Supplementary information

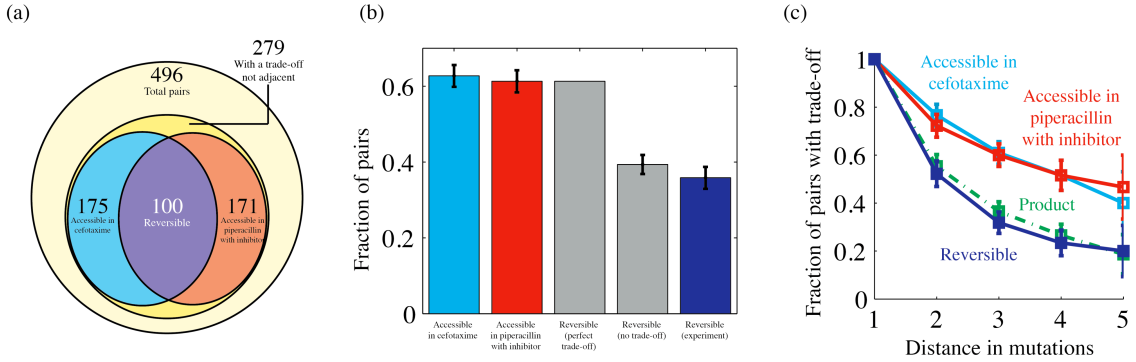

FIG. S1. Allowing neutral mutations to be fixed, reversibility is still comparable to what would be expected if there were no trade-off. (a) (b) (c) Similar to Fig. 3 in the main text, but allowing neutral mutations to fix in both environments. While we allow neutral mutations to fix, we still only consider pairs with a fitness trade-off: we do not count pairs where the start and end alleles confer the same resistance.

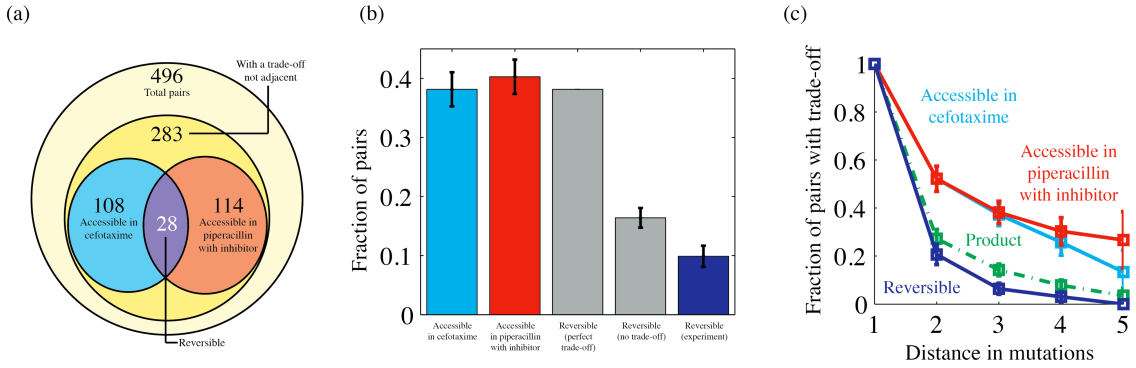

FIG. S2. Using mean log MIC as fitness, reversibility is still comparable to what would be expected if there were no trade-off. (a) (b) (c) Similar to Fig. 3 in the main text, but using mean log MIC instead of median log MIC as fitness. The accessible and reversible fractions are larger in this analysis because there are fewer neutral mutations.

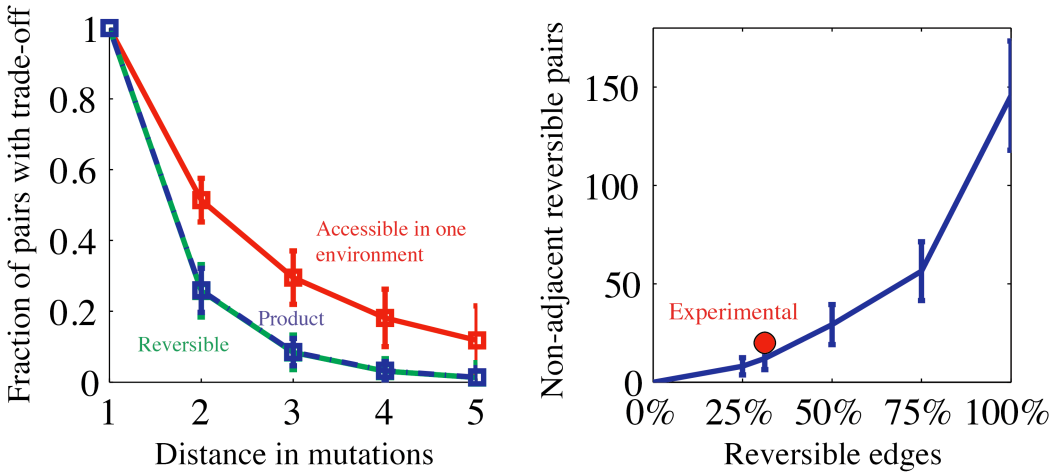

FIG. S3.

