Supplemental Information

Summary

Below find additional treatments of sub-topics which may be relevant to material offered in main text. These include a supplemental discussion, methods and results.

The initial choice of traits for the study of epistasis: drugless growth, IC_{50}, and intracellular abundance

The biology underlying how the measured traits (drugless growth, IC_{50}, intracellular abundance) relate to drug resistance is well-studied and reasonably intuitive. We initially studied all three traits, as all are present in analytical equations that describe the components of drug resistance (Rodrigues *et al.* 2016). Drugless growth rate is synonymous to fitness of an organism in the absence of drug. In order to be resistant to Trimethoprim, a given microbe must demonstrate some baseline ability to grow. In this system, we expect drugless growth to be lower and less variant across genotypic contexts, indicative of a trait with relatively little higher-order epistasis. This expectation comes from our knowledge of the biology of the system: plasmids were used to express the DHFR mutants in the background bacterial strain in order to measure abundance, IC_{50} and drugless growth. In almost all strains, the simple presence of the plasmid was burdensome to the background strain, almost independent of which species of DHFR was being expressed, or what the PQC genetic background was. Consequently, the drugless growth trait provides something analogous to a negative control, a trait that should be relatively bereft of higher-order epistasis, as all bacterial strains carrying plasmid had a similar low growth rate.

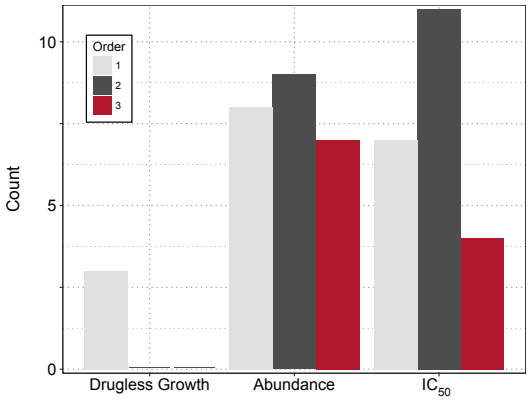

Figure S1. Based on the Bayesian inference criterion (BIC), drugless growth rate has no higher-order interactions, and very few significant main effect drivers. This is in stark contrast to the IC_{50} and abundance phenotypes, both which of contain several higher-order interactions.

Measurement of drugless growth rate

Bacterial cultures were grown overnight (37 °C) in M9 minimal medium were normalized to an OD of 0.1 with fresh medium. When appropriate, GroEL overexpression and/or increase in DHFR concentrations were induced by adding arabinose and IPTG immediately after normalization. After additional growth during 5-6 hours a new normalization to an OD = 0.1 was performed before inoculation of 96-well plates (1/5 dilution) containing M9 medium. Growth was quantified by integration of the area under the growth curve.

Inference of linear genotype-phenotype maps

Epistasis can be erroneously inferred when the underlying genotype does not map linearly to the corresponding phenotypic effects. We checked for potential non-linearity of the genotype-phenotype map by evaluating the relationship between the observed phenotypes and the predicted values from an additive model ($Y \sim S + C + P21L + A26T + L28R$), since deviations from a one-to-one relationship between these two would suggest the presence of 'global' epistasis (as in Sailer and Harms (2017); Otwinowski *et al.* (2018)). We found no strong evidence of non-linear genotype-phenotype maps in either IC_{50} or protein abundance (Figure S2). The best-fit polynomial splines are close to the identity line, especially for IC_{50}.

