

Evolutionary druggability: leveraging low-dimensional fitness landscapes towards new metrics for antimicrobial applications

Reviewed Preprint

Revised by authors after peer review.

[About eLife's process](#)

Reviewed preprint version 2

November 30, 2023 (this version)

Reviewed preprint version 1


August 14, 2023

Sent for peer review

April 25, 2023

Posted to bioRxiv

April 8, 2023

Rafael F. Guerrero, Tandin Dorji, Ra'Mal M. Harris, Matthew D. Shoulders, C. Brandon Ogbunugafor 

Department of Biological Sciences, North Carolina State University • Department of Mathematics and Statistics, University of Vermont, Burlington, VT • Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA • Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT • Santa Fe Institute, Santa Fe, NM • Public Health Modeling Unit, Yale School of Public Health, New Haven, CT

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

The term “druggability” describes the molecular properties of drugs or targets in pharmacological interventions and is commonly used in work involving drug development for clinical applications. There are no current analogues for this notion that quantify the drug-target interaction with respect to a given target variant’s sensitivity across a breadth of drugs in a panel, or a given drug’s range of effectiveness across alleles of a target protein. Using data from low-dimensional empirical fitness landscapes composed of 16 β -lactamase alleles and seven β -lactam drugs, we introduce two metrics that capture (i) the average susceptibility of an allelic variant of a drug target to any available drug in a given panel (“*variant vulnerability*”), and (ii) the average applicability of a drug (or mixture) across allelic variants of a drug target (“*drug applicability*”). Finally, we (iii) disentangle the quality and magnitude of interactions between loci in the drug target and the seven drug environments in terms of their mutation by mutation by environment ($G \times G \times E$) interactions, offering mechanistic insight into the variant variability and drug applicability metrics. Summarizing, we propose that our framework can be applied to other datasets and pathogen-drug systems to understand which pathogen variants in a clinical setting are the most concerning (low variant vulnerability), and which drugs in a panel are most likely to be effective in an infection defined by standing genetic variation in the pathogen drug target (high drug applicability).

eLife assessment

This manuscript introduces two **valuable** new metrics - "variant vulnerability" and "drug applicability" - that would be of use to identify candidate drugs for treating infections while considering longer-term, evolution-based treatment outcomes. Despite the intuitive appeal of the metrics and their potential, the study remains **incomplete**, as it fails to demonstrate the generality of the approach. The work could be improved by analysing a broader range of data in a systematic way and directly tying the metrics to outcomes, which would make it possible to better assess their impact and utility.

Introduction

Evolutionary concepts have long been applied to the problem of antimicrobial resistance. For example, our understanding of how resistance evolves has improved through studies of protein evolution ([Weinreich et al., 2006](#); [Lozovsky et al., 2009](#); [Toprak et al., 2011](#)), reversion ([Tan et al., 2011](#); [Baym et al., 2016](#); [Ogbunugafor and Hartl, 2016](#)), and higher-order interactions between drugs ([Michel et al., 2008](#); [Yeh et al., 2009](#)) or mutations ([Weinreich et al., 2013](#); [Lozovsky et al., 2021](#)). While stepwise, de novo evolution (via mutations and subsequent selection) is a critical force in the evolution of antimicrobial resistance, evolution in natural settings often involves other processes, including horizontal gene transfer ([Bergstrom et al., 2000](#); [Turner et al., 2014](#); [Loftie-Eaton et al., 2017](#)) and selection on standing genetic variation ([Pennings, 2012](#); [Pelleau et al., 2015](#)). Consequently, perspectives that consider variation in pathogens (and their drug targets) are important for understanding treatment at the bedside. Recent studies have made important strides in this arena. Some have utilized large data sets and population genetics theory to measure cross-resistance and collateral sensitivity ([Gjini and Wood, 2021](#); [Ardell and Kryazhimskiy, 2021](#); [Herencias et al., 2021](#); [Nichol et al., 2019](#)). Fewer studies have used evolutionary concepts to establish metrics that apply to the general problem of antimicrobial treatment on standing genetic variation in pathogen populations, or for evaluating the utility of certain drugs' ability to treat the underlying genetic diversity of pathogens ([Yeh et al., 2006](#), [2009](#); [Mira et al., 2021](#)).

Despite these gaps in the world of antimicrobial resistance, a robust literature in immunobiology—specifically concerning the problem of evaluating broadly neutralizing antibodies to counteract the widespread genetic diversity of viral infections—has addressed analogous concepts. Studies have profiled mutations in influenza that allow them to escape antibodies ([Doud et al., 2017](#), [2018](#)), and reconstructed “binding affinity landscapes” that define the constraints around viral escape ([Phillips et al., 2021](#), [2023](#)). These studies, which focus on both the anti-viral antibody and their targets (viruses), demonstrate that metrics capturing the susceptibility of a target, or efficacy of an antimicrobial actor, are highly relevant in our quest to prevent, diagnose, and treat infectious diseases.

This study examines data from protein fitness landscapes across a panel of antimicrobial drugs to develop new metrics for antimicrobial applications. Specifically, we propose novel druggability framings, a concept that relates to the interaction between drugs and their targets. We propose an extension of this concept in the form of two metrics: (i) The average susceptibility of alleles of a drug target across drugs in a panel (“variant vulnerability”) and (ii) the efficacy of drugs across alleles of a given drug target (“drug applicability”). Lastly, we statistically deconstruct the peculiar interactions that underlie these metrics: interactions between drugs and the genetic variants at

their target (β -lactamases in this case). We label this last analysis as measuring the “environmental epistasis” in the drug-target interaction, a concept developed recently to describe molecular evolution in the presence of environmental change ([Lindsey et al., 2013](#)).

Summarizing, we propose that these new metrics can be applied to pathogen-drug datasets to identify challenges associated with specific variants of pathogens, and evaluate the efficacy of certain drug types. Further, our results show how modern concepts in evolutionary genetics, like gene by gene by environment interactions (environmental epistasis) have practical utility in biomedicine.

Results

Here, we describe several analyses, emphasizing two novel metrics we have developed: variant vulnerability and drug applicability. We first plotted the growth rates for the 16 alleles in the TEM-1/TEM-50 fitness landscapes across the seven drug environments ([Figure 1](#)). We then computed the variant vulnerability and drug applicability metrics and deconstructed the metrics by measuring the interactions between mutations and drug environments (including G x G x E interactions).

Variant vulnerability

Low variant vulnerability implies low susceptibility of a genotype to the library of drugs in the panel ([Figure 2](#)). Of the alleles in our set, 0110 (MKS_N) displayed the lowest susceptibility with resistance to 3/7 drugs—ceftazidime, cefotetan, and cefotaxime. Alternatively, the allele with the highest variant vulnerability was a triple mutant, 0111, corresponding to the MKS_D allele.

We can rank alleles from highest to lowest for a different interpretation of the variant vulnerability values ([Table 1](#)). Notably, the allele that had the highest variant vulnerability (0111, MKS_D), and the one with the lowest variant vulnerability (0110, MKS_N) differ by a single mutation (N276D). That a single mutation can significantly impact variant vulnerability highlights how single mutations dramatically affect the topography of the variant vulnerability fitness landscape, with the N276D mutation negatively affecting the susceptibility through interactions with the genetic background. Also, note that the alleles associated with TEM-1 (MEG_N) and TEM-50 (LKS_D) have variant vulnerability values that are close (ranked 12th and 11th respectively; among the lowest values).

How do we interpret the low variant vulnerabilities of TEM-1 and TEM-50—indicative of low susceptibility to the panel of drugs? Whatever the selective pressures that drove the evolution of TEM-1 and TEM-50 in natural settings, these findings suggest that the resulting resistance phenotypes associated with evolved β -lactamases seem to confer general survivability across drugs. This observation is consistent with prior notions of “cross resistance” ([Oz et al., 2014](#); [de Souza Cândido et al., 2019](#); [Colclough et al., 2019](#)).

One-step neighbor variant vulnerability

Our findings regarding large variant vulnerability differences between mutational neighbors like MKS_N (0110; the lowest variant vulnerability value of the 16 alleles) and MKS_D (0111; the highest variant vulnerability value of the 16 alleles) inspired further analysis of the relationship between the variant vulnerability of an allelic variant and its mutant neighbors. As the mutants analyzed in this study compose a combinatorically complete set, they can be depicted in the formal structure of a canonical fitness graph that outlines changes between TEM-1 and TEM-50 ([Figure 1](#)). And while this study focuses on the individual properties of variants/alleles of TEM-1/TEM-50, one might ask how the variant vulnerability values are distributed across the fitness graph, which may

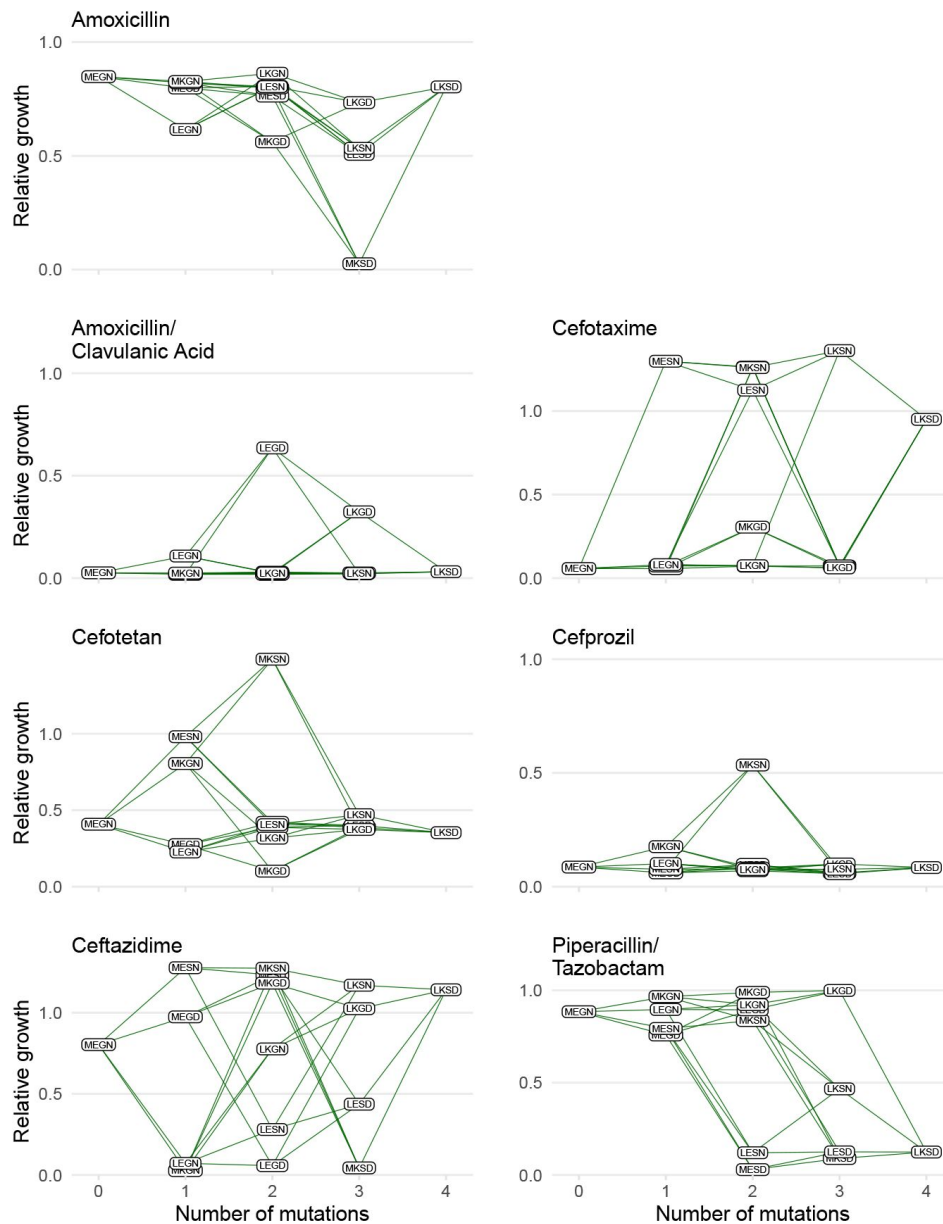


Figure 1.

Graphical depictions of the TEM1/TEM-50 fitness landscape examined in this study. The TEM-50 variant is represented by the quadruple mutant, LKSD, and the wild-type, MEGN, represents TEM-1. The figure highlights the variation in the topography of the various fitness landscapes. Y-axes correspond to the growth rates relative to the wild-type of the allelic variants. The X-axis corresponds to the number of mutations away from the TEM-1 wild-type MEGN.

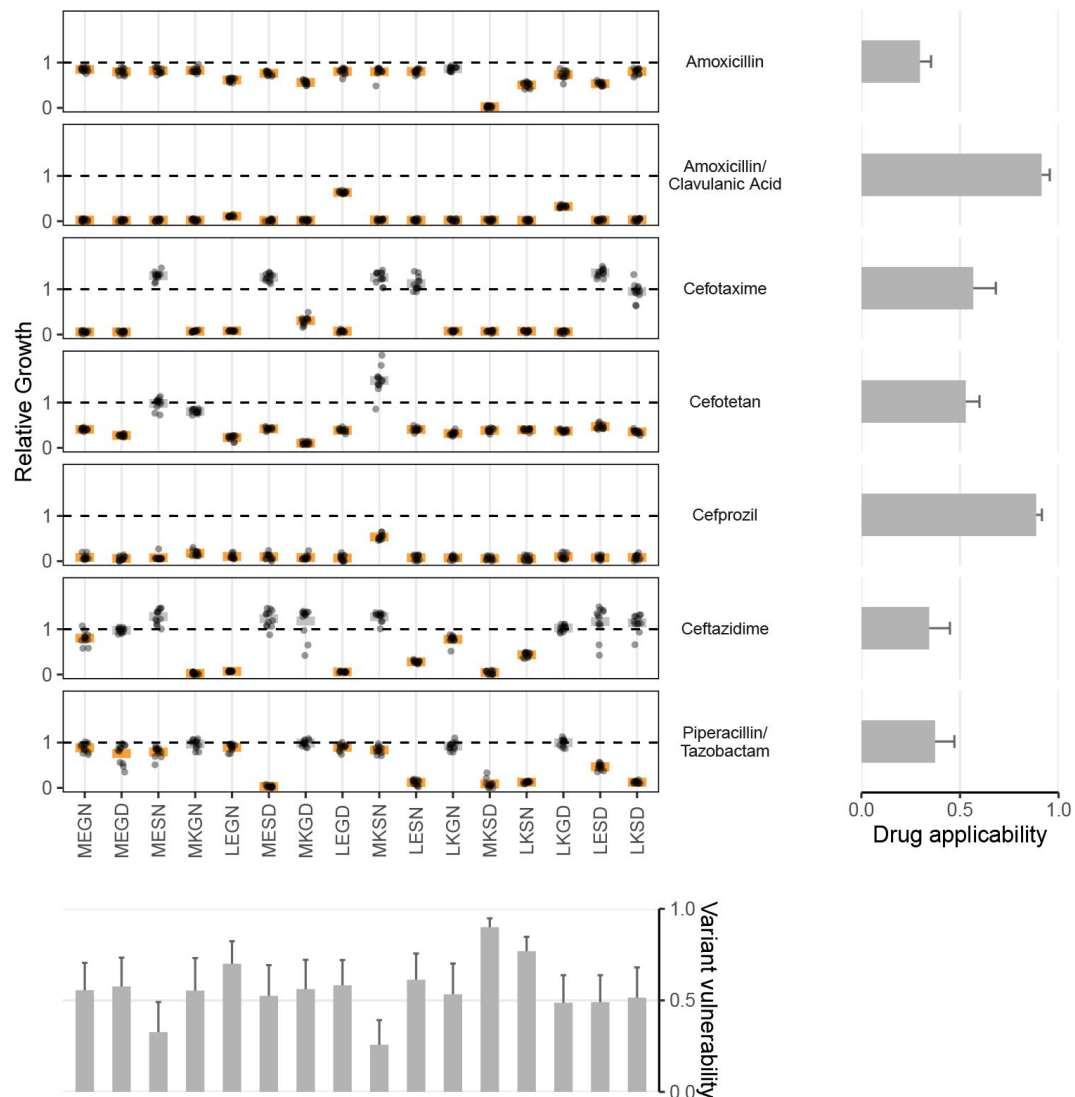


Figure 2.

Data from empirical fitness landscapes can be used to compute two metrics: variant vulnerability (for allelic variants) and drug applicability (for drugs). Growth of allelic variants, relative to control (wild-type), for seven drug treatments. Bars represent the mean value for each group; orange bars indicate significantly lower growth with respect to control (one-tailed t-test, $P < 0.01$), grey bars otherwise. The bars, from left to right, along the bottom correspond to the variant vulnerability of the variants. The bars along the right side, from top to bottom, describe the drug applicability of the seven drugs analyzed in this study. Dashed lines correspond to the TEM-1 growth rate in the putative drug environment.

TEM allelic variant	Binary	Rank (1= highest variant vulnerability)
MKSD	0111	1
LESD	1011	2
LEGN	1000	3
MEGD	0001	4
MKGN	0100	5
LESN	1010	6
LKGN	1100	7
LEGD	1001	8
MKGD	0101	9
LKGD	1101	10
LKSD	1111	11
MEGN	0000	12
MESD	0011	13
LKSN	1110	14
MESN	0010	15
MKSN	0110	16

Source: The authors.

Table 1.

