

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

The evolutionary implications of epistasis and pleiotropy in microbial populations

### Permalink

<https://escholarship.org/uc/item/06v5h54w>

### Author

Ardell, Sarah

### Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

The evolutionary implications of epistasis and pleiotropy in microbial populations

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy

in

Biology with a Specialization in Quantitative Biology

by

Sarah Ardell

Committee in charge:

Professor Sergey Kryazhimskiy, Chair  
Professor Terence Hwa  
Professor Justin Meyer  
Professor Diana Rennison

2024



Copyright

Sarah Ardell, 2024

All rights reserved.

The Dissertation of Sarah Ardell is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

## TABLE OF CONTENTS

|                                                                                                                          |      |
|--------------------------------------------------------------------------------------------------------------------------|------|
| Dissertation Approval Page .....                                                                                         | iii  |
| Table of Contents .....                                                                                                  | iv   |
| List of Figures .....                                                                                                    | vi   |
| List of Tables .....                                                                                                     | viii |
| Acknowledgements .....                                                                                                   | ix   |
| Vita .....                                                                                                               | xi   |
| Abstract of the Dissertation .....                                                                                       | xii  |
| Chapter 1     The population genetics of collateral resistance and sensitivity .....                                     | 1    |
| 1.1     Abstract .....                                                                                                   | 1    |
| 1.2     Introduction .....                                                                                               | 1    |
| 1.3     Results .....                                                                                                    | 5    |
| 1.3.1     Antibiotic resistance mutations in <i>E. coli</i> have diverse collateral effects .....                        | 5    |
| 1.3.2     JDFE determines the pleiotropic outcomes of adaptation .....                                                   | 7    |
| 1.3.3     The population genetics of pleiotropy .....                                                                    | 10   |
| 1.3.4     Robust ranking of drug pairs .....                                                                             | 16   |
| 1.3.5     Measuring JDFEs .....                                                                                          | 17   |
| 1.4     Discussion .....                                                                                                 | 20   |
| 1.5     Materials and Methods .....                                                                                      | 24   |
| 1.5.1     Analysis of knock-out and deep mutational scanning data .....                                                  | 24   |
| 1.5.2     Theory .....                                                                                                   | 25   |
| 1.5.3     Ranking of drug pairs .....                                                                                    | 29   |
| 1.5.4     Generation of JDFEs .....                                                                                      | 29   |
| 1.5.5     Simulations .....                                                                                              | 30   |
| 1.5.6     Sampling beneficial mutants from JDFEs and estimating the collateral risk statistic .....                      | 31   |
| 1.6     Acknowledgements .....                                                                                           | 32   |
| Chapter 2     Environment-independent distribution of mutational effects emerges from microscopic global epistasis ..... | 33   |
| 2.1     Abstract .....                                                                                                   | 33   |
| 2.2     Main Text .....                                                                                                  | 33   |
| 2.2.1     Absence of unconditionally beneficial or deleterious mutations .....                                           | 35   |
| 2.2.2     Consistency of microscopic global epistasis across environments .....                                          | 37   |
| 2.2.3     Distribution of fitness effects .....                                                                          | 41   |
| 2.2.4     Discussion .....                                                                                               | 44   |
| 2.3     Materials and Methods .....                                                                                      | 46   |

|            |                                                                                                                         |     |
|------------|-------------------------------------------------------------------------------------------------------------------------|-----|
| 2.3.1      | Strains and media . . . . .                                                                                             | 46  |
| 2.3.2      | Experimental procedures . . . . .                                                                                       | 47  |
| 2.3.3      | Data Analysis . . . . .                                                                                                 | 52  |
| 2.3.4      | Theoretical calculation of DFE moments . . . . .                                                                        | 71  |
| 2.4        | Supplementary Text . . . . .                                                                                            | 77  |
| 2.4.1      | Justification for using absolute growth rate rather than relative fitness . . . . .                                     | 77  |
| 2.4.2      | Comparison with previous studies and potential limitations . . . . .                                                    | 78  |
| 2.4.3      | Interpretation of the generalized global epistasis equation in terms of the connectedness model . . . . .               | 82  |
| 2.5        | Acknowledgements . . . . .                                                                                              | 85  |
| Chapter 3  | The statistical structure of pleiotropy in the presence of global epistasis and its evolutionary consequences . . . . . | 86  |
| 3.1        | Abstract . . . . .                                                                                                      | 86  |
| 3.2        | Main Text . . . . .                                                                                                     | 87  |
| 3.2.1      | Introduction . . . . .                                                                                                  | 87  |
| 3.2.2      | JDFEs vary widely across strains and environments . . . . .                                                             | 89  |
| 3.2.3      | Generalized global epistasis equation predicts JDFEs . . . . .                                                          | 90  |
| 3.2.4      | Evolutionary implications . . . . .                                                                                     | 97  |
| 3.2.5      | Discussion . . . . .                                                                                                    | 99  |
| 3.3        | Materials and Methods . . . . .                                                                                         | 100 |
| 3.3.1      | Data availability and code . . . . .                                                                                    | 100 |
| 3.3.2      | Analysis of empirical JDFEs . . . . .                                                                                   | 101 |
| 3.3.3      | Theoretical calculation of JDFE correlation and the pleiotropy statistic . . . . .                                      | 102 |
| 3.3.4      | Pivot noise covariance . . . . .                                                                                        | 106 |
| 3.3.5      | Parameterization of the predictions . . . . .                                                                           | 106 |
| 3.3.6      | Non-parametric predicted JDFEs . . . . .                                                                                | 107 |
| 3.3.7      | Evolutionary simulations . . . . .                                                                                      | 108 |
| 3.4        | Acknowledgements . . . . .                                                                                              | 109 |
| Appendix A | Chapter 1 Supplementary Figures . . . . .                                                                               | 110 |
| Appendix B | Chapter 2 Supplementary Figures and Tables . . . . .                                                                    | 116 |
| B.1        | Supplementary tables . . . . .                                                                                          | 131 |
| Appendix C | Chapter 3 Supplementary Figures and Tables . . . . .                                                                    | 135 |
| C.1        | Supplementary Tables . . . . .                                                                                          | 138 |

## LIST OF FIGURES

|             |                                                                                                                                          |     |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1.1. | Fitness effects of gene knock-out mutations in <i>E. coli</i> in the presence of four antibiotics . . . . .                              | 6   |
| Figure 1.2. | Gaussian JDFEs and the resulting fitness trajectories . . . . .                                                                          | 8   |
| Figure 1.3. | Pleiotropy statistics predict the properties of non-home fitness trajectories in simulations . . . . .                                   | 12  |
| Figure 1.4. | Robust ranking of drug pairs . . . . .                                                                                                   | 15  |
| Figure 1.5. | Sampling effects on the ranking of drug pairs . . . . .                                                                                  | 19  |
| Figure 2.1. | Proportions of beneficial and deleterious mutations vary with strain growth rate . . . . .                                               | 36  |
| Figure 2.2. | Distributions of global-epistasis slopes and intercepts and their correlation . . . . .                                                  | 38  |
| Figure 2.3. | Global-epistasis slopes of mutations are nearly invariant across environments. . . . .                                                   | 40  |
| Figure 2.4. | Generalized global epistasis equation captures changes in the DFE across strains and environments. . . . .                               | 43  |
| Figure 3.1. | Variation in empirical JDFEs . . . . .                                                                                                   | 91  |
| Figure 3.2. | Prediction of JDFE correlation . . . . .                                                                                                 | 93  |
| Figure 3.3. | Prediction of the pleiotropy statistic $r_2$ . . . . .                                                                                   | 95  |
| Figure 3.4. | Model of cross-environmental global epistasis . . . . .                                                                                  | 96  |
| Figure 3.5. | Evolutionary implications of cross-environmental global epistasis . . . . .                                                              | 99  |
| Figure A.1. | JDFEs of single point mutations in TEM-1 $\beta$ -lactamase gene in <i>E. coli</i> in the presence of cefotaxime and ampicillin. . . . . | 111 |
| Figure A.2. | JDFEs with equal probability weights in the first and fourth quadrants and the resulting fitness trajectories. . . . .                   | 112 |
| Figure A.3. | Same as Figure 1.3C–H, but with $r_2$ and $D_{22}$ shown on the $x$ -axis. . . . .                                                       | 113 |
| Figure A.4. | Ranking of AMP concentrations according to their risk of collateral resistance . . . . .                                                 | 114 |
| Figure A.5. | Same as Figure 1.5A, but with different thresholds for sampling. . . . .                                                                 | 115 |

|              |                                                                                               |     |
|--------------|-----------------------------------------------------------------------------------------------|-----|
| Figure B.1.  | Experimental setup .....                                                                      | 117 |
| Figure B.2.  | Data quality checks .....                                                                     | 118 |
| Figure B.3.  | Variation in the sign of mutational effects across genetic backgrounds and environments ..... | 120 |
| Figure B.4.  | Properties of global epistasis models .....                                                   | 121 |
| Figure B.5.  | Variable and invariant slope models .....                                                     | 123 |
| Figure B.6.  | Global epistasis model for individual mutations .....                                         | 124 |
| Figure B.7.  | Reshuffling of background strain growth rates across environments .....                       | 125 |
| Figure B.8.  | Reshuffling of the effects of mutations across environments .....                             | 126 |
| Figure B.9.  | Distributions of DFEs similarity statistics .....                                             | 127 |
| Figure B.10. | Robustness of the observed DFE variation with respect to missing measurements .....           | 128 |
| Figure B.11. | QTL analysis .....                                                                            | 129 |
| Figure B.12. | Connectedness model and evolutionary simulations .....                                        | 130 |
| Figure C.1.  | Noise distributions and covariance .....                                                      | 136 |
| Figure C.2.  | Impact of clonal interference and idiosyncratic epistasis .....                               | 137 |

## LIST OF TABLES

|            |                                                                                                                                   |     |
|------------|-----------------------------------------------------------------------------------------------------------------------------------|-----|
| Table B.1. | Statistics of slope and intercept distributions .....                                                                             | 131 |
| Table B.2. | Antibiotic concentrations used in this study, in $\mu\text{g/ml}$ .....                                                           | 132 |
| Table B.3. | Best fit trends in DFE moments .....                                                                                              | 133 |
| Table B.4. | GO term enrichment among all mutations in this study. <i>P</i> -value is reported<br>after the Benjamini-Hochberg correction..... | 134 |
| Table C.1. | Parameters of slope distribution density curve .....                                                                              | 138 |

## ACKNOWLEDGEMENTS

First, I would like to thank my doctoral advisor Dr. Sergey Kryazhimskiy for believing in me from the very beginning. Thank you for being an excellent example of clear thinking and clearer communication, and for being constantly willing to help whenever I asked. Through his mentorship, I have grown as not only a scientist, but as a communicator, a collaborator, and a mentor in my own right.

I would also like to thank my committee members Dr. Justin Meyer, Dr. Diana Rennison, and Dr. Terry Hwa for their excellent feedback and consistent votes of confidence. Their insights have strengthened the rigor and scope of my work. I would especially like to thank Dr. Justin Meyer who frequently went above and beyond to check in and offer his support.

Thank you to all the members of the Kryazhimskiy lab both past and present for their thoughtful feedback and for their enthusiasm for my work. I must especially thank Michael Overton who always had ‘just a minute’ to talk with me and whose selflessness and mentorship skills are major inspirations to me.

I am grateful to the EBE department at UCSD as a whole for cultivating such a positive, collaborative and supportive culture. I am especially grateful to my EBE cohort, soon to be doctors Sahana Kuthyar, Ashley Kim, and Ugbad Farah, who have been with me every step of the way. Time spent with you all has been a huge highlight of this journey. I would also like to thank Megan Demmel and Karagan Smith for their kindness and joy which has made the last few months of this experience all the more memorable. All of these women are constant sources of inspiration to me.

I also must thank many additional friends for their endless patience and for being such excellent cheerleaders. In particular, thank you Haleigh, Ezra, Julie, and Debbie for having endless wells of love to share. To Gabriel, for the many Friday walks, Tuesday movies, and sage advice. And to Alia, for always showing up for me no matter what.

Words cannot capture my gratitude for my partner, Joshua Borin. From being the first test audience for every talk, to planning beautiful trips, to quite literally keeping me alive through



many ailments, Josh's relentless support made grad school not only possible but (often) enjoyable. I cannot thank you enough for your steadfast belief in me.

Finally, I am blessed to have a wealth of family who have been constant sources of motivation, support, and love. My father and step-mother, Jeff and Kim Ardell, offer me the balance between the inspiration to pursue important new science and the wisdom to remember that family is the most important thing of all. To my siblings Christopher, Stacey, Heather, and my niece Saylor, thank you for being such lights in my life. My mother and my step-father, Michelle Duffourc and Robert Schoborg, have never faltered in their support. Thank you, Rob, for seeing me for exactly who I am, and thank you Mom, for always picking up the phone with a smile on your face. I am so grateful for each and every one of you.

Chapter 1, in full, is a reprint of the material as it appears in The population genetics of collateral resistance and sensitivity in *eLife*, 2021. Ardell, Sarah and Kryazhimskiy, Sergey. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in An environment independent distribution of mutational effects emerges from microscopic global epistasis in *bioRxiv*, 2024. Ardell, Sarah; Marstul, Alena; Johnson, Milo; Kryazhimskiy, Sergey. The dissertation author was the primary investigator and co-primary author of this paper.

Chapter 3, is currently being prepared for submission for publication of the material. Ardell, Sarah and Kryazhimskiy, Sergey. The dissertation author was the primary investigator and author of this material.

## VITA

|      |                                                          |
|------|----------------------------------------------------------|
| 2019 | Bachelor of Science, University of California San Diego  |
| 2019 | Course Assistant, University of California Santa Cruz    |
| 2020 | Research Assistant, University of California San Diego   |
| 2024 | Doctor of Philosophy, University of California San Diego |

## PUBLICATIONS

S.M. Ardell, A. Martsul, M. Johnson, S. Kryazhimskiy. Environment independent distribution of mutational effects emerges from microscopic epistasis. In review. (2024).

K.R. Gerbino, J.M. Borin, S.M. Ardell, J.J. Lee, K.D. Corbett, J.R. Meyer. Bacteriophage  $\Phi$ 21's receptor-binding protein evolves new functions through destabilizing mutations that generate non-genetic phenotypic heterogeneity. In review. (2024).

S.M. Ardell and S. Kryazhimskiy. The population genetics of collateral resistance and sensitivity. *eLife*. (2021).

## ABSTRACT OF THE DISSERTATION

The evolutionary implications of epistasis and pleiotropy in microbial populations

by

Sarah Ardell

Doctor of Philosophy in Biology with a Specialization in Quantitative Biology

University of California San Diego, 2024

Professor Sergey Kryazhimskiy, Chair

Evolutionary dynamics depend on how segregating alleles affect organismal fitness. The ability to predict such dynamics would have far reaching applications from human health to agriculture and biotechnology. A major obstacle to understanding them is that the fitness benefit provided by an allele in one genetic background in one environment will often change, and may even turn into a fitness loss, if this allele lands in a different genetic background ( $G \times G$  interactions, or, epistasis), a different environment ( $G \times E$  interactions, or, pleiotropy) or both ( $G \times G \times E$  interactions). Novel technologies have enabled increasingly high-throughput measurements of mutational effects, which allow us to empirically understand some  $G \times G$ ,  $G \times E$  and  $G \times G \times E$  interactions. However, using these measurements directly to predict evolutionary

dynamics is challenging for several reasons. In particular, the set of available beneficial mutations changes as alleles fix in an evolving population, and the environment may change over the course of adaptation. As such, the number of relevant interactions to measure grows exponentially with evolutionary time.

In this dissertation, I develop a unified framework with which to understand the long-term evolutionary consequences of pleiotropy and epistasis using a combination of simulations, mathematical modeling, and empirical data analysis. In Chapter 1, I develop a novel framework called the “Joint Distribution of Fitness Effects” (or, JDFE), which characterizes how mutations in one environment impact fitness in others. I demonstrate how the JDFE determines evolutionary dynamics with an application to the evolution of multi-antibiotic resistance. In Chapter 2, I analyze a large dataset of mutational effects in diverse strains of yeast *Saccharomyces cerevisiae* to uncover simple predictive patterns in epistasis across environments. In Chapter 3, I extend the empirical cross-environmental epistasis model from Chapter 2 to predict how JDFEs change with evolutionary time. I explore the implications of this unified model to the predictability of long-term evolutionary dynamics in variable environments.

# Chapter 1

## The population genetics of collateral resistance and sensitivity

### 1.1 Abstract

Resistance mutations against one drug can elicit collateral sensitivity against other drugs. Multi-drug treatments exploiting such trade-offs can help slow down the evolution of resistance. However, if mutations with diverse collateral effects are available, a treated population may evolve either collateral sensitivity or collateral resistance. How to design treatments robust to such uncertainty is unclear. We show that many resistance mutations in *Escherichia coli* against various antibiotics indeed have diverse collateral effects. We propose to characterize such diversity with a joint distribution of fitness effects (JDFE) and develop a theory for describing and predicting collateral evolution based on simple statistics of the JDFE. We show how to robustly rank drug pairs to minimize the risk of collateral resistance and how to estimate JDFEs. In addition to practical applications, these results have implications for our understanding of evolution in variable environments.

### 1.2 Introduction

The spread of resistance against most antibiotics and the difficulties in developing new ones has sparked considerable interest in using drug combinations and sequential drug treatments

to treat bacterial infections, as well as cancers [1]. Treatments where the drugs are chosen so that resistance against one of them causes the pathogen or cancer population to become sensitive to the other—a phenomenon known as collateral sensitivity—can eliminate the population before multi-drug resistance emerges [1, 2].

The success of a multi-drug treatment hinges on knowing which drugs select for collateral sensitivity against which other drugs. This information is obtained empirically by exposing bacterial and cancer-cell populations to drugs and observing the evolutionary outcomes [3–19]. Prior studies have largely focused on various empirical questions related to the evolution of collateral sensitivity and resistance, such as identifying their genetic basis [3, 7, 13, 16, 17, 20], understanding how collateral outcomes depend on treatment design (e.g sequential versus combination) [8–11, 14, 15, 19–21], or testing whether collateral sensitivity is an evolutionarily stable outcome [12]. However, one important feature of these experimental studies has received little attention, namely, the fact that different experiments often produce collateral sensitivity profiles that are inconsistent with each other [5, 7, 18, 22]. Some inconsistencies can be explained by the fact that resistance mutations vary between bacterial strains, drug dosages, etc. [22–27]. However, wide variation in collateral outcomes is observed even between replicate populations [7, 18, 22, 28]. This variation suggests that bacteria and cancers have access to multiple resistance mutations with diverse collateral effects and that replicate populations accumulate different resistance mutations due to the intrinsic randomness of the evolutionary process [28, 29]. However, the diversity of collateral effects among resistance mutations has rarely if ever been systematically measured. Moreover, few existing approaches for designing robust multi-drug treatments have modelled this mutational diversity explicitly within the population genetics context [7, 28]. Yet, a theory grounded in population genetics could help us understand how the expected collateral outcomes and the uncertainty around these expectations depend on evolutionary parameters and how these expectations and uncertainties change over time.

Here we develop such a theory. Collateral sensitivity and resistance are specific examples of the more general evolutionary phenomenon, pleiotropy, which refers to any situation when

one mutation affects multiple phenotypes [30, 31]. In case of drug resistance evolution, the direct effect of resistance mutations is to increase fitness in the presence of one drug (the “home” environment). In addition, they may also provide pleiotropic gains or losses in fitness in the presence of other drugs (the “non-home” environments) leading to collateral resistance or sensitivity, respectively.

Classical theoretical work on pleiotropy has been done in the field of quantitative genetics [32–37]. In these models, primarily developed to understand how polygenic traits respond to selection in sexual populations, pleiotropy manifests itself as a correlated temporal change in multiple traits in a given environment. The question of how new strongly beneficial mutations that accumulate in asexual populations evolving in one environment affect its fitness in future environments is outside of the scope of these models.

The pleiotropic consequences of adaptation have also been explored in various “fitness landscape” models [7, 28, 38–42]. In particular, Nichol et al [28] specifically addressed the problem of diversity of collateral resistance/sensitivity outcomes in the context of a combinatorially complete fitness landscapes of four mutations in the TEM  $\beta$ -lactamase gene. They found that different *in silico* populations adapting to the same antibiotic often arrive at different fitness peaks which results in different levels of collateral resistance/sensitivity against other drugs. They observed qualitatively similar variability in the collateral outcomes among replicate populations of *E. coli* evolving in the presence of cefotaxime (CTX), although it is unclear whether different populations indeed arrived at different fitness peaks. In general, the fitness landscape approach helps us understand how evolutionary trajectories and outcomes depend on the global structure of the underlying fitness landscape. However, applying this approach in practice is challenging because the global structure of fitness landscapes is unknown and notoriously difficult to estimate, even in controlled laboratory conditions.

Here, we take a different approach which is agnostic with respect to the global structure of the fitness landscape. Instead, we assume only the knowledge of the so-called joint distribution of fitness effects (JDFE), i.e., the probability that a new mutation has a certain pair of fitness

effects in the home and non-home environments [39, 43, 44]. The JDFE is a natural extension of the DFE, the distribution of fitness effects of new mutations, often used in modeling evolution in a single environment [45–53]. Like the DFE, the JDFE is a local property of the fitness landscape which means that it can be, at least in principle, estimated by using a variety of modern high-throughput techniques [53–58]. The downside of this approach is that the JDFE can change over time as the population traverses the fitness landscape [58–60]. However, in the context of collateral drug resistance and sensitivity, we are primarily interested in short time scales over which the JDFE can be reasonably expected to stay approximately constant.

The rest of the paper is structured as follows. First, we use previously published data to demonstrate that the bacterium *Escherichia coli* has access to drug resistance mutations with diverse collateral effects. This implies that, rather than treating collateral effects as deterministic properties of drug pairs, we should think of them probabilistically, in terms of the respective JDFEs. We then show that a naive intuition about how the JDFE determines pleiotropic outcomes of evolution can sometimes fail, and a mathematical model is therefore required. We develop such a model, which reveals two key “pleiotropy statistics” of the JDFE that determine the dynamics of fitness in the non-home condition. Our theory makes quantitative predictions in a variety of regimes if the population genetic parameters are known. However, we argue that in the case of drug resistance evolution the more important problem is to robustly order drug pairs in terms of their collateral sensitivity profiles even if the population genetic parameters are unknown. We develop a metric that allows us to do so. Finally, we provide some practical guidance for estimating the pleiotropy statistics of empirical JDFEs in the context of ranking drug pairs.



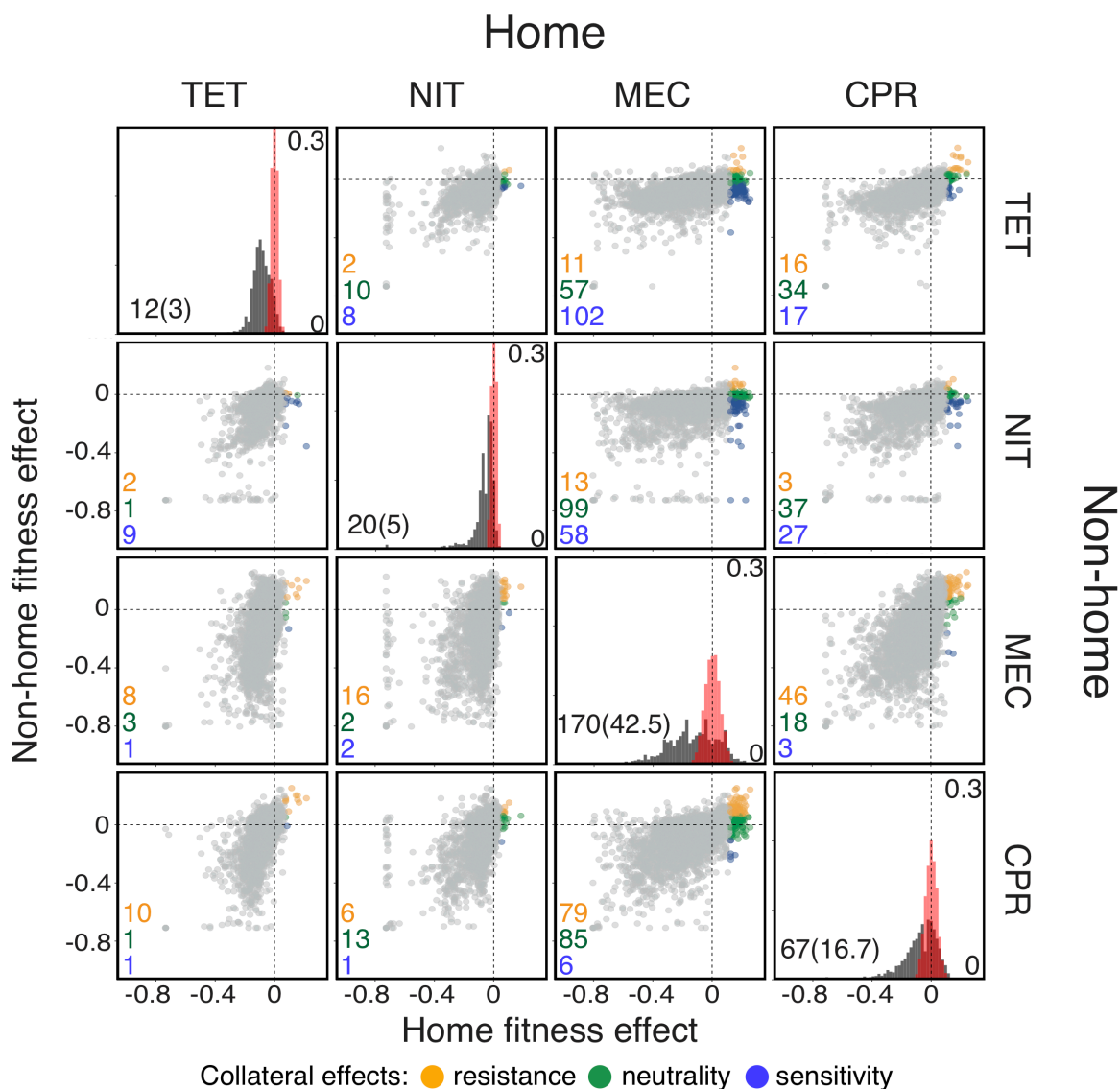
## 1.3 Results

### 1.3.1 Antibiotic resistance mutations in *E. coli* have diverse collateral effects

We begin by demonstrating that JDFE is a useful concept for modeling the evolution of collateral antibiotic resistance and sensitivity. If all resistance mutations against a given drug had identical pleiotropic effects on the fitness of the organism in presence of another drug, the dynamics of collateral resistance/sensitivity could be understood without the JDFE concept. On the other hand, if different resistance mutations have different pleiotropic fitness effects, predicting the collateral resistance/sensitivity dynamics requires specifying the probabilities with which mutations with various home and non-home fitness effects arise in the population. The JDFE specifies these probabilities. Therefore, for the JDFE concept to be useful in the context of collateral resistance/sensitivity evolution, we need to show that resistance mutations against common drugs have diverse collateral effects in the presence of other drugs.

To our knowledge, no data sets are currently publicly available that would allow us to systematically explore the diversity of collateral effects among all resistance mutations against any one drug in any organism. Instead, we examined the fitness effects of 3883 gene knock-out mutations in the bacterium *Escherichia coli*, measured in the presence of six antibiotics [56], as well as the fitness effects of 4997 point mutations in the TEM-1  $\beta$ -lactamase gene measured in the presence of two antibiotics [61].

For four out of six antibiotics used by [56], we find between 12 (0.31 %) and 170 (4.38 %) knock-out mutations that provide some level of resistance against at least one of the antibiotics (false discovery rate (FDR)  $\sim 25\%$ ; Figure ??, see Materials and Methods for details). Plotting on the x-axis the fitness effect of each knock-out mutation in the presence of the drug assumed to be applied first (i.e., the home environment) against its effect in the presence of another drug assumed to be applied later (i.e., the non-home environment, y-axis), we find mutations in all four quadrants of this plane, for all 12 ordered drug pairs (Figure ??). Similarly, we find



**Figure 1.1. Fitness effects of gene knock-out mutations in *E. coli* in the presence of four antibiotics.** Data are from [56]. Each diagonal panel shows the distribution of fitness effects (DFE) of knock-out mutations in the presence of the corresponding antibiotic (equivalent to Figure 1C in [56]). Scale of the y-axis in these panels is indicated inside on the right. The estimated measurement noise distributions are shown in red (see Materials and Methods for details). Note that some noise distributions are vertically cut-off for visual convenience. The number of identified beneficial mutations (i.e., resistance mutations) and the expected number of false positives (in parenthesis) are shown in the bottom left corner. Off-diagonal panels show the fitness effects of knock-out mutations across pairs of drug environments. The x-axis shows the fitness in the environment where selection would happen first (i.e., the “home” environment). Each point corresponds to an individual knock-out mutation. Resistance mutations identified in the home environment are colored according to their collateral effects, as indicated in the legend. The numbers of mutations of each type are shown in the corresponding colors in the bottom left corner of each panel. TET: tetracycline; NIT: nitrofurantoin; MEC: mecillinam; CPR: ciprofloxacin.

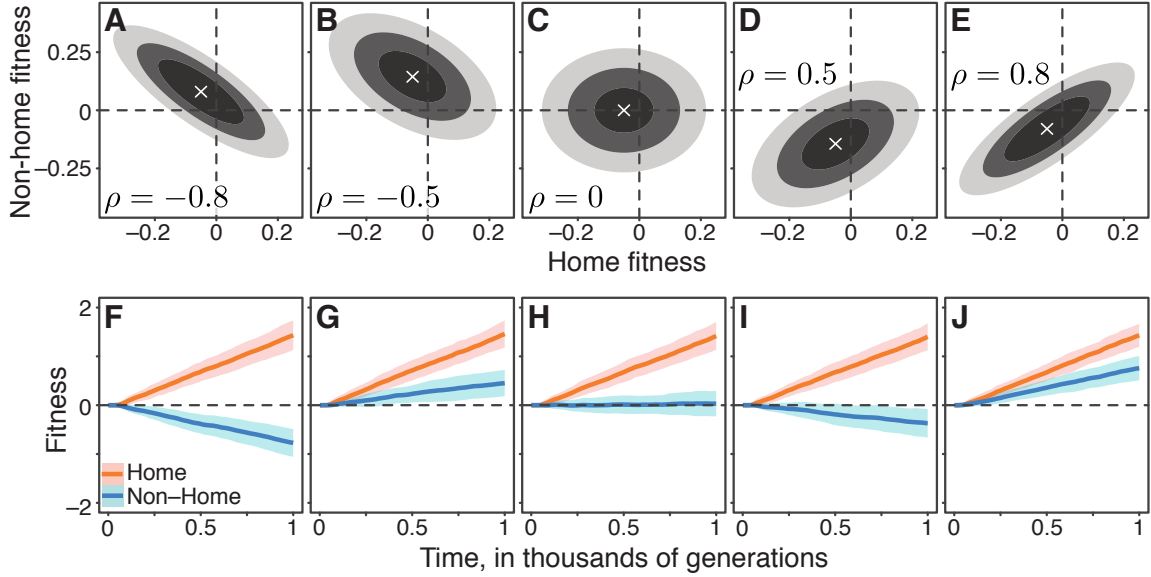
diverse collateral effects among mutations within a single gene (Figure A.1; see Materials and Methods for details).

Since both data sets represent subsets of all resistance mutations, we conclude that *E. coli* likely have access to resistance mutations with diverse pleiotropic effects, such that a fitness gain in the presence of any one drug can come either with a pleiotropic gain or a pleiotropic loss of fitness in the presence of other drugs. Therefore, the JDfE framework is suitable for modeling the evolution of collateral resistance/sensitivity. In the next section, we formally define a JDfE and probe our intuition for how its shape determines the fitness trajectories in the non-home environment.

### 1.3.2 JDfE determines the pleiotropic outcomes of adaptation

For any genotype  $g$  that finds itself in one (“home”) environment and may in the future encounter another “non-home” environment, we define the JDfE as the probability density  $\Phi_g(\Delta x, \Delta y)$  that a new mutation that arises in this genotype has the selection coefficient  $\Delta x$  in the home environment and the selection coefficient  $\Delta y$  in the non-home environment [43]. For concreteness, we define the fitness of a genotype as its malthusian parameter [62]. So, if the home and non-home fitness of genotype  $g$  are  $x$  and  $y$ , respectively, and if this genotype acquires a mutation with selection coefficients  $\Delta x$  and  $\Delta y$ , its fitness becomes  $x + \Delta x$  and  $y + \Delta y$ . This definition of the JDfE can, of course, be naturally extended to multiple non-home environments. In principle, the JDfE can vary from one genotype to another. However, to develop a basic intuition for how the JDfE determines pleiotropic outcomes, we assume that all genotypes have the same JDfE.

The JDfE is a complex object. So, we first ask whether some simple and intuitive summary statistics of the JDfE may be sufficient to predict the dynamics of the non-home fitness of a population that is adapting in the home environment. Intuitively, if there is a trade-off between home and non-home fitness, non-home fitness should decline; if the opposite is true, non-home fitness should increase. Canonically, a trade-off occurs when any mutation that



**Figure 1.2. Gaussian JDFEs and the resulting fitness trajectories.** **A–E.** Contour lines for five Gaussian JDFEs. “x” marks the mean. For all distributions, the standard deviation is 0.1 in both home- and non-home environments. The correlation coefficient  $\rho$  is shown in each panel. **F–J.** Home and non-home fitness trajectories for the JDFEs shown in the corresponding panels above. Thick lines show the mean, ribbons show  $\pm 1$  standard deviation estimated from 100 replicate simulations. Population size  $N = 10^4$ , mutation rate  $U = 10^{-4}$  ( $U_b = 4.6 \times 10^{-5}$ ).

improves fitness in one environment decreases it in the other environment and vice versa [63]. Genotypes that experience such “hard” trade-offs are at the Pareto front [64, 65]. For genotypes that are not at the Pareto front, some mutations that are beneficial in the home environment may be beneficial in the non-home environment and others may be deleterious. In this more general case, trade-offs are commonly quantified by the degree of negative correlation between the effects of mutations on fitness in the two environments [42, 63]. Thus, we might expect that evolution on negatively correlated JDFEs would lead to pleiotropic fitness losses and evolution on positively correlated JDFEs would lead to pleiotropic fitness gains.

To test this intuition, we generated a family of Gaussian JDFEs that varied, among other things, by their correlation structure (Figure 1.2; Materials and Methods). We then simulated the evolution of an asexual population on these JDFEs using a standard Wright-Fisher model (Materials and Methods) and tested whether the trade-off strength, measured by the JDFE’s

correlation coefficient, predicts the dynamics of non-home fitness. Figure 1.2 shows that our naive expectation is incorrect. Positively correlated JDFEs sometimes lead to pleiotropic fitness losses (Figure 1.2D,I), and negatively correlated JDFEs sometimes lead to pleiotropic fitness gains (Figure 1.2B,G). Even if we calculate the correlation coefficient only among mutations that are beneficial in the home environment, the pleiotropic outcomes still do not always conform to the naive expectation, as the sign of the correlation remains the same as for the full JDFEs in all these examples.

There are other properties of the JDFE that we might intuitively expect to be predictive of the pleiotropic outcomes of adaptation. For example, among the JDFEs considered in Figure 1.2, it is apparent that those with similar relative probability weights in the first and fourth quadrants produce similar pleiotropic outcomes. However, simulations with other JDFE shapes show that even distributions that are similar according to this metric can also result in qualitatively different pleiotropic outcomes (Figure A.2).

Overall, our simulations show that JDFEs with apparently similar shapes can produce qualitatively different trajectories of pleiotropic fitness changes (e.g., compare Figure 1.2A,F and Figure 1.2B,G or Figure 1.2D,I and Figure 1.2E,J). Conversely, JDFEs with apparently different shapes can result in rather similar pleiotropic outcomes (e.g., compare Figure 1.2B,G and Figure 1.2E,J or Figure 1.2A,F and Figure 1.2D,I). Thus, while the overall shape of the JDFE clearly determines the trajectory of pleiotropic fitness changes, it is not immediately obvious what features of its shape play the most important role, particularly if the JDFE is more complex than a multivariate Gaussian. In other words, even if we have perfect knowledge of the fitness effects of all mutations in multiple environments, converting this knowledge into a qualitative prediction of the expected direction of pleiotropic fitness change (gain or loss) does not appear straightforward. Therefore, we next turn to developing a population genetics model that would allow us to predict not only the direction of pleiotropic fitness change but also the expected rate of this change and the uncertainty around the expectation.

### 1.3.3 The population genetics of pleiotropy

To systematically investigate which properties of the JDFE determine the pleiotropic fitness changes in the non-home environment, we consider a population of size  $N$  that evolves on a JDFE in the “strong selection weak mutation” (SSWM) regime, also known as the “successional mutation” regime [52, 66–68].

We consider an arbitrary JDFE without epistasis, that is a situation when all genotypes have the same JDFE  $\Phi(\Delta x, \Delta y)$ .

We assume that mutations arise at rate  $U$  per individual per generation. In the SSWM limit, a mutation that arises in the population either instantaneously fixes or instantaneously dies out. Therefore, the population is essentially monomorphic at all times, such that at any time  $t$  we can characterize it by its current pair of fitness values  $(X_t, Y_t)$ . If a new mutation with a pair of selection coefficients  $(\Delta x, \Delta y)$  arises in the population at time  $t$ , it fixes with probability  $\pi(\Delta x) = \frac{1-e^{-2\Delta x}}{1-e^{-2N\Delta x}}$  [69] in which case the population’s fitness transitions to a new pair of values  $(X_t + \Delta x, Y_t + \Delta y)$ . If the mutation dies out, an event that occurs with probability  $1 - \pi(\Delta x)$ , the population’s fitness does not change. This model specifies a continuous-time two-dimensional Markov process.

In general, the dynamics of the probability density  $p(x, y, t)$  of observing the random vector  $(X_t, Y_t)$  at values  $(x, y)$  are governed by an integro-differential forward Kolmogorov equation, which is difficult to solve (Materials and Methods). However, if most mutations that contribute to adaptation have small effects, these dynamics are well approximated by a diffusion equation which can be solved exactly (Materials and Methods). Then  $p(x, y, t)$  is a normal distribution with mean vector

$$m(t) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} NU_b t \quad (1.1)$$

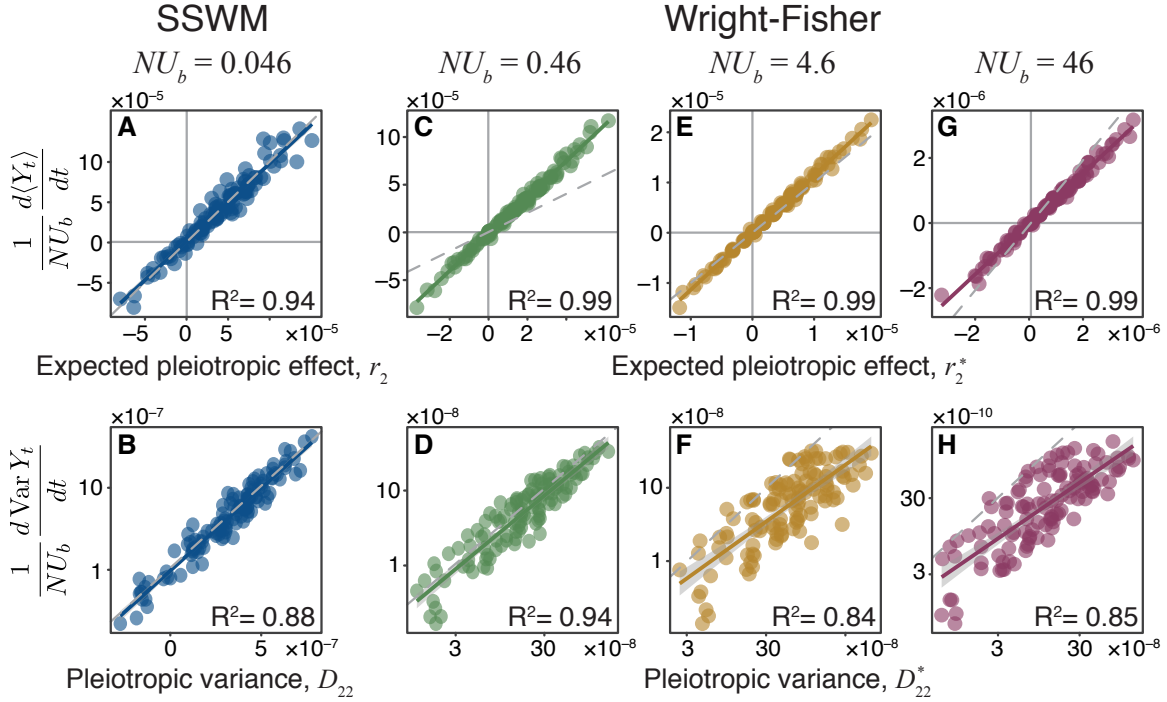
and variance-covariance matrix

$$\sigma^2(t) = \begin{pmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{pmatrix} NU_b t, \quad (1.2)$$

where  $r_1$  and  $r_2$ , given by eq 1.7 and 1.8 in Materials and Methods, are the expected fitness effects in the home and non-home environments for a mutation fixed in the home environment, and  $D_{11}$ ,  $D_{12}$  and  $D_{22}$ , given by eqs 1.9–1.11 in Materials and Methods, are the second moments of this distribution. Here,  $U_b = U \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \Phi(\xi, \eta)$  is the total rate of mutations beneficial in the home environment, and  $x_0$  and  $y_0$  are the initial values of population's fitness in the home and non-home environments.

Eq 1.1 and 1.2 show that the distribution of population's fitness at time  $t$  in the non-home environment is entirely determined by two parameters,  $r_2$  and  $D_{22}$ , which we call the pleiotropy statistics of the JDFE. The expected rate of fitness change in the non-home environment depends on the pleiotropy statistic  $r_2$ , which we refer to as the expected pleiotropic effect. Thus, evolution on a JDFE with a positive  $r_2$  is expected to result in pleiotropic fitness gains and evolution on a JDFE with a negative  $r_2$  is expected to result in pleiotropic fitness losses. Equation 1.2 shows that the variance around this expectation is determined by the pleiotropy variance statistic  $D_{22}$ . Since both the expectation and the variance change linearly with time (provided  $r_2 \neq 0$ ), the change in the non-home fitness in any replicate population will eventually have the same sign as  $r_2$ , but the time scale of such convergence depends on the “collateral resistance risk” statistic  $c = r_2 / \sqrt{D_{22}}$  (Materials and Methods). This observation has important practical implications, and we return to it in the Section “Robust ranking of drug pairs”.

These theoretical results suggest a simple explanation for the somewhat counter-intuitive observations in Figure 1.2. We may intuitively believe that evolution on negatively correlated JDFEs should lead to fitness losses in the non-home environment because on such JDFEs mutations with largest fitness benefits in the home environment typically have negative pleiotropic



**Figure 1.3. Pleiotropy statistics predict the properties of non-home fitness trajectories in simulations.** Each point corresponds to an ensemble of replicate simulation runs with the same population genetic parameters on one of 125 Gaussian JDFEs. **A.** Expected pleiotropic effect  $r_2$  versus the scaled slope of the mean rate of non-home fitness change observed in SSWM simulations. **B.** Pleiotropic variance  $D_{22}$  versus the scaled rate of change in the variance in non-home fitness observed in SSWM simulations. **C, E, G.** Expected pleiotropic effect  $r_2^*$  versus the scaled slope of the mean rate of non-home fitness change observed in Wright-Fisher simulations. **D, F, H.** Pleiotropic variance  $D_{22}^*$  versus the scaled rate of change in the variance in non-home fitness observed in Wright-Fisher simulations simulations. (See Figure A.3 for comparison between simulations and the unadjusted pleiotropy statistics  $r_2$  and  $D_{22}$ .) 1000 replicate simulations were carried out in the SSWM regime. All Wright-Fisher simulations were carried out with  $U = 10^{-4}$  and variable  $N$ , 300 replicate simulations per data point (see Materials and Methods for details). In all panels, the grey dashed line represents the identity (slope 1) line, and the solid line of the same color as the points is the linear regression for the displayed points ( $R^2$  value is shown in each panel;  $P < 2 \times 10^{-16}$  for all regressions).



effects. However, such mutations may be too rare to drive adaptation. At the same time, the more common mutations that do typically drive adaptation may have positive pleiotropic effects, in which case the population would on average gain non-home fitness, as in Figure 1.2B. Our theory shows that to predict the direction of non-home fitness change the frequency of beneficial mutations with various pleiotropic effects and the strength of these effects need to be weighted by the likelihood that these mutations fix. The expected pleiotropic effect  $r_2$  accomplishes this weighting.