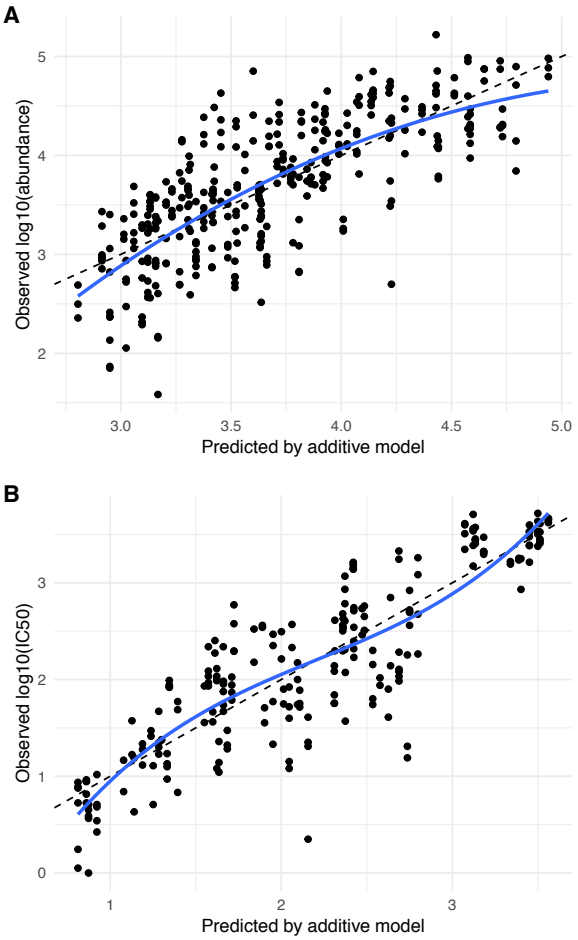

Figure S2. Relationship between the (A) abundance and (B) IC_{50} observed and their values predicted by an additive linear model. The blue lines represent cubic splines fit to the points, the identity line is dashed.

To verify that our overall conclusions are not affected by the choice of phenotypic scale, we linearized the observed abundance (back-transforming from Box-Cox transformation; see Sailer and Harms (2017) Eq.2) and used these values in a new regularized regression analysis. We found no qualitative differences in the results (Figure S3).

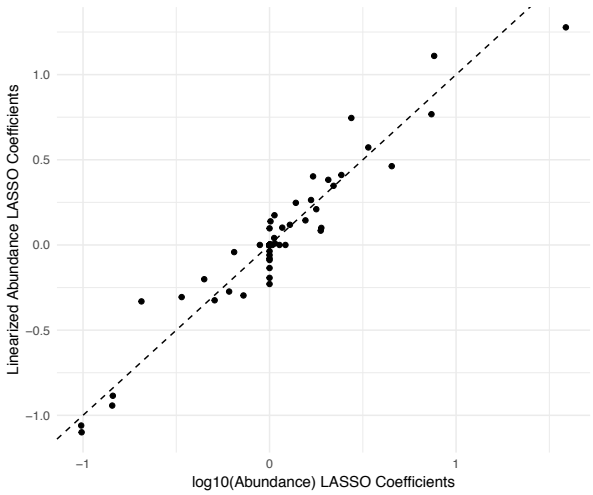

Figure S3. There are only small quantitative differences between the coefficients inferred for abundance (horizontal axis, corresponds to results in main text, Figure 3) and those inferred using a linearized version of the same phenotype.

Notes on the biochemistry and biophysics of the study system

Prior studies have established that the deleterious effect of destabilizing DHFR mutations can be alleviated by the action of the protein quality control(PQC) machinery (Bershtein *et al.* 2013). Specifically, GroEL/ES chaperonins and Lon proteases were shown to be major modulators of the total intracellular DHFR abundance, acting upon partially unfolded protein intermediates to either promote folding or proteolytic degradation, respectively. The impact of PQC background on fitness is particularly relevant in cases where drug-resistance DHFR mutations are associated with stability trade-offs (Rodrigues *et al.* 2016) and in scenarios of horizontal gene transfer (Bershtein *et al.* 2015).

Though the IC_{50} values utilized in this study are laboratory derived, prior studies have identified relationships between IC_{50} and biochemical and biophysical parameters. Rodrigues et al. Rodrigues *et al.* (2016) described such an analytical expression:

$$IC_{50} = \frac{1}{\alpha\beta} \frac{\frac{k_{cat}^n}{k_m^n}}{\frac{k_{cat}^{E.coliWt}}{k_m^{E.coliWt}}} \frac{Abundance}{\gamma ANS} K_i^n - \frac{1}{\alpha} K_i^n$$

We do not use the above equation in this study, and consequently, are providing it in the supplemental information only to highlight that a mathematical relationship has been proposed that links these traits analytically. Mechanistically, we can most simplistically summarize their relationship this way: in order for a population of bacteria to grow in the presence of trimethoprim (which IC_{50} is a presumptive measure of), they must be functional cells that can growth without drug and must produce enough DHFR (the target of Trimethoprim) such that the normal metabolic functions of DHFR are performed. If only small amounts of DHFR are produced, then we can expect low amounts of drug to sufficiently limit growth (low IC_{50}).

