The mutation effect reaction norm (MERN) highlights environmentally dependent mutation effects and genetic interactions

C. Brandon Ogbunugafor
Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520 USA
brandon.ogbunu@yale.edu

***Abstract* – Measuring the fitness effects of mutations and epistasis remain central yet provocative concepts in evolutionary and population genetics. In addition to the baseline complexity that arises from the notion that genetic information can interact in a nonlinear way, recent studies have revealed that interactions can change as a function of environmental context. Here I propose the fusion of measurements of the effect of mutations and physiological epistasis with the reaction norm, a central abstraction used to depict genotype by environment interactions. In doing so, I formalize the notion of a "mutation effect reaction norm" (MERN) as an instrument through which one can analyze or depict the phenotypic consequences of interactions between mutations across environmental contexts. I demonstrate its utility through a discussion of the signature of mutations that undermine reverse evolution of antimicrobial resistance. In closing, I argue that the mutation effect reaction norm may help us resolve the dynamism of evolution across fitness seascapes through specific insight into how mutation effects and interactions are modulated by environmental context.**

I. INTRODUCTION

While breakthroughs in the various subfields of genomics continue to improve our understanding of the architecture of complex traits and phylogenetic relationships between organisms, recent findings have refined our lenses and created new questions. For example, modern perspectives from evolutionary genetics are increasingly driven by notions that complex traits are the product of interactions between many individual SNPs, which inform applications of the infinitesimal model and other iterations of how we study the relationship between genotype and phenotype [1, 2].

For decades, interactions between mutations have played a meaningful role in related discussions surrounding conceptual challenges in population and evolutionary genetics. The term "physiological epistasis" has been used to describe "any situation in which the genotype at one locus modifies the phenotypic expression of the genotype at another locus [3]." Within this definition, an expansive literature exists that has examined epistasis in adaptive landscapes [4, 5, 6]. with respect to protein biophysics [7, 8, 9], in terms of genomic architecture [10, 11], and many other arenas.

The broader notion that mutations may interact with other parcels of genetic information in a spurious or nonlinear fashion casts a shadow over much of modern genetics [12, 13, 14]. In some sense, nonlinear interactions between mutations might challenge simple models of genotype-phenotype mapping, plausibly contributing to phenomenon like phantom heritability [15, 16, 17]. Relatively underexplored in conversations about how epistasis may craft complex phenotypes are theoretical treatments of how environmental gradients may influence the interactions between mutations or SNPs of interest that craft complex traits.

Conveniently, an abstraction exists in the evolutionary biology and ecology canons—the reaction norm (also known as the "norm of reaction")—to describe how the environment shapes the performance (phenotype) of genotypes [18]. The reaction norm is widely applied in evolutionary biology in quantitative genetics [19, 20], in discussions of phenotypic plasticity [21, 22], and other subtopics.

While several studies have examined how environments can tune nonlinear interactions between mutations [23, 24, 25, 26, 27], there have been no formal attempts to integrate how the environment tunes nonlinear interactions between mutations. In this study, I introduce the "mutation effect reaction norm," an abstraction that combines the reaction norm with calculations of mutation effect and physiological epistasis. It demonstrates how mutation effects and epistatic coefficients can change appreciably across environmental contexts of various kinds. To demonstrate its utility, I explore data sets corresponding to a collection of alleles carrying mutations associated with antimicrobial drug resistance. I analyze these data using the mutation effect reaction norm framework and diagnose the signature of the effect of a compensatory ratchet mutation whose effect is specific to environment.

Summarizing, I discuss how this abstraction is relevant for many applications that rely on an understanding of how the effects of individual mutations are influenced by environments, including genetic modification and public health. More broadly, I use the concept to emphasize the importance of more detailed biographies of mutation interactions in our

present and future attempts to rigorously capture the shape of adaptive evolution in light of the mutational steps that define it.

II. METHODS

A. Data Sets

This study utilized two dates sets, each a combinatorially complete adaptive landscapes where a suite of mutations of interest within a locus of interest (a protein in this case) are engineered in all possible combinations. While this data structure is not necessary to measure interactions and epistasic effects, it facilitates the use of certain transparent, established algebraic formulations. The data sets used in this study are as follows:

Traits of interest: growth rate in pyrimethamine and cycloguanil, antifolate drugs used to treat malaria [28]. We utilized a suite of four mutations (N51I, C59R, S108N, I164L) in *Plasmodium falciparum* (a cause of malaria) dihydrofolate reductase (DHFR, an essential enzyme), constructed in combination ($2^4 = 16$ alleles) [29]. All 16 alleles have growth rate values across gradient of drug concentrations (10^{-3} μM to 10^5). These data arise from a set utilized in previous studies that examined the evolution of resistance to antimalarials [30, 29].

B. A note on methods to measure interactions

Many methods exist for measuring the strength of interactions between mutations in empirical datasets. Questions surrounding which methods one should use are not unlike many statistical questions surrounding how to disentangle nonlinear effects in complex systems (biological or other): the shape, scope, and size of the data dictate which analyses are most appropriate. Furthermore, epistasis might be described as an idea whose definitions and implications are based on how it is measured. Available methods consider how noise can conflate the measurement of resistance [31, 32], methods that can responsibly consider non-binary encoding or gaps in data [33], and those that consider the limits of regression methods [34, 35] amongst many others. These methods all have strengths and can be considered more ideal for certain empirical data sets. However, a rigorous treatment of methods used to measure epistasis is beyond the scope of this manuscript.

An example method: The Walsh-Hadamard transform. In this study, we utilize the Walsh-Hadamard transform, which computes a coefficient corresponding to the magnitude and sign of an interaction between mutations with respect to a phenotype of interest. It was pioneered for use in the study of epistasis in a 2013 study that both provided a primer for the calculation and analyzed several combinatorially complete data sets [4]. It has since been further applied to study of higher-order epistasis across a larger sampling of empirical data sets [36].

The Walsh-Hadamard transform implements phenotypic measurements into a vector, then a Hadamard matrix, which is scaled by a diagonal matrix. The result is a set of coefficients that measure the degree to which the genotype-phenotype map, perhaps encoded as an adaptive landscape, is linear, or second order, third, and so forth. For more rigorous discussions of the method, we encourage readers to engage several published manuscripts—especially Weinreich et. al. (2013) [4] and Poelwijk et al. (2016) [37] — each of which explore the methods and their related issues in greater detail. For clarity, I will describe selected aspects of the method in this manuscript.

