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

Rafael F. Guerrero[1] Tandin Dorji[2], Ra'Mal M. Harris[3], Matthew D. Shoulders[3], and C. Brandon Ogbunugafor[3−6*]

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

*Corresponding author:
C. Brandon Ogbunugafor, brandon.ogbunu@yale.edu

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 x G x 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).

Keywords: antimicrobial resistance, environmental epistasis, fitness landscapes, druggability, cross-resistance

1. 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 [1–3], reversion [4–6], and higher-order interactions between drugs [7, 8] or mutations [9, 10]. While stepwise, de novo evolution (via the arrival of 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 [11–13] and selection on standing genetic variation [14, 15]. 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 [16–19]. 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 [8, 20, 21].

These gaps in the world of antimicrobial resistance notwithstanding, there is a robust literature in immunobiology—and specifically with respect to the problem of evaluating broadly neutralizing antibodies to counteract the widespread genetic diversity of viral infections—that has addressed analogous concepts. Studies have profiled mutations in influenza that allow them to escape antibodies [22, 23], and reconstructed "binding affinity landscapes" that define the constraints around viral escape [24, 25]. 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 antimicrobial actor, are highly-relevant in our quest to prevent, diagnose, and treat infectious diseases.

In this study, we examine data from protein fitness landscapes across a panel of antimicrobial drugs to develop new metrics for antimicrobial applications. Specifically, we propose novel framings of the idea of druggability, 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 the presence of environmental change [26].

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

2. Materials and Methods

2.1. Notes on the terminology I

The quantities we define as "variant vulnerability" and "drug applicability" may be seen as broadly related to "drugability" (or "druggability" and other related terms), which are sometimes applied to properties of targets [27, 28] and other times to drug candidates themselves ([29], [30, p.5]). These metrics are generally relegated to molecular properties of drugs and targets, and are most generally deployed in medicinal chemistry. To avoid ambiguity and striving to bridge concepts from chemical biology, pharmacology, and evolutionary systems biology, we have avoided using "druggability" in favor of new terminology. In light of this need, 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 1. Also note that our measure of variant vulnerability has resonance 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 [31, 32]. Some studies highlight that cross-resistance can manifest across genetic variants, with mutations conferring resistance to two drugs simultaneously [20]. While modern studies have added nuance to our understanding of how resistance and cross-resistance can evolve [33], 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).

2.2. 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 can sometimes be used to describe how

evolution moves through “protein space,” an idea pioneered by John Maynard Smith [34, 35], that has since been invoked to describe aspects of protein evolution and engineering [36–38]. 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. We also note that the data used in this study comes from a prior one that talked about these data as an "adaptive landscape [39]." Because of that, we resisted the use of "protein space" or other analogies that similarly describe data of this sort.

Term	**Study definition**
Drugability or Druggability	Generally used to describe the ability to treat a drug target with a small molecule-based drug [27, 28]. 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) [29, 30].
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 [26]. 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.

Table 1: *Terms used or introduced in this study and their definition.* For some terms, references are provided for clarity.

2.3. 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 an *Escherichia coli* isolated from a patient named Temoneria in Greece in 1965 [40]. The evolution and subsequent spread of β-lactamases arose from the widespread clinical use of β-lactam antibiotics. As the clinical treatment landscape involves the use of β-lactam drugs of various kinds (old and new) , the population genetic landscape of β-lactamases is diverse [41]. Among clinical populations of gram-negative microorganisms, the

TEM-1 allele is the most frequently detected antibacterial resistance variant [42]. 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 → Leu-69, Glu → Lys-104, Gly → Ser-238 and Asn → Asp-276 [43]. These mutations confer structural, and functional differences that meaningfully influence the clinical use of antibiotics [44]. 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 [1], evolvability [45], and biophysical tradeoffs [46], and have used techniques like deep-mutational scanning to examine the fitness effects of mutations in high-throughput [47]. 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.

2.4. Data

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 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 [39] and the Supplemental Information), we focus our analyses on drug treatments that had a significant negative effect on the growth of wild-type/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 (concentration in μg/ml): amoxicillin 1024 μg/ ml (β-lactam), amoxicillin/clavulanic acid 1024 μg/ l (β-lactam and β-lactamase inhibitor) cefotaxime 0.123 μg/ml (third-generation cephalosporin), cefotetan 0.125 μg/ml (second-generation cephalosporins), cefprozil 128 μg/ml (second-generation cephalosporin), ceftazidime 0.125 μg/l (third-generation cephalosporin), piperacillin and tazobactam 512/8 μg/ml (penicillin and β-lactamase inhibitor). With these drugs/mixtures, we were able to embody chemical diversity in the panel.

2.5. 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.

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 (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.”

3. Results

Here we describe several analyses, emphasizing two novel metrics that 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).

3.1. 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 (MKSN) 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 MKSD allele.

We can rank alleles from highest to lowest for a different interpretation of the variant vulnerability values (Table 2). Notably, the allele that had the highest variant vulnerability (0111, MKSD), and the one with the lowest variant vulnerability (0110, MKSN) differ by a single mutation (N276D). That a single mutation can have this large an impact on variant vulnerability highlights how single mutations have dramatic effects on 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 (MEGN) and TEM-50 (LKSD) 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 a general survivability across drugs. This observation is consistent with prior notions of "cross resistance" [31–33].