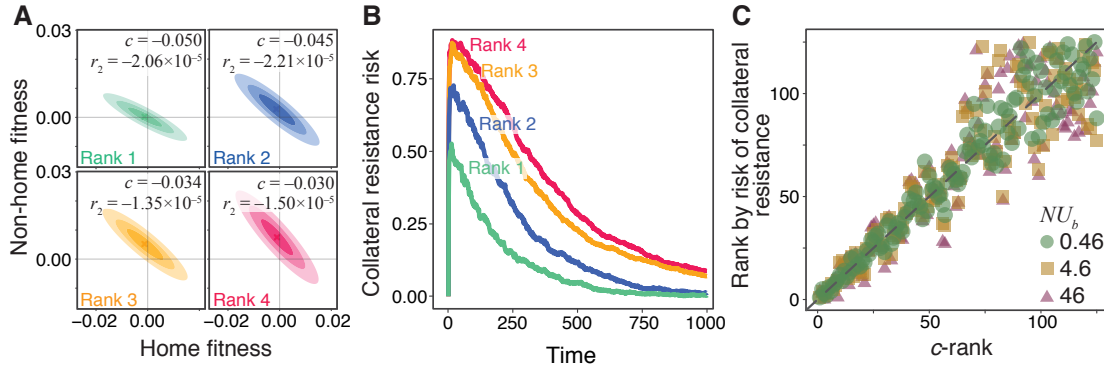
We tested the validity of equations 1.1 and 1.2 by simulating evolution in the SSWM regime on 125 Gaussian JDFEs with various parameters (Materials and Methods) and found excellent agreement 3.3 (Figure 1.3A,B). However, many microbes likely evolve in the “concurrent mutation” regime, i.e., when multiple beneficial mutations segregate in the population simultaneously [67, 70]. As expected, our theory fails to quantitatively predict the pleiotropic fitness trajectories when  $NU_b > 1$  (Figure A.3). However, the expected rate of change of non-home fitness and its variances remain surprisingly well correlated with the pleiotropy statistics  $r_2$  and  $D_{22}$  across various JDFEs (Figure A.3). In other words, we can still use these statistics to correctly predict whether a population would lose or gain fitness in the non-home environment and to order the non-home environments according to their expected pleiotropic fitness changes and variances. We will exploit the utility of such ranking in the next section.

We next sought to expand our theory to the concurrent mutation regime. A key characteristic of adaptation in this regime is that mutations whose fitness benefits in the home environment are below a certain “effective neutrality” threshold are usually outcompeted by superior mutations and therefore fix with lower probabilities than predicted by Kimura’s formula [71, 72]. Reference [72] provides an equation for calculating the fixation probability  $\pi^*(\Delta x)$  for a mutation with home fitness benefit  $\Delta x$  in the concurrent mutation regime (equation (6) in ref [72]). Thus, by replacing  $2\xi$  (the approximate fixation probability in the SSWM regime) in 1.8 and 1.8 with  $\pi^*(\xi)$ , we obtain the adjusted pleiotropy statistics  $r_2^*$  and  $D_{22}^*$  for the concurrent mutation regime (see Materials and Methods for details). Note that in contrast to  $r_2$  and  $D_{22}$ , the adjusted statistics

$r_2^*$  and  $D_{22}^*$  depend on the population genetic parameters  $N$  and  $U_b$ .

To test how well these statistics predict the dynamics of fitness in the non-home environment, we simulated evolution on the same 125 JDFEs using the full Wright-Fisher model with a range of population genetic parameters that span the transition from the successional to the concurrent mutation regimes for 1000 generations. We find that  $r_2^*$  quantitatively predicts the expected rate of non-home fitness change, with a similar accuracy as ref [72] predict the rate of fitness change in the home environment, as long as  $NU_b > 1$  (Figure 1.3C,E,G; compare with Figure A.3A,C,E).  $D_{22}^*$  also predicts the empirically observed variance in non-home fitness trajectories much better than  $D_{22}$ , although this relationship is more noisy than between mean fitness and  $r_2^*$  (Figure 1.3D,F,H; compare with Figure A.3B,D,F). Some of this noise can be attributed to sampling, as we estimate both the mean and the variance from 300 replicate simulation runs, and the variance estimation is more noisy. Even in the absence of sampling noise however, we do not expect that  $D_{22}^*$  would predict the non-home fitness variance perfectly because our theory does not account for the autocorrelation in the fitness trajectories that arise in the concurrent mutation regime but not in the successive mutation regime (see Appendix D in ref [67]). To our knowledge, a rigorous analytical calculation for ensemble variance in fitness even in the home environment is not yet available.

Overall, our theory allows us to quantitatively predict the dynamics of non-home fitness in a range of evolutionary regimes if the JDFE and the population genetic parameters  $N$  and  $U_b$  are known. However, neither the full JDFE nor the population genetic parameters will likely be known in most practical situations, such as designing a drug treatment for a cancer patient. In the next section, we address the question of how to robustly select drug pairs for a sequential treatment, assuming that the pleiotropy statistics  $r_2$  and  $D_{22}$  are known but the population genetic parameters are not. In the Section “Measuring JDFEs” we provide some guidance on how the JDFE can be measured.



**Figure 1.4. Robust ranking of drug pairs.** **A.** Four hypothetical JDFEs, ranked by their  $c$  statistic. For all four JDFEs, the mean and the standard deviation in the home environment are  $-1 \times 10^{-3}$  and 0.01, respectively. The mean and the standard deviation in the non-home environment are  $1 \times 10^{-4}$  and  $5.1 \times 10^{-3}$  (rank 1),  $2.6 \times 10^{-3}$  and  $7.5 \times 10^{-3}$  (rank 2),  $5.1 \times 10^{-3}$  and  $7.5 \times 10^{-3}$  (rank 3),  $7.5 \times 10^{-3}$  and 0.01 (rank 4). Correlation coefficient for all four JDFEs is  $-0.9$ . **B.** Collateral resistance risk over time, measured as the fraction of populations with positive mean fitness in the non-home environment. These fractions are estimated from 1000 replicate Wright-Fisher simulation runs with  $N = 10^4$ ,  $U = 10^{-4}$  ( $NU_b = 0.46$ ). Colors correspond to the JDFEs in panel A. Numbers indicate the  $c$ -rank of each JDFE. **C.** A *priori*  $c$ -rank (x-axis) versus the *a posteriori* rank (y-axis) based on the risk of collateral resistance observed in simulations, for all 125 Gaussian JDFEs and all  $NU_b$  values shown in Figure 1.3. Grey dashed line is the identity line.  $R^2$  values are 0.94, 0.81 and 0.82 for  $NU_b = 0.46, 4.6$  and 46, respectively.  $P < 10^{-15}$  for all regressions.

### 1.3.4 Robust ranking of drug pairs

Consider a hypothetical scenario where a drug treatment is being designed for a patient with a tumor or a bacterial infection. In selecting a drug, it is desirable to take into account not only the standard medical considerations, such as drug availability, toxicity, etc., but also the possibility that the treatment with this drug will fail due to the evolution of resistance. Therefore, it may be prudent to consider a list of drugs pairs (or higher-order combinations), ranked by the propensity of the first drug in the pair to elicit collateral resistance against the second drug in the pair. All else being equal, the drug deployed first should form a high-ranking pair with at least one other secondary drug. Then, if the treatment with the first drug fails, a second one can be deployed with a minimal risk of collateral resistance. Thus, we set out to develop a metric for ranking drug pairs according to this risk.

Clearly, any drug pair with a negative  $r_2$  is preferable over any drug pair with a positive  $r_2$ , since the evolution in the presence of the first drug in a pair with  $r_2 < 0$  is expected to elicit collateral sensitivity against the second drug in the pair but the opposite is true for drug pairs with  $r_2 > 0$ . It is also clear that among two drug pairs with negative  $r_2$ , a pair with a more negative  $r_2$  and lower  $D_{22}$  is preferable over a pair with a less negative  $r_2$  and higher  $D_{22}$  because evolution in the presence of the first drug in the former pair will more reliably lead to stronger collateral sensitivity against the second drug in the pair. The difficulty is in how to compare and rank two drug pairs where one pair has a more negative  $r_2$  but higher  $D_{22}$ . Our theory shows that the chance of emergence of collateral resistance monotonically increases with the collateral risk statistic  $c = r_2/\sqrt{D_{22}}$  (see Materials and Methods). Thus, we propose to rank drug pairs by  $c$  from lowest (most negative and therefore most preferred) to highest (least negative or most positive and therefore least preferred).

To demonstrate the utility of such ranking, consider four hypothetical drug pairs with JDfEs shown in Figure 1.4A. The similarity between their shapes makes it difficult to predict *a priori* which one would have the lowest and highest probabilities of collateral resistance. Thus,

we rank these JDFEs by their  $c$  statistic. To test whether this ranking is accurate with respect to the risk of collateral resistance, we simulate the evolution of a Wright-Fisher population in the presence of the first drug in each pair for 600 generations and estimate the probability that the evolved population has a positive fitness in the presence of the second drug, i.e., the probability that it becomes collaterally resistant (Figure 1.4B). We find that our *a priori* ranking corresponds perfectly to the ranking according to this probability, evidenced by the consistently higher collateral resistance risk for JDFEs with higher  $c$  (Figure 1.4B). Interestingly, the top ranked JDFE does not have the lowest expected pleiotropic effect  $r_2$ . Nevertheless, the fact that the pleiotropic variance statistic  $D_{22}$  for this JDFE is small ensures that the risk of collateral resistance evolution is the lowest. This 1 to 1 rank correlation holds more broadly, for all 125 Gaussian JDFEs and all population genetic parameters considered in the previous section (Figure 1.4C) as well as for the empirical TEM  $\beta$ -lactamase JDFEs (Figure A.4). Overall, we find that we can use the collateral risk statistic  $c$  to robustly rank drug pairs according to the risk of collateral resistance evolution, irrespective of the population genetic parameters.

### 1.3.5 Measuring JDFEs

So far, we assumed that the parameters of the JDFE on which the population evolves are known. In reality, they have to be estimated from data, which opens up at least two practically important questions. The first question is experimental. From what types of data can JDFEs be in principle estimated and how good are different types of data for this purpose? We can imagine, for example, that some properties of JDFEs can be estimated from genome sequencing data [29] or from temporally resolved fitness trajectories [73]. Here we focus on the most direct way of estimating JDFE parameters, from the measurements of the home and non-home fitness effects of individual mutations. The experimental challenge with this approach is to sample those mutations that will most likely contribute to adaptation in the home environment (see “Discussion” for an extended discussion of this problem). Below, we propose two potential strategies for such sampling: the Luria-Delbrück (LD) method and the barcode lineage tracking

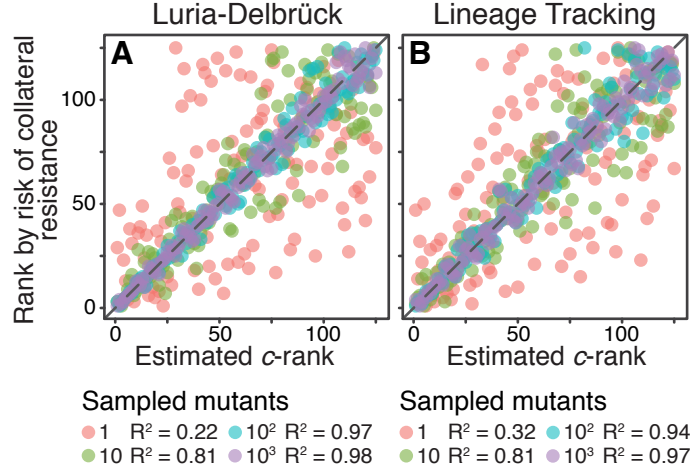
(BLT) method. The second question is statistical: how many mutants need to be sampled to reliably rank drug pairs according to the risk of collateral resistance? We evaluate both proposed methods with respect to this property.

The idea behind the LD method is to expose the population to a given drug at a concentration above the minimum inhibitory concentration (MIC), so that only resistant mutants survive [25]. This selection is usually done on agar plates, so that individual resistant mutants form colonies and can be isolated. The LD method is relatively easy to implement experimentally, but it is expected to work only if the drug concentration is high enough to kill almost all non-resistant cells. In reality, resistant mutants may be selected at concentrations much lower than MIC [74]. Furthermore, mutants selected at different drug concentrations may be genetically and functionally distinct [25, 75] and therefore may have statistically different pleiotropic profiles. As a result, mutants sampled with the LD method may not be most relevant for predicting collateral evolution at low drug concentrations, and other sampling methods may be required for isolating weakly beneficial mutations.

Isolating individual weakly beneficial mutations is more difficult because by the time a mutant reaches a detectable frequency in the population it has accumulated multiple additional driver and passenger mutations [70, 76], all of which can potentially have collateral effects. One way to isolate many single beneficial mutations from experimental populations is by using the recently developed barcode lineage tracking (BLT) method [53, 77]. In a BLT experiment, each cell is initially tagged with a unique DNA barcode. As long as there is no recombination or other DNA exchange, any new mutation is permanently linked to one barcode. A new adaptive mutation causes the frequency of the linked barcode to grow, which can be detected by sequencing. By sampling many random mutants and genotyping them at the barcode locus, one can identify mutants from adapted lineages even if they are rare [77]. As a result, BLT allows one to sample mutants soon after they acquire their first driver mutation, before acquiring secondary mutations.

To evaluate the quality of sampling based on the LD and BLT methods, we consider

the following hypothetical experimental setup.  $K$  beneficial mutants are sampled from each home environment (with either one of the methods), and their home and non-home fitness ( $X_i, Y_i$ ) are measured for each mutant  $i = 1, \dots, K$ . Since we are ultimately interested in ranking drug pairs by their risk of collateral resistance, we estimate the collateral risk statistic  $\hat{c}$  from these fitness data for each drug pair and use  $\hat{c}$  to rank them (see Materials and Methods for details). We compare such *a priori* ranking of 125 hypothetical drug pairs with Gaussian JDFEs used in previous sections with their *a posteriori* ranking based on the risk of collateral resistance observed in simulations.



**Figure 1.5. Sampling effects on the ranking of drug pairs.** Both panels show correlations between the *a priori* estimated  $c$ -rank ( $x$ -axis) of the 125 Gaussian JDFEs and their *a posteriori* rank ( $y$ -axis) based on the risk of collateral resistance observed in simulations (same data as the  $y$ -axis in Figure 1.4C for  $NU_b = 0.46$ ). **A.** The  $c$  statistic is estimated using the Luria-Delbrück method (see text for details). Cutoff for sampling mutations is  $0.5\sigma$ , where  $\sigma$  is the standard deviation of the JDFE in the home environment. See Figure A.5 for other cutoff values. **B.** The  $c$  statistic is estimated using the barcode lineage tracking method with  $N = 10^6$  and  $U = 10^{-4}$  (see text and Materials and Methods for details).  $P < 10^{-6}$  for all regressions.

To model the LD sampling method on a given JDFE, we randomly sample  $K$  mutants whose home fitness exceeds a certain cutoff. To model a BLT experiment, we simulate evolution in the home environment and randomly sample  $K$  beneficial mutants segregating at generation 250 (see Materials and Methods for details). We find that the  $\hat{c}$ -ranking estimated with either LD or BLT methods captures the *a posteriori* ranking surprisingly well, even when the number of

sampled mutants is as low as 10 per drug pair (Figure 1.5). Given that the JDFEs with adjacent ranks differ in  $c$  by a median of only 0.65%, the strong correlations shown in Figure 1.5 suggest that even very similar JDFEs can be differentiated with moderate sample sizes. As expected, this correlation is further improved upon increased sampling, and it is insensitive to the specific home fitness threshold that we use in the LD method (Figure A.5). We conclude that estimating JDFE parameters is in principle feasible with a modest experimental effort, at least for the purpose of ranking drug pairs.

## 1.4 Discussion

We have shown that many resistance mutations against multiple drugs in *E. coli* exhibit a diversity of collateral effects. If this is true more generally, it implies that there is an unavoidable uncertainty in whether any given population would evolve collateral resistance or sensitivity, which could at least in part explain previously observed inconsistencies among experiments. We quantified the diversity of pleiotropic effects of mutations with a joint distribution of fitness effects (JDFE) and developed a population genetic theory for predicting the expected collateral outcomes of evolution and the uncertainty around these expectations. In the successional mutations regime, our theory shows that the average rate at which fitness in the non-home environment is gained or lost during adaptation to the home environment is determined by the pleiotropy statistic  $r_2$  given by equation 1.8. How strongly the non-home fitness in any individual population deviates from this ensemble average is determined by the pleiotropy variance statistic  $D_{22}$  given by equation 1.11. Importantly,  $r_2$  and  $D_{22}$  are properties of the JDFE alone, i.e., they do not depend on the parameters of any specific population. In the concurrent mutations regime, the expected rate of non-home fitness gain or loss and the associated variance are reasonably well predicted by the adjusted pleiotropy statistics  $r_2^*$  and  $D_{22}^*$ . Unlike  $r_2$  and  $D_{22}$ , the adjusted statistics depend on the population size  $N$  and the rate of beneficial mutations  $U_b$ .

To quantitatively predict the probability of evolution of collateral drug resistance in



practice would require the knowledge of both the JDFE for the focal bacterial or cancer-cell population in the presence of the specific pair of drugs and its *in vivo* population genetic parameters. Since estimating the latter parameters is very difficult, it appears unlikely that we would be able to quantitatively predict the dynamics of collateral effects, even if JDFEs were known. A more realistic application of our theory is that it allows us to rank drug pairs according to the risk of collateral resistance even when the population genetic parameters are unknown. Such robust ranking can be computed based on the collateral risk statistic  $c = r_2/\sqrt{D_{22}}$ , a property of the JDFE but not of the evolving population. Drug pairs with positive values of  $c$  have a higher chance of eliciting collateral resistance than collateral sensitivity and should be avoided; drug pairs with more negative values of  $c$  have a lower risk of collateral resistance evolution than those with less negative values.

We have validated our theory *in silico*, but how well it would work *in vivo* (in the clinic) or even *in vitro* (in the lab) is as of yet unclear. A direct way to validate the theory empirically would be to estimate JDFEs for a model organism, such as *E. coli*, in a number of drug pairs, rank these pairs according to our collateral rank statistic and then test this ranking by evolving replicate populations and measuring the empirical distributions of collateral resistance/sensitivity outcomes. To the best of our knowledge, the antibiotic resistance JDFEs among genome-wide mutations have not yet been measured. One could in principle use existing gene knock-out data, such as those obtained by ref [56] (Figure ??), or the data from deep mutational scanning experiments, such as those obtained by ref [61] (Figure A.1), to estimate JDFEs. However, these experiments estimate fitness only for certain subsets of mutations (gene knock-outs or point mutations within a single gene, respectively). Since resistance may arise via other types of mutations [28], these data would give us at best an incomplete picture of actual JDFEs. Our results suggest that JDFEs can be reasonably well estimated by sampling resistance mutants at drug concentrations above MIC or by employing the barcode lineage tracking method.

Another obstacle is that, even though many researchers have experimentally evolved various microbes in the presence of drugs, most experiments have maintained too few replicate

populations to accurately measure the variation in collateral outcomes of evolution. The study by ref [28], with 60 replicates, is a notable exception. In short, a rigorous test of our theory requires new data on the shapes of whole-genome JDFEs as well as higher throughput evolution experiments.

What the most effective ways of measuring JDFEs are and whether it will be possible to measure JDFE *in vivo* are open questions. We speculate that the answers will depend on the shapes of the empirical JDFEs because some shapes may be more difficult to estimate than others. For example, if empirical JDFEs resemble multivariate Gaussian distributions, then we can learn all relevant parameters of such JDFE by sampling a handful of random mutants and measuring their fitness in relevant environments. One can also imagine more complex JDFEs where mutations beneficial in the home environment have a dramatically different distribution of non-home fitness effects than mutations that are deleterious or neutral in the home environment. In this case, very large samples of random mutations would be necessary to correctly predict the pleiotropic outcomes of evolution, so that methods that preferentially sample beneficial mutations may be required. We have considered two such methods, which are experimentally feasible. We have shown that both of them perform extremely well on Gaussian JDFEs in the sense that as few as 10 mutants per drug pair are sufficient to produce largely correct ranking of hypothetical drug pairs. However, it may be difficult to apply these methods *in vivo*, in which case JDFEs may have to be estimated in the lab, with selection pressures reproducing those *in vivo* as accurately as possible.

Our model relies on two important simplifications. It describes the evolution of an asexual population where all resistance alleles arise from *de novo* mutations. In reality, some resistance alleles in bacteria may be transferred horizontally [78]. Understanding collateral resistance evolution in the presence of horizontal gene transfer events would require incorporating JDFE into other models of evolutionary dynamics [e.g., 79]. Another major simplification is in the assumption that the JDFE stays constant as the population adapts. In reality the JDFE will change over time because of the depletion of the pool of adaptive mutations and because of epistasis [59,

60]. How JDFEs vary among genetic backgrounds is currently unknown. Empirically measuring how JDFEs vary across genotypes and theoretically understanding how such variation affects the evolution of pleiotropic outcomes are important open questions.

While we were primarily motivated by the problem of evolution of collateral drug resistance and sensitivity, our theory is applicable more broadly. The shape of JDFE must play a crucial role in determining whether the population evolves towards a generalist or diversifies into multiple specialist ecotypes. Previous literature has viewed this question primarily through the lens of two alternative hypotheses: antagonistic pleiotropy and mutation accumulation [80]. Antagonistic pleiotropy in its strictest sense means that the population is at the Pareto front with respect to the home and non-home fitness, such that any mutation beneficial in the home environment reduces the fitness in the non-home environment [65]. The shape of the Pareto front then determines whether selection would favor specialists or generalists [80, 81]. Alternatively, a population can evolve to become a home-environment specialist even in the absence of trade-offs, simply by accumulating mutations that are neutral in the home environment but deleterious in the non-home environment [82]. More recently, it has been recognized that antagonistic pleiotropy and mutation accumulation are not discrete alternatives but rather extremes of a continuum of models [29, 43, 83]. The JDFE provides a mathematical way to describe this continuum. For example, strict antagonistic pleiotropy can be modeled with a JDFE with zero probability weight in the first quadrant and a bulk of probability in the fourth quadrant. A mutation accumulation scenario can be modeled with a “+”-like JDFE where all mutations beneficial in the home environment are neutral in the non-home environment (i.e., concentrated on the  $x$ -axis) and all or most mutations neutral in the home environment (i.e., those on the  $y$ -axis) are deleterious in the non-home environment. Our theory shows that in fact all JDFEs with negative  $r_2$  lead to loss of fitness in the non-home environment and therefore can potentially promote specialization. While our theory provides this insight, further work is needed to understand how JDFEs govern adaptation to variable environments. This future theoretical work, together with empirical inquiries into the shapes of JDFEs, will not only advance our ability to predict evolution in

practical situations, such as drug resistance, but it will also help us better understand the origins of ecological diversity.

## 1.5 Materials and Methods

### 1.5.1 Analysis of knock-out and deep mutational scanning data

#### Knock-out data

Reference [56] provides growth rate estimates for 3883 gene knock-out mutants of *E. coli* in the presence of six antibiotics. Our goal is to identify those knock-out mutations that provide resistance against one drug and are also collaterally resistant or collaterally sensitive to another drug. However, it is unclear from these original data alone which mutations have statistically significant beneficial and deleterious effects because no measurement noise estimates are provided. To address this problem, we obtained replicate wildtype growth rate measurements in the presence of antibiotics from Guillaume Chevereau and Tobias Bollenbach (available at <https://github.com/ardellsarah/JDFE-project>). In this additional data set, the wildtype *E. coli* strain is measured on average 476 times in the presence of each drug. We estimate the wildtype growth rate  $r_{WT}$  as the mean of these measurements, and we obtain the selection coefficient for all knock-out mutants as  $s_i = r_i - r_{WT}$ . We also obtain the noise distribution  $P_{\text{noise}}(s)$  from the replicate wildtype measurements (shown in red in the diagonal panels in Figure ??). Modeling  $P_{\text{noise}}(s)$  as normal distributions, we obtain the  $P$ -values for each mutation in the presence of each antibiotic.

We then call any knock-out mutant as resistant against a given drug if its selection coefficient in the presence of that drug exceeds a critical value  $s_{\alpha}^{+} > 0$ . We choose  $s_{\alpha}^{+}$  using the Benjamini-Hochberg procedure [84] so that the false discovery rate (FDR) among the identified resistant mutants is  $\alpha \approx 0.25$ . We could not find an  $s_{\alpha}^{+}$  for  $\alpha < 0.25$  for trimethoprim (TMP) and chloramphenicol (CHL), i.e., there were not enough knock-out mutations with positive selection coefficients to reliably distinguish them from measurement errors.

We apply the same procedure to identify mutations that are collaterally resistant and

collaterally sensitive against a second drug among all mutations that are resistant against the first drug, except we aim for  $\text{FDR} < 0.10$ .

### **Deep mutational scanning data**

Reference [61] provide estimates of relative fitness for 4997 point mutations in the TEM-1  $\beta$ -lactamase gene in the presence of cefotaxime (CTX) and four concentrations of ampicillin (AMP). We estimate the selection coefficients from the reported relative fitness values by changing the logarithm from  $\log_{10}$  to natural and dividing it by six, the estimated number of generations that occurred during the 2-hour experiment. The latter is based on the fact that ref [61] chose AMP concentrations which did not significantly alter the *E. coli* doubling time, which we assumed to be 20 minutes. We used the same number of generations for CTX.

Reference [61] reports two replicate measurements per mutant in each concentration of AMP and one measurement per mutant in the presence of CTX. We consider CTX as the home environment and call all mutations with positive measured fitness effects as resistant against CTX. For each such mutation, we use two replicate measurements in each concentration of AMP to estimate its mean fitness effect and the 90% confidence interval around the mean, based on the normal distribution. We call any CTX resistant mutation with the entire confidence interval above (below) zero as collaterally resistant (sensitive) against AMP at that concentration. All remaining CTX resistant mutations are called collaterally neutral.

## **1.5.2 Theory**

### **Successional mutations regime**

We assume that an asexual population evolves according the Wright-Fisher model in the strong selection weak mutation (SSWM) limit [52, 66, 68], also known as the “successional mutations” regime [67]. In this regime, the population remains monomorphic until the arrival of a new mutation that is destined to fix. The waiting time for such new mutation is assumed to be much longer than the time it takes for the mutation to fix, i.e., fixation happens almost instantaneously on this time scale, after which point the population is again monomorphic. If

the per genome per generation rate of beneficial mutations is  $U_b$ , their typical effect is  $s$  and the population size is  $N$ , the SSWM approximation holds when  $NU_b \ll 1/\ln(Ns)$  [67].

We describe our population by a two-dimensional vector of random variables  $(X_t, Y_t)$ , where  $X_t$  and  $Y_t$  are the population's fitness (growth rate or the Malthusian parameter) in the home and non-home environments at generation  $t$ , respectively. We assume that the fitness vector of the population at the initial time point is known and is  $(x_0, y_0)$ . We are interested in characterizing the joint probability density  $p(x, y, t) dx dy = \Pr\{X_t \in [x, x + dx), Y_t \in [y, y + dy)\}$ .

We assume that all genotypes have the same JDFF  $\Phi(\Delta x, \Delta y)$ , i.e., there is no epistasis. In the exponential growth model, the selection coefficient of a mutation is the difference between the mutant and the ancestor growth rates in the home environment, i.e.,  $\Delta x$ . The probability of fixation of the mutant is given by Kimura's formula, which we approximate by  $2\Delta x$  for  $\Delta x > 0$  and zero otherwise [62].

If the total rate of mutations (per genome per generation) is  $U$ , the rate of mutations beneficial in the home environment is given by  $U_b = U f_b$  where  $f_b = \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \Phi(\xi, \eta)$  is the fraction of mutations beneficial in the home environment. Once such a mutation arises, its selection coefficients in the home and non-home environments are drawn from the JDFF of mutations beneficial in the home environment  $\Phi_b(\Delta x, \Delta y) = \Phi(\Delta x, \Delta y)/f_b$ . Then, in the SSWM limit, our population is described by a two-dimensional continuous-time continuous-space Markov chain with the transition rate from state  $(x, y)$  to state  $(x', y')$  given by

$$2NU_b Q(x', y'|x, y) = \begin{cases} 2NU_b (x' - x) \Phi_b(x' - x, y' - y) & \text{if } x' > x, \\ 0 & \text{otherwise.} \end{cases} \quad (1.3)$$

The probability distribution  $p(x, y, t)$  satisfies the integro-differential forward Kolmogorov equation

tion [85]

$$\begin{aligned} & \frac{1}{NU_b} \frac{\partial p}{\partial t}(x, y, t) \\ = & 2 \int_{-\infty}^{\infty} d\eta \int_{-\infty}^{\infty} d\xi \left( p(\xi, \eta, t) Q(x, y | \xi, \eta) - p(x, y, t) Q(\xi, \eta | x, y) \right) \end{aligned} \quad (1.4)$$

with the initial condition

$$p(x, y, 0) = \delta(x - x_0) \delta(y - y_0). \quad (1.5)$$

When beneficial mutations with large effects are sufficiently rare, equation (1.4) can be approximated by the Fokker-Planck equation [85]

$$\frac{1}{NU_b} \frac{\partial p}{\partial t} = -r_1 \frac{\partial p}{\partial x} - r_2 \frac{\partial p}{\partial y} + \frac{D_{11}}{2} \frac{\partial^2 p}{\partial x^2} + D_{12} \frac{\partial^2 p}{\partial x \partial y} + \frac{D_{22}}{2} \frac{\partial^2 p}{\partial y^2}, \quad (1.6)$$

where

$$r_1 = 2 \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \xi^2 \Phi_b(\xi, \eta), \quad (1.7)$$

$$r_2 = 2 \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \eta \xi \Phi_b(\xi, \eta) \quad (1.8)$$

are the expected fitness effects in the home and non-home environments for a mutation fixed in the home environment, and

$$D_{11} = 2 \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \xi^3 \Phi_b(\xi, \eta), \quad (1.9)$$

$$D_{12} = 2 \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \eta \xi^2 \Phi_b(\xi, \eta), \quad (1.10)$$

$$D_{22} = 2 \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \eta^2 \xi \Phi_b(\xi, \eta) \quad (1.11)$$

are the second moments of the distribution of the fitness effects of mutations fixed in the home environment. The solution to equation 1.6 with the initial condition equation 1.5 is a multi-variate normal distribution with the mean vector  $m(t)$  and the variance-covariance matrix  $\sigma^2(t)$  given

by 1.1 and 1.2.

### Concurrent mutations regime

The theory we developed so far for the successional mutations regime breaks down in the concurrent mutations regime, i.e., when multiple adaptive mutations segregate in the population simultaneously [67]. The main effect of competition between segregating adaptive lineages is that many new beneficial mutations arise in relatively low-fitness genetic backgrounds and have almost no chance of surviving competition [67, 71, 72]. As a result, the fixation probability of a beneficial mutation with selective effect  $\Delta x$  in the home environment is no longer  $2\Delta x$ . Instead, beneficial mutations that provide fitness benefits below a certain threshold  $x_c$  behave as if they are effectively neutral (i.e., their fixation probability is close to zero), and most adaptation is driven by mutations with benefits above  $x_c$ , where  $x_c$  depends on the population genetic parameters  $N$  and  $U_b$  as well as the shape of the distribution of fitness effects of beneficial mutations. Reference [72] derived equations that allow us to calculate the effective fixation probability  $\pi^*(\Delta x; N, U_b)$  of a beneficial mutation with the fitness benefit  $\Delta x$  in the home environment in the concurrent mutation regime. Thus, to predict the average rate of non-home fitness change, we replace the SSWM fixation probability  $2\xi$  in equation 1.8 with  $\pi^*(\xi; N, U_b)$  and obtain the adjusted expected pleiotropic effect

$$r_2^*(N, U_b) = \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \eta \pi^*(\xi; N, U_b) \Phi_b(\xi, \eta). \quad (1.12)$$

We similarly obtain the adjusted pleiotropic variance statistic

$$D_{22}^*(N, U_b) = \int_{-\infty}^{\infty} d\eta \int_0^{\infty} d\xi \eta^2 \pi^*(\xi; N, U_b) \Phi_b(\xi, \eta), \quad (1.13)$$

although, as discussed in Section “The population genetics of pleiotropy”, we do not expect  $D_{22}^*$  to capture all of the variation in non-home fitness trajectories.

To calculate  $\pi^*(\Delta x; N, U_b)$  for the Gaussian JDFEs shown in Figure 1.2, we first substitute



equation (20) in ref [72] with  $\beta = 2$  into equations (18), (19) in ref [72] and then numerically solve these equations for  $x_c$  and  $v$  using the FindRoot numerical method in Mathematica. Note that all our Gaussian JDFEs share the same mean and variance in the home environment, so we need to solve these equations only once for each pair of  $N$  and  $U_b$  values. We then substitute the obtained values of  $x_c$  and  $v$  into equations (4) and (9) in ref [72] and calculate  $\pi^*$  by a numerical integration of equation (6) in ref [72] in R (available at <https://github.com/ardellsarah/JDFE-project>).

### 1.5.3 Ranking of drug pairs

According to equation 1.1 and 1.2, both the expected non-home fitness and its variance change linearly with time, so that at time  $t$  the mean is  $Z = c\sqrt{NU_b t}$  standard deviations above  $y_0$  (if  $r_2 > 0$ ) or below  $y_0$  (if  $r_2 < 0$ ), where  $c = r_2/\sqrt{D_{22}}$ . In other words, if  $r_2 > 0$ , the bulk of the non-home fitness distribution eventually shifts above  $y_0$ , and if  $r_2 < 0$ , it shifts below  $y_0$ . All else being equal, a larger value of  $|c|$  implies faster rate of this shift.

The interpretation of these observations in terms of collateral resistance/sensitivity is that adaptation in the presence of the first drug will eventually lead to collateral resistance against the second drug if  $r_2 > 0$  and to collateral sensitivity if  $r_2 < 0$ . Furthermore, all else being equal, collateral sensitivity evolves faster and the chance of evolving collateral resistance is smaller for drug pairs with more negative  $c$  (i.e., larger  $|c|$ ). Thus, we use  $c$  to order drug pairs from the most preferred (those with the most negative values of  $c$ ) to least preferred (those with least negative or positive values of  $c$ ).

### 1.5.4 Generation of JDFEs

#### Gaussian JDFEs

The JDFEs in Figure 1.2 have the following parameters. Mean in the home environment:  $-0.05$ . Standard deviation in both home and non-home environments:  $0.1$ . Means in the non-home environment:  $0.08, 0.145, 0, -0.145, -0.08$  in panels A through E, respectively.

The JDFEs in Figure 1.3 have the following parameters. Mean and standard deviation

in the home environment:  $-0.001$  and  $0.01$ , respectively. The non-home mean varies between  $0.0001$  and  $0.01$ . The non-home standard deviation varies between  $0.0001$  and  $0.01$ . The correlation between home and non-home fitness varies between  $-0.9$  and  $0.9$ , for a total of 125 JDFEs.

### **JDFEs with equal probabilities of pleiotropically beneficial and deleterious mutations**

All JDFEs in Figure A.2 are mixtures of two two-dimensional uncorrelated Gaussian distributions, which have the following parameters. Mean in the home environment:  $0.4$ . Standard deviation in both home and non-home environments:  $0.1$ . Means in the non-home environment:  $0.1$  and  $-0.1$  in panel A,  $0.5$  and  $-0.5$  in panel B,  $0.17$  and  $-0.5$  in panel C, and  $0.5$  and  $-0.17$  in panel D.

## **1.5.5 Simulations**

We carried out two types of simulations, SSWM model simulations and full Wright-Fisher model simulations.

### **Strong selection weak mutation**

The SSWM simulations were carried out using the Gillespie algorithm [86], as follows. We initiate the populations with home and non-home fitness values  $x_0 = 0$  and  $y_0 = 0$ . At each iteration, we draw the waiting time until the appearance of the next beneficial mutation from the exponential distribution with the rate parameter  $NU_b$  and advance the time by this amount. Then, we draw the selection coefficients  $\Delta x$  and  $\Delta y$  of this mutation in the home- and non-home environment, respectively, from the JDFE (a multivariate normal distribution). With probability  $2\Delta x$ , the mutation fixes in the population. If it does, the fitness of the population is updated accordingly.

### **Wright-Fisher model**

We simulate evolution in the home environment according to the Wright-Fisher model with population size  $N$  as follows. We initiate the whole population with a single genotype with

fitness  $x_0 = 0$  and  $y_0 = 0$  in the home and non-home environments. Suppose that at generation  $t$ , there are  $K(t)$  genotypes, such that genotype  $i$  has home- and non-home fitness  $X_i$  and  $Y_i$ , respectively, and it is present at frequency  $f_i(t) > 0$  in the population. We generate the genotype frequencies at generation  $t + 1$  in three steps. In the reproduction step, we draw random numbers  $B'_i(t + 1)$ ,  $i = 1, \dots, K(t)$  from the multinomial distribution with the number of trials  $N$  and success probabilities  $p_i(t) = f_i(t) + f_i(t) (X_i(t) - \bar{X}(t))$ , where  $\bar{X}(t) = \sum_{i=1}^{K(t)} X_i(t) f_i(t)$  is the mean fitness of the population in the home environment at generation  $t$ . In the mutation step, we draw a random number  $M$  of new mutants from the Poisson distribution with parameter  $NU$ , where  $U$  is the total per individual per generation mutation rate. We randomly determine the “parent” genotypes in which each mutation occurs and turn the appropriate numbers of parent individuals into new mutants. We assume that each new mutation creates a new genotype and has fitness effects  $\Delta x$  and  $\Delta y$  in the home and non-home environments.  $\Delta x$  and  $\Delta y$  are drawn randomly from the JDfE  $\Phi(\Delta x, \Delta y)$ . We obtain each mutants fitness by adding these values to the parent genotype’s home and non-home fitness values. In the final step, all genotypes that are represented by zero individuals are removed and we are left with  $K(t + 1)$  genotypes with  $B_i(t + 1) > 0$ ,  $i = 1, \dots, K(t + 1)$  individuals. Then we set  $f_i(t + 1) = B_i(t + 1)/N$ .

### 1.5.6 Sampling beneficial mutants from JDfEs and estimating the collateral risk statistic

We model the LD sampling method by randomly drawing mutants from the JDfE until the desired number  $K$  of mutants whose home fitness exceeds the focal threshold are sampled. We estimate the  $c$  statistic from the pairs of home and non-home fitness effects  $X_i$  and  $Y_i$  of these  $i = 1, \dots, K$  sampled mutants. To do so, we first estimate  $r_2$  and  $D_{22}$  as  $\hat{r}_2 = 1/K \sum_{i=1}^K X_i Y_i$  and  $\hat{D}_{22} = 1/K \sum_{i=1}^K X_i Y_i^2$ . We then calculate  $\hat{c} = \hat{r}_2 / \sqrt{\hat{D}_{22}}$ .

For the BLT sampling method, we simulate the Wright-Fisher model as described above for  $N = 10^6$  and  $U = 10^{-4}$  for 250 generations. At generation 250, we randomly sample existing beneficial mutants proportional to their frequency in the population without replacement (i.e., the

same beneficial mutation is sampled at most once). Sampling more than  $\sim 50$  distinct beneficial mutants from a single population becomes difficult because there may simply be not enough such mutants or some of them may be at very low frequencies. Therefore, if the desired number of mutants to sample exceeds 50, we run multiple replicate simulations and sample a maximum of 100 distinct beneficial mutants per replicate until the desired number of mutants is reached. We then estimate the  $c$  statistics as with the LD method.

## 1.6 Acknowledgements

We thank Shea Summers and Flora Tang for help and input at the initial stages of the project and the Kryazhimskiy and Meyer labs for feedback. We thank Tobias Bollenbach and Guillaume Chevereau for providing wildtype growth rate measurement data. This work was supported by the BWF Career Award at Scientific Interface (Grant 1010719.01), the Alfred P. Sloan Foundation (Grant FG-2017-9227), the Hellman Foundation and NIH (Grants 1R01GM137112 and 1T32GM133351-01).

Chapter 1, in full, is a reprint of the material as it appears in The population genetics of collateral resistance and sensitivity in *eLife* 2021. Ardell, Sarah and Kryazhimskiy, Sergey. The dissertation author was the primary investigator and author of this paper.

## Chapter 2

# Environment-independent distribution of mutational effects emerges from microscopic global epistasis

### 2.1 Abstract

Predicting how new mutations alter phenotypes is difficult because mutational effects vary across genotypes and environments. Recently discovered global epistasis, where the fitness effects of mutations scale with the fitness of the background genotype, can improve predictions, but how the environment modulates this scaling is unknown. We measured the fitness effects of  $\sim 100$  insertion mutations in 42 strains of *Saccharomyces cerevisiae* in six laboratory environments and found that the global-epistasis scaling is nearly invariant across environments. Instead, the environment tunes one global parameter, the background fitness at which most mutations switch sign. As a consequence, the distribution of mutational effects is predictable across genotypes and environments. Our results suggest that the effective dimensionality of genotype-to-phenotype maps across environments is surprisingly low.

### 2.2 Main Text

Adaptive evolution can lead to profound changes in the phenotypes and behaviors of biological systems, sometimes with adverse and sometimes with beneficial consequences for

human health, agriculture and industry [87–90]. However, predicting these changes remains difficult [91–93]. One major challenge is that how new mutations alter phenotypes and fitness of organisms often depends on the genetic background in which they arise ( $G \times G$  interactions or “epistasis”), the environment ( $G \times E$  interactions), or both ( $G \times G \times E$  interactions) [94]. These interactions can alter not only the magnitude but also the sign of mutational effects, causing evolutionary trajectories to become contingent on the initial genotype, environment and random events [95].

Much of prior empirical work focused on characterizing “microscopic”  $G \times G$ ,  $G \times E$  and  $G \times G \times E$  interactions for fitness-related phenotypes, that is, how the effects of individual mutations vary across genetic backgrounds and environments [75, 96–104]. Microscopic interactions determine evolutionary accessibility of mutational paths and the degree of genetic parallelism [68, 94, 96, 105, 106]. However, using this approach for predicting genome evolution is challenging because the number of interactions grows super-exponentially with the number of variable loci [93, 107]. Predicting phenotypic evolution may be more feasible and in many cases more useful [63, 93]. Such predictions rely on coarse-grained, or “macroscopic”, descriptions of  $G \times G$ ,  $G \times E$  and  $G \times G \times E$  interactions which ignore individual mutations and instead describe how the distributions of mutational effects vary across genotypes and environments [68, 94, 108, 109]. While many studies measured the distributions of effects of mutations on fitness, or “DFEs” [48, 53, 58, 110–113], a systematic understanding of macroscopic epistasis is still lacking [58, 109, 113, 114].