The rank order of the alleles with respect to variant vulnerability. An allele with high drug applicability is most susceptible to the collection of drugs in a panel (seven β -lactamase drugs and drug combinations in this study). MEGN (0000) is the genotype known as the TEM-1 variant of β -lactamase. LKSD (1111) is the TEM-50 variant of β -lactamase genotype.

speak to its evolvability. Is there a correlation between the variant vulnerability values and those of one-step mutant neighbors? Alternatively, are variant vulnerability values uncorrelated from that of their neighbors?

To answer these questions, we calculated each allele's average variant vulnerability of one-step mutational neighbors. **Figure 3** depicts the variant vulnerability values for the 16 allelic values (from **Figure 2**), and a scatterplot depicting the relationship between allelic variants and the average variant vulnerability of one-step neighbors (each variant has four neighbors). Correlation between variant vulnerability values and those of their nearest neighbors (linear regression $R^2 = 0.01$, $P = 0.69$). The lack of correlation suggests that the fitness landscape for variant vulnerability is rugged, where individual mutations contribute nonlinearly to the susceptibility of a given allelic variant.

Drug applicability

A high applicability value indicates that a drug is effective across the suite of genetic variants in a population (**Figure 2**). This study corresponds to a combinatorial set of 16 TEM-1/TEM-50 mutants. In **Table 2**, we rank-order the drugs with respect to their drug applicability. The treatment with the highest drug applicability was the amoxicillin/clavulanic acid combination, comprising both a β -lactam (amoxicillin) and a β -lactamase inhibitor (clavulanic acid). This combination affected the growth rate of all 16 allelic variants. The treatment with the lowest drug applicability was amoxicillin alone. This observation fits intuition, as the combination of amoxicillin and clavulanic acid contains both a β -lactam and a β -lactamase inhibitor.

Analysis of environmental epistasis

Finally, we conducted a statistical decomposition of the individual effect sizes associated with individual SNPs, SNP x SNP interactions (epistasis), SNP x environment (plasticity), and SNP x SNP x environment (environmental epistasis) interactions (**Figure 4**). Here we observe that the four-way interaction between all four loci has a powerful positive effect on recovering growth rate defects in three drugs—amoxicillin, cefotaxime, and ceftazidime (5.8, 8.4, and 5.5 standard deviations respectively). Note how the four-way interaction only has a small effect on the drug mixtures that include β -lactamase inhibitors: amoxicillin/clavulanic acid and piperacillin/tazobactam. For amoxicillin/clavulanic acid—the drug regime with the highest drug applicability, only two mutation interactions have a meaningful impact on the effect of those drugs across variants: a pairwise interaction between sites M69L and N276D (labeled A x D in **Figure 4**), and a three-way interaction between M69L, E104K, and N276D (A x B x D). Most interactions either have no effects or negative effects. This observation is consistent with the amoxicillin/clavulanic drug having the highest drug applicability, with few mutation interactions meaningfully providing resistance.

Discussion

In this study, we examine a data set from an empirical fitness landscape of an antimicrobial drug target to develop metrics for identifying (i) which allelic variants are most likely to be resistant to whole panels of drugs, and (ii) which drugs are most likely to be effective against a suite of resistance allelic variants of a given drug target. We then measure the interactions between individual loci and drug environments that underlie these two metrics.

We focused on mutations associated with the TEM-1/TEM-50 class of β -lactamases, associated with resistance in bacteria, and seven β -lactam drugs. We study a combinatorial set of four different mutations, allowing us to compute the epistasis x environment (also known as “environmental

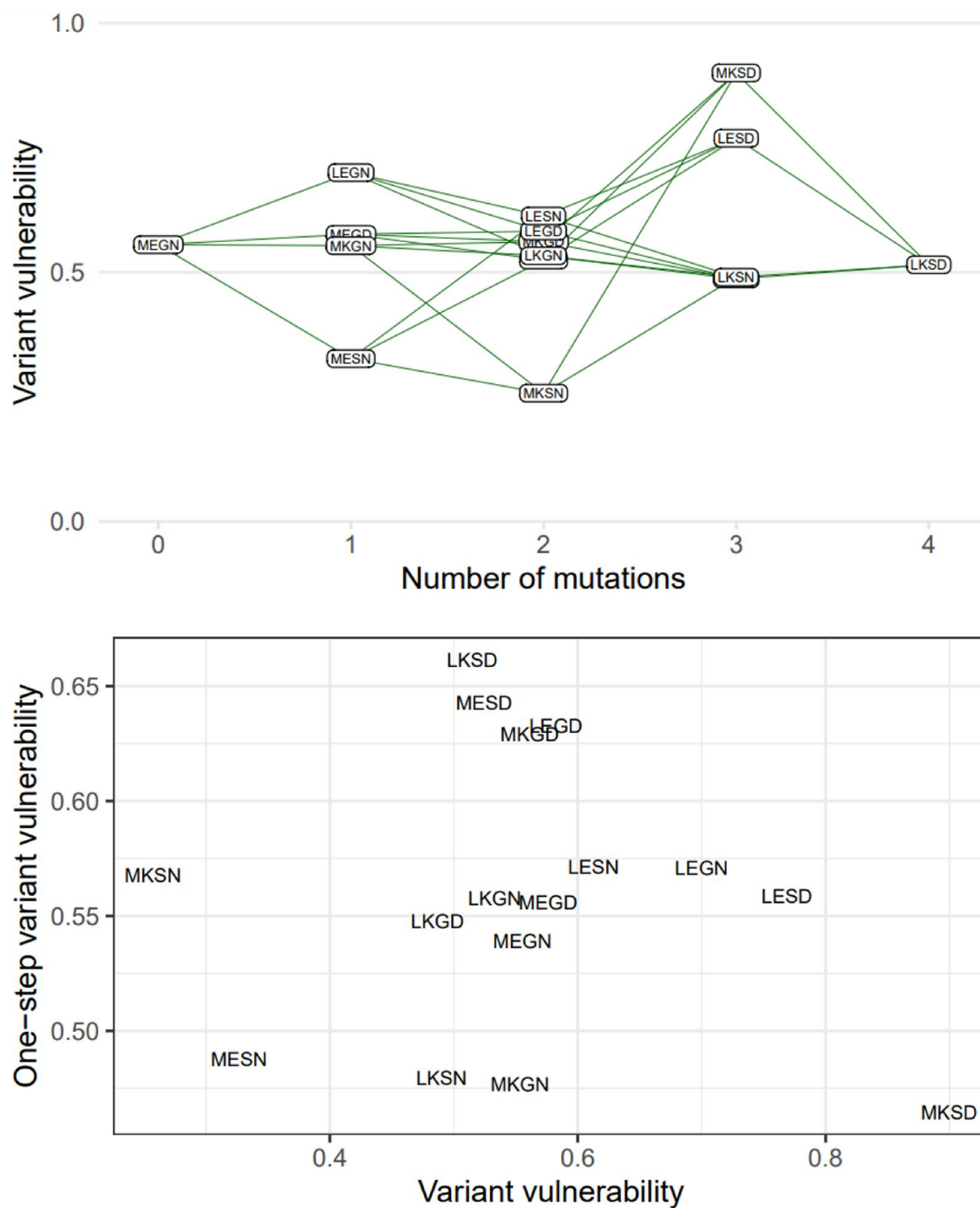


Figure 3.

The variant vulnerability of the 16 allelic variants does not correlate with that of their one-step mutational neighbors. (A) The variant vulnerability values for the 16-allelic variants as a function of the number of mutations away (hamming distance) from TEM-1 (MEGN). (B) The variant vulnerability values of all 16 allelic variants (x-axis) against the average variant vulnerability values of one-step neighbors (each allelic variant has four such neighbors). The analysis demonstrates no correlation, suggesting that variant vulnerability values are distributed nonlinearly across the fitness landscape (linear regression $R^2 = 0.01$, $P = 0.69$).

Antimicrobial	Class	Rank (1 = highest drug applicability)
amoxicillin / clavulanic acid	β -lactam and β -lactamase inhibitor	1
cefprozil	second-generation cephalosporin	2
cefotaxime	third-generation cephalosporin	3
cefotetan	second-generation cephalosporin	4
piperacillin and tazobactam	penicillin and β -lactamase inhibitor	5
ceftazidime	third-generation cephalosporin	6
amoxicillin	β -lactam	7

Source: The authors.

Table 2.

The rank order of the seven drugs with respect to their drug applicability. A drug with a high drug applicability is, on average, highly effective across the range of allelic variants in a set (in this study, a combinatorial set of four mutations that compose the TEM-50 variant of β -lactamase in this study). We emphasize that the findings in this study do not apply to clinical settings nor are they based on clinical data.

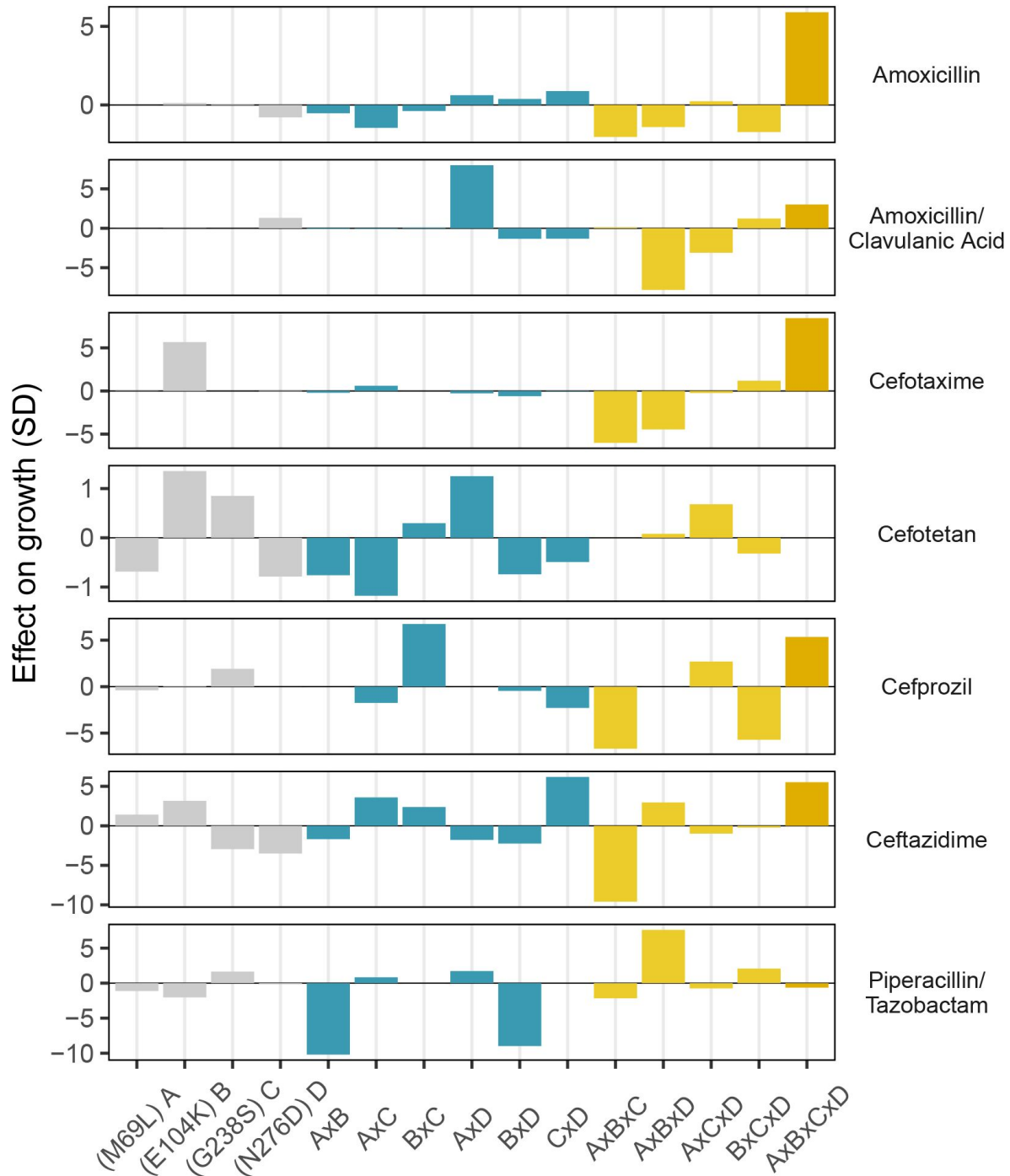


Figure 4.

Environmental epistasis underlies the drug-allele interactions that drive the variant vulnerability and drug applicability. Effect on growth (in standard deviations of the wild-type control values), estimated by LASSO regression, for individual loci and their interactions. [A] corresponds to M69L, [B] to E104K, [C] to G238S, and [D] to N276D. As in a mutation effect reaction norm (Ogbunugafor 2022 [\[1\]](#)), the information describes how the effect of mutations changes across drug environments. This analysis is akin to a large deconstruction of the SNP x SNP x Environment (G x G x E) interactions, also known as environmental epistasis (Lindsey *et al.*, 2013 [\[2\]](#)). In the supplemental information, we provide a version of this figure where all of the coefficients are the same axis.

epistasis”; see ([Lindsey et al., 2013](#)) relationships between alleles.

With respect to variant vulnerability, which measures how susceptible a given allele is across drug types, we observe several intriguing phenomena. Firstly, the allelic variant with the highest variant vulnerability (most susceptible across drug types; MKSD or 0111) and lowest variant vulnerability (MKSX or 0110) are only a single mutation apart (N276D) (**Figure 1**). That a single mutation can have large consequences for treatment across drugs in a panel is not surprising from one perspective—individual mutations can certainly have strong effects. The analysis of environmental epistasis (**Figure 4**) highlights, however, that the dramatic difference in phenotype between MKSD and MKSX is as much about how that mutation effect depends on the other amino acids in the protein background. The context-dependence of mutations contributing to variant vulnerability highlights the continued need to resolve the population genetic particulars of a given infection. For example, in a clinical setting, inaccurate identification of the SNP or amino acid at a single locus (N276D in our study) could create an entirely different picture of the clinical course of infection.

Our findings on drug applicability highlight the complexity underlying drug mechanisms and their consequences. Though all seven drugs/combinations are β -lactams, they have widely varying effects across the 16 alleles. Some of the results are intuitive: for example, the drug regime with the highest drug applicability of the set—amoxicillin/clavulanic acid—is a mixture of a widely used β -lactam (amoxicillin) and a β -lactamase inhibitor (clavulanic acid) (see **Table 2**). We might expect such a mixture to have a broader effect across a diversity of variants. However, this high applicability is hardly a rule, as another mixture in the set, piperacillin/tazobactam, has much lower drug applicability (ranking 5th out of the seven drugs in the set) (**Table 2**). Intuition can over-simplify, with undesirable consequences: not all drug combinations/mixtures are more effective over a larger breadth of alleles. Because this study focused on mutations in combination, rather than drugs, we did not examine the drug applicability of the individual drugs in all the drug combinations (e.g., piperacillin alone and tazobactam alone). However, there is a rich literature covering the specific problem of drug combinations ([Michel et al., 2008](#); [Beppler et al., 2016](#); [Singh and Yeh, 2017](#); [Tekin et al., 2017](#)).

The surprising results in both the variant vulnerability (regarding the ability of allelic variants to grow across niche breadth) and drug applicability (regarding the impact of drugs across allelic variants) highlight the importance of specific interactions between individual loci (individual sites in TEM-50) and drug environment (β -lactam drugs). For example, we observe the strongly positive interaction between mutations at the four sites associated with TEM-50 in several drugs: amoxicillin, cefotaxime, cefprozil, and ceftazidime. Other higher-order interactions are associated with different effects: note the strongly negative effects of the three-way interaction between mutations at loci M69L, E104K, and G238S on ceftazidime growth rate (**Figure 4**). Regarding the main effects, we note the positive effect of the M69L mutation on the growth rate in cefotaxime, cefotetam, and ceftazidime. Alternatively, the G238S, N276D mutations have individual negative effects on ceftazidime. Interestingly, in combination, these mutations strongly positively affect ceftazidime (**Figure 3**). While our study does not delve into the biophysical underpinnings of any of these findings, they do support other studies that speak to the importance of how even slight structural differences between drugs can have strong implications for how they interact with allelic variants of a single protein ([Ogbunugafor et al., 2016](#); [Lozovsky et al., 2021](#); [Mira et al., 2021](#)).

Study limitations

Firstly, this study utilized data from a set of low-dimensional empirical fitness landscapes (16 alleles, across seven drugs/combinations). We emphasize that our aim is not to provide any insight about treatment regimes, or inform the real-world use of any specific antimicrobial. Therefore, any direct clinical applications of our findings are limited. Instead, our study offers an evolutionary perspective and metrics that are largely agnostic with respect to drug type and can

be further applied to drug-pathogen data of various kinds. Given the rise of technologies (e.g., deep mutational scanning) ([Fowler and Fields, 2014](#)) that permit the experimental generation of higher-dimensional fitness landscapes, future studies should apply these metrics and analyses to larger data sets where we might determine the variant vulnerability of thousands of variants, alongside the drug applicability of hundreds of drugs. We hope that we have outlined our methods in a manner that makes such applications relatively easy to consider.

In addition, our study focused on a discrete set of individual drugs and mixtures. A robust literature exists surrounding the importance of drug-drug interactions in antimicrobial resistance ([Michel et al., 2008](#); [Yeh et al., 2009](#); [Beppler et al., 2016](#); [Tekin et al., 2017](#)). Considering how these metrics apply in the context of higher-order combinations of drugs constitutes an area of future inquiry.

