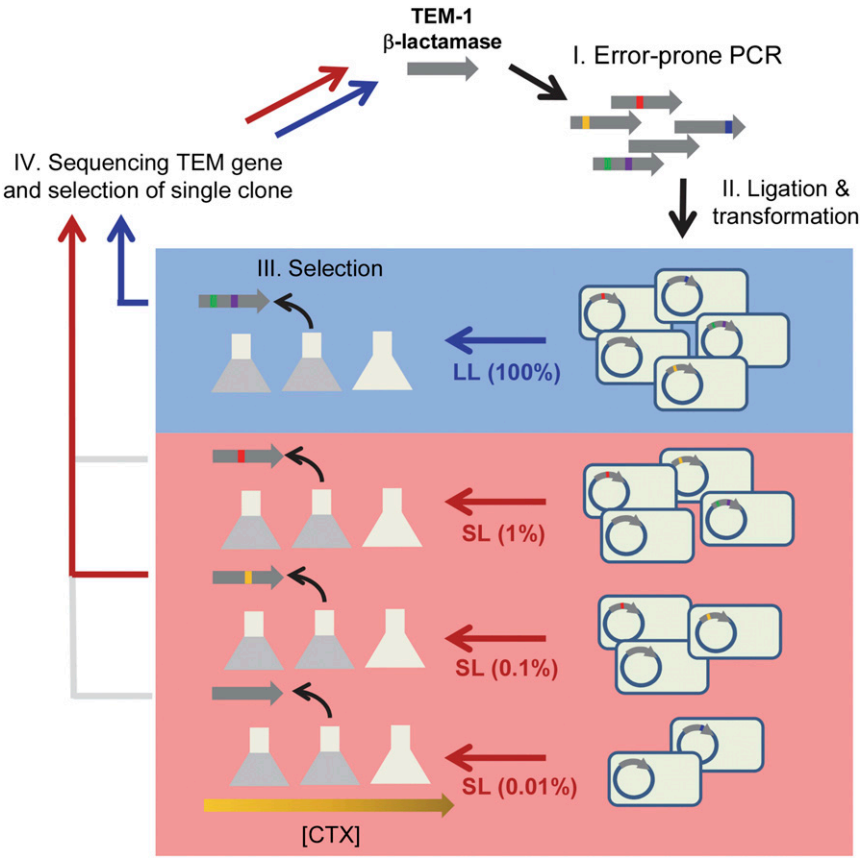

Fig. S1. In vitro evolution experiment with TEM-1 β-lactamase to test for small-mutation supply benefit. Large (LL, blue) and small (SL, red) libraries of TEM-1 mutants are selected at maximal CTX concentrations. LL selection uses 10^5–10^6 mutants per round, SL selection uses random samples of 1%, 0.1%, and 0.01% of undiluted mutant libraries, which are all three subjected to selection at maximal CTX concentrations. From the highest CTX concentration allowing bacterial growth, a single clone of each library is isolated, its TEM gene sequenced, and the gene selected from the smallest library with at least one nonsynonymous mutation is used for the next round of evolution (in the example shown from SL 0.1%). Gray arrows indicate the TEM gene; colored bars inside the gray arrows indicate hypothetical point mutations. Four rounds of LL and SL selection with four lines each were followed by two rounds of LL selection for all eight lines.