Several recent studies found that mutations often make the phenotype of an organism in which they occur less extreme [97, 98, 114–122], an instance of a more general phenomenon of microscopic “global epistasis” [94, 123]. Such epistasis is expected to arise for complex traits, including fitness [124], and it can be used to quantitatively predict the effects of individual mutations in new genetic backgrounds without the full knowledge of their  $G \times G$  interactions, thereby avoiding the combinatorial problem mentioned above [123, 124]. Perhaps more importantly, if most mutations exhibit global epistasis, the distributions of their phenotypic effects

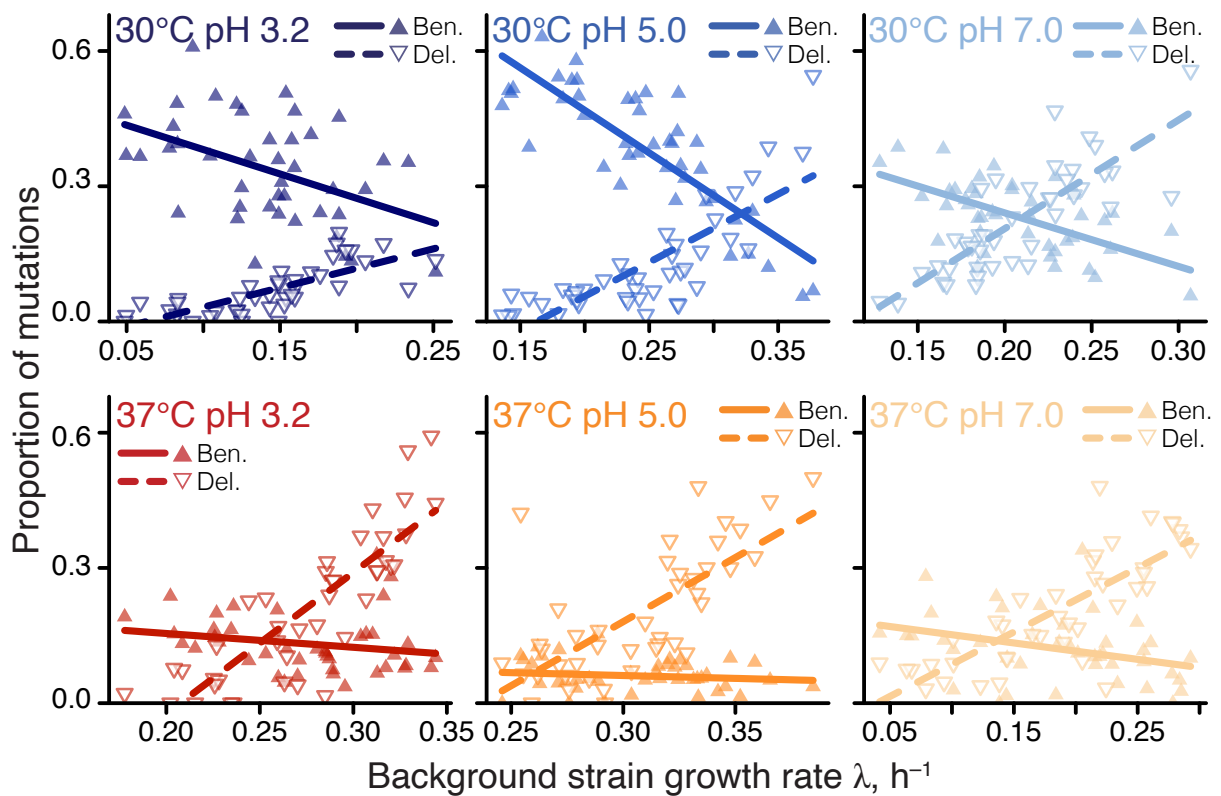
should also have a quantitatively predictable shape [124], which could facilitate evolutionary predictions at the phenotypic level. However, we have limited understanding of how the patterns of microscopic global epistasis vary across environments [104], and no attempt has been made so far to empirically characterize how microscopic  $G \times G$ ,  $G \times E$  and  $G \times G \times E$  interactions constrain the shape of the DFE.

Probing whether global epistasis models can capture  $G \times G$ ,  $G \times E$  and  $G \times G \times E$  interactions at both microscopic and macroscopic levels hinges on measuring the effects of many mutations across multiple genetic backgrounds and environments. To this end, we measured how  $\sim 100$  quasi-random barcoded insertion mutations constructed in our previous study [119] affect growth rate in 42 “background” strains of yeast *Saccharomyces cerevisiae* in six conditions (See Supplementary Materials). All background genotypes are segregants from a cross between two strains of yeast (RM, a vineyard strain, and BY, a lab strain) and differ from each other by  $\sim 2 \times 10^4$  SNPs throughout the genome [125]. Our environments varied by temperature (30°C and 37°C) and pH (3.2, 5.0 and 7.0), two stressors with global effects on yeast physiology [126–128], in a factorial design (see Supplementary Materials and Figure S1). This choice of strains and environments allowed us to explore a different (lower) range of growth rates of the background strains than in previous studies. Unlike previous studies, we kept our cultures growing close to the exponential steady state, which enabled us to infer the effect of each insertion mutation on absolute growth rate (denoted by  $\lambda$ ) from bulk barcode-based competition experiments, with precision of about  $7 \times 10^{-3} h^{-1}$  (see Supplementary Materials).

### 2.2.1 Absence of unconditionally beneficial or deleterious mutations

We first estimated the effects of our mutations on growth rate in different background strains and environments. To this end, following Johnson et al [119], we designated a set of five mutations as a putatively neutral reference and found that the remaining 94 mutations exhibit a range of effects on growth rate relative to this reference (denoted by  $\Delta\lambda$ ), from decreasing it by  $\Delta\lambda = -0.18 h^{-1}$  to increasing it by  $\Delta\lambda = 0.14 h^{-1}$ , with the median effect  $\Delta\lambda \approx 0 h^{-1}$ . We

validated a subset of these estimates with an independent low-throughput competition assay (see Supplementary Materials and Figure B.2C ). We then classified each mutation in each strain and environment as either beneficial or deleterious if the 99% confidence interval around its estimated effect did not overlap zero (see Supplementary Materials). All other mutations were classified as neutral. This procedure yielded conservative calls of mutation sign, with a false discovery rate of 2.4%.



**Figure 2.1. Proportions of beneficial and deleterious mutations vary with strain growth rate** Filled and empty triangles show the proportions of beneficial and deleterious mutations in each background strain as a function of its growth rate, respectively. Lines are the best-fit linear regressions; all are statistically significant ( $P < 0.05$ , t-test) except for beneficial mutations in 37C pH 3.0 and 37C pH 5.0.

We found that the fraction of beneficial and deleterious mutations varied between 0% and 63% and between 0 and 59% per strain, respectively. The high proportions of beneficial mutations was unexpected, but not unprecedented [111, 112]. However, no single mutation was



identified as either beneficial or deleterious in all strains and environments. 94% (88/94) of our mutations are beneficial in at least one strain and condition, and, of those, 96% (85/88) are also deleterious in at least one strain and condition. Even within the same environment, between 33% (31/94) and 63% (59/94) of all mutations change sign across background strains, and between 1% and 41% of mutations change sign across environments in the same strain (Figure B.3). Thus, the vast majority of mutations neither unconditionally increase nor unconditionally decrease growth rate.

A recent global epistasis model suggests that the fitness effects of most mutations should on average linearly decline with the fitness of the background genotype [124]. One striking qualitative prediction of this model is that the proportion of beneficial mutations should increase with the decreasing fitness of the background strain, reaching up to 100% in the lowest-fitness genotypes, whereas the proportion of deleterious mutations should correspondingly decrease. Consistent with this prediction, we found that the proportion of beneficial and deleterious mutations increased and declined in slow-growing strains, respectively, and these relationships were statistically significant in 10/12 cases (Figure 2.1). Thus, global epistasis indeed appears to be a major determinant of the sign of mutations (Figure B.3C).

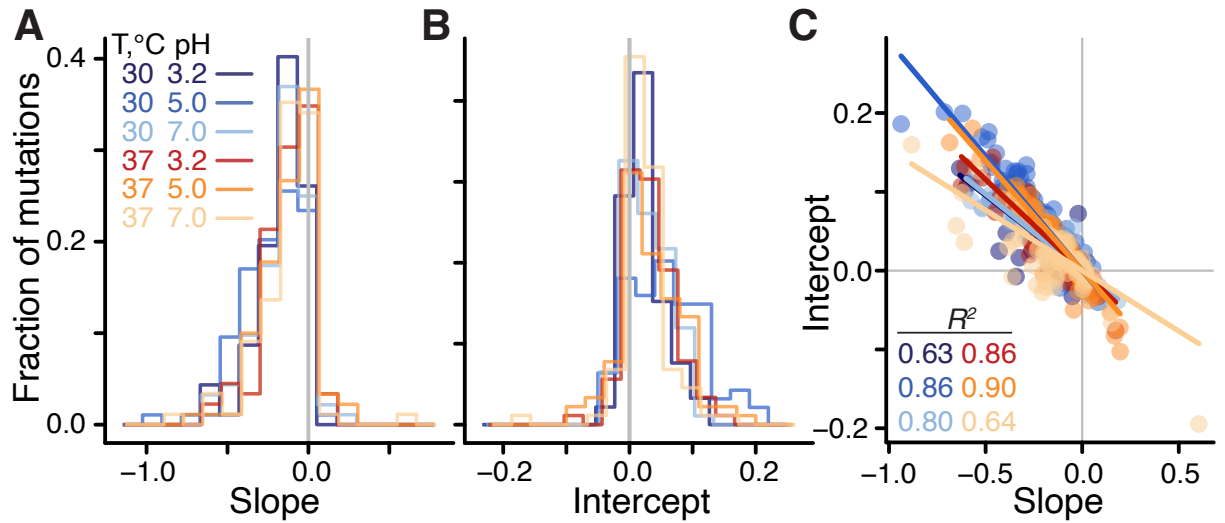
### 2.2.2 Consistency of microscopic global epistasis across environments

To probe the microscopic global epistasis model quantitatively, we modeled the effect  $\Delta\lambda_{mge}$  of each mutation  $m$  on growth rate in strain  $g$  and environment  $e$  as

$$\Delta\lambda_{mge} = a_{me} + b_{me}\lambda_{ge} + \xi_{mge}, \quad (2.1)$$

where  $\lambda_{ge}$  is the growth rate of the background genotype  $g$  in environment  $e$ . The first two terms in equation (2.1) capture global epistasis, a deterministic component which can be used for prediction, and  $\xi_{mge}$  captures the remaining (unpredictable) epistasis, which we refer to as “idiosyncratic” [94, 116].

We found that the linear model (1) was statistically significant for 97% (91/94) of mutations ( $F$ -test,  $P < 0.05$  after Benjamini-Hochberg correction), and explained on average 45% (interquartile interval [30%, 61%]) of variance in the effects of mutations across background strains and environments (see Supplementary Materials). When tested individually, 46% (250/545) of global epistasis slopes  $b_{me}$  are significantly different from zero ( $t$ -test,  $P < 0.05$  after Benjamini-Hochberg correction), with 98% (245/250) of them being negative (Figure 2.2A), consistent with the global epistasis theory [124] and previous observations [116, 119]. 96% (240/250) of significant intercepts  $a_{me}$  are positive (Figure 2.2B), implying that these mutations are expected to be beneficial in a hypothetical non-growing strain, consistent with relationships shown in Figure 2.1.

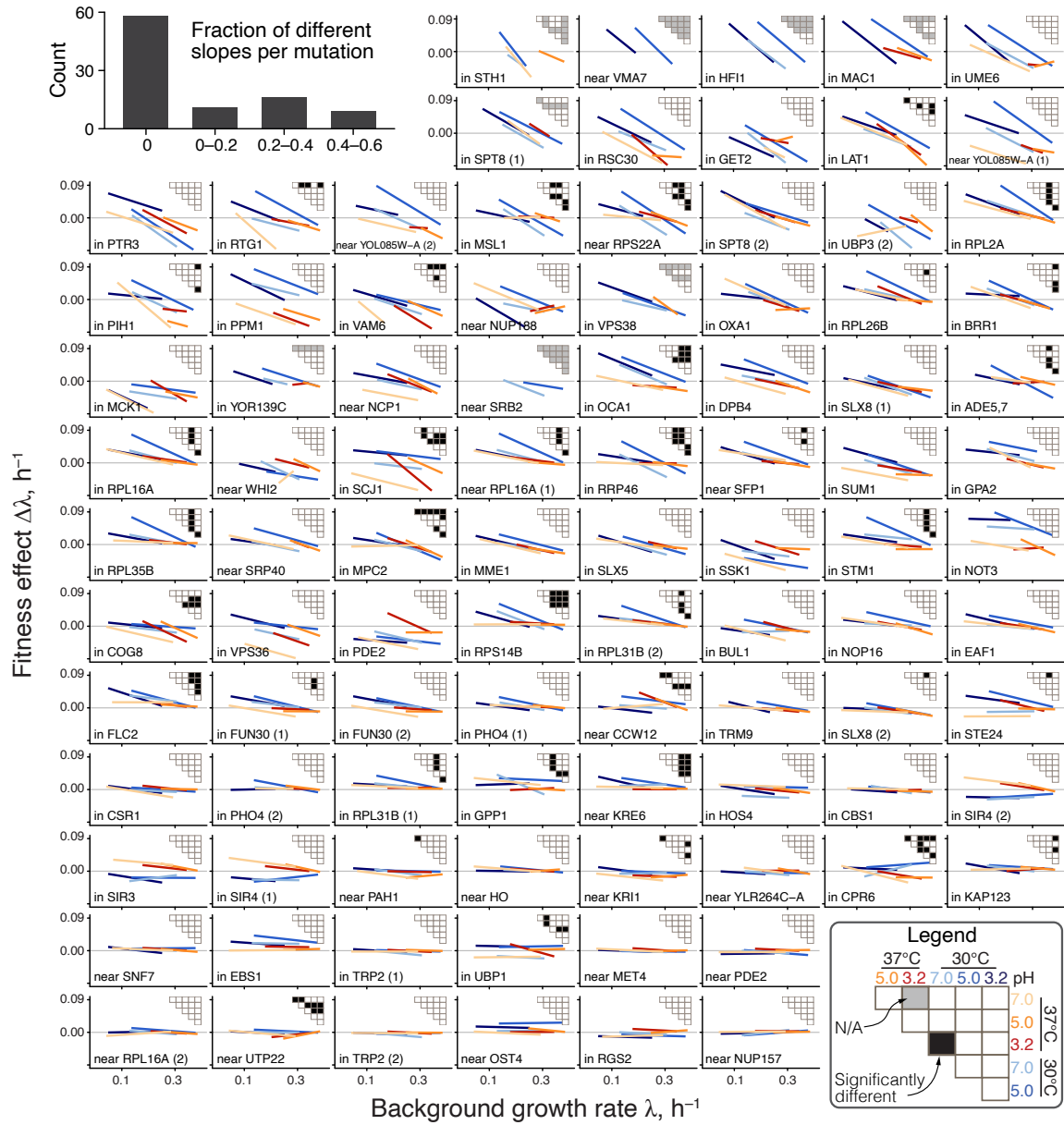


**Figure 2.2. Distributions of global-epistasis slopes and intercepts and their correlation.** **A,B** Histograms of slopes and intercepts estimated from fitting equation (2.1) to data (see Figure B.4 for statistical tests). **C.** Correlation between slopes and intercepts. Each point represents a mutation, colored by environment. Lines are the best fit linear regressions ( $P < 0.01$  for all,  $t$ -test).

We next examined how the global epistasis slopes  $b_{me}$  and intercepts  $a_{me}$  vary across mutations and environments. We found that the distributions of slopes are nearly invariant across environments (Figures 2.2A, B.4A,E and Table B.1), while the distributions of intercepts are more variable (Figure 2.2B, B.4A and Table B.1). Furthermore, slopes and intercepts are strongly

negatively correlated (Figure 2.2C, B.4F), such that mutations with a zero slope have on average a zero intercept, which is consistent with the paucity of unconditionally deleterious and beneficial mutations noted above. Thus, global epistasis slopes and intercepts are mathematically related as  $a_{me} = -\bar{\lambda}_e + \eta_{me}$  where the regression coefficient  $\bar{\lambda}_e > 0$  can be interpreted as the “pivot” growth rate at which a typical mutation switches its sign in environment  $e$  (see Supplementary Materials). Each individual mutation switches its sign at a background growth rate that deviates from  $e$  by  $\eta_{me}/b_{me}$ , and we therefore refer to  $\eta_{me}$  as the “pivot noise”. We find that the distributions of  $\eta_{me}$  in all environments are close to normal, with zero mean and standard deviation of approximately  $1.7 \times 10^{-2}$  (Figure B.4C). On the other hand, pivot growth rates  $e$  differ significantly across environments (see Supplementary Materials, Figure B.4B,G).

There are two possibilities for how the distributions of global-epistasis slopes could be nearly invariant across environments. It could be because slopes of individual mutations are nearly invariant. Alternatively, slopes of individual mutations could change across environments, but their distribution is preserved. To test these hypotheses, we examined how the global-epistasis slope of each individual mutation changes across environments. We found that they are statistically indistinguishable in 87% (1169/1333) of pairwise comparisons, and 62% (58/94) of mutations have statistically indistinguishable slopes in all environments (Figures 2.3,B.5, B.6). Moreover, even when slopes are statistically distinguishable, they are very similar, so that a model with six environment-specific slopes  $b_{me}$  explains only 2% more variance in the effects of mutations compared to an “invariant slope” model where each mutation is characterized by a single environment-independent slope  $b_m$  (45% versus 43% on average; Figure B.5). These results hold even if we only compare non-zero slopes or control for missing measurements (see Supplementary Materials and Figure B.5).



**Figure 2.3. Global-epistasis slopes of mutations are nearly invariant across environments.** Panels show regression lines from fitting equation (2.1) for each mutation, colored by the environment as in previous figures. Data points are shown in Figure B.6. Mutations are displayed in the order of increasing mean slope. Insets show the results of all pairwise slope-comparison tests (legend in lower right). Histogram in top left shows the overall distribution of fractions of significant tests per mutation (see Supplementary Materials).

The near-invariance of global-epistasis slopes of individual mutations could arise trivially if each environment shifted the growth rates of all strains by the same amount while preserving the relative order of their growth rates and the effects of mutations. However, this is not the case. We find that 34 out of 42 (81%) background strains are found among the 50% fastest growing strains in at least one environment and also among the 50% slowest growing strains in at least one other environment (see Supplementary Materials and Figure B.7). Similarly, 50% of mutations on average switch between the top and bottom 50% of mutation effects in the same strain across environments (Figure B.8). Thus, slopes are nearly preserved despite the fact that the relative rank orders of background strains and mutations are reshuffled across environments.

Taken together, the near-invariance of global-epistasis slopes across environments and the linear relationship between slopes and intercepts indicate that the microscopic  $G \times G$ ,  $G \times E$ , and  $G \times G \times E$  interactions for most of our mutations follow a simplified version of equation (2.1),

$$\Delta\lambda_{mge} = b_m(\lambda_{ge} - \bar{\lambda}_e) + \eta_{me} + \xi_{mge} \quad (2.2)$$

Equation (2.2), which we refer to as the “generalized global epistasis equation”, shows that the deterministic effect of the environment on global epistasis is captured by a single effective parameter, the pivot growth rate  $\bar{\lambda}_e$ .

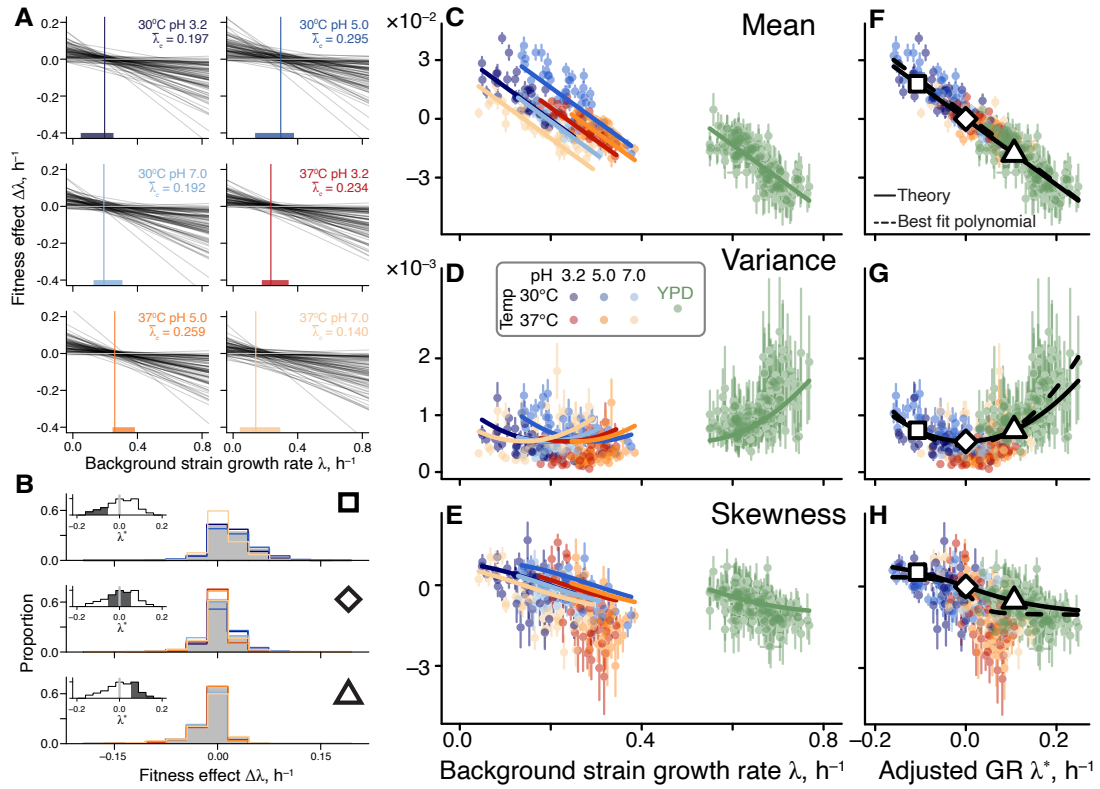
### 2.2.3 Distribution of fitness effects

To understand the implications of equation (2.2) for the macroscopic  $G \times G$ ,  $G \times E$  and  $G \times G \times E$  interactions, we calculated the first three moments of the distribution of fitness effects of mutations (DFE) under the simplifying assumption that pivot noise  $\eta_{me}$  and idiosyncratic epistasis  $\xi_{mge}$  terms are all independent (see Supplementary Materials). The generalized global epistasis equation predicts that the DFE mean ought to decline linearly with the background growth rate, with the same slope across environments, crossing zero at the environment-specific pivot growth rate  $\bar{\lambda}_e$ .

The behavior of higher DFE moments is less obvious. DFE variance is predicted to depend on  $\lambda_{ge}$  quadratically, with the parabola's minimum achieved at the pivot growth rate. To understand this prediction intuitively, consider an idealized case without idiosyncratic epistasis  $\xi_{mge} = 0$  (Figure 2.4A). Since most mutations switch sign near the pivot growth rate, strains growing at this rate have access only to mutations with small effects, and the DFE is narrow. As the background growth rate deviates from  $\bar{\lambda}_e$ , global epistasis lines for individual mutations spread out forming a “bowtie” (Figure 2.4A), and the DFE variance increases in both directions. This general pattern still holds even when  $\xi_{mge} \neq 0$  (see Supplementary Materials).

Finally, DFE skewness is predicted to decline monotonically with  $\lambda_{ge}$ , crossing zero again at the pivot growth rate. More generally, our model makes a qualitative prediction that the DFE varies across strains as a function of their environment-adjusted growth rate  $\lambda_{ge}^* = \lambda_{ge} - \bar{\lambda}_e$  rather than as a function of their absolute growth rate  $\lambda_{ge}$ . Furthermore, when the adjusted growth rate is zero, all odd central moments of the DFE are predicted to vanish and all even central moments are predicted to achieve their minimum (see Supplementary Materials).

To test these predictions, we compared the empirical DFEs across environments. We find that pairs of strains with matched adjusted growth rates have significantly more similar DFEs than pairs of strains with the same absolute growth rate in different environments or the same strain in different environments (Figure 2.4B and B.9), consistent with our predictions. We then plotted the first three moments of the empirical DFEs against the unadjusted and adjusted growth rate in all environments. We find that these moments align well only in the latter plot (Figure 2.4C–H). As predicted, the DFE mean and skewness decline monotonically with the strain's adjusted growth rate  $\lambda_{ge}^*$  and cross zero when  $\lambda_{ge}^* = 0$ , and DFE variance achieves its minimum at  $\lambda_{ge}^* = 0$ . We confirmed that these patterns are not a result of missing measurements (see Supplementary Materials and Figure B.10).



**Figure 2.4. Generalized global epistasis equation captures changes in the DFE across strains and environments.** **A.** The global epistasis contribution to the fitness effect of each mutation (term  $b_m(\lambda_{ge} - \lambda_e)$  in equation (2.2)) as a function of background growth rate term  $\lambda_{ge}$ . Each line represents a mutation. Vertical line indicates the pivot growth rate. Thick colored lines on the x-axis indicate the range of measured background-strain growth rates in each environment. **B.** Estimated DFEs for strains whose adjusted growth rate is negative (top panel), approximately zero (middle panel) and positive (bottom panel). Gray bars show DFEs pooled across all environments, colored lines show DFEs for individual environments (colors are as in previous figures). Insets show the distributions of adjusted growth rates for background strains, with the focal bin shaded. Large square, rhombus and triangle are shown for reference with panels F,G,H. **C,D,E.** DFE moments plotted against the background strain growth rate. Error bars show  $\pm 1$  standard errors (see Supplementary Materials). Solid curves show the theoretical predictions calculated from equation (2.2) and parameterized without the YPD data (see Supplementary Materials). DFE moments are calculated with a median of 74 mutations (interquartile interval [61,79]). **F,G,H.** Same data as in C, D, E, but plotted against the adjusted growth rate. Dashed curves are the best fitting polynomial of the corresponding degree (see Supplementary Materials).

The fact that all of our predictions hold indicates that the linear generalized global epistasis equation with uncorrelated noise terms quantitatively captures a major mode of variation in the DFE shape across genotypes and environments. However, if the environment truly modulates only one effective parameter, the pivot growth rate, then we should be able to predict DFE shapes in any environment, once its pivot growth rate is known. To test this prediction, we turned to our previous work where we measured DFEs of 163 yeast strains (a superset of the 42 strains used in this study) in a rich medium YPD [119]. We estimated the pivot growth rate for this environment as the background growth rate where the DFE mean equals zero (see Supplementary Materials). After growth-rate adjustment, we found that our theoretical predictions quantitatively capture variation in the DFE variance and skewness without any other fitted parameters (green points in Figure 2.4C-H).

#### **2.2.4 Discussion**

Having measured the fitness effects of many mutations across diverse yeast genotypes and environments, we obtained two main results. First, the global-epistasis slopes of individual mutations are nearly invariant across environments. If interpreted through the lens of the Reddy-Desai model [124], this finding suggests that our environmental perturbations largely preserve the overall statistical structure of genetic interactions, consistent with previous work in yeast [129]. Whether this regularity holds for more extreme perturbations or those that target specific cellular processes, such as drugs, remains to be seen [104]. The near-invariance of global-epistasis slopes across environments leads to a simple equation (2.2) that allows us to predict the expected fitness effect of a mutation in a new strain and environment. However, it is important to keep in mind that the actual effect will deviate from this expectation due to idiosyncratic epistasis. In other words, equation (2.2) is statistical in nature and, as such, does not uniquely determine the underlying genotype-to-fitness map (see Supplementary Materials).

Our second main result is the empirical characterization of constraints imposed on the DFE by microscopic global epistasis. We found that the environment controls a single effective



parameter, the pivot growth rate  $\bar{\lambda}_e$ , and a genotype's DFE is set by the deviation of its growth rate from  $\bar{\lambda}_e$ , as shown in Figure 2.4. For example, many of our background strains happen to grow slower than  $\bar{\lambda}_e$  and, as a result, have access to many beneficial mutations (Figure 2.1).

The physiological mechanisms that determine the pivot growth rate are as of yet unclear. According to the Reddy-Desai model, “bowtie” patterns analogous to those shown in Figure 2.4A are expected to arise generically on genotype-to-fitness maps with many epistatic interactions [124]. In this model, a typical mutation switches sign at the average fitness of all genotypes, indicating that the pivot growth rate could be a metric of environmental permissibility. The fact that in our data the pivot growth rate correlates with the average growth rate of our background strains is consistent with this interpretation (Figure B.4G). While the consistency of our results with theory suggests that they should extend to other strains and organisms, types of mutations and environments, their generality must ultimately be established empirically. If they hold broadly, several important implications are conceivable. In genetics, our model can improve predictions of the phenotypic effects of mutations. In conservation biology, the possibility that low-fitness genotypes have access to large supplies of beneficial mutations gives hope that evolutionary rescue may prevent some species extinctions. In evolutionary biology, our results point to the existence of a universal class of distributions of fitness effects of mutations, which could explain why evolutionary dynamics of fitness are so similar and predictable across systems [108, 116, 122, 130]. More broadly, our model can help us better understand these dynamics, including how adaptation to one environment alters fitness in another [109] or how the DFE might change along an evolutionary trajectory (Supplementary Materials). More fundamentally, our results suggest that epistasis reduces the effective dimensionality of genotype-to-phenotype maps [94, 123, 124, 131]. What biological constraints cause this dimensionality reduction, how it emerges and when it breaks down are exciting open questions in systems biology.

## 2.3 Materials and Methods

### 2.3.1 Strains and media

#### Background strains

Our “background strains” of yeast *S. cerevisiae* are a subset of a larger library of segregants that were previously generated from a cross between the lab strain BY and the vineyard strain RM [125] and whose evolutionary properties have been previously characterized [114, 119, 132]. Specifically, our set of 42 background strains (listed in Table S6-Tab 2) is a subset of the strains used in the “Small Library” RB-TnSeq experiment described in Ref. [119]. All the necessary strain details can be found in Refs. [119, 125]. Most importantly, our background strains differ from each other at approximately 25,000 loci and span nearly the full growth rate range in YPD measured in Ref. [119] (see Section 2.3.3). These strains also vary widely in both DFE mean (2.5-fold range) and variance (1.5-fold range) in YPD.

#### RB-TnSeq libraries

Background strains individually transformed with these RB-TnSeq libraries were kindly provided by Michael Desai (Harvard University). The design and construction of the RB-TnSeq libraries is described in detail in Ref. [119]. Briefly, each background strain was transformed twice with the same set of 100 redundantly barcoded transposon-insertion mutations, resulting in two biological replicates for each mutation in each strain, such that, within each transformation, each barcode uniquely tags a particular mutation and background strain (43% of barcodes were used in both transformations). This set of mutations is conditionally random, given two constraints, as described in Ref. [119]. Briefly, (i) mutations whose fitness effects we can measure must be non-lethal and (ii) only those mutations were included in this set that had a measurable fitness effect in at least six out of 18 genetic backgrounds in the rich medium YPD. Thus, the results of this work should be viewed as representing trends in the effects of disruptions of non-essential but functionally important genes.

On average, each mutation was represented by 11 and 37 barcodes in the first and second transformation, respectively (see Table S6-Tab 5). Five mutations (IDs: 91 (nearby COA6), 51 (nearby FIT2), 6 (nearby MET2), 99 (nearby TDA11), 102 (nearby YSP2)) that target intergenic regions were used as a neutral reference, as in Ref. [119]. These mutations were found to have no significant fitness effects in all strains in one environment in Ref. [119]. These reference mutations were represented by on average 19 and 77 barcodes in the two transformations.

### **Media and environments**

Unless otherwise noted, all experiments were performed in synthetic complete medium (SC, 2% dextrose (VWR, #90000-904), 0.67% YNB + nitrogen powder (Sunrise Science Products, #1501-500), 0.2% synthetic complete drop-out powder mixture (Sunrise Science Products, #1300-030)). We added ampicillin (Amp) and tetracycline (Tet) at concentrations given in Table B.2 into the medium to prevent bacterial contamination. Our environmental conditions differed by two factors, temperature (30°C and 37°C) and pH (3.2, 5.0, and 7.0). pH was maintained with the citrate-phosphate buffer [133], which was prepared using 1 M stocks of citric acid (VWR, # 97061-858) and  $K_2HPO_4$  (VWR, #97062-234) following the protocol described in Ref. [134]. Autoclaved media were pH-adjusted by adding the necessary volumes of sterile citric acid and  $K_2HPO_4$  solutions and measuring the pH of the buffered media. If the pH of the buffered media deviated from the desired level, we further adjusted it by adding small volumes of 4 M *HCL* (Sigma-Aldrich #84435). Media was used within two days.

## **2.3.2 Experimental procedures**

### **RB-TnSeq experiment**

To estimate the effects of tn-mutations on the absolute growth rate in 42 background strains in 6 environments, we pooled the tn-mutant libraries of our background strains into multiple pools (see below for details) and maintained these pools in continuous growth in 150 ml of media in 500 ml flasks over the period of 48 hours with dilutions down to  $5 \times 10^7$  cells and

sampling every 12 hours. We carried out two replicate competition assays, one per independently transformed library (see Section 2.3.1).

### **Pre-growth and pooling.**

Prior to pooling, mutagenized strain libraries were defrosted and pre-grown for 24 h in 96-deep-well plates with 1 ml of media in each of our environments. Based on preliminary growth rate estimates, we grouped the background-strain libraries by their growth rate into three groups for the competition assays at 30°C and into two groups for the competition assays at 37°C. Each background strain was represented only in one group per temperature, with the exception of LK5-G01, which was added to each group as cross-group control. Group identities of each background strain can be found in Table S6-Tab 2. Thus, we propagated 6 cultures (3 growth rate groups  $\times$  2 biological replicates) in each of the 30°C environments and 4 cultures (2 growth rate groups  $\times$  2 biological replicates) in each of the 37°C environments for a total of 30 cultures.

After pre-growth, we measured OD600 of each mutant culture and converted it into cell density using a previously obtained calibration curve. Based on these density estimates, we pooled mutant cultures at approximately equal abundances, with a slight over-representation of those background strains whose preliminary growth rate estimates were lower. We then measured the density of each mixed culture again and transferred  $5 \times 10^7$  cells into the corresponding competition flask. The remaining mixed  $T_0$  cultures were frozen using the protocol described below.

### **Growth and dilution.**

Competitions were carried out in 150 ml of media in 500 ml baffled flasks (Pyrex No. 4446-500) in a shaking incubator (Eppendorf New Brunswick I26, 2.5 cm orbit) set to 150 rpm and appropriate temperature. Every 12 h, we estimated the cell density of each culture using plate-reader-based OD600 measurements and a previously obtained calibration curve (Table S6-Tab 3). Then,  $5 \times 10^7$  cells were transferred into the fresh media. When the transfer volume exceeded 1 ml, we adjusted the volume of fresh media to maintain the consistent culture

volume of 150 ml. All cultures were propagated for four growth and dilution cycles (48 hours), yielding five samples per culture.

### **Sampling and storage.**

We pelleted the cells from 50 ml of each cultures remaining after the transfer and froze the cell pellets at  $-70^{\circ}\text{C}$  for subsequent DNA extraction and barcode sequencing.

### **Sequencing library preparation**

We used the YeaStar Genomic Kit Protocol I (Zymo Research, #D2002) to extract gDNA from  $\sim 1$  ml of pelleted yeast cultures. To generate Illumina-ready dual-indexed amplicon library, we used a two-step PCR protocol, modified from Ref. [53] as follows. All primer sequences can be found in Table S5.

#### **First PCR.**

We combined 100 ng of extracted gDNA, 25  $\mu\text{l}$  of OneTaq DNA polymerase Master Mix (New England BioLabs, #M0482L), 0.5  $\mu\text{l}$  of 10  $\mu\text{M}$  oAM-R2P-100-R01 primer, 0.5  $\mu\text{l}$  of 10  $\mu\text{M}$  oAM-R1P-20X-F01 primer, 1  $\mu\text{l}$  of 50  $\mu\text{M}$   $\text{MgCl}_2$ , and molecular biology grade water up to the total volume of 50  $\mu\text{l}$ . We used the following PCR protocol:

1.  $94^{\circ}\text{C}$  for 30 sec.
2.  $94^{\circ}\text{C}$  for 30 sec.
3.  $50.5^{\circ}\text{C}$  for 30 sec.
4.  $68^{\circ}\text{C}$  for 70 sec.
5. Repeat Steps 2–4 for a total of three times
6.  $68^{\circ}\text{C}$  for 5 min.

We purified this PCR product with AMPure XP magnetic beads (Beckman, #A36881) (1:1 ratio).

**Second PCR.**

We combined 15 µl of purified PCR I product, 25 µl of OneTaq DNA polymerase Master Mix, 1 µl of 50 µM MgCl<sub>2</sub>, 1 µl of 10 µM N7XX primer (Nextera), and 1 µl of 10 µM S5XX primer (Nextera), and 7 µl of molecular biology grade H<sub>2</sub>O. We used the following PCR protocol:

1. 94°C for 30 sec.
2. 94°C for 30 sec.
3. 62°C for 30 sec.
4. 68°C for 70 sec.
5. Repeat Steps 2–4 for a total of 24 times
6. 68°C for 5 min.

The final PCR product was purified as above, run on a gel, extracted and purified with QIAquick PCR purification kit (Qiagen, #28106).

**Sequencing.**

We sequenced the libraries with paired-end 150 bp reads on one HiSeq4000 platform and two HiSeq X10 platforms (Illumina).

**Validation of estimated fitness effect of mutations**

We were surprised by the high fraction of beneficial mutations identified in our RB-TnSeq experiment and carried out an additional experiment designed to validate our fitness-effect estimates obtained in the large-scale multi-strain pooled experiment. Specifically, we wanted to validate that the fitness effects of mutations (i) were in fact driven by the tn-insertions themselves rather than some additional unobserved mutations that could have arisen during transformation, and (ii) were not too specific to the environment created in the multi-strain competition pools.

To this end, after preliminary analyses, we selected a set of seven mutations that were expected to span a range of fitness effects: nearby MET2, nearby MET4, in NOT3, in PPM1, in RSC30, nearby TDA11, in MPC2 (Mutation IDs: 6, 10, 117, 66, 71, 99, and 127, Table S6-Tab 4). We generated new RB-TnSeq libraries for these mutations and reconstructed them in five background strains, LK2-D07, LK6-A05, LK5-C04, LK1-H02, and LK1-C09. Two of the selected mutations (nearby MET2 and nearby TDA11, IDs: 6, 99) were used as a neutral reference in the RB-TnSeq experiment. The mutation nearby MET4 (ID: 10) was identified as neutral in the backgrounds and environments used in the validation experiment, and the remaining mutations were identified as non-neutral in at least some of these strains and environments.

The barcoded libraries for individual mutations were generated using the protocol described in Ref. [119]. In total, we created 28 barcoded plasmid libraries (two replicated libraries per mutation, and two different drug resistance markers per mutation and replicate), such that each library contained a unique set of barcodes. We then transformed these libraries into all background strains, with LK2-D07 and LK6-A05 receiving one drug resistant marker (Hygromycin) and the remaining three LK5-C04, LK1-H02, and LK1-C09 receiving the other (Nourseothricin), again following protocols described in Ref. [119]. Transformant colonies were scraped and, after 24 h of additional growth in selective media, cultures were pelleted and frozen in 20% glycerol at  $-70^{\circ}\text{C}$ . To determine which barcodes were associated with each mutation in each background strain, we sequenced each of the 28 yeast mutant libraries at the barcode locus using the same protocols as in Section 2.3.2.

After generating the libraries of individual mutations, we estimated their fitness effects in two environments,  $30^{\circ}\text{C}$  pH 5.0 and  $37^{\circ}\text{C}$  pH 7.0, by competing only mutants that share the same genetic background. We pre-grew 140 cultures ( $5$  background strains  $\times$   $7$  mutations  $\times$   $2$  replicate mutant libraries  $\times$   $2$  environments) in the buffered SC media containing Amp (to avoid bacterial contamination) and either Nat or Hyg (required for selecting for the transformants, see Tables S6-Tab 2 and B.2, and Ref. [119]) and incubated them in the two focal environments ( $30^{\circ}\text{C}$  pH 5.0 and  $37^{\circ}\text{C}$  pH 7.0) for 24 h in test tubes shaken at 220 rpm. After measuring the

concentrations of the grown cultures as described above (see Section 2.3.2), we created one mixed culture per background strain, per mutant library and per environment ( $2 \text{ environments} \times 5 \text{ strains} \times 2 \text{ biological replicates}$ , for a total of 20 mixed cultures) as follows. We added each of the two neutral reference mutation cultures (IDs: 6,99) at frequency 25% each and we added each of the “query” mutant cultures (IDs: 10, 117, 66, 71, and 127) at 10%, such that initial ratio of reference and query mutants in each culture was 1:1. After estimating cell counts in these mixed cultures using OD600 (see Section 2.3.2), we transferred  $5 \times 10^7$  cells to start the competition assay. The assays, sampling and sequencing library preparation were performed following the same protocols as in Section 2.3.2. Fitness effects of mutations estimated in this experiment are provide in Table S6-Tab 8.

### **2.3.3 Data Analysis**

#### **Code**

All analysis code is written in R 4.3.1 and is available at

[https://github.com/ardellsarah/Yeast\\_mutation\\_effects\\_across\\_strains\\_and\\_environments](https://github.com/ardellsarah/Yeast_mutation_effects_across_strains_and_environments).

Packages used are listed in the beginning of all scripts. Computationally intensive analyses were run on the Triton Supercomputing Cluster (TSCC).

#### **Counting barcodes**

##### **Raw barcode counts.**

Barcode counts for the main RB-TnSeq experiment (Section 2.3.2) were obtained using the BarcodeCounter2 package [135] with a pre-determined barcode-mutation association data from Ref. [119].

To determine barcode counts in the validation experiment, we first used the barcode sequencing data for individual mutation libraries (see Section 2.3.2) to associate each barcode sequence with a particular mutation and background strain. To do so, we used regular expressions



to extract all unique barcode sequences and clustered them using the `seq_cluster` function in R's `bioseq` package. We then used `BarcodeCounter2` with the resulting barcode-mutation associations to extract raw barcode counts for each sample file.

### **Filtering.**

Because it is critical to have accurate reference barcode counts for the inference of fitness effects of mutations, we discarded all time points that contained less than 500 reference mutation counts for any given strain, replicate and environment. This filtering removed 1.3% (41,404/3,060,514) of strain-environment-replicate-time point combinations. Then, we retained only those barcodes that were present at three or more time points (in any given condition and replicate) at 5 or more counts at each time point.

### **Estimating growth rates of background strains and fitness effects of mutations**

The central piece of our procedure for estimating the growth rates of background strains and the fitness effects of mutations is the detection of “outlier” barcodes, i.e., those that have abnormally high or low growth rates relative to other barcodes tagging the same mutation. Such outliers arise likely due to the tn-mutants acquiring secondary mutations either during the barcoding step or during the RB-TnSeq experiment. The outlier detection procedure requires preliminary estimates of fitness effects of barcodes tagging each mutation, which in turn requires a robust set of reference barcodes. Thus, our procedure consists of the following steps.

1. Estimate the growth rates of all barcodes.
2. Detect and exclude outlier barcodes for reference mutations to obtain a robust estimate of the growth rate of reference mutants.
3. Obtain preliminary estimates of fitness effects of mutations based on the robust set of reference barcodes.
4. Detect and exclude outlier barcodes for non-reference mutations.

5. Estimate the background growth rates and the fitness effects of mutations in each biological replicate.
6. Pool estimates across replicates.

We describe the outlier detection algorithm at the end of this section. Suffices to say that this algorithm takes as input (i) the set of all barcodes tagging a given mutation (in a given genetic background, biological replicate and environment), (ii) the cell count estimates over time of all these barcoded lineages and (iii) the preliminary estimates of the fitness effects of these lineages, and it outputs a robust “outlier-free” set of barcodes corresponding to the mutation.

Before we describe our estimation procedure, recall that, in a given biological replicate  $r$ , each barcode  $k$  uniquely specifies a particular tn-mutation  $m$  and the background strain  $g$  into which this mutation is introduced (see Section 2.3.1). We denote the set of all barcodes tagging mutation  $m$  in genetic background  $g$  in replicate  $r$  by  $\tilde{S}_{mgr}$ . We also denote the set of five reference mutations by  $\mathcal{S}^{\text{ref}}$  (see Section 2.3.1 and Table S6-Tab 4) and we denote the set of “reference barcodes”, i.e., all barcodes that tag these reference mutations in the genetic background  $g$  and replicate  $r$ , by  $\tilde{S}_{gr}^{\text{ref}} = \bigcup_{m \in \mathcal{S}^{\text{ref}}} \tilde{S}_{mgr}$ . Tilde denotes the fact that these sets potentially include outlier barcodes.

