Quantifying and Understanding Intragenic and Intergenic Epistasis in Yeast

by

Chuan Li

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in the University of Michigan
2017

Doctoral Committee:

Professor George Zhang, Chair
Associate Professor Timothy James
Professor Anuj Kumar
Professor Kerby Shedden
Professor Trisha Wittkopp

Chuan Li

lichuan@umich.edu

ORCID iD: 0000-0002-4381-3891

© Chuan Li 2017

DEDICATION

To my husband Xiaoyu, and my parents

ACKNOWLEDGEMENT

There are numerous people I would like to acknowledge for their guidance, support and friendship throughout my experience as a PhD student. First of all, I would like to express my deepest gratitude to my advisor, Professor George Zhang, for his support and mentorship over the past six years. The guidance and encouragement from him have been a great treasure for me, without which the work could not have been finished. His passion for science and keen sense for research have inspired me a lot. I would also like to thank all my committee members, Professor Trisha Wittkopp, Professor Timothy James, Professor Anuj Kumar and Professor Kerby Shedden for their time and effort serving on my dissertation committee and kindly providing many constructive suggestions and comments.

Next, I would like to thank Calum Maclean, for being a great friend and colleague, tutoring me through many details in experimental work. I would like to thank: Wenfeng Qian, for his encouragements and inspiring ideas as well as his contributions to Chapter 2; Jianrong Yang, for his help in computational work; Zhi Wang, for his contribution to Chapter 4; Wei-chin Ho, who has been very friendly and helped me in various aspects. I am thankful for all current and previous members of Zhang lab: Xiaoshu Chen, Zhengting Zou, Huabin Zhao, Haiqing Xu, Jinrui Xu, Ying Li, Soochin Cho, Diyan Li, Guixia Xu, Xinzhu Wei, Chungoo Park, Bryan Moyers, Nagarjun Vijay, Mengyi Sun, Piaopiao Chen, Zhen Liu, Minhan Yi and Chuan Xu. They have been wonderful colleagues, and my PhD life and research won’t be as interesting and successful without them. I would also like to thank many friends in Wittkopp lab, for the fruitful and happy interactions.

I would like to thank people I met in and out of classrooms for their friendship and support. There are so many names but I can only list a few: Zi Li, Yue Xie, Lihan Xie, Andrea Thomas, Xin Xin, Yuanying Su, Yiwei Zhang, Jingxuan Liu, Jingcheng Wansg, Tara Clancy, Katherine Crocker, Andrew Henderson, Dave Yuan, Jingchun Li, Xiaoxing Wang, Hang Ren, Tamara Milton, Brendan O'Neill and Damian Wassel.

I would also like to thank the National Science Foundation, the Department of Ecology and Evolutionary Biology and Rackham Graduate School for research funding and fellowships. The PhD work couldn’t have been accomplished without their financial support.

Last but not least, my love and gratitude goes to my husband Xiaoyu Wang, my parents Lizhuang Li and Junyi Li. They have been supporting me unconditionally, trusting me wholeheartedly, and providing me with unlimited source of energy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

Epistasis describes a broad range of interactions within and between molecules. However, limited empirical knowledge is currently available for epistasis at large scale. This dissertation focuses on quantifying intragenic and intergenic epistasis on a large scale. For intragenic epistasis, by combining precise gene replacement and next-generation sequencing, I measured fitness for over 65,000 yeast strains each carrying a unique variant of the $tRNA^{Arg}_{CCU}$ gene. I managed to quantify epistasis for 61% of all possible combinations of mutations. Almost half of all mutation pairs exhibit significant epistasis, which has a strong negative bias except when the mutations occur at Watson-Crick paired sites. The strong negative bias is also observed for epistasis on the genetic background with one or multiple existing mutations. To study how the fitness landscape and epistasis vary among environments, I measured fitness landscapes in four environments and found that the same mutation almost always has different fitness effects in different environments, indicating pervasive genotype by environment interactions (G×E). Nevertheless, the observed G×E follows a simple piecewise linear relationship. Given the prediction of fitness, an epistasis prediction is also calculated, and the predictive power is comparatively high. Apart from intragenic epistasis, I also studied genetic incompatibility, a form of intergenic epistatic interactions between otherwise functional genes in their conspecific genetic background, which is commonly considered as the major cause of postzygotic isolation and speciation. Despite repeated efforts, Bateson-Dobzhansky-Muller (BDM) incompatibilities between nuclear genes have never been identified between *S. cerevisiae* and its sister species *S. paradoxus*. Such negative results have led to the belief that simple nuclear BDM incompatibilities do not exist in yeast. I explored an alternative explanation that such incompatibilities exist but were undetectable due to limited statistical power, and discovered that previously employed statistical methods were not ideal and that a redesigned method improves the statistical power. I also determined, under various sample sizes, the probabilities of identifying BDM incompatibilities that cause F1 spore inviability with incomplete penetrance, and confirm that the previously used samples were too small to detect such incompatibilities, calling for an expanded experimental search for yeast BDM incompatibilities. In summary, this dissertation shows that understanding epistasis at large scale is important and can be achieved through several powerful approaches for elucidating the underlying mechanisms governing evolution, such as available evolutionary trajectories and repeatability of evolution.

CHAPTER 1
INTRODUCTION

1.1 Introduction

1.1.1 The history of defining epistasis

The word “epistasis” is widely used to describe a broad range of interactions within and between molecules, after being coined by William Bateson ~100 years ago to refer to the masking of the effects of one locus by another in a dihybrid cross (Bateson 1909). Gradually, this term has been expanded to describe a broad range of complex interactions among genetic loci, including the functional relationship between genes, the genetic ordering of regulatory pathways and the quantitative differences of allele-specific effects (definitions reviewed in Phillips 2008).

Traditionally, given the limitations of available techniques and information, researchers examine one gene or one mutation at a time and tend to associate a gene with a specific function and a mutation with a specific mutant phenotype. However, all genes work in a genomic and cellular context and are almost always working together with other pathways and modules to perform their functions, contributing to the final phenotype. Inevitably, there are interactions with other components of the pathway or other gene products. Indeed, an increasing amount of evidence has shown that epistasis is a prevailing phenomenon and the phenotypic effect of a mutation can depend on its genomic background and the potential interaction with other genes or sites (Moore 2003, Weinreich, Watson et al. 2005, Arias, le Poul et al. 2016).

Since R. A. Fisher used the term “epistacy” to statistically describe deviation from the addition of superimposed effects occurring between different genes in 1918 (Fisher 1918), the term “epistasis” is becoming more frequently used to describe the quantitative deviation of combined effects from the expected phenotype, apart from the original reference to qualitative traits. By comparing the phenotype of a double-mutant organism with the expected phenotype assuming independence, we can categorize epistatic interactions as positive or negative (Mani, St Onge et al. 2008). Positive epistasis refers to cases where the double mutant shows a less extreme phenotype than expected. Positive epistasis can occur when a mutation in one gene fully impairs the function the gene or the whole pathway, or where any further mutations in the gene or the pathway cannot further reduce the functionality. Negative epistasis, on the other hand, refers to cases where the double mutant shows a more extreme phenotype than expected. An enrichment for negative epistasis has been observed in several large-scale studies on intragenic and intergenic interactions (Khan, Dinh et al. 2011, Li, Qian et al. 2016, Puchta, Cseke et al. 2016) and is viewed by some as a by-product of selection on genetic robustness (Azevedo, Lohaus et al. 2006). An extreme case for negative epistasis is synthetic lethality, which describes any combination of two separately non-lethal mutations that leads to inviability.

The definitions of positive and negative epistasis are seemingly straightforward; however, there isn’t a universally applicable way to define the quantitative phenotype measure of the trait and an independence function that predicts the expected phenotype assuming no interaction between the focal sites or genes. There are various ways to define an independence function. For independence functions to characterize the relative-growth rate measure, researchers have used Min, Product, Log and Additive definitions, which can lead to different patterns of positive and negative epistasis even for the same dataset (Mani, St Onge et al. 2008).

Among various existing definitions of independence functions for relative fitness, the two most commonly used are $\varepsilon = f_{AB} - f_A f_B$ assuming a multiplicative model for fitness, and $\varepsilon = \log f_{AB} - \log f_A - \log f_B$ assuming additivity of growth rate, which is proportional to log-fitness (Jasnos and Korona 2007). Moreover, the choice of quantitative phenotype measures can also impact the final result. Take the commonly used yeast colony size as an example. Previous studies have used diameter (Bloom, Ehrenreich et al. 2013), area (Baryshnikova, Costanzo et al. 2010) or 3D volume (Zackrisson, Hallin et al. 2016) of the colony as a proxy for fitness, which will lead to different results in quantifying epistasis. Even relative fitness, which can be easily quantified by many different methods, can take different measures, such as exponential growth rate, number of progeny relative to that of wild-type per generation, etc. Using different definitions and independence functions can sometimes lead to different conclusions for the same dataset, so it is of great importance to choose the definitions that are most biologically and statistically meaningful.

1.1.2 Methodology for quantifying epistasis

There are various methodologies for quantifying intragenic and intergenic epistasis. A classical and straightforward method to measure intragenic epistasis for a few sites is to individually measure the phenotype of all combinations of mutations (Weinreich, Delaney et al. 2006, Ortlund, Bridgham et al. 2007), which can help reveal all possible evolutionary trajectories, gain in-depth understanding of the biological functions and quantify high-order epistatic interactions. However, given that the total number of combinations for n sites is 2^n if we only consider the simplest case of two possible states at each site, this methodology is only applicable to studying a few key sites. If the main focus is on pair-wise interactions at the wild-type genetic background instead of evolutionary trajectories that are a few steps away from the wild-type genotype, then for n sites in a haploid organism such as the yeast lab strain, there are a total of n (n-1)/2 possible pairwise interactions. The advent of high-throughput sequencing techniques and other technologies in measuring phenotype and genotypes at a large scale has allowed quantification of these pair-wise interactions for a whole gene (Li, Qian et al. 2016, Puchta, Cseke et al. 2016, Sarkisyan, Bolotin et al. 2016) or a domain of a protein (Olson, Wu et al. 2014). For instance, using Bar-seq method, fitness can be quantified for tens of thousands of variants simultaneously. Measuring fitness for all single mutants, and the majority of double mutants allows for quantifying epistasis at a large scale.

For interspecific epistasis, the positive and negative epistasis between millions of different loss-of-function mutations has been quantified in yeast (Costanzo, Baryshnikova et al. 2010) and other model organisms (Lehner, Crombie et al. 2006, Byrne, Weirauch et al. 2007, Lin, Wang et al. 2010). The yeast knockout mutant collection (Giaever, Chu et al. 2002) has enabled quantifying intergenic epistasis systematically. For instance, identification of synthetic-lethality for ~6,000 genes using the yeast knockout mutant collection by the synthetic genetic array and synthetic-lethality analysis by microarray has revealed abundant information on the networks of genetically connected genes (reviewed in Ooi, Pan et al. 2006). The most conclusive endeavor is the construction of a genetic interaction network by testing all possible pairwise genetic interactions for most of the ~6,000 genes in yeast, revealing nearly a million interactions involving ~90% of yeast genes (Costanzo, VanderSluis et al. 2016).

1.1.3 Underlying mechanisms for epistasis

Common molecular mechanisms of such intergenic epistasis include interaction between changes in interaction interfaces, functional redundancy and interactions between pathways.

Changes in interaction interfaces apply to proteins that directly interact with one another. For instance, a mutation in one protein can be compensated by a different mutation in its interacting partner, such as in the nematode sex-determining genes (Haag, Wang et al. 2002) and many other genes involved in mating (Clark, Gasper et al. 2009), which commonly lead to coevolution of sequences (Feinauer, Szurmant et al. 2016). Another mechanism, functional redundancy, can lead to negative epistasis between genes, and the redundancy is sometimes a by-product of gene duplication events (Dean, Davis et al. 2008). Intergenic epistasis can sometimes be explained by the interaction between pathways or functional modules. Two or multiple redundant or alternative pathways might be carrying out the same biological functions, so knocking out one of them has mild effects while simultaneously blocking them can cause severe defects (Wang, Lee et al. 2002, Kelley and Ideker 2005). Many other possible mechanisms have been revealed through in-depth case studies, including genetic capacitors and physical constraints (Lehner 2011). However, most of the identified intergenic interactions cannot be easily explained by the above-mentioned mechanisms, and there are many unexpected epistatic effects revealed between seemingly unrelated genes (He, Qian et al. 2010).

As for intragenic epistasis, pervasive negative epistasis is observed in multiple large-scale studies (Bershtein, Segal et al. 2006, Li, Qian et al. 2016, Puchta, Cseke et al. 2016), and the underlying mechanism is largely unexplained. However, many common mechanisms for intramolecular epistasis have been revealed in previous studies, including stability thresholds, conformational epistasis and intramolecular pleiotropy (reviewed in Lehner 2011). Stability threshold has been a common mechanism to explain negative interactions between mutations where the protein has redundant stability, and individual mutations are not too detrimental, as long as the protein has not reached the stability margin. Such extra stability and robustness can promote evolvability (Bloom, Labthavikul et al. 2006). Another mechanism is conformational epistasis, which requires multiple mutations occurring simultaneously for the gene to perform some function. A rigorously tested case about a vertebrate steroid receptor conducted by Ortlund *et al*. (2007) found that two changes at interacting sites are needed to transform a generalized ancestral protein to a more specialized receptor while mutating each site alone destroys the receptor function. Another common scenario is intramolecular pleiotropy, where a functionally beneficial mutation can have a side-effect on stability and can only be accessible when a seemingly neutral mutation provides compensatory stability (Tokuriki, Stricher et al. 2008).

1.1.4 Sign epistasis and its biological implication

Sign epistasis occurs when the fitness effect of a mutation is opposite depending on the presence or absence of another mutation and is of special interest because it greatly reduces the fraction of open paths in adaptation (Weinreich, Delaney et al. 2006). For two mutations, A and B, sign epistasis satisfies $(f_A\text{-}1)\ (f_{AB}/f_B\text{-}1) < 0$ or $(f_B\text{-}1)\ (f_{AB}/f_A\text{-}1) < 0$. A more stringent form of sign epistasis, reciprocal sign epistasis, satisfies $(f_A\text{-}1)\ (f_{AB}/f_B\text{-}1) < 0$ and $(f_B\text{-}1)\ (f_{AB}/f_A\text{-}1) < 0$, and further restricts possible evolutionary paths. From the above definition, we can see that sign epistasis can be either positive or negative epistasis. A very classic case study of sign epistasis involves the lysozyme of modern game birds, with a three-amino-acid triplet being either triplet Thr40, Ile 55 Ser91 or triplet Ser40, Val55, and Thr91. Testing the thermostability of all intermediates shows that each evolutionary pathway connecting two extant triplets includes a variant that is unstable (Malcolm, Wilson et al. 1990). Understanding sign epistasis further emphasizes the fact that effect of a gene mutant is not a stand-alone property for the mutation itself and the importance of studying epistasis to characterize the genotype-phenotype mapping fully.

1.1.5 Epistasis changing with environment and genetic background

Epistasis is not a static trait for a pair of sites; it can vary across different environments and genetic backgrounds. Change of epistasis across genetic background is also commonly referred to as higher-order epistasis. A study using 32 strains carrying all possible combinations of 4 quantitative trait nucleotides (QTN) that govern yeast sporulation efficiency showed clear examples of epistatic interactions being different across various the genetic and environmental background, rendering the phenotypic outcome of QTN unpredictable from the genotype alone (Gerke, Lorenz et al. 2010). Such complex interactions of epistasis changing with environments (You and Yin 2002, Wang, Sharp et al. 2009) and genetic backgrounds (Weinreich, Delaney et al. 2006, Weinreich, Lan et al. 2013) are also broadly observed in other species.

1.1.6 Importance of studying epistasis

Understanding epistasis has important theoretical implications. Knowledge about patterns and mechanisms of epistasis helps people understand the structure and function of genetic pathways and how and why certain evolutionary paths are taken while others are blocked. For instance, knocking out genes separately and measuring the phenotype of double mutant allows for ordering of genes in a regulatory hierarchy (Avery and Wasserman 1992). Such an endeavor at a large scale can help reveal a map of the functional network in a cell (Tong, Evangelista et al. 2001). Without epistasis, evolution should be easy to predict and understand: once one beneficial mutation occurs, it would gradually become fixed after escaping drift, and there would be no preferences in the order of occurrence for multiple mutations, with all evolutionary paths being equally accessible. However, due to the complex interactions, certain evolutionary paths are blocked (Weinreich, Delaney et al. 2006) while other seemingly inaccessible paths become possible (Bloom and Arnold 2009).

Quantifying epistasis also has crucial real-world applications, including understanding quantitative traits (Alvarez-Castro and Carlborg 2007), complex diseases (Moore 2003, Azevedo, Suriano et al. 2006, Nagalakshmi, Wang et al. 2008) and the evolution of antibiotic resistance (Weinreich, Delaney et al. 2006). Moreover, it can help with developing treatments for diseases. For instance, describing synthetic lethality lays the foundation for developing anticancer therapy by targeting a gene that is synthetic lethal to a cancer relevant mutation (Kaelin 2005, Herter-Sprie, Chen et al. 2011). It is also vital to consider epistasis in the area of synthetic biology given that the best amino acid / module might not always be the best in all circumstances, and they always must be evaluated in context (Currin, Swainston et al. 2015). Thus, quantifying and understanding epistasis is a fundamental task in biology.

1.1.7 Difficulties in studying epistasis

Given the theoretical importance and wide application of understanding epistasis, it is surprising that current research on epistasis is still quite rudimentary. This is likely due to the multiple challenges in quantifying epistasis. First, the number of possible interactions is large, increasing polynomially with the number of focal genes / sites when we compare interaction between n parties, or increasing exponentially when we consider all high-order interactions. Secondly, quantifying epistasis requires measurements of three phenotypes accurately, thus requiring a technique with both high-throughput and high precision, which is just starting to become more available. Finally, similar to the fact that the functionality of a gene / site is dependent on its environment, the epistatic interaction between sites / genes may also vary across different genetic backgrounds and environments, making it difficult to draw generally applicable conclusions and make precise predictions across multiple genetic backgrounds and environments.

1.2 Thesis overview

In my thesis, I address several questions in quantifying and understanding epistasis using mostly the budding yeast *Saccharomyces cerevisiae* and its close relatives as my model organisms. These yeast species, especially within the *Saccharomyces sensu stricto* group, are well characterized in terms of their genome sequences (Cherry, Hong et al. 2012). Extensive research has provided us with numerous functional genomic data, including protein-protein interactions (Yu, Braun et al. 2008), expression levels (Nagalakshmi, Wang et al. 2008), gene knock-out effects (Giaever, Chu et al. 2002, Costanzo, Baryshnikova et al. 2010, Ryan, Shapiro et al. 2012, Giaever and Nislow 2014) and population genetics (Strope, Skelly et al. 2015). A well-developed toolkit for genetic analysis is available, such as sporulation, mating, competition, and transformation.

I investigate epistasis at both intramolecular and intermolecular levels. The first half of my thesis focuses on the epistasis within a tRNA gene. I examine how mutations at individual positions of a tRNA gene and their low-order combinations affect the fitness of *Saccharomyces cerevisiae*, from which intragenic epistasis is quantified at a large scale. The second half of my thesis focuses on epistasis between genes. I explain how incompatible gene pairs involved in the postzygotic isolation between two yeast species may be identified effectively by a genomic approach.

In Chapter 2, I measure the fitness landscape of a tRNA gene. Fitness landscapes describe the genotype-fitness relationship and are a major determinant of evolutionary trajectories. The vast genotype space, coupled with the difficulty of measuring fitness, has hindered the empirical determination of fitness landscapes. Combining precise gene replacement and next-generation sequencing, I quantify the Darwinian fitness of over 65,000 yeast strains, each carrying a unique variant of the single-copy $\mathrm{tRNA}_{\mathrm{CCU}}^{\mathrm{Arg}}$ gene at its native genomic location under a high-temperature challenge. Analysis of single and double mutants allows for quantifying epistasis between 61% of all possible mutation pairs, showing that nearly half of all mutation pairs exhibit significant epistasis with a strong negative bias except at paired sites. Similar trends for an enrichment of negative epistasis are also revealed when focusing on mutation pairs on the genetic background with one or two existing mutations.

For Chapter 3, I measured fitness landscapes of the RNA gene in four different environments. I find pervasive genotype by environment interactions (G×E), but they follow a simple piecewise linear relationship. Both fitness and epistasis are largely predictable, lending empirical support in inferring fitness landscape and epistasis across multiple environments in future studies. For instance, the fitness landscape and epistasis of the tRNA gene in any environment can be predicted as long as it has been measured in one environment and the relative gene importance in the two environments compared is known. Epistasis is compared across environments, and the sign of epistasis remains largely unchanged across multiple environment pairs. However, there is an enrichment for switching from positive epistasis to negative epistasis in general as gene importance increases.

In Chapter 4, I switch gears from intragenic epistasis to intergenic epistasis. Genetic incompatibility, a form of epistatic interactions between otherwise functional genes in their conspecific genetic background, is commonly considered as the major cause of postzygotic isolation. The Bateson-Dobzhansky-Muller (BDM) model of reproductive isolation by genetic incompatibility is a widely accepted model of speciation. Despite repeated efforts, BDM incompatibilities between nuclear genes have never been identified between *S. cerevisiae* and its sister species *S. paradoxus*. Such negative results have led to the belief that simple nuclear BDM incompatibilities do not exist in yeast. Here I explore an alternative explanation that such incompatibilities exist but were undetectable due to limited statistical power. I evaluated previous studies and revealed the lack of statistical power due to limited sample size. I modeled the procedures of identifying genetic incompatibilities using Matlab simulation. By designing and comparing different identification strategies, I optimized the identification strategy based on various sample sizes and statistical models.

In summary, research in Chapter 2 and Chapter 3 highlights the quantification and characteristics of intragenic epistasis at a large scale, while Chapter 4 focuses on the identification of intergenic epistasis at the genomic level. Through this research, I was able to test and confirm a series of important evolutionary hypotheses on fitness landscape and epistasis and offer new insights into the underlying mechanism of evolution and strategies in studying evolutionary questions.

1.3 References

Alvarez-Castro, J. M. and O. Carlborg (2007). "A unified model for functional and statistical epistasis and its application in quantitative trait Loci analysis." Genetics **176**(2): 1151-1167.

Arias, M., Y. le Poul, M. Chouteau, R. Boisseau, N. Rosser, M. Thery and V. Llaurens (2016). "Crossing fitness valleys: empirical estimation of a fitness landscape associated with polymorphic mimicry." Proc Biol Sci **283**(1829).

Avery, L. and S. Wasserman (1992). "Ordering Gene-Function - the Interpretation of Epistasis in Regulatory Hierarchies." Trends in Genetics **8**(9): 312-316.

Azevedo, L., G. Suriano, B. van Asch, R. M. Harding and A. Amorim (2006). "Epistatic interactions: how strong in disease and evolution?" Trends Genet **22**(11): 581-585.

Azevedo, R. B. R., R. Lohaus, S. Srinivasan, K. K. Dang and C. L. Burch (2006). "Sexual reproduction selects for robustness and negative epistasis in artificial gene networks (vol 440, pg 87, 2006)." Nature **443**(7111): 598-598.

Baryshnikova, A., M. Costanzo, Y. Kim, H. M. Ding, J. Koh, K. Toufighi, J. Y. Youn, J. W. Ou, B. J. San Luis, S. Bandyopadhyay, M. Hibbs, D. Hess, A. C. Gingras, G. D. Bader, O. G. Troyanskaya, G. W. Brown, B. Andrews, C. Boone and C. L. Myers (2010). "Quantitative analysis of fitness and genetic interactions in yeast on a genome scale." Nature Methods **7**(12): 1017-U1110.

Bateson, W. (1909). Mendel’s Principles of Heredity, Cambridge University Press.

Bershtein, S., M. Segal, R. Bekerman, N. Tokuriki and D. S. Tawfik (2006). "Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein." Nature **444**(7121): 929-932.

Bloom, J. D. and F. H. Arnold (2009). "In the light of directed evolution: Pathways of adaptive protein evolution." Proceedings of the National Academy of Sciences of the United States of America **106**: 9995-10000.

Bloom, J. D., S. T. Labthavikul, C. R. Otey and F. H. Arnold (2006). "Protein stability promotes evolvability." Proc Natl Acad Sci U S A **103**(15): 5869-5874.

Bloom, J. S., I. M. Ehrenreich, W. T. Loo, T. L. V. Lite and L. Kruglyak (2013). "Finding the sources of missing heritability in a yeast cross." Nature **494**(7436): 234-237.

Byrne, A. B., M. T. Weirauch, V. Wong, M. Koeva, S. J. Dixon, J. M. Stuart and P. J. Roy (2007). "A global analysis of genetic interactions in Caenorhabditis elegans." J Biol **6**(3): 8.

Cherry, J. M., E. L. Hong, C. Amundsen, R. Balakrishnan, G. Binkley, E. T. Chan, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. R. Engel, D. G. Fisk, J. E. Hirschman, B. C. Hitz, K. Karra, C. J. Krieger, S. R. Miyasato, R. S. Nash, J. Park, M. S. Skrzypek, M. Simison, S. Weng and E. D. Wong (2012). "Saccharomyces Genome Database: the genomics resource of budding yeast." Nucleic Acids Res **40**(Database issue): D700-705.

Clark, N. L., J. Gasper, M. Sekino, S. A. Springer, C. F. Aquadro and W. J. Swanson (2009). "Coevolution of interacting fertilization proteins." PLoS Genet **5**(7): e1000570.

Costanzo, M., A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Koh, K. Toufighi, S. Mostafavi, J. Prinz, R. P. St Onge, B. VanderSluis, T. Makhnevych, F. J. Vizeacoumar, S. Alizadeh, S. Bahr, R. L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z. Y. Lin, W. Liang, M. Marback, J. Paw, B. J. San Luis, E. Shuteriqi, A. H. Tong, N. van Dyk, I. M. Wallace, J. A. Whitney, M. T. Weirauch, G. Zhong, H. Zhu, W. A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pal, F. P. Roth, G. Giaever, C. Nislow, O. G. Troyanskaya, H. Bussey, G. D. Bader, A. C. Gingras, Q. D. Morris, P. M. Kim, C. A. Kaiser, C. L. Myers, B. J. Andrews and C. Boone (2010). "The genetic landscape of a cell." Science **327**(5964): 425-431.

Costanzo, M., B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons, G. Tan, W. Wang, M. Usaj, J. Hanchard, S. D. Lee, V. Pelechano, E. B. Styles, M. Billmann, J. van Leeuwen, N. van Dyk, Z. Y. Lin, E. Kuzmin, J. Nelson, J. S. Piotrowski, T. Srikumar, S. Bahr, Y. Chen, R. Deshpande, C. F. Kurat, S. C. Li, Z. Li, M. M. Usaj, H. Okada, N. Pascoe, B. J. San Luis, S. Sharifpoor, E. Shuteriqi, S. W. Simpkins, J. Snider, H. G. Suresh, Y. Tan, H. Zhu, N. Malod-Dognin, V. Janjic, N. Przulj, O. G. Troyanskaya, I. Stagljar, T. Xia, Y. Ohya, A. C. Gingras, B. Raught, M. Boutros, L. M. Steinmetz, C. L. Moore, A. P. Rosebrock, A. A. Caudy, C. L. Myers, B. Andrews and C. Boone (2016). "A global genetic interaction network maps a wiring diagram of cellular function." Science **353**(6306).

Currin, A., N. Swainston, P. J. Day and D. B. Kell (2015). "Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently." Chemical Society Reviews **44**(5): 1172-1239.

Dean, E. J., J. C. Davis, R. W. Davis and D. A. Petrov (2008). "Pervasive and Persistent Redundancy among Duplicated Genes in Yeast." Plos Genetics **4**(7).

Feinauer, C., H. Szurmant, M. Weigt and A. Pagnani (2016). "Inter-Protein Sequence Co-Evolution Predicts Known Physical Interactions in Bacterial Ribosomes and the Trp Operon." PLoS One **11**(2): e0149166.

Fisher, R. A. (1918). "The correlation between relatives on the supposition of mendelian inheritance." Transactions of the Royal Society of Edinburgh **52**: 399-433.

Gerke, J., K. Lorenz, S. Ramnarine and B. Cohen (2010). "Gene-Environment Interactions at Nucleotide Resolution." Plos Genetics **6**(9).

Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A. P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kotter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis and M. Johnston (2002). "Functional profiling of the Saccharomyces cerevisiae genome." Nature **418**(6896): 387-391.

Giaever, G. and C. Nislow (2014). "The yeast deletion collection: a decade of functional genomics." Genetics **197**(2): 451-465.

Haag, E. S., S. Wang and J. Kimble (2002). "Rapid coevolution of the nematode sex-determining genes fem-3 and tra-2." Curr Biol **12**(23): 2035-2041.

He, X., W. Qian, Z. Wang, Y. Li and J. Zhang (2010). "Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks." Nat Genet **42**(3): 272-276.

Herter-Sprie, G. S., S. Chen, K. Hopker and H. C. Reinhardt (2011). "[Synthetic lethality as a new concept for the treatment of cancer]." Dtsch Med Wochenschr **136**(30): 1526-1530.

Jasnos, L. and R. Korona (2007). "Epistatic buffering of fitness loss in yeast double deletion strains." Nat Genet **39**(4): 550-554.

Kaelin, W. G., Jr. (2005). "The concept of synthetic lethality in the context of anticancer therapy." Nat Rev Cancer **5**(9): 689-698.

Kelley, R. and T. Ideker (2005). "Systematic interpretation of genetic interactions using protein networks." Nature Biotechnology **23**(5): 561-566.

Khan, A. I., D. M. Dinh, D. Schneider, R. E. Lenski and T. F. Cooper (2011). "Negative Epistasis Between Beneficial Mutations in an Evolving Bacterial Population." Science **332**(6034): 1193-1196.

Lehner, B. (2011). "Molecular mechanisms of epistasis within and between genes." Trends in Genetics **27**(8): 323-331.

Lehner, B., C. Crombie, J. Tischler, A. Fortunato and A. G. Fraser (2006). "Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways." Nat Genet **38**(8): 896-903.

Li, C., W. F. Qian, C. J. Maclean and J. Z. Zhang (2016). "The fitness landscape of a tRNA gene." Science **352**(6287): 837-840.

Lin, A., R. T. Wang, S. Ahn, C. C. Park and D. J. Smith (2010). "A genome-wide map of human genetic interactions inferred from radiation hybrid genotypes." Genome Res **20**(8): 1122-1132.

Malcolm, B. A., K. P. Wilson, B. W. Matthews, J. F. Kirsch and A. C. Wilson (1990). "Ancestral Lysozymes Reconstructed, Neutrality Tested, and Thermostability Linked to Hydrocarbon Packing." Nature **345**(6270): 86-89.

Mani, R., R. P. St Onge, J. L. t. Hartman, G. Giaever and F. P. Roth (2008). "Defining genetic interaction." Proc Natl Acad Sci U S A **105**(9): 3461-3466.

Moore, J. H. (2003). "The ubiquitous nature of epistasis in determining susceptibility to common human diseases." Hum Hered **56**(1-3): 73-82.

Nagalakshmi, U., Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein and M. Snyder (2008). "The transcriptional landscape of the yeast genome defined by RNA sequencing." Science **320**(5881): 1344-1349.

Olson, C. A., N. C. Wu and R. Sun (2014). "A Comprehensive Biophysical Description of Pairwise Epistasis throughout an Entire Protein Domain." Current Biology **24**(22): 2643-2651.

Ooi, S. L., X. Pan, B. D. Peyser, P. Ye, P. B. Meluh, D. S. Yuan, R. A. Irizarry, J. S. Bader, F. A. Spencer and J. D. Boeke (2006). "Global synthetic-lethality analysis and yeast functional profiling." Trends Genet **22**(1): 56-63.

Ortlund, E. A., J. T. Bridgham, M. R. Redinbo and J. W. Thornton (2007). "Crystal structure of an ancient protein: evolution by conformational epistasis." Science **317**(5844): 1544-1548.

Phillips, P. C. (2008). "Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems." Nat Rev Genet **9**(11): 855-867.

Puchta, O., B. Cseke, H. Czaja, D. Tollervey, G. Sanguinetti and G. Kudla (2016). "Network of epistatic interactions within a yeast snoRNA." Science **352**(6287): 840-844.

Ryan, O., R. S. Shapiro, C. F. Kurat, D. Mayhew, A. Baryshnikova, B. Chin, Z. Y. Lin, M. J. Cox, F. Vizeacoumar, D. Cheung, S. Bahr, K. Tsui, F. Tebbji, A. Sellam, F. Istel, T. Schwarzmuller, T. B. Reynolds, K. Kuchler, D. K. Gifford, M. Whiteway, G. Giaever, C.

Nislow, M. Costanzo, A. C. Gingras, R. D. Mitra, B. Andrews, G. R. Fink, L. E. Cowen and C. Boone (2012). "Global gene deletion analysis exploring yeast filamentous growth." Science **337**(6100): 1353-1356.

Sarkisyan, K. S., D. A. Bolotin, M. V. Meer, D. R. Usmanova, A. S. Mishin, G. V. Sharonov, D. N. Ivankov, N. G. Bozhanova, M. S. Baranov, O. Soylemez, N. S. Bogatyreva, P. K. Vlasov, E. S. Egorov, M. D. Logacheva, A. S. Kondrashov, D. M. Chudakov, E. V. Putintseva, I. Z. Mamedov, D. S. Tawfik, K. A. Lukyanov and F. A. Kondrashov (2016). "Local fitness landscape of the green fluorescent protein." Nature **533**(7603): 397-+.

Strope, P. K., D. A. Skelly, S. G. Kozmin, G. Mahadevan, E. A. Stone, P. M. Magwene, F. S. Dietrich and J. H. McCusker (2015). "The 100-genomes strains, an S. cerevisiae resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen." Genome Res **25**(5): 762-774.

Tokuriki, N., F. Stricher, L. Serrano and D. S. Tawfik (2008). "How protein stability and new functions trade off." PLoS Comput Biol **4**(2): e1000002.

Tong, A. H. Y., M. Evangelista, A. B. Parsons, H. Xu, G. D. Bader, N. Page, M. Robinson, S. Raghibizadeh, C. W. V. Hogue, H. Bussey, B. Andrews, M. Tyers and C. Boone (2001). "Systematic genetic analysis with ordered arrays of yeast deletion mutants." Science **294**(5550): 2364-2368.

Wang, A. D., N. P. Sharp, C. C. Spencer, K. Tedman-Aucoin and A. F. Agrawal (2009). "Selection, Epistasis, and Parent-of-Origin Effects on Deleterious Mutations across Environments in Drosophila melanogaster." American Naturalist **174**(6): 863-874.

Wang, L., Y. K. Lee, D. Bundman, Y. Q. Han, S. Thevananther, C. S. Kim, S. S. Chua, P. Wei, R. A. Heyman, M. Karin and D. D. Moore (2002). "Redundant pathways for negative feedback regulation of bile acid production." Developmental Cell **2**(6): 721-731.

Weinreich, D. M., N. F. Delaney, M. A. DePristo and D. L. Hartl (2006). "Darwinian evolution can follow only very few mutational paths to fitter proteins." Science **312**(5770): 111-114.

Weinreich, D. M., Y. H. Lan, C. S. Wylie and R. B. Heckendorn (2013). "Should evolutionary geneticists worry about higher-order epistasis?" Current Opinion in Genetics & Development **23**(6): 700-707.

Weinreich, D. M., R. A. Watson and L. Chao (2005). "Perspective: Sign epistasis and genetic constraint on evolutionary trajectories." Evolution **59**(6): 1165-1174.

