

## **Supplementary Text File 1: The large-effect G238S mutation determines fitness and dictates evolutionary fate**

### **Results**

Simple regression models were made that consider how well fitness can be predicted by (i) CTX resistance, or (ii) the presence of individual beneficial mutations. To consider the relationship between CTX resistance and fitness, Models 1a and 1b were formulated. Model 1a assumes a linear relationship between  $\Delta\text{MIC}$  – the difference in MIC between a genotype and the mean MIC for its node in the landscape – and relative fitness, with a coefficient that scales the relationship. Model 1b limits the range over which the response between  $\Delta\text{MIC}$  and fitness is linear, to capture the effect that beyond a given level of resistance, there is no further effect on fitness. Next to the coefficient to scale the MIC-fitness relationship, this model has a minimum and maximum value of fitness. As an alternative approach, Model 2 divides the genotypes into two fitness classes based on the presence of any one mutation (E104K, M182T or G238S), and assigns each class a fitness value. These models were fitted to the relevant data (Models 1a and 1b:  $\Delta\text{MIC}$  and selection coefficients; Model 2: mutation occurrence and selection coefficients) for each experimental condition separately. Model selection with the Akaike information criterion (AIC) revealed that Model 2, with fitness classes based on the presence of the G238S mutation, provided the best predictions over all conditions (see Table SF1 and Table SF2 below). Support for the different models was comparable in the absence of antibiotics, as the fitness differences between genotypes are minimal. In the presence of CTX, Model 2 was better supported than the other models for 7 out of 8 conditions (Table SF1). Estimated model parameters (Table SF2) also confirm the main trends observed earlier, including a lower fitness of the G238S-carrying variants in structured media at low antibiotic concentration ( $0.02 \mu\text{g mL}^{-1}$  CTX).

Finally, competitive fitness data was used to simulate evolution on TEM-52's fitness landscape, with the goal of predicting evolutionary endpoints. Simulations were kept as close to empirical data as possible, predicting what would happen if competitions proceeded over multiple rounds of passaging. At the start of each passage, the seeding genotype and its three single-mutation neighbours are present at equal frequencies. We used an estimate of the number of generations within a passage and the selection coefficients to predict the final frequencies of genotypes, and then randomly selected a single individual to seed the next round of passaging. TEM-1 was assumed as the starting point for each simulation. As

a contrast to our experimental data, we generated a prediction for conditions in which resistance is the predominant determinant of fitness (see Fig. SF1 below), based on the predicted fitness values from Model 1A for liquid medium with  $0.04 \mu\text{g mL}^{-1}$  of CTX (Table SF2). Under these conditions, the high-resistance triple mutant TEM-52 predominates (Fig. SF1A). Predictions based on our competitive fitness data show different trends. First, in the absence of CTX, all evolutionary endpoints are equally likely, as there are no appreciable differences in fitness (Fig. SF1B, SF1E). Second, in the presence of CTX the evolutionary endpoint depends almost entirely on the G238S mutation: all variants with this mutation are likely endpoints (Fig. SF1C, SF1D and SF1G), except in structured environments with  $0.02 \mu\text{g mL}^{-1}$  of CTX, when all variants without G238S are likely endpoints (Fig. SF1F). Overall, selection of simple models predicting fitness based on our measurements (Table SF1) and simulating multiple passages of evolution under our experimental conditions (Fig. SF1) confirm the expected outcomes based on a first inspection of the data. These results stress the importance of the G238S mutation as the key determinant of fitness in the presence of CTX, and highlight that this mutation stratifies the landscape into low and high fitness variants, which may reverse at low antibiotic concentrations in a structured environment.

## Figures and Tables

**Table SF1: Summary of model selection results.** The Akaike Weight, the likelihood that a model is the best supported model within the set of models tested, is given for each model. CTX indicates the antibiotic concentration ( $\mu\text{g mL}^{-1}$ ) used in the experiment. For model 2, the subscripts indicate the site within the TEM gene used to classify the genotypes. While no model enjoys appreciably higher support in the absence of antibiotics, Model 2<sub>238</sub> is the best supported model for 7 out of 8 conditions with antibiotics. Complete data for the model selection are given in Table SF2.

