Title

Characterization of Epistasis Across Different Environments in Yeast Saccharomyces cerevisiae Using RB-TnSeq Method

Author

Martsul, Alena

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Characterization of Epistasis Across Different Environments in Yeast *Saccharomyces cerevisiae* Using RB-TnSeq Method

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

in

Biology

by

Alena Martsul

Committee in charge:

Professor Sergey Kryazhimskiy, Chair
Professor Vineet Bafna
Professor Lin Chao
Professor Rachel Dutton
Professor Scott Rifkin

2021

The dissertation of Alena Martsul is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

To my Dad, Mom, and Brother.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Sergey Kryazhimskiy, for financial support that made this dissertation possible.

I would like to thank the members of my committee, Professor Lin Chao, Professor Rachel Dutton, Professor Vineet Bafna, Professor Scott Rifkin, for their constructive feedback and comments.

I also would like to thank all the current members of the Kryazhimskiy lab for valuable comments and advice during lab meetings. In particular, I wish to thank Sandeep Venkataram for fruitful scientific discussions and advice on the technical aspects of my project. I also would like to thank Shohreh Sikaroodi for teaching me some of the molecular biology techniques and helping me with my experiments during the first three years of my studies. I also thank Georg Rieckh for being a good friend and supporting me in different situations.

I must also thank Milo Johnson and Michael Desai from Harvard for providing transformed yeast strains and plasmid libraries that were used in my study, for valuable advice on the development of the RB-TnSeq method, and for the feedback on my projects.

The material of this dissertation is based on the following publications.

Chapter 2, in part, is a reprint of Milo Johnson, Alena Martsul, Kryazhimskiy Sergey, Desai Michael, “Higher-fitness yeast genotypes are less robust to deleterious mutations,” *Science*, 2019. The dissertation author was the primary investigator of the material described in chapter 2 and one of the authors of this paper.

Chapter 3, in part, currently being prepared for submission for publication of the material. Alena Martsul, Milo Johnson, and Sergey Kryazhimskiy. The dissertation author was the primary investigator and the author of this material.

VITA

2013	Bachelor of Science, Belarusian State Technological University
2014	Master of Science, Belarusian State Technological University and Norwegian University of Life Sciences
2014-2016	Research Assistant, Belarusian State Technological University
2021	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Milo S. Johnson, Alena Martsul, Sergey Kryazhimskiy, Michael M. Desai, “Higher-Fitness Yeast Genotypes Are Less Robust to Deleterious Mutations,” *Science*, Volume 366, Issue 6464, 2019.

FIELDS OF STUDY

Major Field: Biology

Studies in Ecology, Behavior and Evolution
Professor Sergey Kryazhimskiy

ABSTRACT OF THE DISSERTATION

Characterization of Epistasis Across Different Environments in Yeast *Saccharomyces cerevisiae* using RB-TnSeq Method

by

Alena Martsul

Doctor of Philosophy in Biology

University of California San Diego, 2021

Professor Sergey Kryazhimskiy, Chair

Rapid microbial evolution leads to the emergence of multiple drug resistance and transmission of infectious disease among distinct species. Therefore, to successfully design efficient antibiotic treatment and prevent the spread of infections, it is essential to understand the evolution process. In particular, we must be able to predict the effects of mutations that can arise in the adapting populations.

The fitness effects of many mutations are context-dependent. Interactions between mutations, or epistasis, make the effects of mutations dependent on their genetic background, while gene-environment interactions modulate the effects of mutations in different environments. Thus, knowing the fitness effect of a mutation alone in one genotype and environment is not enough to predict its effect in different conditions.

Recent advances in molecular biology, genetics, and sequencing enabled studying the fitness effects of many mutations across a wide range of different genotypes. While the effects of some mutations are idiosyncratic, the effects of other mutations show a consistent pattern: they become less positive or more negative in genetic backgrounds with higher fitness. The consistency of epistasis for some mutations at least partially resolves the problem of the predictability of the effects of mutations in different genotypes. However, it remains unclear how epistasis changes in different environments.

In this thesis, we use directed mutagenesis and short-term experimental evolution to characterize epistasis in yeast *Saccharomyces cerevisiae* in different environments. Chapter 1 reviews recent findings regarding epistatic and gene-environment interactions. Chapter 2 describes the development of the RB-TnSeq method to generate and measure the effects of many identical mutations in different yeast *S. cerevisiae* genotypes. This chapter also presents a custom bioinformatics analysis pipeline designed to infer the effects of mutations from the next-generation sequencing data. Chapter 3 reports the application of the RB-TnSeq method to characterize epistasis for 100 random mutations across 42 yeast *S. cerevisiae* genotypes in six environments, showing that the negative relationship between the effects of mutations and fitness of their genetic background remains largely consistent across environments. Finally, this chapter discusses our findings in the context of the existing literature in this field and describes the potential future research direction.

Chapter 1 Literature Review

1.1 Introduction

The recent transmission of SARS-CoV-2 virus variants to humans, their rapid change and spread [1–3], as well as the emergence of multiple drug resistance among microbes [4,5] clearly demonstrate that rapid microbial evolution can have a significant impact on different areas of human life [6]. To successfully respond to its consequences and minimize the practices that facilitate the emergence and evolution of undesired microbial traits, it is essential to understand the principles of microbial evolution and to be able to predict the available adaptive trajectories for evolving populations. However, this task is rather challenging. It requires prior knowledge of a sequence of adaptive mutations for an evolving genotype in the set of environmental conditions it encounters.

Laboratory evolution experiments are widely applied to study microbial adaptation [7–9]. Laboratory studies revealed that adaptation represents a multi-step process of monotonic fitness increase [10]. At each step, evolving populations generally have multiple adaptive genetic variants available, and variants that confer a higher fitness advantage have a higher chance to reach fixation [11–15]. Recent methodological advances enabled tracking the fate of genetic variants present in evolving populations and measuring their fitness [11,12,16]. When the selection is well defined, these data allow to link the genotype, phenotype and fitness, to understand the molecular basis of adaptation, and to evaluate the repeatability of evolution at genotype and phenotype levels [8].

However, the outcomes of laboratory evolutionary experiments cannot be directly extended to populations growing in natural conditions. The laboratory studies are typically done with genetically homogeneous populations of model organisms that are maintained in defined and

carefully controlled environments [8,17], while naturally occurring populations differ from the laboratory strains, can harbor standing genetic variation, and are frequently exposed to fluctuations in different environmental parameters [18]. The discrepancies between natural and laboratory conditions as well as the fact that the majority of microbes has not yet been cultivated in the lab [19] demonstrate the need to get a general knowledge of how the effects of mutations depend on their genetic background and environment to be able to predict the evolution of any genotype in any environment.

In the following sections, we review the literature on the dependence of the effects of mutations on their genetic background and on the environment, focus on the studies on how genetic interactions change adaptive landscapes of genotypes and describe the general patterns observed for the effects of mutations across different genotypes. At the end of this chapter, we identify the existing scientific gap on predicting the phenotypic effects of mutations in different genotypes and environments and propose the experimental approach to bridge that.

1.2 Epistasis

1.2.1 Definition and Classification

Epistasis, i.e., genetic interactions between mutations, makes effects of mutations dependent on their genetic background. It is identified by the deviation of the observed fitness of double mutant from the expected value predicted based on the fitness effects of corresponding single mutations. Multiplicative and additive models can be applied to estimate the strength of epistasis [20–23]. The experimental data show that epistatic interactions can alter the magnitude [4,5,24–27] and the sign [5,28–30] of the fitness effects of mutations.

The existing literature applies various terminology to describe epistatic interactions. Synergistic (antagonistic) epistasis is observed when the combined effect of two mutations is

more (less) severe than expected. Negative epistasis between beneficial mutations corresponds to antagonism, while negative epistasis between deleterious mutations is classified as synergism [31]. Sign epistasis is observed when the sign of the fitness effect of a mutation changes depending on the presence of the other alleles in the same genetic background [29,30,32]. Epistatic interactions can occur between the genetic variants within the same gene, i.e., intragenic epistasis, and between different genes, i.e., intergenic epistasis [31]. Epistatic interactions among specific sets of mutations are referred to as microscopic epistasis, while the dependence of the entire distribution of fitness effects of all mutations on its genetic background describes macroscopic epistasis [33,34].

1.2.2 Epistatic Interactions and Genetic Context of Mutations

In this section, we review the experimental findings on epistatic interactions for mutations that were identified in various evolutionary experiments to demonstrate that these interactions are not mutation-specific and are largely dependent on the genetic context of a mutation.

Epistatic interactions for *rpoB* mutations in the RNA polymerase subunit *β* that are usually associated with rifampicin resistance [35,36] have been identified in multiple experiments. In the evolutionary experiment with *Pseudomonas aeruginosa*, magnitude epistasis between *rpoB* mutation and subsequent mutations that arose during evolution restricted lineages harboring this mutation from gaining high fitness compared to genotypes without this mutation [28]. As drug resistance mutations are usually associated with fitness cost [37,38], *rpoB* mutations have been shown to potentiate high effect beneficial mutations to compensate for their cost in *Escherichia coli* [39]. Some *rpoB* alleles have been identified to show either positive or negative epistasis patterns depending on the presence of specific alleles conferring resistance to another drug [4]. The presence of *rpoB* mutation in *E. coli* lineages adapted to growth at high temperature has been associated with a shift in a thermal niche - a characteristic trade-off of one of the alternative

adaptations to that environment [40,41]. Sign epistasis between *rpoB* and *rho* mutations, which are characteristic mutations of two alternative adaptive paths, restricts the presence of both mutations in the same genotype; thus, makes these adaptive pathways mutually exclusive [40]. The evidence of sign epistasis have been identified for other groups of mutations [29].

These data indicate that knowing a mutation itself is not enough to predict its effect on fitness in any genotype. Since epistatic interactions, which modulate the fitness effects of mutations, are dependent on the genetic context, it might be important to identify the underlying molecular mechanisms that cause epistasis to be able to predict the effects of mutations.

1.2.3 Mechanisms That Cause Epistasis

The authors of several studies proposed that effects of mutations act additively on the underlying traits, but the non-linearity of mapping of these traits to larger-scale phenotypes, i.e., fitness, manifests as epistasis [42–44]. We describe the experimental evidence supporting this idea below.

For single molecules, i.e., proteins and RNA molecules, such underlying traits can be the stability of molecules [42,45–48] or the conformation and functionality of their folding structures [48]. The intragenic epistasis can be explained by the exhaustion of the stability threshold by the molecules. After that happens, disruptive effects of mutations tend to become weaker because the functionality of the molecules has already been impaired, which explains positive epistasis between disruptive mutations [49]. Although rare, positive epistasis can also stem from a stabilizing effect of the interactions between mutations that have destabilizing effects when considered individually. In such case, the structure of molecules with multiple mutations is more stable due to the conformational switch in the structure of the molecule that eliminates the unfavorable effects of individual mutations [26,49]. However, before molecules exceed their

stability threshold, the negative epistasis is prevalent among random mutations because the multiple mutations have a stronger destabilizing effect than individual mutations [48].

Enzyme activity [50], affinity of proteins to a target ligand [51], and the performance of metabolic networks [52] are the examples of other traits that have saturation thresholds. In addition, the analysis of metabolic networks provides an insight into the underlying causes of epistasis at the intergenic level. For instance, flux balance analysis of metabolic networks in *E. coli* and *Saccharomyces cerevisiae* showed that negative epistasis frequently arises between genes that perform overlapping functions [24]. Although not common between random disruptive mutations, negative epistasis can occur either between disruptions of nonessential reactions or between mutations in one essential and other nonessential reactions that share a nonessential function [24,53]. No epistasis can be expected between mutations that affect non-related metabolic reactions [24]. On the other hand, positive epistasis is observed for the majority of the mutation pairs which affect both essential and nonessential reactions [24]. When the rates of essential reactions are constrained, the effects of subsequent mutations become negligible because all other reactions in the pathway do not work to their full capacity [54]. In addition, positive epistasis between nonessential reactions can be explained by the availability of alternative reactions that can compensate for the loss of the product from the affected metabolic steps [24,31,53,54].

1.2.4 Predicting the Effects of Mutations Is Challenging

In sum, the above examples demonstrate that effects of mutations are highly dependent on their genetic background and that different molecular mechanisms explain the epistasis observed for the sets of mutations.

Although the knowledge of the underlying molecular mechanisms allows predicting the effects of mutations, the application of such approach is not feasible for any mutation in any

genetic context. First, the traits that are under selection (and molecular mechanisms associated with them) can be obvious in laboratory studies, but it can be difficult to identify them for populations growing in natural conditions as they may be subject to different types of selection acting simultaneously. Second, the identification of traits on which mutations act additively does not provide information on how these traits are mapped to fitness.

The latter question has been a focus of some studies. Specifically, Chou et al. developed a cost-benefit model to explain and predict the effects of combinations of adaptive mutations in *Methylobacterium extorquens* [43,44] and Otwinowski et al. used the data from deep mutational scanning experiments done with different proteins to infer the shape of a nonlinear function that relates the additive trait to fitness [42]. While these studies provide a description of how mutations act on the underlying traits and how these traits are translated to fitness, their findings are limited to the systems that were used in their studies and cannot be applied to predict the effects of mutations in any genetic background.

These arguments suggest that that predicting the effects of mutations based on the knowledge of the underlying molecular traits is currently not feasible and demonstrate the need to find the approach that uses characteristics of mutations and genotypes that are relatively easy to obtain.

1.3 Mutations Have Different Effects in Different Environments

The interactions between genes and environments (GxE) modify the effects of mutations in addition to the epistatic interactions [55–60]. In this section, we review the mechanisms describing how the GxE interactions can change the effects of mutations and discuss the ways to predict the effects of mutations in different environments.

1.3.1 Description of Different Types of Gene-Environment Interactions

Environments can modulate the magnitude of fitness effects of mutations by altering the rates of various biochemical and biophysical processes inside the cell [61,62]. Chu et al. showed that at higher, but relatively benign temperatures many beneficial mutations have stronger effects on fitness in *E. coli* [62]. The study by Kishony and Leibler suggested that environments with strong impacts on cell physiology, e.g., low temperature, tend to alleviate the effects of deleterious mutations [63]. A hypothesis linking the cell growth to the rate of the most limiting metabolic reaction can explain these observations [64]. According to it, environments and mutations impact different reactions, the resulting growth rate of a mutant in a given environment should highly depend on the rate of the slowest impacted reaction [63,64]. Thus, in the study by Chu et al., warmer temperatures could have eliminated the growth constraints imposed by low-temperature environments which resulted in mutations having larger effects on fitness in those environments [62]. In the study by Kishony and Leibler, the limiting impact of the external environment was much stronger than that of the deleterious mutations. Thus, cellular growth was mostly limited by environmental stress, and the effects of mutations were weaker compared to environments with no stress [63]. In addition, stressful environments can induce the expression of genes relative to their expression in benign conditions [61,65]. Therefore, environmental changes can either reveal the disruptive effects of mutations in previously silent genes or make their effects stronger [57].

Different environments select for different phenotypes. Therefore, some mutations with beneficial or deleterious effects in one environment can have the opposite effects in other environments. The recent study by Chen and Zhang shows that adaptive mutations that were identified in the evolutionary experiment in one environment tend to have comparable effects on fitness in similar environments, while in antagonistic environments these mutations tend to have opposite effects [66]. The evidence of antagonistic pleiotropy has also been reported in other studies [67–70]. However, only 28 % of natural genetic polymorphisms [67] and mutations in

about 14% of nonessential genes [68] in *S. cerevisiae* have been identified to be subject to antagonistic pleiotropy. In comparison, about 58 % of natural genetic polymorphisms in yeast have been reported to exhibit GxE interactions [67], which suggests that the effects of mutations change differently with respect to their environment.

Environment-related changes in the gene interaction networks at least partially explain why GxE interactions differ between individual mutations and why effects of mutations depend on the environment. Although the study by Costanzo et al. demonstrates that gene interaction networks are robust to environmental changes [71], the authors of several other studies report the opposite observations. In particular, Bandyopadhyay et al. studied changes in the gene interaction network in response to the DNA damage. They show that pairs of genes that change their interaction profiles between environments are usually involved in cellular processes modified upon the transition from one environment to another [72]. Such genes gain new functions and change the magnitude of the interactions with other genes upon the shift in the environment [73], which makes it clear why the effects of mutations in such genes are environment-dependent. However, the proportion of the genes that show differential epistatic interactions across environments is low [71], probably, because the functions of the majority of the genes are associated with “housekeeping” in cells, e.g., chromatin organization, [72] and remain important and unchanged regardless of the environment type.

1.3.2 Predicting the Effects of Mutations in Different Environments

As there is a variety of the molecular mechanisms that explain epistasis for mutations, there are different types of gene-environment interactions that describe the changes in the fitness effects of mutations across environments. This suggests that predicting the effects of mutations by considering epistatic and GxE interactions individually is unrealistic for most mutations. Therefore, looking for the universal trends on how epistatic interactions change across

environments can be promising, as this approach allows to simultaneously study the dependence of the effects of mutations on both parameters. In the following sections, we review the current literature on epistasis patterns and propose the idea of how studying epistasis across environments can help to predict the phenotypic effects of mutations.

1.4 Epistasis and Adaptive Landscapes

The adaptive landscape represents a multidimensional map that associates genotypes with their fitness levels. Studying adaptive landscapes provides an understanding of which of the available adaptive trajectories lead to a fitness optimum for a given genotype in the specified set of environmental conditions. In addition, it shows how epistatic interactions between mutations may restrict or favor various adaptive trajectories for evolving populations and how they shape the fitness landscape.

There are two main approaches to study how epistasis shapes adaptive landscapes. The first approach focuses on mutational neighborhood of a particular genotype and involves generating large sets of genotypes that represent single-, double-, and sometimes higher-order mutants of the focal genotype. This approach is generally applied to study local fitness landscapes of proteins of the RNA molecules, to identify the prevalence of the specific type of epistasis in the fitness landscape and to characterize the causes of such epistatic interactions. In particular, it was applied to describe the local fitness landscape of the tRNA gene in *S. cerevisiae* [47,48] and of green fluorescence protein from *Aequorea victoria* [46], to characterize the fitness map of mutations in the immunoglobulin binding fragment (Ig-G) of protein G [49], to study the prevalent type of epistasis among mutations in loci commonly associated with resistance to HIV-1 protease inhibitors [26]. Interestingly, these studies report quite opposite observations for different molecules. Specifically, the studies by Li et al. show the prevalence of the negative epistasis in the fitness landscape of the tRNA gene and explain it by the lower functionality of tRNA molecules

with multiple mutations [48]. On the other hand, Zhang et al. demonstrate that positive epistasis is prevalent in the fitness landscape of the HIV-1 protease and associate that with the possibility of physical interactions between the mutated sites [26]. Thus, these observations show that the prevalent pattern of epistasis depends on the type of molecule studied and generally cannot be extended to the fitness landscapes of other, substantially different, molecules.

The second approach focuses on the interactions between mutations that appeared at different adaptive steps and studies how epistasis between them changes the adaptive trajectories for evolving genotypes. This approach is based on constructing and analyzing mutants that represent various combinations of small sets of pre-selected mutations that were previously identified in evolutionary experiments. For example, Khan et al. studied the nature of epistasis among the first five beneficial mutations that fixed in the long-term evolutionary experiment with *E.coli* (the experiment is described in [39,74]) and whose spread coincided with the period of the fastest adaptation [27]. They reconstructed 2^5 genotypes and tested how fitness effect of each of the five mutations depend on its genetic background. The same genotypes were used to compare the topology of the fitness landscapes in two environments imposing different demands for respiration [75] and in additional eight environments [76]. Four beneficial mutations identified in the evolutionary experiment with *Methylobacterium extorquens* that grew on methanol as a sole carbon source were reconstructed in 120 possible combinations to study the consequences of epistasis on adaptation [43,44]. The same number of mutational trajectories consisting of five mutations which collectively confer a resistance to *β*-lactam antibiotics were evaluated for evolutionary accessibility in *E. coli* [30]. Another study reconstructed fitness landscapes of three inverse *lac* repressor variants and used them to analyze high-order interactions between genetic and environmental changes [77]. The common observation of the above studies is that epistatic interactions shape the fitness landscape, and these interactions depend on the environment. However, these studies do not provide any general patterns describing how epistatic interactions change across environments.

1.5 Emerging Patterns for the Effects of Mutations

Diminishing-returns epistasis (DRE) is a phenomenon that describes an inverse dependence of effects of beneficial mutations on the fitness of their genetic background. The negative linear relationship between the effects of mutations and their background fitness has been observed in evolution experiments done with various model organisms, for example, with bacteria *E.coli* [39,78–80] and *P. aeruginosa* [28,81], with yeast *S. cerevisiae* [82,83], with bacteriophages [84] and multicellular fungus [85]. This demonstrates that epistatic interactions for beneficial mutations conform to a general pattern and suggests that the effects of these mutations might be statistically predictable [86].

The dependence of the effects of some beneficial mutations on the background fitness was confirmed by directly generating and measuring the effects of these mutations in genetic backgrounds of various fitness [27,44,75,76,83]. This suggests that those mutations are subject to global epistasis, a phenomenon described in [83] which refers to the dependence of the effects of mutations on the fitness of their genetic background and not the genotype itself. A recent study by Wei and Zhang showed that a large fraction of mutations is subject to global epistasis, as between 66% to 92% of tested naturally occurring genetic polymorphisms exhibit diminishing-returns epistasis across 47 different environments [87]. Taken together, this suggests that global epistasis is widespread among beneficial mutations (but not all mutations exhibit it). Different types of mutations can be subject to it, and some mutations exhibit diminishing-returns epistasis only in some environmental conditions. These conclusions are relevant to only beneficial mutations.

Similar linear relationship has been observed for deleterious mutations. In our recent study, we found that random mutations tend to have stronger deleterious effects in more-fit genotypes, i.e., exhibit increasing cost (IC) epistasis, as the distributions of the fitness effects (DFE) of the same set of mutations had lower mean and higher variance in genotypes with higher

fitness [88]. Analogous observations were reported for deleterious mutations in the mutation accumulation experiments done with near absence of selection [89,90].

1.6 Scientific Gap and Approach to Solving It

The recently proposed connectedness model [91] and idiosyncratic epistasis theory [92] suggest that the observed patterns for beneficial and deleterious mutations both represent the same trend. That trend is a consequence of widespread epistasis among individual loci, and many mutations can be subject to it. In turn, this implies that it possible to predict the effect of a mutation given the fitness of its background genotype.

However, these models do not give any predictions about how epistasis changes in different environments. For example, it is unclear whether the same relationship between the effects of mutations and fitness of their background genotype holds in environments in which mutations switch the sign of their fitness effect from positive to negative. In other words, whether the fitness of the genetic background can be used as a predictor of the effects of mutations in any environment it is not known.

To bridge that gap, we propose focus on the effects of mutations in the single-step mutational neighborhood of genotypes. For that, we suggest generating and measuring the fitness effects of many identical mutations across genotypes of various fitness in environments that differ in the similarity to each other. Using this approach, we can first evaluate how the entire DFE of genotypes changes with respect to their fitness and can compare how these relationships change across different environments. In case these relationships are not affected by the environment, we can expect to observe consistent patterns even between environments that have no similarity to each other. If the difference is identified, we can use this data to examine the epistasis patterns for individual mutations and test how these patterns change in different environments and how they explain the relationships observed to the entire DFE.

1.7 Thesis Structure

In Chapter 2, we report the development of the RB-TnSeq method to generate and measure the fitness effects of many identical insertion mutations in different genotypes of yeast *S. cerevisiae*. The application of this method also allows to simultaneously estimate fitness of the selected genotypes. In addition, we describe the custom bioinformatics pipeline that was designed to analyze the next-generation sequencing data from the RB-TnSeq experiment.

In Chapter 3, we apply the RB-TnSeq method to insert and measure fitness effects of 100 identical mutations across 42 yeast *S. cerevisiae* segregant strains across six different environmental conditions. We compare the similarity of environments, demonstrate that patterns of macroscopic epistasis change across environments and examine what drives the observed differences at the level of DFE by comparing epistasis for individual mutations. Lastly, we discuss our findings in the context of earlier studies in this field and make conclusions.

1.8 References

1. Holmes EC, Goldstein SA, Rasmussen AL, Robertson DL, Crits-Christoph A, Wertheim JO, Anthony SJ, Barclay WS, Boni MF, Doherty PC, Farrar J, Georghegan JL,Weiss SR, Worobey M, Andersen KG, Garry RF, Rambaut A. The origins of SARS-CoV-2: A critical review. Cell. 2021;184: 4848–4856. doi:10.1016/j.cell.2021.08.017

2. Morens DM, Breman JG, Calisher CH, Doherty PC, Hahn BH, Keusch GT, Kramer LD, LeDuc JW, Monath TP, Taubenberger JK. The Origin of COVID-19 and Why It Matters. The American Journal of Tropical Medicine and Hygiene. 2020;103: 955–959. doi:10.4269/ajtmh.20-0849

3. Latinne A, Hu B, Olival KJ, Zhu G, Zhang L, Li H, Chmura AA, Field HE, Epstein JH, Li B, Zhang W, Shi ZL, Daszak P. Origin and cross-species transmission of bat coronaviruses in China. Nat Commun. 2020;11: 4235. doi:10.1038/s41467-020-17687-3

4. Trindade S, Sousa A, Xavier KB, Dionisio F, Ferreira MG, Gordo I. Positive Epistasis Drives the Acquisition of Multidrug Resistance. Zhang J, editor. PLoS Genet. 2009;5: e1000578. doi:10.1371/journal.pgen.1000578

5. Ward H, Perron GG, Maclean RC. The cost of multiple drug resistance in *Pseudomonas aeruginosa*. Journal of Evolutionary Biology. 2009;22: 997–1003. doi:10.1111/j.1420-9101.2009.01712.x

6. Chakraborty I, Maity P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. Science of The Total Environment. 2020;728: 138882. doi:10.1016/j.scitotenv.2020.138882

7. Adams J, Rosenzweig F. Experimental microbial evolution: history and conceptual underpinnings. Genomics. 2014;104: 393–398. doi:10.1016/j.ygeno.2014.10.004

8. Gresham D, Dunham MJ. The enduring utility of continuous culturing in experimental evolution. Genomics. 2014;104: 399–405. doi:10.1016/j.ygeno.2014.09.015

9. Winkler JD, Kao KC. Recent advances in the evolutionary engineering of industrial biocatalysts. Genomics. 2014;104: 406–411. doi:10.1016/j.ygeno.2014.09.006

10. Weinreich DM. The Rank Ordering of Genotypic Fitness Values Predicts Genetic Constraint on Natural Selection on Landscapes Lacking Sign Epistasis. Genetics. 2005;171: 1397–1405. doi:10.1534/genetics.104.036830

11. Kinnersley M, Schwartz K, Yang D-D, Sherlock G, Rosenzweig F. Evolutionary dynamics and structural consequences of de novo beneficial mutations and mutant lineages arising in a constant environment. BMC Biol. 2021;19: 20. doi:10.1186/s12915-021-00954-0

12. Cooper VS. Experimental Evolution as a High-Throughput Screen for Genetic Adaptations. Gales AC, editor. mSphere. 2018;3. doi:10.1128/mSphere.00121-18

13. Kvitek DJ, Sherlock G. Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment. Zhang J, editor. PLoS Genet. 2013;9: e1003972. doi:10.1371/journal.pgen.1003972

14. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, Desai MM. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013;500: 571–574. doi:10.1038/nature12344

15. Maddamsetti R, Lenski RE, Barrick JE. Adaptation, Clonal Interference, and Frequency-Dependent Interactions in a Long-Term Evolution Experiment with *Escherichia coli*. Genetics. 2015;200: 619–631. doi:10.1534/genetics.115.176677

16. Levy SF, Blundell JR, Venkataram S, Petrov DA, Fisher DS, Sherlock G. Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature. 2015;519: 181–186. doi:10.1038/nature14279

17. Lenski RE. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J. 2017;11: 2181–2194. doi:10.1038/ismej.2017.69

18. Koskella B, Vos M. Adaptation in Natural Microbial Populations. Annu Rev Ecol Evol Syst. 2015;46: 503–522. doi:10.1146/annurev-ecolsys-112414-054458

19. Schmeisser C, Steele H, Streit WR. Metagenomics, biotechnology with non-culturable microbes. Appl Microbiol Biotechnol. 2007;75: 955–962. doi:10.1007/s00253-007-0945-5

20. Wade MJ, Winther RG, Agrawal AF, Goodnight CJ. Alternative definitions of epistasis: dependence and interaction. Trends in Ecology & Evolution. 2001;16: 498–504. doi:10.1016/S0169-5347(01)02213-3

21. Mani R, St.Onge RP, Hartman JL, Giaever G, Roth FP. Defining genetic interaction. Proceedings of the National Academy of Sciences. 2008;105: 3461–3466. doi:10.1073/pnas.0712255105

22. Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Human Molecular Genetics. 2002;11: 2463–2468. doi:10.1093/hmg/11.20.2463

23. Gao H, Granka JM, Feldman MW. On the Classification of Epistatic Interactions. Genetics. 2010;184: 827–837. doi:10.1534/genetics.109.111120

24. He X, Qian W, Wang Z, Li Y, Zhang J. Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks. Nat Genet. 2010;42: 272–276. doi:10.1038/ng.524

25. Bonhoeffer S. Evidence for Positive Epistasis in HIV-1. Science. 2004;306: 1547–1550. doi:10.1126/science.1101786

26. Zhang T, Dai L, Barton JP, Du Y, Tan Y, Pang W, et al. Predominance of positive epistasis among drug resistance-associated mutations in HIV-1 protease. Zhang J, editor. PLoS Genet. 2020;16: e1009009. doi:10.1371/journal.pgen.1009009

27. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative Epistasis Between Beneficial Mutations in an Evolving Bacterial Population. Science. 2011;332: 1193–1196. doi:10.1126/science.1203801

28. Gifford DR, Toll-Riera M, MacLean RC. Epistatic interactions between ancestral genotype and beneficial mutations shape evolvability in *Pseudomonas aeruginosa*: BRIEF COMMUNICATION. Evolution. 2016;70: 1659–1666. doi:10.1111/evo.12958

29. Kvitek DJ, Sherlock G. Reciprocal Sign Epistasis between Frequently Experimentally Evolved Adaptive Mutations Causes a Rugged Fitness Landscape. Zhang J, editor. PLoS Genet. 2011;7: e1002056. doi:10.1371/journal.pgen.1002056

30. Weinreich DM. Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. Science. 2006;312: 111–114. doi:10.1126/science.1123539

31. de Visser JAGM, Cooper TF, Elena SF. The causes of epistasis. Proc R Soc B. 2011;278: 3617–3624. doi:10.1098/rspb.2011.1537

32. Bridgham JT. Evolution of Hormone-Receptor Complexity by Molecular Exploitation. Science. 2006;312: 97–101. doi:10.1126/science.1123348

33. Jerison ER, Desai MM. Genomic investigations of evolutionary dynamics and epistasis in microbial evolution experiments. Current Opinion in Genetics & Development. 2015;35: 33–39. doi:10.1016/j.gde.2015.08.008

34. Good BH, Desai MM. The Impact of Macroscopic Epistasis on Long-Term Evolutionary Dynamics. Genetics. 2015;199: 177–190. doi:10.1534/genetics.114.172460

35. Klein JL, Brown TJ, French GL. Rifampin Resistance in *Mycobacterium kansasii* Is Associated with *rpoB* Mutations. Antimicrob Agents Chemother. 2001;45: 3056–3058. doi:10.1128/AAC.45.11.3056-3058.2001

36. Shea J, Halse TA, Kohlerschmidt D, Lapierre P, Modestil HA, Kearns CH, Dworkin FF,Diekema DJ. Low-Level Rifampin Resistance and *rpoB* Mutations in Mycobacterium tuberculosis: an Analysis of Whole-Genome Sequencing and Drug Susceptibility Test Data in New York. Diekema DJ, editor. J Clin Microbiol. 2021;59. doi:10.1128/JCM.01885-20

37. Wi YM, Greenwood-Quaintance KE, Brinkman CL, Lee JYH, Howden BP, Patel R. Rifampicin resistance in Staphylococcus epidermidis: molecular characterisation and fitness cost of rpoB mutations. International Journal of Antimicrobial Agents. 2018;51: 670–677. doi:10.1016/j.ijantimicag.2017.12.019

38. Reynolds MG. Compensatory Evolution in Rifampin-Resistant *Escherichia coli*. Genetics. 2000;156: 1471–1481. doi:10.1093/genetics/156.4.1471

39. Barrick JE, Kauth MR, Strelioff CC, Lenski RE. Escherichia coli rpoB Mutants Have Increased Evolvability in Proportion to Their Fitness Defects. Molecular Biology and Evolution. 2010;27: 1338–1347. doi:10.1093/molbev/msq024

40. Rodriguez-Verdugo A, Carrillo-Cisneros D, Gonzalez-Gonzalez A, Gaut BS, Bennett AF. Different tradeoffs result from alternate genetic adaptations to a common environment.