You, L. C. and J. Yin (2002). "Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage T7." Genetics **160**(4): 1273-1281.

Yu, H., P. Braun, M. A. Yildirim, I. Lemmens, K. Venkatesan, J. Sahalie, T. Hirozane-Kishikawa, F. Gebreab, N. Li, N. Simonis, T. Hao, J. F. Rual, A. Dricot, A. Vazquez, R. R. Murray, C. Simon, L. Tardivo, S. Tam, N. Svrzikapa, C. Fan, A. S. de Smet, A. Motyl, M. E. Hudson, J. Park, X. Xin, M. E. Cusick, T. Moore, C. Boone, M. Snyder, F. P. Roth, A. L. Barabasi, J. Tavernier, D. E. Hill and M. Vidal (2008). "High-quality binary protein interaction map of the yeast interactome network." Science **322**(5898): 104-110.

Zackrisson, M., J. Hallin, L. G. Ottosson, P. Dahl, E. Fernandez-Parada, E. Landstrom, L. Fernandez-Ricaud, P. Kaferle, A. Skyman, S. Stenberg, S. Omholt, U. Petrovic, J. Warringer and A. Blomberg (2016). "Scan-o-matic: High-Resolution Microbial Phenomics at a Massive Scale." G3-Genes Genomes Genetics **6**(9): 3003-3014.

CHAPTER 2

DESCRIBING FITNESS LANDSCAPE ALLOWS FOR QUANTIFYING INTRAGENIC EPISTASIS AT A LARGE SCALE

2.1 Abstract

Fitness landscapes describe the genotype-fitness relationship and represent major determinants of evolutionary trajectories. However, the vast genotype space, coupled with the difficulty of measuring fitness, has hindered the empirical determination of fitness landscapes. Combining precise gene replacement and next-generation sequencing, we quantify Darwinian fitness under a high-temperature challenge for over 65,000 yeast strains each carrying a unique variant of the single-copy $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene at its native genomic location. Approximately 1% of single point mutations in the gene are beneficial, while 42% are deleterious. Fitness is broadly correlated with the predicted fraction of correctly folded tRNA molecules, revealing a biophysical basis of the fitness landscape. Almost half of all mutation pairs exhibit significant epistasis, which has a strong negative bias except when the mutations occur at Watson-Crick paired sites. The strong negative bias is also observed when focusing on mutations on genetic background with multiple existing mutations.

2.2 Introduction

Fitness landscapes can inform on the direction and magnitude of natural selection and elucidate evolutionary trajectories (de Visser and Krug 2014), but their empirical determination requires the formidable task of quantifying the fitness of an astronomically large number of possible genotypes. Past studies were limited to relatively few genotypes (Weinreich, Delaney et al. 2006, Lind, Berg et al. 2010). Next-generation DNA sequencing (NGS) has permitted the analysis of many more genotypes (Pitt and Ferre-D'Amare 2010, Hietpas, Jensen et al. 2011, Melamed, Young et al. 2013, Findlay, Boyle et al. 2014, Guy, Young et al. 2014, Melnikov, Rogov et al. 2014, Olson, Wu et al. 2014, Bank, Hietpas et al. 2015), but research has focused on biochemical functions (Pitt and Ferre-D'Amare 2010, Hinkley, Martins et al. 2011, Melamed, Young et al. 2013, Findlay, Boyle et al. 2014, Guy, Young et al. 2014, Melnikov, Rogov et al. 2014, Olson, Wu et al. 2014) rather than fitness. In the few fitness landscapes reported, only a small fraction of sites or combinations of mutations per gene were examined (Hietpas, Jensen et al. 2011, Findlay, Boyle et al. 2014, Melnikov, Rogov et al. 2014, Bank, Hietpas et al. 2015).

We combine gene replacement in *Saccharomyces cerevisiae* with an NGS-based fitness assay to determine the fitness landscape of a tRNA gene. tRNAs carry amino acids to ribosomes for protein synthesis, and mutations can cause diseases such as cardiomyopathy and deafness (Abbott, Francklyn et al. 2014). tRNA genes are typically shorter than 90 nucleotides, allowing coverage by a single Illumina sequencing read. We focus on $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$, which recognizes the arginine codon AGG via its anticodon 5’-CCU-3’. $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ is encoded by a single-copy nonessential gene in *S. cerevisiae* (Bloom-Ackermann, Navon et al. 2014), because AGG is also recognizable by $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{UCU}}$ via wobble pairing. Deleting the $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene (**Figure S1**; **Table 2-**

1) reduces growth rates in both fermentable (YPD) and non-fermentable (YPG) media, a problem exacerbated by high temperature (**Figure S2**).

2.3 Results

We chemically synthesized the 72-nucleotide $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene with a mutation rate of 3% per site (1% per alternate nucleotide) at 69 sites; for technical reasons, we kept the remaining three sites invariant. Using these variants, we constructed a pool of >10^5 strains, each carrying a $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene variant at its native genomic location (**Figure S1, 2-3**). Six parallel competitions of this strain pool were performed in YPD at 37°C for 24 hours. The $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene amplicons from the common starting population (T_0) and those from six replicate competitions (T_{24}) were sequenced with 100-nucleotide paired-end NGS (**Figure 2-1**; **Table 2-2**). Genotype frequencies were highly correlated between two T_0 technical repeats (Pearson's correlation r = 0.99997; **Figure S3A**) and among six T_{24} biological replicates (average r = 0.9987; **Figure S3B**). Changes in genotype frequencies between T_0 and T_{24} were used to determine the Darwinian fitness of each genotype relative to the wild-type. For our fitness estimation, we considered 65,537 genotypes with read counts ≥ 100 at T_0. In theory, a cell that does not divide has a fitness of 0.5 (Qian, Ma et al. 2012). Because $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ mutations are unlikely to be fatal, we set genotype fitness at 0.5 when the estimated fitness is < 0.5 (due to stochasticity). Fitness values from these *en masse* competitions agreed with those obtained from growth curve and pairwise competition (**Figure S4**), as reported previously (Qian, Ma et al. 2012). We observed strong fitness correlations across diverse environments for a subset of genotypes examined (**Figure S5**), suggesting that our fitness landscape is broadly relevant.

We estimated the fitness (f) of all 207 possible mutants that differ from the wild-type by one point mutation (N1 mutants), and calculated the average mutant fitness at each site (**Figure**

2-2A). Average fitness decreased to < 0.75 by mutation at nine key sites, including all three anticodon positions (**Table S3**), three TΨC loop sites, one D stem site, and two paired TΨC stem sites (**Figure 2-2A**). The TΨC loop and stem sites are components of the B Box region of the internal promoter, with C55 essential for both TFIIIC transcription factor binding and Pol III transcription (Hiraga, Botsios et al. 2012). In addition, some sites such as T54 are ubiquitously post-transcriptionally modified (Phizicky and Hopper 2010). By contrast, the average mutant fitness is ≥ 0.95 at 30 sites (**Figure 2-2A**). Overall, mutations in loops are more deleterious than in stems ($P = 0.01$, Mann-Whitney *U* test), although this difference becomes insignificant after excluding the anticodon ($P = 0.09$). Unsurprisingly, different mutations at a site have different fitness effects (**Figure S6**). For example, mutation C11T in the D stem is tolerated ($f_{\mathrm{C11T}} \pm \mathrm{SE} = 1.006 \pm 0.036$), but C11A and C11G are not ($f_{\mathrm{C11A}} = 0.676 \pm 0.030$ and $f_{\mathrm{C11G}} = 0.661 \pm 0.035$); likely due to G:U paring in RNA.

The fitness distribution of N1 mutants shows a mean of 0.89 and a peak at 1 (**Figure 2-2B**). Only 1% of mutations are significantly beneficial (nominal $P < 0.05$; *t*-test based on the six replicates), whereas 42% are significantly deleterious. We estimated the fitness of 61% of all possible genotypes carrying two mutations (N2 mutants), and observed a left-shifted distribution peaking at 0.50 and 0.67 (**Figure 2-2C**). We also estimated the fitness of 1.6% of genotypes with three mutations (N3 mutants); they exhibited a distribution with only one dominant peak at 0.5, indicating that many triple mutations completely suppress yeast growth in the *en masse* competition (**Figure 2-2D**). The fitness distribution narrows and shifts further toward 0.5 in strains carrying more than three mutations (**Figure 2-2E**).

Fitness landscapes allow predicting evolution, because sites where mutations are on average more harmful should be evolutionarily more conserved. We aligned 200 non-redundant

$\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene sequences across the eukaryotic phylogeny. The percentage of sequences having the same nucleotide as yeast at a given site is negatively correlated with the average fitness upon mutation at the site (Spearman's ρ = -0.61, $P = 2\times10^{-8}$; **Figure 2-2F**). Among N1 mutants, the number of times that a mutant nucleotide appears in the 200 sequences is positively correlated with the fitness of the mutant (ρ = 0.51, $P = 2\times10^{-15}$; **Figure 2-2G**). Furthermore, mutations observed in other eukaryotes have smaller fitness costs in yeast than those unobserved in other eukaryotes ($P = 9\times10^{-6}$, Mann–Whitney *U* test).

Two mutations may interact with each other, creating epistasis ε, with functional and evolutionary implications (Phillips 2008). We estimated ε within the tRNA gene from the fitness of 12,985 N2 mutants and 207 N1 mutants (**Figure 2-3A**). ε is negatively biased, with only 34% positive values ($P < 10^{-300}$, binomial test; **Figures 2-3B**, **S7A**, **S8**). Forty-five percent of ε values differ significantly from 0 (nominal $P < 0.05$, *t*-test based on the six replicates), among which 86% are negative ($P < 10^{-300}$, binomial test; **Figures 2-3B**, **S7A**, **S8**). Consistent with the overall negative ε, the mean fitness of N2 mutants (0.75) is lower than that predicted from N1 mutants assuming no epistasis (0.81) (**Figure 2-2E**). Interestingly, as the first mutation becomes more deleterious, the mean epistasis between this mutation and the next mutation becomes less negative and in some cases even positive (**Figures 2-3C**, **S9**), similar to between-gene epistasis involving an essential gene (He, Qian et al. 2010). Consequently, the larger the fitness cost of the first mutation, the smaller the mean fitness cost of the second mutation (**Figures 2-3D**, **S10**). Pairwise epistasis involving three or four mutations is also negatively biased (**Figure S11**). Consistently, N3 to N8 mutants all show lower average fitness than expected assuming no epistasis (**Figure 2-2E**).

The distribution of epistasis between mutations at paired sites is expected to differ from the above general pattern, because different Watson-Crick (WC) pairs may be functionally similar (Meer, Kondrashov et al. 2010). We estimated the fitness of 71% of all possible N2 mutants at WC paired sites. Among the 41 cases that switched from one WC pair to another, 23 (56%) have positive ε (**Figure 2-3E**). Among the 80 N2 mutants that destroyed WC pairing, 39 (49%) showed positive ε (**Figure 2-3F**). The ε values are more positive for each of these two groups than for N2 mutants where the two mutations do not occur at paired sites ($P = 7\times10^{-6}$ and 2.6×10^{-3}, respectively, Mann-Whitney U test). Furthermore, ε is significantly more positive in the 41 cases with restored WC pairing than the 80 cases with destroyed pairing ($P = 0.04$). These two trends also apply to cases with significant epistasis (corresponding $P = 3\times10^{-5}$, 0.01, and 0.01, respectively; **Figures 2-3EF**, **S7BC)**. Nevertheless, epistasis is not always positive between paired sites, likely because base pairing is not the sole function of the nucleotides at paired sites. We observed 160 cases of significant sign epistasis , which is of special interest because it may block potential paths for adaptation (Weinreich, Delaney et al. 2006). We also detected ε with opposite signs in different genetic backgrounds, a high-order epistasis (**Table S4**).

A tRNA can fold into multiple secondary structures. We computationally predicted the proportion of $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ molecules that are potentially functional (i.e., correctly folded, no anticodon mutation) for each genotype (P_{func}). Raising P_{func} increases fitness ($\rho = 0.40$, $P < 10^{-300}$) albeit with diminishing returns (**Figure 2-4A)**, and this correlation holds after controlling for mutation number ($\rho = 0.26$, 0.37, and 0.24 for N1, N2, and N3 mutants, respectively). Because computational prediction of RNA secondary structures is only moderately accurate, the P_{func}–fitness correlation demonstrates an important role of P_{func} in shaping the tRNA fitness landscape. Nonetheless, after controlling for P_{func}, mutant fitness still correlates with mutation number (ρ = -0.51, $P < 10^{-300}$; see also LOESS regressions for N1, N2, and N3 mutants in **Figure 2-4B**), suggesting that other factors also impact fitness.

To investigate whether P_{func} explains epistasis, we computed epistasis using the fitness of N1 and N2 mutants predicted from their respective P_{func}–fitness regression curves (**Figure 2-4B**), and observed a significant correlation between the predicted and observed epistasis (ρ = 0.04, $P = 2.7\times10^{-5}$). The weakness of this correlation is at least partly due to the fact that epistasis is computed from three fitness measurements (or predictions) and therefore associated with a considerable error. There is a similar bias in predicted epistasis toward negative values (**Figure 2-4C**), but further analyses suggest that it probably arises from factors other than tRNA folding. These results regarding P_{func} and epistasis are not unexpected given that a tRNA site can be involved in multiple molecular functions (Phizicky and Hopper 2010, Hiraga, Botsios et al. 2012).

2.4 Discussion

In summary, we described the *in vivo* fitness landscape of a yeast tRNA gene under a high-temperature challenge. Broadly consistent with the neutral theory, beneficial mutations are rare (1%), relative to deleterious (42%) and (nearly) neutral mutations (57%). We found widespread intragenic epistasis between mutations, consistent with studies of smaller scales (de Visser and Krug 2014).

Intriguingly, 86% of significant epistasis is negative, indicating that the fitness cost of the second mutation is on average greater than that of the first. A bias toward negative epistasis was also observed in protein genes and RNA molecules (Bershtein, Segal et al. 2006, Melamed, Young et al. 2013, Olson, Wu et al. 2014, Bank, Hietpas et al. 2015, Puchta, Cseke et al. 2016), suggesting that this may be a general trend. Variation in fitness is partially explained by the predicted fraction of correctly folded tRNA molecules, suggesting general principles underlying complex fitness landscapes. Our tRNA variant library provides a resource in which various mechanisms contributing to its fitness landscape can be evaluated and the methodology developed here is applicable to the study of fitness landscapes of longer genomic segments including protein genes.

2.5 Materials and Methods

2.5.1 Media

Standard YPD (1% yeast extract, 2% peptone, 2% glucose) and YPG (1% yeast extract, 2% peptone, 3% glycerol) media were used as indicated. These two media differ in the carbon source, with YPD providing glucose as a fermentable carbon source and YPG providing glycerol as a non-fermentable carbon source. Complete Supplement Media (CSM) used contained 0.017% yeast nitrogen base without amino acids, 0.5% ammonium sulfate, 2% glucose, with addition of appropriate CSM drop-out mix as outlined by the manufacturer (Clontech).

2.5.2 Assessing the fitness effects of $tRNA^{Arg}_{CCU}$ gene deletion across environments

Two different environments are needed in this experiment. The first is a permissive environment, for collection of transformants, in which growth rate differences among tRNA variant-carrying cells are minimized in order to maximize equal representation in the initial tRNA gene variant pool. A second selective environment is then required where the fitness variation among cells carrying different tRNA gene variants is maximized, allowing a fitness landscape to be determined.

To identify these two environments, we first replaced the single-copy wild-type $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene (standard gene name *HSX1*) with *LEU2* in the haploid strain BY4742 (*MATα; his3Δ 1; leu2Δ 0; lys2Δ 0; ura3Δ 0; hsx1::LEU2*) (**Figure S1**), followed by confirmation by Sanger sequencing. Growth curves for the wild-type strain and the strain lacking the $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene were then determined by optical density at 600 nm every 15 minutes for 24 hours using a Synergy H1 Microplate Reader across multiple environments, consisting of combinations of two media (YPD and YPG) and three temperatures (30, 35 and 37°C) (**Figure S2**). The highest slope of the growth curve during the log phase was calculated for each growth curve following a previously established method (Zorgo, Gjuvsland et al. 2012). For the reasons outlined in the previous paragraph, YPD at room temperature was chosen as the condition for transformation, while YPD at 37°C was chosen as the condition for fitness landscape determination.

2.5.3 Chemical synthesis of yeast tRNA gene variants

The yeast $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene was chemically synthesized by IDT (https://www.idtdna.com/site). IDT cannot synthesize oligonucleotides longer than 100 nucleotides with sequence variations that require manual mixing of nucleotides. With this limit of the total length and the need for constant regions at the two ends of the oligonucleotides for polymerase chain reaction (PCR), 69 variable sites are allowed. That is, the first nucleotide and last two nucleotides (counting from the 5’ end) of the 72-nucleotide gene were invariant and synthesized according to the wild-type sequence. At each of the 69 variable sites, the probability of incorporating the wild-type nucleotide was set at 0.97, while the probability of incorporating each of the other three nucleotides was 0.01. The DNA sequence synthesized is TTCAACCAAGTTGGttccgtttgcgtaatggtaacgcgtctccctcctaaggagaagactgcgggttcgagtcccgtacggaa

CGTTGATTATTTTTTTT, where capital letters indicate the nucleotides at invariant sites while lower-case letters indicate the nucleotides with 97% probability at variable sites. The underlined region corresponds to the tRNA gene, whereas the flanking non-underlined regions are used for fusion PCR and homologous recombination (**Figure S1**). Using 97% wild-type nucleotides at each variable position maximizes the fraction of variants carrying two mutations, facilitating the study of pairwise epistasis. In the pool of tRNA gene variants synthesized, the fractions of molecules with 0, 1, 2, 3, 4 and >4 mutations are expected to be 12%, 26%, 27%, 19%, 10%, and 6%, respectively, while the possible numbers of variants with 0, 1, 2, 3, and 4 mutations are 1, 207, 2.1×10^4, 1.4×10^6, and 7.0×10^7, respectively. Sanger sequencing of 24 randomly picked variants confirmed that the synthesis was as expected and contained no indel. The mutation rate was estimated to be 3.2±0.29%, not significantly different from the expected value of 3%, and different base changes were roughly equally frequent.

2.5.4 Construction of the tRNA gene variant strain pool

The pool of synthesized single-stranded oligonucleotides were amplified by PCR and then fused with the *URA3* marker gene by PCR (**Figure S1**). High fidelity AccuPrime™ *Pfx* DNA polymerase was used in all PCR reactions. The tRNA gene deletion strain (*MATα; his3Δ 1; leu2Δ 0; lys2Δ 0; ura3Δ 0; hsx1::LEU2*) was then transformed with the *tRNA-URA3* variant cassette to integrate a single tRNA gene variant and to simultaneously remove *LEU2* at the native tRNA gene locus. Over 100,000 colonies were collected from CSM minus uracil plates by washing with sterile water. The large number of colonies collected ensured the inclusion of a large number of tRNA gene variants. Pooled variants were stored in 20% glycerol at -80°C.

2.5.5 Competition

A frozen sample of cells carrying tRNA gene variants was removed from storage at -80°C and allowed to revive at 30°C in YPD for 3 hours. Six replicate competitions were then started by dilution of this common starter population into six 50 ml Falcon tubes, each containing 25 ml of YPD at 37°C at an initial OD660 = 0.1. Each culture was maintained at 250 RPM in a shaking incubator and diluted to OD660 = 0.1 through transfer to fresh 25 ml of YPD media every 12 hours, at which time population aliquots were also frozen in 20% glycerol at -80°C. The competitions lasted for 24 hours.

2.5.6 Library preparation, Illumina sequencing, and read mapping

DNA was extracted from thawed population aliquots of interest. We amplified the tRNA gene from cell populations using two rounds of PCR to ensure that only those tRNA gene variants that are inserted at the native location were amplified (**Figure S1**; **Table S1**). Two technical repeats of the starting population before competition (T_0) and six biological replicates from the populations after 24 hours in competition (T_{24}) were subjected to 100-nucleotide paired-end Illumina sequencing. Paired reads for the tRNA gene sequence are required to be identical to be counted. Read counts are combined across technical repeats or biological replicates in subsequent analyses unless otherwise noted. To ensure relative accuracy in fitness estimation, 65,537 genotypes with a total of at least 100 reads in the two technical repeats at T_0 were analyzed.

2.5.7 PCR and sequencing errors

The error rate for Illumina sequencing is 3×10^{-4} per site per read (http://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf). Thus, due to sequencing error, a genotype is expected to lose $U = [1-(1-3\times10^{-4})^{2\times69}]M_0$ read pairs, where M_0 is the true number of read pairs of the genotype. Because the fractional loss $U/M_0 = 0.04$ is a constant for all genotypes including the wild-type in each sample, the loss of reads due to sequencing error does not affect fitness estimation. Sequencing error also causes the genotype to gain on average $V = (3\times10^{-4}/3)^2M_1 = 10^{-8}M_1$ read pairs, where M_1 is the total number of reads for all neighbors of the focal genotype (i.e., the genotypes that differ from the focal genotype by one nucleotide). Thus, the fractional gain of read pairs for the genotype is expected to be $V/M_0 = 10^{-8}M_1/M_0$, which has virtually no impact on fitness estimation in our study. For instance, at T_0, for the wild-type, M_1/M_0 is expected to be ~2; for an N1 genotype, M_1/M_0 is expected to be 99; for an N2 genotype, M_1/M_0 is expected to be 2×99 = 198; and so on. Hence, the fractional gain of read pairs is $<10^{-5}$ for genotypes with no more than 10 mutations.

We similarly estimated the impact of PCR error. AccuPrime$^{\text{TM}}$ *Pfx* DNA polymerase used in PCR has a very low error rate of 2.9×10^{-6} per nucleotide incorporated (https://tools.thermofisher.com/content/sfs/brochures/711-021834%20AccuPrime%20Brochu.pdf). Each of the two PCRs used in sequencing library preparation had 30 cycles. Because later cycles are inefficient, we considered effectively 50 cycles total for the two PCRs. Thus, due to PCR error, a genotype is expected to lose $U = (2.9\times10^{-6}\times69\times50)\ M_0$ molecules, where M_0 is the true number of DNA molecules of the genotype, 69 is the sequence length in nucleotides, and 50 is the total number of PCR cycles. Because the fractional loss $U/M_0 = 0.01$ is a constant for all genotypes in each sample, the loss of molecules due to PCR error does not affect fitness estimation. Sequencing error also causes the genotype to gain on average $V = 2.9\times10^{-6}\times50/3\ M_1 = 4.8\times10^{-5}M_1$ molecules, where M_1 is the total number of molecules for all neighbors of the focal genotype. Thus, the fractional gain of molecules for the genotype is expected to be $V/M_0 = 4.8\times10^{-5}M_1/M_0$, which has little impact on fitness estimation in our study. As mentioned, at T_0, for the wild-type, M_1/M_0 is expected to be ~2; for an N1 genotype, M_1/M_0 is expected to be 99; for an N2 genotype, M_1/M_0 is expected to be 2×99 = 198; and so on. Hence, the fractional gain in the number of molecules is < 0.024 for genotypes with no more than 5 mutations.

To independently validate the above calculations that are based on the published sequencing and PCR error rates, we estimated the upper bound rate of error caused by PCR and sequencing. It is very unlikely for any N1 genotype at T_{24} to have all of its reads arising from its neighbors by PCR and sequencing errors. Thus, by assuming that all these reads are from errors, we can estimate the upper bound error rate. We calculated the frequency of each N1 genotype in each replicate in T_{24} and identified the smallest frequency among the total of 207×6=1242 frequencies and the corresponding genotype. We then divided the total number of read pairs for this genotype in the other five replicates by the total number of read pairs for its neighbors in these five replicates. The result, 5.0×10^{-5}, is an upper bound estimate of the probability that a genotype "mutates" to a specific neighbor, or V defined earlier. Interestingly, this upper bound estimate of V from the sum of PCR error and the much lower sequencing error is virtually identical to that calculated based on the published PCR error rate. Together, these analyses suggest that PCR and sequencing errors have minimal impacts on our fitness estimation.

2.5.8 Number of generations of competition

During competition, cell populations were diluted every 12 hours, with OD660 recorded before and after dilution. From the sequencing results, we calculated the frequency of the wild-type at the beginning of the competition (F_0) and at 24 hours in competition (F_{24}). The number of wild-type generations for the 24 hours is then $G = \log_2(dgF_{24}/\mathrm{F}_0) = 11.5$, where d is the dilution factor and g is the ratio between the cell number calculated from OD660 (Zorgo, Gjuvsland et al. 2012) at 24 hours and that at the beginning of the competition.

2.5.9 Estimating relative fitness from read frequency changes

Using a method previously designed for the fitness estimation of gene deletion strains (Bar-seq) (Smith, Heisler et al. 2009), we estimated the relative fitness of each strain by using the 72-nucleotide tRNA sequence that it carries as the barcode to directly determine its abundance within the population at each time point. The per generation Darwinian fitness of a variant relative to the wild-type is

$$\text{Fitness} = \left(\frac{\text{\# of reads for the variant at } T_{24} \Big/ \text{\# of reads for the variant at } T_0}{\text{\# of reads for the wild-type at } T_{24} \Big/ \text{\# of reads for the wild-type at } T_0} \right)^{1/G}$$

, where G =11.5 is the number of wild-type generations in 24 hours. By definition, the wild-type fitness is 1. If the read number drops to 0 in all six replicates or if fitness drops under 0.5, fitness was assigned to be 0.5, representing no cell division for this variant. There are six biological replicates, and we used a *t*-test to examine if the fitness of a variant is significantly different from 1 at a nominal *P* value of 5%. Comparison of read frequencies obtained from populations before and after competition corrects for frequency differences among genotypes in the starting population and potential biases in variant-specific PCR amplification efficiency, sequencing library preparation efficiency, as well as any Illumina sequencing efficiency and accuracy differences that may exist.

Because the tRNA gene deletion strain can grow (**Figure S2**), the gene is nonessential. However, the deletion strain was not in the pool of genotypes that underwent Bar-seq, so we could not directly compare a genotype with the deletion strain to examine the potential existence of dominant negative effects. For a genotype to have a computed fitness <0.5, its frequency relative to the wild-type (w) must decrease by at least $2^{11.5}$= 2896 folds from T_0 to T_{24}. For an average N2 mutant, the expected w at T_0 is 10^{-4}. So, the expected w is $<3\times10^{-8}$ in T_{24} if its fitness is <0.5. Given the bottleneck population size of $\sim3\times10^{7}$ (at dilution) and final population size of $\sim1.7\times10^{9}$ in the competition, such a small w means that the corresponding cell number is very low. Thus, the fate of the genotype depends largely on genetic drift. In other words, the formula for estimating mutant fitness in the previous paragraph, which ignores genetic drift, would not work well. Given the known function of tRNAs, the most likely reason for potential dominant negative effects would be anticodon mutations. Yet, most mutants with anticodon mutations have fitness >0.5 (**Table S3**). Thus, dominant negative effects probably do not exist here, but further studies are certainly required to confirm this point.

2.5.10 Fitness estimation from growth curves

In order to verify our *en masse* fitness estimates, we isolated 55 strains from the variant pool with distinct tRNA gene sequences based on Sanger sequencing. The growth rates of the 55 strains were measured using Bioscreen C OD reader at 37°C in YPD. Cells were grown at room temperature overnight until saturation, and then diluted by a factor of 50 to roughly OD600 = 0.1. OD measurements at wide band (450-580 nm) were taken every 20 minutes for 48 hours. Proliferation efficiency and maximum growth rate were calculated following standard procedures from measurement 10 to 72 (Warringer, Zörgö et al. 2011). Two biological replications in fitness measurement were performed per genotype.

2.5.11 Fitness estimation by pairwise competition

To further confirm our *en masse* fitness estimates, we performed pairwise competition assays of the 55 strains against a fluorescent reference strain. The reference strain, YCM2644 (*MATα; his3Δ 1; leu2Δ 0; lys2Δ 0; URA3; ho::TDH3p-VenusYFP-HygMX4*), was constructed by replacing the *HO* gene in a strain that carries the native tRNA gene and *URA3* with a cassette comprised of a yellow florescent protein (YFP) gene and a hygromycin resistance gene. We also competed between the reference stain and one of the constructed variant strains that happens to carry the wild-type tRNA gene (referred to as the wild-type variant). The competition procedure followed that in the main experiment. Samples were collected at 0 and 24 hours, and the fitness of each variant strain relative to that of the wild-type variant was calculated following an established protocol (He, Qian et al. 2010). Three biological replications in fitness measurement were performed.

2.5.12 Comparing growth rates across multiple environments

We measured the growth rates of the aforementioned 55 strains using Bioscreen C OD reader in four environments at 30°C: YPD, YPD with 7% EtOH, YPD with 3% DMSO, and YPD with 0.85M NaCl. Cells were grown at room temperature overnight until saturation, and then diluted by a factor of 100 to roughly OD600 = 0.1. OD measurements at wide band (450-580 nm) were taken every 20 minutes for 24 hours. Maximum growth rates were calculated as described above. Three biological replications were performed. Growth rates in these four environments and those in YPD at 37°C were compared for the 55 strains.

2.5.13 Phylogenetic data of the tRNA genes

From GtRNAdb (http://gtrnadb.ucsc.edu), we downloaded 1098 eukaryotic tRNA genes with anticodon CCU. A total of 416 distinct sequences were aligned using the cmalign program in the Infernal package (http://infernal.janelia.org/), which aligns tRNA sequences based on both the primary sequence and the secondary structure. The region corresponding to the 72-nucleotide segment of yeast tRNA was extracted for further analysis. To acquire a good representation of the tRNA sequence variation over evolutionary time and avoid oversampling from certain well-studied groups of organisms, we calculated the pairwise sequence distances among the tRNA genes, and randomly removed one of the two sequences with the smallest distance until the distance between any two sequences in the dataset is at least 7 nucleotides (~10%). A total of 200 sequences remained in this dataset. We also examined a subset of 23 sequences, each having at least 20 nucleotide differences from any other sequences in the subset, and obtained qualitatively similar results as those from the 200 sequences.

2.5.14 Estimating epistasis from fitness values

Epistasis is defined as $\varepsilon = f_{AB} - f_A f_B$, where f_{AB} is the fitness of a N2 mutant and f_A and f_B are the fitness of the two corresponding N1 mutants. ε is computed by $f_{AB} - 0.5$ when $f_A f_B < 0.5$. The overall distribution of ε is unaffected when we exclude 854 cases in which f_{AB} or $f_A f_B$ is ≤ 0.5 (**Figure S7A**). To examine if ε differs significantly from 0, we estimated ε from each of the six biological replicates and conducted a *t*-test.

In the tRNA under study, there are 20 paired stem sites, with 1 being wobble pairing (GU) and the rest 19 being WC pairing. Pairwise epistasis is more likely to be positive when two mutations occur at paired sites than when they do not occur at paired sites. It is easy to understand why epistasis is positive when the second mutation restores pairing after the first mutation breaks it. Further, even when pairing is not restored by the second mutation, epistasis could be positive, likely because the second mutation does no more harm to pairing when the pairing has been broken by the first mutation. Nevertheless, even at paired sites, epistasis is not always positive, suggesting that base pairing is not the sole function of the nucleotides at paired sites such that the second mutation, regardless of whether it restores pairing, could do additional harm.

Sign epistasis occurs when the fitness effect of a mutation is opposite depending on the presence or absence of another mutation. That is, sign epistasis satisfies $(f_A\text{-}1)(f_{AB}/f_B\text{-}1) < 0$ or $(f_B\text{-}1)(f_{AB}/f_A\text{-}1) < 0$. Reciprocal sign epistasis satisfies $(f_A\text{-}1)(f_{AB}/f_B\text{-}1) < 0$ and $(f_B\text{-}1)(f_{AB}/f_A\text{-}1) < 0$. Statistical significance was determined as described above. In total, 160 cases of significant sign epistasis were found, 6 of which were reciprocal sign epistasis. Interestingly, 75 of the 160 cases involved T8C and 54 involved A70C, suggesting that sign epistasis is highly concentrated at a few sites. The reason why T8C and A70C are concentrated in significant sign epistasis cases is that they are the only significantly beneficial mutations in the wild-type background, but they apparently are deleterious in many N1 backgrounds. Among the 160 sign epistasis cases, 9 involve paired sites in stems, significantly more than the chance expectation ($P < 0.05$, χ^2 test). This is not unexpected, because a mutation at a stem site could either destroy or restore a base pair depending on the presence or absence of the pairing before the mutation.

We predicted the mean fitness of mutants carrying *n* mutations from the fitness of N1 mutants, under the assumption of no epistasis (red circles in **Figure 2-2E**). For each mutant with *n* mutations, the predicted fitness is the product of the fitness of the constituent N1 mutants or

0.5 if the product is < 0.5. We then averaged the predicted fitness for all mutants with *n* mutations.

2.5.15 Varying pairwise epistasis in different genetic backgrounds

To examine whether the sign of pairwise epistasis varies depending on the genetic background, we compared the epistasis between mutations A and B in the wild-type and that in the N1 mutant carrying mutation C. Epistasis in the wild-type is calculated by $\varepsilon_{AB}|WT = f_{AB} - f_A f_B$, while epistasis in the N1 mutant is calculated by $\varepsilon_{AB}|C = f_{ABC}/f_C - (f_{AC}/f_C)(f_{BC}/f_C)$. The expected fitness $f_A f_B$ and $(f_{AC}/f_C)(f_{BC}/f_C)$ are set to be 0.5 if smaller than 0.5.

2.5.16 An alternative measure of pairwise epistasis

We also used an alternative definition of epistasis based on an additive model of the logarithm of fitness, $\varepsilon' = \ln(f_{AB}) - \ln f_A - \ln f_B$, to calculate all pairwise epistasis (**Figure S8**), but found the general pattern unchanged.

2.5.17 Structural stability and fraction of correctly folded tRNA molecules

The function/fitness relevance of mutational impacts on protein structure stability is well known (Jacquier, Birgy et al. 2013), and we here examine the importance of RNA structure stability to fitness. The secondary structure of the wild-type tRNA follows http://lowelab.ucsc.edu/GtRNAdb/Sacc_cere/Sacc_cere-structs.html. For each tRNA variant, a series of suboptimal secondary structures and the corresponding minimum free energy (*E*) were predicted from the "subopt" function in the Vienna RNA package (Lorenz, Bernhart et al. 2011) at 37°C. The predicted proportion of functional tRNA molecules is estimated by $P_{\text{func}} = [\sum_i (J_i e^{\frac{-E_i}{kT}})] / (\sum_i e^{\frac{-E_i}{kT}})$. Here, i refers to the ith considered secondary structure, J_i is an identity function, taking the value of 0 if the structure i is not functional and 1 if it is functional, E_i is the minimum free energy of the ith structure, k = 0.001987 kcal/mol/K is the Boltzmann constant, and T = 310 K is the absolute temperature corresponding to 37°C. We considered only those structures whose E values are smaller than 3 kcal/mol + the E value of the most stable structure for the mutant concerned. If the wild-type structure is not included within the 3 kcal/mol range, we add it (with the constraint of the mutant sequence) to the list of predicted structures, with E predicted by the energy_of_struct function of the Vienna package. A structure is considered functional when it satisfies two criteria. First, no base pairing occurs at any position of the anticodon and no mutation occurs at any position of the anticodon. Second, the distance between the structure considered and the wild-type structure does not exceed d = 2. The distance was calculated by the RNAdistance function in the Vienna package. We varied the parameter d between 0 and 16 and found the result qualitatively unchanged.