As the data utilized by the Walsh-Hadamard transform are combinatorially complete, we can represent the presence or absence of a given mutation by a 0 or 1 at a given locus.

For example, we can represent a wildtype variant of a gene as 0000. In this one scenario, the mutations at each of four sites (e.g. the four mutations corresponding to antifolate resistance *Plasmodium falciparum* dihydrofolate reductase) [38, 39], N51I, C59R, S108N, and I164L. The quadruple mutant, IRNL, would be encoded as 1111 in this scenario.

The full data set consists of a vector of phenotypic values (growth rate in the presence of two different antifolates) for all possible combinations of mutations (for 16 alleles in total):

NCSI, NCSL, NCNL, NCNL, NRSI, NRNI, NRNL, ICSI, ICSL, ICNI, ICNL, IRSI, IRSL, IRNI, IRNL

These can be depicted in binary as the following:

0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111

This vector of phenotypes corresponding to each of those variants can be multiplied by a (16 × 16) square matrix, which is the product of a diagonal matrix V and a Hadamard matrix H. These are defined recursively:

$$V_{n+1} = \begin{pmatrix} \frac{1}{2}V_n & 0 \\ 0 & -V_n \end{pmatrix}, V_0 = 1 \quad (1)$$

$$H_{n+1} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}, H_0 = \quad (2)$$

n is the number of loci (n = 4 in this *Plasmodium falciparum* DHFR setting corresponding to resistance mutations). This multiplication gives the following:

$$\gamma = VH_\kappa \quad (3)$$

Where V and H are the matrices described in equations 1 and 2 above and γ is the Walsh coefficient, the measure of the interaction between mutations. Using this formulation, we compute γ values for every possible interaction between bits in each string.

In addition to the aforementioned references where this approach were described [4, 37], this method is also detailed in the Supporting Information, which also contains a spreadsheet that outlines the calculation and provides a means for even inexperienced users to calculate interaction coefficients.

Having outlined the method used to calculate the Walsh-Hadamard coefficient, I must be clear about the interpretation. The Walsh coefficient corresponds to the average effect of a given mutation effect (first order, pairwise, etc.) across all genetic backgrounds. Negative values for an effect suggest that the average effect is negative in a given setting, positive if it has a beneficial effect on a phenotype (e.g., antibiotic resistance).

As stated, this iteration of the Walsh-Hadamard transform relies on the existence of combinatorial data sets, where mutations are constructed in all possible combinations. Another limitation of the Walsh-Hadamard transform is that it can only accommodate two variants (amino acid substitutions in this case) per site. For example, if we wanted to measure the higher-order interactions between 5 mutations within a locus, we would need $2^L = 32$ individual measurements, with L corresponding to the number of different mutations whose effects we were interested in disentangling. Another established limitation of the method is that it doesn't formally incorporate experimental noise. Consequently, its resultant measurement is more consistent with an average of the effect between mutations at the representative loci. Though these limitations reveal that the Walsh-Hadamard transform might be boutique, it still applies to many real-world data sets, and provides insight into the particulars of how mutations interact.

C. Calculations of higher-order epistasis

Previous studies have inquired about how higher-order epistasis manifests across empirical fitness landscapes [4, 36, 25]. "Order" corresponds to the number of actors involved in an interaction. "First-order" would correspond to the interaction between single mutations and genetic backgrounds, second-order or "pairwise" interactions would apply to pairs of mutations, and so forth.

We can calculate higher-order epistasis using several minor modifications to the Walsh-Hadamard transform method outlined above, very similar to how prior studies carried this out [4].

For example, in a combinatorially complete data set comprising 16 alleles, we can depict the interactions between individual loci and genetic background with binary representation, with each 0 or 1 representing a mutation:

0000: zeroth order interaction

0001, 0010, 0100, 1000 are first-order interactions. For example, this translates to the average effect of the N51I mutation across all possible genetic backgrounds (composed of combinations of the other three loci), those between the C59R mutation and all possible genetic backgrounds, between S108N mutation and all possible genetic backgrounds, and between I164L and the other loci.

Relatedly: 0011 0101, 0110, 1001, 1100 are second-order or pairwise interactions; 0111, 1010, 1011, 1101, 1110 are third-order interactions; and 1111 is a fourth-order interaction, the interaction between the four mutations that constitute the quadruple mutant, IRNL. For even more clarity, one can replace the 0s with asterisks (*) to emphasize that those sites represent mutation interaction effects across all possible combinations of those sites. For example, the pairwise effect coefficient corresponding to "0110" truly means the average effect of the C59R and S018N mutations, across all other backgrounds. We can depict this effect as "*11*".

One can average the interaction coefficients within an order to facilitate comparisons between orders (e.g., are third order effects stronger than pairwise effects across environments?).

Though the data in this study are not normalized, it is often prudent to compute a normalized version of the epistatic coefficients, and then take the absolute value. The normalization standardizes the value relative to those in the set, which is useful to contextualize effects relative to the entire set. For a given epistatic coefficient γ, we define the normalized epistatic coefficient E, as in prior studies of in silico adaptive landscapes [40]:

$$E_i = \frac{|\gamma_i|}{\sum_j |\gamma_j|} \tag{4}$$

In this study, we only use the absolute values of the epistatic coefficients for all analyses, without normalization.

While the absolute mean masks negative values, this calculation offers a window into the magnitude of nonlinearity in a genotype-phenotype map, across environments. One can take the average value of the epistatic coefficients across orders (average all pairwise, three-way interactions, etc.), which facilitates comparisons between orders (in the case of the mutation effect reaction norm, across environments).

We label these order-averaged effects with the term "absolute mean." This provides mean values for each order, which calculates the overall contribution of, for example, 1st order effects and higher-order (3rd order, 4th order, etc.) effects.

And we can examine how the order of effects changes across environmental gradients, representing a kind of mutation effect reaction norm for higher-order epistasis.

III. RESULTS

A. The reaction norm demonstrates the growth rate of alleles across drug environments

Figure 1 depicts reaction norms for a combinatorial set of 16 alleles. The data demonstrate growth rates as a function of concentrations of two different drugs: pyrimethamine (1A, B) and cycloguanil (1C, D).