Proceedings of the National Academy of Sciences. 2014;111: 12121–12126. doi:10.1073/pnas.1406886111

41. Rodríguez-Verdugo A, Gaut BS, Tenaillon O. Evolution of Escherichia coli rifampicin resistance in an antibiotic-free environment during thermal stress. BMC Evol Biol. 2013;13: 50. doi:10.1186/1471-2148-13-50

42. Otwinowski J, McCandlish DM, Plotkin JB. Inferring the shape of global epistasis. Proc Natl Acad Sci USA. 2018;115: E7550–E7558. doi:10.1073/pnas.1804015115

43. Chou H-H, Chiu H-C, Delaney NF, Segre D, Marx CJ. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. Science. 2011;332: 1190–1192. doi:10.1126/science.1203799

44. Chou H-H, Delaney NF, Draghi JA, Marx CJ. Mapping the Fitness Landscape of Gene Expression Uncovers the Cause of Antagonism and Sign Epistasis between Adaptive Mutations. Achtman M, editor. PLoS Genet. 2014;10: e1004149. doi:10.1371/journal.pgen.1004149

45. Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS. Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. Nature. 2006;444: 929–932. doi:10.1038/nature05385

46. Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, Memedov IZ, Tawfik DS, Lukyanov KA, Kondrashow FA. Local fitness landscape of the green fluorescent protein. Nature. 2016;533: 397–401. doi:10.1038/nature17995

47. Li C, Zhang J. Multi-environment fitness landscapes of a tRNA gene. Nat Ecol Evol. 2018;2: 1025–1032. doi:10.1038/s41559-018-0549-8

48. Li C, Qian W, Maclean CJ, Zhang J. The fitness landscape of a tRNA gene. Science. 2016;352: 837–840. doi:10.1126/science.aae0568

49. Olson CA, Wu NC, Sun R. A Comprehensive Biophysical Description of Pairwise Epistasis throughout an Entire Protein Domain. Current Biology. 2014;24: 2643–2651. doi:10.1016/j.cub.2014.09.072

50. Dean AM. Selection and neutrality in lactose operons of Escherichia coli. Genetics. 1989;123: 441.

51. de Visser JAGM, Krug J. Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet. 2014;15: 480–490. doi:10.1038/nrg3744

52. Moreno-Sánchez R, Saavedra E, Rodríguez-Enríquez S, Olín-Sandoval V. Metabolic Control Analysis: A Tool for Designing Strategies to Manipulate Metabolic Pathways. Journal of Biomedicine and Biotechnology. 2008;2008: 1–30. doi:10.1155/2008/597913

53. Segrè D, DeLuna A, Church GM, Kishony R. Modular epistasis in yeast metabolism. Nat Genet. 2005;37: 77–83. doi:10.1038/ng1489

54. Lehner B. Molecular mechanisms of epistasis within and between genes. Trends in Genetics. 2011;27: 323–331. doi:10.1016/j.tig.2011.05.007

55. Remold SK, Lenski RE. Contribution of individual random mutations to genotype-by-environment interactions in Escherichia coli. Proceedings of the National Academy of Sciences. 2001;98: 11388–11393. doi:10.1073/pnas.201140198

56. Korona R. Genetic load of the yeast *Sacchatomyces cerevisiae* under diverse environmental conditions. Evolution. 1999;53: 1966–1971. doi:10.1111/j.1558-5646.1999.tb04577.x

57. Szafraniec K, Borts RH, Korona R. Environmental stress and mutational load in diploid strains of the yeast Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences. 2001;98: 1107–1112. doi:10.1073/pnas.98.3.1107

58. Vassilieva LL, Hook AM, Lynch M. The fitness effects of spontaneous mutations id *Caenorhabditis elegans*. Evolution. 2000;54: 1234–1246. doi:10.1111/j.0014-3820.2000.tb00557.x

59. Bataillon T, Zhang T, Kassen R. Cost of Adaptation and Fitness Effects of Beneficial Mutations in *Pseudomonas fluorescens*. Genetics. 2011;189: 939–949. doi:10.1534/genetics.111.130468

60. Ketola T, Saarinen K. Experimental evolution in fluctuating environments: tolerance measurements at constant temperatures incorrectly predict the ability to tolerate fluctuating temperatures. J Evol Biol. 2015;28: 800–806. doi:10.1111/jeb.12606

61. Miller MJ, Xuong NH, Geiduschek EP. Quantitative analysis of the heat shock response of Saccharomyces cerevisiae. J Bacteriol. 1982;151: 311–327. doi:10.1128/jb.151.1.311-327.1982

62. Chu X, Zhang D, Buckling A, Zhang Q. Warmer temperatures enhance beneficial mutation effects. J Evol Biol. 2020;33: 1020–1027. doi:10.1111/jeb.13642

63. Kishony R, Leibler S. Environmental stresses can alleviate the average deleterious effect of mutations. J Biol. 2003;2: 14. doi:10.1186/1475-4924-2-14

64. Jasnos L, Tomala D, Paczesniak D, Korona R. Interactions between stressful environment and gene deletions alleviate the expected average loss of fitness in yeast. Genetics. 2008;178: 2105–11. doi:doi: 10.1534/genetics.107.084533.

65. Craig EA, Gambill BD, Nelson RJ. Heat shock proteins: molecular chaperones of protein biogenesis. Microbiol Rev. 1993;57: 402–414. doi:10.1128/mr.57.2.402-414.1993

66. Chen P, Zhang J. Antagonistic pleiotropy conceals molecular adaptations in changing environments. Nat Ecol Evol. 2020;4: 461–469. doi:10.1038/s41559-020-1107-8

67. Wei X, Zhang J. The Genomic Architecture of Interactions Between Natural Genetic Polymorphisms and Environments in Yeast Growth. Genetics. 2017;205: 925–937. doi:10.1534/genetics.116.195487

68. Qian W, Ma D, Xiao C, Wang Z, Zhang J. The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. Cell Reports. 2012;2: 1399–1410. doi:10.1016/j.celrep.2012.09.017

69. Jerison ER, Kryazhimskiy S, Desai MM. Pleiotropic consequences of adaptation across gradations of environmental stress in budding yeast. arXiv:14097839 [q-bio]. 2014 [cited 18 Jul 2021]. Available: http://arxiv.org/abs/1409.7839

70. Wytock TP, Zhang M, Jinich A, Fiebig A, Crosson S, Motter AE. Extreme Antagonism Arising from Gene-Environment Interactions. Biophysical Journal. 2020;119: 2074–2086. doi:10.1016/j.bpj.2020.09.038

71. Costanzo M, Hou J, Messier V, Nelson J, Rahman M, VanderSluis B, Wang W, Pons C, Ross C, Usaj Matej, Koch EN, Aloy P, Andrews B. Environmental robustness of the global yeast genetic interaction network. Science. 2021;372: eabf8424. doi:10.1126/science.abf8424

72. Bandyopadhyay S, Mehta M, Kuo D, Sung M-K, Chuang R, Jaehnig EJ, Huh W-K, Aebersold R, Keogh M-C, Krogan NJ, Ideker T. Rewiring of Genetic Networks in Response to DNA Damage. Science. 2010;330: 1385–1389. doi:10.1126/science.1195618

73. Harrison R, Papp B, Pal C, Oliver SG, Delneri D. Plasticity of genetic interactions in metabolic networks of yeast. Proceedings of the National Academy of Sciences. 2007;104: 2307–2312. doi:10.1073/pnas.0607153104

74. de Visser JAG, Lenski RE. Long-term experimental evolution in Escherichia coli. XI. Rejection of non-transitive interactions as cause of declining rate of adaptation. BMC Evol Biol. 2002;2: 19. doi:10.1186/1471-2148-2-19

75. Flynn KM, Cooper TF, Moore FB-G, Cooper VS. The Environment Affects Epistatic Interactions to Alter the Topology of an Empirical Fitness Landscape. Fay JC, editor. PLoS Genet. 2013;9: e1003426. doi:10.1371/journal.pgen.1003426

76. Hall AE, Karkare K, Cooper VS, Bank C, Cooper TF, Moore FB -G. Environment changes epistasis to alter trade-offs along alternative evolutionary paths. Evolution. 2019;73: 2094–2105. doi:10.1111/evo.13825

77. de Vos MGJ, Poelwijk FJ, Battich N, Ndika JDT, Tans SJ. Environmental Dependence of Genetic Constraint. Hoekstra HE, editor. PLoS Genet. 2013;9: e1003580. doi:10.1371/journal.pgen.1003580

78. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature. 2009;461: 1243–1247. doi:10.1038/nature08480

79. Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. Second-Order Selection for Evolvability in a Large Escherichia coli Population. Science. 2011;331: 1433–1436. doi:10.1126/science.1198914

80. Pearson VM, Miller CR, Rokyta DR. The Consistency of Beneficial Fitness Effects of Mutations across Diverse Genetic Backgrounds. Kassen R, editor. PLoS ONE. 2012;7: e43864. doi:10.1371/journal.pone.0043864

81. MacLean RC, Perron GG, Gardner A. Diminishing Returns From Beneficial Mutations and Pervasive Epistasis Shape the Fitness Landscape for Rifampicin Resistance in *Pseudomonas aeruginosa*. Genetics. 2010;186: 1345–1354. doi:10.1534/genetics.110.123083

82. Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, Lazar V, Hryan M, Kemmeren P, Pai C, Barton NH. The Genomic Landscape of Compensatory Evolution. Barton NH, editor. PLoS Biol. 2014;12: e1001935. doi:10.1371/journal.pbio.1001935

83. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science. 2014;344: 1519–1522. doi:10.1126/science.1250939

84. Rokyta DR, Abdo Z, Wichman HA. The Genetics of Adaptation for Eight Microvirid Bacteriophages. J Mol Evol. 2009;69: 229–239. doi:10.1007/s00239-009-9267-9

85. Schoustra S, Hwang S, Krug J, de Visser JAGM. Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. Proc R Soc B. 2016;283: 20161376. doi:10.1098/rspb.2016.1376

86. Couce A, Tenaillon OA. The rule of declining adaptability in microbial evolution experiments. Front Genet. 2015;6: 99. doi:10.3389/fgene.2015.00099

87. Wei X, Zhang J. Patterns and Mechanisms of Diminishing Returns from Beneficial Mutations. Agashe D, editor. Molecular Biology and Evolution. 2019;36: 1008–1021. doi:10.1093/molbev/msz035

88. Johnson MS, Martsul A, Kryazhimskiy S, Desai MM. Higher-fitness yeast genotypes are less robust to deleterious mutations. Science. 2019;366: 490–493. doi:10.1126/science.aay4199

89. Maisnier-Patin S, Roth JR, Fredriksson Å, Nyström T, Berg OG, Andersson DI. Genomic buffering mitigates the effects of deleterious mutations in bacteria. Nat Genet. 2005;37: 1376–1379. doi:10.1038/ng1676

90. Perfeito L, Sousa A, Bataillon T, Gordo I. RATES OF FITNESS DECLINE AND REBOUND SUGGEST PERVASIVE EPISTASIS. Evolution. 2014;68: 150–162. doi:10.1111/evo.12234

91. Reddy G, Desai MM. Global epistasis emerges from a generic model of a complex trait. eLife. 2021;10: e64740. doi:10.7554/eLife.64740

92. Lyons DM, Zou Z, Xu H, Zhang J. Idiosyncratic epistasis creates universals in mutational effects and evolutionary trajectories. Nat Ecol Evol. 2020;4: 1685–1693. doi:10.1038/s41559-020-01286-y

Chapter 2 Development of the RB-TnSeq Method

In this chapter, we report the development of the RB-TnSeq method that was applied to simultaneously generate and measure the fitness effects of many identical mutations in different genetic backgrounds and environments. In addition, we introduce the custom bioinformatics pipeline that was developed to analyze the sequencing data from the RB-TnSeq experiment. We also describe the setup of the confirmation experiment that was used to validate the reproducibility of the measurements obtained in the RB-TnSeq experiment.

2.1 Selection of Strains and Mutations

2.1.1. Strains

We tested the performance of the RB-TnSeq method on different types of yeast *Saccharomyces cerevisiae* strains. In the preliminary experiments, we used strains from the experiment described in [1]. The genomes of these strains differed by three to seven mutations, on average, which translated into fitness differences between them of a maximum of 1.5 % in various environments. The high similarity between these strains also manifested in the high similarity between the fitness effects of mutations across these genotypes (Fig. 2.1). As a result, this made it impossible to identify a statistically significant difference in the mean and the variance of the distributions of fitness effects of mutations (DFE) and compare the epistasis patterns across genotypes.

To ensure a sufficient signal difference between genotypes in this experiment, we selected a set of 42 yeast *S. cerevisiae* segregants from a cross between a prototrophic MAT**a** BY parent and a prototrophic MATα *hoΔ::HphMX4 flo8Δ::NatMX4 AMN1–BY* RM parent [2]. In our previous study, these segregants were identified to significantly differ in fitness and to have DFEs with

statistically different values of the mean and variance in YPD (20 g of peptone,10 g of yeast extract, 20 g of dextrose in 1 l of H_2O) media [3] (Fig. 2.2, Table S2.1).

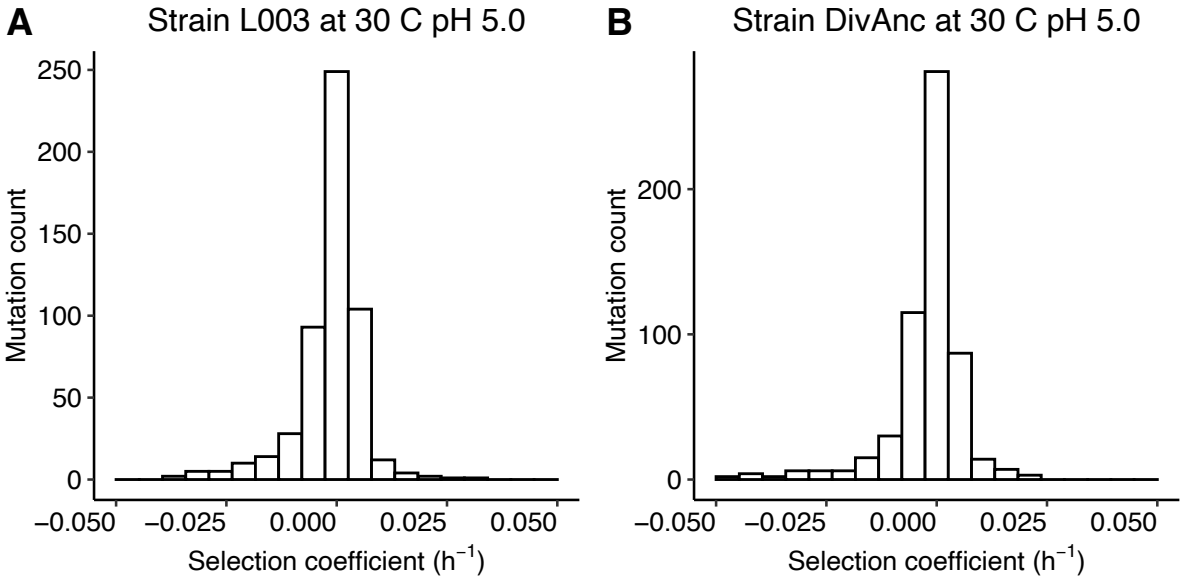

Figure 2.1: Distributions of Fitness Effects of Random RB-TnSeq Mutations in Strains Which Differ in Fitness by 1.0 % at 30 °C pH 5.0.

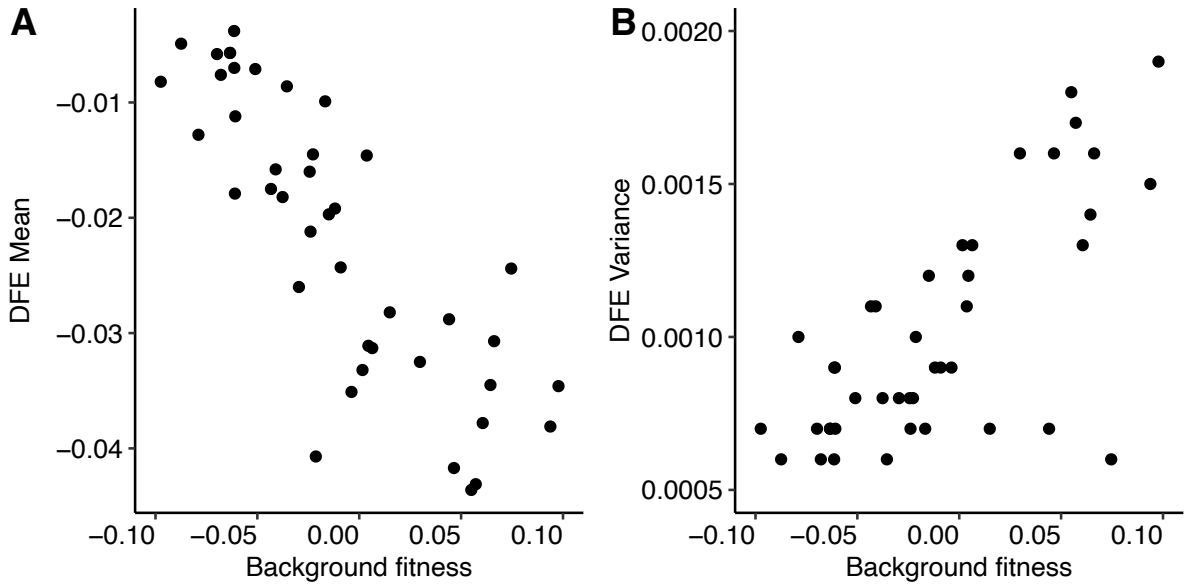

Figure 2.2: The Characteristics of the Selected Yeast Segregants Based on the Measurements in the YPD Media [3].

2.1.2 Mutations

In the preliminary experiments, we tested the fitness effects of 781 random RB-TnSeq mutations in different yeast genotypes. We found that 560 out of 781 mutations were neutral (Fig. 2.1) regardless of the genotype or the environment which confirms the observations of the earlier studies about the prevalence of mutations with neutral effects among sets of random mutations [4,5]. Having a high proportion of neutral mutations among the set of RB-TnSeq mutations would lead to getting the weaker signal difference between genotypes and losing a lot of sequencing read coverage on non-informative mutations. To overcome these limitations, we narrowed down the set of the RB-TnSeq mutations to 100. These mutations were selected out of the pool of 1180 mutations because they were identified to have strong effects on fitness across many yeast segregants in YPD media [3].

2.2 Design of the RB-TnSeq Libraries

The Desai lab at Harvard kindly provided us with yeast segregants that were already transformed with RB-TnSeq libraries to generate a set of 100 identical RB-TnSeq mutations in each genotype. Each segregant was transformed twice which allowed us to use mutants generated in separate transformations as biological replicates.

On average, each mutation was represented by 12 barcodes in transformation 1 and by 27 barcodes in transformation 2. Each barcode could be mapped onto one unique location in the genome and represented only one mutation. There were five out of 100 mutations that were specially generated in the intergenic regions. These mutations were used as a neutral reference in [3], and we used them as a reference in our experiment to estimate the fitness of segregants and selection coefficients of the generated mutations. Reference mutations were represented by a higher number of barcodes: 25 and 62 in transformation 1 and transformation 2, respectively. A high barcode redundancy of mutations allowed us to identify and exclude barcode lineages with

adaptive mutations that appeared during fitness assays or represented the existing genetic variation in the populations of yeast segregants, or lineages that got multiple insertions within the same genome (Fig. 2.3). In addition, each segregant was transformed with unique barcode libraries which allowed us to distinguish mutants of different segregants and perform fitness assays in bulk.

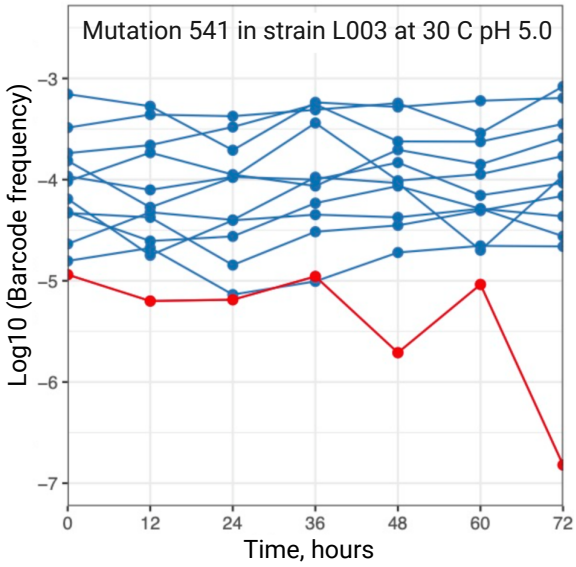

Figure 2.3: The Example of the Outlier Barcode Lineage That Was Identified in the Preliminary Experiment With Strain L003 at 30 °C pH 5.0. Each line represents a separate barcode lineage. All these lineages are associated with the same mutation. The lineage shown in red was identified as an outlier and was removed from the analysis.

2.3 Competition Assay Setup

2.3.1 Competition Assay Format

In preliminary experiments, we did fitness assays in 96 well plates, 50 ml flasks with 10 ml of media, and in 500 ml flasks with 150 ml of media. While it was easier to maintain cultures in 96 well plates and required small volumes of media, frequent cross-contamination between cultures from different wells resulted in a low correlation between estimates from replicates and, consequently, inaccuracy of the obtained measurements. The problem of cross-contamination

was resolved by doing competition assays in flasks; however, this did not account for all variance between estimates in our data. We found that low bottleneck sizes of in the competition assays done in plates ($1 \cdot 10^4$ cells) and in 50 ml flasks ($5 \cdot 10^5$ cells) lead to low cell representation of each barcode in a competition assay (less than 50 cells per each barcode at initial time point). This frequently resulted in the inability to measure the effects of mutations with deleterious effects because the barcodes associated with such mutations were removed from the competition assays due to dilution or were represented by only a few cells. To get accurate estimates for the effects of deleterious mutations, we increased the bottleneck size in the competition assays to $5 \cdot 10^7$ cells (in 500 ml flasks with 150 ml of media). In our experimental setup with 42 yeast segregants and 100 mutations represented by many barcodes, this translated into at least 1000 cells representing each barcode. As we demonstrate in chapter 3, doing competition experiments with large bottleneck size allowed us to get accurate and reproducible measurements of fitness effects.

2.3.2 Media and Environments

Fitness assays were performed in synthetic complete media (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 antibiotics Amp (ampicillin) and Tet (tetracycline) (Table S2.2) into SC media to prevent bacterial contamination.

We used six environmental conditions for fitness assays with temperature (30°C and 37°C) and pH (3.2, 5.0, and 7.0) as stress factors. Our goal was to compare the measurements from the environments that differ in the degree of similarity between each other. Therefore, we selected environments that represent a gradient of pH values and two temperature levels. We used pH 3.2 (instead of pH 3.0) because at pH 3.0 most of the segregants either had very low fitness or exhibited no growth. The constant pH of the media was maintained using citric

phosphate buffer because this buffer has a wide buffering range (3.0 – 8.0) [6]. To prepare the buffer, we used 1 M stocks of citric acid (VWR, #97061-858) and K_2HPO_4 (VWR, #97062-234) and followed the protocol described in [7]. The adjustments of the pH were performed for the autoclaved media by adding the necessary volumes of the sterile citric acid and K_2HPO_4 solutions and measuring the pH of the buffered media. When 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). After pH adjustment, the media was used within the next two days.

2.3.3 Competition Assays

The strain 24-hour pre-growth was performed in 96 deep well plates with 1 ml of buffered media in each environment. The subsequent steps of propagation were done in 500 ml flasks with 150 ml of media in shaking incubators (at 150 rpm). To be able to simultaneously maintain the growth of 504 cultures (mutants of 42 segregants x 6 environments x 2 transformations), we arranged the segregants into three and two groups in 30 °C and 37 °C environments, respectively, based on their fitness estimates obtained in the preliminary experiment. Each segregant was represented only in one group. However, a segregant LK5-G01 was intentionally added to each group, and we used the fitness estimates for this segregant to control for the similarity between the assay conditions.

After 24 h of the pre-growth, we measured the population densities of mutant cultures of each segregant and used these data to calculate the volumes needed to pool segregants within each group. To convert optical density values (OD600) into cell concentration, we applied the equation (2.1) that was obtained from building a calibration curve representing a correspondence of the OD600 values to cell concentration.

$$Log_{10}(OD600) = 0.9224 \times Log_{10}\left(Cell\ concentration, \frac{cells}{ml}\right) - 6.8814 \qquad (2.1)$$

Although we aimed to set the same initial abundances for segregants with similar fitness within each group, we intentionally added higher cell numbers for populations of low fitness segregants to avoid their loss due to the competition. After pooling the cultures, we again measured the density of the resulting cultures and completed the first transfer of $5 \cdot 10^7$ cells (a bottleneck size at each transfer) into 150 ml of fresh buffered media in 500 ml flasks to start time point one. We froze the remaining mixed cultures, marking it as a time point zero. We propagated cultures with shaking (at 150 rpm) with a 12-hour transfer schedule. After each transfer, we pelleted cells from 50 ml of the remaining cultures and froze the cell pellets at -80 °C. To maintain the consistency of the bottleneck size, we measured the OD600 of the cultures after 12 hours of growth and calculated the necessary transfer volume using the equation 2.1. When the transfer volume exceeded 1 ml, we adjusted the volume of fresh media to maintain the consistent culture volume of 150 ml in each assay.