3.2. One-step neighbor variant vulnerability

. Our findings regarding large variant vulnerability differences between mutational neighbors like MKSN (0110; the lowest variant vulnerability value of the 16 alleles) and MKSD (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 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 the average variant vulnerability of one-step mutational neighbors for each allele. Figure 3 depicts the variant vulnerability values for the 16 allelic values (from Figure 2), as well as 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 a rugged one, where individual mutations contribute nonlinearly to the susceptibility of a given allelic variant.

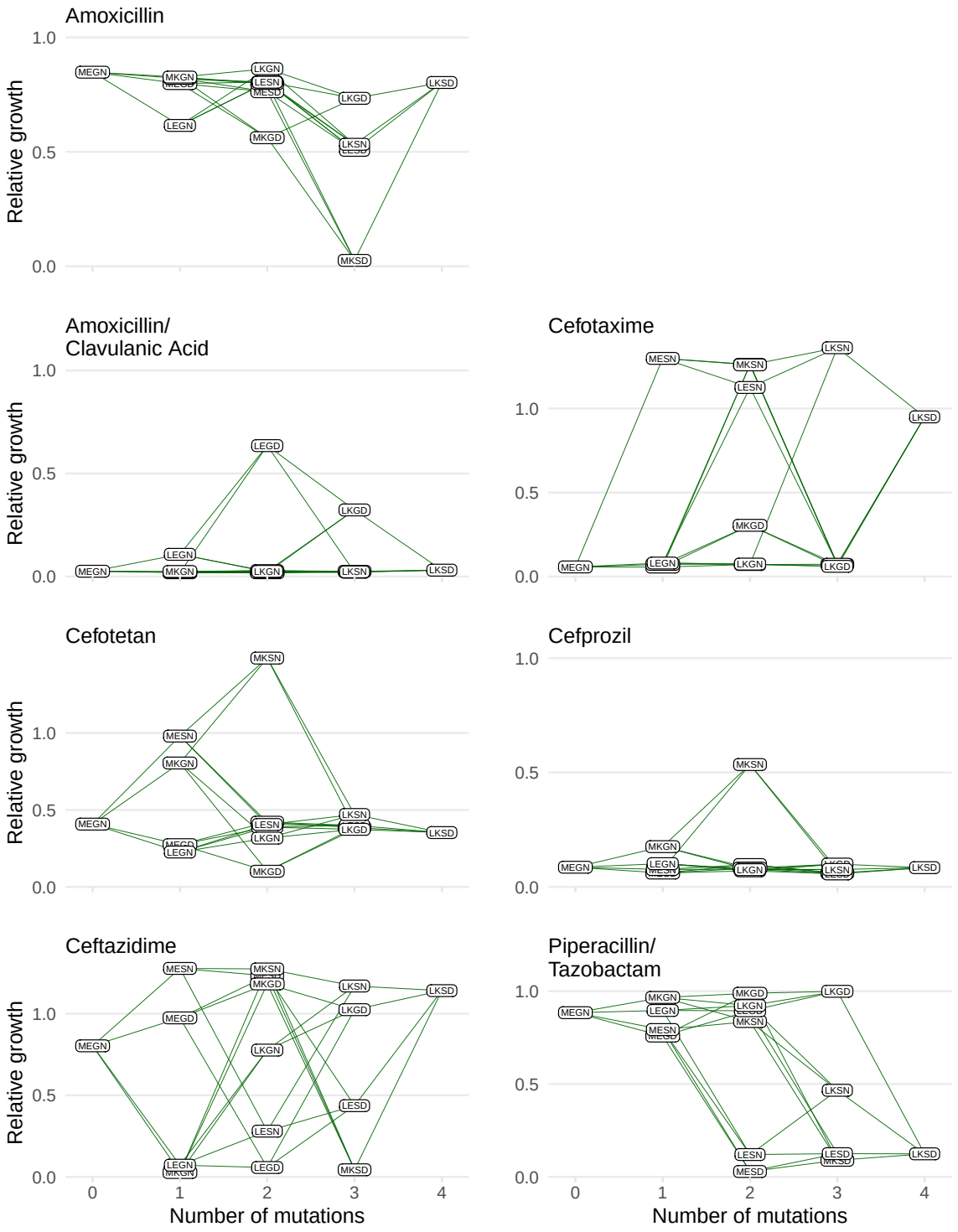

Figure 1: *Hypercube descriptions of the TEM1/TEM-50 fitness landscape examined in this study.* The TEM-50 variant is represented by the quadruple mutant, LKSD, and the wild-type, MEGN, represents TEM-1. The figure highlights the variation in the topography of the various fitness landscapes. Y-axes correspond 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.