### Estimation of barcode growth rates across time intervals.

We first calculate the frequency of each barcode  $k$  (reference or non-reference) in each replicate  $r$  at each time point  $t$  by dividing its read count by the total count of all barcodes present in the same flask at that time. To estimate the number of cells  $N_{kret}$  that carry barcode  $k$  in replicate  $r$  in environment  $e$  at the sampling time  $t$ , we multiply barcode frequency by the total number of cells present in the flask at that time point (Table S6-Tab 3). We estimate the growth rate  $\lambda_{kret}$  of the barcode lineage  $k$  as

$$\lambda_{kret} = \frac{1}{\Delta t} \ln \frac{N_{kret+1}}{N_{kret}},$$

where  $\Delta t$  is the time between transfers (typically 12 h, Table S6-Tab 3).

### Detection and exclusion of outlier reference barcodes.

We obtain a preliminary estimate of the effect of each reference barcode  $k$  in background  $g$ , environment  $e$  and replicate  $r$  as

$$\widetilde{\Delta\lambda}_{kre} = \frac{1}{M_{kre}} \sum_t \left( \lambda_{kret} - \tilde{\lambda}_{gret} \right), \quad k \in \tilde{S}_{gr}^{\text{ref}},$$

where  $\tilde{\lambda}_{gret} = \text{Median} \{ \lambda_{kret} : k \in \tilde{S}_{gr}^{\text{ref}} \}$  and  $M_{kre}$  is the number of time intervals where barcode  $k$  is observed in environment  $e$  in replicate  $r$ . We use these estimates to apply our outlier detection algorithm (see below) and generate a robust set of reference barcodes  $S_{gre}^{\text{ref}}$  for each background strain  $g$  in each environment  $e$  and biological replicate  $r$ .

### Preliminary estimates of fitness effects of mutations.

For any barcode  $k$  tagging a non-reference mutation  $m$  in genotype  $g$ , we calculate its fitness effect in environment  $e$  and replicate  $r$  as

$$\Delta\lambda_{kre} = \frac{1}{M_{kre}} \sum_t \Delta\lambda_{kret}, \quad \text{for any } k \in \tilde{S}_{mgr}, \quad (2.3)$$

where  $M_{kre}$  is the number of time intervals when this barcode is observed in environment  $e$  in replicate  $r$ , and

$$\begin{aligned} \Delta\lambda_{kret} &= \lambda_{kret} - \lambda_{gret}, \quad \text{for any } k \in \tilde{S}_{mgr}, \\ \lambda_{gret} &= \text{Median} \left\{ \lambda_{kret} : k \in S_{gre}^{\text{ref}} \right\} \end{aligned} \quad (2.4)$$

are robust estimates of the fitness effect of barcode  $k$  and growth rate of the background strain  $g$ , respectively, both at time interval  $t$  in environment  $e$  and replicate  $r$ . We then obtain a preliminary estimate of the fitness effect of each mutation  $m$  in genetic background  $g$  in environment  $e$  and

biological replicate  $r$  as

$$\widetilde{\Delta\lambda}_{mgre} = \text{Median} \{ \Delta\lambda_{kre} : k \in \tilde{S}_{mgr} \}. \quad (2.5)$$

### Detection and exclusion of outlier barcodes for non-reference mutations.

We use preliminary fitness effect estimates given by equation (2.5) and apply our outlier detection algorithm (see below) to generate a clean set of barcodes  $S_{mgre}$  for each mutation  $m$  in each background strain  $g$  in each environment  $e$  and biological replicate  $r$ .

### Estimation of background growth rates and fitness effects of mutations in each biological replicate.

We estimate the growth rate of each background strain  $g$  in each environment  $e$  and biological replicate  $r$  as

$$\lambda_{gre} = \frac{1}{M_{gre}} \sum_t \sum_{k \in S_{gre}^{\text{ref}}} \lambda_{kret}, \quad (2.6)$$

where  $M_{gre}$  is the number of barcode-time interval combinations at which  $\lambda_{kret}$  are estimated.

We estimate the fitness effect of each mutation  $m$  in each genetic background  $g$  environment  $e$  and biological replicate  $r$  as

$$\Delta\lambda_{mgre} = \frac{1}{M_{mgre}} \sum_t \sum_{k \in S_{mgr}} \Delta\lambda_{kret}, \quad (2.7)$$

where  $\Delta\lambda_{kret}$  are given by equation (2.4) and  $M_{mgre}$  is the number of barcode-time interval combinations at which  $\lambda_{kret}$  are estimated.

### Pooling estimates across neutral references.

We estimate how the growth rate estimates of background strains correlate between estimates from single references and all four other reference mutations combined. We find they are highly correlated (Figure B.2F), so we pool the data from all reference mutations together to obtain final growth rate estimates of each background.

### Pooling estimates across biological replicates.

We estimate how the fitness effects of mutations obtained using equation (2.7) correlate across replicates. Since the replicates are highly correlated (B.2A-B), we pool the data from both replicates to obtain our final estimates of background growth rates and their standard errors,

$$\lambda_{ge} = \frac{1}{M_{ge}} \sum_{r,t} \sum_{k \in S_{gre}^{\text{ref}}} \lambda_{kret}, \quad (2.8)$$

$$\sigma_{ge}^{\lambda} = \left[ \frac{1}{M_{ge}(M_{ge} - 1)} \sum_{r,t} \sum_{k \in S_{gre}^{\text{ref}}} (\lambda_{kret} - \lambda_{ge})^2 \right]^{\frac{1}{2}}, \quad (2.9)$$

where  $M_{ge}$  is the number of barcode-replicate-time combinations at which  $\lambda_{kret}$  are estimated for background  $g$  in environment  $e$ . We estimate the fitness effects of mutations and their standard errors

$$\Delta\lambda_{mge} = \frac{1}{N_{mge}} \sum_{r,t} \sum_{k \in S_{mgr}} \Delta\lambda_{kret}, \quad (2.10)$$

$$\sigma_{mge}^{\Delta\lambda} = \left[ \frac{1}{M_{mge}(M_{mge} - 1)} \sum_{r,t} \sum_{k \in S_{mgre}} (\Delta\lambda_{kret} - \Delta\lambda_{mge})^2 + \left( \sigma_{ge}^{\lambda} \right)^2 \right]^{\frac{1}{2}}, \quad (2.11)$$

where  $M_{mge}$  is the number of barcode-replicate-time combinations at which  $\Delta\lambda_{kret}$  are estimated for mutation  $m$  in background strain  $g$  and environment  $e$ .

The distributions of standard errors for background growth rates and fitness effects are shown in Figure B.2G. The average standard error of the background growth rate is

$$\sigma_{bg} = 1.6 \times 10^{-3} \text{ h}^{-1} \quad (2.12)$$

and the average standard error of fitness effect is

$$\sigma_{mut} = 7.0 \times 10^{-3} \text{ h}^{-1}. \quad (2.13)$$

The comparison of these values shows that the noise in non-reference barcodes is typically more than 4-fold higher than noise in reference barcodes.

### **Detection of outlier barcodes.**

The goal of this procedure is to detect those barcodes whose frequencies either rise or fall unexpectedly quickly compared to the other barcodes tagging the same mutation in the same genetic background. To this end, we follow the method developed in Ref. [119], which takes as input the set of all barcodes  $k$  tagging a given mutation (in a given genetic background, biological replicate and environment), the corresponding cell counts  $N_{kt}$  at each time sampling  $t$  and the preliminary estimates  $\widetilde{\Delta\lambda}_k$  of the fitness effects of all these barcoded lineages. Briefly, we calculate the “within-mutation” frequencies  $f_{kt}$  at time  $t$  as  $f_{kt} = N_{kt} / (\sum_{k'} N_{k't})$ . All barcoded lineages tagging the same mutation should grow at the same rate. Therefore, we expect all frequencies  $f_{kt}$  to be constant over time, barring demographic and sampling noise. To determine which frequency trajectories are inconsistent with this neutral expectation, we create a “within-mutation-neutral reference” (WMNR) set of barcodes whose preliminary fitness effect  $\widetilde{\Delta\lambda}_k$  is within 0.01 of the median fitness effect of all these barcodes. Then, we fit two models to all frequency trajectories  $f_{kt}$  for those barcodes that are not in the WMNR set. In the neutral model, each query barcode’s trajectory does not systematically change relative to the pooled WMNR trajectory. In the model with selection, query barcode’s frequency can systematically increase or decrease. We find the log-likelihood of the observed barcode trajectory given each of the two models and calculate the likelihood ratio (LR) statistic. We conservatively exclude all barcodes with the LR statistic values greater than 40, corresponding to a  $P$ -value from a  $\chi^2$  distribution with 1 d.f.  $< 10^{-9}$ . This algorithm returns a “clean” subset of barcodes that tag a given mutation (in a given genetic background, environment and replicate).

Using this method, we exclude a total of 2.4% (16,799/683,754) of barcode-replicate combinations.

## Calling beneficial and deleterious mutations

To call each mutation as either beneficial, deleterious or neutral in each genetic background and environment, we construct the 99% confidence interval using a normal distribution with the mean equal to  $\Delta\lambda_{mge}$  (equation (2.10)) and variance equal to the standard error of the mean  $\sigma_{mge}^{\Delta\lambda}$  (equation (2.11)). Mutations whose entire confidence interval is below zero are called deleterious and those whose entire confidence interval is above zero are called beneficial. All other mutations, i.e., those whose confidence interval spans zero, are called neutral. We identify a total of 6971 non-neutral mutation-genotype-environment combinations out of 17306 tested. If all mutations were truly neutral, we would expect to call 1% or  $\sim 173$  of them non-neutral by chance, yielding the false discovery rate of  $173/6971 = 2.5\%$ .

## Growth rates of background strains in YPD

Johnson et al estimated the mean, variance and skewness of tn-insertion DFEs in 163 yeast background strains  $g$  in rich YPD medium, as well as the fitness  $s_g$  of these strains relative to a common reference strain [119]. They also separately measured exponential growth rate  $\lambda_g$  for a subset of their background strains. Using this subset of strains, we find a very good linear relationship between  $\lambda_g$  and  $s_g$  in YPD ( $P = 9.632 \times 10^{-9}$ ,  $R^2 = 0.97$ ),

$$\lambda_g = 0.9729 s_g + 0.6732. \quad (2.14)$$

We use equation (2.14) to estimate the growth rate in YPD for all 163 strains. We estimate the pivot growth rate in YPD as the growth rate at which DFE mean is predicted to cross 0 from a linear model of DFE mean against estimated background growth rate.

## Variation of growth rates and mutational effects across environments

To assess how the growth rate rank order of background strains varies across environments, we first find the median growth rate of all strains in each environment. We then call a

strain as “above median” if its growth rate is above this median by at least one unit of its standard error  $\sigma_{ge}^\lambda$  (equation (2.9)). Analogously, we call a strain as “below median” if its growth rate is below the median by at least  $\sigma_{ge}^\lambda$ . Any strain that is identified at least once as above median growth rate and at least once as below median growth rate was labelled as “Rank change” in Figure B.7A).

We assessed rank-order changes of mutations across environments within each strain using an analogous procedure (Figure B.8).

### Models of global epistasis

We fit equations (2.1) and (2.2) in the main text using `lm` function in R using the mean mutation effect and background strain growth rates. In addition to the estimates of slopes and intercepts, these fitted models return model  $P$ -values and adjusted  $R^2$  values. After Benjamini-Hochberg  $P$ -value correction, we find that equation (2.1) is significant ( $P < 0.05$ ) for 91/94 (97%) mutations and equation (2.2) is significant for 90/94 (96%) mutations.

### Accounting for measurement errors in the calculation of global epistasis slopes.

As pointed out by Berger and Postma [136], negative slopes in equation (2.1) in the main text (for any given mutation  $m$  and environment  $e$ ) can arise spuriously due to the fact that measurements errors in  $\lambda_{ge}$  and  $\Delta\lambda_{mge}$  are correlated. We take two approaches to assess the potential impact of these correlated errors.

First, using the same approach as in Ref. [136], we compute the corrected correlation coefficient between  $\Delta\lambda_{mge}$  and  $\lambda_{ge}$  across background genotypes for a fixed mutation  $m$  in a fixed environment  $e$  as

$$\rho'(\Delta\lambda_{mge}, \lambda_{ge}) = \frac{\text{Cov}(\Delta\lambda_{mge}, \lambda_{ge}) + \sigma_{bg}^2}{\sqrt{(\text{Var}(\lambda_{ge}) - \sigma_{bg}^2)(\text{Var}(\Delta\lambda_{mge}) - \sigma_{bg}^2 - \sigma_{mut}^2)}}. \quad (2.15)$$

Here  $\lambda_{mge}$  and  $\Delta\lambda_{mge}$  are given by equations (2.8) and (2.10), respectively;  $\sigma_{bg}^2$  and  $\sigma_{mut}^2$  are the



measurement noise variance for growth rate of the background and mutant strains, respectively, which can be calculated from expressions (2.12) and (2.13), respectively; and

$$\begin{aligned}\langle X \rangle &= \frac{1}{K} \sum_g X_g, \\ \text{Var}(X) &= \frac{1}{K-1} \sum_g (X_g - \langle X \rangle)^2, \\ \text{Cov}(X, Y) &= \frac{1}{K-1} \sum_g (X_g - \langle X \rangle)(Y_g - \langle Y \rangle).\end{aligned}$$

are the estimates of the mean, variance and covariance, taken over all background genotypes  $g$ .

We find that the uncorrected correlation coefficient

$$\rho(\Delta\lambda_{mge}, \lambda_{ge}) = \frac{\text{Cov}(\Delta\lambda_{mge}, \lambda_{ge})}{\sqrt{\text{Var}(\lambda_{ge}) \text{Var}(\Delta\lambda_{mge})}}$$

deviates very little from the corrected correlation coefficient  $\rho'$  given by equation (2.15) (see Figure B.2D).

Second, for each mutation in each environment, we estimate the slope  $b$  of the regression of the growth rate  $\lambda$  of background strain against the mutant growth rate  $\lambda^+$  (we drop the indices for convenience). We then estimate the mutation's global epistasis slope as  $b - 1$ . We find that 255/546 (47%) global epistasis slopes estimated in this way are significantly different from zero and very close to slopes estimated using equation (2.2) in the main text (Figure B.2E).

These analyses confirm that the global epistasis trends that we observe are not spurious.

### **Comparing distribution of slopes and intercepts across environments.**

We tested how the distributions of fitted slopes and intercepts vary across environments using three different pairwise tests (Figure B.4A).

1. A Kolmogorov–Smirnov test assess the overall differences between two distributions.
2. A paired t-test assess the differences between the means of the two distributions.

3. An F-test assess the difference between the variances of the two distributions.

All tests were performed using the stats package in R, and the raw  $P$ -values were adjusted using the Benjamini-Hochberg multiple testing correction. Adjusted  $P$ -values are reported in Figure B.4A.

### **Comparison of the variable slopes and invariant slopes models for individual mutations.**

In addition to the full “variable slopes” model (equation (2.1) in the main text) in which a mutation can have different slopes in different environments, we also fit an “invariant slopes” model to our data in which every mutation has a single environment-invariant slope.

We compared that variable and invariant slopes models using the likelihood ratio test and found that, for the majority of mutations, we could not reject the invariant slopes model in favor of the variable slopes model which has 5 more parameters (Figure B.5A). We then calculated the adjusted  $R^2_{\text{adj}}$  for both the invariant and variable slopes models for each mutation using the R function `lm`,

$$R^2_{\text{adj}} = 1 - \left( (1 - R^2) \frac{n - 1}{n - k - 1} \right),$$

where  $R^2$  is the standard coefficient of determination,  $n$  is the number of observations and  $k$  is the number of predictors. In this case,  $n$  is the number of unique genotype-environment combinations in which the mutation is measured, and  $k$  is twice the number of environments in which the mutation is measured (variable slopes model) or the number of such environments plus one (invariant slopes model). This adjustment helps identify potential over-fitting by penalizing a high number of parameters relative to the number of observations.

To assess if the similarity of the variant and invariant slopes models could be driven by missing measurements, we apply the same procedure on the reduced data set which contains, for each mutation, only those strains for which that mutation has a measured effect in all environments. We find that the results are very similar in both data sets (Figure B.5D-E).

### **Analysis of microscopic epistasis slopes.**

To determine whether global-epistasis slopes for a given mutation are statistically distinguishable across environments, we estimate these slopes as described above using the `lm` function in R. Along with the maximum likelihood estimates of the slopes, this function returns the standard errors of these estimates. Then, for each pairwise slope comparison, we calculate the difference between the slopes and estimate the associated error variance as the square root of the summed squared errors. Assuming that errors are normally distributed, we calculate the  $P$ -value for the observed error. We then apply the Benjamini-Hochberg multiple testing correction for all pairwise comparisons for a given mutation and obtain adjusted  $P$ -values. A pair of slopes is then called significantly different if the adjusted  $P$ -value is below 0.05.

To assess if the cross-environmental similarity of global epistasis slopes could be driven by missing measurements, we apply the same procedure on the reduced data set which contains, for each mutation, only those strains for which that mutation has a measured effect in all environments. We find that, similar to the full data, 87% of pairwise slope comparisons are statistically indistinguishable (Figure B.5F).

We also tested whether the slope similarity was driven primarily by “flat” mutations, i.e., those where slopes were not significantly different from zero. To this end, we removed those 258/546 (47%) of mutation-environment combinations for which the slope of the fit of (1) in the main text was not significantly different from zero. Of the remaining 415 pairwise slope comparisons, 356 (86%) were not significantly different, suggesting that the “flat” mutations are not the main driver of the observed similarity of slopes across environments.

### **Determining the pivot growth rate and its variation across environments.**

There are two ways of determining the pivot growth rate. First, for each mutation  $m$  in environment  $e$ , we can find the background growth rate  $\bar{\lambda}_{me}$  at which the mutation’s expected fitness effect is zero. Using equation (2.1) in the main text (with  $b_{me} = b_m$ ), we obtain an

equation for  $\bar{\lambda}_{me}$ ,

$$0 = a_{me} + b_m \bar{\lambda}_{me},$$

which yields  $\bar{\lambda}_{me} = -a_{me}/b_m$ . We can then obtain the pivot growth rate  $\bar{\lambda}_e$  for environment  $e$  by averaging  $\bar{\lambda}_{me}$  over all mutations  $m$ . The drawback of this approach is that it produces a value of  $\bar{\lambda}_e$  regardless of whether different mutations actually switch sign at similar background growth rates or not.

The second approach, which is the one we have taken, starts by asking whether  $a_{me}$  and  $b_m$  are correlated. A priori,  $a_{me}$  and  $b_m$  need not be correlated, or, if they are correlated, we have no a priori constraints on the slope or intercept of this correlation. In fact, we find that the relationship between  $a_{me}$  and  $b_m$  is very well described by the statistical model with a zero intercept and slope  $c_e < 0$  (see Figure 2.2C in the main text),

$$a_{me} = c_e b_m + \eta_{me}. \quad (2.16)$$

Substituting equation (2.16) into equation (2.1) in the main text (with  $b_{me} = b_m$ ) yields

$$\Delta\lambda_{mge} = b_m (\lambda_{ge} + c_e) + \xi_{mge} + \eta_{me},$$

which shows that  $\bar{\lambda}_e = -c_e$ , i.e.,  $c_e$  is the growth rate at which a typical mutation switches sign. Thus, we estimate the pivot growth rate  $\bar{\lambda}_e$  in each environment  $e$  as the negative of the best-fit ordinary linear regression slope in equation (2.16) using the `lm` package in R.

To test if pivot growth rates significantly differ across environments, we follow a procedure analogous to the one used to test pairwise differences in microscopic epistasis slopes (see above), and we find that pivot growth rates indeed significantly differ in almost all comparisons (Figure B.4B). Additionally, we find that the pivot growth rate estimated with the invariant slopes  $b_m$  well correlate with those estimated using variable slopes  $b_{me}$  (Figure B.5C).

## Variance partitioning for the sign of mutations

Let  $Y_{mge}$  be the observed sign of mutation  $m$  in genetic background  $g$  in environment  $e$ , such that  $Y_{mge} = \pm 1$ . The total variance in the observed signs of mutations is

$$V^{\text{tot}} = \frac{1}{K-1} \sum_{m,g,e} (Y_{mge} - \bar{Y})^2,$$

where

$$\bar{Y} = \frac{1}{K} \sum_{m,g,e} Y_{mge}$$

is the average sign of the mutational effect and  $K = \sum_{m,g,e} 1$  is the total number of mutations measured across all genotypes and environments.

Let the true fitness effect of mutation  $m$  in genetic background  $g$  in environment  $e$  (without measurement noise) be

$$s_{mge} = G_{mge} + \xi_{mge},$$

where  $G_{mge} = a_{me} + b_{me} \lambda_{ge}$  is the global epistasis term,  $\lambda_{ge}$  is the growth rate of background strain  $g$  in environment  $e$ , and  $\xi_{mge}$  is the idiosyncratic epistasis term (see equation (2.1) in the main text). Then, the probability that the observed sign of this mutation is positive is

$$p_{mge}^{\text{gl+id}} = \Pr(Y_{mge} = 1) = \frac{1}{\sqrt{2\pi\sigma_{mge}^2}} \int_0^\infty \exp\left[-\frac{(x - s_{mge})^2}{2\sigma_{mge}^2}\right] dx, \quad (2.17)$$

where  $\sigma_{mge}^2$  is the variance of the measurements noise for mutation  $m$  in background  $g$  in environment  $e$ . The super-index gl + id indicates that this probability takes into account both global and idiosyncratic epistasis. We estimate  $p_{mge}^{\text{gl+id}}$  using equation (2.17) with  $s_{mge}$  and  $\sigma_{mge}$  given by equations (2.10) and (2.11). This allows us to calculate the expected sign of mutation  $m$  in background  $g$  in environment  $e$ ,

$$\bar{Y}_{mge}^{\text{gl+id}} = 2p_{mge}^{\text{gl+id}} - 1$$

and estimate the variance in the mutational sign attributed solely to measurement noise,

$$V^n = \frac{1}{K-1} \sum_{m,g,e} \left( Y_{mge} - \bar{Y}_{mge}^{\text{gl+id}} \right)^2. \quad (2.18)$$

In a model without idiosyncratic epistasis, the probability  $p_{mge}^{\text{gl}}$  that the observed sign of the mutational effect is positive can be estimated using the same equation (2.17), but with

$$\begin{aligned} s_{mge} &= G_{mge}, \\ \sigma_{mge}^2 &= \frac{1}{M_{me}(M_{me}-1)} \sum_{r,t,g'} \sum_{k \in S_{mg're}} (\Delta\lambda_{kret} - G_{mg'e})^2, \end{aligned}$$

where  $\Delta\lambda_{kret}$  are estimated with equation (2.4) and  $M_{me}$  is the number of barcode-genotype-replicate-time combinations at which  $\Delta\lambda_{kret}$  are estimated for mutation  $m$  and environment  $e$ .

Thus, the variance in the mutational signs attributed to both idiosyncratic epistasis and measurement noise is

$$V^{\text{id+n}} = \frac{1}{K-1} \sum_{m,g,e} \left( Y_{mge} - \bar{Y}_{mge}^{\text{gl}} \right)^2,$$

where

$$\bar{Y}_{mge}^{\text{gl}} = 2p_{mge}^{\text{gl}} - 1$$

is the expected sign of mutation  $m$  in background  $g$  in environment  $e$  in the model without idiosyncratic epistasis. Thus, we can partition the total variance  $V^t$  into the global, idiosyncratic and noise components as follows. Variance  $V^n$  given by equation (2.18) is attributed to noise, variance

$$V^{\text{id}} = V^{\text{id+n}} - V^n$$

is attributed to idiosyncratic epistasis, and variance

$$V^{\text{gl}} = V^{\text{tot}} - V^{\text{id+n}}$$

is attributed to global epistasis.

### Empirical DFEs

For YPD, we use the estimates of DFE moments and their corresponding standard errors obtained in Ref. [119]. In each other environment  $e$  and for each background strain  $g$ , we use the fitness effect estimates  $\Delta\lambda_{mge}$  obtained from equation (2.10) to estimate the mean  $\langle\Delta\lambda\rangle_{ge}$ , variance  $\text{Var}_{ge}[\Delta\lambda]$  and skewness  $\text{Skew}_{ge}[\Delta\lambda]$  of the empirical DFEs as

$$\begin{aligned}\langle\Delta\lambda\rangle_{ge} &= \frac{1}{K_{ge}} \sum_m \Delta\lambda_{mge}, \\ \text{Var}_{ge}[\Delta\lambda] &= \frac{1}{K_{ge}-1} \sum_m \left( \Delta\lambda_{mge} - \langle\Delta\lambda\rangle_{ge} \right)^2, \\ \text{Skew}_{ge}[\Delta\lambda] &= \frac{\frac{1}{K_{ge}} \sum_m \left( \Delta\lambda_{mge} - \langle\Delta\lambda\rangle_{ge} \right)^3}{(\text{Var}_{ge}[\Delta\lambda])^{\frac{3}{2}}}.\end{aligned}$$

Here,  $K_{ge}$  is the number of mutations whose effects were estimated in background strain  $g$  in environment  $e$ . To estimate the uncertainty in these estimates, we resampled 70 random mutations from each empirical DFE with replacement (bootstrapping). For each resampled mutation, we drew its fitness effect from a normal distribution with mean  $\Delta\lambda_{mge}$  (given by equation 2.10) and standard deviation  $\sigma_{mge}^{\Delta\lambda}$  (given by equation 2.11). We performed 300 iterations of this procedure.

### Sensitivity of DFE moment estimates to missing measurements.

In pooled cultures, slow growing mutants may go extinct during the competition assay, which could prevent us from estimating their effects and lead to biases in our estimates of the DFE moments. In particular, highly deleterious mutations missing from the data could lead us to overestimate the DFE mean, underestimate the DFE variance and overestimate DFE skewness. Furthermore, we expect that these biases would be stronger in slower growing background strains, which could produce spurious declines in DFE mean and skewness with the background-strain growth rate.

As described above, we sought to mitigate this potential issue experimentally by compet-

ing our mutants in groups with similar growth rates (see Section 2.3.2). However, these spurious effects may still be present in our data. To investigate how severe these effects might be, we plotted the number of mutations for which we have reliable fitness-effect estimates against the respective background strain growth rate. We found that the number of mutations per strain varies little with the background-strain growth rate in 30°C environments (Figure B.10B), but it does vary in the expected direction in 37°C environments. However, since this analysis does not reveal the fitness effects of missing mutations, this observation alone does not imply that our estimates of DFE moments are biased for slow-growing strains in 37°C environments. Thus, we carried out three additional analyses to further probe the extent of these potential biases.

First, we imputed the effect of the missing mutation measurements in the 37°C environments to assess if we expect to have missed many deleterious mutations. We are only able to do this for those missing mutation-environment-strain combinations in which (i) there are sufficient measurements of the mutation in the focal environment to allow the estimation of microscopic epistasis regression, and (ii) we have an estimate of the background strain growth rate in the focal environment. Out of a total of 4252 missing mutation-environment-strain combinations, 3568 (84%) are imputable. We then calculate the predicted effect of the missing mutation-strain-environment combination in the absence of idiosyncratic epistasis using (1) in the main text with the fitted parameters and the estimated growth rate of the background strain and use it as the imputed value. We find that, while the distributions of imputed and measured fitness effects (combined across all three 37°C environments) are statistically different, the average imputed fitness effect is only by  $0.002 \text{ h}^{-1}$  ( $P = 7 \times 10^{-7}$ , t-test; Figure B.10C), which is much smaller than  $0.026 \text{ h}^{-1}$ , the standard deviation of the distribution of measured effects.

Second, we eliminated all strains from our analysis for which we measured less than 60 mutations (bottom 24%) and replotted DFE moments for this reduced number of strains (Figure B.10D–F) and found that all the trends found in the full data set remain in this reduced data set (compare Figure B.10D–F and 2.4F–H in the main text).

Third, in each environment, we identified the set of 40 mutations all of which were



measured in the maximum number of strains in that environment, which ranged from 16 to 40 strains, depending on environment. We then restricted our analysis of DFE moments to these strains and mutations, thereby creating a reduced data set without any missing measurements. We found that the DFE moments recapitulate the trends reported for the full data set (compare Figures B.10G–I and 2.4F–H in the main text).

Based on these analyses, we conclude that the trends in the DFE moments that we report are not spurious results of missing measurements.

### **Variation of the DFEs with adjusted growth rate.**

We carried a series of pairwise DFE comparisons across strains and/or environments. To this end, we created matched pairs of background strains using three methods:

1. *Adjusted growth rate matching.* We matched each background strain  $g_1$  in environment  $e_1$  with another background strain  $g_2 \neq g_1$  in a different environment  $e_2 \neq e_1$ , such that the adjusted growth rate  $\lambda_{g_2 e_2}^*$  of the latter strain was most similar to the former strain among all other strain-environment combinations.
2. *Raw growth rate matching.* We carried out the same procedure as above except we matched raw growth rates.
3. *Strain matching.* We matched each background strain  $g_1$  in environment  $e_1$  with itself in another environment  $e_2$  such that  $\lambda_{g_1 e_2}$  was the closest to  $\lambda_{g_1 e_1}$ .

Then, we compared the DFEs of the two matched strain-environment combinations using four metrics of similarity shown in Figure B.9.

### **QTL analysis**

With only 42 background strains, we have little power to identify QTLs *de novo*. Instead, we assess whether certain loci identified in a previous study [125] help explain some of the idiosyncratic epistasis. Specifically, Bloom et al identified 37 QTLs as having significant

explanatory power for the growth rate of strains in environments with pH and high temperature stress (e.g., environments YNB pH 3, YNB pH 8 and YPD 37C). Since testing the effect of all 37 loci jointly with 42 strains would lead to overfitting, we chose to restrict our analysis to the top 10 QTLs on different chromosomes<sup>1</sup> that explain most variance in growth rate in any of the environments. The list of QTLs that we chose to test is provided in Table S6-Tab 2. For each QTL, we found an individual site variable in our strains that was closest to the position of the QTL peak to represent that QTL. We then tested whether these QTL-representative sites have significant explanatory power when added to our statistical model.

First, we used ANOVA to determine whether these 10 loci explain variance in background strain growth rate in each environment at a significance level of  $P < 0.05$ . We found that the maximum number of significant loci in any environment was two, and that the total variance in strain growth rate attributable to these QTLs ranged between 0 and 52% (Figure B.11A). The strongest QTL, which explains 51.7% and 52.4% of growth rate variance in 37°C, pH 3.2 and 5.0 environments, is associated with the KRE33 locus on Chromosome 14, consistent with previous studies [119, 132]. We then used ANOVA to determine whether the addition of these 10 loci to the generalized global epistasis model (equation (2.2) in the main text) significantly reduced the proportion of unexplained variance in fitness effect for each mutation. If the contribution of a locus was significant at  $P$ -value 0.05, then we also calculated the fraction of variance explained by this locus above and beyond the generalized global epistasis model.

Out of a total of 93 mutations in which our global epistasis models explained some variance in their fitness effects, we found three mutations (in genes UBP3 (ID 95), in GET2 (ID 61), and nearby WHI2 (ID 57)) where 27%, 23%, and 20% of overall variance was jointly explained by the 10 candidate QTLs, corresponding to 50%, 39%, and 37% of the explained variance in each mutation, respectively. For the remaining 97% (90/93) mutations, the 10 candidate loci together explain less than 11% of overall variance, with a median of only 2.1% (interquartile interval [0.8%, 3.7%]), corresponding to a median of 4.4% of the total explained

---

<sup>1</sup>This was done to avoid testing correlated effects of linked QTLs.

variance within each mutation (interquartile interval [1.3%, 9.6%], Supplementary Table S6-Tab 9).

### **GO-term enrichment analysis**

In an attempt to identify possible mechanistic signals underpinning our observations, we performed a GO-term enrichment analysis. First, we tested whether our set of tn-insertions was significantly enriched for any GO terms, using the `clusterProfiler` tool and the *S. cerevisiae* reference database from R's Bioconductor suite. We called a GO term as enriched if it passed the 0.05 significance threshold after the Benjamini-Hochberg multiple testing correction. We found 11 significantly enriched GO terms, including regulation of transcription by RNA polymerase II and double-strand break repair via nonhomologous end joining, with the enrichment ranging from 3- to 10-fold compared to the random expectation (Table B.4). Thus, our set of mutations is not entirely random, which is expected given that we have excluded lethal and unconditionally neutral mutations (see Section 2.3.1).

We then asked whether mutations with certain global microscopic epistasis slopes were enriched for some GO terms relative to the overall set. To this end, we separated mutations into four groups, based on the quantiles of the distribution microscopic epistasis slopes  $b_m$  shown in Figure 2.2A in the main text. Then, for each group, we generated the empirical null distribution of each GO-term count by randomly resampling GO terms 10,000 times from the set of terms for all mutations. This allowed us to calculate the empirical  $P$ -value for all observed GO-term counts in each group. After applying the Benjamini-Hochberg multiple testing correction, we found no terms in any group that were either significantly enriched or depleted relative to the full set of mutations.

### **2.3.4 Theoretical calculation of DFE moments**

Here we derive the moments of the distribution of fitness effects (DFE) of mutations from the generalized global epistasis equation (equation (2.2) in the main text). Since we consider the

environment fixed, we drop the subindex  $e$ . To simplify notations, we will denote the adjusted growth rate of the background strain  $g$  by  $F_g \equiv \lambda_g^*$  and we denote the fitness effect of a mutation in this background by  $s_g \equiv \Delta\lambda_g$ . To derive DFE moments, we assume that the effects of mutations  $s$  are drawn from a continuous distribution defined by equation

$$s_g = bF_g + \eta + \xi_g, \quad (2.19)$$

which is the continuous analog of equation (2.2) in the main text. Here  $b$  and  $\eta$  are the slope and the y-intercept of the focal mutation, and  $\xi_g$  is the idiosyncratic epistasis of this mutation in the background strain  $g$ . We assume that  $b$  and  $\eta$  are independent, and that  $b$  has probability density  $p_{sl}$ ,  $\eta$  is normally distributed with zero mean and variance  $\sigma_{\text{pivot}}^2$ . We also assume that  $\xi_g$  is normally distributed with zero mean and variance  $\sigma_{\text{id}}^2$  (which can in principle depend on  $b$ , see [124]). Thus, conditional on  $b$  and  $\eta$ , the fitness effects  $s_g$  of the mutation in the background  $g$  is a normal random variable with distribution with mean  $\langle s|b, \eta \rangle = bF_g + \eta$  and variance  $\sigma_{\text{id}}^2$ . Then, the DFE  $p_g(s)$  in the background  $g$  with adjusted growth rate  $F_g$  is given by

$$p_g(s) = \int_{-\infty}^{\infty} db p_{sl}(b) \int_{-\infty}^{\infty} d\eta N(\eta; 0, \sigma_{\text{pivot}}^2) N(s; bF_g + \eta, \sigma_{\text{id}}^2), \quad (2.20)$$

where  $N(x; \mu, \sigma^2)$  is the normal probability density with mean  $m$  and variance  $\sigma^2$ . Since

$$N(s; bF_g + \eta, \sigma_{\text{id}}^2) = N(\eta; s - bF_g, \sigma_{\text{id}}^2)$$

and since

$$N(x; \mu_1, \sigma_1^2) N(x; \mu_2, \sigma_2^2) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) N(x, \mu_3, \sigma_3^2),$$

where  $\mu_3 = \frac{\sigma_1^{-2}}{\sigma_1^{-2} + \sigma_2^{-2}} \mu_1 + \frac{\sigma_2^{-2}}{\sigma_1^{-2} + \sigma_2^{-2}} \mu_2$  and  $\sigma_3^2 = \frac{1}{\sigma_1^{-2} + \sigma_2^{-2}}$  for any  $\mu_1, \mu_2, \sigma_1^2$  and  $\sigma_2^2$ , the integral with respect to  $\eta$  can be taken, such that the expression (2.20) simplifies to

$$p_g(s) = \int_{-\infty}^{\infty} db p_{\text{sl}}(b) N(s; bF_g, \tilde{\sigma}^2), \quad (2.21)$$

where we denoted  $\tilde{\sigma}^2 = \sigma_{\text{pivot}}^2 + \sigma_{\text{id}}^2$ .

Equation (2.21) allows us to compute the DFE mean as

$$\langle s \rangle = \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \int_{-\infty}^{\infty} ds s N(s; bF_g, \tilde{\sigma}^2) = \langle b \rangle F_g \quad (2.22)$$

and higher central moments of the DFE as

$$M^{(n)}[s] = \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \int_{-\infty}^{\infty} ds (s - \langle b \rangle F_g)^n N(s; bF_g, \tilde{\sigma}^2) \quad (2.23)$$

Using expression (2.23) and the fact that

$$s - \langle b \rangle F_g = (s - bF_g) + (b - \langle b \rangle) F_g$$

we obtain the following explicit expressions for the DFE variance  $\text{Var}[s]$  and its third central moment  $M^{(3)}[s]$ ,

$$\begin{aligned} \text{Var}[s] &\equiv M^{(2)}[s] = \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \left[ \tilde{\sigma}^2 + (b - \langle b \rangle)^2 F_g^2 \right] \\ &= \text{Var}[b] F_g^2 + \sigma_{\text{pivot}}^2 + \int_{-\infty}^{\infty} p_{\text{sl}}(b) \sigma_{\text{id}}^2 db, \end{aligned} \quad (2.24)$$

$$\begin{aligned} M^{(3)}[s] &= \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \left[ 3\tilde{\sigma}^2 (b - \langle b \rangle) F_g + (b - \langle b \rangle)^3 F_g^3 \right] \\ &= M^{(3)}[b] F_g^3 + 3F_g \int_{-\infty}^{\infty} p_{\text{sl}}(b) \sigma_{\text{id}}^2 (b - \langle b \rangle) db. \end{aligned} \quad (2.25)$$

Here  $\text{Var}[b]$  and  $M^{(3)}[b]$  are the variance and the third central moment of the distribution  $p_{\text{sl}}(b)$

of global epistasis slopes, respectively.

We find that the variance of residuals and slopes are correlated (Figure B.4D). Thus, we set  $\sigma_{\text{id}}^2 = -\alpha b$ , and the expressions (2.24) and (2.25) become

$$\text{Var}[s] = -\alpha \langle b \rangle + \sigma_{\text{pivot}}^2 + \text{Var}[b] F_g^2, \quad (2.26)$$

$$M_g^{(3)}[s] = -3\alpha \text{Var}[b] F_g + M^{(3)}[b] F_g^3. \quad (2.27)$$

and the skewness of the DFE is given by

$$\text{Skew}_g[s] = \frac{M_g^{(3)}[s]}{(\text{Var}_g[s])^{3/2}} = \frac{-3\alpha F_g / \sqrt{\text{Var}[b]} + \text{Skew}[b] F_g^3}{\left( \frac{-\alpha \langle b \rangle + \sigma_{\text{pivot}}^2}{\text{Var}[b]} + F_g^2 \right)^{3/2}}. \quad (2.28)$$

Equations (2.22), (2.26) and (2.28) show that when  $F_g = 0$ , DFE mean and skewness are zero and the DFE variance achieves its minimal value  $\sigma_{\text{id}}^2$ . Furthermore, since  $\text{Skew}[b] < 0$ , DFE skewness monotonically declines from  $-\text{Skew}[b] > 0$  to  $\text{Skew}[b] < 0$ .

To show that all odd moments of the DFE vanish when  $F_g = 0$ , we notice that, when  $F_g = 0$ , equation (2.21) simplifies to

$$p_g(s) = N(s; 0, \tilde{\sigma}^2), \quad (2.29)$$

i.e., the DFE is a normal distribution, which implies that all its odd central moments vanish.

To show that all even moments of the DFE reach their minimum at  $F_g = 0$ , we differentiate expression (2.23) with respect to  $F_g$  at  $F_g = 0$  and obtain

$$\left. \frac{dM^{(n)}[s]}{dF} \right|_{F_g=0} = \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \left[ -n \langle b \rangle M^{(n-1)}[s|b] + \frac{b}{\tilde{\sigma}^2} M^{(n+1)}[s|b] \right].$$

where  $M^{(n)}[s|b] = \int_{-\infty}^{\infty} s^n N(s; 0, \tilde{\sigma}^2) ds$  is the  $n$ th central moment of a normal distribution with mean zero and variance  $\tilde{\sigma}^2$ . Therefore, when  $n$  is even,  $\left. \frac{dM^{(n)}[s]}{dF} \right|_{F_g=0}$  vanishes because all odd

central moments of a normal distribution are zero. To see that  $M^{(n)}[s]$  achieve their minimum at  $F_g = 0$  for any even  $n$ , we find that the second derivative of  $M^{(n)}[s]$  at  $F_g = 0$  is given by

$$\begin{aligned} & \left. \frac{d^2 M^{(2k)}[s]}{dF^2} \right|_{F_g=0} \\ &= \int_{-\infty}^{\infty} db p_{sl}(b) \left[ n(n-1) \langle b \rangle^2 M^{(n-2)}[s|b] - \frac{b}{\tilde{\sigma}^2} (b + 2n \langle b \rangle) M^{(n)}[s|b] + \frac{b^2}{\tilde{\sigma}^4} M^{(n+2)}[s|b] \right]. \end{aligned}$$

The even moments  $M^{(n)}[s|b]$  of the normal distribution  $N(s; 0, \tilde{\sigma}^2)$  can be expressed as  $\tilde{\sigma}^n (n-1)!!$  where  $(n-1)!! = (n-1)(n-3) \cdots 3 \cdot 1$  is the double factorial. Therefore, we have

$$\left. \frac{d^2 M^{(2k)}[s]}{dF^2} \right|_{F_g=0} = (n-1)!! n \int_{-\infty}^{\infty} p_{sl}(b) \tilde{\sigma}^{n-2} (b - \langle b \rangle)^2 db > 0$$

for any even  $n$ , which implies that  $M^{(2k)}[s]$  indeed achieves its minimum at  $F_g = 0$ .

### Parameterization of DFE moments.

The predicted DFE moment curves in Figure 2.4 C–H are calculated with equations (2.22), (2.26), and (2.28) for DFE mean, variance and skewness, respectively. These equations depend on the moments of the distribution of slopes  $p_{sl}(b)$  and parameter  $\alpha$ , which determines how variance in idiosyncratic epistasis depends on the global-epistasis slope. We estimated the moments of  $p_{sl}(b)$  from the empirical distribution of global epistasis slopes obtained from fitting equation (2.2) in the main text (see Figure 2.2A). We estimated the value of the parameter  $\alpha = 1.37 \times 10^{-3}$  by regressing the variance  $\sigma_{id}^2$  of the idiosyncratic epistasis residuals  $\xi_{mge}$  against the microscopic epistasis slopes  $b_m$ , with the regression constrained to pass through the origin (see Figure B.4D).

To quantify how well these theoretical predictions fit our data, we compare them to the best-fit polynomial of the corresponding degree (Figure 2.4 F–H). Specifically, for mean and variance, we use the `lm` function in R to fit a first and second degree polynomial, respectively. For skewness, we utilize the predicted variance values from the fit of the second degree polynomial

to parameterize the denominator of the rational function in equation (2.28). We then fit the best fit third degree polynomial in the numerator. The fitted and theoretically predicted parameters are shown in Table B.3.



## 2.4 Supplementary Text

### 2.4.1 Justification for using absolute growth rate rather than relative fitness

As far as we know, almost all previously reported instances of global epistasis were observed with respect to competitive fitness in batch culture (but see Ref. [104]). In contrast, we designed our study so that we can estimate not only competitive fitness but also the absolute growth rates of our background strains and mutants. In some ways, these two choices are equivalent. Specifically, (i) the fitness effects of mutations given by equation (2.10) in fact depend only on the relative barcode frequencies (see equation (2.4)); and (ii) the competitive fitness of background strains relative to each other can be calculated by taking differences between their estimated absolute growth rates. Thus, most of our results would not be impacted by our choice. However, we see three advantages of using the absolute growth rate.