We did fitness assays in two replicates, where each replicate represented mutants generated in a separate transformation. There were 18 pooled cultures at 30 °C and 12 pooled cultures at 37°C, and we took measurements at five time points for each culture.

2.4 Sequencing Library Preparation

We used the YeaStar Genomic Kit Protocol I (Zymo Research, #D2002) to extract gDNA from ~ 1 ml of pelleted yeast cultures. We used a two-step PCR protocol to generate the Illumina-ready dual-indexed amplicon library tagged with UMIs (unique molecular identifiers) and with dual barcodes. For PCR we modified the protocol described in [8]. For each sample, at a first PCR step, we combined 100 ng of extracted gDNA, 25 µl OneTaq DNA polymerase Master Mix (New England BioLabs, #M0482L), 0.5 µl 10 µM oAM-R2P-100-R01 primer, 0.5 µl 10 µM oAM-R1P-20X-F01 primer (Table S2.3), 1 µl 50 µM $MgCl_2$, and molecular biology grade H_2O up to the total volume of 50 µl. Ran the following PCR protocol: 1). 94 °C 0:30, 2). 94 °C 0:30, 3). 50.5 °C 0:30,

4). 68 °C 1:00, 5). GO TO step 2 3 times, 6). 68 °C 5:00. Then, we purified this PCR product with AMPure XP magnetic beads (Beckman, #A36881) (1:1 ratio). To set up a second PCR step, we used 15 µl of purified PCR I product, 25 µl OneTaq DNA polymerase Master Mix, 1 µl 50 µM $MgCl_2$, 1 µl 10 µM N7XX primer (Nextera), and 1 µl 10 µM S5XX primer (Nextera), and 7 µl molecular biology grade H_2O. Ran the following protocol: 1). 94 °C 0:30, 2). 94 °C 0:30, 3). 62 °C 0:30, 4). 68 °C 1:10, 5). GO TO step 2 24 times, 6). 68 °C 5:00. After, we purified a second PCR product with AMPure XP magnetic beads (1:1 ratio) and did gel extraction and purification using QIAquick PCR purification kit (Qiagen, #28106). We sequenced final libraries with paired-end 150 bp reads on one HiSeq4000 platform and two NovaSeq platforms (Illumina).

The use of UMIs allowed us to identify barcodes that were subject to multiple amplification during the PCR. The addition of dual inline indices to the DNA fragments during first PCR step allowed to identify and remove contamination reads that are introduced at sequencing step as a result of index hopping (Fig. 2.4).

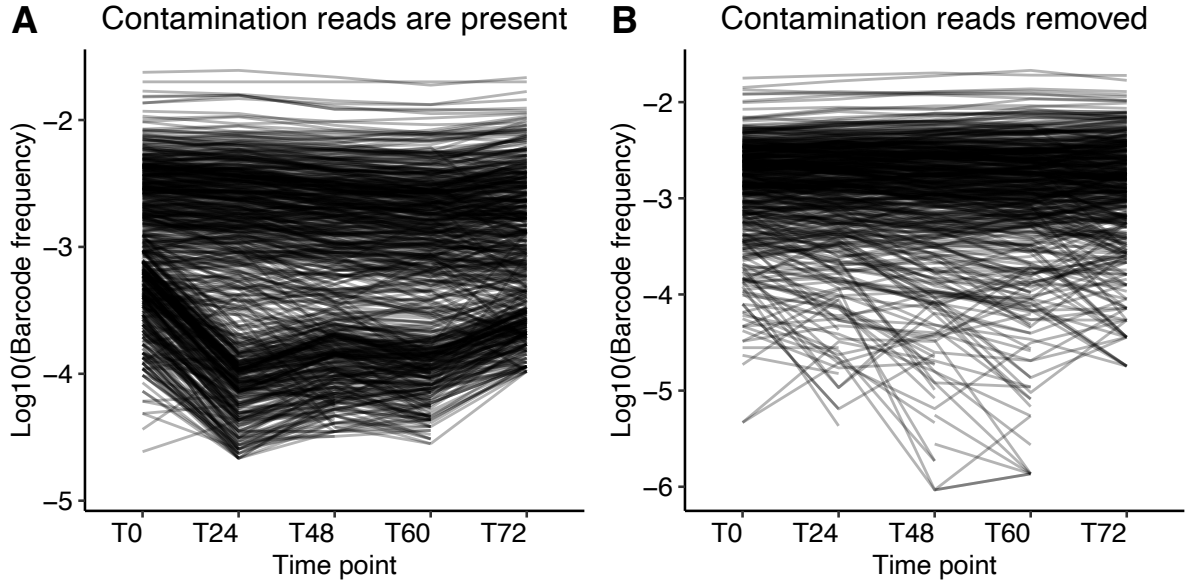

Figure 2.4: Barcode Frequencies Representing Strain L013 at 30 °C pH 5.0 Before and After Contamination Reads Were Removed. These data were obtained in the preliminary experiment.

2.5 Data Analysis Methods

2.5.1 Extracting Barcodes and Associating Them With Specific Mutations

Before extracting barcodes, we filtered the sequencing reads based on the sequence of the inline barcodes that were attached during the first round of PCR. That allowed us to identify and remove contamination reads that are a result of index hopping at a sequencing step. The proportion of such contamination reads reached up to 15 % for some samples. Therefore, their identification and removal was a crucial step that enabled us to get accurate estimates.

To extract the barcodes from sequencing reads, we applied the regular expressions function from the python regex module (https://pypi.org/project/regex/). Since the structure of the sequencing reads is known, we specified the sequences of the flanking regions for each barcode (they were identical for each read) and extracted the 26 - 28 nucleotide region in between them that supposedly corresponded to the barcode sequence. To associate the extracted barcodes with the known barcode sequences, we applied a custom-designed function that identified the similarity between the extracted and reference barcodes based on the Hamming distance calculated for these barcodes and allowing for one nucleotide mismatch. Since each reference barcode was associated with a specific mutation, matching the extracted barcode to the reference barcode enabled us to add the corresponding information about the mutation. The proportion of barcodes that could not be matched to the reference barcode library ranged between 7 % and 10 %. We assume that these barcodes correspond to the low frequency barcodes that were present in the transposon libraries but were not associated with a particular mutation.

To control for amplification bias during the library preparation, we also extracted UMI sequences associated with each barcode. UMI sequences corresponded to the first eight nucleotides in the read sequence.

2.5.2 Determining Barcode Counts

To estimate the counts for barcodes, we calculated the number of unique UMI sequences associated with each barcode. However, we identified that the cases of biased barcode amplification were rare, which suggests that we could skip this step and could simply count the number of occurrences for each barcode.

2.5.3 Estimating Fitness Effects From Barcode Counts

We excluded barcodes with less than 10 reads and mutations represented by less than three barcodes from the analysis. We normalized the frequency of each barcode by the mean frequency of barcodes representing five mutations that were used as a neutral reference. Frequencies were calculated based on the read counts obtained for a population of the corresponding segregant. This normalization step corrected for the difference in the read counts between different sequencing samples and accounted for the changes in the frequency of populations of segregants between time points.

We used the following formula to estimate selection coefficient for each barcode at each time interval:

$$s_i = \frac{1}{T} Ln \frac{F_{T+1}}{F_T}, \quad (2.2)$$

where T -time between transfers (12 h), F_{T+1} – normalized frequency of a barcode at time point T+1, F_T – normalized frequency of a barcode at time point T, s_i- selection coefficient for a barcode at time interval i, h^{-1}.

We calculated the mean and the variance for the barcode selection coefficient estimate using the following formula:

$$\overline{s_{bc_k}} = \frac{1}{N} \sum_{i=1}^{N} s_i, \quad (2.3)$$

$$Var_k = \frac{1}{N} \left[\frac{1}{N-1} \sum_{i=1}^{N} \left(s_i - \overline{s_{bc_k}} \right)^2 \right], \quad (2.4)$$

where $\overline{s_{bc_k}}$- mean selection coefficient for a barcode k, h^{-1}, N - number of time intervals used for calculation, Var_k- variance in fitness effect estimate for the barcode k, h^{-2}.

2.5.4 Identifying Outlier Barcodes

We identified that the selection coefficient for some barcodes differed significantly from the selection coefficients of most barcodes associated with the same mutation. The observed deviations can be explained by either the presence of new adaptive mutations, pre-existing genetic variation in the populations of segregants before transformation, or by the insertion of multiple RB-TnSeq mutations within the same genome. The redundancy of barcodes representing each mutation allowed us to identify and remove most of such outliers.

We firstly classified barcodes into two groups. For a particular mutation, we calculated the median for the selection coefficients of the barcodes associated with that mutation and found barcodes with selection coefficient estimates within 0.01 of the calculated median value. We classified these barcodes as reliable, while barcodes with selection coefficients outside of that range – as potential outliers.

For a given mutation, we calculated a confidence interval for its selection coefficient measurement based on the si estimates for reliable barcodes at all time intervals. Specifically, we applied qnorm() function in R to find boundary values for the normal distribution of s_i estimates with the mean of s_{av} and the standard deviation of sd_s. Since the s_i values constituting the distribution were estimated at several time intervals, we adjusted the margins of the 95 % confidence interval based on the number of the unique time intervals used for the calculation. Barcodes from the “potential outliers” group, the mean selection coefficients, $\overline{s_{bc_k}}$, of which exceeded the limits of the calculated confidence interval for the selection coefficient of the mutation were identified as outliers and were removed from the analysis.

Procedure to calculate the limits for the confidence interval:

1). $max_{conf} = \sqrt[N]{(1 - 0.025)}$,

2). $Upper = qnorm(max_{conf}, s_{av}, sd_s, lower.tail = T)$,

3). $Lower = 2 \cdot s_{av} - Upper$,

where upper and lower represent limits of the confidence interval, max_{conf} -adjusted boundary of the 95 % confidence interval based on N - the number of the time intervals used for calculation, sav and sd_s – the mean and the standard deviation of the estimates of selection coefficients for "reliable" barcodes from N different time intervals.

Using this procedure, we identified and removed 15.9 % (161249 out of 1013089) outlier barcodes. We compared the efficiency of this procedure to the barcode excluding method described in [37]. The application of that method resulted in excluding a larger fraction of barcodes (23 %, 233011 out of 1013089) without significantly improving the Pearson's correlation coefficients for the estimates of fitness effects from different transformations. The latter conclusion is based on the comparison of the distributions of the obtained Pearson's correlation coefficients for estimates in different segregants and environments after the application of these filtering methods. Therefore, to proceed we used the barcode exclusion procedure that removed fewer barcodes.

2.5.5 Inferring Selection Coefficients for Mutations

After filtering out the outliers, we used the inverse-variance weighting method to calculate the selection coefficient for each mutation in each genotype and environment. The application of this method ensured that the barcodes with less-variant fitness estimates contribute more to the final estimate of the selection coefficient for a mutation and, therefore, improved the accuracy of the resulting measurements.

$$\overline{s_{mut}} = \frac{\sum_{k=1}^{M} \frac{\overline{s_{bc_k}}}{Var_k}}{\sum_{k=1}^{M} \frac{1}{Var_k}}, \quad (2.5)$$

where $\overline{s_{mut}}$ – selection coefficient for a mutation, h^{-1}, M- number of barcodes representing a given mutation, $\overline{s_{bc_k}}$- selection coefficient for a barcode k, h^{-1}, Var_k – the variance associated with the $\overline{s_{bc_k}}$ estimate, h^{-2}.

In section 3.2.1, we show that Pearson's correlation coefficients for the fitness effects of mutations estimated from transformation 1 and transformation 2 experiments varied between 0.60 to 0.95 for most segregants. The high correlation between estimates obtained from replicate assays allowed us to pool these data. Specifically, the final estimates for selection coefficients of mutations were obtained by applying the inverse-variance weighting method (equation 2.5) for all barcodes associated with the same mutation from both transformations.

2.5.6 Calculating the Mean and the Variance for the DFEs and Their Confidence Intervals

On average, we obtained the measurements for fitness effects of 82 mutations in every genotype in five out of six environments. In 37 °C pH 7.0 environment, the average number of mutations per genotype was 68. Presumably, because in this environment all strains had low fitness and genotypes with highly deleterious mutations were removed due to the competition and dilution.

To calculate a 95 % confidence interval for the mean of the DFE in each segregant we calculated the standard deviation for the bootstrap distribution of 10000 mean values calculated over the same number of mutations as the original DFE and obtained by sampling the mutations with replacement. The limits of the confidence interval were calculated using the following formula:

$$DFE\ mean = mean \pm 2 \times STD, \tag{2.6}$$

where mean – the calculated mean for the DFE, STD – the standard deviation of the generated bootstrap distribution.

To calculate a 95 % confidence interval for the variance of the DFE for each segregant, we assumed the effects of mutations comprising the DFE are normally distributed and approximated the distribution for the variance of the DFE with a chi-square distribution with n_{mut} - 1 degrees of freedom. Specifically, we performed the following calculations in R:

$$Var_{min} = \frac{(n_{mut} - 1) \times Var_{DFE}}{qchisq\big(0.975, (n_{mut} - 1)\big)}, \tag{2.7}$$

$$Var_{max} = \frac{(n_{mut} - 1) \times Var_{DFE}}{qchisq\big(0.025, (n_{mut} - 1)\big)}, \tag{2.8}$$

where n_{mut} – number of mutations in the given DFE, Var_{DFE} – calculated variance for the given DFE, h^{-2}, qchisq – a function that takes the probability value as an argument (0.975 or 0.025), along with degrees of freedom (n_{mut} -1), and returns a chi-square value corresponding to that probability.

2.5.7 Estimating Fitness of Segregants

To estimate fitness of segregants from fitness assays done in bulk, we used the frequency of neutral reference mutations and the number of cells in each culture before and after a 12-hour growth cycle and did the following calculations:

$$W_i = \frac{1}{12} Ln\left(\frac{F_T \times N_T}{F_{T-1} \times C_{T-1} \times V_{Tr}}\right), \tag{2.9}$$

where W_i – fitness estimated at time interval [T-1; T], 12- number of hours between transfers, F_T – frequency of five neutral reference mutations at time point T (estimated based on sequencing data), N_T – number of cells in the 150 ml of culture at time point T after a 12-hour growth cycle (estimated by measuring OD600 of the culture), F_{T-1} – frequency of five neutral reference mutations at time point T-1 (estimated based on sequencing data), C_{T-1} – concentration of cells in the culture after 12-hour growth cycle at T-1(estimated by measuring OD600 of the culture),

cells/ml, V_{Tr} – volume in ml of the culture with concentration C_{T-1} that was transferred to start culture corresponding to time point T.

For each segregant, we obtained fitness estimates at four time intervals. Generally, the estimates obtained at the time interval one were significantly different from the remaining estimates for most segregants as for fitness of segregants, so for selection coefficients of mutations. These differences are likely caused by the transition of cells to the growth in flasks after 24 h of incubation in a 96 deep well plates at the first transfer cycle and, therefore, by them being in a different growth regime than at the remaining time intervals. We removed the estimates of background fitness and selection coefficients of mutations at time interval one from the subsequent analysis to eliminate the variation in the data caused by the difference in growth regime of the cells.

2.6 Confirmation of the RB-TnSeq Measurements

The goal of the confirmation experiment described in this section was to test the reproducibility of the selection coefficient estimates obtained in the RB-TnSeq experiment. For that, we additionally generated separate RB-TnSeq libraries for seven mutations (nearby MET2, nearby MET4, in PPM1, in RCS30, nearby TDA11, in NOT3, in MPC2), reconstructed them in five segregants (LK1- C09, LK1-H02, LK2-B07, LK2-D05, and LK6-A05), and estimated their effects in two environments (30 °C pH 5.0 and 37 °C pH 7.0). Two of the selected mutations (nearby MET2 and nearby TDA11) were used as a neutral reference in the RB-TnSeq experiment, the mutation nearby MET4 was identified to be neutral in all tested genetic backgrounds and environments, the remaining mutations were identified to have a significant positive effect on fitness in most genetic backgrounds in 30 °C environments and to have neutral or negative effects at 37 °C. By confirming the fitness effects measurements for these mutations, we aimed to show that the changes of the sign of fitness effects of many mutations observed at 30 °C and 37 °C

environments in the RB-TnSeq experiment do not stem from the experimental or the data analysis bias. A detailed description of the RB-TnSeq experiment is given below.

2.6.1 Generation of the Barcoded Transposon Libraries

To validate the estimates of selection coefficients from the RB-TnSeq experiment, we additionally generated sets of barcode libraries for seven mutations. We followed the same protocol as was used to generate barcode libraries that were used to create mutants for the RB-TnSeq experiment [3].

Specifically, from the Desai lab, we received plasmid libraries of transposons which were flanked by unique fragments of the genomic DNA and were associated with individual mutations. These transposon libraries were identical to those used to create RB-TnSeq mutations; however, they did not have barcodes and drug resistance markers that would allow to select, differentiate, and track the mutants. To create the plasmids with all the necessary fragments, we extracted plasmids with transposons from the overnight cultures of *Escherichia coli* grown on Luria Broth (LB) media with drug kanamycin (Kan, 40 ug/ml), measured plasmid concentrations by Nanodrop, and did SalI-HF restriction according to the recipe provided for that enzyme. Simultaneously, we extracted pUC-19 Hyg and pUC-19 yNat plasmids from the corresponding overnight cultures of *E. coli* grown in LB with hygromycin (Hyg, 200 ug/ml) or nourseothricin sulfate (clonNat, 20 ug/ml), respectively. After measuring the concentrations of these plasmids, we set up the PCR reactions to extract the drug marker containing fragments and to add nucleotide overhangs that were necessary for the following Gibson Assembly step. For the PCR, we mixed 0.75 uM of the Tn_BC_Amp_Connectorator_2 primer, 0.75 uM of the tTEF_to_Tn7R primer (Table S2.3), 12.5 ul of 2X KAPA Mix, 2 ng of the corresponding pUC-19 Hyg (or yNat) plasmids and H_2O to the total volume of 25 ul. Ran the following protocol: 1). 98 °C 5:00, 2). 94 °C 0:10, 3). 63 °C 0:30, 4). 72 °C 0:25, 5). GO TO step 2 12 times, 6). 72 °C 5:00. For Gibson Assembly, mixed 2 ul of SalI-HF

restriction product (with concentration 20 ng/ul), 1 ul of the PCR product (normalized to 50 ng/ul), 5 ul Gibson mix, and 2 ul of H_2O. Incubated the mixture at 50 °C for one hour. For each mutation, did two Gibson reactions: one – with drug Hyg, another – with clonNat. After, did purification of Gibson Assembly reaction products by adding 10 ml of AMPure XP magnetic beads and eluted the purified product in 45 ul of 10 mM Tris-HCl buffer. We used chemically competent *E. coli* DH10B cells (Thermo Scientific, #EC0113) to deliver plasmids into the cells. Specifically, we thaw 50 ul aliquots of competent cells on ice, added 2 ul of the purified Gibson Assembly product to each aliquot, incubated the resulting mixtures on ice for 30 min, heat shocked the cells for 30 seconds at 42 °C, incubated the cells on ice for 2 min, then mixed them in 1 ml of S.O.C. medium (New England BioLabs, #B9020S) and incubated at 37 °C for one hour with shaking (225 rpm). Next, plated 100 ul of the culture on LB media with appropriate drugs (Table S2.2) and added the remaining part to 10 ml of liquid LB media with drugs for additional selection. After 24 hours of incubation, checked the colonies on plates and picked 10-15 colonies from each plate to create a barcode library for each mutation. We inoculated 5 ml of fresh LB media with those colonies and incubated for 24 h at 37 °C. Then, spun down the resulting cultures and froze at -80 °C. We used those frozen stocks for transformation. For the outgrowth cultures that were selected in 10 ml of LB, we concentrated them by centrifuging, created stocks in 20 % glycerol, and kept in -80 °C as a backup.

In total, we created 28 barcoded plasmid libraries. These libraries represented seven individual mutations, and each library was associated with one mutation only. Since yeast segregants used in the experiments differed in the initial drug resistance (Table S2.1), we created separate libraries with clonNat and Hyg drug markers for each mutation. Each library was represented in two replicates, replicate libraries contained unique sets of barcodes that did not overlap with barcodes of any of the 28 libraries.

2.6.2 Extraction and Digestion of the Barcoded Plasmid Libraries

We used the digested barcoded plasmids to create insertion mutations in yeast genomes by the means of homologous recombination which happened during the transformation step.

We extracted barcoded plasmids from overnight *E. coli* cultures grown in 3 ml of LB media with appropriate drugs (Table S2.2) using QIAprep spin mini prep kit (# 27104), measured plasmid concentrations by Nanodrop, and did NotI (New England BioLabs, #R0189) digestions. For that, mixed 1 ug of plasmids, 5 ul of 10x restriction buffer, 1 ul of NotI enzyme and nuclease-free H_2O up to the total volume of 50 ml. Incubated at 37 °C for 1 hour, then inactivated the enzyme by keeping the digestion product at 65 °C for 10 min. We used the product of the digestion reactions for transformations.

2.6.3 Optimization of Yeast Transformation

We tested two transformation protocols which are described below. The protocol 1 (described in [9]) was used in our preliminary experiments with yeast strains from [1]. This protocol is 2.5 hours longer than the protocol 2, requires the preparation of the additional solutions, and yields the maximum of 2000-3000 transformants from 10 ml of yeast culture. We identified that this protocol is inefficient for transforming yeast segregants from [2], as it gives 10-30 colonies from 10 ml of yeast culture. To transform segregants, we optimized the transformation protocol developed by Gietz and Schiestl [10]. The optimized protocol 2 yielded a lawn of transformant colonies from the same volume of yeast culture. However, we have not tested the efficiency of this protocol on other yeast strains, e.g., from [1]. The descriptions of both protocols are listed below.

2.6.3.1 Transformation Protocol 1

We prepared overnight cultures of strains in 3 ml of YPD. In the morning, these cultures were diluted 1:100 in a fresh YPD medium and incubated for the next six hours at 30 °C with shaking (220 rpm). After, the cultures were concentrated by centrifuge (2000 G). For each transformation, we used cells from 10 ml of yeast culture. The cells were washed three times in sterile distilled H_2O by resuspending them in 5 ml of H_2O and spinning them down in the centrifuge for 5 min. After, we did two washes in 1 ml of lithium acetate mix (100 mM lithium acetate, 10 mM Tris-HCl, 1mM EDTA), spun cells down, and decanted the liquid. Next, resuspended cells in 70 ul of lithium acetate mix, added to them 10 ul of boiled salmon sperm DNA (10 mg/ml) and 50 ul of digested plasmids from step 2.6.1, and briefly vortexed the resulting mixture. Then, we added 660 ul of polyethylene glycol mix (40 % polyethylene glycol 3350, 100 mM lithium acetate, 10 mM Tris-HCl, 1 mM EDTA) and mixed the cells by pipetting 10-15 times. After, we incubated the cells at 30 °C for 30 min, heat shocked them at 42 °C for 60 min, and sedimented them by centrifuging for 3 min. Cells were resuspended in 1 ml of YPD and incubated at 30 °C (220 rpm) for 2 hours. After the outgrowth, cells were pelleted and plated onto YPD plates with corresponding drugs (clonNat or Hyg, Table S2.2).

2.6.3.2 Transformation Protocol 2

We grew yeast strains overnight in 3 ml of YPD, diluted these cultures 1:25 in a fresh YPD medium, incubated with shaking for four hours, and used 5 ml of the resulting culture for each transformation. We spun cells down in a centrifuge (3000 G) for 3 min and washed the cells twice in 1 ml of sterile H_2O by resuspending the cell pellet in H_2O, spinning the cells down, and removing the liquid. Then, we added to the cells 324 ul of the transformation mix (240 ul of 50 % polyethylene glycol 3350, 36 ul of 1 mM lithium acetate, 50 ul of boiled salmon sperm DNA (2 mg/ml) and 34 ul of plasmid digestion product from step 2.6.1. After vigorous vortexing, the

cultures were incubated at 42 °C for 1 hour, spun down in a centrifuge (30 sec at 13k G), separated from the liquid, and resuspended in 1 ml of fresh YPD. The resuspended cultures were incubated at 30 °C (220 rpm) for 2 hours and plated on drug selection plates.

After 48 hours of incubation, we scraped the colonies of transformants and did additional 24 h of selection in liquid media with drugs in the same concentration as was used to prepare the solid media. The resulting cultures were pelleted and frozen in 20 % glycerol at - 80 °C.

2.6.4 Competition Assays and Sequencing Library Preparation

Following the protocol described in 2.6.3.2, for each segregant we obtained 14 mutant pools. Each mutant pool was associated with an individual mutation, and each mutation was represented by two pools with distinct sets of barcodes, which were treated as replicates.

To pre-grow mutant cultures for the competition assays, we started 140 (5 segregants x 7 mutations x 2 replicates x 2 environments) in the buffered SC media containing Amp and either clonNat or Hyg drugs (Table S2.2) and incubated them in the appropriate environmental conditions (30 °C pH 5.0 and 37 °C pH 7.0) for 24 h in tubes with shaking (220 rpm). After measuring the concentrations of the grown cultures, the mutants of for the same type of segregant were pooled in the following way: the initial frequency of two neutral reference mutations was 25 % each, the frequency of the each of the remaining mutations was 10 %. There were 20 different pooled cultures in total.

We measured the OD600 of the pooled cultures and transferred $5 \cdot 10^7$ cells to start time point one of the competition assays. Competition assays were performed following the same protocol as described in section 2.3.3. The only difference was that assays were done separately for each segregant, and each assay included only seven mutations.

The remaining parts of the experiment and all the data analysis were performed the same way as described for the RB-TnSeq experiment.

We compared the obtained measurements from the confirmation to the measurements from the RB-TnSeq experiment (Fig. 2.5). The correlation between these estimates is high (R^2=0.76), which suggests that the developed RB-TnSeq method produces robust and reproducible results. In addition, these data confirm that mutations with significant positive effects on fitness at 30 °C become neutral or deleterious in the same genetic backgrounds at 37 °C. This implies that our findings that some mutations have antagonistic effects between 30 °C and 37 °C environments are real.

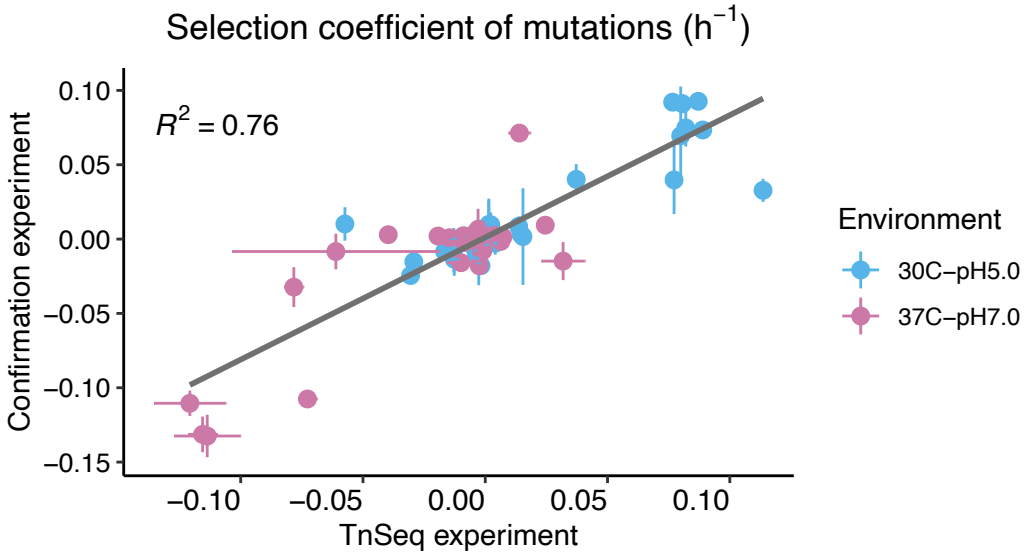

Figure 2.5: Correlation Between Selection Coefficient Estimates Obtained in the RB-TnSeq and Confirmation Experiments. Each dot represents a particular mutation in a particular genotype in one of the two environments.

2.7 Acknowledgements

Chapter 2, in part, is a reprint of Milo Johnson, Alena Martsul, Kryazhimskiy Sergey, Desai Michael, “Higher-fitness yeast genotypes are less robust to deleterious mutations,” *Science*, 2019. The dissertation author was the primary investigator of the material described in chapter 2 and one of the authors of this paper.