CTX	Medium	Time	Akaike Weight				
			Model 1a	Model 1b	Model 2 <sub>104</sub>	Model 2 <sub>182</sub>	Model 2 <sub>238</sub>
0	Liquid	24	0.358	0.068	0.229	0.191	0.154
		48	0.424	0.100	0.002	0.002	0.472
	Solid	24	0.238	0.050	0.164	0.087	0.460
		48	0.190	0.551	0.038	0.029	0.193
0.02	Liquid	24	0.019	0.113	0.000	0.000	0.868
		48	0.019	0.108	0.000	0.000	0.874
	Solid	24	0.227	0.528	0.000	0.000	0.245
		48	0.118	0.169	0.000	0.000	0.713
0.04	Liquid	24	0.000	0.000	0.000	0.000	1.000
		48	0.000	0.000	0.000	0.000	1.000
	Solid	24	0.000	0.010	0.000	0.000	0.990
		48	0.000	0.011	0.000	0.000	0.988

**Table SF2: Model selection results.** For each experimental condition (CTX concentration in  $\mu\text{g mL}^{-1}$ , medium type and time of sampling), the five models were fitted. For model 2, the subscripts indicate the site within the TEM gene used to classify the genotypes. Estimated model parameters are given (see Methods section for an explanation of model parameters), as well as the log likelihood, Akaike information criterion (AIC), the difference in AIC compared to the best supported model ( $\Delta\text{AIC}$ ), and the likelihood that a model is the best supported model within the set of models tested (Akaike Weight, AW). The table is divided over three pages, with results for experimental conditions with antibiotics on the next pages. Note that because we obtained a small residual sum of squares, log likelihood values are positive and AIC values are negative. Overall, Model 2<sub>238</sub> is the best-supported model over the different datasets. Note that for intermediate CTX concentrations ( $0.02 \mu\text{g mL}^{-1}$ ) on solid media, estimated model parameters indicate that high resistance alleles appear to be disadvantaged.

CTX	Medium	Time	Model	Parameters	LL	AIC	$\Delta\text{AIC}$	AW
0	Liquid	24	1a	$\alpha = 0$	185.672	-369.344	0.000	0.358
			1b	$\alpha = 0.001, \gamma_{\min} = -0.001, \gamma_{\max} = 0$	186.010	-366.020	3.323	0.068
			2 <sub>104</sub>	$S_1 = 0.001, S_2 = 0$	186.224	-368.448	0.896	0.229
			2 <sub>182</sub>	$S_1 = -0.001, S_2 = 0$	186.043	-368.087	1.257	0.191
			2 <sub>238</sub>	$S_1 = -0.001, S_2 = 0$	185.826	-367.652	1.692	0.154
		48	1a	$\alpha = 0.001$	185.191	-368.383	0.212	0.424
			1b	$\alpha = 0.001, \gamma_{\min} = -0.004, \gamma_{\max} = 0.002$	185.749	-365.498	3.098	0.100
			2 <sub>104</sub>	$S_1 = 0, S_2 = 0$	180.713	-357.425	11.170	0.002
			2 <sub>182</sub>	$S_1 = 0, S_2 = 0$	180.713	-357.425	11.170	0.002
			2 <sub>238</sub>	$S_1 = -0.002, S_2 = 0.002$	186.298	-368.595	0.000	0.472
	Solid	24	1a	$\alpha = 0$	177.534	-353.068	1.316	0.238
			1b	$\alpha = -0.001, \gamma_{\min} = -0.001, \gamma_{\max} = 0.002$	177.983	-349.965	4.418	0.050
			2 <sub>104</sub>	$S_1 = -0.001, S_2 = 0.001$	178.162	-352.324	2.059	0.164
			2 <sub>182</sub>	$S_1 = 0, S_2 = 0$	177.530	-351.061	3.323	0.087
			2 <sub>238</sub>	$S_1 = 0.001, S_2 = -0.001$	179.192	-354.383	0.000	0.460
		48	1a	$\alpha = -0.002$	148.803	-295.606	2.134	0.190
			1b	$\alpha = -0.032, \gamma_{\min} = -0.008, \gamma_{\max} = 0.003$	151.870	-297.740	0.000	0.551
			2 <sub>104</sub>	$S_1 = 0.001, S_2 = -0.004$	148.191	-292.383	5.357	0.038
			2 <sub>182</sub>	$S_1 = -0.004, S_2 = 0.001$	147.930	-291.861	5.879	0.029
			2 <sub>238</sub>	$S_1 = 0.003, S_2 = -0.006$	149.819	-295.639	2.101	0.193