1. Using absolute growth rate makes our findings independent of any particular reference strain and therefore potentially more generalizable to other systems.
2. To the best of our knowledge, almost all previous observations of global epistasis were made with respect to competitive fitness. However, competitive fitness in batch-culture conditions depends (non-linearly) on multiple components, such as the duration of the lag phase, exponential growth rate, survival during starvation, etc. [65, 137–139]. Thus, it is conceivable that epistasis could arise at the level of competitive fitness even among mutations that exhibit no epistasis for any of the individual components of fitness. Our observation that global epistasis holds at the level of growth rate strongly suggests that it is an intrinsic property of the cell. This narrows the scope of possible mechanistic explanations for this phenomenon and suggests that it would be relevant in any environment where growth is an important component of fitness.
3. Absolute growth rate provides us with a concrete threshold for strain viability. In other

words, there is no a priori minimum on a relative growth rate scale, whereas the absolute scale naturally cuts off at zero. This fact allows us to attribute meaning to the intercepts of microscopic global epistasis regressions and to the location of the pivot growth rate relative to zero. In particular, a model in which global epistasis lines intersect below viability would have been consistent with our a priori intuition that most mutations are unconditionally deleterious and that DFEs for all viable genotypes are negatively skewed.

## 2.4.2 Comparison with previous studies and potential limitations

As mentioned in the main text, many studies carried out in several microbial species found evidence of microscopic global epistasis [97, 98, 114–119, 121, 122]. There are also a few recent studies that found evidence for macroscopic global epistasis [58, 113, 114, 119]. It is instructive to compare our results with these previous works. We focus here on the four most relevant studies, Refs. [58, 104, 113, 114], grouped by the types of mutations that they have measured.

### **Diaz-Colunga et al, Ref. [104].**

Diaz-Colunga et al examined how the patterns of microscopic global epistasis within a combinatorially complete set of mutations in the *DHFR* gene (from *Plasmodium falciparum*, integrated into the genome of *S. cerevisiae*) vary across environments where the concentration of the antimalarial drug pyrimethamine varies over five orders of magnitude. They found that the patterns of global epistasis vary across this wide range of drug concentrations. However, interestingly, these patterns remain largely invariant for concentrations below 1  $\mu\text{M}$ , despite the fact that the growth rates of mutants and their rank order change over this range [140]. There are many differences between our experimental design and that of Ref. [104], e.g., types and number of mutations, background strains, type and degree of environmental variation. So, it is difficult to draw definitive conclusions. However, one intriguing possibility is that the near-invariance of the global epistasis patterns across environments has to do with how distributed or localized the environmental perturbation is. For example, it is possible that at low concentrations (below

1  $\mu\text{M}$ ) the drug exhibits a global toxic effect on the cell, similar to our temperature and pH perturbations, and global epistasis patterns remain largely invariant. At high concentrations (above 1  $\mu\text{M}$ ), the toxicity may come primarily from the drug's interaction with its target protein, dihydrofolate reductase enzyme, and the patterns of epistasis become highly idiosyncratic.

**Aggeli et al, Ref.[58]**

Aggeli et al used the barcode-lineage tracking approach [53] to measure how the beneficial part of the DFE (bDFE) varies across four strains of *S. cerevisiae*, one ancestral strain and three adapted mutants each carrying a single beneficial mutation. They found that the bDFE in all three adapted strains is shifted left relative to that of the ancestral strain, which is qualitatively consistent with our results.

**Johnson and Desai, Ref. [114] and Couce et al, Ref. [113].**

Both Johnson and Desai and Couce et al use the transposon-insertion mutagenesis followed by sequencing (TnSeq) approach to measure how the DFE changes along an evolutionary trajectory over about 10 to 15 thousand generations. While conceptually very similar, these studies differ in two aspects. First, Johnson and Desai sampled strains from six time points of the long-term evolution experiment carried out in budding yeast *S. cerevisiae* in two environments, YPD 30°C and SC 37°C<sup>2</sup>, whereas Couce et al sampled strains from two to three time points of the Lenski's long-term evolution experiment (LTEE) in bacterium *Escherichia coli*. The second difference is that Johnson and Desai used the same TnSeq approach as we employed in the present study, i.e., they measured the effects of the same  $\sim 100$  tn-insertions, whereas Couce et al used "saturated" TnSeq approach, with  $\sim 10^5$  tn-insertions covering 78% of all *E. coli* genes.

Differences in experimental design constrained the power of the two studies in different ways. While both studies found the vast majority of tn-insertions to be deleterious, Couce et al were able make statistical statements about beneficial mutations whereas Johnson and Desai focused only deleterious mutations. Johnson and Desai found evidence for global microscopic

---

<sup>2</sup>Unlike the environments in the present work, the SC 37°C medium in Ref. [114] was not buffered.

epistasis in both environments, although the fraction of variation explained by background fitness was generally smaller than in Ref. [119] where the background strains were derived from the RM×BY cross, as in the present work. Couce et al found significant differences in the effects of many tn-insertions across genetic backgrounds (i.e., evidence of microscopic epistasis), but they likely did not have enough power to detect any global component in it. However, microscopic global epistasis has been previously observed among adaptive mutations in the LTEE [97, 123].

Focusing now on macroscopic epistasis, Couce et al found that evolved clones had a clearly depleted bDFE compared to the LTEE ancestor, which is consistent with our model as well as with previous observations of “the rule of declining adaptability” [94, 141]. However, they detected no systematic shifts in the overall DFE over the course of evolution, which is inconsistent with our results. Johnson and Desai found fitness-dependent shifts in the DFE among clones evolved in the YPD 30°C consistent with our results and Ref. [119]. However, despite the presence of global microscopic epistasis, they found no consistent fitness-dependent shifts in the DFE among clones evolved in SC 37°C, similar to the results by Couce et al.

One possible technical reason that could explain the lack of a detectable shift in the DFE in the Couce et al study is that their DFEs appear to be dominated by neutral or nearly neutral mutations, similar to the “large library” investigated in Ref. [119]. If the majority of mutations are in fact unconditionally neutral (or at least neutral in all investigated strains), detecting shifts in the distribution of fitness effects of non-neutral mutations becomes statistically challenging. However, this cannot explain the why Johnson and Desai failed to observe DFE shifts in the SC 37°C environment because they specifically focused on non-neutral mutations and had power to detect such shifts.

Patterns of global epistasis could be fundamentally different between *S. cerevisiae* and *E. coli* or they might differ between exponential growth (used in our study) and batch culture fitness (used in both Couce et al and Johnson et al). Yet, neither of these explain the puzzling differences between yeast strains evolved in two environments observed by Johnson and Desai.

Perhaps a more likely biological explanation for the aforementioned discrepancies is that

epistasis involving mutations sampled along an evolutionary trajectory may be different than epistasis involving random genetic variants [142]. The background strains used in this study are hybrids derived from the RM×BY cross and may represent more or less random genetic variation<sup>3</sup>, or a least variation that has not been under selection in our experimental conditions. In contrast, background genotypes in the studies by Johnson and Desai and by Couce et al were sampled from adapting populations and thus represent selected (i.e., non-random) genetic variation. Moreover, mutations that drive adaptation to different environments may have different physiological effects and, as a result, exhibit different types of epistasis. Indeed, Johnson and Desai speculate that adaptation to YPD 30°C may be primarily driven by mutations in core cellular processes related to growth while adaptation to SC 37°C may be driven at least in part by mutations improving survival during heat stress; they then argue that these types of mutations could plausibly exhibit different types epistasis, which could explain the observed discrepancy between conditions.

Another intriguing possibility is that evolved genotypes could have experienced second-order selection for higher evolvability and/or lower robustness. In fact, such selection has been documented in the LTEE [143]. Generally, second-order selection favors strains with DFEs that have heavier beneficial tails and/or weaker deleterious tails. Thus, even if the overall fitness landscape is characterized by the statistical trends shown in Figure 2.4 in the main text, second-order selection could in principle favor genotypes that deviate from these trends.<sup>4</sup> Of course, the degree of such bias would be determined by the amount of DFE variation available in the population and on the strength of second-order selection, both of which could vary between environments. This could explain difference between the results of our study and those of Couce

---

<sup>3</sup>According to the Reddy-Desai model, the pivot growth rate equals the average fitness of all genotypes [124]. The fact that the average growth rate of our background strains correlates with the pivot growth rate (Figure B.4G) suggests that our background genotypes are not particularly exceptional.

<sup>4</sup>Our focus in this paper has been on the deterministic fitness-correlated trends in the shape of the DFE, but Figure 2.4 D shows that there is also a considerable amount of variation around these trends on which second-order selection could potentially act. Furthermore, Johnson and Desai find that clones sampled from the same population at the same time point sometimes have significantly different DFEs, which further supports the existence of meaningful DFE variation on which second-order selection could act.

et al and Johnson and Desai.

Overall, this discussion highlights that extrapolating our results to predict the outcomes of evolution may not be straightforward. We interpret our results as revealing the overall statistical structure of genotype-to-fitness maps. How evolution navigates populations through these maps, especially in the presence of second-order selection or environmental variability, is not yet understood. These questions represent an interesting area for future research.

### 2.4.3 Interpretation of the generalized global epistasis equation in terms of the connectedness model

It is important to keep in mind that the generalized global epistasis equation (2.2) in the main text must be interpreted statistically, not deterministically. While it predicts the expected effect of a mutation in a given strain background and environment, there is considerable uncertainty around this expectation due to idiosyncratic epistasis and pivot noise. The fact that this equation is a statistical description of the underlying genotype-to-fitness map (GFM) also means that it does not uniquely specify a GFM; instead, it specifies a (potentially large) family of GFMs with certain statistical properties. However, it is relatively straightforward to find a GFM from this family, i.e., a GFM that quantitatively matches the slopes and intercepts observed in the our data.

To do so, we use the connectedness model developed by Reddy and Desai [124]. In this model, an organism's genotype is represented by a series of diallelic loci with alleles denoted as 1 or  $-1$ . Each locus  $i$  is parameterized by a single parameter  $\mu_i$  which determines the likelihood of that locus participating in any given pathway, such that loci participating in the same pathway interact with each other. Reddy and Desai show that when the number of loci is large, the global epistasis slope for a “directionless” mutation at locus  $i$  (i.e., either a flip from  $-1$  to 1 or vice versa) is given by

$$b_i = -\frac{2\mu_i}{1 + \mu_i}. \quad (2.30)$$

To fit the  $\mu_i$  parameters to our data, we consider a genome with 94 loci, corresponding

to the tn-insertion locations. Since this model considers directionless mutations, we follow Reddy and Desai and first calculate the global epistasis slopes in the concatenated data set which contains both the effects of tn-insertions and the corresponding effects of tn-reversions [124]. The distribution of slopes in concatenated data in 30°C pH 5.0 is shown in Figure B.12A and the corresponding ‘bowtie’ plot is shown in Figure B.12D. We then use equation (2.30) to fit the parameter  $\mu_i$  for each locus using these empirical slopes.

To define a specific GFM, we utilize the code provided with the manuscript [124] with a number of pathways  $M = 200$ . We briefly detail how the GFM is generated in their code below. Each pathway is populated with loci by randomly sampling with the likelihood of ‘addition’ of each locus to each path =  $\mu_i$ . The members of a path are resampled if less than one or more than the max allowed epistatic order (14) loci are selected.

After this assignment, an average pathway contains about 11 loci. To calculate fitness of each genotype, all pairwise and higher order epistatic interaction coefficients are sampled from a standard normal distribution for each pathway up to the interaction order of the number of loci in the pathway (Max = 14). Then, the terms are normalized such that the total variance of background genotypes = 1. The overall fitness of the genotype is calculated as the sum of the fitness contribution from each pathway.

Because the fitness values produced by this model have no biological meaning (they have an arbitrary scale and the mean fitness across all genotypes is arbitrarily set to 0), we linearly shift and scale them to match our data. Specifically, we shift all fitness values by the pivot growth rate (such that the mean fitness across all genotypes equals the pivot growth rate). We also scale all fitness values by factor 0.2, which defines the variance in fitness and thereby allows us to match the empirical slope-intercept correlation (Figure B.12B–D).

### **Evolutionary simulations.**

The generalized global epistasis equation alone allows us to simulate evolution at the fitness level. For example, several previous studies have shown that global epistasis is sufficient

to recapitulate fitness trajectories observed in evolution experiments [68, 108, 124]. However, if we want to understand how evolution unfolds at the genetic level, we need an explicit GFM model, such as the one connectedness model described above. For example, we can interpret the state  $+1$  in the connectedness model as the presence of a functional gene (or as the absence of the tn-insertion at this locus) and the state  $-1$  as the absence of a functional gene, so that a  $+1$  to  $-1$  bit flips roughly corresponds to a tn-insertion (or other type of gene disruption) and the  $-1$  to  $+1$  corresponds to a reversions of tn-insertions (or, for example, the addition of the gene via horizontal gene transfer), and ask how many functional genes is an organism expected to retain, if the fitness effects of these mutations obey the statistics of global epistasis.

To answer this question, we set the pivot growth rate in our GFM to 0.245 which is equal to that of the 37°C pH 5.0 environment in our experiment in the connectedness model with 94 loci described above. We then run 50 replicate Wright-Fisher simulations with population size  $N = 10^4$  and mutation rate  $U = 10^{-5}$  in the strong selection weak mutation regime, starting either with a random low-fitness genotype (initial fitness = 0.02, with 53/94 initially intact genes) or with a genotype with all genes intact (i.e., a genotype with with all loci at  $+1$ ), which has initial fitness of 0.18. As expected, we observe qualitatively similar decelerating fitness trajectories in simulations with both initial genotypes (Figure B.12F,H). However, the genome dynamics are different. In populations that start with a random low-fitness genotype, the number of disrupted genes quickly plateaus at about 44/94 (47%) on average, whereas populations that start with all intact genes, the number of gene disruptions continues to accumulate over thousands of generations and appears to plateau at a lower number (Figure B.12G,I). These basic simulations demonstrate how our results can be used to understand evolutionary dynamics on fitness landscapes with a realistic statistical structure.



## 2.5 Acknowledgements

We thank the members of the Kryazhimskiy and Meyer labs, Michael Desai, Terry Hwa and Erik Hom for insightful discussions, Gautam Reddy and three reviewers for constructive feedback on the manuscript. Computational work was assisted by the UC San Diego Triton Computing Cluster. This work was supported by an NIH T32 program (5T32GM133351-02, to S.M.A), the Curci Foundation (to S.M.A), the Christopher Wills endowed fellowship (to A.M), an NSF Postdoctoral Research Fellowships in Biology Program (Grant No. 2109800, to M.S.J.), an NSF Graduate Research Fellowship (to M.S.J.), The Burroughs Wellcome Fund (Career Award at Scientific Interface (1010719.01, to S.K.), The Alfred P. Sloan Foundation (FG-2017-9227, to S.K.), The Hellman Foundation (Hellman Fellowship, to S.K.), The National Institutes of Health (Grant No. 1R01GM137112 and R35GM153242, to S.K.).

Chapter 2, in full, is a reprint of the material as it appears in An environment independent distribution of mutational effects emerges from microscopic global epistasis in bioRxiv 2024. Ardell, Sarah; Marstul, Alena; Johnson, Milo; Kryazhimskiy, Sergey. The dissertation author was the primary investigator and co-primary author of this paper.

## Chapter 3

# The statistical structure of pleiotropy in the presence of global epistasis and its evolutionary consequences

### 3.1 Abstract

The mutations which promote adaptation in a given population's 'home' environment often have diverse effects in other 'non-home' environments. These so-called 'pleiotropic' outcomes of evolution have important implications for the ability of populations to survive in natural fluctuating environments. Substantial work has examined how constraints on the epistatic effects of mutations generate constraints on the shape of the evolutionary fitness trajectory in single environments. Yet, it remains unclear if these same constraints will limit the space of possible pleiotropic outcomes of evolution. Here, we develop a low dimensional model that well captures how the effects of mutations change across strains and environments using data on the effects of 94 mutations each in 42 yeast strains in 6 different environments. From this model we discover simple rules which govern the likely pleiotropic outcomes of evolution.

## 3.2 Main Text

### 3.2.1 Introduction

Most organisms live in environments that vary across space and time. The ability of populations to adapt to such changes underpins many important phenomena, from the emergence of multi-drug resistance in bacteria and cancers [115, 144], to the persistence of populations in the face of climate change [145–147]. As such, predicting how a population’s adaptation to one (“home”) environment impacts its fitness in potential future (“non-home”) environments is a major problem in evolutionary biology with potentially far-reaching practical applications.

Despite decades of research [44, 148–150], our understanding of evolution in variable environments remains incomplete. For example, theoretical work frequently assumes that adaptation in variable environments must be constrained by ‘trade-offs’, where fitness gains in one condition come at a cost of fitness in others [151–153]. Results with this assumption have been used to explain important phenomena such restricted host-ranges and the existence of highly specialized organisms [154–156]. While the number of empirical examples of tradeoffs is extensive (Reviewed in [157, 158]), with classic examples including antibiotic resistance versus fitness in the absence of drugs (Reviewed in [159]) and growth rate in some nutrient conditions versus transition rate to new ones [160–162], they are certainly not ubiquitous. Numerous examples of the opposite, referred to as ‘correlated fitness gains’ or ‘evolved generalism’ also exist and sometimes even for the very same conditions where tradeoffs were previously identified [29, 163–167]. Further, recent work suggests that specialism may evolve in the absence of tradeoffs [80, 83, 168].

One major challenge to achieving a more holistic understanding of how adaptation to one home environment will affect fitness in another ‘non-home’ environment (which we refer to as the ‘pleiotropic outcomes of evolution’ for the remainder of the chapter) is that they depend on many interacting factors, such as the genetic background of the starting population [95, 114], the specifics of the environments [169–173], and the inherent stochasticity of evolution [28, 29, 73,

114, 174]. With such complexity, it is impossible to directly measure the pleiotropic outcomes of evolution with sufficient replication for all potential conditions of interest. Instead, to predict them, it will likely be necessary to identify patterns or constraints on how mutations that are beneficial in one environment affect fitness of the organism in other environments.

Significant effort has been made to assess whether heuristic metrics of environmental dissimilarity may predict the likely pleiotropic outcome of evolution, especially in high-throughput microbial systems, but have yielded conflicting results. For example, some studies suggest that microbial populations are more prone to exhibit tradeoffs for antibiotic pairs with different classes of action [5, 14, 164, 175]. Yet, others show these instances of cross-sensitivity are not reliably predicted by dissimilarity in antibiotic class [7, 22, 176]. Even if it were possible to identify robust heuristic predictors of pleiotropy for one evolutionary time scale, an additional challenge is that pleiotropic outcomes of evolution may change over time. For example, tradeoff-promoting mutations become more common over evolutionary time in some systems [171, 172], but they appear to become less common in others [168, 177, 178]. Thus, it is unclear if the ability to predict the pleiotropic outcomes of evolution for a single time scale may be informative of shorter or longer term evolutionary dynamics.

We have recently developed a general framework for thinking about these questions that centers on the concept of the Joint Distribution of Fitness Effect (JDFE) [109]. The JDFE is a generalization of the canonical one-dimensional distribution of fitness effects (DFE) to multiple environments, and it characterizes how mutations available to a given organism affect its fitness in two (or more) environments. In our previous work, we developed a theory for understanding how the shape of the JDFE determines the pleiotropic outcome of evolution. However, one major challenge for this framework is that JDFEs can vary between genetic backgrounds and environments, and it is currently unclear what kinds of JDFE shapes exist in nature, out of the vast space of all potentially conceivable shapes.

Here, we leverage a high-throughput mutational dataset recently collected by our lab to characterize the space of JDFEs that are observed in the lab in yeast *Saccharomyces cerevisiae*.

We then extend a model of global epistasis [179] to derive a class of JDFE shapes that captures a large fraction of variation among the observed JDFEs. We then simulate evolution on these JDFEs and characterize the typical pleiotropic outcomes of adaptation.

### 3.2.2 JDFEs vary widely across strains and environments

To understand how JDFEs vary across strains and environment pairs, we leveraged a dataset characterized in our previous paper [179]. Briefly, we measured the effects on fitness of  $\sim 100$  barcoded transposon insertion mutations each in a panel of 42 strains of *Saccharomyces cerevisiae* in 6 environmental conditions. The environments vary by temperature (30°C and 37°C) and pH (3.2, 5.0, and 7.0) in a factorial design. This dataset allows us to explore up to 630 unique JDFEs. However, we restrict our analysis to the subset of 345 high-quality JDFEs that are constructed based on at least 60 mutations. Previously, we demonstrated that marginal DFEs vary widely in their mean, variance and skewness across strains and environments. As such, we began to explore JDFEs through a parameter independent of the marginal DFEs – the JDFE correlation. We find that JDFE correlations range from  $-0.56$  to  $0.85$  (Figure 3.1A–C). The average span of JDFE correlations for the same strain in different environments covers 53% (interquartile interval 48–64%) of this range, and the average span of JDFE correlations for the same environment pair and different strains covers 46% (interquartile interval 39–53%).

We tested whether this variation in the JDFE correlation can be attributed to various genetic and environmental factors. We find that the identity of the strain does not significantly explain any variance in the JDFE correlation (ANOVA,  $P = 0.94$ ), and pH pair (same or different) explains only 11% (ANOVA,  $P < 10^{-5}$ ). On the other hand, the temperature pair (same or different) explains 55% of variance in the JDFE correlation ( $P < 10^{-5}$ ). Specifically, 97% (149/153) of JDFEs within a single temperature are positively correlated, versus only 67% (128/192) of JDFEs spanning different temperatures.

However, the value of the correlation coefficient alone does not have obvious implications for the pleiotropic outcomes of evolution. To illustrate this, consider an extreme case with a

JDFE with correlation 1. One might expect that a population evolving in environment 1 with this JDFE would necessarily incur a pleiotropic fitness gain in environment 2 in the next evolutionary step. Yet, if all mutations in environment 2 are deleterious, then the next available evolutionary step on this JDFE must be a pleiotropic fitness loss (or, tradeoff). As such, we next quantified the relative proportion of mutations which are the same (and different) sign in each JDFE.

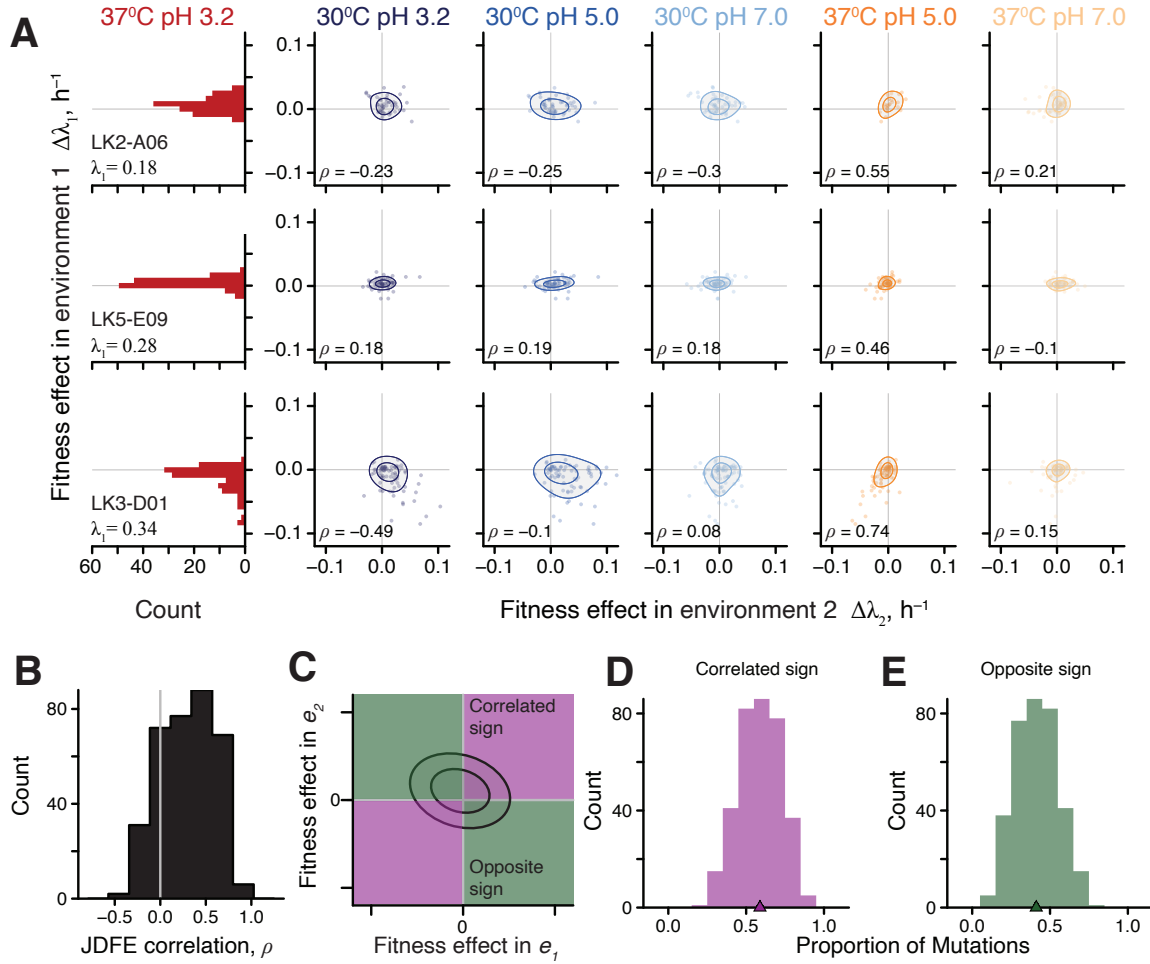
The frequency of mutations with correlated and opposite qualitative effects varies widely (Figure 3.1 C–F). On average, 58% (range 24-90%) of mutations are of the same sign, and, correspondingly, 41% (range 10-76%) are of different sign. The background strain explains no ( $P = 0.84$ ) variation in the relative frequency of these mutation classes, and pH pair explains 8% ( $P < 10^{-5}$ ). The temperature pair explains 40% of variation in the frequency of these mutational classes ( $P < 10^{-5}$ ).

Overall, parameters of the JDFE change significantly both across non-home environments for the same strain in a single home environment and across strains for the same home and non-home environment pair. Though the temperature pair appears to be an important controlling factor, these results suggest that predicting the shape of JDFEs based solely on heuristics will likely be insufficient to predict the pleiotropic outcomes of evolution.

### 3.2.3 Generalized global epistasis equation predicts JDFEs

Because heuristic measures leave substantial variance unexplained in the shape of JDFEs, we next attempt to build a model to quantitatively capture and predict how JDFEs will change with background strain and environment. In our previous work, we have shown that some key features of the DFE in individual environments can be predicted based on a simple model of microscopic global epistasis [179]. Specifically, we found that the fitness effect  $\Delta\lambda_{mge}$  of mutation  $m$  in background strain  $g$  in environment  $e$  can be expressed with the generalized global epistasis equation

$$\Delta\lambda_{mge} = b_m\lambda_g^* + \eta_{me} + \xi_{mge} \quad (3.1)$$



**Figure 3.1. Variation in empirical JDFEs based on data from [179] A.**

All JDFEs for 3 strains with environment 1 as 37C pH 3.2. Column 1 shows the marginal DFEs for each strain in environment 1. Remaining columns show JDFEs across five other environments for each strain. Pearson correlation coefficient indicated at the bottom left of each panel. JDFEs are colored by environment 2. **B.** Histogram of JDFE correlations for all strains and all unique environment pairs. Grey line at 0. **C.** JDFE schematic highlighting mutation classes of interest. **D.** Histogram of the frequency of mutations which have the same sign out of all mutations measured in both environments in a given strain. **E.** Analogous to D for mutations which have the opposite sign.

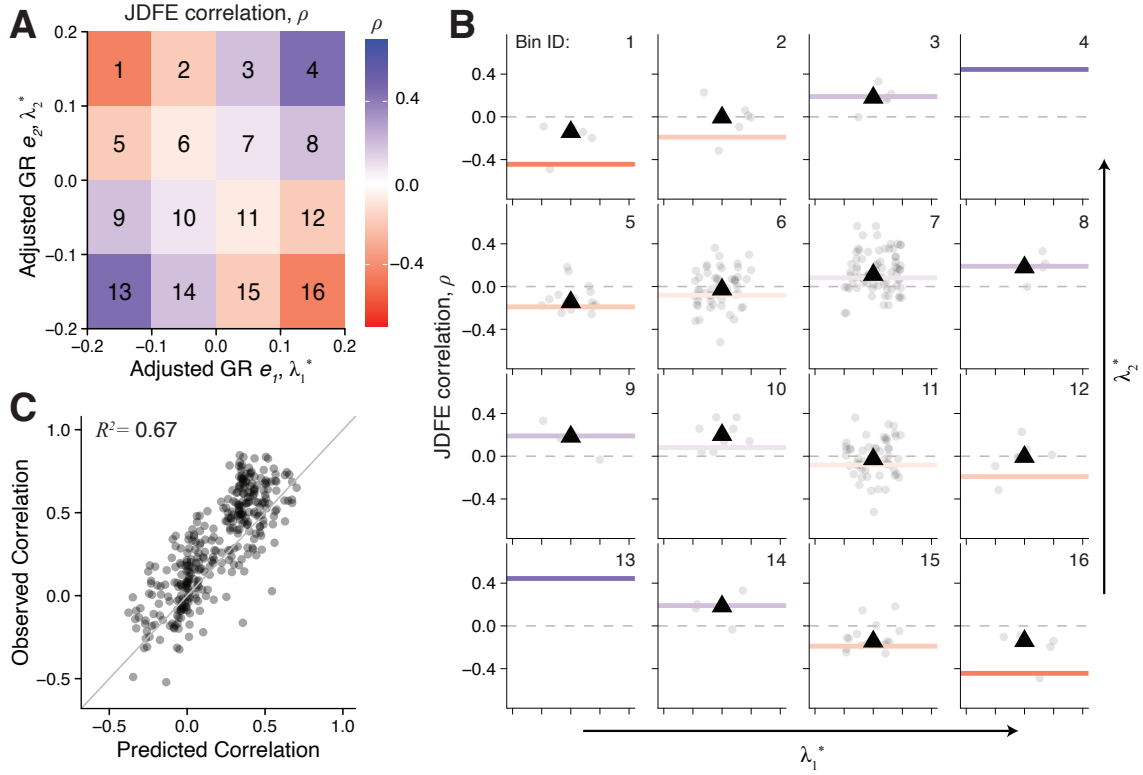
where  $b_m$  is the environment-independent global epistasis slope of mutation  $m$ ,  $\lambda_g^* = \lambda_{ge} - \bar{\lambda}_e$  with  $\lambda_{ge}$  being the growth rate of background strain  $g$  in environment  $e$ , and  $\bar{\lambda}_e$  being the pivot growth rate in environment  $e$ , i.e., the growth rate at which the fitness effect of a typical mutation changes sign.  $\eta_{me}$  and  $\xi_{mge}$  are stochastic noise terms that we model as environment-independent Gaussian random variables with a zero mean (Figure C.1). More specifically,  $\eta_{me}$  is the “pivot noise”, i.e., the deviation of the y-intercept from zero, and  $\xi_{mge}$  represents the unpredictable variation in the effect of a mutation across genetic backgrounds often referred to as “idiosyncratic epistasis”.

We have previously shown that we can use equation 3.1 to predict the moment of DFEs [179]. Here, we apply the same approach to understand the variation in the JDFEs across strains and environment pairs. As in the previous section, we first focus on the JDFE correlation, a parameter that is independent of the features of individual marginal DFEs and is symmetric with respect to the two environments. From the generalized global epistasis equation, we derive the JDFE correlation under two simplifying assumptions: that there is no idiosyncratic epistasis ( $\xi_{mge} = 0$ ) and pivot noise values ( $\eta_{me}$ ) are uncorrelated across environments. The first assumption is empirically untrue; however, we demonstrate that its impact on evolutionary dynamics appears to be minimal (Figure C.2). We find the second to be empirically true for almost half (7/15) of our tested environment pairs (Figure C.1). With these assumptions, we derive that JDFE correlation  $\rho_{ge_1e_2}$  for genotype  $g$  and environments  $e_1$  and  $e_2$  is predicted to be:

$$\rho_{ge_1e_2} = \frac{\lambda_{ge_1}^* \lambda_{ge_2}^* \text{Var}(b_m)}{\sqrt{\text{Var}(DFE)_{ge_1} \text{Var}(DFE)_{ge_2}}} \quad (3.2)$$

Where the numerator is the predicted JDFE covariance, and  $\lambda_{ge_1}^*$  and  $\lambda_{ge_2}^*$  are the adjusted growth rate of strain  $g$  in environments 1 and 2, respectively.  $\text{Var}(b_m) = 0.022$  is the empirical variance of the cross-environmental slope distribution. To yield the predicted JDFE correlation, this covariance is scaled by the predicted DFE variances in each environment  $\text{Var}(DFE)_{ge_1}$ ,  $\text{Var}(DFE)_{ge_2}$ , given by equation (2.26) in Chapter 2.





**Figure 3.2. Prediction of JDFE correlation**

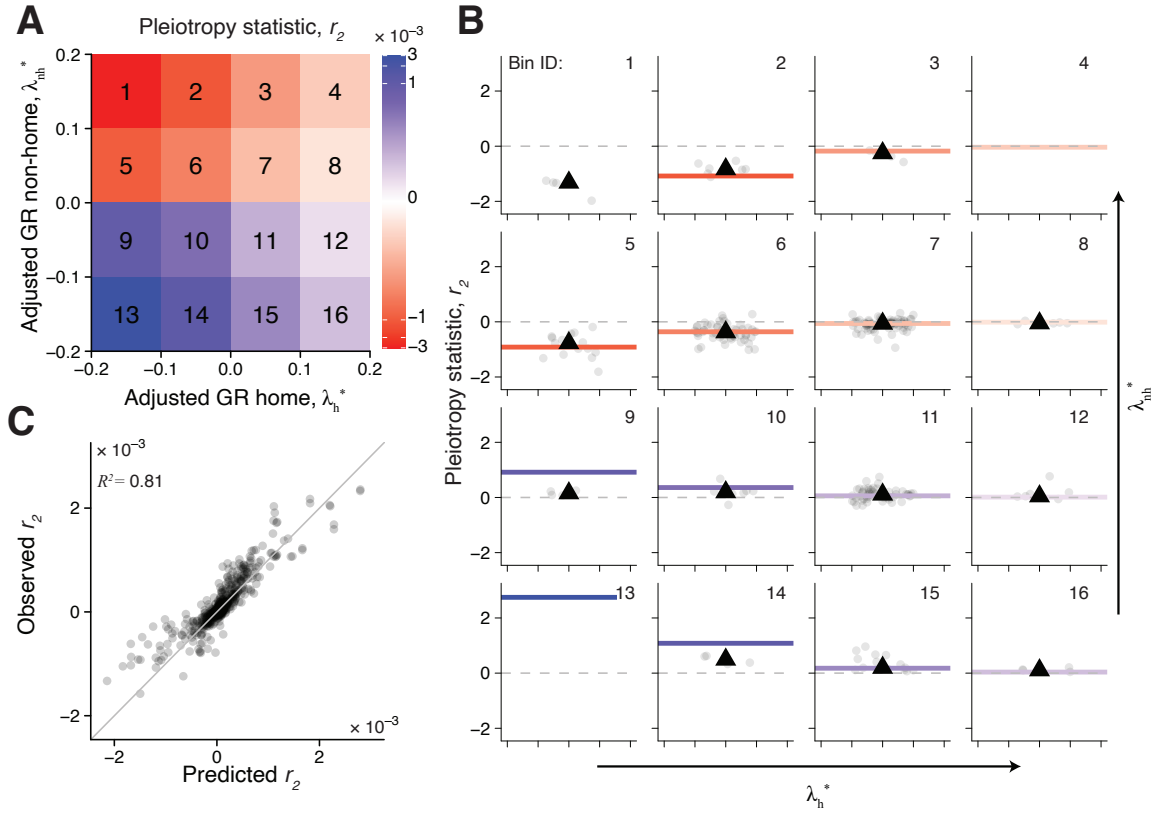
**A.** Predicted JDFE correlation of a hypothetical strain at the midpoint of 16 growth rate bins based on environment 1 and 2 adjusted growth rates. **B.** JDFE correlation values in each GR bin in (A). Gray points represent the correlation coefficients of empirical JDFEs for all strain–e1–e2 combinations in the bin. Black triangles are the means of the observed points. Solid lines are the JDFE correlation predicted at the midpoint of the bin. **C.** Overall predicted vs observed JDFE correlation. (Linear regression,  $P < 10^{-5}$ ).

To assess how well equation (3.2) captures the variation in the empirically observed JDFE correlation coefficients, we binned all our measured JDFEs into 16 growth rate bins by their adjusted growth rate in environments 1 and 2 (Figure 3.2A) and compared the observed correlations of JDFEs within each bin with those predicted by equation (3.2) at the bin's midpoint. Equation (3.2) predicts that when  $\lambda_{ge_2}^*$  is positive, the JDFE correlation increases with  $\lambda_{ge_1}^*$  from negative to positive, switching sign at  $\lambda_{ge_1}^* = 0$  (Figure 3.2A,C, bins 1–8). Analogously, when  $\lambda_{ge_2}^*$  is negative, JDFE correlation decreases with  $\lambda_{ge_1}^*$  from positive to negative, switching sign again at  $\lambda_{ge_1}^* = 0$  (Figure 3.2A,B, bins 9–16). And indeed, we observe these trends in our empirical JDFEs (Figure 3.2B). Overall, we find that eq (2) explains 67% of variance in observed JDFE correlations (Figure 3.2C).

While the first moment and the covariance matrix help us understand how the overall JDFE shape varies across genetic backgrounds and environment pairs, these statistics alone are insufficient to predict the pleiotropic outcomes of evolution [109]. In our previous work, we have demonstrated that the quantity that can be directly calculated from the JDFE referred to as the “pleiotropy statistic”, or  $r_2$ , allows us to predict how the fitness in a “non-home” environment changes as a population adapts to a “home” environment [109]. Positive (negative)  $r_2$  values imply that on average, a population adapting to the home environment will gain (lose) fitness in the non-home environment. Again, we use the generalized global epistasis equation under the same assumptions as above to calculate how the pleiotropy statistic  $r_{2,ge_h e_{nh}}$  for strain  $g$ , home environment  $e_h$  and non-home environment  $e_{nh}$  depends on the adjusted growth rates of the background strain in the two environments.

$$r_{2,ge_h e_{nh}} = x\lambda_{ge_h}^* \lambda_{ge_{nh}}^* + y\lambda_{ge_{nh}}^* \quad (3.3)$$

Here,  $x$  and  $y$  are functions of the home adjusted growth rate  $\lambda_{ge_h}^*$  and statistics of the slope distribution (See Supplement). Importantly, unlike the JDFE correlation, the pleiotropy statistic is not symmetric with respect to the two environments. As such, we consider the set of 690



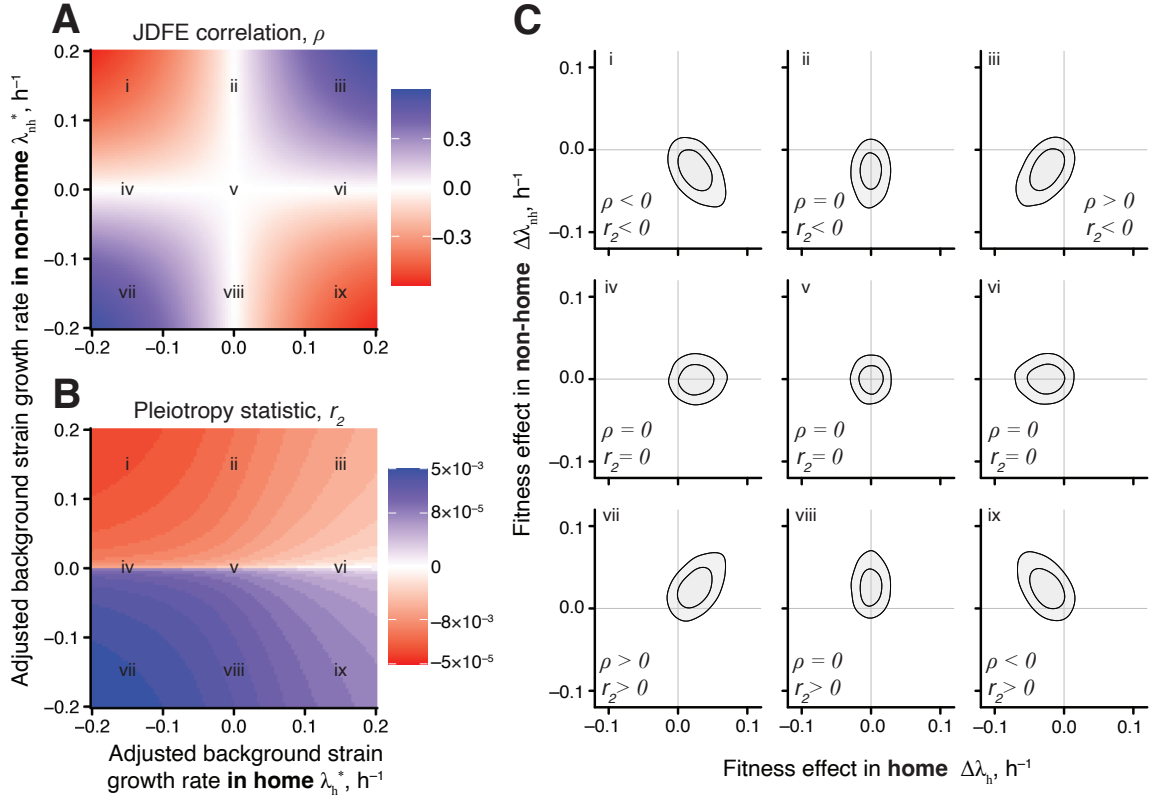
**Figure 3.3. Prediction of the pleiotropy statistic  $r_2$**

**A.** Predicted  $r_2$  at the midpoint of 16 growth rate bins based on environment 1 and 2 adjusted growth rates. **B.** JDFE  $r_2$  in each GR bin in (A). Gray points represent the  $r_2$  of empirical JDFEs for all strain–e1–e2 combinations in the bin. Black triangles are the means of the observed points. Solid lines are the  $r_2$  values predicted at the midpoint of the bin. **C.** Overall predicted vs observed  $r_2$ . (Linear regression,  $P < 10^{-5}$ ).

(=  $345 \times 2$ ) JDFEs which comprise all environment pairs in both orientations (e.g.,  $(e_1, e_2)$  and  $(e_2, e_1)$ ).

We assessed the goodness of fit of this model using the same 16 growth rate bins as in Figure 3.2. As expected we found that if  $\lambda_{ge_{nh}}^*$  is positive (Bins 1-8), then  $r_{2,ge_{he_{nh}}}$  is largely negative and increases toward 0 with  $\lambda_{ge_h}^*$ . Similarly, if  $\lambda_{ge_{nh}}^*$  is negative (Bins 9-16), then  $r_{2,ge_{he_{nh}}}$  is largely positive and decreases toward 0 with  $\lambda_{ge_h}^*$  (Figure 3.3B). Our predictions capture 81% of variance in the quantitative value of  $r_{2,ge_{he_{nh}}}$  (Figure 3.3 C) and have 85% accuracy in predicting the sign of  $r_{2,ge_{he_{nh}}}$ .

With empirical support for our model, we turned to exploring its predictions across



**Figure 3.4. Model of cross-environmental global epistasis**

**A.** High resolution predicted JDGE correlation space for variation in adjusted growth rate values in 2 environments. **B.** High resolution predicted  $r_2$  space for variation in adjusted growth rate values in 2 environments. **C.** Predicted JDGEs for the 9 growth rate values indicated in (A-B).

a broader scope of growth rates. Overall, our model predicts that JDGEs will be positively correlated if growth rates are on the same side of the pivot GR in each environment and negative otherwise. The magnitude of the correlations become stronger with more extreme growth rates (Figure 3.4A). The pleiotropy statistic is predicted to be negative whenever the adjusted growth rate in the non-home environment is positive and negative otherwise, and it declines in magnitude toward zero as the adjusted growth rate in the home environment increases (Figure 3.4B).