2.8 Supplementary Materials

Table S2.1: Characteristics of the Selected Segregants

Segregant	Barcode number transformation 1	Barcode number transformation 2	Drug resistance	Relative fitness *	DFE mean*	DFE variance*
LK1-A02	625	2683	Hyg	0.0746	-0.0244	0.0006
LK1-B03	752	3359	Hyg	-0.0613	-0.0070	0.0009
LK1-B05	1062	3458	Hyg	-0.0510	-0.0071	0.0008
LK1-C09	1019	2765	Hyg	0.0978	-0.0346	0.0019
LK1-F05	1813	4336	clonNat	0.0465	-0.0417	0.0016
LK1-G03	778	3224	Hyg	0.0441	-0.0288	0.0007
LK1-H02	856	3075	NA	-0.0634	-0.0057	0.0007
LK2-A06	529	832	NA	-0.0874	-0.0049	0.0006
LK2-A10	560	2776	Hyg	-0.0698	-0.0058	0.0007
LK2-A12	720	4854	clonNat	-0.0296	-0.0260	0.0008
LK2-B07	496	5450	clonNat	-0.0038	-0.0351	0.0009
LK2-D05	863	2214	NA	-0.0410	-0.0158	0.0011
LK2-D07	801	487	clonNat	-0.0091	-0.0243	0.0009
LK2-E08	1077	1393	clonNat	-0.0119	-0.0192	0.0009
LK2-F04	1693	3856	NA	0.0644	-0.0345	0.0014
LK3-B08	617	4976	NA	0.0016	-0.0332	0.0013
LK3-B10	819	1213	clonNat	0.0606	-0.0378	0.0013
LK3-C04	840	1727	NA	-0.0213	-0.0407	0.0010
LK3-C11	418	2638	clonNat	-0.0243	-0.0160	0.0008
LK3-C12	597	3641	Hyg	0.0045	-0.0311	0.0012
LK3-D01	523	2530	NA	0.0662	-0.0307	0.0016
LK3-G08	743	2068	clonNat	0.0938	-0.0381	0.0015
LK3-H11	475	895	clonNat	-0.0376	-0.0182	0.0008
LK4-B01	401	2351	clonNat	-0.0610	-0.0179	0.0009
LK4-B12	1255	2410	Hyg	-0.0227	-0.0145	0.0008
LK4-D11	2020	671	clonNat	-0.0608	-0.0112	0.0007
LK4-H11	869	2099	Hyg	-0.0433	-0.0175	0.0011

Table S2.1 Continued

Segregant	Barcode number transformation 1	Barcode number transformation 2	Drug resistance	Relative fitness *	DFE mean*	DFE variance*
LK5-C01	566	2788	Hyg	-0.0355	-0.0086	0.0006
LK5-C04	401	2917	NA	0.0550	-0.0436	0.0018
LK5-C11	445	3302	NA	-0.0789	-0.0128	0.0010
LK5-D09	631	5370	clonNat	0.0037	-0.0146	0.0011
LK5-E09	768	4067	NA	-0.0149	-0.0197	0.0012
LK5-F07	560	4306	clonNat	0.0572	-0.0431	0.0017
LK5-F08	710	2841	Hyg	0.0150	-0.0282	0.0007
LK5-G01	727	3844	Hyg	-0.0974	-0.0082	0.0007
LK5-G03	403	2597	Hyg	-0.0614	-0.0038	0.0006
LK5-H12	819	2657	clonNat	-0.0239	-0.0212	0.0007
LK6-A05	615	902	clonNat	0.0064	-0.0313	0.0013
LK6-B07	935	2991	Hyg	-0.0167	-0.0099	0.0007
LK6-B10	586	1368	clonNat	0.0298	-0.0325	0.0016
LK6-D06	1106	4005	Hyg	-0.0634	-0.0057	0.0007
LK6-D07	517	3454	Hyg	-0.0679	-0.0076	0.0006

* Measurements obtained in YPD media in the experiment described in [3].

clonNat - stands for drug Nourseothricin, Hyg - drug Hygromycin B, NA - no drug resistance. Barcode numbers show the total number of unique barcodes that represent 100 mutations. In transformation 1, each mutation was represented by 12 unique barcodes, and in transformation 2 by 27 barcodes, on average.

Table S2.2: Antibiotic Concentrations Used in the Study

Antibiotic	Concentration for *E. coli* selection (ug/ml)	Concentration for *S. cerevisiae* selection (ug/ml)
Kanamycin (Kan)	40	NA
Ampicillin (Amp)	100	100 (used to prevent bacterial growth)
Nourseothricin sulfate (clonNat)	20	20
Hygromycin (Hyg)	200	300

Table S2.3: Primers Used in the Study

Purpose	Name	Sequence
First round barcode PCR	oAM-R1P-200-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNC**TTACTCAGGC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-201-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NN**GGACGCATC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-202-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NN**AACGTGC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-203-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNT**CGTC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-204-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNT**AGCAGTCGTC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-205-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNA**CTTCTAGC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-206-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNA**TAGAGTGATACC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-207-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNC**AATTGTGCGC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-208-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NNGC**TTACTAGTC**CGTAACGTAGGTCTCTGACG
	oAM-R1P-209-F01	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNN NN**CTGCTTTAACAGC**CGTAACGTAGGTCTCTGACG
	oAM-R2P-100-R01	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNN NNNGACACAATGCTTGCTGATAAATCTGGAGCC
PCR for barcoding Gibson Assembly	Tn_BC_Amp_Connectorator_2	CAAACTGGACAAAAAAGATCCGTAACGTAGGTCTCTGACG CNNNNCANNNNCANNNNCANNNNCANNNNTTCTACGGGG TCTGACGC
	tTEF_to_Tn7R	TTGAACTGAACAAAATAGATCCGTGGATGGCGGCGTTAG

* N sequence fragments in the first round barcode PCR primers correspond to eight nucleotides of the UMI sequence; fragments of primers shown in bold represent inline barcode sequences that were used to identify and remove contamination reads from the sequencing step.

2.9 References

1. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science. 2014;344: 1519–1522. doi:10.1126/science.1250939

2. Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, Kruglyak L. Finding the sources of missing heritability in a yeast cross. Nature. 2013;494: 234–237. doi:10.1038/nature11867

3. Johnson MS, Martsul A, Kryazhimskiy S, Desai MM. Higher-fitness yeast genotypes are less robust to deleterious mutations. Science. 2019;366: 490–493. doi:10.1126/science.aay4199

4. Qian W, Ma D, Xiao C, Wang Z, Zhang J. The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. Cell Reports. 2012;2: 1399–1410. doi:10.1016/j.celrep.2012.09.017

5. Harrison R, Papp B, Pal C, Oliver SG, Delneri D. Plasticity of genetic interactions in metabolic networks of yeast. Proceedings of the National Academy of Sciences. 2007;104: 2307–2312. doi:10.1073/pnas.0607153104

6. Common Buffers and Stock Solutions. Current Protocols in Food Analytical Chemistry. 2001;00. doi:10.1002/0471142913.fax02as00

7. Gomori G. Preparation of buffers for use in enzyme studies. Methods in Enzymology. Elsevier; 1955. pp. 138–146. doi:10.1016/0076-6879(55)01020-3

8. Levy SF, Blundell JR, Venkataram S, Petrov DA, Fisher DS, Sherlock G. Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature. 2015;519: 181–186. doi:10.1038/nature14279

9. Lang GI. Yeast transformation protocol. 2017. Available: https://glanglab.com/protocols.html

10. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007;2: 31–34. doi:10.1038/nprot.2007.13

Chapter 3 Investigating Epistasis Patterns in Different Environments

The rapid evolution of microorganisms underpins the emergence of multiple drug resistance [1–3], the transmission of infectious diseases among species [4,5], and has a significant impact on human life [6]. To successfully prevent the negative consequences of microbial evolution, it is important to understand its principles and be able to predict the available adaptive trajectories for evolving populations. However, this task is rather challenging, as it requires prior knowledge of the effects of adaptive mutations that will sequentially appear in an evolving genotype in the set of environmental conditions it encounters.

Epistasis, i.e., genetic interactions between mutations, makes effects of mutations dependent on their genetic background. Epistatic interactions can modify the magnitude [1,7–9] and change the sign [10–13] of the fitness effect of a mutation. Moreover, these interactions can be caused by various mechanisms, ranging from the exhaustion of stability threshold for some molecules [14–16] to the variation in the ability to compensate for the functions of the disrupted genes by the remaining functional genes within a metabolic network [7]. Since the genetic context of a mutation and, perhaps, the causes of epistasis change in different genotypes, mutations can exhibit different types of epistatic interactions depending on their genetic background, e.g., positive, negative, and sign types of epistasis were identified for *rpoB* mutations [3,12,17–19]. In addition to epistasis, environments can modify the effects of mutations in different ways: by changing the rate of biochemical and biophysical processes inside the cell [20–22], by altering the underlying gene interaction networks [23,24], or by selecting distinct phenotypes [25]. Therefore, one might expect the effects of mutations to be highly idiosyncratic and unpredictable.

However, a negative linear relationship between the effects of beneficial mutations and background fitness, i.e., diminishing-returns epistasis, has been observed in evolution experiments done with various model organisms, e.g., with bacteria *Escherichia coli* [18,26,27]

and *Pseudomonas aeruginosa* [10,28], with yeast *Saccharomyces cerevisiae* [29,30], with bacteriophages [31] and multicellular fungus [32]. The dependence of the effects of some beneficial mutations on the background fitness was confirmed by directly generating and measuring the effects of these mutations in genetic backgrounds of various fitness [9,30,33–35]. This suggests that those mutations are subject to global epistasis, a phenomenon described in [30] which refers to the dependence of the effects of mutations on the fitness of their genetic background and not the genotype itself. A recent study by Wei and Zhang showed that a large fraction of mutations is subject to global epistasis, as between 66% to 92% of tested naturally occurring genetic polymorphisms exhibit diminishing-returns epistasis across 47 different environments [36]. Taken together, this suggests that global epistasis is widespread among beneficial mutations (but not all mutations exhibit it). Different types of mutations can be subject to it, and some mutations exhibit diminishing-returns epistasis only in some environmental conditions. However, these conclusions are relevant to only beneficial mutations, and it is unclear whether similar patterns could be observed for deleterious mutations.

In our recent study, we found that random mutations tend to have stronger deleterious effects in more-fit genotypes, i.e., exhibit increasing cost epistasis, as the distributions of the fitness effects (DFE) of the same set of mutations have lower mean and higher variance in genotypes with higher fitness [37]. This suggests mutations with deleterious effects are also affected by global epistasis and their effects scale negatively with the background fitness.

The recently suggested connectedness model [38] and idiosyncratic epistasis theory [39] suggest that the observed patterns for beneficial and deleterious mutations both represent the same trend. This trend is a consequence of widespread epistasis among individual loci. On the maximally idiosyncratic fitness landscape, the fitness of neighboring genotypes tends to be the same regardless of the fitness of the focal genotype. Thus, the fitness difference between the focal genotype and its neighbors tends to be more negative or less positive for focal genotypes with higher fitness [39]. However, these models do not give any predictions about how the

relationships describing the dependence of the effects of mutations on the fitness of their genetic background change in different environments. For example, it is unclear whether the observed patterns for the DFE described in [37] will be the same in environments in which most of the tested mutations change the sign of their fitness effect. For individual mutations, it is not known whether fitness of the genotype can be used as a predictor of their effects in different environments.

To bridge that gap, we propose to focus on a single-step mutational neighborhood of genotypes. For that, we suggest generating and measuring the fitness effects of many identical mutations across genotypes of various fitness in environments that differ in the degree of similarity to each other. Using this approach, we can evaluate how the shape of the entire DFE changes with respect to the fitness of genotypes in different environments. If the relationship is not affected by the environment, we can expect to observe consistent patterns even between environments that have no similarity to each other. If the difference is identified, we can use this data to examine the epistasis patterns for individual mutations and identify how these patterns change in different environments and how they explain the relationships observed to the entire DFE.

The structure of this chapter is as follows. First, we apply the RB-TnSeq method (described in the previous chapter) to measure the effects of 100 identical mutations in 42 yeast *S. cerevisiae* strains and to evaluate the fitness of these yeast strains in six different environments. Second, we use the obtained fitness measurements for yeast genotypes and estimated selection coefficients of mutations to evaluate the similarity of the selected environments. Third, we examine how the shape of the DFE changes with respect to the background fitness in different environments. Fourth, we examine epistasis patterns for individual mutations in different environments and explain how they contribute to the observed patterns for the entire DFE. Lastly, we discuss our findings in the context of earlier studies in this field and make conclusions.

3.1 Results

3.1.1 Fitness Profiling of the RB-TnSeq Mutants and Yeast Strains

We assayed libraries of hundred insertion mutants in 42 yeast *S. cerevisiae* segregants that were derived from a cross between a laboratory (BY) and a wine strain (RM) [40,41] across six different environments (Fig. 3.1). To ensure sufficient signal difference between strains, we selected a subset of strains from our previous experiment [37], where these strains have been identified to differ in fitness and in robustness to disruptive effects of random mutations (Table S2.1). Each of the selected strains was transformed twice with a distinct random barcode transposon library to create a uniquely labeled set of hundred identical mutants in each genotype. Each mutant was represented by at least four different barcodes (Table S2.1) which allowed us to identify and eliminate mutant lineages with multiple insertions within the same genome or which got adaptive mutations (see Chapter 2). By the design of the mutant libraries, five out of 100 mutations assayed were neutral, and we used their frequencies to estimate the fitness of the selected genotypes and as a reference to infer selection coefficients for the generated mutations.

We performed competition assays in bulk with a 12-hour transfer schedule (see Chapter 2) to measure selection coefficients of the generated mutations and to estimate the fitness of the selected yeast segregants in the exponential phase. Competition assays were done in six environments with temperature (30 ℃ and 37 ℃) and pH (3.2, 5.0, and 7.0) as stress factors. To ensure sufficient representation for each barcode lineage in a competition assay, we divided mutants into three groups and two groups in 30 ℃ and 37 ℃ environments, respectively, based on the fitness measurements for the selected untransformed segregants obtained in our preliminary experiment. Mutants of each segregant were represented in only one group; however, mutants of segregant LK5-G01 were intentionally added to each assay group. The measurements for LK5-G01 were used as an internal control for reproducibility between assays.

We performed independent competition assays for mutants obtained from different transformations to validate the reproducibility of the selection coefficient and background fitness measurements. For each transformation, these measurements were calculated separately. The Pearson's correlation coefficients for the selection coefficients of mutations ranged from 0.60 to 0.95 (Fig. S3.1A - S3.1F) for mutants of most of the segregants after removing the barcode lineages with low read coverage and/or with multiple mutations (see Chapter 2). This enabled us to merge these data and calculate the selection coefficients based on the estimates from both transformations. However, for 37 out of (42 genotypes x 6 environments) 252 total cases, we identified low correlations and excluded these data from the subsequent analysis.

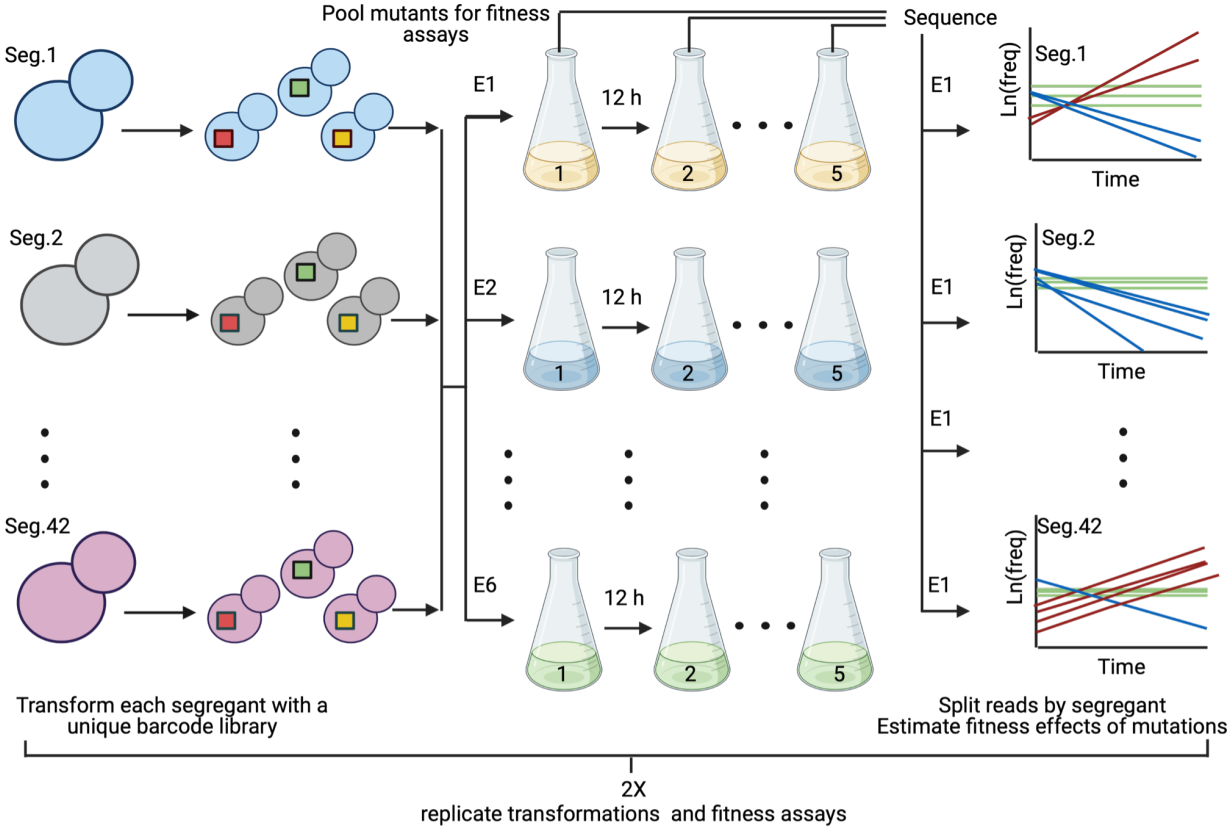

Figure 3.1: Schematic Representation for the Experimental Setup.

To validate the quantification of fitness estimates, we first evaluated the correlation between background fitness measurements from replicate assays. The Pearson's correlation coefficients ranged from 0.86 to 0.99 (Fig. S3.2). Then, we tested for the consistency of assay conditions by comparing the fitness measurements for the control segregant LK5-G01 between assays done with different groups of mutants. There was no significant difference between these estimates (Fig. S3.3).

Next, we tested whether the obtained fitness measurements from assays done in bulk reflect growth rates of segregants measured individually. For that, we conducted individual growth rate measurement assays in two replicates for a set of 15 segregants in 30 ℃ pH 5.0 and 37 ℃ pH 5.0 environments. In general, fitness estimates from the experiment done in bulk reflected the trend observed for growth rate measurements done individually (Fig. S3.4). However, at 30 ℃ pH 5.0 bulk fitness measurements for many segregants were slightly lower than individual measurements, while 37 ℃ pH 5.0 bulk measurements were higher for all tested segregants. We assume that differences in population densities between bulk and individual measurements can account for the observed discrepancies. In 30 ℃ pH 5.0, the OD600 of cultures at the end of the 12-hour growth cycle ranged between 0.30 and 0.70 for different groups of segregants. However, we found that some segregants transition to stationary phase at lower population densities (e.g., OD600=0.63 for the population of LK3-D01 in 37 ℃ pH 5.0) (Fig. S3.5). Therefore, we assume the discrepancy between measurements in 30 ℃ pH 5.0 may reflect the fact that populations of some segregants transitioned to a stationary growth phase before the end of the 12-hour cycle and, consequently, we measured growth not in pure exponential phase in bulk assays. On the other hand, higher population densities in bulk assays could change the composition of the media by accumulating metabolites in higher concentrations. Such differences in the chemical composition of the media could either stimulate or restrain the growth of segregants compared to the conditions in individual assays.

In a separate set of experiments, we confirmed that populations of some segregants exhibited a gradual decline in fitness until they eventually were removed from the assay due to dilution in some environments (Fig. S3.6). Additionally, such decline in fitness was observed in stressful environments before [42]; therefore, it does not indicate flaws in our experimental setup or analysis. We found that the majority of segregants that exhibit fitness decline also had a low correlation between the measurements of selection coefficients between mutations. Therefore, we excluded most of them in the previous step. At this step, we additionally excluded the data for five segregants in environments where they exhibit fitness decline from the subsequent analysis.

In the following parts of the chapter, we use the term “fitness” which represents the growth of a population calculated per hour and is estimated in experiments done in bulk. This parameter is different from the growth rate of a population in an exponential phase.

3.1.2 Selected Environments Differ in the Similarity to Each Other

The six environments were intentionally chosen to differ in the number of simultaneously acting stress factors to create a variation in the degree of similarity for the tested environments. We assume that this variation can lead to a distinction in the gene-environment (GxE) interactions which can manifest in the difference in the fitness effects of mutations between these conditions. We performed all assays in SC (synthetic complete) media, where 30 ℃ pH 5.0 environment supposedly represents no-stress (i.e., favorable) growth conditions, 30 ℃ pH 3.2, 30 ℃ pH 7.0, 37 ℃ pH 5.0 - single-stress and 37 ℃ pH 3.2 and 37 ℃ pH 7.0 - double-stress conditions. We expect the effects of mutations in double-stress environments to differ from the fitness effects in no-stress conditions.

To identify the degree of similarity between the selected environments, we calculated Pearson’s correlation coefficients between fitness estimates for segregants without mutations. The same approach was previously applied in [25,40,43]. Any two environments where

segregants' fitness estimates are negatively correlated represent dissimilar (antagonistic) environments, while conditions in which fitness estimates are positively correlated represent similar (concordant) environments [25]. We found that environments tend to be more similar within the same temperature level, while different pH of the media does not result in the significant dissimilarity between environments (Fig. 3.2). Analogous observations were previously reported for fitness estimates of 1000 segregant strains, 42 of which we used in our experiment, in media of different compositions but at a similar temperature and pH levels as in our study [40]. However, we identified a few pairs of environments that deviated from the above pattern. In particular, pH 7.0 environment was similar to pH 3.2 (r=0.53, P=3.1e-4) and pH 5.0 (r=0.78, P=9.5e-10) environments at 30 ℃, while at 37 ℃ the correlation became insignificant for pairs of this environment with pH 3.2 (r=-0.03, P=0.83) and pH 5.0 environments (r=-0.13, P=0.41), respectively. In addition, pH 7.0 environments exhibited a high correlation between fitness estimates regardless of the temperature level (r=0.44, P=0.003).

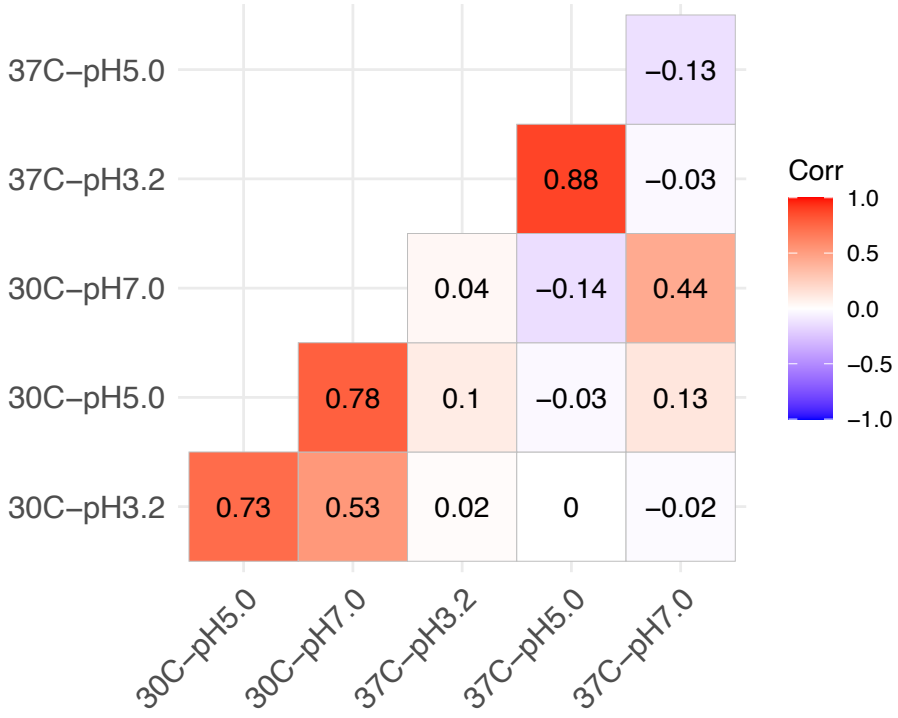

Figure 3.2: Pearson's Correlation Coefficients for Fitness Estimates of Segregants Between Different Pairs of Environments.

Next, we compared the effects of all mutations between six environments to test whether the identified lack of similarity between 30 ℃ and 37 ℃ environments translates into the difference in fitness effects of mutations between these conditions (Fig. 3.3, Table 3.1). We found that mutations have more positive and/or less negative effects on fitness in 30 ℃ pH 3.2 and 30 ℃ pH 5.0 environments. However, the effects of mutations in 30 ℃ pH 7.0 are more similar to the estimates obtained in the 37 ℃ environments, where mutations have more negative effects. Out of two environments with pH 7.0, mutations generally have more negative effects on fitness at 37 ℃. These observations confirm the observations for the similarity for the pairs of environments based on fitness estimates and suggest that 30 ℃ environments differ from 37 ℃ environments.

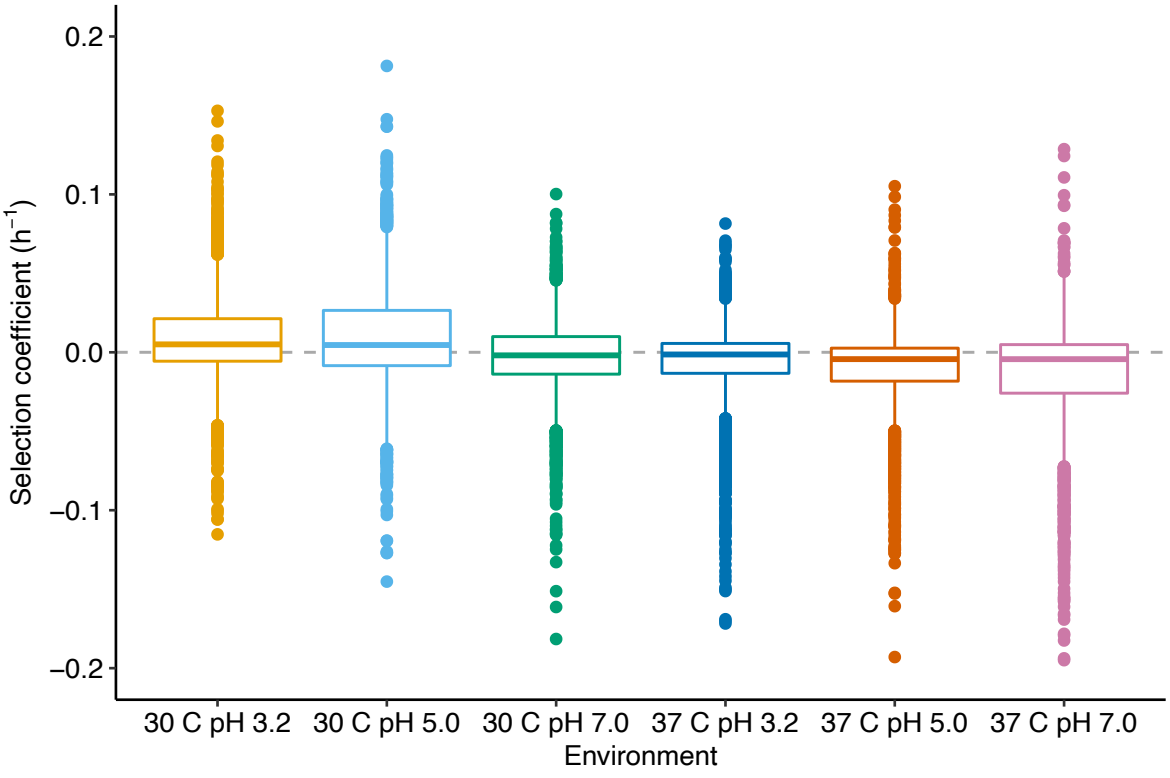

Figure 3.3: Distributions of Fitness Effects of All Tested Mutations in Different Environments.

Table 3.1: Median Values for the Distributions of the Effects of Mutations in Different Environments.