Regarding antifolates and the evolution of resistance

The study focused on dihydrofolate reductase, an essential enzyme and target of antibiotics. Though the focus of the study was more general (about resolving epistastic effects across genotypic contexts), the specific biological problem of antifolate resistance did warrant a more detailed examination, which we provide here. Antifolates are used clinically as treatments for a wide range of diseases, ranging from bacteria, to protozoal diseases and as anticancer agents (Bershtein *et al.* 2015; Schnell *et al.* 2004; Kompis *et al.* 2005; Liu *et al.* 2013). These compounds interfere with one of two steps in the *de novo* biosynthetic pathway of tetrahydrofolate (THF), essential for the production of purines and of several amino acids. The genetic basis for antifolate resistance evolution in bacteria lies in a small number of missense mutations in several genes, one of which is dihydrofolate reductase (DHFR). Previous studies had identified that three mutations (A26T, P21L, L28R) that are often found in present various combinations and have an effect on trimethoprim resistance (an antifolate) in *E. coli* (Toprak *et al.* 2012).

Regarding the implications of the results for the study of antibiotic resistance

We should very briefly highlight the results in light of their implications for the study of drug resistance. As previously described, the study system was bacterial DHFRs, the protein target of antifolate drugs. For future efforts at resistance management, we should be clear about what contextual details influence the phenotypic consequences of resistance-associated SNPs before we fully conclude how a given set of SNPs drives resistance evolution in nature. For example, we should be mindful of how off-target mutations (like the ones that constitute species background and proteostasis machinery in this study) might influence patterns of resistance. Related questions about the evolution of resistance are the object of current inquiry from several of the co-authors of this study.

Study limitations

As with any study making a general claim about an important problem (epistasis in this case), there is the potential critique that the results of a study "do not generalize." We remind the authors of such criticism that the study system focused on traits related to antibiotic resistance, a phenotype with biomedical implications. That being the case, even if the methods and results were only relevant to the problem of antibacterial resistance to antifolate compounds (and did not generalize further), we would consider the findings to be relevant for several scientific and biomedical communities. We are, however, confident that the methods and results are reflective of phenomena present in complex traits across the biosphere.

Minor notes on methods to detect epistasis

Several studies that measure epistasis utilize data sets where multiple mutations are constructed in all possible combinations, often in the guise of a graph called a fitness (or adaptive) landscape (Greene and Crona 2014; Ferretti *et al.* 2016; Ogbunugafor *et al.* 2016; Sailer and Harms 2017; Weinreich *et al.* 2018). For data sets where variation at sites of interest is biallelic, these combinatorial sets are composed of 2^L mutations, where L is the number of different loci being examined. The mutations that compose the combinatorial set might have originated from experimental evolution (Chou *et al.* 2014; Toprak *et al.* 2012) or from field surveys (Projecto-Garcia *et al.* 2013; Domyan *et al.* 2014; Natarajan *et al.* 2018). Regardless of their source, several methods have been introduced to detect the presence of higher-order interactions between mutations in data sets similar to the one in this study. One notable method involves the Fourier-Walsh transformation to generate terms corresponding to epistatic interactions between biallelic sites in a fitness graph (Weinreich *et al.* 2013;

Poelwijk *et al.* 2016; Weinreich *et al.* 2018). Others have used regression methods similar to the ones in this study (Otwinowski and Plotkin 2014; Poelwijk and Ranganathan 2017).

As with several of these methods, the epistatic decomposition methods utilized in this study incorporate experimental noise, do not require biallelic loci, and can accommodate missing data. The relaxing of the biallelic loci constraint is especially important for this study: while the individual SNP loci in the data set can be characterized as biallelic (P21L, A26T, L28R), the species context (*Escherichia coli*, *Chlamydia muridarum*, and *Listeria grayi*) and protein quality control genetic background (wild-type, GroEL+, and Δ*lon*) are each composed of three variants per locus.