The dynamism of the reaction norms is further encapsulated by depictions of the respective rank orders of alleles in the presence of the two antimalarial drugs, pyrimethamine and cycloguanil (FIG. 1B, D). That the rank order of alleles changes rapidly at some concentrations is a signature of epistasis present in the system, as rank order reflects nonlinear interactions between the mutations that compose the allelic background [41]. Specifically, note the rapid rank order changes occurring at certain drug concentrations (roughly from $10^{-2}\mu M$ to 10^{2} μM, on both pyrimethamine and cycloguanil). The data that compose these reaction norms were previously examined with respect to how drug environments create different evolutionary dynamics (across drug type and concentration [30, 29].

B. Exploring how the phenotypic effects of mutations changes according to environment

Mutation effect reaction norms display interactions between mutations for two similar antifolate drugs (pyrimethamine and cycloguanil) along continuous environmental dimensions (FIG. 2A and 2C). Using this approach, we can observe how mutation effects are tuned differently across environments that are at least somewhat similar. That is, pyrimethamine and cycloguanil are both antifolate drugs used to treat malaria infections and have slightly different patterns of mutation effects across contexts, all coefficients converging towards [0] as drug concentrations increase. That is, at a high enough drug concentration, no alleles grow (see the reaction norm results, Figure 1), and so the mutation interactions that compose the alleles in the reaction norm have small effects at the highest drug concentrations.

The subtle differences between pyrimethamine and cycloguanil are related to differences in the topographies of their respective fitness landscapes with respect to the suite of alleles associated with varying antifolate resistance. These features were captured in the reaction norm, and in prior studies of adaptive landscape by environment interactions in these data sets [29]. For clarity, the average effects of single mutations are emphasized in the Figures 2A and 2C (thicker lines), as they are the coefficients whose interpretation are the most intuitive: the single mutation effect lines provide an average description of how impactful each of the four individual loci across the drug environments. The values of the remainder of the interactions arise from the formula outlined in the Walsh-Hadamard transform calculation and are more challenging to summarize verbally, as their value is computed from the diagonal matrix outlined in the Methods and Supporting Information.

To offer a clearer understanding of how environments shape higher-order interactions, I provide a mutation effect reaction norm corresponding to the normalized absolute mean values of mutation effects, organized by order (Fig. 2B and 2D). These represent magnitude differences between orders of effects and communicate the overall presence of higher-order interactions across environmental gradients (drug type and concentration in this case).

C. The mutation effect reaction norm highlights the specific signature of compensatory ratchet mutations

Of particular interest are the effects of S108N in *P. falciparum* dihydrofolate reductase (Fig. 2A, 2C). The effects of an orthologous mutation was described in a study of reverse evolution of antifolate resistance in *Plasmodium vivax* [42]. Note that in both pyrimethamine and cycloguanil, the mutation effect has a similar pattern: a negative effect at low drug concentrations, with a sign change (from negative to positive) as drug concentrations increase towards 100 μg/ml, after which it then decreases as the drug concentration increases. This reflects a single mutation whose interaction effect changes from negative at low drug concentrations to positive effects at higher drug concentrations.

This "flip" of the sign of a mutation effect is reflective of a mutation that could be described as compensatory. That is, the mutation corresponding to S108N is conditionally beneficial, offering positive epistatic interaction in drug environments with higher concentrations (Fig. 2A), where many alleles are growing poorly. These mutations restore growth in genetic backgrounds that are growing poorly. These mutations also serve as ratchets that undermine the reversal of evolution (from the IRNL quadruple mutant towards the NCSI wildtype in this setting).

Fig. 3 further describes how mutation interactions involving the S018N mutation can influence evolution. In 3A, we observe that alleles that contain S018N have a significantly higher growth rate, across all drug concentrations of both pyrimethamine (Kruskal-Wallis: 5A, pyrimethamine, p = 0.0002;). Notably, the S108N mutation is an orthologue of a mutation, S117N, that has been described as a "pivot" mutation, that both dictates the direction of adaptive evolution, and precludes reversal [42].

Fig. 3B is a hypergraph summary of the predicted evolutionary trajectories in pyrimethamine. Predictions were made from the rank orders of alleles outlined in Figure 1. That is, starting from NCSI, evolution follows a path of increasing growth rate, a proxy for reproductive fitness in this setting. Figure 3B depicts "forward" evolution starting from the wild type (NCSI) allele evolving at 10^{6} μM Pyrimethamine, as summarized in previ-

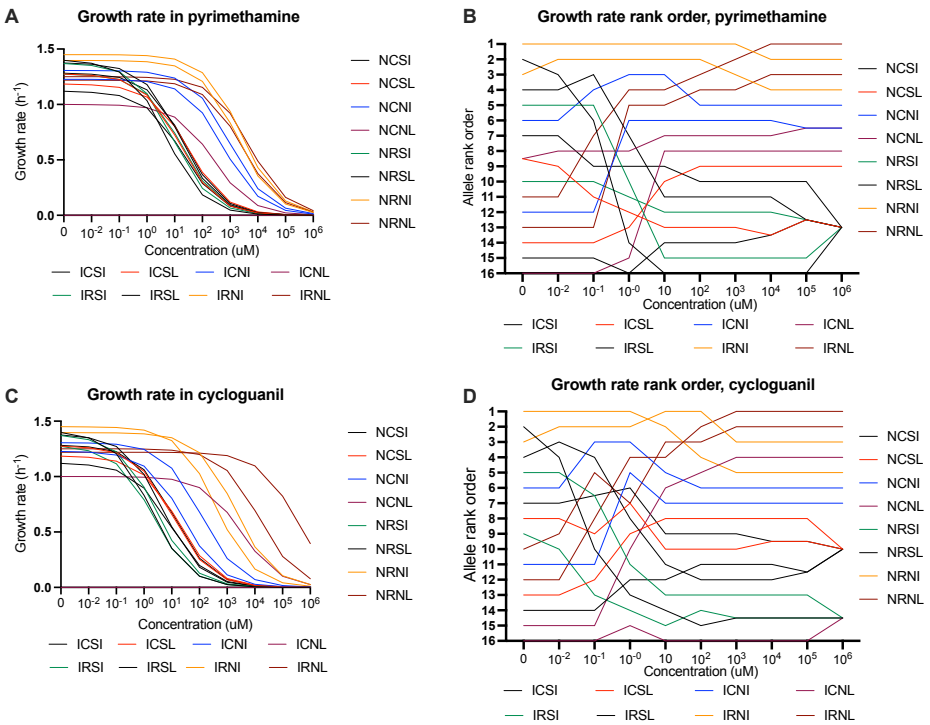

Figure 1: Reaction norms. The reaction norm for growth rates corresponding to *Plasmodium falciparum* carrying 16 different alleles of dihydrofolate reductase associated with resistance to antifolates, across drug environments. (A) Reaction norm for growth rate of alleles across drug concentrations of pyrimethamine. (B) Rank order of alleles across drug concentrations in pyrimethamine. (C) Reaction norm for growth rate of alleles across drug concentrations of cycloguanil, and (D) Rank order of alleles across drug concentrations in cycloguanil. Despite the similarities between pyrimethamine and cycloguanil, the two have differences in the shape of their reaction norm and rank orders of the same set of alelles.