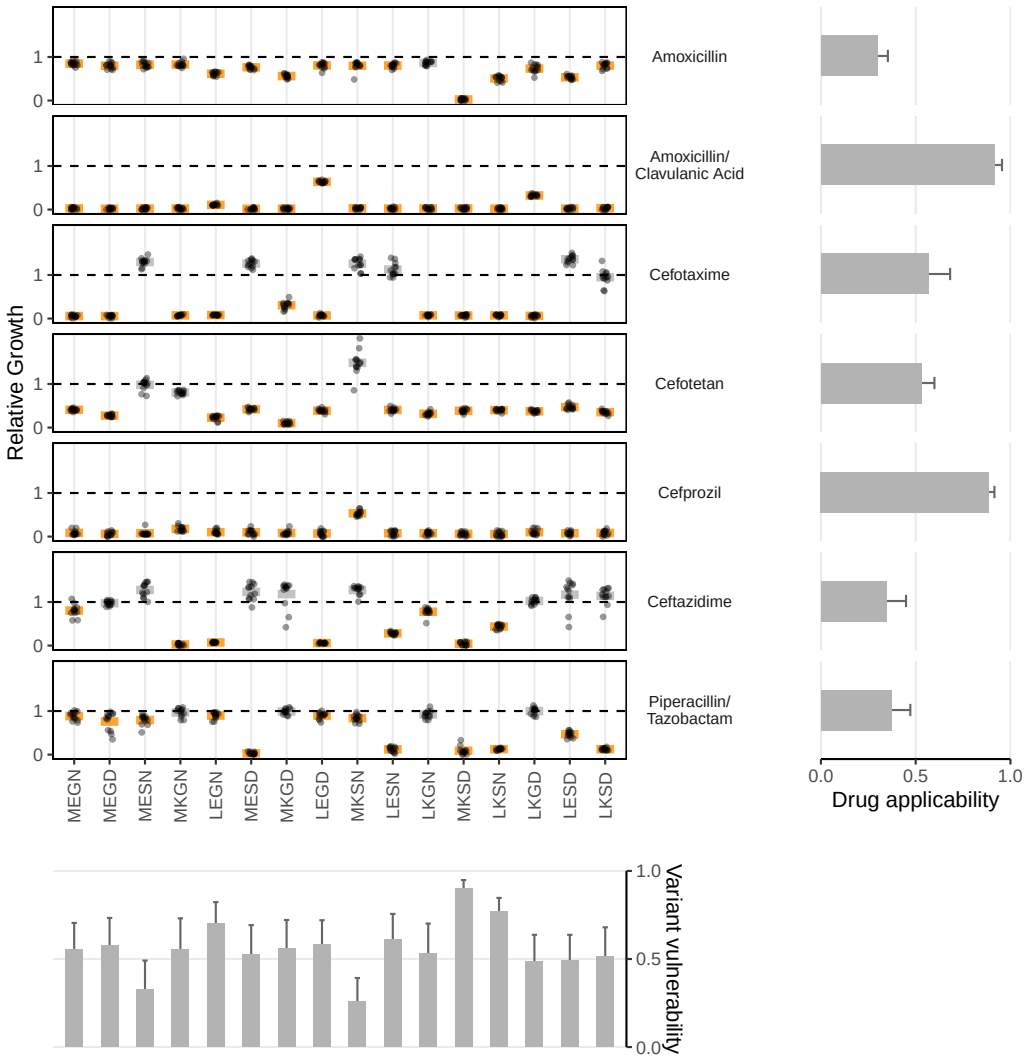

Figure 2: *Data from empirical fitness landscapes can be used to compute two metrics: variant vulnerability (for allelic variants) and drug applicability (for drugs).* Growth of allelic variants, relative to control (wild-type), for seven drug treatments. Bars represent the mean value for each group; 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 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

Table 2: *The rank order of the alleles with respect to variant vulnerability.* An allele with a 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 genotype know as the TEM-50 variant of β-lactamase.