Taken together, we can intuitively understand the variation in the JDGE shapes as follows (Figure 3.4C). First, consider a series of strains whose adjusted growth rate in the non-home environment are zero and whose adjusted growth rate in the home environment increases from negative to positive (Figure 3.4C, iv-v). All these strains would have a narrow DFE in the

non-home environment, and their DFE in the home-environment would change from broad, right-skewed to narrow and symmetric, to broad and left-skewed. As long as the adjusted growth rate in the non-home environment is zero, the JDFE is predicted to have zero correlation coefficient and a zero pleiotropy statistic (Figure 3.4C, iv-vi). Now consider a series of strains whose adjusted growth rate in the non-home environment is positive (Figure 3.4C, i-iii). All these strains would have a broad negatively skewed DFE in the non-home environment, and their DFE in the home-environment would change, as in the previous case, from broad, right-skewed to narrow and symmetric, to broad and left-skewed as the home adjusted growth rate increases. In this case, however, the the JDFE correlation is expected to change from negative to positive as home fitness increases, crossing zero at  $\lambda_{g_{nh}}^* = 0$  (Figure 3.4ii-iii), while the pleiotropy statistic is expected to remain negative. The variation in the JDFE in case when the non-home adjusted growth rate is negative can be understood analogously (Figure 3.4viii-ix).

Importantly, this argument demonstrates that the sign of the pleiotropy statistic (which determines the outcomes of evolution) may differ from the sign of the JDFE correlation. For example, the JDFE depicted in Figure 3.4Ciii, despite being strongly positively correlated, has more density in quadrant 4 of the space than it does in quadrant 1 (e.g., it contains more tradeoff promoting mutations than pleiotropic fitness gain mutations), generating  $r_2 < 0$ . This observation underscores that, counter to a naive intuition, the JDFE correlation is a poor proxy for the chance of observing an evolutionary tradeoff.

### 3.2.4 Evolutionary implications

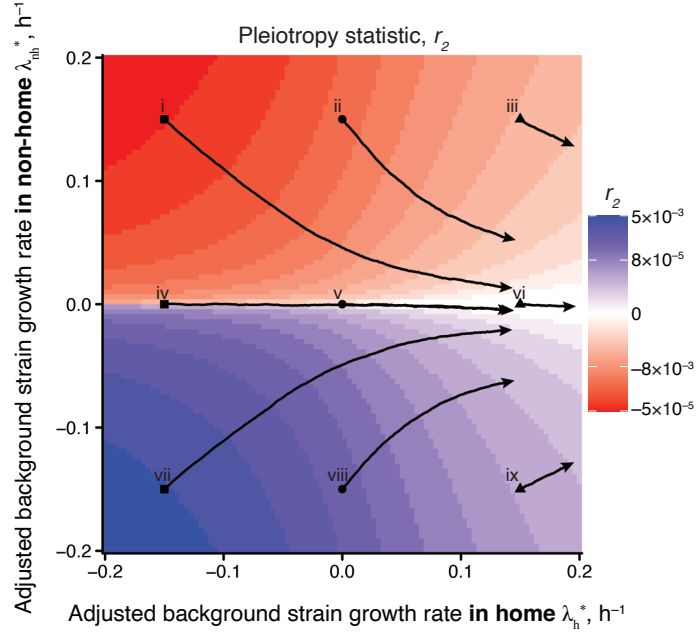
The phase diagram shown in Figure 3.4B allows us to understand how the fitness of the population in both environments is expected to change as the population adapts to the home environment. As the population adapts, the proportion of mutations which are beneficial in the home environment will decrease and adaptation will slow down, leading to the phenomenon of “declining adaptability” which has been extensively described in the empirical literature (reviewed in [141]). The dynamics of fitness in the non-home environment are more interesting.

These dynamics are expected to be qualitatively different depending on the initial adjusted growth rate in the non-home environment. Specifically, if the initial adjusted non-home fitness is positive, it is expected to decline over time, whereas if the initial non-home fitness is negative, it is expected to increase over time. In other words, our model predicts that non-home fitness will converge toward the zero adjusted growth rate as evolution proceeds in the home environment.

To test these predictions, we simulated evolution in the home environment in the strong selection weak mutation regime, where we sampled new mutations from JDFEs predicted by our model (see Methods for details). To gain intuition, we examined the mean fitness trajectories over 4000 generations from 300 replicate simulations each for a broad range of starting fitnesses in the adjusted growth rate space (Figure 3.5A). As expected, this model generates declining adaptability in the home environment, where populations that start at higher home fitness values incur less fitness gain in 4000 generations than the populations that start at lower fitness in the home (Figure 3.5, Figure C.2).

Additionally, mean non-home fitness trajectories indeed move toward adjusted growth rate 0. Populations beginning at the non-home pivot growth rate demonstrate pleiotropic neutrality over evolutionary time (Figure 3.5, middle). Populations beginning below the non-home pivot GR demonstrate pleiotropic fitness gains over evolutionary time (Figure 3.5, bottom), and populations beginning above the non-home pivot GR demonstrate pleiotropic fitness losses (or, evolutionary ‘tradeoffs’) as evolution proceeds.

We tested the robustness of this result in two ways. We find that evolutionary trajectories follow the same qualitative patterns in the presence of idiosyncratic epistasis, though the rate of adaptation marginally increases in the home environment (Figure C.2A). Second, we find that evolution in the clonal interference regime, where multiple mutations may segregate in a population at the same time, also results in the same qualitative evolutionary outcomes (Figure C.2B).



**Figure 3.5. Evolutionary implications of cross-environmental global epistasis**

Simulations of fitness trajectories with global epistasis on the pleiotropy statistic landscape starting from each point in Figure 3.4. Population size  $N = 10^5$ , mutation rate  $U = 10^{-6}$ . Each line is the average of 1000 replicate simulations of 4000 generations each.

### 3.2.5 Discussion

From analyzing experimentally measured JDFEs and extending previous theoretical work on global epistasis, we arrived at two major conclusions. First, the shape of the JDFE is determined largely by the adjusted growth rate of the strain in each environment, regardless of the specifics of the environment pair so long as pivot noise is uncorrelated. Second, the evolutionary dynamics in the non-home environment are predicted to be a simple function of the fitness in the non-home environment, where the ‘pivot growth rate’ in each environment is a stable point toward which pleiotropic evolutionary trajectories are expected to converge so long as the assumptions of our model hold.

These results offer a natural resolution to the apparently conflicting empirical results on the likelihood of tradeoff promoting mutations to either increase or decrease with evolutionary time. If the initial GR in the non-home is above the pivot GR for that environment, then most

initial mutations beneficial in the home will be deleterious (or, promote a tradeoff) in the non-home. However, as adaptation proceeds, the JD<sub>FE</sub> will become less negatively correlated and have less density of tradeoff-promoting mutations. Similarly, if the initial non-home GR is below the pivot GR, then most adaptive mutations in the home are predicted to also be adaptive in the non-home. As adaptations process, the density of tradeoff promoting mutations increases. Overall, our work suggests that the likelihood of observing a tradeoff over evolutionary time is largely a function of the initial population’s starting point with respect to the pivot growth rate in the non-home environment of interest.

These results also offer a surprising new perspective on the ability of populations to survive environmental perturbations. A priori, we may have expected that, as a population evolves to a new environment, it will randomly walk in the non-home fitness space. In this case, evolution in a home environment would result in fundamentally unpredictable outcomes in the non-home. Alternatively, we may have expected that sufficient specialization to any given home environment must necessarily result in the loss of the ability to survive in others. In this case, would expect the range of environments a population can survive in to predictably decrease with evolutionary time in a constant environment. However, our model suggests that for any population adapting to one environment, there exists a class of environments for which non-home fitness is expected to predictably approach the single, deterministic ‘pivot growth rate’ value. Because the pivot growth rate is above 0 (at least in the conditions we have tested), this result suggests that microbial populations may retain the ability to survive in diverse environments despite thousands of generations of adaptation to a different condition.

### **3.3 Materials and Methods**

#### **3.3.1 Data availability and code**

All experimental data analyzed in this chapter, as well as all parameters for equation 3.1 in the main text, were previously reported in [179]. All analysis code was written in R and will



be made available upon publication.

### 3.3.2 Analysis of empirical JDFEs

The main data in [179] contains the fitness effects of 94 mutations each measured in  $\sim 42$  strains of *S. cerevisiae* in  $\sim 6$  environments. Here, we use this data to explore how JDFEs vary with background strain and environmental pairs.

We construct pairwise JDFEs as the collection of mutations measured in each strain in both environments for each environmental pair. Because not all mutations are measured in all strains and all environments, some JDFEs contain few mutations. To ensure our results are not skewed by spurious values calculated from low-density JDFEs, we subset our data to only retain JDFEs with 60 or more mutations measured in both environments.

Importantly, some JDFE statistics (such as correlation) are symmetric with regard to the identity of the environments (e.g.,  $\rho_g(1, 2) = \rho_g(2, 1)$ ). In this case, we retain only 'unique' environmental pairs in our data (e.g., if the data contains environmental pair  $(e_1, e_2)$ , it does not contain  $(e_2, e_1)$ ), for a total of 15 unique environmental pairs. For other statistics, including the pleiotropy statistic  $r_2$ , the order of the environment matters (e.g.,  $r_{2,g}(1, 2) \neq r_{2,g}(2, 1)$ ). In these cases, we retain all environmental pairs in our analyses and refer to environments as 'home' and 'non-home' instead of '1' and '2', for a total of 30 possible environment pairs. We retain a total of 345 JDFEs in the unique environment pair data and 690 in the full data.

JDFE correlation is calculated as the standard Pearson correlation coefficient  $\rho$ . The proportion of mutations of the same or opposite effect is calculated based on the sign of the average estimated mutation effect. The pleiotropy statistic  $r_2$  is calculated as in [109]:

$$r_2 = 2 \int_0^\infty s_h ds_h \int_{-\infty}^\infty s_{nh} p(s_h, s_{nh}) ds_{nh} \quad (3.4)$$

Where,  $s_h$  and  $s_{nh}$  are the fitness effects of a mutation in the home and non-home environments, respectively, and are equivalent to  $\Delta\lambda_h$  and  $\Delta\lambda_{nh}$ .  $p(s_h, s_{nh})$  represents the JDFE

and defines the probability of sampling a mutation with effect  $(s_h, s_{nh})$ . Intuitively,  $r_2$  quantifies the expected value on fitness of the next evolutionary step in the non-home. See Ref [109] for details.

### 3.3.3 Theoretical calculation of JDGE correlation and the pleiotropy statistic

#### Notations

We use notations adopted in the Supplementary Materials to [179]. Specifically, we denote the adjusted growth rate (GR) of the background strain  $g$  in environment  $e$  by  $F_{ge} \equiv \lambda_{ge}^*$  and we denote the fitness effect of a mutation in this background and environment by  $s_{ge} \equiv \Delta\lambda_{ge}$ . We will further assume that the effects of mutations  $s_{ge}$  are drawn from a continuous distribution defined by the generalized global epistasis equation

$$s_{ge} = bF_{ge} + \eta_e + \xi_{ge},$$

where  $b$  and  $\eta$  are the slope and the y-intercept of the focal mutation in environment  $e$  (slope is environment-independent), and  $\xi_{ge}$  is the idiosyncratic epistasis of this mutation in the background strain  $g$  in environment  $e$ . We assume that  $b$  and  $\eta_e$  are independent, and that  $b$  has probability density  $p_{sl}$ ,  $\eta_e$  is normally distributed with zero mean and variance  $\sigma_{pi}^2$ .

#### Derivations

Here we assume that  $\xi_{ge} = 0$ . Then, the DFE  $p_{ge}(s)$  in environment  $e$  is given by

$$p_{ge}(s) = \int_{-\infty}^{\infty} db p_{sl}(b) N\left(s - bF_{ge}; 0, \sigma_{pi}^2\right) = \int_{-\infty}^{\infty} db p_{sl}(b) N\left(s; bF_{ge}, \sigma_{pi}^2\right).$$

Now consider environments 1 and 2. We will assume that  $\eta = (\eta_1, \eta_2)$  are jointly normally distributed with zero mean and covariance matrix

$$\Sigma_{\text{pi}} = \begin{pmatrix} \sigma_{\text{pi}}^2 & c_{\text{pi}} \\ c_{\text{pi}} & \sigma_{\text{pi}}^2 \end{pmatrix}.$$

Then, the JDFE  $p_g(s)$  for strain  $g$  in environments 1 and 2 is given by

$$p_g(s) = \int_{-\infty}^{\infty} db p_{\text{sl}}(b) N(s; bF_g, \Sigma_{\text{pi}}), \quad (3.5)$$

where  $s = (s_1, s_2)$  is the vector of selection coefficients in environments 1 and 2 and  $F_g = (F_{g1}, F_{g2})$  is the vector of adjusted GRs of strain  $g$  in these two environments. As before,  $\langle s_e \rangle = \langle b \rangle F_{ge}$ .

This allows us to calculate the JDFE covariance as

$$\begin{aligned} \text{Cov}(s_1, s_2) &= \langle (s_1 - \langle s_1 \rangle) (s_2 - \langle s_2 \rangle) \rangle \\ &= \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \int_{\mathbb{R}^2} ds N(s; bF_g, \Sigma_{\text{pi}}) (s_1 - \langle b \rangle F_{g1}) (s_2 - \langle b \rangle F_{g2}) \\ &= \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \int_{\mathbb{R}^2} ds N(s; bF_g, \Sigma_{\text{pi}}) (s_1 - bF_{g1}) (s_2 - bF_{g2}) \\ &+ \int_{-\infty}^{\infty} db p_{\text{sl}}(b) (bF_{g1} - \langle b \rangle F_{g1}) (bF_{g2} - \langle b \rangle F_{g2}) \int_{\mathbb{R}^2} ds N(s; bF_g, \Sigma_{\text{pi}}) \\ &= c_{\text{pi}} + \text{Var}[b] F_{g1} F_{g2} \end{aligned} \quad (3.6)$$

We can obtain the predicted JDFE correlation from the JDFE covariance as:

$$\text{Cor}(s_1, s_2) = \frac{\text{Cov}(s_1, s_2)}{\sqrt{\text{Var}(s_1)\text{Var}(s_2)}} \quad (3.7)$$

where the expressions for the expected values of  $\text{Var}(s_1)$  and  $\text{Var}(s_2)$  given  $F_{g1}$  and  $F_{g2}$  are reported in [179].

To understand how adaptation in environment 1 affects fitness in environment 2, we calculate  $r_2$ ,

$$\begin{aligned} r_2 &= 2 \int_0^\infty s_1 ds_1 \int_{-\infty}^\infty s_2 p(s) ds_2 \\ &= 2 \int_{-\infty}^\infty db p_{sl}(b) \int_0^\infty ds_1 s_1 p(s_1|b) \int_{-\infty}^\infty ds_2 s_2 p(s_2|s_1, b), \end{aligned}$$

where

$$p(s_1|b) = N\left(s_1; bF_{g1}, \sigma_{pi}^2\right)$$

and

$$p(s_2|s_1, b) = N\left(s_2; \mu_2, \sigma_2^2\right)$$

with

$$\begin{aligned} \mu_2 &= bF_{g2} + \frac{c_{pi}}{\sigma_{pi}^2} (s_1 - bF_{g1}) \\ \sigma_2^2 &= \sigma_{pi}^2 - \frac{c_{pi}^2}{\sigma_{pi}^2}. \end{aligned}$$

is the conditional distribution of  $s_2$  given  $s_1$ . Thus, we have

$$r_2 = 2 \int_{-\infty}^\infty db p_{sl}(b) \int_0^\infty ds_1 s_1 N\left(s_1; bF_{g1}, \sigma_{pi}^2\right) \left( bF_{g2} + \frac{c_{pi}}{\sigma_{pi}^2} (s_1 - bF_{g1}) \right).$$

Since

$$\int_0^\infty x N(x; \mu, \sigma^2) dx = \frac{\mu}{2} \left( 1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right) \right) + \sqrt{\frac{\sigma^2}{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}}$$

and

$$\int_0^\infty x^2 N(x; \mu, \sigma^2) dx = \frac{\mu^2 + \sigma^2}{2} \left( 1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right) \right) + \mu \sqrt{\frac{\sigma^2}{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}},$$

we have

$$\begin{aligned}
r_2 &= 2 \int_{-\infty}^{\infty} db p_{\text{sl}}(b) \left[ b \left( F_{g2} - \frac{c_{\text{pi}}}{\sigma_{\text{pi}}^2} F_{g1} \right) \left( \frac{b F_{g1}}{2} \left( 1 + \operatorname{erf} \left( \frac{b F_{g1}}{\sqrt{2\sigma_{\text{pi}}^2}} \right) \right) + \sqrt{\frac{\sigma_{\text{pi}}^2}{2\pi}} e^{-\frac{(b F_{g1})^2}{2\sigma_{\text{pi}}^2}} \right) \right. \\
&\quad \left. + \frac{c_{\text{pi}}}{\sigma_{\text{pi}}^2} \left( \frac{(b F_{g1})^2 + \sigma_{\text{pi}}^2}{2} \left( 1 + \operatorname{erf} \left( \frac{b F_{g1}}{\sqrt{2\sigma_{\text{pi}}^2}} \right) \right) + b F_{g1} \sqrt{\frac{\sigma_{\text{pi}}^2}{2\pi}} e^{-\frac{(b F_{g1})^2}{2\sigma_{\text{pi}}^2}} \right) \right] \\
&= x F_{g1} F_{g2} + y F_{g2} + z
\end{aligned} \tag{3.8}$$

where

$$\begin{aligned}
x &= \int_{-\infty}^{\infty} b^2 \left( 1 + \operatorname{erf} \left( \frac{b F_{g1}}{\sqrt{2\sigma_{\text{pi}}^2}} \right) \right) p_{\text{sl}}(b) db \\
y &= \sqrt{\frac{2\sigma_{\text{pi}}^2}{\pi}} \int_{-\infty}^{\infty} b e^{-\frac{(b F_{g1})^2}{2\sigma_{\text{pi}}^2}} p_{\text{sl}}(b) db \\
z &= c_{\text{pi}} \int_{-\infty}^{\infty} \left( 1 + \operatorname{erf} \left( \frac{b F_{g1}}{\sqrt{2\sigma_{\text{pi}}^2}} \right) \right) p_{\text{sl}}(b) db.
\end{aligned}$$

To analyze expression eq: 3.8 , first note that  $x \geq 0$  and  $y \geq 0$  for any  $F_{g1}$ , and the sign of  $z$  depends on the sign of  $c_{\text{pi}}$ . Let's set  $c_{\text{pi}} = 0$ . First, consider  $F_{g1} \rightarrow -\infty$ . Then

$$\operatorname{erf} \left( \frac{b F_{g1}}{\sqrt{2\sigma_{\text{pi}}^2}} \right) \rightarrow 1$$

since most  $b < 0$ . Thus,  $x \approx 2\operatorname{Var}[b]$  and  $y \approx 0$ . Therefore, when  $F_g$  is very small,  $r_2 \approx 2\operatorname{Var}[b] F_{g1} F_{g2}$ , i.e., it is positive for  $F_{g2} < 0$  and negative for  $F_{g2} > 0$ . When  $F_{g1} = 0$ , we have  $r_2 = y F_{g2}$ , and again it is positive for  $F_{g2} < 0$  and negative for  $F_{g2} > 0$ . Finally, when  $F_{g1} \rightarrow +\infty$ ,

$$\operatorname{erf} \left( \frac{b F_{g1}}{\sqrt{2\sigma_{\text{pi}}^2}} \right) \rightarrow -1.$$

Thus,  $x \approx 0$  and  $y \approx 0$ . Therefore, determining the sign of  $r_2$  is not straightforward and requires a deeper analysis.

### 3.3.4 Pivot noise covariance

In the main text, we assume the pivot noise covariance  $c_{pi} = 0$  for simplicity. We tested the validity of this assumption with the empirical  $\eta$  values reported for all mutations in all environments in Ref [179]. If the  $\eta$  values do not covary for a given environment pair (1,2), then the best fit linear regression of  $\eta_1$  and  $\eta_2$  will have a slope not significantly different from 0. We find that this is indeed the case in 7/15 of our unique environment pairs (Figure C.1), where the environments with significant pivot noise covariance tend to differ only by pH and not temperature.

### 3.3.5 Parameterization of the predictions

We assessed the predictive power of eqs: 3.7 and 3.8 in two ways. First, we tested the overall predictive power of the models by comparing the quantitative predicted and observed JDfE correlation and  $r_2$  for each strain-environment pair combination.

Second, we tested if measured JDfE correlation and  $r_2$  values on average follow the predicted trends in the statistics as a function of  $F_{g1}$  and  $F_{g2}$  values. To do this, we categorized all JDfEs into 16 evenly spaced bins in the 2-dimensional adjusted growth rate space (Figure 3.2A in the main text). We calculated the expected JDfE correlation and  $r_2$  for a hypothetical strain with growth rate  $(F_{g1}, F_{g2})$  at the midpoint of each bin. We then compared how the median JDfE correlation and  $r_2$  among strains in each bin compared to the predicted value at the bin's midpoint.

We detail the specifics of the implementation and parameterization for each statistic below.

## JDFE correlation

To predict JDFE covariance (eq 3.6), we estimate the empirical variance of the distribution of slopes reported in [179], and we assume  $c_{pi} = 0$  throughout the main text. To predict JDFE correlation, we use the predicted covariance as well as the DFE variance predictions reported in Supplementary materials of [179].

## The pleiotropy statistic

The parameterization of eq 3.8 is slightly more complex. Because  $p_{sl}(b)$  is not easily characterized by a parametric distribution, we took a numeric approach to approximating the values  $x$ ,  $y$ , and  $z$ .

First, we fit a  $5^{th}$  degree polynomial to the best fit kernel-density curve created from the empirical distribution of global-epistasis slopes ( $R^2 = 0.999$ ) (Figure C.1, Table C.1). We then sample with replacement 100,000 slopes from a sequence of 5,000,000 evenly spaced values across the range of empirically observed slopes  $(-0.69, 0.02)$  with the sampling probability of any value within the range being proportional to the probability density at that value. With this set of re-sampled slopes, we numerically compute the values of  $x$ ,  $y$ , and  $z$  with the additional parameter  $\sigma_{pi}^2 = 3.89 \times 10^{-4}$ , which is the variance of the empirical pivot noise ( $\eta$ ) distribution across all environments.

### 3.3.6 Non-parametric predicted JDFEs

To obtain a full predicted JDFE for any pair of adjusted growth rate values  $F_{g1}, F_{g2}$ , we must estimate the pairwise effect of each mutation  $m$  for an arbitrary number of mutations. To do this, we must have estimates for all relevant parameters of the generalized global epistasis equation (3.1). In the main text Figure 3.4, we sampled the effects of  $N = 10000$  mutations.

We obtain  $N$  mutation-specific  $b$  values by sampling with our non-parametric probability density curve (see section above).

We then sample  $2N$   $\eta_{me}$  values, which are randomly assigned to each mutation in each

environment. These values are sampled from a Gaussian distribution with mean = 0 and variance  $= \sigma_{pi}^2 = 3.89 \times 10^{-4}$ . Throughout the main text, we assume that idiosyncratic epistasis  $\xi = 0$ . As such, we do not sample  $\xi$  values when constructing predicted JDfEs, though we test the impact of this assumption on evolutionary trajectories below.

With sampled parameters of the generalized global epistasis equation for an arbitrary number of mutations, we can then calculate the predicted fitness effects  $\Delta\lambda_1$  and  $\Delta\lambda_2$  for each mutation given the adjusted growth rates  $F_{g1}, F_{g2}$  in each environment as  $\Delta\lambda_{mge} = b_m F_{ge} + \eta_{me}$ . We use these predicted values to construct the non-parametric JDfE.

### 3.3.7 Evolutionary simulations

We simulate evolution in the Strong Selection Weak Mutation regime using the Gillespie algorithm [86]. Briefly, we sample the waiting time until the next mutation from an exponential distribution with rate  $= NU$ , where  $N$  is the population size and  $U$  is the mutation rate. In our simulations,  $NU = 0.1$ . We then estimate the current JDfE using the procedure detailed in the previous section. We randomly sample 1 mutation with fitness in the home and non-home  $(\Delta\lambda_h, \Delta\lambda_{nh})$ , from this JDfE. This mutation then has a fixation probability of:

$$\begin{cases} 0 & \Delta\lambda_h < 0 \\ 2\Delta\lambda_h & 0 < \Delta\lambda_h \leq 0.5 \\ 1 & \Delta\lambda_h > 0.5 \end{cases}$$

If the mutation fixes, we update fitness in the home and non-home. We continue the cycle until the population reaches the desired total number of generations.

#### Idiosyncratic epistasis

In our previous work, we emphasized that the rates and values of idiosyncratic epistasis are non-negligible and cause genuine deviation in our experimental data from model predictions. As such, we next assessed if evolutionary simulations including idiosyncratic epistasis would



deviate from our model predicted trajectories.

To incorporate idiosyncratic epistasis, we add one additional step to our non-parametric JDFE sampling method. We sample  $\xi_{mge}$  values from a normal distribution with mean = 0 and variance =  $-b_m\alpha$ , where  $\alpha = 1.37 \times 10^{-3}$  is the slope of the relationship between the variance of idiosyncratic epistasis residuals  $\sigma_{id}^2$  and the microscopic epistasis slopes  $b_m$  (See Figure B.4 in Chapter 2), as reported in [179].  $\Delta\lambda_{mge}$  values are then calculated with the generalized global epistasis equation (3.1) in the main text.

Simulations of evolution on JDFEs with idiosyncratic epistasis result in the same average evolutionary trajectory as simulations without, with the exception that the rate of adaptation in the home is marginally higher in the presence of idiosyncratic epistasis.

### **Clonal interference**

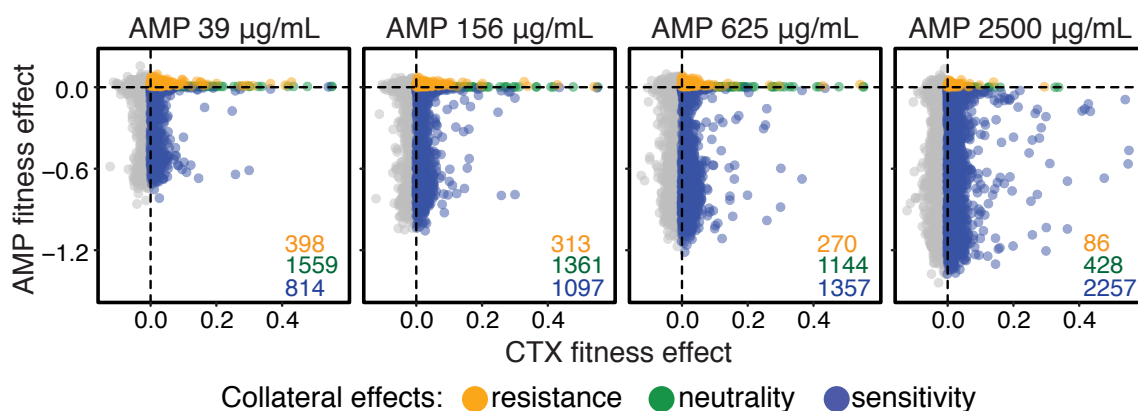
Many microbes likely evolve in a 'concurrent mutation' regime, as opposed to the 'successional mutation' regime we model above. In this regime, multiple mutations may be segregating and competing in a population at any given time. To test if competition among mutant lineages alters the average evolutionary trajectories, we updated the Wright-Fisher simulation framework we previously used in [109] with the ability to resample JDFEs for each mutant lineage. We run these simulations with population size =  $10^6$  and mutation rate =  $10^{-4}$ , and we find that the qualitative outcomes mirror those we observe in the SSWM (Figure C.1), though the adaptation rate is faster and the trajectories are substantially noisier for the same number of replicates.

## **3.4 Acknowledgements**

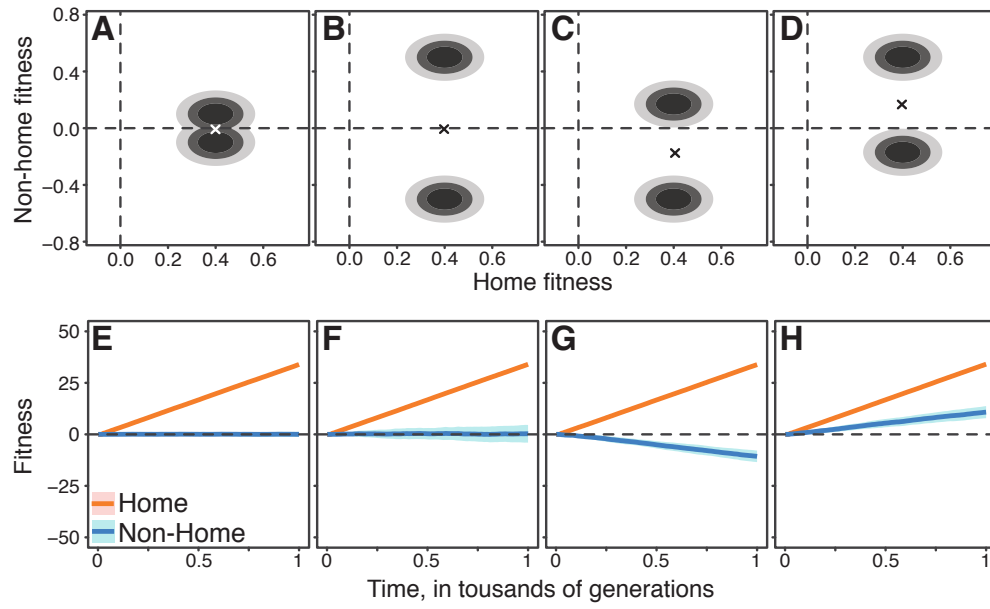
Chapter 3, is currently being prepared for submission for publication of the material. Ardell, Sarah and Kryazhimskiy, Sergey. The dissertation author was the primary investigator and author of this material.

# **Appendix A**

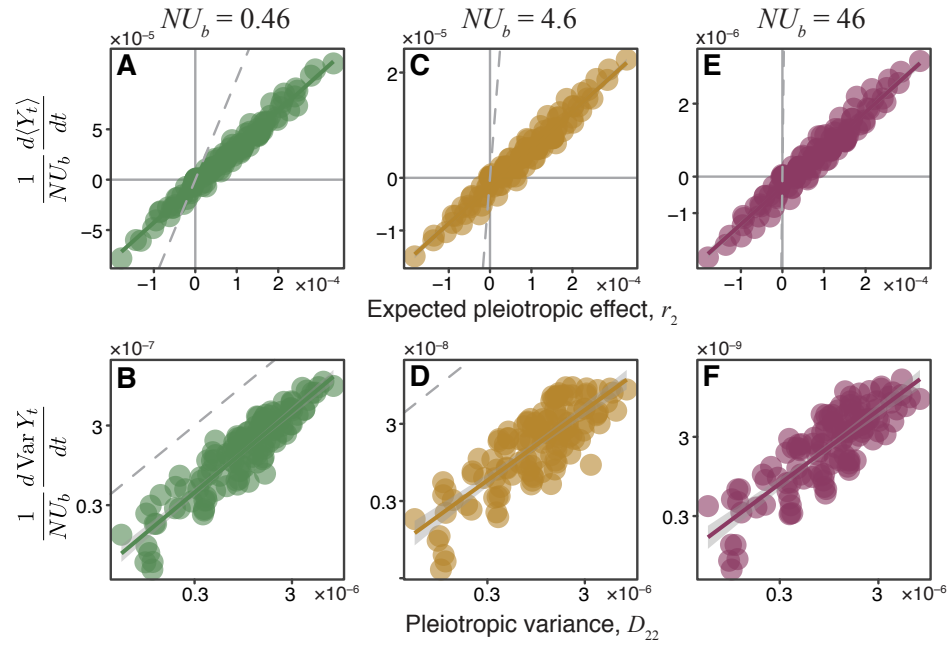
## **Chapter 1 Supplementary Figures**



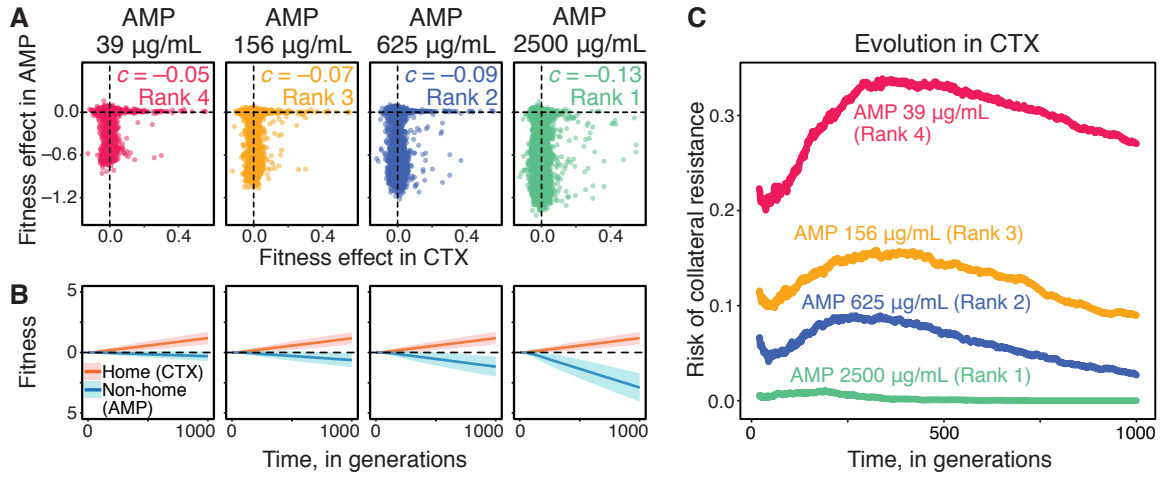
**Figure A.1. JDFEs of single point mutations in the TEM-1  $\beta$ -lactamase gene in *E. coli* in the presence of cefotaxime and ampicillin.** Data from ref [61]. Panels show data for different concentrations of ampicillin, as indicated. Fitness is measured as the change in the log ratio of the mutant to wildtype frequency during growth in the presence of the drug. Cefotaxime (CTX) is chosen as the home environment (see Materials and Methods for details). Each point represents a single point mutation and is colored by its (collateral) fitness effect in the presence of ampicillin, as indicated in the legend. The numbers of mutations with positive fitness in the presence of cefotaxime with different collateral effects are shown in the lower right corner of each panel.



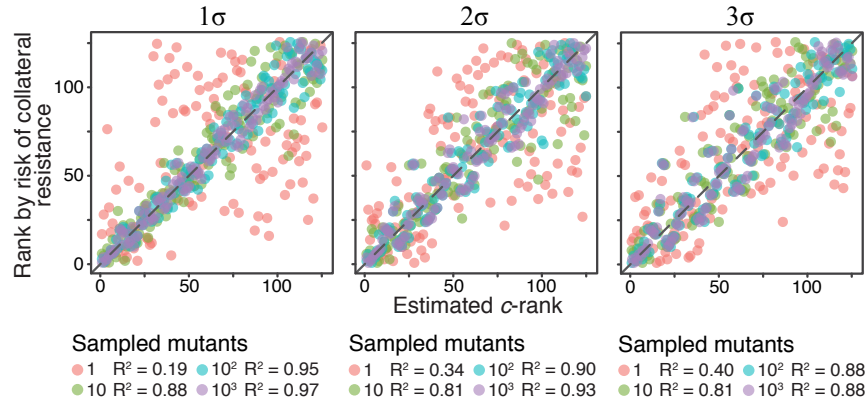
**Figure A.2. JDFEs with equal probability weights in the first and fourth quadrants and the resulting fitness trajectories.** Same as Figure 1.2, but for JDFEs with equal probability weights in the first and fourth quadrants. See Materials and Methods for details.



**Figure A.3.** Same as Figure 1.3C–H, but with  $r_2$  and  $D_{22}$  shown on the  $x$ -axis. Same as Figure 1.2, but for JDFEs with equal probability weights in the first and fourth quadrants. See Materials and Methods for details.



**Figure A.4. Ranking of AMP concentrations according to their risk of collateral resistance, based on ref [61] data.** **A.** Same as Figure A.1 but colored by  $c$ -rank, as indicated. **B.** Home and non-home fitness trajectories for the JDFEs shown in the corresponding panels in A. Notations are as in Figure 1.2. Population size  $N = 10^4$ , mutation rate  $U = 10^{-4}$ . **C.** Collateral resistance risk over time estimated from 5000 replicate simulations shown in B. Colors correspond to the ranks in A.

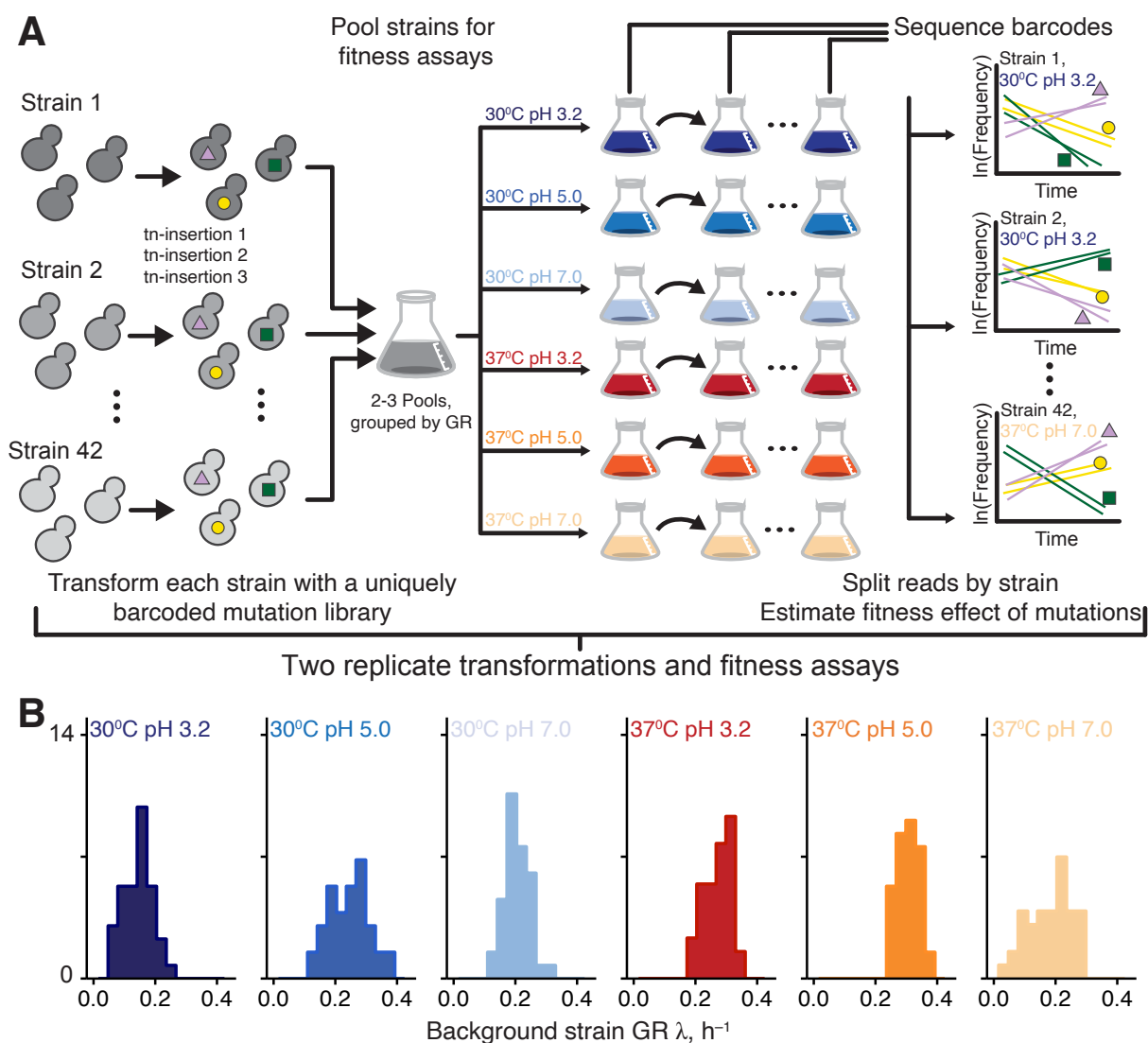


**Figure A.5. Same as Figure 1.5A, but with different thresholds for sampling.** Same as Figure 1.5A, but with different thresholds for sampling mutations, as indicated above each panel ( $\sigma$  is the standard deviation of the JDFA in the home environment). See Materials and Methods for details.  $P < 10^{-6}$  for all regressions.

## **Appendix B**

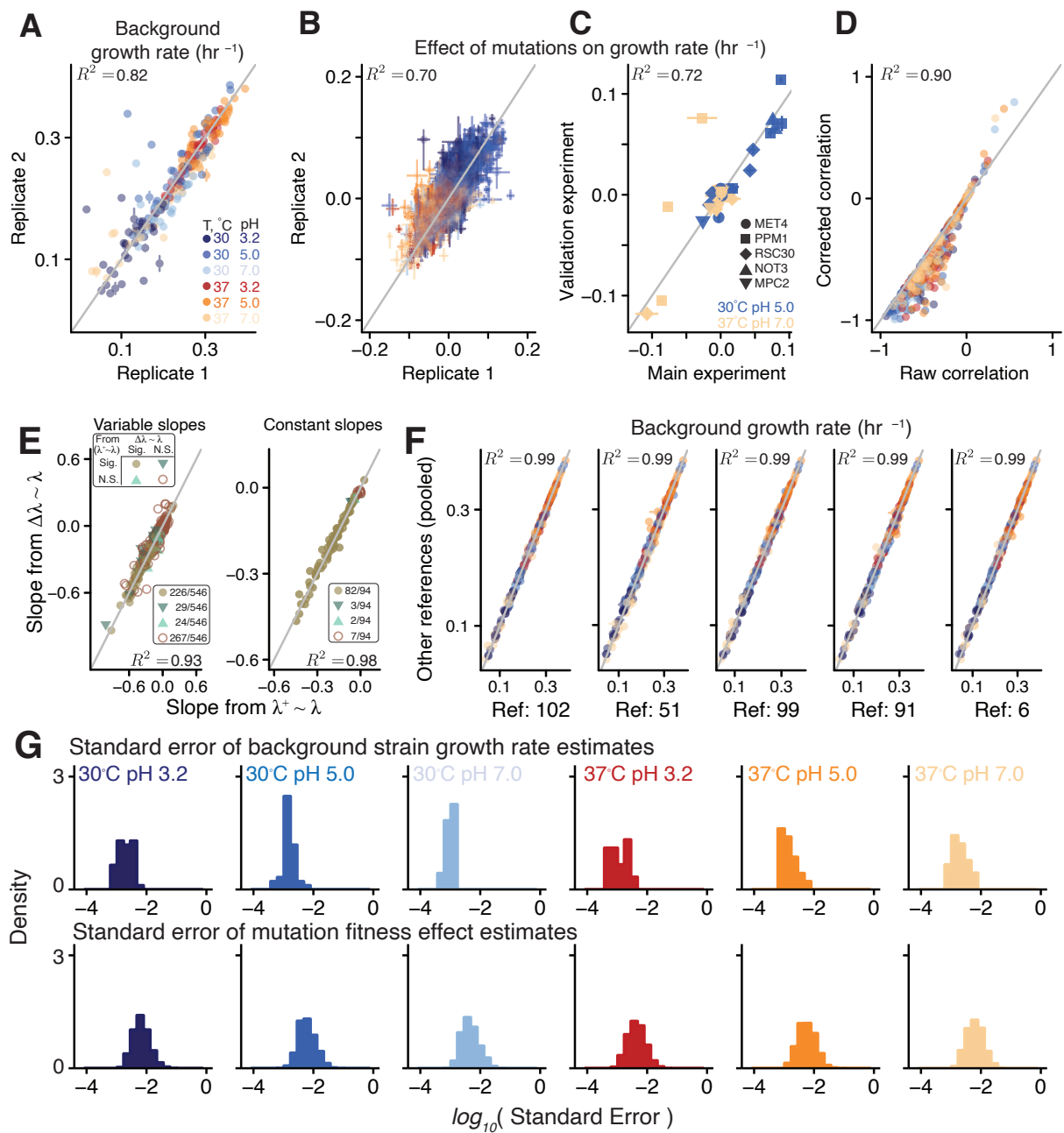
### **Chapter 2 Supplementary Figures and Tables**

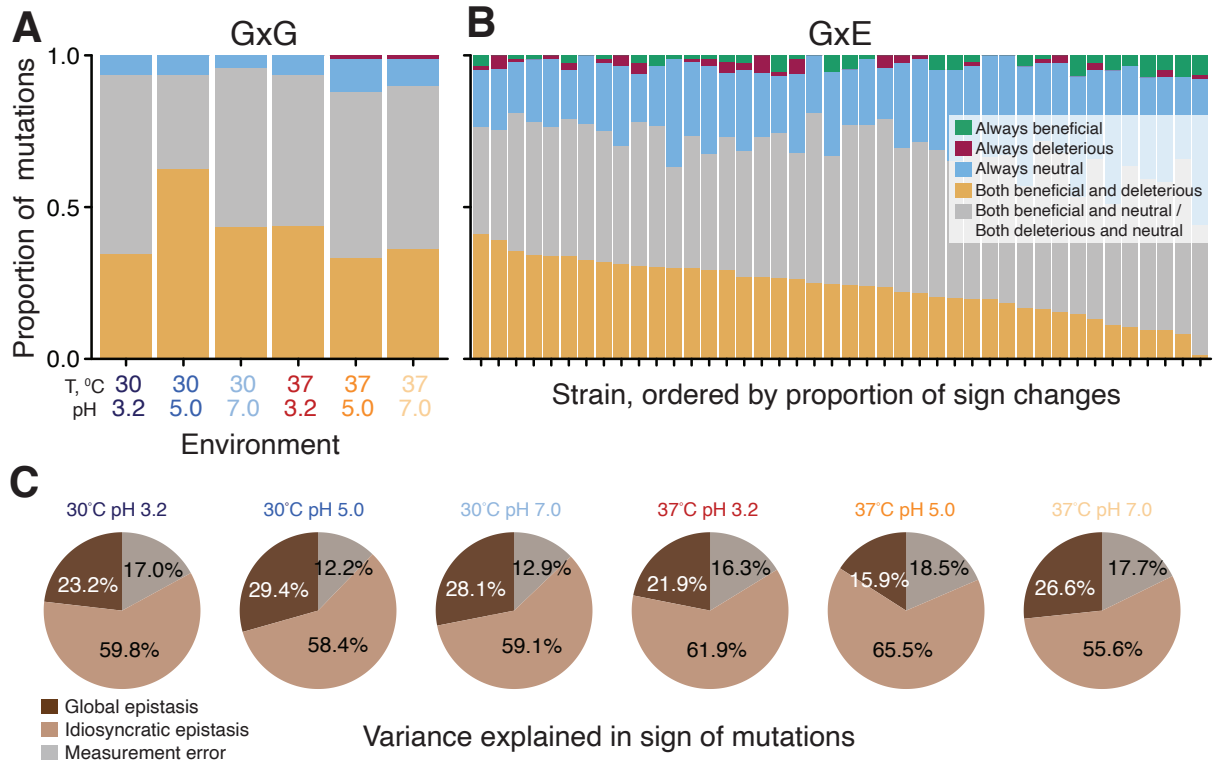




**Figure B.1. Experimental setup.** **A.** Schematic of the experiment. **B.** Distribution of growth rates of background strains in all environments.