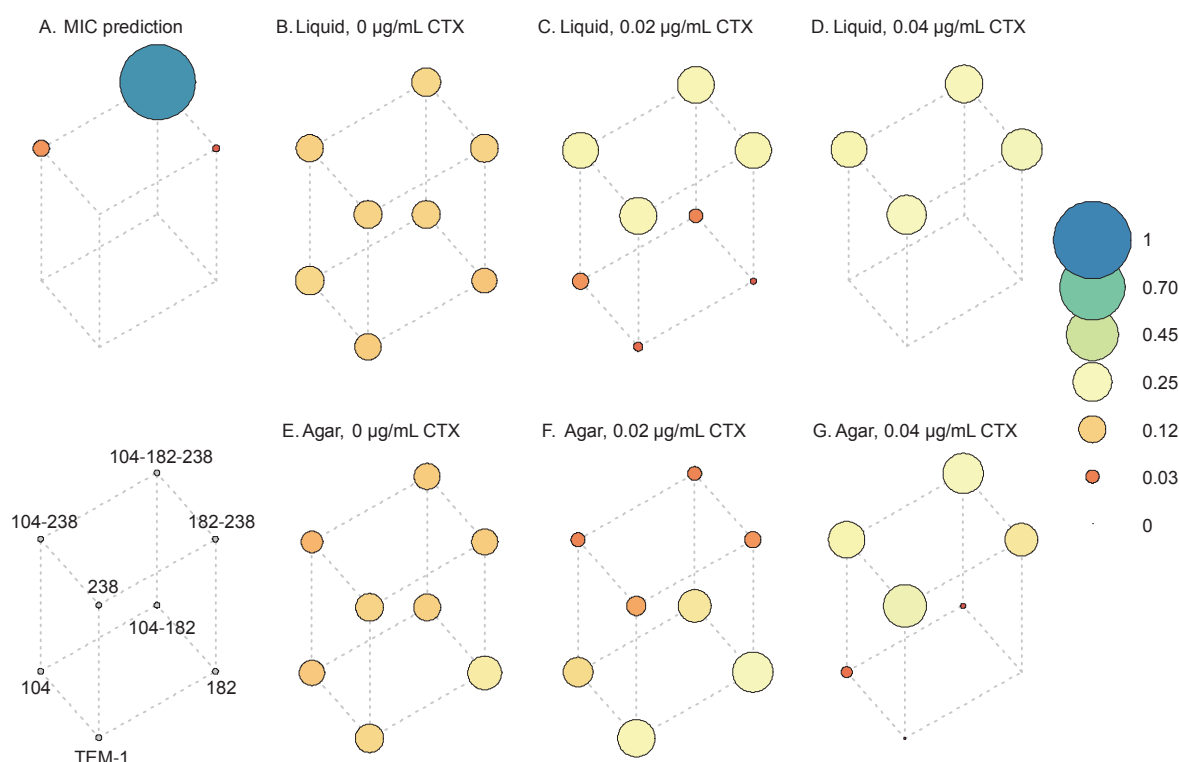
**Table SF2: Model selection results (continued section 2).**