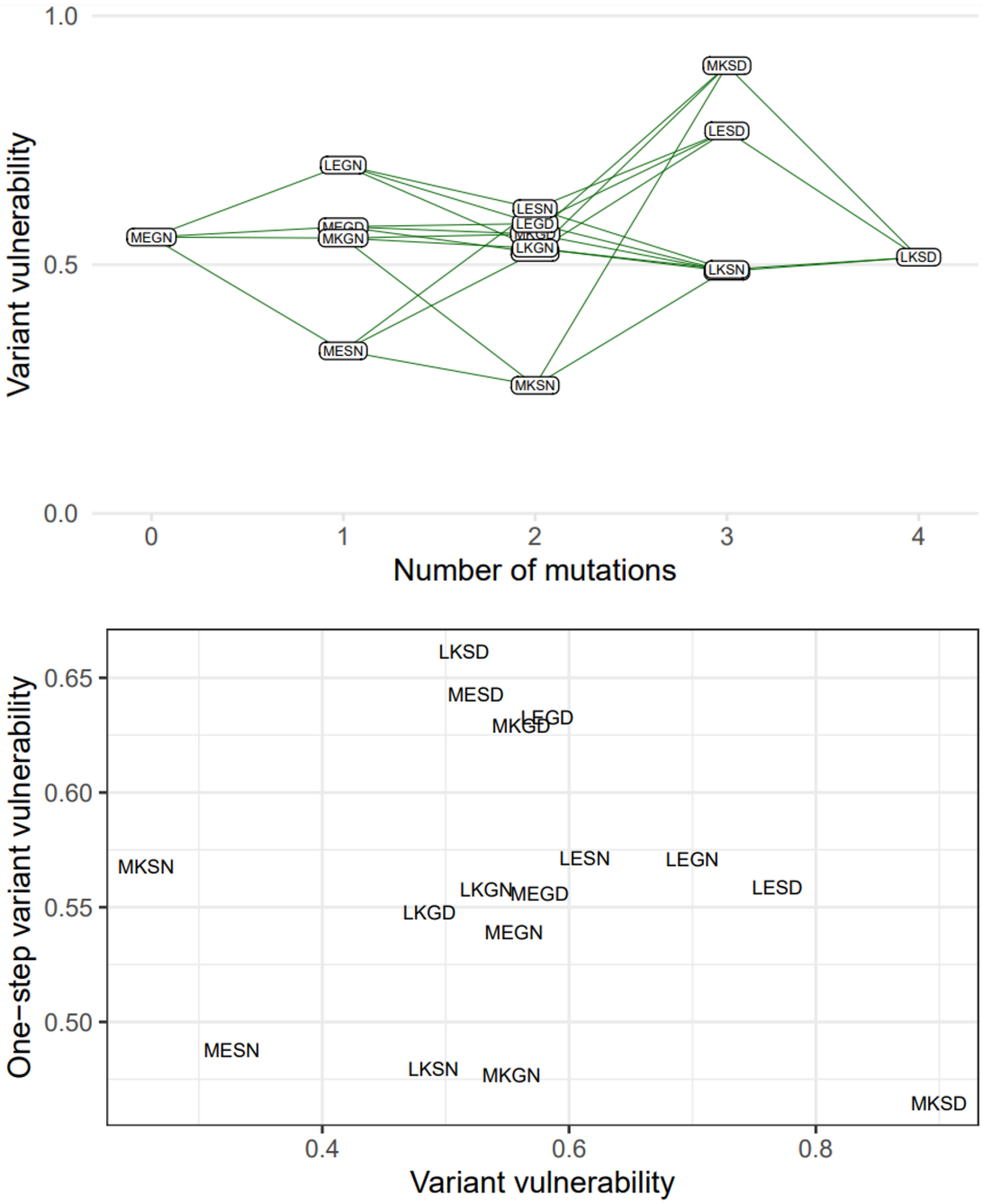

Figure 3: *The variant vulnerability of the 16 allelic variants do not correlate with that of their one-step mutational neighbors.* (A) The variant vulnerability values for the 16-allelic variants as a function of the number of mutations away (hamming distance) from TEM-1 (MEGN). (B) The variant vulnerability 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 $R^2 = 0.01$, $P = 0.69$).

3.3. Drug applicability

A high applicability value indicates that a drug is effective across the suite of genetic variants in a population (Figure 2). In this study, that corresponds to a combinatorial set of 16 TEM-1/TEM-50 mutants. In Table 3, 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.

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

Table 3: *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 based on clinical data.

3.4. 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 to it.