Environment	Number of mutations	Median selection coefficient, %	Bootstrap 95% CI, %
30 ℃ pH 3.2	3268	0.503	[0.422; 0.584]
30 ℃ pH 5.0	3217	0.461	[0.348; 0.573]
30 ℃ pH 7.0	3848	-0.194	[-0.244; -0.144]
37 ℃ pH 3.2	2986	-0.137	[-0.185; -0.089]
37 ℃ pH 5.0	3172	-0.432	[-0.503; -0.361]
37 ℃ pH 7.0	2363	-0.440	[-0.530; -0.350]

The observation that many mutations became beneficial in 30 ℃ environments might be surprising given that we measured effects of randomly generated mutations and not of those that were initially identified to be beneficial in evolutionary experiments and then were reconstructed in different genetic backgrounds, as it was done in studies described in [9,30,33–35,44]. However, 100 mutations tested in this study are not completely random. These mutations were selected from a big set of random mutations because they had non-neutral fitness effects across many genetic backgrounds in favorable growth conditions for yeast [37]. Additionally, as we identified above, 30 ℃ and 37 ℃ environments are not similar to each other. Therefore, some of the mutations can have antagonistic effects between these conditions [25,45]. Lastly, we confirmed fitness estimates for a set of mutations that switched the sign of their fitness effect between 30 ℃ and 37 ℃ environments in a separate experiment (Fig. S3.7). Thus, the shift towards more positive fitness effects at 30 ℃ is not an artifact of our experimental design or analysis.

In this part, we showed that the selected environments vary in the degree of similarity to each other which is reflected in the difference in the correlations between fitness estimates for segregants and in the difference in prevalent sign of the fitness effects of mutations. This experimental setup provides a great opportunity to compare how epistasis changes across environments. If we find consistent epistasis patterns even among dissimilar environments, that

will suggest that these patterns generally persist regardless of the environment type. On the other hand, epistasis patterns can hold only in very similar environments; therefore, comparing the patterns across environments with various degrees of similarity would reveal how different two environments should be to change epistasis patterns. Additionally, this setup allows us to test whether patterns of global epistasis change for beneficial and deleterious mutations.

3.1.3 DFEs Have Different Shapes in Dissimilar Environments

To test whether epistasis changes across environments, we examined how the shape of the distributions of fitness effects (DFEs) of mutations change with respect to the fitness of their genetic background. In particular, we focused on the relationships between the mean and the variance of the DFE and the fitness of the genetic background.

We identified that the negative relationship between the mean of the DFE and the fitness of the genetic background is consistent regardless of the environment (Fig. 3.4A, Table 3.2). These data also confirm the observations by Johnson et al. in YPD media [37] for deleterious mutations. In addition, these relationships are not affected by the prevalent sign of the fitness effects of mutations. In particular, the slopes of the regression lines are very similar to each other in 30 ℃ pH 5.0 and 37 ℃ pH 5.0 environments (Table 3.2), where mutations have more positive and more negative effects, respectively.

However, we found that the regression lines representing different environments are shifted relative to each other. Because the range of fitness values of the segregants varies among six environments (Fig. S3.8), the observed shifts could be associated with the rise or the decline in the fitness of segregants between environments. To test that, we compared the x-axis intercepts for the regression lines describing the relationship between the mean of the DFE and background fitness in different environments. These values correspond to the fitness of the genetic background at which mutations have a neutral effect on fitness (F0). If shifts result from

Figure 3.4: Relationship Between the Mean and the Variance of the DFE and Fitness of the Genetic Background in Different Environments.

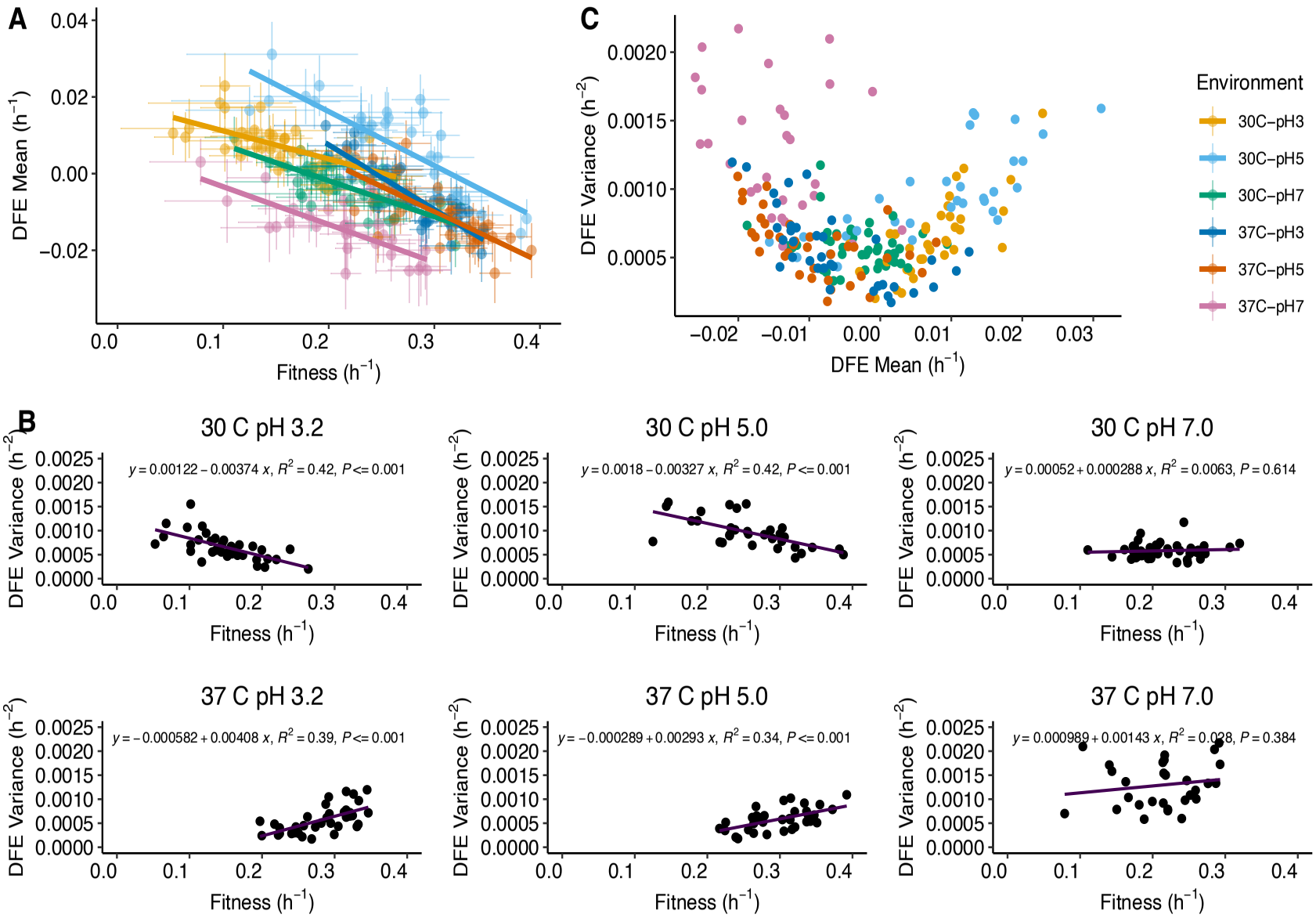
A
DFE Mean (h^{-1})
Fitness (h^{-1})
C
DFE Variance (h^{-2})
DFE Mean (h^{-1})
Environment
30C–pH3
30C–pH5
30C–pH7
37C–pH3
37C–pH5
37C–pH7
B
30 C pH 3.2
$y = 0.00122 - 0.00374\,x$, $R^2 = 0.42$, $P <= 0.001$
30 C pH 5.0
$y = 0.0018 - 0.00327\,x$, $R^2 = 0.42$, $P <= 0.001$
30 C pH 7.0
$y = 0.00052 + 0.000288\,x$, $R^2 = 0.0063$, $P = 0.614$
37 C pH 3.2
$y = -0.000582 + 0.00408\,x$, $R^2 = 0.39$, $P <= 0.001$
37 C pH 5.0
$y = -0.000289 + 0.00293\,x$, $R^2 = 0.34$, $P <= 0.001$
37 C pH 7.0
$y = 0.000989 + 0.00143\,x$, $R^2 = 0.028$, $P = 0.384$

Table 3.2: Characteristics of the Regression Lines for the Relationships Between the Mean of the DFE and the Background Fitness in Different Environments.

Environment	Slope; Bootstrap 95% CI	F0; Bootstrap 95% CI, h^{-1}
30 ℃ pH 3.2	-0.073; [-0.110, -0.036]	0.252; [0.244, 0.259]
30 ℃ pH 5.0	-0.141; [-0.178, -0.103]	0.314; [0.303; 0.326]
30 ℃ pH 7.0	-0.093; [-0.122; -0.064]	0.180; [0.173; 0.187]
37 ℃ pH 3.2	-0.168; [-0.201, -0.135]	0.244; [0.234; 0.254]
37 ℃ pH 5.0	-0.131; [-0.167, -0.094]	0.224; [0.212; 0.236]
37 ℃ pH 7.0	-0.098; [-0.128, -0.068]	0.065; [0.058; 0.072]

comparing the values of DFE mean at different intervals of background fitness, we expect to see similar F0 values for all environments. However, we found that F0 values vary between environments, e.g., 30 ℃ pH 5.0 and 37 ℃ pH 5.0 (Table 3.2). This suggests that the observed shifts at least partially result from a form of gene-gene-environment (GxGxE) interactions that are distinct from diminishing-returns or increasing cost forms of epistasis (GxG) between more- or less-fit genotypes.

Next, we examined the relationship between the variance of the DFE and the background fitness (Fig. 3.4B). The type of these relationships was determined by the prevalent sign of the fitness effects of mutations. In the 37 ℃ environments, where deleterious mutations were prevalent, we found a positive relationship between these parameters which is also consistent with previous observations for the DFE of random mutations in YPD media [37]. In environments with many beneficial mutations, e.g., 30 ℃ pH 5.0 and 30 ℃ pH 3.2, this relationship turned out to be negative. In pH 7.0 environments, there was no relationship between these parameters. The lack of correlation at 37 ℃ pH 7.0 is largely caused by the insufficient power of our experiment in that environment. We removed data for 15 out of 42 segregants because these segregants had a very low fitness and, consequently, low read coverage at the sequencing step and low

correlation between selection coefficient estimates between transformations (Fig. S3.1A-S3.1F). In addition, at 37 ℃ pH 7.0, the populations of the remaining segregants had the lowest fitness out of all six environments. Their DFEs were represented by a lower number of distinct mutants (68 vs 82 in other environments, on average) and were dominated by mutant lineages with more benign fitness defects as genotypes with strong deleterious mutations were removed due to dilution. Therefore, these data are not sufficient to make valid conclusions about the relationship between the variance of the DFE and background fitness in 37 ℃ pH 7.0 environment.

The relationship between the mean and the variance of the DFE has a U-shape based on the measurements across all environments (Fig. 3.4C). The observation that the change in the variance patterns corresponds to the switch in the sign of the mean of the DFE suggests that the difference in the observed patterns is largely driven by mutations changing their sign from positive to negative or vice versa. Our previous observations that mutations change their effects between more dissimilar environments (e.g., 30 ℃ and 37 ℃) and that patterns observed for variance remain consistent between pairs of similar environments also support that hypothesis. We examine this in detail in the next section.

In this part, we showed that mutations have lower effects on fitness in more-fit genetic backgrounds regardless of the environment type. This suggests that the fitness effects of at least some mutations can be determined by the fitness of their genetic background and not the genotype itself. We found evidence of a special type of GxGxE interactions which modifies the F0 values between environments. We also showed that the shape of the DFE for less- and more- fit genotypes varies between environments that are not similar to each other. That suggests that a single-step mutational neighborhood of a genotype changes in different environments. At 37 ℃, more-fit genotypes are located at a local fitness peak. They have many deleterious mutations in their single-step mutational neighborhood, as shown by their DFEs with lower mean and higher variance. At 30 ℃, more-fit segregants reside on a fitness plateau, as the variance of their DFEs is low and the mean is close to zero.

3.1.4 Various Types of Gene-Gene and Gene-Environment Interactions Account for the Differences in the Shape of the DFE Across Environments

Since DFE consists of mutations, we next examined how epistasis patterns for individual mutations change across environments to understand what drives the observed differences in the shape of the DFE.

We did the analysis of variance (one-way ANOVA) and fit linear regression models to the relationships between the effect of individual mutations and fitness of their genetic background in different environments to identify mutations that have distinct effects across environments and mutations effects of which change with respect to the background fitness. We identified 21 out of 100 mutations with nearly constant effects, as their fitness effects did not differ across environments and did not depend on the fitness of their genetic background (significance level, α=0.05) (e.g., nearby MET4) (Fig. 3.5, Fig. S3.9A-S3.9E). There were 9 out of 100 mutations for which we obtained measurements in only a small proportion of segregants (e.g., in UBP3 Fig. S3.9A-S3.9E); therefore, we couldn't confidently classify these mutations and excluded them. Among 70 mutations with distinct effects, most mutations had significantly different effects between 30 ℃ and 37 ℃ environments, while only a low fraction of mutations showed different fitness effects within the same temperature level. For example, fitness effects of 45 mutations differed significantly between 37 ℃ pH 5.0 and 30 ℃ pH 5.0 environments, while effects of only 11 mutations differed between 37 ℃ pH 5.0 and 37 ℃ pH 3.2 (Tukey's HSD test; significance level, α=0.05). Most mutations changed the sign of their fitness effect from positive at 30 ℃ to neutral or negative at 37 ℃ (e.g., in OCA1); however, we observed the opposite trend for a few mutations (e.g., in SIR3 and SIR4). We identified 23 out of 70 mutations whose effects did not depend on the background fitness in any of the six environments (e.g., in UBP1). The effects of the remaining 47 mutations changed either negatively (42 mutations, e.g., in LAT1) or positively (5 mutations, e.g., in SIR3) with respect to the background fitness in most of the environments.

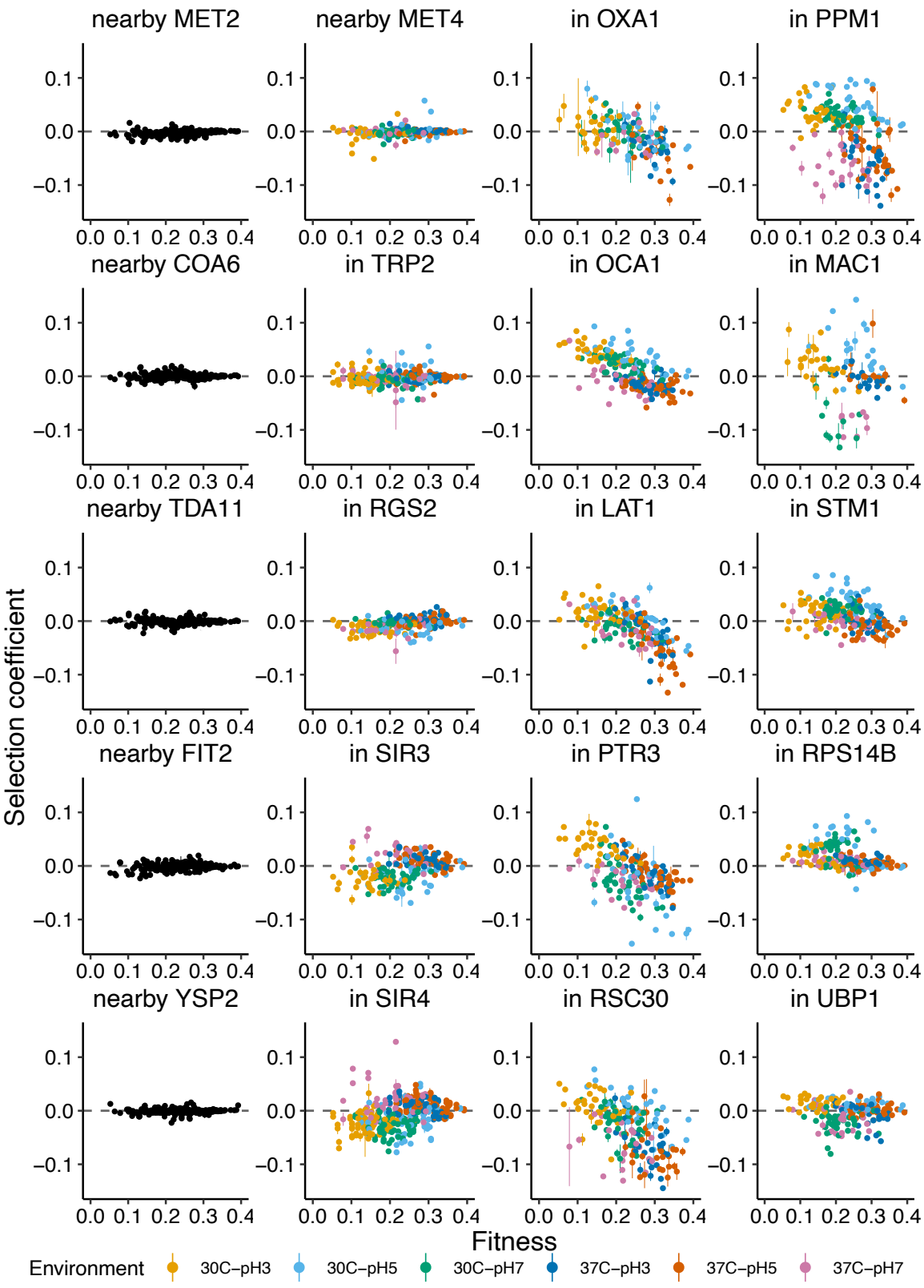

Figure 3.5: Epistasis Patterns for Individual Mutations in Different Environments. Mutations shown in black were used as a neutral reference.

The slope of the regression lines describing the relationship between the effect of these 47 mutations and fitness of their genetic background (and also showing the magnitude of global epistasis affecting the mutated locus [38]) differed among mutations (Fig. S3.10).

Next, we examined how the patterns of global epistasis change for 47 mutations across environments. As we mentioned above, we did not observe a statistically significant relationship between the effects of a mutation and its genetic background for 47 mutations in all six environments. Because the span of range of segregants' fitness changed between environments (Fig. S3.8) and some mutations were subject to a weaker degree of global epistasis, the lack of significant correlations for some mutations in some environments could stem from us measuring weak relationships over a narrow range of fitness values. To test that, an additional experiment is needed where we can measure the effects of these 47 mutations in segregants that span over a wide fitness range in many different environments. With these data, we cannot make strong conclusions about what causes the lack of the relationship observed for some of the mutations in some environments.

Next, we tested how the average magnitude of global epistasis affecting these 47 loci changes across environments by comparing the distributions of the slopes calculated for individual mutations in different environments (Fig. S3.10). We found that the mean values of these distributions were similar (Table. 3.3). However, the average magnitudes of global epistasis in 37 ℃ pH 5.0 and 30 ℃ pH 5.0 environments were more negative than in the rest of the environments.

We further fit four different models to the relationship between the fitness effect of a mutation and its genetic background to identify how exactly different environments affect these relationships for individual mutations (Fig. 3.6). Model 1 that assumed that both slope and the intercept of the relationship remain constant regardless of the environment type had the lowest explanatory power of the four models for any mutation (16 % on average). The highest variance

Table 3.3: Mean Values for the Distributions of Slopes for Mutations Effects of Which Scale With Respect to Their Background Fitness.

Environment	Mean; Bootstrap 95% CI
30 ℃ pH 3.2	-0.124; [-0.168, -0.080]
30 ℃ pH 5.0	-0.196; [-0.238, -0.154]
30 ℃ pH 7.0	-0.119; [-0.188, -0.050]
37 ℃ pH 3.2	-0.186; [-0.275, -0.096]
37 ℃ pH 5.0	-0.220; [-0.273, -0.165]
37 ℃ pH 7.0	-0.108; [-0.175, -0.041]

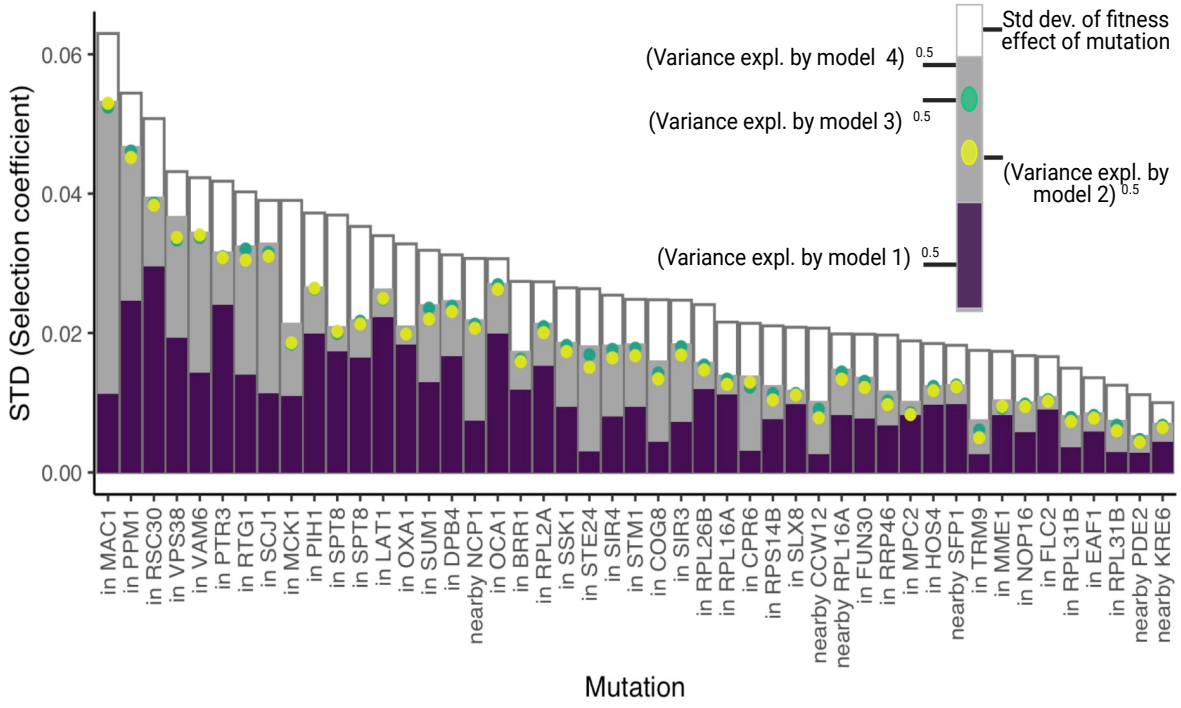

Figure 3.6: Fractions of Variance Explained by Each of the Four Tested Models. For each mutation, the standard deviation of the fitness effect across segregants and the square root of the variance explained by each of the four models are shown.

fraction was explained by model 4 assuming that environments change both parameters (48 % on average). Models assuming that environments modify only the slope (model 2) or only the intercept (model 3) had similar explanatory power to model 4 (on average 43 % and 40 %, respectively). For 24 out of 47 mutations, model 3 had higher explanatory power by 6.3 % on average (Fig. 3.6; Fig. S3.9A- S3.9E), while for the remaining mutations models 2 and 3 explained similar fractions of variance. This suggests that for 24 mutations, the degree of global epistasis did not significantly change across environments; however, the environment modified the F0 (fitness of the background genotype at which a mutation has a neutral effect on fitness) by shifting the intercept of the regression lines along the y-axis.

In this part, we examined how epistasis patterns of individual mutations change across environments. We observed various types of interactions between mutations and environments, ranging from mutations having nearly constant effects across different genetic backgrounds to differences in F0 for mutations that are subject to global epistasis. We identified 47 out of 100 mutations that were subject to global epistasis. Although the magnitude of global epistasis affecting individual mutated loci varied, its mean value based on slopes of 47 mutations was not significantly different between the environments. That explains the consistency of slopes observed for the mean of the DFE (Fig. 3.4A).

Most of the 47 mutations had positive effects on fitness in genotypes with low fitness, were neutral at the intermediate fitness levels, and became deleterious in genotypes with high fitness. In addition to that, the changes in the sign of the fitness effect for these mutations usually coincided with the transition from 30 ℃ to 37 ℃ environments, where many segregants exhibited the highest fitness. We also observed antagonistic effects for most mutations which effects did not depend on fitness between pairs of 30 ℃ and 37 ℃ environments. These findings explain the changes in the patterns for the variance of the DFE between these environments (Fig. 3.4B and 3.4C).

We identified 23 mutations effects of which did not depend on the background fitness but changed across environments. This at least partially accounts for the shifts observed for the relationships between the mean of the DFE and background fitness in Fig. 3.4A. Additionally, we identified 24 mutations that were subject to global epistasis, and for which environments changed the relationship between their fitness effect and fitness of their genetic background by modifying the F0. Although not prevalent among mutations, this type of gene-gene-environment interactions (GxGxE) could also contribute to the observed shifts in Fig. 3.4A.

3.2 Discussion

In this study, we systematically compared the single-step mutational neighborhood of 42 yeast *S. cerevisiae* segregants in six different environments. We found that DFEs of more-fit genotypes are represented by mutations with less beneficial and/or more deleterious effects on fitness regardless of the environment type. However, these relationships are subject to a specific type of GxGxE and GxE interactions which modify the effects of mutations and change the intercept of the regression lines describing these relationships in different environments. We observed three different patterns for the relationship between the variance of the DFE and the fitness of the genetic background. These patterns are consistent among highly similar environments and change to the opposite for pairs of antagonistic environments, where most mutations change either the magnitude or the sign of their fitness effects. We identified various GxE, GxG, and GxGxE interactions for individual mutations which constitute the DFE and explained how these interactions contribute to the observed patterns for the mean and the variance of the DFE in different environments.

The results of this project confirm the observations of our previous study where we compared the effects of the same set of mutations across multiple genotypes in YPD media [37]. While we report similar relationships for the mean and the variance of the DFE and the

background fitness in environments where most mutations have negative effects, our study expands these findings by showing that the relationship between the variance of the DFE is determined by the prevalent sign of the fitness effects of mutations and changes between environments with no similarity. In addition, we show that the relationship between the mean of the DFE and the background fitness is consistently negative and is not affected by the type of environment; however, the fitness of genetic backgrounds at which most mutations have neutral effects on fitness (F0) is modulated by the environment.

These patterns are largely driven by 47 out of 100 mutations effects of which change with background fitness (i.e., are subject to global epistasis). Similar to [38], we show that individual mutations are affected by various degrees of global epistasis, as shown by the difference in the strength of the relationship between fitness effect and fitness of genetic background for individual mutations. We also report that these mutations generally exhibit consistent patterns of global epistasis regardless of the environment type and the sign of the fitness effect of mutations. However, the magnitude of the relationships between the effect of a mutation and the fitness of its genetic background varied for some mutations in some of the environments. We assume that such variation is a consequence of us measuring weak relationships over a narrow range of background fitness values in some of the environments; in particular, 37 ℃ pH 3.2 and 37 ℃ pH 5.0. We discuss the limitations of our study later in this section.

We report 42 and 5 mutations effects of which changed negatively and positively with respect to background fitness, respectively. Interestingly, most of these mutations changed the sign of their fitness effect between more- and less-fit genotypes. They had positive effects in less-fit backgrounds, were neutral in the genotypes of intermediate fitness, and became deleterious in more-fit genotypes (opposite for mutations that scaled positively with background fitness). Our observations align with the findings for other mutations in earlier studies. In particular, the effects of *hfg* mutations affecting small RNA regulation in *E. coli* have been shown to be remarkably sensitive to the growth rate. In the same genetic background, these mutations were beneficial in

populations with slow growth, but became neutral or deleterious in populations growing faster [46]. The consistent negative relationship between fitness effects of mutations and their genetic background was also observed for four of the adaptive mutations that were identified in the long-term evolutionary experiment with *E.coli* [26,9,33,34]. The same relationship was observed for some of the beneficial mutations from the evolutionary experiment with *Methylobacterium extorquens* in media with methanol a sole carbon source [35,44], for natural genetic polymorphisms in yeast *S.cerevisiae* in 47 different environments [36], for beneficial mutations reconstructed in yeast of various fitness [30], and in our previous experiment in which we measured the effects of random mutations in yeast genotypes of various fitness [37]. The evidence for a positive relationship is currently limited to mutations enhancing cobalt uptake in *Methylobacterium sp*. [47], for *pkyF* mutations in experiments done with different organisms [9,33,44,34], and for nine mutations in our previous study [37]**.** Together, these observations suggest that different mutations can be subject to global epistasis, these mutations are common, the patterns remain consistent across environments and do not change when mutations occur in a genetic background that belongs to a different species.

Our observation that most of the 47 mutations change the sign of their fitness effect allows us to rule out some of the models that were developed to explain the negative relationship between the effects of mutations and the fitness of their genetic background. Specifically, the modular life model explains diminishing returns epistasis for beneficial mutations but fails to explain increasing-cost epistasis when mutations become deleterious [36]. Similarly, metabolic control theory predicts weaker deleterious effects for mutations in already disrupted biological pathways, assuming that fitness correlates with the flux through that pathway [7]; but it does not explain diminishing returns epistasis for beneficial mutations.