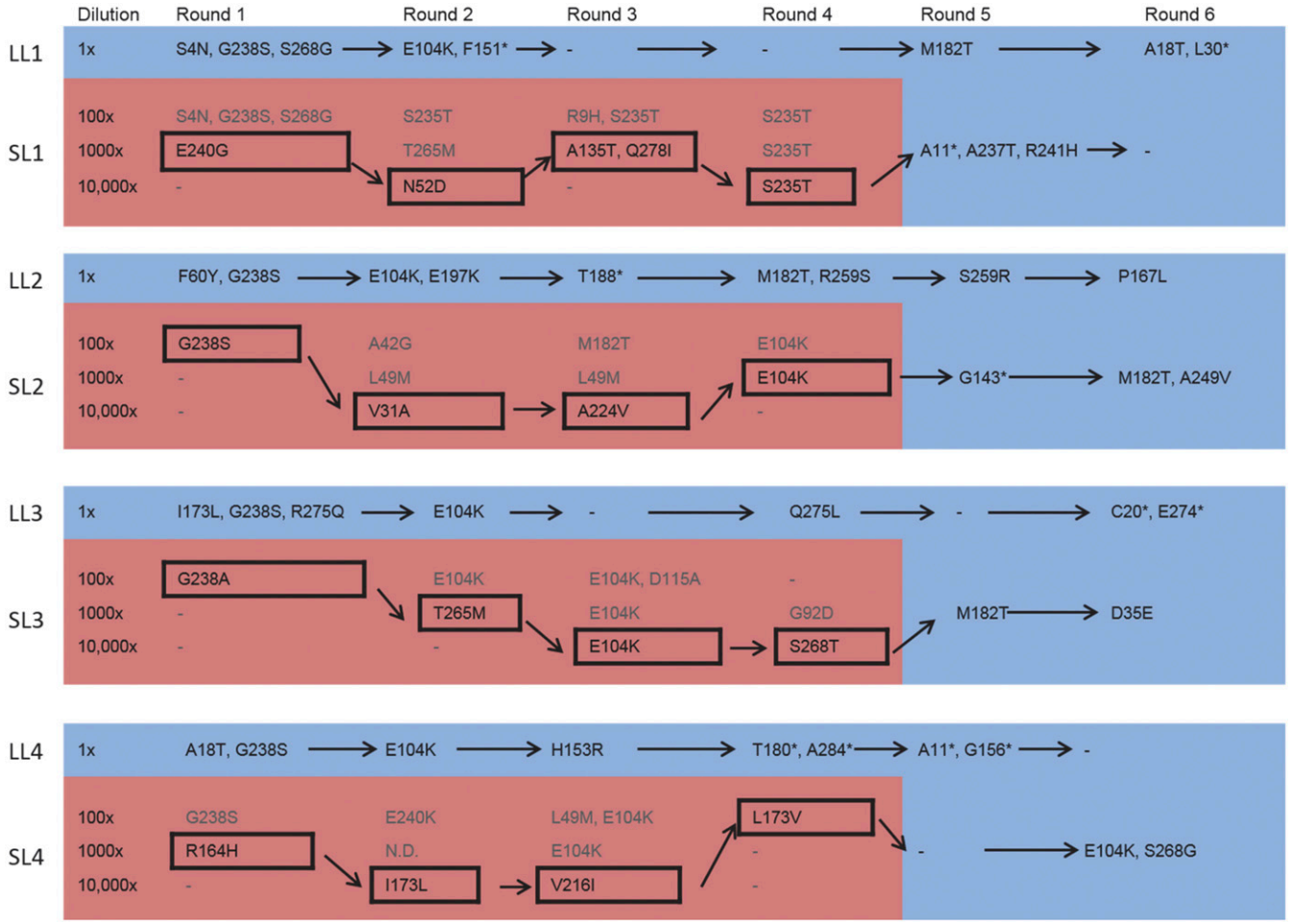

Fig. S2. Overview of mutant library sizes and selected mutations. Entire libraries were used in LL lines, whereas for SL lines we continued with the smallest of three libraries showing at least one nonsynonymous mutation, always from the highest CTX concentration allowing bacterial growth. Blue indicates LL; red indicates SL selection conditions; black boxes show the library size used for SL lines; "-" indicates that no mutations was selected for that line and round.