2.5.18 LOESS regression and prediction of epistasis from P_{func}

LOESS regression in R was used to summarize the relationship (and 95% confidence interval) between fitness and P_{func} for N1, N2, and N3 mutants, respectively. The span parameter α was set at 1 and all other parameters were as default. For each N2 mutant, epistasis is predicted in the following manner. First, the P_{func} of the N2 mutant is computed as described above. Second, the corresponding fitness is predicted using the LOESS curve for N2 mutants. Third, P_{func} is computed for each of the two corresponding N1 mutants. Fourth, fitness is predicted for each of the two N1 mutants from the LOESS curve for N1 mutants. Fifth, epistasis is then calculated based on the three predicted fitness values as if they are observed fitness values.

When we predicted epistasis using a single LOESS curve, the predicted epistasis is positively biased. For instance, when only the N1 LOESS curve is used, the mean predicted epistasis is 0.07. When only the N2 curve is used, the mean predicted epistasis is 0.16. When a combined LOESS curve for N1 and N2 mutants is used, the mean predicted epistasis is 0.16. We also found that epistasis in P_{func} in overall positive.

2.6 References

Abbott, J. A., C. S. Francklyn and S. M. Robey-Bond (2014). "Transfer RNA and human disease." Front Genet **5**: 158.

Bank, C., R. T. Hietpas, J. D. Jensen and D. N. Bolon (2015). "A systematic survey of an intragenic epistatic landscape." Mol Biol Evol **32**(1): 229-238.

Bershtein, S., M. Segal, R. Bekerman, N. Tokuriki and D. S. Tawfik (2006). "Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein." Nature **444**(7121): 929-932.

Bloom-Ackermann, Z., S. Navon, H. Gingold, R. Towers, Y. Pilpel and O. Dahan (2014). "A comprehensive tRNA deletion library unravels the genetic architecture of the tRNA pool." PLoS Genet **10**(1): e1004084.

de Visser, J. A. and J. Krug (2014). "Empirical fitness landscapes and the predictability of evolution." Nat Rev Genet **15**(7): 480-490.

Findlay, G. M., E. A. Boyle, R. J. Hause, J. C. Klein and J. Shendure (2014). "Saturation editing of genomic regions by multiplex homology-directed repair." Nature **513**(7516): 120-123.

Guy, M. P., D. L. Young, M. J. Payea, X. Zhang, Y. Kon, K. M. Dean, E. J. Grayhack, D. H. Mathews, S. Fields and E. M. Phizicky (2014). "Identification of the determinants of tRNA function and susceptibility to rapid tRNA decay by high-throughput in vivo analysis." Genes Dev **28**(15): 1721-1732.

He, X., W. Qian, Z. Wang, Y. Li and J. Zhang (2010). "Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks." Nat Genet **42**(3): 272-276.

Hietpas, R. T., J. D. Jensen and D. N. Bolon (2011). "Experimental illumination of a fitness landscape." Proc Natl Acad Sci U S A **108**(19): 7896-7901.

Hinkley, T., J. Martins, C. Chappey, M. Haddad, E. Stawiski, J. M. Whitcomb, C. J. Petropoulos and S. Bonhoeffer (2011). "A systems analysis of mutational effects in HIV-1 protease and reverse transcriptase." Nat Genet **43**(5): 487-489.

Hiraga, S., S. Botsios, D. Donze and A. D. Donaldson (2012). "TFIIIC localizes budding yeast ETC sites to the nuclear periphery." Mol Biol Cell **23**(14): 2741-2754.

Jacquier, H., A. Birgy, H. Le Nagard, Y. Mechulam, E. Schmitt, J. Glodt, B. Bercot, E. Petit, J. Poulain, G. Barnaud, P. A. Gros and O. Tenaillon (2013). "Capturing the mutational landscape of the beta-lactamase TEM-1." Proc Natl Acad Sci U S A **110**(32): 13067-13072.

Lind, P. A., O. G. Berg and D. I. Andersson (2010). "Mutational robustness of ribosomal protein genes." Science **330**(6005): 825-827.

Lorenz, R., S. H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler and I. L. Hofacker (2011). "ViennaRNA Package 2.0." Algorithms Mol Biol **6**: 26.

Meer, M. V., A. S. Kondrashov, Y. Artzy-Randrup and F. A. Kondrashov (2010). "Compensatory evolution in mitochondrial tRNAs navigates valleys of low fitness." Nature **464**(7286): 279-282.

Melamed, D., D. L. Young, C. E. Gamble, C. R. Miller and S. Fields (2013). "Deep mutational scanning of an RRM domain of the Saccharomyces cerevisiae poly(A)-binding protein." RNA **19**(11): 1537-1551.

Melnikov, A., P. Rogov, L. Wang, A. Gnirke and T. S. Mikkelsen (2014). "Comprehensive mutational scanning of a kinase in vivo reveals substrate-dependent fitness landscapes." Nucleic Acids Res **42**(14): e112.

Olson, C. A., N. C. Wu and R. Sun (2014). "A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain." Curr Biol **24**(22): 2643-2651.

Phillips, P. C. (2008). "Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems." Nat Rev Genet **9**(11): 855-867.

Phizicky, E. M. and A. K. Hopper (2010). "tRNA biology charges to the front." Genes Dev **24**(17): 1832-1860.

Pitt, J. N. and A. R. Ferre-D'Amare (2010). "Rapid construction of empirical RNA fitness landscapes." Science **330**(6002): 376-379.

Puchta, O., B. Cseke, H. Czaja, D. Tollervey, G. Sanguinetti and G. Kudla (2016). "Network of epistatic interactions within a yeast snoRNA." Science **352**(6287): 840-844.

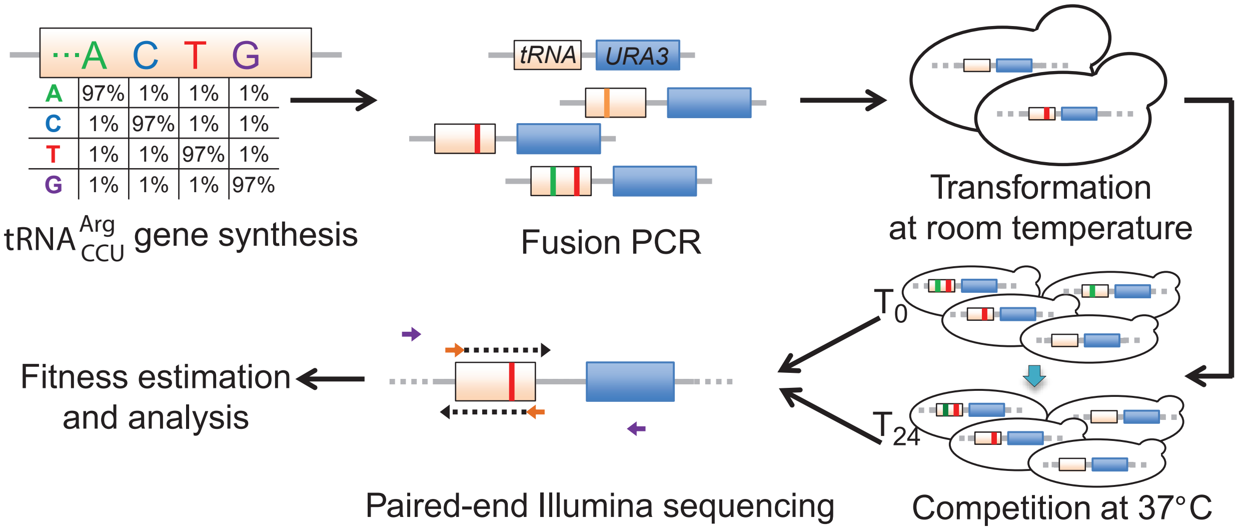
Qian, W., D. Ma, C. Xiao, Z. Wang and J. Zhang (2012). "The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast." Cell Rep **2**(5): 1399-1410.

Smith, A. M., L. E. Heisler, J. Mellor, F. Kaper, M. J. Thompson, M. Chee, F. P. Roth, G. Giaever and C. Nislow (2009). "Quantitative phenotyping via deep barcode sequencing." Genome Res **19**(10): 1836-1842.

Warringer, J., E. Zörgö, F. A. Cubillos, A. Zia, A. Gjuvsland, J. T. Simpson, A. Forsmark, R. Durbin, S. W. Omholt, E. J. Louis, G. Liti, A. Moses and A. Blomberg (2011). "Trait Variation in Yeast Is Defined by Population History." PLoS Genet **7**(6): e1002111.

Weinreich, D. M., N. F. Delaney, M. A. Depristo and D. L. Hartl (2006). "Darwinian evolution can follow only very few mutational paths to fitter proteins." Science **312**(5770): 111-114.

Zorgo, E., A. Gjuvsland, F. A. Cubillos, E. J. Louis, G. Liti, A. Blomberg, S. W. Omholt and J. Warringer (2012). "Life history shapes trait heredity by accumulation of loss-of-function alleles in yeast." Mol Biol Evol **29**(7): 1781-1789.



A	97%	1%	1%	1%
C	1%	97%	1%	1%
T	1%	1%	97%	1%
G	1%	1%	1%	97%

Figure 2-1. Determining the fitness landscape of the yeast $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene. Chemically synthesized $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene variants are fused with the marker gene *URA3* before placed at the native $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ locus. The tRNA variant-carrying cells are competed. Fitness of each $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ genotype relative to wild-type is calculated from the relative frequency change of paired-end sequencing reads covering the tRNA gene variant during competition.

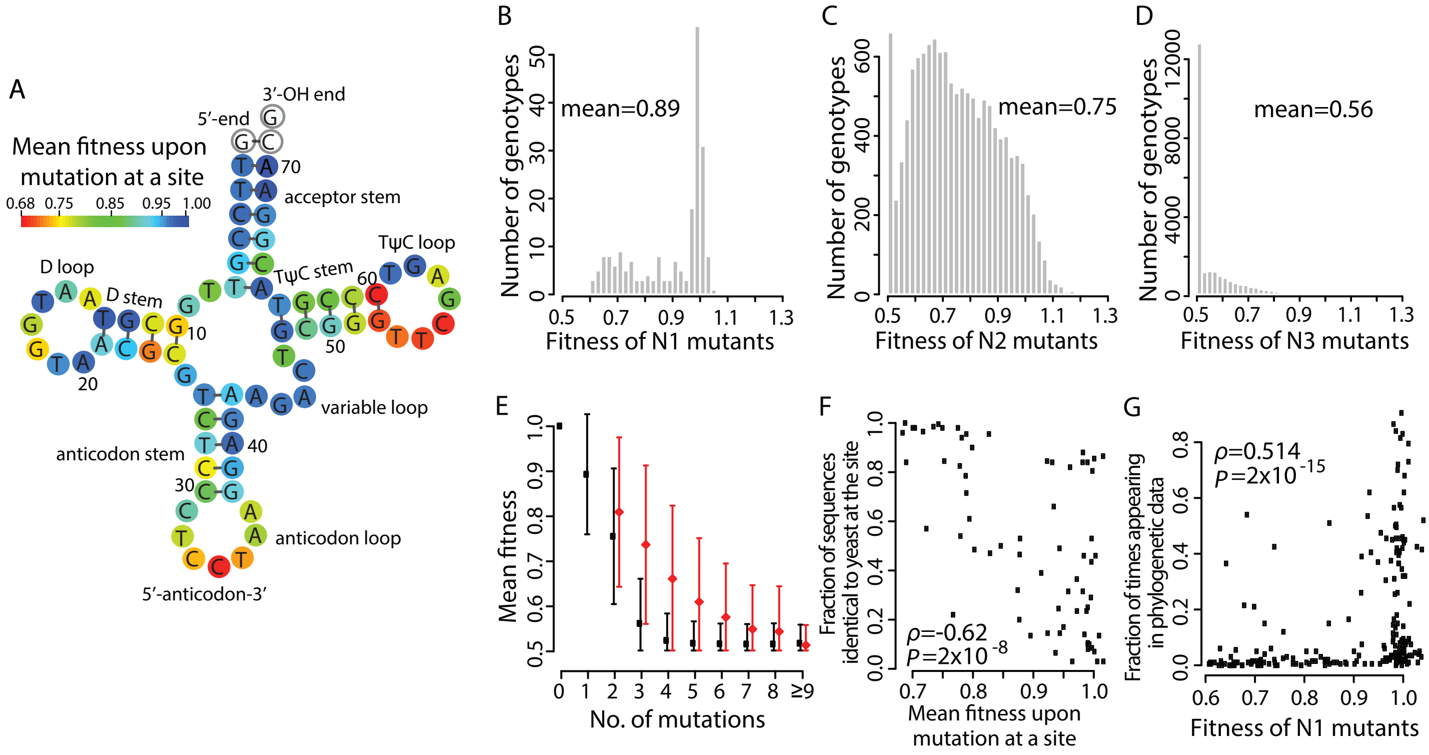


Figure 2-2. Yeast $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ **gene fitness landscape**. (**A**) Average fitness upon a mutation at each site. White circles indicate invariant sites. (**B**-**D**) Fitness distributions of (**B**) N1, (**C**) N2, and (**D**) N3 mutants, respectively. (**E**) Mean observed fitness (black circles) decreases with mutation number. Red circles show mean expected fitness without epistasis (right shifted for viewing). Error bars show one standard deviation. (**F**) Fraction of the 200 eukaryotic $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ genes with the same nucleotide as yeast at a given site decreases with the average fitness upon mutation at the site in yeast. Each dot represents one of the 69 examined tRNA sites. (**G**) Fraction of times that a mutant nucleotide appears in the 200 sequences increases with the fitness of the mutant in yeast. Each dot represents a N1 mutant. In (F) and (G), *ρ*, rank correlation coefficient; *P*, *P*-value from *t*-tests.

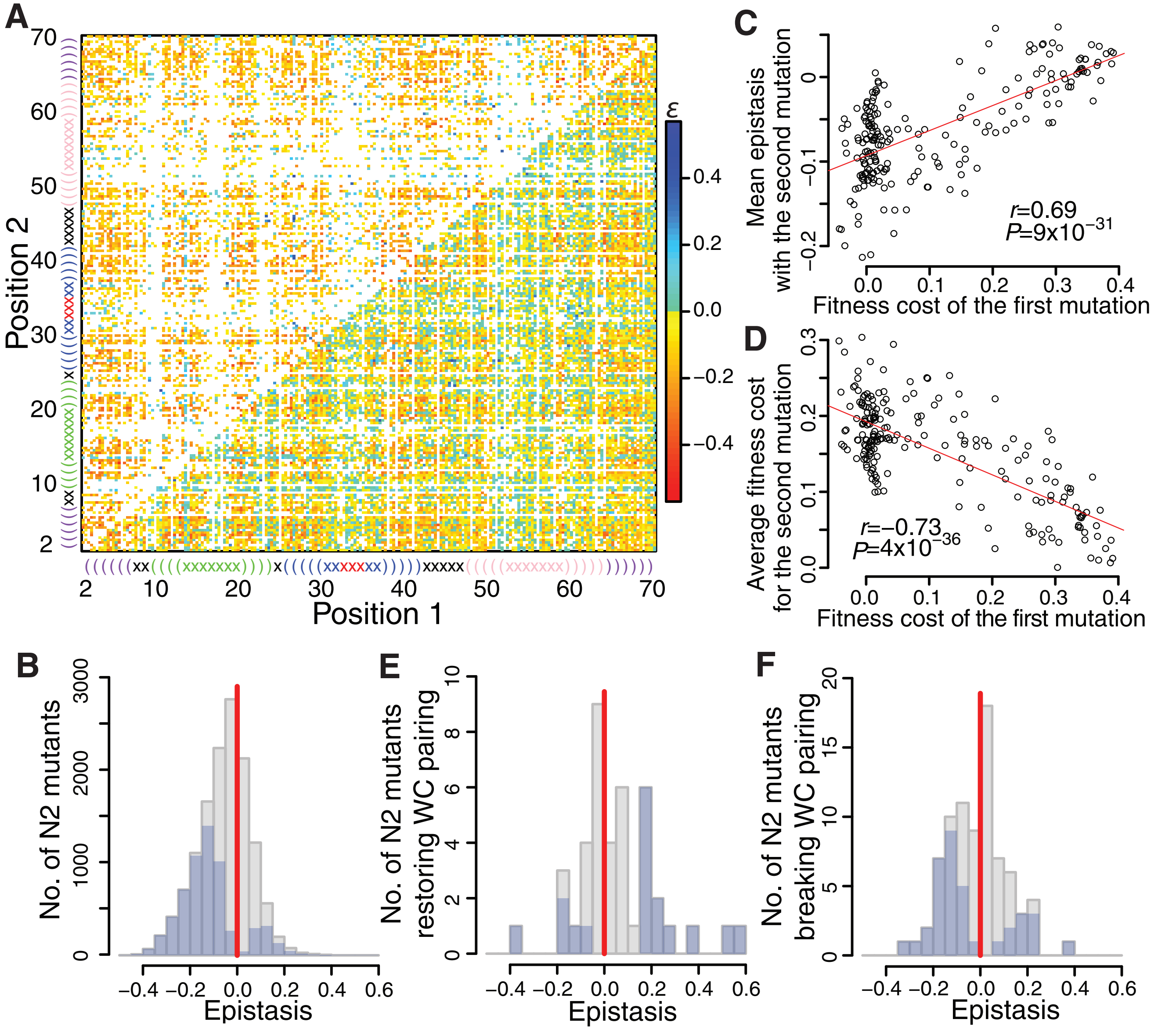


Fig. 2-3. Epistasis (ε) in fitness between point mutations in the $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene is negatively biased. (**A**) Epistasis between point mutations. Lower-right triangle shows all pairwise epistasis (white = not estimated), while upper-left triangle shows statistically significant epistasis (white = no estimation or insignificant). $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ secondary structure is plotted linearly. Parentheses and crosses show stem and loop sites, respectively. Same color indicates sites in the same loop/stem. Each site has three mutations. (**B**) Distributions of pairwise epistasis (gray) and statistically significant pairwise epistasis (blue) among 12,985 mutation pairs. (**C**) Mean epistasis between first and second mutations increases with the fitness cost of the first mutation. (**D**) Mean fitness cost of the second mutation decreases with the fitness cost of the first mutation. In (C) and (D), Pearson's correlation (*r*), associated *P* value, and the linear regression (red) are shown. (**E-F**) Distributions of epistasis (gray) and statistically significant epistasis (blue) between pairs of mutations that (**E**) convert a Watson-Crick (WC) base pair to another WC pair or (**F**) break a WC pair in stems. In (B), (E), and (F), the vertical red line shows zero epistasis.

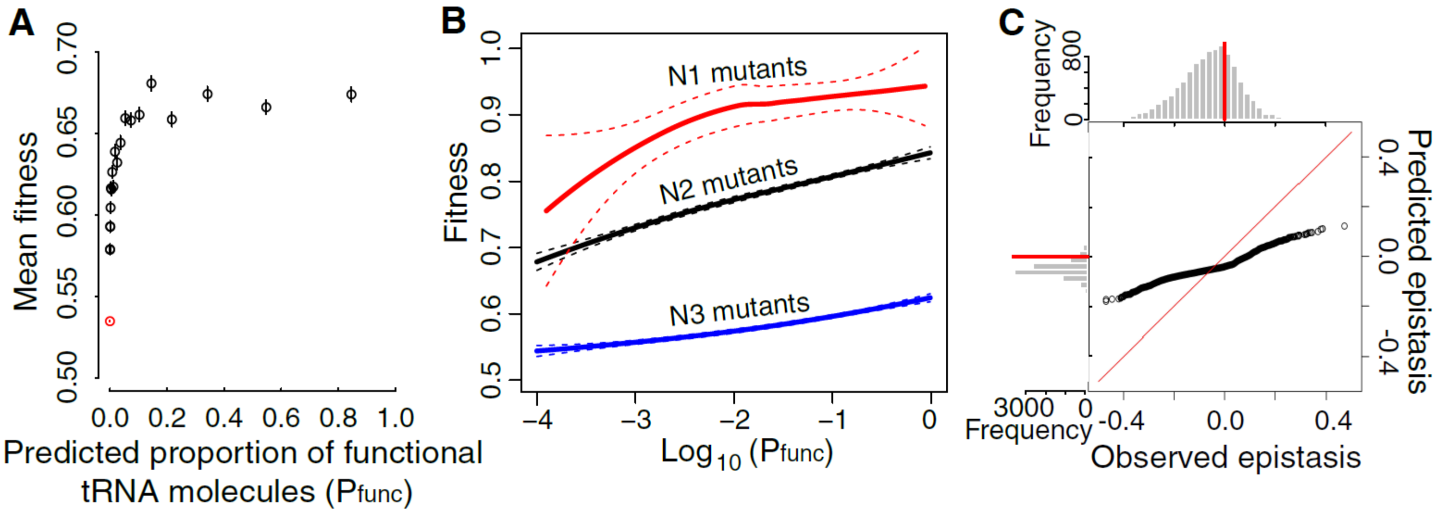


Figure 2-4. tRNA folding offers a mechanistic explanation of the fitness landscape. (**A**) Relationship between the predicted proportion of tRNA molecules that are functional (*P*func) for a genotype and its fitness. Genotypes (with $P_{\text{func}} \geq 10^{-4}$) are ranked by *P*func and grouped into 20 equal-size bins; mean *P*func and mean fitness ± SE of each bin are presented. The red dot represents all variants with *P*func < 10^{-4}. (**B**) LOESS regression curves between *P*func and fitness for N1, N2, and N3 mutants, respectively, with dashed lines indicating 95% confidence intervals. (**C**) Quantile-quantile plot between epistasis predicted from *P*func values using N1 and N2 LOESS curves and observed epistasis. The *i*th dot from the left shows the *i*th smallest predicted epistasis value (y-axis) and *i*th smallest observed epistasis value (x-axis). Red diagonal line shows the ideal situation of y = x. Above and left of the plot are frequency distributions of observed and predicted epistasis, respectively. Red horizontal and vertical lines indicate zero epistasis.

CHAPTER 3
PREDICTING FITNESS LANDSCAPE AND EPISTASIS ACROSS FOUR ENVIRONMENTS

3.1 Abstract

To study how the fitness landscape of a tRNA gene varies among environments, I measured the landscape in four environments with high precision using high-throughput barcode sequencing. I found that the same mutation almost always has different fitness effects in different environments, indicating pervasive genotype by environment interactions (G×E). Nevertheless, the observed G×E follows a simple piecewise linear relationship in which the fitness effect of a (deleterious) mutation in an environment is proportional to gene importance in the environment, while beneficial mutations have more similar fitness advantages across environments. This rule allows predicting the fitness landscape of the tRNA gene in any environment as long as the fitness landscape in one environment and the relative gene importance in the two environments have been measured. Our high-throughput mapping reveals relatively simple rules underlying the seemingly complex tRNA fitness landscapes, giving hopes for understanding and predicting fitness landscapes of other genes. Given the prediction on fitness, epistasis values are predicted with high predictive power. Change of epistasis sign across environments also appears at a low frequency, partially predicted by the proposed model.

3.2 Introduction

The effect of a new mutation on a gene can be positive, negative, or neutral in a specific environment, and may vary in a different environment. Organisms are constantly adapting to their local environment, exploring the fitness landscape by a random walk. During the random work on a fitness landscape towards higher fitness, sometimes local fitness peaks are separated by deep fitness valleys. Organisms need to travel across these fitness valleys to reach a higher peak. If we do not consider the extreme case of small population size that forfeits selection when crossing fitness valleys (Weissman, Feldman et al. 2010), gradual stepwise improvements alone cannot bridge such fitness valleys (Lindstrom, Alatalo et al. 1999). Epistasis, however, is one possibility that can make crossing fitness valleys possible when the genetic background changes (Bloom, Gong et al. 2010, Chang and Torbett 2011). Moreover, many case studies have revealed another possibility through the change of the shape of the fitness landscape across environments (Arias, le Poul et al. 2016, Steinberg and Ostermeier 2016). Having different alleles favorited by selection in different environments can lead to maintenance of genetic diversity or even reproductive isolation between subpopulations in various environments (Mitchell-Olds, Willis et al. 2007). Such mechanism has been widely referred to as one of the contributing factors when explaining allopatric / parapatric speciation and the evolution of genetic incompatibility (Schluter 2001), but the there is a scarcity of evidence available for the general existence and the magnitude of such changes.

Evaluating gene by environment interaction at a large scale asks for measuring fitness landscapes across multiple environments. Quantifying fitness landscape in a single environment has been a formidable task by itself due to the huge genotype space and complex epigenetic interactions among sites. However, even after obtaining the fitness landscape in one particular environment, how much information we can infer from it and further impose on another environment is still unclear. Many individual case studies have documented strong genotype by environment interactions (Arntz, Delucia et al. 2000, Hietpas, Bank et al. 2013), raising the possibility that genotype by environment interactions are abundant and gene-specific, therefore being unpredictable. The variation of fitness landscape across environments and the relationship and predictive power across environments are currently unknown. In this chapter, I will present a fitness landscape data set collected in four environments, where can explore the relationship of fitness between environment pairs, and find out the outliers that deviate from the general relationship.

Similar to fitness, epistasis can also be environment dependent. A previous study on five mutations revealed that the overall patterns of epistasis are negative, but sign and magnitude of epistasis among generally beneficial mutations vary widely even across similar external environments (Flynn, Cooper et al. 2013). It is of general interests to quantify changes of epistasis across environments and revealing the underlying trends for such sign and magnitude changes.

In this chapter, I first establish a simple piecewise linear model for predicting fitness in one environment from the fitness landscape in another environment. The model performance is evaluated based on the bias and deviation of the model prediction across the observed fitness range. This model outperforms two alternative models, especially for predicting beneficial mutations. Based on the biological replicate information, less than 0.2% of fitness values are identified as outliers for model prediction. Lastly, given the predicted fitness values, I further evaluated the predictive power for estimating epistasis across various environments. Epistasis can also change sign depending on the environment, and I characterize the change of epistasis sign for six environment pairs. A few global beneficially mutations are also identified and confirmed with growth rate measurements.

3.3 Results

3.3.1 Describing the fitness landscape and epistasis in four environments

In a pilot study, the growth rate is shown to be highly correlated for 55 mutant strains using Bioscreen C measurements (**Figure S5**). Using the methodology described in Chapter 2 (Li, Qian et al. 2016), I determine the $\text{tRNA}_{\text{CCU}}^{\text{Arg}}$ gene fitness landscape in four environments, including 23°C in YPD (hereafter 23°C), 30°C in YPD (hereafter 30°C), 30°C in YPD with 3% DMSO added (hereafter DMSO) and 37°C in YPD (hereafter 37°C). Among these environments, 30°C is the optimal growth condition for the lab strain BY4742, with the other conditions having low temperature, high temperature, and oxidative stress challenges (Sadowska-Bartosz, Paczka et al. 2013). The four environments are chosen because there is a gradient of factors changing among these environments, allowing us to test different hypotheses. Specifically, for the temperature that affects tRNA folding, we have 37°C > 30°C = DMSO > 23°C, while for the wild-type growth rate, we have 30°C > 37°> DMSO > 23°C.

The distribution of fitness varies across environments, with 37°C being the most severe condition and 23°C being the mildest condition (**Figure 3-1A**). Because $\text{tRNA}_{\text{CCU}}^{\text{Arg}}$ mutations are unlikely to be fatal, we set genotype fitness at 0.5 when the estimated fitness is < 0.5 (due to stochasticity). The average fitness of all N1 mutants is 0.976, 0.962, 0.957 and 0.934 for 23°C, 30°C, DMSO and 37°C, respectively. Average fitness values at different positions also vary across environments (**Figure 3-1B**). Fitness in different environments, although distributed over various ranges, are highly correlated (Correlation for N1 mutants shown in **Figure 3-2**). The same trend is also observed for higher order mutants, though less significant, probably due to higher measurement error for these mutants. In all four environments, epistasis was found to be negative in general (**Figure 3-3**) except when focusing on paired sites.

3.3.2 Pervasive GxE and Epistasis by environment interactions

To elucidate the general relationship between fitness across different environments, I use LOESS regression to visualize the relationship better(Figure 3-4A, B). Fitness across different environment are linearly correlated, with different slopes for beneficial and deleterious mutations. For environment pairs with a similar distribution of fitness, a smaller fraction of sites shows significant differences in fitness across environments (t-test, nominal $P<0.05$). 4.92% of mutants showed significant differences between 30°C and DMSO. For environment pairs with drastically different fitness distributions, a much higher fraction of mutants shows significant differences in fitness, as high as 39.94% between 23°C and 37°C. The fraction in all environmental pairs is summarized in the upper left triangle of **Figure 3-4C**. An alternative method to quantify GxE interaction is to focus on its magnitude. Without GxE interaction, the fitness should have the same magnitude across environments. Given a large number of variants sampled, the fraction of fitness with a higher magnitude than the other environment should be very close to 50% (95% CI: 50.0% ±0.8%). However, across all environment pairs, we found the fraction to be deviating significantly from 50%. The most significant pair is 23°C and 37°C, and over 90% of variants showed a higher fitness in the former environment. Percentages for all environment pairs is summarized in the lower right triangle of **Figure 3-4C**. Similarly, I quantified epistasis by environment (GxGxE) interaction using the abovementioned method (**Figure 3-4D**). A substantial fraction of epistasis is significantly different across environment pairs except for between 30°C and DMSO. Moreover, all environment pairs have a non-random distribution in terms of the magnitude change of epistasis across environments.

3.3.3 A piecewise linear model predicts fitness across environments

Knowing the predictability of fitness across environments are important because it allows for inferring fitness landscape in a new environment. As an example, I use fitness measurements at 30°C to predict that at 23°C. Similar results were obtained for other environment pairs. To ensure comparatively low measurement errors, I further restrict analysis to variants with read number at T0 higher than 500, a total of 18,902 variants, including 207 N1 mutants (100% of all possible variants), 5,754 N2 mutants (27.3%), 6,099 N3 mutants (0.44%) and 3,884 N4 mutants, etc. I first proposed and compared two piecewise robust linear model assuming linearity for fitness (formula [1]) and log (fitness) (formula [2]), respectively, as shown below. $\widehat{f_Y}$ represents the fitness to be estimated in the new environment, and f_X is the observed fitness in the previously measured environment. The two parameters, k_1 and k_2, are separately estimated for beneficial and deleterious mutations.

$$\widehat{f_Y} - 1 = \begin{cases} k_1 \times (f_X - 1) & if\ f_X \geq 1 \\ k_2 \times (f_X - 1) & if\ f_X < 1 \end{cases} \qquad [1]$$

$$\log(\widehat{f_Y}) = \begin{cases} k_1 \times \log(f_X) & if\ f_X \geq 1 \\ k_2 \times \log(f_X) & if\ f_X < 1 \end{cases} \qquad [2]$$

For lethal mutants or mutations that almost completely stop cell growth, it is biologically impossible to quantify the severity of these mutations, and technically difficult to measure its fitness accurately using Bar-seq. In all four environments, a large fraction of high-order mutants has fitness equal to 0.5, corresponding to cases where no read is recovered by sequencing at the end of competition. For these variants, measurement accuracy is limited, so I use a total of 5,960

N1 and N2 variants to train the linear model, but the model is used to predicted fitness for all variants. Both linear models and all data points are shown in Figure **3-5A**.

The most commonly used yeast gene importance data are the measures of relative growth rates of 5,936 single-gene deletion yeast strains collections (Giaever, Chu et al. 2002) in the YPD media or the presence of chemical or environmental stress conditions (Hillenmeyer, Fung et al. 2008). Similarly here, I define gene importance as the relative fitness of a deletion strain in that environment, which is proportional to the slope of the linear model (k_2) for deleterious mutations. Compared with deleterious mutations, the slope for beneficial mutations (k_1) is closer to 1 (Figure 3-5A). The likely explanation is that deleterious mutations are usually loss-of-function mutations, so its fitness effect is proportional to gene importance, while some of the beneficial mutations can be gain-of-function mutations, which might not be proportional to gene importance in each environment, but confers a universal advantage. The model applies well to N1, N2, and higher order mutants1 (**Figure 3-5A**), without any obvious bias for any particular groups of variants or across certain fitness ranges.

3.3.4 Evaluating model performance based on correlation, bias, and deviation

The bias and deviation for the linear model (**Figure 3-5A, B**) across the observed fitness range are calculated as below. The variable f_Y is the observed fitness in the new environment and $\widehat{f_Y}$ is the predicted fitness in the new environment.

$$Bias = \widehat{f_Y} - f_Y \qquad [3]$$

$$Deviation = |\widehat{f_Y} - f_Y| \qquad [4]$$

Points with fitness equal to 0.5 in either environment are removed because of low predictive power and high measurement error. There is not much systematic prediction error in our model except for variants with extremely low fitness values (**Figure 3-6B**). This bias is expected because when fitness values are low, the corresponding reads at the pool after competition can be completely uncovered in the sequencing pool for some of the biological replicates, where I assign fitness as 0.5 for these biological replicates. The deviation of fitness prediction is the lowest when the fitness is higher than 0.6 but becomes higher for lower fitness values, which is consistent with the technical measurement errors quantified by standard deviation across five biological replicates (dashed lines in **Figure 3-6C**). The predicted and observed fitness values are highly correlated assuming linearity for log fitness (Spearman's $\rho = 0.872$, Pearson's correlation $r = 0.866$). The correlation is the highest for N1 mutants with the least measurement error (Spearman's $\rho = 0.932$, Pearson's correlation $r = 0.982$), and is also very high for N2 mutants (Spearman's $\rho = 0.906$, Pearson's correlation $r = 0.912$). The performances for models assuming linearity for fitness and log fitness are similar, but the one for log fitness shows slightly higher correlation, lower overall bias, and deviation, as well as lower mean squared error across multiple environment pairs, so I used it throughout the rest of the chapter.

3.3.5 Comparison with two alternative models

I further compare the above-mentioned piecewise linear model with two alternative models, one simple linear model with a single slope for both beneficial and deleterious mutation, and a quadratic model. Both models are also required to go through P(1, 1). The three models and all data points are plotted in **Figure 3-5D**. The bias and deviation of the alternative models were shown in **Figure 3-5E, F**. Because most of the training dataset are deleterious, the model performance is largely unchanged for deleterious mutations. Both alternative models perform poorly in predicting fitness for beneficial mutations (**Figure 3-5E, F**) with large deviations and biases compare with the piecewise linear model.

3.3.6 Identifying outliers of the linear model

Using the piecewise linear model, I predicted fitness at 23°C using the fitness values of five biological replicates at 30°C individually. Points with fitness all reaching 0.5 in either condition were removed because of their high measurement errors. A t-test was used to compare the predicted and observed fitness of the five biological replicates at 23°C, and a cutoff of P = 0.05 for FDR corrected p-value was used. Among the 16,509 predictions, 18 outliers (0.11%) were detected, 9 (4.3%) of them were N1 mutants, and 2 of them were N2 mutants (0.034%). Using P = 0.05 for cutoff without FDR correction, a total of 1,157 (7.0%) mutants were identified as outliers of model prediction. The smaller fraction of outliers further indicates the good performance of the model.

3.3.7 Change of epistasis sign across multiple environments

Epistasis is found to be highly correlated between environment pairs. However, epistasis can sometimes change signs across environments. **Table 3-1** listed the number of cases that change the sign for epistasis that is significantly different from 0 for the two environments compared. The sign of epistasis remains largely unchanged across environment pairs. When the gene importance is similar in two environments (30°C and DMSO), very few switches in the sign of epistasis are observed, as predicted by the model. When gene importance increases, there is an enrichment for switching from positive epistasis to negative epistasis in general. For instance, a total of 58 pairs of mutations showing positive epistasis at 23°C become negative epistasis at 37°C, with the former environment having the lowest gene importance and the latter having the highest gene importance.