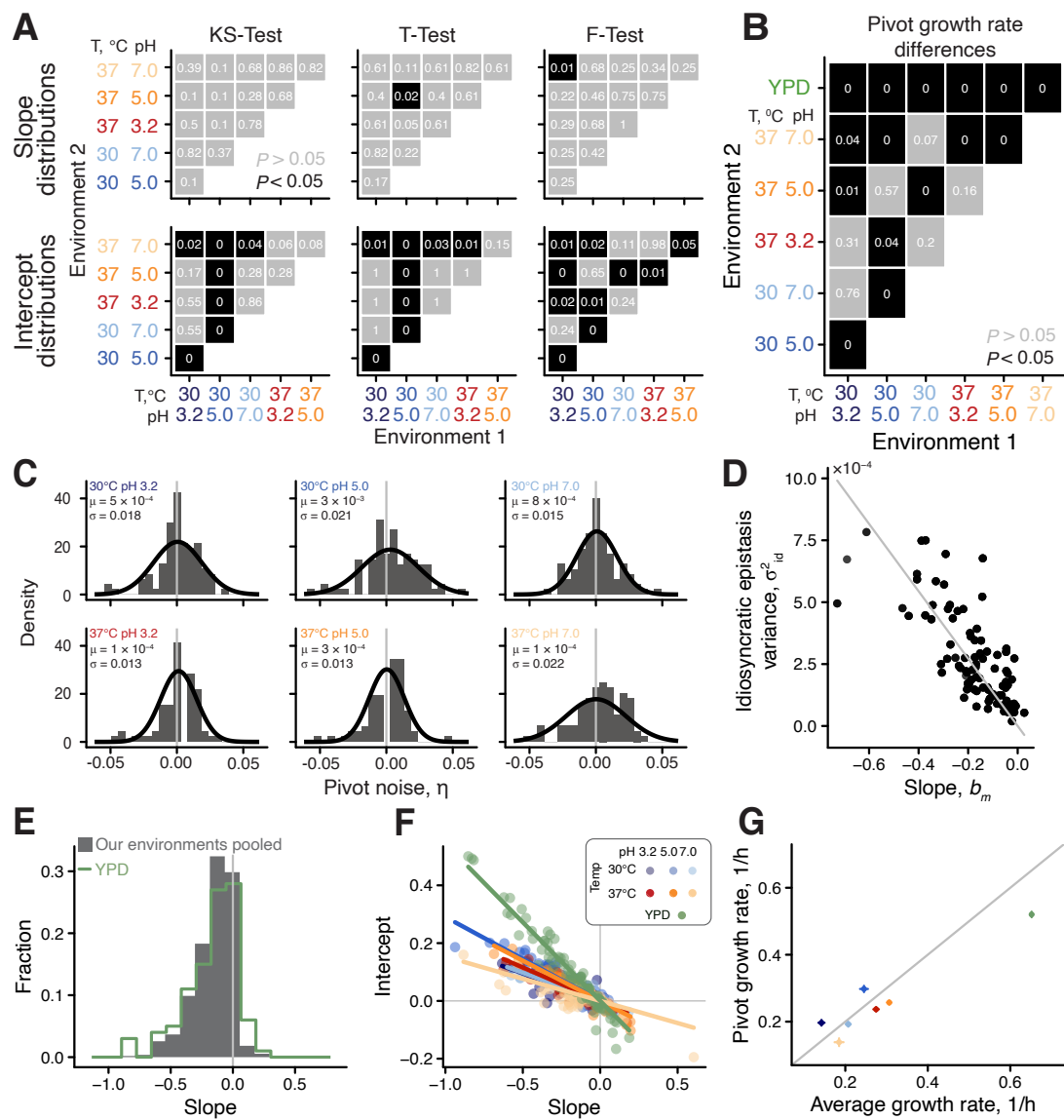
**Figure B.2. (Next page) Data quality checks.** **A.** Correlation between growth rates of background strains estimated in two biological replicates. **B.** As in panel A, but for the fitness effects of mutations. In both panels, error bars represent  $\pm 1$  standard error. **C.** Correlation between fitness effect estimates in the high-throughput RB-TnSeq experiment and the validation experiment (see Section 2.3.2). **D.** Raw and corrected estimates of the correlation coefficient between background growth rate and fitness effect for each mutation in each environment. **E.** Correlation between two approaches to estimate microscopic global epistasis slopes (see Section 2.3.3). Points are colored by the significance ( $P < 0.05$ ) of the estimated slope. **F.** Correlation between background strain growth rates estimated from each neutral reference separately against the growth rate estimated from all other references pooled. Colors as in A. In all panels, grey line is the diagonal,  $R^2$  is reported for linear regression ( $P < 0.01$  for all regressions). **G.** Distribution of the standard error of background strain growth rates (top) and mutation effects (bottom) in all environments.

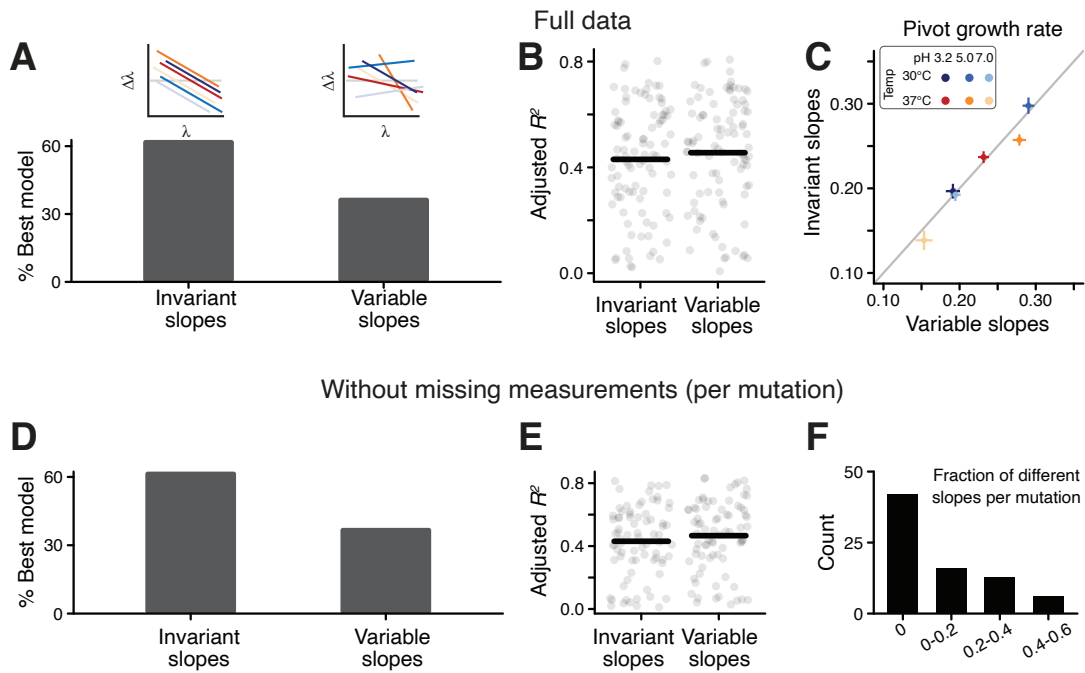




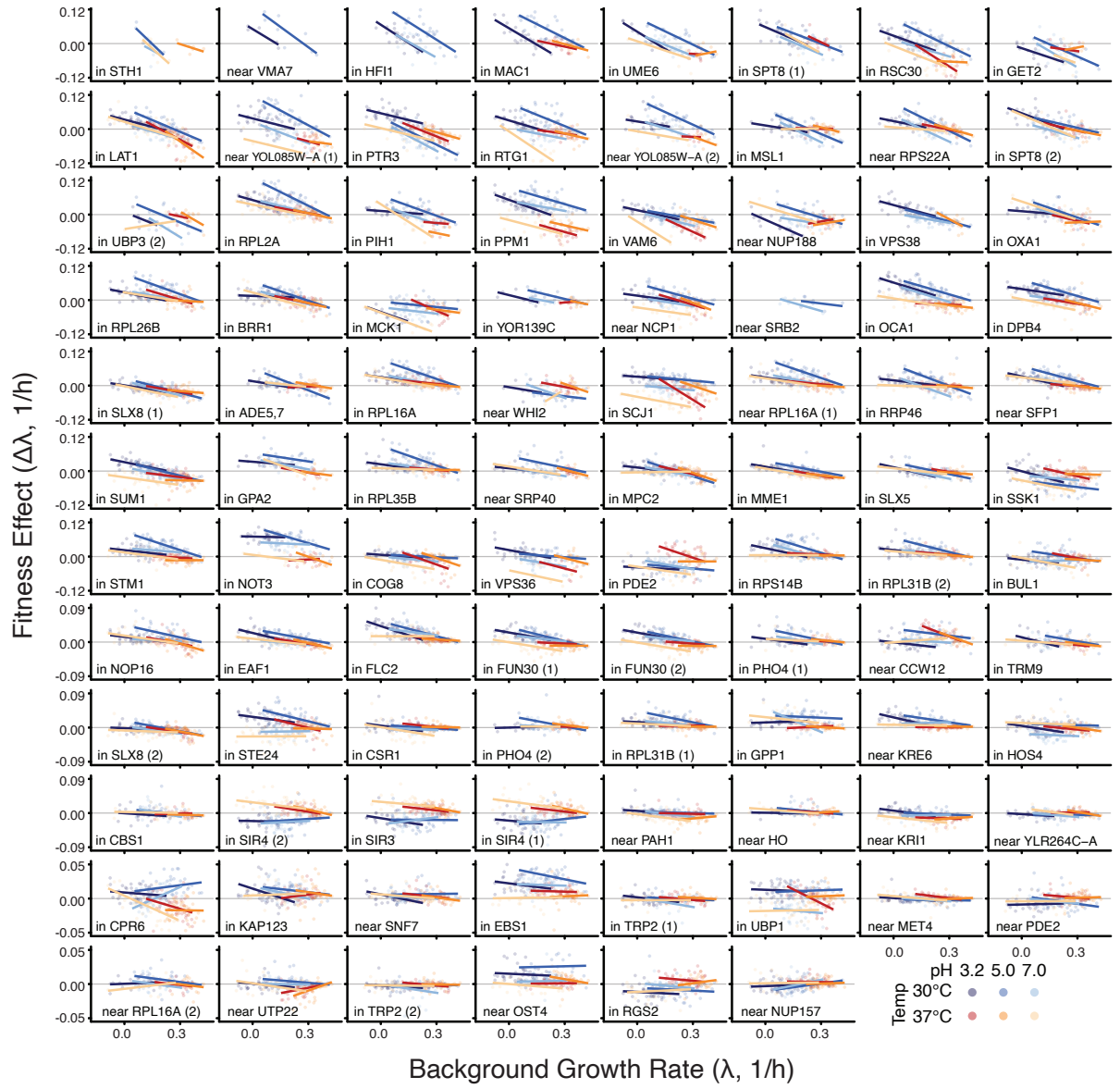
**Figure B.3. Variation in the sign of mutational effects across genetic backgrounds ( $G \times G$  interactions) and across environments ( $G \times E$  interactions).** **A.** Proportions of mutations that do and do not change sign across background strains in each environment. **B.** Proportions of mutations that do and do not change sign across environments in each background strain. **C.** The proportion of variance in the observed sign of mutations in each environment explained by measurement noise, global and idiosyncratic epistasis (see Section 2.3.3 for details).

**Figure B.4. (Next page) Properties of global epistasis models.** **A.** Comparison of distributions of slopes (top) and intercepts (bottom) across environments using three metrics (see Section 2.3.3). **B.** Comparison of the pivot growth rates across environments. In both panels, the number in each tile is the  $P$ -value (after Benjamini-Hochberg correction), with values  $P < 0.01$  indicated with '0'. Tiles with  $P < 0.05$  are colored black. **C.** Distribution of the pivot noise term  $\eta$  in each environment. Mean and variance of the distribution are labelled in each panel, and the best fit normal distribution is overlayed. **D.** Relationship between global epistasis slope and the variance of idiosyncratic epistasis residuals from the fit of equation (2.2) in the main text. Grey line is best fit linear regression through the origin. **E.** Distribution of microscopic epistasis slopes in all our measured environments (pooled, grey) and in YPD (data from Ref. [119]). **F.** Slope-intercept relationship (as in Figure 2 in the main text), but including YPD data. **G.** The relationship between the average background strain growth rate and the pivot growth rate in each environment. Colors as in Panel F. Error bars are  $\pm 1$  standard error.



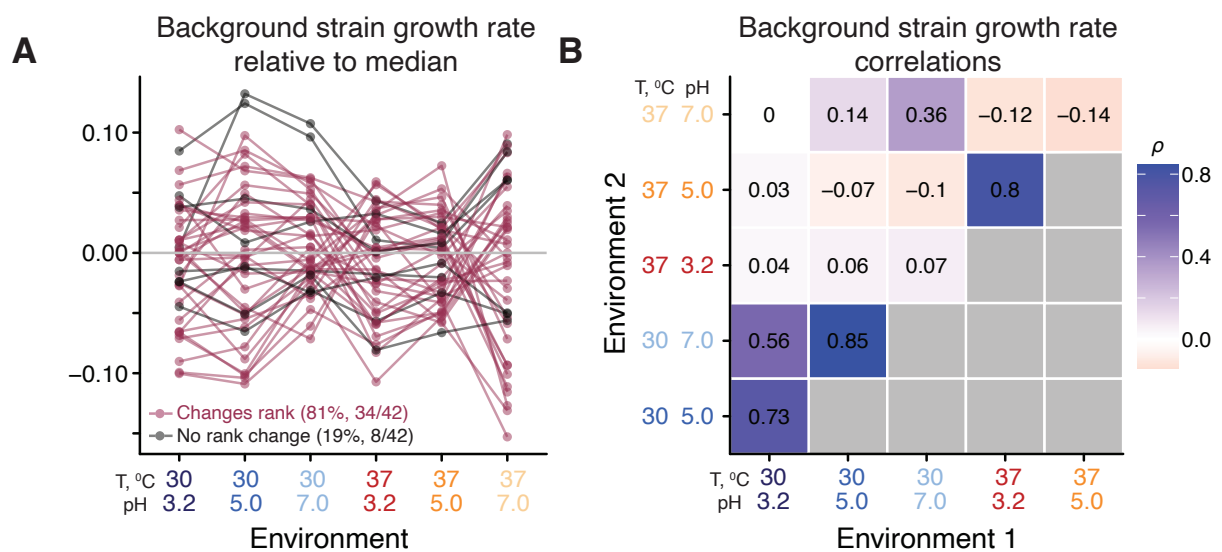


**Figure B.5. Variable and invariant slope models** **A.** Bar graph showing the percent of mutations best fit by the invariant slopes and variable slopes models (see Section 2.3.3). Schematic example of each model is shown on top. **B.** The adjusted  $R^2$  for the fits of both models for all mutations. **C.** Relationship between the pivot growth rates estimated from each model. Error bars are  $\pm 1$  standard error. **D, E.** As (A, B) but with a reduced data set with no missing measurements for each mutation (see Section 2.3.3). **F.** The fraction of significantly different slopes per mutation across environments in pairwise tests with the reduced data set (as Figure 3 in the main text).

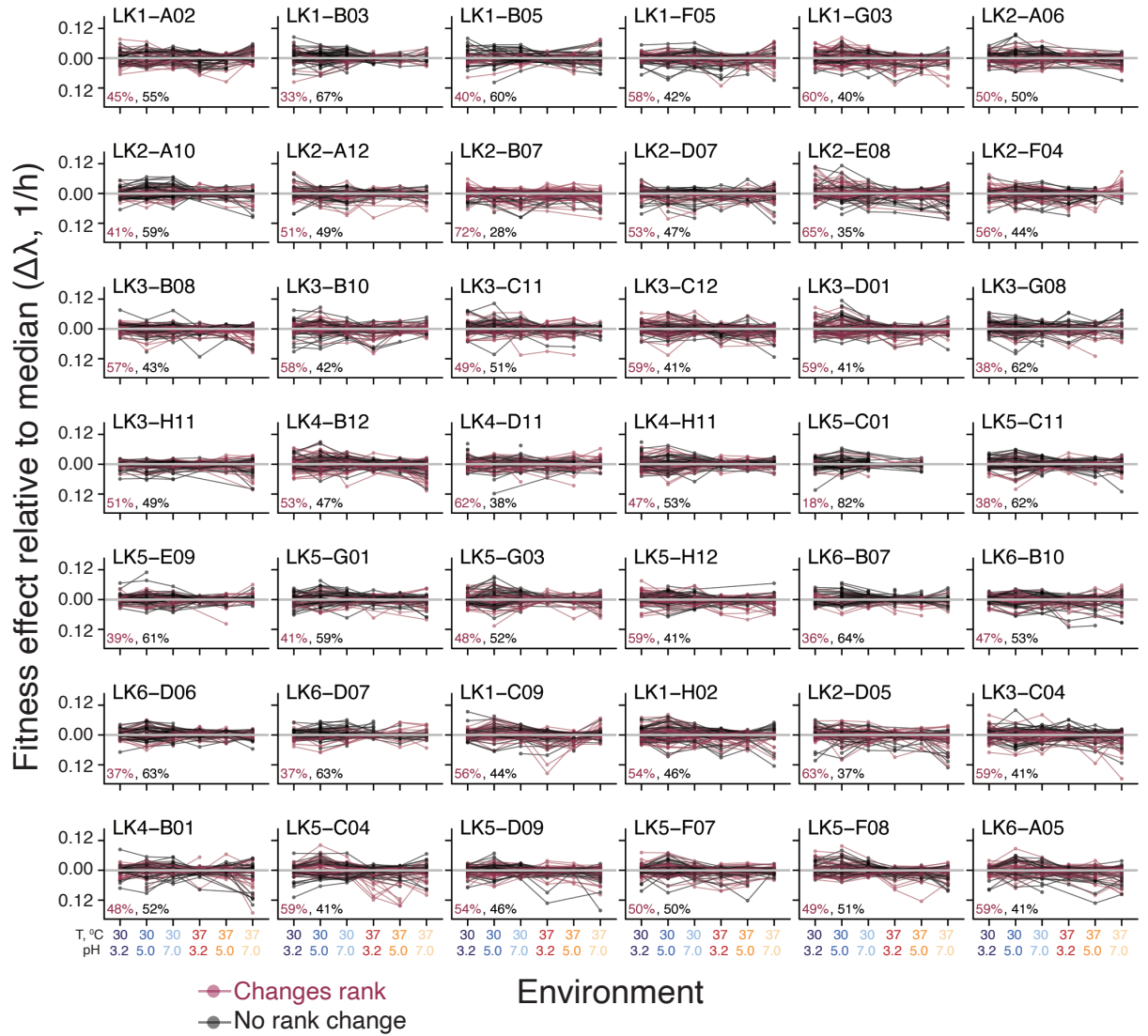


**Figure B.6. Global epistasis model for individual mutations.** Same as Figure 3, but with data points. Mutations are ordered by slope, as in Figure 3. y-axis varies across rows.

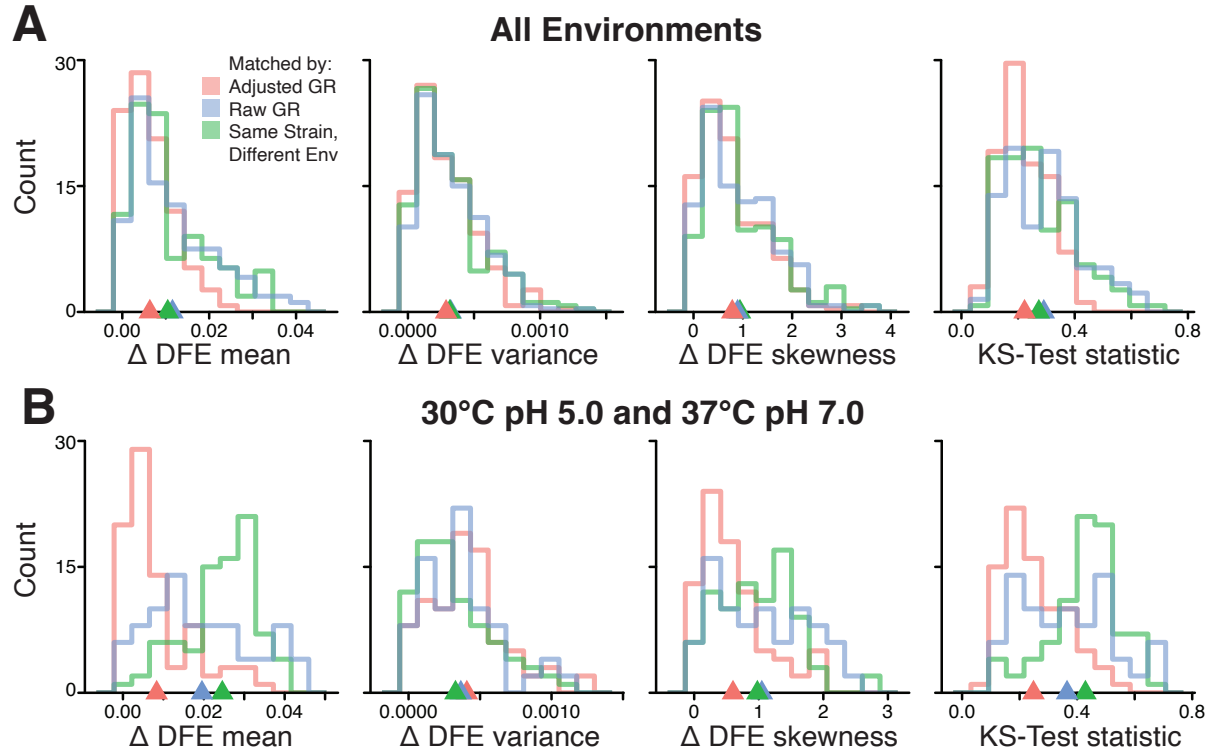




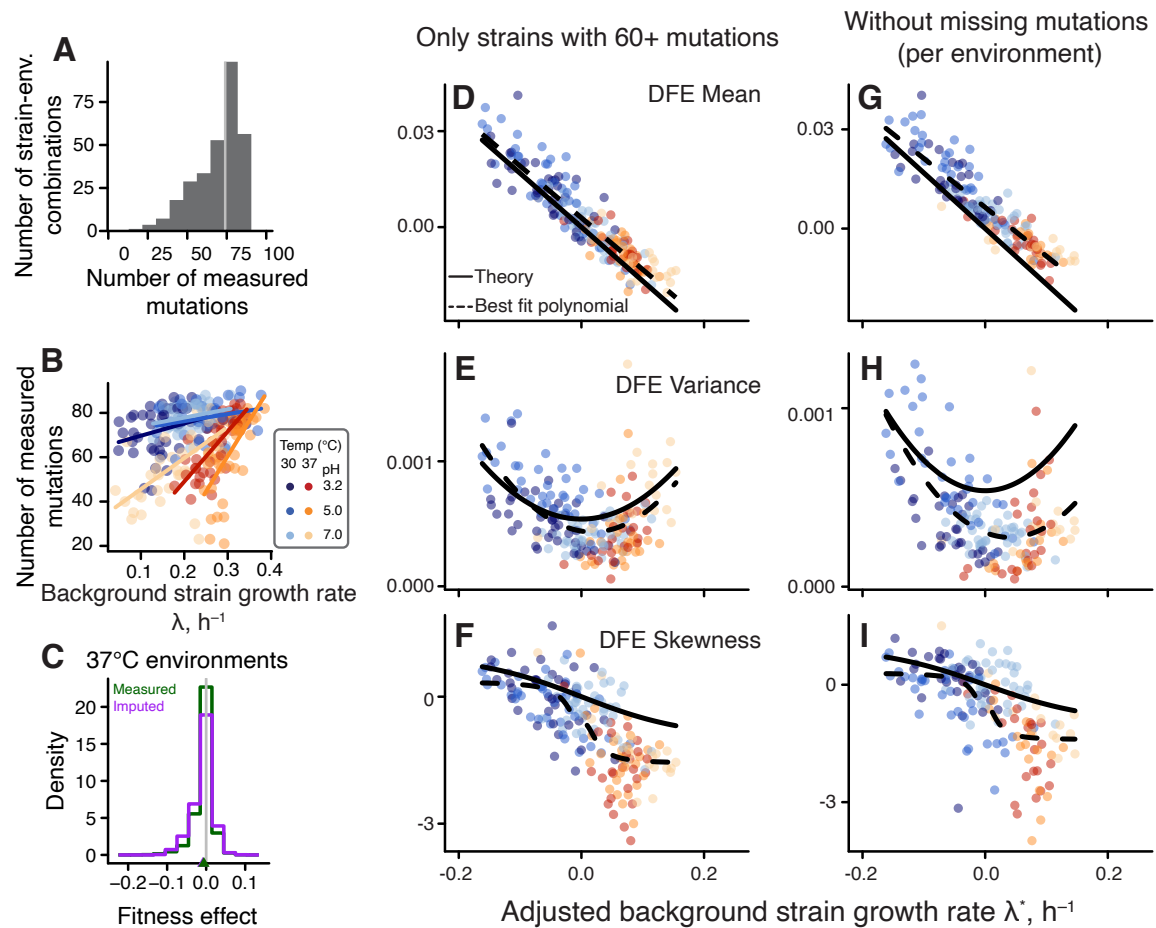
**Figure B.7. Reshuffling of background strain growth rates across environments.** **A.** Strain growth rate relative to the median growth rate in each environment. Lines connect the same strain across environments and are colored maroon if the strain is on different sides of the median in different environments (see Section 2.3.3). **B.** The correlation of background strain growth rate across all pairs of environments.



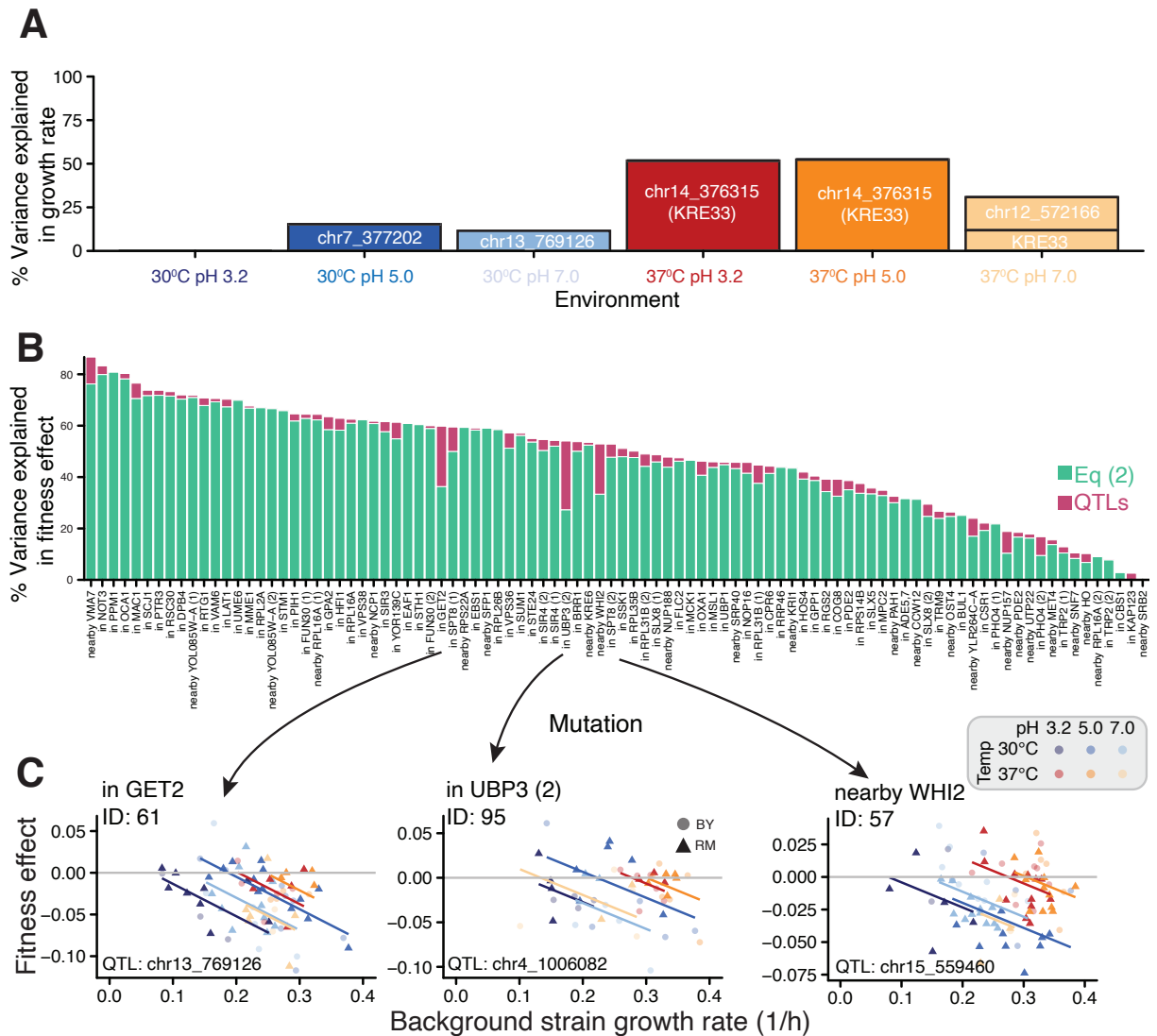
**Figure B.8. Reshuffling of the effects of mutations across environments.** Each panel corresponds to a background strain and shows the effect of all mutations relative to the median in each environment. Lines connect the same mutation across environments and are colored maroon if the mutation is on different sides of the median in different environments (see Section 2.3.3).



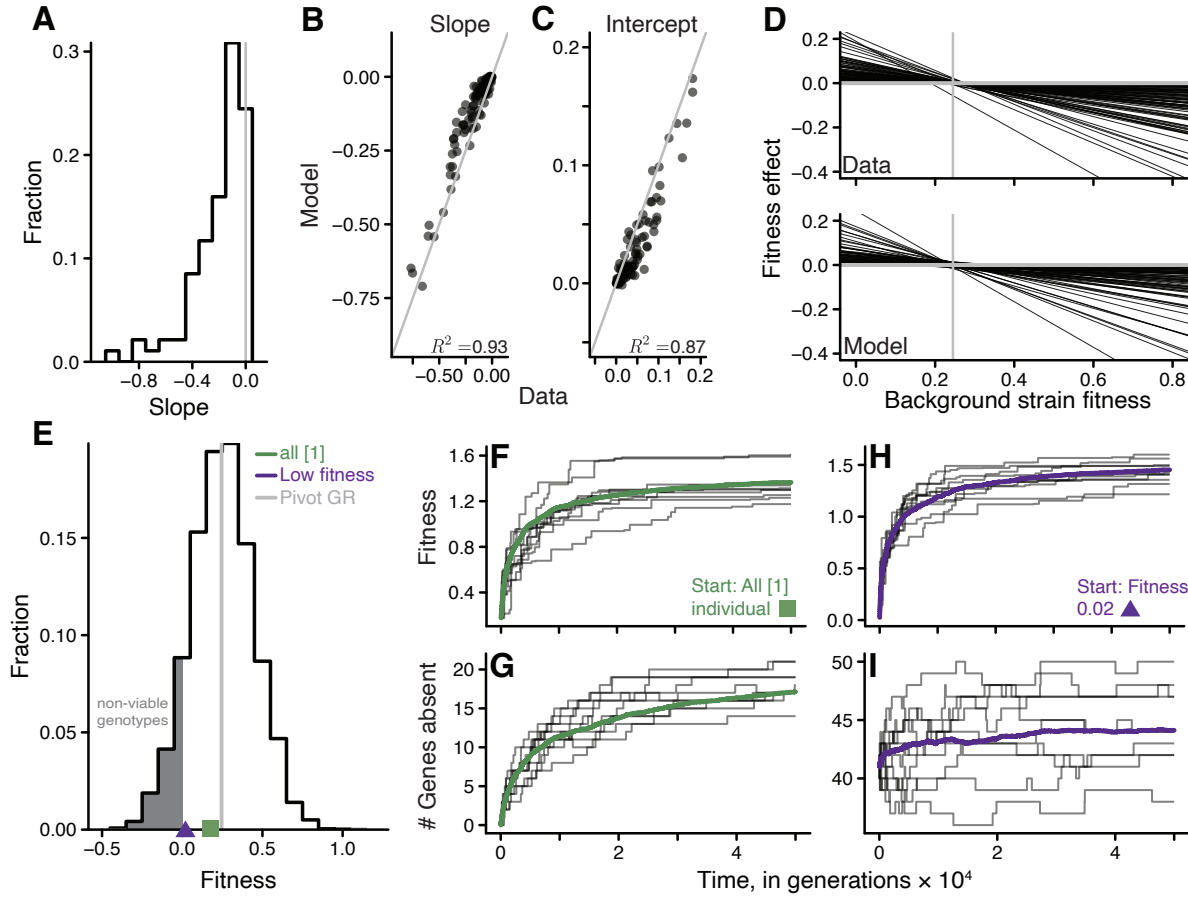
**Figure B.9. Distributions of DFEs similarity statistics.** **A.** Distribution of four metrics of DFE similarity for all pairs of strains from any two different environments and matched either by their adjusted growth rate (red), raw growth rate (blue), or strain identity (green). See Section 2.3.3 for details. For all metrics, lower values mean more similar DFEs. Triangle shows the mean of the corresponding colored distribution. **B.** Same as A but with the strains sampled from the two most dissimilar environments, 30°C pH 5.0 and 37°C pH 7.0.



**Figure B.10. Robustness of the observed DFE variation with respect to missing measurements.** **A.** Distribution of the number of mutations measured per DFE. **B.** Relationship between the number of mutations in each DFE and the background growth rate. Lines represent the best fit linear regression. **C.** Distribution of all measured mutation effects and the imputed missing mutation effects in 37°C environments (see Section 2.3.3). **D-F.** Same as Figure 4 but excluding all strains whose DFE contains less than 60 mutations. **G-I.** Same as Figure 4 but based on a reduced data set without missing measurements (see Section 2.3.3). Colors for D-I as in Panel B.



**Figure B.11. QTL analysis.** **A.** Percent of variance in background strain growth rate explained by a model with 10 tested loci (see Section 2.3.3). Loci that explain a significant ( $P < 0.05$ ) amount of variance are labelled, other loci are not shown. **B.** The percent variance in fitness effect of each mutation explained by 10 candidate loci combined (pink) above and beyond the variance explained by the generalized global epistasis equation (2.2) in the main text (teal). Mutations are ordered by the total explained variance. **C.** Three example mutations, with lines representing the best fit generalized global epistasis model, colored by environment. Point shape represents the allele, either BY (circles) or RM (triangles), at the locus explaining most variation for that mutation (locus indicated in the bottom left of each panel).



**Figure B.12. Connectedness model and evolutionary simulations.** **A.** Distribution of global epistasis slopes in the concatenated data from 30°C pH 5.0 environment (see Section 2.4.3). **B, C.** Correlation between the empirical global epistasis slopes and intercepts and those estimated from the connectedness model. Grey line is the identity line and  $R^2$  from the best fit linear regression is reported ( $P < 0.01$ ). **D.** “Bowtie” plot of all mutation regressions from the data (top) and the connectedness model (bottom). **E.** Fitness distribution of 5,000 random genotypes sampled from the connectedness model. Gray line indicates the pivot growth rate. Green rhombus indicates the genotype with +1 at all sites representing all genes being intact. Purple triangle indicates a low-fitness genotype. **F.** Fitness in Wright-Fisher simulations with  $NU = 0.1$  (Section 2.4.3), initiated with a genotype with all intact genes (see panel E). Thick green line is mean over 50 replicates, with 10 random individual trajectories shown in gray. **G.** Count of intact genes from the simulations described in panel F. **H, I.** Same as panels F, G, but with the initial low-fitness genotype (panel E) that initially has 52/94 intact genes.

## B.1 Supplementary tables

**Table B.1.** Statistics of slope and intercept distributions.

| Environment |     | Slopes |       |        | Intercepts |        |        |
|-------------|-----|--------|-------|--------|------------|--------|--------|
| Temp, °C    | pH  | Mean   | Stdev | Skew   | Mean       | Stdev  | Skew   |
| 30          | 3.2 | −0.160 | 0.143 | −1.40  | 0.0326     | 0.0313 | 0.824  |
| 30          | 5.0 | −0.214 | 0.186 | −0.922 | 0.0640     | 0.0556 | 0.404  |
| 30          | 7.0 | −0.161 | 0.157 | −0.824 | 0.0316     | 0.0337 | 0.230  |
| 37          | 3.2 | −0.136 | 0.151 | −1.09  | 0.0311     | 0.0378 | 0.581  |
| 37          | 5.0 | −0.121 | 0.157 | −0.839 | 0.0292     | 0.0498 | 0.324  |
| 37          | 7.0 | −0.137 | 0.189 | −0.731 | 0.0169     | 0.0401 | −0.790 |

**Table B.2.** Antibiotic concentrations used in this study, in  $\mu\text{g/ml}$ .

| <b>Antibiotic</b>    | <b><i>E. coli</i></b> | <b><i>S. cerevisiae</i></b> |
|----------------------|-----------------------|-----------------------------|
| Kanamycin (Kan)      | 40                    | N/A                         |
| Ampicilin (Amp)      | 100                   | 100                         |
| Nourseothricin (Nat) | 20                    | 20                          |
| Hygromycin (Hyg)     | 200                   | 300                         |



**Table B.3.** Best fit trends in DFE moments. “Pwr” refers to the power of to each the predictor (background growth rate) is raised in each term of the polynomial (see 2.3.4). “Factor” is the scaling factor by which all entries in the row are multiplied. “Reduced data 1” is the data set with only DFEs made of  $> 60$  mutations. “Reduced data 2” is the data set with no missing mutations per environment. “Est” refers to the best-fit parameter value.

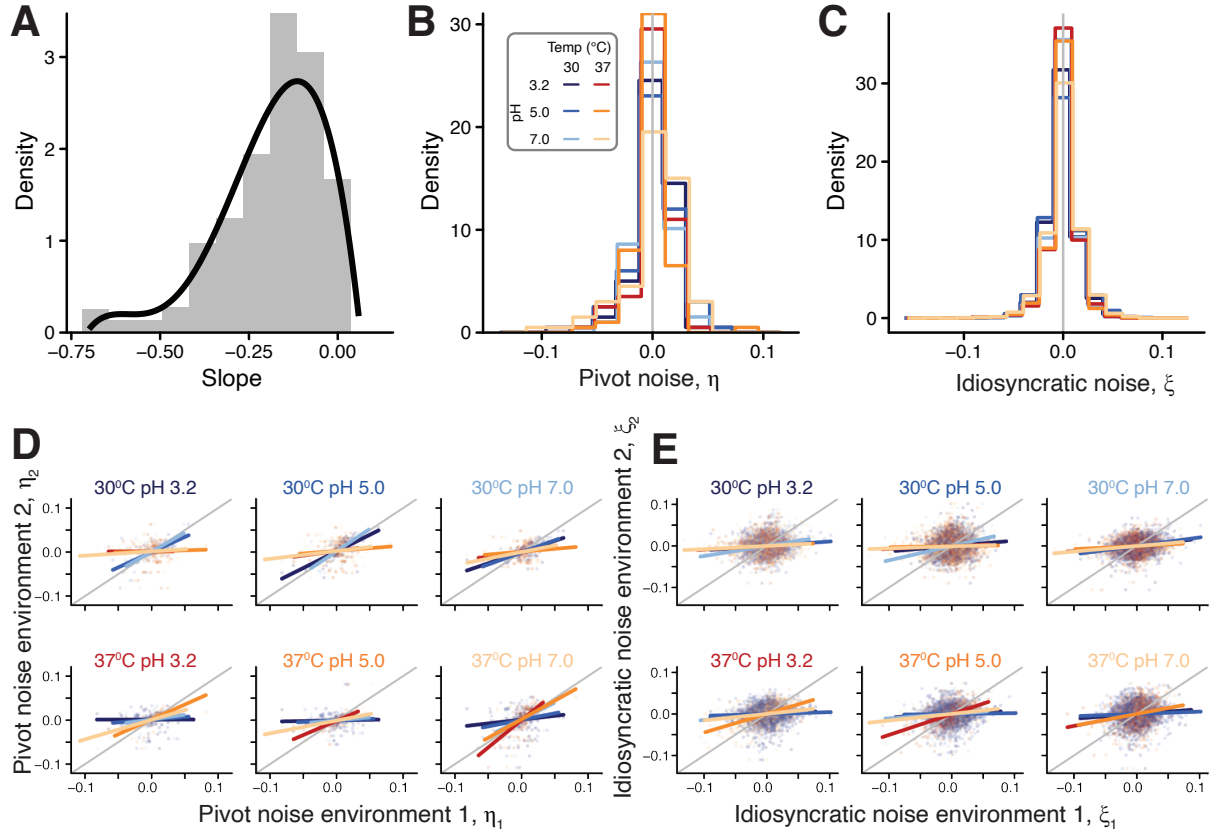
| Moment | Pwr | Factor       | Theory | Full data |                | Reduced data 1 |               | Reduced data 2 |               |
|--------|-----|--------------|--------|-----------|----------------|----------------|---------------|----------------|---------------|
|        |     |              |        | Est.      | 95% CI         | Est.           | 95% CI        | Est.           | 95% CI        |
| Mean   | 0   | $10^{-3}$    | 0      | 2.2       | [1.5, 2.9]     | 2.9            | [2.2, 3.6]    | 6.4            | [5.7, 7.2]    |
|        | 1   | $(-10^{-1})$ | 1.7    | 1.7       | [1.8, 1.7]     | 1.6            | [1.7, 1.5]    | 1.5            | [1.6, 1.4]    |
| Var    | 0   | $10^{-4}$    | 5.4    | 4.6       | [4.3, 5.0]     | 4.4            | [4.0, 4.9]    | 3.0            | [2.6, 3.4]    |
|        | 1   | $10^{-4}$    | 0      | 2.0       | [-2.5, 6.5]    | -7.8           | [-12, -3.2]   | -13            | [-18, -9]     |
|        | 2   | $10^{-2}$    | 1.7    | 2.4       | [2.1, 2.8]     | 2.1            | [1.6, 2.7]    | 1.7            | [1.2, 2.2]    |
| Skew   | 0   | $(-1)$       | 0      | 0.39      | [0.52, 0.27]   | 0.41           | [0.56, 0.26]  | 0.27           | [0.47, 0.07]  |
|        | 1   | $10^{-2}$    | -3.1   | -4.6      | [-5.7, -3.4]   | -5.4           | [-6.9, 3.9]   | -2.5           | [-3.7, -1.5]  |
|        | 2   | $10^{-1}$    | 0      | 0.02      | [-1.5, 1.5]    | -0.7           | [-2.4, 1.0]   | -0.13          | [-1.6, 1.3]   |
|        | 3   | 1            | -1.12  | -1.21     | [-2.10, -0.33] | -1.52          | [-3.07, 0.03] | -1.03          | [-2.31, 0.25] |

**Table B.4.** GO term enrichment among all mutations in this study. *P*-value is reported after the Benjamini-Hochberg correction.