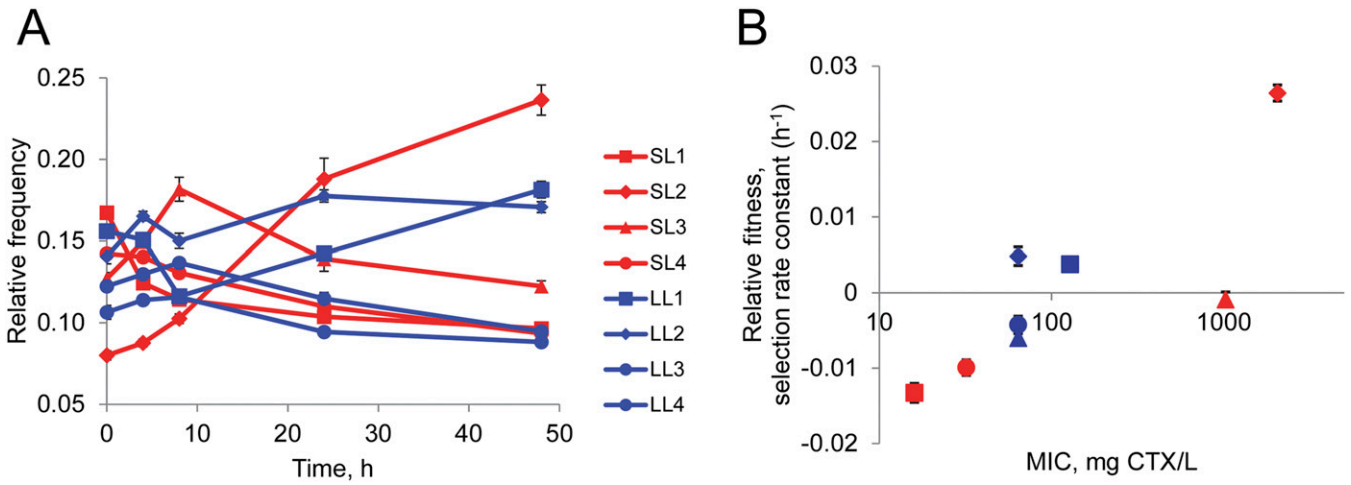

Fig. S3. Relationship between MIC and competitive fitness under selection conditions of the LL lines. (*A*) Frequency estimates of the eight final alleles of the evolution experiment at five time points inferred from Illumina sequencing at ~500,000-fold coverage. (*B*) Correlation between CTX resistance and competitive fitness for the eight final alleles ($r = 0.806$, $P = 0.0157$). Selection rate constants were calculated from the change in frequency of each allele (*A*) relative to that of the seven others across 48 h. Error bars show SEM based on eight replicate competition assays.