To see if the model predicts such switch of signs, I use environment pair 30°C and 23°C as an example. Epistasis at 23°C can be calculated from three individual mutant fitness values predicted from the model. According to the model prediction, 0 pairs of mutation would show positive epistasis at 30°C and negative epistasis at 23°C, while 14 pairs of mutation would show positive epistasis at 23°C and negative epistasis at 30°C. Among the 14 predicted pairs, seven were confirmed by the Bar-seq experiment (Bootstrap $P < 10^{-6}$), with a total of 206 pairs show this pattern. Among all significant epistatic pairs measured, 0 pairs of mutation showed positive epistasis at 30°C and negative epistasis at 23°C, while seven pairs showed positive epistasis at 23°C and negative epistasis at 30°C, however, none of these seven pairs overlap with those above 14 predicted pairs. Therefore, the switch of epistasis sign is partially predicted by the model.

3.3.8 Inferring fitness landscape in a new environment from a few measurements

All analyses mentioned above for fitness prediction focus on building and evaluating the model when a large number of fitness values from both environments are already available. In reality, when inferring fitness landscape in a new environment from a measured environment, fitness measurements in the new environment will not be available. Therefore, I further evaluate the power when collecting only a few data points of fitness. The performance of the methodology for sampling a small number of points across different fitness range is shown in **Figure 3-6**. Sampling even one single data point that is not too close to wild-type genotype will allow recovery of the relationship comparatively accurately, while sampling a few more data points will allow even better prediction that is close to having all the fitness information.

3.3.9 Beneficial mutations and its structural basis

2.03%, 1.75%, 1.77% and 1.85% of variants confer over 5% benefits at 23°C, 30°C, DMSO and 37°C, respectively, while 1.22% (230) variants are over 5% more beneficial than the wild-type genotype across all four environments. Among them, three were N1 mutations, all mutating from G to T at position 38, 39 and 68, respectively. Interestingly, the most beneficial mutation, G38T, appears 72.6% in a pre-compiled eukaryotic $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ dataset (see Chapter 2), compared with 14.7% for the wild-type G nucleotide in yeast. The other two highly beneficial mutations, although appearing at a lower frequency compared with the wild-type nucleotide, are also observed multiple times in other eukaryotic species. I then include all N1 mutations that are universally more beneficial than wild-type genotype in the four environments and found that they appear significantly more frequently in other eukaryotes compare with the rest of mutations (25.1% and 12.0% respectively, $P = 3\times10^{-7}$, Mann–Whitney *U* test). Both beneficial mutations and deleterious mutations appear less frequently than the wild-type nucleotides on average, but the former group appear at a much higher frequency in other eukaryotes (12.1% and 48.0% less than the wild-type nucleotides on average at the corresponding positions for beneficial and deleterious mutations, respectively, $P = 3\times10^{-7}$, Mann–Whitney *U* test). Moreover, these mutants have a significantly higher predicted fraction of functional molecules (31.9% and 2.8% for beneficial and deleterious N1 mutations, $P = 2\times10^{-28}$, Mann–Whitney *U* test).

I also compared the cases where the fitness effect differs across environments. When requiring over 5% fitness changes for beneficial and deleterious effects, a total of 27 cases were found to be switching from beneficial at 23°C to deleterious at 37°C, while none happened in the opposite direction. Similarly, 8 cases were found to be switching from beneficial at 23°C to deleterious at 30°C with none in the opposite direction; 7 such cases was found to be switching from beneficial at 30°C to deleterious at 37°C with none in the opposite direction. This phenomenon can be potentially explained by the hypothesis that one mutation conferring beneficial functionality at a lower temperature may no longer be beneficial due to higher sensitivity to disruptions in the structure by mutations at higher temperatures. Comparing any environment pairs involving DMSO, there is at most 2 cases in either direction. Reducing the cutoff of fitness change to 2% generates similar results.

3.3.10 Confirmation of the model in other published datasets

While previous studies have been extensively focused on the fraction of beneficial and deleterious mutations, I reanalyzed some of the previous studies using a simple linear model. Beneficial mutations aren't separate fitted because of limited sample size and high measurement noise. For instance, our model fits well with a large-scale dataset concerning the fitness of Hsp90 across four environments at two temperatures (30°C and 36°C) and two salinity levels (C for low salinity and S for high salinity level) (Figure 3-7) (Hietpas, Bank et al. 2013). Our model well captures the fitness distribution across environments. While previous studies lack a general null hypothesis for the distribution of fitness across environment and focus on the frequency of fitness at different ranges, our model provides a clear prediction and explained the majority of observed variance (Spearman's ρ = 0.934, Pearson's correlation r = 0.952 for predicted and observed fitness at 36C corresponding to the left column of **Figure 3-7A**, and Spearman's ρ = 0.854, Pearson's correlation r = 0.966 for predicted and observed fitness at 36S corresponding to the right column of **Figure 3-7B**). To show linearity of the model, I plotted the observed log fitness versus the predicted log fitness (**Figure 3-7 C, D**), and the expected log fitness versus the residual log fitness (**Figure 3-7 E, F**). The balanced distribution of points on both side of the dashed trend line indicates that the points indeed follows a simple linear relationship, with a higher deviation at lower fitness range because of higher measurement error.

3.4 Discussion

In summary, there are wide-spread gene-environment interactions, but the effect of these interactions turned out to be largely predictable. Our high-throughput mapping and modeling reveal relatively simple rules underlying the seemingly complex tRNA fitness landscapes and gene-environment interactions, giving hopes for understanding and predicting fitness landscapes of other genes across multiple environments. Having a simple model for prediction can also help us to identify candidate sites that are differentially interacting with various environments, which could potential help to reveal more complex gene by environment interactions.

A substantial fraction of GxGxE interaction is also observed and largely explainable by our model, which provides a convenient null model to help reveal the interaction of sites that are different from the general trend. The observation that most of the changes in fitness and epistasis across environments can be explained by gene importance is not totally surprising. In a modular view of the gene network, each module contributes differently to the overall fitness across each environment. Different deleterious mutations might be destroying the general functionality of the module by the same proportion, but the organismal fitness differs across environments because the importance of the whole functional module is environment-dependent.

There are a few caveats of the study. Firstly, this study focuses on a single tRNA gene in *S. cerevisiae* across four environments, rather than focusing on a genome-wide scale or assessing the complex and variable natural environments. Whether the trends observed here are widely applicable to other genes, organisms and environment pairs are currently unknown, and more case studies are highly desirable for confirmation. Secondly, the model is very simplistic, more sophisticated models considering site information might be able to achieve even better prediction. However, this information is likely to be gene specific and is not available for most candidate genes. Therefore, I didn’t discuss such information except when discussing model outliers. Thirdly, although the sample size for this study is comparatively large, more than half of the possible N2 mutations and most of the high-order mutations are not present. However, I able to show that our model applies to mutants carrying different numbers of mutations and across different fitness ranges. The variants were chemically synthesized and collected in a condition where each variant grows similarly, so the representation of deviation is roughly random in our dataset. Therefore, I conclude that the trend is widely applicable even for unobserved variants.

3.5 Materials and Methods

3.5.1 Measuring fitness in multiple environments

The competition was conducted as in Chapter 2. In short, cells cultured in four conditions were from the same founding population and subject to growth for ~13 generations. Cells were then harvested and lysed to extract DNA. Two rounds of PCR amplification were conducted to amplify the $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene incorporated at the correct genomic location and add adaptors to the gene to be sequenced. Three lanes of Hi-seq sequencing were conducted, and the frequency change of each variance before and after the competition is used to calculate fitness for each variant.

Five biological replicates were conducted at 23°C and 30°C, and three biological replicates were conducted at 37°C and in the DMSO condition, respectively. I also sequenced twice the $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene amplicon from the common starting population (T0) of the competitions. A perfect match between the fully overlapping paired-end reads was required in estimating genotype frequencies. The change in relative genotype frequency between the pool before and after the competition assay was used to determine the fitness of each genotype relative to the wild-type genotype. The fitness for each mutant is calculated as the average of fitness across multiple biological replicates. To ensure relatively accurate fitness estimation, I focus on 23,284 genotypes with read counts ≥ 100 at T0.

3.5.2 Quantifying GxE and GxGxE interaction

The fraction of sites showing GxE interaction is quantified from two perspectives. Firstly, I focus on each individual observation for fitness measurements across the two environments. A t-test is used to compare the fitness measured from biological replicates. The first methodology avoids the noise coming from measurement error by taking advantage of information from biological replicate. With a higher measurement accuracy, we expect to observe a higher fraction of sites showing significant GxE interaction. The second methodology is from a population perspective. If there is no GxE interaction, then fitness should be equal across the two environments. I focus on variants showing beneficial or deleterious effects in both environments, and counted the number of cases showing higher magnitude in one environment versus the other. Due to measurement noise, the fitness values won't be exactly equal, but the fraction of fitness to be higher in either environment should be roughly equal, following a binomial distribution. Therefore, the confidence interval can be calculated as

$$\hat{p} \pm 1.96\sqrt{p(1-p)/n} \qquad [5]$$

where p is 50% for random noise to contribute in either direction and n is the number of variants measured in both environments. Deviation from the confidence interval indicate GxE interaction, and a larger deviation shows a higher proportion of GxE interactions.

Similarly, the above-mentioned metric can be used to quantify GxGxE interaction. For the first methodology, epistasis can be quantified as $\varepsilon = f_{AB} - f_A \times f_B$ for each biological replicate, and subsequently compared using t-test across environments. For the second methodology, a single epistasis value calculated from the average fitness is used to compare the change of magnitude.

3.5.3 Building piece-wise robust linear model for fitness prediction

I built the piecewise robust linear model using the rlm function of the MASS package in R. The robust linear model optimizes to a majority best fit, as opposed to least-square error, reducing the possible extreme influences from outliers. The model is calculated separately for deleterious and beneficial mutations. The model was tested assuming linearity for fitness or log(fitness). In each scenario, the model is forced to go through P(1, 1) for fitness, because the relative fitness is defined as 1 in every environment for wild-type fitness.

3.5.4 Model evaluation

The model performance is measured from two perspectives, model bias, and model deviation. Bias is calculated by the median difference between the predicted fitness and observed fitness in one environment, while deviation is quantified by the median absolute difference between the predicted fitness and observed fitness in one environment. A LOWESS curve fitting (local polynomial regression) is used to get the general trend of the bias and deviation difference across different fitness ranges with f=0.2. Deviation for measurements in each environment was calculated as the mean of deviation for all biological replicates.

3.5.5 Two alternative models

Two alternative models were also built for comparison assuming linearity for log fitness. The first alternative model is a simple robust linear model that go through P(1, 1) for fitness, without having a separate calculation for beneficial and deleterious mutations. Another alternative model is a quadratic model that go through P(1, 1). Both models were evaluated using the methodology mentioned in 3.5.3.

3.5.6 Predicting epistasis

Expected fitness values $\hat{f}_{\mathrm{A}}, \hat{f}_{B}$ and $\hat{f}_{\mathrm{AB}}$were individually predicted from observed fitness values f_{A}, f_{B} and f_{AB} in another environment, and the predicted epistasis is calculated as $\hat{\varepsilon} = \hat{f}_{\mathrm{AB}} - \hat{f}_{\mathrm{A}}\hat{f}_{\mathrm{B}}$. Bias and deviation of the model is evaluated as in section 3.4.2.

3.6 References

Arias, M., Y. le Poul, M. Chouteau, R. Boisseau, N. Rosser, M. Thery and V. Llaurens (2016). "Crossing fitness valleys: empirical estimation of a fitness landscape associated with polymorphic mimicry." Proc Biol Sci **283**(1829).

Arntz, A. M., E. H. Delucia and N. Jordan (2000). "Fitness effects of a photosynthetic mutation across contrasting environments." Journal of Evolutionary Biology **13**(5): 792-803.

Bloom, J. D., L. I. Gong and D. Baltimore (2010). "Permissive secondary mutations enable the evolution of influenza oseltamivir resistance." Science **328**(5983): 1272-1275.

Chang, M. W. and B. E. Torbett (2011). "Accessory mutations maintain stability in drug-resistant HIV-1 protease." J Mol Biol **410**(4): 756-760.

Flynn, K. M., T. F. Cooper, F. B. Moore and V. S. Cooper (2013). "The environment affects epistatic interactions to alter the topology of an empirical fitness landscape." PLoS Genet **9**(4): e1003426.

Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A. P. Arkin, A. Astromoff, M. El Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kotter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Y. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. H. Yang, G. Yen, E. Youngman, K. X. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis and M. Johnston (2002). "Functional profiling of the Saccharomyces cerevisiae genome." Nature **418**(6896): 387-391.

Hietpas, R. T., C. Bank, J. D. Jensen and D. N. A. Bolon (2013). "Shifting Fitness Landscapes in Response to Altered Environments." Evolution **67**(12): 3512-3522.

Hillenmeyer, M. E., E. Fung, J. Wildenhain, S. E. Pierce, S. Hoon, W. Lee, M. Proctor, R. P. St Onge, M. Tyers, D. Koller, R. B. Altman, R. W. Davis, C. Nislow and G. Giaever (2008). "The chemical genomic portrait of yeast: Uncovering a phenotype for all genes." Science **320**(5874): 362-365.

Li, C., W. Qian, C. J. Maclean and J. Zhang (2016). "The fitness landscape of a tRNA gene." Science **352**(6287): 837-840.

Lindstrom, L., R. V. Alatalo, J. Mappes, M. Riipi and L. Vertainen (1999). "Can aposematic signals evolve by gradual change?" Nature **397**(6716): 249-251.

Mitchell-Olds, T., J. H. Willis and D. B. Goldstein (2007). "Which evolutionary processes influence natural genetic variation for phenotypic traits?" Nat Rev Genet **8**(11): 845-856.

Sadowska-Bartosz, I., A. Paczka, M. Molon and G. Bartosz (2013). "Dimethyl sulfoxide induces oxidative stress in the yeast Saccharomyces cerevisiae." Fems Yeast Research **13**(8): 820-830.

Schluter, D. (2001). "Ecology and the origin of species." Trends Ecol Evol **16**(7): 372-380.

Steinberg, B. and M. Ostermeier (2016). "Environmental changes bridge evolutionary valleys." Sci Adv **2**(1): e1500921.

Weissman, D. B., M. W. Feldman and D. S. Fisher (2010). "The rate of fitness-valley crossing in sexual populations." Genetics **186**(4): 1389-1410.

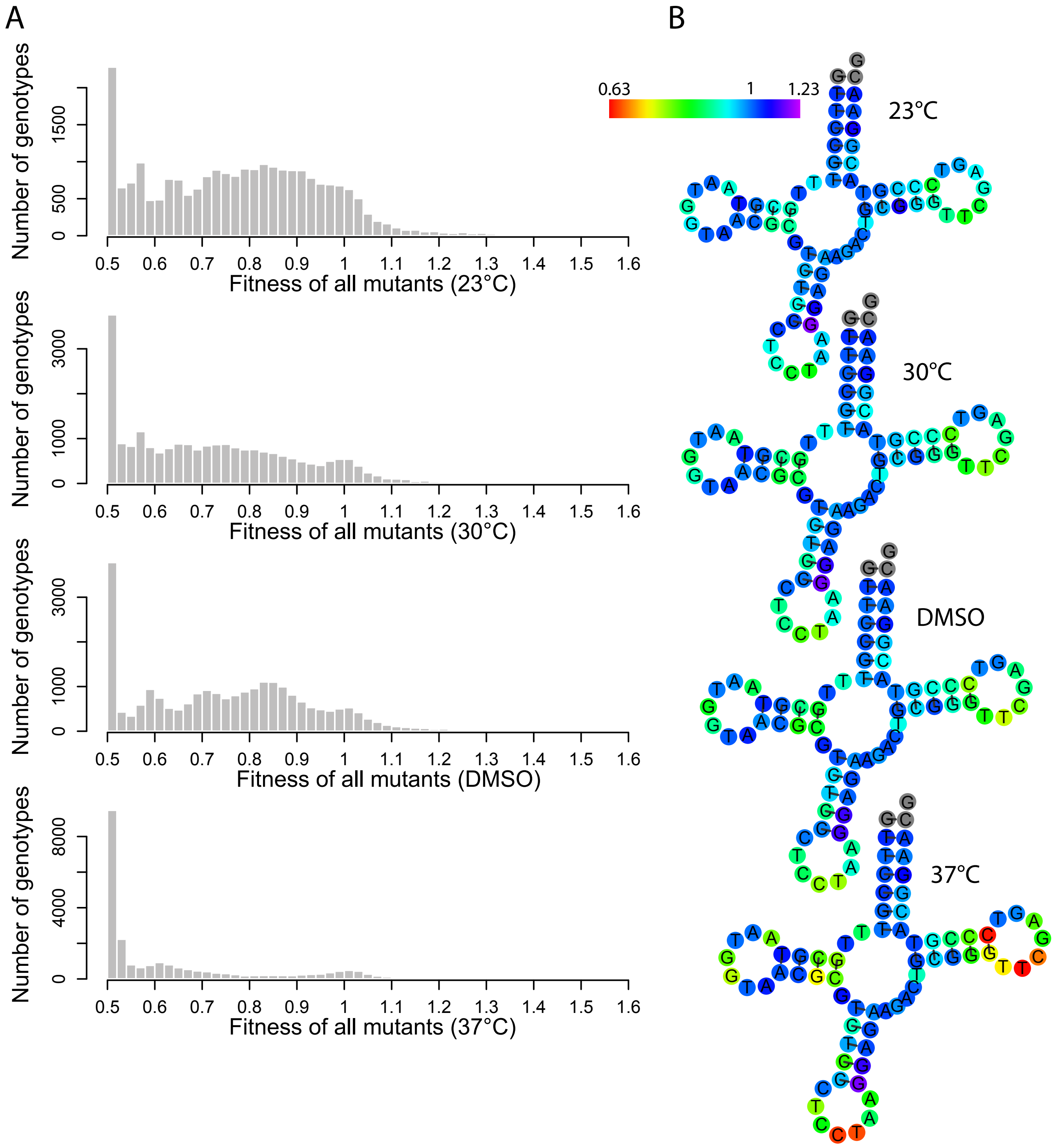


Figure 3-1. **Yeast** tRNA$^{Arg}_{CCU}$ **gene fitness landscape**. (**A**) Fitness distribution of all mutants in the four environments. **(B)** Average fitness upon a mutation at each site. Grey circles indicate invariant sites.

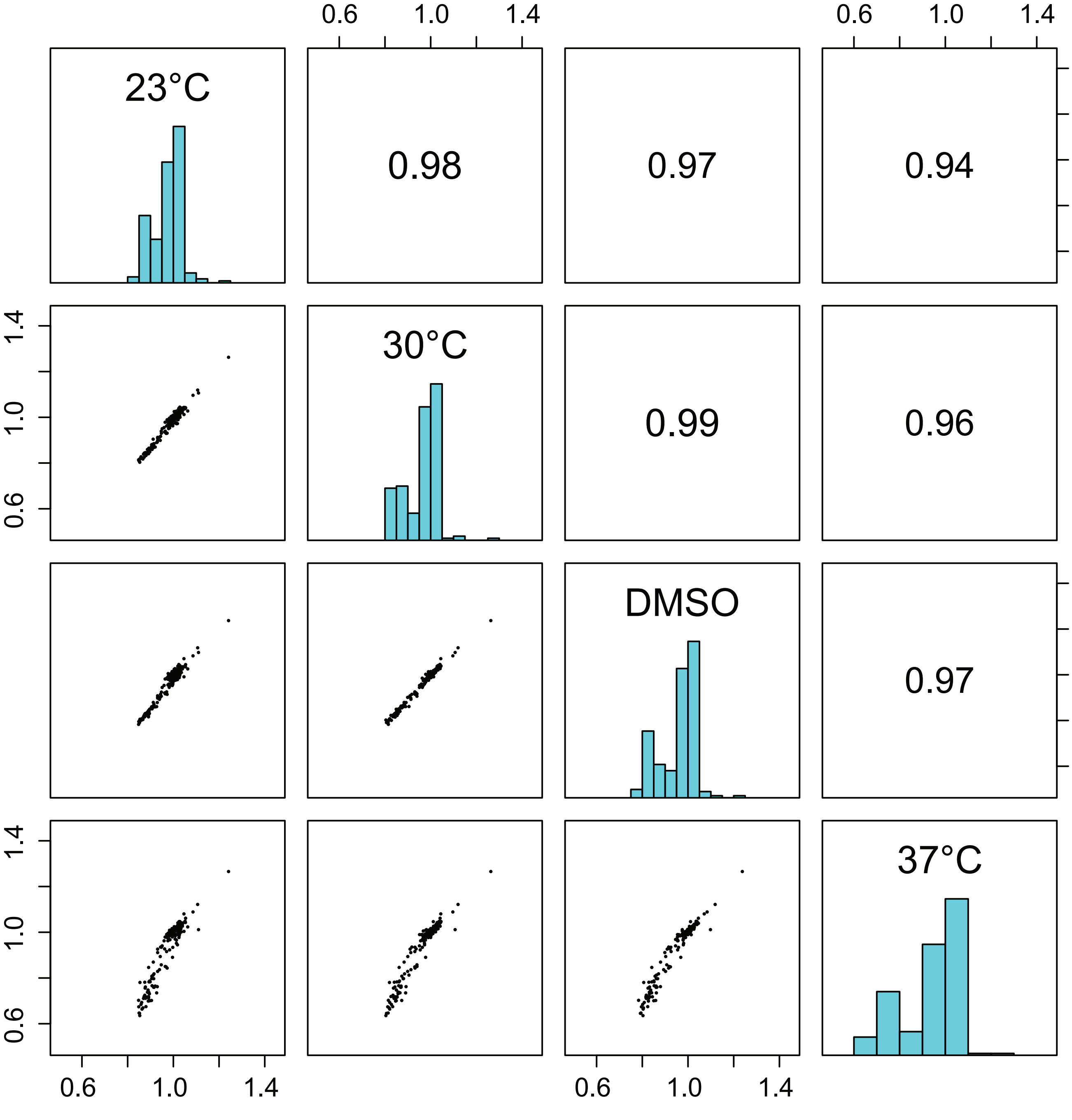


Figure 3-2. **Fitness in different environments is highly correlated**. Distribution and correlation of fitness for all variants carrying one mutation are plotted. Fitness values of all 207 possible single mutant genotypes in each environment are plotted against those in the other three environments, with Pearson's correlation coefficients indicated. Environments used and the distribution of fitness for single mutants in the environments are shown in the diagonal panels.

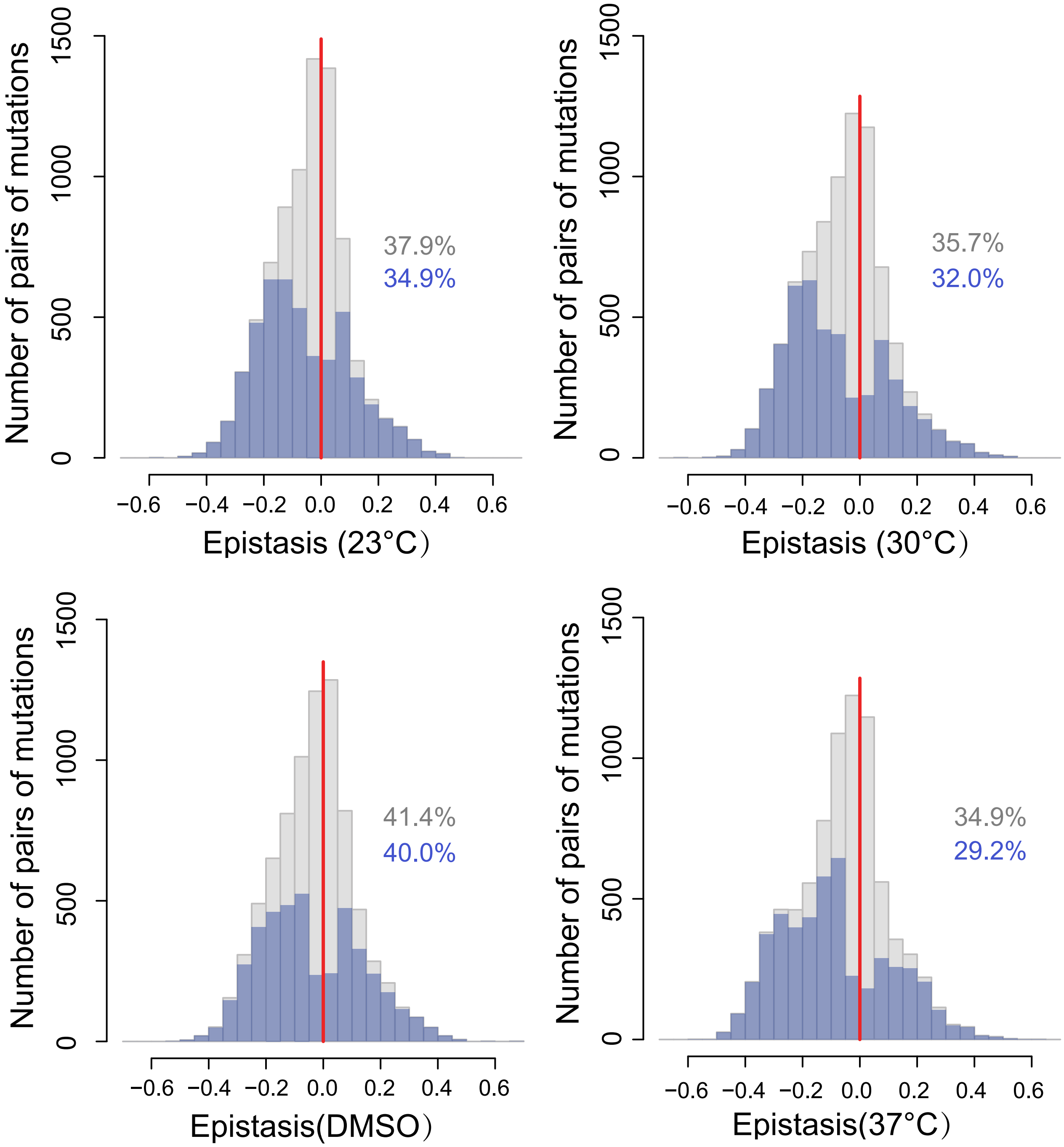


Figure 3-3. **Epistasis between point mutations in the tRNA gene is negatively biased in all four environments**. Frequency distributions of pairwise epistasis (gray) and statistically significant pairwise epistasis (blue) among 8,101 pairs of point mutations studied. The percentages show the fractions of positive epistasis in that environment. The gray and blue numbers show the fraction of positive epistasis in each environment for all epistasis and significant epistasis, respectively.

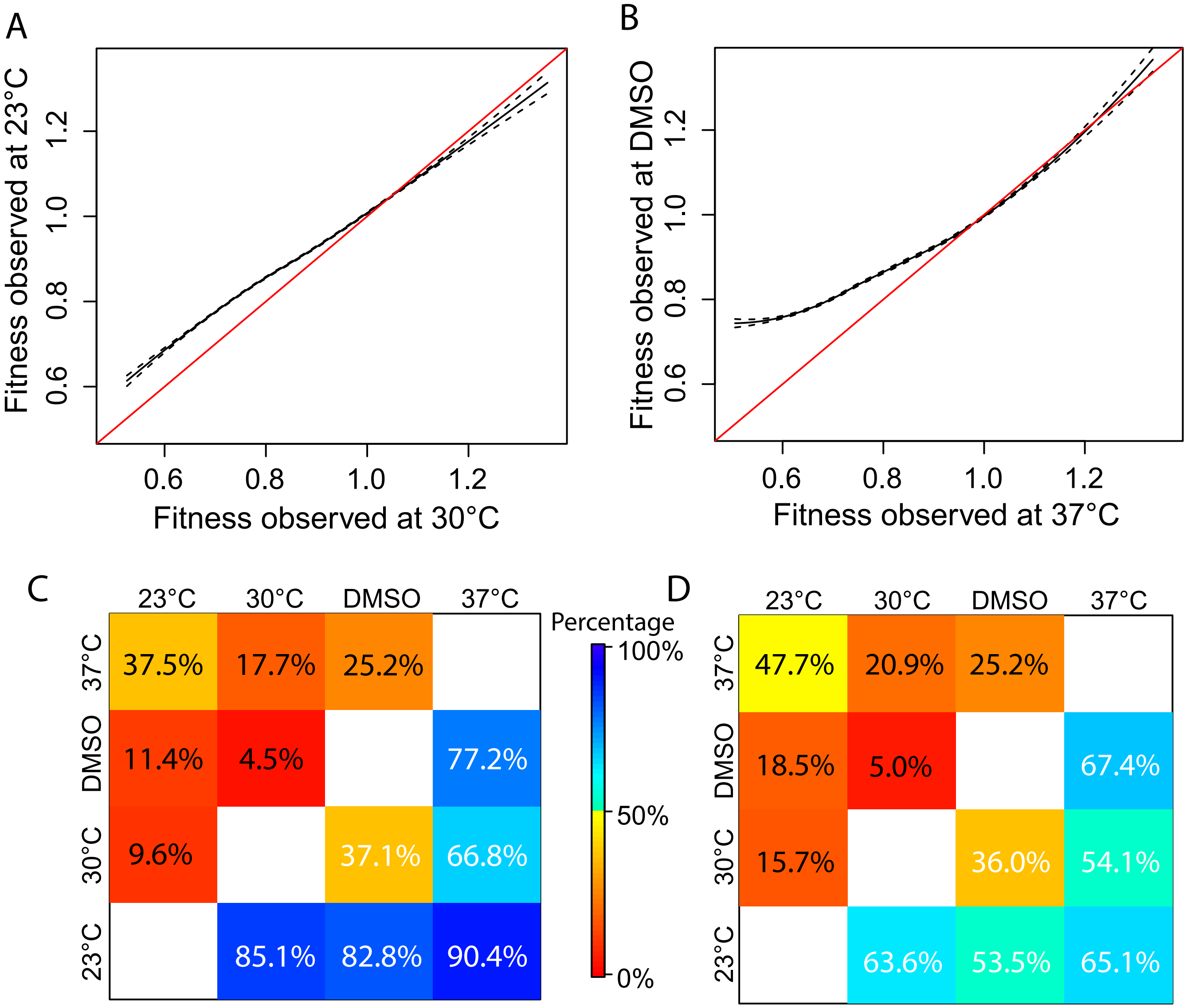


Figure 3-4. **Pervasive GxE and GxGxE interactions**. (**A**) The relationship between fitness at 23 °C and 30 °C is summarized by a LOESS regression line (95% confidence interval shown by the dashed line). (**B**) The LOESS regression line is plotted between fitness at 37 °C and DMSO. (**C**) Summary of GxE interaction between environment pairs. The upper left triangle shows the fraction of fitness that is significantly different between the two environments (t-test, *P*<0.05). The lower right triangle shows the percentage of fitness that is higher in magnitude than the other environment. The 95% confidence interval for the expected percentage assuming binomial distribution is 50.0% ±0.8%. (**D**) Similar percentages are plotted for GxGxE interaction. The upper left triangle shows the fraction of epistasis that is significantly different between the two environments (t-test, *P*<0.05). The lower right triangle shows the percentage of epistasis that is higher in magnitude than the other environment. The 95% confidence interval assuming binomial distribution is 50.0% ±1.5%.

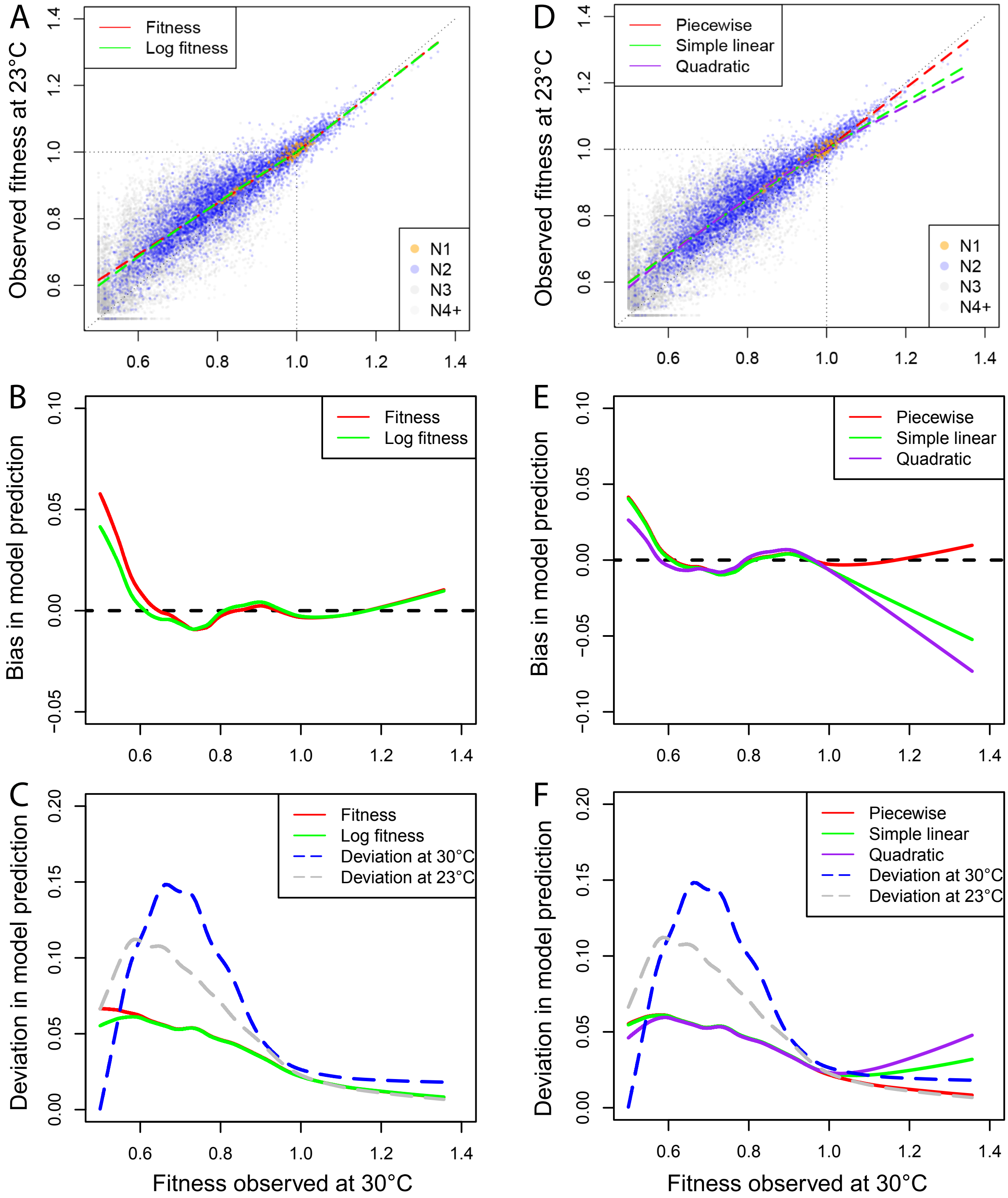


Figure 3-5. **The piece-wise linear model outperforms alternative models**. (**A**) Dots with different colors show fitness values in two environments for variants carrying a different number of mutations. The models are plotted as dashed lines. The dotted lines show Y=X, X=1, and Y=1. (**B**) The bias for prediction excluding fitness of 0.5 in either environment. (**C**) The deviation for both models are plotted as the solid line, and the dashed lines show the deviation of measurements in each environment. (**D**) The piecewise linear model is compared with two alternative models. (**E**) The bias of alternative models for prediction excluding fitness of 0.5 in either environment. (**F**) The deviations for both alternative models is plotted together with the piece-wise linear model.