ous studies [29]. In addition, Figure 3B shows the predicted "reverse" evolution trajectory, when the IRNL quadruple mutant evolves in a drugless environment. The preferred trajectories for both forward and reverse evolution in pyrimethamine—Forward: NCSI → NCNI → NRNI →IRNI→IRNL; Reverse IRNL→IRNI all steps in the preferred pathways—forward and reverse contain the S108N mutation. The mutation plays a central role in dictating the direction of evolution in both high drug (10^6 μM pyrimethamine μM) and the drugless environment.

IV. DISCUSSION

In this study, I introduce an abstraction called the mutation effect reaction norm, which depicts how interactions between mutations and other genetic mutations/genetic backgrounds can change across environmental contexts. It represents a fusion of the norm of reaction, a representation of genotype-phenotype maps across environments, with calculations of mutation effects and epistasis. Rather than simply representing this information, I have attempted to emphasize how it is useful in understanding the dynamics of adaptation and reversal. To demonstrate the utility of the abstraction, I use empirical fitness landscapes as test data sets. I apply previously developed mathematical methods introduced to measure higher-order epistasis on binary-encoded combinatorially complete data sets, across categorical and both continuous environments. The purpose of this examination is to consolidate these perspectives into a single abstraction that simplifies the notion that epistatic effects can be depicted across environments.

A. *Tracking mutation effects across environmental contexts*

With this mutation effect reaction norm, several subquestions can be examined, such as how multi-dimensional environments tune the phenotypic effects of mutations. We observe this through comparing the topography of the mutation effect reaction norms for a suite of SNPs associated with resistance to pyrimethamine and cycloguanil (one dimension: antifolate compounds used to treat malaria that are similar in structure), across a range of drug concentrations (second dimension). In this case, we observe how and why the fitness landscapes (composed of different haplotypes) differ for the two landscapes. Though the two drugs have a similar

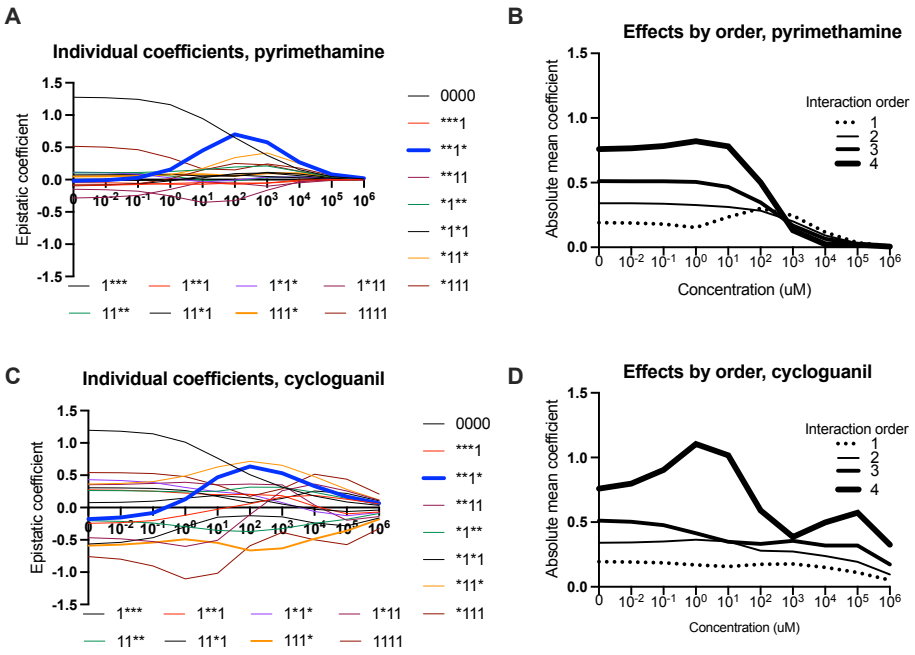

Figure 2: Mutation effect reaction norms (MERN). The mutation effect reaction norm corresponding to the strength of interactions. (A) The MERN depicts interactions between individual loci associated with resistance within the *Plasmodium falciparum* dihydrofolate reductase across a breadth of concentrations of pyrimethamine (x-axis, removed for clarity). The binary notation corresponds to interactions between one of four individual loci within the *P. falciparum* dihydrofolate reductase. (B) MERN corresponding to a transformation of the data in (A), whereby the absolute mean of values of all the effects of a certain order are combined, which provides a perspective on how higher-order epistasis varies across environmental context (see Methods for details). (C) A MERN for individual loci interactions effects on growth rate across cycloguanil concentrations (x-axis, removed for clarity), as discussed in A.(D) A MERN depiction of higher-order epistasis for resistance across a set of cycloguanil environments. Note that both A and C, the single mutation effects corresponding to **1* is emphasized emphasized with a thicker line. This effect, corresponding to the average effect of the S108N mutation across genetic backgrounds and drug environments, is of special interest, as discussed in the main text

mechanism of action and are similar in size (pyrimethamine molecular weight = 248.7; cycloguanil molecular weight = 251.7), these differences have consequences for patterns of resistance in nature [28], specifically because interactions between mutations in the *P. falciparum* dihydrofolate reductase protein backbone differ as a function of both drug (pyrimethamine or cycloguanil) and drug concentration.