Lastly, we use the β -lactam/ β -lactamase system in our study. There are hundreds of other possible drug-target environments to examine. Thankfully, the arithmetic behind the proposed metrics (and the computation of the environmental epistasis that underlines them) is simple, and outlined in the Methods section. To demonstrate how this metric could be applied to different drug classes, we have analyzed a data set of 16 alleles of the *P. falciparum* ortholog of dihydrofolate reductase, in the presence of two different antifolate drugs: pyrimethamine and cycloguanil (See supplemental information). This analysis also demonstrates how the metrics can be computed across a broader range of drug dosages. Note that these data have been examined in prior analyses of fitness landscape topography and epistasis ([Ogbunugafor et al., 2016](#); [Ogbunugafor, 2022](#); [Diaz-Colunga et al., 2022](#)).

Ideas, speculation, and future directions

Returning to a prior point, we must highlight the resonance between the lens offered from this study, and existing perspectives in the virus-antibody realm. In that case, studies have examined “binding affinity landscapes” to extract metrics that are akin to our drug applicability ([Doud et al., 2017](#), 2018), and have characterized virus mutations and epistatic interactions ([Phillips et al., 2021](#), 2023) that are analogous to our variant vulnerability measure. In the virus-antibody problem, the goals are somewhat similar: (i) to identify the characteristics of anti-infectious disease agents that make them likely to be effective against a breadth of variants and (ii) to identify the genetic architecture or signature of pathogens that render these agents ineffective. While the biological problems of antimicrobials-microbes and antibody-virus binding are far different and warrant separate discussions, some methods or metrics can be employed across these domains. At the very least, the convergence in perspectives reflects an urgency to develop new evolutionary-inspired strategies to combat pathogen evolution. This is and will remain a frontier of both evolutionary biology and biomedicine for the foreseeable future. Furthermore, the problem of antimicrobial resistance can more directly employ higher-throughput methods and new technologies to elaborate on variant vulnerability and drug applicability metrics.

Future work may seek to elaborate on these and other studies that highlight the role of environments in crafting plasticity and adaptive evolution ([Cvijović et al., 2015](#); [Raghavan et al., 2021](#); [Kinsler et al., 2020](#)). Such efforts may include studies of “fitness seascapes” and evolution in fluctuating environments, as others have examined ([Mustonen and Lässig, 2009](#); [Cvijović et al., 2015](#); [King et al., 2022](#)). These efforts should make use of modern modeling approaches which can rigorously simulate fitness landscapes in realistic, population-level contexts ([Cárdenas et al., 2022](#)).

In sum, the findings are relevant to the evolution of antimicrobial resistance and, more broadly, evolutionary medicine. For example, our study supports prior efforts that highlight the importance of higher-order interactions in the evolution of antimicrobial resistance ([Weinreich et al., 2006](#); [Lozovsky et al., 2021](#)) and the existence of G x G x E interactions (“environmental epistasis”) as a meaningful characteristic of fitness landscapes in microbial systems ([Lindsey et al., 2013](#); [Ogbunugafor, 2022](#)). Generally, this study supports recent efforts that employ evolutionary perspectives to understand antimicrobial resistance ([Michel et al., 2008](#); [Yeh et al., 2009](#); [Rosenbloom et al., 2012](#); [Baym et al., 2016](#); [Beppler et al., 2016](#); [Gjini and Wood, 2021](#); [Ardell and Kryazhimskiy, 2021](#)). Like many others, we believe that we are only at the beginning of a larger, holistic effort to supplement existing metrics and concepts developed within clinical medicine, all towards more effective therapies that can improve outcomes at the bedside.

Materials and Methods

Notes on the terminology I

The quantities we define as “variant vulnerability” and “drug applicability” may be seen as descendants of “drugability” (or “druggability” and other related terms), which are sometimes applied to properties of targets ([Halgren, 2009](#); [Jubb et al., 2015](#)) and other times to drug candidates themselves ([Keller et al., 2006](#)), ([Benet et al., 2016](#), p.5). These metrics are generally relegated to molecular properties of drugs and targets, and are most often deployed in medicinal chemistry. To avoid ambiguity and in striving to bridge concepts from chemical biology, pharmacology, and evolutionary systems biology, we have avoided using either term (druggability or drugability) in favor of new terminology. In light of this, we have selected “variant vulnerability” to apply to allelic variants and “drug applicability” to apply to drugs. We provide study definitions of these and other terms and metrics in [Table 3](#). Also note that our measure of variant vulnerability resonates with the concept of cross-resistance, which applies to pathogens that are resistant by a single mechanism to all or most antimicrobials belonging to a class ([de Souza Cândido et al., 2019](#); [Colclough et al., 2019](#)). Some studies highlight that cross-resistance can manifest across genetic variants, with mutations conferring resistance to two drugs simultaneously ([Yeh et al., 2006](#)). While modern studies have added nuance to our understanding of how resistance and cross-resistance can evolve ([Oz et al., 2014](#)), ambiguity remains surrounding how it is used and, more importantly, how it is measured. The variant vulnerability metric improves upon notions of cross-resistance in two important ways: (i) it incorporates genetic variation into our understanding of how pathogens can be resistant to a panel of drugs, and (ii) it provides a quantitative metric for such cross-resistance (cross-resistance is typically used qualitatively).

Notes on the terminology II

In addition to the above conversation surrounding the terms “drugability” and “druggability,” it is also important to briefly clarify the uses of the term “fitness landscape” vs. “adaptive landscape.” These terms can generally be used interchangeably and sometimes to describe how evolution moves through protein space, an idea pioneered by John Maynard Smith ([Maynard Smith, 1970](#); [Ogbunugafor, 2020](#)), that has since been invoked to describe aspects of protein evolution and engineering ([Romero and Arnold, 2009](#); [Arnold, 2011](#), [2019](#)). We use fitness landscape in this study, partly because we are talking about a phenotype—growth rate—that is a component of bacterial fitness in the presence of antibiotics (there are, of course, other proxies, like competitive

Term	Study definition
Drugability or Druggability	Generally used to describe the ability to treat a drug target with a small molecule-based drug (<i>Halgren, 2009; Jubb et al., 2015</i>). Sometimes refers to the ability of a drug to be used against a single or set of drug targets, and/or used to describe whether a small molecule fits into a compound class that is or can be adapted to have pharmacologic properties useful in the clinic (e.g., solubility) (<i>Keller et al., 2006; Benet et al., 2016</i>).
Cross-resistance	Phenomenon whereby pathogens are resistant to all or most antimicrobials belonging to a given class via a single mechanism.
Susceptibility	$1 - w$, where w is the mean growth of the allelic variant under drug conditions relative to the mean growth of the wild-type control.
Environmental epistasis	A term that comes from a prior study on G x G x E interactions (<i>Lindsey et al., 2013</i>). Analysis of the various G x G x E interactions that drive the variant vulnerability and drug applicability metrics.
Variant vulnerability	The average susceptibility of an allele across drugs in a panel.
Drug applicability	The average susceptibility of all variants to an antibiotic.

Source: The authors.

Table 3.

Terms used or introduced in this study and their definition. For some terms, references are provided for clarity.

fitness). We also note that the data used in this study comes from a prior study that discussed these data ([Mira et al., 2015](#)). Because of that, we resisted the use of “protein space” or other analogies that similarly describe data of this sort.

Note on the study system

This study utilizes the β -lactam/ β -lactamase target-pairs for purposes of illustration. β -lactamases are a class of enzymes produced by bacteria that break down the β -lactam ring of antibiotics, making them ineffective. TEM -type β -lactamases are the main culprit of β -lactam resistance in gram-negative microorganisms. TEM was first identified in *Escherichia coli* isolated from a patient named Temoneria in Greece in 1965 ([Steward et al., 2000](#)). The evolution and subsequent spread of β -lactamases arose from the widespread clinical use of β -lactam antibiotics. As the clinical treatment landscape involves using β -lactam drugs of various kinds (old and new), the population genetic landscape of β -lactamases is diverse ([Blazquez et al., 2000](#)). Among clinical populations of gram-negative microorganisms, the TEM-1 allele is the most frequently detected antibacterial resistance variant ([Mroczkowska and Barlow, 2008](#)). Many of the most common alleles of other TEM genes are mutants of TEM-1, with TEM-50 a derivative of TEM-1 separated by four mutations: Met \rightarrow Leu-69, Glu \rightarrow Lys-104, Gly \rightarrow Ser-238 and Asn \rightarrow Asp-276 ([Sirost et al., 1997](#)). These mutations confer structural, and functional differences that meaningfully influence the clinical use of antibiotics ([Pimenta et al., 2014](#)). In the decades since β -lactamases were identified for their public health relevance, they have become one of the key protein systems for evolutionary biochemistry and biophysics studies. Pioneering studies have examined concepts like the adaptive trajectory ([Weinreich et al., 2006](#)), evolvability ([Stiffler et al., 2015](#)), and biophysical tradeoffs ([Knies et al., 2017](#)), and have used techniques like deep-mutational scanning to examine the fitness effects of mutations in high-throughput ([Rollins et al., 2019](#)). While this study focuses on data from the β lactam/ β -lactamase interaction, the methods and understandings can broadly apply to other study systems (e.g., dihydrofolate reductase and antifolate drugs).

In the current study, we describe the set of allelic variants using their single amino acid abbreviations: M69L, E104K, G238S, N276D. We can also utilize binary notation describing the mutants (as do many studies of low-dimension fitness landscapes), with the TEM-1 locus corresponding to a 0, and a 1 corresponding to a mutation associated with TEM-50. For example, the TEM-1 allele, MEGH, can also be described as 0000, and the quadruple mutant as LKSD or 1111.

Data

The data we examine comes from a past study of a combinatorial set of four mutations associated with TEM-50 resistance to β -lactam drugs ([Mira et al., 2015](#)). This past study measured the growth rates of these four mutations in combination, across 15 different drugs (see supplemental information). We examined these data, identifying a subset of structurally similar β -lactams that also included β -lactams combined with β -lactamase inhibitors, cephalosporins and penicillins.

From the original data set (see ([Mira et al., 2015](#)) and the supplemental information), we focus our analyses on drug treatments that had a significant negative effect on the growth of wildtype/TEM-1 strains (one-tailed t-test of wild-type treatment vs. control, $P < 0.01$). After identifying the data from the set that fit our criteria, we were left with seven drugs or combinations: amoxicillin 1024 $\mu\text{g/ml}$ (β -lactam), amoxicillin/clavulanic acid 1024 $\mu\text{g/ml}$ (β -lactam and β -lactamase inhibitor) cefotaxime 0.123 $\mu\text{g/ml}$ (third-generation cephalosporin), cefotetan 0.125 $\mu\text{g/ml}$ (second-generation cephalosporins), cefprozil 128 $\mu\text{g/ml}$ (second-generation cephalosporin), ceftazidime 0.125 $\mu\text{g/ml}$ (third-generation cephalosporin), piperacillin and tazobactam 512/8 $\mu\text{g/ml}$ (penicillin and β -lactamase inhibitor). With these drugs/mixtures, we were able to embody chemical diversity in the panel. While drug combinations are composed of mixtures, this study did not measure all combinations in their components, but analyzed results as measured in [Mira et al., 2015](#) ([Mira et al., 2015](#)).

Lastly, we studied a single system so that we could focus our analyses and have a substantive discussion about the biological details of the protein-drug interactions. However, the broader concepts and metrics offered in this study can be applied more broadly. In the supplemental material, we have analyzed a different dataset, corresponding to a different target-drug: the *Plasmodium falciparum* ortholog of dihydrofolate reductase (DHFR), and various concentrations of two antifolate drugs, pyrimethamine and cycloguanil. These data offer a few additional benefits: while this set examines fewer drugs in number, they do contain a number of drug concentrations, and so their analyses demonstrates how one can examine the problem across different drug dosages.

Metric calculations

To estimate the two metrics we are interested in, we must first quantify the susceptibility of an allelic variant to a drug. We define susceptibility as $1 - w$, where w is the mean growth of the allelic variant under drug conditions relative to the mean growth of the wild-type/TEM-1 control. If a variant is not significantly affected by a drug (i.e., growth under the drug is not statistically lower than the growth of wild-type/TEM-1 control, by t-test P-value < 0.01), its susceptibility is zero. Values in these metrics are summaries of susceptibility: the variant vulnerability of an allelic variant is its average susceptibility across drugs in a panel, and the drug applicability of an antibiotic is the average susceptibility of all variants to it.

We further explored the interactions across this fitness landscape and panels of drugs in two additional ways. First, we calculated the variant vulnerability for 1-step neighbors, which is the mean variant vulnerability of all alleles one mutational step away from a focal variant. This metric explains how the variant vulnerability values are distributed across a fitness landscape. Second, we estimated statistical interaction effects on bacterial growth through LASSO regression (implemented in the ‘glmnet’ R package; Friedman et al 2010). For each drug, we fit a model of relative growth as a function of M69L x E104K x G238S x N276D (i.e., including all interaction terms between the four amino acid substitutions). The effect sizes of the interaction terms from this regularized regression analysis allow us to infer higher-order dynamics for susceptibility. We label this calculation as an analysis of “environmental epistasis.”

Note on methods used to measure epistasis

Here we will briefly summarize methods used to study epistasis on fitness landscapes. Several studies of combinatorially-complete fitness landscapes use some variation of the Fourier Transform or Taylor formulation. One in particular, the Walsh-Hadamard Transform has been used to measure epistasis across many study systems ([Weinreich et al., 2013](#) [2018](#); [Ogbunugafor, 2022](#)). Furthermore, studies have reconciled these methods with others, or expanded upon the Walsh-Hadamard Transform in ways that can accommodate incomplete data sets ([Faure et al., 2023](#); [Doro and Herman, 2022](#)).

The method that we have utilized, the LASSO regression, determines effect sizes for all interactions (alleles and drug environments). It has been utilized for data sets of similar size and structure, on alleles resistant to trimethoprim, an antifolate antibiotic ([Guerrero et al., 2019](#)). Among many benefits, the method can accommodate gaps in data and responsibly incorporates experimental noise into the calculation.

Acknowledgements

The authors would like to acknowledge seminar invitations from the Massachusetts Institute of Technology, University of California, San Diego, Brown University, and the Innovative Genomics Institute (University of California, Berkeley), where iterations of the ideas in this manuscript were discussed. The authors acknowledge support from the National Institutes of Health grants

R35GM136–354 (M.D.S. and RMH) and R01AI168166 (M.D.S. and C.B.O.), R35GM147107 (R.F.G.), and the National Science Foundation’s Division of Environmental Biology Award Number 2142719 (C.B.O.). The authors would also like to thank the Martin Luther King Jr Visiting Professors and Scholars Program at the Massachusetts Institute of Technology for support (C.B.O.). The authors would like to thank the organizers and participants in the 2022 workshop entitled “Reimagining the Central Dogma” at The Foundations Institute, University of California, Santa Barbara, where ideas related to those covered in this manuscript were discussed. Lastly, the authors thank K. Kabengele and S. Scarpino for their helpful feedback on the manuscript.

Data Availability

Data and code are available at: <https://github.com/OgPlexus/evodrugability>

Appendix 1

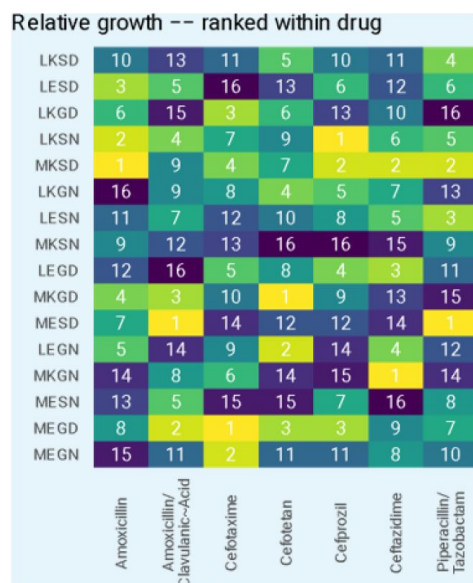
Supplemental Material

Guerrero RF, Dorji T, Harris RM, Shoulders MD, Ogbunugafor CB. Evolutionary druggability: leveraging low-dimensional fitness landscapes towards new metrics for antimicrobial applications. For publication in eLife. 2023

In the supplemental material, we offer additional analysis and selected discussion points.