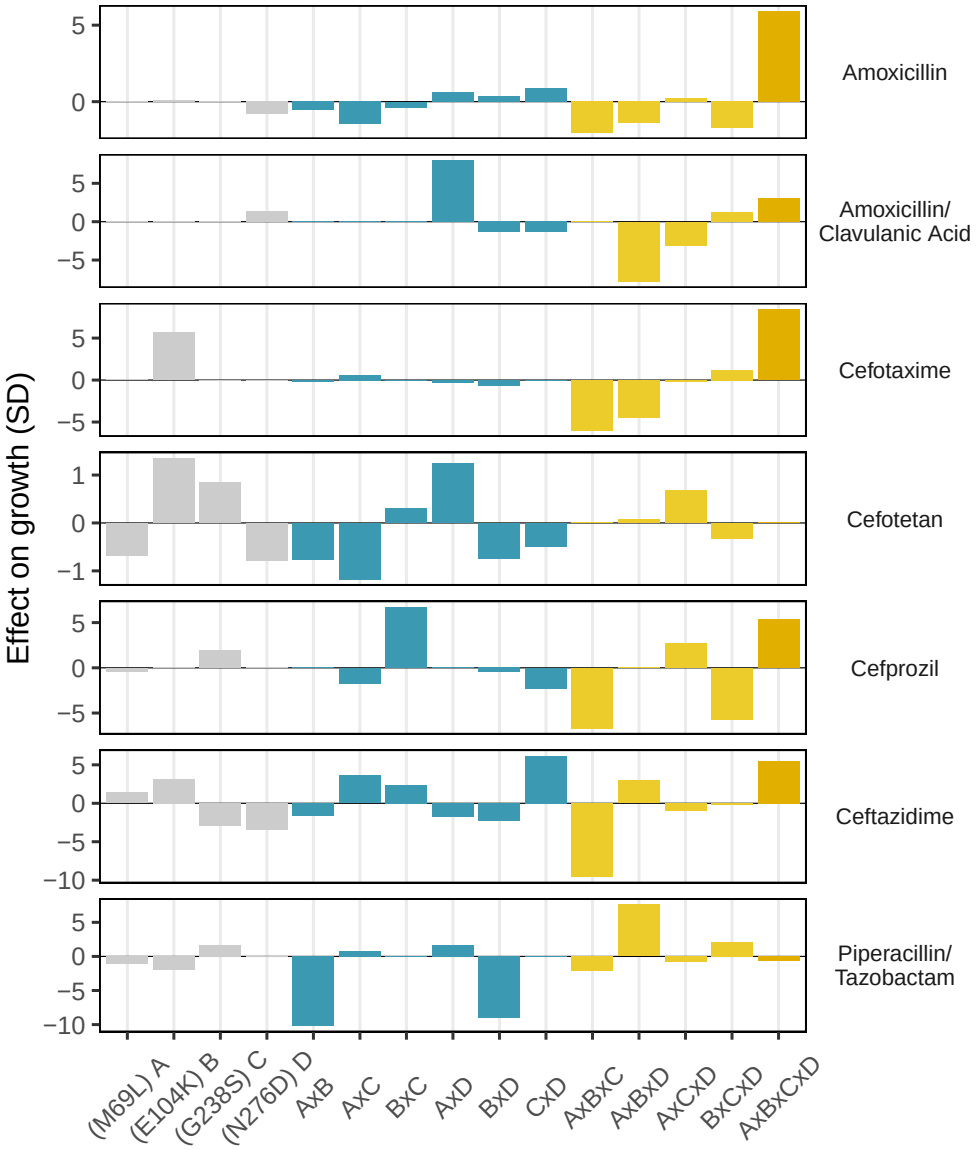

Figure 4: *Environmental epistasis underlies the drug-allele interactions that drive the variant vulnerability and drug applicability.* Effect on growth (in standard deviations of the wild-type control values), estimated by LASSO regression, for individual loci and their interactions. [A] corresponds to M69L, [B] to E104K, [C] to G238S, and [D] to N276D. As in a mutation effect reaction norm (Ogbunugafor 2022), 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 [26].

4. 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 various interactions between individual loci and drug environments that underlie these two metrics.

We focused on a set of 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 epistasis"; see [26]) 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 (MKSN 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 MKSN 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 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). Intuition can over-simplify, with undesirable consequences: not all drug combinations/mixutres 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 [7, 48–50].

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 site 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, G238S on ceftazidime growth rate (Figure 4). With regards to main effects, we note the positive effect of the M69L mutation on 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 ceftazidine (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 [10, 21, 51].

4.1. 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, nor to inform the real-world use of any specific antimicrobial. Therefore, any direct clinical applications of our findings are limited. Rather, 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) [52] 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. Our hope is 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 [7, 8, 48, 50]. Considering how these metrics apply in the context of higher-order combinations of drugs constitutes an area of future inquiry.

4.2. 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 [22, 23], and have characterized virus mutations and epistatic interactions [24, 25] 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, perhaps there are methods or metrics that 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 drugs-microbes 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 [53–55]. Such efforts may include studies of "fitness seascapes" and evolution in fluctuating environments, as has been examined by others [56–58]. 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.

4.3. Conclusions

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 [1, 10] and the existence of G x G x E interactions ("environmental epistasis") as a meaningful characteristic of fitness landscapes in microbial systems [26, 59]. Generally, this study supports recent efforts that employ evolutionary perspectives to understand antimicrobial resistance [5, 7, 8, 16, 17, 48, 60]. Moreover, it demonstrates the utility in combining perspectives—evolutionary genetics, systems biology and medicinal chemistry in this study—to understand complex biological systems. Lastly, this study highlights the increasingly relevant role of environment and context in crafting the interactions that define genotype-phenotype maps.

Acknowledgments

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 R35GM136354 (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 would like to thank K. Kabengele and S. Scarpino for helpful feedback on the manuscript.

Author contributions

Project conception: R.F.G., C.B.O. Collected data: R.F.G., T.D., R.M.H, M.D.S., C.B.O. Analyzed data: R.F.G., T.D., C.B.O. Integrated and interpreted results: R.F.G., M.D.S., C.B.O. Supervision: C.B.O., M.D.S. Writing: R.F.G., M.D.S., C.B.O.

Data availability

Data and code are available at: `https://github.com/OgPlexus/evodruggability`

References

[1] D. M. Weinreich, N. F. Delaney, M. A. DePristo, D. L. Hartl, Darwinian evolution can follow only very few mutational paths to fitter proteins, Science 312 (5770) (2006) 111–114. doi:10.1126/science.1123539.

[2] E. R. Lozovsky, T. Chookajorn, K. M. Brown, M. Imwong, P. J. Shaw, S. Kamchonwongpaisan, D. E. Neafsey, D. M. Weinreich, D. L. Hartl, Stepwise acquisition of pyrimethamine resistance in the malaria parasite, Proceedings of the National Academy of Sciences 106 (29) (2009) 12025–12030. doi:10.1073/pnas.0905922106.