However, several existing models explain both the diminishing-returns and increasing cost epistasis types. The recently proposed connectedness model [38] and idiosyncratic epistasis theory [39] demonstrate that on an idiosyncratic fitness landscape, the fitness difference between

a neighboring genotype and the focal genotype is expected to be less positive or more negative for more-fit genotypes; therefore, suggest that these types of epistasis are a consequence of widespread epistatic interactions between individual genes. A mechanistic cost-benefit model developed by Chou et al. and optimized by Chiu et al. explains both types of epistasis by describing the underlying molecular mechanisms that cause mutations having different effects depending on the fitness of their genetic background [48,49]. However, none of the three models makes any predictions on how the fitness level of genetic background at which mutations become neutral (F0) change in different environments. We identified 24 mutations that can be subject to GxGxE interactions that change the F0 for these mutations in different environments. Similar environment-induced shifts were reported in [50] and for beneficial mutations in [36], where they were attributed to the negative relationship between the quality of the environment and selection coefficients of mutations. Due to the limitations of our study, which we discuss later in this section, we cannot make any conclusions whether our data confirm that relationship. However, the evidence of such shifts in our study and in several other studies suggests that the existing models can be further improved to account for the environment-related changes in F0.

Our findings that the shape of the DFE changes between environments with no similarity (Fig. 3.4) together with observations that segregants change their fitness across environments (Fig. 3.2, similar observations were described in [40,43]) suggest that the single-step mutational neighborhood of a genotype changes upon an environmental shift. This implies that populations located on a local fitness peak and have no beneficial mutations available can move to a different area of the fitness landscape where they can regain access to beneficial mutations in a different environment [14]. These findings have important implications for the selection of combinations of drugs used for the treatment of infectious disease, as administration of drug pairs that result in the change of the DFE can facilitate the evolution of multiple drug resistance [51]. However, we show that the changes in the DFE shape across environments are largely driven by mutations that are either subject to global epistasis or to GxE interactions that result in these mutations

having antagonistic effects on fitness in dissimilar environments. A recent study by Chen and Zhang showed that a large fraction of mutations with antagonistic effects are being purged from populations evolving in fluctuating dissimilar environments [25]. Therefore, further research is needed to understand how the evolvability of genotypes changes upon environmental shifts.

3.3 Limitations of This Study

Our study has a few limitations. First, we failed to measure the absolute fitness of segregants; therefore, we cannot fully compare the relationships observed for mutations that are subject to global epistasis to the relationships obtained from other studies which used absolute fitness, e.g. [46]. Second, the fitness of the selected segregants spun over a narrow fitness range in some of the environments. Comparing the observed relationship over a larger range would allow us to make stronger conclusions, for example about the consistency of epistasis patterns across environments for individual mutations. Moreover, the fitness range of segregants differed significantly between 30°C and 37 ℃ environments, where most mutations had opposite effects on fitness. Therefore, we cannot evaluate the contribution made by environment type and by background fitness to the observed switch in the fitness effects of mutations that are subject to global epistasis. Additionally, we identified that some mutations can be subject to GxGXE interactions. Comparing their effects in genotypes that spun over the same large fitness range in many different environments can help to make stronger conclusions about these mutations. Lastly, the set of mutations tested in our experiment can be biased towards mutations that are subject to global epistasis, as they were selected in the previous experiment from a pool of more than 1100 mutations because they had significant effects on fitness in a large number of genetic backgrounds [37]. It is likely that the frequency of such mutations available to adapting populations can be different than in our experiment; therefore, one can expect to get relationships

of lower strength between the mean of the DFE and background fitness in experiments done with random sets of mutations.

3.4 Conclusions

The goal of this study was to examine how epistasis patterns change for the entire DFE within single-step mutational proximity of a focal genotype and for individual mutations with respect to background fitness in different environments. We showed that the DFEs of more-fit genotypes are represented by mutations with less beneficial and/or more deleterious effects regardless of the environment type. However, the shape of the DFE changes between dissimilar environments depending on the prevalent sign of the fitness effects of mutations. Although various types of GxE, GxG, GxGxE interactions contribute to it, these patterns are largely driven by 47 mutations that are subject to global epistasis which magnitude mostly remains similar regardless of the environment type. For some of these mutations, environments seem to change the fitness of the genetic background at which mutations switch the sign of their fitness effect. Our results suggest that a genotype can be located on a local fitness peak with no beneficial mutations available in one environment and transition to a bottom of a fitness peak with many beneficial mutations available in a different environment. This has important implications, for example, in the selection of the combinations of drugs used for the treatment of infectious diseases. The knowledge of conditions in which a mutational neighborhood of a genotype changes can help to prevent the evolution of multiple drug resistance in infectious strains. However, we show that such changes are driven by mutations with antagonistic effects; therefore, it is unclear whether these mutations will reach fixation or be purged from the evolving populations. Further research is needed to answer this question.

3.5 Acknowledgements

Chapter 3, in part, currently being prepared for submission for publication of the material. Alena Martsul, Milo Johnson, and Sergey Kryazhimskiy. The dissertation author was the primary investigator and the author of this material.

3.6 Supplementary Materials

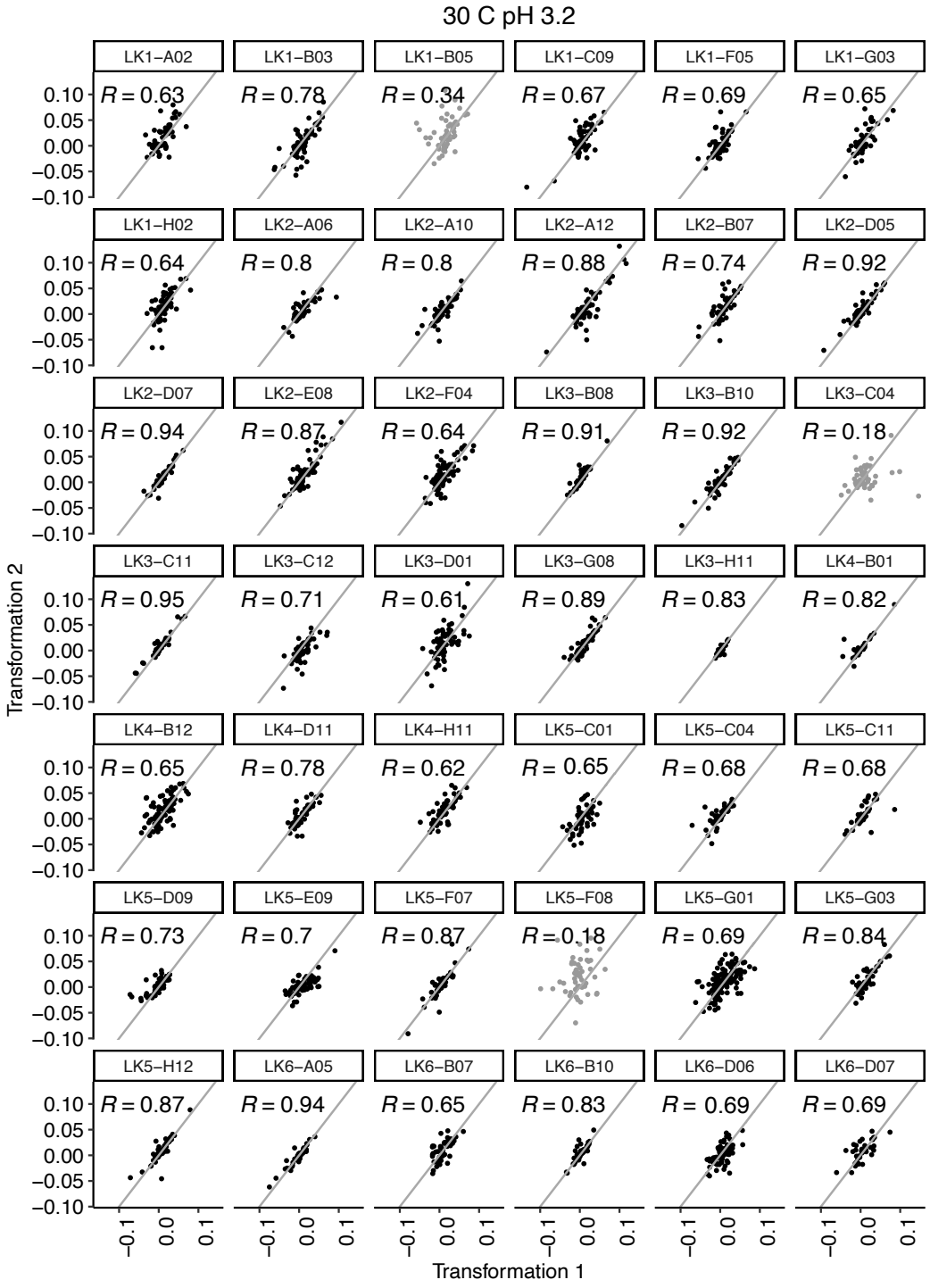

Figure S3.1A: Pearson's Correlation Coefficients for Estimates of Selection Coefficients of Mutations From Two Different Transformations in 30°C pH 3.2. Data for segregants shown in grey were excluded from the analysis.

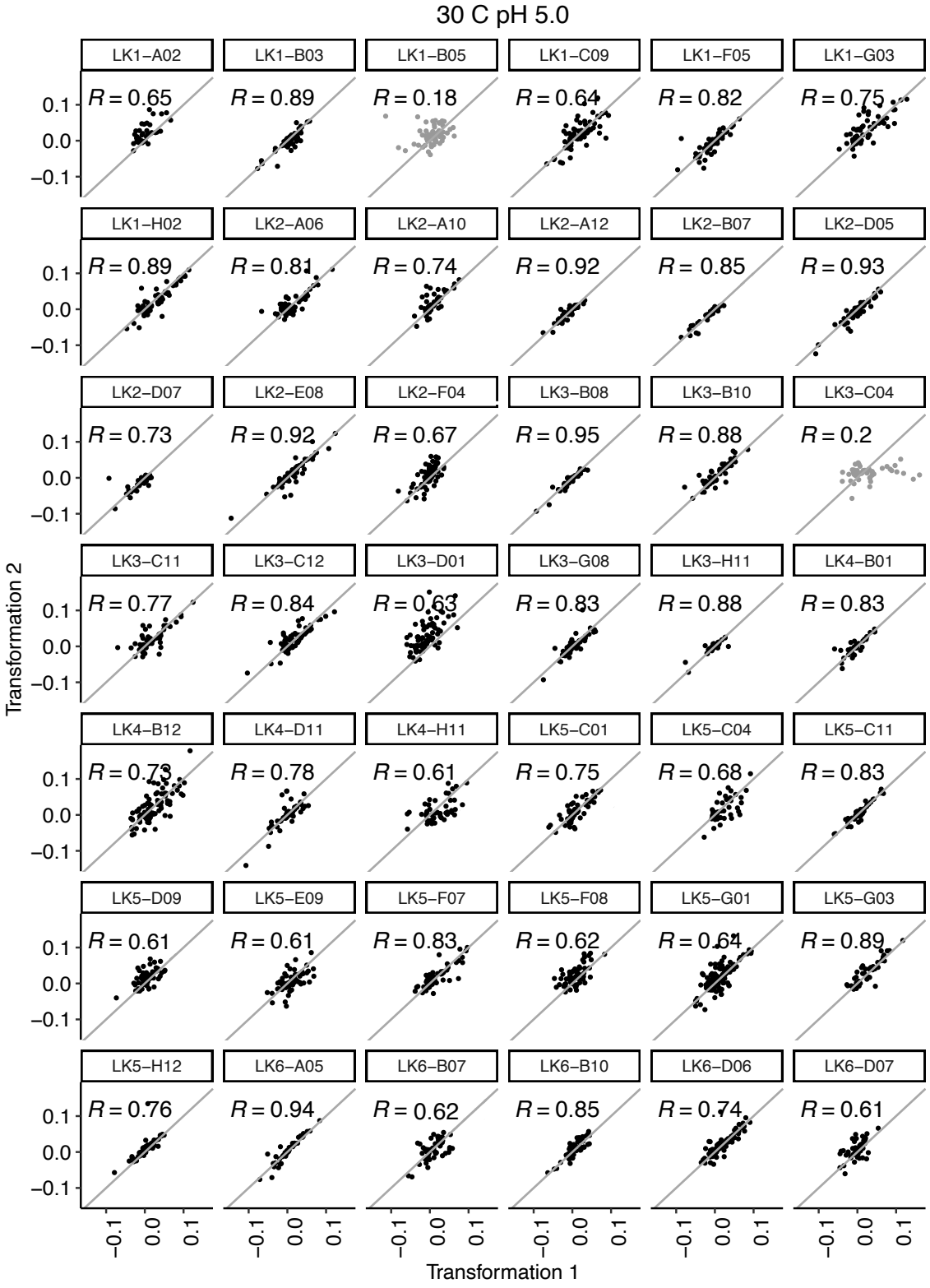

Figure S3.1B: Pearson's Correlation Coefficients for Estimates of Selection Coefficients of Mutations From Two Different Transformations in 30°C pH 5.0. Data for segregants shown in grey were excluded from the analysis.

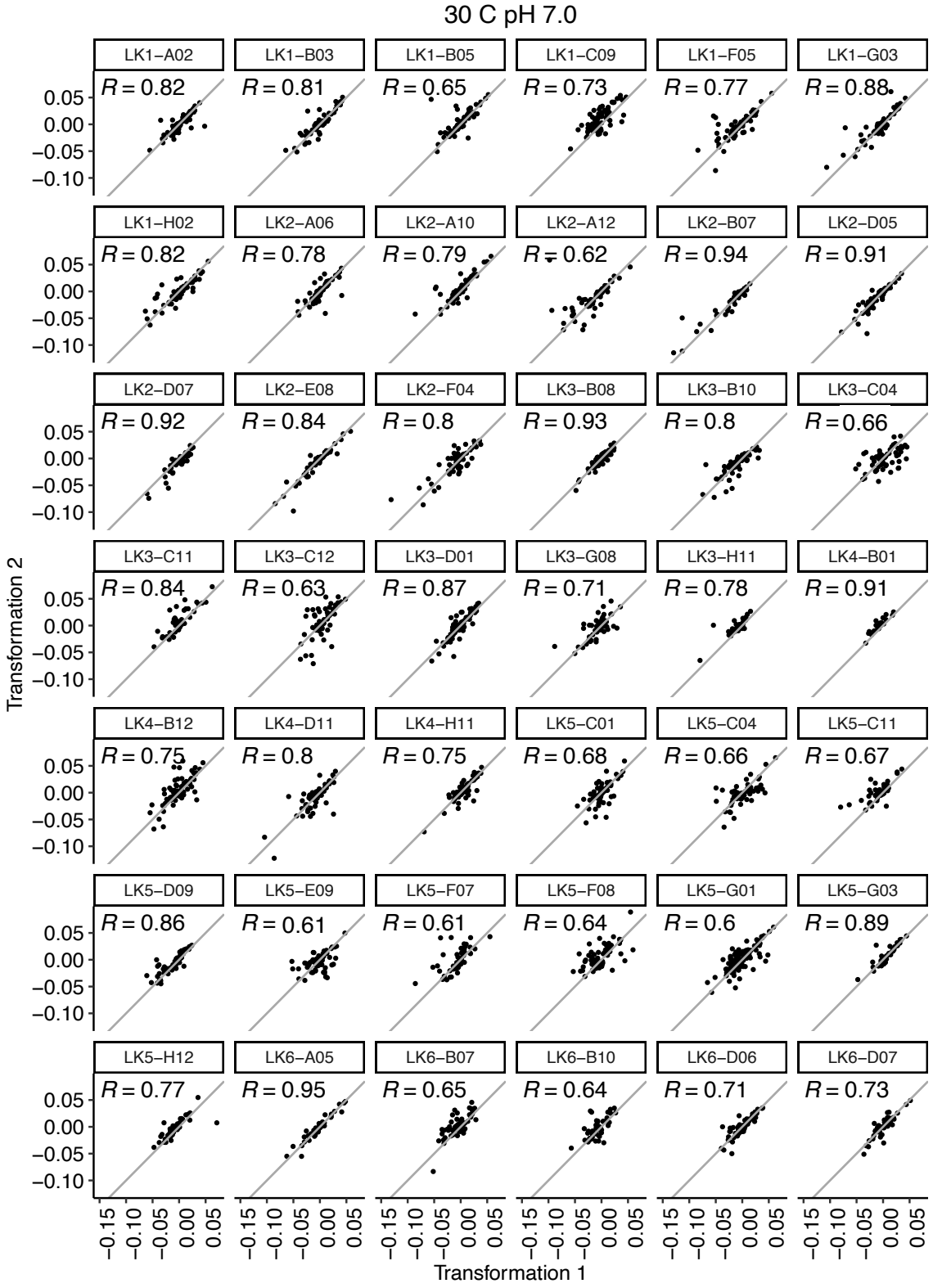

Figure S3.1C: Pearson's Correlation Coefficients for Estimates of Selection Coefficients of Mutations From Two Different Transformations in 30°C pH 7.0. Data for segregants shown in grey were excluded from the analysis.

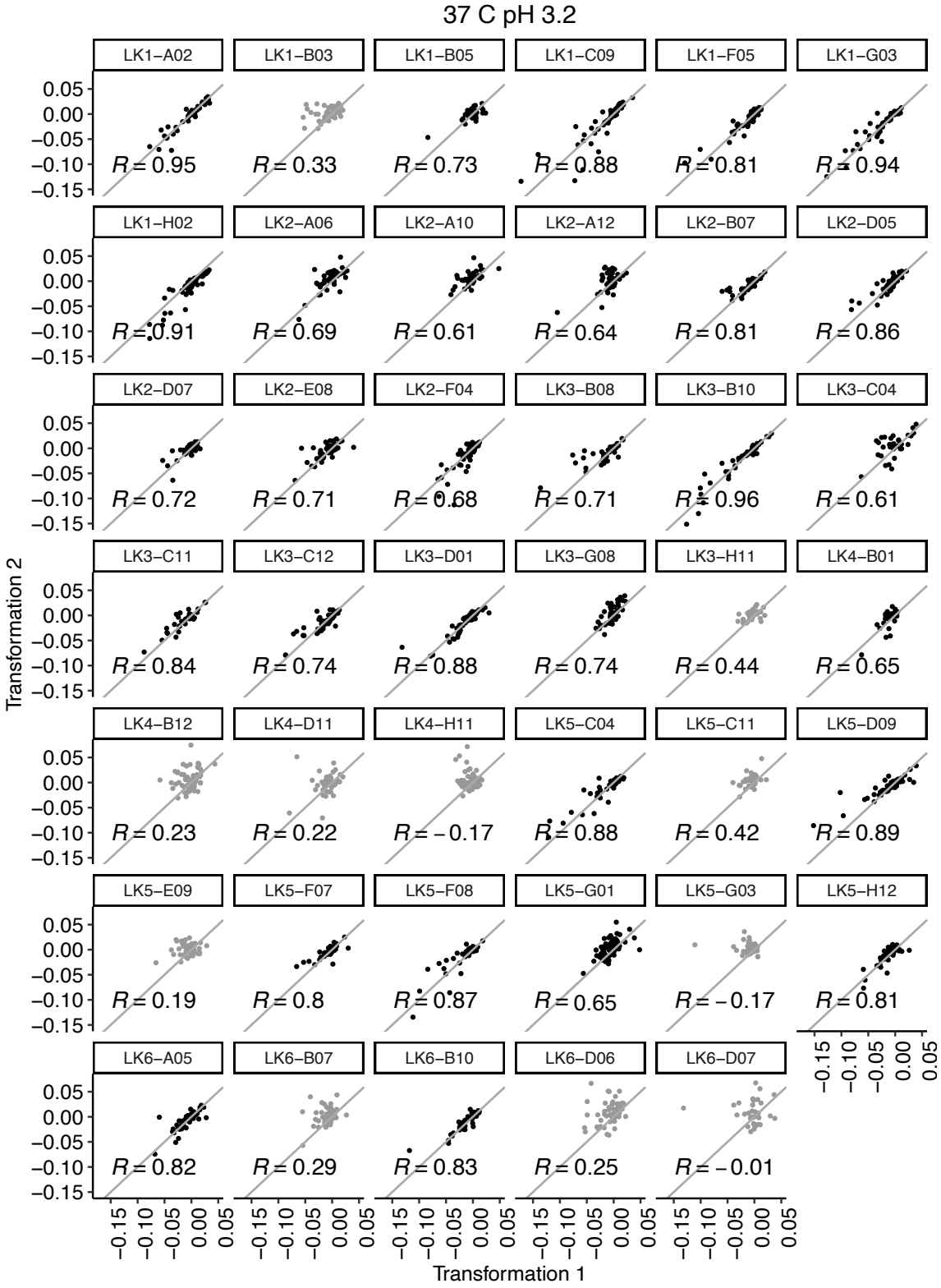

Figure S3.1D: Pearson's Correlation Coefficients for Estimates of Selection Coefficients of Mutations From Two Different Transformations in 37°C pH 3.2. Data for segregants shown in grey were excluded from the analysis.

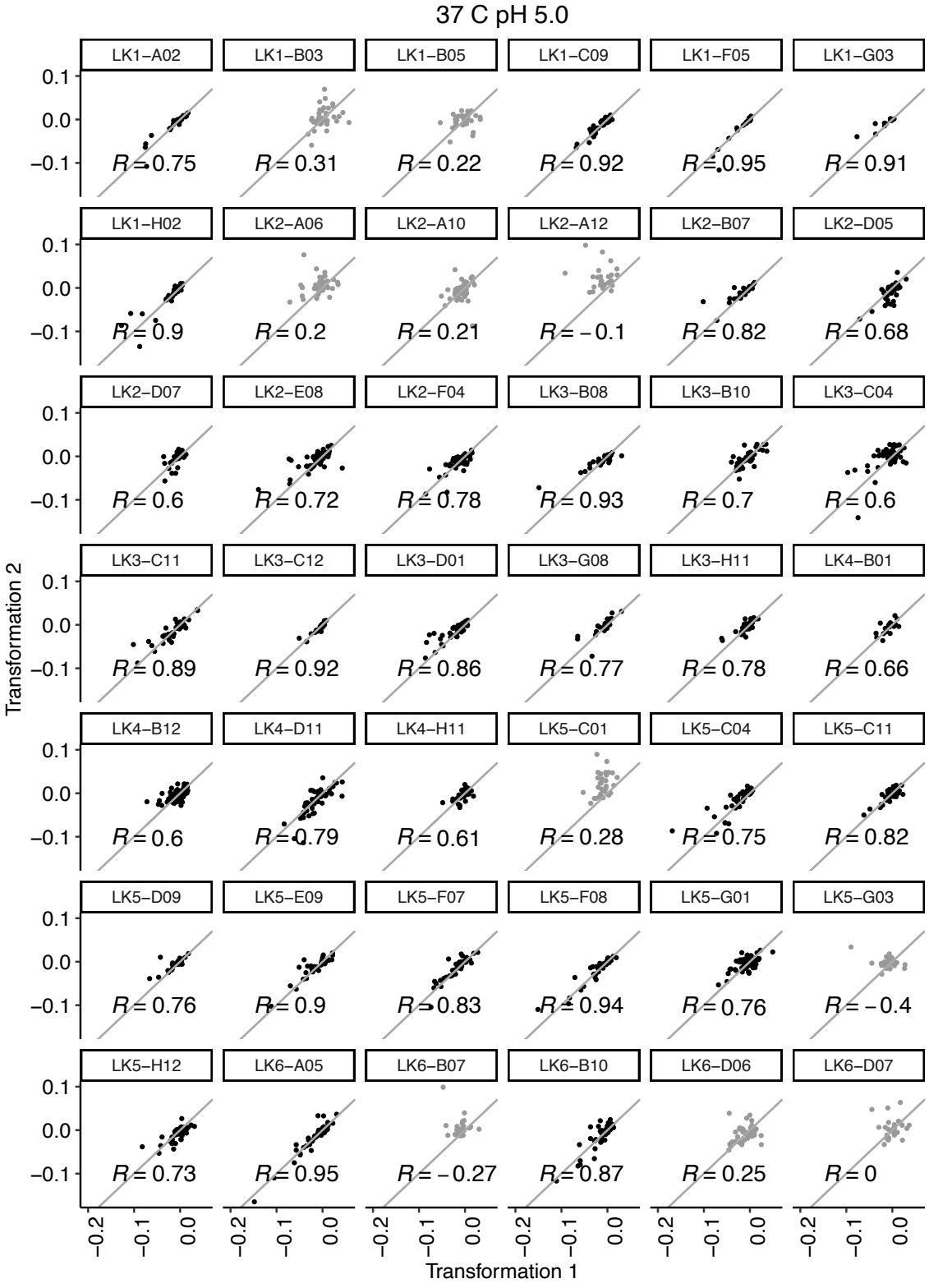

Figure S3.1E: Pearson's Correlation Coefficients for Estimates of Selection Coefficients of Mutations From Two Different Transformations in 37°C pH 5.0. Data for segregants shown in grey were excluded from the analysis.

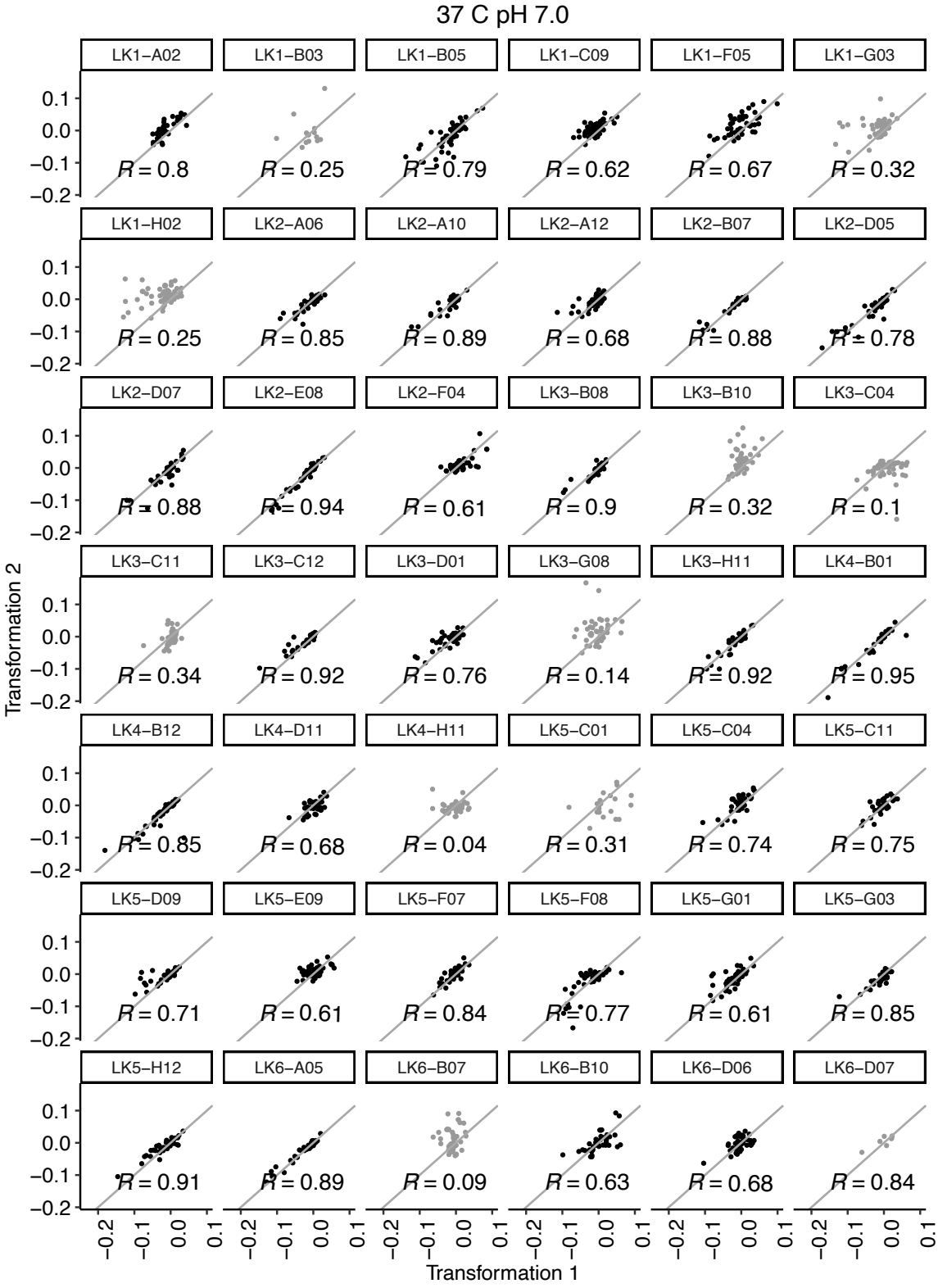

Figure S3.1F: Pearson's Correlation Coefficients for Estimates of Selection Coefficients of Mutations From Two Different Transformations in 37°C pH 7.0. Data for segregants shown in grey were excluded from the analysis.