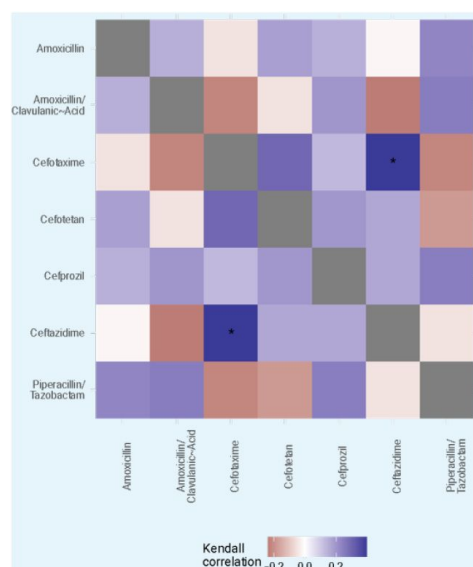
1. In the main text **Figure 1**, we offer graphs of fitness landscapes corresponding to the 16 TEM mutants of β lactamase, across 7 different drug environments.
Here we offer the rank order changes of alleles for the 7 different fitness landscapes. These rank orders further highlight the dynamics of the topography.
2. We offer an analysis of the topography of those landscapes, based on the Kendall rank order test. This tests the hypothesis that there is no correlation (concordance or discordance) between the topographies of the fitness landscapes.
3. To test the hypothesis that the variant vulnerability values differ, we calculated a paired t-test. These are paired by haplotype/allelic variant, so the comparisons are change in growth between drugs for each haplotype.
4. While the study focused on the β lactamase– β lactam system (widely adopted in examinations of the evolution of drug resistance), the metrics are presented such that they might be applied broadly. To demonstrate another use case, we offer a set of 16 alleles of the *Plasmodium falciparum* ortholog of dihydrofolate reductase. These constitute a combinatorial set of mutations, with growth rate across a breadth of concentrations of two antifolate drugs: pyrimethamine (PYR), and cycloguanil (CYC).
In Figure S4 below, find a figure where we’ve computed both the variant variability y-axis) and the drug applicability (x-axis) across the 16 alleles and across 9 concentrations of 2 drugs. Importantly, this analysis demonstrates how the drug applicability metric can be computed across drug concentrations. Previous work has demonstrated that drug concentration can be a meaningful influence on the dynamics of drug resistance in antifolates:
 - Ogbunugafor CB, Wylie CS, Diakite I, Weinreich DM, Hartl DL. Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance. *PLoS computational biology*. 2016 Jan 25;12(1):e1004710.

As we observe, drug concentration does have a meaningful impact on the metrics. In the main text, we focused on single high doses of each of the 7 drug/mixtures. But the metrics can just as readily be computed across different drug concentrations.



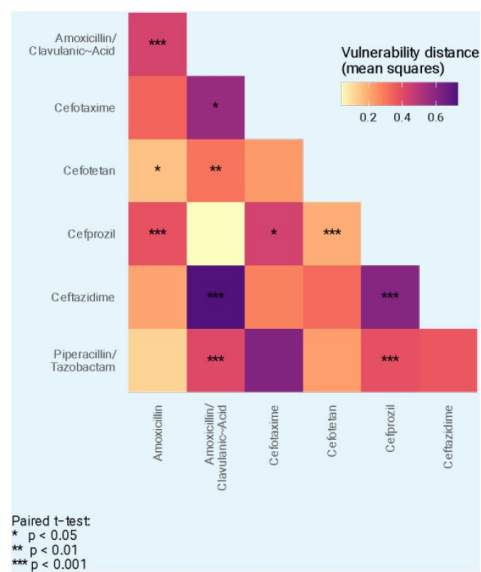
Appendix 1—figure 1.

Rank orders for the alleles in the fitness landscapes outlined in the main text **Figure 1**



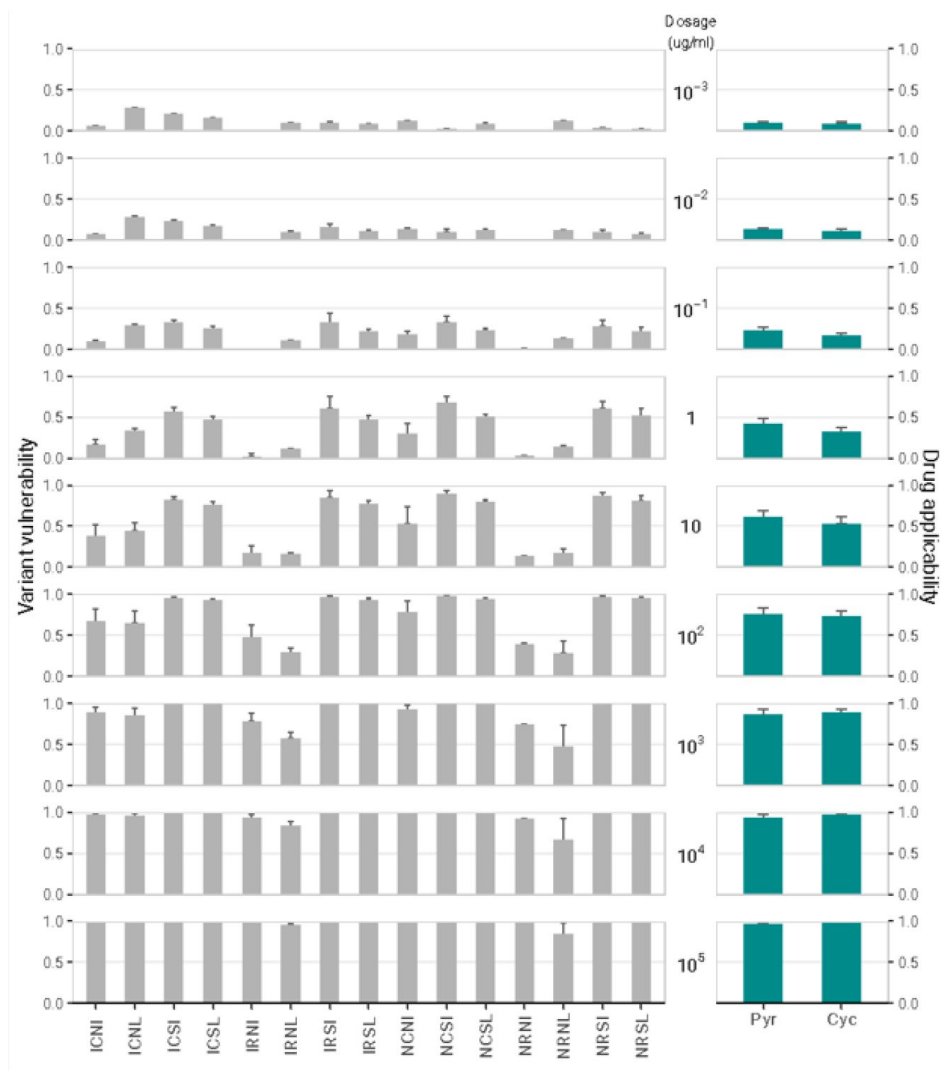
Appendix 1—figure 2.

We offer an analysis of the topography of those landscapes, based on the Kendall rank order test. This tests the hypothesis that there is no correlation (concordance or discordance) between the topographies of the fitness landscapes



Appendix 1—figure 3.

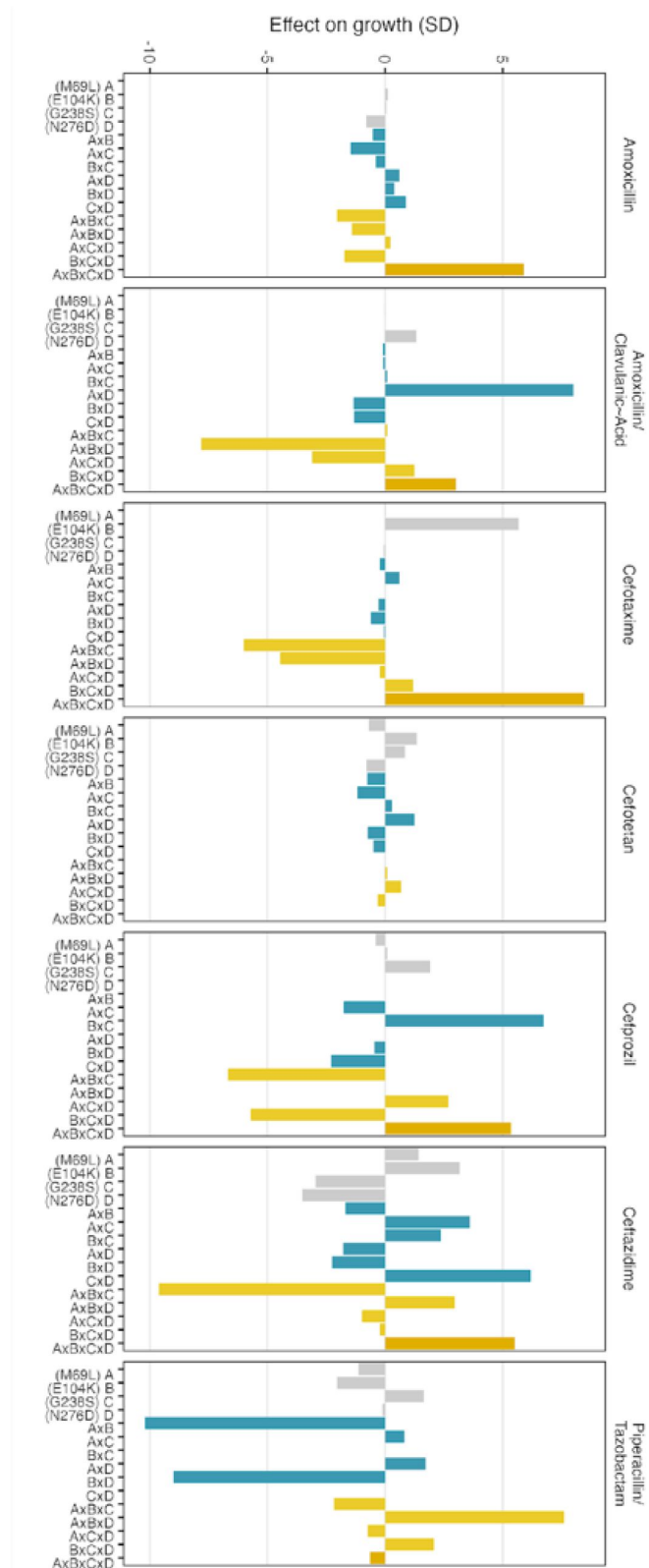
To test the hypothesis that the variant vulnerability values differ, we calculated a paired t-test. These are paired by haplotype/allelic variant, so the comparisons are change in growth between drugs for each haplotype



Appendix 1—figure 4.

Variant vulnerability and drug applicability values for 16 allelic variants of dihydrofolate reductase (*Plasmodium falciparum*), across 9 concentrations of 2 different antifolate drugs

5. Alternative version of main text **Figure 4** [↗](#), with all coefficients on the same axis, to permit absolute comparisons.



Appendix 1—figure 5.

As stated in them main text: "Environmental epistasis underlies the drug-allele interactions that drive the variant vulnerability and drug applicability". Effect on growth (in standard deviations of the wild-type control values), estimated by LASSO regression, for individual loci and their interactions. [A] corresponds to M69L, [B] to E104K, [C] to G238S, and [D] to N276D?

References

- Ardell SM, Kryazhimskiy S. (2021) **The population genetics of collateral resistance and sensitivity** *eLife* **10** <https://doi.org/10.7554/elife.73250>
- Arnold FH (2011) **The Library of Maynard-Smith: my search for meaning in the protein universe** *Microbe* **6**
- Arnold FH (2019) **Innovation by evolution: bringing new chemistry to life (Nobel Lecture)** *Angewandte Chemie International Edition* **58**:14420–14426
- Baym M, Stone LK, Kishony R. (2016) **Multidrug evolutionary strategies to reverse antibiotic resistance** *Science* **351** <https://doi.org/10.1126/science.aad3292>
- Benet LZ, Hosey CM, Ursu O, Oprea TI (2016) **BDDCS, the Rule of 5 and drugability** *Advanced Drug Delivery Reviews* **101**:89–98 <https://doi.org/10.1016/j.addr.2016.05.007>
- Beppler C, Tekin E, Mao Z, White C, McDiarmid C, Vargas E, Miller JH, Savage VM, Yeh PJ (2016) **Uncovering emergent interactions in three-way combinations of stressors** *Journal of The Royal Society Interface* **13** <https://doi.org/10.1098/rsif.2016.0800>
- Bergstrom CT, Lipsitch M, Levin BR (2000) **Natural Selection, Infectious Transfer and the Existence Conditions for Bacterial Plasmids** *Genetics* **155**:1505–1519 <https://doi.org/10.1093/genetics/155.4.1505>
- Blazquez J, Morosini MI, Negri MC, Baquero F. (2000) **Selection of Naturally Occurring Extended-Spectrum TEM β -Lactamase Variants by Fluctuating β -Lactam Pressure** *Antimicrobial Agents and Chemotherapy* **44**:2182–2184 <https://doi.org/10.1128/aac.44.8.2182-2184.2000>
- Cárdenas P, Corredor V, Santos-Vega M. (2022) **Genomic epidemiological models describe pathogen evolution across fitness valleys** *Science advances* **8**
- Colclough A, Corander J, Sheppard SK, Bayliss SC, Vos M. (2019) **Patterns of cross-resistance and collateral sensitivity between clinical antibiotics and natural antimicrobials** *Evolutionary Applications* **12**:878–887 <https://doi.org/10.1111/eva.12762>
- Cvijović I, Good BH, Jerison ER, Desai MM (2015) **Fate of a mutation in a fluctuating environment** *Proceedings of the National Academy of Sciences* **112**:E5021–E5028
- Cvijović I, Good BH, Jerison ER, Desai MM (2015) **Fate of a mutation in a fluctuating environment** *Proceedings of the National Academy of Sciences* **112** <https://doi.org/10.1073/pnas.1505406112>
- Diaz-Colunga J, Sanchez A, Ogbunugafor CB (2022) **Environmental modulation of global epistasis is governed by effective genetic interactions** *bioRxiv* :2022–11
- Doro S, Herman MA (2022) **On the Fourier transform of a quantitative trait: Implications for compressive sensing** *Journal of Theoretical Biology* **540**

- Doud MB, Hensley SE, Bloom JD (2017) **Complete mapping of viral escape from neutralizing antibodies** *PLOS Pathogens* **13** <https://doi.org/10.1371/journal.ppat.1006271>
- Doud MB, Lee JM, Bloom JD (2018) **How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin** *Nature Communications* **9** <https://doi.org/10.1038/s41467-018-03665-3>
- Faure AJ, Lehner B, Miró Pina V, Serrano Colome C, Weghorn D. (2023) **An extension of the Walsh-Hadamard transform to calculate and model epistasis in genetic landscapes of arbitrary shape and complexity** *bioRxiv* :2023–3
- Fowler DM, Fields S. (2014) **Deep mutational scanning: a new style of protein science** *Nature Methods* **11**:801–807 <https://doi.org/10.1038/nmeth.3027>
- Gjini E, Wood KB (2021) **Price equation captures the role of drug interactions and collateral effects in the evolution of multidrug resistance** *eLife* **10** <https://doi.org/10.7554/elife.64851>
- Guerrero RF, Scarpino SV, Rodrigues JV, Hartl DL, Ogbunugafor CB (2019) **Proteostasis Environment Shapes Higher-Order Epistasis Operating on Antibiotic Resistance** *Genetics* **212**:565–575 <https://doi.org/10.1534/genet-ics.119.302138>
- Halgren TA (2009) **Identifying and Characterizing Binding Sites and Assessing Druggability** *Journal of Chemical Information and Modeling* **49**:377–389 <https://doi.org/10.1021/ci800324m>
- Herencias C, Rodríguez-Beltrán J, León-Sampedro R, Alonso-del Valle A, Palkovičová J, Cantón R, San Millán Á. (2021) **Collateral sensitivity associated with antibiotic resistance plasmids** *Elife* **10**
- Jubb H, Blundell TL, Ascher DB (2015) **Flexibility and small pockets at protein–protein interfaces: New insights into druggability** *Progress in Biophysics and Molecular Biology* **119**:2–9 <https://doi.org/10.1016/j.pbiomolbio.2015.01.009>
- Keller TH, Pichota A, Yin Z. (2006) **A practical view of ‘druggability’** *Current Opinion in Chemical Biology* **10**:357–361 <https://doi.org/10.1016/j.cbpa.2006.06.014>
- King ES, Maltas J, Weaver DT, Barker-Clarke R, Pelesko J, Dolson E, Scott JG (2022) **Fitness seascapes are necessary for realistic modeling of the evolutionary response to drug therapy** *bioRxiv* <https://doi.org/10.1101/2022.06.10.495696>
- Kinsler G, Geiler-Samerotte K, Petrov DA (2020) **Fitness variation across subtle environmental perturbations reveals local modularity and global pleiotropy of adaptation** *Elife* **9**
- Knies J, Cai F, Weinreich DM (2017) **Enzyme efficiency but not thermostability drives cefotaxime resistance evolution in TEM-1 β -lactamase** *Molecular Biology and Evolution* **34** <https://doi.org/10.1093/molbev/msx053>
- Lindsey HA, Gallie J, Taylor S, Kerr B. (2013) **Evolutionary rescue from extinction is contingent on a lower rate of environmental change** *Nature* **494**:463–467 <https://doi.org/10.1038/nature11879>