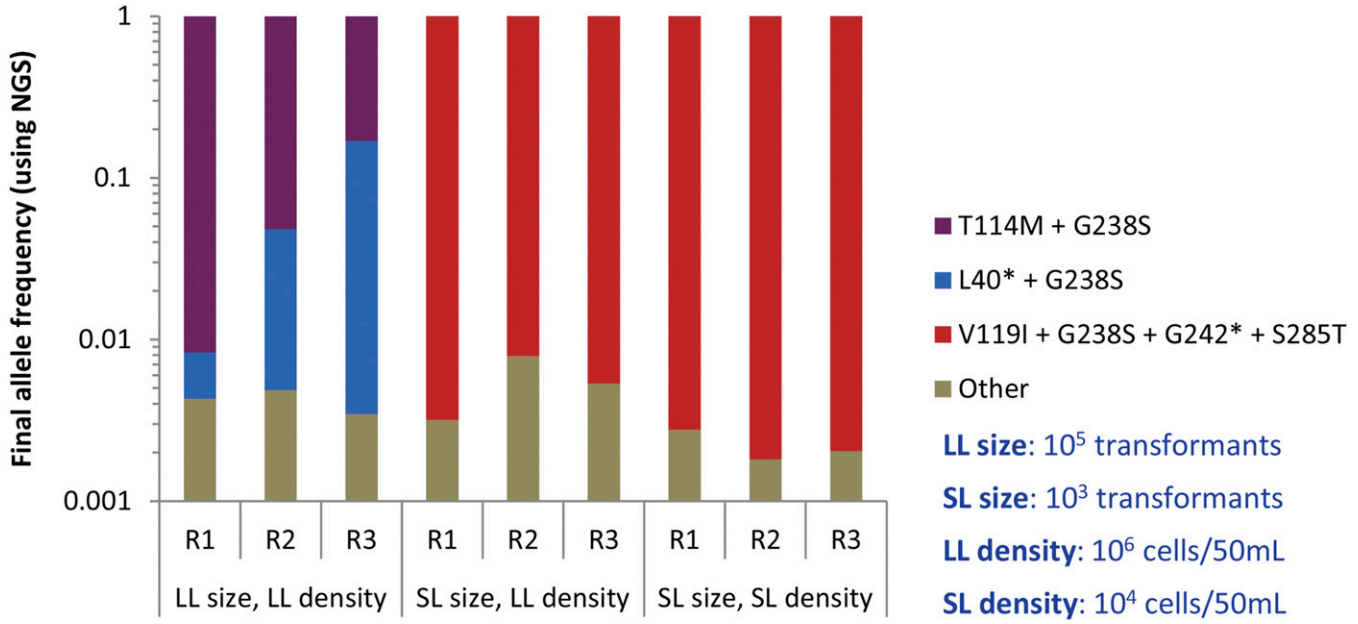

Fig. S4. Available genetic variation, not initial cell density, determines the outcome of selection. Shown are results from the selection of a single error-prone library of TEM-1 β-lactamase alleles at the highest possible [CTX] under three conditions: (*i*) LL mutation supply size and LL cell density; (*ii*) SL mutation supply size and LL cell density; and (*iii*) SL mutation supply size and SL cell density.