Moving past the specific case of antimicrobial resistance, these findings speak to concepts that are of central importance in evolutionary theory. For example, the mutation effect reaction norm may inform models of how adaptive evolution occurs fluctuating environments, as context-specific interactions can create opportunity and constraint in different settings. These ideas that are especially germane to modern efforts to transcend the study of static fitness landscapes, toward examination of fitness seascapes, defined by dynamic and fluctuating environments [43]

B. The mutation effect reaction and the dynamics of adaptation and reversal

The study's examination of S108N's interactions serves as an example of how tracking the mutation effects allows one to identify the (i) quantitative range of environments in which a mutation has effects, (ii) the genotypic context in which a given mutation is compensatory and (iii) the degree to which the effect is compensatory or not. I further argue that the story that we've revealed about S108N likely applies to many compensatory ratchet mutations: their effects are often not binary—as in, they are beneficial in an environment or not—but rather, have stories that are more nuanced. This is consistent with studies of bacterial translation machinery have reflected the dubiousness of compensatory mutations in evolution [44].

The term "compensatory" implies that a mutation appears on genetic backgrounds where its effects are restorative. That is, one might believe that compensatory mutations must oc-

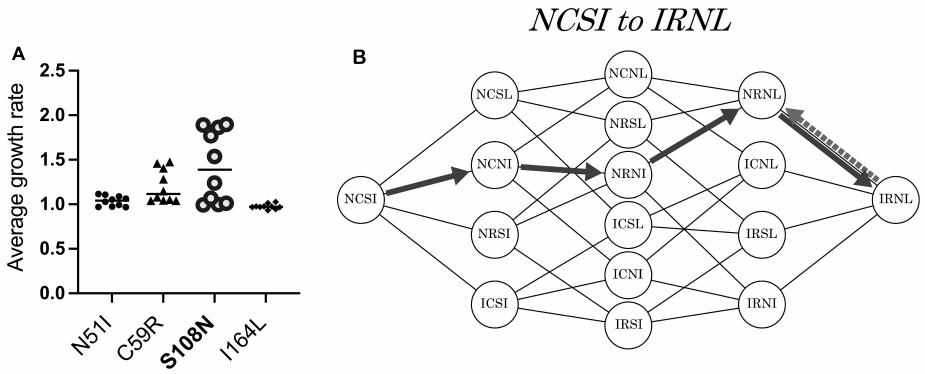

Figure 3: The signature of a compensatory ratchet mutation. Here we depict how the S108N mutation's effect across environments plays a critical role in "forward" evolution and undermines "reverse" evolution. (A) In pyrimethamine, alleles containing the S018N mutation grow significantly better than any set of alleles containing any other single mutation (Kruskal-Wallis: 5A, pyrimethamine, p = 0.0002). (B) The hypercube represents the combinatorial set of 16 alleles as described in the Methods. Based on the rank-order in Figure 1, the predicted pathways of stepwise evolution from the wild-type genotype (NCSI) through the fitness landscape at a high drug concentration (black arrows) at 106 uM, and reversal in a drugless environment (dashed arrow). In this scenario, the compensatory nature of the S108N mutation provides a compensatory ratchet, that helps evolution evolve up a fitness peak in the high drug concentration (black arrows) but prevents it from reversing towards wildtype (NCSI) in a drugless environment (dashed arrow).

cur downstream in an adaptive landscape. As summarized in the adaptive landscape in Figure 3B, the S108N mutation can occur early in an adaptive trajectory (e.g., first step). Consequently, the language of epistatic "ratchet" applies as well, even though studies that have outlined them often describe ratchets in terms more biophysical than are explored in this study [45, 46]. For that reason, we can call mutations like S108N—whose environment specific effects (discussed in Figs. 2 and 3) impede reverse evolution—"compensatory ratchets." The signature of a compensatory ratchet mutation is critical for understanding the conditions that foster or prevent reverse evolution in evolutionary dynamics. Dollo's Law is a classical concept from paleontology that postulates that reverse evolution is unlikely at the macroevolutionary scale [47, 48]. Several studies have addressed the molecular underpinnings of reverse evolution, and highlighted how "hidden randomness" [49], epistasis [45, 42, 50], or population genetic parameters like mutation rate [51] can constrain evolution and reversal.

These findings highlight why reversal can be unlikely in circumstances where phenotypic effects of mutations (and interactions) are specific to certain environmental contexts: compensatory mutations can have environment-specific interactions with other mutations, that impede reversibility. This has practical relevance and informs public health and biomedical approaches to addressing the antimicrobial resistance problem. Studies have revealed why resistance management approaches that attempt to drive populations of resistant microbes "back" to more susceptible forms can be challenging [52, 53, 54]. And other studies have proposed methods to reverse the effects of resistance [55].

V. CONCLUSION

Table 1 describes the differences between the reaction norm and the mutation effect reaction norm, including how one can interpret the information contained within them. This mutation effect reaction norm abstraction offers a different view of gene by environment interactions, by deconstructing genotype-phenotype maps in terms of the often peculiar interactions between genes and mutations across environments. This can offer a "mutation-centric" view of genotype-phenotype maps where the object of interest is not entire haplotypes, but rather, the individual interactions between the mutations that compose those haplotypes. While evolutionary geneticists have long appreciated the importance of the measuring the effects of individual mutations,

I argue that the mutation effect reaction norm represents another level of nuance, whereby more detailed biographies of the mutations are available.

From the data that composes a norm of reaction, we can compute the fitness effects, interaction terms and epistatic coefficients. From this, we can extract meaningful information about how evolution may take place across the environments that define reaction norm, or how we might expect an evolved or engineered mutation to perform across relevant environmental contexts.

It is important to note that multiple methods exist to measure interactions between mutations (see Methods). In this study, I utilized a single established method and data structure. But the mutation effect reaction norm could conceivably be computed with metrics used to measure mutation effects or epistasis in any one of several ways, for problem cases more complicated than the one outlined in this study. For example, tests of marginal epistasis across large genomic data sets effects could also be used to measure interactions across environmental contexts that affect genome-wide signatures of epistasis [56]. Methods like these are the key to broadening the reach of the mutation effect reaction norm to larger scale evolutionary genetics scales and datasets.

Finally, I argue that the mutation effect reaction norm has important practical implications across several domains. By understanding how mutation effects and epistasis are driven by context, we can appreciate their role in obfuscating results from experimental studies of mutations [57], or in terms of how the context specificity of mutations may complicate genetic modification efforts [58]. And the implications for predictive evolution are fairly obvious: the environmental mediator of nonlinear genetic interactions frames the topography of genotype-phenotype space, and by extension, how we expect the process of adaptive evolution to occur [59, 60, 61]. In addition, recent studies that support the role of contingency in evolution might be explained by the specificity of mutation effects, adaptive landscape topography or environmental context [62, 63]. In the public health arena, the abstraction has obvious connections to modern efforts to explain and predict the evolution of antibiotic resistance, often through measuring the effects of mutations[64, 65, 66, 67].