(**Left**) Reversible fraction at each distance is the product of accessible fractions in two uncorrelated fitness landscapes. Simulation on 5,000 randomly generated (therefore on average uncorrelated) pairs of fitness landscapes generates the mean accessible and reversible fractions among pairs with a trade-off. Error bars, standard deviations; n = 5,000. The decline of accessibility and reversibility as distance increases is a general property of fitness landscapes. At each distance, the reversible fraction is within two times the standard error of the product of accessible fractions in the two environments.

(**Right**) Experimental reversibility is similar to reversibility on landscapes with the same degree of local correlation. As a measure of local correlation, we analyzed the percentage of 'reversible' edges on the hybercube, that is, the percentage of edges along which selection acts in opposite directions in the two environments (corresponding to the probability that a mutation has opposite effects on a given genetic background). Each point on the blue curve shows the mean reversibility level for randomly generated landscape-pairs that have 0%, 25%, 31.25%, 50%, 75% or 100% reversible edges. Error bars: standard deviations; n = 1000 for 31.25%, n = 200 otherwise. Our experimental hypercube has 31.25% reversible edges, and its reversibility is represented by the red dot. The mean number of non-adjacent reversible pairs is 12.0 ± 5.7 (standard deviation; n = 1000) for landscapes with 31.25% reversible pairs. Thus our experimental result (20 pairs) is within two times the standard deviation from the mean level for these landscapes.

TABLE I. Resistance of each allele to different antibiotics, measured by logarithm base √2 of minimum inhibitory concentration (MIC [μg/ml]).

Strain	Cefotaxime		Weinreich	Piperacillin		Piperacillin with inhibitor		
	median	mean	median	median	mean	median	mean	
- - - - -	-8.00	-8.00 ± 0.50	-7.00	24.00	24.00 ± 0.50	10.00	10.63 ± 0.74	n= 12
- - - - +	-1.00	-0.67 ± 0.60	1.00	17.00	17.00 ± 0.50	0.00	0.33 ± 0.42 †	
- - - + -	-8.00	-7.67 ± 0.60	-8.00	24.00	24.00 ± 0.50	16.25	16.19 ± 0.64	n= 8
- - - + +	13.00	12.67 ± 0.60	10.00	24.00	24.00 ± 0.50	1.00	0.83 ± 0.30 †	
- - + - -	-6.00	-6.00 ± 0.50	-6.00	24.00	24.00 ± 0.50	16.00	16.23 ± 0.81	n= 11
- - + - +	15.00	14.67 ± 0.60	17.00	18.00	18.00 ± 0.50	0.00	0.17 ± 0.30 †	
- - + + -	-5.00	-5.33 ± 0.60	-5.00	24.00	24.00 ± 0.50	23.00	22.80 ± 0.36 †	n= 5
- - + + +	20.00	20.00 ± 0.50	17.00	24.00	23.67 ± 0.60	1.00	1.00 ± 0.25 †	
- + - - -	-6.00	-6.33 ± 0.60	-7.00	24.00	24.00 ± 0.50	13.00	13.06 ± 0.63	n= 16
- + - - +	13.00	13.33 ± 0.60	9.00	24.00	23.67 ± 0.60	0.50	0.50 ± 0.25 †	
- + - + -	-7.00	-7.33 ± 0.60	1.00	24.00	24.00 ± 0.50	13.50	14.04 ± 0.70	n= 12
- + - + +	15.00	15.00 ± 0.50	17.00	23.00	23.00 ± 0.50	0.50	0.33 ± 0.30 †	
- + + - -	0.00	0.33 ± 0.60	1.00	24.00	24.00 ± 0.50	16.00	16.23 ± 0.70	n= 13
- + + - +	20.00	20.00 ± 0.50	22.00	23.00	23.00 ± 0.50	0.50	0.33 ± 0.30 †	
- + + + -	-1.00	-1.33 ± 0.60	-1.00	24.00	24.00 ± 0.50	15.00	15.60 ± 0.62	n= 5
- + + + +	21.00	21.00 ± 0.50	23.00	24.00	23.67 ± 0.60	1.00	1.00 ± 0.25 †	
+ - - - -	-7.00	-6.67 ± 0.60	-7.00	24.00	24.00 ± 0.50	19.25	19.08 ± 0.77	n= 6
+ - - - +	0.00	0.33 ± 0.60	1.00	19.00	18.67 ± 0.60	0.00	0.17 ± 0.30 †	
+ - - + -	-8.00	-7.67 ± 0.60	-7.00	24.00	24.00 ± 0.50	23.50	23.60 ± 0.27 †	n= 5
+ - - + +	14.00	14.00 ± 0.50	17.00	24.00	24.00 ± 0.50	1.50	1.67 ± 0.30 †	
+ - + - -	-4.00	-4.00 ± 0.50	-5.00	24.00	24.00 ± 0.50	23.50	23.67 ± 0.27 †	n= 6
+ - + - +	15.00	15.00 ± 0.50	17.00	17.00	17.00 ± 0.50	0.50	0.67 ± 0.30 †	
+ - + + -	-5.00	-4.67 ± 0.60	-5.00	24.00	24.00 ± 0.50	24.00	23.83 ± 0.27 †	n= 6
+ - + + +	20.00	20.00 ± 0.50	22.00	24.00	24.00 ± 0.50	1.50	1.83 ± 0.42 †	
+ + - - -	-8.00	-7.67 ± 0.60	-7.00	24.00	24.00 ± 0.50	22.50	21.98 ± 0.61	n= 29
+ + - - +	15.00	15.00 ± 0.50	17.00	24.00	24.00 ± 0.50	1.00	1.17 ± 0.30 †	
+ + - + -	-8.00	-7.67 ± 0.60	-7.00	24.00	24.00 ± 0.50	18.00	18.22 ± 0.64	n= 9
+ + - + +	15.00	15.33 ± 0.60	17.00	24.00	24.00 ± 0.50	1.50	1.33 ± 0.30 †	
+ + + - -	2.00	1.67 ± 0.60	2.00	24.00	24.00 ± 0.50	23.50	23.75 ± 0.28 †	n= 8
+ + + - +	20.00	20.33 ± 0.60	21.00	24.00	23.67 ± 0.60	1.50	1.67 ± 0.30 †	
+ + + + -	2.00	1.67 ± 0.60	1.00	24.00	24.00 ± 0.50	24.00	23.93 ± 0.28 †	n= 7
+ + + + +	22.00	22.00 ± 0.50	24.00	24.00	24.00 ± 0.50	2.00	1.83 ± 0.30 †	