- Loftie-Eaton W, Bashford K, Quinn H, Dong K, Millstein J, Hunter S, Thomason MK, Merrih H, Ponciano JM, Top EM (2017) **Compensatory mutations improve general permissiveness to antibiotic resistance plasmids** *Nature Ecology & Evolution* **1**:1354–1363 <https://doi.org/10.1038/s41559-017-0243-2>
- Lozovsky ER, Chookajorn T, Brown KM, Imwong M, Shaw PJ, Kamchonwongpaisan S, Neafsey DE, Weinreich DM, Hartl DL (2009) **Stepwise acquisition of pyrimethamine resistance in the malaria parasite** *Proceedings of the National Academy of Sciences* **106**:12025–12030 <https://doi.org/10.1073/pnas.0905922106>
- Lozovsky ER, Daniels RF, Heffernan GD, Jacobus DP, Hartl DL (2021) **Relevance of Higher-Order Epistasis in Drug Resistance** *Molecular Biology and Evolution* **38**:142–151 <https://doi.org/10.1093/molbev/msaa196>
- Maynard Smith J. (1970) **Natural selection and the concept of a protein space** *Nature* **225**:563–564
- Michel JB, Yeh PJ, Chait R, Moellering RC, Kishony R. (2008) **Drug interactions modulate the potential for evolution of resistance** *Proceedings of the National Academy of Sciences* **105**:14918–14923 <https://doi.org/10.1073/pnas.0800944105>
- Mira PM, Meza JC, Nandipati A, Barlow M. (2015) **Adaptive Landscapes of Resistance Genes Change as Antibiotic Concentrations Change** *Molecular Biology and Evolution* **32**:2707–2715 <https://doi.org/10.1093/mol-bev/msv146>
- Mira PM, Østman B, Guzman-Cole C, Sindi S, Barlow M. (2021) **Adaptive Processes Change as Multiple Functions Evolve** *Antimicrobial Agents and Chemotherapy* **65** <https://doi.org/10.1128/aac.01990-20>
- Mroczkowska JE, Barlow M. (2008) **Fitness Trade-Offs in blaTEM Evolution** *Antimicrobial Agents and Chemotherapy* **52**:2340–2345 <https://doi.org/10.1128/aac.00018-08>
- Mustonen V, Lässig M. (2009) **From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation** *Trends in Genetics* **25**:111–119 <https://doi.org/10.1016/j.tig.2009.01.002>
- Nichol D, Rutter J, Bryant C, Hujer AM, Lek S, Adams MD, Jeavons P, Anderson AR, Bonomo RA, Scott JG (2019) **Antibiotic collateral sensitivity is contingent on the repeatability of evolution** *Nature communications* **10**
- Ogbunugafor CB (2020) **A reflection on 50 years of John Maynard Smith’s “Protein Space”** *Genetics* **214**:749–754
- Ogbunugafor CB (2022) **The mutation effect reaction norm (mu-rn) highlights environmentally dependent mutation effects and epistatic interactions** *Evolution* **76**:37–48 <https://doi.org/10.1111/evo.14428>
- Ogbunugafor CB, Hartl D. (2016) **A pivot mutation impedes reverse evolution across an adaptive landscape for drug resistance in Plasmodium vivax** *Malaria Journal* **15** <https://doi.org/10.1186/s12936-016-1090-3>
- Ogbunugafor CB, Wylie CS, Diakite I, Weinreich DM, Hartl DL (2016) **Adaptive Landscape by Environment Interactions Dictate Evolutionary Dynamics in Models of Drug Resistance** *PLOS Computational Biology* **12** <https://doi.org/10.1371/journal.pcbi.1004710>

- Oz T *et al.* (2014) **Strength of Selection Pressure Is an Important Parameter Contributing to the Complexity of Antibiotic Resistance Evolution** *Molecular Biology and Evolution* **31**:2387–2401 <https://doi.org/10.1093/molbev/msu191>
- Pelleau S *et al.* (2015) **Adaptive evolution of malaria parasites in French Guiana: Reversal of chloroquine resistance by acquisition of a mutation in *pfcrt*** *Proceedings of the National Academy of Sciences* **112**:11672–11677 <https://doi.org/10.1073/pnas.1507142112>
- Pennings PS (2012) **Standing Genetic Variation and the Evolution of Drug Resistance in HIV** *PLoS Computational Biology* **8** <https://doi.org/10.1371/journal.pcbi.1002527>
- Phillips AM, Lawrence KR, Moulana A, Dupic T, Chang J, Johnson MS, Cvijovic I, Mora T, Walczak AM, Desai MM (2021) **Binding affinity landscapes constrain the evolution of broadly neutralizing anti-influenza antibodies** *eLife* **10** <https://doi.org/10.7554/elife.71393>
- Phillips AM, Maurer DP, Brooks C, Dupic T, Schmidt AG, Desai MM (2023) **Hierarchical sequence-affinity landscapes shape the evolution of breadth in an anti-influenza receptor binding site antibody** *eLife* **12** <https://doi.org/10.7554/elife.83628>
- Pimenta AC, Fernandes R, Moreira IS (2014) **Evolution of Drug Resistance: Insight on TEM β -Lactamases Structure and Activity and β -Lactam Antibiotics** *Mini-Reviews in Medicinal Chemistry* **14**:111–122 <https://doi.org/10.2174/1389557514666140123145809>
- Raghavan S *et al.* (2021) **Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer** *Cell* **184**:6119–6137
- Rollins NJ, Brock KP, Poelwijk FJ, Stiffler MA, Gauthier NP, Sander C, Marks DS (2019) **Inferring protein 3D structure from deep mutation scans** *Nature Genetics* **51**:1170–1176 <https://doi.org/10.1038/s41588-019-0432-9>
- Romero PA, Arnold FH (2009) **Exploring protein fitness landscapes by directed evolution** *Nature reviews Molecular cell biology* **10**:866–876
- Rosenbloom DIS, Hill AL, Rabi SA, Siliciano RF, Nowak MA (2012) **Antiretroviral dynamics determines HIV evolution and predicts therapy outcome** *Nature Medicine* **18**:1378–1385 <https://doi.org/10.1038/nm.2892>
- Singh N, Yeh PJ (2017) **Suppressive drug combinations and their potential to combat antibiotic resistance** *The Journal of Antibiotics* **70**:1033–1042 <https://doi.org/10.1038/ja.2017.102>
- Sirot D, Recule C, Chaibi EB, Bret L, Croize J, Chanal-Claris C, Labia R, Sirot J. (1997) **A complex mutant of TEM-1 beta-lactamase with mutations encountered in both IRT-4 and extended-spectrum TEM-15, produced by an Escherichia coli clinical isolate** *Antimicrobial Agents and Chemotherapy* **41**:1322–1325 <https://doi.org/10.1128/aac.41.6.1322>
- de Souza Cândido E, de Barros E, Cardoso MH, Franco OL (2019) **Bacterial cross-resistance to anti-infective compounds. Is it a real problem?** *Current Opinion in Pharmacology* **48**:76–81 <https://doi.org/10.1016/j.coph.2019.05.004>
- Steward CD, Wallace D, Hubert SK, Lawton R, Fridkin SK, Gaynes RP, McGowan JE, Tenover FC (2000) **Ability of laboratories to detect emerging antimicrobial resistance in nosocomial pathogens: a survey of Project ICARE laboratories** *Diagnostic Microbiology and Infectious Disease* **38**:59–67 [https://doi.org/10.1016/s0732-8893\(00\)00161-9](https://doi.org/10.1016/s0732-8893(00)00161-9)

- Stiffler MA, Hekstra DR, Ranganathan R. (2015) **Evolvability as a Function of Purifying Selection in TEM-1 β -Lactamase** *Cell* **160**:882–892 <https://doi.org/10.1016/j.cell.2015.01.035>
- Tan L, Serene S, Chao HX, Gore J. (2011) **Hidden Randomness between Fitness Landscapes Limits Reverse Evolution** *Physical Review Letters* **106** <https://doi.org/10.1103/physrevlett.106.198102>
- Tekin E, Savage VM, Yeh PJ (2017) **Measuring higher-order drug interactions: A review of recent approaches** *Current Opinion in Systems Biology* **4**:16–23 <https://doi.org/10.1016/j.coisb.2017.05.015>
- Toprak E, Veres A, Michel JB, Chait R, Hartl DL, Kishony R. (2011) **Evolutionary paths to antibiotic resistance under dynamically sustained drug selection** *Nature Genetics* **44**:101–105 <https://doi.org/10.1038/ng.1034>
- Turner PE, Williams ESCP, Okeke C, Cooper VS, Duffy S, Wertz JE (2014) **Antibiotic resistance correlates with transmission in plasmid evolution** *Evolution* **68**:3368–3380 <https://doi.org/10.1111/evo.12537>
- Weinreich DM, Delaney NF, DePristo MA, Hartl DL (2006) **Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins** *Science* **312**:111–114 <https://doi.org/10.1126/science.1123539>
- Weinreich DM, Lan Y, Jaffe J, Heckendorn RB (2018) **The Influence of Higher-Order Epistasis on Biological Fitness Landscape Topography** *Journal of Statistical Physics* **172**:208–225 <https://doi.org/10.1007/s10955-018-1975-3>
- Weinreich DM, Lan Y, Wylie CS, Heckendorn RB (2013) **Should evolutionary geneticists worry about higher-order epistasis?** *Current Opinion in Genetics & Development* **23**:700–707 <https://doi.org/10.1016/j.gde.2013.10.007>
- Yeh P, Tschumi AI, Kishony R. (2006) **Functional classification of drugs by properties of their pairwise interactions** *Nature Genetics* **38**:489–494 <https://doi.org/10.1038/ng1755>
- Yeh PJ, Hegreness MJ, Aiden AP, Kishony R. (2009) **Drug interactions and the evolution of antibiotic resistance** *Nature Reviews Microbiology* **7**:460–466 <https://doi.org/10.1038/nrmicro2133>

Article and author information

Rafael F. Guerrero

Department of Biological Sciences, North Carolina State University
ORCID iD: [0000-0002-8451-3609](https://orcid.org/0000-0002-8451-3609)

Tandin Dorji

Department of Mathematics and Statistics, University of Vermont, Burlington, VT

Ra'Mal M. Harris

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA

Matthew D. Shoulders

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA
ORCID iD: [0000-0002-6511-3431](https://orcid.org/0000-0002-6511-3431)

C. Brandon Ogbunugafor

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA,
Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, Santa Fe
Institute, Santa Fe, NM, Public Health Modeling Unit, Yale School of Public Health, New Haven,
CT

For correspondence: brandon.ogbunu@yale.edu
ORCID iD: [0000-0002-1581-8345](https://orcid.org/0000-0002-1581-8345)

Copyright

© 2023, Guerrero et al.

This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor

Sara Mitri

University of Lausanne, Switzerland

Senior Editor

George Perry

Pennsylvania State University, United States of America

Reviewer #1 (Public Review):

The manuscript by Guerrero and colleagues introduces two new metrics that extend the concept of "druggability" - loosely speaking, the potential suitability of a particular drug, target, or drug-target interaction for pharmacological intervention - to collections of drugs and genetic variants. The study draws on previously measured growth rates across a combinatorially complete mutational landscape involving 4 variants of the TEM-50 (beta lactamase) enzyme, which confers resistance to commonly used beta-lactam antibiotics. To quantify how growth rate - in this case, a proxy for evolutionary fitness - is distributed across allelic variants and drugs, they introduce two concepts: "variant vulnerability" and "drug applicability".

Variant vulnerability is the mean vulnerability (1-normalized growth rate) of a particular variant to a library of drugs, while drug applicability measures the mean across the collection of genetic variants for a given drug. The authors rank the drugs and variants according to these metrics. They show that the variant vulnerability of a particular mutant is uncorrelated with the vulnerability of its one-step neighbors, and analyze how higher-order combinations of single variants (SNPs) contribute to changes in growth rate in different drug environments.

The work addresses an interesting topic and underscores the need for evolution-based metrics to identify candidate pharmacological interventions for treating infections. The authors are clear about the limitations of their approach - they are not looking for immediate clinical applicability - and provide simple new measures of druggability that incorporate an

evolutionary perspective, an important complement to the orthodoxy of aggressive, kill-now design principles.

As I said in my initial review, I think the work could be improved with additional analysis that tie the new metrics to evolutionary outcomes. Without this evidence, or some other type of empirical or theoretical support for the utility of these metrics, I am not fully convinced that these concepts have substantial impact. The new metrics could indeed be useful--and they have intuitive appeal--but the current revisions stop short of demonstrating that these intuitive notions hold up under "realistic" conditions (whether in simulation, theory, or experiment).

Reviewer #2 (Public Review):

In the main text, the authors apply their metrics to a data set that was published by Mira et al. in 2015. The data consist of growth rate measurements for a combinatorially complete set of 16 genetic variants of the antibiotic resistance enzyme beta-lactamase across 10 drugs and drug combinations at 3 different drug concentrations, comprising a total of 30 different environmental conditions. In my previous report I had asked the authors to specify why they selected only 7 out of 30 environments for their analysis, with only one concentration for drug, but a clear explanation is still lacking. In the Data section of Material and Methods, the authors describe their criterion for data selection as follows: "we focus our analyses on drug treatments that had a significant negative effect on the growth of wildtype/TEM-1 strains". However, in Figure 2 it is seen that, even for the selected data sets, not all points are significant compared to wild type (grey points). So what criterion was actually applied?

In effect, for each chosen drug or drug combination, the authors choose the data set corresponding to the highest drug concentration. As a consequence, they cannot assess to what extent their metrics depend on drug concentration. This is a major concern, since Mira et al. concluded in their study that the differences between growth rate landscapes measured at different concentrations were comparable to the differences between drugs. I argued before that, if the new metrics display a significant dependence on drug concentration, this would considerably limit their usefulness. The authors challenge this, saying in their rebuttal that "no, that drug concentration would be a major actor in the value of the metrics does not limit the utility of the metric. It is simply another variable that one can consider when computing the metrics." While this is true in principle, I don't think any practicing scientist would disagree with the statement that the existence of additional confounding factors (in particular if they are unknown) reduces the usefulness of a quantitative metric.

As a consequence of the small number of variant-drug-combinations that are used, the conclusions that the authors draw from their analysis are mostly tentative. For example, on line 123 the authors write that the observation that the treatment of highest drug applicability is a combination of two drugs "fits intuition". In the Discussion this statement is partly retracted with reference to the piperacillin/tazobactam-combination which has low drug applicability. Being based on only a handful of data points, both observations are essentially anecdotal and it is unclear what the reader is supposed to learn.

To assess the environment-dependent epistasis among the genetic mutations comprising the variants under study, the authors decompose the data of Mira et al. into epistatic interactions of different orders. This part of the analysis is incomplete in two ways. First, in their study, Mira et al. pointed out that a fairly large fraction of the fitness differences between variants that they measured were not statistically significant. This information has been removed in the depiction of the Mira et al. fitness landscapes in Figure 1 of the present manuscript, and it does not seem to be reflected in the results of the interaction analysis in Figure 4. Second, the

interpretation of the coefficients obtained from the epistatic decomposition depends strongly on the formalism that is being used. In a note added on page 15 of the revised manuscript, the authors write that they have used the LASSO regression for their analysis and refer the reader to a previous publication (Guerrero et al. 2019) which however (as far as I could see) also does not fully explain how the method works. To give an example of the difficulty of interpreting the data in Figure 4 without further information: The substitution C (G238S) is well known to have a strong positive effect in cefotaxime, but the corresponding coefficient is essentially zero. So whatever the LASSO regression does, it cannot simply measure the effect on growth.

Reviewer #3 (Public Review):

The authors introduce two new concepts for antimicrobial resistance borrowed from pharmacology, "variant vulnerability" (how susceptible a particular resistance gene variant is across a class of drugs) and "drug applicability" (how useful a particular drug is against multiple allelic variants). They group both terms under an umbrella term "drugability". They demonstrate these features for an important class of antibiotics, the beta-lactams, and allelic variants of TEM-1 beta-lactamase. In the revised version, they investigate a second drug class that targets dihydrofolate reductase in *Plasmodium* (the causative agent of malaria).

The strength of the result is in its conceptual advance and that the concepts seem to work for beta-lactam resistance and DHFR inhibitors in a protozoan. However, I do not necessarily see the advance of lumping both terms under "drugability", as this adds an extra layer of complication in my opinion.

I think that the utility of the terms will be more comprehensively demonstrated by using examples across a breadth of drug classes and/or resistance genes. For instance, another good bacterial model with published data might have been trimethoprim resistance, which arises through point mutations in the *folA* gene (although, clinical resistance tends to be instead conferred by a suite of horizontally acquired dihydrofolate reductase genes, which are not so closely related as the TEM variants explored here).

The impact of the work on the field depends on a more comprehensive demonstration of the applicability of these new concepts to other drugs. This would be demonstrated in future work.

Author Response

The following is the authors' response to the original reviews.

Firstly, we must take a moment to express our sincere gratitude to editorial board for allowing this work to be reviewed, and to the peer reviewers for taking the time and effort to review our manuscript. The reviews are thoughtful and reflect the careful work of scientists who undoubtedly have many things on their schedule. We cannot express our gratitude enough. This is not a minor sentiment. We appreciate the engagement.

Allow us to briefly highlight some of the changes made to the revised manuscript, most on behalf of suggestions made by the reviewers:

1. A supplementary figure that includes the calculation of drug applicability and variant vulnerability for a different data set—16 alleles of dihydrofolate reductase, and two antifolate compounds used to treat malaria—pyrimethamine and cycloguanil.
2. New supplementary figures that add depth to the result in Figure 1 (the fitness graphs): we demonstrate how the rank order of alleles changes across drug environments and offer a statistical comparison of the equivalence of these fitness landscapes.

3. A new subsection that explains our specific method used to measure epistasis.
4. Improved main text with clarifications, fixed errors, and other addendums.
5. Improved referencing and citations, in the spirit of better scholarship (now with over 70 references).

Next, we'll offer some general comments that we believe apply to several of the reviews, and to the eLife assessment. We have provided the bulk of the responses in some general comments, and in response to the public reviews. We have also included the suggestions and made brief comments to some of the individual recommendations.