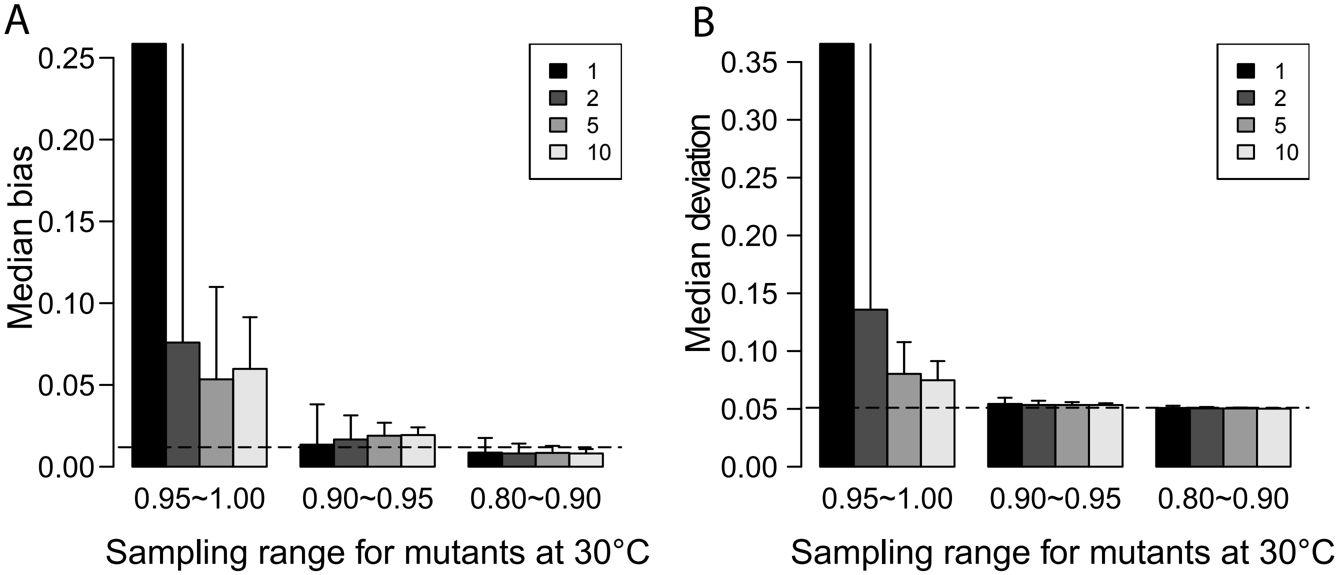


Figure 3-6. **Sampling a few data points can effectively recover the fitness pattern**. (**A**) Median bias for model prediction when sampling different numbers of N1 training samples at the designated sampling range. The figure legends show the number of points sampled. The dotted line shows the median bias when the model is built based on all the training data. The error bar of 0.81 and 0.30 for the first two bars are not shown for scaling reason. (**B**) Median deviation for model prediction when sampling different numbers of N1 training samples at the designated sampling range. The figure legends show the number of points sampled. The dotted line shows the median bias when the model is built based on all the training data. The error bar of 0.77 and 0.28 for the first two bars are not shown for scaling reason.

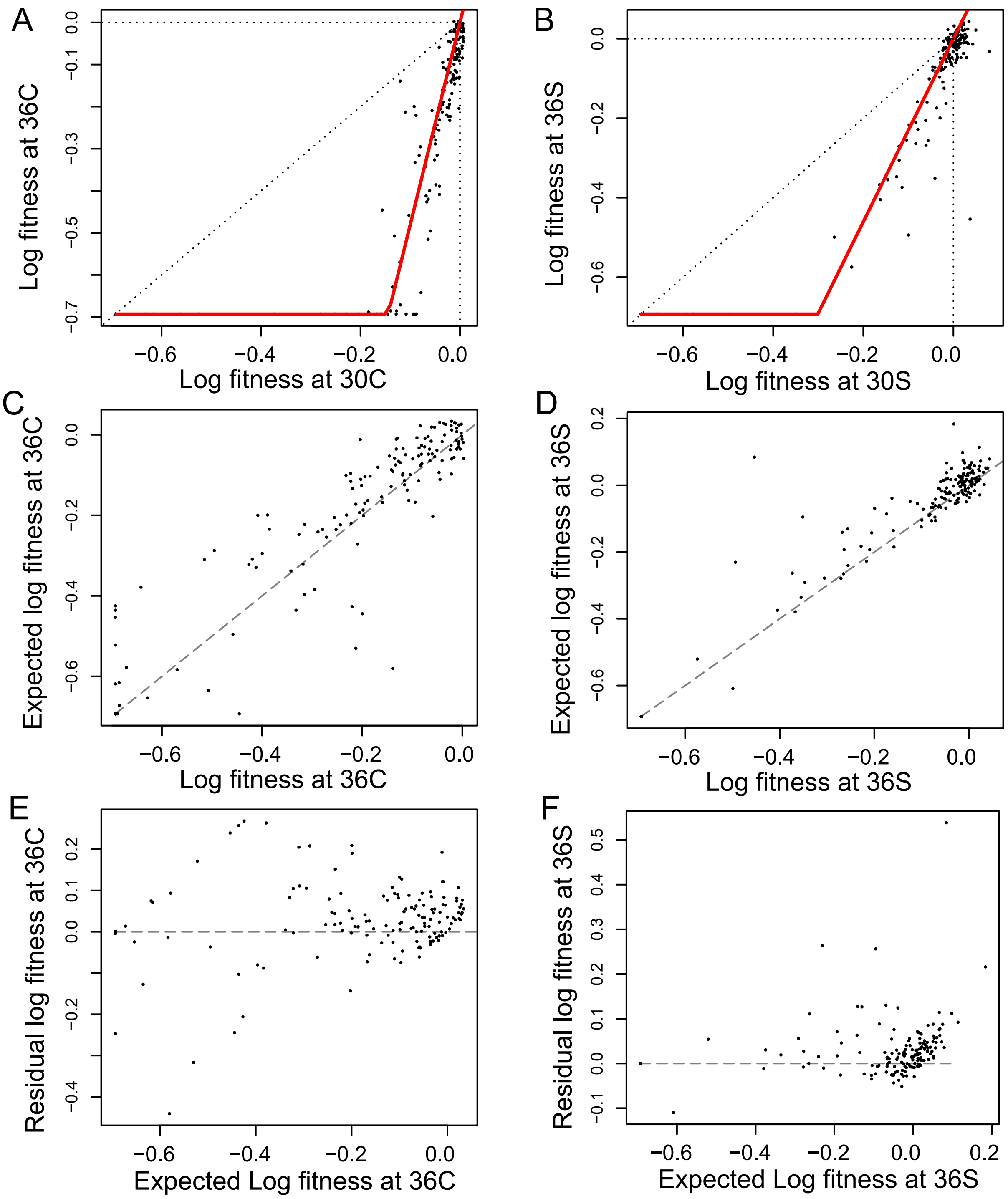


Figure 3-7. **Confirmation of model using HSP90 fitness datasets**. (**A**) Log fitness at 30°C with and without elevated salinity follows a simple linear relationship. (**C**) Observed and predicted log fitness values are roughly symmetric around the diagonal line (shown as dotted line). (**E**) The residuals are roughly symmetrical around the horizontal line when correlated predicted, indicating linearity between the two conditions. (**B, D, F**) Similar patterns are observed when comparing log fitness under at 30°C and 37°C with elevated salinity.

Table 3-1. Change of epistasis sign across different environments.

		ε at 23°C		
		Positive ε	Negative ε	Total
ε at 30°C	Positive ε	399	**0**	399
	Negative ε	**7**	2517	2524
	Total	406	2517	2923

		ε at 23°C		
		Positive ε	Negative ε	Total
ε at 30°C+3%DMSO	Positive ε	466	**0**	466
	Negative ε	**7**	2176	2183
	Total	473	2176	2649

		ε at 23°C		
		Positive ε	Negative ε	Total
ε at 37°C	Positive ε	266	**6**	272
	Negative ε	**58**	2341	2399
	Total	324	2347	2671

		ε at 30°C		
		Positive ε	Negative ε	Total
ε at 30°C+3%DMSO	Positive ε	404	**1**	405
	Negative ε	**0**	2149	2149
	Total	404	2150	2554

		ε at 30°C		
		Positive ε	Negative ε	Total
ε at 37°C	Positive ε	249	**8**	257
	Negative ε	**31**	2330	2361
	Total	280	2338	2618

		ε at 30°C+3%DMSO		
		Positive ε	Negative ε	Total
ε at 37°C	Positive ε	279	**1**	280
	Negative ε	**70**	2030	2100
	Total	349	2031	2380

CHAPTER 4

TOWARD GENOMEWIDE IDENTIFICATION OF GENETIC INCOMPATIBILITIES IN YEAST

4.1 Abstract

Genetic incompatibility, a form of epistatic interactions between otherwise functional genes in their conspecific genetic background, is commonly considered as the major cause of postzygotic isolation. The Bateson-Dobzhansky-Muller (BDM) model of reproductive isolation by genetic incompatibility is a widely accepted model of speciation. Because of the exceptionally rich biological information about the budding yeast *Saccharomyces cerevisiae*, the identification of BDM incompatibilities in yeast would greatly deepen our understanding of the molecular genetic basis of reproductive isolation and speciation. However, despite repeated efforts, BDM incompatibilities between nuclear genes have never been identified between *S. cerevisiae* and its sister species *S. paradoxus*. Such negative results have led to the belief that simple nuclear BDM incompatibilities do not exist in yeast. Here we explore an alternative explanation that such incompatibilities exist but were undetectable due to limited statistical power. We discover that previously employed statistical methods were not ideal and that a redesigned method improves the statistical power. We determine, under various sample sizes, the probabilities of identifying BDM incompatibilities that cause F1 spore inviability with incomplete penetrance, and confirm that the previously used samples were too small to detect such incompatibilities. Our findings call for an expanded experimental search for yeast BDM incompatibilities, which has become possible with the decreasing cost of genome sequencing.

The improved methodology developed here is in principle applicable to other organisms and can help detect epistasis in general.

4.2 Introduction

Speciation, the “mystery of mysteries” in Darwin’s words (Darwin 1859), is one of the most important processes in evolution, responsible for the generation of the tremendous biodiversity on Earth. Important as it is, speciation is not well understood at the genetic level. For example, it is unknown how many genetic changes underlie the formation of a new species in nature, and the relative roles of natural selection and genetic drift in causing these changes are still debated (Schluter 2009, Nei and Nozawa 2011). A key step in speciation is the establishment of reproductive isolation, which can occur prezygotically or postzygotically (Coyne and Orr 2004). While forms of prezygotic isolation can involve a variety of spatial, behavioral, mechanical and temporal isolation, postzygotic isolation is commonly considered to be majorly caused by genetic incompatibility, a form of deleterious epistatic interaction among multiple genes that have evolved separately for an extended period of time in two different genetic backgrounds. Specifically, the simplest form of the Bateson-Dobzhansky-Muller (BDM) model for two loci asserts that a genetic change at locus *A* in one population and a genetic change at locus *B* in another population may be incompatible when residing in the same genome upon the hybridization between individuals of the two populations, which could result in postzygotic incompatibility and lead to inviability, infertility, or inferiority (Orr 1996). Although this model is generally accepted, only a small number of genes in a few species pairs have been identified to be genetically incompatible (Wu and Ting 2004, Maheshwari and Barbash 2011, Nosil and Schluter 2011). One classical example involves the melanoma formation in the hybrids of *Xiphophorus* species. Normally, the *Tu* locus controls the formation of spots composed of black pigment cells. In interspecific hybrids between the platyfish *X. maculatus* and swordtail *X. helleri*, these spots sometimes spontaneously develop into malignant melanomas (Wittbrodt, Adam et al. 1989). A two-locus BDM model can explain this phenomenon: overexpression of *Tu*, which has been identified to be *Xmrk* on the X chromosome, causes melanomas to form (Adam, Dimitrijevic et al. 1993), while an autosomal repressor gene mapped near *cdkn2a/b* negatively regulates *Tu* (Schartl, Walter et al. 2013). The hybrids that have *Tu* but not the repressor will develop melanomas (Meierjohann, Schartl et al. 2004). There is, however, much disagreement on the existence of such major BDM incompatibilities and their role in speciation in general (Liti, Barton et al. 2006, Maheshwari and Barbash 2011). Identifying such genes and studying their functions and evolution could help settle this debate and uncover the molecular genetic basis of reproductive isolation and speciation. Because BDM incompatibilities are expected to accumulate with the divergence of two species, identifying such incompatibilities from closely related species is most relevant to understanding speciation (Nosil and Schluter 2011).

For four reasons, the budding yeast *Saccharomyces cerevisiae* (*Sc*) and its sister species *S. paradoxus* (*Sp*) are ideal for identifying BDM incompatibilities and studying their roles in speciation. First, *S. cerevisiae* is one of the best-studied eukaryotes, with abundant information on its genetics, genomics, physiology, cell biology, and molecular biology. Numerous genetic tools and molecular methods are readily available for further study. Its short generation time allows rapid genetic analysis and its small genome (~12 million bases) makes genotyping and fine genetic mapping easier than in most of other species. Second, separated ~10 million years ago (Kawahara and Imanishi 2007) and with ~85% genome sequence identity (Kellis, Patterson et al. 2003), *Sc* and *Sp* are relatively closely related. The two species can readily mate with each other (Murphy, Kuehne et al. 2006); yet, their postzygotic isolation is strong, with *Sc*-*Sp* hybrids producing only ~1% viable spores (Hunter, Chambers et al. 1996). Third, the genomes of the two species are essentially collinear with no gross chromosomal rearrangements and no reciprocal translocation; only four inversions and three segmental duplications exist (Kellis, Patterson et al. 2003). This fact eliminates chromosomal rearrangement as a major contributor to their postzygotic isolation. Fourth, the genotypes and phenotypes of yeast haploids can be directly analyzed, avoiding the need to generate homozygotes from the spores produced by F1 hybrids. Note that F1 hybrids are not suitable for identifying genetic incompatibilities unless they are dominant, while a previous study has excluded the existence of dominant genetic incompatibilities underlying the infertility of the hybrid between *Sc* and *Sp* (Greig, Borts et al. 2002). One complication of the yeast system is that a large fraction of spores produced by *Sc*-*Sp* hybrids are killed by aneuploidy (Hunter, Chambers et al. 1996). At least one recombination is usually required for correct segregation of homologous chromosomes during meiosis. In the *Sc*-*Sp* hybrid, the sequence differences between homologous chromosomes cause the mismatch repair system to suppress recombination, resulting in a high frequency of aneuploidy (Chambers, Hunter et al. 1996). Deleting the mismatch repair gene *MSH2* increases the recombination rate in the hybrid from 5.4 to 35.6 crossovers per meiosis (Kao, Schwartz et al. 2010). Consequently, F1 spore viability rises to ~10% (Kao, Schwartz et al. 2010).

Research in the last decade has focused on understanding the genetic basis of *Sc*-*Sp* F1 hybrid infertility, which is equivalent to F1 spore inviability. In spite of the multiple advantages of the study system and repeated efforts (Greig, Borts et al. 2002, Greig 2007, Kao, Schwartz et al. 2010, Xu and He 2011), no nuclear-nuclear genetic incompatibilities have been identified for *Sc*-*Sp* F1 infertility, although a mitochondrial-nuclear incompatibility has been reported for F2 hybrid infertility (Chou, Hung et al. 2010). Two general strategies have been used to identify nuclear-nuclear genetic incompatibilities between *Sc* and *Sp*. The first approach is to replace chromosomes in *Sc* with their *Sp* homologs one at a time. If interchromosomal incompatibilities exist, one would observe a reduction in strain fertility, viability, or growth rate upon a chromosomal replacement. The fact that such replacements were made for at least 9 of the 16 chromosomes demonstrates the lack of BDM incompatibility for F1 spore viability in the 9 chromosomes (Greig 2007). This result, however, does not exclude the possibility of incompatibilities for F1 spore growth rate or higher-order incompatibilities for viability. Note that even when an interchromosomal incompatibility is detected using this approach, further work is needed to localize the incompatible genes.

The second approach is to identify genetic incompatibilities in F1 spores by linkage analysis. Briefly, if the *Sc* allele at locus *A* (A_{Sc}) is incompatible with the *Sp* allele at locus *B* (B_{Sp}), spores of the genotype $A_{Sc}B_{Sp}$ may have reduced viability and thus may be underrepresented among viable F1 spores. This decrease in frequency also applies to pairs of markers closely linked to A_{Sc} and B_{Sp}, respectively. Thus, it is possible to use existing genetic markers such as single nucleotide differences (SNDs) between the two species to map BDM incompatibilities. This approach is virtually identical to mapping genetic interaction or intergenic epistasis. Because of the large number of marker pairs to be tested, the statistical power is expected to be low.

Two groups have used the second approach above to look for incompatibilities between *Sc* and *Sp* that kill F1 spores with 100% penetrance, but with no success (Kao, Schwartz et al. 2010, Xu and He 2011). The negative result has led to the suggestion that simple two-locus BDM incompatibilities do not exist in yeast and are unimportant to yeast speciation (Kao, Schwartz et al. 2010). However, for two reasons, genetic incompatibility need not have 100% penetrance. First, an incompatibility may only increase the probability of spore inviability rather than killing the spore deterministically, because spore viability is likely to be a complex trait controlled by multiple genes. Second, a high-order incompatibility behaves like a two-locus incompatibility with incomplete penetrance. For instance, a three-locus incompatibility with 100% penetrance behaves exactly as a two-locus incompatibility with 50% penetrance. Given the possibility of incomplete penetrance, one wonders what conclusion about the genetic incompatibility between *Sc* and *Sp* can be drawn from the existing data of the linkage analysis. To answer this question, it becomes necessary to understand the properties of this linkage analysis. Here we use computer simulation to inspect the statistical properties of the linkage analysis, under the scenario that two-locus genetic incompatibilities cause F1 spore inviability with incomplete penetrance, which, as aforementioned, includes the possibility of multiple-locus incompatibility. We show that the previously designed statistical method is not ideal and propose a modified method that improves the statistical power. We find previously used sample sizes too small to detect genetic incompatibilities and offer guidelines for future experimental searches of the BDM incompatibilities between *Sc* and *Sp*. The methodology simulated here can be readily applied to determine the sample size and power for studying of BDM incompatibilities in other species pairs, and may also be broadly applied to guide other types of interchromosomal epistasis mapping.

4.3 Methods

4.3.1 General strategy of simulating the identification of BDM incompatibilities

Based on theoretical predictions and experimental results (Welch 2004, Wu and Ting 2004, Lee, Chou et al. 2008), we assume that genetic incompatibility is asymmetric. That is, if A_{Sc} and B_{Sp} are incompatible, A_{Sp} and B_{Sc} can still be compatible (**Figure 4-1A**). We define I as the probability that an F1 spore dies due to an incompatible allelic pair. We consider the use of *msh2* mutants of both *Sc* and *Sp* in this study (Kao, Schwartz et al. 2010) such that spore deaths have three potential causes: random death, aneuploidy, and genetic incompatibility. Random death refers to spore death caused by deleterious mutations, meiotic errors, or environmental factors, and is assumed to have the same rate in the parental species and their hybrid.

The following steps outline the procedure of simulating spore production (**Figure 4-1B**). First, to simulate the hybridization between the two yeast species, we set the *in silico* genome to contain 16 chromosomes with lengths following those of *Sc*. SND density was set to be one per seven nucleotides based on the 85% sequence identity between the two species. We assume *N* pairs of incompatibilities and randomly assign them to the existing SNDs. The effects of these *N* pairs of incompatibilities on F1 spore inviability were either set to be equal or set to follow a certain distribution. The number of crossovers generated during the meiosis of F1 hybrids followed a Poisson distribution with a mean of 35.6 per meiosis (Kao, Schwartz et al. 2010) and the crossovers were randomly assigned to the genome. Meiotic gene conversion and variable recombination rates across the genome are not considered. After meiosis, four spores are generated. We then calculate spore viability as described in the next section and stochastically determine viable spores based on their viabilities.

In the actual experiment, the viable spores may be genotyped by restriction enzyme digestion (Xu and He 2011), microarray-based SND typing (Kao, Schwartz et al. 2010), or genome sequencing. Here we use 1207 SNDs (one per 10 kb) as markers in linkage analysis. Using more markers does not improve the precision or power of identifying BDM incompatibilities because of limited recombination in *msh2 Sc*-*Sp* hybrids: 10,000 nucleotides correspond to 1.5 cM. Use of one marker per 10 kb means that the expected mapping resolution is at best 2.5 kb.

Our preliminary analysis revealed that any BDM incompatibility between two intrachromosomal loci is difficult to detect due to a strong linkage. Hence, we examine the frequencies of spores for every pair of interchromosomal SND markers. That is, for markers *A* and *B* that are located on different chromosomes, we obtain the numbers of spores with the genotypes of $A_{Sc}B_{Sc}$ (*a*), $A_{Sp}B_{Sc}$ (*b*), $A_{Sc}B_{Sp}$,(*c*), and $A_{Sp}B_{Sp}$ (*d*), respectively. These numbers form a 2 × 2 table (**Figure 4-1C**), from which three statistics are calculated: chi-squared value, *G*-test statistic, and odds ratio (*OR*) (see below). Because of viability differences among the four genotypes, the incompatible genotype should have a reduced frequency, compared with its expected value.

In theory, when the sample size is sufficiently large, we should be able to recover the preassigned incompatible allelic pairs. After acquiring a statistic of genetic incompatibility for each pair of markers, we determine statistical significance using a familywise 5% type-I error rate (see below). We then attempt to estimate the chromosomal segments encompassing the incompatibility genes (see below).

4.3.2 Calculating spore viability

In our simulation, random death, aneuploidy, and BDM incompatibility are the three causes of F1 spore inviability. We set the random death rate to be R = 1-0.804 = 0.196, based on the fact that *S. cerevisiae* and *S. paradoxus msh2* mutants have spore viabilities of 84.0% and 80.4%, respectively (Hunter, Chambers et al. 1996). It has been estimated that aneuploidy occurs at a frequency of 0.29 per viable *msh2 Sc-Sp* hybrid spore (Kao, Schwartz et al. 2010), but it is unknown what the corresponding fraction is in dead spores. The impact of aneuploidy on spore viability is complicated. While losing a chromosome is lethal, gaining an extra chromosome could be beneficial if it masks the deleterious effect of genetic incompatibility. We set the probability of spore inviability due to aneuploidy to be either G = 0% or 50% to obtain a minimal and a more realistic estimate of the required sample size for identifying BDM incompatibilities, respectively. Inviability caused by aneuploidy is applied to pairs of sister spores because nondisjunction typically occurs in meiosis I of the hybrid (Hunter, Chambers et al. 1996). We assume no epistasis among incompatible gene pairs. Let T be the fraction of viable spores produced by F1 hybrids, N be the number of BDM incompatibility pairs between *Sc* and *Sp*, and I_k be the probability of spore death caused by the kth pair of incompatibility or penetrance. We have

$$T = (1-R)(1-G)\prod_{k=1}^{N}[0.75+0.25(1-I_k)]\,. \qquad [1]$$

In the simple case of $I_k = I$ for all k values, we have

$$T = (1-R)(1-G)[0.75+0.25(1-I)]^N\,. \qquad [2]$$

4.3.3 Statistics characterizing genetic incompatibility

Genetic incompatibility between A_{Sc} and B_{Sp} leads to a reduction in the frequency of $A_{Sc}B_{Sp}$, compared with its expected value. This signal can be detected in multiple ways. Because of strong linkage within a chromosome, we only evaluate pairs of markers that reside on different chromosomes. In a previous study (Kao, Schwartz et al. 2010), a chi-squared test was used to test if the frequency of a recombinant equals the product of corresponding allele frequencies. For example, if the A_{Sc} and B_{Sc} frequencies among viable F1 spores are 0.3 and 0.5, respectively, the expected frequency of viable $A_{Sc}B_{Sc}$ spores is 0.3×0.5 = 0.15. Chi-squared is then calculated by summing over all genotypes the squared difference between the expected and observed numbers of a genotype divided by the expected number. This test is nondirectional in the sense that it does not distinguish whether the recombinants are overrepresented or underrepresented. Besides the chi-squared test, the *G*-test of independence may be used to test the goodness of fit of the observed genotype frequencies to their expected values. *G*-test is designed for cases where the margins of a 2 × 2 table are not fixed by investigators whereas the total number in the four cells of the table is fixed (Sokal and Rohlf 1995). We conduct the *G*-test with Williams's correction (Sokal and Rohlf 1995). In addition, we calculate an odds ratio by dividing the product of the numbers of the two parental genotypes by that of the two recombinant genotypes: $OR = (a \times d)/(b \times c)$ (**Figure 4-1C**).

Because multiple pairs of markers are tested in an experiment, we evaluate the significance of the above statistics by controlling the familywise type-I error rate. We first randomly shuffle each of the 16 chromosomes among spores and then find the highest statistic among all pairs of markers. We conduct this shuffling 100 times and rank the resulting 100 highest statistics. The 5th largest number among these 100 numbers is chosen as the critical value corresponding to a familywise type-I error rate of 5%.

After applying the cutoff, we group statistically significant pairs of markers as follows. Let us use the odds ratio as an example, but the same procedure applies to the other statistics used. First, we find the maximal odds ratio, and take a step of seven markers on each side of each focal marker to obtain the initial square of close linkage. The number seven is chosen by considering the tradeoff between grouping markers showing signals of different incompatibilities and dividing markers showing the signal of the same incompatibility. Second, we keep expanding the square with a step size of one marker until it is no longer significant or it reaches an end of the chromosome. Third, if two squares overlap with each other, we ignore the square with the lower maximal odds ratio. Fourth, we repeat these steps until all significant pairs of markers are included in the squares. Fifth, the marker pair of the maximal odds ratio of each square is recorded. If two marker pairs in the same square tie for the maximal odds ratio, we record the locations of their midpoints.

A preassigned BDM incompatible pair is considered to be correctly identified when both causal SNDs are within seven markers from the maximum in an aforementioned square. Sensitivity is calculated as the fraction of true incompatible pairs identified. False discovery rate is calculated as the total number of false discoveries divided by the total number of discoveries. When no discovery is made in all simulations, false discovery rate is defined as 0. Genomic distance is calculated as the average distance between the two identified markers and their respective causal SNDs. Standard errors of sensitivity, false discovery rate, and genomic distance estimates are estimated using 1000 bootstrap samples.

4.4 Results

4.4.1 Odds ratio outperforms other statistics in identifying genetic incompatibility

Following Kao and colleagues (Kao, Schwartz et al. 2010), we use *msh2* mutants of *Sc* and *Sp* in our simulation of identifying BDM incompatibilities, unless otherwise noted. Based on theoretical predictions and experimental results (Welch 2004, Wu and Ting 2004, Lee, Chou et al. 2008), we assume that genetic incompatibility is asymmetrical. That is, if A_{Sc} and B_{Sp} are incompatible, A_{Sp} and B_{Sc} can still be compatible (**Figure 4-1A**). It is difficult to detect BDM incompatibility between two loci that reside in the same chromosome because of limited recombination in the hybrid. Hence, we only examine pairs of markers located on different chromosomes. That is, for markers *A* and *B* that are located on different chromosomes, we obtain the numbers of spores with the genotypes of $A_{Sc}B_{Sc}$ (*a*), $A_{Sp}B_{Sc}$ (*b*), $A_{Sc}B_{Sp}$ (*c*), and $A_{Sp}B_{Sp}$ (*d*), respectively, which form a 2×2 table (**Figure 4-1C**). Because of viability differences among the four genotypes, the incompatible genotype should have a reduced frequency, compared with its expected value (**Figure 4-1B**). In theory, when the sample size is sufficiently large, we should be able to detect such incompatible allelic pairs.

We calculate three statistics using the 2×2 table: chi-squared, *G*-test statistic, and odds ratio $OR = (ad)/(bc)$ (see Materials and Methods), and evaluate their relative performances in identifying preassigned incompatibilities by simulation. The chi-squared statistic was previously used in this context (Kao, Schwartz et al. 2010), but this statistic does not differentiate between overrepresentation and underrepresentation of a genotype relative to its expectation and thus may be less specific. Because the chi-squared test is an approximation of the *G*-test, they have similar properties, although *G*-test may be more precise. By contrast, a lower-than-expected *OR* indicates overrepresentation of $A_{Sp}B_{Sc}$ and/or $A_{Sc}B_{Sp}$, whereas a higher-than-expected *OR* indicates depletion of these genotypes, which is predicted under genetic incompatibility. After acquiring a statistic of genetic incompatibility for each interchromosomal marker pair, we determine statistical significance using a familywise 5% type-I error rate to control multiple testing. We then identify the chromosomal segments that are likely to encompass the incompatibility genes (see Materials and Methods).

Because the incompatible marker pairs are preassigned in the simulation, we can evaluate how well the three statistics perform in terms of the (i) sensitivity, (ii) false discovery rate, and (iii) mean genomic distance between the identified markers and the preassigned incompatible SNDs. For each parameter set, we conduct 400 simulation replications and pool the data in our analysis. Sensitivity is the fraction of all preassigned incompatible pairs that are recovered by the analysis. False discovery rate is the number of false discoveries divided by the total number of discoveries. The standard errors of these estimates are estimated by bootstrapping the pooled data 1000 times. There are 12.07 million nucleotides × 15% = 1.8105 million SNDs between *Sp* and *Sc*. We randomly assigned *N* pairs of single nucleotide differences to form *N* incompatibility pairs. In mapping these incompatibilities, however, we use only 1207 markers, or one marker per 10,000 nucleotides, because the use of more markers does not increase mapping resolution due to limited recombination (see Materials and Methods).

We start the simulation with the following parameters. We assume no contribution of aneuploidy to spore inviability, and set $N = 10$ pairs of incompatibilities that have equal effects on inviability. Given the known viability of *msh2* hybrid spores, the 10 pairs each contribute $I = 0.75$ to spore inviability. That is, a spore with one pair of incompatibility is 25% as viable as a spore without any incompatibility. The 10 pairs of incompatibilities (i.e., 20 causal SNDs) are randomly distributed in the 16 yeast chromosomes. The number of viable spores genotyped is $M = 200$. When *OR* is used, the sensitivity is 40%, significantly greater than that of chi-squared (28%) or *G*-test statistic (30%) (**Figure 4-2A**). The false discovery rate under *OR* is 24%, not significantly different from that under the other two statistics (22% and 23%, respectively) (**Figure 4-2B**). The mean genomic distance between the identified marker and the preassigned incompatibility loci is 18.3 kb under *OR*, significantly smaller than that under the other two statistics (19.3 and 19.1 kb, respectively) (**Figure 4-2C**).

If the differences among the three methods are simply due to the fact that chi-squared and *G*-test statistic cannot distinguish whether parental or nonparental types are in excess, we could use the directional information from *OR* and consider only those chi-squared or *G*-test statistic values when *OR* > 1. While such modified chi-squared and *G*-test statistic outperform their original versions in sensitivity, they are still worse than *OR* (**Figure 4-2A**). In terms of the false discovery rate, the modified versions appear worse than the original versions (**Figure 4-2B**). In terms of the genomic distance, the modified versions are similar to the original versions (**Figure 4-2C**). We subsequently confirmed the advantage of *OR* over chi-squared and *G*-test statistic in multiple conditions, by varying *N*, *M*, and the influence of aneuploidy (*G*) (**Table 4-1**). When the genetic incompatibility is symmetrical, however, the advantage of *OR* over chi-squared and *G*-test statistic disappears (**Table 4-2**).

4.4.2 Previous studies were underpowered

To understand why previous experimental searches of nuclear BDM incompatibilities between *Sc* and *Sp* were unsuccessful, we perform a simulation following the scheme of a previous experiment study, which genotyped 58 spores from F1 with *MSH2* and 48 spores from F1 lacking *MSH2* (Kao, Schwartz et al. 2010). Before we started the simulation, we confirmed that no pair of markers in that study (Kao, Schwartz et al. 2010) showed significant *OR* in the familywise test. The simulation parameters used for *msh2* spores are the same as described above. For mismatch repair proficient spores, the random death rate is set to be *R* = 0.05 (Greig, Borts et al. 2002). Given the observed viability of 1% among these spores, the contribution of aneuploidy to spore inviability (*G*) is calculated using Eq. 2 to be 91.54% and 95.77%, for the corresponding numbers of 0% and 50% in *msh2* spores, respectively. To be consistent with the previous study (Kao, Schwartz et al. 2010), we used the density of 1 marker per 2 kb and subsequently combined markers showing no recombination in all 106 spores as a single marker. Using 1 marker per 10 kb yielded similar results.

Starting with different pairs of incompatibilities in the simulation, we calculate the corresponding probabilities of nondiscovery, which is the probability that no marker pair has an *OR* that deviates significantly from the expectation at the familywise 5% level. We first assume equal effects on spore viability from all pairs of incompatibilities. When aneuploidy does not reduce *msh2* spore viability, at least 8 pairs of incompatibilities are required to explain the observed spore inviability. We found the probability of nondiscovery to exceed 0.05 in all cases except when n = 8 (**Figure 4-3A**). If aneuploidy reduces *msh2* spore viability by 50% and correspondingly reduces the viability of *MSH2* spores, there should be at least 5 pairs of incompatibilities. Under this assumption, we found the probability of nondiscovery to exceed 0.05 in all cases except when n = 5 (**Figure 4-3B**). Thus, it is possible for the previous experiment to have missed all incompatibilities. Our analysis tends to overestimate the power of the previous study, because segments in spores with aneuploidy were ignored in the experimental study (Kao, Schwartz et al. 2010) such that the actual sample size is smaller than the number of sampled spores. Furthermore, we have not considered genotyping errors, which would further decrease the statistical power. It might seem counter-intuitive that the more pairs of genetic incompatibility there are, the more difficult it is to identify any of them. The underlying reason is that the total contribution of all incompatibility pairs on inviability is fixed in this simulation and that all pairs are assumed to contribute equally. Thus, having a larger number of incompatible pairs means a smaller contribution from each pair.

Because multiple pairs of genetic incompatibility are unlikely to have equal effect sizes on spore viability, it would be more realistic to consider unequal effect sizes. The difficulty, however, is that there is no prior knowledge on the effect size distribution. Because BDM incompatibilities may be similar to loss-of-function mutations (Maheshwari and Barbash 2011), we assume that the effect size distribution follows the distribution of the deleterious fitness effects of single-nonessential-gene deletions in yeast (Qian, Ma et al. 2012). We randomly sample *I* from this distribution until the total incompatibility explains the observed spore inviability. The mode of the number of incompatible pairs required to explain the observed spore inviability is 150 (**Figure 4-3C**) and 100 (**Figure 4-3D**) when the contribution of aneuploidy to *msh2* spore inviability is 0 and 50%, respectively. The corresponding distributions of *I* under the two scenarios used in this simulation study are presented in **Figure 4-3C** and **Figure 4-3D**, respectively, and the probability of nondiscovery is 79% (**Figure 4-3A**) and 77% (**Figure 4-3B**), respectively.

Because Kao et al.'s (2009) study was the largest experiment for identifying BDM incompatibilities between *Sc* and *Sp*, our results suggest that all previous studies on the subject were not sufficiently powerful to detect BDM incompatibilities between the two yeasts.