†Measured to a factor of $\sqrt[4]{2}$. Otherwise, measured to a factor of √2.

All alleles were transformed into *E. coli* strain DH5α through clone vector pBR322. First, a series of √2-fold dilutions of cefotaxime and piperacillin were made in LB. Each strain was cultured at 37°C in 1 ml LB with 5 μg/ml piperacillin for 8 h with 500 rpm shaking, and then diluted to ~ 2×10^6 cells/ml in 200 μl antibiotic solution. For the inhibitor environment, 0.5 μg/ml clavulanic acid was also added. Bacterial solutions were then cultured at 37°C in 96-well plates for 20 h with 500 rpm shaking. Finally, MIC was determined by the lowest concentration that ceased bacterial growth (optical density < 0.3). MIC measurement of each strain was done in triplicate. Some strains exhibited large

variations, so they were measured again in triplicate with a higher resolution ($\sqrt[4]{2}$-fold antibiotic dilutions).

Five + or - represent the presence or absence of, in order, g4205a, A42G, E104K, M182T and G238S (D. M. Weinreich, N. F. Delaney, M. A. DePristo, and D. L. Hartl, *Science* **312**, 111-114 (2006)). For each mean value, the square of the error is estimated by summing the square of the standard errors (n = 3 unless specified) and the square of the average error due to the digitization of measurement. For values measured to a factor of $\sqrt[4]{2}$ (labeled with †), the digital error is estimated to be 0.25. Otherwise, the value is measured to a factor of $\sqrt{2}$, and its digital error is estimated to be 0.50. For the piperacillin with inibitor environment, we use piperacillin with 0.5 μg/ml clavulanic acid.

For MIC measurements in cefotaxime, results from Weinreich et al. (measured at 35°C by visual turbidity inspection) are shown as a comparison. DNA sequencing was conducted for every strain to confirm that our genotypes are the same as Weinreich et al. The relative large discrepancy for strain -+-+- is probably due to the strain's temperature sensitivity (we used 37°C, which is standard for bacterial culture). We repeated this experiment a full 12 times, and obtained consistent results.

TABLE II. Among 279 non-adjacent pairs with a trade-off, only 20 are reversible.

More fit in piperacillin with inhibitor	More fit in cefotaxime	Distance in mutations
- - -+ -	-+ - - - *	2
- - -+ -	-+ - -+ *	3
- - -+ -	-+ -++ *	2
- -++ -	-++ -+	3
- -++ -	-++++ *	2
-+ -+ -	-+ - -+ *	2
-++ - -	-++++	2
-+++ -	-++ -+	2
+ - -+ -	+ - - -+ *	2
+ - -+ -	++ - -+	3
+ - -+ -	++ -++	2
+ -++ -	+ -+ -+ *	2
++ - - -	-+ - -+ *	2
++ - - -	-+ -++	3
++ - - -	+ - - -+ *	2
++ - - -	++ -++ §	2
++ -+ -	-+ - - - *	2
++ -+ -	-+ - -+ *	3
++ -+ -	-+ -++ *	2
++ -+ -	++ - -+ §	2

*At least one path is accessible in both directions.
§Reversible only by gaining and subsequently losing an additional mutation.

Five + or - represent the presence or absence of, in order, g4205a, A42G, E104K, M182T and G238S. In 12 of the 20 reversible pairs, at least one path is accessible in both directions. Two pairs are reversible only by gaining and subsequently losing an additional mutation.