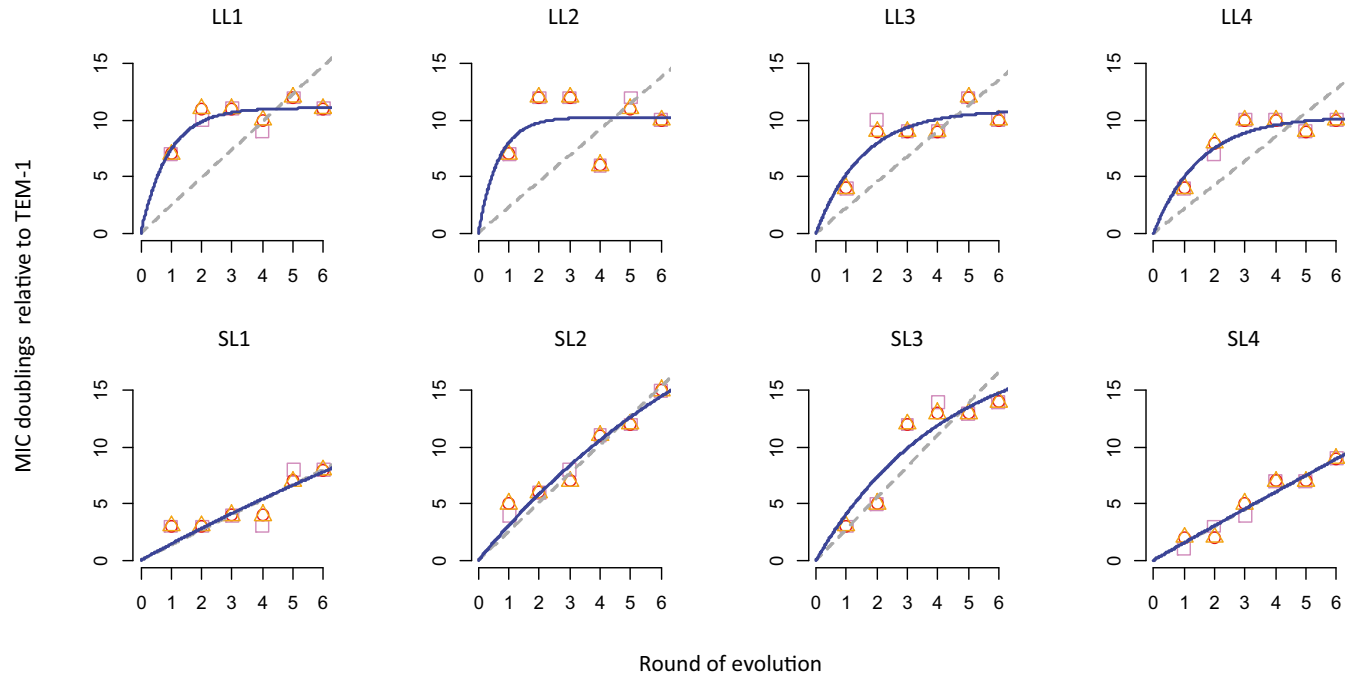

Fig. S5. Test of tradeoff between short- and long-term adaptation. Linear and exponential models were fitted to the MIC-doubling data to test for differences in adaptive dynamics. For all panels, passage number is the abscissa and MIC doublings is the ordinate. The observed MIC values of replicates 1, 2, and 3 are given by the purple squares, red squares, or orange triangles, respectively. The dashed gray line indicates the fitted linear model, and the solid blue line indicates the fitted exponential model. Estimated model parameters are given in Table S1. The exponential model is best supported for all LL lines, but only for two SL lines (SL2 and SL3). Both initial improvement (parameter c of the exponential model; $t_6 = 3.85$, $P = 0.008$) and deceleration (parameter k of the exponential model; $t_6 = -5.94$, $P < 0.001$) are greater for LL than SL lines.

Line	1	2	3	4	5		1	2	3	4	5
			S4*								
										A11T	
	V23*										
		P27S			P27S						P27L
							V31A	V31A	V31A	V31A	V31A
											L57F
	E104K	E104K	E104K	E104K	E104K		E104K	E104K	E104K	E104K	E104K
	M182T	M182T	M182T	M182T	M182T				M182T		M182T
		L190*									
							A224V	A224V	A224V	A224V	A224V
	G238S	G238S	G238S	G238S	G238S		G238S	G238S	G238S	G238S	G238S
			E212K								
								P252*			
										T271N	
	E274D										
				R277I							
									E281D		
medianMIC R1	128	256	512	512	512		256	256	1024	256	512
medianMIC R2	128	512	512	1024	256		64	256	1024	64	256
medianMIC R3	512	512	1024	512	256		256	256	256	128	512

Fig. S6. Three-round replay evolution experiments. Evolution with threefold mutant E104K/M182T/G238S (*Left*) and fourfold mutant V31A/E104K/A224V/G238S (*Right*), representing allele SL2 after four rounds of evolution under LL conditions. Ancestral mutations are gray; mutations found in the first round are yellow; mutations found in the second round are orange. Three rounds of evolution were applied, but no mutations were found in the third round.

Table S1. Model fitting to MIC doubling data

	Linear model		Exponential model				Partial *F* test	
Line	$c \pm$ SEM	RSS[†]	$c \pm$ SEM	$k \pm$ SEM	RSS[†]	t_{50}[‡]	*F*	*P*
LL1	2.47 ± 0.164	244.986	7.44 ± 0.26	0.33 ± 0.03	14.686	0.62	250.879	<0.001**
LL2	2.30 ± 0.252	383.370	8.05 ± 0.44	0.22 ± 0.06	88.948	0.45	52.954	<0.001**
LL3	2.26 ± 0.137	138.015	5.23 ± 0.40	0.52 ± 0.04	21.816	1.06	85.217	<0.001**
LL4	2.13 ± 0.149	121.254	4.96 ± 0.29	0.52 ± 0.03	9.624	1.05	185.528	<0.001**
SL1	1.34 ± 0.030	17.999	1.45 ± 0.22	0.96 ± 0.07	17.716	15.77	0.252	0.622
SL2	2.56 ± 0.051	21.373	3.12 ± 0.22	0.90 ± 0.04	14.594	6.31	7.412	0.014*
SL3	2.77 ± 0.123	80.462	4.11 ± 0.34	0.79 ± 0.04	38.901	2.92	17.093	<0.001**
SL4	1.50 ± 0.026	5.250	1.53 ± 0.14	0.99 ± 0.04	5.227	68.97	0.064	0.803

*Significant at the 0.05 level. **Significant at the 0.001 level.
[†]Residual sum of squares.
[‡]Time in passages until MIC has increased to half of its maximum value as predicted by the model $t_{50} = -\ln(2)/\ln(k)$.