[3] E. Toprak, A. Veres, J.-B. Michel, R. Chait, D. L. Hartl, R. Kishony, Evolutionary paths to antibiotic resistance under dynamically sustained drug selection, Nature Genetics 44 (1) (2011) 101–105. doi:10.1038/ng.1034.

[4] L. Tan, S. Serene, H. X. Chao, J. Gore, Hidden randomness between fitness landscapes limits reverse evolution, Physical Review Letters 106 (19). doi:10.1103/physrevlett.106.198102.

[5] M. Baym, L. K. Stone, R. Kishony, Multidrug evolutionary strategies to reverse antibiotic resistance, Science 351 (6268). doi:10.1126/science.aad3292.

[6] C. B. Ogbunugafor, D. Hartl, A pivot mutation impedes reverse evolution across an adaptive landscape for drug resistance in plasmodium vivax, Malaria Journal 15 (1). doi:10.1186/s12936-016-1090-3.
URL https://doi.org/10.1186/s12936-016-1090-3

[7] J.-B. Michel, P. J. Yeh, R. Chait, R. C. Moellering, R. Kishony, Drug interactions modulate the potential for evolution of resistance, Proceedings of the National Academy of Sciences 105 (39) (2008) 14918–14923. doi:10.1073/pnas.0800944105.

[8] P. J. Yeh, M. J. Hegreness, A. P. Aiden, R. Kishony, Drug interactions and the evolution of antibiotic resistance, Nature Reviews Microbiology 7 (6) (2009) 460–466. doi:10.1038/nrmicro2133.

[9] D. M. Weinreich, Y. Lan, C. S. Wylie, R. B. Heckendorn, Should evolutionary geneticists worry about higher-order epistasis?, Current Opinion in Genetics & Development 23 (6) (2013) 700–707. doi:10.1016/j.gde.2013.10.007.

[10] E. R. Lozovsky, R. F. Daniels, G. D. Heffernan, D. P. Jacobus, D. L. Hartl, Relevance of higher-order epistasis in drug resistance, Molecular Biology and Evolution 38 (1) (2021) 142–151. doi:10.1093/molbev/msaa196.

[11] C. T. Bergstrom, M. Lipsitch, B. R. Levin, Natural selection, infectious transfer and the existence conditions for bacterial plasmids, Genetics 155 (4) (2000) 1505–1519. doi:10.1093/genetics/155.4.1505.

[12] P. E. Turner, E. S. C. P. Williams, C. Okeke, V. S. Cooper, S. Duffy, J. E. Wertz, Antibiotic resistance correlates with transmission in plasmid evolution, Evolution 68 (12) (2014) 3368–3380. doi:10.1111/evo.12537.

[13] W. Loftie-Eaton, K. Bashford, H. Quinn, K. Dong, J. Millstein, S. Hunter, M. K. Thomason, H. Merrikh, J. M. Ponciano, E. M. Top, Compensatory mutations improve general permissiveness to antibiotic resistance plasmids, Nature Ecology & Evolution 1 (9) (2017) 1354–1363. doi:10.1038/s41559-017-0243-2.

[14] P. S. Pennings, Standing genetic variation and the evolution of drug resistance in HIV, PLoS Computational Biology 8 (6) (2012) e1002527. doi:10.1371/journal.pcbi.1002527.

[15] S. Pelleau, E. L. Moss, S. K. Dhingra, B. Volney, J. Casteras, S. J. Gabryszewski, S. K. Volkman, D. F. Wirth, E. Legrand, D. A. Fidock, D. E. Neafsey, L. Musset, 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 (37) (2015) 11672–11677. doi:10.1073/pnas.1507142112.

[16] E. Gjini, K. B. Wood, Price equation captures the role of drug interactions and collateral effects in the evolution of multidrug resistance, eLife 10 (2021) e64851. doi:10.7554/elife.64851. URL https://doi.org/10.7554/elife.64851

[17] S. M. Ardell, S. Kryazhimskiy, The population genetics of collateral resistance and sensitivity, eLife 10 (2021) e73250. doi:10.7554/elife.73250.

[18] C. Herencias, J. Rodríguez-Beltrán, R. León-Sampedro, A. Alonso-del Valle, J. Palkovičová, R. Cantón, Á. San Millán, Collateral sensitivity associated with antibiotic resistance plasmids, Elife 10 (2021) e65130.

[19] D. Nichol, J. Rutter, C. Bryant, A. M. Hujer, S. Lek, M. D. Adams, P. Jeavons, A. R. Anderson, R. A. Bonomo, J. G. Scott, Antibiotic collateral sensitivity is contingent on the repeatability of evolution, Nature communications 10 (1) (2019) 334.

[20] P. Yeh, A. I. Tschumi, R. Kishony, Functional classification of drugs by properties of their pairwise interactions, Nature Genetics 38 (4) (2006) 489–494. doi:10.1038/ng1755.

[21] P. M. Mira, B. Østman, C. Guzman-Cole, S. Sindi, M. Barlow, Adaptive processes change as multiple functions evolve, Antimicrobial Agents and Chemotherapy 65 (4). doi:10.1128/aac.01990-20.