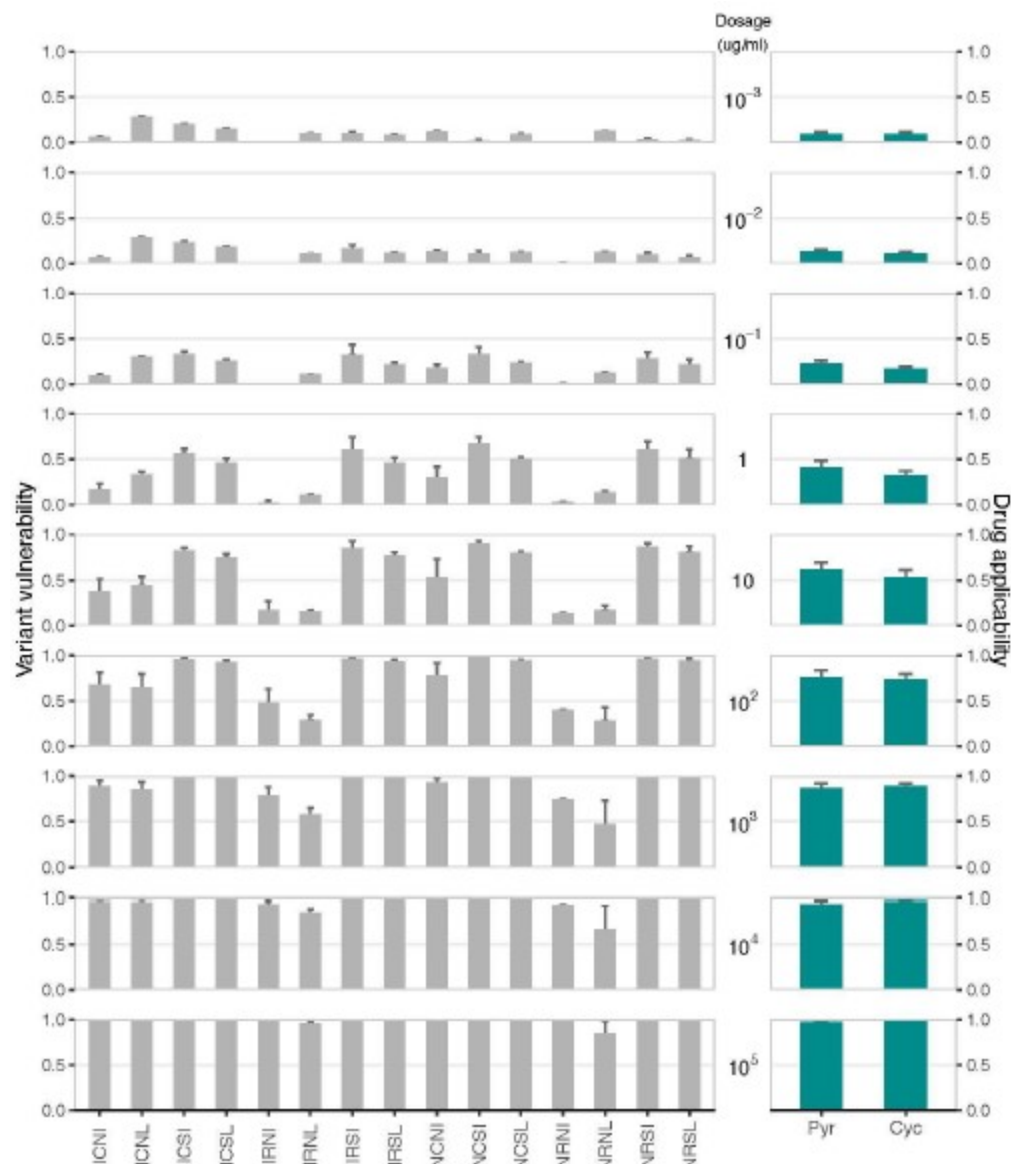
On the completeness of our analysis

In our response, we'll address the completeness issue first, as iterations of it appear in several of the reviews, and it seems to be one of the most substantive philosophical critiques of the work (there are virtually no technical corrections, outside of a formatting and grammar fixes, which we are grateful to the reviewers for identifying).

To begin our response, we will relay that we have now included an analysis of a data set corresponding to mutants of a protein, dihydrofolate reductase (DHFR), from *Plasmodium falciparum* (a main cause of malaria), across two antifolate drugs (pyrimethamine and cycloguanil). We have also decided to include this new analysis in the supplementary material (see Figure S4).

Author response image 1.

Drug applicability and variant vulnerability for 16 alleles of dihydrofolate reductase.



Here we compute the variant vulnerability and drug applicability metrics for two drugs, pyrimethamine (PYR) and cycloguanil (CYC), both antifolate drugs used to treat malaria. This is a completely different system than the one that is the focus of the submitted paper, for a different biomedical problem (antimalarial resistance), using different drugs, and targets. Further, the new data provide information on both drugs of different kinds, and drug concentrations (as suggested by Reviewer #1; we've also added a note about this in the new supplementary material). Note that these data have already been the subject of detailed analyses of epistatic effects, and so we did not include those here, but we do offer that reference:

- Ogbunugafor CB. The mutation effect reaction norm (mu-rn) highlights environmentally dependent mutation effects and epistatic interactions. *Evolution*. 2022 Feb 1;76(s1):37-48.
- Diaz-Colunga J, Sanchez A, Ogbunugafor CB. Environmental modulation of global epistasis is governed by effective genetic interactions. *bioRxiv*. 2022:202211.

Computing our proposed metrics across different drugs is relatively simple, and we could have populated our paper with suites of similar analyses across data sets of various kinds.

Such a paper would, in our view, be spread too thin—the evolution of antifolate resistance and/or antimalarial resistance are enormous problems, with large literatures that warrant focused studies. More generally, as the reviewers doubtlessly understand, simply analyzing more data sets does not make a study stronger, especially one like ours, that is using empirical data to both make a theoretical point about alleles and drugs and offer a metric that others can apply to their own data sets.

Our approach focused on a data set that allowed us to discuss the biology of a system: a far stronger paper, a far stronger proof-of-concept for a new metric. We will revisit this discussion about the structure of our study. But before doing so, we will elaborate on why the “more is better” tone of the reviews is misguided.

We also note that study where the data originate (Mira et al. 2015) is focused on a single data set of a single drug-target system. We should also point out that Mira et al. 2015 made a general point about drug concentrations influencing the topography of fitness landscapes, not unlike our general point about metrics used to understand features of alleles and different drugs in antimicrobial systems.

This isn't meant to serve as a feeble appeal to authority – just because something happened in one setting doesn't make it right for another. But other than a nebulous appeal to the fact that things have changed in the 8 years since that study was published, it is difficult to argue why one study system was permissible for other work but is somehow “incomplete” in ours. Double standards can be appropriate when they are justified, but in this case, it hasn't been made clear, and there is no technical basis for it.

Our study does what countless other successful ones do: utilizes a biological system to make a general point about some phenomena in the natural world. In our case, we were focused on the need for more evolution-inspired iterations of widely used concepts like druggability. For example, a recent study of epistasis focused on a single set of alleles, across several drugs, not unlike our study:

- Lozovsky ER, Daniels RF, Heffernan GD, Jacobus DP, Hartl DL. Relevance of higher-order epistasis in drug resistance. *Molecular biology and evolution*. 2021 Jan;38(1):142-51.

Next, we assert that there is a difference between an eagerness to see a new metric applied to many different data sets (a desire we share, and plan on pursuing in the future), and the notion that an analysis is “incomplete” without it. The latter is a more serious charge and suggests that the researcher-authors neglected to properly construct an argument because of gaps in the data. This charge does not apply to our manuscript, at all. And none of the reviewers effectively argued otherwise.

Our study contains 7 different combinatorially-complete datasets, each composed of 16 alleles (this not including the new analysis of antifolates that now appear in the revision). One can call these datasets “small” or “low-dimensional,” if they choose (we chose to put this front-and-center, in the title). They are, however, both complete and as large or larger than many datasets in similar studies of fitness landscapes:

- Knies JL, Cai F, Weinreich DM. Enzyme efficiency but not thermostability drives cefotaxime resistance evolution in TEM-1 β -lactamase. *Molecular biology and evolution*. 2017 May 1;34(5):1040-54.

- Lozovsky ER, Daniels RF, Heffernan GD, Jacobus DP, Hartl DL. Relevance of higher-order epistasis in drug resistance. *Molecular biology and evolution*. 2021 Jan;38(1):142-51.

- Rodrigues JV, Bershtein S, Li A, Lozovsky ER, Hartl DL, Shakhnovich EI. Biophysical principles predict fitness landscapes of drug resistance. *Proceedings of the National Academy of Sciences*. 2016 Mar 15;113(11):E1470-8.

- Ogbunugafor CB, Eppstein MJ. Competition along trajectories governs adaptation rates towards antimicrobial resistance. *Nature ecology & evolution*. 2016 Nov 21;1(1):0007.
- Lindsey HA, Gallie J, Taylor S, Kerr B. Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature*. 2013 Feb 28;494(7438):463-7.

These are only five of very many such studies, some of them very well-regarded.

Having now gone on about the point about the data being “incomplete,” we’ll next move to the more tangible comment-criticism about the low-dimensionality of the data set, or the fact that we examined a single drug-drug target system (β lactamases, and β -lactam drugs).

The criticism, as we understand it, is that the authors could have analyzed more data,

This is a common complaint, that “more is better” in biology. While we appreciate the feedback from the reviewers, we notice that no one specified what constitutes the right amount of data. Some pointed to other single data sets, but would analyzing two different sets qualify as enough? Perhaps to person A, but not to persons B - Z. This is a matter of opinion and is not a rigorous comment on the quality of the science (or completeness of the analysis).

- Should we analyze five more drugs of the same target (beta lactamases)? And what bacterial orthologs?
- Should we analyze 5 antifolates for 3 different orthologs of dihydrofolate reductase?
- And in which species or organism type? Bacteria? Parasitic infections?
- And why only infectious disease? Aren’t these concepts also relevant to cancer? (Yes, they are.)
- And what about the number of variants in the aforementioned target? Should one aim for small combinatorially complete sets? Or vaster swaths of sequence space, such as the ones generated by deep mutational scanning and other methods?

I offer these options in part because, for the most part, were not given an objective suggestion for appropriate level of detail. This is because there is no answer to the question of what size of dataset would be most appropriate. Unfortunately, without a technical reason why a data set of unspecified size [X] or [Y] is best, then we are left with a standard “do more work” peer review response, one that the authors are not inclined to engage seriously, because there is no scientific rationale for it.

The most charitable explanation for why more datasets would be better is tied to the abstract notion that seeing a metric measured in different data sets somehow makes it more believable. This, as the reviewers undoubtedly understand, isn’t necessarily true (in fact, many poor studies mask a lack of clarity with lots of data).

To double down on this take, we’ll even argue the opposite: that our focus on a single drug system is a strength of the study.

The focus on a single-drug class allows us to practice the lost art of discussing the peculiar biology of the system that we are examining. Even more, the low dimensionality allows us to discuss—in relative detail—individual mutations and suites of mutations. We do so several times in the manuscript, and even connect our findings to literature that has examined the biophysical consequences of mutations in these very enzymes.

(For example: Knies JL, Cai F, Weinreich DM. Enzyme efficiency but not thermostability drives cefotaxime resistance evolution in TEM-1 β -lactamase. *Molecular biology and evolution*. 2017 May 1;34(5):1040-54.)

Such detail is only legible in a full-length manuscript because we were able to interrogate a system in good detail. That is, the low-dimensionality (of a complete data set) is a strength, rather than a weakness. This was actually part of the design choice for the study: to offer a new metric with broad application but developed using a system where the particulars could be interrogated and discussed.

Surely the findings that we recover are engineered for broader application. But to suggest that we need to apply them broadly in order to demonstrate their broad impact is somewhat antithetical to both model systems research and to systems biology, both of which have been successful in extracting general principles for singular (often simple) systems and models.

An alternative approach, where the metric was wielded across an unspecified number of datasets would lend to a manuscript that is unfocused, reading like many modern machine learning papers, where the analysis or discussion have little to do with actual biology. We very specifically avoided this sort of study.

To close our comments regarding data: Firstly, we have considered the comments and analyzed a different data set, corresponding to a different drug-target system (antifolate drugs, and DHFR). Moreover, we don't think more data has anything to do with a better answer or support for our conclusions or any central arguments. Our arguments were developed from the data set that we used but achieve what responsible systems biology does: introduces a framework that one can apply more broadly. And we develop it using a complete, and well-vetted dataset. If the reviewers have a philosophical difference of opinion about this, we respect it, but it has nothing to do with our study being "complete" or not. And it doesn't speak to the validity of our results.

Related: On the dependence of our metrics on drug-target system

Several comments were made that suggest the relevance of the metric may depend on the drug being used. We disagree with this, and in fact, have argued the opposite: the metrics are specifically useful because they are not encumbered with unnecessary variables. They are the product of rather simple arithmetic that is completely agnostic to biological particulars.

We explain, in the section entitled "Metric Calculations:

"To estimate the two metrics we are interested in, we must first quantify the susceptibility of an allelic variant to a drug. We define susceptibility as $1 - w$, where w is the mean growth of the allelic variant under drug conditions relative to the mean growth of the wild-type/TEM-1 control. If a variant is not significantly affected by a drug (i.e., growth under drug is not statistically lower than growth of wild-type/TEM-1 control, by t-test P -value < 0.01), its susceptibility is zero. Values in these metrics are summaries of susceptibility: the variant vulnerability of an allelic variant is its average susceptibility across drugs in a panel, and the drug applicability of an antibiotic is the average susceptibility of all variants to it."

That is, these can be animated to compute the variant vulnerability and drug applicability for data sets of various kinds. To demonstrate this (and we thank the reviewers for suggesting it), we have analyzed the antifolate-DHFR data set as outlined above.

Finally, we will make the following light, but somewhat cynical point (that relates to the "more data" more point generally): the wrong metric applied to 100 data sets is little more than 100 wrong analyses. Simply applying the metric to a wide number of datasets has nothing to do with the veracity of the study. Our study, alternatively, chose the opposite approach: used a data set for a focused study where metrics were extracted. We believe this to be a much more rigorous way to introduce new metrics.

On the Relevance of simulations

Somewhat relatedly, the eLife summary and one of the reviewers mentioned the potential benefit of simulations. Reviewer 1 correctly highlights that the authors have a lot of experience in this realm, and so generating simulations would be trivial. For example, the authors have been involved in studies such as these:

- Ogbunugafor CB, Eppstein MJ. Competition along trajectories governs adaptation rates towards antimicrobial resistance. *Nature ecology & evolution*. 2016 Nov 21;1(1):0007.
- Ogbunugafor CB, Wylie CS, Diakite I, Weinreich DM, Hartl DL. Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance. *PLoS computational biology*. 2016 Jan 25;12(1):e1004710.
- Ogbunugafor CB, Hartl D. A pivot mutation impedes reverse evolution across an adaptive landscape for drug resistance in *Plasmodium vivax*. *Malaria Journal*. 2016 Dec;15:1-0.

From the above and dozens of other related studies, we've learned that simulations are critical for questions about the end results of dynamics across fitness landscapes of varying topography. To simulate across the datasets in the submitted study would be a small ask. We do not provide this, however, because our study is not about the dynamics of *de novo* evolution of resistance. In fact, our study focuses on a different problem, no less important for understanding how resistance evolves: determining static properties of alleles and drugs, that provide a picture into their ability to withstand a breadth of drugs in a panel (variant vulnerability), or the ability of a drug in a panel to affect a breadth of drug targets.

The authors speak on this in the Introduction:

“While stepwise, *de novo* evolution (via mutations and subsequent selection) is a key force in the evolution of antimicrobial resistance, evolution in natural settings often involves other processes, including horizontal gene transfer and selection on standing genetic variation. Consequently, perspectives that consider variation in pathogens (and their drug targets) are important for understanding treatment at the bedside. Recent studies have made important strides in this arena. Some have utilized large data sets and population genetics theory to measure cross-resistance and collateral sensitivity. Fewer studies have made use of evolutionary concepts to establish metrics that apply to the general problem of antimicrobial treatment on standing genetic variation in pathogen populations, or for evaluating the utility of certain drugs' ability to treat the underlying genetic diversity of pathogens”

That is, the proposed metrics aren't about the dynamics of stepwise evolution across fitness landscapes, and so, simulating those dynamics don't offer much for our question. What we have done instead is much more direct and allows the reader to follow a logic: clearly demonstrate the topography differences in Figure 1 (And Supplemental Figure S2 and S3 with rank order changes).

Relative growth -- ranked within drug

LKSD	10	13	11	5	10	11	4
LESD	3	5	16	13	6	12	6
LKGD	6	15	3	6	13	10	16
LKSN	2	4	7	9	1	6	5
MKSD	1	9	4	7	2	2	2
LKGN	16	9	8	4	5	7	13
LESN	11	7	12	10	8	5	3
MKSN	9	12	13	16	16	15	9
LEGD	12	16	5	8	4	3	11
MKGD	4	3	10	1	9	13	15
MESD	7	1	14	12	12	14	1
LEGN	5	14	9	2	14	4	12
MKGN	14	8	6	14	15	1	14
MESN	13	5	15	15	7	16	8
MEGD	8	2	1	3	3	9	7
MEGN	15	11	2	11	11	8	10

Amoxicillin
Amoxicillin/
Clavulanic-Acid
Cefotaxime
Cefotetan
Cefprozil
Ceftazidime
Piperacillin/
Tazobactam

These results tell the reader what they need to know: that the topography of fitness landscapes changes across drug types. Further, we should note that Mira et al. 2015 already told the basic story that one finds different adaptive solutions across different drug environments. (Notably, without computational simulations).