But the applications of the mutation effect reaction norm transcend the antimicrobial realm. For example, it may allow us to better identify mutations in pathogens that are associated with emergence such as the ones identified in the form of a "watchlist" of mutations [68]. This notion has become especially relevant in the context of the COVID-19 pandemic. A year after the start of the pandemic, several variants of concern (VoC) began to circulate and define a new wave of the pandemic globally. These VoCs are the product of suites of mutations, almost surely some of which interact in a nonlinear fashion and have complicated our attempts to resolve which mutations are sole signatures of pathogenic potential (specifically, increased transmission and/or possible escape from vaccine-induced immunity) [69]. I argue that the key to properly characterizing these mutations resides in an understanding of how environmental context shapes their effect. That is, host structure, demographics (e.g., age), and other factors may influence how a given SARS-CoV-2 mutation interacts with others, creating a variant of concern.

Notably absent from my introduction of the mutation effect reaction norm are mathematical formalisms that more directly link the shape of underlying reaction norms to that of the mutation effect reaction norms. While this study proposes the utility for an abstraction that depicts how environments tune interactions between mutations, it doesn't firmly ground them in quantitative genetic terms, like the reaction norm has been in the past [70, 19, 71]. The existence of that studies have utilized rank orders of genotypes to infer genetic interactions [41, 33], those that have examined the fate of mutations in fluctuating environments [72] and others that mathematically resolve the shape of fitness landscapes [73] suggest that such formalisms may exist. An analytical reconciliation between these existing treatments constitutes an obvious future direction of inquiry.

These gaps notwithstanding, the mutation effect reaction norm may allow us to tell more meaningful stories of mutation effects, that describe many aspects of how a mutation or SNP of interest may have arisen, and/or how we would expect it to influence phenotypes of interest. While true prediction in evolutionary genetics remains mostly a fantasy, future efforts to properly predict evolution must include a consideration for how environments tune interactions between mutations.

ACKNOWLEDGEMENTS

I would like to thank R. Oomen, M. Miller-Dickson, R. Guerrero, D. Weinreich, and P. Pennings for helpful discussions on this topic. I would like to thank M. Rebolleda-Gomez, R. Shaw, the participants, and other contributors to the symposium entitled "Society for the Study of Evolution (SSE) at 75 years: continuity and change in evolutionary research" at the 2021 Evolution Meetings. In addition, I would like to thank the organizers and participants in the 2020 Population, Evolutionary, and Quantitative Genetics (PEQG) symposium at the Genetics Society of America. And I would like to acknowledge seminar invitations from the University of Michigan; North Carolina State University; the University of California, Santa Cruz; Princeton University; and the University of Chicago, where iterations of the ideas in this manuscript were discussed. I would like to acknowledge NSF Grant 1736253 "Using Biophysical Protein Models to Map Genetic Variation to Phenotypes."

Term	Data	Interpretation
Reaction norm or norm of reaction	The performance, phenotype or trait value for different alleles, strains, mutants, variants, or forms, or populations, across an environmental gradient (continuous or other)	Used to identify and measure the effects of genes, environments, gene x environment interactions, phenotypic plasticity, and other related properties in quantitative genetics, evolutionary genetics, and ecology.
Mutation effect reaction norm (MERN)	The phenotypic effect of individual or collections of mutations across a set of environments (continuous or other)	Used to measure how environments influence the effect of individual mutations, the interaction between suites of mutations, and epistatic interactions. It can help to explain how the environmental sculpts the topography of adaptive landscapes via environment-dependent mutation effects.

Table 1: Reaction norm vs. Mutation effect reaction norm. This table briefly describes the differences between the two abstractions

SUPPORTING INFORMATION AND DATA ARCHIVING

Data and other materials can be found on Figshare: https://doi.org/10.6084/m9.figshare.16661371

REFERENCES

[1] N.H. Barton, A.M. Etheridge, and A. Véber. The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, 118:50–73, 2017.

[2] Evan A. Boyle, Yang I. Li, and Jonathan K. Pritchard. An expanded view of complex traits: From polygenic to omnigenic. *Cell*, 169(7):1177–1186, 2017.

[3] Timothy B. Sackton and Daniel L. Hartl. Genotypic context and epistasis in individuals and populations. *Cell*, 166(2):279–287, 2016.

[4] Daniel M Weinreich, Yinghong Lan, C Scott Wylie, and Robert B. Heckendorn. Should evolutionary geneticists worry about higher-order epistasis? *Current Opinion in Genetics & Development*, 23(6):700–707, 2013.

[5] Jeremy A. Draghi and Joshua B. Plotkin. Selection biases the prevalence and type of epistasis along adaptive trajectories. *Evolution*, 67(11):3120–3131, 2013.

[6] Hsin-Hung Chou, Nigel F. Delaney, Jeremy A. Draghi, and Christopher J. Marx. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. *PLoS Genetics*, 10(2):e1004149, 2014.

[7] Shimon Bershtein, Michal Segal, Roy Bekerman, Nobuhiko Tokuriki, and Dan S. Tawfik. Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature*, 444(7121):929–932, 2006.

[8] Tyler N. Starr and Joseph W. Thornton. Epistasis in protein evolution. *Protein Science*, 25(7):1204–1218, 2016.

[9] Amit Kumar, Chandrasekhar Natarajan, Hideaki Moriyama, Christopher C. Witt, Roy E. Weber, Angela Fago, and Jay F. Storz. Stability-mediated epistasis restricts accessible mutational pathways in the functional evolution of avian hemoglobin. *Molecular Biology and Evolution*, 34(5):1240–1251, 2017.

[10] Santiago F. Elena and Richard E. Lenski. Epistasis between new mutations and genetic background and a test of genetic canalization. *Evolution*, 55(9):1746–1752, 2001.

[11] R. Sanjuan and S. F. Elena. Epistasis correlates to genomic complexity. *Proceedings of the National Academy of Sciences*, 103(39):14402–14405, 2006.

[12] J M Cheverud and E J Routman. Epistasis and its contribution to genetic variance components. *Genetics*, 139(3):1455–1461, 1995.