[22] M. B. Doud, S. E. Hensley, J. D. Bloom, Complete mapping of viral escape from neutralizing antibodies, PLOS Pathogens 13 (3) (2017) e1006271. doi:10.1371/journal.ppat.1006271.

[23] M. B. Doud, J. M. Lee, J. D. Bloom, How single mutations affect viral escape from broad and narrow antibodies to h1 influenza hemagglutinin, Nature Communications 9 (1). doi:10.1038/s41467-018-03665-3.

[24] A. M. Phillips, K. R. Lawrence, A. Moulana, T. Dupic, J. Chang, M. S. Johnson, I. Cvijovic, T. Mora, A. M. Walczak, M. M. Desai, Binding affinity landscapes constrain the evolution of broadly neutralizing anti-influenza antibodies, eLife 10 (2021) e71393. doi:10.7554/elife.71393.

[25] A. M. Phillips, D. P. Maurer, C. Brooks, T. Dupic, A. G. Schmidt, M. M. Desai, Hierarchical sequence-affinity landscapes shape the evolution of breadth in an anti-influenza receptor binding site antibody, eLife 12 (2023) e83628. doi:10.7554/elife.83628. URL https://doi.org/10.7554/elife.83628

[26] H. A. Lindsey, J. Gallie, S. Taylor, B. Kerr, Evolutionary rescue from extinction is contingent on a lower rate of environmental change, Nature 494 (7438) (2013) 463–467. doi:10.1038/nature11879.

[27] T. A. Halgren, Identifying and characterizing binding sites and assessing druggability, Journal of Chemical Information and Modeling 49 (2) (2009) 377–389. doi:10.1021/ci800324m.

[28] H. Jubb, T. L. Blundell, D. B. Ascher, Flexibility and small pockets at protein–protein interfaces: New insights into druggability, Progress in Biophysics and Molecular Biology 119 (1) (2015) 2–9. doi:10.1016/j.pbiomolbio.2015.01.009.

[29] T. H. Keller, A. Pichota, Z. Yin, A practical view of 'druggability', Current Opinion in Chemical Biology 10 (4) (2006) 357–361. doi:10.1016/j.cbpa.2006.06.014.

[30] L. Z. Benet, C. M. Hosey, O. Ursu, T. I. Oprea, BDDCS, the rule of 5 and drugability, Advanced Drug Delivery Reviews 101 (2016) 89–98. doi:10.1016/j.addr.2016.05.007.

[31] E. de Souza Cândido, E. de Barros, M. H. Cardoso, O. L. Franco, Bacterial cross-resistance to anti-infective compounds. is it a real problem?, Current Opinion in Pharmacology 48 (2019) 76–81. doi:10.1016/j.coph.2019.05.004.

[32] A. Colclough, J. Corander, S. K. Sheppard, S. C. Bayliss, M. Vos, Patterns of cross-resistance and collateral sensitivity between clinical antibiotics and natural antimicrobials, Evolutionary Applications 12 (5) (2019) 878–887. doi:10.1111/eva.12762.

[33] T. Oz, A. Guvenek, S. Yildiz, E. Karaboga, Y. T. Tamer, N. Mumcuyan, V. B. Ozan, G. H. Senturk, M. Cokol, P. Yeh, E. Toprak, Strength of selection pressure is an important parameter contributing to the complexity of antibiotic resistance evolution, Molecular Biology and Evolution 31 (9) (2014) 2387–2401. doi:10.1093/molbev/msu191.

[34] J. Maynard Smith, Natural selection and the concept of a protein space, Nature 225 (5232) (1970) 563–564.

[35] C. B. Ogbunugafor, A reflection on 50 years of john maynard smith's "protein space", Genetics 214 (4) (2020) 749–754.

[36] P. A. Romero, F. H. Arnold, Exploring protein fitness landscapes by directed evolution, Nature reviews Molecular cell biology 10 (12) (2009) 866–876.

[37] F. H. Arnold, The library of maynard-smith: my search for meaning in the protein universe, Microbe 6 (7) (2011) 316.

[38] F. H. Arnold, Innovation by evolution: bringing new chemistry to life (nobel lecture), Angewandte Chemie International Edition 58 (41) (2019) 14420–14426.

[39] P. M. Mira, J. C. Meza, A. Nandipati, M. Barlow, Adaptive landscapes of resistance genes change as antibiotic concentrations change, Molecular Biology and Evolution 32 (10) (2015) 2707–2715. doi:10.1093/molbev/msv146.

[40] C. D. Steward, D. Wallace, S. K. Hubert, R. Lawton, S. K. Fridkin, R. P. Gaynes, J. E. McGowan, F. C. Tenover, Ability of laboratories to detect emerging antimicrobial resistance in nosocomial pathogens: a survey of project ICARE laboratories, Diagnostic Microbiology and Infectious Disease 38 (1) (2000) 59–67. doi:10.1016/s0732-8893(00)00161-9.