In summary, we attempted to provide a rigorous, clean, and readable study that introduced two new metrics. Appeals to adding extra analysis would be considered if they augmented the study's goals. We do not believe this to be the case.

Nonetheless, we must reiterate our appreciation for the engagement and suggestions. All were made with great intentions. This is more than one could hope for in a peer review exchange. The authors are truly grateful.

eLife assessment

The work introduces two valuable concepts in antimicrobial resistance: "variant vulnerability" and "drug applicability", which can broaden our ways of thinking about microbial infections through evolution-based metrics. The authors present a compelling

analysis of a published dataset to illustrate how informative these metrics can be, study is still incomplete, as only a subset of a single dataset on a single class of antibiotics was analyzed. Analyzing more datasets, with other antibiotic classes and resistance mutations, and performing additional theoretical simulations could demonstrate the general applicability of the new concepts.

The authors disagree strongly with the idea that the study is ‘incomplete,’ and encourage the editors and reviewers to reconsider this language. Not only are the data combinatorially complete, but they are also larger in size than many similar studies of fitness landscapes. Insofar as no technical justification was offered for this “incomplete” summary, we think it should be removed. Furthermore, we question the utility of “theoretical simulations.” They are rather easy to execute but distract from the central aims of the study: to introduce new metrics, in the vein of other metrics—like druggability, IC50, MIC—that describe properties of drugs or drug targets.

Public Reviews:

Reviewer #1 (Public Review):

The manuscript by Geurrero and colleagues introduces two new metrics that extend the concept of "druggability"- loosely speaking, the potential suitability of a particular drug, target, or drug-target interaction for pharmacological intervention-to collections of drugs and genetic variants. The study draws on previously measured growth rates across a combinatorially complete mutational landscape involving 4 variants of the TEM-50 (beta lactamase) enzyme, which confers resistance to commonly used beta-lactam antibiotics. To quantify how growth rate - in this case, a proxy for evolutionary fitness - is distributed across allelic variants and drugs, they introduce two concepts: "variant vulnerability" and "drug applicability".

Variant vulnerability is the mean vulnerability (1-normalized growth rate) of a particular variant to a library of drugs, while drug applicability measures the mean across the collection of genetic variants for a given drug. The authors rank the drugs and variants according to these metrics. They show that the variant vulnerability of a particular mutant is uncorrelated with the vulnerability of its one-step neighbors and analyze how higher-order combinations of single variants (SNPs) contribute to changes in growth rate in different drug environments.

The work addresses an interesting topic and underscores the need for evolutionbased metrics to identify candidate pharmacological interventions for treating infections. The authors are clear about the limitations of their approach - they are not looking for immediate clinical applicability - and provide simple new measures of druggability that incorporate an evolutionary perspective, an important complement to the orthodoxy of aggressive, kill-now design principles. I think the ideas here will interest a wide range of readers, but I think the work could be improved with additional analysis - perhaps from evolutionary simulations on the measured landscapes - that tie the metrics to evolutionary outcomes.

The authors greatly appreciate these comments, and the proposed suggestions by reviewer 1. We have addressed most of the criticisms and suggestions in our comments above.

Reviewer #2 (Public Review):

The authors introduce the notions of "variant vulnerability" and "drug applicability" as metrics quantifying the sensitivity of a given target variant across a panel of drugs and the effectiveness of a drug across variants, respectively. Given a data set comprising a measure of drug effect (such as growth rate suppression) for pairs of variants and drugs,

the vulnerability of a variant is obtained by averaging this measure across drugs, whereas the applicability of a drug is obtained by averaging the measure across variants.

The authors apply the methodology to a data set that was published by Mira et al. in 2015. The data consist of growth rate measurements for a combinatorially complete set of 16 genetic variants of the antibiotic resistance enzyme betalactamase across 10 drugs and drug combinations at 3 different drug concentrations, comprising a total of 30 different environmental conditions. For reasons that did not become clear to me, the present authors select only 7 out of 30 environments for their analysis. In particular, for each chosen drug or drug combination, they choose the data set corresponding to the highest drug concentration. As a consequence, they cannot assess to what extent their metrics depend on drug concentration. This is a major concern since Mira et al. concluded in their study that the differences between growth rate landscapes measured at different concentrations were comparable to the differences between drugs. If the new metrics display a significant dependence on drug concentration, this would considerably limit their usefulness.

The authors appreciate the point about drug concentration, and it is one that the authors have made in several studies.

The quick answer is that whether the metrics are useful for drug type-concentration A or B will depend on drug type-concentration A or B. If there are notable differences in the topography of the fitness landscape across concentration, then we should expect the metrics to differ. What Reviewer #2 points out as a “major concern,” is in fact a strength of the metrics: it is agnostic with respect to type of drug, type of target, size of dataset, or topography of the fitness landscape. And so, the authors disagree: no, that drug concentration would be a major actor in the value of the metrics does not limit the utility of the metric. It is simply another variable that one can consider when computing the metrics.

As discussed above, we have analyzed data from a different data set, in a different drug-target problem (DHFR and antifolate drugs; see supplemental information). These demonstrate how the metric can be used to compute metrics across different drug concentrations.

As a consequence of the small number of variant-drug combinations that are used, the conclusions that the authors draw from their analysis are mostly tentative with weak statistical support. For example, the authors argue that drug combinations tend to have higher drug applicability than single drugs, because a drug combination ranks highest in their panel of 7. However, the effect profile of the single drug cefprozil is almost indistinguishable from that of the top-ranking combination, and the second drug combination in the data set ranks only 5th out of 7.

We reiterate our appreciation for the engagement. Reviewer #2 generously offers some technical insight on measurements of epistasis, and their opinion on the level of statistical support for our claims. The authors are very happy to engage in a dialogue about these points. We disagree rather strongly, and in addition to the general points raised above (that speak to some of this), will raise several specific rebuttals to the comments from Reviewer #2.

For one, the Reviewer #2 is free to point to what arguments have “weak statistical support.” Having read the review, we aren’t sure what this is referring to. “Weak statistical support” generally applies to findings built from underpowered studies, or designs constructed in manner that yield effect sizes or p-values that give low confidence that a finding is believable (or is replicable). This sort of problem doesn’t apply to our study for various reasons, the least

of which being that our findings are strongly supported, based on a vetted data set, in a system that has long been the object of examination in studies of antimicrobial resistance.

For example, we did not argue that magnetic fields alter the topography of fitness landscapes, a claim which must stand up to a certain sort of statistical scrutiny. Alternatively, we examined landscapes where the drug environment differed statistically from the non-drug environment and used them to compute new properties of alleles and drugs.

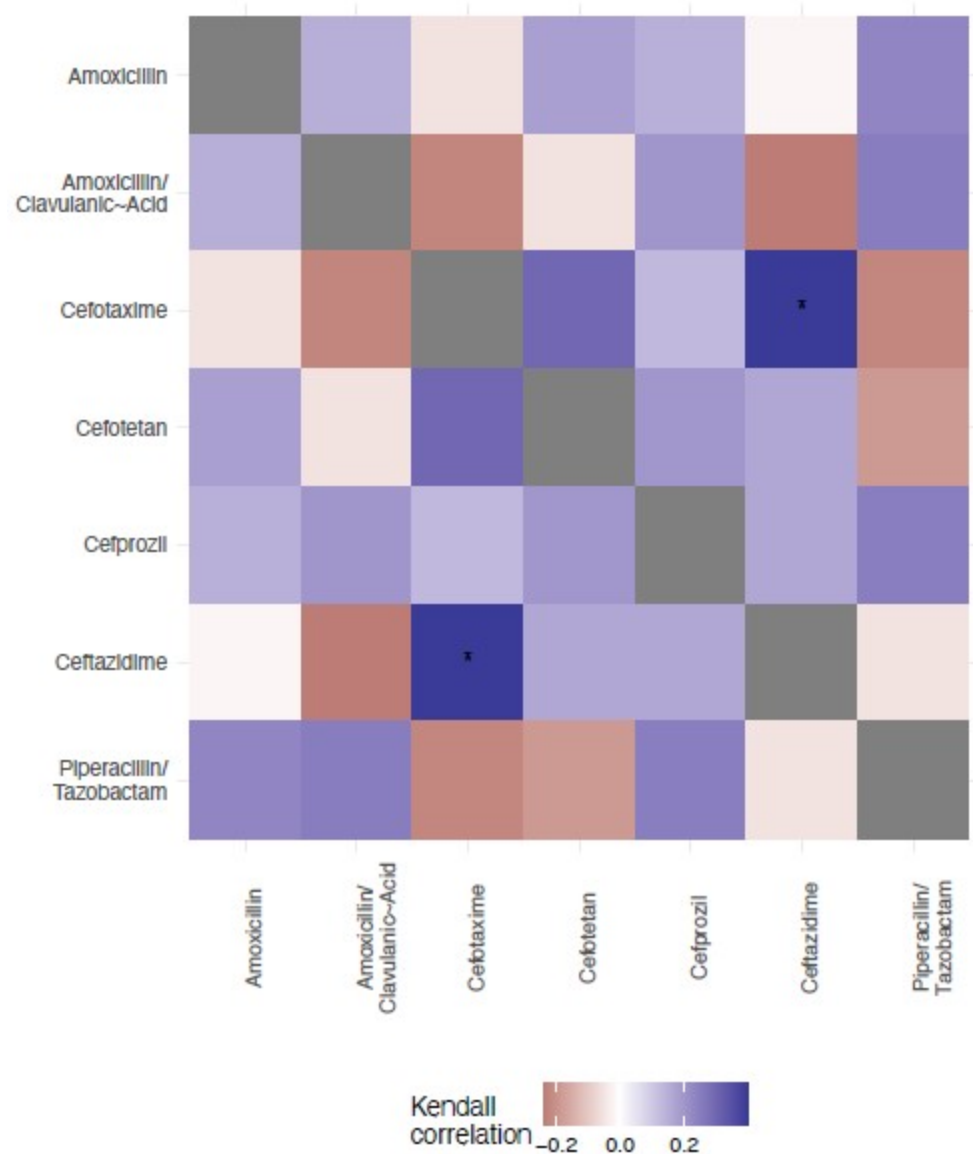
We can imagine that the reviewer is referring to the low-dimensionality of the fitness landscapes in the study. Again: the features of the dataset are a detail that the authors put into the title of the manuscript. Further, we emphasize that it is not a weakness, but rather, allows the authors to focus, and discuss the specific biology of the system. And we responsibly explain the constraints around our study several times, though none of them have anything to do with “weak statistical support.”

Even though we aren’t clear what “weak statistical support” means as offered by Reviewer 2, the authors have nonetheless decided to provide additional analyses, now appearing in the new supplemental material.

We have included a new Figure S2, where we offer an analysis of the topography of the 7 landscapes, based on the Kendall rank order test. This tests the hypothesis that there is no correlation (concordance or discordance) between the topographies of the fitness landscapes.

Author response image 3.

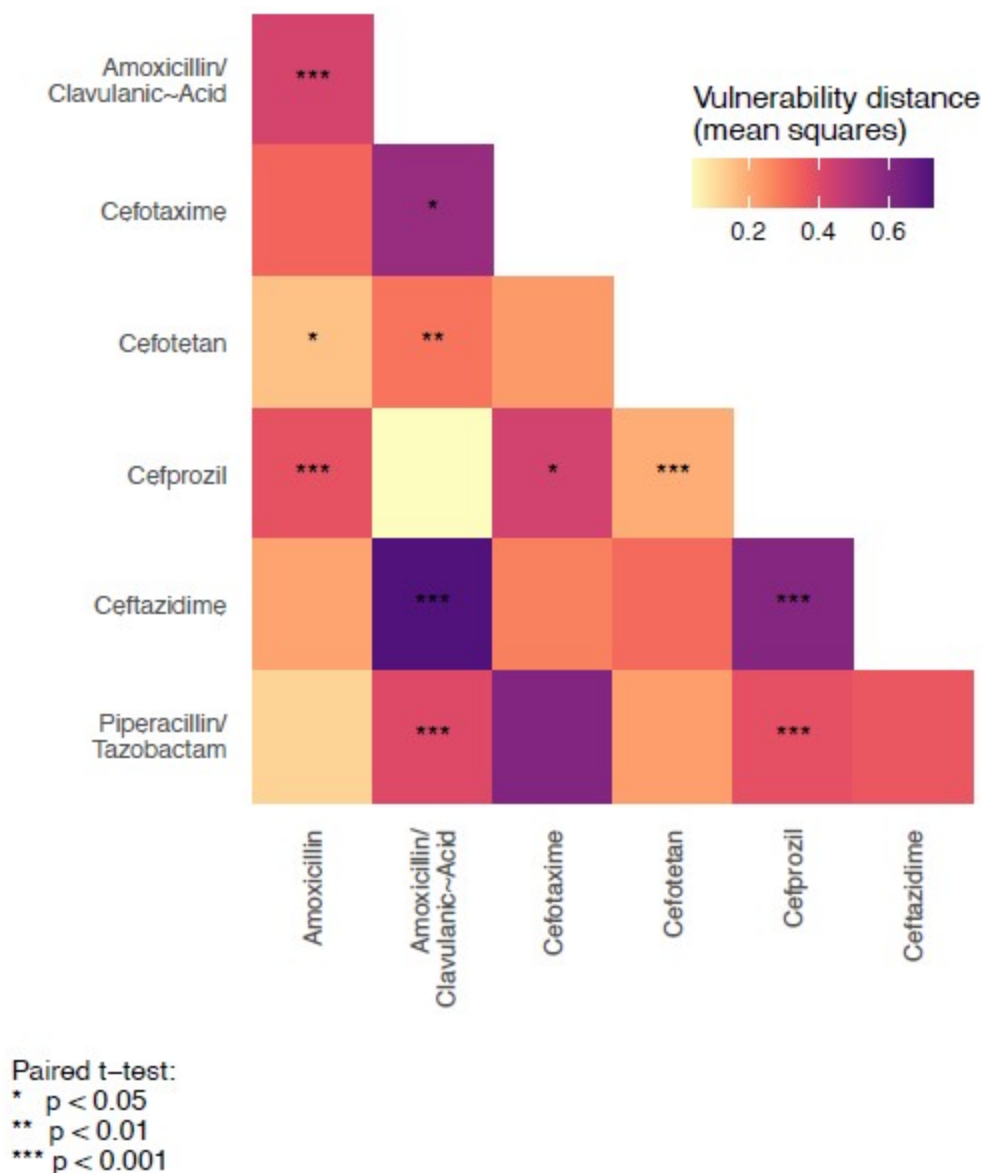
Kendall rank test for correlation between the 7 fitness landscapes.



In Figure S3, we test the hypothesis that the variant vulnerability values differ. To do this, we calculate a paired t-test. These are paired by haplotype/allelic variant, so the comparisons are change in growth between drugs for each haplotype.

Author response image 4.

Paired t-tests for variant vulnerability.



To this point raised by Reviewer #2:

“For example, the authors argue that drug combinations tend to have higher drug applicability than single drugs, because a drug combination ranks highest in their panel of 7. However, the effect profile of the single drug cefprozil is almost indistinguishable from that of the top-ranking combination, and the second drug combination in the data set ranks only 5th out of 7.”

Our study does not argue that drug combinations are necessarily correlated with a higher drug applicability. Alternatively, we specifically highlight that one of the combinations does not have a high drug applicability:

“Though all seven drugs/combinations are β -lactams, they have widely varying effects across the 16 alleles. Some of the results are intuitive: for example, the drug regime with the highest drug applicability of the set—amoxicillin/clavulanic acid—is a mixture of a widely used β -lactam (amoxicillin) and a β -lactamase inhibitor (clavulanic acid) (see Table 3). We might expect such a mixture to have a broader effect across a diversity of variants. This high

applicability is hardly a rule, however, as another mixture in the set, piperacillin/tazobactam, has a much lower drug applicability (ranking 5th out of the seven drugs in the set) (Table 3)."

In general, we believe that the submitted paper is responsible with regards to how it extrapolates generalities from the results. Further, the manuscript contains a specific section that explains limitations, clearly and transparently (not especially common in science). For that reason, we'd encourage reviewer #2 to reconsider their perspective. We do not believe that our arguments are built on "weak" support at all. And we did not argue anything particular about drug combinations writ large. We did the opposite— discussed the particulars of our results in light of the biology of the system.

Thirdly, to this point:

"To assess the environment-dependent epistasis among the genetic mutations comprising the variants under study, the authors decompose the data of Mira et al. into epistatic interactions of different orders. This part of the analysis is incomplete in two ways. First, in their study, Mira et al. pointed out that a fairly large fraction of the fitness differences between variants that they measured were not statistically significant, which means that the resulting fitness landscapes have large statistical uncertainties. These uncertainties should be reflected in the results of the interaction analysis in Figure 4 of the present manuscript."

The authors are uncertain with regards to the "uncertainties" being referred to, but we'll do our best to understand: our study utilized the 7 drug environments from Mira et al. 2015 with statistically significant differences between growth rates with and without drug. And so, this point about how the original set contained statistically insignificant treatments is not relevant here. We explain this in the methods section:

"The data that we examine comes from a past study of a combinatorial set of four mutations associated with TEM-50 resistance to β -lactam drugs [39]. This past study measured the growth rates of these four mutations in combination, across 15 different drugs (see Supplemental Information)."