[13] H. J. Cordell. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics*, 11(20):2463–2468, 2002.

[14] Xionglei He. The biology complicated by genetic analysis. *Molecular Biology and Evolution*, 33(9):2177–2181, 2016.

[15] David Haig. Does heritability hide in epistasis between linked SNPs? *European Journal of Human Genetics*, 19(2):123–123, 2010.

[16] O. Zuk, E. Hechter, S. R. Sunyaev, and E. S. Lander. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, 109(4):1193–1198, 2012.

[17] Gibran Hemani, Sara Knott, and Chris Haley. An evolutionary perspective on epistasis and the missing heritability. *PLoS Genetics*, 9(2):e1003295, 2013.

[18] Rebekah A Oomen and Jeffrey Alexander Hutchings. *Evolution of Reaction Norms*. Oxford University Press, Oxford, England, 2020.

[19] Richard Gomulkiewicz and Mark Kirkpatrick. Quantitative genetics and the evolution of reaction norms. *Evolution*, 46(2):390, 1992.

[20] Rebekah A. Oomen and Jeffrey A. Hutchings. Genetic variability in reaction norms in fishes. *Environmental Reviews*, 23(3):353–366, 2015.

[21] Carl Schlichting and Massimo Pigliucci. *Phenotypic Evolution: A Reaction Norm Perspective*. Sinauer, Cary, NC, 1998.

[22] Carl D. Schlichting and Harry Smith. Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, 16(3):189–211, 2002.

[23] Kenneth M. Flynn, Tim F. Cooper, Francisco B-G. Moore, and Vaughn S. Cooper. The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. *PLoS Genetics*, 9(4):e1003426, 2013.

[24] Anne E. Hall, Kedar Karkare, Vaughn S. Cooper, Claudia Bank, Tim F. Cooper, and Francisco B.-G. Moore. Environment changes epistasis to alter trade-offs along alternative evolutionary paths. *Evolution*, 73(10):2094–2105, 2019.

[25] Elena R Lozovsky, Rachel F Daniels, Gavin D Heffernan, David P Jacobus, and Daniel L Hartl. Relevance of higher-order epistasis in drug resistance. *Molecular Biology and Evolution*, 38(1):142–151, 2020.

[26] Lingchong You and John Yin. Dependence of epistasis on environment and mutation severity as revealed byin SilicoMutagenesis of phage t7. *Genetics*, 160(4):1273–1281, 2002.

[27] Chuan Li and Jianzhi Zhang. Multi-environment fitness landscapes of a tRNA gene. *Nature Ecology & Evolution*, 2(6):1025–1032, 2018.

[28] D. S. Peterson, W. K. Milhous, and T. E. Wellems. Molecular basis of differential resistance to cycloguanil and pyrimethamine in plasmodium falciparum malaria. *Proceedings of the National Academy of Sciences*, 87(8):3018–3022, April 1990.

[29] C. Brandon Ogbunugafor, C. Scott Wylie, Ibrahim Diakite, Daniel M. Weinreich, and Daniel L. Hartl. Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance. *PLOS Computational Biology*, 12(1):e1004710, 2016.

[30] C. Brandon Ogbunugafor and Margaret J. Eppstein. Competition along trajectories governs adaptation rates towards antimicrobial resistance. *Nature Ecology & Evolution*, 1(1):1–8, 2016.

[31] Zachary R Sailer and Michael J Harms. Detecting high-order epistasis in nonlinear genotype-phenotype maps. *Genetics*, 205(3):1079–1088, 2017.

[32] Zachary R. Sailer and Michael J. Harms. Molecular ensembles make evolution unpredictable. *Proceedings of the National Academy of Sciences*, 114(45):11938–11943, 2017.

[33] Kristina Crona. Rank orders and signed interactions in evolutionary biology. *eLife*, 9:e51004, 2020.

[34] J. Otwinowski and J. B. Plotkin. Inferring fitness landscapes by regression produces biased estimates of epistasis. *Proceedings of the National Academy of Sciences*, 111(22):E2301–E2309, 2014.

[35] Zachary R. Sailer and Michael J. Harms. Uninterpretable interactions: epistasis as uncertainty. *bioRxiv*, preprint:1–32, 2018.

[36] Daniel M. Weinreich, Yinghong Lan, Jacob Jaffe, and Robert B. Heckendorn. The influence of higher-order epistasis on biological fitness landscape topography. *Journal of Statistical Physics*, 172(1):208–225, 2018.

[37] Frank J. Poelwijk, Vinod Krishna, and Rama Ranganathan. The context-dependence of mutations: A linkage of formalisms. *PLOS Computational Biology*, 12(6):e1004771, 2016.

[38] João V. Rodrigues, Shimon Bershtein, Anna Li, Elena R. Lozovsky, Daniel L. Hartl, and Eugene I. Shakhnovich. Biophysical principles predict fitness landscapes of drug resistance. *Proceedings of the National Academy of Sciences*, 113(11):E1470–E1478, 2016.

[39] Rafael F Guerrero, Samuel V Scarpino, João V Rodrigues, Daniel L Hartl, and C Brandon Ogbunugafor. Proteostasis environment shapes higher-order epistasis operating on antibiotic resistance. *Genetics*, 212(2):565–575, 2019.

[40] Victor A. Meszaros, Miles D. Miller-Dickson, and C. Brandon Ogbunugafor. Lexical landscapes as large in silico data for examining advanced properties of fitness landscapes. *PLOS ONE*, 14(8):e0220891, 2019.

[41] Daniel M Weinreich. The rank ordering of genotypic fitness values predicts genetic constraint on natural selection on landscapes lacking sign epistasis. *Genetics*, 171(3):1397–1405, 2005.

[42] C. Brandon Ogbunugafor and Daniel Hartl. A pivot mutation impedes reverse evolution across an adaptive landscape for drug resistance in plasmodium vivax. *Malaria Journal*, 15(1):1–20, 2016.

[43] Ville Mustonen and Michael Lässig. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends in Genetics*, 25(3):111–119, 2009.

[44] Sandeep Venkataram, Ross Monasky, Shohreh H. Sikaroodi, Sergey Kryazhimskiy, and Betul Kacar. Evolutionary stalling and a limit on the power of natural selection to improve a cellular module. *Proceedings of the National Academy of Sciences*, 117(31):18582–18590, 2020.