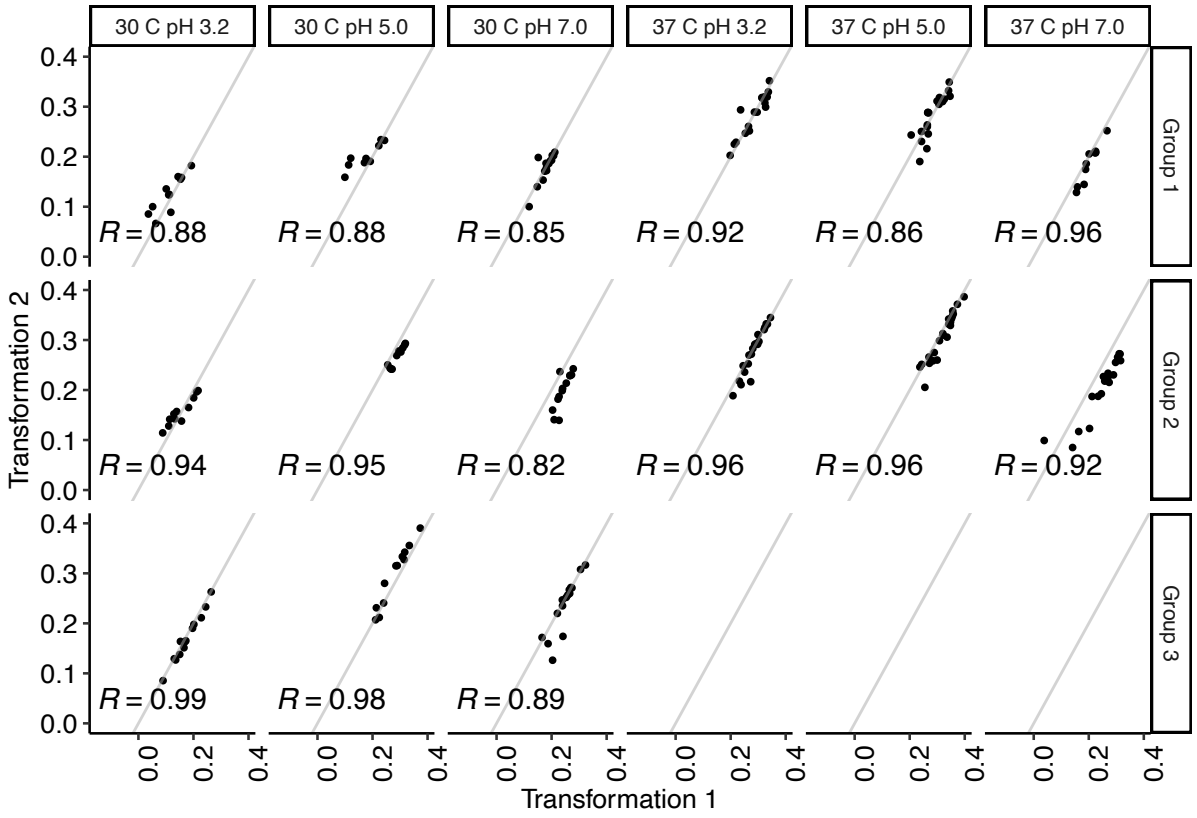

Figure S3.2: Pearson's Correlation Coefficients for Fitness Estimates of Segregants in Six Environments. Segregants were divided into three groups in 30°C environments and into two groups in 37°C environments based on their fitness estimates from our preliminary experiment.

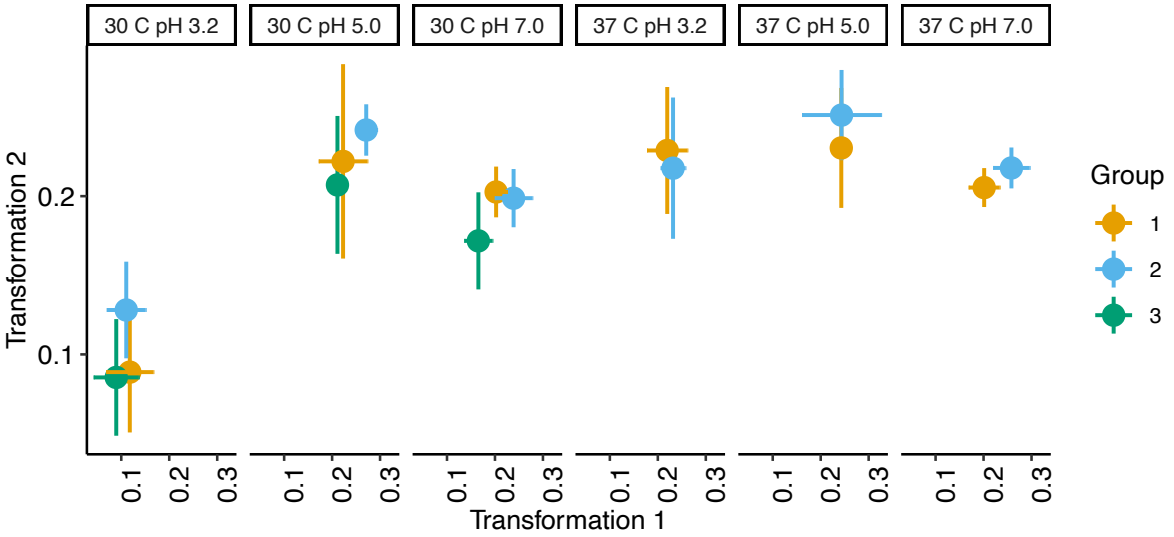

Figure S3.3: Fitness Estimates for a Control Segregant LK5-G01 That Was Present in All Assays. Colors represent different groups of segregants. Segregants were divided into three groups in 30°C environments and into two groups in 37°C environments.

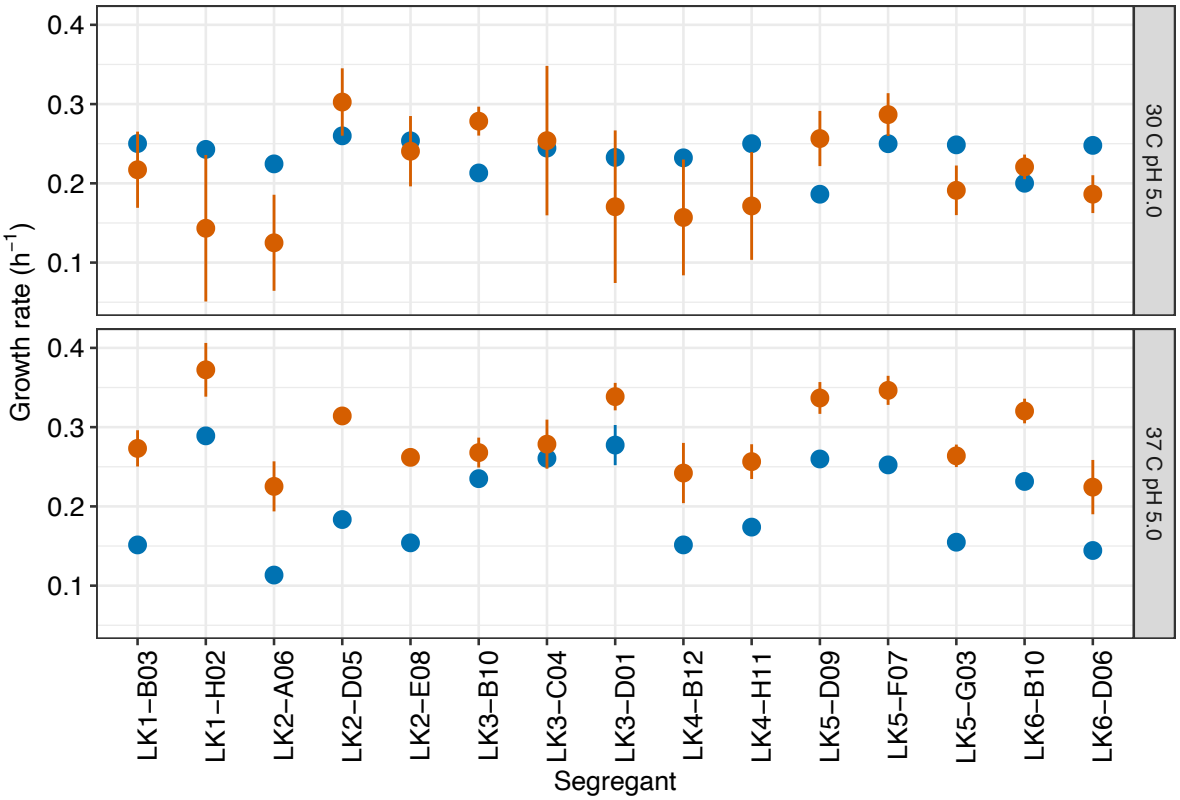

Figure S3.4: Fitness Estimates for Fifteen Segregants From Assays Done in Bulk and Individually. Orange color represents estimates from the assays done in bulk, blue color shown the estimates from the individual assays. For details, see section 3.6.1.1.

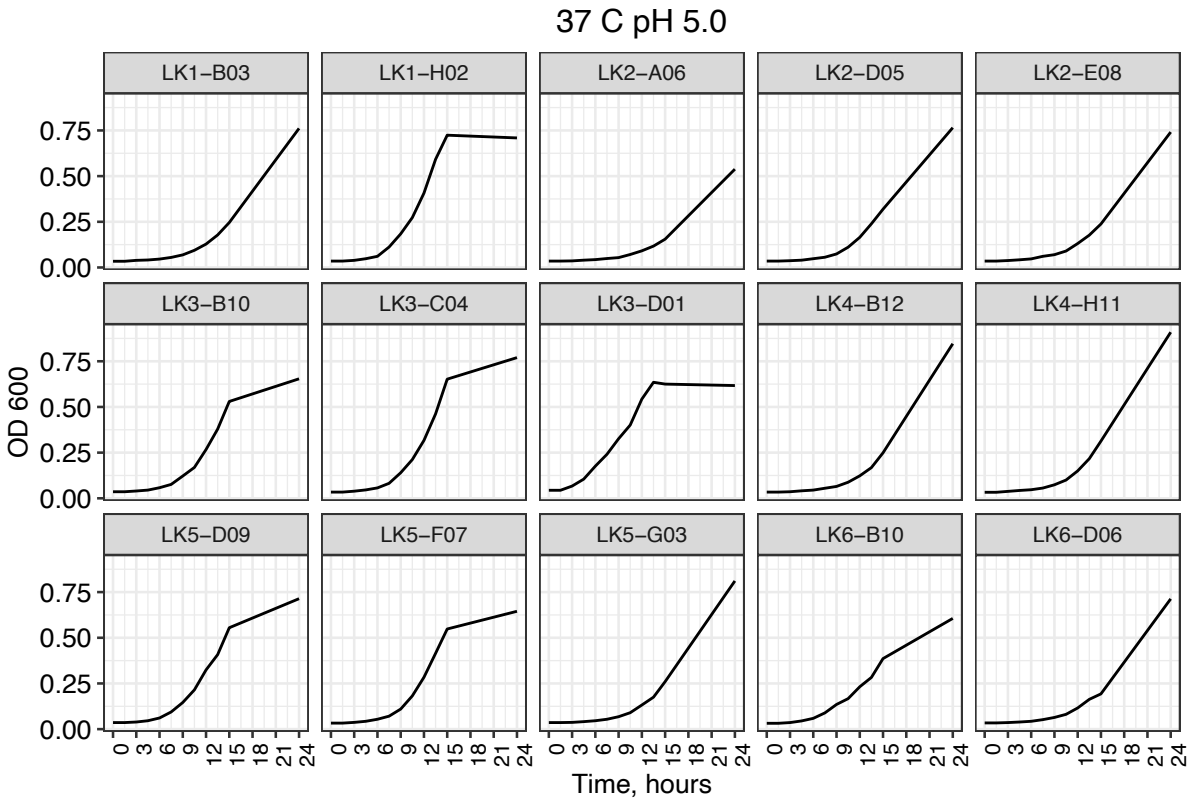

Figure S3.5: Growth Curves for Fifteen Segregants in 37°C pH 5.0.

Experiments were done in 500 ml flasks with 150 ml of media. The populations of segregants were individually precultured for 24 h in SC media at 37 °C pH 5.0. At T0, $5 \cdot 10^7$ cells were transferred from each pre-growth culture to the fresh media in flasks to start the experiment. Measurements were taken every 1.5 hours for initial 15 hours starting from T0. The last measurement was taken 24 h after the start of the experiment. For more details, see section 3.6.1.2.

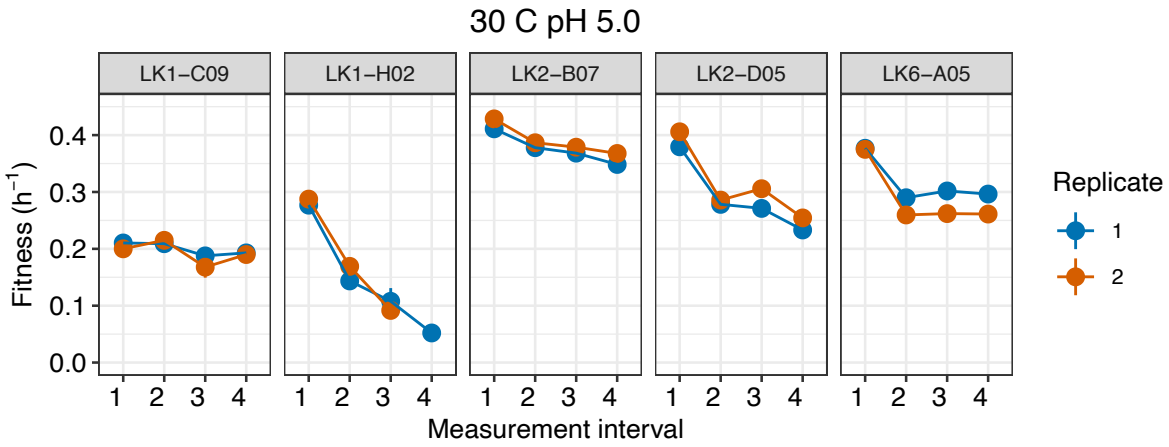

Figure S3.6: Fitness Estimates for Five Segregants Obtained in the Confirmation Experiment Done at 30 °C pH 5.0.

In the TnSeq experiment, we identified that the fitness of some segregants gradually declined over all measurement intervals. The goal of this confirmation experiment was to test whether such decline indicates a possible flaw in the measurements obtained from experiments done in bulk or whether the observed declines are real. In the confirmation experiment, we individually measured fitness for five segregants in two environments following the same protocol as in the RB-TnSeq experiment. The dots on the graph represent the estimated fitness at each time interval. These data show that the fitness of LK1-H02 gradually declined over four time intervals, which is also supported by lower OD values for the population of this segregant at later time points (data not shown). This suggests that the observed decline in fitness is real and does not indicate flaws in the RB-TnSeq method. For details, see section 3.6.1.3.

Additionally, the fitness of the majority of segregants at the first time interval differed from the rest of the estimates. We observed this as in this confirmation experiment, so in the RB-TnSeq experiment. We associate this difference with segregants transitioning to growth in fresh media in flasks after growing for 24 hours in deep-well plates or tubes. For all segregants, we excluded measurements obtained from the first time interval from the analysis.

3.6.1 Description of the Confirmation Experiments

3.6.1.1 Measuring Growth Rates of Segregants

To test whether fitness measurements obtained in the RB-TnSeq experiment represent the growth rate of segregants, we separately measured the growth rate for a set of 15 segregants in two environments, 30 °C pH 5.0 and 37 °C pH 5.0, and compared those measurements to the estimates from the RB-TnSeq experiment. Specifically, we selected 15 segregants that varied in fitness based on our estimates from the RB-TnSeq experiment. The cultures of these segregants were individually pre-grown in a 96 deep-well plates with 1 ml of media at two environmental conditions, 30 °C pH 5.0 and 37 °C pH 5.0, for 24 h following the same protocol as was used in the RB-TnSeq experiment (see Chapter 2).To start time point zero, we measured the concentration of each pre-growth culture and transferred $5 \cdot 10^7$ cells (a bottleneck size used in the RB-TnSeq experiment) into a 500 ml flask with 150 ml of fresh media. The populations of each segregant were maintained in separate flasks. After, we measured the OD600 of the diluted cultures, marking it as a measurement at time point zero. The measurements were taken consecutively for the next 12 hours with a 1.5 h interval. There were two replicate assays for each population. Since it was impossible to maintain more than 15 flasks at the same time, we did replicate assays and measurements in different environments on different days.

To estimate the growth rate, we converted the obtained OD600 measurements into the cell numbers (equation 2.1) and fit them into the exponential growth equation (3.1).

$$R_i = \frac{1}{\Delta T} Ln\left(\frac{N_T}{N_{T-1}}\right), \quad (3.1)$$

where R_i -growth rate estimate for segregant i, h^{-1}, ΔT- time interval (T-(T-1)) at which culture grew in exponential phase based on the OD600 measurements, N_T – number of cells at time point T, N_{T-1} – number of cells at T-1.

The obtained growth rate measurements and their comparison to the estimates from the RB-TnSeq experiment are shown in Fig. S3.4.

3.6.1.2 Measuring Growth Curves

To measure the growth curves, we followed the same procedure as is described in section 3.6.1.1. We used the same set of 15 segregants and did experiment in two replicates. However, the measurements were taken in only one environment, 37 °C pH 5.0. This environment was intentionally chosen because most segregants showed the highest fitness in it. The latter fact ensured that we could capture the transition of populations to a stationary growth phase by taking measurements over 24 hours. We took measurements every 1.5 hours starting from time point zero until time point 15 h. The last measurement was taken 24 h after we measured time point zero. No measurements were taken between time points 15 h and 24 h. The obtained measurements are shown in Fig. S3.5.

3.6.1.3 Confirming the Estimates of Fitness Effects of Mutations Obtained in the RB-TnSeq Experiment

We carried out an additional set of confirmation experiments to test for the reproducibility of fitness effect estimates for mutations from the RB-TnSeq experiment. For that, we picked five segregants, created seven identical mutations in each segregant, and estimated the effects of the generated mutations in 30 °C pH 5.0 and 37 °C pH 7.0 (see section 2.6 of the Chapter 2 for more details).

The set of the seven mutations consisted of two mutations that were used as a neutral reference in the RB-TnSeq experiment, one mutation that was identified to be neutral based on the estimates from the RB-TnSeq experiment. The remaining four mutations had significant effect

on fitness in most of the five segregants and changed the sign of their fitness effect from positive to negative between 30 °C pH 5.0 and 37 °C pH 7.0. environments.

For each mutation, we generated transposon libraries with 10-15 distinct barcodes. Since the selected segregants had different drug resistance markers, we ended up generating 28 unique barcode libraries. These libraries differed in the drug resistance marker: Hygromycin (14 libraries) or Nourseothricin (14 libraries) and in the mutations they are associated with (seven mutations). There were two distinct barcode libraries with the same drug marker for the same mutation which we used to create replicates.

We transformed each segregant with seven libraries twice to generate mutants (see chapter 2 for the transformation protocol). After selecting the mutants on the media with the corresponding drug, we made liquid cultures for individual mutants of each segregant (5 segregants x 7 mutations x 2 replicates x 2 environments = 140 cultures total) in SC media and incubated them in the appropriate environmental conditions (30 °C pH 5.0 and 37 °C pH 7.0) for 24 h in tubes. After, we measured the concentrations of the grown cultures and pooled mutants of the same segregant from the same set of transformations in the following way: the initial frequency of two neutral reference mutations was 25 % each, while each of the remaining mutations had a frequency 10 %. There were 20 different pooled cultures in total. Next, we measured the OD600 of the pooled cultures and transferred $5 \cdot 10^7$ cells to start time point one of the competition assays. Competition assays were performed following the same protocol as in the RB-TnSeq experiment. The only difference was that assays were done separately for each segregant, and they included only seven mutations.

We compared the obtained measurements from this confirmation to the measurements from the RB-TnSeq experiment (Fig. S3.7). The correlation between these estimates is high (R^2=0.76), which means that the obtained measurements are reproducible.

The data from this experiment enabled us also to evaluate the fitness of individual segregants. Since in the bulk assays in the RB-TnSeq experiment many factors could contribute to the observed gradual fitness decline for some segregants, we decided to test whether similar declines are present when segregants grow individually. In this experiment, we observed such decline for LK1-H02 in 30 °C pH 5.0 (Fig. S3.6) and for LK1-H02 and LK1-C09 in 37 °C pH 7.0. We assume that the identified declines are a result of these strains having low fitness and inability to persist in the assay with a given dilution rate.

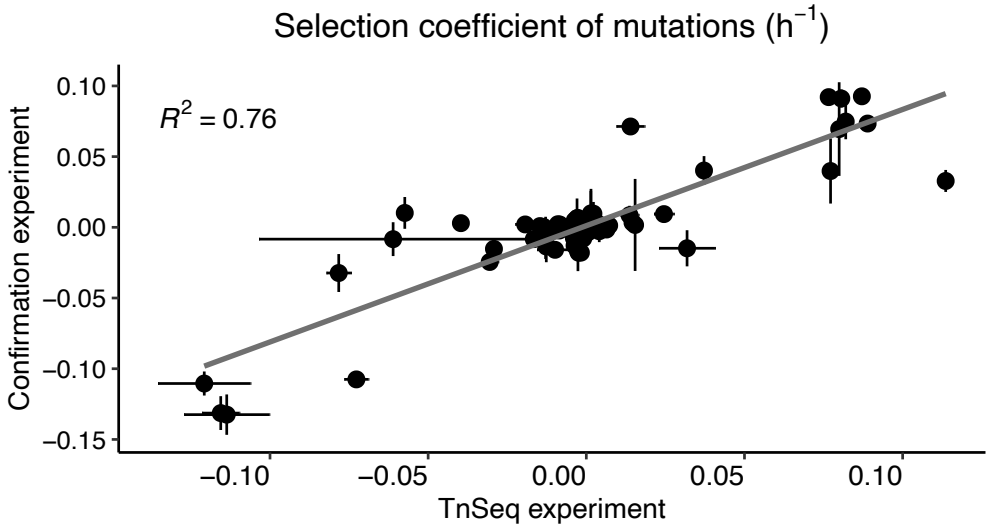

Figure S3.7: Correlation Between Selection Coefficient Estimates for Mutations Between the RB-TnSeq Experiment and the Confirmation Experiment.

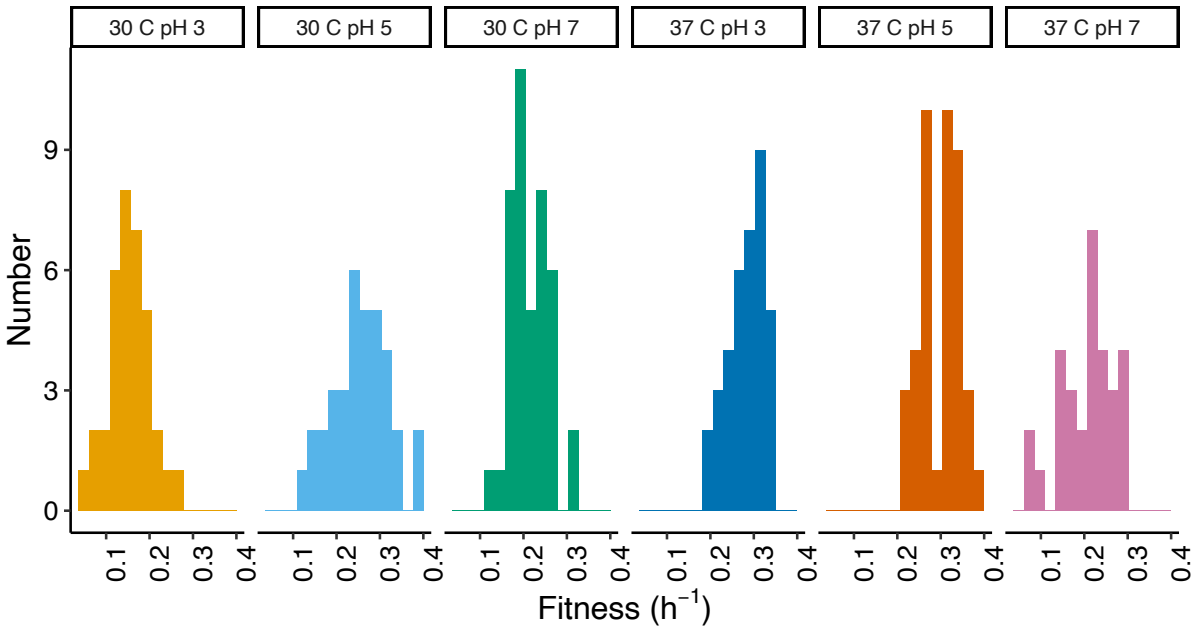

Figure S3.8: Distributions of Fitness of Segregants in Different Environments. In 37°C pH 7.0 the distribution contains the data for only 27 segregants because the other 15 segregants showed a consistent decline in their fitness and we excluded these data from the analysis.

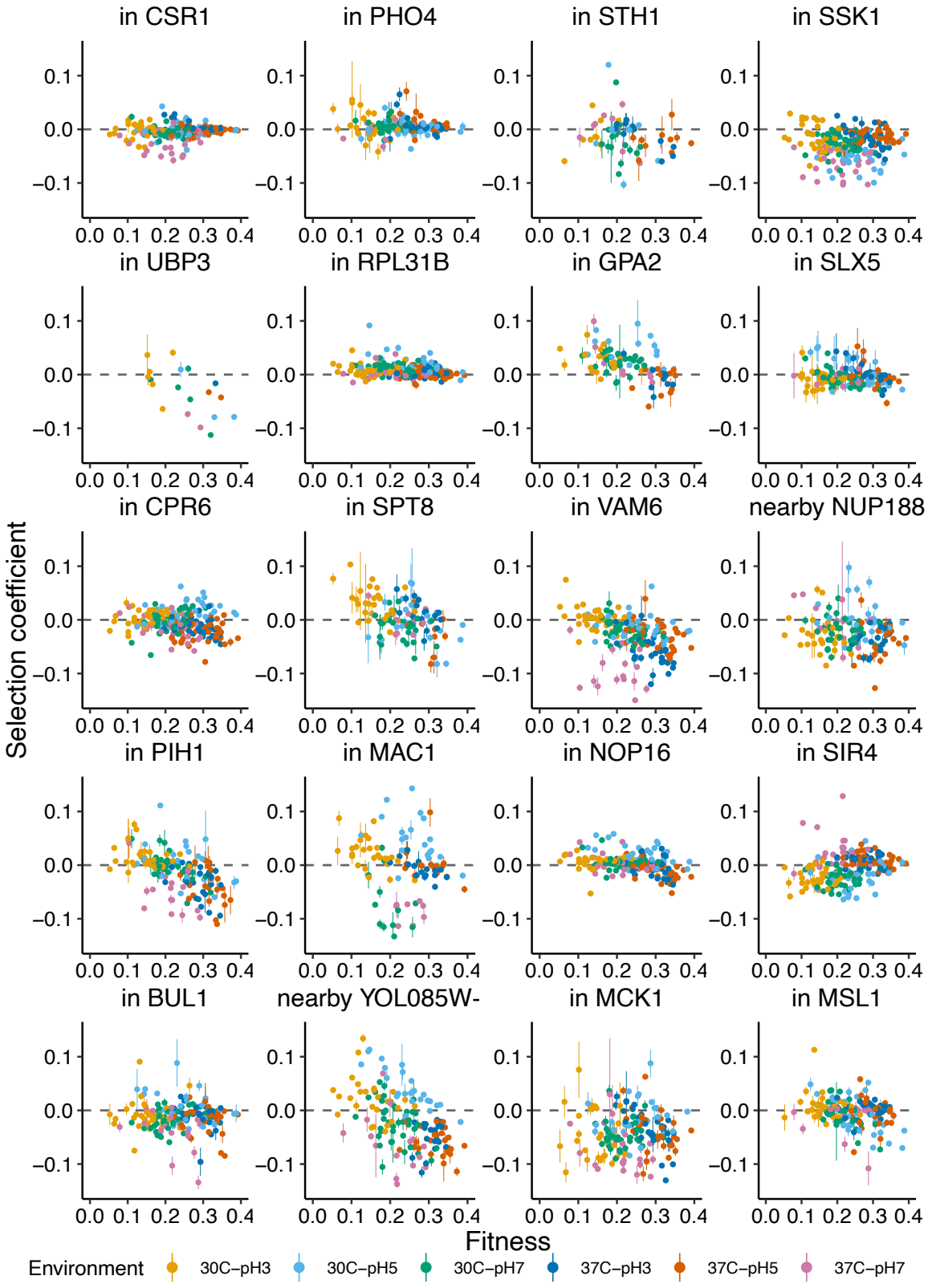

Figure S3.9A: Fitness Effects of Individual Mutations in Different Genotypes and Environments.