We go on to say the following:

"We examined these data, identifying a subset of structurally similar β -lactams that also included β -lactams combined with β -lactamase inhibitors, cephalosporins and penicillins. From the original data set, we focus our analyses on drug treatments that had a significant negative effect on the growth of wild-type/TEM-1 strains (one-tailed ttest of wild-type treatment vs. control, $P < 0.01$). After identifying the data from the set that fit our criteria, we were left with seven drugs or combinations (concentration in $\mu\text{g/ml}$): amoxicillin 1024 $\mu\text{g/ml}$ (β -lactam), amoxicillin/clavulanic acid 1024 $\mu\text{g/ml}$ (β -lactam and β -lactamase inhibitor) cefotaxime 0.123 $\mu\text{g/ml}$ (third-generation cephalosporin), cefotetan 0.125 $\mu\text{g/ml}$ (second-generation cephalosporins), cefprozil 128 $\mu\text{g/ml}$ (second-generation cephalosporin), ceftazidime 0.125 $\mu\text{g/ml}$ (third-generation cephalosporin), piperacillin and tazobactam 512/8 $\mu\text{g/ml}$ (penicillin and β -lactamase inhibitor). With these drugs/mixtures, we were able to embody chemical diversity in the panel."

Again: The goal of our study was to develop metrics that can be used to analyze features of drugs and targets and disentangle these metrics into effects.

Second, the interpretation of the coefficients obtained from the epistatic decomposition depends strongly on the formalism that is being used (in the jargon of the field, either a Fourier or a Taylor analysis can be applied to fitness landscape data). The authors need to specify which formalism they have employed and phrase their interpretations accordingly.

The authors appreciate this nuance. Certainly, how to measure epistasis is a large topic of its own. But we recognize that we could have addressed this more directly and have added text to this effect.

In response to these comments from Reviewer #2, we have added a new section focused on these points (reference syntax removed here for clarity; please see main text for specifics):

“The study of epistasis, and discussions regarding the means to detect and measure now occupies a large corner of the evolutionary genetics literature. The topic has grown in recent years as methods have been applied to larger genomic data sets, biophysical traits, and the “global” nature of epistatic effects. We urge those interested in more depth treatments of the topic to engage larger summaries of the topic.”

“Here will briefly summarize some methods used to study epistasis on fitness landscapes. Several studies of combinatorially-complete fitness landscapes use some variation of Fourier Transform or Taylor formulation. One in particular, the Walsh-Hadamard Transform has been used to measure epistasis across a wide number of study systems. Furthermore, studies have reconciled these methods with others, or expanded upon the Walsh-Hadamard Transform in a way that can accommodate incomplete data sets. These methods are effective for certain sorts of analyses, and we strongly urge those interested to examine these studies.”

“The method that we’ve utilized, the LASSO regression, determines effect sizes for all interactions (alleles and drug environments). It has been utilized for data sets of similar size and structure, on alleles resistant to trimethoprim. Among many benefits, the method can accommodate gaps in data and responsibly incorporates experimental noise into the calculation.”

As Reviewer #2 understands, there are many ways to examine epistasis on both high and low-dimensional landscapes. Reviewer #2 correctly offers two sorts of formalisms that allow one to do so. The two offered by Reviewer #2, are not the only means of measuring epistasis in data sets like the one we have offered. But we acknowledge that we could have done a better job outlining this. We thank Reviewer #2 for highlighting this, and believe our revision clarifies this.

Reviewer #3 (Public Review):

The authors introduce two new concepts for antimicrobial resistance borrowed from pharmacology, “variant vulnerability” (how susceptible a particular resistance gene variant is across a class of drugs) and “drug applicability” (how useful a particular drug is against multiple allelic variants). They group both terms under an umbrella term “drugability”. They demonstrate these features for an important class of antibiotics, the beta-lactams, and allelic variants of TEM-1 beta-lactamase.

The strength of the result is in its conceptual advance and that the concepts seem to work for beta-lactam resistance. However, I do not necessarily see the advance of lumping both terms under “drugability”, as this adds an extra layer of complication in my opinion.

Firstly, the authors greatly appreciate the comments from Reviewer #3. They are insightful, and prescriptive. And allow us to especially thank reviewer 3 for supplying a commented PDF with some grammatical and phrasing suggestions/edits. This is much appreciated. We have examined all these suggestions and made changes.

In general, we agree with the spirit of many of the comments. In addition to our prior comments on the scope of our data, we’ll communicate a few direct responses to specific points raised.

I also think that the utility of the terms could be more comprehensively demonstrated by using examples across different antibiotic classes and/or resistance genes. For instance, another good model with published data might have been trimethoprim resistance, which arises through point mutations in the folA gene (although, clinical resistance tends to be instead conferred by a suite of horizontally acquired dihydrofolate reductase genes, which are not so closely related as the TEM variants explored here).

1. In our new supplemental material, we now feature an analysis of antifolate drugs, pyrimethamine and cycloguanil. We have discussed this in detail above and thank the reviewer for the suggestion.
2. Secondly, we agree that the study will have a larger impact when the metrics are applied more broadly. This is an active area of investigation, and our hope is that others apply our metrics more broadly. But as we discussed, such a desire is not a technical criticism of our own study. We stand behind the rigor and insight offered by our study.

The impact of the work on the field depends on a more comprehensive demonstration of the applicability of these new concepts to other drugs.

The authors don't disagree with this point, which applies to virtually every potentially influential study. The importance of a single study can generally only be measured by its downstream application. But this hardly qualifies as a technical critique of our study and does not apply to our study alone. Nor does it speak to the validity of our results. The authors share this interest in applying the metric more broadly.

Reviewer #1 (Recommendations For The Authors):

- *The main weakness of the work, in my view, is that it does not directly tie these new metrics to a quantitative measure of "performance". The metrics have intuitive appeal, and I think it is likely that they could help guide treatment options—for example, drugs with high applicability could prove more useful under particular conditions. But as the authors note, the landscape is rugged and intuitive notions of evolutionary behavior can sometimes fail. I think the paper would be much improved if the authors could evaluate their new metrics using some type of quantitative evolutionary model. For example, perhaps the authors could simulate evolutionary dynamics on these landscapes in the presence of different drugs. Is the mean fitness achieved in the simulations correlated with, for example, the drug applicability when looking across an ensemble of simulations with the same drug but varied initial conditions that start from each individual variant? Similarly, if you consider an ensemble of simulations where each member starts from the same variant but uses a different drug, is the average fitness gain captured in some way by the variant vulnerability? All simulations will have limitations, of course, but given that the landscape is fully known I think these questions could be answered under some conditions (e.g. strong selection weak mutation limit, where the model could be formulated as a Markov Chain; see 10.1371/journal.pcbi.1004493 or doi: 10.1111/evo.14121 for examples). And given the authors' expertise in evolutionary dynamics, I think it could be achieved in a reasonable time. With that said, I want to acknowledge that with any new "metrics", it can be tempting to think that "we need to understand it all" before it is useful, and I don't want to fall into that trap here.*

The authors respect and appreciate these thoughtful comments.

As Reviewer #1 highlighted, the authors are experienced with building simulations of evolution. For reasons we have outlined above, we don't believe they would add to the arc of the current story and may encumber the story with unnecessary distractions. Simulations of evolution can be enormously useful for studies focused on particulars of the dynamics of evolution. This submitted study is not one of those. It is charged with identifying features of alleles and drugs that capture an allele's vulnerability to treatment (variant vulnerability) and a drug's effectiveness across alleles (drug applicability). Both features integrate aspects of variation (genetic and environmental), and as such, are improvements over both metrics used to describe drug targets and drugs.

- The new metrics rely on means, which is a natural choice. Have the authors considered how variance (or other higher moments) might also impact evolutionary dynamics? I would imagine, for example, that the ultimate outcome of a treatment might depend heavily on the shape of the distribution, not merely its mean. This is also something one might be able to get a handle on with simulations.*

These are relevant points, and the authors appreciate them. Certainly, moments other than the mean might have utility. This is the reason that we computed the one-step neighborhood variant vulnerability—to see if the variant vulnerability of an allele was related to properties of its mutational neighborhood. We found no such correlation. There are many other sorts of properties that one might examine (e.g., shape of the distribution, properties of mutational network, variance, fano factor, etc). As we don't have an informed reason to pursue any of this in lieu of others, we are pleased to investigate this in the future.

Also, while we've addressed general points about simulations above, we want to note that our analysis of environmental epistasis does consider the variance. We urge Reviewer #1 to see our new section on "Notes on Methods Used to Measure Epistasis" where we explain some of this and supply references to that effect.

- As I understand it, the fitness measurements here are measures of per capita growth rate, which is reasonable. However, the authors may wish to briefly comment on the limitations of this choice-i.e. the fact that these are not direct measures of relative fitness values from head-to-head competition between strains.*

Reviewer #1 is correct: the metrics are computed from means. As Reviewer 1 definitely understands, debates over what measurements are proper proxies for fitness go back a long time. We added a slight acknowledgement about the existence of multiple fitness proxies in our revision.

- The authors consider one-step variant vulnerability. Have the authors considered looking at 2-step, 3-step, etc analogs of the 1-step vulnerability? I wonder if these might suggest potential vulnerability bottlenecks associated with the use of a particular drug/drug combo or trajectories starting from particular variants.*

This is an interesting point. We provided one-step values as a means of interrogating the mutational neighborhood of alleles in the fitness landscape. While there could certainly be other pattern-relationships between the variant vulnerability and features of a fitness landscape (as the reviewer recognizes), we don't have a rigorous reason to test them, other than an appeal to "I would be curious if [Blank]." As in, attempting to saturate the paper with these sorts of examinations might be fun, could turn up an interesting result, but this is true for most studies.

To highlight just how serious we are about future questions along these lines, we'll offer one specific question about the relationship between metrics and other features of alleles or landscapes. Recent studies have examined the existence of "evolvability-enhancing mutations," that propel a population to high-fitness sections of a fitness landscape:

- Wagner, A. Evolvability-enhancing mutations in the fitness landscapes of an RNA and a protein. Nat Commun 14, 3624 (2023). <https://doi.org/10.1038/s41467023-39321-8>

One present and future area of inquiry involves whether there is any relationship between metrics like variant vulnerability and these sorts of mutations.

We thank Reviewer 1 for engagement on this issue.

- *Fitness values are measured in the presence of a drug, but it is not immediately clear how the drug concentrations are chosen and, more importantly, how the choice of concentration might impact the landscape. The authors may wish to briefly comment on these effects, particularly in cases where the environment involves combinations of drugs. There will be a "new" fitness landscape for each concentration, but to what extent do the qualitative features change--or whatever features drive evolutionary dynamics--change?*

This is another interesting suggestion. We have analyzed a new data set for dihydrofolate reductase mutants that contains a range of drug concentrations of two different antifolate drugs. The general question of how drug concentrations change evolutionary dynamics has been addressed in prior work of ours:

- Ogbunugafor CB, Wylie CS, Diakite I, Weinreich DM, Hartl DL. Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance. PLoS computational biology. 2016 Jan 25;12(1):e1004710.
- Ogbunugafor CB, Eppstein MJ. Competition along trajectories governs adaptation rates towards antimicrobial resistance. Nature ecology & evolution. 2016 Nov 21;1(1):0007.

There are a very large number of environment types that might alter the drug availability or variant vulnerability metrics. In our study, we used an established data set composed of different alleles of a Beta lactamase, with growth rates measured across a number of drug environments. These drug environments consisted of individual drugs at certain concentrations, as outlined in Mira et al. 2015. For our study, we examined those drugs that had a significant impact on growth rate.

For a new analysis of antifolate drugs in 16 alleles of dihydrofolate reductase (*Plasmodium falciparum*), we have examined a breadth of drug concentrations (Supplementary Figure S4). This represents a different sort of environment that one can use to measure the two metrics (variant vulnerability or drug applicability). As we suggest in the manuscript, part of the strength of the metric is precisely that it can incorporate drug dimensions of various kinds.

- *The metrics introduced depend on the ensemble of drugs chosen. To what extent are the chosen drugs representative? Are there cases where nonrepresentative ensembles might be advantageous?*

The authors thank the reviewer for this. The general point has been addressed in our comments above. Further, the general question of how a study of one set of drugs applies to other drugs applies to every study of every drug, as no single study interrogates every sort of drug ensemble. That said, we've explained the anatomy of our metrics, and have outlined

how it can be directly applied to others. There is nothing about the metric itself that has anything to do with a particular drug type – the arithmetic is rather vanilla.

Reviewer #2 (Recommendations For The Authors):

1. Regarding my comment about the different formalisms for epistatic decomposition analysis, a key reference is

Poelwijk FJ, Krishna V, Ranganathan R (2016). The Context-Dependence of Mutations: A Linkage of Formalisms. *PLoS Comput Biol* 12(6): e1004771.

The authors appreciate this, are fans of this work, and have cited it in the revision.

An example where both Fourier and Taylor analyses were carried out and the different interpretations of these formalisms were discussed is

Unraveling the causes of adaptive benefits of synonymous mutations in TEM-1 β lactamase. Mark P. Zwart, Martijn F. Schenk, Sungmin Hwang, Bertha Koopmanschap, Niek de Lange, Lion van de Pol, Tran Thi Thuy Nga, Ivan G. Szendro, Joachim Krug & J. Arjan G. M. de Visser *Heredity* 121:406-421 (2018)

The authors are grateful for these references. While we don't think they are necessary for our new section entitled "Notes on methods used to detect epistasis," we did engage them, and will keep them in mind for other work that more centrally focuses on methods used to detect epistasis. As the author acknowledges, a full treatment of this topic is too large for a single manuscript, let alone a subsection of one study. We have provided a discussion of it, and pointed the readers to longer review articles that explore some of these topics in good detail:

- C. Bank, Epistasis and adaptation on fitness landscapes, *Annual Review of Ecology, Evolution, and Systematics* 53 (1) (2022) 457–479.
- T. B. Sackton, D. L. Hartl, Genotypic context and epistasis in individuals and populations, *Cell* 166 (2) (2016) 279–287.
- J. Diaz-Colunga, A. Skwara, J. C. C. Vila, D. Bajic, Á. Sánchez, Global epistasis and the emergence of ecological function, *BioRxiv*

1. Although the authors label Figure 4 with the term "environmental epistasis", as far as I can see it is only a standard epistasis analysis that is carried out separately for each environment. The analysis of environmental epistasis should instead focus on which aspects of these interactions are different or similar in different environments, for example, by looking at the reranking of fitness values under environmental changes [see Ref.[26] as well as more recent related work, e.g. Gorter et al., *Genetics* 208:307-322 (2018); Das et al., *eLife*9:e55155 (2020)]. To some extent, such an analysis was already performed by Mira et al., but not on the level of epistatic interaction coefficients.

The authors have provided a new analysis of how fitness value rankings have changed across drug environments, often a signature of epistatic effects across environments (Supplementary Figure S1).

We disagree with the idea that our analysis is not a sort of environmental epistasis; we resolve coefficients between loci across different environments. As with every interrogation of G x E effects (G x G x E in our case), what constitutes an "environment" is a messy conversation. We have chosen the route of explaining very clearly what we mean:

“We further explored the interactions across this fitness landscape and panels of drugs in two additional ways. First, we calculated the variant vulnerability for 1-step neighbors, which is the mean variant vulnerability of all alleles one mutational step away from a focal variant. This metric gives information on how the variant vulnerability values are distributed across a fitness landscape. Second, we estimated statistical interaction effects on bacterial growth through LASSO regression. For each drug, we fit a model of relative growth as a function of M69L x E104K x G238S x N276D (i.e., including all interaction terms between the four amino acid substitutions). The effect sizes of the interaction terms from this regularized regression analysis allow us to infer higher-order dynamics for susceptibility. We label this calculation as an analysis of “environmental epistasis.”

As the grammar for these sorts of analyses continues to evolve, the best one can do is be clear about what they mean. We believe that we communicated this directly and transparently.

1. As a general comment, to strengthen the conclusions of the study, it would be good if the authors could include additional data sets in their analysis.

The authors appreciate this comment and have given this point ample treatment. Further, other main conclusions and discussion points are focused on the biology of the system that we examined. Analyzing other data sets may demonstrate the broader reach of the metrics, but it would not alter the strength of our own conclusions (or if they would, Reviewer #2 has not told us how).

1. There are some typos in the units of drug concentrations in Section 2.4 that should be corrected.

The authors truly appreciate this. It is a great catch. We have fixed this in the revised manuscript.

Reviewer #3 (Recommendations For The Authors):

I would suggest demonstrating the concepts for a second drug class, and suggest folA variants and trimethoprim resistance, for which there is existing published data similar to what the authors have used here (e.g. Palmer et al. 2015, <https://doi.org/10.1038/ncomms8385>)

The authors appreciate this insight. As previously described, we have analyzed a data set of folA mutants for the Plasmodium falciparum ortholog of dihydrofolate reductase, and included these results in new supplemental material. Please see the supplementary material.

There are some errors in formatting and presentation that I have annotated in a separate PDF file (https://elifep-files/2023/04/11/00117789/00117789_0_attach_8_30399_convrt.pdf), as the absence of line numbers makes indicating specific things exceedingly difficult.

The authors apologize for the lack of line numbers (an honest oversight), but moreover, are tremendously grateful for this feedback. We have looked at the suggested changes carefully and have addressed many of them. Thank you.

One thing to note: we have included a version of Figure 4 that has effects on the same axes. It appears in the supplementary material (Figure S4).

In closing, the authors would like to thank the editors and three anonymous reviewers for engagement and for helpful comments. We are confident that the revised manuscript

qualifies as a substantive revision, and we are grateful to have had the opportunity to participate.