[45] Jamie T. Bridgham, Eric A. Ortlund, and Joseph W. Thornton. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature*, 461(7263):515–519, 2009.

[46] Moshe Ben-David, Misha Soskine, Artem Dubovetskyi, Kesava-Phaneendra Cherukuri, Orly Dym, Joel L Sussman, Qinghua Liao, Klaudia Szeler, Shina Caroline Lynn Kamerlin, and Dan S Tawfik. Enzyme evolution: An epistatic ratchet versus a smooth reversible transition. *Molecular Biology and Evolution*, 37(4):1133–1147, 2019.

[47] Stephen Jay Gould. Dollo on dollo's law: Irreversibility and the status of evolutionary laws. *Journal of the History of Biology*, 3(2):189–212, 1970.

[48] Rachel Collin and Maria Pia Miglietta. Reversing opinions on dollo's law. *Trends in Ecology & Evolution*, 23(11):602–609, 2008.

[49] Longzhi Tan, Stephen Serene, Hui Xiao Chao, and Jeff Gore. Hidden randomness between fitness landscapes limits reverse evolution. *Physical Review Letters*, 106(19):1–4, 2011.

[50] Jay F Storz. Compensatory mutations and epistasis for protein function. *Current Opinion in Structural Biology*, 50:18–25, 2018.

[51] Pleuni S. Pennings, C. Brandon Ogbunugafor, and Ruth Hershberg. Reversion is most likely under high mutation supply, when compensatory mutations don't fully restore fitness costs. *bioRxiv*, preprint:1–9, 2020.

[52] Maria Sjölund, Karin Wreiber, Dan I. Andersson, Martin J. Blaser, and Lars Engstrand. Long-term persistence of resistant enterococcus species after antibiotics to eradicate helicobacter pylori. *Annals of Internal Medicine*, 139(6):483–487, 2003.

[53] Dan I. Andersson and Diarmaid Hughes. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology*, 8(4):260–271, 2010.

[54] M. Sundqvist, P. Geli, D. I. Andersson, M. Sjolund-Karlsson, A. Runehagen, H. Cars, K. Abelson-Storby, O. Cars, and G. Kahlmeter. Little evidence for reversibility of trimethoprim resistance after a drastic reduction in trimethoprim use. *Journal of Antimicrobial Chemotherapy*, 65(2):350–360, 2009.

[55] M. Baym, L. K. Stone, and R. Kishony. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science*, 351(6268):aad3292–aad3292, 2015.

[56] Lorin Crawford, Ping Zeng, Sayan Mukherjee, and Xiang Zhou. Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. *PLOS Genetics*, 13(7):e1006869, 2017.

[57] Tom J. Little and Nick Colegrave. Caging and uncaging genetics. *PLOS Biology*, 14(7):e1002525, 2016.

[58] Peter B. Otoupal, Keesha E. Erickson, Antoni Escalas-Bordoy, and Anushree Chatterjee. CRISPR perturbation of gene expression alters bacterial fitness under stress and reveals underlying epistatic constraints. *ACS Synthetic Biology*, 6(1):94–107, 2016.

[59] Adam C. Palmer and Roy Kishony. Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. *Nature Reviews Genetics*, 14(4):243–248, 2013.

[60] J. Arjan G.M. de Visser and Joachim Krug. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*, 15(7):480–490, 2014.

[61] Allison J. Lopatkin and James J. Collins. Predictive biology: modelling, understanding and harnessing microbial complexity. *Nature Reviews Microbiology*, 18(9):507–520, 2020.

[62] Mateusz Kędzior, Amanda K. Garcia, Meng Li, Arnaud Taton, Zachary R. Adam, Jodi N. Young, and Betül Kaçar. Molecular foundations of precambrian uniformitarianism. *bioRxiv*, preprint:1–24, 2021.

[63] Victoria Cochran Xie, Jinyue Pu, Brian PH Metzger, Joseph W Thornton, and Bryan C Dickinson. Contingency and chance erase necessity in the experimental evolution of ancestral proteins. *eLife*, 10:e67336, 2021.

[64] Adam C. Palmer, Erdal Toprak, Michael Baym, Seungsoo Kim, Adrian Veres, Shimon Bershtein, and Roy Kishony. Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. *Nature Communications*, 6(1):1–8, 2015.

[65] Jennifer Knies, Fei Cai, and Daniel M. Weinreich. Enzyme efficiency but not thermostability drives cefotaxime resistance evolution in TEM-1 β-lactamase. *Molecular Biology and Evolution*, 34(5):1040–1054, 2017.

[66] Erida Gjini and Kevin B Wood. Price equation captures the role of drug interactions and collateral effects in the evolution of multidrug resistance. *eLife*, 10:e64851, 2021.

[67] Sarah M. Ardell and Sergey Kryazhimskiy. The population genetics of collateral resistance and sensitivity. *bioRxiv*, preprint:1–39, 2020.

[68] Craig R. Miller, Erin L. Johnson, Aran Z. Burke, Kyle P. Martin, Tanya A. Miura, Holly A. Wichman, Celeste J. Brown, and F. Marty Ytreberg. Initiating a watch list for ebola virus antibody escape mutations. *PeerJ*, 4:e1674, 2016.

[69] Sarah P. Otto, Troy Day, Julien Arino, Caroline Colijn, Jonathan Dushoff, Michael Li, Samir Mechai, Gary Van Domselaar, Jianhong Wu, David J.D. Earn, and Nicholas H. Ogden. The origins and potential future of SARS-CoV-2 variants of concern in the evolving COVID-19 pandemic. *Current Biology*, 31(14):R918–R929, 2021.

[70] G. de Jong. Quantitative genetics of reaction norms. *Journal of Evolutionary Biology*, 3(5-6):447–468, 1990.

[71] Carl D. Schlichting and Massimo Pigliucci. Gene regulation, quantitative genetics and the evolution of reaction norms. *Evolutionary Ecology*, 9(2):154–168, 1995.

[72] Ivana Cvijović, Benjamin H. Good, Elizabeth R. Jerison, and Michael M. Desai. Fate of a mutation in a fluctuating environment. *Proceedings of the National Academy of Sciences*, 112(36):E5021–E5028, 2015.

[73] Daniel M Weinreich, Suzanne Sindi, and Richard A Watson. Finding the boundary between evolutionary basins of attraction, and implications for wright's fitness landscape analogy. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01001, 2013.