4.4.3 Sample sizes required for identifying BDM incompatibilities

How many viable spores should be genotyped in order to identify BDM incompatibilities with a reasonable success rate? Here we again assume the exclusive use of *msh2* strains in the experiment. Under the assumption of no effect from aneuploidy on viability, we examine the sceneries of N = 8, 10, and 15 incompatible pairs with equal effects, respectively. We use the sample size of M = 100, 200, 400, and 800 spores, respectively. In the case of N = 8, the probability of nondiscovery is negligible even when M = 100 (**Figure 4-4A**). In the case of N = 10 and 15, the probability of nondiscovery declines quickly as M increases from 100 to 200 and 400 (**Figure 4-4A**). As expected, the total number of discoveries increases with the sample size M (**Figure 4-4B**), so does the sensitivity (**Figure 4-4C**). By contrast, the false discovery rate and the mean genomic distance between the causal SNDs and the identified markers generally decline with *M* (**Figure 4-4D**). We also examined the situation when the probability of *msh2* spore inviability due to aneuploidy is 50%, and obtained overall similar results (**Figure 4-4F-J**). **Figure 4-5** shows randomly picked examples of our simulation results under various *M* when *N* is fixed at 10 and *G* at 0. Because one incompatibility pair happens to reside on the same chromosome, the maximal number of pairs detectable is 9. It is clear how increasing the sample size increases the power of detection.

To obtain a more realistic estimate of the required sample size for detecting incompatibilities, we use the aforementioned unequal effect sizes depicted in Figure 4-3C and D, respectively. Because, under this model, most incompatibilities have small effects, which are hard to detect, we focus on incompatibilities with *I* > 0.2 and its subset that has *I* > 0.4, respectively, when evaluating sensitivity, false discovery rate, and genomic distance. The probability of nondiscovery, however, is evaluated as originally defined. As aforementioned, when there is no contribution of aneuploidy to *msh2* spore inviability, 150 incompatibility pairs are required to explain the observed spore inviability. Among them, 10 pairs have *I* > 0.2, four of which have *I* > 0.4 (**Figure 4-3C**). When there is a 50% contribution of aneuploidy to *msh2* spore inviability, 100 incompatibility pairs are required to explain the observed spore inviability. Among them, six pairs have *I* > 0.2, two of which have *I* > 0.4 (**Figure 4-3D**). Our simulation (**Figure 4-6**) shows that a much larger sample is required for successful detection of BDM incompatibilities under unequal effect sizes than under equal effect sizes. For example, when *M* = 1600, the probability of nondiscovery becomes negligible (**Figure 4-6A**, **E**). With such a large sample, the sensitivity is ~40% for *I* > 0.2 and ~80% for *I* > 0.4 (**Figure 4-6B**, **F**) and the false discovery rate is ~30% for $I > 0.2$ and ~50% for $I > 0.4$ (**Figure 4-6C**, **G**). The genomic distance is between 30 and 40 kb for both $I > 0.2$ and $I > 0.4$ (**Figure 4-6D**, **H**).

4.5 Discussion

In this chapter, we demonstrate that odds ratio outperforms chi-squared and *G*-test statistic in detecting asymmetrical BDM incompatibility through linkage analysis. Our simulation suggests that the existence of two-locus BDM incompatibility between *Sc* and *Sp* cannot be excluded and its nondiscovery in previous yeast experiments could be due to the limited sample size and low statistical power. Our study provides important guidelines for designing experiments for identifying yeast BDM incompatibilities and for interpreting potential experimental outcomes. More generally, it highlights the importance of understanding the statistical properties of an experimental method (e.g., sensitivity and false discovery rate) in order to use it efficiently and interpret the result correctly.

We made several assumptions in our simulation that are worth discussion. First, for simplicity, we assumed that recombination rates are equal throughout the genome and ignored recombination hot/cold spots and interferences between crossovers (Mancera, Bourgon et al. 2008). This assumption should not affect the overall results because of the relatively low marker density used (one per 10 kb). But recombination rate variation would make the genomic distances between the causal SNDs and the identified markers more variable across the genome. Second, due to the lack of prior knowledge on the distribution of *I*, we assumed either equal *I* values for different incompatibility pairs or unequal *I* values that follow a specific distribution mimicking the fitness effects of gene deletions. We believe that the result from the unequal *I* are closer to the truth than that from the equal *I*. Third, we assumed that BDM incompatibility is asymmetrical, which is in accordance with the theory and most of the incompatible pairs identified so far (Wu and Beckenbach 1983, Meierjohann, Schartl et al. 2004, Welch 2004). Nevertheless, our test still works even when it is symmetrical. Fourth, it is unclear how much aneuploidy affects viability in *msh2* spores, and we used 0% and 50%, respectively, in our study to have a sense of the range of possible outcomes. Fifth, we assumed no error in genotyping the spores. Although genotyping errors would reduce the statistical power, we expect the genotyping error rate to be low, especially when high-coverage next-generation DNA sequencing is used. Moreover, due to low recombination, nearby SNDs can be used for correction of sequencing errors at specific positions. Sixth, we did not explicitly study high-order incompatibility, but because high-order incompatibility is equivalent to two-locus incompatibility with incomplete penetrance, our results apply to high-order incompatibility. For example, I = 0.5 in a two-locus incompatibility (**Figure 4-3**) is equivalent to a three-locus incompatibility with 100% penetrance.

In our simulation, we used one marker per 10 kb to look for BDM incompatibility. Although next-generation sequencing-based genotyping will offer much more markers, the extra markers do not enhance the mapping resolution, because the low recombination rate in *msh2* F1 makes all markers within a 10 kb segment almost completely linked. Due to this property, pairs of incompatible genes that are located on the same chromosome are difficult to detect and therefore are not examined in our simulation. Intrachromosomal incompatible gene pairs are expected to constitute only 7.54% of all incompatible pairs if incompatibility genes are uniformly distributed in the genome.

We found that, by the current method, much larger samples than previously used are required for identifying yeast BDM incompatibilities with incomplete penetrance. Given the rapid increase in DNA sequencing capacity and decline in sequencing cost, genotyping ~1000 spores is no longer out of reach. In fact, a recent study sequenced the genomes of 1000 F2 individuals from a genetic cross between two yeast strains in order to map quantitative traits (Bloom, Ehrenreich et al. 2013). Our simulation shows that by genotyping 800 to 1600 F1 spores, the chance of identifying genetic incompatibilities with relatively high penetrance (>20%) is not small.

Given the power of today's DNA sequencing capacity, an alternative strategy of identifying BDM incompatibility may be used. This strategy involves two steps. First, because an incompatibility allele (e.g., A_{Sc} in Figure 4-1A) has a fitness of 1-0.25I, relative to its alternative (e.g., A_{Sp}), it is relatively easy to identify it by sequencing a pool of viable F1 spores en masse. Second, after identifying low-fitness alleles, one can then look for their incompatible partners by sequencing individual spores. Because of the reduced number of marker pairs to be tested, the sample size required in the second step will be much smaller. A critical requirement in this design is to minimize the competition among spores in mitotic growth before sequencing them en masse, because allelic differences in growth rate between *Sc* and *Sp* that are unrelated to the incompatibility for spore viability may be common.

Although *Sc* and *Sp* are used here to parameterize our simulation study, our methodology and results are useful for mapping recessive genetic incompatibilities in other species when the haploid stage can be assayed, including species with haplontic or haploid-diploid life cycles and diplontic species that can undergo homozygous diploidization. Because BDM incompatibility is a type of intergenic epistasis, our methods and results also apply to other types of genomic detection of epistasis.

4.6 References

Adam, D., N. Dimitrijevic and M. Schartl (1993). "Tumor suppression in Xiphophorus by an accidentally acquired promoter." Science **259**(5096): 816-819.

Bloom, J. S., I. M. Ehrenreich, W. T. Loo, T. L. Lite and L. Kruglyak (2013). "Finding the sources of missing heritability in a yeast cross." Nature **494**(7436): 234-237.

Chambers, S. R., N. Hunter, E. J. Louis and R. H. Borts (1996). "The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss." Mol Cell Biol **16**(11): 6110-6120.

Chou, J. Y., Y. S. Hung, K. H. Lin, H. Y. Lee and J. Y. Leu (2010). "Multiple molecular mechanisms cause reproductive isolation between three yeast species." PLoS Biol **8**(7): e1000432.

Coyne, J. A. and H. A. Orr (2004). Speciation. Sunderland, Sinauer Associates.

Darwin, C. (1859). On the origin of species by means of natural selection. London, John Murray.

Greig, D. (2007). "A screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast." PLoS Genet **3**(2): e21.

Greig, D., R. H. Borts, E. J. Louis and M. Travisano (2002). "Epistasis and hybrid sterility in Saccharomyces." Proc Biol Sci **269**(1496): 1167-1171.

Hunter, N., S. R. Chambers, E. J. Louis and R. H. Borts (1996). "The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid." EMBO J **15**(7): 1726-1733.

Kao, K. C., K. Schwartz and G. Sherlock (2010). "A genome-wide analysis reveals no nuclear dobzhansky-muller pairs of determinants of speciation between S. cerevisiae and S. paradoxus, but suggests more complex incompatibilities." PLoS Genet **6**(7): e1001038.

Kawahara, Y. and T. Imanishi (2007). "A genome-wide survey of changes in protein evolutionary rates across four closely related species of Saccharomyces sensu stricto group." BMC Evol Biol **7**: 9.

Kellis, M., N. Patterson, M. Endrizzi, B. Birren and E. S. Lander (2003). "Sequencing and comparison of yeast species to identify genes and regulatory elements." Nature **423**(6937): 241-254.

Lee, H. Y., J. Y. Chou, L. Cheong, N. H. Chang, S. Y. Yang and J. Y. Leu (2008). "Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species." Cell **135**(6): 1065-1073.

Liti, G., D. B. Barton and E. J. Louis (2006). "Sequence diversity, reproductive isolation and species concepts in Saccharomyces." Genetics **174**(2): 839-850.

Maheshwari, S. and D. A. Barbash (2011). "The genetics of hybrid incompatibilities." Annu Rev Genet **45**: 331-355.

Mancera, E., R. Bourgon, A. Brozzi, W. Huber and L. M. Steinmetz (2008). "High-resolution mapping of meiotic crossovers and non-crossovers in yeast." Nature **454**(7203): 479-U471.

Meierjohann, S., M. Schartl and J. N. Volff (2004). "Genetic, biochemical and evolutionary facets of Xmrk-induced melanoma formation in the fish Xiphophorus." Comp Biochem Physiol C Toxicol Pharmacol **138**(3): 281-289.

Murphy, H. A., H. A. Kuehne, C. A. Francis and P. D. Sniegowski (2006). "Mate choice assays and mating propensity differences in natural yeast populations." Biol Lett **2**(4): 553-556.

Nei, M. and M. Nozawa (2011). "Roles of mutation and selection in speciation: from Hugo de Vries to the modern genomic era." Genome Biol Evol **3**: 812-829.

Nosil, P. and D. Schluter (2011). "The genes underlying the process of speciation." Trends Ecol Evol **26**(4): 160-167.

Orr, H. A. (1996). "Dobzhansky, Bateson, and the genetics of speciation." Genetics **144**(4): 1331-1335.

Qian, W., D. Ma, C. Xiao, Z. Wang and J. Zhang (2012). "The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast." Cell Rep **2**(5): 1399-1410.

Schartl, M., R. B. Walter, Y. Shen, T. Garcia, J. Catchen, A. Amores, I. Braasch, D. Chalopin, J. N. Volff, K. P. Lesch, A. Bisazza, P. Minx, L. Hillier, R. K. Wilson, S. Fuerstenberg, J. Boore, S. Searle, J. H. Postlethwait and W. C. Warren (2013). "The genome of the platyfish, Xiphophorus maculatus, provides insights into evolutionary adaptation and several complex traits." Nat Genet.

Schluter, D. (2009). "Evidence for ecological speciation and its alternative." Science **323**(5915): 737-741.

Sokal, R. R. and F. J. Rohlf (1995). Biometry, W. H. Freeman and company.

Welch, J. J. (2004). "Accumulating Dobzhansky-Muller incompatibilities: reconciling theory and data." Evolution **58**(6): 1145-1156.

Wittbrodt, J., D. Adam, B. Malitschek, W. Maueler, F. Raulf, A. Telling, S. M. Robertson and M. Schartl (1989). "Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in Xiphophorus." Nature **341**(6241): 415-421.

Wu, C. I. and A. T. Beckenbach (1983). "Evidence for Extensive Genetic Differentiation between the Sex-Ratio and the Standard Arrangement of DROSOPHILA PSEUDOOBSCURA and D. PERSIMILIS and Identification of Hybrid Sterility Factors." Genetics **105**(1): 71-86.

Wu, C. I. and C. T. Ting (2004). "Genes and speciation." Nat Rev Genet **5**(2): 114-122.

Xu, M. and X. He (2011). "Genetic incompatibility dampens hybrid fertility more than hybrid viability: yeast as a case study." PLoS One **6**(4): e18341.

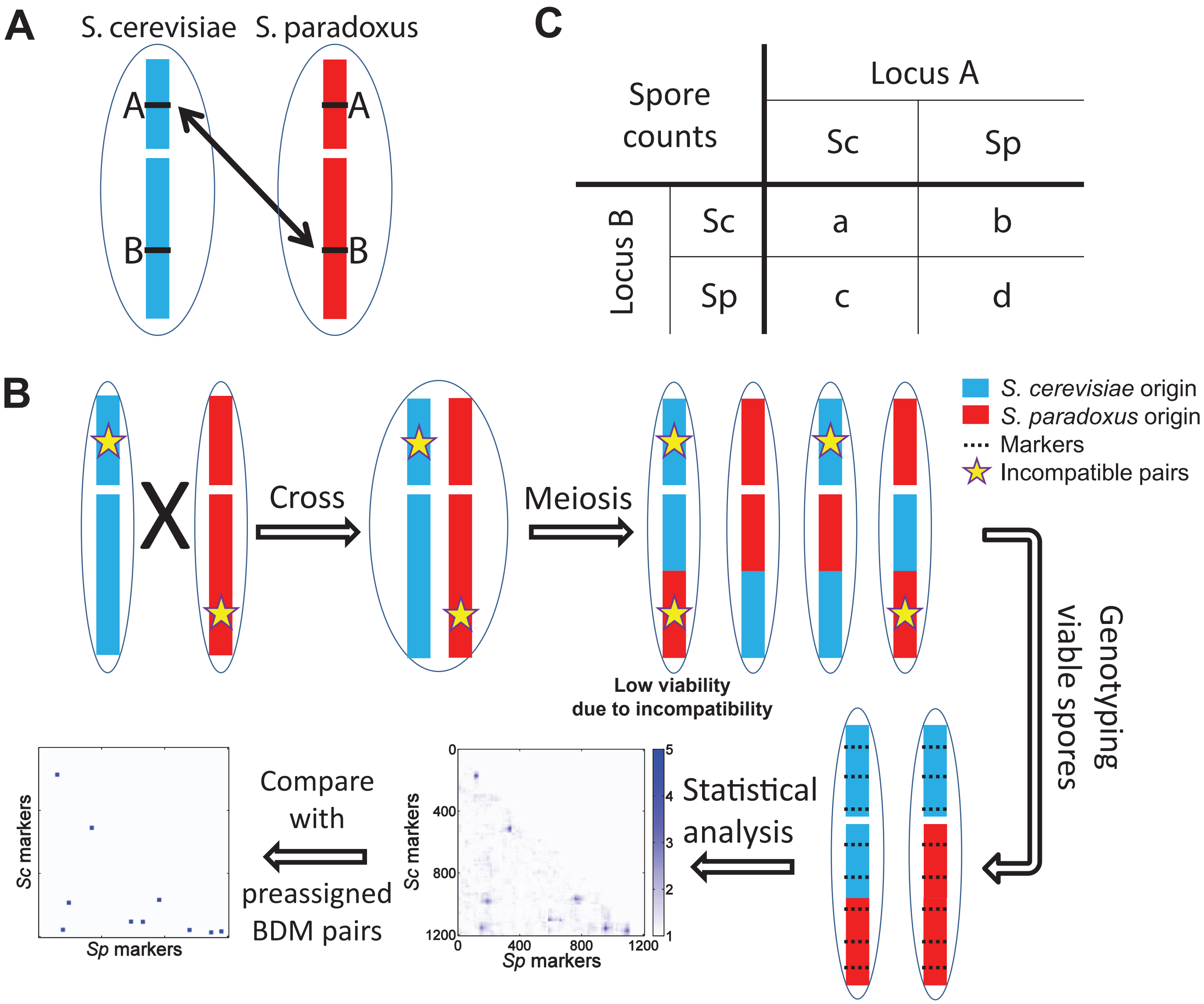


Figure 4-1. General strategy of simulating the identification of BDM incompatibilities between *S. cerevisiae* (*Sc*) and *S. paradoxus* (*Sp*). (**A**) The *Sc* allele at locus *A* and the *Sp* allele at locus *B* are incompatible, leading to reduced viability when in the same spore. (**B**) Procedure for detecting BDM incompatibility between *Sc* and *Sp*. (**C**) A 2×2 table for spore counts of each marker pair. Several statistics for genetic incompatibility are computed using these counts.

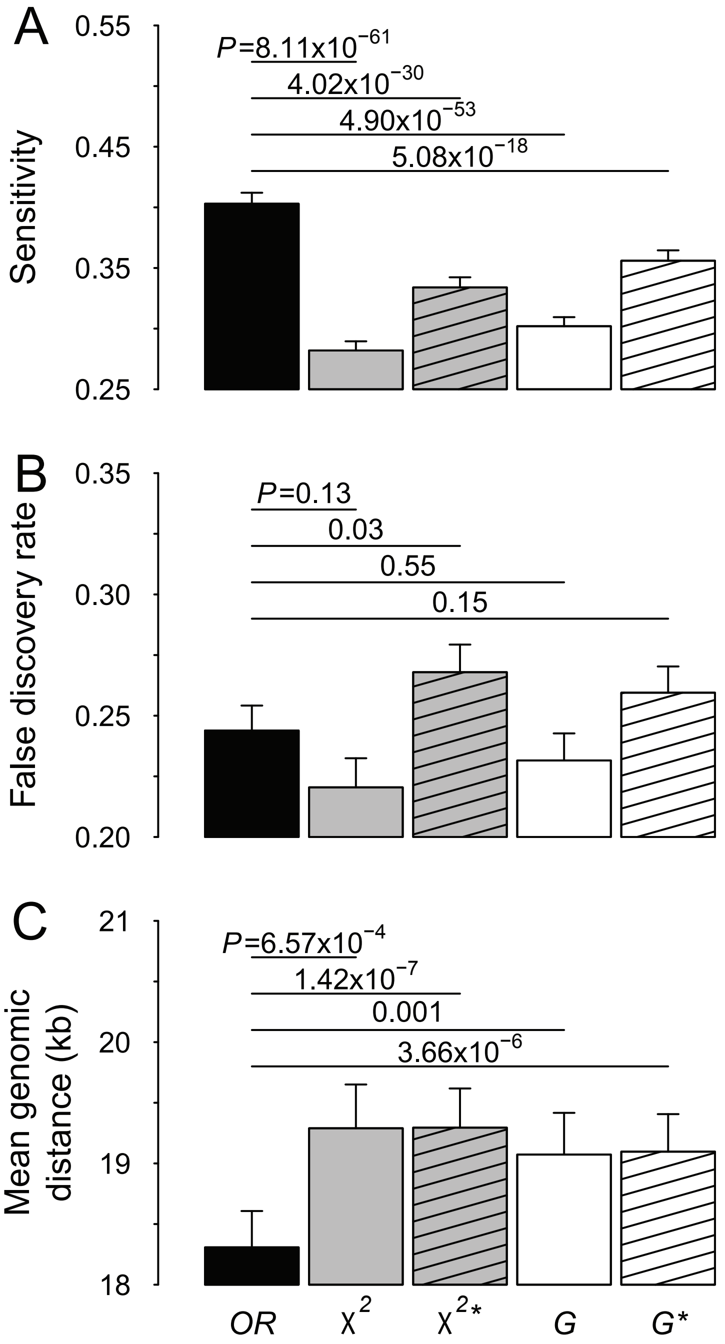


Figure 4-2. Performances of odds ratio, chi-squared, and *G*-test statistic for detecting BDM incompatibilities. Data shown are from 400 simulations of 10 incompatible pairs with equal *I* and no contribution of aneuploidy to spore inviability. The sample size is 200 viable spores. *OR*, χ^2, and *G* represent odds ratio, chi-squared, and *G*-test statistic, respectively. χ^{2*} and G^* respectively consider χ^2 and *G* only when $OR > 1$. Standard error is estimated by 1000 bootstrap replications. (**A**) Sensitivity of the five tests. *P*-values are from paired *t*-test. (**B**) False discovery rates of the five tests. (**C**) Average genomic distance between preassigned incompatibilities and the identified significant markers.

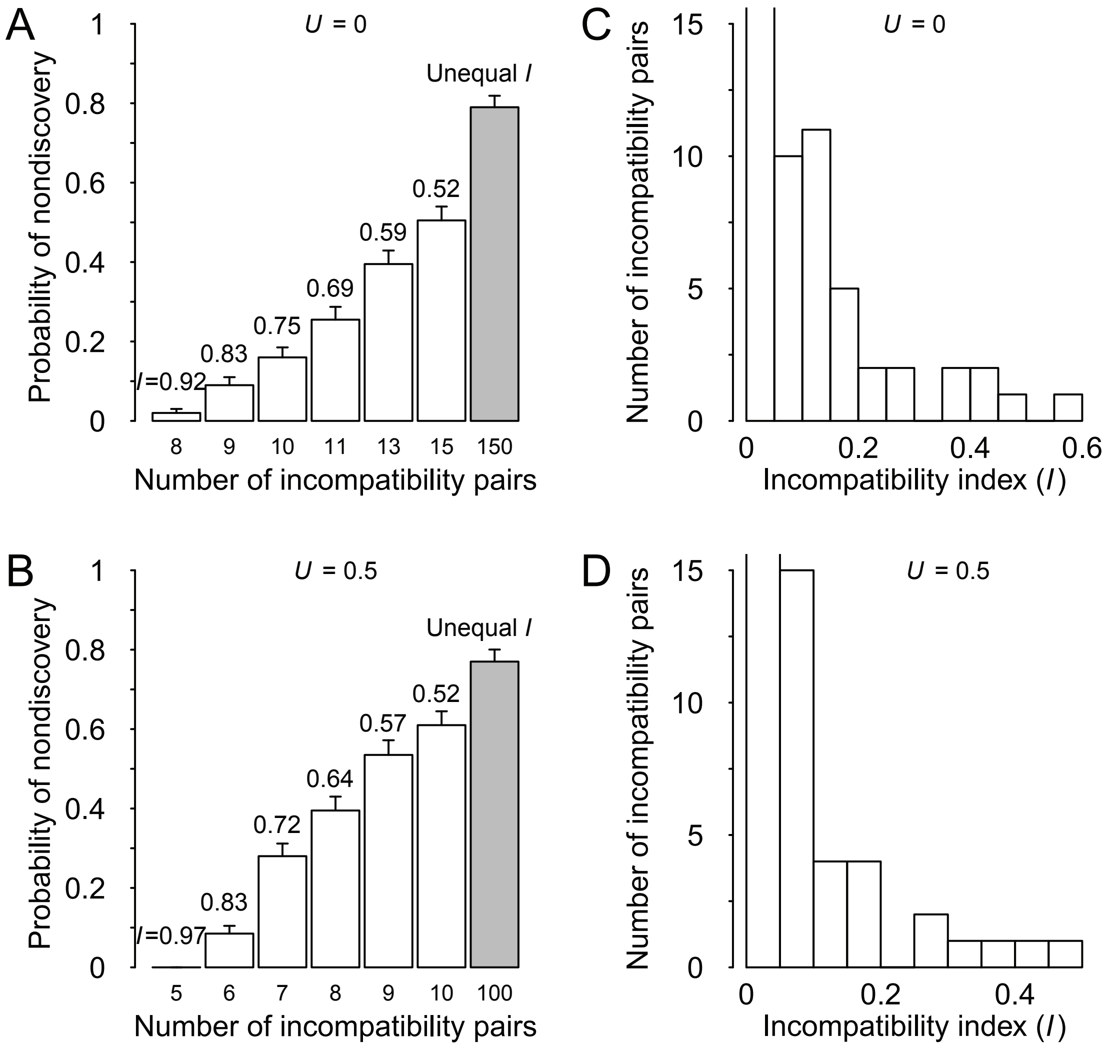


Figure 4-3. Sample size in Kao et al. (2009) is too small to detect BDM incompatibilities with incomplete penetrance. Data shown are from 200 simulations for each parameter set used. (**A**) Probability of nondiscovery in Kao et al's study when aneuploidy is assumed to cause no *msh2* spore inviability (*G* = 0). White bars show the results for various pairs of equal-effect (i.e., equal-penetrance) incompatibilities, while the grey bar shows the result for unequal effects of 150 incompatibility pairs as described in panel C. (**B**) Probability of nondiscovery in Kao et al's study when aneuploidy is assumed to cause *G* = 50% inviability to *msh2* spores. White bars show the results for various pairs of equal-effect incompatibilities, while the grey bar shows the result for unequal effects of 100 incompatibility pairs as described in panel D. (**C**) Distribution of the effect sizes (i.e., penetrances) of 150 BDM incompatibility pairs (under *G* = 0) considered for the grey bar of panel A. (**D**) Distribution of the effect sizes of 100 BDM incompatibility pairs (under *G* = 50%) considered for the grey bar of panel C.

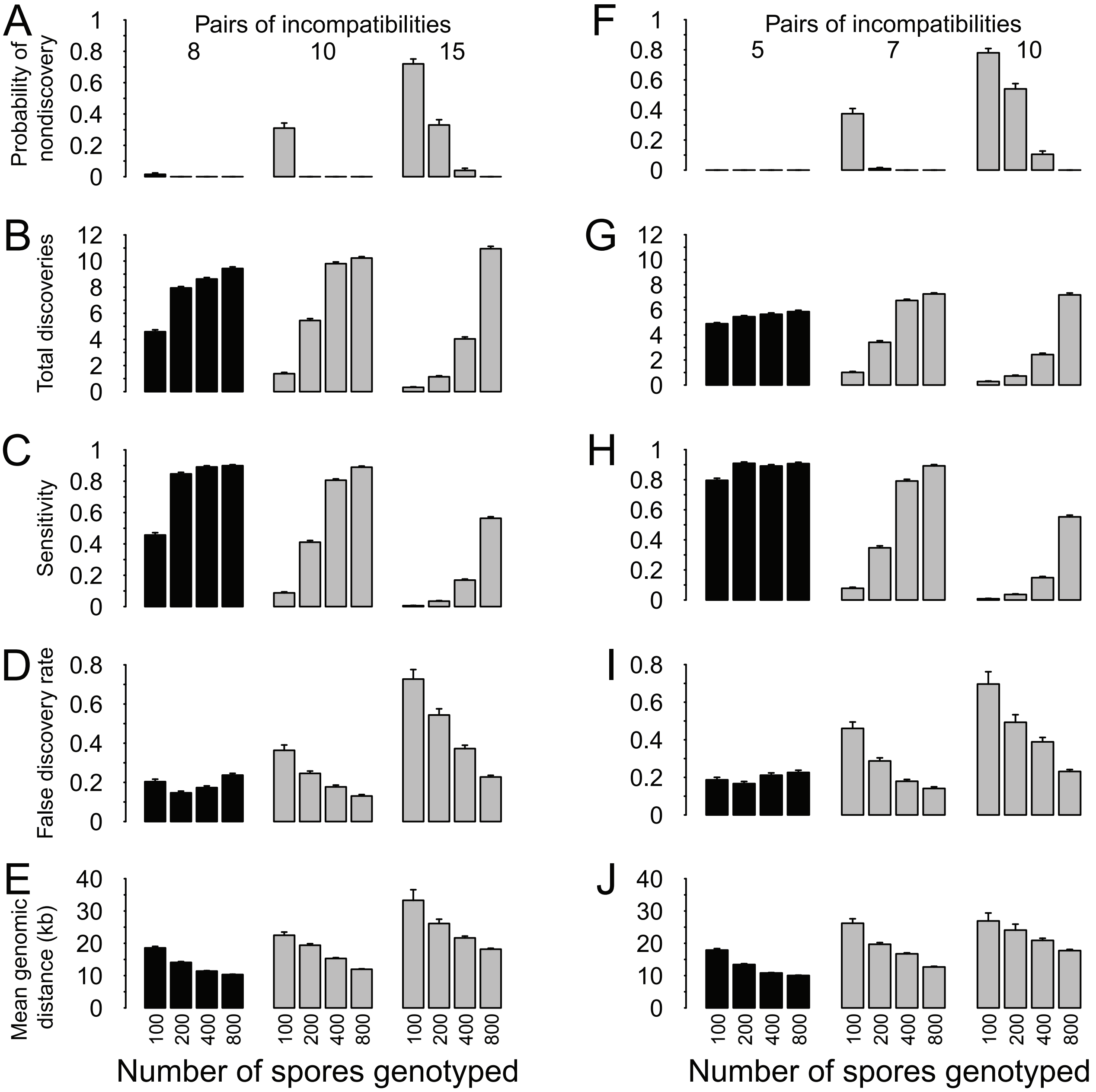


Figure 4-4. Genotyping more F1 spores improves the efficiency of identifying BDM incompatibilities with equal effect sizes. (**A**) Probability of nondiscovery, (**B**) Number of total discoveries, (**C**) sensitivity, (**D**) false discovery rate, and (**E**) mean genomic distance between the preassigned and identified incompatibilities, when aneuploidy is assumed to have no impact on spore inviability. (**F**) Probability of nondiscovery, (**G**) Number of total discoveries, (**H**) sensitivity, (**I**) false discovery rate, and (**J**) mean genomic distance between the preassigned and identified incompatibilities, when aneuploidy is assumed to cause a 50% probability of spore inviability. Data shown are from 200 simulations per parameter set.

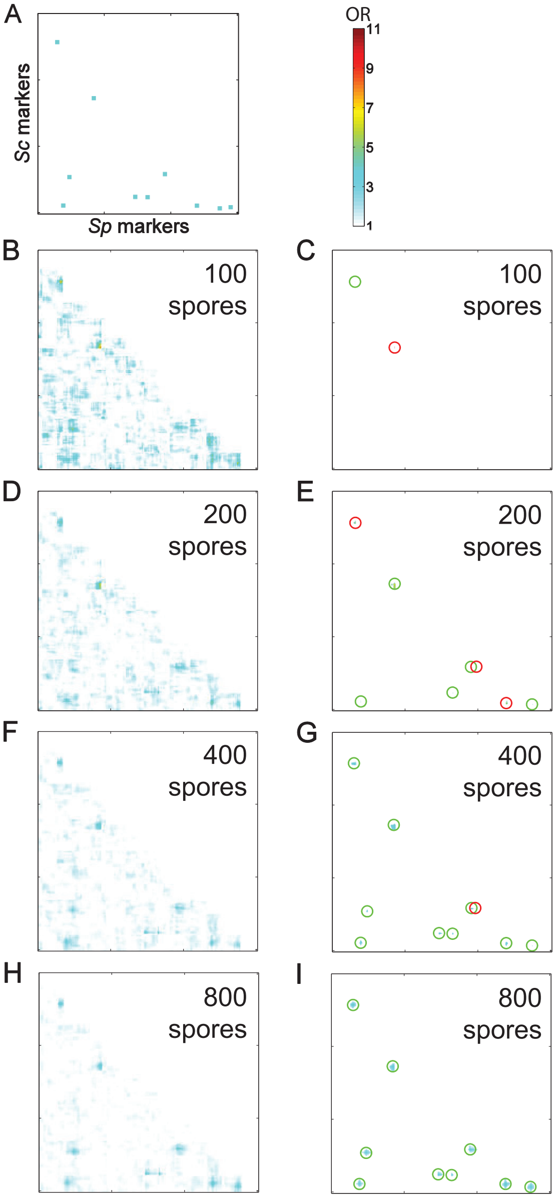


Figure 4-5. An example showing the benefit of using large samples in identifying genetic incompatibilities. (**A**) Genomic positions of 10 pairs of randomly placed equal-size genetic incompatibilities in the simulation. Genomic positions are defined by marker numbers on both axes. Note that one pair of incompatibility near marker #1200 on both axes are located in the same chromosome and therefore are undetectable in our study because only interchromosomal marker pairs are examined. Color shows the expected odds ratio. Spore viability is assumed to be immune to aneuploidy. (**B, D, F, H**) Odds ratios (*OR*s) for all interchromosomal marker pairs when the sample size (number of viable *msh2* spores genotyped) is (**B**) 100, (**D**) 200, (**F**) 400, and (**H**) 800, respectively. *OR* < 1 is not shown. (**C, E, G, I**) Interchromosomal marker pairs whose *OR* values are significant at the familywise 5% level, when the sample size is (**C**) 100, (**E**) 200, (**G**) 400, and (**I**) 800, respectively. The identified incompatibilities are circled, with the correct identifications in green and incorrect identifications in red. Note that an incompatible pair is considered to be correctly identified only when both loci of a preassigned pair are within 7 markers (i.e., 70 kb) from an identified *OR* peak. X and Y labels in (**B-I**) are the same as in (**A**).

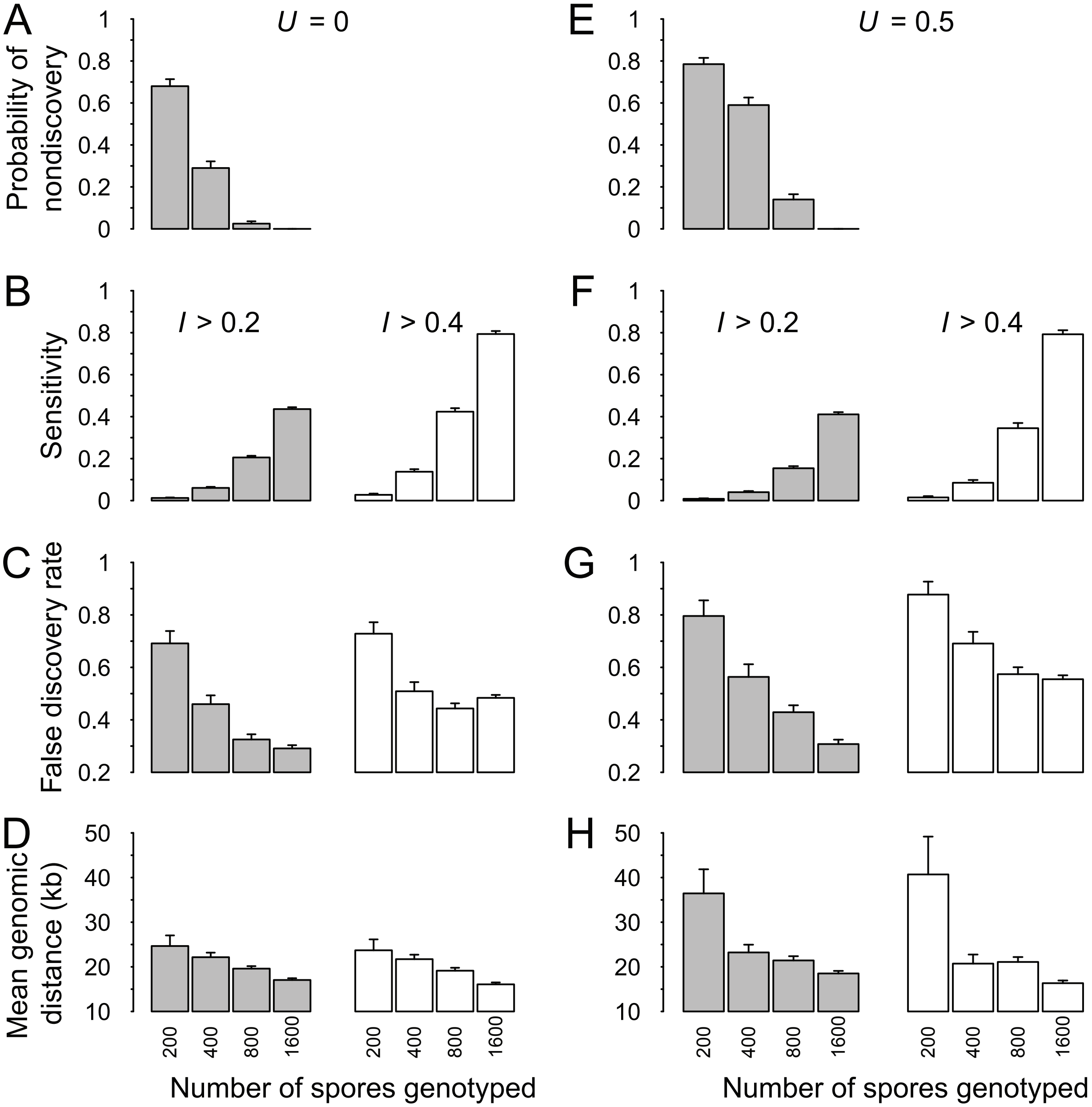


Figure 4-6. Genotyping more F1 spores improves the efficiency of identifying BDM incompatibilities with unequal effect sizes. (**A**) Probability of nondiscovery, (**B**) sensitivity, (**C**) false discovery rate, and (**D**) mean genomic distance between the preassigned and identified incompatibilities, when aneuploidy is assumed to have no impact on spore inviability. The effect sizes of the 150 incompatibility pairs are shown in Fig. 3C. We only show results for incompatibilities with $I > 0.2$ and $I > 0.4$, respectively. Probability of nondiscovery refers to the probability of no significant marker pair regardless of effect size. (**E**) Probability of nondiscovery, (**F**) sensitivity, (**G**) false discovery rate, and (**H**) mean genomic distance between the preassigned and identified incompatibilities, when aneuploidy is assumed to cause a 50% probability of spore inviability. The effect sizes of the 100 incompatibility pairs are shown in Fig. 3D. Data shown are from 200 simulations per parameter set.