CTX	Medium	Time	Model	Parameters	LL	AIC	$\Delta$ AIC	AW
0.02	Liquid	24	1a	$\alpha = 0.022$	96.152	-190.304	7.647	0.019
			1b	$\alpha = 0.039, \gamma_{\min} = -0.088, \gamma_{\max} = 0.042$	99.939	-193.878	4.073	0.113
			2 <sub>104</sub>	$S_1 = -0.037, S_2 = 0.006$	91.017	-178.035	19.917	0.000
			2 <sub>182</sub>	$S_1 = -0.015, S_2 = -0.015$	88.968	-173.936	24.016	0.000
			2 <sub>238</sub>	$S_1 = -0.060, S_2 = 0.030$	100.976	-197.951	0.000	0.868
		48	1a	$\alpha = 0.021$	96.877	-191.754	7.674	0.019
			1b	$\alpha = 0.038, \gamma_{\min} = -0.085, \gamma_{\max} = 0.041$	100.619	-195.238	4.189	0.108
			2 <sub>104</sub>	$S_1 = -0.036, S_2 = 0.006$	91.635	-179.270	20.158	0.000
			2 <sub>182</sub>	$S_1 = -0.014, S_2 = -0.015$	89.604	-175.209	24.219	0.000
			2 <sub>238</sub>	$S_1 = -0.059, S_2 = 0.030$	101.714	-199.427	0.000	0.874
	Solid	24	1a	$\alpha = -0.008$	145.397	-288.793	1.685	0.227
			1b	$\alpha = -0.011, \gamma_{\min} = -0.022, \gamma_{\max} = 0.034$	148.239	-290.478	0.000	0.528
			2 <sub>104</sub>	$S_1 = 0, S_2 = -0.002$	127.080	-250.160	40.318	0.000
			2 <sub>182</sub>	$S_1 = 0, S_2 = -0.003$	127.151	-250.301	40.177	0.000
			2 <sub>238</sub>	$S_1 = 0.015, S_2 = -0.017$	146.471	-288.942	1.536	0.245
		48	1a	$\alpha = -0.014$	125.708	-249.416	3.598	0.118
			1b	$\alpha = -0.018, \gamma_{\min} = -0.036, \gamma_{\max} = 0.051$	128.067	-250.134	2.880	0.169
			2 <sub>104</sub>	$S_1 = 0, S_2 = -0.008$	110.242	-216.483	36.531	0.000
			2 <sub>182</sub>	$S_1 = -0.005, S_2 = -0.004$	110.003	-216.005	37.009	0.000
			2 <sub>238</sub>	$S_1 = 0.022, S_2 = -0.031$	128.507	-253.014	0.000	0.713

**Table SF2: Model selection results (continued section 3).**

CTX	Medium	Time	Model	Parameters	LL	AIC	$\Delta$ AIC	AW
0.04	Liquid	24	1a	$\alpha = 0.114$	51.964	-101.927	31.132	0.000
			1b	$\alpha = 0.504, \gamma_{\min} = -0.306, \gamma_{\max} = 0.148$	59.293	-112.585	20.473	0.000
			2 <sub>104</sub>	$S_1 = -0.121, S_2 = -0.124$	43.049	-82.098	50.961	0.000
			2 <sub>182</sub>	$S_1 = -0.125, S_2 = -0.120$	43.049	-82.098	50.961	0.000
			2 <sub>238</sub>	$S_1 = -0.355, S_2 = 0.110$	68.529	-133.059	0.000	1.000
		48	1a	$\alpha = 0.110$	52.438	-102.875	31.493	0.000
			1b	$\alpha = 0.518, \gamma_{\min} = -0.296, \gamma_{\max} = 0.145$	59.818	-113.635	20.733	0.000
			2 <sub>104</sub>	$S_1 = -0.114, S_2 = -0.123$	43.746	-83.492	50.875	0.000
			2 <sub>182</sub>	$S_1 = -0.121, S_2 = -0.117$	43.742	-83.485	50.883	0.000
			2 <sub>238</sub>	$S_1 = -0.346, S_2 = 0.109$	69.184	-134.368	0.000	1.000
	Solid	24	1a	$\alpha = 0.066$	67.319	-132.638	17.594	0.000
			1b	$\alpha = 0.122, \gamma_{\min} = -0.266, \gamma_{\max} = 0.116$	73.540	-141.081	9.151	0.010
			2 <sub>104</sub>	$S_1 = -0.100, S_2 = -0.026$	59.474	-114.949	35.283	0.000
			2 <sub>182</sub>	$S_1 = -0.059, S_2 = -0.067$	58.607	-113.215	37.017	0.000
			2 <sub>238</sub>	$S_1 = -0.196, S_2 = 0.070$	77.116	-150.232	0.000	0.990
		48	1a	$\alpha = 0.055$	70.571	-139.141	15.567	0.000
			1b	$\alpha = 0.115, \gamma_{\min} = -0.223, \gamma_{\max} = 0.071$	75.876	-145.753	8.955	0.011
			2 <sub>104</sub>	$S_1 = -0.093, S_2 = -0.018$	64.285	-124.570	30.138	0.000
			2 <sub>182</sub>	$S_1 = -0.053, S_2 = -0.058$	63.085	-122.170	32.538	0.000
			2 <sub>238</sub>	$S_1 = -0.167, S_2 = 0.056$	79.354	-154.708	0.000	0.988