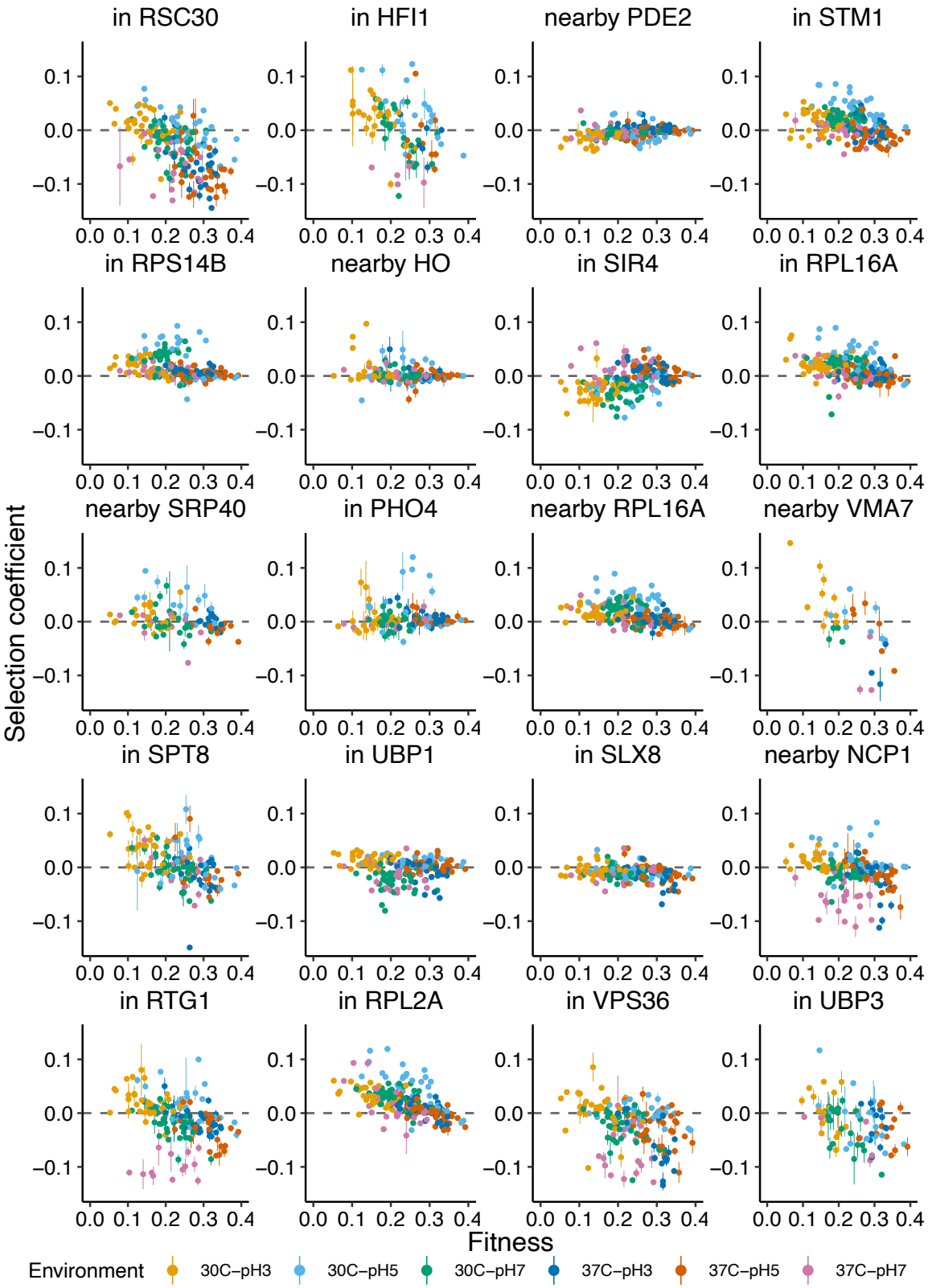

Figure S3.9B: Fitness Effects of Individual Mutations in Different Genotypes and Environments.

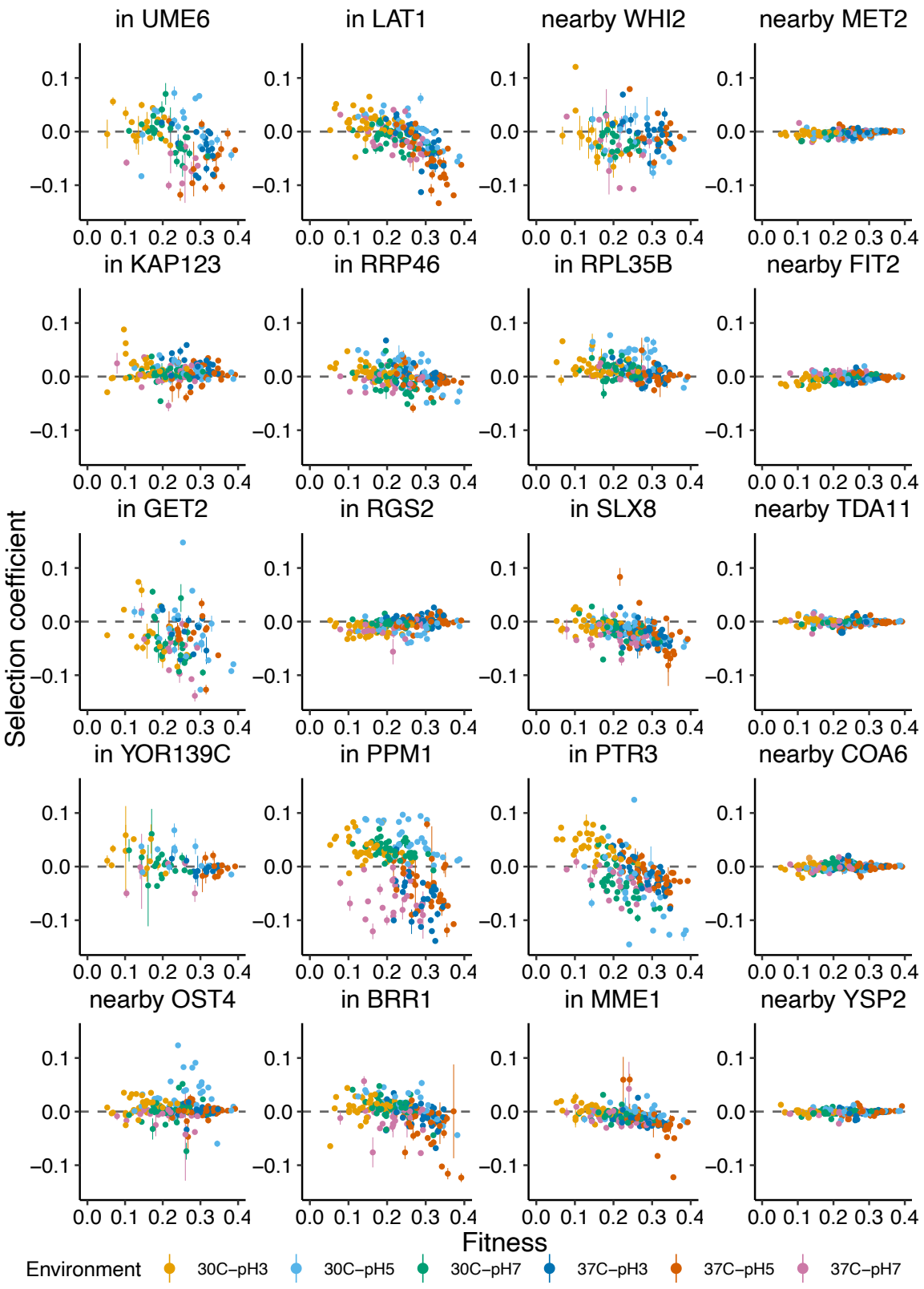

Figure S3.9C: Fitness Effects of Individual Mutations in Different Genotypes and Environments.

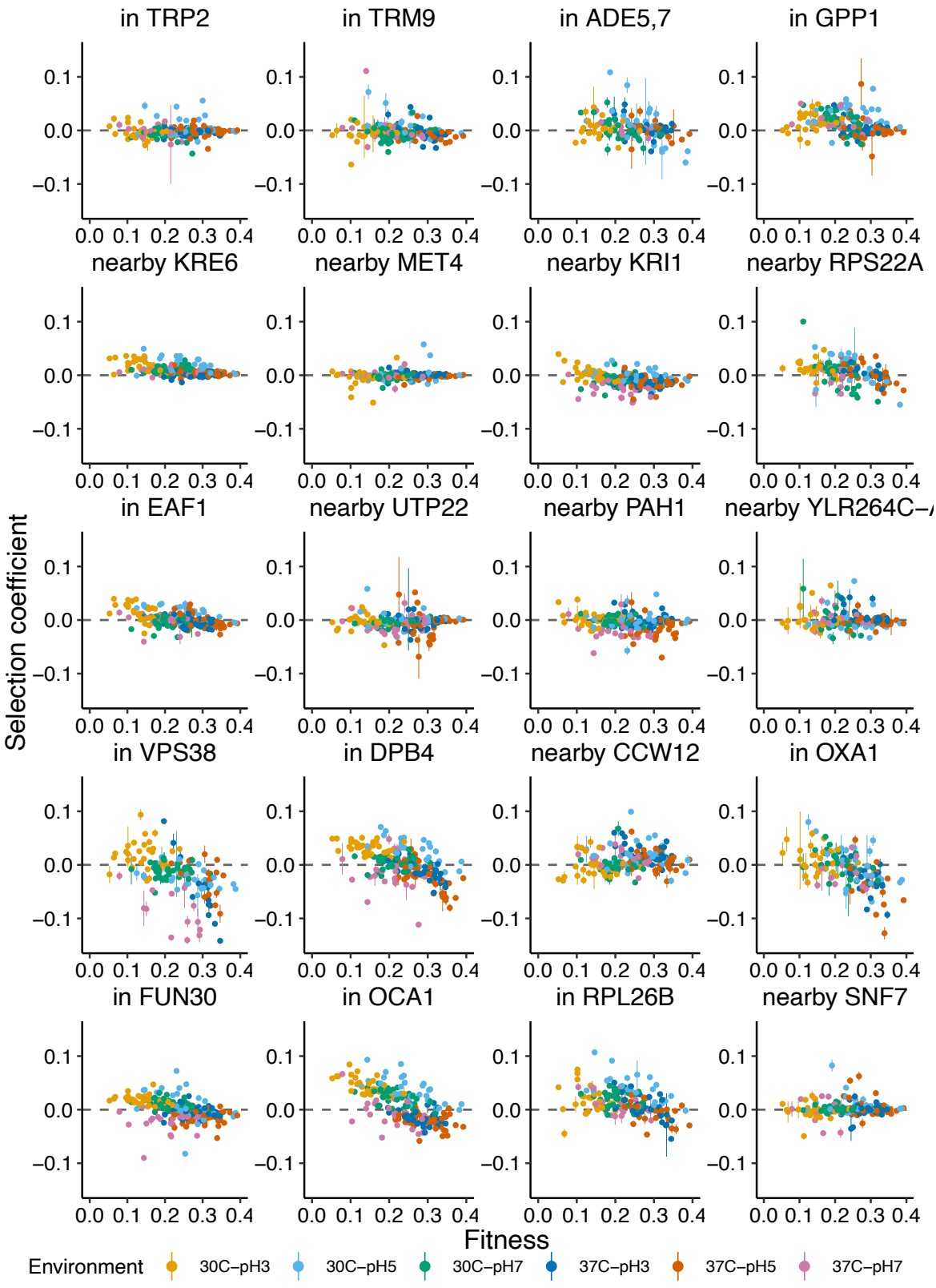

Figure S3.9D: Fitness Effects of Individual Mutations in Different Genotypes and Environments.

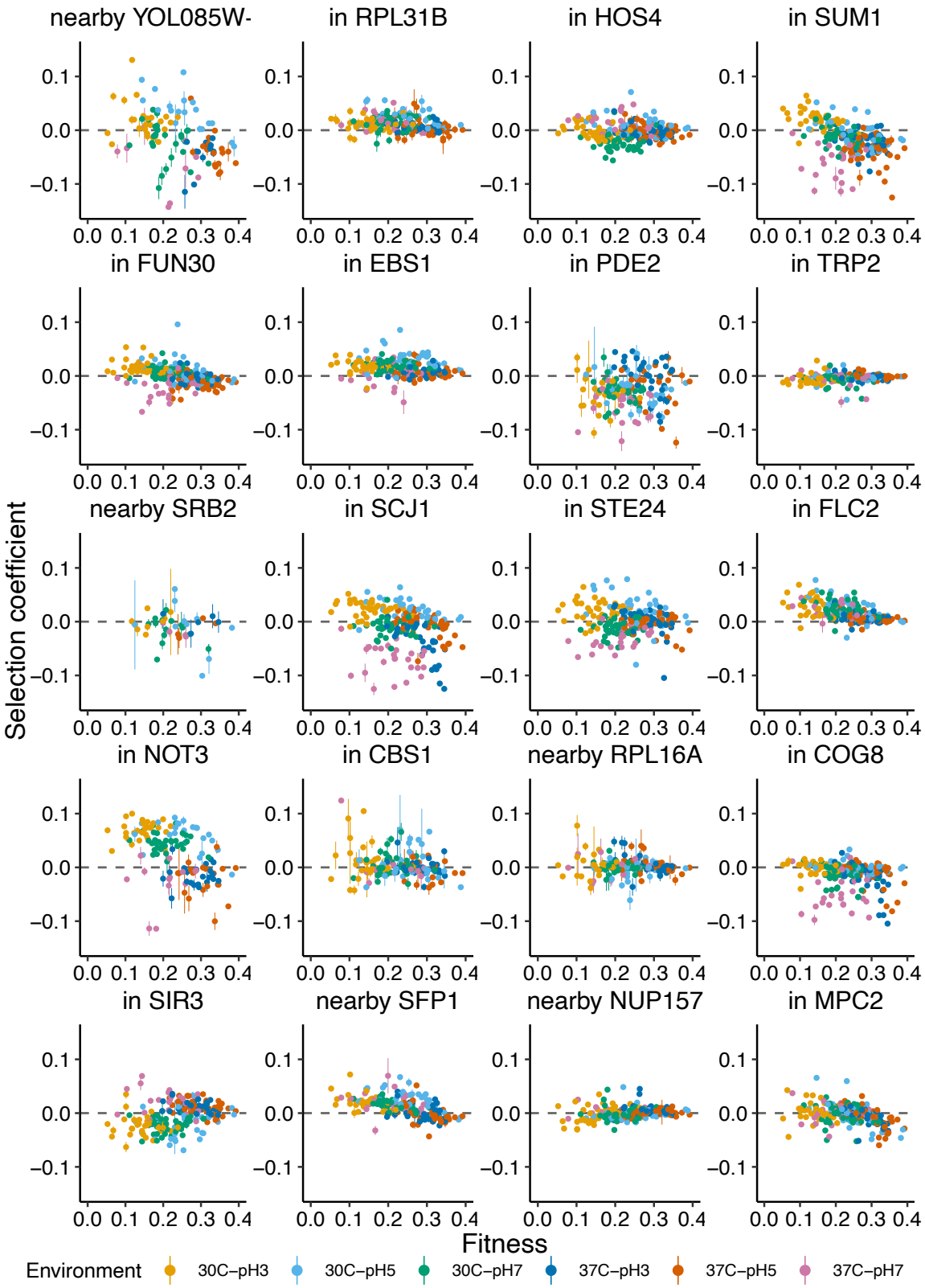

Figure S3.9E: Fitness Effects of Individual Mutations in Different Genotypes and Environments.

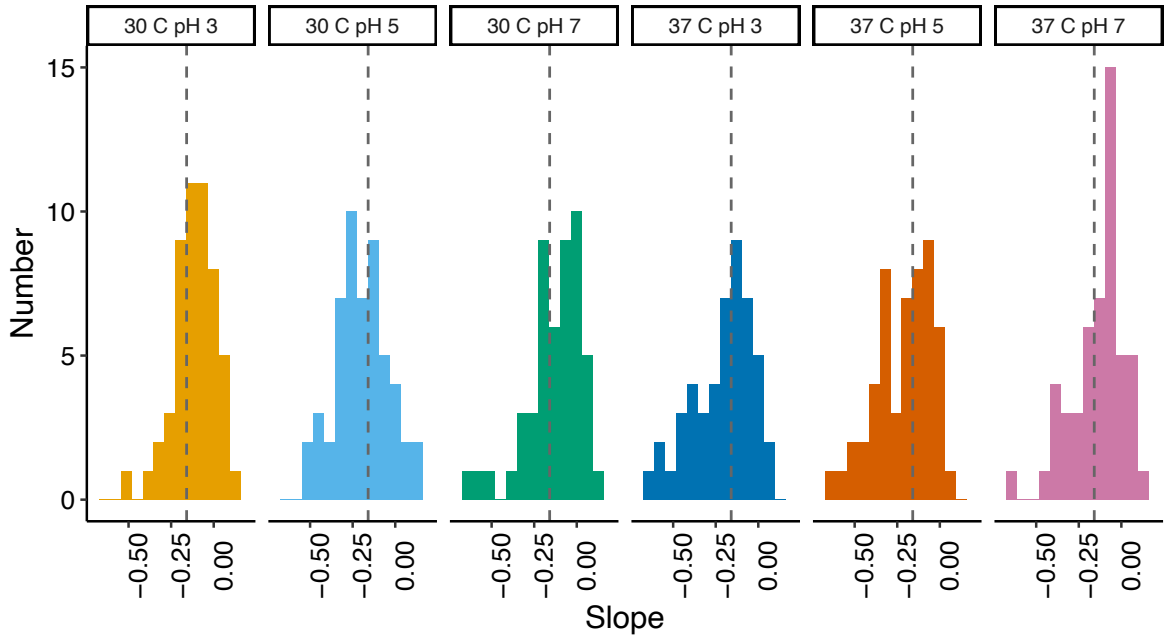

Figure S3.10: Distributions of Slopes Calculated for Regression Lines Describing the Relationship Between Fitness Effects of Individual Mutations and Fitness of Their Genetic Background.

These distributions show the data for 47 mutations for which we identified significant relationships between their effect and background fitness in at least three environments. The slope represents the magnitude of global epistasis affecting the mutated locus [38]. The dashed lines represent mean slope for each distribution.

3.7 References

1. Trindade S, Sousa A, Xavier KB, Dionisio F, Ferreira MG, Gordo I. Positive Epistasis Drives the Acquisition of Multidrug Resistance. Zhang J, editor. PLoS Genet. 2009;5: e1000578. doi:10.1371/journal.pgen.1000578

2. Ward H, Perron GG, Maclean RC. The cost of multiple drug resistance in *Pseudomonas aeruginosa*. J Evol Biol. 2009;22: 997–1003. doi:10.1111/j.1420-9101.2009.01712.x

3. Reynolds MG. Compensatory Evolution in Rifampin-Resistant *Escherichia coli*. Genetics. 2000;156: 1471–1481. doi:10.1093/genetics/156.4.1471

4. Latinne A, Hu B, Olival KJ, Zhu G, Zhang L, Li H. Chmura AA, Field HE, Epstein JH, Li B, Zhang W, Shi ZL, Daszak P. Origin and cross-species transmission of bat coronaviruses in China. Nat Commun. 2020;11: 4235. doi:10.1038/s41467-020-17687-3

5. Morens DM, Breman JG, Calisher CH, Doherty PC, Hahn BH, Keusch GT, Kramer LD, LeDuc JW, Monath TP, Taubenberger JK. The Origin of COVID-19 and Why It Matters. Am J Trop Med Hyg. 2020;103: 955–959. doi:10.4269/ajtmh.20-0849

6. Chakraborty I, Maity P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. Sci Total Environ. 2020;728: 138882. doi:10.1016/j.scitotenv.2020.138882

7. He X, Qian W, Wang Z, Li Y, Zhang J. Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks. Nat Genet. 2010;42: 272–276. doi:10.1038/ng.524

8. Bonhoeffer S. Evidence for Positive Epistasis in HIV-1. Science. 2004;306: 1547–1550. doi:10.1126/science.1101786

9. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative Epistasis Between Beneficial Mutations in an Evolving Bacterial Population. Science. 2011;332: 1193–1196. doi:10.1126/science.1203801

10. Gifford DR, Toll-Riera M, MacLean RC. Epistatic interactions between ancestral genotype and beneficial mutations shape evolvability in *Pseudomonas aeruginosa*: BRIEF COMMUNICATION. Evolution. 2016;70: 1659–1666. doi:10.1111/evo.12958

11. Weinreich DM. Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. Science. 2006;312: 111–114. doi:10.1126/science.1123539

12. Rodriguez-Verdugo A, Carrillo-Cisneros D, Gonzalez-Gonzalez A, Gaut BS, Bennett AF. Different tradeoffs result from alternate genetic adaptations to a common environment. Proc Natl Acad Sci. 2014;111: 12121–12126. doi:10.1073/pnas.1406886111

13. Kvitek DJ, Sherlock G. Reciprocal Sign Epistasis between Frequently Experimentally Evolved Adaptive Mutations Causes a Rugged Fitness Landscape. Zhang J, editor. PLoS Genet. 2011;7: e1002056. doi:10.1371/journal.pgen.1002056

14. Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, Memedov IZ, Tawfik DS, Lukyanov KA, Kondrashow FA. Local fitness landscape of the green fluorescent protein. Nature. 2016;533: 397–401. doi:10.1038/nature17995

15. Olson CA, Wu NC, Sun R. A Comprehensive Biophysical Description of Pairwise Epistasis throughout an Entire Protein Domain. Curr Biol. 2014;24: 2643–2651. doi:10.1016/j.cub.2014.09.072

16. Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS. Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. Nature. 2006;444: 929–932. doi:10.1038/nature05385

17. Wi YM, Greenwood-Quaintance KE, Brinkman CL, Lee JYH, Howden BP, Patel R. Rifampicin resistance in Staphylococcus epidermidis: molecular characterisation and fitness cost of rpoB mutations. Int J Antimicrob Agents. 2018;51: 670–677. doi:10.1016/j.ijantimicag.2017.12.019

18. Barrick JE, Kauth MR, Strelioff CC, Lenski RE. Escherichia coli rpoB Mutants Have Increased Evolvability in Proportion to Their Fitness Defects. Mol Biol Evol. 2010;27: 1338–1347. doi:10.1093/molbev/msq024

19. Rodríguez-Verdugo A, Gaut BS, Tenaillon O. Evolution of Escherichia coli rifampicin resistance in an antibiotic-free environment during thermal stress. BMC Evol Biol. 2013;13: 50. doi:10.1186/1471-2148-13-50

20. Chu X, Zhang D, Buckling A, Zhang Q. Warmer temperatures enhance beneficial mutation effects. J Evol Biol. 2020;33: 1020–1027. doi:10.1111/jeb.13642

21. Kishony R, Leibler S. Environmental stresses can alleviate the average deleterious effect of mutations. J Biol. 2003;2: 14. doi:10.1186/1475-4924-2-14

22. Jasnos L, Tomala D, Paczesniak D, Korona R. Interactions between stressful environment and gene deletions alleviate the expected average loss of fitness in yeast. Genetics. 2008;178: 2105–11. doi:doi: 10.1534/genetics.107.084533.

23. Bandyopadhyay S, Mehta M, Kuo D, Sung M-K, Chuang R, Jaehnig EJ, Huh W-K, Aebersold R, Keogh M-C, Krogan NJ, Ideker T. Rewiring of Genetic Networks in Response to DNA Damage. Science. 2010;330: 1385–1389. doi:10.1126/science.1195618

24. Harrison R, Papp B, Pal C, Oliver SG, Delneri D. Plasticity of genetic interactions in metabolic networks of yeast. Proc Natl Acad Sci. 2007;104: 2307–2312. doi:10.1073/pnas.0607153104

25. Chen P, Zhang J. Antagonistic pleiotropy conceals molecular adaptations in changing environments. Nat Ecol Evol. 2020;4: 461–469. doi:10.1038/s41559-020-1107-8

26. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature. 2009;461: 1243–1247. doi:10.1038/nature08480

27. Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. Second-Order

Selection for Evolvability in a Large Escherichia coli Population. Science. 2011;331: 1433–1436. doi:10.1126/science.1198914

28. MacLean RC, Perron GG, Gardner A. Diminishing Returns From Beneficial Mutations and Pervasive Epistasis Shape the Fitness Landscape for Rifampicin Resistance in *Pseudomonas aeruginosa*. Genetics. 2010;186: 1345–1354. doi:10.1534/genetics.110.123083

29. Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, Lazar V, Hryan M, Kemmeren P, Pai C, Barton NH. The Genomic Landscape of Compensatory Evolution. Barton NH, editor. PLoS Biol. 2014;12: e1001935. doi:10.1371/journal.pbio.1001935

30. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science. 2014;344: 1519–1522. doi:10.1126/science.1250939

31. Rokyta DR, Abdo Z, Wichman HA. The Genetics of Adaptation for Eight Microvirid Bacteriophages. J Mol Evol. 2009;69: 229–239. doi:10.1007/s00239-009-9267-9

32. Schoustra S, Hwang S, Krug J, de Visser JAGM. Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. Proc R Soc B Biol Sci. 2016;283: 20161376. doi:10.1098/rspb.2016.1376

33. Flynn KM, Cooper TF, Moore FB-G, Cooper VS. The Environment Affects Epistatic Interactions to Alter the Topology of an Empirical Fitness Landscape. Fay JC, editor. PLoS Genet. 2013;9: e1003426. doi:10.1371/journal.pgen.1003426

34. Hall AE, Karkare K, Cooper VS, Bank C, Cooper TF, Moore FB -G. Environment changes epistasis to alter trade-offs along alternative evolutionary paths. Evolution. 2019;73: 2094–2105. doi:10.1111/evo.13825

35. Chou H-H, Delaney NF, Draghi JA, Marx CJ. Mapping the Fitness Landscape of Gene Expression Uncovers the Cause of Antagonism and Sign Epistasis between Adaptive Mutations. Achtman M, editor. PLoS Genet. 2014;10: e1004149. doi:10.1371/journal.pgen.1004149

36. Wei X, Zhang J. Patterns and Mechanisms of Diminishing Returns from Beneficial Mutations. Agashe D, editor. Mol Biol Evol. 2019;36: 1008–1021. doi:10.1093/molbev/msz035

37. Johnson MS, Martsul A, Kryazhimskiy S, Desai MM. Higher-fitness yeast genotypes are less robust to deleterious mutations. Science. 2019;366: 490–493. doi:10.1126/science.aay4199

38. Reddy G, Desai MM. Global epistasis emerges from a generic model of a complex trait. eLife. 2021;10: e64740. doi:10.7554/eLife.64740

39. Lyons DM, Zou Z, Xu H, Zhang J. Idiosyncratic epistasis creates universals in mutational effects and evolutionary trajectories. Nat Ecol Evol. 2020;4: 1685–1693. doi:10.1038/s41559-020-01286-y

40. Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, Kruglyak L. Finding the sources of missing heritability in a yeast cross. Nature. 2013;494: 234–237. doi:10.1038/nature11867

41. Ehrenreich IM, Torabi N, Jia Y, Kent J, Martis S, Shapiro JA, Gresham D, Caudy AA, Kruglyak L. Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature. 2010;464: 1039–1042. doi:10.1038/nature08923

42. Szafraniec K, Borts RH, Korona R. Environmental stress and mutational load in diploid strains of the yeast Saccharomyces cerevisiae. Proc Natl Acad Sci. 2001;98: 1107–1112. doi:10.1073/pnas.98.3.1107

43. Jerison ER, Kryazhimskiy S, Desai MM. Pleiotropic consequences of adaptation across gradations of environmental stress in budding yeast. ArXiv14097839 Q-Bio. 2014 [cited 18 Jul 2021]. Available: http://arxiv.org/abs/1409.7839

44. Chou H-H, Chiu H-C, Delaney NF, Segre D, Marx CJ. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. Science. 2011;332: 1190–1192. doi:10.1126/science.1203799

45. Qian W, Ma D, Xiao C, Wang Z, Zhang J. The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. Cell Rep. 2012;2: 1399–1410. doi:10.1016/j.celrep.2012.09.017

46. Maharjan R, McKenzie C, Yeung A, Ferenci T. The basis of antagonistic pleiotropy in hfq mutations that have opposite effects on fitness at slow and fast growth rates. Heredity. 2013;110: 10–18. doi:10.1038/hdy.2012.46

47. Chou H-H, Berthet J, Marx CJ. Fast Growth Increases the Selective Advantage of a Mutation Arising Recurrently during Evolution under Metal Limitation. Matic I, editor. PLoS Genet. 2009;5: e1000652. doi:10.1371/journal.pgen.1000652

48. Chiu H-C, Marx CJ, Segrè D. Epistasis from functional dependence of fitness on underlying traits. Proc R Soc B Biol Sci. 2012;279: 4156–4164. doi:10.1098/rspb.2012.1449

49. Chou H-H, Chiu H-C, Delaney NF, Segre D, Marx CJ. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. Science. 2011;332: 1190–1192. doi:10.1126/science.1203799

50. Otwinowski J, McCandlish DM, Plotkin JB. Inferring the shape of global epistasis. Proc Natl Acad Sci. 2018;115: E7550–E7558. doi:10.1073/pnas.1804015115

51. Ardell SM, Kryazhimskiy S. The Population Genetics of Collateral Resistance and Sensitivity. Evolutionary Biology; 2020 Aug. doi:10.1101/2020.08.25.267484