[41] J. Blazquez, M.-I. Morosini, M.-C. Negri, F. Baquero, Selection of naturally occurring extended-spectrum TEM β-lactamase variants by fluctuating β-lactam pressure, Antimicrobial Agents and Chemotherapy 44 (8) (2000) 2182–2184. doi:10.1128/aac.44.8.2182-2184.2000.

[42] J. E. Mroczkowska, M. Barlow, Fitness trade-offs in blaTEM evolution, Antimicrobial Agents and Chemotherapy 52 (7) (2008) 2340–2345. doi:10.1128/aac.00018-08.

[43] D. Sirot, C. Recule, E. B. Chaibi, L. Bret, J. Croize, C. Chanal-Claris, R. Labia, J. Sirot, 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 (6) (1997) 1322–1325. doi:10.1128/aac.41.6.1322.

[44] A. Pimenta, R. Fernandes, I. Moreira, Evolution of drug resistance: Insight on TEM β-lactamases structure and activity and β-lactam antibiotics, Mini-Reviews in Medicinal Chemistry 14 (2) (2014) 111–122. doi:10.2174/1389557514666140123145809.

[45] M. A. Stiffler, D. R. Hekstra, R. Ranganathan, Evolvability as a function of purifying selection in TEM-1 β-lactamase, Cell 160 (5) (2015) 882–892. doi:10.1016/j.cell.2015.01.035.

[46] J. Knies, F. Cai, D. M. Weinreich, Enzyme efficiency but not thermostability drives cefotaxime resistance evolution in TEM-1 β-lactamase, Molecular Biology and Evolution 34 (2017) 1040—-1054. doi:10.1093/molbev/msx053.

[47] N. J. Rollins, K. P. Brock, F. J. Poelwijk, M. A. Stiffler, N. P. Gauthier, C. Sander, D. S. Marks, Inferring protein 3d structure from deep mutation scans, Nature Genetics 51 (7) (2019) 1170–1176. doi:10.1038/s41588-019-0432-9.

[48] C. Beppler, E. Tekin, Z. Mao, C. White, C. McDiarmid, E. Vargas, J. H. Miller, V. M. Savage, P. J. Yeh, Uncovering emergent interactions in three-way combinations of stressors, Journal of The Royal Society Interface 13 (125) (2016) 20160800. doi:10.1098/rsif.2016.0800.

[49] N. Singh, P. J. Yeh, Suppressive drug combinations and their potential to combat antibiotic resistance, The Journal of Antibiotics 70 (11) (2017) 1033–1042. doi:10.1038/ja.2017.102.

[50] E. Tekin, V. M. Savage, P. J. Yeh, Measuring higher-order drug interactions: A review of recent approaches, Current Opinion in Systems Biology 4 (2017) 16–23. doi:10.1016/j.coisb.2017.05.015.

[51] C. B. Ogbunugafor, C. S. Wylie, I. Diakite, D. M. Weinreich, D. L. Hartl, Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance, PLOS Computational Biology 12 (1) (2016) e1004710. doi:10.1371/journal.pcbi.1004710.

[52] D. M. Fowler, S. Fields, Deep mutational scanning: a new style of protein science, Nature Methods 11 (8) (2014) 801–807. doi:10.1038/nmeth.3027.

[53] I. Cvijović, B. H. Good, E. R. Jerison, M. M. Desai, Fate of a mutation in a fluctuating environment, Proceedings of the National Academy of Sciences 112 (36) (2015) E5021–E5028.

[54] S. Raghavan, P. S. Winter, A. W. Navia, H. L. Williams, A. DenAdel, K. E. Lowder, J. Galvez-Reyes, R. L. Kalekar, N. Mulugeta, K. S. Kapner, et al., Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer, Cell 184 (25) (2021) 6119–6137.

[55] G. Kinsler, K. Geiler-Samerotte, D. A. Petrov, Fitness variation across subtle environmental perturbations reveals local modularity and global pleiotropy of adaptation, Elife 9 (2020) e61271.

[56] V. Mustonen, M. Lässig, From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation, Trends in Genetics 25 (3) (2009) 111–119. doi:10.1016/j.tig.2009.01.002.

[57] I. Cvijović, B. H. Good, E. R. Jerison, M. M. Desai, Fate of a mutation in a fluctuating environment, Proceedings of the National Academy of Sciences 112 (36). doi:10.1073/pnas.1505406112.

[58] E. S. King, J. Maltas, D. T. Weaver, R. Barker-Clarke, J. Pelesko, E. Dolson, J. G. Scott, Fitness seascapes are necessary for realistic modeling of the evolutionary response to drug therapy, bioRxivdoi:10.1101/2022.06.10.495696.

[59] C. B. Ogbunugafor, The mutation effect reaction norm (mu-rn) highlights environmentally dependent mutation effects and epistatic interactions, Evolution 76 (S1) (2022) 37–48. doi:10.1111/evo.14428.

[60] D. I. S. Rosenbloom, A. L. Hill, S. A. Rabi, R. F. Siliciano, M. A. Nowak, Antiretroviral dynamics determines HIV evolution and predicts therapy outcome, Nature Medicine 18 (9) (2012) 1378–1385. doi:10.1038/nm.2892.