Table 4-1. Odds ratio outperforms other statistics in detecting asymmetrical genetic incompatibilities. The results are from 400 simulations for each parameter set, with * showing $P < 0.05$ and $^{+}$ showing $P < 0.005$ when comparing the performance of a statistic with that of odds ratio by a paired *t*-test.

Parameters				**Sensitivity (%)**					**False discovery rate (%)**					**Genomic distance (kb)**				
G[1]	N[2]	I[3]	M[4]	OR[5]	χ^2[6]	G[7]	$\chi^{2}*$[8]	$G*$[9]	OR	χ^2	G	$\chi^{2}*$	$G*$	OR	χ^2	G	$\chi^{2}*$	$G*$
0	8	0.92	100	45.4	22.8^{+}	30.6^{+}	27.8^{+}	35.7^{+}	19.4	26.3^{+}	23.5^{*}	30.4^{+}	27.1^{+}	18.7	20.2^{*}	20.1^{+}	20.4^{+}	20.3^{+}
0	8	0.92	200	85.8	77.3^{+}	81.2^{+}	79.4^{+}	82.2^{+}	15.4	15.2	14.7	17.8^{+}	17.2^{+}	14.3	15.9^{+}	15.3^{+}	15.8^{+}	15.3^{+}
0	8	0.92	400	88.9	88.8	88.9	88.8	88.8	17.7	12.9^{+}	13.3^{+}	15.7^{+}	16.0^{*}	11.4	12.4^{+}	12.0^{+}	12.4^{+}	12.0^{+}
0	8	0.92	800	88.4	88.7^{*}	88.6	88.5	88.5	23.1	18.5^{+}	18.5^{+}	20.4^{+}	20.6^{+}	10.3	10.5^{+}	10.4^{*}	10.5^{+}	10.4^{*}
0	10	0.75	100	8.0	4.1^{+}	4.9^{+}	5.7^{+}	6.7^{+}	37.1	45.8^{+}	43.4^{+}	47.6^{+}	45.7^{+}	25.0	24.1	24.2	24.0	24.1
0	10	0.75	200	40.3	28.2^{+}	30.2^{+}	33.4^{+}	35.6^{+}	24.4	22.0	23.2	26.8^{*}	25.9	18.3	19.3^{+}	19.1^{+}	19.3^{+}	19.1^{+}
0	10	0.75	400	81.8	75.8^{+}	76.9^{+}	77.5^{+}	78.3^{+}	17.2	17.0	16.5	19.0^{+}	18.6^{*}	15.3	16.4^{+}	16.2^{+}	16.4^{+}	16.2^{+}
0	10	0.75	800	88.1	87.9	88.0	87.9	88.1	15.2	12.9^{+}	12.9^{+}	15.3	15.2	12.0	12.7^{+}	12.6^{+}	12.7^{+}	12.6^{+}
0	15	0.52	100	0.7	0.2^{+}	0.3^{+}	0.4^{*}	0.4^{*}	62.3	82.4	78.9	78.5	78.5	30.2	43.4	38.7	35.9	33.9
0	15	0.52	200	3.8	2.4^{+}	2.6^{+}	3.2^{+}	3.2^{+}	48.4	49.3	47.7	54.5	55.1	25.2	26.2	26.3	26.5	26.1
0	15	0.52	400	17.1	11.8^{+}	11.9^{+}	14.5^{+}	14.8^{+}	35.7	35.4	35.4	37.9^{+}	37.5^{+}	21.8	21.8	21.7	22.1	22.1
0	15	0.52	800	56.8	49.3^{+}	50.0^{+}	53.0^{+}	53.4^{+}	23.1	21.2^{+}	21.1^{+}	24.0^{+}	23.8^{*}	18.3	18.8^{+}	18.7^{+}	18.7^{+}	18.7^{+}
0.5	5	0.97	100	77.5	36.0^{+}	48.9^{+}	42.2^{+}	53.9^{+}	18.0	25.3^{+}	24.0^{+}	27.9^{+}	26.3^{+}	18.1	19.4^{+}	20.2^{+}	19.7^{+}	20.6^{+}
0.5	5	0.97	200	90.8	87.9^{+}	90.0^{+}	88.2^{+}	90.0^{+}	16.6	14.0^{*}	13.4^{+}	16.6	17.2	13.2	14.7^{+}	14.3^{+}	14.7^{+}	14.3^{+}
0.5	5	0.97	400	89.6	89.8	89.8	89.7	89.8	17.4	13.1^{+}	13.5^{+}	15.9^{*}	16.3	10.7	11.5^{+}	11.1^{+}	11.5^{+}	11.1^{+}
0.5	5	0.97	800	90.2	90.3	90.3	90.3	90.2	20.8	17.0^{+}	17.0^{+}	19.7	19.8	9.8	9.9	9.8	9.9	9.8
0.5	7	0.72	100	7.2	3.1^{+}	3.6^{+}	4.2^{+}	4.6^{+}	40.8	48.2^{*}	46.0^{*}	53.2^{+}	53.1^{+}	23.9	27.8^{*}	27.2^{*}	25.8	25.5
0.5	7	0.72	200	33.5	18.8^{+}	20.8^{+}	23.2^{+}	25.0^{+}	26.7	31.8^{+}	30.6^{+}	33.9^{+}	32.8^{+}	19.6	20.3	20.3	20.3^{*}	19.9
0.5	7	0.72	400	77.7	68.6^{+}	70.1^{+}	71.8^{+}	72.6^{+}	18.7	17.9	17.9	20.8^{+}	20.7^{+}	16.6	17.4^{+}	17.1^{+}	17.4^{+}	17.2^{+}
0.5	7	0.72	800	89.2	88.6^{+}	88.8^{*}	88.5^{+}	88.7^{+}	14.4	13.2	13.2	15.5^{*}	15.3^{*}	12.8	13.6^{+}	13.4^{+}	13.6^{+}	13.4^{+}
0.5	10	0.52	100	0.9	0.4^{+}	0.5^{+}	0.7	0.7	66.7	75.0	76.2^{*}	79.7	79.1	27.2	32.9	31.8	30.0	29.3
0.5	10	0.52	200	3.4	1.8^{+}	1.8^{+}	2.5^{+}	2.7^{+}	53.8	53.9	55.7	62.0^{*}	60.5	23.3	25.3^{*}	25.4^{*}	24.3	24.3
0.5	10	0.52	400	17.2	11.2^{+}	11.6^{+}	14.2^{+}	14.6^{+}	35.6	37.7	37.1	39.9^{+}	39.4^{*}	20.6	21.1	21.0	20.9	21.0
0.5	10	0.52	800	57.4	49.4^{+}	50.0^{+}	53.2^{+}	53.6^{+}	24.2	23.5	23.4^{*}	26.5^{+}	26.4^{+}	18.1	18.5^{*}	18.4	18.6^{+}	18.5^{+}

[1] Probability of aneuploidy-induced inviability.

[2] Number of preassigned BDM incompatibility pairs.

[3] Probability of spore death caused by one pair of incompatibility.

[4] Total number of genotyped spores.

[5] Odds ratio.

[6] Chi-squared statistic.

[7] *G*-test statistic.

[8] Chi-squared statistic only when $OR > 1$.

[9] *G*-test statistic only when $OR > 1$.

Table 4-2. Odds ratio does not outperform other statistics in detecting symmetrical genetic incompatibilities. The results are from 400 simulations for each parameter set, with * showing $P < 0.05$ and $^+$ showing $P < 0.005$ when comparing the performance of a statistic with that of odds ratio using a paired *t*-test.

Parameters				Sensitivity (%)					False discovery rate (%)					Genomic distance (kb)				
G[1]	N[2]	I[3]	M[4]	OR[5]	χ^2[6]	G[7]	χ^2*[8]	$G*$[9]	OR	χ^2	G	χ^2*	$G*$	OR	χ^2	G	χ^2*	$G*$
0	4	0.81	100	88.1	88.0	87.9	88.1	88.0	21.0	18.4$^+$	18.5$^+$	21.5	21.2	14.9	14.8	14.8	14.8	14.8
0	4	0.81	200	90.3	90.2	90.2	90.2	90.2	24.5	21.6$^+$	21.8$^+$	24.4	24.4	11.7	11.7	11.7	11.7	11.7
0	4	0.81	400	91.2	91.2	91.2	91.2	91.2	26.3	24.3$^+$	24.3$^+$	26.2	26.3	10.0	10.0	10.0	10.0	10.0
0	4	0.81	800	90.1	90.1	90.1	90.1	90.1	29.0	27.1$^+$	27.1$^+$	29.0	29.0	10.1	10.1*	10.1*	10.1*	10.1*
0	6	0.59	100	31.3	28.2$^+$	28.1$^+$	32.2$^+$	32.2$^+$	32.3	27.4$^+$	27.1$^+$	32.3	32.5	21.4	20.9	20.9	21.2	21.2
0	6	0.59	200	76.6	74.5$^+$	74.6$^+$	76.5	76.5	21.0	18.3$^+$	18.3$^+$	20.9	20.9	16.9	17.0	17.0	17.0	17.0
0	6	0.59	400	89.5	89.6	89.7*	89.5	89.6	16.5	13.8$^+$	13.7$^+$	16.6	16.6	12.8	12.8	12.8	12.8	12.8
0	6	0.59	800	91.1	91.1	91.1	91.1	91.1	19.2	16.0$^+$	16.0$^+$	19.2	19.2	10.8	10.8	10.8	10.8	10.8
0	8	0.46	100	6.2	5.0$^+$	4.9$^+$	6.6*	6.3	50.3	45.6*	46.4	47.9*	48.3*	22.7	22.7	22.9	22.8	22.6
0	8	0.46	200	32.4	28.9$^+$	28.8$^+$	32.7	32.6	31.8	29.2$^+$	29.3$^+$	31.5	31.5	20.8	20.9	20.9	20.8	20.8
0	8	0.46	400	76.8	74.7$^+$	74.8$^+$	76.7	76.7	21.4	18.6$^+$	18.6$^+$	21.0*	21.1*	16.8	16.8	16.8	16.8	16.8
0	8	0.46	800	89.3	89.4*	89.4*	89.3	89.3	15.6	13.6$^+$	13.7$^+$	15.5	15.5	13.2	13.2	13.2	13.2	13.2
0.5	3	0.74	100	81.1	80.2*	80.2	81.8	81.8	22.2	18.5$^+$	18.5$^+$	22.0	22.3	16.4	16.4	16.3	16.4	16.4
0.5	3	0.74	200	89.8	90.0	90.0	90.0	90.0	17.9	14.8$^+$	14.7$^+$	17.8	17.7	13.1	13.0	13.0	13.0	13.0
0.5	3	0.74	400	90.5	90.5	90.5	90.5	90.5	21.4	18.3$^+$	18.2$^+$	21.4	21.5	10.7	10.7	10.7	10.7	10.7
0.5	3	0.74	800	89.8	89.8	89.8	89.8	89.8	22.9	20.4$^+$	20.4$^+$	22.9	22.9	10.1	10.1	10.1	10.1	10.1
0.5	4	0.59	100	28.7	26.1$^+$	26.0$^+$	29.2	29.0	35.6	30.7*	30.6*	35.0	35.1	20.1	19.9	20.0	20.0	20.0
0.5	4	0.59	200	78.6	76.7$^+$	76.6$^+$	78.6	78.7	21.8	19.3$^+$	19.1$^+$	22.1	22.0	17.0	17.1	17.1	17.0	17.0
0.5	4	0.59	400	90.0	90.1	90.1	90.0	90.0	16.0	14.3*	14.2$^+$	16.0	16.1	13.0	12.9	12.9	12.9*	12.9*
0.5	4	0.59	800	91.1	91.2	91.2	91.1	91.1	16.1	12.8$^+$	12.8$^+$	16.1	16.1	10.8	10.8	10.8	10.8	10.8
0.5	6	0.41	100	3.3	2.5$^+$	2.5$^+$	3.4	3.5	61.4	56.1	57.4	61.3	60.6	22.4	21.7	21.8	22.5	22.2
0.5	6	0.41	200	17.5	14.5$^+$	14.5$^+$	17.8*	17.7	38.7	35.3	35.3	38.4	38.3	20.9	20.2	20.2	21.0	21.0
0.5	6	0.41	400	61.0	57.6$^+$	57.7$^+$	61.3	61.4*	24.0	21.6$^+$	21.6$^+$	23.7	23.6	18.5	18.5	18.5	18.6	18.6
0.5	6	0.41	800	87.4	87.3	87.3	87.4	87.4	16.8	14.7$^+$	14.7$^+$	16.7	16.7	14.2	14.2	14.2	14.2	14.2

[1] Probability of aneuploidy-induced inviability.
[2] Number of preassigned BDM incompatibility pairs.
[3] Probability of spore death caused by one pair of incompatibility.
[4] Total number of genotyped spores.
[5] Odds ratio.
[6] Chi-squared statistic.
[7] *G*-test statistic.
[8] Chi-squared statistic only when $OR > 1$.
[9] *G*-test statistic only when $OR > 1$.

CHAPTER 5

CONCLUSIONS

5.1 Summary

In the past decade, research work has been gradually shifting its focus from elucidating functionality of individual mutation/gene and by changing one site or knocking out a gene at a time to studying their interactions, i.e. intragenic and intergenic epistasis, by quantifying their individual and collective effects. Intragenic epistasis is shown to restrict evolutionary paths for the accumulation of several consecutive beneficial mutations in multiple cases (Weinreich, Delaney et al. 2006, Bridgham, Ortlund et al. 2009, Salverda, Dellus et al. 2011, Toprak, Veres et al. 2011), while opening up seemingly impossible paths or increasing the accessibility of fitness peaks in other cases (Bloom and Arnold 2009, Abed, Pizzorno et al. 2011, Palmer, Toprak et al. 2015). For intergenic epistasis, many cases of non-additive effects have been recorded between genes encoding multi-component proteins, components of a metabolic network, genes in the same biochemical, developmental and signaling pathway, etc. However, such intergenic interactions can also occur between seemingly unrelated genes (He, Qian et al. 2010). Such endeavors in quantifying epistasis reflect the increasing interest in studying the complex interactions in the biological system, which is having a significant impact on scientists' viewpoint for biological functions, from a more isolated standpoint to a more panoramic picture of interactions and networks.

Although many case studies of epistasis have been reported, the general trend of epistasis is still unclear. To name a few questions, is intragenic epistasis more likely to be positive or negative in general? How similar is epistasis across changing environments and how predictable is epistasis across environments? What fraction of reproductive isolation is caused by genetic incompatibility, a type of intergenic epistasis? Characterizing inter- and intragenic epistasis at a large scale to gain a more holistic understanding will surely offer in-depth and broad insights into the abovementioned questions and help the community to better understand evolutionary processes, such as speciation, and repeatability of evolution, etc. The past few years has witnessed a series of studies in studying epistasis at large scale (Olson, Wu et al. 2014, Costanzo, VanderSluis et al. 2016, Puchta, Cseke et al. 2016, Skwark, Croucher et al. 2017).

My research work is built on these previous research work and other relevant studies and aims at directly answering or facilitating future answering of the aforementioned questions. My focus lies in characterizing epistasis at a large scale at both intragenic and intergenic level to shed light on possible evolutionary trajectories and the speciation process in the evolutionary history. I study the overall distribution of intragenic epistasis (Chapter 2), its interaction with environments (Chapter 3), and the optimized strategy to genome-widely identify genetic incompatibility, a type of strong intergenic epistasis (Chapter 4).

My research provides effective strategies for characterizing epistasis. For example, in Chapter 2, I combine mass competition with high-throughput barcode sequencing to characterizing fitness for over 65,000 mutants simultaneously, allowing for quantifying epistasis for over half of all possible 21,115 mutation pairs. In Chapter 3, I build a model for predicting fitness across environments and suggest a simple and accurate way to infer fitness landscape and epistasis in a new environment. In Chapter 4, I come up with better test statistic and a more effective methodology to detect signals for incompatible gene pairs. These strategies are expected to be used by more research groups in the near future to generate more valuable datasets to gain an in-depth understanding of epistasis and its underlying biological mechanisms.

Many conclusions of my dissertation work have been further confirmed by other studies. In Chapter 2, I reveal the overall enrichment of negative intragenic epistasis. This finding and many of my other main conclusions are echoed by a co-published paper on *Science* (Puchta, Cseke et al. 2016), and also confirmed by multiple other recent studies afterwards (Sarkisyan, Bolotin et al. 2016, Hopf, Ingraham et al. 2017). My research also generated high-quality fitness measurements, providing a valuable large-scale dataset for testing a series of hypotheses and evolutionary theories (Hopf, Ingraham et al. 2017). Moreover, in Chapter 4, my strategy for better identifying BDM incompatibility has later been successfully implemented by another research group (Duncan Greig, poster at 2015 SMBE conference).

All of my studies emphasize the importance of computer simulation before conducting the actual large-scale survey. In Chapter 2 and 3, careful analysis is conducted before experiments to ensure that the fraction of double mutants in the synthesized variants is maximized, enough colonies are being collected, and the optimal competition durations are chosen for sequencing to maximize the power of experiments to draw meaningful conclusions. In Chapter 4, I evaluate the power of previous studies and highlight the importance of choosing the adequate sample size and testing strategies using computer simulations.

5.2 Implications

My findings have multiple implications in the field of evolutionary genetics. First, my results illustrate the importance of understanding epistasis to gain an in-depth understanding of molecular evolution by revealing a general negative trend for intragenic epistasis. There are many evolutionary theories that depend on the overall pattern of negative epistasis. For instance, previous theoretical work has predicted that negative epistasis is favored if the selection is efficient compared to drift (Gros, Le Nagard et al. 2009). Such negative epistasis is associated with a higher level of tolerance to mutations, i.e. robustness, and is therefore selected by natural selection. Moreover, negative epistasis enhances the ability of natural selection to remove deleterious mutations in sexual populations compared asexual populations. Our observation further confirms the maintenance of sexual reproduction despite its high costs (Lucchesi 1978), consistent with previous simulation results (Azevedo, Lohaus et al. 2006). Also, the hypothesis of reduction in mutational load by truncation selection against deleterious mutations also relies on the overall negative epistasis (Crow and Kimura 1979).

Second, my research emphasizes the importance of understanding the structural basis of a gene in order to understand its fitness landscape. For instance, in Chapter 2, I aimed at revealing the link between tRNA folding and tRNA fitness. No direct evidence at a large scale was previously available between folding stability and fitness for RNA genes. Indeed, when I focus on N1, N2 and N3 mutants separately, the folding stability of each tRNA variant represented by its minimum free energy predicted by Vienna RNA package shows weak or no significant correlation with the fitness of the strain carrying this tRNA variant, unless we focus on all variants together. In the latter scenario, the seemingly strong correlation between the structural stability and the fitness value might not indicate a causal relationship but instead is likely to be an artefact of both measures being strongly correlated with the number of mutations in the molecule. That is, a higher number of mutation would in general lead to further destruction of structure and further reduction in fitness, but there might be no causal relationship between the two metrics. Such pattern was also observed in other studies (Puchta, Cseke et al. 2016). The lack of correlation is somewhat expected because tRNAs are almost always folded stably with a strong secondary structure and comparatively low minimum free energy. To better understand how tRNA folding affects fitness, I come up with a model to calculate the fraction of functional molecules based on the relative stability of tRNAs folding in multiple functional or nonfunctional structures, which turned out to be highly overall correlated with fitness (Spearman's rho=0.51 after excluding internal promoter sites), and significant correlations for N1, N2 and N3 mutants respectively. These results highlight the importance of understanding the structural basis of the candidate molecule to gain in-depth knowledge.

Moreover, my research has potential implications for synthetic biology. When designing a functional molecule, if the design is done in a piecewise fashion without taking epistasis into consideration, the final outcome is highly likely to be different from the expectation assuming no interdependence. The pervasiveness of epistasis has been observed in previous case studies (Williams and Lovell 2009, Lunzer, Golding et al. 2010), and also confirmed at a large scale in Chapter 2. In Chapter 3, a simple piecewise linear model is built to predict fitness and epistasis across multiple environments, which could be helpful to predict the performance of molecular machinery across multiple environments without extensive and laborious re-measurements.

5.3 Future directions

My research covers several interesting topics on intra- and inter-genic epistasis. There are many more vital questions to be addressed in the near future.

It is of great importance to reveal biological explanations of the patterns of the fitness landscape. For the fitness landscape of the tRNA gene, I use the fraction of functional molecules to partially explain the fitness differences. Meanwhile, other factors, such as the gene expression level, secondary and 3-dimentional structure and other sequence features to attach amino acids, bind ribosome or recognize codon could be measured to build a full model for predicting fitness. Such efforts would also be vital for other studies measuring fitness landscapes at a large scale.

For intragenic epistasis, although an enrichment of negative epistasis is observed by multiple large-scale studies for both beneficial and deleterious mutations, the underlying mechanism for this phenomenon is still unclear. An increasing number of large-scale fitness landscape data sets are coming out, and it would be interesting and important to understand the phenomenon from a structural, expression-related, catalytic, population genetic or modular aspect.

For intergenic epistasis, the exact casual gene conferring reproductive isolation during incipient speciation still need to be mapped at the genome scale even after the segments of the genomic region containing these genes could be identified. Given the recent application of CRISPR-Cas9 in fine mapping (Sadhu, Bloom et al. 2016), the identification of such candidate genes is expected in the near future. Because of the tremendous functional genomic data available in *S. cerevisiae* and its close relatives, it can serve as an excellent model eukaryote for studying the genetic basis of the speciation process.

My research focuses mainly on two-way interactions between genes and mutations. However, there could be high-order epistasis playing important roles (Weinreich, Lan et al. 2013). Some of our methodologies can also be readily applied for studying high-order epistasis. For instance, by adjusting the fraction of the mutant base in the synthesized gene in Chapter 2 and increasing the throughput of sequencing, analysis can be done on a large scale for quantifying high order interactions. For intergenic epistasis, because high-order incompatibility with complete penetrance is effectively equivalent to three pairs of two-locus incompatibility with incomplete penetrance, our results in Chapter 4 can be directly applied to identify high-order incompatibility. More algorithms (Guo, Meng et al. 2014) and large-scale datasets would be beneficial in identifying such high-order epistasis.

Another important aspect is to study epistasis of sites in different genes. Most of the previous studies on intergenic epistasis focus on deletion effects, while studies of intragenic epistasis focus exclusively on a single gene. Studying epistasis of interacting interfaces and other functionally related genes would provide valuable information on the co-evolution of these genes.

Finally, it is always important to validate whether the conclusions obtained in one unicellular organism are directly applicable to other organisms and whether the patterns observed in one protein or RNA molecule are broadly true in other functional molecules. More extensive case studies for fitness landscapes and incompatibility mapping is highly essential before drawing a general conclusion for all molecules or organisms.

5.4 References

Abed, Y., A. Pizzorno, X. Bouhy and G. Boivin (2011). "Role of Permissive Neuraminidase Mutations in Influenza A/Brisbane/59/2007-like (H1N1) Viruses." Plos Pathogens **7**(12).

Azevedo, R. B., R. Lohaus, S. Srinivasan, K. K. Dang and C. L. Burch (2006). "Sexual reproduction selects for robustness and negative epistasis in artificial gene networks." Nature **440**(7080): 87-90.

Bloom, J. D. and F. H. Arnold (2009). "In the light of directed evolution: Pathways of adaptive protein evolution." Proceedings of the National Academy of Sciences of the United States of America **106**: 9995-10000.

Bridgham, J. T., E. A. Ortlund and J. W. Thornton (2009). "An epistatic ratchet constrains the direction of glucocorticoid receptor evolution." Nature **461**(7263): 515-519.

Costanzo, M., B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons, G. Tan, W. Wang, M. Usaj, J. Hanchard, S. D. Lee, V. Pelechano, E. B. Styles, M. Billmann, J. van Leeuwen, N. van Dyk, Z. Y. Lin, E. Kuzmin, J. Nelson, J. S. Piotrowski, T. Srikumar, S. Bahr, Y. Chen, R. Deshpande, C. F. Kurat, S. C. Li, Z. Li, M. M. Usaj, H. Okada, N. Pascoe, B. J. San Luis, S. Sharifpoor, E. Shuteriqi, S. W. Simpkins, J. Snider, H. G. Suresh, Y. Tan, H. Zhu, N. Malod-Dognin, V. Janjic, N. Przulj, O. G. Troyanskaya, I. Stagljar, T. Xia, Y. Ohya, A. C. Gingras, B. Raught, M. Boutros, L. M. Steinmetz, C. L. Moore, A. P. Rosebrock, A. A. Caudy, C. L. Myers, B. Andrews and C. Boone (2016). "A global genetic interaction network maps a wiring diagram of cellular function." Science **353**(6306).

Crow, J. F. and M. Kimura (1979). "Efficiency of truncation selection." Proc Natl Acad Sci U S A **76**(1): 396-399.

Gros, P. A., H. Le Nagard and O. Tenaillon (2009). "The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation." Genetics **182**(1): 277-293.

Guo, X., Y. Meng, N. Yu and Y. Pan (2014). "Cloud computing for detecting high-order genome-wide epistatic interaction via dynamic clustering." Bmc Bioinformatics **15**.

He, X., W. Qian, Z. Wang, Y. Li and J. Zhang (2010). "Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks." Nat Genet **42**(3): 272-276.

Hopf, T. A., J. B. Ingraham, F. J. Poelwijk, C. P. Scharfe, M. Springer, C. Sander and D. S. Marks (2017). "Mutation effects predicted from sequence co-variation." Nat Biotechnol **35**(2):

128-135.

Lucchesi, J. C. (1978). "Gene dosage compensation and the evolution of sex chromosomes." Science **202**(4369): 711-716.

Lunzer, M., G. B. Golding and A. M. Dean (2010). "Pervasive cryptic epistasis in molecular evolution." PLoS Genet **6**(10): e1001162.

Olson, C. A., N. C. Wu and R. Sun (2014). "A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain." Curr Biol **24**(22): 2643-2651.

Palmer, A. C., E. Toprak, M. Baym, S. Kim, A. Veres, S. Bershtein and R. Kishony (2015). "Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes." Nat Commun **6**: 7385.

Puchta, O., B. Cseke, H. Czaja, D. Tollervey, G. Sanguinetti and G. Kudla (2016). "Network of epistatic interactions within a yeast snoRNA." Science **352**(6287): 840-844.

Sadhu, M. J., J. S. Bloom, L. Day and L. Kruglyak (2016). "CRISPR-directed mitotic recombination enables genetic mapping without crosses." Science **352**(6289): 1113-1116.

Salverda, M. L., E. Dellus, F. A. Gorter, A. J. Debets, J. van der Oost, R. F. Hoekstra, D. S. Tawfik and J. A. de Visser (2011). "Initial mutations direct alternative pathways of protein evolution." PLoS Genet **7**(3): e1001321.

Sarkisyan, K. S., D. A. Bolotin, M. V. Meer, D. R. Usmanova, A. S. Mishin, G. V. Sharonov, D. N. Ivankov, N. G. Bozhanova, M. S. Baranov, O. Soylemez, N. S. Bogatyreva, P. K. Vlasov, E. S. Egorov, M. D. Logacheva, A. S. Kondrashov, D. M. Chudakov, E. V. Putintseva, I. Z. Mamedov, D. S. Tawfik, K. A. Lukyanov and F. A. Kondrashov (2016). "Local fitness landscape of the green fluorescent protein." Nature **533**(7603): 397-+.

Skwark, M. J., N. J. Croucher, S. Puranen, C. Chewapreecha, M. Pesonen, Y. Y. Xu, P. Turner, S. R. Harris, S. B. Beres, J. M. Musser, J. Parkhill, S. D. Bentley, E. Aurell and J. Corander (2017). "Interacting networks of resistance, virulence and core machinery genes identified by genome-wide epistasis analysis." PLoS Genet **13**(2): e1006508.

Toprak, E., A. Veres, J. B. Michel, R. Chait, D. L. Hartl and R. Kishony (2011). "Evolutionary paths to antibiotic resistance under dynamically sustained drug selection." Nat Genet **44**(1): 101-105.

Weinreich, D. M., N. F. Delaney, M. A. DePristo and D. L. Hartl (2006). "Darwinian evolution can follow only very few mutational paths to fitter proteins." Science **312**(5770): 111-114.

Weinreich, D. M., Y. H. Lan, C. S. Wylie and R. B. Heckendorn (2013). "Should evolutionary geneticists worry about higher-order epistasis?" Current Opinion in Genetics & Development **23**(6): 700-707.

Williams, S. G. and S. C. Lovell (2009). "The effect of sequence evolution on protein structural divergence." Mol Biol Evol **26**(5): 1055-1065.

APPENDIX

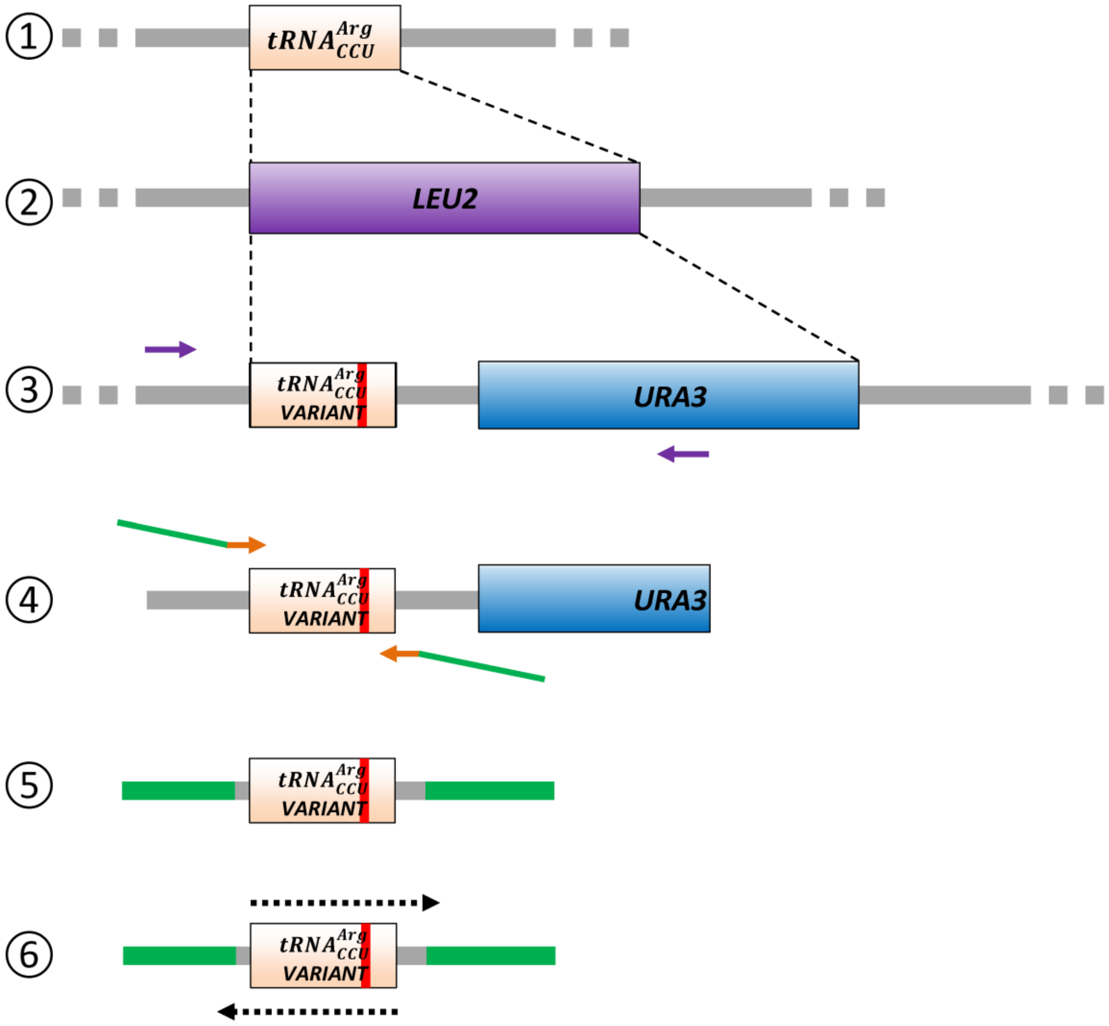


Figure S1. Schematics of experimental procedures.

① The wild-type $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene at its native genomic position.

② The wild-type $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene is replaced with *LEU2*. The dotted black lines show the region replaced. This is the strain referred to as the $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene deletion strain in the paper.

③ *LEU2* is replaced with a tRNA gene cassette composed of a $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ gene variant and *URA3*.

④ Genomic region amplified by the first round of PCR with the purple primer pair shown in ③. The purple primer pair shown in ③ only amplifies tRNA gene cassettes that are located at the correct genomic position.

⑤ tRNA gene variant amplified from ④ using the orange primer pair shown in ④. Adapters for Illumina sequencing are shown by green lines and are part of the primers.

⑥ The tRNA gene variant is sequenced using 100-nucleotide paired-end Illumina sequencing, indicated by black dotted arrows.