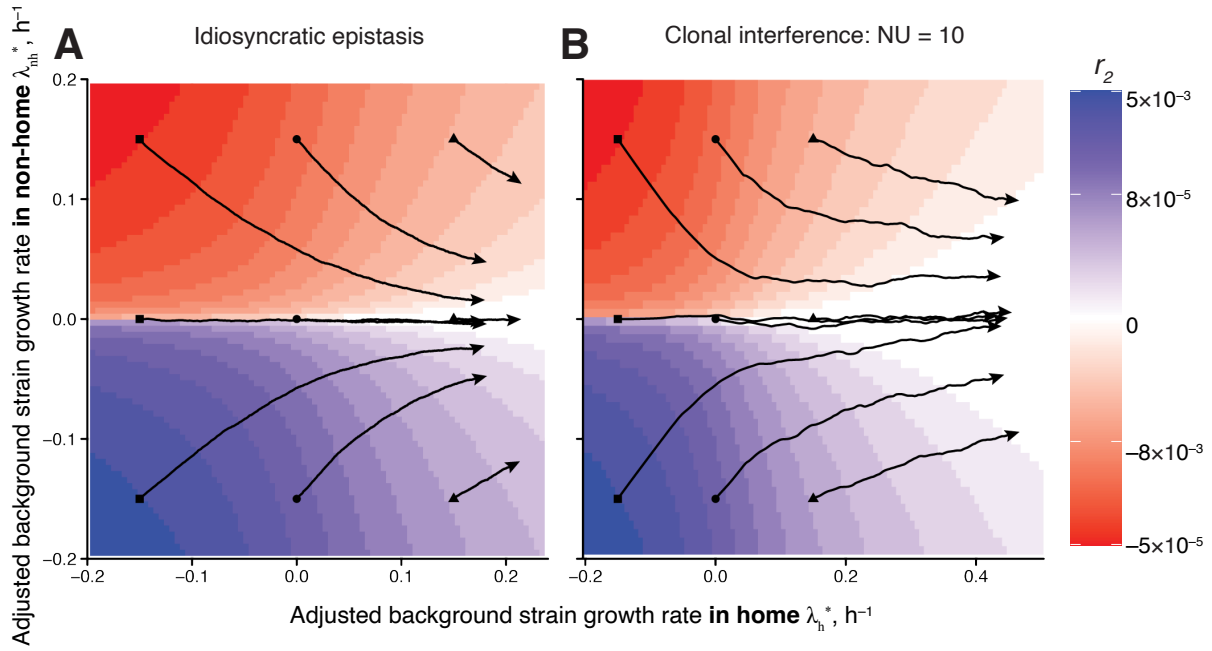
| <b>GO Term</b>                                                                | <b><i>P</i>-value</b> | <b>Fold enrichment</b> |
|-------------------------------------------------------------------------------|-----------------------|------------------------|
| double-strand break repair via<br>nonhomologous end joining                   | 0.05                  | 10.0                   |
| nucleotide transport                                                          | 0.04                  | 8.5                    |
| organophosphate ester transport                                               | 0.05                  | 5.7                    |
| chromatin remodeling                                                          | 0.04                  | 4.4                    |
| nucleobase-containing compound<br>transport                                   | 0.05                  | 3.8                    |
| chromatin organization                                                        | 0.04                  | 3.9                    |
| protein-DNA complex organization                                              | 0.04                  | 3.2                    |
| positive regulation of cellular biosynthetic<br>process                       | 0.05                  | 2.7                    |
| positive regulation of biosynthetic process                                   | 0.05                  | 2.7                    |
| positive regulation of<br>nucleobase-containing compound<br>metabolic process | 0.05                  | 2.6                    |
| regulation of transcription by RNA<br>polymerase II                           | 0.04                  | 2.8                    |

## **Appendix C**

### **Chapter 3 Supplementary Figures and Tables**



**Figure C.1. Noise distributions and covariance.** **A.** Global distribution of microscopic epistasis slopes as reported in [179]. Black curve is the fitted 5<sup>th</sup> degree polynomial describing the probability density. **B.** Distribution of pivot noise  $\eta$  values, colored by environment. **C.** Distribution of idiosyncratic noise  $\xi$  values, colored by environment. Legend as in B. **D.** Covariance of  $\eta$  values across environments. Each panel is one 'environment 1' and each point is the matched  $\eta$  value for the same mutation, colored by environment 2. Lines are best fit linear regressions, where  $P > 0.05$  for all environment pairs of opposite temperatures apart from 37°CpH 7.0 with 30°CpH 5.0 and 7.0.  $P < 0.05$  for all environment pairs of the same temperature **E.** As in E. but for  $\xi$  values. All regressions are significant  $P < 0.05$  apart from 30°C pH 5.0 and 37°C pH 5.0



**Figure C.2. Impact of clonal interference and idiosyncratic epistasis on evolutionary trajectories** Same as Figure 3.5 but **A.** with the addition of idiosyncratic epistasis (see Materials and Methods) or **B.** in the clonal interference regime (see Materials and Methods). Each line represents the average of 100 replicate simulations of 4000 generations each.

## C.1 Supplementary Tables

**Table C.1.** Parameters of slope distribution density curve

| Power | Value   | Std. Err |
|-------|---------|----------|
| 0     | 1.74    | 0.005    |
| 1     | -19.40  | 0.167    |
| 2     | -111.12 | 1.750    |
| 3     | -155.22 | 7.029    |
| 4     | -10.46  | 11.882   |
| 5     | 68.85   | 7.124    |

# Bibliography

- [1] Pál, C., Papp, B., & Lázár, V. (2015). Collateral sensitivity of antibiotic-resistant microbes. *Trends in microbiology*, 23(7), 401–407.
- [2] Pluchino, K. M., Hall, M. D., Goldsborough, A. S., Callaghan, R., & Gottesman, M. M. (2012). Collateral sensitivity as a strategy against cancer multidrug resistance. *Drug Resistance Updates*, 15(1-2), 98–105.
- [3] Roemhild, R., Linkevicius, M., & Andersson, D. I. (2020). Molecular mechanisms of collateral sensitivity to the antibiotic nitrofurantoin. *PLoS Biology*, 18(1), e3000612.
- [4] Jensen, P., Holm, B., Sorensen, M, Christensen, I., & Sehested, M. (1997). In vitro cross-resistance and collateral sensitivity in seven resistant small-cell lung cancer cell lines: Preclinical identification of suitable drug partners to taxotere, taxol, topotecan and gemcitabin. *British journal of cancer*, 75(6), 869–877.
- [5] Imamovic, L., & Sommer, M. O. (2013). Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. *Science translational medicine*, 5(204), 204ra132.
- [6] Lázár, V., Martins, A., Spohn, R., Daruka, L., Grézal, G., Fekete, G., Számel, M., Jangir, P. K., Kintses, B., Csörgő, B., Nyerges, Á., Györkei, Á., Kincses, A., Dér, A., Fruzsing, W. R., Deli, M. A., Urbán, E., Hegedüs, Z., Olajos, G., ... Pál, C. (2018). Antibiotic-resistant bacteria show widespread collateral sensitivity to antimicrobial peptides. *Nature microbiology*, 3(6), 718–731.
- [7] Maltas, J., & Wood, K. B. (2019). Pervasive and diverse collateral sensitivity profiles inform optimal strategies to limit antibiotic resistance. *PLoS Biology*, 17(10), e3000515.

- [8] Batra, A., Roemhild, R., Rousseau, E., Franzenburg, S., Niemann, S., & Schulenburg, H. (2021). High potency of sequential therapy with only beta-lactam antibiotics. *Elife*, *10*, e68876.
- [9] Sanz-García, F., Sánchez, M. B., Hernando-Amado, S., & Martínez, J. L. (2020). Evolutionary landscapes of pseudomonas aeruginosa towards ribosome-targeting antibiotic resistance depend on selection strength. *International journal of antimicrobial agents*, *55*(6), 105965.
- [10] Schenk, M. F., Witte, S., Salverda, M. L., Koopmanschap, B., Krug, J., & de Visser, J. A. G. (2015). Role of pleiotropy during adaptation of tem-1  $\beta$ -lactamase to two novel antibiotics. *Evolutionary applications*, *8*(3), 248–260.
- [11] Lázár, V., Pal Singh, G., Spohn, R., Nagy, I., Horváth, B., Hrtyan, M., Busa-Fekete, R., Bogos, B., Méhi, O., Csörgő, B., & Pósfai, G. (2013). Bacterial evolution of antibiotic hypersensitivity. *Molecular systems biology*, *9*(1), 700.
- [12] Barbosa, C., Römhild, R., Rosenstiel, P., & Schulenburg, H. (2019). Evolutionary stability of collateral sensitivity to antibiotics in the model pathogen pseudomonas aeruginosa. *Elife*, *8*, e51481.
- [13] Hernando-Amado, S., Sanz-García, F., & Martínez, J. L. (2020). Rapid and robust evolution of collateral sensitivity in pseudomonas aeruginosa antibiotic-resistant mutants. *Science advances*, *6*(32), eaba5493.
- [14] Kim, S., Lieberman, T. D., & Kishony, R. (2014). Alternating antibiotic treatments constrain evolutionary paths to multidrug resistance. *Proceedings of the National Academy of Sciences*, *111*(40), 14494–14499.
- [15] Jahn, L. J., Simon, D., Jensen, M., Bradshaw, C., Ellabaan, M. M. H., & Sommer, M. O. A. (2021). Compatibility of evolutionary responses to constituent antibiotics drive resistance evolution to drug pairs. *Molecular biology and evolution*, *38*(5), 2057–2069.



- [16] Kavanaugh, L. G., Flanagan, J. N., & Steck, T. R. (2020). Reciprocal antibiotic collateral sensitivity in burkholderia multivorans. *International journal of antimicrobial agents*, 56(1), 105994.
- [17] Laborda, P., Martínez, J. L., & Hernando-Amado, S. (2021). Convergent phenotypic evolution towards fosfomycin collateral sensitivity of pseudomonas aeruginosa antibiotic-resistant mutants. *Microbial Biotechnology*.
- [18] Oz, T., Guvenek, A., Yildiz, S., Karaboga, E., Tamer, Y. T., Mumcuyan, N., Ozan, V. B., Senturk, G. H., Cokol, M., Yeh, P., & Toprak, E. (2014). Strength of selection pressure is an important parameter contributing to the complexity of antibiotic resistance evolution. *Molecular biology and evolution*, 31(9), 2387–2401.
- [19] Munck, C., Gumpert, H. K., Wallin, A. I. N., Wang, H. H., & Sommer, M. O. (2014). Prediction of resistance development against drug combinations by collateral responses to component drugs. *Science translational medicine*, 6(262), 262ra156–262ra156.
- [20] Lázár, V., Nagy, I., Spohn, R., Csörgő, B., Györkei, Á., Nyerges, Á., Horváth, B., Vörös, A., Busa-Fekete, R., Hrtyan, M., & Bogos, B. (2014). Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nature communications*, 5(1), 1–12.
- [21] Bergstrom, C. T., Lo, M., & Lipsitch, M. (2004). Ecological theory suggests that antimicrobial cycling will not reduce antimicrobial resistance in hospitals. *Proceedings of the National Academy of Sciences*, 101(36), 13285–13290.
- [22] Barbosa, C., Trebosc, V., Kemmer, C., Rosenstiel, P., Beardmore, R., Schulenburg, H., & Jansen, G. (2017). Alternative evolutionary paths to bacterial antibiotic resistance cause distinct collateral effects. *Molecular biology and evolution*, 34(9), 2229–2244.
- [23] Mira, P. M., Meza, J. C., Nandipati, A., & Barlow, M. (2015). Adaptive landscapes of resistance genes change as antibiotic concentrations change. *Molecular biology and evolution*, 32(10), 2707–2715.

- [24] Das, S. G., Direito, S. O., Waclaw, B., Allen, R. J., & Krug, J. (2020). Predictable properties of fitness landscapes induced by adaptational tradeoffs. *eLife*, 9, e55155.
- [25] Pinheiro, F., Warsi, O., Andersson, D. I., & Lässig, M. (2021). Metabolic fitness landscapes predict the evolution of antibiotic resistance. *Nature Ecology & Evolution*, 1–11.
- [26] Card, K. J., Thomas, M. D., Graves, J. L., Barrick, J. E., & Lenski, R. E. (2021). Genomic evolution of antibiotic resistance is contingent on genetic background following a long-term experiment with *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 118(5).
- [27] Gjini, E., & Wood, K. B. (2021). Price equation captures the role of drug interactions and collateral effects in the evolution of multidrug resistance. *Elife*, 10, e64851.
- [28] Nichol, D., Rutter, J., Bryant, C., Hujer, A. M., Lek, S., Adams, M. D., Jeavons, P., Anderson, A. R. A., Bonomo, R. A., & Scott, J. G. (2019). Antibiotic collateral sensitivity is contingent on the repeatability of evolution. *Nat. Commun.*, 10(1), 334.
- [29] Jerison, E. R., Nguyen Ba, A. N., Desai, M. M., & Kryazhimskiy, S. (2020). Chance and necessity in the pleiotropic consequences of adaptation for budding yeast. *Nature Ecology & Evolution*, 4(4), 601–611.
- [30] Wagner, G. P., & Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: The evolvability of complex organisms. *Nature Reviews Genetics*, 12(3), 204–213.
- [31] Paaby, A. B., & Rockman, M. V. (2013). The many faces of pleiotropy. *Trends in genetics*, 29(2), 66–73.
- [32] Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.
- [33] Rose, M. R. (1982). Antagonistic pleiotropy, dominance, and genetic variation. *Heredity*, 48(1), 63–78.
- [34] Barton, N. (1990). Pleiotropic models of quantitative variation. *Genetics*, 124(3), 773–782.

- [35] Slatkin, M., & Frank, S. A. (1990). The quantitative genetic consequences of pleiotropy under stabilizing and directional selection. *Genetics*, 125(1), 207–213.
- [36] Jones, A. G., Arnold, S. J., & Bürger, R. (2003). Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution*, 57(8), 1747–1760.
- [37] Johnson, T., & Barton, N. (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1459), 1411–1425.
- [38] Connallon, T., & Clark, A. G. (2015). The distribution of fitness effects in an uncertain world. *Evolution*, 69(6), 1610–1618.
- [39] Martin, G., & Lenormand, T. (2015). The fitness effect of mutations across environments: Fisher’s geometrical model with multiple optima. *Evolution*, 69(6), 1433–1447.
- [40] Harmand, N., Gallet, R., Jabbour-Zahab, R., Martin, G., & Lenormand, T. (2017). Fisher’s geometrical model and the mutational patterns of antibiotic resistance across dose gradients. *Evolution*, 71(1), 23–37.
- [41] Wang, S., & Dai, L. (2019). Evolving generalists in switching rugged landscapes. *PLoS Comput Biology*, 15(10), e1007320.
- [42] Tikhonov, M., Kachru, S., & Fisher, D. S. (2020). A model for the interplay between plastic tradeoffs and evolution in changing environments. *Proceedings of the National Academy of Sciences*, 117(16), 8934–8940.
- [43] Jerison, E. R., Kryazhimskiy, S., & Desai, M. M. (2014). Pleiotropic consequences of adaptation across gradations of environmental stress in budding yeast. *arXiv:1409.7839*. <https://arxiv.org/abs/1409.7839>
- [44] Bono, L. M., Smith, L. B., Jr, Pfennig, D. W., & Burch, C. L. (2017). The emergence of performance trade-offs during local adaptation: Insights from experimental evolution. *Mol. Ecol.*, 26(7), 1720–1733.

- [45] King, J. L. The role of mutation in evolution. In: *Proceedings of the sixth berkeley symposium on mathematical statistics and probability*. 5. University of California Press Berkeley. 1972, 69–100.
- [46] Ohta, T. (1987). Very slightly deleterious mutations and the molecular clock. *Journal of Molecular Evolution*, 26(1-2), 1–6.
- [47] Orr, H. A. (2003). The distribution of fitness effects among beneficial mutations. *Genetics*, 163(4), 1519–1526.
- [48] Kassen, R., & Bataillon, T. (2006). Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nat. Genet.*, 38(4), 484–488.
- [49] Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8), 610–618.
- [50] Martin, G., & Lenormand, T. (2008). The distribution of beneficial and fixed mutation fitness effects close to an optimum. *Genetics*, 179(2), 907–916.
- [51] MacLean, R. C., & Buckling, A. (2009). The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. *PLoS Genet*, 5(3), e1000406.
- [52] Kryazhimskiy, S., Tkačik, G., & Plotkin, J. B. (2009). The dynamics of adaptation on correlated fitness landscapes. *Proceedings of the National Academy of Sciences*, 106(44), 18638–18643.
- [53] Levy, S. F., Blundell, J. R., Venkataram, S., Petrov, D. A., Fisher, D. S., & Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, 519(7542), 181–186.
- [54] Qian, W., Ma, D., Xiao, C., Wang, Z., & Zhang, J. (2012). The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell reports*, 2(5), 1399–1410.
- [55] Van Opijnen, T., Bodi, K. L., & Camilli, A. (2009). Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature methods*, 6(10), 767–772.

- [56] Chevereau, G., Dravecká, M., Batur, T., Guvenek, A., Ayhan, D. H., Toprak, E., & Bollenbach, T. (2015). Quantifying the determinants of evolutionary dynamics leading to drug resistance. *PLoS Biology*, *13*(11), e1002299.
- [57] Blundell, J. R., Schwartz, K., Francois, D., Fisher, D. S., Sherlock, G., & Levy, S. F. (2019). The dynamics of adaptive genetic diversity during the early stages of clonal evolution. *Nature ecology & evolution*, *3*(2), 293–301.
- [58] Aggeli, D., Li, Y., & Sherlock, G. (2021). Changes in the distribution of fitness effects and adaptive mutational spectra following a single first step towards adaptation. *Nature Communications*, *12*(1), 5193.
- [59] Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., & Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature*, *551*(7678), 45–50.
- [60] Venkataram, S., Monasky, R., Sikaroodi, S. H., Kryazhimskiy, S., & Kaçar, B. (2020). 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.
- [61] Stiffler, M. A., Hekstra, D. R., & Ranganathan, R. (2015). Evolvability as a function of purifying selection in TEM-1  $\beta$ -lactamase. *Cell*, *160*(5), 882–892.
- [62] Crow, J. F., & Kimura, M. (1972). *An introduction to population genetics theory*. Harper & Row Ltd.
- [63] Roff, D. A. (2007). A centennial celebration for quantitative genetics. *Evolution*, *61*(5), 1017–1032.
- [64] Shoal, O., Sheftel, H., Shinar, G., Hart, Y., Ramote, O., Mayo, A., Dekel, E., Kavanagh, K., & Alon, U. (2012). Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science*, *336*(6085), 1157–1160.
- [65] Li, Y., Petrov, D. A., & Sherlock, G. (2019). Single nucleotide mapping of trait space reveals pareto fronts that constrain adaptation. *Nat Ecol Evol*, *3*(11), 1539–1551.
- [66] Orr, H. A. (2000). The rate of adaptation in asexuals. *Genetics*, *155*(2), 961–968.

- [67] Desai, M. M., & Fisher, D. S. (2007). Beneficial mutation–selection balance and the effect of linkage on positive selection. *Genetics*, 176(3), 1759–1798.
- [68] Good, B. H., & Desai, M. M. (2015). The impact of macroscopic epistasis on long-term evolutionary dynamics. *Genetics*, 199(1), 177–190.
- [69] Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics*, 47(6), 713.
- [70] Lang, G. I., Rice, D. P., Hickman, M. J., Sodergren, E., Weinstock, G. M., Botstein, D., & Desai, M. M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*, 500(7464), 571–574.
- [71] Schiffels, S., Szöllősi, G. J., Mustonen, V., & Lässig, M. (2011). Emergent neutrality in adaptive asexual evolution. *Genetics*, 189(4), 1361–1375.
- [72] Good, B. H., Rouzine, I. M., Balick, D. J., Hallatschek, O., & Desai, M. M. (2012). Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proceedings of the National Academy of Sciences*, 109(13), 4950–4955.
- [73] Bakerlee, C. W., Phillips, A. M., Nguyen Ba, A. N., & Desai, M. M. (2021). Dynamics and variability in the pleiotropic effects of adaptation in laboratory budding yeast populations. *Elife*, 10.
- [74] Andersson, D. I., & Hughes, D. (2014). Microbiological effects of sublethal levels of antibiotics. *Nature Reviews Microbiology*, 12(7), 465–478.
- [75] Lindsey, H. A., Gallie, J., Taylor, S., & Kerr, B. (2013). Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature*, 494(7438), 463–467.
- [76] Nguyen Ba, A. N., Cvijović, I., Echenique, J. I. R., Lawrence, K. R., Rego-Costa, A., Liu, X., Levy, S. F., & Desai, M. M. (2019). High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature*, 575(7783), 494–499.
- [77] Venkataram, S., Dunn, B., Li, Y., Agarwala, A., Chang, J., Ebel, E. R., Geiler-Samerotte, K., Hérissant, L., Blundell, J. R., Levy, S. F., & Fisher, D. (2016). Development of a

- comprehensive genotype-to-fitness map of adaptation-driving mutations in yeast. *Cell*, 166(6), 1585–1596.
- [78] Sun, D., Jeannot, K., Xiao, Y., & Knapp, C. W. (2019). Horizontal gene transfer mediated bacterial antibiotic resistance. *Frontiers in microbiology*, 10, 1933.
  - [79] Neher, R. A., Shraiman, B. I., & Fisher, D. S. (2010). Rate of adaptation in large sexual populations. *Genetics*, 184(2), 467–481.
  - [80] Visher, E., & Boots, M. (2020). The problem of mediocre generalists: Population genetics and eco-evolutionary perspectives on host breadth evolution in pathogens. *Proceedings of the Royal Society B*, 287(1933), 20201230.
  - [81] Levins, R. (1968). *Evolution in changing environments*. Princeton University Press.
  - [82] Kawecki, T. J. (1994). Accumulation of deleterious mutations and the evolutionary cost of being a generalist. *The American Naturalist*, 144(5), 833–838.
  - [83] Bono, L. M., Draghi, J. A., & Turner, P. E. (2020). Evolvability costs of niche expansion. *Trends in Genetics*, 36(1), 14–23.
  - [84] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289–300.
  - [85] Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry* (Vol. 1). Elsevier.
  - [86] Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4), 403–434.
  - [87] Koel, B. F., Burke, D. F., Bestebroer, T. M., van der Vliet, S., Zondag, G. C. M., Vervaet, G., Skepner, E., Lewis, N. S., Spronken, M. I. J., Russell, C. A., Eropkin, M. Y., Hurt, A. C., Barr, I. G., de Jong, J. C., Rimmelzwaan, G. F., Osterhaus, A. D. M. E., Fouchier, R. A. M., & Smith, D. J. (2013). Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science*, 342(6161), 976–979.

- [88] Dixon, G. B., Davies, S. W., Aglyamova, G. A., Meyer, E., Bay, L. K., & Matz, M. V. (2015). Genomic determinants of coral heat tolerance across latitudes. *Science*, 348(6242), 1460–1462.
- [89] Promdonkoy, P., Mhuantong, W., Champreda, V., Tanapongpipat, S., & Runguphan, W. (2020). Improvement in d-xylose utilization and isobutanol production in *S. cerevisiae* by adaptive laboratory evolution and rational engineering. *J. Ind. Microbiol. Biotechnol.*, 47(6-7), 497–510.
- [90] Kreiner, J. M., Latorre, S. M., Burbano, H. A., Stinchcombe, J. R., Otto, S. P., Weigel, D., & Wright, S. I. (2022). Rapid weed adaptation and range expansion in response to agriculture over the past two centuries. *Science*, 378(6624), 1079–1085.
- [91] Papp, B., Notebaart, R. A., & Pál, C. (2011). Systems-biology approaches for predicting genomic evolution. *Nat. Rev. Genet.*, 12(9), 591–602.
- [92] Munday, P. L., Warner, R. R., Monro, K., Pandolfi, J. M., & Marshall, D. J. (2013). Predicting evolutionary responses to climate change in the sea. *Ecol. Lett.*, 16(12), 1488–1500.
- [93] Lässig, M., Mustonen, V., & Walczak, A. M. (2017). Predicting evolution. *Nat Ecol Evol*, 1(3), 77.
- [94] Johnson, M. S., Venkataram, S., & Kryazhimskiy, S. (2023). Best practices in designing, sequencing, and identifying random dna barcodes. *Journal of Molecular Evolution*, 1–18.
- [95] Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life’s tape. *Science*, 362(6415).
- [96] Weinreich, D. M., Delaney, N. F., Depristo, M. A., & Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, 312(5770), 111–114.
- [97] Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., & Cooper, T. F. (2011). Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, 332(6034), 1193–1196.



- [98] Chou, H.-H., Chiu, H.-C., Delaney, N. F., Segrè, D., & Marx, C. J. (2011). Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, 332(6034), 1190–1192.
- [99] Gong, L. I., Suchard, M. A., & Bloom, J. D. (2013). Stability-mediated epistasis constrains the evolution of an influenza protein. *Elife*, 2, e00631.
- [100] Palmer, A. C., Toprak, E., Baym, M., Kim, S., Veres, A., Bershtein, S., & Kishony, R. (2015). Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. *Nat. Commun.*, 6, 7385.
- [101] Flynn, J. M., Rossouw, A., Cote-Hammarlof, P., Fragata, I., Mavor, D., Hollins, C., 3rd, Bank, C., & Bolon, D. N. (2020). Comprehensive fitness maps of hsp90 show widespread environmental dependence. *Elife*, 9.
- [102] Bakerlee, C. W., Ba, A. N. N., Shulgina, Y., Echenique, J. I. R., & Desai, M. M. (2022). Idiosyncratic epistasis leads to global fitness-correlated trends. *Science*, 376(6593), 630–635.
- [103] Gupta, A., Zaman, L., Strobel, H. M., Gallie, J., Burmeister, A. R., Kerr, B., Tamar, E. S., Kishony, R., & Meyer, J. R. (2022). Host-parasite coevolution promotes innovation through deformations in fitness landscapes. *Elife*, 11.
- [104] Diaz-Colunga, J., Sanchez, A., & Ogbunugafor, C. B. (2023a). Environmental modulation of global epistasis in a drug resistance fitness landscape. *Nature Communications*, 14(1), 8055.
- [105] Payne, J. L., & Wagner, A. (2018). The causes of evolvability and their evolution. *Nat. Rev. Genet.*, 20(1), 24–38.
- [106] Bank, C. (2022). Epistasis and adaptation on fitness landscapes. *Annu. Rev. Ecol. Evol. Syst.*, 53(1), 457–479.
- [107] Bank, C., Matuszewski, S., Hietpas, R. T., & Jensen, J. D. (2016). On the (un)predictability of a large intragenic fitness landscape. *Proceedings of the National Academy of Sciences*, 113(49), 14085–14090.

- [108] Wisner, M. J., Ribeck, N., & Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science*, 342(6164), 1364–1367.
- [109] Ardell, S. M., & Kryazhimskiy, S. (2021). The population genetics of collateral resistance and sensitivity. *Elife*, 10.
- [110] Domingo-Calap, P., Cuevas, J. M., & Sanjuán, R. (2009). The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet.*, 5(11), e1000742.
- [111] Kemble, H., Eisenhauer, C., Couce, A., Chapron, A., Magnan, M., Gautier, G., Le Nagard, H., Nghe, P., & Tenaillon, O. (2020). Flux, toxicity, and expression costs generate complex genetic interactions in a metabolic pathway. *Sci Adv*, 6(23), eabb2236.
- [112] Sane, M., Diwan, G. D., Bhat, B. A., Wahl, L. M., & Agashe, D. (2023). Shifts in mutation spectra enhance access to beneficial mutations. *Proc. Natl. Acad. Sci. U. S. A.*, 120(22), e2207355120.
- [113] Couce, A., Limdi, A., Magnan, M., Owen, S. V., Herren, C. M., Lenski, R. E., Tenaillon, O., & Baym, M. (2024). Changing fitness effects of mutations through long-term bacterial evolution. *Science*, 383(6681), eadd1417.
- [114] Johnson, M. S., & Desai, M. M. (2022). Mutational robustness changes during long-term adaptation in laboratory budding yeast populations. *Elife*, 11, e76491.
- [115] MacLean, R. C., Hall, A. R., Perron, G. G., & Buckling, A. (2010). The population genetics of antibiotic resistance: Integrating molecular mechanisms and treatment contexts. *Nature Reviews Genetics*, 11(6), 405–414.
- [116] Kryazhimskiy, S., Rice, D. P., Jerison, E. R., & Desai, M. M. (2014). Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science*, 344(6191), 1519–1522.
- [117] Schoustra, S., Hwang, S., Krug, J., & de Visser, J. A. G. M. (2016). Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. *Proc. Biol. Sci.*, 283(1837).

- [118] Wünsche, A., Dinh, D. M., Satterwhite, R. S., Arenas, C. D., Stoebel, D. M., & Cooper, T. F. (2017). Diminishing-returns epistasis decreases adaptability along an evolutionary trajectory. *Nat Ecol Evol*, 1(4), 61.
- [119] Johnson, M. S., Martsul, A., Kryazhimskiy, S., & Desai, M. M. (2019). Higher-fitness yeast genotypes are less robust to deleterious mutations. *Science*, 366(6464), 490–493.
- [120] Wei, X., & Zhang, J. (2019). Patterns and mechanisms of diminishing returns from beneficial mutations. *Mol. Biol. Evol.*, 36(5), 1008–1021.
- [121] Lukačšínová, M., Fernando, B., & Bollenbach, T. (2020). Highly parallel lab evolution reveals that epistasis can curb the evolution of antibiotic resistance. *Nat. Commun.*, 11(1), 3105.
- [122] Persson, K., Stenberg, S., Tamás, M. J., & Warringer, J. (2022). Adaptation of the yeast gene knockout collection is near-perfectly predicted by fitness and diminishing return epistasis. *G3*, 12(11).
- [123] Diaz-Colunga, J., Skwara, A., Gowda, K., Diaz-Uriarte, R., Tikhonov, M., Bajic, D., & Sanchez, A. (2023b). Global epistasis on fitness landscapes. *Philosophical Transactions of the Royal Society B*, 378(1877), 20220053.
- [124] Reddy, G., & Desai, M. M. (2021). Global epistasis emerges from a generic model of a complex trait. *Elife*, 10.
- [125] Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V., & Kruglyak, L. (2013). Finding the sources of missing heritability in a yeast cross. *Nature*, 494(7436), 234–237.
- [126] Mira, N. P., Teixeira, M. C., & Sá-Correia, I. (2010). Adaptive response and tolerance to weak acids in *Saccharomyces cerevisiae*: A genome-wide view. *OMICS*, 14(5), 525–540.
- [127] Ariño, J. (2010). Integrative responses to high ph stress in *S. cerevisiae*. *OMICS*, 14(5), 517–523.
- [128] Morano, K. A., Grant, C. M., & Moye-Rowley, W. S. (2012). The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics*, 190(4), 1157–1195.

- [129] Costanzo, M., Hou, J., Messier, V., Nelson, J., Rahman, M., VanderSluis, B., Wang, W., Pons, C., Ross, C., Ušaj, M., San Luis, B.-J., Shuteriqi, E., Koch, E. N., Aloy, P., Myers, C. L., Boone, C., & Andrews, B. (2021). Environmental robustness of the global yeast genetic interaction network. *Science*, 372(6542).
- [130] La Rosa, R., Rossi, E., Feist, A. M., Johansen, H. K., & Molin, S. (2021). Compensatory evolution of *Pseudomonas aeruginosa*'s slow growth phenotype suggests mechanisms of adaptation in cystic fibrosis. *Nat. Commun.*, 12(1), 1–15.
- [131] Kryazhimskiy, S. (2021). Emergence and propagation of epistasis in metabolic networks. *Elife*, 10, e60200.
- [132] Jerison, E. R., Kryazhimskiy, S., Mitchell, J. K., Bloom, J. S., Kruglyak, L., & Desai, M. M. (2017). Genetic variation in adaptability and pleiotropy in budding yeast. *eLife*, 6, e27167.
- [133] Wrolstad, R., Acerr, T., Decker, E., Penner, M., Reid, D., Schwartz, S., Shoemaker, C., Smith, D., & Sporns, P. (2001). Common buffers and stock solutions. *Current Protocols in Food Analytical Chemistry*, John Wiley & Sons, Inc.
- [134] Gomori, G. (2010). Preparation of buffers for use in enzyme studies. *Biochemistry and Molecular Biology*, 719.
- [135] Venkataram, S., & Kinsler, G. (2021). Barcodecounter2.
- [136] Berger, D., & Postma, E. (2014). Biased estimates of diminishing-returns epistasis? empirical evidence revisited. *Genetics*, 198(4), 1417–1420.
- [137] Vasi, F., Travisano, M., & Lenski, R. E. (1994). Long-term experimental evolution in escherichia coli. ii. changes in life-history traits during adaptation to a seasonal environment. *The american naturalist*, 144(3), 432–456.
- [138] Li, Y., Venkataram, S., Agarwala, A., Dunn, B., Petrov, D. A., Sherlock, G., & Fisher, D. S. (2018). Hidden complexity of yeast adaptation under simple evolutionary conditions. *Current Biology*, 28(4), 515–525.

- [139] Atolia, E., Cesar, S., Arjes, H. A., Rajendram, M., Shi, H., Knapp, B. D., Khare, S., Aranda-Díaz, A., Lenski, R. E., & Huang, K. C. (2020). Environmental and physiological factors affecting high-throughput measurements of bacterial growth. *MBio*, *11*(5), 10–1128.
- [140] Ogbunugafor, C. B., Wylie, C. S., Diakite, I., Weinreich, D. M., & Hartl, D. L. (2016). Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance. *PLoS computational biology*, *12*(1), e1004710.
- [141] Couce, A., & Tenaillon, O. A. (2015). The rule of declining adaptability in microbial evolution experiments. *Front. Genet.*, *6*, 99.
- [142] Draghi, J. A., & Plotkin, J. B. (2013). Selection biases the prevalence and type of epistasis along adaptive trajectories. *Evolution*, *67*(11), 3120–3131.
- [143] Woods, R. J., Barrick, J. E., Cooper, T. F., Shrestha, U., Kauth, M. R., & Lenski, R. E. (2011). Second-order selection for evolvability in a large escherichia coli population. *Science*, *331*(6023), 1433–1436.
- [144] Hausser, J., & Alon, U. (2020). Tumour heterogeneity and the evolutionary trade-offs of cancer. *Nat. Rev. Cancer*, *20*(4), 247–257.
- [145] Walworth, N. G., Zakem, E. J., Dunne, J. P., Collins, S., & Levine, N. M. (2020). Microbial evolutionary strategies in a dynamic ocean. *Proc. Natl. Acad. Sci. U. S. A.*, *117*(11), 5943–5948.
- [146] Salamin, N., Wüest, R. O., Lavergne, S., Thuiller, W., & Pearman, P. B. (2010). Assessing rapid evolution in a changing environment. *Trends Ecol. Evol.*, *25*(12), 692–698.
- [147] Willi, Y., & Van Buskirk, J. (2022). A review on trade-offs at the warm and cold ends of geographical distributions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *377*(1848), 20210022.
- [148] Futuyma, D. J., & Moreno, G. (1988). THE EVOLUTION OF ECOLOGICAL SPECIALIZATION. *Annu. Rev. Ecol. Evol. Syst.*, *19*(Volume 19, 1988), 207–233.

- [149] Van Tienderen, P. H. (1991). EVOLUTION OF GENERALISTS AND SPECIALISTS IN SPATIALLY HETEROGENEOUS ENVIRONMENTS. *Evolution*, 45(6), 1317–1331.
- [150] Kassen, R. (2002). The experimental evolution of specialists, generalists, and the maintenance of diversity. *J. Evol. Biol.*, 15(2), 173–190.
- [151] Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 31(1), 343–366.
- [152] Hedrick, P. W. (1986). GENETIC POLYMORPHISM IN HETEROGENEOUS ENVIRONMENTS: A DECADE LATER. *Annu. Rev. Ecol. Evol. Syst.*, 17(Volume 17, ), 535–566.
- [153] Felsenstein, J. (1976). THE THEORETICAL POPULATION GENETICS OF VARIABLE SELECTION AND MIGRATION. *Annu. Rev. Genet.*, 10(Volume 10, 1976), 253–280.
- [154] Werner, E. E., & Anholt, B. R. (1993). Ecological consequences of the Trade-Off between growth and mortality rates mediated by foraging activity. *Am. Nat.*
- [155] Tilman, D. (1990). Constraints and tradeoffs: Toward a predictive theory of competition and succession. *Oikos.*, 58(1), 3–15.
- [156] Vincent, T. L., Scheel, D., Brown, J. S., & Vincent, T. L. (1996). Trade-offs and coexistence in consumer-resource models: It all depends on what and where you eat. *The American Naturalist*, 148(6), 1038–1058.
- [157] Stearns, S. C. (1989). Trade-offs in life-history evolution. *Functional ecology*, 3(3), 259–268.
- [158] Zhu, M., & Dai, X. (2024). Shaping of microbial phenotypes by trade-offs. *Nat. Commun.*, 15(1), 4238.
- [159] Andersson, D. I., & Hughes, D. (2010). Antibiotic resistance and its cost: Is it possible to reverse resistance? *Nature Reviews Microbiology*, 8(4), 260–271.
- [160] Basan, M., Honda, T., Christodoulou, D., Hörl, M., Chang, Y.-F., Leoncini, E., Mukherjee, A., Okano, H., Taylor, B. R., Silverman, J. M., Sanchez, C., Williamson, J. R., Paulsson,

- J., Hwa, T., & Sauer, U. (2020). A universal trade-off between growth and lag in fluctuating environments. *Nature*, 584(7821), 470–474.
- [161] Venturelli, O. S., Zuleta, I., Murray, R. M., & El-Samad, H. (2015). Population diversification in a yeast metabolic program promotes anticipation of environmental shifts. *PLoS Biol.*, 13(1), e1002042.
- [162] Wang, J., Atolia, E., Hua, B., Savir, Y., Escalante-Chong, R., & Springer, M. (2015). Natural variation in preparation for nutrient depletion reveals a cost-benefit tradeoff. *PLoS Biol.*, 13(1), e1002041.
- [163] Bennett, A. F., & Lenski, R. E. (2007). An experimental test of evolutionary trade-offs during temperature adaptation. *Proc. Natl. Acad. Sci. U. S. A.*, 104 Suppl 1(Suppl 1), 8649–8654.
- [164] Normark, B. H., & Normark, S. (2002). Evolution and spread of antibiotic resistance. *J. Intern. Med.*, 252(2), 91–106.
- [165] Meyer, J. R., Dobias, D. T., Weitz, J. S., Barrick, J. E., Quick, R. T., & Lenski, R. E. (2012). Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science*, 335(6067), 428–432.
- [166] Hall, A. R., Scanlan, P. D., & Buckling, A. (2011). Bacteria-phage coevolution and the emergence of generalist pathogens. *Am. Nat.*, 177(1), 44–53.
- [167] Meyer, J. R., Agrawal, A. A., Quick, R. T., Dobias, D. T., Schneider, D., & Lenski, R. E. (2010). Parallel changes in host resistance to viral infection during 45,000 generations of relaxed selection. *Evolution*, 64(10), 3024–3034.
- [168] Remold, S. (2012). Understanding specialism when the jack of all trades can be the master of all. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), 4861–4869.
- [169] Meyer, J. R., Dobias, D. T., Medina, S. J., Servilio, L., Gupta, A., & Lenski, R. E. (2016). Ecological speciation of bacteriophage lambda in allopatry and sympatry. *Science*, 354(6317), 1301–1304.

- [170] Kinsler, G., Geiler-Samerotte, K., & Petrov, D. A. (2020). Fitness variation across subtle environmental perturbations reveals local modularity and global pleiotropy of adaptation. *eLife*, 9, e61271.
- [171] Schick, A., Bailey, S. F., & Kassen, R. (2015). Evolution of fitness Trade-Offs in locally adapted populations of *pseudomonas fluorescens*. *Am. Nat.*, 186 Suppl 1, S48–59.
- [172] Satterwhite, R. S., & Cooper, T. F. (2015). Constraints on adaptation of *escherichia coli* to mixed-resource environments increase over time. *Evolution*, 69(8), 2067–2078.
- [173] Doud, M. B., Gupta, A., Li, V., Medina, S. J., De La Fuente, C. A., & Meyer, J. R. (2024). Competition-driven eco-evolutionary feedback reshapes bacteriophage lambda’s fitness landscape and enables speciation. *Nat. Commun.*, 15(1), 863.
- [174] Spor, A., Kvitek, D. J., Nidelet, T., Martin, J., Legrand, J., Dillmann, C., Bourgaïs, A., de Vienne, D., Sherlock, G., & Sicard, D. (2014). Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. *Evolution*, 68(3), 772–790.
- [175] Wang, Y. C., & Lipsitch, M. (2006). Upgrading antibiotic use within a class: Tradeoff between resistance and treatment success. *Proc. Natl. Acad. Sci. U. S. A.*, 103(25), 9655–9660.
- [176] Greulich, P., Scott, M., Evans, M. R., & Allen, R. J. (2015). Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics. *Mol. Syst. Biol.*, 11(3), 796.
- [177] Lenski, R. E. (1988). EXPERIMENTAL STUDIES OF PLEIOTROPY AND EPISTASIS IN *ESCHERICHIA COLI*. II. COMPENSATION FOR MALADAPTIVE EFFECTS ASSOCIATED WITH RESISTANCE TO VIRUS T4. *Evolution*, 42(3), 433–440.
- [178] Bono, L. M., Gensel, C. L., Pfennig, D. W., & Burch, C. L. (2015). Evolutionary rescue and the coexistence of generalist and specialist competitors: An experimental test. *Proc. Biol. Sci.*, 282(1821).



- [179] Ardell, S., Martsul, A., Johnson, M. S., & Kryazhimskiy, S. (2024). Environment-independent distribution of mutational effects emerges from microscopic epistasis. *bioRxiv*.