**Supplementary Figure SF1: Simulation of evolution by serial passages.** The cube represents different genotypes on the three-mutation TEM landscape, with genotype noted in the legend in the bottom left. The circle size and color represent the probability that a genotype will be the final genotype selected after 100 passages of 48 h duration, as indicated by the legend on the far right. These simulations capture conditions identical to our setup for the competition experiments, but running over multiple passages instead of single passage. Each passage is initiated by the starting genotype and the three Hamming distance = 1 mutants at equal frequencies, and the empirical estimates of selection coefficients are used to predict population composition at 48 h. A single allele is randomly drawn and used to seed the next passage as a starting genotype.



## Methods

### Models of fitness

We explored three simple models predicting fitness from differences in MIC or from the presence of mutations in TEM. Models 1a and 1b are based on the difference in  $\log[\text{MIC}]$  for a genotype compared to the mean of the four genotypes corresponding to a node ( $\Delta\text{MIC}$ ). Model 1a assumes a linear relationship between  $\Delta\text{MIC}$  and the predicted selection coefficient  $S$ , such that for the  $i^{\text{th}}$  allele  $S_i = \alpha \cdot$

$\Delta MIC_i$ , where  $\alpha$  is a constant estimated from the data for each experimental condition. Model 1b assumes a linear relationship that is constrained to a minimum  $\gamma_{\min}$ , such that if  $\alpha \cdot \Delta MIC_i > \gamma_{\min}$  then  $S_i = \gamma_{\min}$ , as well as a maximum  $\gamma_{\max}$ . Model 1a has one free parameter ( $\alpha$ ), whereas Model 1b has three parameters ( $\alpha$ ,  $\gamma_{\min}$  and  $\gamma_{\max}$ ). Model 2 assigns alleles to two classes depending on the presence of the mutations E104K, M182T or G238S, and assigns each fixed selection coefficients  $S_1$  and  $S_2$ . Model parameters were estimated by minimizing the negative log likelihood (NLL) with grid searches over a broad parameter space, and the NLL was calculated from the residual sum of squares.

### Simulations of serial passaging

To evaluate multi-passage evolution on this landscape, for each condition we performed 1000 independent 100-passage simulations, with the ancestral TEM as the starting genotype. Each passage is initiated by the initial genotype or that sampled in the previous passage, together with the three genotypes with a Hamming distance = 1 in the landscape. The final frequency  $f$  of a genotype  $i$  is  $f_i = (1 + S_i)^g$ , where  $S_i$  is the empirically estimated selection coefficient for that genotype and  $g$  is the number of generations for each passage. Based on the cell counts from flow cytometry, we estimated 15.77 generations had occurred by 48 h, a number that was consistent over the different antibiotic and media treatments (standard error of the mean = 0.21). We therefore assumed  $g = 16$  for all simulations. At the end of growth, one individual is randomly chosen using a pseudorandom number generator (*sample()* in R) to initiate the next passage, with the probability of sampling being weighed by the normalized frequency of each genotype. After 100 passages, we consider the final genotype selected as the evolutionary endpoint for that replicate.

### Code accessibility

Scripts used for statistical models of fitness (Table SF1 and SF2) and for simulations (Fig. SF1) are similarly available via Zenodo in folder “Scripts for supp text file 1” (1)

### References

1. Farr AD, Pesce D, Das SG, Zwart MP, de Visser JAGM. 2022. Supplementary data for "The fitness of beta-lactamase mutants depends nonlinearly on resistance level at sublethal antibiotic concentrations." doi:10.5281/zenodo.6818117, Zenodo.