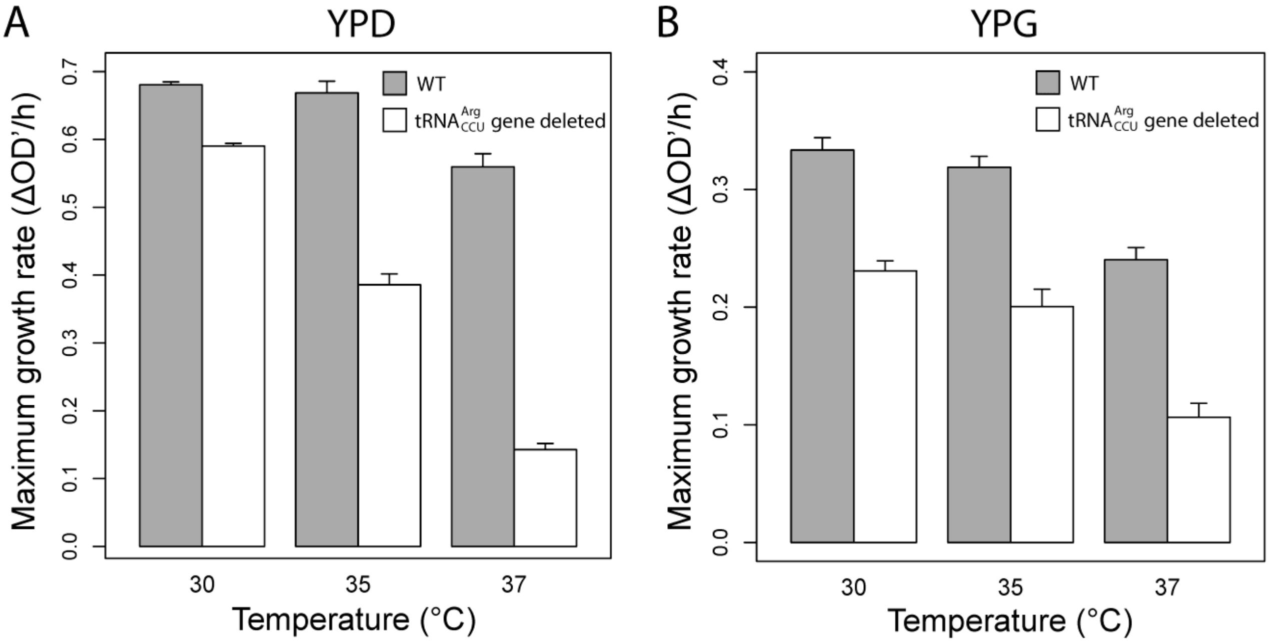


Figure S2. Maximum growth rates of the yeast wild-type strain (gray bars) and $tRNA^{Arg}_{CCU}$ gene deletion strain (white bars) in two media at three temperatures. (**A**) Mitotic growth rates in the fermentable medium YPD. (**B**) Mitotic growth rates in the non-fermentable medium YPG. Growth rates are measured by the maximum increase in OD' per hour in mid-log phase. OD', converted from optical density (OD) at 600 nm by the formula OD+0.8324×OD^3, is approximately proportional to cell density.

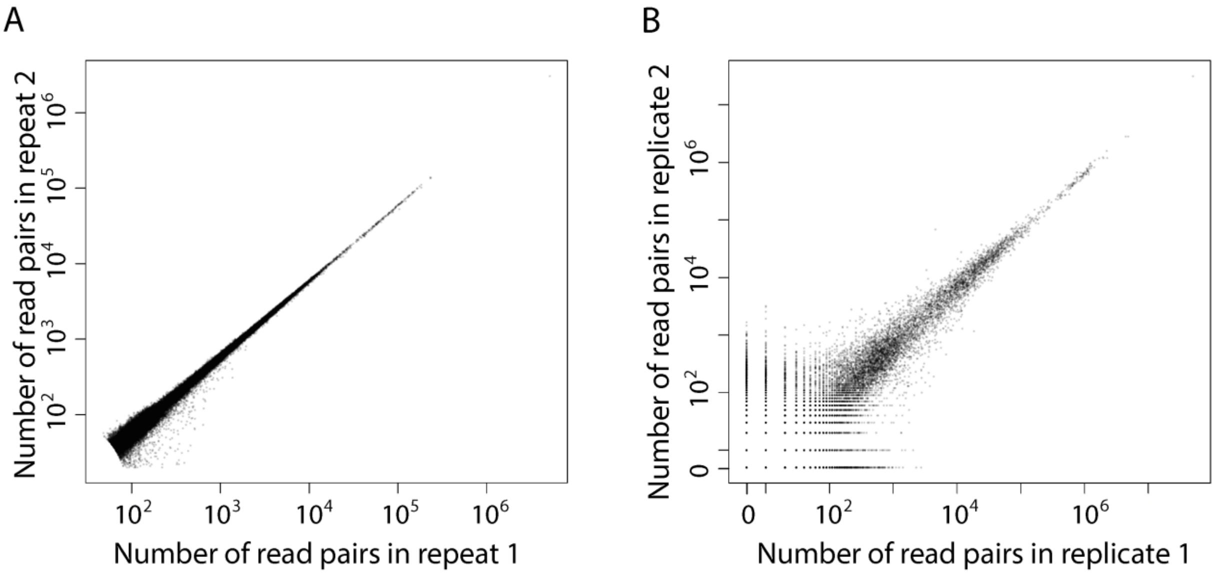


Figure S3. Numbers of read pairs across genotypes are highly correlated between technical repeats and between biological replicates. (**A**) Comparison in read pair number between two technical repeats at T_0 across genotypes. Each dot represents a genotype and only those genotypes with a total of ≥100 read pairs are considered. Pearson's correlation coefficient $r = 0.99997$. (**B**) Comparison in read pair number between biological replicates 1 and 2 at T_{24} across genotypes. Each dot represents a genotype and only those genotypes with a total of ≥100 read pairs at T_0 are considered. Pearson's correlation coefficient $r = 0.9997$. The mean $r = 0.9987$ for the 15 pairs of biological replicates.

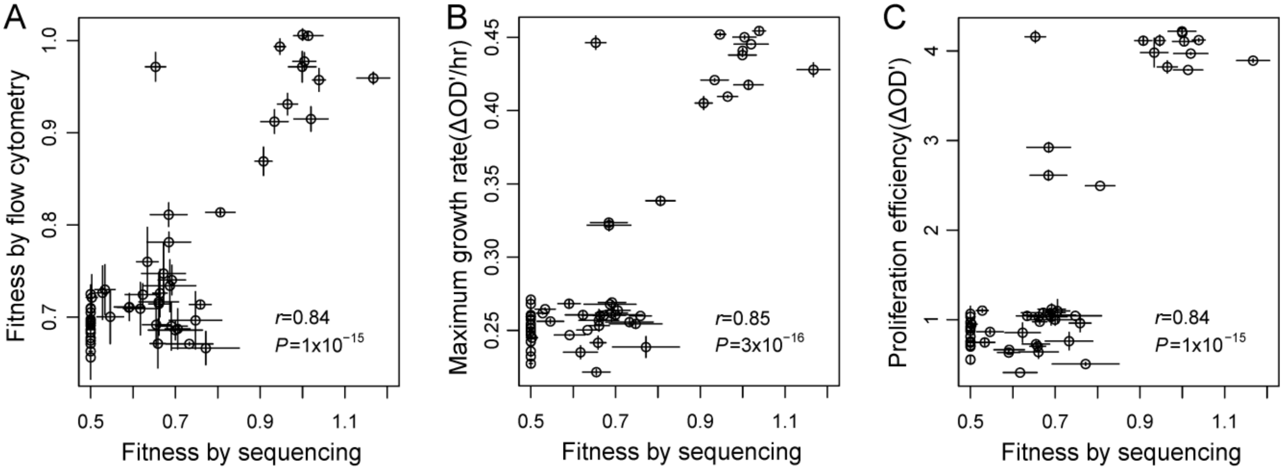


Figure S4. Comparison of three fitness estimation methods for 55 strains carrying different tRNA variants. For the sequencing method, fitness values shown are per generation fitness relative to the wild-type. For the flow cytometry method, fitness is measured by pairwise competition followed by flow cytometry, and the values shown are per generation fitness relative to the wild-type. For the growth curve method, fitness is measured by either the maximum growth rate in mid-log phase (ΔOD'/hr) or proliferation efficiency (OD' change in the first 48 hours of growth). OD', converted from optical density (OD) by the formula $OD+0.8324 \times OD^3$, is approximately proportional to cell density. (**A**) Fitness estimated by sequencing is correlated with that estimated by flow cytometry. (**B**) Fitness estimated by sequencing is correlated with that estimated by maximum growth rate. (**C**) Fitness estimated by sequencing is correlated with that estimated by proliferation efficiency. Error bars show one standard error.

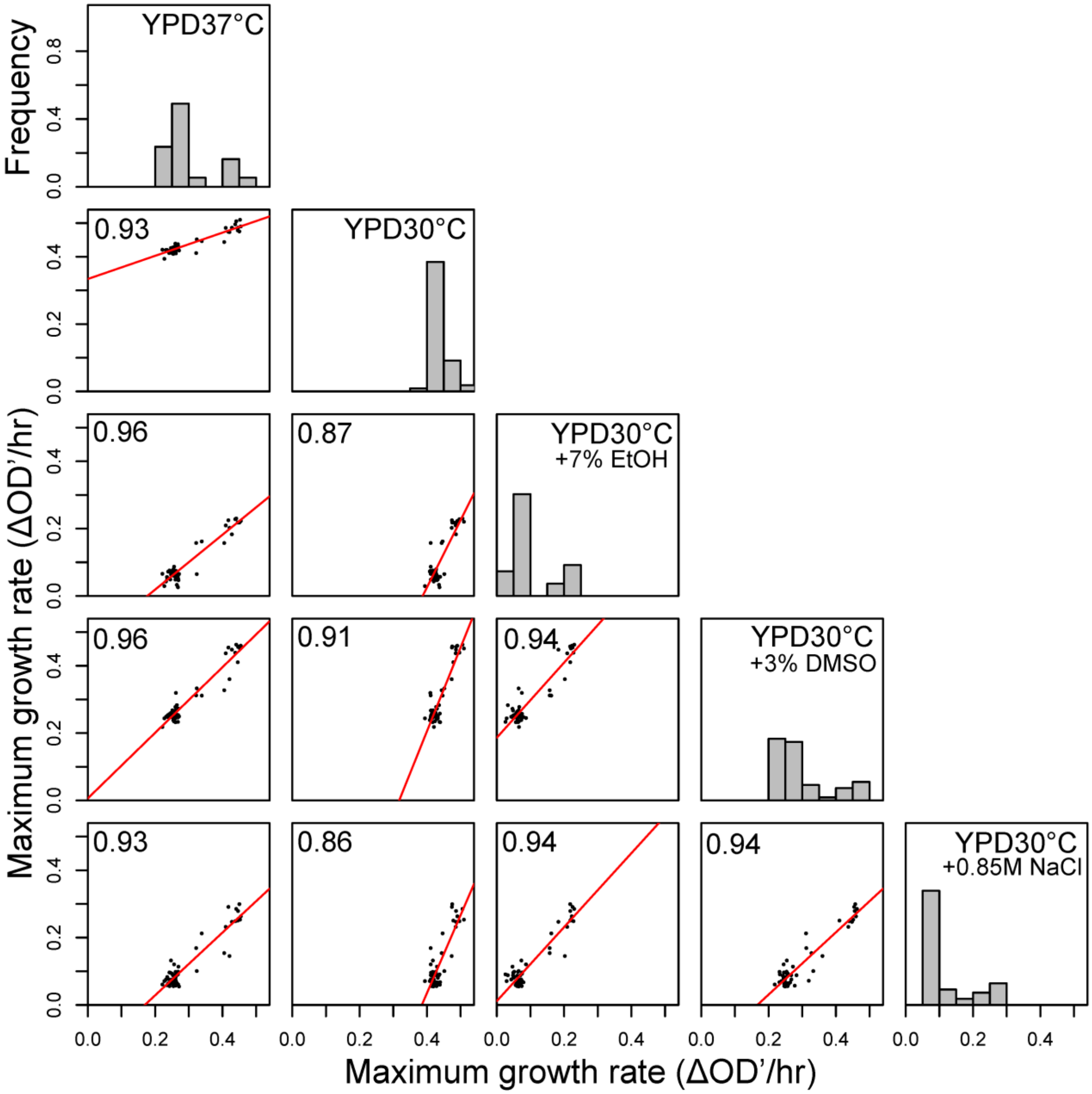


Figure S5. Correlations of maximum growth rates of 55 strains carrying different tRNA variants among five different environments. Environments used and the frequency distribution of maximum growth rates in the environments are shown in the diagonal panels. In lower left panels are maximum growth rates in each environment plotted against those in each of the other four environments, along with Pearson's correlation coefficients and red linear regression lines.

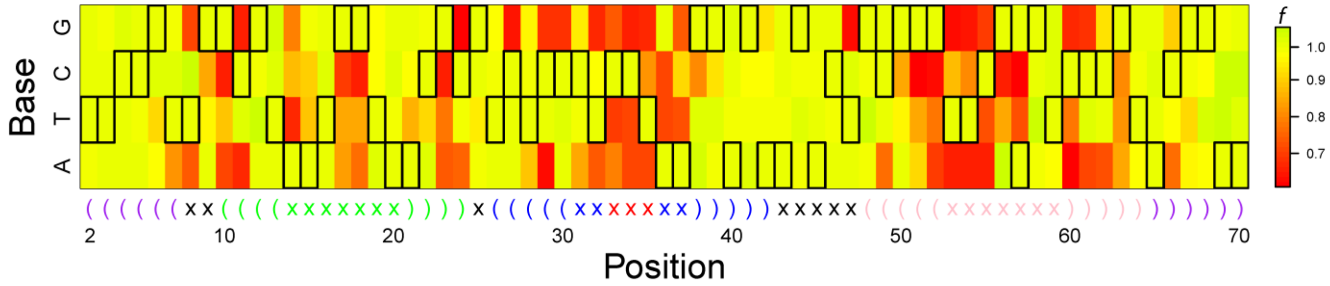


Figure S6. Heat map showing the fitness of all N1 mutants. At each site, the wild-type nucleotide is boxed. The tRNA secondary structure is plotted linearly with parentheses showing sites in stems and crosses showing sites in loops; the same color is used for sites in the same loop or stem. The anticodon is shown by the three red crosses.

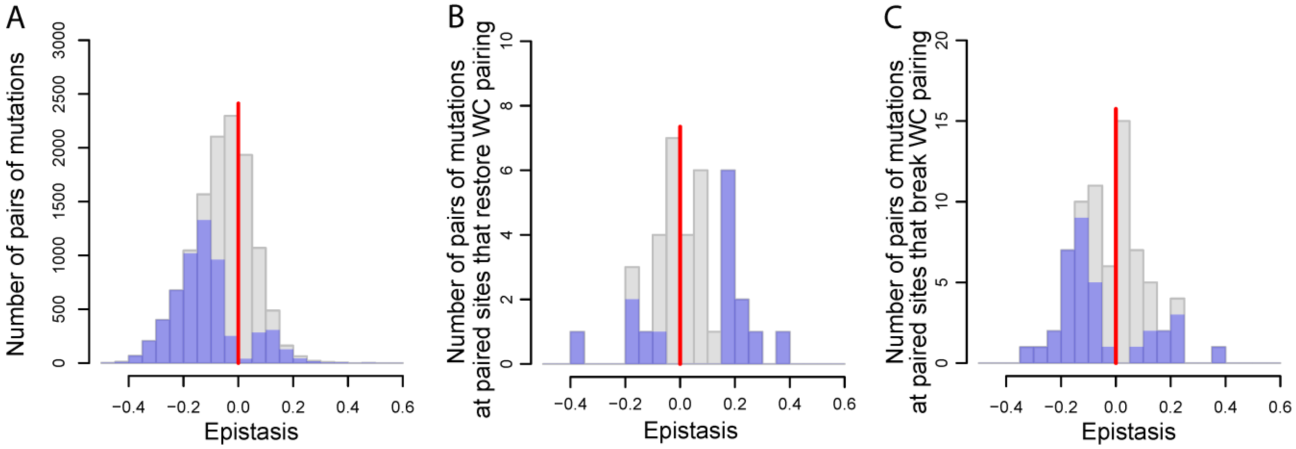


Figure S7. Distribution of epistasis between mutations after the removal of all 854 cases with observed or expected fitness reaching the lower limit of 0.5. (**A**) Frequency distributions of pairwise epistasis (gray) and statistically significant pairwise epistasis (blue) among 12,475 pairs of point mutations studied. (**B**) Frequency distributions of epistasis (gray) and statistically significant epistasis (blue) between pairs of mutations that convert a Watson-Crick (WC) base pair to another WC pair. (**C**) Frequency distributions of epistasis (gray) and statistically significant epistasis (blue) between pairs of mutations that break a WC pair. The vertical red line shows $\varepsilon = 0$. Median ε is significantly greater in (B) and (C) than in (A) when all epistasis cases (P = 3×10^{-5} and 0.01, respectively; Mann-Whitney U test) or only significant epistasis cases (P = 3×10^{-4} and 0.02, respectively) are considered. Median ε is significantly greater in (C) than in (B) when all epistasis cases (P = 0.049) or only significant epistasis cases (P = 0.043) are considered.

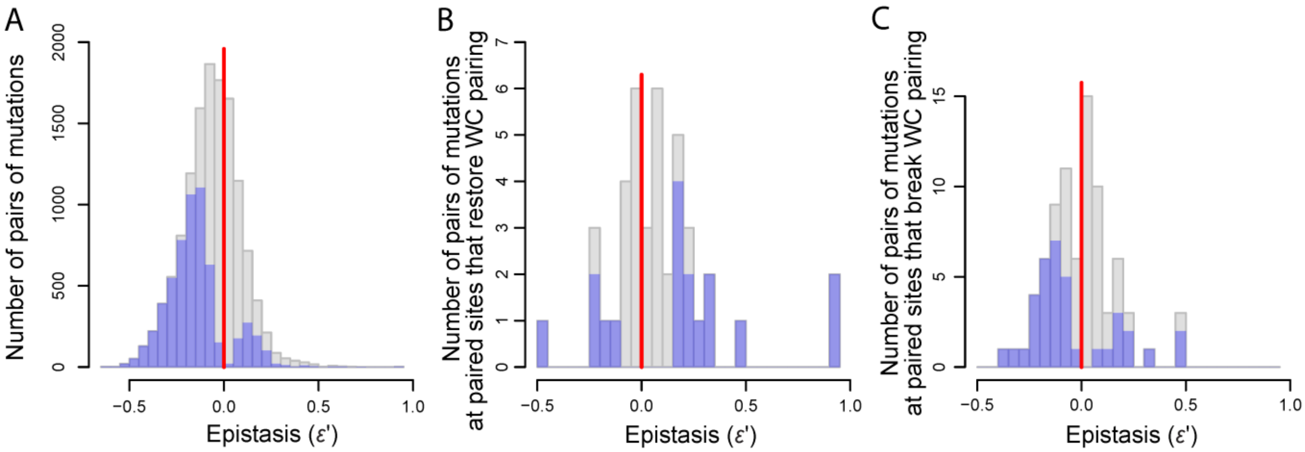


Figure S8. Distribution of pairwise epistasis $\varepsilon' = \ln(f_{AB}) - \ln f_A - \ln f_B$. (**A**) Frequency distributions of pairwise epistasis (gray) and statistically significant pairwise epistasis (blue) among all 12,985 pairs of point mutations studied. (**B**) Frequency distributions of epistasis (gray) and statistically significant epistasis (blue) between pairs of mutations that convert a Watson-Crick (WC) base pair to another WC pair. (**C**) Frequency distributions of epistasis (gray) and statistically significant epistasis (blue) between pairs of mutations that break a WC pair.

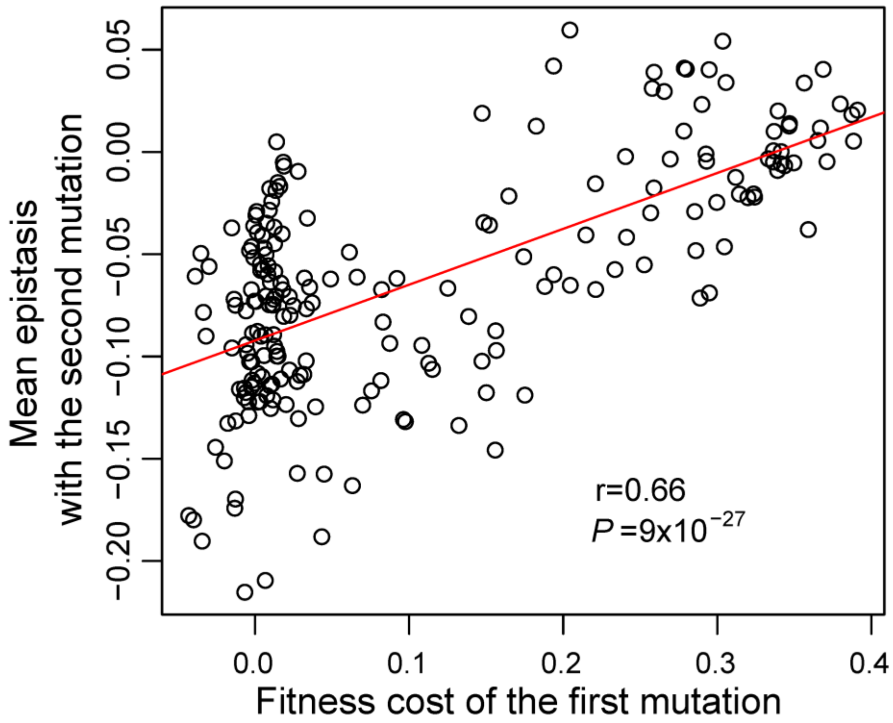


Figure S9. Correlation between the fitness cost of the first mutation and the mean epistasis with the second mutation, after the removal of N2 mutants whose expected or observed fitness is ≤ 0.5. Red line shows the linear regression.

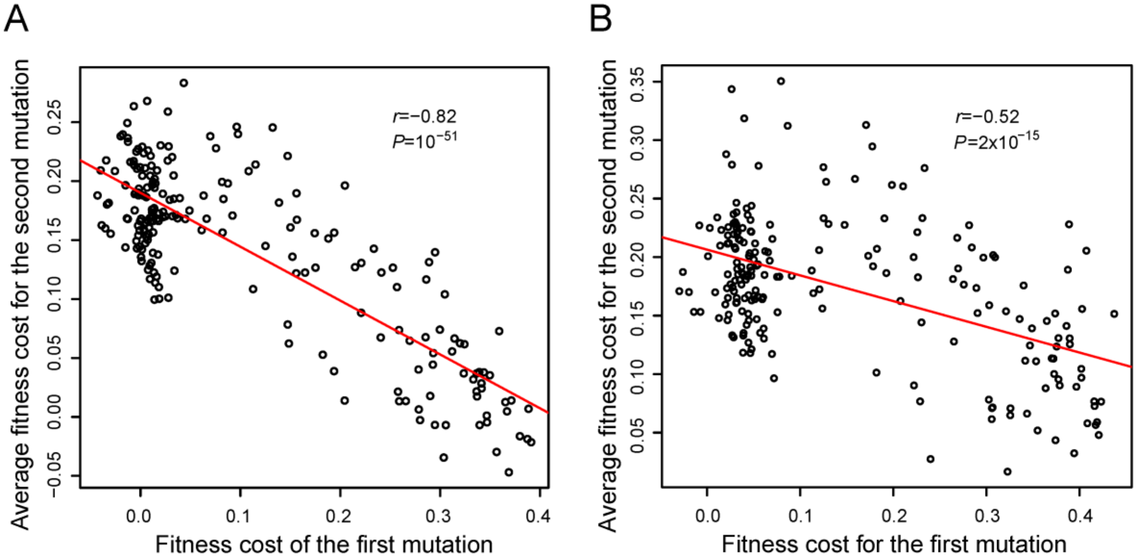


Figure S10. Negative correlation between the fitness cost of the first mutation and the mean fitness cost of the second mutation is not artifactual. (**A**) Correlation between the fitness cost of the first mutation and the mean fitness cost of the second mutation, after the removal of all N2 mutants whose expected or observed fitness is ≤ 0.5. Red line shows the linear regression. The fitness cost of the first mutation is calculated by 1-f_A, where f_A is the fitness of a N1 mutant carrying the A mutation. The mean fitness cost of the second mutation, given the first mutation A, is calculated by 1- (mean f_{AB})/f_A, where the subscript AB refers to a genotype carrying the A mutation as well as another mutation (including the reversion of the A mutation) and f_{AB} is the average fitness of all such genotypes. (**B**) Correlation between the fitness cost of the first mutation and the mean fitness cost of the second mutation, where the x-axis and y-axis are based on fitness data from different biological replicates. Because measurement errors of f_A could lead to an artefactual negative correlation between the estimated fitness costs of the first and second mutations, we used three biological replicates to estimate f_A for the x-axis and used the other three biological replicates to estimate f_{AB} and f_A for computing the y-axis value, thus removing such potential artifacts. All 20 possible combinations of such sampling were used to calculate the correlation, which ranges from -0.51 to -0.76, with a mean of -0.61. The slope ranges from -0.397 to -0.220, with a mean of -0.293.

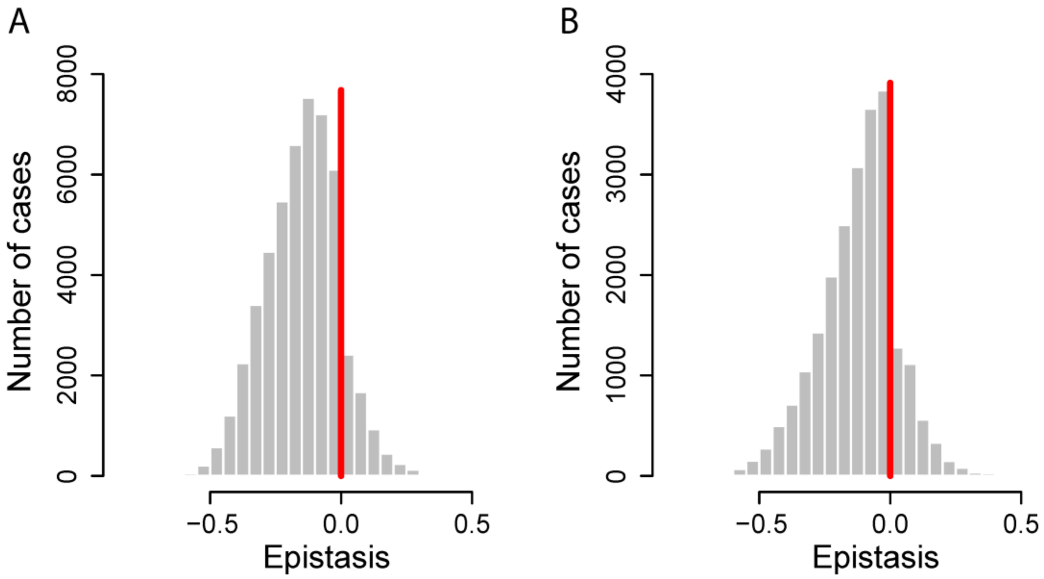


Figure S11. Frequency distributions of pairwise epistasis involving more than two sites. (**A**) Frequency distribution of epistasis between mutations AB and mutation C, defined as $\varepsilon = f_{ABC} - f_{AB}f_C$, is overall negatively biased. 6,234 cases with epistasis = 0 (both expected and observed fitness = 0.5) are not shown. (**B**) Epistasis between mutations AB and mutations CD, defined as $\varepsilon = f_{ABCD} - f_{AB}f_{CD}$, is overall negatively biased. 9,343 cases with epistasis = 0 (both expected and observed fitness = 0.5) are not shown. Here A, B, C, and D refer to four point mutations, relative to the wild-type. The red vertical lines show zero epistasis.

Table S1. Primers used

Purpose of the primers	Forward primer (5’-3’)	Reverse primer (5’-3’)
Replacing the endogenous $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene with *LEU2*	TCGTAATAATATTACTATGCAACTTAGGTACCTCATATTTCTTAGAGTTCAACCAAGTTGAAGAGTTCGAATCTCTTAGCAACCA	ATATGAACCTTCAACTAGTTATTACCACTGTGGCACTCTTTCTGCGGTAAGATTATCTCACTCCATCAAATGGTCAGGTCATTGA
Amplifying the chemically synthesized $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene variants	CGAAGTTTATTCATTCAATTTGAAGTGCTTCGTAATAATATTACTATGCAACTTAGGTACCTCATATTTCTTAGAGTTCAACCAAGTTGG	TCGCAAGGTAATATCGTCTGAATTTTTTCTATAAAGAAACGAAAAAAAAAAAAATAATCAACG
Amplifying *URA3*	TTGATTATTTTTTTTTTTTTCGTTTCTTTATAGAAAAAATTCAGACGATATTACCTTGCGAAGCTTTTCAATTCAATTCATCATTT	ATATAATATGAACCTTCAACTAGTTATTACCACTGTGGCACTCTTTCTGCGGTAAGATTATCTCAGGGTAATAACTGATATAATTAAATT
Fusing $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene variants with *URA3*	CGAAGTTTATTCATTCAATTTGAAG	ATATAATATGAACCTTCAACTAGTTA
First round of PCR for library preparation	GGGGTTCATTACAGCAGCTT	TGTGCTCCTTCCTTCGTTCT
Second round of PCR for library preparation*	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNAGTTCAACCAAGTTGG	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAAAAAAAAATAATCAACG

*The underlined sequence in the reverse primer at the second round of PCR corresponds to the index sequence for multiplex sequencing.

Table S2. Illumina read numbers from each sample

Time (hrs)	Sample description	Raw read number	Read pair number after filtering	Percentage used
0	Repeat 1	174,956,172	74,749,170	0.854
	Repeat 2	113,885,948	43,808,386	0.769
24	Replicate 1	45,821,636	19,957,042	0.871
	Replicate 2	29,756,408	12,921,113	0.868
	Replicate 3	51,889,890	22,273,492	0.858
	Replicate 4	119,335,654	53,881,470	0.903
	Replicate 5	67,091,400	29,939,362	0.892
	Replicate 6	82,746,144	37,263,697	0.901

Table S3. Fitness of mutants carrying anticodon mutations

Anticodon	Corresponding codon (by Watson-Crick pairing)	Corresponding amino acid	Fitness
CCC	GGG	G	0.82
ACT	AGT	S	0.80
AGT	ACT	T	0.76
GCT	AGC	S	0.74
CCA	TGG	W	0.71
CAT	ATG	M	0.71
CTT	AAG	K	0.71
TCT	AGA	R	0.69
CCG	CGG	R	0.68
CGT	ACG	T	0.65
TGT	ACA	T	0.65
AAT	ATT	I	0.62
CGC	GCG	A	0.61
CTC	GAG	E	0.61
GCC	GGC	G	0.60
TCC	GGA	G	0.58
TAT	ATA	I	0.57
GTT	AAC	N	0.56
CAC	GTG	V	0.52
ATT	AAT	N	0.50
GGT	ACC	T	0.50
ACC	GGT	G	0.50
GCA	TGC	C	0.50
CTG	CAG	Q	0.50
TCA	TGA	Stop	0.50
CTA	TAG	Stop	0.50

Table S4. Examples of pairwise epistasis whose sign depends on the genetic background

Mut C	Mut A	Mut B	f_C	f_A	f_B	f_{AC}	f_{BC}	f_{AB}	f_{ABC}	ε_{AB}|WT	ε_{AB}|C
A15T	G50A	C55G	0.88	1.00	0.70	0.60	0.51	0.71	0.67	-0.18	0.25
G18C	G12A	C29T	0.66	1.00	0.99	0.59	0.55	0.74	0.60	-0.25	0.16
A36C	C27T	G68C	0.70	1.00	0.97	0.54	0.59	0.76	0.60	-0.21	0.21
C33G	A15T	G50A	0.74	0.88	1.00	0.55	0.61	0.71	0.62	-0.18	0.23
C60G	G9C	G67A	0.66	0.85	0.95	0.53	0.58	0.58	0.57	-0.23	0.16
A15C	C5G	C22A	0.90	0.99	1.00	0.63	0.70	0.77	0.63	-0.23	0.16
C33G	A15T	C62G	0.74	0.88	0.92	0.55	0.60	0.65	0.60	-0.17	0.22
A36C	C30T	A65C	0.70	1.01	1.01	0.57	0.60	0.84	0.62	-0.19	0.19
A14C	T7C	G63C	0.87	1.01	0.81	0.68	0.54	0.61	0.57	-0.20	0.16
G48C	C49T	C55T	0.98	1.00	0.72	0.67	0.65	0.79	0.66	-0.18	0.18
C31A	A15T	C49G	0.85	0.88	0.99	0.59	0.64	0.67	0.58	-0.20	0.16
A15T	T16A	C27G	0.88	0.99	0.64	0.53	0.55	0.70	0.59	-0.17	0.17
C31G	G9C	C30T	0.94	0.85	1.01	0.63	0.66	0.79	0.63	-0.16	0.17
A14C	C4G	G48C	0.87	1.01	0.98	0.57	0.66	0.82	0.58	-0.16	0.17
C62G	A20T	G63C	0.92	0.99	0.81	0.68	0.59	0.65	0.60	-0.15	0.15
A40T	A14G	G58C	0.99	0.80	0.99	0.90	0.94	0.99	0.71	0.20	-0.16
C49T	C34T	A65C	1.00	0.71	1.01	0.87	0.94	0.88	0.61	0.17	-0.20
G39A	C5T	C55T	0.99	0.99	0.72	0.94	0.87	0.86	0.61	0.15	-0.22
C49G	C46G	C62A	0.99	0.99	0.72	0.97	0.84	0.87	0.61	0.16	-0.22
G44A	C5T	C55T	0.98	0.99	0.72	1.08	0.78	0.86	0.63	0.15	-0.23
T54C	C5A	G44C	0.81	1.00	0.98	0.80	0.72	0.98	0.56	0.18	-0.20
C49G	T19C	G67T	0.99	0.97	0.92	1.14	0.70	1.07	0.60	0.18	-0.21
A20C	A14G	C46G	1.01	0.80	0.99	0.81	1.02	0.99	0.64	0.20	-0.18
A65T	C46T	C49A	0.99	0.99	0.76	0.86	0.88	0.96	0.55	0.21	-0.21
G48C	C4T	G25A	0.98	0.99	0.98	0.90	0.95	1.14	0.62	0.18	-0.26
A42G	A14G	C46G	0.93	0.80	0.99	0.78	0.89	0.99	0.52	0.20	-0.24
C31A	A45T	A70T	0.85	0.99	1.01	0.93	0.92	1.02	0.79	0.18	-0.26
T28A	C5T	G12C	0.89	0.99	0.98	0.82	0.86	1.04	0.55	0.16	-0.28
G50A	T32A	A69T	1.00	0.74	1.03	0.78	1.00	0.96	0.53	0.19	-0.25
G50A	C46G	C62A	1.00	0.99	0.72	1.04	0.88	0.87	0.60	0.16	-0.30
G67A	G48A	G52A	0.95	0.98	0.70	0.98	0.91	0.83	0.65	0.17	-0.30
A65G	A20G	G23A	1.00	1.00	0.74	0.89	0.99	0.91	0.57	0.17	-0.30
C46G	C31T	G63A	0.99	0.99	0.85	0.97	0.80	1.14	0.61	0.30	-0.18
C55T	G9A	C22A	0.72	0.99	1.00	0.92	0.65	0.92	0.63	0.20	-0.28
C55T	C4G	C5T	0.72	1.01	0.99	0.75	0.75	0.86	0.54	0.15	-0.32
G52A	C46A	G63C	0.70	1.00	0.81	0.82	0.65	1.00	0.56	0.18	-0.30
G23A	C29T	C30A	0.74	0.99	1.00	0.79	0.74	0.92	0.56	0.19	-0.31
C46G	T3A	A14G	0.99	1.00	0.80	0.93	0.98	0.99	0.62	0.20	-0.31
A14G	T2C	C5G	0.80	1.00	0.99	0.74	0.86	1.00	0.56	0.21	-0.30

All listed epistasis values are statistically significant (nominal $P < 0.05$, *t*-test from the six biological replicates).